



(12) 发明专利

(10) 授权公告号 CN 112634911 B

(45) 授权公告日 2022. 11. 11

(21) 申请号 202011519961.5

G10L 15/22 (2006.01)

(22) 申请日 2020.12.21

G10L 15/25 (2013.01)

(65) 同一申请的已公布的文献号

G10L 15/26 (2006.01)

申请公布号 CN 112634911 A

G10L 17/20 (2013.01)

(43) 申请公布日 2021.04.09

审查员 尚在颖

(73) 专利权人 思必驰科技股份有限公司

地址 215123 江苏省苏州市苏州工业园区

新平街388号腾飞创新园14栋

(72) 发明人 宋洪博 石韡斯 樊帅 朱成亚

邓建凯

(74) 专利代理机构 北京商专永信知识产权代理

事务所(普通合伙) 11400

专利代理师 黄谦 车江华

(51) Int. Cl.

G10L 17/22 (2013.01)

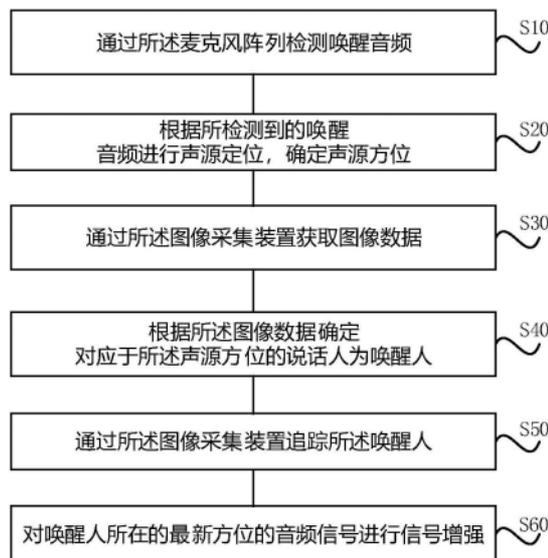
权利要求书2页 说明书7页 附图4页

(54) 发明名称

人机对话方法、电子设备及计算机可读存储介质

(57) 摘要

本发明公开一种人机对话方法,用于电子设备,电子设备配置有麦克风阵列和图像采集装置,该方法包括:通过麦克风阵列检测唤醒音频;根据所检测到的唤醒音频进行声源定位,确定声源方位;通过图像采集装置获取图像数据;根据图像数据确定对应于声源方位的说话人为唤醒人;通过图像采集装置追踪唤醒人;对唤醒人所在的最新方位的音频信号进行信号增强。本发明电子设备在被用户唤醒之后会的对话过程中持续基于图像采集装置来追踪当前对话任务中的唤醒人,时刻掌握当前唤醒人的方位,从而在对话过程中时刻对唤醒人实际所在方位的音频信号进行增强处理,提高了用于语音识别的音频信号的信噪比,提升了语音识别的准确率,确保了用户体验。



1. 一种人机对话方法,用于电子设备,所述电子设备配置有麦克风阵列和图像采集装置,所述方法包括:

通过所述麦克风阵列检测唤醒音频,其中,所述唤醒音频能够使所述电子设备从待机状态或者休眠状态进入工作状态;

根据所检测到的唤醒音频进行声源定位,确定声源方位;

通过所述图像采集装置获取图像数据;

根据所述图像数据确定对应于所述声源方位的说话人为唤醒人;

通过所述图像采集装置追踪所述唤醒人;

对所述唤醒人所在的最新方位的音频信号进行信号增强。

2. 根据权利要求1所述的方法,其特征在于,对所述唤醒人所在的最新方位的音频信号进行增强包括:

对所述唤醒人所在的最新方位的音频信号进行语音端点检测;

当所述音频信号中存在多段语音片段时,获取所述多段语音片段各自的起始时间和截止时间;

获取所述唤醒人所在的最新方位时的最新图像数据;

根据所述最新图像数据确定所述唤醒人开始说话的时间和截止说话的时间;

根据所述唤醒人开始说话的时间和截止说话的时间从所述多段语音片段中匹配出对应的语音片段;

