



(12) 发明专利

(10) 授权公告号 CN 111046216 B

(45) 授权公告日 2024. 02. 09

(21) 申请号 201911248130.6

(22) 申请日 2019.12.06

(65) 同一申请的已公布的文献号

申请公布号 CN 111046216 A

(43) 申请公布日 2020.04.21

(73) 专利权人 广州国音智能科技有限公司

地址 510000 广东省广州市黄埔区黄埔东

路3110号三社大楼二楼2009房

(72) 发明人 彭辉 黎智勇 许敏强

(74) 专利代理机构 深圳市世纪恒程知识产权代

理事务所 44287

专利代理师 宋朝政

(51) Int. Cl.

G06F 16/61 (2019.01)

G06F 16/68 (2019.01)

(56) 对比文件

CN 106250400 A, 2016.12.21

CN 108268667 A, 2018.07.10

CN 109284416 A, 2019.01.29

CN 104090902 A, 2014.10.08

CN 108538311 A, 2018.09.14

US 2005182627 A1, 2005.08.18

审查员 陈树华

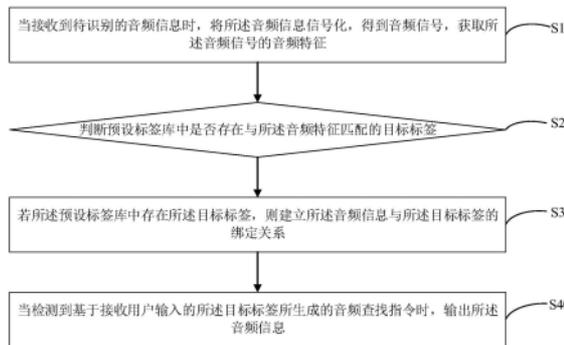
权利要求书2页 说明书10页 附图3页

(54) 发明名称

音频信息存取方法、装置、设备及计算机可读存储介质

(57) 摘要

本发明公开了一种音频信息存取方法,该方法包括:当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征;判断预设标签库中是否存在与所述音频特征匹配的目标标签;若所述预设标签库中存在所述目标标签,则建立所述音频信息与所述目标标签的绑定关系;当基于所述目标标签接收到音频查找指令时,输出所述音频信息。本发明还公开了一种音频信息存取装置、设备及可读存储介质。本发明通过获取将音频信息信号化处理得到的音频信号的音频特征,并在预设标签库中存在与音频特征匹配的目标标签时,建立音频信息与目标标签的绑定关系,以使基于目标标签检索时,输出音频信息。实现了对音频信息的管理。



1. 一种音频信息存取方法,其特征在于,所述音频信息存取方法包括以下步骤:

当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征;

判断预设标签库中是否存在与所述音频特征匹配的目标标签;

若所述预设标签库中存在所述目标标签,则建立所述音频信息与所述目标标签的绑定关系;

当检测到基于接收用户输入的所述目标标签所生成的音频查找指令时,输出所述音频信息;

其中,所述获取所述音频信号的音频特征的步骤包括:

当得到音频信号之后,获取所述音频信号的总时长和帧长;

判断所述音频信号是否具有周期性;

若所述音频信号具有周期性,则根据所述总时长和所述帧长,计算所述音频信号的过零率,并将所述过零率作为第一音频特征;

将所述第一音频特征与预设标签库中的性别标签匹配得到第一匹配值,将大于第一预设阈值的所述第一匹配值对应的性别标签作为目标标签;

若所述音频信号具有周期性,则获取所述音频信号的信号强度;

根据所述帧长和所述信号强度,计算所述音频信号单帧能量,并将所述单帧能量作为第二音频特征;

将所述第二音频特征与预设标签库中的年龄段匹配得到第二匹配值,将大于第二预设阈值的第二匹配值对应的年龄段标签作为目标标签。

2. 如权利要求1所述的音频信息存取方法,其特征在于,所述当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征的步骤之后,包括:

获取所述音频信号的音频特征之后,判断所述音频特征中是否包含非语音特征;

若所述音频特征中包含非语音特征,则获取预设场景集合中与所述非语音特征适配的目标场景,并建立所述音频信息与所述目标场景的绑定关系;

若所述音频特征中不包含所述非语音特征,则执行判断预设标签库中是否存在与所述音频特征匹配的目标标签的步骤。

3. 如权利要求1所述的音频信息存取方法,其特征在于,所述判断所述音频信号是否具有周期性的步骤之后,还包括:

若所述音频信号不具有周期性,则获取所述音频信号的第三音频特征,并判断所述第三音频特征中是否包含非语音特征;

若所述第三音频特征中包含所述非语音特征,则判定所述音频信号对应的音频信息为无效音频。

4. 如权利要求2所述的音频信息存取方法,其特征在于,所述若所述音频特征中包含非语音特征,则获取预设场景集合中与所述非语音特征适配的目标场景,并建立所述音频信息与所述目标场景的绑定关系的步骤包括:

获取音频特征中的非语音特征,判断预设场景模型库中是否存在与所述非语音特征匹配的目标场景模型;

若存在所述目标场景模型,则将预设场景集合中与所述目标场景模型对应的场景作为目标场景,并建立音频信息与所述目标场景的绑定关系。

5. 一种音频信息存取装置,其特征在于,所述音频信息存取装置包括:

接收获取模块,用于当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征;

判断模块,用于判断预设标签库中是否存在与所述音频特征匹配的目标标签;

建立模块,用于若所述预设标签库中存在所述目标标签,则建立所述音频信息与所述目标标签的绑定关系;

输出模块,用于当检测到基于接收用户输入的所述目标标签所生成的音频查找指令时,输出所述音频信息;

