



(12) 发明专利

(10) 授权公告号 CN 114022901 B

(45) 授权公告日 2024. 07. 16

(21) 申请号 202111298178.5

G06F 18/24 (2023.01)

(22) 申请日 2021.11.04

G06F 18/2321 (2023.01)

G01S 13/88 (2006.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 114022901 A

(56) 对比文件

CN 113349761 A, 2021.09.07

CN 112505684 A, 2021.03.16

(43) 申请公布日 2022.02.08

(73) 专利权人 东南大学

审查员 武茹茹

地址 210000 江苏省南京市麒麟科创园智

识路26号启迪城立业园04幢

(72) 发明人 王帅 曹东江 梅洛瑜 裴思远

(74) 专利代理机构 南京众联专利代理有限公司

32206

专利代理师 薛雨妍

(51) Int. Cl.

G06V 40/10 (2022.01)

G06V 10/762 (2022.01)

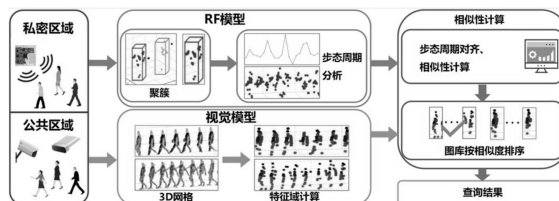
权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种基于视觉和射频感知的跨模态ReID方法

(57) 摘要

本发明公开了一种基于视觉和射频感知的跨模态ReID方法,对数据进行处理来消除时空上的一致性对相似性计算的影响,具体包括以下步骤:步骤1:对视觉数据,建立3D网格,然后计算特征域;步骤2:对射频数据,首先进行聚类,然后进行步态周期的分析;步骤3:将处理好的视觉和射频数据进行步态周期的对齐;步骤4:评估相似性,评估结果按相似性降序排列。本发明解决了视觉和射频感知在数据形式,时间、空间上的一致性,达到了90%以上的总体准确率。



1. 一种基于视觉和射频感知的跨模态ReID方法,其特征在于:
包括以下步骤:

步骤1:对视觉数据,建立3D网格,然后计算特征域;

步骤2:对射频数据,首先进行聚类,然后进行步态周期的分析;

步骤3:将处理好的视觉和射频数据进行步态周期的对齐;

步骤4:评估相似性,评估结果按相似性降序排列;所述步骤4还包括以下具体步骤,

步骤4.1. 姿态相似性计算

动态位置的相似性使用雷达点集和特征域点集的IoU来描述;其公式为:

$$IoU(t) = \frac{|P_m(t) \cap P_s(t)|}{|P_m(t) \cup P_s(t)|}$$

其中 $P_m(t)$ 为t时刻雷达点集, $P_s(t)$ 为t时刻特征域点集,相交定义为两个点之间的距离小于毫米波雷达的精度;

进一步地,动态位置相似度可以用雷达点集和特征域点集相交部分的平均距离来给出更精确的度量,平均距离越小,相似度越高;其公式为:

$$Dis(t) = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M d(m_i(t), s_j(t)),$$

$$m_i \in P_m(t), s_j \in P_s(t)$$

其中 $d(m_i(t), s_j(t))$ 表示雷达点 m_i 和特征域点 s_j 之间的欧氏距离, N 为雷达点的数量, M 为特征域点的数量;如果两个点不是相交的,则他们之间距离定义为0;

步骤4.2. 信号强度相似性计算;信号强度相似性用信号强度的差值的倒数和表示;其公式为:

$$Intensity(t) = \sum_{i=1}^N \frac{1}{|I_m(m_i(t)) - I_s(m_i(t))|}$$

其中 $I_m(m_i(t))$ 表示雷达点 $m_i(t)$ 的信号强度, $I_s(m_i(t))$ 表示通过特征域估计的在雷达点 $m_i(t)$ 附近的信号强度;估计方式为:计算雷达点 $m_i(t)$ 附近特征域占比;

步骤4.3. 基于人体部位权重的相似性优化;身体各部位在人体重识别中表现不同;不同的人在行走时的躯干姿态差异较小,而四肢的活动幅度差异较大;通过单独计算各身体部位的动态位置相似性和信号强度相似性,然后根据各身体部位在重识别中的显著性给予对应权重,达到优化相似性计算的目的;

步骤4.4. 综合相似性计算;

