



(12)发明专利

(10)授权公告号 CN 105898219 B

(45)授权公告日 2019.05.21

(21)申请号 201610258363.4

G10L 15/06(2013.01)

(22)申请日 2016.04.22

G10L 15/10(2006.01)

(65)同一申请的已公布的文献号

申请公布号 CN 105898219 A

(56)对比文件

JP 特开2004-153732 A,2004.05.27,

JP 特开2007-228459 A,2007.09.06,

(43)申请公布日 2016.08.24

CN 105407316 A,2016.03.16,

(73)专利权人 北京小米移动软件有限公司

CN 105489217 A,2016.04.13,

地址 100085 北京市海淀区清河中街68号

华润五彩城购物中心二期9层01房间

审查员 盛磊

(72)发明人 贾伟光 马娜 侯恩星

(74)专利代理机构 北京博思佳知识产权代理有

限公司 11415

代理人 殷亚平 陈蕾

(51)Int.Cl.

H04N 7/18(2006.01)

G10L 15/02(2006.01)

权利要求书3页 说明书12页 附图7页

(54)发明名称

对象监控方法及装置

(57)摘要

本公开是关于一种对象监控方法及装置,其中,对象监控方法包括:当采集到的声音中存在目标声音时,判断目标声音的发声原因;向移动终端推送发声信息,该发声信息携带有发声原因。本公开实施例,通过判断采集到的目标声音的发声原因,并向移动终端推送携带有发声原因的发声信息,使得用户可以获得被监控对象的发声原因,达到很好的监控目的。



1. 一种对象监控方法,其特征在于,所述方法应用于监控设备上,所述方法包括:  
当采集到的声音中存在目标声音时,判断所述目标声音的发声原因;  
向移动终端推送发声信息,所述发声信息携带有所述发声原因;  
所述判断所述目标声音的发声原因,包括:  
对所述目标声音进行特征提取,得到目标声音特征信息;  
将所述目标声音特征信息与预存的经训练得到的多个声音模板进行匹配,并根据匹配结果得到所述目标声音的发声原因,所述多个声音模板用于识别不同原因引起的目标声音。
2. 根据权利要求1所述的对象监控方法,其特征在于,在所述向移动终端推送发声信息之后,所述方法还包括:  
接收所述移动终端发送的视频获取指令;  
根据所述视频获取指令向所述移动终端推送视频信息,以使所述移动终端显示所述视频信息。
3. 根据权利要求2所述的对象监控方法,其特征在于,在所述根据所述视频获取指令向所述移动终端推送视频信息之前,所述方法还包括:  
获取所述目标声音的发声位置;  
控制摄像头转向所述发声位置对应的方向,并采集所述视频信息。
4. 根据权利要求1所述的对象监控方法,其特征在于,所述对所述目标声音进行特征提取,得到目标声音特征信息,包括:  
对所述目标声音进行压缩处理,得到压缩声音信息;  
对所述压缩声音信息进行特征提取,得到所述目标声音特征信息。
5. 根据权利要求1所述的对象监控方法,其特征在于,所述判断所述目标声音的发声原因,包括:  
向服务器发送所述目标声音;  
接收所述服务器返回的所述目标声音的发声原因。
6. 根据权利要求2所述的对象监控方法,其特征在于,所述方法还包括:  
若检测到所述监控设备未接入网络,则广播报文,以使所述移动终端显示所述监控设备未接入网络的提示信息;  
接收所述移动终端广播的路由器的名称和密码,并向所述路由器发送连接请求,以连接到所述路由器上。
7. 根据权利要求1所述的对象监控方法,其特征在于,所述方法还包括:  
针对每个声音模板,对当前声音模板进行训练。
8. 根据权利要求7所述的对象监控方法,其特征在于,所述对当前声音模板进行训练,包括:  
采集预设数量的目标声音特征样本;  
将所述预设数量的目标声音特征样本输入当前声音模板,得到所述预设数量的目标声音特征样本相对于当前声音模板的失真距离,所述失真距离用于表示所述预设数量的目标声音特征样本与通过所述当前声音模板输出的目标声音之间的差距;  
若所述失真距离小于预设距离,则停止对所述当前声音模板进行训练;

若所述失真距离大于等于预设距离,则继续对所述当前声音模板进行训练,直至所述失真距离小于预设距离。

9. 根据权利要求1至8任一项所述的对象监控方法,其特征在于,所述发声信息还携带有所述目标声音。

10. 一种对象监控装置,其特征在于,所述装置应用于监控设备上,所述装置包括:

判断模块,被配置为当采集到的声音中存在目标声音时,判断所述目标声音的发声原因;

第一推送模块,被配置为向移动终端推送发声信息,所述发声信息携带有所述判断模块判断出的所述发声原因;

所述判断模块包括:

提取子模块,被配置为对所述目标声音进行特征提取,得到目标声音特征信息;

匹配子模块,被配置为将所述提取子模块提取到的所述目标声音特征信息与预存的经训练得到的多个声音模板进行匹配,并根据匹配结果得到所述目标声音的发声原因,所述多个声音模板用于识别不同原因引起的目标声音。

11. 根据权利要求10所述的对象监控装置,其特征在于,所述装置还包括:

接收模块,被配置为在所述第一推送模块向移动终端推送发声信息之后,接收所述移动终端发送的视频获取指令;

第二推送模块,被配置为根据所述接收模块接收的所述视频获取指令向所述移动终端推送视频信息,以使所述移动终端显示所述视频信息。

12. 根据权利要求11所述的对象监控装置,其特征在于,所述装置还包括:

获取模块,被配置为在所述第二推送模块根据所述视频获取指令向所述移动终端推送视频信息之前,获取所述目标声音的发声位置;

转向采集模块,被配置为控制摄像头转向所述获取模块获取的所述发声位置对应的方向,并采集所述视频信息。

13. 根据权利要求10所述的对象监控装置,其特征在于,所述提取子模块包括:

压缩单元,被配置为对所述目标声音进行压缩处理,得到压缩声音信息;

提取单元,被配置为对所述压缩单元得到的所述压缩声音信息进行特征提取,得到所述目标声音特征信息。

14. 根据权利要求10所述的对象监控装置,其特征在于,所述判断模块包括:

发送子模块,被配置为向服务器发送所述目标声音;

接收子模块,被配置为接收所述服务器返回的所述目标声音的发声原因。

15. 根据权利要求11所述的对象监控装置,其特征在于,所述装置还包括:

检测广播模块,被配置为若检测到所述监控设备未接入网络,则广播报文,以使所述移动终端显示所述监控设备未接入网络的提示信息;

接收连接模块,被配置为接收所述移动终端广播的路由器的名称和密码,并向所述路由器发送连接请求,以连接到所述路由器上。

16. 根据权利要求10所述的对象监控装置,其特征在于,所述装置还包括:

训练模块,被配置为针对每个声音模板,对当前声音模板进行训练。

17. 根据权利要求16所述的对象监控装置,其特征在于,所述训练模块包括:

采集子模块,被配置为采集预设数量的目标声音特征样本;

输入子模块,被配置为将所述采集子模块采集到的所述预设数量的目标声音特征样本输入当前声音模板,得到所述预设数量的目标声音特征样本相对于当前声音模板的失真距离,所述失真距离用于表示所述预设数量的目标声音特征样本与通过所述当前声音模板输出的目标声音之间的差距;

第一训练子模块,被配置为若所述输入子模块得到的所述失真距离小于预设距离,则停止对所述当前声音模板进行训练;

第二训练子模块,被配置为若所述输入子模块得到的所述失真距离大于等于预设距离,则继续调用所述训练模块对所述当前声音模板进行训练,直至所述失真距离小于预设距离。

18. 根据权利要求10至17任一项所述的对象监控装置,其特征在于,所述发声信息还携带有所述目标声音。

19. 一种对象监控装置,其特征在于,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为:

当采集到的声音中存在目标声音时,判断所述目标声音的发声原因;

向移动终端推送发声信息,所述发声信息携带有所述发声原因;

所述判断所述目标声音的发声原因,包括:

对所述目标声音进行特征提取,得到目标声音特征信息;

将所述目标声音特征信息与预存的经训练得到的多个声音模板进行匹配,并根据匹配结果得到所述目标声音的发声原因,所述多个声音模板用于识别不同原因引起的目标声音。

## 对象监控方法及装置

### 技术领域

[0001] 本公开涉及移动互联网技术领域,尤其涉及一种对象监控方法及装置。

### 背景技术

[0002] 随着摄像技术的发展,出现了智能摄像机,智能摄像机是采用当今最新的科技成果、最精尖的制造工艺,综合多年的宝贵经验研制成功的摄像机。

[0003] 随着网络化、数字化、高清化的广泛应用,人们对智能化的需求日益增多,智能摄像机的应用越来越广,其种种特性使之适用于各行各业活动目标的监控。例如,可以对智能大厦、城市道路、电力部门、机场和车站等进行监控等。

[0004] 由此可见,智能摄像机可用于监控各种对象,因此,如何更好地利用智能摄像机为用户提供监控服务,是目前急需解决的技术问题之一。

### 发明内容

[0005] 为克服相关技术中存在的问题,本公开提供一种对象监控方法及装置。

[0006] 根据本公开实施例的第一方面,提供一种对象监控方法,包括:

[0007] 当采集到的声音中存在目标声音时,判断所述目标声音的发声原因;

[0008] 向移动终端推送发声信息,所述发声信息携带有所述发声原因。

[0009] 在一实施例中,在所述向移动终端推送发声信息之后,所述方法还包括:

[0010] 接收所述移动终端发送的视频获取指令;

[0011] 根据所述视频获取指令向所述移动终端推送视频信息,以使所述移动终端显示所述视频信息。

[0012] 在一实施例中,在所述根据所述视频获取指令向所述移动终端推送视频信息之前,所述方法还包括:

[0013] 获取所述目标声音的发声位置;

[0014] 控制摄像头转向所述发声位置对应的方向,并采集所述视频信息。

[0015] 在一实施例中,所述判断所述目标声音的发声原因,包括:

[0016] 对所述目标声音进行特征提取,得到目标声音特征信息;

[0017] 将所述目标声音特征信息与预存的多个声音模板进行匹配,并根据匹配结果得到所述目标声音的发声原因,所述多个声音模板用于识别不同原因引起的目标声音。

[0018] 在一实施例中,所述对所述目标声音进行特征提取,得到目标声音特征信息,包括:

[0019] 对所述目标声音进行压缩处理,得到压缩声音信息;

[0020] 对所述压缩声音信息进行特征提取,得到所述目标声音特征信息。

[0021] 在一实施例中,所述判断所述目标声音的发声原因,包括:

[0022] 向服务器发送所述目标声音;

[0023] 接收所述服务器返回的所述目标声音的发声原因。

- [0024] 在一实施例中,所述方法还包括:
- [0025] 若检测到所述当前监控设备未接入网络,则广播报文,以使所述移动终端显示所述当前监控设备未接入网络的提示信息;
- [0026] 接收所述移动终端广播的路由器的名称和密码,并向所述路由器发送连接请求,以连接到所述路由器上。
- [0027] 在一实施例中,所述方法还包括:
- [0028] 针对每个声音模板,对当前声音模板进行训练。
- [0029] 在一实施例中,所述对当前声音模板进行训练,包括:
- [0030] 采集预设数量的目标声音特征样本;
- [0031] 将所述预设数量的目标声音特征样本输入当前声音模板,得到所述预设数量的目标声音特征样本相对于当前声音模板的失真距离,所述失真距离用于表示所述预设数量的目标声音特征样本与通过所述当前声音模板输出的目标声音之间的差距;
- [0032] 若所述失真距离小于预设距离,则停止对所述当前声音模板进行训练;
- [0033] 若所述失真距离大于等于预设距离,则继续对所述当前声音模板进行训练,直至所述失真距离小于预设距离。
- [0034] 在一实施例中,所述发声信息还携带有所述目标声音。
- [0035] 根据本公开实施例的第二方面,提供一种对象监控装置,包括:
- [0036] 判断模块,被配置为当采集到的声音中存在目标声音时,判断所述目标声音的发声原因;
- [0037] 第一推送模块,被配置为向移动终端推送发声信息,所述发声信息携带有所述判断模块判断出的所述发声原因。
- [0038] 在一实施例中,所述装置还包括:
- [0039] 接收模块,被配置为在所述第一推送模块向移动终端推送发声信息之后,接收所述移动终端发送的视频获取指令;
- [0040] 第二推送模块,被配置为根据所述接收模块接收的所述视频获取指令向所述移动终端推送视频信息,以使所述移动终端显示所述视频信息。
- [0041] 在一实施例中,所述装置还包括:
- [0042] 获取模块,被配置为在所述第二推送模块根据所述视频获取指令向所述移动终端推送视频信息之前,获取所述目标声音的发声位置;
- [0043] 转向采集模块,被配置为控制摄像头转向所述获取模块获取的所述发声位置对应的方向,并采集所述视频信息。
- [0044] 在一实施例中,所述判断模块包括:
- [0045] 提取子模块,被配置为对所述目标声音进行特征提取,得到目标声音特征信息;
- [0046] 匹配子模块,被配置为将所述提取子模块提取到的所述目标声音特征信息与预存的多个声音模板进行匹配,并根据匹配结果得到所述目标声音的发声原因,所述多个声音模板用于识别不同原因引起的目标声音。
- [0047] 在一实施例中,所述提取子模块包括:
- [0048] 压缩单元,被配置为对所述目标声音进行压缩处理,得到压缩声音信息;
- [0049] 提取单元,被配置为对所述压缩单元得到的所述压缩声音信息进行特征提取,得

到所述目标声音特征信息。

[0050] 在一实施例中,所述判断模块包括:

[0051] 发送子模块,被配置为向服务器发送所述目标声音;

[0052] 接收子模块,被配置为接收所述服务器返回的所述目标声音的发声原因。

[0053] 在一实施例中,所述装置还包括:

[0054] 检测广播模块,被配置为若检测到所述当前监控设备未接入网络,则广播报文,以使所述移动终端显示所述当前监控设备未接入网络的提示信息;

[0055] 接收连接模块,被配置为接收所述移动终端广播的路由器的名称和密码,并向所述路由器发送连接请求,以连接到所述路由器上。

[0056] 在一实施例中,所述装置还包括:

[0057] 训练模块,被配置为针对每个声音模板,对当前声音模板进行训练。

[0058] 在一实施例中,所述训练模块包括:

[0059] 采集子模块,被配置为采集预设数量的目标声音特征样本;

[0060] 输入子模块,被配置为将所述采集子模块采集到的所述预设数量的目标声音特征样本输入当前声音模板,得到所述预设数量的目标声音特征样本相对于当前声音模板的失真距离,所述失真距离用于表示所述预设数量的目标声音特征样本与通过所述当前声音模板输出的目标声音之间的差距;

[0061] 第一训练子模块,被配置为若所述输入子模块得到的所述失真距离小于预设距离,则停止对所述当前声音模板进行训练;

[0062] 第二训练子模块,被配置为若所述输入子模块得到的所述失真距离大于等于预设距离,则继续调用所述训练模块对所述当前声音模板进行训练,直至所述失真距离小于预设距离。

[0063] 在一实施例中,所述发声信息还携带有所述目标声音。

[0064] 根据本公开实施例的第三方面,提供一种对象监控装置,包括:

[0065] 处理器;

