



(12) 发明专利

(10) 授权公告号 CN 110909647 B

(45) 授权公告日 2021.11.02

(21) 申请号 201911118462.2

(22) 申请日 2019.11.15

(65) 同一申请的已公布的文献号
申请公布号 CN 110909647 A

(43) 申请公布日 2020.03.24

(73) 专利权人 清华大学
地址 100084 北京市海淀区清华园1号

(72) 发明人 侯静轩 龙瀛 陈龙

(74) 专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201
代理人 廖元秋

(51) Int.Cl.
G06K 9/00 (2006.01)

(56) 对比文件

US 2019287264 A1,2019.09.19

CN 105931168 A,2016.09.07

Seok-Han Lee et al..A Single-View
Based Framework for Robust Estimation of
Heights and Positions of Moving People.
《2010 Digest of technical papers
international conference on consumer
electronics》.2010,

唐丽玉.视频图像中监控目标的空间定位方
法.《福州大学学报(自然科学版)》.2014,第42卷
(第1期),

审查员 郑岩

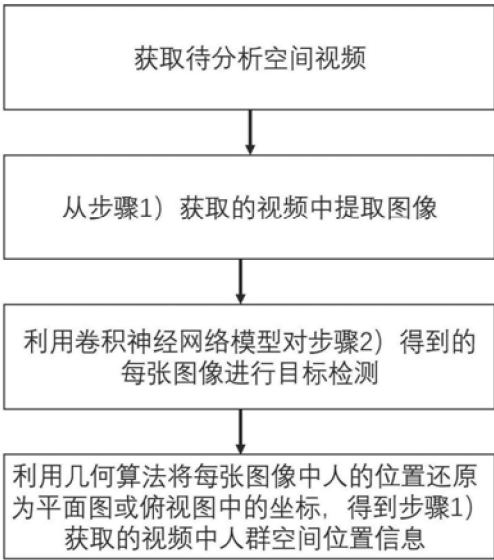
权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种获取视频中人群空间位置的方法

(57) 摘要

本发明提出一种获取视频中人群空间位置的方法,属于城市空间分析技术领域。该方法首先获取待分析空间视频,并提取相应的图像;利用卷积神经网络模型对每张图像进行目标检测;对目标检测结果中保留识别结果为人的目标识别框,利用几何算法将每张图像中人的位置还原为平面图或俯视图中的坐标,得到视频中人群空间位置信息。本发明具有能够快速、准确的获取人群的实际地理位置信息的优点,且在城市空间分析和城乡规划评估等方面有较好的应用前景。



1. 一种获取视频中人群空间位置的方法,其特征在于,包括以下步骤:

1) 获取待分析空间视频;

获取反映待分析空间中人群活动的待分析空间视频,对所述视频有以下要求:无镜头失真、拍摄视角固定、视频能够覆盖完整的空间、视频中在待分析空间中活动的人为完整图像且视频中需要被识别的人大于2像素;视频长度不小于15分钟;

2) 从步骤1) 获取的视频中提取图像;

按照设定的帧数间隔从步骤1) 获取的视频中提取图像;所述帧数间隔计算表达式如下:

$$TI = GW / AS * FN$$

其中,TI为帧数间隔;GW为待分析空间对应的平面图或俯视图中分析网格的边长,AS为人平均运动速度,FN为视频拍摄帧率;

3) 利用卷积神经网络模型对步骤2) 得到的每张图像进行目标检测;

利用ssd_resnet_50_fpn作为目标检测的卷积神经网络模型,所述模型的输入是步骤2) 得到的每张图像,模型的输出包括:输入图像对应的识别范围,该图像中每个识别范围的置信度,每个识别范围中目标的类别以及该图像对应的帧数序号;其中每个识别范围对应一个目标识别框,每个识别范围包含四个数值,分别代表在输入图像目标识别平面直角坐标系下该识别范围对应的目标识别框的左下角X坐标、左下角Y坐标、右上角X坐标和右上角Y坐标,记为(Xa,Ya,Xc,Yc),每个坐标值取值范围均为0至1;其中,所述目标识别平面直角坐标系的原点为图像左上角顶点,X轴从上至下取值范围为0到1,Y轴从左至右取值范围为0到1;

4) 利用几何算法将每张图像中人的位置还原为平面图或俯视图中的坐标,得到步骤1) 获取的视频中人群空间位置信息;具体步骤如下:

4-1) 利用步骤3) 的结果,对每张图像,保留卷积神经网络模型输出中目标类别为1的目标识别框,其中1代表行人;

4-2) 获取待分析空间的平面图或俯视图,在平面图或俯视图中,建立待分析空间最小包络正方形,该正方形的四个顶点分别记为左下角顶点PA、右下角顶点PB、右上角顶点PC和左上角顶点PD;采用平面图或俯视图的分析网格平面直角坐标系分别标记PA、PB、PC和PD的X、Y坐标;其中,所述平面图或俯视图的分析网格平面直角坐标系原点为该平面图或俯视图中最小包络正方形左下角顶点PA,水平从左到右为X轴正方向、垂直从下到上为Y轴正方向;

4-3) 任意选取一张步骤2) 中获取的图像,在该图像中标记步骤4-2) 中得到待分析空间最小包络正方形在该图像中对应的四边形的四个顶点,分别记为左下角顶点A、右下角顶点B、右上角顶点C和左上角顶点D,并获取A、B、C、D在以该图像左下角作为坐标原点、水平从左到右为X轴正方向、垂直从下到上为Y轴正方向的该图像分析网格平面直角坐标系中X、Y坐标;

4-4) 使用几何算法按照待分析空间的真实面积将步骤2) 得到的每张图像中的待分析空间划分为 4^n 个分析网格;获取该图像中每个分析网格的四个顶点在图像分析网格平面直角坐标系中的坐标;

其中,n的计算表达式如下:

$$n = \lceil (L/GW)^{0.25} \rceil - 1$$

式中, L 为每张图像中待分析空间最小包络正方形的边长;

4-5) 对步骤2) 获取的每张图像,计算经过步骤4-1) 保留的每个识别框中底边中点在该图像目标识别平面直角坐标系中的坐标,并将该坐标转换为图像分析网格平面直角坐标系对应的坐标;

其中,目标识别平面直角坐标系中目标识别框底边中点坐标为:

$$[X_c, (Y_a+Y_c)/2]$$

转换后,图像分析网格平面直角坐标系中目标识别框底边中点坐标为:

$$[(Y_a+Y_c)/2, 1-X_c]$$

4-6) 统计所有图像中每个识别框底边中点在各个分析网格中出现的次数,得到每个分析网格对应的统计结果;

4-7) 在平面图或俯视图分析网格平面直角坐标系中,根据 PA 、 PB 、 PC 和 PD 的 X 、 Y 坐标,使用几何算法将平面图或俯视图中的最小包络正方形化分为 4^n 个分析网格;

4-8) 将步骤4-6) 得到的每个分析网格对应的统计结果赋值到步骤4-7) 得到的平面图或俯视图中的 4^n 个分析网格中;

4-9) 按照各分析网格中底边中点数量从低到高的顺序利用从浅到深的颜色分别对步骤4-4) 和步骤4-8) 获得的分析网格进行填色,得到视频中人群空间位置的可视化结果;

步骤4-6) 获得的统计结果和步骤4-9) 获得的可视化结果即为从视频中获取的人群空间位置信息。

一种获取视频中人群空间位置的方法

技术领域

[0001] 本发明属于城市空间分析技术领域,特别提出一种获取视频中人群空间位置的方法。

背景技术

[0002] 信息技术领域的不断发展为大规模的数据处理提供了条件,特别是近年来深度学习技术的发展极大地推动了计算机视觉领域的研究,使得从常见的录影设备中获得的数据,如照片、视频等,有了新的分析手段。视频图像中蕴含了非常丰富环境特征和人群行为信息,利用视频分析人群行为可以帮助我们更直观地理解人群的行为特征。如果能够获取视频中人群空间位置,将极大的提升城市空间分析的精度和效率。

[0003] 现有的通过视频获取人群信息的方法为,在研究区域内布置监控摄像头,并使用卷积神经网络中的目标检测算法,对监控摄像头所录视频内人群行为进行分析。在此基础上通过将人群信息使用热力图形式进行可视化,可获取人群对于空间的使用情况。

