



(12) 发明专利

(10) 授权公告号 CN 108734204 B

(45) 授权公告日 2021.03.16

(21) 申请号 201810401924.0

H04M 1/72454 (2021.01)

(22) 申请日 2018.04.28

H04M 1/72406 (2021.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 108734204 A

(56) 对比文件

CN 107403084 A, 2017.11.28

CN 107423730 A, 2017.12.01

(43) 申请公布日 2018.11.02

CN 106227851 A, 2016.12.14

(73) 专利权人 武汉大学
地址 430072 湖北省武汉市武昌区珞珈山
武汉大学

US 2008064334 A1, 2008.03.13

Yaacob N I 等. Feature selection for gait recognition.《Humanities, Science & Engineering Research》.2012,

(72) 发明人 邹勤

审查员 罗畅

(74) 专利代理机构 武汉科皓知识产权代理事务
所(特殊普通合伙) 42222

代理人 齐晨涵

(51) Int. Cl.

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

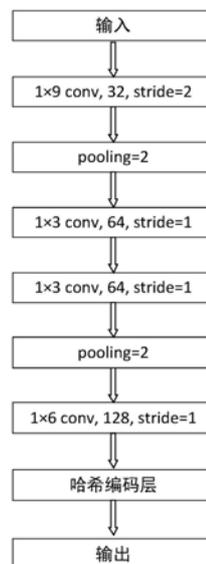
权利要求书1页 说明书5页 附图1页

(54) 发明名称

基于步态的大规模手机用户快速身份识别方法

(57) 摘要

本发明公开了一种基于步态的大规模手机用户快速身份识别方法,通过采集步态样本、构建步态检索样本集、对步态样本集进行哈希编码、对步态样本集进行特征描述、识别未知身份步态,实现步态的识别。该方法通过快速检索到一批相似用户,然后进行精细比较,缩短了识别时间,且利用所设计的哈希编码和特征描述可以实现快捷的同时,准确率也有提高。



1. 一种基于步态的大规模手机用户快速身份识别方法,其特征在于,包括以下步骤:

s1采集步态样本:利用手机采集大规模用户的运动数据,分离出步行时产生的数据并构建步态样本集;

s2构建步态检索样本集:基于步骤s1构建的大规模步态样本集,构建步态检索样本集,具体过程如下:

根据步态样本数据集计算相似度矩阵,规定同一用户的样本相似度为1,不同用户的样本相似度为0;构建三元组步态样本,即每个样本包含三个步态样本,其中两个样本相似度为1,另一个样本与它们不相似;将所有三元组步态样本构成步态检索样本集

s3对步态样本集进行哈希编码:构建深度学习哈希网络,并利用所述大规模步态检索样本集进行训练,得到哈希编码器,利用该哈希编码器对步态样本集进行哈希编码;

s4对步态样本集进行特征描述:构建深度学习分类网络,并利用所述大规模步态样本集对深度学习分类网络进行训练,得到步态特征描述器,并利用该特征描述器对所述步态样本集进行步态特征描述;

s5识别未知身份步态:利用所述的哈希编码器对待识别步态进行编码,并计算与已知身份的步态哈希编码的海明距离,得到最相似的K个用户;利用所述的特征描述器对待识别步态进行特征描述,并与所述的K个用户的步态特征描述进行比对,将最大相似度值对应的用户身份作为未知身份步态样本的身份。

2. 根据权利要求1基于步态的大规模手机用户快速身份识别方法,其特征在于:所述步骤s1中的运动数据包含但不限于步行,还包括跑、跳、乘车颠簸和静坐;所述运动数据采用手机APP程序采集并通过无线通讯传输到服务器。

3. 根据权利要求1基于步态的大规模手机用户快速身份识别方法,其特征在于:所述步骤s1中构建步态样本数据时,步态样本的采用按时间分割的方式。

4. 根据权利要求1基于步态的大规模手机用户快速身份识别方法,其特征在于:所述步骤s1中分离出步行时产生的数据具体为:根据步行的周期特点,采用加速度的幅度阈值、跨度阈值来对步态数据进行过滤。

