



(12) 发明专利申请

(10) 申请公布号 CN 112312064 A

(43) 申请公布日 2021. 02. 02

(21) 申请号 202011204878.9

(22) 申请日 2020.11.02

(71) 申请人 腾讯科技(深圳)有限公司
地址 518000 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72) 发明人 李杨

(74) 专利代理机构 北京市立方律师事务所
11330

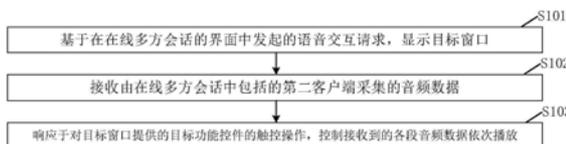
代理人 张筱宁

(51) Int. Cl.
H04N 7/15 (2006.01)

权利要求书3页 说明书18页 附图7页

(54) 发明名称
语音交互方法及相关设备

(57) 摘要
本申请涉及云计算技术领域,提供了一种语音交互方法及相关设备,该方法包括:基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口;接收由在线多方会话中包括的第二客户端采集的音频数据;响应于对目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。本申请的实施有利于简便语音发言的操作,避免多用户同时发言的发言语音冲突问题。



1. 一种语音交互方法,应用于第一客户端,其特征在于,包括:
基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口;
接收由所述在线多方会话中包括的第二客户端采集的音频数据;
响应于对所述目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。

2. 根据权利要求1所述的方法,其特征在于,所述基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口,包括:

响应于在在线多方会话的界面上对开启语音交互功能控件的触控操作,向服务器发起语音交互请求;

显示用于控制音频数据进行播放的目标窗口;

其中,所述目标窗口包括以下至少一项:

显示区域,用于显示当前接收到的音频数据;

设置区域,用于设置每一段音频数据的语音交互时长;

其中,所述目标功能控件至少包括跳过音频数据、清空音频数据、暂停播放、调整播放顺序和停止语音交互中对应的一个功能控件。

3. 根据权利要求1所述的方法,其特征在于,所述接收由所述在线多方会话中包括的第二客户端采集的音频数据,包括:

接收服务器发送的由所述在线多方会话中包括的第二客户端采集的音频数据;

基于接收各段音频数据的时间,以时间轴的形式在所述目标窗口中显示各段音频数据;

所述音频数据的显示内容由以下至少一项组合而成:各段音频数据分别对应第二客户端的账户名称、语音波形略缩图和时长。

4. 根据权利要求3所述的方法,其特征在于,所述基于接收各段音频数据的时间,以时间轴的形式在所述目标窗口中显示各段音频数据,包括:

若当前正在接收的音频数据的时长不小于预设时长,将该音频数据基于接收的时间写入播放队列,并继续接收该音频数据;

若当前正在接收的音频数据的时长小于预设时长,删除该音频数据;

基于所述播放队列以时间轴的形式在所述目标窗口中显示各段音频数据。

5. 根据权利要求2所述的方法,其特征在于,所述响应于对所述目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放,包括以下至少一项:

响应于对所述跳过音频数据功能控件的触控操作,暂停播放当前正在播放的音频数据,并播放下一段音频数据;

响应于对所述清空音频数据功能控件的触控操作,删除以下至少一项:当前正在播放的音频数据、当前选定的至少一段音频数据和当前接收到的所有音频数据;

响应于对所述暂停播放功能控件的触控操作,暂停播放所述音频数据;

响应于对所述调整播放顺序功能控件的触控操作,调整当前选定的至少一段音频数据的播放顺序。

6. 根据权利要求2所述的方法,其特征在于,所述控制接收到的各段音频数据依次播放之后,还包括:

响应于对所述停止语音交互功能控件的触控操作,向所述服务器发起停止语音交互请求,以停止所述在线多方会话中正在进行的语音交互行为;

其中,停止的语音交互行为包括以下至少一项:停止播放所述音频数据、关闭所述目标窗口、删除当前接收的所有音频数据。

7. 一种语音交互方法,应用于第二客户端,其特征在于,包括:

在在线多方会话的界面显示第一提示信息;所述第一提示信息用于提醒用户进行语音交互;

响应于语音交互指令,采集音频数据并上传该音频数据,以使该音频数据与所述在线多方会话中的其他待播放的音频数据一并进行依次播放。

8. 根据权利要求7所述的方法,其特征在于,所述方法还包括:

播放当前接收到的目标音频数据,所述目标音频数据包括第一客户端控制所接收到的音频数据依次播放的音频数据。

9. 根据权利要求7所述的方法,其特征在于,所述响应于语音交互指令包括以下至少一项:

响应于对预设的硬件按键的长按操作,开启音频数据的采集功能;

