



(12)发明专利

(10)授权公告号 CN 110008766 B

(45)授权公告日 2020.04.24

(21)申请号 201910231661.8

G06K 9/62(2006.01)

(22)申请日 2019.03.26

(56)对比文件

(65)同一申请的已公布的文献号

CN 108871332 A,2018.11.23,

申请公布号 CN 110008766 A

CN 102542289 A,2012.07.04,

(43)申请公布日 2019.07.12

CN 108764399 A,2018.11.06,

US 2018/0189529 A1,2018.07.05,

(73)专利权人 南京大学

审查员 杨战鹏

地址 210023 江苏省南京市栖霞区仙林大

道163号南京大学

(72)发明人 刘佳 陈力军 郁裕杰 陈星宇

张晓聪 李珍珠 严颖丽 边星宇

许林 龚华林 程唯

(74)专利代理机构 江苏圣典律师事务所 32237

代理人 胡建华 于瀚文

(51)Int.Cl.

G06K 7/00(2006.01)

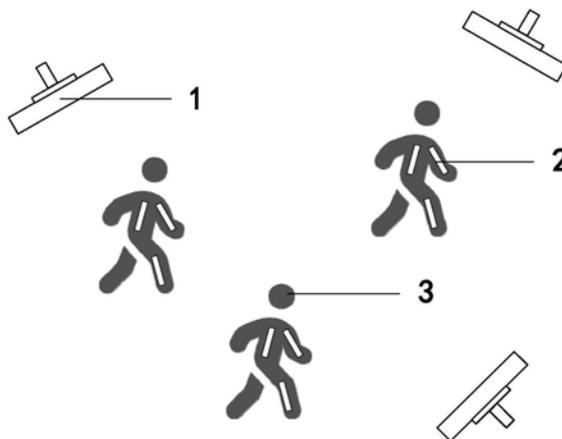
权利要求书2页 说明书7页 附图2页

(54)发明名称

一种基于射频识别技术的室内人数统计方法

(57)摘要

本发明公开了一种基于射频识别技术的室内人数统计方法,在室内部署若干RFID阅读器以保证读取范围恰当得覆盖室内区域,然后身上携带一定量RFID标签的人在该场景中自由活动。每一个标签会在一定时间段内被若干天线读取到,从时间维度出发,可以得到一个读取天线的集合序列。对于同一个人身上标签,它们对应时间的读取天线集合序列越相近,根据这一特性并利用一种基于密度和距离的聚类算法分析可以统计场景中的人数情况。



1. 一种基于射频识别技术的室内人数统计方法,其特征在于,包括如下步骤:

步骤1,在室内环境中部署 t_1 个RFID阅读器和天线;

步骤2,携带RFID标签的人在室内环境中任意活动,设定所有人身上RFID标签的总数为 n ,使用RFID阅读器读取分散在每个人身上的RFID标签信号;

步骤3,对步骤2中采集的RFID标签信号进行数据预处理;

步骤4,对预处理后的数据进行聚类分析,获取到每一个RFID标签的特征参数;

步骤5,根据步骤4获取到的特征参数利用数字异常值方法获取聚类中心和非聚类中心的分界阈值,从而找出聚类中心,通过聚类中心的个数得到人数。

2. 根据权利要求1所述的方法,其特征在于,步骤1中,布置 t_1 个RFID阅读器和天线以保证读取范围能够覆盖室内区域。

3. 根据权利要求2所述的方法,其特征在于,步骤2中,连接在RFID阅读器上的天线读取 n 个RFID标签,从而每一个RFID阅读器能够获取读取范围内 n 个RFID标签的一系列三元信号组 $\{epc, t, a\}$,包括标签的EPC号 epc 、读取到标签的时间 t 、天线编号 a ,所述EPC号即电子产品码。

4. 根据权利要求3所述的方法,其特征在于,步骤3包括:将步骤2中每一个RFID阅读器获取到的一系列三元信号组按照标签的EPC号进行分组,即相同的EPC号分到一组中,对于每一个RFID标签的三元信号组,以时间 t_0 为单位将整个读取时间 T 进行切片,然后将每个时间片读取到该RFID标签的天线编号组成集合 s_i ,从而得到每一个RFID标签的特征向量 $v = [s_1, s_2, s_3, \dots, s_n]$, $n = \frac{T}{t_0}$,第 n 个RFID标签的特征向量记为 v_n ,最后,将空集合占比超过比例 r 的特征向量过滤掉。