综合相似性计算公式为:

$$similarity(P_m, P_s) = \begin{bmatrix} w_{iou} \\ w_{dis} \\ w_{int} \end{bmatrix}^T \begin{bmatrix} IoU_1(t) & \dots & IoU_K(t) \\ Dis_1(t) & \dots & Dis_K(t) \\ Int_1(t) & \dots & Int_K(t) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_K \end{bmatrix}$$

其中 w_{iou} , w_{dis} 和 w_{int} 是上述三种相似度度量标准的权重,中间的矩阵的每一列代表一个

身体部位在上述三种相似度量标准下的得分, w_1, w_2, \dots, w_k 代表个身体部位的权重。

2. 根据权利要求1所述的一种基于视觉和射频感知的跨模态ReID方法, 其特征在于: 所述步骤1: 采取使用视觉感知数据模拟射频感知数据的方式实现数据形态的一致化; 首先根据人对射频信号的反射特性提出了“镜面模型”: 由于人体表面的粗糙程度远小于信号波长, 而且人体皮肤含水量高, 人体对毫米波的反射呈现出镜面反射的效果; 其次, 由于毫米波雷达天线尺寸较小, 只有正对天线反射的信号会被天线接收到; 根据以上两条性质, 可以利用RGB-D相机获得的3D网格模型计算会被毫米波雷达感知到的区域, 称为“特征域”; 具体来说, 对每一个3D网格, 计算其法线向量, 计算其与连线的向量, 计算上述两个向量的夹角, 对夹角足够小的网格, 将其归为特征域; t 时刻的特征域 $P_s(t)$ 的严格定义为:

$$P_s(t) = \{s(t) \in P(t) \mid \arccos \frac{(s(t) - x_m)n_s(t)}{|s(t) - x_m||n_s(t)|} < \epsilon\}$$

其中 $P(t)$ 为 t 时刻3D网格集, $s(t)$ 为 $P(t)$ 的元素, x_m 为雷达坐标, $n_s(t)$ 为 $s(t)$ 的法向量。

3. 根据权利要求1所述的一种基于视觉和射频感知的跨模态ReID方法, 其特征在于: 所述步骤2中使用DBSCAN算法对射频数据进行聚类;

由于毫米波雷达产生的数据在 x 轴和 y 轴形成的水平面上的投影较为清晰而在 z 轴代表的纵向空间上较为模糊, 在使用DBSCAN算法时, 对使用的欧氏距离进行了修改; 通过权重调整, 弱化了 z 轴数据对聚类的影响; 修改后的距离定义为:

$$D(p^i, p^j) = (p_x^i - p_x^j)^2 + (p_y^i - p_y^j)^2 + \alpha(p_z^i - p_z^j)^2$$

其中 p^i, p^j 为两个不同的点, α 为弱化 z 轴影响的系数。

4. 根据权利要求1所述的一种基于视觉和射频感知的跨模态ReID方法, 其特征在于: 所述步骤3: RGB-D相机和毫米波雷达捕捉到的人体行走姿态在时间和空间上是不一致的; 为了提高相似性计算的有效性, 需要消除时空上的不一致性; 我们观察到人的行走是一个周期性的行为, 两脚间距离随时间呈周期性变化; 本发明以两脚间距离最大的一帧为关键帧, 对RGB-D相机和毫米波雷达捕捉到的行走记录进行统一的划分, 两个关键帧之间为一个步态周期; 通过将关键帧进行对齐的方式, 实现时间上的一致性; 为实现空间上的一致性, 关键帧对齐时的匹配原则为: 到对应传感器的距离接近的关键帧相对齐。

一种基于视觉和射频感知的跨模态ReID方法

技术领域

[0001] 本发明涉及一种基于步态识别的跨模态身份识别技术,属于计算机视觉领域,尤其涉及一种基于视觉和射频感知的跨模态ReID方法。

背景技术

[0002] 基于相机图像的人物重识别问题作为一个特殊的人物检索问题,被广泛研究。人物重识别的目标是:用户给出针对某一个人的查询,系统反馈此人是否在特定时间出现在同一台相机或另一台相机的画面中。查询的人物一般以图片的形式给出,相机录制的图片和视频被称为候选图库。