5. 根据权利要求1基于步态的大规模手机用户快速身份识别方法,其特征在于:所述深度学习哈希网络采用一维卷积构建。

6. 根据权利要求5基于步态的大规模手机用户快速身份识别方法,其特征在于:所述深度学习哈希网络有8层,具体结构如下:第1层是输入层,即步态样本数据,第2层是卷积层,采用 1×9 的卷积核,卷积核个数为32,步长为2,第3层是池化层,池化尺寸为 1×2 ,第4、5层是两个卷积层,采用 1×3 的卷积核,卷积核个数为64,步长为1,第6层是池化层,池化尺寸为 1×2 ,第7层是卷积层,卷积核个数为128,步长为1,第8层是一个全连接层,称为哈希层,第9层是输出层。

7. 根据权利要求1基于步态的大规模手机用户快速身份识别方法,其特征在于:所述步骤s4中的深度学习分类网络采用双向双层长短时记忆递归神经网络作为分类网络。

基于步态的大规模手机用户快速身份识别方法

技术领域

[0001] 本发明涉及人工智能和信息安全领域,具体涉及一种基于步态的大规模手机用户快速身份识别方法。

技术背景

[0002] 近十年来,随着通信技术的发展,智能手机迅速发展并且在社会大众中迅速普及,手机已成为人们社会生活中最重要的工具之一。截至2017年12月,我国手机网民规模达7.53亿;以手机为中心的智能设备,已成为物联网的基础,它使构筑个性化、智能化的应用场景成为可能。然而,移动互联网服务场景不断丰富、移动终端规模加速提升、移动数据量持续扩大,也给移动互联网的安全管理带来了新的更高要求和严峻挑战。以智能手机为终端载体的移动互联网安全是国家网络安全的重要组成,手机用户的身份识别是构建移动互联网安全的重要基础。

[0003] 步态作为一种非接触的生物特征,是由人在行走时各个肢体器官协调动作的模式构成。与其他生物特征相比,步态支持动态采集、容易被用户接受、并且难以被模仿,使得它具有多种优势。而新一代智能手机普遍集成了高灵敏度的惯性传感器,如加速度传感器、陀螺仪等,为采集用户行走时的步态信息提供了广泛支持,从而为基于步态识别的手机用户身份识别提供了可能。

[0004] 但是,当用户规模巨大时,例如百万、千万甚至上亿的用户对象,传统的步态分类与识别方法速度很慢,难以满足实时手机用户识别的需求。因此,研发基于步态的大规模手机用户快速身份识别具有重要的意义和价值。

发明内容

[0005] 本发明针对现有技术的不足,提出了一种基于步态的大规模手机用户快速身份识别方法,用于快速的手机用户身份识别,其特征在于包含以下步骤:

[0006] s1采集步态样本:利用手机采集大规模用户的运动数据;并从运动数据中分离出步行时产生的数据,构建大规模步态样本集;

[0007] s2构建步态检索样本集:基于步骤s1构建的大规模步态样本集,构建步态检索样本集;

[0008] s3对步态样本集进行哈希编码:构建深度学习哈希网络,利用步骤s2构建的步态检索样本集对深度学习哈希网络进行训练,得到深度哈希模型;并从深度哈希模型中取出从输入层到哈希层的部分,作为哈希编码器;利用该哈希编码器对步骤s1构建的步态样本集中的步态样本进行哈希编码,从而将输入的步态样本编码成只有0和1组成的向量;

[0009] s4对步态样本集进行特征描述:构建深度学习分类网络,利用步骤s1构建的大规模步态样本集对深度学习分类网络进行训练,得到步态分类模型;从所述步态分类模型中取出从输入层到最后一个隐含层的部分,作为步态特征描述器;利用该步态特征描述器对步态样本集中的所有样本进行步态特征描述;

[0010] s5识别未知身份步态:对于待识别的未知身份步态样本 i ,利用步骤4所得的哈希编码器进行编码得到 H_i ;计算 H_i 与步骤5所得的已知身份的步态哈希编码的海明距离(Hamming Distance),得到最相似的 K 个用户。利用所述的步态特征描述器,对待识别的未知身份步态样本 i 进行步态特征描述得到 F_i ,并分别计算 F_i 和最相似的 K 个用户的步态特征之间的相似度;并对相似度值进行排序,将最大相似度值对应的用户身份作为未知身份步态样本的身份。