5. 根据权利要求4所述的方法,其特征在于,步骤4包括:针对每两个特征向量 v_1 和 v_2 ,使用豪斯多夫距离定义特征向量每一位之间的距离,然后采用欧式距离来计算每两个特征向量之间的距离 d_{ij} ,利用 d_{ij} ,采用一种基于密度和距离的聚类算法分析,得到每一个特征向量的局部密度 ρ 和特征距离 δ 。

6. 根据权利要求5所述的方法,其特征在于,步骤4中,所述局部密度采用高斯内核计算方法:

$$\rho_i = \sum_{j \in I_s \setminus \{i\}} e^{-\left(\frac{d_{ij}}{d_c}\right)^2},$$

其中 $S = \{x_i\}_{i=1}^N$ 为待聚类的数据集, $I_s = \{1, 2, \dots, N\}$ 为待聚类的数据集对应的下标集, N 表示所有的标签个数, S 中的数据即对应步骤3中的每一个RFID标签的特征向量, x_i 为第 i 个标签对应的特征向量, e 是自然常数, ρ_i 是根据高斯内核计算出来的第 i 个标签的局部密度,参数 d_c 为截断距离, $d_c > 0$ 。

7. 根据权利要求6所述的方法,其特征在于,步骤4中,所述特征距离计算方式如下:

设 $\{q_i\}_{i=1}^N$ 表示局部密度 $\{\rho_i\}_{i=1}^N$ 的一个降序排列下标序,即它满足:

$$\rho_{q_1} \geq \rho_{q_2} \geq \dots \geq \rho_{q_N},$$

在 ρ_i 进行降序排列的情况下, q_N 为第 N 个局部密度 ρ_{q_N} 的下标;

进行如下定义：

$$\delta_{q_i} = \begin{cases} \min_{\substack{q_j \\ j < i}} \{d_{q_i q_j}\}, & i \geq 2; \\ \max_{j \geq 2} \{\delta_{q_j}\}, & i = 1. \end{cases}$$

δ_{q_i} 表示第 q_i 个标签的特征距离， δ_{q_j} 表示第 q_j 个标签的特征距离， $d_{q_i q_j}$ 表示第 q_i 个标签和第 q_j 个标签之间的距离，最终每一个RFID标签的特征向量都得到了一个二元组 $\{\rho_i, \delta_i\}$ ； δ_i 表示第 i 个标签的特征距离。

8. 根据权利要求7所述的方法，其特征在于，步骤5包括：计算一个将局部密度和特征距离综合考虑的量 $\gamma_i = \rho_i * \delta_i^2$ ，则 γ_i 越大，则表示第 i 个特征向量越可能是聚类中心，将每一个特征向量的 γ_i 值按升序排列，用数字异常值方法从小到大分析每一个 γ_i 值，直到找到第一个异常值 γ_0 ，以 γ_0 为阈值，取 $\gamma_i > \gamma_0$ 的特征向量为聚类中心，聚类中心的个数即为室内环境中的人数。

一种基于射频识别技术的室内人数统计方法

技术领域

[0001] 本发明涉及一种室内人数统计方法,特别是一种基于射频识别技术的室内人数统计方法。

背景技术

[0002] RFID(Radio Frequency Identification,射频识别)技术是基于无线通信技术的自动识别技术,它的基本原理是利用射频信号与空间耦合传输特性,自动化地识别被识别物体携带的信息,它最大的优点是非接触识别,并且可以识别多目标。RFID技术目前被越来越广泛地应用于各行各业,包括供应链管理、仓储盘点、电子支付、安全访问控制、目标监测与追踪等。

[0003] 人数统计技术已经逐步应用于人流量较大的多种公共场所,例如超市、学校、地铁和汽车站等。协助安全人员在遇到紧急情况时可以更加及时的处理,保证公民的人身和财产安全。同时还可以帮助管理人员及时的处理公共设施的调配问题,使得公共资源能够合理分配,社会生活有序快速进行。所以这些场景下的人流量数据能够为公共安全、市场决策和资源分配等领域提供有用的信息。

