

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 1 月 6 日 (06.01.2022)



(10) 国际公布号
WO 2022/001925 A1

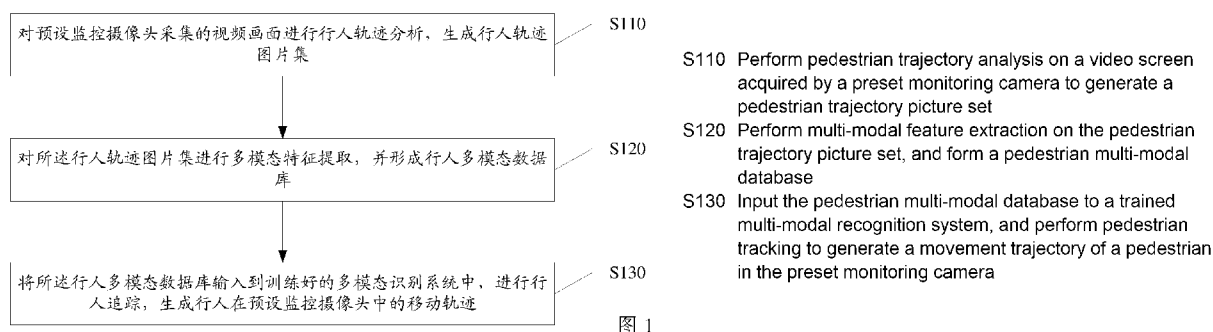
- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2021/102652
- (22) 国际申请日: 2021 年 6 月 28 日 (28.06.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010603573.9 2020 年 6 月 29 日 (29.06.2020) CN
- (71) 申请人: 中兴通讯股份有限公司 (**ZTE CORPORATION**) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。
- (72) 发明人: 窦笑 (**DOU, Xiao**); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。 申光 (**SHEN, Guang**); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。 侯春华 (**HOU, Chunhua**); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦,

Guangdong 518057 (CN)。 李东方 (**LI, Dongfang**); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。

- (74) 代理人: 北京天昊联合知识产权代理有限公司 (**TEE & HOWE INTELLECTUAL PROPERTY ATTORNEYS**); 中国北京市东城区东长安街 1 号东方广场东方经贸城西一办公楼 5 层 1、6-12 室崔利梅/卢吉辉, Beijing 100738 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,

(54) **Title:** PEDESTRIAN TRACKING METHOD AND DEVICE, AND COMPUTER READABLE STORAGE MEDIUM

(54) 发明名称: 行人追踪方法和设备, 及计算机可读存储介质



(57) **Abstract:** A pedestrian tracking method and device, and a computer readable storage medium, relating to the technical field of communications. The pedestrian tracking method comprises: performing pedestrian trajectory analysis on a video screen acquired by a preset monitoring camera to generate a pedestrian trajectory picture set (S110); performing multi-modal feature extraction on the pedestrian trajectory picture set, and forming a pedestrian multi-modal database (S120); and inputting the pedestrian multi-modal database to a trained multi-modal recognition system, and performing pedestrian tracking to generate a movement trajectory of a pedestrian in the preset monitoring camera (S130).

(57) **摘要:** 一种行人追踪方法和设备, 以及计算机可读存储介质, 涉及通信技术领域。行人追踪方法包括: 对预设监控摄像头采集的视频画面进行行人轨迹分析, 生成行人轨迹图片集 (S110); 对行人轨迹图片集进行多模态特征提取, 并形成行人多模态数据库 (S120); 以及, 将行人多模态数据库输入到训练好的多模态识别系统中, 进行行人追踪, 生成行人在预设监控摄像头中的移动轨迹 (S130)。

NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

行人追踪方法和设备，及计算机可读存储介质

本申请要求在 2020 年 6 月 29 日提交中国专利局、申请号为 202010603573.9 的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本申请涉及通信技术领域。

背景技术

视频监控已经遍布在我们生活的各个角落中，并且人脸识别技术已经非常成熟。然而，在实际的安防应用场景中，并非所有的摄像头都能够拍摄到清晰的人脸，由于头发、口罩和帽子等遮挡，很难通过人脸识别系统判定行人身份。不仅如此，在实际的应用场景中，一个摄像头往往无法覆盖所有的区域，而且多个摄像头之间一般也没有重叠，因此跨镜追踪检索系统对人员进行锁定和查找显得十分必要。

目前，跨镜追踪技术在产业界和学术界都受到广泛关注，并取得显著进展，从政策方面看，公安部推出平安城市的概念，并且发布了多项预研课题，相关的行业标准也在紧锣密鼓的制定当中。

发明内容

本申请实施例的一个方面提供一种行人追踪方法，包括：对预设监控摄像头采集的视频画面进行行人轨迹分析，生成行人轨迹图片集；对行人轨迹图片集进行多模态特征提取，并形成行人多模态数据库；以及，将行人多模态数据库输入到训练好的多模态识别系统中，进行行人追踪，生成行人在预设监控摄像头中的移动轨迹。

本申请实施例的另一个方面提供一种行人追踪设备，包括：存储器、处理器、存储在存储器上并可在处理器上运行的程序，以及用于实现处理器和存储器之间的连接通信的数据总线；程序被处理器执行，以实现本申请实施例提供的行人追踪方法的至少一个步骤。

本申请实施例的再一个方面提供一种计算机可读存储介质，其上存储有一个或者多个程序，一个或者多个程序可被一个或者多个处理器执行，以实现本申请实施例提供的行人追踪方法的至少一个步骤。