[0004] 目前通过视频获取的人群信息仅能在显示视频中人群位置,而无法将人群信息精确转换至实际的地理空间中。这是由于目前图像处理中使用的透视法无法还原因为透视导致的空间变形。以现有的图像变现方法为例,对于一个形状为正正方形的广场,其四个顶点A、B、C、D的投影坐标分别为 $[0,0]$ 、 $[0,1]$ 、 $[1,1]$ 和 $[1,0]$,其中点E的坐标为 $[0.5,0.5]$ 。假设在使用摄像设备对广场进行拍摄后,正方形的广场根据单点透视的原理变形为梯形,顶点A、B、C、D在图像上的坐标变为 $[0,0]$ 、 $[0,1]$ 、 $[0.8,0.6]$ 和 $[0.2,0.6]$,通过对角线交点法可知中点E的坐标变为 $[0.5,0.375]$ 。现在Photoshop等图像软件使用的拉伸复原法为将图像的边框与原空间进行匹配,及保持A、B点位置不变,将C、D点的纵坐标增大66.7%,横坐标至0.5的距离增加66.7%,使A、B、C、D在图像上从 $[0,0]$ 、 $[0,1]$ 、 $[0.8,0.6]$ 和 $[0.2,0.6]$ 复原为 $[0,0]$ 、 $[0,1]$ 、 $[1,1]$ 和 $[1,0]$ 。但此时中的E的坐标为 $[0.5,0.625]$,而不是正确的 $[0.5,0.5]$ 。因此使用此方法获取人群位置会存在很大误差。而空间规划和设计所需要的人群信息均是依托于地理坐标系或投影坐标系,无法获取图像或视频中人群在空间中的准确位置,将导致视频中的人群信息难以被用于支撑相关工作。

发明内容

[0005] 本发明的目的是为克服已有技术的不足之处,提出一种获取视频中人群空间位置的技术。本发明具有能够快速、准确的获取人群的实际地理位置信息的优点,且在城市空间分析和城乡规划评估等方面有较好的应用前景。

[0006] 本发明提出一种获取视频中人群空间位置的方法,其特征在于,包括以下步骤:

[0007] 1) 获取待分析空间视频;

[0008] 获取反映待分析空间中人群活动的待分析空间视频,对所述视频有以下要求:无镜头失真、拍摄视角固定、视频能够覆盖完整的空间、视频中在待分析空间中活动的人为完整图像且视频中需要被识别的人大于2像素;视频长度不小于15分钟;

[0009] 2) 从步骤1) 获取的视频中提取图像;

[0010] 按照设定的帧数间隔从步骤1) 获取的视频中提取图像;所述帧数间隔计算表达式如下:

[0011] $TI = GW / AS * FN$

[0012] 其中, TI为帧数间隔;GW为待分析空间对应的平面图或俯视图中分析网格的边长, AS为人平均运动速度, FN为视频拍摄帧率;

[0013] 3) 利用卷积神经网络模型对步骤2) 得到的每张图像进行目标检测;

[0014] 利用ssd_resnet_50_fpn作为目标检测的卷积神经网络模型,所述模型的输入是步骤2) 得到的每张图像,模型的输出包括:输入图像对应的识别范围,该图像中每个识别范围的置信度,每个识别范围中目标的类别以及该图像对应的帧数序号;其中每个识别范围对应一个目标识别框,每个识别范围包含四个数值,分别代表在输入图像目标识别平面直角坐标系下该识别范围对应的目标识别框的左下角X坐标、左下角Y坐标、右上角X坐标和右上角Y坐标,记为(X_a, Y_a, X_c, Y_c),每个坐标值取值范围均为0至1;其中,所述目标识别平面直角坐标系的原点为图像左上角顶点,X轴从上至下取值范围为0到1,Y轴从左至右取值范围为0到1;

[0015] 4) 利用几何算法将每张图像中人的位置还原为平面图或俯视图中的坐标,得到步骤1) 获取的视频中人群空间位置信息;具体步骤如下:

[0016] 4-1) 利用步骤3) 的结果,对每张图像,保留卷积神经网络模型输出中目标类别为1的目标识别框,其中1代表行人;

[0017] 4-2) 获取待分析空间的平面图或俯视图,在平面图或俯视图中,建立待分析空间最小包络正方形,该正方形的四个顶点分别记为左下角顶点PA、右下角顶点PB、右上角顶点PC和左上角顶点PD;采用平面图或俯视图的分析网格平面直角坐标系分别标记PA、PB、PC和PD的X、Y坐标;其中,所述平面图或俯视图的分析网格平面直角坐标系原点为该平面图或俯视图中最小包络正方形左下角顶点PA,水平从左到右为X轴正方向、垂直从下到上为Y轴正方向;

