

基于目标跟踪与轨迹聚类的行人移动数据挖掘方法研究

赛 斌, 曹自强, 谭跃进, 吕 欣

(国防科技大学 系统工程学院, 长沙 410073)

摘 要 随着物联网、大数据、人工智能等技术在安防领域不断取得突破性进展, 公共视频监测系统近年来得到飞跃式发展. 基于监控设备产生海量的非结构化视频数据, 通过对监控视频中的行人轨迹进行分析和研究, 可以挖掘出其中蕴含的行为模式, 这对人群行为研究有着重要的研究价值. 本文使用基于目标检测的多目标跟踪算法对地铁站出口, 商场出口等场景中的行人移动轨迹进行提取, 并在此基础上对行人的轨迹模式进行分析. 针对行人轨迹的特点, 在基于点密度聚类算法的基础上, 提出并实现了基于轨迹相似度的轨迹聚类方法. 结果表明, 该方法能够有效的提取行人轨迹, 并且从大规模轨迹数据中提取出轨迹模式.

关键词 目标检测; 多目标跟踪; 轨迹聚类; 轨迹模式; 人群行为

Pedestrian data mining with object tracking and trajectory clustering

SAI Bin, CAO Ziqiang, TAN Yuejin, LÜ Xin

(College of Systems Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract With the Internet of Things, big data, artificial intelligence making breakthroughs in the field of security, public video monitoring systems have developed quickly in recent years. The equipment generates massive amount of unstructured data, through analysis and research on pedestrian trajectory of video data, it can be found that the hidden behavior patterns contained which have an important research value. The article uses the multiple object tracking algorithm based on object detection to extract and describe the pedestrian movement trajectory in the surveillance video of subway station and mall exits, and then analyzed the trajectory pattern of pedestrians on the basis of trajectory. Aiming at the characteristics of pedestrian trajectory, a trajectory clustering method based on trajectory similarity was designed and implemented on the basis of point density clustering algorithm. The results showed that the method can effectively extract pedestrian trajectories, and extract trajectory patterns from large types of trajectory data.

Keywords object detection; multiple object tracking; trajectory clustering; trajectory pattern; crowd behavior

1 引言

随着城市建设和社会经济的发展, 人群集体活动的需求越来越频繁, 如大型集会, 演唱会以及宗教活动

收稿日期: 2019-12-31

作者简介: 赛斌 (1991–), 男, 回族, 硕士研究生, 研究方向: 大数据挖掘与应用; 曹自强 (1996–), 男, 硕士研究生, 研究方向: 大数据挖掘与应用; 谭跃进 (1958–), 男, 教授, 博士生导师, 博士, 研究方向: 复杂系统理论、武器装备体系论证; 通信作者: 吕欣 (1984–), 男, 教授, 博士, 研究方向: 大数据挖掘、人类行为动力学分析.

基金项目: 国家自然科学基金 (71771213, 91846301, 71790615, 71774168); 湖南省科技计划项目 (2017RS3040, 2018JJ1034)

Foundation item: National Natural Science Foundation of China (71771213, 91846301, 71790615, 71774168); Science and Technology Program of Hunan (2017RS3040, 2018JJ1034)

中文引用格式: 赛斌, 曹自强, 谭跃进, 等. 基于目标跟踪与轨迹聚类的行人移动数据挖掘方法研究 [J]. 系统工程理论与实践, 2021, 41(1): 231–239.

英文引用格式: Sai B, Cao Z Q, Tan Y J, et al. Pedestrian data mining with object tracking and trajectory clustering[J]. Systems Engineering — Theory & Practice, 2021, 41(1): 231–239.

等. 由于缺少相应的技术手段对大规模人群进行监控预警, 大规模人群聚集而导致的悲剧事故层出不穷. 通过检索紧急灾难数据库^[1] (emergency events database, EM-DAT) 可以发现, 从 1900 至 2019 年 10 月 31 日, 在亚洲地区总共发生死亡 10 人以上的踩踏事故 18 起, 死亡 3247 人. 其中比较严重事故有: 2014 年 12 月 31 日, 在上海外滩陈毅广场发生的踩踏事故导致 36 人死亡; 2015 年 9 月 24 日, 沙特阿拉伯麦加发生踩踏事故, 致使 2236 人死亡. 由此可见, 通过对人群活动进行实时监控和分析, 对于公共安全和城市应急规划十分重要^[2].

近年来, 随着智能监控、智慧城市、智能交通的建设, 监控设备每天能够产生海量的非结构化视频数据, 在这其中蕴含着丰富的人群行为信息. 针对这种情况, 有学者基于深度学习的框架提出相应的算法方案研究视频中的人群行为, 多目标跟踪^[3]正是其中的一种方法. 多目标跟踪 (multiple object tracking) 的主要任务是对视频中的多个目标物体进行识别和跟踪. 随着深度学习的发展, 以 YOLO^[4] 为代表的 “one-stage” 和以 Faster R-CNN^[5] 为代表的 “two-stage” 的目标检测算法, 不断提高着目标检测的速度和精度. 由于在实时性和准确性上有着良好的表现, 基于目标检测的目标跟踪算法 (tracking-by-detection) 在多目标跟踪领域逐渐成为主流.