5 附图说明

图 1 是本申请实施例提供的行人追踪方法的一种流程图。

图 2 是本申请实施例提供的行人追踪方法的一种流程图。

图 3 是本申请实施例提供的行人追踪系统的一种结构框图。

10 具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

在后续的描述中，使用用于表示元件的诸如“模块”、“部件”或“单元”的后缀仅为了有利于本申请的说明，其本身没有特有的意义。因此，“模块”、“部件”或“单元”可以混合地使用。

跨境追踪检索系统使用最多的就是行人重识别。在该领域，大多研究者一般采用基于行人图片特征来定位和检索行人的方案，这样对行人特征的鲁棒性提出了很高的要求，但是真实的场景往往都非常复杂，比如无正脸照、姿态变换、服装变换、遮挡、光线、摄像头分辨率较低和室内外环境变化等，这些因素通常会导致行人检索与追踪的失败。

本申请提出了一种基于多模态的跨境追踪检索系统，以多目标行人追踪为基础，结合行人重识别网络、行人质量分析、行人属性分析、人脸识别和摄像头的时间和空间位置信息，来进一步提高跨境追踪检索的准确率和速度。

如图 1 所示，本申请实施例提供了一种行人追踪方法，包括以下步骤 S110-步骤 130。

在步骤 S110 中，对预设监控摄像头采集的视频画面进行行人轨迹分析，生成行人轨迹图片集。

在步骤 S120 中，对行人轨迹图片集进行多模态特征提取，并形

成行人多模态数据库。

在步骤 S130 中，将行人多模态数据库输入到训练好的多模态识别系统中，进行行人追踪，生成行人在预设监控摄像头中的移动轨迹。

5 在一种可实施方式中，该行人追踪方法还可包括：接收目标行人轨迹，提取目标行人的多模态特征，并在行人多模态数据库中查找与目标行人的多模态特征匹配的第一行人轨迹；将目标行人轨迹和第一行人轨迹合并生成第二行人轨迹，在行人多模态数据库中查询与第二行人轨迹匹配的行人轨迹；以及，根据与第二行人轨迹匹配的行人轨迹，生成目标行人在预设监控摄像头中的移动轨迹。

10 在一种可实施方式中，该行人追踪方法还可包括：从行人轨迹图片集选取质量参数在预设范围内的图像，并对选取的质量参数在预设范围内的图像进行特征提取。

在一种可实施方式中，可根据训练集对多模态识别系统中各模态参数的影响因子进行调整，得到训练好的多模态识别系统。

15 在一种可实施方式中，行人轨迹图片集中的图片名称可包括：轨迹标识（ID）、视频帧号、图片拍摄时间，和/或地点信息。

在一种可实施方式中，生成行人在预设监控摄像头中的移动轨迹，可包括：根据监控摄像机分布拓扑的图结构进行行人的移动规律分析。

20 具体地，可将监控摄像头的时空拓扑关系与目标的外观表现模型匹配算法进行结合，使用监控摄像头拓扑的图结构分析行人移动和转移的规律，从而对行人的跨镜追踪进行时空约束。如果跟踪目标在某个节点（摄像头）处消失，则在其相邻几步可达的节点处进行目标检测，再进行匹配与关联。

25 更进一步地，空间关系定义了节点之间是否建立边，以及边的方向。图模型的建立过程中，如果两个节点之间在物理空间位置上一步可达，即不经过其他任何节点，则为它们之间建立一条边。

在实际应用系统中，采用统计学习方法为目标的运动建立时间约束，从而定义节点之间合理的权值。想要获取一组摄像头节点数据的统计规律往往比较困难，这是由许多因素决定的：包括目标的运动

30

规律、摄像机的地理位置、监控周边交通环境变化等。在本申请实施例中，对所有观测时间进行聚类并计算每个类内的方差；根据摄像头相对坐标及路线情况初始化权重，根据行人重识别比对情况进行修正。

考虑到行人无法同时出现在多个摄像头以及从一个摄像头移动到另一个摄像头需要考虑其时间统计规律，可以利用该时空约束显著减少待查询的样本量，减少了查询时间并提高了检索性能。

结合摄像头的空间经纬坐标及可行走路线的空间约束，可以估计出摄像头节点间的连接关系及初始移动时间。后续结合行人重识别的间隔时间进行不断的修正，可以获得摄像头网络拓扑的边权值。

后续查询时，首先根据待查询轨迹确定以此节点为中心的摄像机网络拓扑中的临近节点，再结合边权值限定临近节点中查询数据的时间范围。在每个临近节点的对应时间范围内进行轨迹匹配。

在临近节点 A 中匹配到目标后，则以该节点为新的网络中心，在摄像机网络拓扑中的临近节点中继续查询，并更新行人行进轨迹和出现的时间节点。在查询结束后完成行人行进轨迹的绘制。

若在推荐的时间范围内没有匹配到目标，则在扩大的时间范围内进一步查询，若仍没有查询到，则在以该节点为中心的下一层临近节点进行查询。

在一种可实施方式中，多模态特征可包括以下一种或多种：行人特征、人脸特征和行人属性特征。行人特征可包括以下一种或多种：高矮胖瘦的体型特征和姿势特征。人脸特征信息可包括以下一种或多种：脸型特征、脸部表情特征和肤色特征。行人属性信息可包括以下一种或多种：发型长短、发型颜色、服装款式、服装颜色以及携带物品。