[0004] 目前已有的室内人数统计方法以视频监控系统为主,对环境的光照条件比较敏感,在监控场景环境复杂、目标严重遮挡的情况下检测效果差强人意。

发明内容

[0005] 发明目的:本发明所要解决的技术问题是针对现有技术的不足,提供一种基于射频识别技术的室内人数统计方法,在人流量较大室内场馆中部署RFID阅读器和天线,持续读取在场景中自由活动的人员身上携带的若干标签,利用读取到的标签信号分析出每一个标签的特征向量,从而通过聚类统计出场景中的人员个数。随着RFID技术的普及,现在的服装、钥匙、卡片等生活用品都被嵌入了RFID标签,几乎每一个出门的时候身上会携带1个或以上的标签,这个现状为本发明提供了先决条件。

[0006] 为了解决上述技术问题,本发明公开了一种基于射频识别技术(Radio Frequency Identification,RFID)的室内人数统计方法,具体步骤如下:

[0007] 步骤1,在室内环境中部署 t_1 个RFID阅读器和天线;

[0008] 步骤2,携带 n 个(一般为3~5个)RFID标签的人在室内环境中任意活动,使用RFID阅读器读取分散在每个人身上的RFID标签信号;

[0009] 步骤3,对步骤2中采集的RFID标签信号进行数据预处理;

[0010] 步骤4,对预处理后的数据进行聚类分析,获取到每一个RFID标签的特征参数;

[0011] 步骤5,根据步骤4获取到的特征参数利用数字异常值方法获取聚类中心和非聚类中心的分界阈值,从而找出聚类中心,通过聚类中心的个数间接得到人数。

[0012] 步骤1中,合理地布置 t_1 个RFID阅读器和天线以保证读取范围能够覆盖室内区域,一般来说单个天线可以覆盖60平米左右的面积,也要根据场馆的实际情况做相应的调整,

比如600平米的场馆, t_1 可以取值为10。

[0013] 步骤2中, 连接在RFID阅读器上的天线读取每个人身上携带的 n 个RFID标签, 从而每一个RFID阅读器能够获取读取范围内 n 个RFID标签的一系列三元信号组 $f_{epc, t, a}$, 包括标签的EPC号 epc 、读取到标签的时间 t 、天线编号 a 。所述EPC号即电子产品码 (Electronic Product Code, EPC)

[0014] 步骤3包括: 将步骤2中每一个RFID阅读器获取到的一系列三元信号组按照标签的EPC号进行分组, 即相同的EPC号分到一组中 (这里按照标签进行分组, 所以分组的目的是为了后面对同一个标签的所有数据进行处理后得到反映它特点的一个特征向量 v)。对于每一个RFID标签的三元信号组, 以时间 t_0 (取值一般在0.5秒到1秒之间) 为单位将整个读取时间 T 进行切片, 然后将每个时间片读取到该RFID标签的天线编号组成集合 s_i , 从而得到每一个RFID标签的特征向量 $v = [s_1, s_2, s_3, \dots, s_n]$, $n = \frac{T}{t_0}$, (这里的 v 对应的是一个标签的数据, 它里面的每一个 s_i , 则对应这个标签在不同时间段被天线读取的情况。举个例子来说, 比如 $s_1 = \{1, 2\}$, 那就表示在第一个时间段内, 这个标签被1号天线和2号天线读取到了)。第 n 个RFID标签的特征向量记为 v_n , 最后, 将空集合占比超过比例 r (取值一般在65%到70%左右) 的特征向量过滤掉。

[0015] 步骤4包括: 针对每两个特征向量 v_1 和 v_2 , 使用豪斯多夫距离定义特征向量每一位之间的距离, 然后采用欧式距离来计算每两个特征向量之间的距离 d_{ij} , (设定 $v_1 = [s_{11}, s_{12}, s_{13}, \dots, s_{1n}]$, $v_2 = [s_{21}, s_{22}, s_{23}, \dots, s_{2n}]$, 豪斯多夫距离是用于计算 v_1 和 v_2 之间每一位的距离, 即计算 v_1 中的 s_{1n} 和 v_2 中的 s_{2n} 之间的距离, 最后再用欧式距离计算 v_1 和 v_2 的距离) 利用 d_{ij} , 采用一种基于密度和距离的聚类算法分析, 得到每一个特征向量的局部密度 ρ 和特征距离 δ 。

[0016] 步骤4中, 所述局部密度采用高斯内核Gaussian Kernel计算方法:

$$[0017] \quad \rho_i = \sum_{j \in I_s \setminus \{i\}} e^{-\left(\frac{d_{ij}}{d_c}\right)^2},$$