[0003] 步态携带了可用于区分人物身份的独一无二的信息,因为每个人行走时的姿态都有其独特性。基于步态的识别系统从数据序列中提取行走的特征(步长、摆臂幅度、步频等)并将其用于身份识别。由于行走是一个具有周期性的动作,一个步态周期被定义为两个相同步态之间的时期。行走特征的提取需要作用于完整的步态周期。

[0004] 但不是所有场景下都可以获得图像数据,在一些涉及隐私的场景下,人们更倾向于使用射频技术代替相机进行监控。在现实世界中,人们会在在公众场合与私密场合之间来回走动,因此需要一种技术方案,实现跨视觉与射频数据的人物重识别任务。

[0005] 视觉感知与射频感知在数据形态上有巨大差异:视觉数据可以详细的描绘出人的各个身体部位及其运动姿态,而射频数据是离散而稀疏的点云。因此很难对两种数据直接进行相似性计算。

[0006] 视觉感知与射频感知在时间和空间上是不同步的:射频感知由于其数据形态带来的良好私密性,在隐私敏感度较高的场景中被用于代替视觉感知,常见的场景是视觉感知被运用于公共区域,射频感知被运用于私密区域,人在公共区域与私密区域之间来回移动,就在不同时间和空间上留下了视觉感知数据或射频感知数据。因此需要对数据进行处理来消除时空上的不一致性对相似性计算的影响。

发明内容

[0007] 为解决上述问题,本发明公开了一种基于视觉和射频感知的跨模态ReID方法,对数据进行处理来消除时空上的不一致性对相似性计算的影响。

[0008] 具体为:一种基于视觉和射频感知的跨模态ReID方法,包括以下步骤:

[0009] 步骤1:对视觉数据,建立3D网格,然后计算特征域;

[0010] 步骤2:对射频数据,首先进行聚类,然后进行步态周期的分析;

[0011] 步骤3:将处理好的视觉和射频数据进行步态周期的对齐;

[0012] 步骤4:评估相似性,评估结果按相似性降序排列。

[0013] 进一步优选的,步骤1通过“镜面模型”实现数据形态一致化

[0014] 视觉感知获得的数据比射频感知更为详细,为便于计算相似性,本发明采取使用视觉感知数据模拟射频感知数据的方式实现数据形态的一致化。首先根据人对射频信号

(毫米波)的反射特性提出了“镜面模型”:由于人体表面的粗糙程度远小于信号波长,而且人体皮肤含水量高,人体对毫米波的反射呈现出镜面反射的效果。其次,由于毫米波雷达天线尺寸较小,只有正对天线反射的信号会被天线接收到。根据以上两条性质,可以利用RGB-D相机获得的3D网格模型计算会被毫米波雷达感知到的区域,称为“特征域”。具体来说,对每一个3D网格,计算其法线向量,计算其与连线的向量,计算上述两个向量的夹角,对夹角足够小的网格,将其归为特征域。 t 时刻的特征域 $P_s(t)$ 的严格定义为:

$$[0015] \quad P_s(t) = \{s(t) \in P(t) \mid \arccos \frac{(s(t) - x_m)n_s(t)}{|s(t) - x_m||n_s(t)|} < \epsilon\}$$

[0016] 其中 $P(t)$ 为 t 时刻3D网格集, $s(t)$ 为 $P(t)$ 的元素, x_m 为雷达坐标, $n_s(t)$ 为 $s(t)$ 的法向量。

[0017] 进一步优选的,步骤2使用DBSCAN算法对射频数据进行聚类

[0018] 由于毫米波雷达产生的数据在 x 轴和 y 轴形成的水平面上的投影较为清晰而在 z 轴代表的纵向空间上较为模糊,在使用DBSCAN算法时,对使用的欧氏距离进行了修改。通过权重调整,弱化了 z 轴数据对聚类的影响。修改后的距离定义为:

$$[0019] \quad D(p^i, p^j) = (p_x^i - p_x^j)^2 + (p_y^i - p_y^j)^2 + \alpha(p_z^i - p_z^j)^2$$

[0020] 其中 p^i, p^j 为两个不同的点, α 为弱化 z 轴影响的系数。

[0021] 进一步优选的,步骤3通过“关键帧对齐”消除时空不一致性

[0022] RGB-D相机和毫米波雷达捕捉到的人体行走姿态在时间和空间上是不一致的。为了提高相似性计算的有效性,需要消除时空上的不一致性。我们观察到人的行走是一个周期性的行为,两脚间距离随时间呈周期性变化。本发明以两脚间距离最大的一帧为关键帧,对RGB-D相机和毫米波雷达捕捉到的行走记录进行统一的划分,两个关键帧之间为一个步态周期。通过将关键帧进行对齐的方式,实现时间上的一致性;为实现空间上的一致性,关键帧对齐时的匹配原则为:到对应传感器的距离接近的关键帧相对齐。