本申请实施例还提供一种行人追踪设备，包括存储器、处理器、存储在存储器上并可在处理器上运行的程序，以及用于实现处理器和存储器之间的连接通信的数据总线。程序被处理器执行时，可实现本申请实施例提供的行人追踪方法的至少一个步骤，例如，图 1 所示的步骤。

本申请实施例还提供一种计算机可读存储介质，其上存储有一

个或者多个程序，该一个或者多个程序可被一个或者多个处理器执行，以实现本申请实施例提供的行人追踪方法的至少一个步骤，例如，图 1 所示的步骤。

本申请实施例公开了一种利用多模态信息融合来对不同摄像头
5 下同行人进行检索与追踪的系统，如图 2 所示，该系统实施行人追踪方法可包括以下步骤 S1-步骤 S6。

在步骤 S1 中，获取监控区域中不同摄像头的视频。

在步骤 S2 中，对获取到的离线视频进行行人检测，并完成行人
10 轨迹提取，相应的行人轨迹图片集中的图片名以轨迹 ID、视频帧号
和对应时间、地点复合命名（例 0001_00025_202003210915_NJ），
保存在以轨迹 ID 命名的子文件夹下。

在步骤 S3 中，通过行人质量分析，提取行人轨迹图片集中质量
参数在预设范围内的图像；在一种可实施方式中，可选取图片质量较
好且时间较分散的 5 张图片，作为最佳 5 张（top5）行人轨迹。

15 在步骤 S4 中，分别利用行人重识别网络、人脸识别网络和行人
属性网络提取 top5 行人轨迹的行人特征、人脸特征（若未检测到人
脸数据置为空）和行人属性特征，特征提取完成，把三种特征和（轨
迹 ID、视频帧号、时间、地点）保存到数据库中。

在一种可实施方式中，行人重识别网络是利用计算机视觉技术
20 判断图像或者视频序列中是否存在特定行人的技术，通过行人重识别
网络确定行人特征。

行人属性网络用于提取行人属性，行人属性是关于行人外貌的
语义性描述，人体不同的部位具有不同的属性，例如，人体头部相关
属性有“短发”及“长发”等；衣服款式相关属性有“长袖”、“短
25 袖”、“连衣裙”和“短裤”等；携带物相关属性有“双肩包”、“单
肩包”、“手提包”和“无携带”等。在不同环境和场合可以对行人
属性进行选择 and 细分，使之有利于行人的再识别。行人属性信息与人的
外观信息相关联的，是更为具体的语义信息，在进行行人比对或行人
检索时，根据行人属性的相似情况可以过滤不相关数据。

30 在步骤 S5 中，采用一批人工标注好的测试集来优化多模态权重

参数。

在步骤 S6 中，输出最终检测结果。

本申请实施例提供的方案相对于其它只基于行人图片特征来定位和检索的方案，融合了人脸、行人属性、时间和空间等多模态信息，使得检索更具备鲁棒性，更能够适应复杂的真实场景。

图 3 是本申请实施例提供的基于多模态检索的跨镜行人追踪系统的一种结构框图。如图 3 所示，该系统可包括：数据获取与轨迹提取模块、最佳轨迹提取模块、特征提取和多模态信息入库模块、多模态权重参数调整模块，和检索交互和结果展示模块。

数据获取与轨迹提取模块，从监控的视频单元中获取离线视频图像，每个监控单元只负责自己区域的数据保存和提取，保存到指定文件夹下，对已保存的视频进行轨迹追踪提取，并对行人图片自动标注轨迹 ID、图片帧号、时间信息和位置信息。

最佳轨迹提取模块，从行人轨迹中筛选出行人质量较完整的，同时，具有较大时间间隔的 5 张行人轨迹图片。

特征提取和多模态信息入库模块是提取行人、人脸和行人属特征，并将这三种特征和轨迹 ID、行人轨迹图片时间和空间信息，保存到行人多模态数据库中。

多模态权重参数调整模块是使用一批标注的测试集来优化多模态参数值的权重，最终达到针对不同的数据集有自己最优的模式参数。

检索交互和结果展示模块能够提供界面化操作轨迹搜索轨迹和图片搜索轨迹，并能够显示最优轨迹和每个摄像头下的最优轨迹排名，并能够通过轨迹中的图片帧号，可以搜索到视频中该轨迹，并进行实时播放。

本申请实施例提供的基于多模态的跨镜追踪检索系统，可获取监控区域的离线视频，对视频中的行人进行行人检索，采用轨迹追踪算法进行行人轨迹提取，并对每张图片完成以轨迹 ID、视频帧号 and 对应时间、地点复合命名，通过行人质量分析提取轨迹中的最佳 5 张行人图片。对其所有轨迹图片进行人脸、行人和行人属性特征进行提取，特征提取完成后将所有的多模态信息入库。使用测试集对多模

态系统参数进行自适应调整，最后完成跨镜头的行人轨迹搜索，并将结果显示在界面上。该方法相对人工检索方法，很大程度上减少了工作量，高效率的同时还有较高的准确率，该方案可实现跨镜头的行人检索，为智慧安防和平安城市提供了相当有力的支持。

5 请参见图 2，下面对上述各步骤进行详细的描述。

 在步骤 S1 中，确定检索区域，并获取该区域监控的离线视频，该区域可以是商场、写字楼、居民小区和社区等相对固定的场所，并且该离线视频应该是某一时间段的，至少是同一天的监控视频。将视频保存到本地，并标记摄像头 ID、位置和起始时间，在本实施例中，
10 选取三个不同角度的镜头，摄像头 ID 为 C0、C1 和 C2。