[0018] 4-3) 任意选取一张步骤2) 中获取的图像,在该图像中标记步骤4-2) 中得到待分析空间最小包络正方形在该图像中对应的四边形的四个顶点,分别记为左下角顶点A、右下角顶点B、右上角顶点C和左上角顶点D,并获取A、B、C、D在以该图像左下角作为坐标原点、水平从左到右为X轴正方向、垂直从下到上为Y轴正方向的该图像分析网格平面直角坐标系中X、Y坐标;

[0019] 4-4) 使用几何算法按照待分析空间的真实面积将步骤2) 得到的每张图像中的待分析空间划分为 4^n 个分析网格;获取该图像中每个分析网格的四个顶点在图像分析网格平面直角坐标系中的坐标;

[0020] 其中,n的计算表达式如下:

[0021] $n = \lfloor (L/GW)^{0.25} \rfloor - 1$

[0022] 式中,L为每张图像中待分析空间最小包络正方形的边长;

[0023] 4-5) 对步骤2) 获取的每张图像,计算经过步骤4-1) 保留的每个识别框中底边中点在该图像目标识别平面直角坐标系中的坐标,并将该坐标转换为图像分析网格平面直角坐标系对应的坐标;

[0024] 其中,目标识别平面直角坐标系中目标识别框底边中点坐标为:

[0025] $[X_c, (Y_a+Y_c)/2]$

[0026] 转换后,图像分析网格平面直角坐标系中目标识别框底边中点坐标为:

[0027] $[(Y_a+Y_c)/2, 1-X_c]$

[0028] 4-6) 统计所有图像中每个识别框底边中点在各个分析网格中出现的次数,得到每个分析网格对应的统计结果;

[0029] 4-7) 在平面图或俯视图分析网格平面直角坐标系中,根据PA、PB、PC和PD的X、Y坐标,使用几何算法将平面图或俯视图中的最小包络正方形化分为 4^n 个分析网格;

[0030] 4-8) 将步骤4-6)得到的每个分析网格对应的统计结果赋值到步骤4-7)得到的平面图或俯视图中的 4^n 个分析网格中;

[0031] 4-9) 按照各分析网格中底边中点数量从低到高的顺序利用从浅到深的颜色分别对步骤4-4)和步骤4-8)获得的分析网格进行填色,得到视频中人群空间位置的可视化结果;

[0032] 步骤4-6)获得的统计结果和步骤4-9)获得的可视化结果即为从视频中获取的人群空间位置信息。

[0033] 本发明的特点及有益效果:

[0034] 相比于现有获取视频内空间技术的有偏性,使用本发明不存在图像中透视变形导致的位置偏差,有效的弥补了现有方法的缺点,其优点体现在以下三个方面:本发明使图像中的人群信息具有了地理空间属性,因此能够使用图像数据真实地反映人群行为,且可反向检查验证;结合本发明获取的人群行为数据与空间测绘数据,可进一步将空间要素落位,进而研究人群行为特征受环境空间的影响方式及程度,反映空间特征以及人群与空间要素的互动关系,弥补相关研究方法的不足;本方法极大的提高了传统城市空间分析方法在数据获取方面的时间粒度和空间粒度,使得人们的在精细粒度上的空间使用情况得以量化,能够用于对特定空间的精细化研究。

[0035] 本发明在城市空间分析领域的应用前景主要可体现在三个方面:通过细化空间网络和缩小图像采集间隔,可识别空间内人群行为轨迹,并用于研究群体在某一空间内的行为路线规律;通过长时间的摄影可计算每个网格的累积人次,可判断该研究范围内人群活跃的空间;本发明将平面要素与转换后的空间网格进行嵌入分析,相较于目前多基于人群数据的研究而言,本发明具有较强的空间关联性,可反映真实空间要素,因此可从环境行为学的角度探索人群活动与环境要素之间的关系,有利于拓展其基础理论。

附图说明

[0036] 图1是本发明方法的整体流程图;

[0037] 图2是本发明实施例中自行采集分析空间视频时拍摄位置示例图;