[0023] 进一步优选的,步骤4通过多种度量标准综合计算相似性

[0024] 步骤4.1.姿态相似性计算

[0025] 动态位置的相似性使用雷达点集和特征域点集的IoU来描述。其公式为:

$$[0026] \quad \text{IoU}(t) = \frac{|P_m(t) \cap P_s(t)|}{|P_m(t) \cup P_s(t)|}$$

[0027] 其中 $P_m(t)$ 为 t 时刻雷达点集, $P_s(t)$ 为 t 时刻特征域点集,相交定义为两个点之间的距离小于毫米波雷达的精度(0.047m)。

[0028] 进一步地,动态位置相似度可以用雷达点集和特征域点集相交部分的平均距离来给出更精确的度量,平均距离越小,相似度越高。其公式为:

$$[0029] \quad \text{Dis}(t) = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M d(m_i(t), s_j(t)),$$

[0030] $m_i \in P_m(t), s_j \in P_s(t)$

[0031] 其中 $d(m_i(t), s_j(t))$ 表示雷达点 m_i 和特征域点 s_j 之间的欧氏距离, N 为雷达点的数

量, M 为特征域点的数量。如果两个点不是相交的, 则他们之间距离定义为 0。

[0032] 步骤 4.2. 信号强度相似性计算

[0033] 信号强度相似性用信号强度的差值的倒数和表示。其公式为:

$$[0034] \quad \text{Intensity}(t) = \sum_{i=1}^N \frac{1}{|I_m(m_i(t)) - I_s(m_i(t))|}$$

[0035] 其中 $I_m(m_i(t))$ 表示雷达点 $m_i(t)$ 的信号强度, $I_s(m_i(t))$ 表示通过特征域估计的在雷达点 $m_i(t)$ 附近的信号强度。估计方式为: 计算雷达点 $m_i(t)$ 附近特征域占比。

[0036] 步骤 4.3. 基于人体部位权重的相似性优化

[0037] 身体各部位在人体重识别中表现不同。不同的人在行走时的躯干姿态差异较小, 而四肢的活动幅度差异较大。通过单独计算各身体部位的动态位置相似性和信号强度相似性, 然后根据各身体部位在重识别中的显著性给予对应权重, 达到优化相似性计算的目的。

[0038] 步骤 4.4. 综合相似性计算

[0039] 综合相似性计算公式为:

$$[0040] \quad \text{similarity}(P_m, P_s) = \begin{bmatrix} w_{iou} \\ w_{dis} \\ w_{int} \end{bmatrix}^T \begin{bmatrix} IoU_1(t) & \dots & IoU_K(t) \\ Dis_1(t) & \dots & Dis_K(t) \\ Int_1(t) & \dots & Int_K(t) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_K \end{bmatrix}$$

[0041] 其中 w_{iou} , w_{dis} 和 w_{int} 是上述三种相似度量标准的权重, 中间的矩阵的每一列 (如 $[IoU_1(t) \quad Dis_1(t) \quad Int_1(t)]^T$) 代表一个身体部位在上述三种相似度量标准下的得分, w_1, w_2, \dots, w_K 代表个身体部位的权重。

[0042] 本发明的有益效果:

[0043] 解决了视觉和射频感知在数据形式, 时间、空间上的不一致性, 达到了 90% 以上的总体准确率。

附图说明

[0044] 图 1、本发明的总体流程图;

[0045] 图 2、镜面模型示意图; (a) 说明只有法向量朝着雷达的区域其反射信号会被雷达收到; (b) 是镜面模型运用在人体上的示意图;

[0046] 图 3、RGB-D 相机和毫米波雷达观测到的两脚间距离的周期性变化;

[0047] 图 4、不同人行走时身体各部位的姿态重合情况比较; 躯干的重合率显著高于四肢, 这表明躯干不适用于身份识别。

具体实施方式

[0048] 下面结合附图和具体实施方式, 进一步阐明本发明, 应理解下述具体实施方式仅用于说明本发明而不适用于限制本发明的范围。需要说明的是, 下面描述中使用的词语“前”、“后”、“左”、“右”、“上”和“下”指的是附图中的方向, 词语“内”和“外”分别指的是朝向或远离特定部件几何中心的方向。