 在步骤 S2 中，对各个摄像头下的离线视频进行行人检测、轨迹追踪。相应的行人轨迹图片集中图片以轨迹 ID、视频帧号 and 对应时间、地点复合命名（例 0001_00025_202003210915_NJ），保存在以轨迹 ID 命名的子文件夹下。这里行人检测模型采用的是 SSD (Single Shot MultiBox Detector，深度学习目标检测算法) 算法来获取当前
15 帧行人的位置框和边界框，采用匈牙利追踪算法获取行人轨迹。

 在步骤 S3 中，对上步获取的行人轨迹，使用行人质量分析模型，这里采用的是人体骨骼关键点检测算法，通过骨骼关键点的个数来判断行人的完整性，若图片中行人骨骼关键点的个数等于预设值，则判断获取的该行人的图片信息完整。在本实施例中，选取的关键点包括：
20 头部、肩部、手掌、脚掌。对于有较多行人图片的轨迹，提取该轨迹中质量较好且时间较分散的 5 张图片，作为 top5 行人轨迹。

 在步骤 S4 中，分别利用行人重识别网络、人脸识别网络和行人属性网络提取 top5 行人轨迹的行人特征、人脸特征（若未检测到人脸数据置为空）和行人属性特征，特征提取完成，把三种特征和（轨迹 ID、视频帧号、时间、地点）保存到行人多模态数据库中。
25

 在步骤 S5 中，因为跨镜追踪的数据集对场景要求非常严格，在网络上没有任何资源可供使用，所以采用了自己构建的训练集，通过提取三个不同镜头下的离线监控视频，三个摄像头分别命名为 C0、
30 C1 和 C2。然后对离线视频进行行人多目标检测跟踪、审核和人工标

注。查询 (query) 数据归属于 C0 摄像头, 被查询轨迹 (gallery) 数据分属于两个摄像头 C1 和 C2。使用该标注好的训练集优化多模态识别系统, 经过上述 S4 步骤, 完成一系列的多模态信息入库 (该数据库归属于多模态权重优化数据库, 与之前 S4 步生成的信息检索数据库不存在冲突), 每条轨迹包含 5 张图片, 在进行人脸、行人特征及行人属性比对时将采取批量特征比对的方式进行, 然后采用 C0→C1, C0→C2 的检索方式, 最终对 C0→C1, C0→C2, 统计检索命中率。然后动态的调整多模态权重参数, 重新进行 C0→C1, C0→C2, 统计检索命中率。当检索命中率达到最高时认为, 当前的多模态参数为最优多模态参数, 既完成多模态参数特征的优化调整。

在步骤 S6 中, 最终采用优化好的多模态权重参数, 对 S4 步已经生成的信息检索数据库, 进行跨镜下的行人检索, 输出最终检测结果。根据本申请实施例, 最终能够提供界面化操作轨迹搜索轨迹和图片搜索轨迹, 并能够显示最优轨迹和每个摄像头下的最优轨迹排名, 并能够通过轨迹中的图片帧号, 可以搜索到视频中该轨迹, 并进行实时播放。

本申请实施例提供的这种基于多模态的跨镜追踪检索系统, 可应用于如下两种场景: 行人轨迹搜索和行人图片搜索。利用数据库中的轨迹 ID, 行人特征、人脸特征, 行人属性和摄像头位置信息, 来进行轨迹和图片快速准确的检索, 利用不同特征之间的约束, 来达到精确匹配的效果。

轨迹匹配的任务目的是: 任意选中一个已抽取的轨迹, 根据多模态特征进行检索, 在同一个视频内以及视频间匹配出与其相关的所有轨迹。具体实施可包括以下步骤 S11-步骤 S15。

在步骤 S11 中, 确定检索区域, 并获取该区域监控的离线视频, 该区域可以是商场、写字楼、居民小区和社区等相对固定的场所, 并且该离线视频应该是某一时间段的, 至少是同一天的监控视频。将视频保存到本地, 并标记摄像头 ID、位置和起始时间, 在本实施例中, 选取三个不同角度的镜头, 摄像头 ID 为 C0、C1 和 C2。

在步骤 S12 中, 对各个摄像头下的离线视频进行行人检测、轨

迹追踪。相应的行人轨迹图片以轨迹 ID、视频帧号和对应时间、地点复合命名（例 0001_00025_202003210915_NJ），保存在以轨迹 ID 命名的子文件夹下。这里行人检测模型采用的是 SSD 算法来获取当前帧行人的位置框和边界框，采用匈牙利追踪算法获取行人轨迹。

5 在步骤 S13 中，对上步获取的行人轨迹，使用行人质量分析模型，这里采用的是人体骨骼关键点检测算法，通过骨骼关键点的个数来判别行人的完整性，对于有较多行人图片的轨迹，提取该轨迹中质量较好且时间较分散的 5 张图片，作为 top5 行人轨迹。

10 在步骤 S14 中，分别利用行人重识别网络、人脸识别网络和行人属性网络提取 top5 行人轨迹的行人特征、人脸特征（若未检测到人脸数据置为空）和行人属性特征，特征提取完成，把三种特征和（轨迹 ID、视频帧号、时间、地点）保存到信息检索数据库中。