目前, 已经有学者针对监控视频中的人群行为展开了研究. 例如, 国内清华大学刘晓冬^[6]等人开发了一套集运动目标检测、目标跟踪、目标分类为一体的智能监控系统. 湖南大学万琴^[7]等人提出的这一种固定监控场景的运动检测和目标跟踪方法. 国外学者 Nikouei^[8]等人提出了一种轻量级 CNN 算法, 将边缘网络服务应用到智能监控中, 具有运算速度快, 内存占用小等优点. 但是在面对复杂场景, 光线变化, 遮挡等复杂情况时, 上述系统仍然存在着准确率低的问题.

针对人群行为分析问题, 相关学者建立了各种模型^[9,10]分析行人的动力学行为, 其中 Henderson^[11-14]将气体动力学和流体动力学模型应用到人群行走规律的分析当中, 在假设动量守恒和能量守恒的前提下, 用二维气体的麦克斯韦-玻尔兹曼速度分布进行描述. Helbing^[15]改进了流体动力学模型, 在不进行任何假设的前提下, 用改进的类玻尔兹曼的气体动力学模型来描述不同的行人团体. 后来 Helbing, Johansson^[16,17]等针对流体动力学模型对于无法从微观层面描述个体行为的问题, 提出了社会力模型, 将行人的社会力分为自驱动力、排斥力和吸引力, 但存在着公式复杂, 参数校准难等问题^[18]. Moussaïd^[19]等使用行为启发式模型挖掘其中潜在的规律, 该模型在通过捕捉人群行为进行分析, 并且加入了行人的视觉信息^[20,21], 比较简单高效, 但无法对外力进行描述. Porter^[22]等人根据各个模型的特点, 提出了 IM 模型框架, 该模型可以很好地集成社会力模型、行为启发式模型以及材料科学理论, 发挥各模型的优势, 大大简化了运算量. 各类动力学模型对人群整体行为可以有很好的描述效果, 但是对个体的描述比较困难, 且数据获取方式主要以仿真实验为主.

实际的监控视频中记录了大量的轨迹数据, 为单独个体行为模式的挖掘提供了丰富的数据基础, 另一方面, 通过对大类具有相似行为的时空轨迹数据进行聚类, 从中可以挖掘出隐藏在海量行人轨迹中的轨迹特征和规律, 对人群行为监控, 异常行为预测和公共安全等有着重要意义, 但由于行人在移动时有随意性大的特点, 对于单个轨迹的挖掘有较大的困难. 聚类^[23]是挖掘轨迹模式的一种常用方法, 因此在行为模式挖掘过程中, 需要对轨迹数据首先进行轨迹聚类^[24], 即将具有相似行为的时空轨迹划分到一个轨迹簇内, 利用轨迹聚类的结果, 可以充分挖掘特定区域人群的行为模式.

在对现有基于深度学习框架基础上的算法进行学习和研究之后, 本文选用准确率和实时性较好的 YOLO v3^[25] 目标检测算法作为行人轨迹提取的基础, 并在此基础上加入 Deep SORT^[26] 多目标跟踪法, 解决了 YOLO v3 无法匹配近邻帧的问题, 以保证能够在复杂场景下对行人进行检测和跟踪, 获取到视频中行人的每一帧的坐标信息, 从而可以对行人的轨迹信息进行下一步的分析. 在聚类方法的选择上, 针对行人轨迹的特点, 选取 DBSCAN 聚类算法^[27]并进行改进, 对轨迹数据进行处理避免以轨迹噪声导致聚类结果不精确的问题, 并在此基础上实现了基于轨迹相似度的轨迹聚类算法, 框架结构如图 1 所示. 实验表明, 该方法可以对行人轨迹数据进行聚类, 并从中提取出典型轨迹模式, 在此基础上, 可以进一步分析行人的行为特征, 进而实现对人群的行人行为监控、异常行为预警等功能.