所述音频信息存取装置,还用于当得到音频信号之后,获取所述音频信号的总时长和帧长;判断所述音频信号是否具有周期性;若所述音频信号具有周期性,则根据所述总时长和所述帧长,计算所述音频信号的过零率,并将所述过零率作为第一音频特征;将所述第一音频特征与预设标签库中的性别标签匹配得到第一匹配值,将大于第一预设阈值的所述第一匹配值对应的性别标签作为目标标签;若所述音频信号具有周期性,则获取所述音频信号的信号强度;根据所述帧长和所述信号强度,计算所述音频信号单帧能量,并将所述单帧能量作为第二音频特征;将所述第二音频特征与所述预设标签库中的年龄段匹配得到第二匹配值,将大于第二预设阈值的第二匹配值对应的年龄段标签作为目标标签。

6. 一种音频信息存取设备,其特征在于,所述音频信息存取设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的音频信息存取程序,所述音频信息存取程序被所述处理器执行时实现如权利要求1至4中任一项所述的音频信息存取方法的步骤。

7. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有音频信息存取程序,所述音频信息存取程序被处理器执行时实现如权利要求1至4中任一项所述的音频信息存取方法的步骤。

音频信息存取方法、装置、设备及计算机可读存储介质

技术领域

[0001] 本发明涉及语音处理领域,尤其涉及一种音频信息存取方法、装置、及计算机可读存储介质。

背景技术

[0002] 随着语音处理技术的快速发展,对因各种原因产生的大量声音信息进行管理成为了相关领域需要解决的技术问题。

[0003] 在公安系统内部存在着大量的音频信息,这些音频信息包括与案件相关的音频资料,也包括根据某些需要采集而来的音频文件。现有的对这些音频信息的管理方法中,只能通过人工识别,并逐个为之添加标签,这无疑增加了管理的工作量,且音频信息之间存在信息量多少和信息复杂度的问题,这又增加了人工处理的困难度。因此,自动为音频信息添加对应的标签,通过标签对音频信息查看,提高存取音频信息的效率,成为了亟待解决的技术问题。

发明内容

[0004] 本发明的主要目的在于提供一种音频信息存取方法,旨在解决人工管理音频信息时,存取音频信息效率低的技术问题。

[0005] 此外,为实现上述目的,本发明还提供一种音频信息存取方法,所述音频信息存取方法包括以下步骤:

[0006] 当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征;

[0007] 判断预设标签库中是否存在与所述音频特征匹配的目标标签;

[0008] 若所述预设标签库中存在所述目标标签,则建立所述音频信息与所述目标标签的绑定关系;

[0009] 当检测到基于接收用户输入的所述目标标签所生成的音频查找指令时,输出所述音频信息。

[0010] 可选地,所述当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征的步骤之后,包括:

[0011] 获取所述音频信号的音频特征之后,判断所述音频特征中是否包含非语音特征;

[0012] 若所述音频特征中包含非语音特征,则获取预设场景集合中与所述非语音特征适配的目标场景,并建立所述音频信息与所述目标场景的绑定关系;

[0013] 若所述音频特征中不包含所述非语音特征,则执行判断预设标签库中是否存在与所述音频特征匹配的目标标签的步骤。

[0014] 可选地,所述获取所述音频信号的音频特征的步骤包括:

[0015] 当得到音频信号之后,获取所述音频信号的总时长和帧长;

[0016] 判断所述音频信号是否具有周期性;

[0017] 若所述音频信号具有周期性,则根据所述总时长和所述帧长,计算所述音频信号的过零率,并将所述过零率作为第一音频特征。

[0018] 可选地,所述若所述音频信号具有周期性的步骤之后,包括:

[0019] 若所述音频信号具有周期性,则获取所述音频信号的信号强度;

[0020] 根据所述帧长和所述信号强度,计算所述音频信号单帧能量,并将所述单帧能量作为第二音频特征。

[0021] 可选地,所述判断所述音频信号是否具有周期性的步骤之后,还包括:

[0022] 若所述音频信号不具有周期性,则获取所述音频信号的第三音频特征,并判断所述第三音频特征中是否包含非语音特征;

[0023] 若所述第三音频特征中包含所述非语音特征,则判定所述音频信号对应的音频信息为无效音频。

[0024] 可选地,所述若所述音频信号具有周期性的步骤之后,包括:

[0025] 若所述音频信号具有周期性,则获取第一音频特征,将所述第一音频特征与预设标签库中的性别标签匹配得到第一匹配值,将大于第一预设阈值的第一匹配值对应的性别标签作为目标标签;

[0026] 若所述音频信号具有周期性,则获取第二音频特征,将所述第二音频特征与所述预设标签库中的年龄段匹配得到第二匹配值,将大于第二预设阈值的第二匹配值对应的年龄段标签作为目标标签。

[0027] 可选地,所述若所述音频特征中包含非语音特征,则获取预设场景集合中与所述非语音特征适配的目标场景,并建立所述音频信息与所述目标场景的绑定关系的步骤包括:

[0028] 获取音频特征中的非语音特征,判断预设场景模型库中是否存在与所述非语音特征匹配的目标场景模型;

[0029] 若存在所述目标场景模型,则将预设场景集合中与所述目标场景模型对应的场景作为目标场景,并建立音频信息与所述目标场景的绑定关系。

[0030] 此外,为实现上述目的,本发明还提供一种音频信息存取设备,所述音频信息存取设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的音频信息存取程序,所述音频信息存取程序被所述处理器执行时实现如上述的音频信息存取方法的步骤。

[0031] 此外,为实现上述目的,本发明还提供一种计算机可读存储介质,所述计算机可读存储介质上存储有音频信息存取程序,所述音频信息存取程序被处理器执行时实现如上述的音频信息存取方法的步骤。