对匹配出的语音片段进行信号增强。

3. 根据权利要求1所述的方法,其特征在于,通过所述图像采集装置获取图像数据包括:当所述麦克风阵列检测唤醒音频时,通过所述图像采集装置获取图像数据。

4. 根据权利要求3所述的方法,其特征在于,根据所述图像数据确定对应于所述声源方位的说话人为唤醒人包括:

根据所述图像数据确定对应于所述声源方位的说话人;

当所确定的说话人唯一时,确定该唯一说话人为唤醒人;

当所确定的说话人为多个说话人时,获取预设时间之前所述图像采集装置采集到的历史图像数据;

根据所述历史图像数据确定所述多个说话人中张口说话的人为唤醒人。

5. 根据权利要求4所述的方法,其特征在于,根据所述历史图像数据确定所述多个说话人中张口说话的人为唤醒人包括:

根据所述历史图像数据确定所述多个说话人各自的嘴唇动作数据;

根据所述多个说话人各自的嘴唇动作数据确定所述多个说话人各自的说话口型;

确定多个说话口型中与说出唤醒语的口型一个口型为目标口型;

确定所述目标口型所对应的说话人为唤醒人。

6. 根据权利要求1-5中任一项所述的方法,其特征在于,还包括:

在通过所述图像采集装置追踪所述唤醒人的过程中,判断所述唤醒人是否朝向所述电子设备;

如果是,则继续维持与所述唤醒人之间的对话任务;

如果不是,则结束与所述唤醒人之间的对话任务。

7. 根据权利要求6所述的方法,其特征在于,当所述唤醒人朝向所述电子设备,但超过预设时间未与电子设备进行对话,则结束与所述唤醒人之间的对话任务。

8. 根据权利要求1-5任一项所述的方法,其特征在于,还包括:对语音增强之后的音频信号进行语音识别。

9. 一种电子设备,其包括:至少一个处理器,以及与所述至少一个处理器通信连接的存储器,其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-8中任意一项所述方法的步骤。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现权利要求1-8中任意一项所述方法的步骤。

## 人机对话方法、电子设备及计算机可读存储介质

### 技术领域

[0001] 本发明涉及人机对话技术领域,尤其涉及一种人机对话方法、电子设备及计算机可读存储介质。

### 背景技术

[0002] 人机对话智能语音机器人已慢慢走入我们的生活,给我们的生活带来及极大的便利。为了提升对用户语音信号的识别准确率,现有技术中通常采用声源定位的方式确定用户所在方位,然后对该方位的音频信号进行信号增强处理,并抑制其它方位的音频信号,从而获得高信噪比的用于语音识别的音频信号。

[0003] 然而发明人在实现本发明的过程中发现,用户在于智能语音机器人对话过程中可能会走动,从而相对于智能语义机器人的相对方位也就发生了变化,如果此时仍然对原方位的音频信号进行增强,对其它方位音频信号进行抑制的话,将会极大的影响对当前用户语音识别的准确率,严重影响用户体验。

### 发明内容

[0004] 本发明实施例提供一种人机对话方法、电子设备及计算机可读存储介质,用于至少解决上述技术问题之一。

[0005] 第一方面,本发明实施例提供一种人机对话方法,用于电子设备,所述电子设备配置有麦克风阵列和图像采集装置,所述方法包括:

[0006] 通过所述麦克风阵列检测唤醒音频;

[0007] 根据所检测到的唤醒音频进行声源定位,确定声源方位;

[0008] 通过所述图像采集装置获取图像数据;

[0009] 根据所述图像数据确定对应于所述声源方位的说话人为唤醒人;

[0010] 通过所述图像采集装置追踪所述唤醒人;

[0011] 对所述唤醒人所在的最新方位的音频信号进行信号增强。

[0012] 第二方面,本发明实施例提供一种计算机可读存储介质,所述存储介质中存储有一个或多个包括执行指令的程序,所述执行指令能够被电子设备(包括但不限于计算机,服务器,或者网络设备等)读取并执行,以用于执行本发明上述任一项人机对话方法。