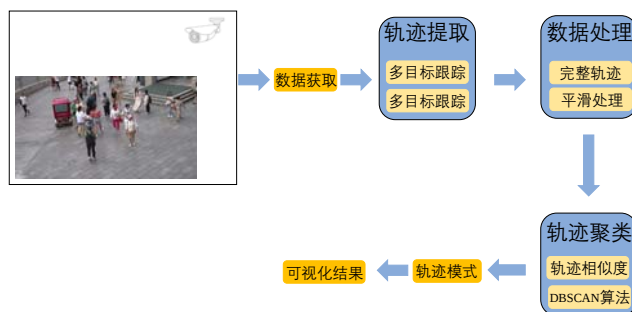


图 1 基于视频监控数据的行人轨迹提取及模式挖掘框架图

2 目标跟踪

2.1 目标检测

YOLO^[4,25,28] 是 2015 年由华盛顿大学的 Redmon 和 Farhadi 提出, 其全称为 “You Only Look Once”, 经过一系列的改进优化, 最新的 YOLO v3 算法于 2018 年提出. 该系列算法将目标检测问题设计为回归问题, 其中包括对多个边框和相关类别概率的回归. YOLO 只使用了一个神经网络和一次评价, 就可以从整幅的输入图像进行边框和类别概率的预测. 正是因为全流程只设计了一个网络, 所以它可以直接进行端对端的优化, 从而使得 YOLO 在运行速度方便比较占优势, 同时能够保持较高的检测平均准确率 (mAP).

YOLO 算法先将输入图像重新设置为固定大小的图片. 将重置后的图片作为卷积网络的输入, 运行卷积网络, 得到检测结果的边界框、坐标位置和置信度, 最后使用非极大值抑制对边界框进行筛选. YOLO v3-416 在 COCO^[29] 数据集上能够在 29 ms 内完成检测任务, 同时准确率能够维持在 55.3, 实时性和准确性具有较为明显的优势.

2.2 基于目标检测的多目标跟踪方法

随着深度学习在目标检测领域的广泛应用, 基于 Tracking-by-Detection 方式的多目标跟踪算法越来越成为主流. SORT^[3] 算法是在 2016 年由 Bewley, Ge 等人提出, 其全称为 simple online and realtime tracking, 在多目标跟踪过程中将卡尔曼滤波算法和匈牙利算法结合到一起, 其最快速度达到了 260 Hz. 2018 年基于 SORT 的改进算法 Deep SORT 被提出, 在多目标跟踪过程中, 结合目标运动和面特征信息提取, 并使用 CNN 网络在大规模行人数据集进行训练提取特征, 增强了算法的鲁棒性, 提升了跟踪效果, 表 1 为 SORT 和 Deep SORT 在公开数据集 MOT16 上的结果. Deep SORT 相较于 SORT 减少了大约 45% 的 ID switches 情况, 在高帧率视频中达到很好的跟踪效果.

表 1 在 MOT16 数据集的跟踪结果

跟踪算法	MOTA	MOTP	MT	ML	ID	Runtime
SORT	59.8	79.1	25.4	22.7	781	40 Hz
Deep SORT	61.4	79.1	32.8	18.2	1432	60 Hz

由于 YOLO v3 是基于单帧图像进行目标检测, 缺少上下帧的匹配信息, 并不能锁定检测目标进行跟踪. 因此, 本文在基于 YOLO v3 的目标检测算法的基础上, 加入 Deep SORT 跟踪器.

Deep SORT 算法主要分为卡尔曼滤波跟踪和匈牙利匹配两个部分. 目标检测算法对行人目标进行检测之后, 将目标位置和大小信息送入到跟踪器, 卡尔曼滤波其利用检测信息对目标进行跟踪预测, 得到相应行人下一帧的运动信息. 此时, 会得到跟踪器的预测结果和目标检测算法的检测结果, 如果同时输出两类信息, 则会导致重复检测, 降低算法信效率. 因此通过计算预测位置和检测位置之间的马氏空间距离, 以及检测框和预测框之间的匹配相似度, 利用匈牙利匹配算法进行预测匹配, 输出最终的跟踪结果. YOLO-Deep SORT 算法进行检测跟踪的基本思路如图 2 所示.

YOLO 检测算法的检测结果和加入 Deep SORT 的跟踪结果如图 3 所示, 分别截取第 100, 10005 和 30017 帧进行对比. 可以看到, 在加入 Deep SORT 多目标跟踪之后, 能够对多个行人目标进行持续跟踪并且记录轨迹.