[0011] 进一步地,所述步骤s1中的运动数据包含但不限于步行,还包括跑、跳、走、乘车颠簸和静坐;所述运动数据采用手机APP程序采集并通过无线通讯传输到服务器。

[0012] 进一步地,所述步骤s1中构建步态样本数据时,步态样本的采用按时间分割的方式。

[0013] 进一步地,所述步骤s1中分离出步行时产生的数据具体为:根据步行的周期特点,采用加速度的幅度阈值、跨度阈值来对步态数据进行过滤。

[0014] 进一步地,所述步骤s2中构建步态检索样本集具体过程为:根据步态样本数据集计算相似度矩阵,规定同一用户的样本相似度为1,不同用户的样本相似度为0;构建三元组步态样本,即每个样本包含三个步态样本,其中两个样本相似度为1,另一个样本与它们不相似;将所有三元组步态样本构成步态检索样本集。

[0015] 进一步地,所述深度哈希网络采用一维卷积构建。

[0016] 更进一步地,所述深度哈希网络有8层,具体结构如下:第1层是输入层,即步态样本数据,第2层是卷积层,采用 1×9 的卷积核,卷积核个数为32,步长为2,第3层是池化层,池化尺寸为 1×2 ,第4、5层是两个卷积层,采用 1×3 的卷积核,卷积核个数为64,步长为1,第6层是池化层,池化尺寸为 1×2 ,第7层是卷积层,卷积核个数为128,步长为1,第8层是一个全连接层,称为哈希层,第9层是输出层。

[0017] 进一步地,所述步骤s4中的深度学习分类网络采用双向双层长短时记忆递归神经网络作为分类网络。

[0018] 本发明的优点在于:

[0019] 1) 本发明针对大规模用户环境下手机用户身份难以快速识别的问题,设计了一种先快速步态检索后精细步态比较的身份识别策略,能大幅提高手机用户身份识别的速度。首先在快速步态检索中,利用深度哈希网络对步态数据进行二进制编码,从而可利用亦或操作实现海明距离的快速计算,快速检索到一批相似用户,即相似用户集;然后在精细步态比较中,利用预先训练好的步态特征描述器对相似用户集中的步态样本进行步态特征描述,计算待识别步态样本特征与相似用户集步态特征的相似度,从而判别出待识别步态样本的用户身份。

[0020] 2) 本发明方法在进行手机用户步态识别和身份识别时,设计了步行数据和非步行数据的分离,从而使所发明方法可以进行非约束条件下的手机用户步态识别和身份识别;而传统的利用手机的步态识别和身份方法需要用户在特定区间行走一段距离,这很大程度上限制了识别算法的使用环境;而本发明不需要对用户活动进行约束,只需要利用手机软件采集一段较长时间内用户的所有活动数据,用户可以做步行、跑步、乘车、静坐以及其他一切活动,从而可以大大扩展步态识别算法的应用场景,并且实现数据的隐蔽性采集,使得本发明方法既适合用于大规模手机用户的身份识别与监控,也适用于手机易主的检测。

[0021] 3) 本发明在构建步态样本数据集时,步态样本的采用按时间分割的方式,而不是传统的按步伐分割的方式,因为步伐分割的精确性难移保证,会影响后续分类识别的准确度。而采用按时间分割的方式,并通过大数据量样本的深度学习,可以增强步态特征描述的通用性,提高步态识别的准确度。

附图说明

[0022] 图1为发明的深度哈希网络模型图。

具体实施方式

[0023] 下面结合附图和具体实施例对本发明作进一步的详细描述:

[0024] 一种基于步态的大规模手机用户快速身份识别,包括以下步骤:

[0025] 步骤s1:利用手机采集N个用户的运动数据。采集所用手机传感器包含但不限于加速度计和陀螺仪,还可以包含气压计、磁力计等;用户的运动包含但不限于步行,还可以包含跑步、乘车颠簸、静坐等;手机可以任意姿态放在口袋里或拿在手上;运动数据采用手机App程序采集并通过无线通讯传输到服务器;

[0026] 优选地,手机传感器选用加速度计和陀螺仪;