[0013] 第三方面,提供一种电子设备,其包括:至少一个处理器,以及与所述至少一个处理器通信连接的存储器,其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行本发明上述任一项人机对话方法。

[0014] 第四方面,本发明实施例还提供一种计算机程序产品,所述计算机程序产品包括存储在存储介质上的计算机程序,所述计算机程序包括程序指令,当所述程序指令被计算机执行时,使所述计算机执行上述任一项人机对话方法。

[0015] 本发明实施例的有益效果在于:电子设备在被用户唤醒之后会的对话过程中持续

基于图像采集装置来追踪当前对话任务中的唤醒人,时刻掌握当前唤醒人的方位,从而在对话过程中时刻对唤醒人实际所在方位的音频信号进行增强处理,提高了用于语音识别的音频信号的信噪比,提升了语音识别的准确率,确保了用户体验。

### 附图说明

[0016] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0017] 图1为本发明的人机对话方法的一实施例的流程图;

[0018] 图2为本发明的人机对话方法的另一实施例的流程图;

[0019] 图3为本发明的人机对话方法的再一实施例的流程图;

[0020] 图4为本发明的人机对话方法的又一实施例的流程图;

[0021] 图5为本发明的人机对话方法的另一实施例的流程图;

[0022] 图6为本发明的电子设备的一实施例的结构示意图。

### 具体实施方式

[0023] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0024] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0025] 本发明可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、元件、数据结构等等。也可以在分布式计算环境中实践本发明,在这些分布式计算环境中,通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0026] 在本发明中,“模块”、“装置”、“系统”等指应用于计算机的相关实体,如硬件、硬件和软件的组合、软件或执行中的软件等。详细地说,例如,元件可以、但不限于运行于处理器的过程、处理器、对象、可执行元件、执行线程、程序和/或计算机。还有,运行于服务器上的应用程序或脚本程序、服务器都可以是元件。一个或多个元件可在执行的过程和/或线程中,并且元件可以在一台计算机上本地化和/或分布在两台或多台计算机之间,并可以由各种计算机可读介质运行。元件还可以根据具有一个或多个数据包的信号,例如,来自一个与本地系统、分布式系统中另一元件交互的,和/或在因特网的网络通过信号与其它系统交互的数据的信号通过本地和/或远程过程来进行通信。

[0027] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”,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、物品或者设备所固有

的要素。在没有更多限制的情况下,由语句“包括……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0028] 本发明实施例提供一种人机对话方法,用于电子设备,该电子设备配置有麦克风阵列和图像采集装置。示例性地,电子设备可以是配置有麦克风阵列和图像采集装置的智能终端设备。例如,可以是智能手机、智能音箱、智能客服机器人等,本发明对此不作限定。

[0029] 如图1所示,为本发明的人机对话方法的一实施例的流程图,在该实施例中该方法包括:

[0030] S10、通过所述麦克风阵列检测唤醒音频。电子设备处于待机状态或者休眠状态,通过麦克风阵列监测唤醒音频,当检测到唤醒音频时则进入工作装填。例如,对于一款在商场使用的名为“xx”的智能客服机器人,用户只需要喊他的名字“xx您好”,该智能客服机器人就会被唤醒进入对话状态。

[0031] S20、根据所检测到的唤醒音频进行声源定位,确定声源方位。

[0032] 示例性地,智能客服机器人根据检测到的唤醒音频进行特征提取,并根据所提取的特征进行声源定位,从而确定说话人所在方位。

[0033] 示例性地,通过麦克风阵列获取的音频特征,可以计算出发生源和麦克风阵列的角度,通过beamforming技术可以增强唤醒人所在区域的音频新信号,抑制其他方向区域的信号,从而实现其他方向噪声屏蔽。

[0034] S30、通过所述图像采集装置获取图像数据。示例性地,当所述麦克风阵列检测唤醒音频时,通过所述图像采集装置获取图像数据。图像数据可以是连续的多帧图像。