[0066] 用于存储处理器可执行指令的存储器;

[0067] 其中,处理器被配置为:

[0068] 当采集到的声音中存在目标声音时,判断所述目标声音的发声原因;

[0069] 向移动终端推送发声信息,所述发声信息携带有所述发声原因。

[0070] 本公开的实施例提供的技术方案可以包括以下有益效果:通过判断采集到的目标声音的发声原因,并向移动终端推送携带有发声原因的发声信息,使得用户可以获得被监控对象的发声原因,达到很好的监控目的。

[0071] 通过向移动终端推送视频信息,使得移动终端可以显示该视频信息,从而方便用户远程查看该视频信息。

[0072] 通过获取目标声音的发声位置来采集视频信息,实现方式简单。

[0073] 通过将提取到的目标声音特征信息与预存的已经训练好的多个声音模板进行匹配,得到目标声音的发声原因,实现方式简单,得到的发声原因准确率高。

[0074] 通过对目标声音进行压缩处理,达到降低系统的运算和存储负担以及去除目标声音相关性的目的。

[0075] 通过与服务器进行交互也可以判断目标声音的发声原因,由此可见,可以通过多种方式判断目标声音的发声原因。

[0076] 通过与移动终端进行交互以接入路由器,从而为二者建立绑定关系提供条件,进而为后续进行监控提供条件。

[0077] 通过对声音模板进行训练,为判断目标声音的发声原因提供条件。

[0078] 通过将预设数量的目标声音特征样本输入当前声音模板,得到失真距离,并将失真距离与预设距离进行比较,来完成声音模板的训练过程,实现方式简单,训练出的声音模板可以准确地识别出目标声音的发声原因。

[0079] 通过描述发声信息还携带有目标声音,展示了向移动终端推送的发声信息的多样性。

[0080] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本公开。

### 附图说明

[0081] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本发明的实施例,并与说明书一起用于解释本发明的原理。

[0082] 图1是根据一示例性实施例示出的一种对象监控方法的流程图。

[0083] 图2是根据一示例性实施例示出的另一种对象监控方法的流程图。

[0084] 图3是根据一示例性实施例示出的一种对象监控方法的场景图。

[0085] 图4是根据一示例性实施例示出的一种判断目标声音的发声原因的流程图。

[0086] 图5是根据一示例性实施例示出的一种对当前声音模板进行训练的流程图。

[0087] 图6是根据一示例性实施例示出的一种对象监控装置的框图。

[0088] 图7是根据一示例性实施例示出的另一种对象监控装置的框图。

[0089] 图8是根据一示例性实施例示出的另一种对象监控装置的框图。

[0090] 图9是根据一示例性实施例示出的另一种对象监控装置的框图。

[0091] 图10是根据一示例性实施例示出的另一种对象监控装置的框图。

[0092] 图11是根据一示例性实施例示出的另一种对象监控装置的框图。

[0093] 图12是根据一示例性实施例示出的另一种对象监控装置的框图。

[0094] 图13是根据一示例性实施例示出的另一种对象监控装置的框图。

[0095] 图14是根据一示例性实施例示出的另一种对象监控装置的框图。

[0096] 图15是根据一示例性实施例示出的一种适用于对象监控装置的框图。

### 具体实施方式

[0097] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[0098] 图1是根据一示例性实施例示出的一种对象监控方法的流程图,如图1所示,该对象监控方法可应用于监控设备上,该监控设备可以包括但不限于智能摄像机,该方法包



括以下步骤S101-S102:

[0099] 在步骤S101中,当采集到的声音中存在目标声音时,判断该目标声音的发声原因。

[0100] 其中,可以通过内置在摄像机中的麦克风采集被监控对象的声音,被监控对象可以包括但不限于儿童、老人等,目标声音可以指哭声等。

[0101] 在该实施例中,智能摄像机在采集到目标声音例如哭声后,可以通过提取哭声特征信息,并将提取的哭声特征信息与已经训练好的多个声音模板进行匹配来得到哭声的发声原因。

[0102] 另外,智能摄像机也可以通过与服务器进行交互来获得哭声的发声原因。例如,智能摄像机可以向服务器发送目标声音,服务器在接收到目标声音后也可以通过提取哭声特征信息,并将提取的哭声特征信息与已经训练好的多个声音模板进行匹配来得到哭声的发声原因,然后向智能摄像机返回哭声的发声原因。

[0103] 在步骤S102中,向移动终端推送发声信息,发声信息携带有发声原因。

[0104] 在该实施例中,在判断目标声音的发声原因之后,可以向移动终端例如手机推送携带有发声原因的发声信息,以提示用户儿童哭泣的原因。

[0105] 另外,该发声信息中也可以携带有目标声音,例如可以向手机推送儿童的哭声及哭泣的原因。

[0106] 上述对象监控方法实施例,通过判断采集到的目标声音的发声原因,并向移动终端推送携带有发声原因的发声信息,使得用户可以获得被监控对象的发声原因,达到很好的监控目的。

[0107] 图2是根据一示例性实施例示出的另一种对象监控方法的流程图,如图2所示,在上述步骤S102之后,该方法还可以包括如下步骤:

[0108] 在步骤S103中,接收移动终端发送的视频获取指令。

[0109] 在该实施例中,当前监控设备例如内置有无线保真(WIFI)芯片的智能摄像机在向移动终端推送发声信息之后,如果用户希望查看被监控对象的视频信息,例如儿童的视频信息,则可以通过移动终端上的应用程序(APP)向智能摄像机发送视频获取指令。

[0110] 在该实施例中,智能摄像机与上述APP可以通过同一用户账号建立绑定关系。而在二者建立绑定关系之前,智能摄像机可以与该移动终端接入到同一路由器上,智能摄像机接入路由器的过程可以为:

[0111] 智能摄像机检测到自己未接入网络,则广播报文,以使移动终端显示智能摄像机未接入网络的提示信息,用户在看到该提示信息后,可以通过移动终端向智能摄像机广播自己所连接的路由器的名称和密码,智能摄像机在接收到路由器的名称和密码后,可以向该路由器发送连接请求,以连接到该路由器上。

[0112] 由于智能摄像机和移动终端连接到同一路由器即位于同一局域网内,因此,APP可以通过扫描获得智能摄像机的绑定状态,并显示智能摄像机的绑定状态。其中,绑定状态包括已绑定和未绑定两种状态。

[0113] 当智能摄像机的绑定状态为未绑定时,如果用户想控制该智能摄像机,则可以触发绑定操作,然后该APP向服务器发送携带智能摄像机的标识和用户账号例如小米账号的请求,这样,服务器可以根据该请求将智能摄像机与用户账号进行绑定。

[0114] 当该智能摄像机与移动终端APP通过同一用户账号绑定后,用户可以通过该移动

终端APP向智能摄像机发送控制指令例如视频获取指令。

[0115] 在步骤S104中,根据该视频获取指令向移动终端推送视频信息,以使移动终端显示视频信息。

[0116] 在该实施例中,智能摄像机可以获取目标声音的发声位置,并控制摄像头转向发声位置对应的方向,然后采集视频信息。该视频信息采集方式简单,易于实现。

[0117] 在智能摄像机采集好视频信息之后,若接收到来自移动终端的视频获取指令,则可以向该移动终端推送视频信息,以方便用户远程查看被监控对象例如儿童的视频信息。

[0118] 下面结合图3对本公开进行示例性说明,如图3所示,智能摄像机31通过内置的麦克风采集到声音中存在哭声后,通过提取哭声特征信息,并将提取的哭声特征信息与已经训练好的多个声音模板进行匹配来得到哭声的发声原因。由于智能摄像机31和手机32连接到同一路由器33即位于同一局域网内,因此,智能摄像机31与位于手机32中的APP可以通过当前用户账号建立绑定关系,故智能摄像机31可以通过服务器34向手机32推送发声信息,该发声信息携带有哭声的发声原因,此时若用户希望查看智能摄像机31拍摄的视频信息,则可以通过手机32向智能摄像机31发送视频获取指令,智能摄像机31在接收到该视频获取指令后,可以向手机32推送获取的视频信息,使得用户可以远程查看视频信息。