[0038] 图3是本发明实施例中利用几何算法将图像中人的位置还原为平面图中的坐标中将图像中透视变形的空间等分为实际面积、形状相同的四个初始分析网格方法的示例图;

[0039] 图4是本发明实施例中获取视频中人群空间位置结果可视化后的示例图。

具体实施方式

[0040] 本发明提出一种获取视频中人群空间位置的方法,下面结合附图及具体实施例对本发明进一步详细说明如下。

[0041] 本发明提出一种获取视频中人群空间位置的方法,将经过透视变形的人的位置还原至平面图中;该方法整体流程如图1所示,具体步骤如下:

[0042] 1) 获取待分析空间视频;

[0043] 获取能够反映待分析空间中人群活动的待分析空间视频。本发明对所述视频有以下要求:无镜头失真、拍摄视角固定、视频能够覆盖完整的空间、视频中在待分析空间中活动的人能够从头至脚获取完整图像以及视频中需要被识别的人应大于2像素。视频长度不少于15分钟,所述视频的获取可通过调取监控录像或自行采集。

[0044] 其中自行采集分析空间视频需要进行摄影设备选择、拍摄位置选择和拍摄时间设置,具体如下:1) 摄影设备选择方面,本发明要求选择无透视畸变的摄影设备,焦距可采用35-300毫米,而对于摄影设备、镜头的品牌和视频格式无特殊要求。2) 拍摄位置选择分为“高位”与“低位”两种(实施例中自行采集分析空间视频时拍摄位置示例图如图2所示,图2(a)为“低位”拍摄、图2(b)为“高位”拍摄,图中虚线框内为摄影设备位置)。“低位”指将设备放置于墙顶端或平房屋顶等距地面处置高度6米以内的位置,使得摄影设备能够完整拍摄分析空间。“高位”指将设备设置于距地面垂直高度高于6米的位置,通常为楼道或房顶上,使得摄影设备能够完整拍摄分析空间。为了空间中所有人位置均能显示在视频中而不会因前后遮挡难以识别,当存在“高位”拍摄的条件时均应将摄影设备设置于高位。当不具备“高位”拍摄的设备且“高位”拍摄时人的位置会被树叶等高处的障碍物遮挡时且当拍摄空间长度小于20米时可选择“低位”拍摄。3) 拍摄时间方面,需按拍摄地点的日出日落情况进行选择日出后、日落前,晴天或阴天无雨的时间段进行拍摄,且每次拍摄不小于15分钟。

[0045] 2) 从步骤1) 获取的视频中提取图像;

[0046] 从视频中提取图像时需要根据分析网格长度、人平均运动速度及所拍摄视频的帧率计算帧数间隔。具体参数计算公式为:

[0047] $TI = GW / AS * FN$

[0048] 其中TI指帧数间隔,代表从构成视频的全部图像中每隔多少图像中抽取出一帧用于后面的分析;GW指待分析空间对应的平面图或俯视图中分析网格的边长(单位:米);AS指人平均运动速度;FN指视频拍摄帧率。设定分析网格时,要尽量使在每帧图像中,每个网格仅有一单独出行的人。自然情况下陌生人之间的最小距离为0.9米,因此网格边长不应过显著大于0.9米,所以将GW设定为0.9米。行人的平均步行速度为1.2米/秒,因此AS设定为1.2米/秒。本实施例中使用常规摄像设备进行视频拍摄,采用帧率为24的普通拍摄模式,因此FN设定为24,则 $TI = 0.9 / 1.2 * 24 = 18$,即从视频中每18帧中抽取一帧用于后续步骤。

[0049] 3) 利用卷积神经网络模型对步骤2) 得到的每张图像进行目标检测;

[0050] 本发明选择了ssd_resnet_50_fpn作为目标检测的卷积神经网络模型,兼具目标识别准确率和识别速度。所述模型的输入是步骤2) 得到的每张图像,模型的输出包括:输入图像对应的若干个识别范围,该图像中每个识别范围的置信度,每个识别范围中目标的类别以及帧数;其中每个识别范围对应一个目标识别框,每个识别范围包含四个数值,分别代表在输入图像目标识别所用平面直角坐标系(该坐标系原点为图像左上角顶点,从原点出