响应于在所述在线多方会话的界面上对发言功能控件的触控操作,开启音频数据的采集功能。

10. 根据权利要求9所述的方法,其特征在于,所述在在线多方会话的界面显示第一提示信息之前,还包括接收根据第一客户端在在线多方会话的界面中发起的语音交互请求下发的语音交互权限;所述语音交互权限中携带有预设的语音交互时长;所述采集音频数据并上传该音频数据,还包括以下至少一项:

当正在采集的音频数据的时长达所述预设的语音交互时长时,中断对该音频数据的采集和/或关闭所述采集功能;

当正在采集音频数据时,响应于停止语音交互指令,中断对该音频数据的采集和/或关闭所述采集功能,若该音频数据的时长小于预设时长,显示语音交互失败的提示信息。

11. 根据权利要求7所述的方法,其特征在于,所述响应于语音交互指令,采集音频数据并上传该音频数据,还包括:

显示提示窗口,所述提示窗口包括用于显示所采集音频数据时长的提示框。

12. 一种语音交互装置,应用于第一客户端,其特征在于,包括:

窗口显示模块,用于基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口;

接收模块,用于接收由所述在线多方会话中包括的第二客户端采集的音频数据;

控制模块,用于响应于对所述目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。

13. 一种语音交互装置,应用于第二客户端,其特征在于,包括:

显示模块,用于在在线多方会话的界面显示第一提示信息;所述第一提示信息用于提醒用户进行语音交互;

采集模块,用于响应于语音交互指令,采集音频数据并上传该音频数据,以使该音频数据与所述在线多方会话中的其他待播放的音频数据一并进行依次播放。

14. 一种电子设备,其特征在于,其包括:

一个或多个处理器;

存储器;

一个或多个应用程序,其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序配置用于:执行根据权利要求1~6和/或7~11中任一项所述的语音交互方法。

15. 一种计算机可读存储介质,其特征在于,所述存储介质存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现如权利要求1~6和/或7~11中任一项所述的语音交互方法。

语音交互方法及相关设备

技术领域

[0001] 本申请涉及云计算技术领域,具体而言,本申请涉及一种语音交互方法及相关设备。

背景技术

[0002] 在线会话包括在线教育、在线会议等,是一种基于网络进行远程交流的行为。随着互联网技术的发展,在线会话已慢慢渗透于人们的日常生活。与传统需要参与会话的人员同在一个实体空间进行交流的模式相比,在线会话突破了时空的界限,节省了各个会话参与者的时间成本。

[0003] 在线会话中重要的一环是语音发言,语音发言是指会话参与者通过语言与其他会话参与者进行交流。现有技术中,实现语音发言有两种模式:一种是会话参与者直接开启麦克风连麦进行发言,但是该模式在同一会话环境中参与者较多时,容易出现多个参与者同时开麦进行发言的情况,导致发言冲突;另一种是会话参与者发起发言请求经会话主持人同意后参与者连麦发言,但是该模式在每个参与者发言时均需要主持人进行操作,降低了交流的效率。

发明内容

[0004] 本申请提供了一种语音交互方法及相关设备,可以解决上述至少一种技术问题。所述技术方案如下:

[0005] 第一方面,提供了一种语音交互方法,应用于第一客户端,包括:基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口;接收由所述在线多方会话中包括的第二客户端采集的音频数据;响应于对所述目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。

[0006] 结合第一方面,本申请在第一方面的第一种实施方式中,所述基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口,包括:响应于在在线多方会话的界面上对开启语音交互功能控件的触控操作,向服务器发起语音交互请求;显示用于控制音频数据进行播放的目标窗口;其中,所述目标窗口包括以下至少一项:显示区域,用于显示当前接收到的音频数据;设置区域,用于设置每一段音频数据的语音交互时长;其中,所述目标功能控件,用于控制音频数据进行播放;所述目标功能控件至少包括跳过音频数据、清空音频数据、暂停播放、调整播放顺序和停止语音交互中对应的一个功能控件。

[0007] 结合第一方面,本申请在第一方面的第二种实施方式中,所述接收由所述在线多方会话中包括的第二客户端采集的音频数据,包括:接收服务器发送的由所述在线多方会话中包括的至少两个第二客户端才的音频数据;基于接收各段音频数据的时间,以时间轴的形式在所述目标窗口中显示各段音频数据;所述音频数据的显示内容由以下至少一项组合而成:各段音频数据分别对应第二客户端的账户名称、语音波形略缩图和时长。