[0119] 上述实施例,通过向移动终端推送视频信息,使得移动终端可以显示该视频信息,从而方便用户远程查看该视频信息。

[0120] 图4是根据一示例性实施例示出的一种判断目标声音的发声原因的流程图,该实施例从监控设备例如智能摄像机侧进行描述,也可以从服务器侧进行描述,即监控设备和服务器均可以通过以下方式判断目标声音的发声原因,如图4所示,该方法包括如下步骤:

[0121] 在步骤S401中,针对每个声音模板,对当前声音模板进行训练。

[0122] 在该实施例中,首先需要对多个声音模板中的每个声音模板进行训练,以达到将目标声音输入声音模板后,可以识别出发声原因。其中,多个声音模板用于识别不同原因引起的目标声音。

[0123] 其中,如图5所示,对当前声音模板进行训练的过程可以包括如下步骤:

[0124] 在步骤S4011中,采集预设数量的目标声音特征样本。

[0125] 在步骤S4012中,将预设数量的目标声音特征样本输入当前声音模板,得到预设数量的目标声音特征样本相对于当前声音模板的失真距离。

[0126] 失真距离用于表示预设数量的目标声音特征样本与通过当前声音模板输出的目标声音之间的差距。

[0127] 在步骤S4013中,判断失真距离是否小于预设距离,若失真距离小于预设距离,则停止对当前声音模板进行训练,若失真距离大于等于预设距离,则转向执行步骤S4011。

[0128] 其中,失真距离越小,则表明预设数量的目标声音特征样本与通过当前声音模板输出的目标声音之间的差距越小。

[0129] 当失真距离小于预设距离,则完成对当前声音模板的训练,否则,继续训练,直至失真距离小于预设距离。

[0130] 在步骤S402中,对目标声音进行特征提取,得到目标声音特征信息。

[0131] 在该实施例中,若目标声音例如哭声数据量比较大,则可以先对目标声音进行压缩处理,以降低系统的运算和存储负担,另外,压缩后的目标声音可以去除目标声音的相关

性,例如去除哭声的相关性,以达到用较少的特征信息来表达哭声信号的目的。当然,若目标声音数据量不大且已去除相关性,则不需要进行压缩处理。其中,可以采用已有的压缩技术例如无损压缩技术对目标声音进行压缩处理。

[0132] 在该实施例中,在对目标声音进行压缩处理,得到压缩声音信息之后,可以对压缩声音信息进行特征提取,得到目标声音特征信息。

[0133] 在步骤S403中,将目标声音特征信息与预存的多个声音模板进行匹配,并根据匹配结果得到目标声音的发声原因。

[0134] 在该实施例中,可以将提取到的目标声音特征信息与预存的已经训练好的多个声音模板进行匹配,从而得到目标声音的发声原因。

[0135] 上述实施例,通过将提取到的目标声音特征信息与预存的已经训练好的多个声音模板进行匹配,得到目标声音的发声原因,实现方式简单,得到的发声原因准确率高。

[0136] 与前述对象监控方法实施例相对应,本公开还提供了对象监控装置实施例。

[0137] 图6是根据一示例性实施例示出的一种对象监控装置的框图,如图6所示,该对象监控装置包括:判断模块61和第一推送模块62。

[0138] 判断模块61被配置为当采集到的声音中存在目标声音时,判断目标声音的发声原因。

[0139] 其中,判断模块61可以通过内置在摄像机中的麦克风采集被监控对象的声音,被监控对象可以包括但不局限于儿童、老人等,目标声音可以指哭声等。

[0140] 在该实施例中,智能摄像机在采集到目标声音例如哭声后,可以通过提取哭声特征信息,并将提取的哭声特征信息与已经训练好的多个声音模板进行匹配来得到哭声的发声原因。

[0141] 另外,智能摄像机也可以通过与服务器进行交互来获得哭声的发声原因。例如,智能摄像机可以向服务器发送目标声音,服务器在接收到目标声音后也可以通过提取哭声特征信息,并将提取的哭声特征信息与已经训练好的多个声音模板进行匹配来得到哭声的发声原因,然后向智能摄像机返回哭声的发声原因。

[0142] 第一推送模块62被配置为向移动终端推送发声信息,发声信息携带有判断模块61判断出的发声原因。

[0143] 在该实施例中,在判断目标声音的发声原因之后,可以向移动终端例如手机推送携带有发声原因的发声信息,以提示用户儿童哭泣的原因。

[0144] 另外,该发声信息中也可以携带有目标声音,例如可以向手机推送儿童的哭声及哭泣的原因。

[0145] 如图6所示的装置用于实现上述如图1所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0146] 上述实施例,通过判断采集到的目标声音的发声原因,并向移动终端推送携带有发声原因的发声信息,使得用户可以获得被监控对象的发声原因,达到很好的监控目的。

[0147] 图7是根据一示例性实施例示出的另一种对象监控装置的框图,如图7所示,在上述图6所示实施例的基础上,该装置还可包括:接收模块63和第二推送模块64。

[0148] 接收模块63,被配置为在第一推送模块62向移动终端推送发声信息之后,接收移动终端发送的视频获取指令。

[0149] 在该实施例中,当前监控设备例如内置有无线保真(WIFI)芯片的智能摄像机在向移动终端推送发声信息之后,如果用户希望查看被监控对象的视频信息,例如儿童的视频信息,则可以通过移动终端上的应用程序(APP)向智能摄像机发送视频获取指令。

[0150] 当该智能摄像机与移动终端APP通过同一用户账号绑定后,用户可以通过该移动终端APP向智能摄像机发送控制指令例如视频获取指令。

[0151] 第二推送模块64,被配置为根据接收模块63接收的视频获取指令向移动终端推送视频信息,以使移动终端显示视频信息。

[0152] 如图7所示的装置用于实现上述如图2所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0153] 上述实施例,通过向移动终端推送视频信息,使得移动终端可以显示该视频信息,从而方便用户远程查看该视频信息。

[0154] 图8是根据一示例性实施例示出的另一种对象监控装置的框图,如图8所示,在上述图7所示实施例的基础上,该装置还可以包括:获取模块65和转向采集模块66。

[0155] 获取模块65被配置为在第二推送模块64根据视频获取指令向移动终端推送视频信息之前,获取目标声音的发声位置。

[0156] 转向采集模块66被配置为控制摄像头转向获取模块65获取的发声位置对应的方向,并采集视频信息。

[0157] 在该实施例中,智能摄像机可以获取目标声音的发声位置,并控制摄像头转向发声位置对应的方向,然后采集视频信息。该视频信息采集方式简单,易于实现。

[0158] 如图8所示的装置用于实现上述如图2所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0159] 上述实施例,通过获取目标声音的发声位置来采集视频信息,实现方式简单。

[0160] 图9是根据一示例性实施例示出的另一种对象监控装置的框图,如图9所示,在上述图6所示实施例的基础上,判断模块61可包括:提取子模块611和匹配子模块612。

[0161] 提取子模块611被配置为对目标声音进行特征提取,得到目标声音特征信息。

[0162] 匹配子模块612被配置为将提取子模块611提取到的目标声音特征信息与预存的多个声音模板进行匹配,并根据匹配结果得到目标声音的发声原因,多个声音模板用于识别不同原因引起的目标声音。