[0035] S40、根据所述图像数据确定对应于所述声源方位的说话人为唤醒人。

[0036] 示例性地,图像数据中有多个人像,确定其中相对于智能客服机器人的方位与步骤S20中所确定的说话人方位相同的人像为唤醒人。

[0037] S50、通过所述图像采集装置追踪所述唤醒人。

[0038] 示例性地,在确定唤醒人之后,从图像数据中提取该唤醒人的特征信息(例如,人脸特征信息),在后续的对话过程中基于该特征信息实时的通过图像采集装置追踪唤醒人。例如,对于智能客服机器人,用户在说完一句话之后可能想左侧走了一步,这时通过图像采集装置就能够追踪到用户的行走,并重新确定出该用户的当前最新方位。

[0039] S60、对所述唤醒人所在的最新方位的音频信号进行信号增强。

[0040] 本实施例中,电子设备在被用户唤醒之后会的对话过程中持续基于图像采集装置来追踪当前对话任务中的唤醒人,时刻掌握当前唤醒人的方位,从而在对话过程中时刻对唤醒人实际所在方位的音频信号进行增强处理,提高了用于语音识别的音频信号的信噪比,提升了语音识别的准确率,确保了用户体验。

[0041] 如图2所示为本发明的人机对话方法的另一实施例的流程图,在该实施例中,对所述唤醒人所在的最新方位的音频信号进行增强包括:

[0042] S61、对所述唤醒人所在的最新方位的音频信号进行语音端点检测。

[0043] S62、当所述音频信号中存在多段语音片段时,获取所述多段语音片段各自的起始时间和截止时间。

[0044] 示例性地,根据VAD内核抛出的音频特征,可以标记出音频中silence的时间段和speech的时间段,从而可以获取到从silence变为speech的时间点 $T_a$ ,和从speech变为

silence的时间点Tb。

[0045] S63、获取所述唤醒人所在的最新方位时的最新图像数据。

[0046] S64、根据所述最新图像数据确定所述唤醒人开始说话的时间和截止说话的时间。

[0047] 示例性地,根据图像数据获取人脸关键点坐标,截取其中嘴部的关键点,根据上下嘴唇坐标计算,嘴部张开大小,从而确认唤醒人是否在张开嘴说话,如果是则标记为speech状态,并记录时间点T1;上下嘴唇距离小于某距离,则认为嘴部为闭合状态,标记为silence,并记录时间点T2。

[0048] 本发明实施例中通过摄像头获取的图像信息,可以获取到设备唤醒时刻的图像帧,根据图像中的人脸特征和唇部特征,可以确认唤醒设备的唤醒人人脸,根据人脸位置,确认唤醒人所在beam,增强该beam的音频信号。

[0049] S65、根据所述唤醒人开始说话的时间和截止说话的时间从所述多段语音片段中匹配出对应的语音片段。

[0050] 示例性地,一段音频中可能包含多段speech数据,参考视觉的T1、T2时间点则可以确认哪段speech数据是唤醒人输入的音频,只将该段音频送识别,则有效屏蔽了其他噪声造成的误识别。

[0051] S66、对匹配出的语音片段进行信号增强。

[0052] 本实施例中,当同一段音频中存在多个说话人的多段语音时,根据多段语音分别的起始和截止时间,与唤醒人的嘴唇动作的起始和截止时间进行匹配,从而筛选出对应于所述唤醒人的语音片段进行增强。

[0053] 实际情况中异口同声(同时开始、同时结束的概率将微乎其微)几乎是不可能发生的。因此,当在同一段音频数据中存在有多段语音时,本发明实施例可以根据图像数据中获得的嘴唇动作数据来确定目标说话人说话的起始和结束时间点,并从多段语音中获得相匹配的一段为目标语音。

[0054] 如图3所示为本发明的人机对话方法的另一实施例的流程图,在该实施例中,根据所述图像数据确定对应于所述声源方位的说话人为唤醒人包括:

[0055] S41、根据所述图像数据确定对应于所述声源方位的说话人;

[0056] S42、当所确定的说话人唯一时,确定该唯一说话人为唤醒人;