15 在步骤 S15 中，在该步完成轨迹匹配和图匹配。在选取视频内匹配和视频间匹配优先级时，考虑到视频内的图像是同源数据，在进行匹配时更能保证准确性，所以优先处理视频内的轨迹匹配。同时在选用特征的作用优先级时，考虑到人脸特征是行人最为鲁棒的特征，优先进行人脸特征的比对，根据 S14 步骤中存储的结构性特征，并根据不同特征作用的先后、视频内轨迹匹配及视频间轨迹匹配的优先级，匹配工作的流程包括以下 1) -3)。

20 1) 首先进行视频内的轨迹匹配，首先使用目标轨迹的人脸特征，与其它含有人脸特征的轨迹进行批量化的特征比对，若能完成特征匹配，并在行人特征比对和行人属性的批量特征比对时有一定的相关性，则认为轨迹能够匹配成功。其次将匹配到的轨迹与目标轨迹结合作为第二行人轨迹，在剩余的轨迹内进行查询，查询算法为使用行人特征和行人属性特征的批量特征比对，在该过程中将使用重排序算法进行

25 轨迹匹配。该过程充分结合初步查询的稳定轨迹，使第二行人轨迹中包含更多姿态和角度的样本，使该过程的查询更加的稳定。至此完成了视频内的轨迹匹配。

30 2) 然后进行视频间的轨迹匹配，与视频内的轨迹匹配较类似，也要首先使用 query 中的人脸特征与时空约束中的临近节点中的行

人轨迹进行批量特征匹配。然后利用结合了初次匹配到的样本作为 query 在临近节点中进行再次查询。不同的是，考虑到跨镜过程中数据源的变化，该过程中特征比对的阈值将适当降低。

3) 以行人图片搜索行人图片，对传入的待查询图片要进行结构化的特征提取，通过进行行人检测、行人特征提取、行人属性识别及人脸检测、人脸特征提取完成其特征结构化。在进行以图搜图时，首先在可疑节点的视频轨迹内进行查询，查询到目标轨迹后，后续将使用轨迹匹配的算法进行进一步查询。

若在可疑节点没有查询到目标行人，则进一步扩大范围，在临近节点内进行全量查询，若能确定大致时间范围，则可疑进一步增加检索的效率和准确性。

至此，已完成以行人轨迹匹配搜索轨迹和以行人图片搜索行人图片的过程。

综上所述，本申请实施例提供的技术方案可包括以下步骤 11) - 步骤 16)。

在步骤 11) 中，在进行行人轨迹提取时，对行人图片自动标注轨迹 ID、时间信息和位置信息，在后续跨镜检索的时候可以利用轨迹 ID、图片帧号、时间和空间信息。同时，该系统还能够同时对多个镜头下的多个视频进行轨迹提取。

在步骤 12) 中，行人质量分析采用的是人体骨骼关键点检测技术，通过该技术选取几个关键点，通过检测到的关键点个数来判断行人的完整性，会输出一个完整性得分，通过该得分提取一些质量较差的行人图片（这样会去除一些遮挡较为严重的行人图片），然后在利用图片的时间信息选取轨迹中 5 张时间间隔较大的图片（因为相邻帧图片之间的行人姿态变化较小，多个不同姿态同一个行人的轨迹特征更具有判别性，同时五张图片可以减少轨迹匹配的计算量）。

在步骤 13) 中，本申请结合了行人多目标跟踪、关键点骨骼检测、行人重识别、行人属性结构化、人脸识别和摄像机的拓扑时空约束等多模态信息的融合来实现跨摄像头下的行人追踪检索方案。

在步骤 14) 中，利用多模态信息融合后的特征来更好的实现跨

镜头下行人轨迹搜索行人轨迹，行人图片搜索行人轨迹的目标，最终实现跨镜追踪。

在步骤 15) 中，在进行目标行人的跨镜追踪时，可能面临巨大的数据量，进行全量搜索查询几乎是不可能的。本系统将摄像机的时空拓扑关系与目标的外观表现模型匹配算法进行结合，使用摄像机网络拓扑的图结构分析行人移动和转移的规律，从而对行人的跨镜追踪进行时空约束。

在步骤 16) 中，采用一批标注好的实际场景的数据集来优化各个模态之间的权重参数来达到跨镜追踪检索的最优效果。

本申请实施例提供的行人追踪方法和设备，以及计算机可读存储介质，可以自动完成跨镜追踪检索，打破单一固定摄像头的视角局限，同时也避免了手动回放大量监控视频来搜寻检索目标，极大地提高了检索效率，并且提高了追踪范围。多模态信息的利用，该跨镜检索特征综合了多种模态信息，包括人脸、行人、属性和时空信息，形成了多模态特征的互补，综合后的特征更具有特征判别性，在进行跨镜追踪检索时具有更好的鲁棒性，同时提高了检索精度。该系统能够通过测试集自适应地调整多模态权重参数，这样在很大的程度上解决了摄像头的跨域问题，通过参数的调整能够更好的适应不同的监控场景。该系统具有较好的人机交互界面，能够界面化的配置摄像头位置信息，模态权重参数信息，能够通过按键操作，行人追踪，特征提取和特征入库，界面化操作轨迹检索轨迹，行人图片搜索轨迹，界面显示最优轨迹，不同摄像头下的轨迹搜索排名，并能够进行轨迹播放。数据库信息可视化，非常便于操作使用。