[0032] 本发明实施例提出的一种音频信息存取方法、装置、设备及可读存储介质。本发明实施例中当音频信息存取程序接收到待识别(即,待管理)的音频信息时,对所述音频信息做信号化处理,处理完成即得到音频信号,音频信息存取程序获取音频信号的音频特征后,进而判断预设标签库中是否存在与获取到的音频特征匹配的目标标签,若判断的结果为,预设标签库中存在与获取到的音频特征匹配的目标标签,则音频信息存取程序将直接建立目标标签与待识别的音频信息之间的绑定关系,绑定关系建立后,当音频信息存取程序接收到用户手动输入或系统自动输入的,基于标签的音频查找指令时,音频信息存取

程序将直接输出与标签之间存在绑定关系的音频信息。实现了对音频信息存取效率的提高。

附图说明

- [0033] 图1为本发明实施例提供的音频信息存取设备一种实施方式的硬件结构示意图；
[0034] 图2为本发明音频信息存取方法第一实施例的流程示意图；
[0035] 图3为本发明音频信息存取方法第二实施例的流程示意图；
[0036] 图4为本发明音频信息存取方法第三实施例的流程示意图；
[0037] 图5为本发明音频信息存取方法第四实施例的流程示意图；
[0038] 图6为本发明装置一实施例的功能模块示意图。
[0039] 本发明目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

- [0040] 应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。
- [0041] 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”的后缀仅为为了有利于本发明的说明，其本身没有特定的意义。因此，“模块”、“部件”或“单元”可以混合地使用。
- [0042] 本发明实施例音频信息存取终端(又叫终端、设备或者终端设备)可以是PC，也可以是智能手机、平板电脑和便携计算机等具有显示功能和音频处理功能的可移动式终端设备。
- [0043] 如图1所示，该终端可以包括：处理器1001，例如CPU，网络接口1004，用户接口1003，存储器1005，通信总线1002。其中，通信总线1002用于实现这些组件之间的连接通信。用户接口1003可以包括显示屏(Display)、输入单元比如键盘(Keyboard)，可选用户接口1003还可以包括标准的有线接口、无线接口。网络接口1004可选的可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器1005可以是高速RAM存储器，也可以是稳定的存储器(non-volatile memory)，例如磁盘存储器。存储器1005可选的还可以是独立于前述处理器1001的存储装置。
- [0044] 可选地，终端还可以包括摄像头、RF(Radio Frequency, 射频)电路，传感器、音频电路、WiFi模块等等。其中，传感器比如光传感器、运动传感器以及其他传感器。具体地，光传感器可包括环境光传感器及接近传感器，其中，环境光传感器可根据环境光线的明暗来调节显示屏的亮度，接近传感器可在移动终端移动到耳边时，关闭显示屏和/或背光。作为运动传感器的一种，重力加速度传感器可检测各个方向上(一般为三轴)加速度的大小，静止时可检测出重力的大小及方向，可用于识别移动终端姿态的应用(比如横竖屏切换、相关游戏、磁力计姿态校准)、振动识别相关功能(比如计步器、敲击)等；当然，移动终端还可配置陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器，在此不再赘述。
- [0045] 本领域技术人员可以理解，图1中示出的终端结构并不构成对终端的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。
- [0046] 如图1所示，作为一种计算机存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及音频信息存取程序。

[0047] 在图1所示的终端中,网络接口1004主要用于连接后台服务器,与后台服务器进行数据通信;用户接口1003主要用于连接客户端(用户端),与客户端进行数据通信;而处理器1001可以用于调用存储器1005中存储的音频信息存取程序,所述音频信息存取程序被处理器执行时实现下述实施例提供的音频信息存取方法中的操作。

[0048] 基于上述设备硬件结构,提出了本发明音频信息存取方法的实施例。

[0049] 参照图2,在本发明音频信息存取方法的第一实施例中,所述音频信息存取方法包括:

[0050] 步骤S10,当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征。

[0051] 本实施例中音频信息存取方法应用于音频信息存取设备(又叫终端、设备或者终端设备),音频信息存取设备包括PC、智能手机、平板电脑和便携计算机等具有显示功能和音频处理功能的终端设备。

[0052] 本实施例中音频信息存取方法能够实现对音频信息的标签化管理,以及当用户搜索目标标签时,能够查找到与目标标签对应的音频信息。本实施例中音频信息是指,用户获取到的待管理的音频,音频存在的形式可以是任何格式的包含声音信息且可被播放的文件。将接收到的音频信息信号化操作的目的是,将无形的声音转化成有形的声波,这有助于通过对声波的分析获得声音的特征,进而根据声音特征对音频信息进行管理。已知,将音频信息信号化得到的音频信号实质上是反映声音特征的波形图,声波按照默认的时间线展开,通过获取音频信号在整个时间线上所体现出的某些特征,可以将不同的音频信息进行区分,给不同的音频信息添加标识即实现了对音频信息的管理,而音频信号在整个时间线上所体现出的某些特征即是本实施例中的音频特征。

[0053] 本实施例以人的声音(即语音)来举例说明,可以理解的是,对于语音信息的管理具有代表性且更易理解,但并不代表本实施例中音频信息存取方法仅能用于语音信息的管理。

[0054] 步骤S20,判断预设标签库中是否存在与所述音频特征匹配的目标标签。

[0055] 可知地,对音频信息存取的目的是,解决在音频信息较多时人工对其进行管理的复杂和不便。本实施例中预设标签库中存储有各种标签,存放这些标签的目的是为了给不同音频特征的音频信息添加对应的标签,标签的作用除了用于区分不同的音频信息外,还可以在标签内存储与音频信息有关的信息,例如,音频信息内的人物性别,年龄分布以及音频发生的场景(声音背景),以便于根据标签查找含有该标签的音频信息,可以理解的是,一段完整的音频信息可以对应多个标签,当音频信息较多时,一个标签也会对应多个音频信息。