[0057] S43、当所确定的说话人为多个说话人时,获取预设时间之前所述图像采集装置采集到的历史图像数据。

[0058] 示例性地,预设时间设置为使得所获得的历史图像数据中包含了唤醒人说出唤醒语时的图像数据。例如,预设时间可以是0.5s,本发明对此不作限定。

[0059] S44、根据所述历史图像数据确定所述多个说话人中张口说话的人为唤醒人。

[0060] 本实施例中,有其他人和唤醒人站在同一方位时,仅仅根据声源定位已经不在能够确定哪个人是真正的唤醒人,从而进一步的获取当前时刻之前的历史图像数据,通过分析历史图像数据来确定在同一方位上的多个人中在相应时间说过话的人,并将该说过话的人确定为唤醒人。

[0061] 如图4所示为本发明的人机对话方法的另一实施例的流程图,在该实施例中,根据所述历史图像数据确定所述多个说话人中张口说话的人为唤醒人包括:

[0062] S441、根据所述历史图像数据确定所述多个说话人各自的嘴唇动作数据;

[0063] S442、根据所述多个说话人各自的嘴唇动作数据确定所述多个说话人各自的说话口型；

[0064] S443、确定多个说话口型中与说出唤醒语的口型一个口型为目标口型；

[0065] S444、确定所述目标口型所对应的说话人为唤醒人。

[0066] 本实施例中，站在同一方位的多个人中除了真正的唤醒人之外，其他人在唤醒人说出唤醒语的同时可能也正在跟别人交谈（即，正在说话）。尤其是商场这种人流量交大的场所，相对于智能客服机器人的同一个方位上同时有多个人，并且一个在于智能客服机器人进行交互，另外还有人可能正在与其他人沟通。因此，这时基于之前实施例中的方法就能够检测到两个同时开口说话的人，这时就无法确定出真正的唤醒人了。因此本实施例中进一步的根据所获得的图像数据来确定说话人的嘴唇动作所对应的口型，并确定与说出唤醒语应当具备的口型相同的那个人为唤醒人。

[0067] 在一些实施例中，本发明的人机对话方法还包括：在通过所述图像采集装置追踪所述唤醒人的过程中，判断所述唤醒人是否朝向所述电子设备；如果是，则继续维持与所述唤醒人之间的对话任务；如果否，则结束与所述唤醒人之间的对话任务。

[0068] 在本实施中，如果判断所述唤醒人朝向所述电子设备，则表明当前目标说话人（即，唤醒人）仍在沟通中，否则判定说话人已经不在需要沟通，结束对其它方位声音的抑制。

[0069] 在一些实施例中，当所述唤醒人朝向所述电子设备，但超过预设时间未与电子设备进行对话，则结束与所述唤醒人之间的对话任务。

[0070] 本实施例中，虽然目标说话人是朝向终端设备的，但是超过预设时间未与终端设备进行沟通，则判定说话人已经不在需要沟通，结束对其它方位声音的抑制。

[0071] 在一些实施例中，本发明的人机对话方法还包括：对语音增强之后的音频信号进行语音识别。

[0072] 如图5所示为本发明的人机对话方法的另一实施例的流程图，在该实施例中，包括以下步骤：

[0073] 1. 说话人通过唤醒语唤醒设备，设备根据音频特征，计算说话人的方位。设备需要具备摄像头和麦克风阵列，并且排放在相同位置，在定位说话人时，音频方位和视觉方位可以对齐。

[0074] 2. 说话人在唤醒设备的同时，人脸识别引擎通过分析图像数据，定位摄像头前的全部人脸，并分别标记人脸ID，确定说话人的人脸坐标。

[0075] 3. 根据第2步通过音频特征计算出的说话人方位，和第3步通过图像特征计算出的说话人方位，标记出说话人ID，该ID记录为唤醒人。

[0076] 4. 唤醒人被确定后，通过图像追踪唤醒人方位，并通过beamforming技术，增强说话人所在方位的音频信号，其他方位的音频被抑制，减少其他方向的音频造成的误识别。

[0077] 5. 根据确定的人脸关键点81点，获取到唇部18点的坐标，根据唇部上下嘴唇坐标，判断嘴部是否在说话；对输入音频进行标记，标记开始说话的时间点和结束说话的时间点。