[0049] 如图 1-4 所示, 本实施例提供一种基于视觉和射频感知的跨模态 ReID 方法, 其特征

在于:包括以下步骤:

[0050] 步骤1:对视觉数据,建立3D网格,然后计算特征域;采取使用视觉感知数据模拟射频感知数据的方式实现数据形态的一致化;首先根据人对射频信号(毫米波)的反射特性提出了“镜面模型”:由于人体表面的粗糙程度远小于信号波长,而且人体皮肤含水量高,人体对毫米波的反射呈现出镜面反射的效果;其次,由于毫米波雷达天线尺寸较小,只有正对天线反射的信号会被天线接收到;根据以上两条性质,可以利用RGB-D相机获得的3D网格模型计算会被毫米波雷达感知到的区域,称为“特征域”;具体来说,对每一个3D网格,计算其法线向量,计算其与连线的向量,计算上述两个向量的夹角,对夹角足够小的网格,将其归为特征域。 t 时刻的特征域 $P_s(t)$ 的严格定义为:

$$[0051] \quad P_s(t) = \{s(t) \in P(t) \mid \arccos \frac{(s(t) - x_m)n_s(t)}{|s(t) - x_m||n_s(t)|} < \epsilon\}$$

[0052] 其中 $P(t)$ 为 t 时刻3D网格集, $s(t)$ 为 $P(t)$ 的元素, x_m 为雷达坐标, $n_s(t)$ 为 $s(t)$ 的法向量。

[0053] 步骤2:对射频数据,首先进行聚类,然后进行步态周期的分析;使用DBSCAN算法对射频数据进行聚类;

[0054] 由于毫米波雷达产生的数据在 x 轴和 y 轴形成的水平面上的投影较为清晰而在 z 轴代表的纵向空间上较为模糊,在使用DBSCAN算法时,对使用的欧氏距离进行了修改。通过权重调整,弱化了 z 轴数据对聚类的影响。修改后的距离定义为:

$$[0055] \quad D(p^i, p^j) = (p_x^i - p_x^j)^2 + (p_y^i - p_y^j)^2 + \alpha(p_z^i - p_z^j)^2$$

[0056] 其中 p^i, p^j 为两个不同的点, α 为弱化 z 轴影响的系数。

[0057] 步骤3:将处理好的视觉和射频数据进行步态周期的对齐;RGB-D相机和毫米波雷达捕捉到的人体行走姿态在时间和空间上是不一致的。为了提高相似性计算的有效性,需要消除时空上的不一致性。我们观察到人的行走是一个周期性的行为,两脚间距离随时间呈周期性变化。本发明以两脚间距离最大的一帧为关键帧,对RGB-D相机和毫米波雷达捕捉到的行走记录进行统一的划分,两个关键帧之间为一个步态周期;通过将关键帧进行对齐的方式,实现时间上的一致性;为实现空间上的一致性,关键帧对齐时的匹配原则为:到对应传感器的距离接近的关键帧相对齐。

[0058] 步骤4:评估相似性,评估结果按相似性降序排列。

[0059] 包括:步骤4.1.姿态相似性计算

[0060] 动态位置的相似性使用雷达点集和特征域点集的IoU来描述。其公式为:

$$[0061] \quad \text{IoU}(t) = \frac{|P_m(t) \cap P_s(t)|}{|P_m(t) \cup P_s(t)|}$$

[0062] 其中 $P_m(t)$ 为 t 时刻雷达点集, $P_s(t)$ 为 t 时刻特征域点集,相交定义为两个点之间的距离小于毫米波雷达的精度(0.047m);

[0063] 进一步地,动态位置相似度可以用雷达点集和特征域点集相交部分的平均距离来给出更精确的度量,平均距离越小,相似度越高。其公式为:

$$[0064] \quad \text{Dis}(t) = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M d(m_i(t), s_j(t)),$$

[0065] $m_i \in P_m(t), s_j \in P_s(t)$

[0066] 其中 $d(m_i(t), s_j(t))$ 表示雷达点 m_i 和特征域点 s_j 之间的欧氏距离, N 为雷达点的数量, M 为特征域点的数量。如果两个点不是相交的,则他们之间距离定义为0。