[0056] 本实施例给出一种具体应用场景,已知一段音频信息,经过对其信号化处理得到音频信号,通过获取该音频信号的音频特征,确定该音频信号对应的标签为性别:女性,年龄段:16至30岁,音频背景:马路边,确定该音频信号对应的标签后,音频信息存取程序将这些标签与这段音频信息绑定,绑定成功并保存后,当用户按照女性,16至30岁和马路边之中的一个或多个标签进行查找时,都将搜索到这段音频信息。

[0057] 步骤S30,若所述预设标签库中存在所述目标标签,则建立所述音频信息与所述目标标签的绑定关系。

[0058] 本实施例中给音频信息添加标签的目的是,给不同的音频信息添加标识以实现对音频信息的管理,所以给音频信息添加标签的前提是,根据音频信息的音频特征可以将每个音频信息进行区分。已知,因为音频信息的获取途径是多种多样的,所以不免会存在无法与任何标签匹配的音频信息,比如,一段只有噪音的音频信息,当出现这种特殊的音频信息时,预设标签库中查找不到与之匹配的目标标签,这时,音频信息存取程序将会输出“不存在标签”的提示信息。其中,提示信息的形式和内容本实施例不做详述和限制。可知地,当预设标签库中存在与音频信息匹配的目标标签时,音频信息存取程序将会建立音频信息与目标标签的绑定关系,以使用户后期通过目标标签查找到对应的音频信息。目标标签可能为一个,也可能为多个。

[0059] 步骤S40,当检测到基于接收用户输入的所述目标标签所生成的音频查找指令时,输出所述音频信息。

[0060] 建立音频信息与目标标签绑定关系的最终目的是,方便用户后续在数量较多的音频信息中查找某一类或某一个音频信息。用户可以通过多种标签的任意组合查找用户想要查找的一类音频,已知,标签存在类别,当用户选择同一类别的多个标签时,音频信息存取程序将默认查找与每个标签匹配的音频信息。当用户选择不同类别的多个标签时,音频信息存取程序将默认查找同时与所有标签匹配的音频信息。

[0061] 本实施例给出一种具体应用场景,若标签的类型包括性别,年龄和场景,其中,性别包括男性和女性两种,年龄包括多个年龄段,这里不做详述,场景包括公路,机场和火车站等。当用户输入的搜索标签是男性和女性时,音频信息存取程序将会默认查找所有与男性标签匹配的音频信息,和所有与女性标签匹配的音频信息。而当用户输入的标签是男性标签,16至30岁标签和机场标签时,音频信息存取程序将会查找同时与上述三个标签匹配的音频信息。可知地,用户在输入年龄标签时,还可以输入一个确切的年龄,音频信息存取程序会根据用户输入的年龄查找对应的年龄段。

[0062] 在本实施例中通过先将待识别的音频信息信号化处理得到音频信号,进而通过获取到音频信号的音频特征,并判断在预设标签库中是否存在与获取到的音频特征匹配的目标标签,通过判断排除一些没有管理价值的音频信息(比如噪音),当预设标签库中存在与获取到的音频特征匹配的目标标签时,音频信息存取程序将建立音频信息与目标标签的绑定关系,绑定关系建立完成后,当用户输入标签查找对应的音频信息时,音频信息存取程序将会查找与用户输入的标签存在绑定关系的音频信息,并将之输出。当用户需要管理的音频信息较多时,通过对每个音频信息添加与之匹配的标签,实现了对音频信息查找效率的提高。

[0063] 进一步地,参照图3,在本发明上述实施例的基础上,提出了本发明音频信息存取方法的第二实施例。

[0064] 本实施例是第一实施例中步骤S10之后的步骤,本实施例与本发明上述实施例的区别在于:

[0065] 步骤S50,获取所述音频信号的音频特征之后,判断所述音频特征中是否包含非语音特征。

[0066] 步骤S60,若所述音频特征中包含非语音特征,则获取预设场景集合中与所述非语音特征适配的目标场景,并建立所述音频信息与所述目标场景的绑定关系。

[0067] 步骤S70,若所述音频特征中不包含所述非语音特征,则执行判断预设标签库中是否存在与所述音频特征匹配的目标标签的步骤。

[0068] 已知,音频是所有声音的统称,音频包括语音(即人的声音)和其他声音(即非语音),本实施例将音频特征区分为语音特征和非语音特征的目的除了便于区分管理外,还由于人声的特别之处,即语音在信号表征上具有一定的规律性,与声纹鉴定等技术相结合具有易辨识的特点。而非语音在一定程度上是噪音,信号表征上不具有规律性,且本实施例中音频信息存取方法中标签的设置也与音频特征中是否包含语音特征有关,即若音频信息的音频特征中包含语音特征,则可以与该音频信息匹配的标签就与人有关,若音频信息的音频特征中包含非语音特征,则可以与该音频信息匹配的标签就与声音本身或者声音发生的场景有关。已知,存在一段音频信息中同时包含语音特征和非语音特征的情况,这种情况下,为了便于区分管理,可以采用音频分离技术将语音从音频信息中分离出来,这样音频信息就只包含一种音频特征了。

[0069] 本实施例中目标场景除了与声音发生的背景有关,还可以跟发出音频的物品有关,例如,钢琴声,雨声,风声和燃烧声等灵活地设置更多的标签,然后对标签进行分类更利于后期用户根据标签查找对应的音频信息。

[0070] 具体地,步骤S60细化的步骤,还包括:

[0071] 步骤a1,获取音频特征中的非语音特征,判断预设场景模型库中是否存在与所述非语音特征匹配的目标场景模型。

[0072] 步骤a2,若存在所述目标场景模型,则将预设场景集合中与所述目标场景模型对应的场景作为目标场景,并建立音频信息与所述目标场景的绑定关系。

[0073] 本实施例中预设场景模型库中存储有多个场景模型,这些场景模型有些是与音频发生的背景有关,例如汽车喇叭声模型和飞机引擎声模型,它们分别对应的目标场景是马路边和飞机场;有些模型是与音频发生的情境有关,例如,雨声模型和风声模型,它们分别对应的情境是下雨天和起风日,无论是音频发生的背景还是音频发生的情景,所有与语音无关的声音在本实施例中被统称为场景。通过判断预设场景模型库中是否存在与音频信息中的非音频特征匹配的目标场景模型,若预设场景模型库中不存在与音频信息中的非音频特征匹配的目标场景模型,则说明待管理的音频信息没有管理价值,或者场景模型不足,用户可以选择舍弃该音频信息,或者添加更多的场景模型以适应不同的管理需求。可知地,当预设场景模型库中不存在与音频信息中的非音频特征匹配的目标场景模型时,音频信息存取程序将输出提示信息,提示用户删除该音频信息,或者暂时搁置以添加更多的场景模型。若存在目标场景模型,则音频信息存取程序将会把目标场景模型对应的场景作为目标场景。

[0074] 在本实施例中通过将音频信息区分语音和非语音,并建立对应的标签,实现了对音频信息更加方便的管理。

[0075] 进一步地,参照图4,在本发明上述实施例的基础上,提出了本发明音频信息存取方法的第三实施例。

[0076] 本实施例是第一实施例中步骤S10细化的步骤,本实施例与本发明上述实施例的区别在于:

[0077] 步骤S11,当得到音频信号之后,获取所述音频信号的总时长和帧长。

[0078] 步骤S12,判断所述音频信号是否具有周期性。

[0079] 步骤S13,若所述音频信号具有周期性,则根据所述总时长和所述帧长,计算所述音频信号的过零率,并将所述过零率作为第一音频特征。

[0080] 当接收到待管理的音频信息之后,音频信息存取程序会将该音频信息信号化得到音频信号,具体做法可以是,利用某些声音处理软件得到该音频在声波形式上的表征。已知,声波在时间线上具有连续性,震动点在零点线上线波动形成声波的基本单位一周期,本实施例中将一个周期的声波称为一帧,单位时间内声波的周期数与声音的频率有关,通过获取音频信号的总时长和帧长(即周期数),可以计算该音频信息的声音频率。通过判断音频信号是否具有周期性,可以判定该音频信号具有语音特征还是具有非语音特征,这决定了后续对应标签的绑定操作。判断音频信号是否具有周期性的目的是,通过判断声波在整个时间线上是否表现出周期性(即一定的规律性),来判定该音频信息具有语音特征还是非语音特征。若声波在整个时间线上表现出了一定的周期性,则音频信息具有语音特征,若声波在整个时间线上未表现出了一定的周期性,则音频信息具有非语音特征。若音频信息存取程序判定的结果为,音频信号具有周期性,则根据获取到的音频信号的总时长和帧长,计算出音频信号的过零率。音频信号的过零率是指,声波的震动点通过零点线的频率,过零率与声音的频率有关,频率越高过零率越大。将过零率作为第一音频特征在一定程度上反映了声音发起人的性别特征。

[0081] 具体地,步骤S12之后的步骤,还包括:

[0082] 步骤b1,若所述音频信号具有周期性,则获取所述音频信号的信号强度。

[0083] 步骤b2,根据所述帧长和所述信号强度,计算所述音频信号单帧能量,并将所述单帧能量作为第二音频特征。

[0084] 步骤b3,若所述音频信号不具有周期性,则获取所述音频信号的第三音频特征,并判断所述第三音频特征中是否包含非语音特征。

[0085] 步骤b4,若所述第三音频特征中包含所述非语音特征,则判定所述音频信号对应的音频信息为无效音频。

[0086] 在音频信号具有周期性的基础上,通过进一步获取音频信号的信号强度和再一次获取音频信号的帧长,可以计算出该音频信号每帧所具有的能量,本实施例中将音频信号每帧所具有的能量称之为单帧能量,将单帧能量作为第一音频特征可以在一定程度上反映声音发起人的年龄特征。已知,声音具有能量,且声音能量的大小与声音的强度成正相关,通过获取音频信号的总能量和音频信号的帧长,很容易计算音频信号的单帧能量,即总能量除以帧长。

[0087] 当音频信号不具有周期性的时候,即初步确定音频信息中不包含语音特征。已知音频信息中不包含语音特征并不代表音频信息中就一定包含非语音特征,所以获取音频信号的音频特征后,进一步判断音频特征中是否包含非语音特征就显得很有必要,若音频特征中既不包含语音特征,又不包含非语音特征,则音频信息存取程序将判定该音频信息为无效音频,即,该音频信息中可能不包含有效声音,也可能仅存在无法识别的杂音。

[0088] 在本实施中通过计算音频信号的过零率和单帧能量,详细说明了音频特征的组成部分,也进一步说明了标签与音频特征的对应关系。

[0089] 进一步地,参照图5,在本发明上述实施例的基础上,提出了本发明音频信息存取

方法的第四实施例。

[0090] 本实施例是第三实施例中步骤b1之后的步骤,本实施例与本发明上述实施例的区别在于:

[0091] 步骤c1,若所述音频信号具有周期性,则获取第一音频特征,将所述第一音频特征与预设标签库中的性别标签匹配得到第一匹配值,将大于第一预设阈值的第一匹配值对应的性别标签作为目标标签。

[0092] 步骤c2,若所述音频信号具有周期性,则获取第二音频特征,将所述第二音频特征与预设标签库中的年龄段匹配得到第二匹配值,将大于第二预设阈值的第二匹配值对应的年龄段标签作为目标标签。