[0078] 6. VAD内核根据音频特征对输入音频进行标记，标记说话的开始时间点和结束时间点。

[0079] 7. 将视觉和音频特性标记的时间点进行对齐，不匹配的音频段丢弃，匹配的音频

段送识别。

[0080] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作合并,但是本领域技术人员应该知悉,本发明并不受所描述的动作顺序的限制,因为依据本发明,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本发明所必须的。在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中沒有详述的部分,可以参见其他实施例的相关描述。

[0081] 在一些实施例中,本发明实施例提供一种非易失性计算机可读存储介质,所述存储介质中存储有一个或多个包括执行指令的程序,所述执行指令能够被电子设备(包括但不限于计算机,服务器,或者网络设备)读取并执行,以用于执行本发明上述任一项人机对话方法。

[0082] 在一些实施例中,本发明实施例还提供一种计算机程序产品,所述计算机程序产品包括存储在非易失性计算机可读存储介质上的计算机程序,所述计算机程序包括程序指令,当所述程序指令被计算机执行时,使所述计算机执行上述任一项人机对话方法。

[0083] 在一些实施例中,本发明实施例还提供一种电子设备,其包括:至少一个处理器,以及与所述至少一个处理器通信连接的存储器,其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行人机对话方法。

[0084] 在一些实施例中,本发明实施例还提供一种存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现人机对话方法。

[0085] 图6是本申请另一实施例提供的执行人机对话方法的电子设备的硬件结构示意图,如图6所示,该设备包括:

[0086] 一个或多个处理器610以及存储器620,图6中以一个处理器610为例。

[0087] 执行人机对话方法的设备还可以包括:输入装置630和输出装置640。

[0088] 处理器610、存储器620、输入装置630和输出装置640可以通过总线或者其他方式连接,图6中以通过总线连接为例。

[0089] 存储器620作为一种非易失性计算机可读存储介质,可用于存储非易失性软件程序、非易失性计算机可执行程序以及模块,如本申请实施例中的人机对话方法对应的程序指令/模块。处理器610通过运行存储在存储器620中的非易失性软件程序、指令以及模块,从而执行服务器的各种功能应用以及数据处理,即实现上述方法实施例人机对话方法。

[0090] 存储器620可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据人机对话装置的使用所创建的数据等。此外,存储器620可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实施例中,存储器620可选包括相对于处理器610远程设置的存储器,这些远程存储器可以通过网络连接至人机对话装置。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0091] 输入装置630可接收输入的数字或字符信息,以及产生与人机对话装置的用户设置以及功能控制有关的信号。输出装置640可包括显示屏等显示设备。

[0092] 所述一个或者多个模块存储在所述存储器620中,当被所述一个或者多个处理器610执行时,执行上述任意方法实施例中的人机对话方法。

[0093] 上述产品可执行本申请实施例所提供的方法,具备执行方法相应的功能模块和有益效果。未在本实施例中详尽描述的技术细节,可参见本申请实施例所提供的方法。

[0094] 本申请实施例的电子设备以多种形式存在,包括但不限于:

[0095] (1) 移动通信设备:这类设备的特点是具备移动通信功能,并且以提供话音、数据通信为主要目标。这类终端包括:智能手机(例如iPhone)、多媒体手机、功能性手机,以及低端手机等。

[0096] (2) 超移动个人计算机设备:这类设备属于个人计算机的范畴,有计算和处理功能,一般也具备移动上网特性。这类终端包括:PDA、MID和UMPC设备等,例如iPad。

[0097] (3) 便携式娱乐设备:这类设备可以显示和播放多媒体内容。该类设备包括:音频、视频播放器(例如iPod),掌上游戏机,电子书,以及智能玩具和便携式车载导航设备。

[0098] (4) 其他具有数据交互功能的电子装置。

[0099] 以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0100] 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各实施方式可借助软件加通用硬件平台的方式来实现,当然也可以通过硬件。基于这样的理解,上述技术方案本质上或者说对相关技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在计算机可读存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行各个实施例或者实施例的某些部分所述的方法。