发, X轴从上至下取值范围为0到1, Y轴从左至右取值范围为0到1, 则图像中每个像素的X轴坐标取值范围为0到1、Y轴坐标轴从左至右取值范围为0到1) 下该识别范围对应的目标识别框的左下角X坐标、左下角Y坐标、右上角X坐标和右上角Y坐标, 记为 (Xa, Ya, Xc, Yc); 每个识别范围的置信度, 取值范围0至1, 表示该识别范围 (即目标识别框) 中的目标识别为正确的可能性; 每个目标识别框中的目标类别, 通过数字表示该目标识别框所识别目标的类别, 其中1代表行人、2代表自行车、3代表机动车 (本实施例共80个类别, 具体对应表可于一下链接中查询: <https://github.com/pjreddie/darknet/blob/master/data/coco.names>); 帧数, 指输入图像对应的帧的序号数。

[0051] 本实施例模型输出的结果如表1所示

[0052] 表1本实施例中卷积神经网络模型输出表

识别范围x	置信度	目标类别	帧数
0.41, 0.35, 0.55, 0.51	0.99	3	1
0.54, 0.48, 0.74, 0.52	0.99	1	1
0.68, 0.14, 0.90, 0.22	0.98	1	2
0.67, 0.41, 0.76, 0.44	0.96	1	2

[0054] 表1中帧数值为1和2, 代表对两张图像进行了目标识别。帧数为1的图像中共有两个目标识别框, 其中识别范围为 (0.41, 0.35, 0.55, 0.51) 的识别框内的目标类别为3, 即机动车, 识别置信度为99%; 识别范围为 (0.54, 0.48, 0.74, 0.52) 的识别框内的目标类别为1, 即行人, 识别置信度为99%。帧数为2的图像中共有两个目标识别框, 识别范围为 (0.68, 0.14, 0.90, 0.22) 的识别框内的目标类别为1即行人, 识别置信度为98%; 识别范围为 (0.67, 0.41, 0.76, 0.44) 的识别框内的目标类别为1即行人, 识别置信度为96%。

[0055] 4) 利用几何算法将每张图像中人的位置还原为平面图或俯视图中的坐标, 得到步骤1) 获取的视频中人群空间位置信息; 具体步骤如下:

[0056] 4-1) 利用步骤3) 的结果, 对每张图像, 保留卷积神经网络模型输出中目标类别为1的目标识别框;

[0057] 4-2) 获取待分析空间的平面图或俯视图 (平面图可由为人工测绘获得或从设计院等有测绘能力的单位处获取, 俯视图可从卫星影像或垂直视角的航拍图中截取), 在平面图或俯视图中, 建立待分析空间最小包络正方形, 该正方形的四个顶点分别记为左下角顶点PA、右下角顶点PB、右上角顶点PC和左上角顶点PD; 采用平面图或俯视图的分析网格平面直角坐标系 (该坐标系原点为该平面图或俯视图中最小包络正方形左下角顶点PA, X轴坐标从左到右数值增加、Y轴坐标从下到上数值增加, 长度单位为米) 分别标记待分析空间最小包络正方形四个顶点PA、PB、PC和PD的X、Y坐标;

[0058] 4-3) 任意选取一张步骤2) 中获取的图像, 在该图像中标记步骤4-2) 中得到待分析空间最小包络正方形在该图像中对应的四边形的四个顶点, 分别记为左下角顶点A、右下角顶点B、右上角顶点C和左上角顶点D, 并获取A、B、C、D在以该图像左下角作为坐标原点、水平从左到右为X轴正方向、垂直从下到上为Y轴正方向的该图像分析网格平面直角坐标系中X、Y坐标;

[0059] 4-4) 使用几何算法按照待分析空间的真实面积将步骤2) 得到的每张图像中的待分析空间划分为 4^n 个分析网格 (n为正整数); 获取该图像中每个分析网格的四个顶点在图

像分析网格平面直角坐标系中的坐标。

[0060] 其中,按照真实面积将图像中的分析空间划分为等大的分析网格的几何算法,首先将图像中透视变形的分析空间等分为在步骤4-2)中获取的平面图或俯视图中几何上全等、即对应的空间的真实面积相等的四个初始矩形分析网格,通过递归运算再将分析网格从4个细分为16、64…直至 4^n 个矩形分析网格。