[0093] 本实施例中,在音频特征中具有语音特征的基础上,对音频特征的组成部分进行细分,得到第一音频特征和第二音频特征。已知,过零率与声音的频率有关,而声音发起人的性别在一定程度上与声音的频率有关,女性发出的声音一般频率较高,而男性的声音一般频率较低,但这个情况不是绝对的。又可知,单帧能量可以在一定程度上反映声音发起人的年龄状态,年龄较低和较高的人发出的声音单帧能量较小,而正值年轻状态的人发出的声音单帧能量较大,但相似年龄的人之间不存在这种确定的单帧能量的差异,所以,对声音发起人的年龄区分范围较大,本实施例中音频信息存取程序可以将年龄划分为多个区间,具体区间范围本实施例不做详述和限制。

[0094] 当计算得到的过零率与某个性别标签的匹配值大于第一预设阈值,则该性别标签为与音频信息匹配的目标标签,当单帧能量与某个年龄段标签的匹配值大于第二预设阈值,则该年龄段标签为与音频信息匹配的目标标签。

[0095] 在本实施例中音频特征与标签匹配得到匹配值,并通过判断匹配值,进而确定与音频信息匹配的目标标签。

[0096] 本发明还提供一种音频信息存取装置,所述音频信息存取装置包括:

[0097] 接收获取模块,用于当接收到待识别的音频信息时,将所述音频信息信号化,得到音频信号,获取所述音频信号的音频特征;

[0098] 判断模块,用于判断预设标签库中是否存在与所述音频特征匹配的目标标签;

[0099] 建立模块,用于若所述预设标签库中存在所述目标标签,则建立所述音频信息与所述目标标签的绑定关系;

[0100] 输出模块,用于当检测到基于接收用户输入的所述目标标签所生成的音频查找指令时,输出所述音频信息。

[0101] 可选地,所述音频信息存取装置,还包括:

[0102] 第一获取单元,用于获取所述音频信号的音频特征之后,判断所述音频特征中是否包含非语音特征;

[0103] 第一建立单元,用于若所述音频特征中包含非语音特征,则获取预设场景集合中与所述非语音特征适配的目标场景,并建立所述音频信息与所述目标场景的绑定关系;

[0104] 执行单元,用于若所述音频特征中不包含所述非语音特征,则执行判断预设标签库中是否存在与所述音频特征匹配的目标标签的步骤。

[0105] 可选地,所述获取模块包括:

[0106] 第二获取单元,用于当得到音频信号之后,获取所述音频信号的总时长和帧长;

- [0107] 第一判断单元,用于判断所述音频信号是否具有周期性;
- [0108] 第一计算单元,用于若所述音频信号具有周期性,则根据所述总时长和所述帧长,计算所述音频信号的过零率,并将所述过零率作为第一音频特征。
- [0109] 可选地,所述获取模块,还包括:
- [0110] 第三获取单元,用于若所述音频信号具有周期性,则获取所述音频信号的信号强度;
- [0111] 第二计算单元,用于根据所述帧长和所述信号强度,计算所述音频信号单帧能量,并将所述单帧能量作为第二音频特征。
- [0112] 可选地,所述获取模块,还包括:
- [0113] 第四获取单元,用于若所述音频信号不具有周期性,则获取所述音频信号的第三音频特征,并判断所述第三音频特征中是否包含非语音特征;
- [0114] 判定单元,用于若所述第三音频特征中包含所述非语音特征,则判定所述音频信号对应的音频信息为无效音频。
- [0115] 可选地,所述获取模块,还包括:
- [0116] 第一匹配单元,用于若所述音频信号具有周期性,则获取第一音频特征,将所述第一音频特征与预设标签库中的性别标签匹配得到第一匹配值,将大于第一预设阈值的第一匹配值对应的性别标签作为目标标签;
- [0117] 第二匹配单元,用于若所述音频信号具有周期性,则获取第二音频特征,将所述第二音频特征与所述预设标签库中的年龄段匹配得到第二匹配值,将大于第二预设阈值的第二匹配值对应的年龄段标签作为目标标签。
- [0118] 可选地,所述第一建立单元包括:
- [0119] 第二判断单元,用于获取音频特征中的非语音特征,判断预设场景模型库中是否存在与所述非语音特征匹配的目标场景模型;
- [0120] 第二建立单元,用于若存在所述目标场景模型,则将预设场景集合中与所述目标场景模型对应的场景作为目标场景,并建立音频信息与所述目标场景的绑定关系。
- [0121] 上述各程序模块所执行的方法可参照本发明方法各个实施例,此处不再赘述。
- [0122] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者系统不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者系统所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。
- [0123] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。
- [0124] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件商品的形式体现出来,该计算机软件商品存储在如上所述的一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机和平板电脑等)执行本发明各个实施例所述的方法。
- [0125] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发