[0027] 从步骤1采集的运动数据中分离出步行数据,构建大规模步态样本集。在进行步行数据和非步行数据的分离时,根据步行的周期特点,采用加速度的幅度阈值、跨度阈值来对步态数据进行过滤;通常加速度是三轴的,用 A_x , A_y 和 A_z 表示,计算它们的平方和开方值得到 A_{xyz} ,在 A_{xyz} 构成的连续曲线上,步态对应的幅度介于一定范围之内,峰值和峰值之间的间隔也在一定的幅度之内;同一用户的运动数据,先计算其平均的 A_{xyz} 值 A_{ave} ,以及平均的峰值跨度 L_{ave} ,然后利用幅度阈值 A_{min} 和 A_{max} ,跨度阈值 L_{min} 和 L_{max} 对步行数据进行提取,即同时满足介于 A_{min} 和 A_{max} , L_{min} 和 L_{max} 之间的运动数据才被认为是步行数据;对提取的步行数据分割成步态样本,分割采用时间定长的方式,即每一个时间间隔的数据作为一个样本;

[0028] 优选地,设置幅度阈值 $A_{min}=0.7 \times A_{ave}$, $A_{max}=1.3 \times A_{ave}$,设置跨度阈值 $L_{min}=0.7 \times L_{ave}$, $L_{max}=1.3 \times L_{ave}$;

[0029] 优选地,步态样本采用2秒时间间隔内的数据作为一个步态样本,插值后长度为128,插值方法采用线性插值法;

[0030] 优选地,用户数量 $N=10000$,每个用户收集到的步态样本数为1000;

[0031] 步骤s2构建步态检索样本集:基于步骤s1构建的大规模步态样本集,构建步态检索样本集;根据步态样本数据集计算相似度矩阵,规定同一用户的样本相似度为1,不同用户的样本相似度为0;构建三元组步态样本,即每个样本包含三个步态样本,其中两个样本相似度为1,另一个样本与它们不相似;将所有三元组步态样本构成步态检索样本集;

[0032] 优选地,步态检索样本数量为10000000;

[0033] 步骤s3对步态样本集进行哈希编码:构建深度学习哈希网络,利用步骤s2构建的步态检索样本集对深度学习哈希网络进行训练,得到深度哈希模型;并从深度哈希模型中取出从输入层到哈希层的部分,作为哈希编码器;哈希编码器的作用是将输入步态样本编码成只有0和1组成的向量;

[0034] 优选地,深度学习哈希网络主要利用一维卷积来构建,总共有8层,具体为:第1层

是输入层,即步态样本数据,第2层是卷积层,采用 1×9 的卷积核,卷积核个数为32,步长为2,第3层是池化层,池化尺寸为 1×2 ,第4、5层是两个卷积层,采用 1×3 的卷积核,卷积核个数为64,步长为1,第6层是池化层,池化尺寸为 1×2 ,第7层是卷积层,卷积核个数为128,步长为1,第8层是一个全连接层,称为哈希层,节点数为 V ,第9层是输出层。

[0035] 哈希层节点数可以采用多种设置方式,比如24、32、50、48、64等,本发明中,优选地,设置哈希层节点数 $V=48$,激活函数选择Sigmoid函数;

[0036] 利用所得的哈希编码器对步骤s1构建的步态样本集中的步态样本进行哈希编码;

[0037] 步骤s4对步态样本集进行特征描述:构建深度学习分类网络,利用步骤s1构建的大规模步态样本集对深度学习分类网络进行训练,得到步态分类模型;

[0038] 优选地,采用双向双层长短时记忆递归神经网络作为分类网络,隐含节点数设置为1024,输出层节点数与步骤s1中构建的步态样本集中的类别数相同;

[0039] 构建步态特征描述器;具体操作是,从所得的步态分类模型中取出从输入层到最后一个隐含层的部分,作为步态特征描述器;

[0040] 利用所得的步态特征描述器,对步骤s1构建的步态样本集中的所有样本进行步态特征描述;

[0041] 步骤s5识别未知身份步态:对于待识别的未知身份步态样本 i ,利用步骤s3所得的哈希编码器进行编码得到 H_i ;计算 H_i 与步骤s3所得的已知身份的步态哈希编码的海明距离(Hamming Distance),得到最相似的 K 个用户;