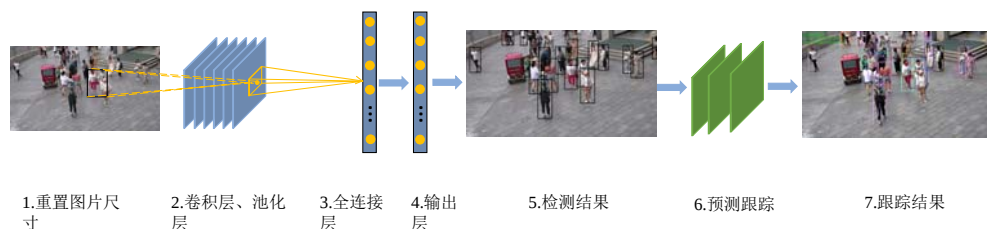


图 2 YOLO 目标检测的基本思路

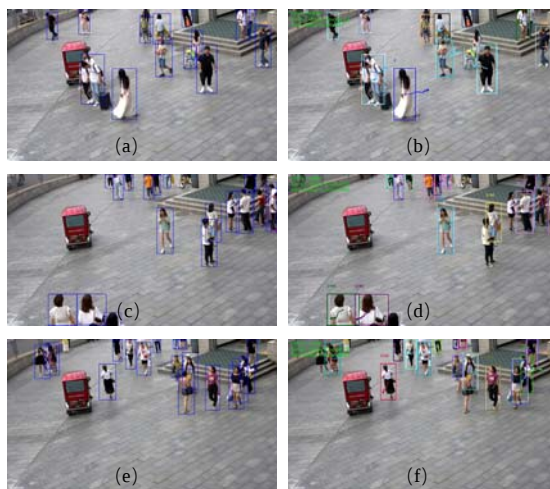


图 3 YOLO v3 和 YOLO-Deep SORT 结果对比

3 基于 Hausdroff 距离的轨迹聚类

轨迹聚类的目的在于发现该区域不同类别行人移动规律的特征, 从而挖掘出频繁项轨迹模式. 在进行轨迹聚类过程中, 聚类的对象为行人轨迹, 为解决轨迹间相似度的计算, 本文采用 Hausdroff 距离对轨迹整体相似度进行计算, 提出基于 Hausdroff 距离的 DBSCAN 聚类算法. DBSCAN 是一种基于点密度的聚类算法, 该算法具有噪声数据不敏感的特性并能划分出数据分布中高密度区域, 适用于轨迹聚类, 更容易从人群轨迹密集的区域发现簇, 进行轨迹聚类.

3.1 轨迹描述

视频中提取到的行人轨迹是一条按照时间序列排列的坐标序列, 坐标序列的个数可以表示轨迹的长度, 通过帧信息计算可以提取轨迹的时间信息. 定义轨迹集合^[30]为: $T = \{T_1, T_2, \dots, T_p\}$, 其中 p 为轨迹的个数. 目标 i 的移动轨迹表示为: $T_i = \{(x_1, y_1, t_1), (x_2, y_2, t_2), \dots, (x_J, y_J, t_J)\}$, 其中 x, y 为目标在视频单帧画面中的像素坐标信息, t_i 代表帧数, 即时间信息, J 为轨迹 i 的长度.

3.2 相似度度量

由于行人在行走时有较大的随意性, 因此每个行人产生的轨迹数据长度各不相同. 基于行人轨迹长度不一致的特点, 我们采用 Hausdroff 距离作为轨迹相似度的一种度量. Hausdroff 距离是对两个点集之间密集度的一种度量, 是点集之间距离的一种表达形式, 可以对任意两条轨迹进行计算且不需要考虑长度差异, 其定义为:

$$H(A, B) = \max[h(A, B), h(B, A)], \quad (1)$$

其中, $h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$, $h(B, A) = \max_{b \in B} \min_{a \in A} \|b - a\|$, 过 Hausdroff 距离能够体现轨迹之间的密集程度, 两条轨迹间的 Hausdroff 距离越小, 代表轨迹相似度越高.

3.3 算法描述

针对轨迹聚类的情况, 我们对 DBSCAN 算法中的相关概念进行调整, 引入核心轨迹、边界轨迹以及噪声轨迹的概念, 以满足轨迹聚类的要求, 下面给出相关定义.

a) MinTys: 判定是否为核心轨迹周围轨迹的数量阈值.

b) 核心轨迹: 如果某一轨迹在一定距离范围的领域内, 其邻居轨迹数量大于 MinTys, 则该轨迹为核心轨迹. 轨迹的领域由距离的度量方法和指定的距离参数决定.

c) 边界轨迹: 边界轨迹不属于核心轨迹, 但是位于核心轨迹的领域.

d) 噪声轨迹: 在轨迹集合中, 既不属于核心轨迹也不属于边界轨迹的轨迹.

基于 Hausdorff 距离的 DBSCAN 算法, 也只需要确定两个参数: 邻域半径 ε 和核心轨迹最小阈值 MinTys. 算法伪代码如算法 1 所示.

Algorithm 1 轨迹聚类算法.

输入:

轨迹段集合 $TS = \{TR_1, TR_2, \dots, TR_n\}$;

邻域半径 ε 和最小阈值 MinTys;

输出:

轨迹聚类结果集合 $O = \{C_1, C_2, \dots, C_m\}$;