[0018] 其中 $S = \{x_i\}_{i=1}^N$ 为待聚类的数据集, $I_s = \{1, 2, \dots, N\}$ 为待聚类的数据集对应的下标集, N 表示所有的标签个数, S 中的数据即对应步骤3中的每一个RFID标签的特征向量, x_i 为第 i 个标签对应的特征向量, e 是自然常数, ρ_i 是根据高斯内核计算出来的第 i 个标签的局部密度, 参数 d_c 为截断距离, $d_c > 0$, 需由用户事先指定, 在本发明中一般选取一个 d_c 使得每个数据点的平均邻居个数为数据点总数的10%左右 (可以根据实际效果做相应的调整), 这里的邻居, 指的是与之距离不超过 d_c 意义下的邻居;

[0019] 步骤4中, 所述特征距离计算方式如下:

[0020] 设 $\{q_i\}_{i=1}^N$ 表示局部密度 $\{\rho_i\}_{i=1}^N$ 的一个降序排列下标序, 即它满足:

$$[0021] \quad \rho_{q_1} \geq \rho_{q_2} \geq \dots \geq \rho_{q_N},$$

[0022] 在 ρ_i 进行降序排列的情况下, q_N 为第 N 个局部密度 ρ_{q_N} 的下标;

[0023] 进行如下定义:

$$[0024] \quad \delta_{q_i} = \begin{cases} \min_{j < i} \{d_{q_i q_j}\}, & i \geq 2; \\ \max_{j \geq 2} \{\delta_{q_j}\}, & i = 1. \end{cases}$$

[0025] δ_{q_i} 表示第 q_i 个标签的特征距离 δ , δ_{q_j} 表示第 q_j 个标签的特征距离 δ , $d_{q_i q_j}$ 表示第 q_i 个标签和第 q_j 个标签之间的距离。最终每一个RFID标签的特征向量都得到了一个二元组 $\{\rho_i, \delta_i\}$, δ_i 表示第 i 个标签的特征距离。

[0026] 步骤5包括:计算一个将局部密度 ρ 值和特征距离 δ 值综合考虑的量 $\gamma_i = \rho_i * \delta_i^2$, 则 γ_i 越大, 则表示第 i 个特征向量越可能是聚类中心, 将每一个特征向量的 γ_i 值按升序排列, 用数字异常值方法从小到大分析每一个 γ_i 值, 直到找到第一个异常值 γ_0 , 以 γ_0 为阈值, 取 $\gamma_i > \gamma_0$ 的特征向量为聚类中心, 聚类中心的个数即为室内环境中的人数。

[0027] 本发明利用同一个人携带的多个标签在部署RFID阅读器的场景中的读取特征表现出一定的相似性, 通过聚类的方式统计出场景中的真实人数, 提供了一种安全、可靠的室内人数统计方法。与现有的室内人数统计方法相比, 本发明最大的特点就是利用RFID技术的非视距感知能力, 只需要在场景部署一定量的RFID阅读器, 伴随着现在大家身上携带了越来越多的标签, 本发明可以便捷高效地统计室内场景下的人数情况, 在更好地保护公民隐私地前提下, 为公共安全、市场决策和资源分配等领域提供有用的信息

[0028] 有益效果:通过本方法实现了室内人数统计的目的, 只需在室内需要监测人数的场景中部署一定量的RFID阅读器, 结合大家身上携带的若干标签, 就可以准确地进行室内地人数统计。相对于视频监控系统对于照明环境, 目标遮挡等问题的敏感, RFID的非视距识别能力带来了更好的统计效果, 同时它也能更好地保障公共环境下活动人员地隐私安全。

附图说明

[0029] 下面结合附图和具体实施方式对本发明做更进一步的具体说明, 本发明的上述或其他方面的优点将会变得更加清楚。

[0030] 图1是本发明的部署示意图。

[0031] 图2是实施例的流程图。

具体实施方式

[0032] 下面结合附图及实施例对本发明做进一步说明。

[0033] 本发明公开了一种基于射频识别技术(Radio Frequency Identification, RFID)的室内人数统计方法, 具体步骤如下:

[0034] 步骤1, 在室内环境中部署 t_1 个RFID阅读器和天线;

[0035] 步骤2, 携带 n 个(一般为3~5个)RFID标签的人在室内环境中任意活动, 使用RFID阅读器读取分散在每个人身上的RFID标签信号;