[0008] 结合第一方面的第二种实施方式,本申请在第一方面的第三种实施方式中,所述

基于接收各段音频数据的时间,以时间轴的形式在所述目标窗口中显示各段音频数据,包括:若当前正在接收的音频数据的时长不小于预设时长,将该音频数据基于接收的时间写入播放队列,并继续接收该音频数据;若当前正在接收的音频数据的时长小于预设时长,删除该音频数据;基于所述播放队列以时间轴的形式在所述目标窗口中显示各段音频数据。

[0009] 结合第一方面的第一种实施方式,本申请在第一方面的第四种实施方式中,所述响应于对所述目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放,包括以下至少一项:响应于对所述跳过音频数据功能控件的触控操作,暂停播放当前正在播放的音频数据,并播放下一段音频数据;响应于对所述清空音频数据功能控件的触控操作,删除以下至少一项:当前正在播放的音频数据、当前选定的至少一段音频数据和当前接收到的所有音频数据;响应于对所述暂停播放功能控件的触控操作,暂停播放所述音频数据;

[0010] 响应于对所述调整播放顺序功能控件的触控操作,调整当前选定的至少一段音频数据的播放顺序。

[0011] 结合第一方面的第一种实施方式,本申请在第一方面的第五种实施方式中,所述控制接收到的各段音频数据依次播放之后,还包括:响应于对所述停止语音交互功能控件的触控操作,向所述服务器发起停止语音交互请求,以停止所述在线多方会话中正在进行的语音交互行为;其中,停止的语音交互行为包括以下至少一项:停止播放所述音频数据、关闭所述目标窗口、删除当前接收的所有音频数据。

[0012] 第二方面,提供了一种语音交互方法,应用于第二客户端,包括:在在线多方会话的界面显示第一提示信息;所述第一提示信息用于提醒用户进行语音交互;响应于语音交互指令,采集音频数据并上传该音频数据,以使该音频数据与所述在线多方会话中的其他待播放的音频数据一并进行依次播放。

[0013] 结合第二方面,本申请在第二方面的第一种实施方式中,所述方法还包括:播放当前接收到的目标音频数据,所述目标音频数据包括第一客户端控制所接收到的音频数据依次播放的音频数据。

[0014] 结合第二方面,本申请在第二方面的第二种实施方式中,所述响应于语音交互指令包括以下至少一项:响应于对预设的硬件按键的长按操作,开启音频数据的采集功能;响应于在所述在线多方会话的界面上对发言功能控件的触控操作,开启音频数据的采集功能。

[0015] 结合第二方面的第二种实施方式,本申请在第二方面的第三种实施方式中,所述在在线多方会话的界面显示第一提示信息之前,还包括接收根据第一客户端在在线多方会话的界面中发起的语音交互请求下发的语音交互权限;所述语音交互权限中携带有预设的语音交互时长;所述采集音频数据并上传该音频数据,还包括以下至少一项:当正在采集的音频数据的时长达所述预设的语音交互时长时,中断对该音频数据的采集和/或关闭所述采集功能;当正在采集音频数据时,响应于停止语音交互指令,中断对该音频数据的采集和/或关闭所述采集功能,若该音频数据的时长小于预设时长,显示语音交互失败的提示信息。

[0016] 结合第二方面,本申请在第二方面的第四种实施方式中,所述响应于语音交互指令,采集音频数据并上传该音频数据,还包括:显示提示窗口,所述提示窗口包括用于显示

所采集音频数据时长的提示框。

[0017] 第三方面,提供了一种语音交互装置,应用于第一客户端,包括:窗口显示模块,用于基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口;接收模块,用于接收由所述在线多方会话中包括的第二客户端采集的音频数据;控制模块,用于响应于对所述目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。

[0018] 结合第三方面,本申请在第三方面的第一种实施方式中,所述窗口显示模块包括:请求单元,用于响应于在在线多方会话的界面上对开启语音交互功能控件的触控操作,向服务器发起语音交互请求;第一显示单元,用于显示用于控制音频数据进行播放的目标窗口;其中,所述目标窗口包括以下至少一项:显示区域,用于显示当前接收到的音频数据;设置区域,用于设置每一段音频数据的时长;其中,所述目标功能控件,用于控制音频数据进行播放;所述目标功能控件至少包括跳过音频数据、清空音频数据、暂停播放、调整播放顺序和停止语音交互中对应的一个功能控件。

[0019] 结合第三方面,本申请在第三方面的第二种实施方式中,所述接收模块,包括:接收单元,用于接收服务器发送的由所述在线多方会话中包括的第二客户端采集的音频数据;第二显示单元,用于基于接收各段音频数据的时间,以时间轴的形式在所述目标窗口中显示各段音频数据;所述音频数据的显示内容由以下至少一项组合而成:各段音频数据分别对应第二客户端的账户名称、语音波形略缩图和时长。