/* 第 1 阶段: 轨迹 Hausdorff 距离计算

```

1: for each  $TR_i, TR_j \in TS \ \& \ i \neq j$  do
2:   compute  $Haus(TR_i, TR_j)$ ;
3:   if  $Haus(TR_i, TR_i) \geq \varepsilon$  then
4:      $N_\varepsilon(TR_i) \leftarrow TR_j$ 
5:     /* 设定  $TR_j$  的领域集合 */
6:   end if
7: end for
/* 第 2 阶段: 轨迹进行聚类 */
8: Set clusterID = 0;
9: Mark all trajectory as unclassified;
10: for each  $TR_i \in TS \ \& \ TR_i$  is unclassified do
11:   if  $|N_\varepsilon(TR_i)| > MinTys$  then
12:     for each  $TR_j \in N_\varepsilon(TR_i)$  do
13:       if  $TR_j$  is unclassified then
14:         Assign clusterID to  $\forall TR \in N_\varepsilon(TR_i)$ 
15:         Add  $N_\varepsilon(TR_i) - TR_i$  to Queue  $Q$ 
16:         while  $Q \neq \emptyset$  do
17:           for each  $Q_i \in Q$  do
18:             if  $|N_\varepsilon(Q_i)| \geq MinTys$  then
19:               for each  $X \in N_\varepsilon(Q_i)$  do
20:                 if  $X$  is noise then
21:                   Assign clusterID to  $X$ 
22:                 else
23:                   Add  $X$  to QUEUE  $Q$ 
24:                 end if
25:               end for
26:             else
27:               Remove  $Q_i$  from Queue  $Q$ 
28:               clusterID ++
29:             end if
30:           end for
31:         end while
32:       else
33:         Make  $TR_i$  as noise
34:       end if
35:     end for
36:   end if
37: end for

```

4 实验结果及分析

4.1 实验设计

本次实验使用 Sony A7 R III 全画幅微单数码相机, 采用固定机位拍摄的方式对两个地铁出入口和一个十字路口, 三个场景进行 20 分钟的视频拍摄, 拍摄时间为 2019 年 9 月 3 日。如图 4, 在本次实验的场景中, 视频监控设备设置在距离地面有一定距离的位置, 可以确保拍摄到的行人目标不会太小, 避免 YOLO 对于小目标检测不敏感的缺陷。采集到分辨率为 1920×1080 , 帧速率 25, 码流 50M 的视频。在轨迹提取和聚类实验环节, 硬件环境为: Windows7 操作系统, Inter Core-i7 处理器 (2.80 GHz), 内存 16 GB, NVIDIA 940MX(2 GB) 显卡, Python3.6 作为编译语言。



图 4 拍摄场景

4.2 完整轨迹筛选

利用基于 YOLO v3 的 Deep SORT 算法对采集到的视频进行行人轨迹提取, 算法将视频中目标移动的坐标进行记录, 得到一系列的坐标点, 并将分属于同一个目标的坐标带连在一起成为一条轨迹。其中场景一采集到 3009 条行人轨迹, 场景二采集到 1470 条, 场景三为 1632 条。其中场景一采集到 3009 条行人轨迹, 场景二采集到 1470 条, 场景三为 1632 条, 所有轨迹均采用 2.1 节中的描述方法表示。场景一中的地铁口邻近商场, 且规格较大, 行人流量较多, 主要的出入口有右上角 (地铁站出口), 顶部 (路口), 右边 (路口) 以及底部 (商场门口)。场景二和场景三邻近十字路口, 附近商场和出入口较少, 因此相较于场景一, 总体人流量较小。场景二包括左上角 (路口), 右边 (路口) 以及左下角 (路口) 三个出入口, 场景三包括左上角 (地铁站出口), 右上角 (路口) 以及底部 (路口) 三个出入口。在提取过程除严重遮挡情况外, 均能成功跟踪行人并记录轨迹数据。三个场景的提取过程和结果如图 5 所示。

提取到轨迹数据后, 我们对轨迹的起点和终点进行了密度图分析, 图 6 展示了三个场景的起点和终点分布, 其中左侧为起点分布图, 右侧为终点分布图, 从图中可以看出起始点和终止点主要分布在视频画面的四周。Deep SORT 在目标跟踪过程中, 由于遮挡导致目标丢失, 当在再次检测到该目标后, 算法误认为是新发现的行人目标, 因此会在画面中产生起始点和终止点。为确保提取的轨迹数据能够完整的表示某一行人在该区域完整的移动轨迹, 本文对轨迹数据进行了筛选, 删除掉起始点和终止点不在视频四边产生并且记录帧数小于 100 的轨迹, 最终得到 504, 222, 400 条完整轨迹。