[0036] 步骤3, 对步骤2中采集的RFID标签信号进行数据预处理;

[0037] 步骤4, 对预处理后的数据进行聚类分析, 获取到每一个RFID标签的特征参数;

[0038] 步骤5, 根据步骤4获取到的特征参数利用数字异常值方法获取聚类中心和非聚类中心的分界阈值, 从而找出聚类中心, 通过聚类中心的个数间接得到人数。

[0039] 步骤1中,合理地布置 t_1 个RFID阅读器和天线以保证读取范围能够覆盖室内区域,一般来说单个天线可以覆盖60平米左右的面积,也要根据场馆的实际情况做相应的调整,比如600平米的场馆, t_1 可以取值为10。

[0040] 步骤2中,连接在RFID阅读器上的天线读取每个人身上携带的 n 个RFID标签,从而每一个RFID阅读器能够获取读取范围内 n 个RFID标签的一系列三元信号组 $\{epc, t, a\}$,包括标签的EPC号 epc 、读取到标签的时间 t 、天线编号 a 。所述EPC号即电子产品码(Electronic Product Code, EPC)

[0041] 步骤3包括:将步骤2中每一个RFID阅读器获取到的一系列三元信号组按照标签的EPC号进行分组,即相同的EPC号分到一组中(这里按照标签进行分组,所以分组的目的是为了后面对同一个标签的所有数据进行处理后得到反映它特点的一个特征向量 v)。对于每一个RFID标签的三元信号组,以时间 t_0 (取值一般在0.5秒到1秒之间)为单位将整个读取时间 T 进行切片,然后将每个时间片读取到该RFID标签的天线编号组成集合 s_i ,从而得到每一个RFID标签的特征向量 $v = [s_1, s_2, s_3, \dots, s_n]$, $n = \frac{T}{t_0}$, (这里的 v 对应的是一个标签的数据,它里面的每一个 s_i ,则对应这个标签在不同时间段被天线读取的情况。举个例子来说,比如 $s_1 = \{1, 2\}$,那就表示在第一个时间段内,这个标签被1号天线和2号天线读取到了)。第 n 个RFID标签的特征向量记为 v_n ,最后,将空集合占比超过比例 r (取值一般在65%到70%左右)的特征向量过滤掉。

[0042] 步骤4包括:针对每两个特征向量 v_1 和 v_2 ,使用豪斯多夫距离定义特征向量每一位之间的距离,然后采用欧式距离来计算每两个特征向量之间的距离 d_{ij} , (设定 $v_1 = [s_{11}, s_{12}, s_{13}, \dots, s_{1n}]$, $v_2 = [s_{21}, s_{22}, s_{23}, \dots, s_{2n}]$,豪斯多夫距离是用于计算 v_1 和 v_2 之间每一位的距离,即计算 v_1 中的 s_{1n} 和 v_2 中的 s_{2n} 之间的距离,最后再用欧式距离计算 v_1 和 v_2 的距离)利用 d_{ij} ,采用一种基于密度和距离的聚类算法分析,得到每一个特征向量的局部密度 ρ 和特征距离 δ 。

[0043] 步骤4中,所述局部密度采用高斯内核Gaussian Kernel计算方法:

$$[0044] \quad \rho_i = \sum_{j \in I_s \setminus \{i\}} e^{-\left(\frac{d_{ij}}{d_c}\right)^2},$$

[0045] 其中 $S = \{x_i\}_{i=1}^N$ 为待聚类的数据集, $I_s = \{1, 2, \dots, N\}$ 为待聚类的数据集对应的下标集, N 表示所有的标签个数, S 中的数据即对应步骤3中的每一个RFID标签的特征向量, x_i 为第 i 个标签对应的特征向量, e 是自然常数, ρ_i 是根据高斯内核计算出来的第 i 个标签的局部密度,参数 d_c 为截断距离, $d_c > 0$,需由用户事先指定,在本发明中一般选取一个 d_c 使得每个数据点的平均邻居个数为数据点总数的10%左右(可以根据实际效果做相应的调整),这里的邻居,指的是与之距离不超过 d_c 意义下的邻居;

[0046] 步骤4中,所述特征距离计算方式如下:

[0047] 设 $\{q_i\}_{i=1}^N$ 表示局部密度 $\{\rho_i\}_{i=1}^N$ 的一个降序排列下标序,即它满足:

$$[0048] \quad \rho_{q_1} \geq \rho_{q_2} \geq \dots \geq \rho_{q_N},$$

[0049] 在 ρ_i 进行降序排列的情况下, q_N 为第 N 个局部密度 ρ_{q_N} 的下标;

[0050] 进行如下定义:

$$[0051] \quad \delta_{q_i} = \begin{cases} \min_{j < i} \{d_{q_i q_j}\}, & i \geq 2; \\ \max_{j \geq 2} \{\delta_{q_j}\}, & i = 1. \end{cases}$$

[0052] δ_{q_i} 表示第 q_i 个标签的特征距离 δ , δ_{q_j} 表示第 q_j 个标签的特征距离 δ , $d_{q_i q_j}$ 表示第 q_i 个标签和第 q_j 个标签之间的距离。最终每一个RFID标签的特征向量都得到了一个二元组 $\{\rho_i, \delta_i\}$, δ_i 表示第 i 个标签的特征距离。

[0053] 步骤5包括:计算一个将局部密度 ρ 值和特征距离 δ 值综合考虑的量 $\gamma_i = \rho_i * \delta_i^2$, 则 γ_i 越大, 则表示第 i 个特征向量越可能是聚类中心, 将每一个特征向量的 γ_i 值按升序排列, 用数字异常值方法从小到大分析每一个 γ_i 值, 直到找到第一个异常值 γ_0 , 以 γ_0 为阈值, 取 $\gamma_i > \gamma_0$ 的特征向量为聚类中心, 聚类中心的个数即为室内环境中的人数。

[0054] 实施例

[0055] 聚类是把相似的对象通过静态分类的方法分成不同的组别或者更多的子集(subset), 这样让在同一个子集中的成员对象都有相似的一些属性。在本发明中, 同一个人身上的标签在场景中的活动的时候会出现相似的属性, 以时间 t_0 为单位将整个读取时间 $T = 10$ 分钟进行切片, 由于场景中有60个人, 每人携带3~5个标签, 所以环境中大约有200多个标签, 考虑到阅读器的读取效率选择 $t_0 = 0.5s$, 可以保证一个时间片内每一个标签可以被有效读取到。然后将每个时间片读取到该标签的天线编号组成集合 s_i , 从而得到标签的特征向量 $v = [s_1, s_2, s_3, \dots, s_n]$ ($n = \frac{T}{t_0} = 1200$)可作为标签的属性。

[0056] 为了将标签聚类, 需要定义两个标签特征向量之间的距离。豪斯多夫距离是描述两组点集之间相似程度的一种量度, 体现了两个点集之间的最大不匹配程度, 它是两个点集之间距离的一种定义形式: 设定有两组集合 $A = \{a_1, \dots, a_p\}$, $B = \{b_1, \dots, b_q\}$, a_p 表示集合 A 中第 p 个点, b_q 表示集合 B 中第 q 个点, 则这两个点集合之间的豪斯多夫距离 $H(A, B)$ 定义为:

$$[0057] \quad H(A, B) = \max(h(A, B), h(B, A))$$

$$[0058] \quad 1)$$

[0059] 其中 $h(A, B)$ 、 $h(B, A)$ 为狭义的豪斯多夫距离, 他们的计算方式如下, a 为集合 A 中的点, b 为集合 B 中的点:

$$[0060] \quad h(A, B) = \max_{a \in A} (\min_{b \in B} (\|a - b\|)) \quad 2)$$

$$[0061] \quad h(B, A) = \max_{b \in B} (\min_{a \in A} (\|b - a\|)) \quad 3)$$

[0062] 使用豪斯多夫距离得到标签特征向量每一位之间的距离后, 运用欧式距离便可计算得到任意两个特征向量 $v_1 = [x_1, x_2, x_3, \dots, x_n]$ 和 $v_2 = [y_1, y_2, y_3, \dots, y_n]$ 的距离 d_{ij} 。

$$[0063] \quad d_{ij} = \sqrt{\sum_{i=1}^n H(x_i, y_i)^2} \quad 4)$$

[0064] 然后采用一种基于密度和距离的聚类算法分析,得到每一个向量的局部密度 ρ_i 和特征距离 δ_i 。局部密度采用高斯内核Gaussian Kernel计算方法:

$$[0065] \quad \rho_i = \sum_{j \in I_s \setminus \{i\}} e^{-\left(\frac{d_{ij}}{d_c}\right)^2} \quad 5)$$