[0101] 最后应说明的是:以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

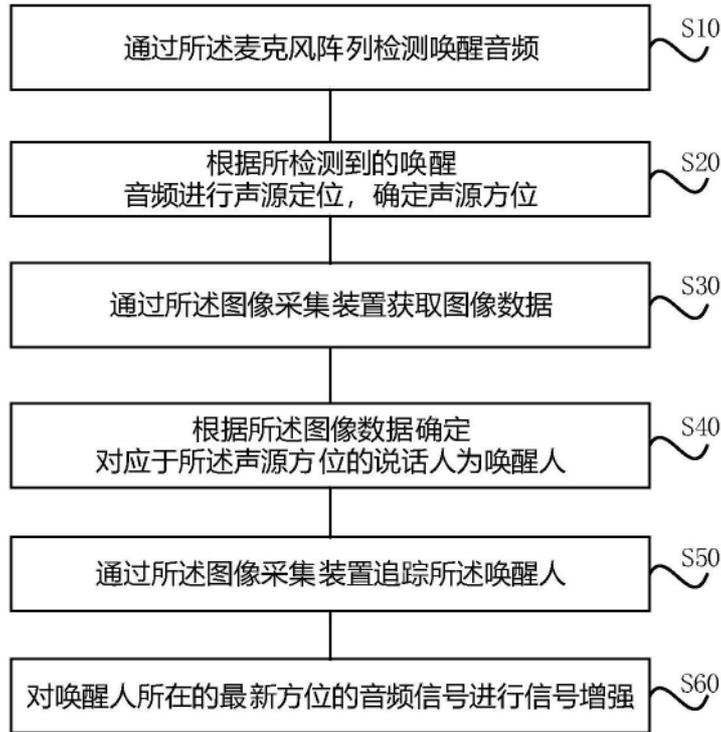


图1

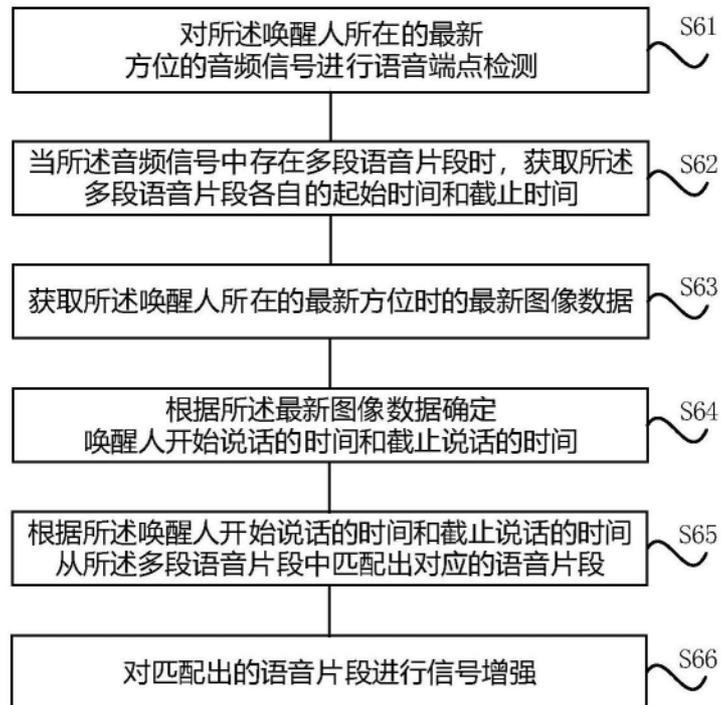


图2