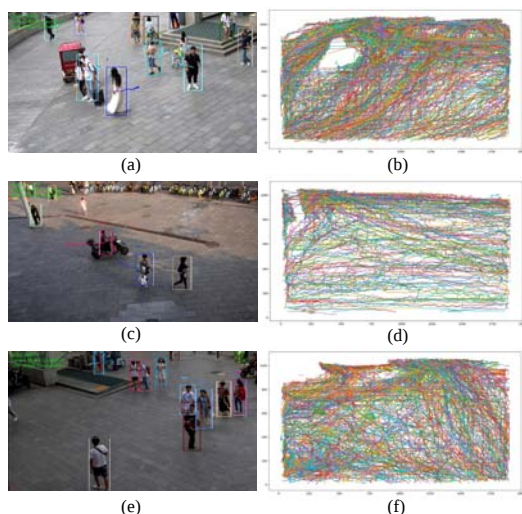


图 5 YOLO-Deep SORT 算法对行人轨迹进行提取

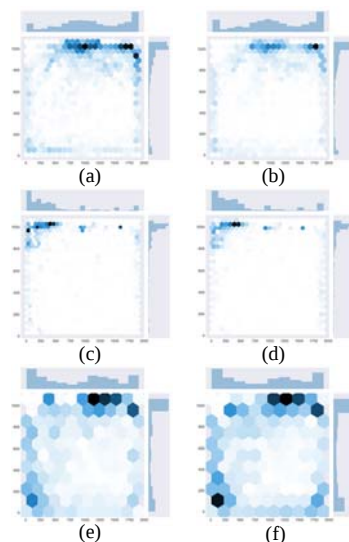


图 6 起点和终点密度分布图

4.3 轨迹平滑处理

由于 Deep SORT 算法提取的是视频中行人目标跟踪框的中心点坐标, 在实际的跟踪过程中, 由于目标跟踪框会根据行人形态发生尺寸变化, 导致中心点产生飘逸, 不符合人行行走时的真实情况. 图 7 中的完整轨迹有很多波动, 因此需要对轨迹数据进行平滑处理, 以提高轨迹相似度的计算和聚类结果准确性.

在实际场景中行人的移动速度接近于匀速, 并且采集到的数据为时间序列顺序, 所以本文采用滑动平均平滑引用对采集到的轨迹数据进行平滑处理, 在对轨迹平滑处理的过程中, 尽可能使处理后的轨迹数据保留原始视频中的行人信息.

通过滑动平均平滑将邻域内的数据点进行平均处理, 并代替中心点值, 消除序列中的扰动, 反映轨迹序列的趋势. 记 $t(i)$ 为原始轨迹数据, 则进行滑动平均平滑后的轨迹为:

$$T_s(i) = \frac{1}{n}(t(i+N) + t(i+(N-1)) + \cdots + t(i)), \quad (2)$$

其中 N 为滑动时间窗口长度. 由于拍摄时选择 25 帧速率进行拍摄, 所以在滑动平均平滑过程中, 令 $N = 25$, 对轨迹数据进行平滑处理. 图 7 为三个场景的平滑处理效果对比图, 左侧为完全轨迹, 右侧为平滑处理后的效果. 各个场景中的轨迹消除了波动, 并且没有产生大幅度的偏移, 能够准确反映行人的移动趋势.

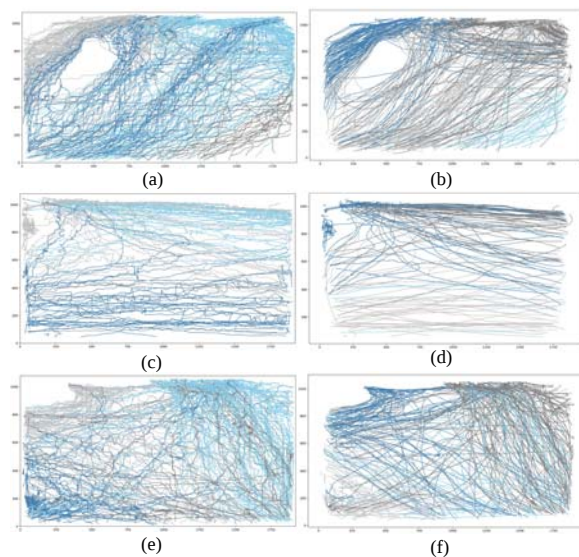


图 7 轨迹提取平滑处理效果对比

4.4 聚类结果

为了比较不同参数的聚类结果的质量, 通过计算聚类轨迹簇中各轨迹间的平方误差来评价聚类结果^[31]:

$$QMeasure = \sum_{i=1}^{num_{clus}} \left(\frac{1}{2|C_i|} \sum_{x \in C_i} \sum_{y \in C_i} dist(x, y)^2 \right). \quad (3)$$

其中, $QMeasure$ 值越小表示轨迹簇内轨迹相似度越高, 聚类质量越好. $dist(x, y)$ 为轨迹簇中任意两条轨迹的 Hausdorff 距离. 在结合聚类评价公式和人工经验的基础上, 经过多次试验和优化, 最终确定场景一中 $eps=1750$, $MinTys=5$; 场景二 $eps=1650$, $MinTys=6$; 场景三 $eps=1850$, $MinTys=5$. 如图 8 所示, 为优化过程的可视化结果, 其中 x 轴代表 eps 取值, y 轴为 $MinTys$, z 轴为 $QMeasure$ 值, $QMeasure$ 均能在上述参数值情况下取得最小值.