[0061] 划分分析网格的具体方法为:对于图像中需要划分的四边形ABCD,取对角线AC与BD的交点R;对边AD、BC延长线相交H;对边AB、DC延长线交点E;延长ER,与AD、BC相较于点F、G;延长HR,与DC、AB相交于点I、J;ABCD分为AJRF、JBGR、RGCI、FRID四个初始分析网格(本实施例示意图如图3所示)。

[0062] 递归运算中n的取值公式为:

[0063] $n = \lfloor (L/GW)^{0.25} \rfloor - 1$

[0064] 式中L为每张图像中待分析空间最小包络正方形的边长;对于每张图像L的取值相等。

[0065] 4-5)对步骤2)获取的每张图像,计算经过步骤4-1)保留的每个识别框中底边中点在该图像目标识别平面直角坐标系中的坐标,并将该坐标转换为图像分析网格平面直角坐标系对应的坐标;

[0066] 其中,目标识别平面直角坐标系中目标识别框底边中点坐标为:

[0067] $[X_c, (Y_a + Y_c) / 2]$

[0068] 转换后,图像分析网格平面直角坐标系中目标识别框底边中点坐标为:

[0069] $[(Y_a + Y_c) / 2, 1 - X_c]$

[0070] 式中, X_c 、 Y_a 、 Y_c 为目标识别框平面直角坐标系中目标识别框的右上角顶点的X坐标、左下角顶点的Y坐标和右上角顶点的Y坐标。

[0071] 4-6)统计所有图像中每个识别框底边中点在各个分析网格中出现的次数,得到每个分析网格对应的统计结果;

[0072] 4-7)在平面图或俯视图分析网格平面直角坐标系中,根据PA、PB、PC和PD的X、Y坐标,重复步骤4-4),使用几何算法将平面图或俯视图中的最小包络正方形化分为 4^n 个分析网格;

[0073] 4-8)将步骤4-6)得到的每个网格对应的统计结果赋值到步骤4-7)得到的平面图或俯视图中对应的 4^n 个分析网格中;

[0074] 4-9)按照各分析网格中底边中点数量从低到高的顺序利用从浅到深的颜色分别对步骤4-4)和步骤4-8)获得的分析网格进行填色,得到视频中人群空间位置的可视化结果;图4是本发明实施例中获取视频中人群空间位置结果可视化后的示例图,图4(a)为在图像上进行可视化的示例、图4(b)为在平面图上进行可视化示例,网格中人群使用次数由少到多使用颜色由浅至深显示。