[0067] 步骤4.2. 信号强度相似性计算;信号强度相似性用信号强度的差值的倒数和表示;其公式为:

$$[0068] \quad \text{Intensity}(t) = \sum_{i=1}^N \frac{1}{|I_m(m_i(t)) - I_s(m_i(t))|}$$

[0069] 其中 $I_m(m_i(t))$ 表示雷达点 $m_i(t)$ 的信号强度, $I_s(m_i(t))$ 表示通过特征域估计的在雷达点 $m_i(t)$ 附近的信号强度;估计方式为:计算雷达点 $m_i(t)$ 附近特征域占比。

[0070] 步骤4.3. 基于人体部位权重的相似性优化;身体各部位在人体重识别中表现不同。不同的人在行走时的躯干姿态差异较小,而四肢的活动幅度差异较大。通过单独计算各身体部位的动态位置相似性和信号强度相似性,然后根据各身体部位在重识别中的显著性给予对应权重,达到优化相似性计算的目的。

[0071] 步骤4.4. 综合相似性计算;

[0072] 综合相似性计算公式为:

$$[0073] \quad \text{similarity}(P_m, P_s) = \begin{bmatrix} w_{iou} \\ w_{dis} \\ w_{int} \end{bmatrix}^T \begin{bmatrix} IoU_1(t) & \dots & IoU_K(t) \\ Dis_1(t) & \dots & Dis_K(t) \\ Int_1(t) & \dots & Int_K(t) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_K \end{bmatrix}$$

[0074] 其中 w_{iou}, w_{dis} 和 w_{int} 是上述三种相似度度量标准的权重,中间的矩阵的每一列(如 $[IoU_1(t) \quad Dis_1(t) \quad Int_1(t)]^T$)代表一个身体部位在上述三种相似度度量标准下的得分, w_1, w_2, \dots, w_K 代表个身体部位的权重。

[0075] 实施例1:个性化定制服务

[0076] 通过匹配视觉和射频感知的数据,可以实现在射频传感器捕捉到人物时,通过匹配获得的视觉信息,获得更丰富的信息(如性别、年龄、外表),有助于在不侵犯用户隐私的情况下提供更个性化的服务(如根据衣物调整空调温度)。

[0077] 实施例2:身份识别与追踪

[0078] 在公众场合,摄像头被用于对场合提供监控,但私密场合中摄像头的存在并不合适。通过本发明,只需在私密场合装上射频传感器,当人们在各个场合之间穿梭时,就可以实现无间断的轨迹追踪。

[0079] 实施例3:家庭监控

[0080] 近年来家庭摄像头的兴起给许多家庭带来了便利,利用家庭摄像头可以对呆在家中的老人、小孩和宠物的动向进行监控,在紧急情况发生时可以第一时间救助。但家庭摄像头存在被攻击、入侵的安全隐患,随之带来了隐私泄露的风险。本发明让可以家庭摄像机被射频设备(如毫米波雷达)代替,在继承家庭监控功能的同时规避了隐私泄露的风险