[0042] 优选地,从所得的已知身份的步态哈希编码中个,对所有 N 个用户中的每个用户取出1个哈希编码,分别计算 H_i 与它们之间的距离,将距离值最小的 K 个用户构成相似用户集;

[0043] 优选地,设置 $K=100$;

[0044] 利用步骤s4所得的步态特征描述器,对未知身份步态样本 i 进行步态特征描述得到 F_i ,并分别计算 F_i 和所得的最相似的 K 个用户的步态特征之间的相似度;

[0045] 对所得的相似度值进行排序,将最大相似度值对应的用户身份作为未知身份步态样本的身份。

[0046] 本发明针对大规模用户环境下手机用户身份难以快速识别的问题,设计了一种先快速步态检索后精细步态比较的身份识别策略,能大幅提高手机用户身份识别的速度。首先在快速步态检索中,利用深度哈希网络对步态数据进行二进制编码,从而可利用亦或操作实现海明距离的快速计算,快速检索到一批相似用户,即相似用户集;然后在精细步态比较中,利用预先训练好的步态特征描述器对相似用户集中的步态样本进行步态特征描述,计算待识别步态样本特征与相似用户集步态特征的相似度,从而判别出待识别步态样本的用户身份。

[0047] 本发明方法在进行手机用户步态识别和身份识别时,设计了步行数据和非步行数据的分离,从而使所发明方法可以进行非约束条件下的手机用户步态识别和身份识别;而传统的利用手机的步态识别和身份方法需要用户在特定区间行走一段距离,这很大程度上限制了识别算法的使用环境;而本发明不需要对用户活动进行约束,只需要利用手机软件采集一段较长时间内用户的所有活动数据,用户可以做步行、跑步、乘车、静坐以及其他一切活动,从而可以大大扩展步态识别算法的应用场景,并且实现数据的隐蔽性采集,使得本发明方法既适合用于大规模手机用户的身份识别与监控,也适用于手机易主的检测。

[0048] 发明在构建步态样本数据集时,步态样本的采用按时间分割的方式,而不是传统的按步伐分割的方式,因为步伐分割的精确性难移保证,会影响后续分类识别的准确度。而采用按时间分割的方式,并通过大数据量样本的深度学习,可以增强步态特征描述的通用性,提高步态识别的准确度。

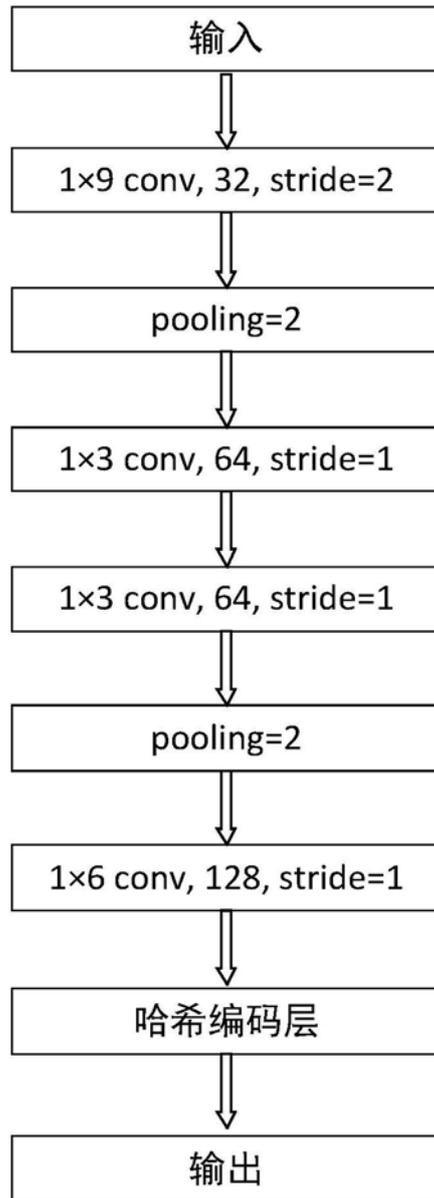


图1