[0066] 其中 $S = \{x_i\}_{i=1}^N$ 为待聚类的数据集, $I_s = \{1, 2, \dots, N\}$ 为对应的下标集, S 中的数据即对应步骤3中的每一个RFID标签的特征向量, e 是自然常数, N 表示所有的标签个数, x_i 为第 i 个标签对应的特征向量, ρ_i 是根据高斯内核计算出来的第 i 个标签的局部密度,参数 d_c 为截断距离, $d_c > 0$,需由用户事先指定,在本例选取一个 d_c 使得每个数据点的平均邻居个数为数据点总数的10%左右(可以根据实际效果做相应的调整),这里的邻居,指的是与之距离不超过 d_c 意义下的邻居。

[0067] 距离计算方式如下:设 $\{q_i\}_{i=1}^N$ 表示局部密度 $\{\rho_i\}_{i=1}^N$ 的一个降序排列下标序,换句话说 ρ_{q_N} 即表示在 ρ_i 进行降序排列的情况下, q_N 为第 N 个局部密度的下标,即它满足:

$$[0068] \quad \rho_{q_1} \geq \rho_{q_2} \geq \dots \geq \rho_{q_N} \quad 6)$$

[0069] 则可定义:

$$[0070] \quad \delta_{q_i} = \begin{cases} \min_{\substack{q_j \\ j < i}} \{d_{q_i q_j}\}, & i \geq 2; \\ \max_{\substack{q_j \\ j \geq 2}} \{\delta_{q_j}\}, & i = 1. \end{cases} \quad 7)$$

[0071] 至此,每一个标签的特征向量都得到了一个二元组 $\{\rho_i, \delta_i\}$ 。 δ_{q_i} 即表示第 q_i 个标签的特征距离 δ , δ_{q_j} 表示第 q_j 个标签的特征距离 δ , $d_{q_i q_j}$ 就是指第 q_i 个标签和第 q_j 个标签之间的距离。计算一个将 ρ_i 值和 δ_i 值综合考虑的量 $\gamma_i = \rho_i * \delta_i^2$,显然 γ_i 越大,这个特征向量越可能是聚类中心。将每一个特征向量的 γ_i 值按升序排列,用数字异常值方法从小到大分析每一个 γ_i 值。

$$[0072] \quad \gamma_i > Q3 + k * IQR \cup \gamma_i < Q1 - k * IQR$$

$$[0073] \quad 8)$$

[0074] 其中 $Q1, Q3$ 分别是第一四分位数和第三四分位数,即按从小到大的顺序排序后,处于25%、75%的数字, $IQR = Q3 - Q1$, k 为四分位数乘数值,本实施例取 $k = 7$ 。直到找到第一个异常值 γ_0 ,以 γ_0 为阈值,取 $\gamma_i > \gamma_0$ 的向量为聚类中心,其个数即为该场景中的人数。

[0075] 如图1所示,本实施例包括天线1、标签2、自由活动的人3。合理地布置3个RFID阅读器和6个天线以保证读取范围恰当得覆盖室内区域,身上携带3~5个标签的人在场景中自由活动。

[0076] 本实施例提供了一种基于RFID技术的室内人数统计方法,流程如图2所示,过程如下:

[0077] 1、RFID设备读取标签信号

[0078] 本实施例中使用的RFID阅读器是Impinj R420,连接在RFID阅读器上的天线读取

一个大厅中的60个人,每人携带3~5个的RFID标签,从而每一个RFID阅读器可以获取读取范围内若干RFID标签的一系列三元信号组 $\{epc, t, a\}$,包括标签的EPC号epc、读取到标签的时间t、天线编号a。

[0079] 2、数据预处理

[0080] 将每一个RFID阅读器获取到一系列三元信号组按照标签的EPC号进行分组,即相同的EPC号分到一组中。对于每一个标签的三元信号组,以时间 t_0 为单位将整个读取时间 $T=10$ 分钟进行切片,由于场景中有60个人,每人携带3~5个标签,所以环境中大约有200多个标签,考虑到阅读器的读取效率选择 $t_0=0.5s$,可以保证一个时间片内每一个标签可以被有效读取到。然后分析每一段时间得到每个时间片读取到该标签的天线编号集合 s_i ,从而得到每一个标签的特征向量 $v = [s_1, s_2, s_3, \dots, s_n]$ ($n = \frac{T}{t_0} = 1200$)。最后,将空集合占比超过比例 $r=65\%$ 的特征向量过滤掉。