明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

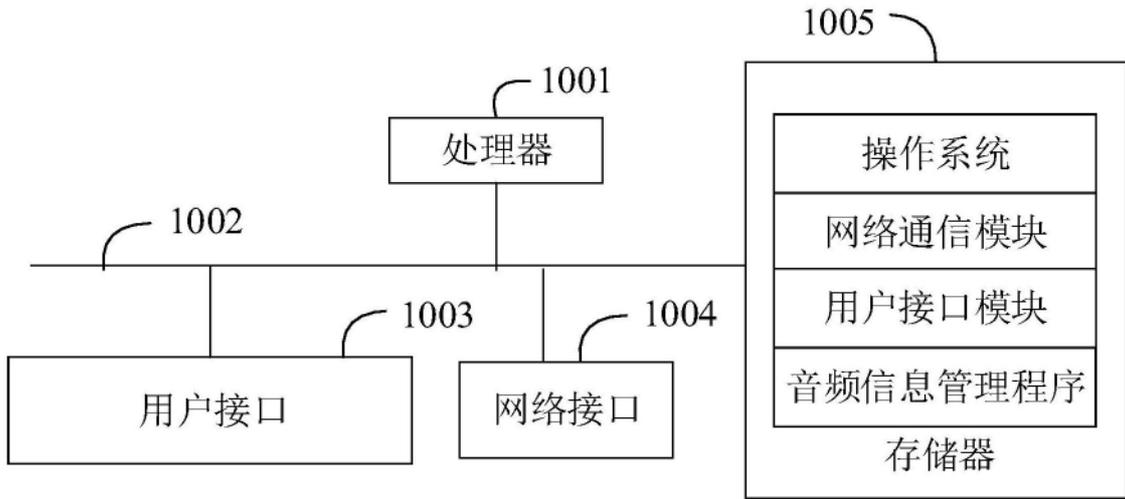


图1

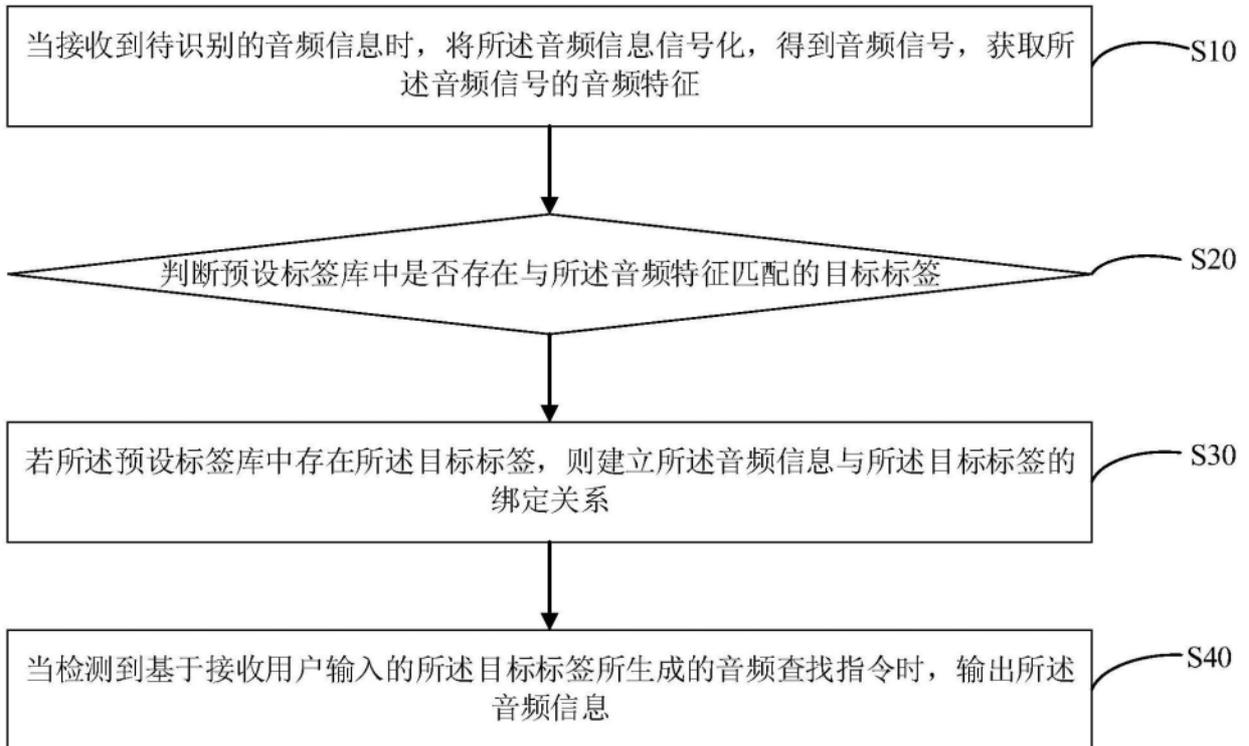


图2