[0163] 在该实施例中,可以将提取到的目标声音特征信息与预存的已经训练好的多个声音模板进行匹配,从而得到目标声音的发声原因。

[0164] 如图9所示的装置用于实现上述如图4所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0165] 上述实施例,通过将提取到的目标声音特征信息与预存的已经训练好的多个声音模板进行匹配,得到目标声音的发声原因,实现方式简单,得到的发声原因准确率高。

[0166] 图10是根据一示例性实施例示出的另一种对象监控装置的框图,如图10所示,在上述图9所示实施例的基础上,提取子模块611可包括:压缩单元6111和提取单元6112。

[0167] 压缩单元6111被配置为对目标声音进行压缩处理,得到压缩声音信息。

[0168] 提取单元6112被配置为对压缩单元6111得到的压缩声音信息进行特征提取,得到目标声音特征信息。

[0169] 如图10所示的装置用于实现上述如图4所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0170] 在该实施例中,若目标声音例如哭声数据量比较大,则可以先对目标声音进行压缩处理,以降低系统的运算和存储负担,另外,压缩后的目标声音可以去除目标声音的相关性,例如去除哭声的相关性,以达到用较少的特征信息来表达哭声信号的目的。当然,若目标声音数据量不大且已去除相关性,则不需要进行压缩处理。其中,可以采用已有的压缩技术例如无损压缩技术对目标声音进行压缩处理。

[0171] 上述实施例,通过对目标声音进行压缩处理,达到降低系统的运算和存储负担以及去除目标声音相关性的目的。

[0172] 图11是根据一示例性实施例示出的另一种对象监控装置的框图,如图11所示,在上述图6所示实施例的基础上,判断模块61可包括:发送子模块613和接收子模块614。

[0173] 发送子模块613被配置为向服务器发送目标声音。

[0174] 接收子模块614被配置为接收服务器返回的目标声音的发声原因。

[0175] 对象监控装置例如智能摄像机可以通过与服务器进行交互来获得哭声的发声原因。例如,智能摄像机可以向服务器发送目标声音,服务器在接收到目标声音后也可以通过提取哭声特征信息,并将提取的哭声特征信息与已经训练好的多个声音模板进行匹配来得到哭声的发声原因,然后向智能摄像机返回哭声的发声原因。

[0176] 如图11所示的装置用于实现上述如图1所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0177] 上述实施例,通过与服务器进行交互也可以判断目标声音的发声原因,由此可见,可以通过多种方式判断目标声音的发声原因。

[0178] 图12是根据一示例性实施例示出的另一种对象监控装置的框图,如图12所示,在上述图7所示实施例的基础上,该装置还可包括:检测广播模块67和接收连接模块68。

[0179] 检测广播模块67被配置为若检测到当前监控设备未接入网络,则广播报文,以使移动终端显示当前监控设备未接入网络的提示信息。

[0180] 接收连接模块68被配置为接收移动终端广播的路由器的名称和密码,并向路由器发送连接请求,以连接到路由器上。

[0181] 在该实施例中,智能摄像机与上述APP可以通过同一用户账号建立绑定关系。而在二者建立绑定关系之前,智能摄像机可以与该移动终端接入到同一路由器上,智能摄像机接入路由器的过程可以为:

[0182] 智能摄像机检测到自己未接入网络,则广播报文,以使移动终端显示智能摄像机未接入网络的提示信息,用户在看到该提示信息后,可以通过移动终端向智能摄像机广播自己所连接的路由器的名称和密码,智能摄像机在接收到路由器的名称和密码后,可以向该路由器发送连接请求,以连接到该路由器上。

[0183] 由于智能摄像机和移动终端连接到同一路由器即位于同一局域网内,因此,APP可以通过扫描获得智能摄像机的绑定状态,并显示智能摄像机的绑定状态。其中,绑定状态包括已绑定和未绑定两种状态。

[0184] 当智能摄像机的绑定状态为未绑定时,如果用户想控制该智能摄像机,则可以触发绑定操作,然后该APP向服务器发送携带智能摄像机的标识和用户账号例如小米账号的

请求,这样,服务器可以根据该请求将智能摄像机与用户账号进行绑定。

[0185] 如图12所示的装置用于实现上述如图2所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0186] 上述实施例,通过与移动终端进行交互以接入路由器,从而为二者建立绑定关系提供条件,进而为后续进行监控提供条件。

[0187] 图13是根据一示例性实施例示出的另一种对象监控装置的框图,如图13所示,在上述图9所示实施例的基础上,该装置还可包括:训练模块69。

[0188] 训练模块69被配置为针对每个声音模板,对当前声音模板进行训练。

[0189] 在该实施例中,首先需要对多个声音模板中的每个声音模板进行训练,以达到将目标声音输入声音模板后,可以识别出发声原因。其中,多个声音模板用于识别不同原因引起的目标声音。

[0190] 如图13所示的装置用于实现上述如图4所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0191] 上述实施例,通过对声音模板进行训练,为判断目标声音的发声原因提供条件。

[0192] 图14是根据一示例性实施例示出的另一种对象监控装置的框图,如图14所示,在上述图13所示实施例的基础上,训练模块69可包括:采集子模块691、输入子模块692、第一训练子模块693和第二训练子模块694。

[0193] 采集子模块691被配置为采集预设数量的目标声音特征样本。

[0194] 输入子模块692被配置为将采集子模块691采集到的预设数量的目标声音特征样本输入当前声音模板,得到预设数量的目标声音特征样本相对于当前声音模板的失真距离,失真距离用于表示预设数量的目标声音特征样本与通过当前声音模板输出的目标声音之间的差距。

[0195] 第一训练子模块693被配置为若输入子模块692得到的失真距离小于预设距离,则停止对当前声音模板进行训练。

[0196] 第二训练子模块694被配置为若输入子模块692得到的失真距离大于等于预设距离,则继续调用训练模块对当前声音模板进行训练,直至失真距离小于预设距离。

[0197] 其中,失真距离越小,则表明预设数量的目标声音特征样本与通过当前声音模板输出的目标声音之间的差距越小。

[0198] 当失真距离小于预设距离,则完成对当前声音模板的训练,否则,继续训练,直至失真距离小于预设距离。

[0199] 如图14所示的装置用于实现上述如图5所示的方法流程,涉及到的相关内容描述相同,此处不赘述。

[0200] 上述实施例,通过将预设数量的目标声音特征样本输入当前声音模板,得到失真距离,并将失真距离与预设距离进行比较,来完成声音模板的训练过程,实现方式简单,训练出的声音模板可以准确地识别出目标声音的发声原因。

[0201] 关于上述实施例中的装置,其中各个模块、子模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0202] 图15是根据一示例性实施例示出的一种适用于对象监控装置的框图。例如,装置1500可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗

设备,健身设备,个人数字助理,飞行器等。

[0203] 参照图15,装置1500可以包括以下一个或多个组件:处理组件1502,存储器1504,电源组件1506,多媒体组件1508,音频组件1510,输入/输出(I/O)的接口1512,传感器组件1514,以及通信组件1516。

[0204] 处理组件1502通常控制装置1500的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理元件1502可以包括一个或多个处理器1520来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件1502可以包括一个或多个模块,便于处理组件1502和其他组件之间的交互。例如,处理部件1502可以包括多媒体模块,以方便多媒体组件1508和处理组件1502之间的交互。

[0205] 存储器1504被配置为存储各种类型的数据以支持在设备1500的操作。这些数据的示例包括用于在装置1500上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器1504可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0206] 电源组件1506为装置1500的各种组件提供电力。电源组件1506可以包括电源管理系统,一个或多个电源,及其他与为装置1500生成、管理和分配电力相关联的组件。

[0207] 多媒体组件1508包括在所述装置1500和用户之间的提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件1508包括一个前置摄像头和/或后置摄像头。当设备1500处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0208] 音频组件1510被配置为输出和/或输入音频信号。例如,音频组件1510包括一个麦克风(MIC),当装置1500处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器1504或经由通信组件1516发送。在一些实施例中,音频组件1510还包括一个扬声器,用于输出音频信号。