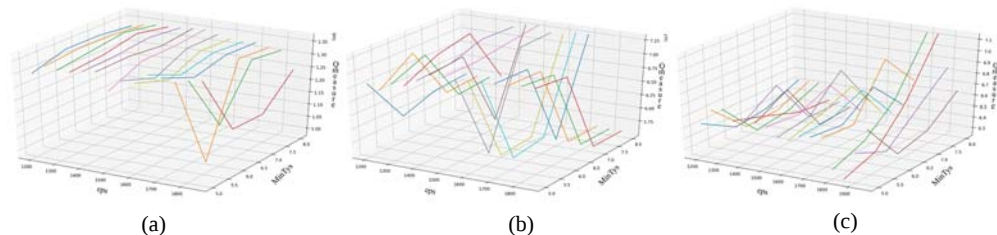


图 8 轨迹聚类参数优化

图 9 为各场景中轨迹数据集的聚类结果以及轨迹簇数量的分布. 其中场景一得到 6 个聚类结果, 场景二 5 个, 场景三 8 个, 噪声轨迹的占比分布为 11.7%, 9.9%, 21.25%. 该方法能够成功地从处理过后的轨迹数据集重提取出行人的轨迹模式, 与采集到的的视频进行对照, 提取到的轨迹模式能够准确的刻画出该区域行人的轨迹模式和行为规律. 以场景一为例 (图 10(a)), 该算法准确地提取出该场景中的五个主要行人通道: 紫色轨迹表示直接靠左边穿越整个区域的轨迹, 红色轨迹为左下角至右上方的轨迹簇, 主要包括地铁口进出行人和靠近右下方穿越整个区域的轨迹簇, 蓝色为从画面上方进入到地铁口的人群轨迹簇, 绿色轨迹是画面右边至地铁口路径的行人轨迹簇, 橙色为画面上方经过地铁站出口至画面右侧的轨迹簇.

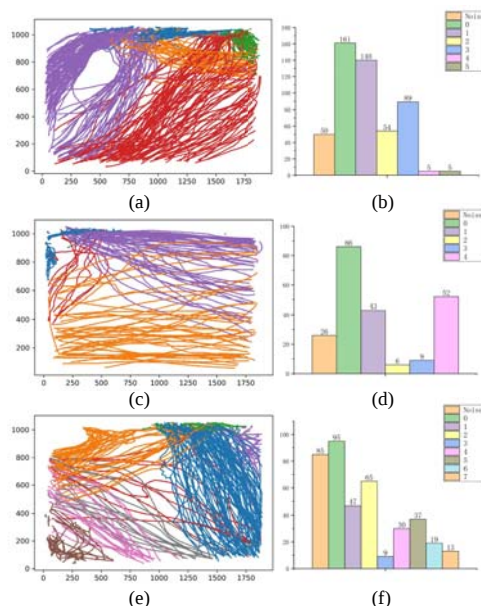


图 9 聚类结果和分布

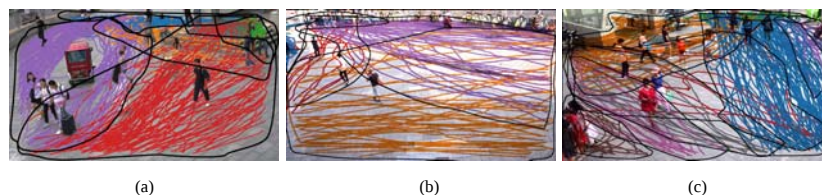


图 10 轨迹聚类效果

5 结束语

本文设计并实现了对公共区域人群的行人轨迹提取方法, 并且用于行人行为模式挖掘. 该方法的实现过程主要分为轨迹提取, 轨迹平滑和轨迹聚类三个阶段. 在轨迹提取阶段, 在 YOLO v3 目标检测算法的基础上加入 Deep SORT 多目标跟踪方法, 有效解决了 YOLO v3 无法匹配邻近帧信息的问题, 提出采集行人轨迹的有效方法; 在轨迹处理阶段, 主要对采集到的原始轨迹数据进行平滑处理和相似度计算; 轨迹聚类阶段, 改进基于密度的聚类算法对轨迹进行聚类并挖掘出轨迹模式. 实验结果表明, 该方法可以有效地对监控视频中的行人目标进行轨迹提取和聚类分析, 具有较强的鲁棒性和可移植性. 该方法不仅可以应用到公共安全事件预警和城市规划, 而且提取和挖掘方法对于公共安全公开数据集建设、大规模群体行为研究都具有重要意义.

参考文献

- [1] EM-DAT. The international disaster database[EB/OL]. 2012. <https://www.emdat.be/>.
- [2] 吕欣. 大数据技术在应急救援领域的应用及展望 [J]. 中国计算机学会通讯, 2018, 14: 56-62.
Lü X. Application and prospect of big data technology in emergency rescue field[J]. Communications of the CCF, 2018, 14: 56-62.