[0075] 步骤4-6)获得的统计结果和步骤4-9)中获得的可视化结果即为从视频中获取的人群空间位置信息。

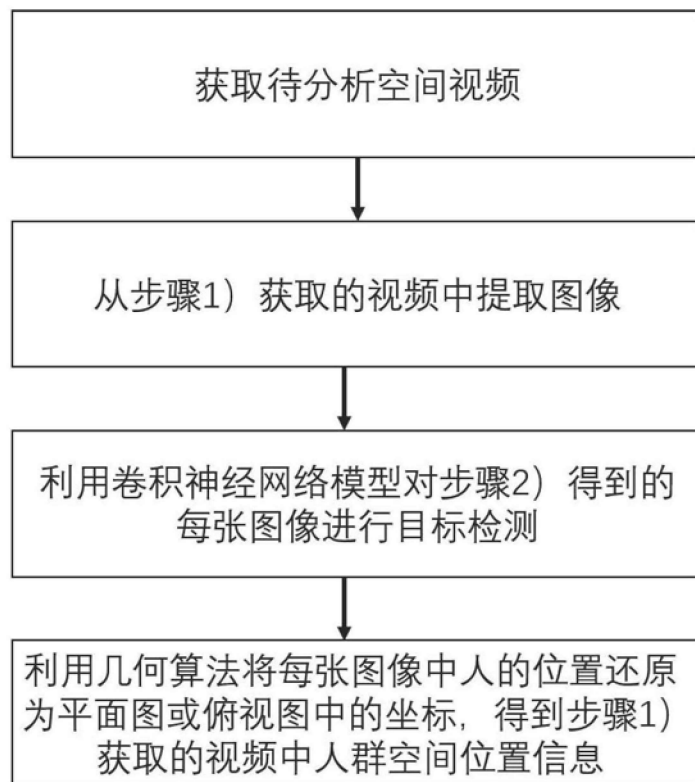


图1

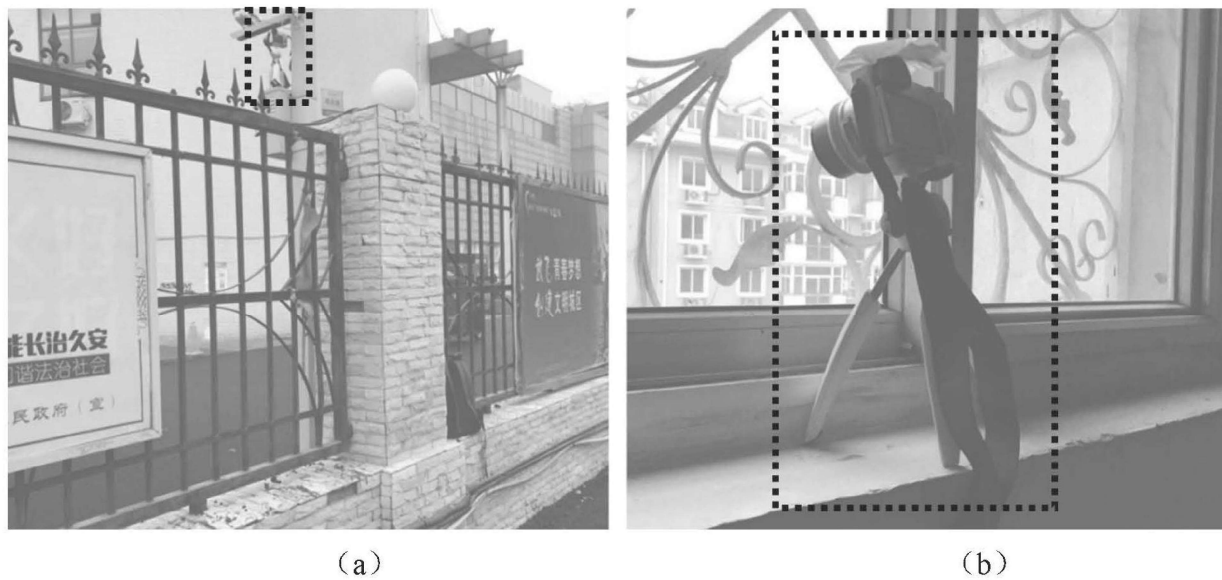


图2

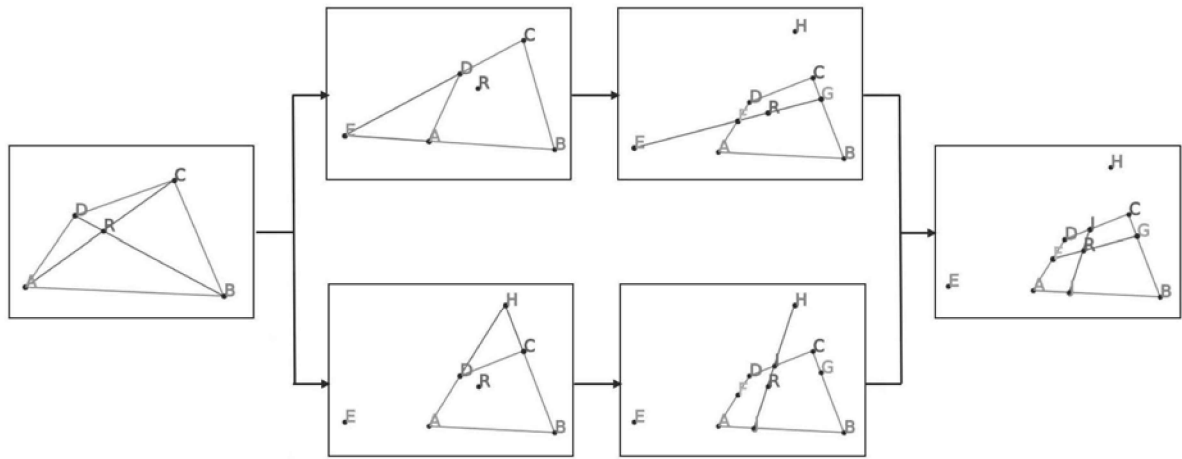


图3



(a)



(b)

图4