[0209] I/O接口1512为处理组件1502和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0210] 传感器组件1514包括一个或多个传感器,用于为装置1500提供各个方面的状态评估。例如,传感器组件1514可以检测到设备1500的打开/关闭状态,组件的相对定位,例如所述组件为装置1500的显示器和小键盘,传感器组件1514还可以检测装置1500或装置1500一个组件的位置改变,用户与装置1500接触的存在或不存在,装置1500方位或加速/减速和装置1500的温度变化。传感器组件1514可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件1514还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件1514还可以包括加速度传感

器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0211] 通信组件1516被配置为便于装置1500和其他设备之间有线或无线方式的通信。装置1500可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信部件1516经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信部件1516还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0212] 在示例性实施例中,装置1500可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0213] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器1504,上述指令可由装置1500的处理器1520执行以完成上述方法。例如,所述非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0214] 本领域技术人员在考虑说明书及实践这里公开的公开后,将容易想到本公开的其它实施方案。本申请旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由下面的权利要求指出。

[0215] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。



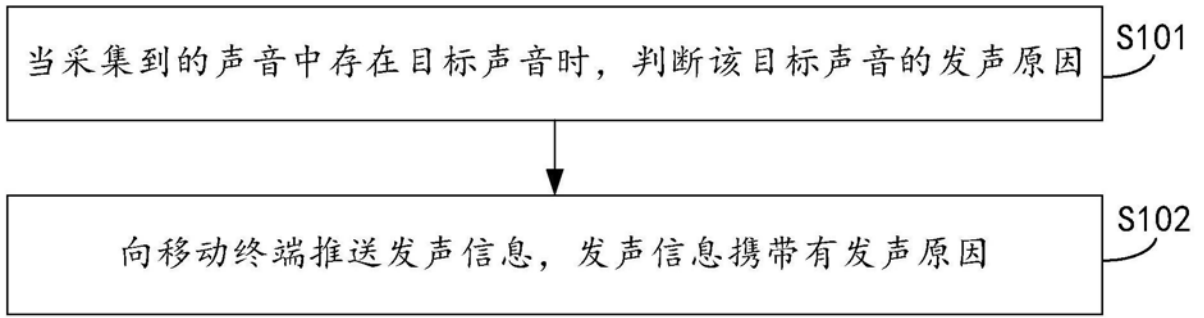