- [3] Bewley A, Ge Z, Ott L, et al. Simple online and realtime tracking[C]// 2016 IEEE International Conference on Image Processing (ICIP), 2016: 3464–3468.
- [4] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779–788.
- [5] Girshick R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440–1448.
- [6] 刘晓冬, 苏光大. 一种可视化智能户外监控系统 [J]. 中国图象图形学报: A 辑, 2000, 5(12): 1024–1029.
Liu X D, Su G D. A visual intelligent outdoor surveillance system[J]. Journal of Image and Graphics: A, 2000, 5(12): 1024–1029.
- [7] 万琴, 王耀南. 一种多运动目标检测, 跟踪方法研究与实现 [J]. 计算机应用研究, 2007, 24(1): 199–202.
Wan Q, Wang Y N. Research and implementation of detecting and tracking multiple moving objects method[J]. Application Research of Computers, 2007, 24(1): 199–202.
- [8] Nikouei S Y, Chen Y, Faughnan T R. Smart surveillance as an edge service for real-time human detection and tracking[C]// 2018 IEEE/ACM Symposium on Edge Computing (SEC), 2018: 336–337.
- [9] Barbosa H, Barthelemy M, Ghoshal G, et al. Human mobility: Models and applications[J]. Physics Reports, 2018, 734: 1–74.
- [10] Yan X Y, Wang W X, Gao Z Y, et al. Universal model of individual and population mobility on diverse spatial scales[J]. Nature Communications, 2017, 8(1): 1639.
- [11] Henderson L. The statistics of crowd fluids[J]. Nature, 1971, 229(5284): 381–383.
- [12] Henderson L, Lyons D. Sexual differences in human crowd motion[J]. Nature, 1972, 240(5380): 353–355.
- [13] Henderson L F. On the fluid mechanics of human crowd motion[J]. Transportation Research, 1974, 8(6): 509–515.
- [14] Henderson L, Jenkins D. Response of pedestrians to traffic challenge[J]. Transportation Research, 1974, 8(1): 71–74.
- [15] Helbing D. A fluid dynamic model for the movement of pedestrians[J]. arXiv preprint cond-mat/9805213, 1998.
- [16] Helbing D. Interrelations between stochastic equations for systems with pair interactions[J]. Physica A: Statistical Mechanics and its Applications, 1992, 181(1–2): 29–52.
- [17] Johansson A, Helbing D, Shukla P K. Specification of the social force pedestrian model by evolutionary adjustment to video tracking data[J]. Advances in Complex Systems, 2007, 10(Supp2): 271–288.
- [18] Frank G A, Dorso C O. Room evacuation in the presence of an obstacle[J]. Physica A: Statistical Mechanics and its Applications, 2011, 390(11): 2135–2145.
- [19] Moussaïd M, Helbing D, Theraulaz G. How simple rules determine pedestrian behavior and crowd disasters[J]. Proceedings of the National Academy of Sciences, 2011, 108(17): 6884–6888.
- [20] Turner A, Penn A. Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment[J]. Environment and Planning B: Planning and Design, 2002, 29(4): 473–490.
- [21] Batty M. Predicting where we walk[J]. Nature, 1997, 388(6637): 19–20.
- [22] Porter E, Hamdar S H, Daamen W. Pedestrian dynamics at transit stations: An integrated pedestrian flow modeling approach[J]. Transportmetrica A: Transport Science, 2018, 14(5–6): 468–483.
- [23] 韩家炜, 坎伯. 数据挖掘: 概念与技术 [M]. 2 版. 北京: 机械工业出版社, 2012.
- [24] 吉根林, 赵斌. 时空轨迹大数据模式挖掘研究进展 [J]. 数据采集与处理, 2015, 30(1): 47–58.
Ji G L, Zhao B. Research progress in pattern mining for big spatio-temporal trajectories[J]. Journal of Data Acquisition and Processing, 2015, 30(1): 47–58.
- [25] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv Preprint arXiv: 1804.02767, 2018.
- [26] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric[C]// 2017 IEEE International Conference on Image Processing (ICIP), 2017: 3645–3649.
- [27] Murça M C R, Hansman R J, Li L, et al. Flight trajectory data analytics for characterization of air traffic flows: A comparative analysis of terminal area operations between New York, Hong Kong and Sao Paulo[J]. Transportation Research Part C: Emerging Technologies, 2018, 97: 324–347.
- [28] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517–6525. <https://doi.org/10.1109/CVPR.2017.690>
- [29] Chen X, Fang H, Lin T Y, et al. Microsoft coco captions: Data collection and evaluation server[J]. arXiv Preprint arXiv: 1504.00325, 2015.
- [30] 康凯. 视频监控轨迹数据模式学习及行为分析 [D]. 北京: 北京交通大学, 2014.
Kang K. Motion pattern study and analysis from video monitoring trajectory[D]. Beijing: Beijing Jiaotong University, 2014.
- [31] Lee J G, Han J, Whang K Y. Trajectory clustering: A partition-and-group framework[C]// Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, 2007: 593–604.