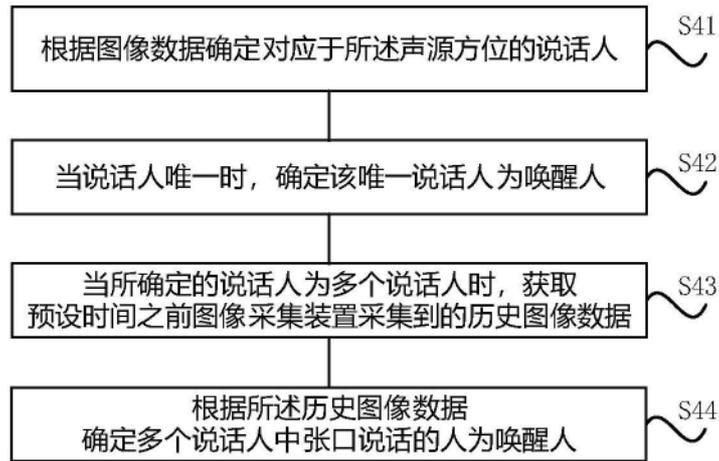


图3

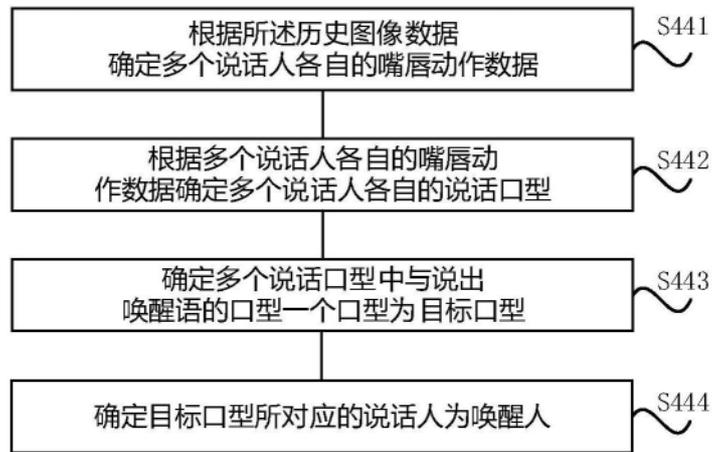


图4

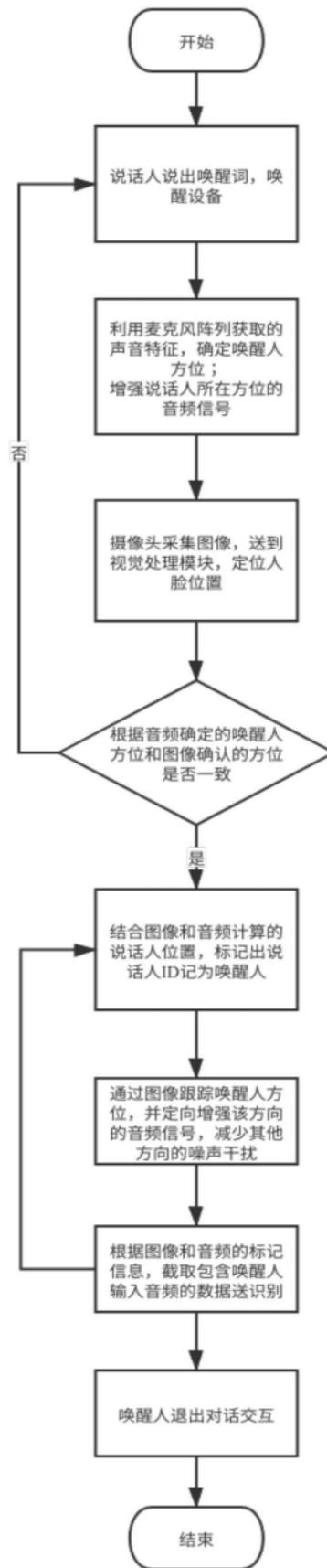


图5

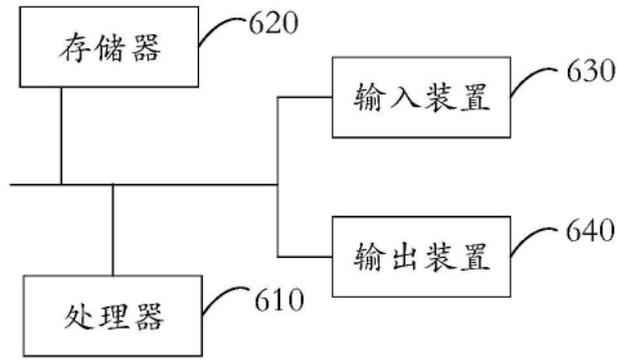


图6