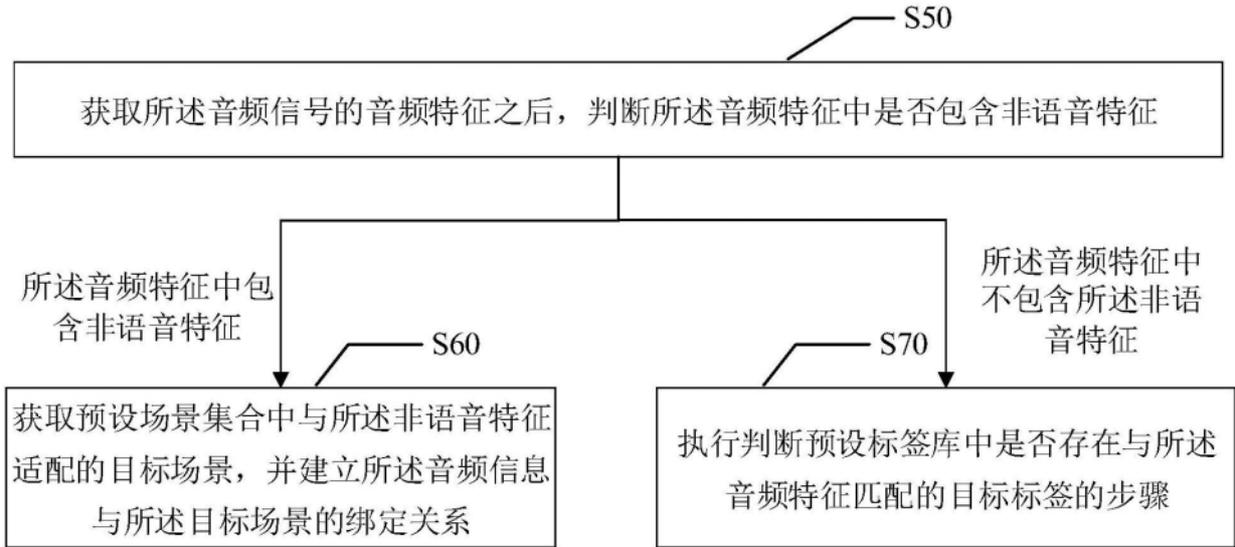


图3

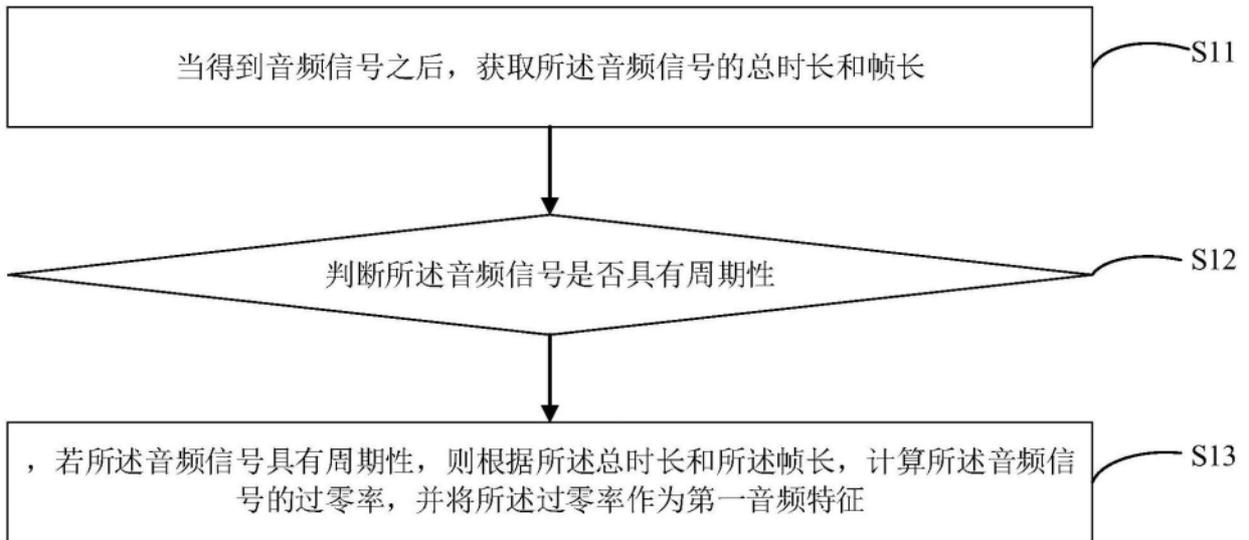


图4

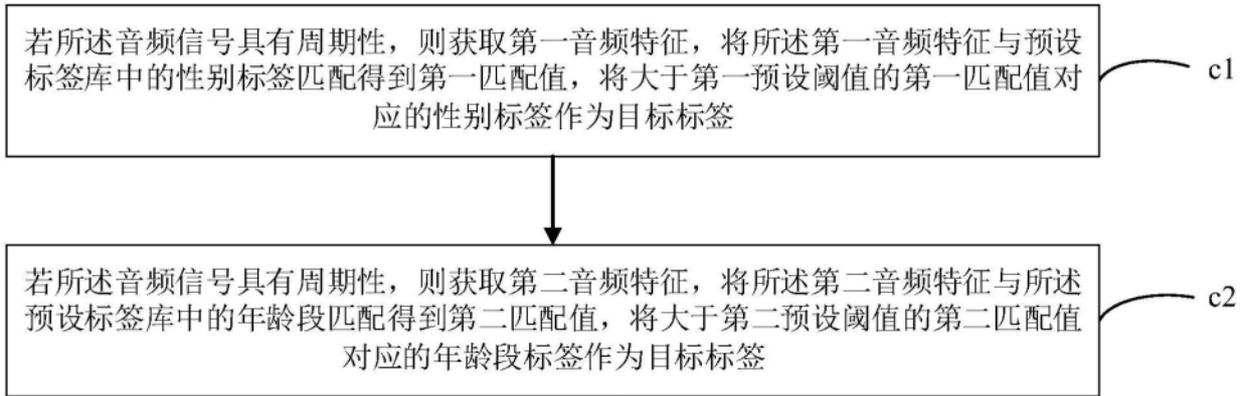


图5

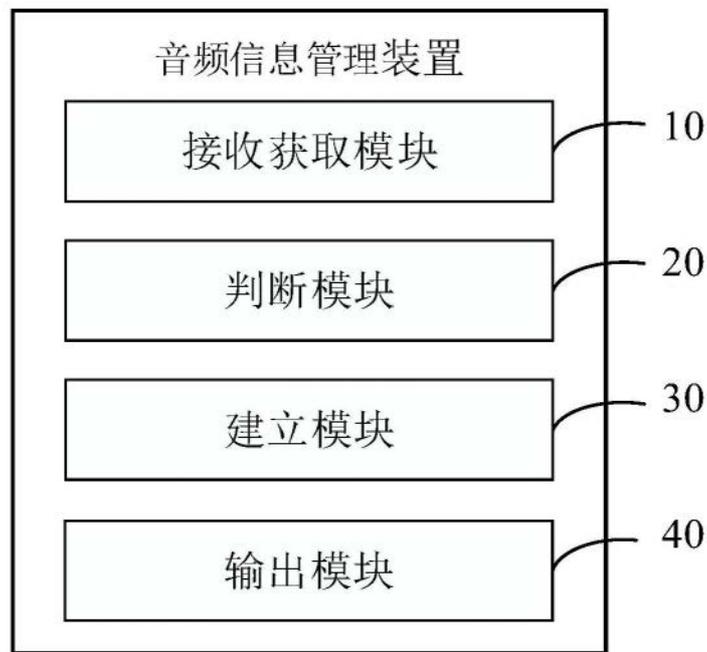


图6