图1

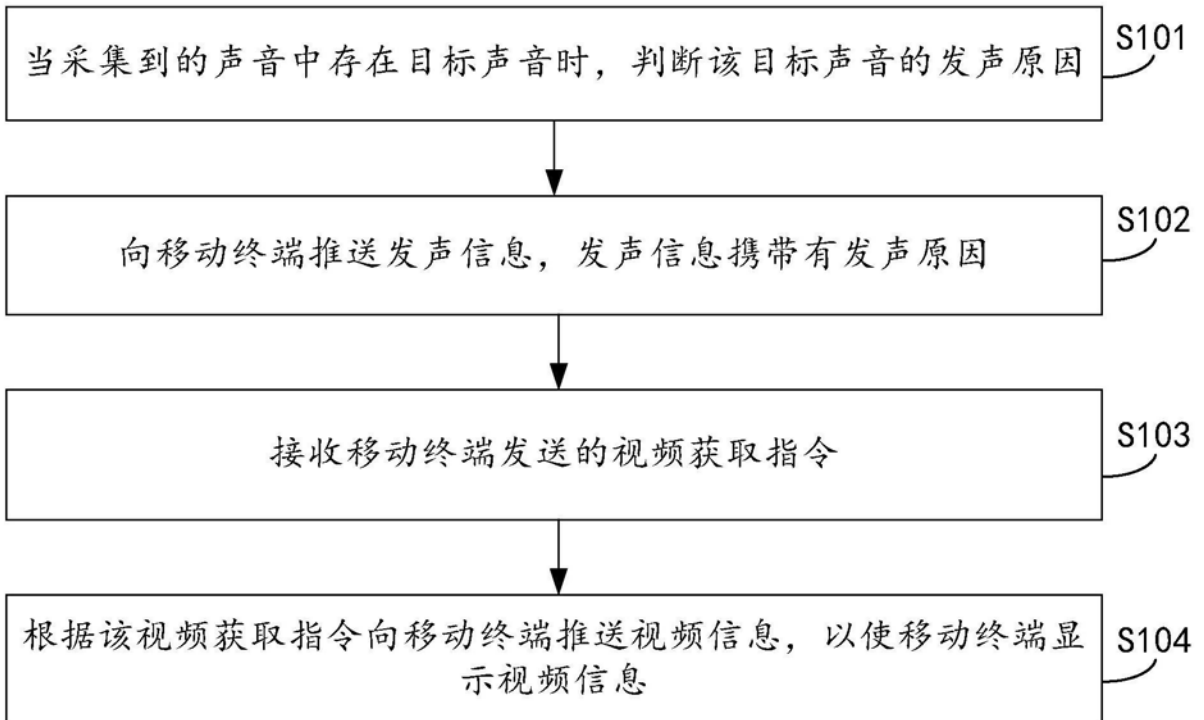


图2

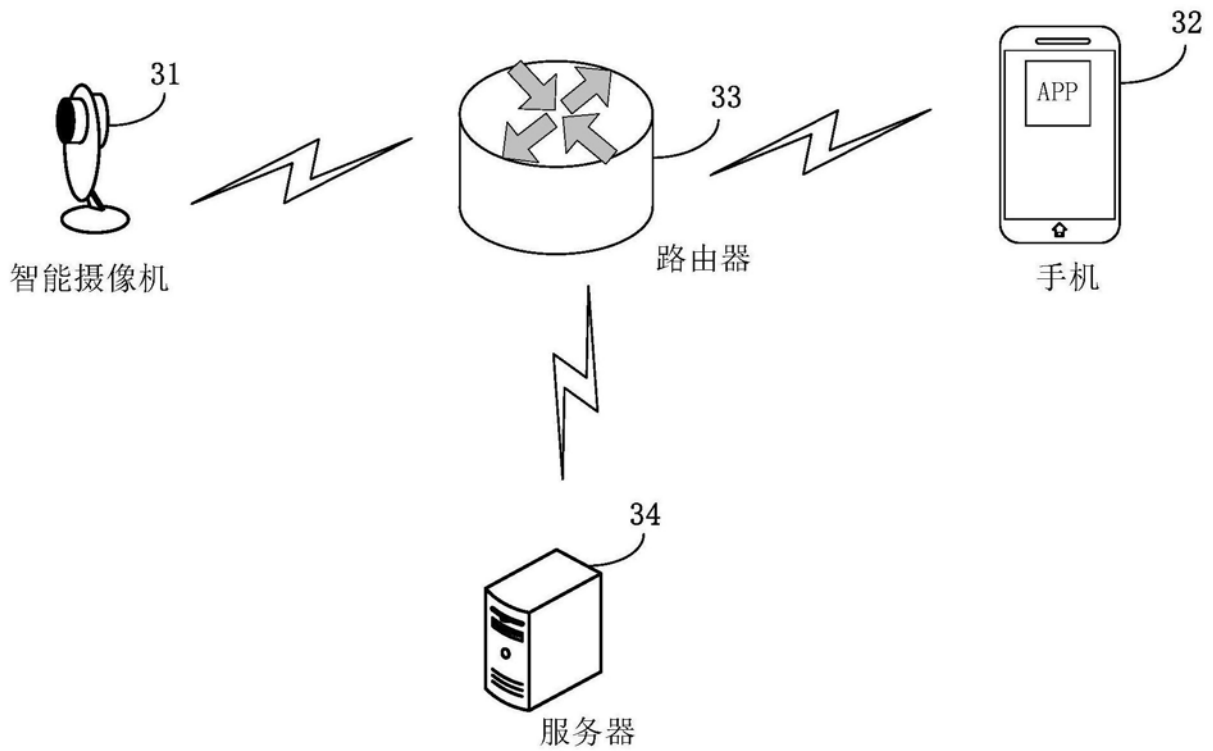


图3

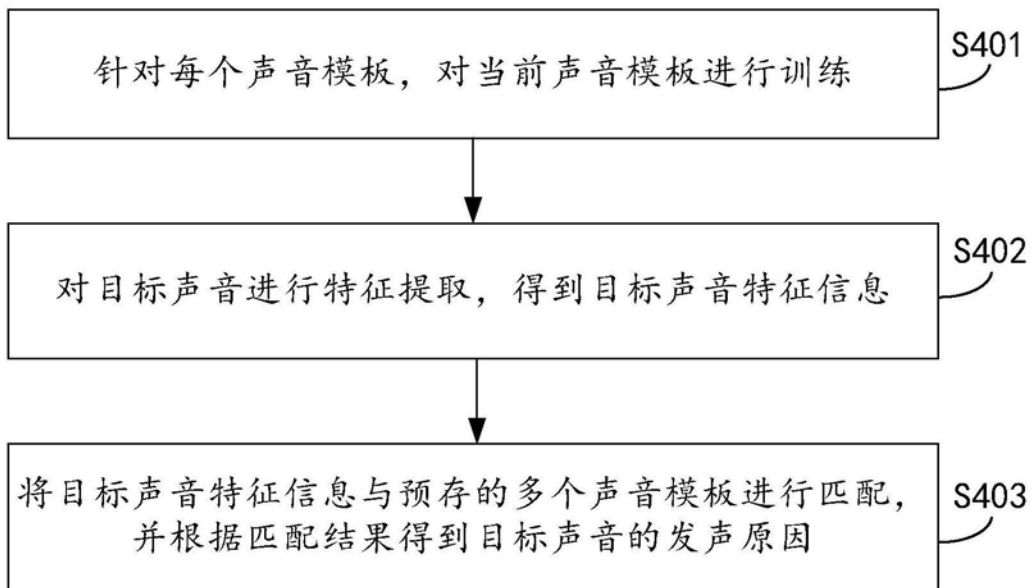


图4

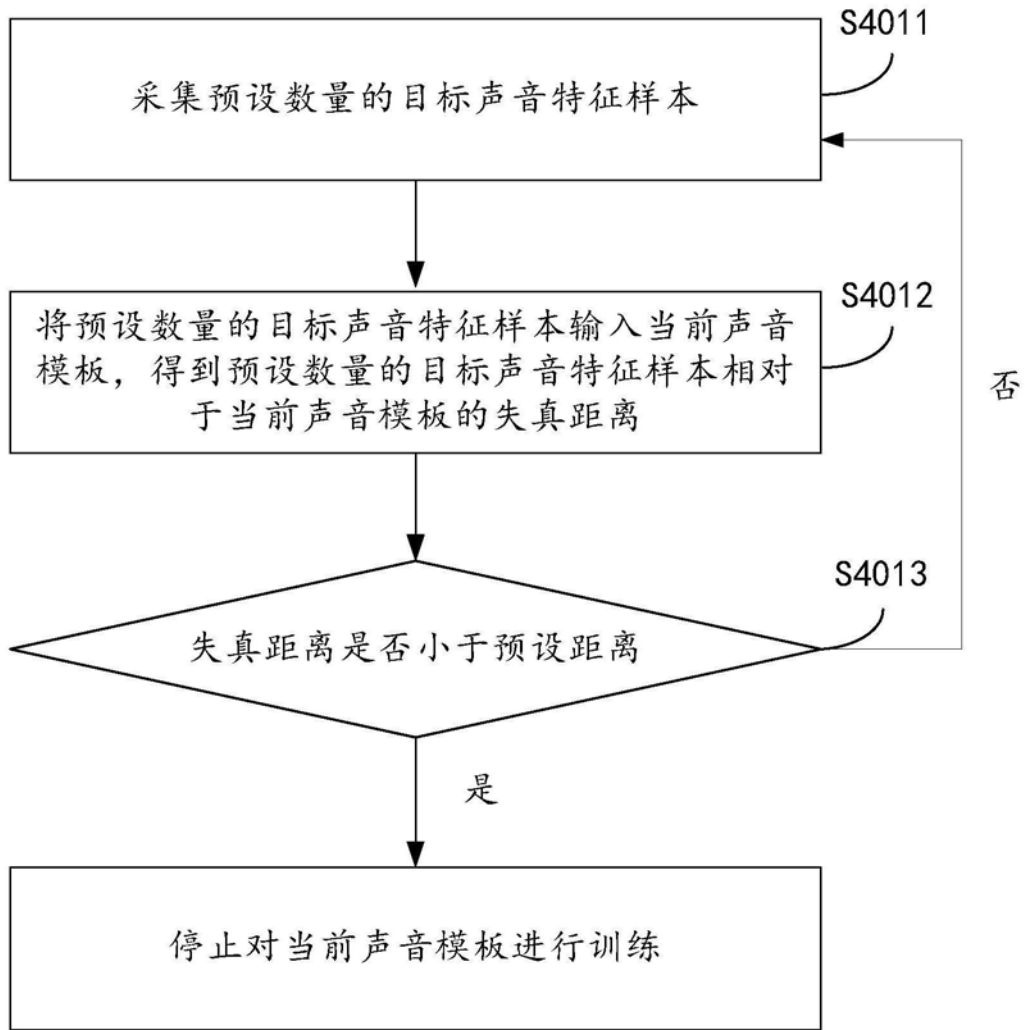


图5



图6

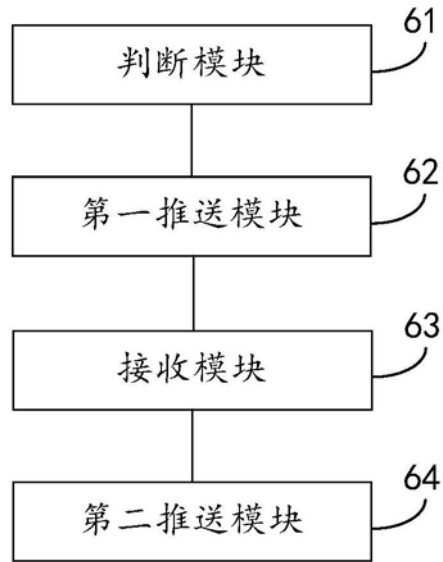


图7



图8

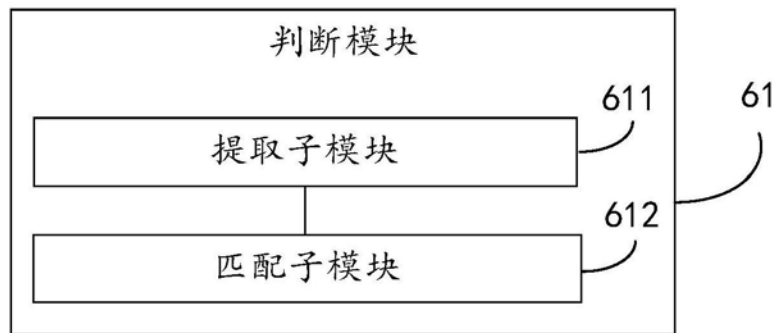


图9



图10

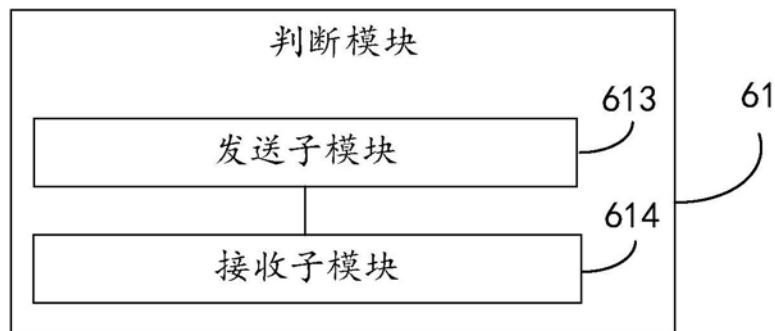


图11

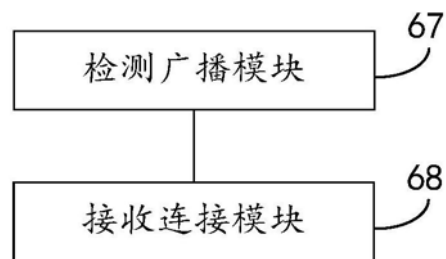


图12

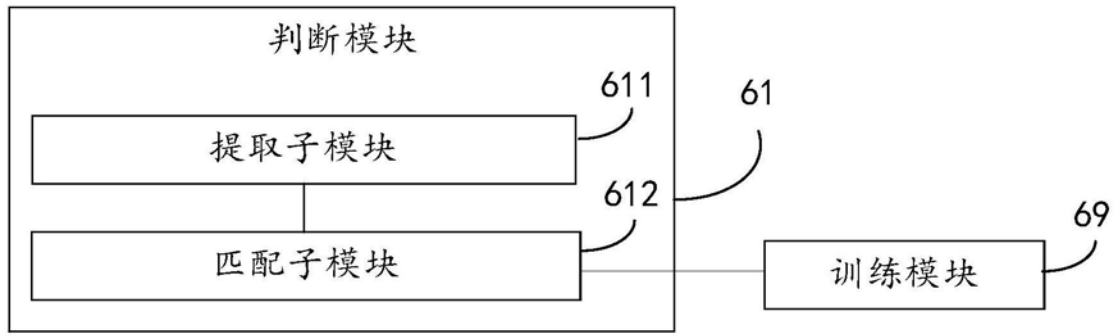


图13

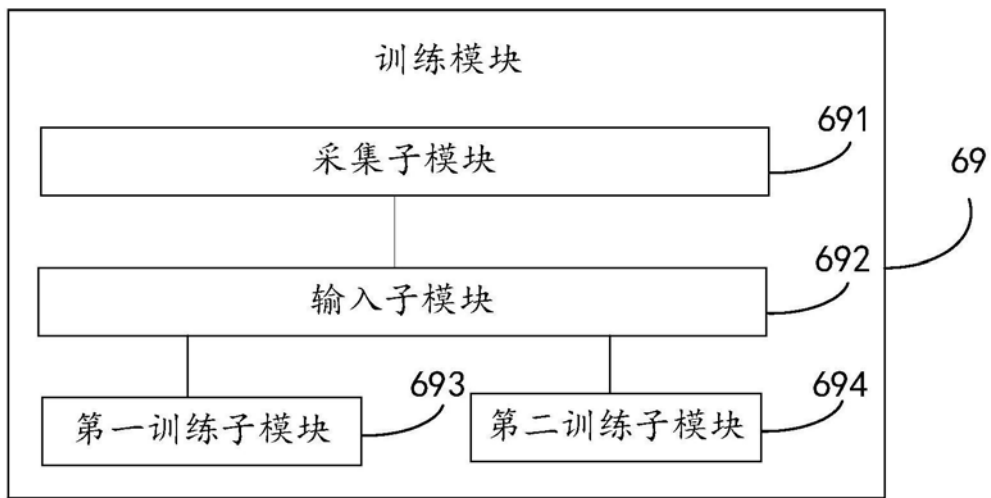


图14

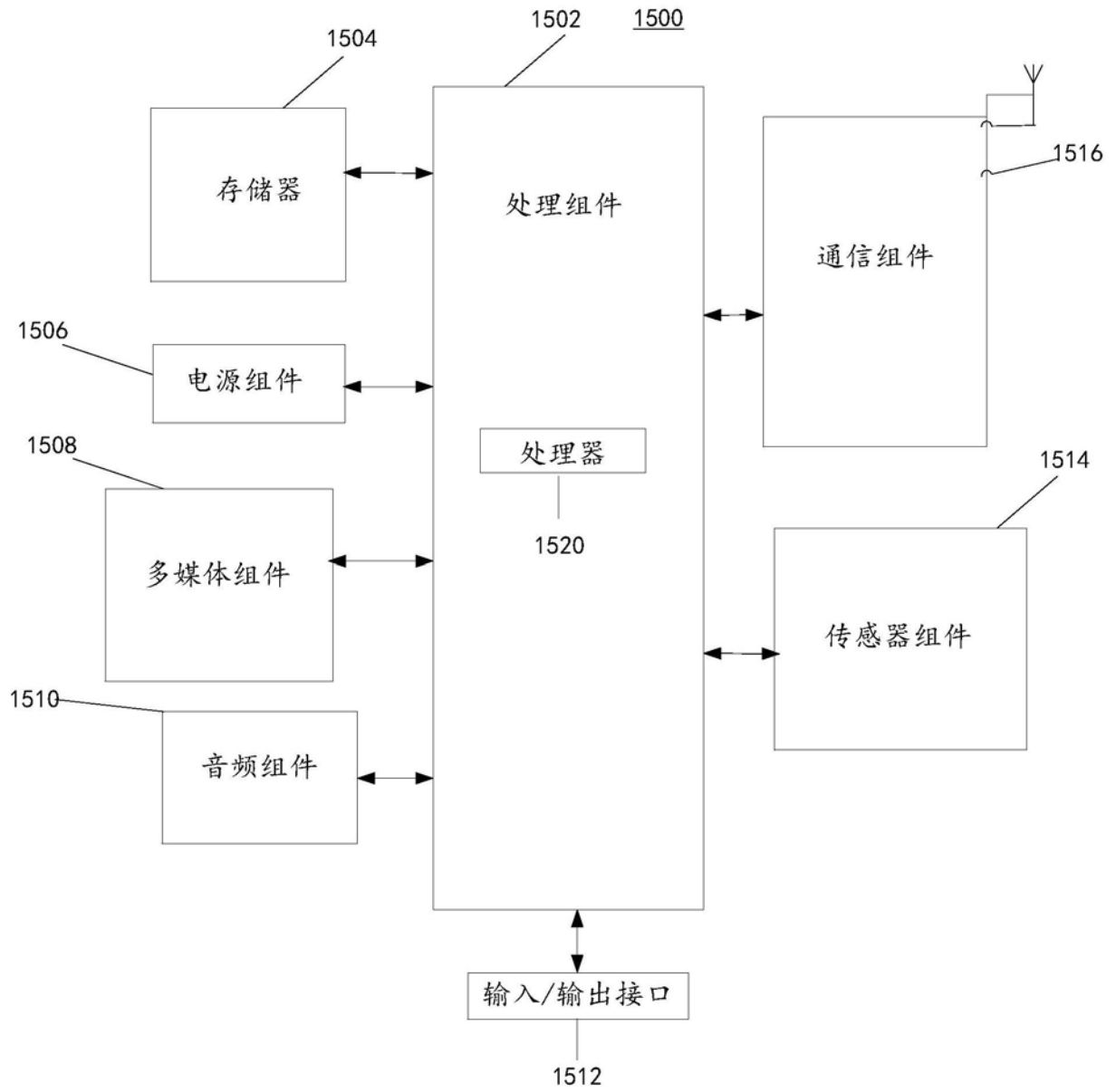


图15