本领域普通技术人员可以理解，上文中所公开方法中的全部或某些步骤、系统、设备中的功能模块/单元可以被实施为软件、固件、硬件及其适当的组合。

在硬件实施方式中，在以上描述中提及的功能模块/单元之间的划分不一定对应于物理组件的划分；例如，一个物理组件可以具有多个功能，或者一个功能或步骤可以由若干物理组件合作执行。某些物理组件或所有物理组件可以被实施为由处理器，如中央处理器、数字

信号处理器或微处理器执行的软件，或者被实施为硬件，或者被实施为集成电路，如专用集成电路。这样的软件可以分布在计算机可读介质上，计算机可读介质可以包括计算机存储介质（或非暂时性介质）和通信介质（或暂时性介质）。如本领域普通技术人员公知的，术语计算机存储介质包括在用于存储信息（诸如计算机可读指令、数据结构、程序模块或其他数据）的任何方法或技术中实施的易失性和非易失性、可移除和不可移除介质。计算机存储介质包括但不限于随机存取存储器（Random Access Memory, RAM）、只读存储器（Read-Only Memory, ROM）、带电可擦可编程只读存储器（Electrically Erasable Programmable Read Only Memory, EEPROM）、闪存或其他存储器技术、CD-ROM、数字多功能盘（DVD）或其他光盘存储、磁盒、磁带、磁盘存储或其他磁存储装置、或者可以用于存储期望的信息并且可以被计算机访问的任何其他的介质。此外，本领域普通技术人员公知的是，通信介质通常包含计算机可读指令、数据结构、程序模块或者诸如载波或其他传输机制之类的调制数据信号中的其他数据，并且可包括任何信息递送介质。

以上参照附图说明了本申请的优选实施例，并非因此局限本申请的权利范围。本领域技术人员不脱离本申请的范围和实质内所作的任何修改、等同替换和改进，均应在本申请的权利范围之内。

权利要求

1、一种行人追踪方法，包括：

对预设监控摄像头采集的视频画面进行行人轨迹分析，生成行人轨迹图片集；

5 对所述行人轨迹图片集进行多模态特征提取，并形成行人多模态数据库；以及

将所述行人多模态数据库输入到训练好的多模态识别系统中，进行行人追踪，生成行人在所述预设监控摄像头中的移动轨迹。

10 2、根据权利要求 1 所述的方法，还包括：

接收目标行人轨迹，提取所述目标行人的多模态特征，并在所述行人多模态数据库中查找与所述目标行人的多模态特征匹配的第一行人轨迹；

15 将所述目标行人轨迹和所述第一行人轨迹合并生成第二行人轨迹，在所述行人多模态数据库中查询与所述第二行人轨迹匹配的行人轨迹；以及

根据与所述第二行人轨迹匹配的行人轨迹，生成所述目标行人在所述预设监控摄像头中的所述移动轨迹。

20 3、根据权利要求 1 所述的方法，还包括：从所述行人轨迹图片集选取质量参数在预设范围内的图像，并对所述选取的质量参数在预设范围内的图像进行特征提取。

25 4、根据权利要求 1 所述的方法，其中，所述训练好的多模态识别系统是根据训练集对所述多模态识别系统中各模态参数的影响因子进行调整得到的。

30 5、根据权利要求 1 所述的方法，其中，所述行人轨迹图片集中的图片名称包括：轨迹标识 ID、视频帧号、所述图片拍摄时间和地点信息。

6、根据权利要求 1 所述的方法，其中，生成所述行人在所述预设监控摄像头中的所述移动轨迹，包括：

根据所述监控摄像机分布拓扑的图结构进行所述行人的移动规律分析。

7、根据权利要求 1-6 中任一项所述的方法，其中，所述多模态特征包括以下一种或多种：行人特征、人脸特征和行人属性特征。

8、根据权利要求 7 所述的方法，其中，所述行人特征包括以下一种或多种：高矮胖瘦的体型特征和姿势特征；

所述人脸特征信息包括以下一种或多种：脸型特征、脸部表情特征和肤色特征；和/或

所述行人属性信息包括以下一种或多种：发型长短、发型颜色、服装款式、服装颜色以及携带物品。

9、一种行人追踪设备，包括：存储器、处理器、存储在所述存储器上并可在所述处理器上运行的程序，以及用于实现所述处理器和所述存储器之间的连接通信的数据总线；其中，所述程序被所述处理器执行时实现根据权利要求 1-8 中任一项所述的行人追踪方法。

10、一种计算机可读存储介质，其上存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现根据权利要求 1-8 中任一项所述的行人追踪方法。