[0081] 3、聚类算法分析

[0082] 针对每两个特征向量 v_1 和 v_2 ,使用豪斯多夫距离定义向量之间每一位的距离,然后采用欧式距离来计算每两个向量之间的距离 d_{ij} 。利用这个距离,采用一种基于密度和距离的聚类算法分析,得到每一个向量的局部密度 ρ 和特征距离 δ ,其中截断距离 d_c 取所有距离中第20大的值。

[0083] 4、统计场景中的人数

[0084] 计算一个将 ρ 值和 δ 值综合考虑的量 $\gamma = \rho * \delta^2$,显然 γ 越大,这个向量越可能是聚类中心。将每一个向量的 γ 值按升序排列,用数字异常值方法从小到大分析每一个 γ 值,直到找到第一个异常值 γ_0 ,以 γ_0 为阈值,取 $\gamma > \gamma_0$ 的向量为聚类中心,其个数即为该场景中的人数。

[0085] 本发明提供了一种基于射频识别技术的室内人数统计方法,具体实现该技术方案的方法和途径很多,以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。本实施例中未明确的各组成部分均可用现有技术加以实现。

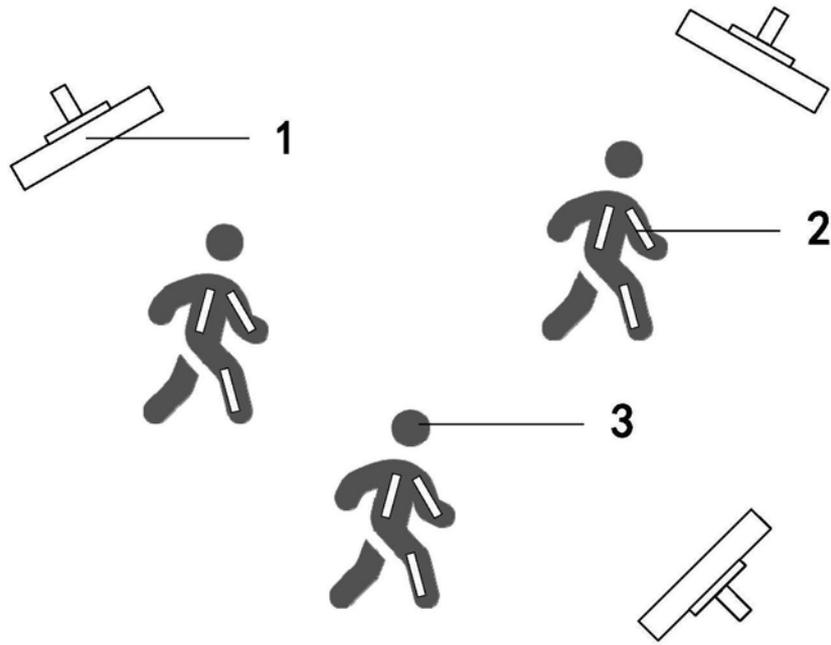


图1

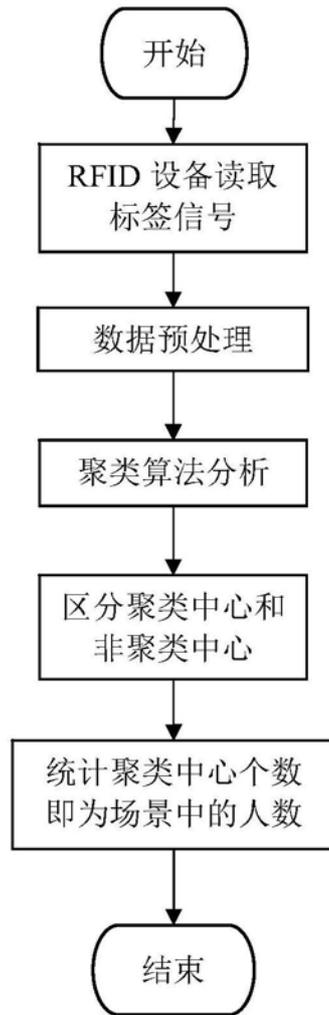


图2