[0081] 实施例4:安防联动

[0082] 居民在家中安装的射频设备与附近街道的监控摄像机联动,当射频设备捕捉到私闯民宅的现象时,可以通过与附近街道的监控比对,帮助调查者快速锁定入侵者。

[0083] 本发明方案所公开的技术手段不仅限于上述实施方式所公开的技术手段,还包括由以上技术特征任意组合所组成的技术方案。

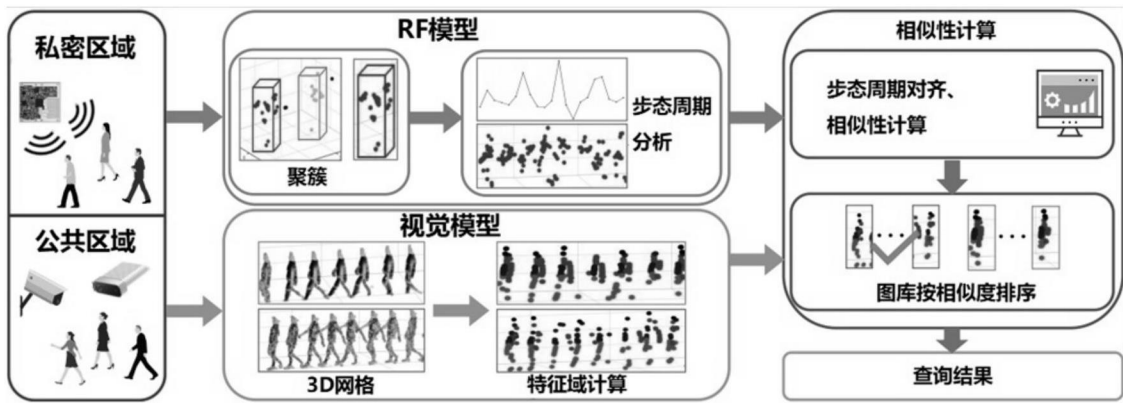


图1

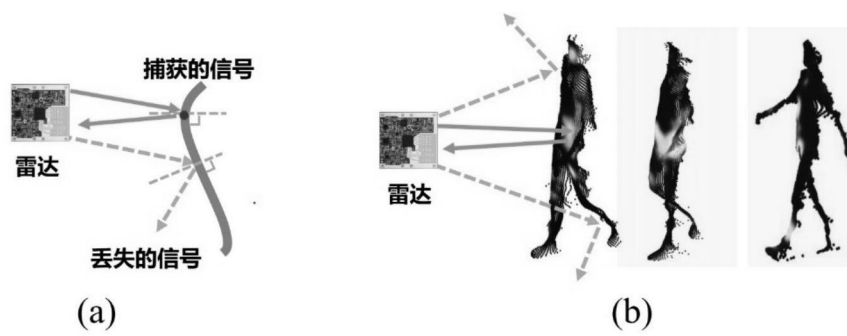


图2

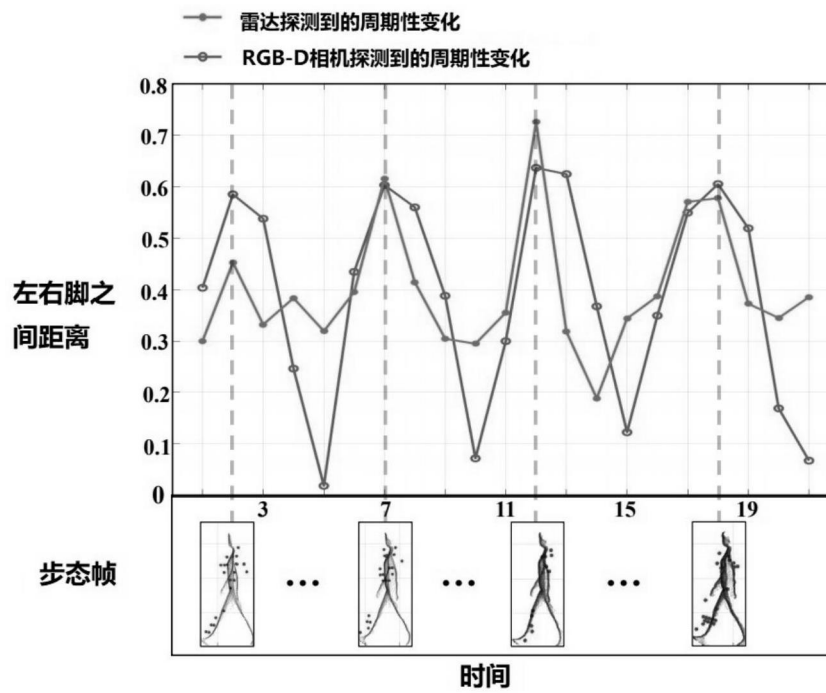


图3

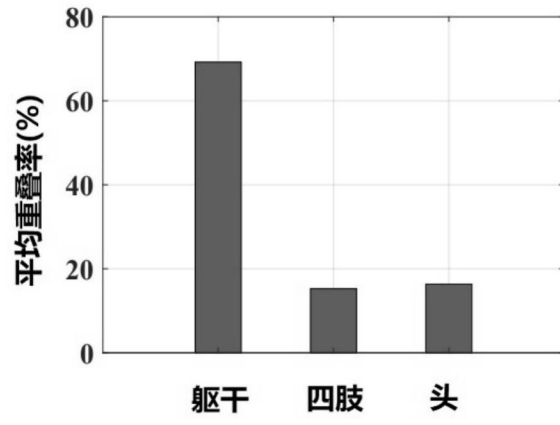


图4