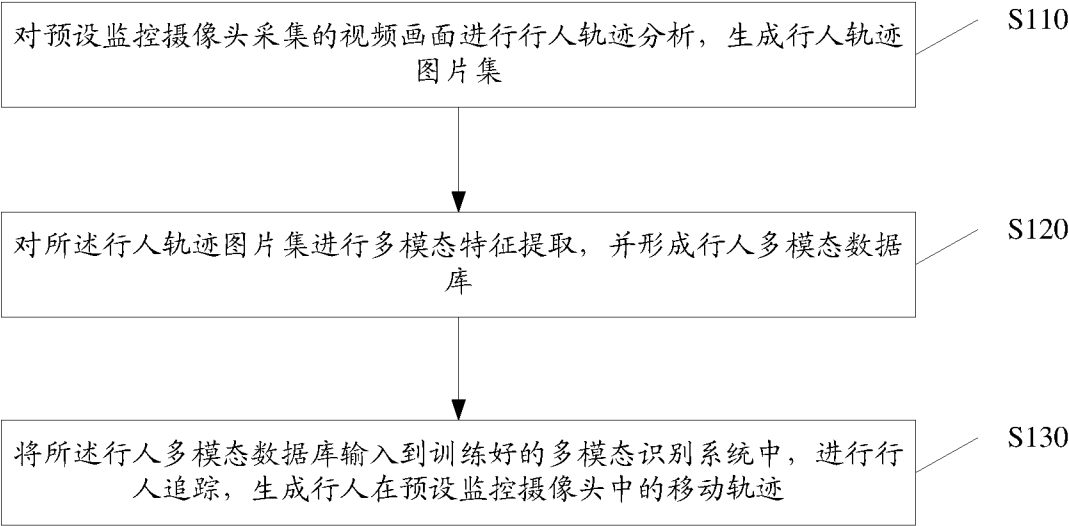


图 1

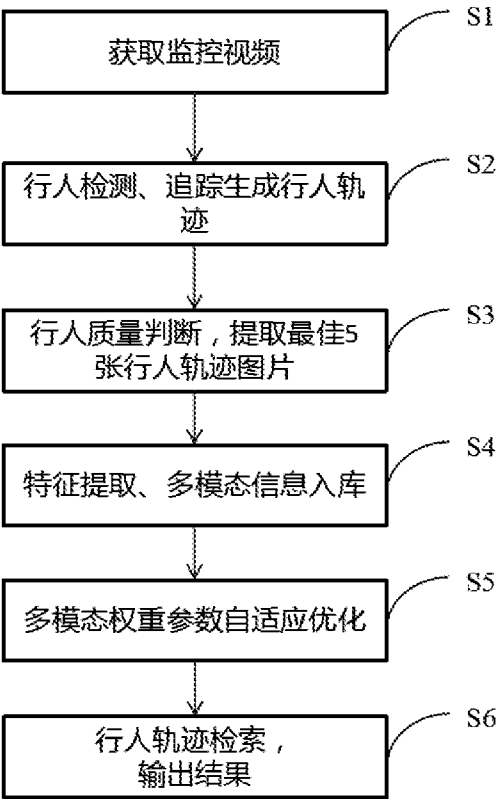


图 2

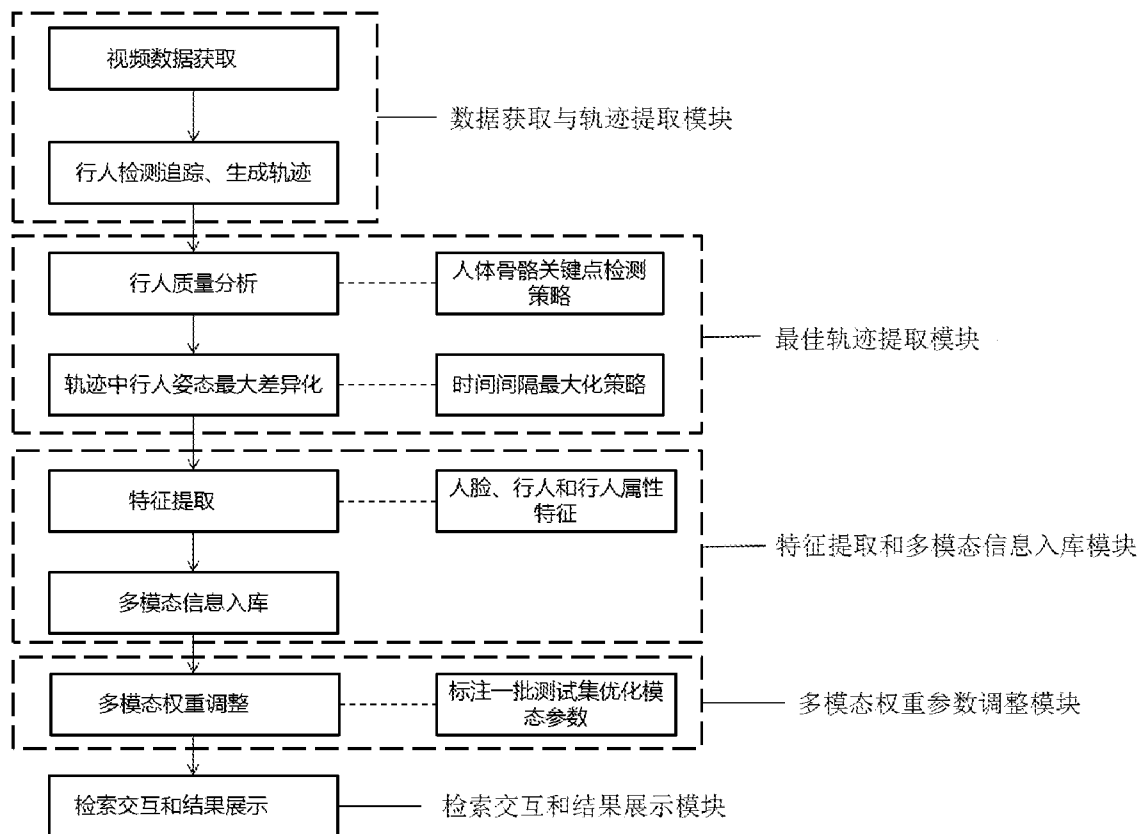


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/102652

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, IEEE, CNKI: 行人, 轨迹, 跟踪, 追踪, 采集, 拍摄, 视频, 图片, 多模态, 特征, 提取, 匹配, 合并, 监控, 移动, 摄像机, pedestrian, trajectory, tracking, acquisition, take, video, picture, multi-modal, feature, extract, match, merging, monitor, moving, camera

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 108629791 A (BEIJING KUANGSHI TECHNOLOGY CO., LTD. et al.) 09 October 2018 (2018-10-09) description, paragraphs [0005]-[0017] and [0045]-[0048]	1-10
Y	CN 110188691 A (WATRIX AI TECHNOLOGY (BEIJING) CO., LTD.) 30 August 2019 (2019-08-30) description, paragraphs [0004]-[0012] and [0057]-[0060]	1-10
A	CN 107657232 A (NANTONG UNIVERSITY) 02 February 2018 (2018-02-02) entire document	1-10
A	CN 108229456 A (SHENZHEN SENSETIME TECHNOLOGY CO., LTD.) 29 June 2018 (2018-06-29) entire document	1-10
A	US 10095954 B1 (VERINT SYSTEMS LTD.) 09 October 2018 (2018-10-09) entire document	1-10
A	US 2010266159 A1 (NEC SOFT, LTD.) 21 October 2010 (2010-10-21) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

09 September 2021

Date of mailing of the international search report

29 September 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/102652

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
CN	108629791	A	09 October 2018	None		
CN	110188691	A	30 August 2019	None		
CN	107657232	A	02 February 2018	None		
CN	108229456	A	29 June 2018	None		
US	10095954	B1	09 October 2018	None		
US	2010266159	A1	21 October 2010	JP	2010257441	A 11 November 2010

国际检索报告

国际申请号

PCT/CN2021/102652

A. 主题的分类 G06K 9/00 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																							
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNPAT, WPI, EPDOC, IEEE, CNKI: 行人, 轨迹, 跟踪, 追踪, 采集, 拍摄, 视频, 图片, 多模态, 特征, 提取, 匹配, 合并, 监控, 移动, 摄像机, pedestrian, trajectory, tracking, acquisition, take, video, picture, multi-modal, feature, extract, match, merging, monitor, moving, camera																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 108629791 A (北京旷视科技有限公司 等) 2018年 10月 9日 (2018 - 10 - 09) 说明书第[0005]-[0017], [0045]-[0048]段</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 110188691 A (银河水滴科技北京有限公司) 2019年 8月 30日 (2019 - 08 - 30) 说明书第[0004]-[0012], [0057]-[0060]段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 107657232 A (南通大学) 2018年 2月 2日 (2018 - 02 - 02) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 108229456 A (深圳市商汤科技有限公司) 2018年 6月 29日 (2018 - 06 - 29) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 10095954 B1 (VERINT SYSTEMS LTD.) 2018年 10月 9日 (2018 - 10 - 09) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2010266159 A1 (NEC SOFT, LTD.) 2010年 10月 21日 (2010 - 10 - 21) 全文</td> <td>1-10</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 108629791 A (北京旷视科技有限公司 等) 2018年 10月 9日 (2018 - 10 - 09) 说明书第[0005]-[0017], [0045]-[0048]段	1-10	Y	CN 110188691 A (银河水滴科技北京有限公司) 2019年 8月 30日 (2019 - 08 - 30) 说明书第[0004]-[0012], [0057]-[0060]段	1-10	A	CN 107657232 A (南通大学) 2018年 2月 2日 (2018 - 02 - 02) 全文	1-10	A	CN 108229456 A (深圳市商汤科技有限公司) 2018年 6月 29日 (2018 - 06 - 29) 全文	1-10	A	US 10095954 B1 (VERINT SYSTEMS LTD.) 2018年 10月 9日 (2018 - 10 - 09) 全文	1-10	A	US 2010266159 A1 (NEC SOFT, LTD.) 2010年 10月 21日 (2010 - 10 - 21) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
Y	CN 108629791 A (北京旷视科技有限公司 等) 2018年 10月 9日 (2018 - 10 - 09) 说明书第[0005]-[0017], [0045]-[0048]段	1-10																					
Y	CN 110188691 A (银河水滴科技北京有限公司) 2019年 8月 30日 (2019 - 08 - 30) 说明书第[0004]-[0012], [0057]-[0060]段	1-10																					
A	CN 107657232 A (南通大学) 2018年 2月 2日 (2018 - 02 - 02) 全文	1-10																					
A	CN 108229456 A (深圳市商汤科技有限公司) 2018年 6月 29日 (2018 - 06 - 29) 全文	1-10																					
A	US 10095954 B1 (VERINT SYSTEMS LTD.) 2018年 10月 9日 (2018 - 10 - 09) 全文	1-10																					
A	US 2010266159 A1 (NEC SOFT, LTD.) 2010年 10月 21日 (2010 - 10 - 21) 全文	1-10																					
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																			
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																						
国际检索实际完成的日期 2021年 9月 9日		国际检索报告邮寄日期 2021年 9月 29日																					
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国 北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 石爽 电话号码 86-(10)-53961413																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/102652

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	108629791	A	2018年 10月 9日	无	
CN	110188691	A	2019年 8月 30日	无	
CN	107657232	A	2018年 2月 2日	无	
CN	108229456	A	2018年 6月 29日	无	
US	10095954	B1	2018年 10月 9日	无	
US	2010266159	A1	2010年 10月 21日	JP 2010257441 A	2010年 11月 11日