[0020] 结合第三方面的第二种实施方式,本申请在第三方面的第三种实施方式中,所述第二显示单元包括:第一接收子单元,用于若当前正在接收的音频数据的时长不小于预设时长,将该音频数据基于接收的时间写入播放队列,并继续接收该音频数据;第二接收子单元,用于若当前正在接收的音频数据的时长小于预设时长,删除该音频数据;显示子单元,用于基于所述播放队列以时间轴的形式在所述目标窗口中显示各段音频数据。

[0021] 结合第三方面的第一种实施方式,本申请在第三方面的第四种实施方式中,所述控制模块至少包括以下一个单元:第一暂停单元,用于响应于对所述跳过音频数据功能控件的触控操作,暂停播放当前正在播放的音频数据,并播放下一段音频数据;删除单元,用于响应于对所述清空音频数据功能控件的触控操作,删除以下至少一项:当前正在播放的音频数据、当前选定的至少一段音频数据和当前接收到的所有音频数据;第二暂停单元,用于响应于对所述暂停播放功能控件的触控操作,暂停播放所述音频数据;调整单元,用于响应于对所述调整播放顺序功能控件的触控操作,调整当前选定的至少一段音频数据的播放顺序。

[0022] 结合第三方面的第一种实施方式,本申请在第三方面的第五种实施方式中,还包括停止模块,用于响应于对所述停止语音交互功能控件的触控操作,向所述服务器发起停止语音交互请求,以停止所述在线多方会话中正在进行的语音交互行为;其中,停止的语音交互行为包括以下至少一项:停止播放所述音频数据、关闭所述目标窗口、删除当前接收的所有音频数据。

[0023] 第四方面,提供了一种语音交互装置,应用于第二客户端,包括:显示模块,用于在在线多方会话的界面显示第一提示信息;所述第一提示信息用于提醒用户进行语音交互;采集模块,用于响应于语音交互指令,采集音频数据并上传该音频数据,以使该音频数据与所述在线多方会话中的其他待播放的音频数据一并进行依次播放。

[0024] 结合第四方面,本申请在第四方面的第一种实施方式中,所述装置还包括播放模块,用于播放当前接收到的目标音频数据,所述目标音频数据包括第一客户端控制所接收到的音频数据依次播放的音频数据。

[0025] 结合第四方面,本申请在第四方面的第二种实施方式中,所述采集模块至少包括以下一个单元:第一开启单元,用于响应于对预设的硬件按键的长按操作,开启音频数据的采集功能;第二开启单元,用于响应于在所述在线多方会话的界面上对发言功能控件的触控操作,开启音频数据的采集功能。

[0026] 结合第四方面的第二种实施方式,本申请在第四方面的第三种实施方式中,所述显示模块还用于接收根据第一客户端在在线多方会话的界面中发起的语音交互请求下发的语音交互权限;所述语音交互权限中携带有预设的语音交互时长;所述采集模块还包括以下至少一个单元:第一处理单元,用于当正在采集的音频数据的时长达所述预设的语音交互时长时,中断对该音频数据的采集和/或关闭所述采集功能;第二处理单元,用于当正在采集音频数据时,响应于停止语音交互指令,中断对该音频数据的采集和/或关闭所述采集功能,若该音频数据的时长小于预设时长,显示语音交互失败的提示信息。

[0027] 结合第四方面,本申请在第四方面的第四种实施方式中,所述采集模块还用于显示提示窗口,所述提示窗口包括用于显示所采集音频数据时长的提示框。

[0028] 第五方面,提供了一种电子设备,其包括:一个或多个处理器;存储器;一个或多个应用程序,其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序配置用于:执行第一方面、第一方面任一种实施方式、第二方面和/或第二方面任一种实施方式所述的语音交互方法。

[0029] 第六方面,提供了一种计算机可读存储介质,所述存储介质存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现第一方面、第一方面任一种实施方式、第二方面和/或第二方面任一种实施方式所述的语音交互方法。

[0030] 本申请提供的技术方案带来的有益效果是:

[0031] (1) 本申请提供的语音交互方法及相关设备中,在线会话的控制员(下称第一用户)可采用第一客户端实现语音交互;具体地,控制员可通过第一客户端在在线多方会话的界面中发起的语音交互请求,即在会话过程中,第一客户端可控制语言发言的时间节点,且通过一次请求即可开启参加当前在线多方会话所有用户的语言交互权限,简便了第一用户的操作;进而基于语音交互请求显示目标窗口;在第一客户端接收到由在线多方会话中包括的第二客户端采集的音频数据时,第一用户可通过目标窗口提供的目标功能控件控制接收到的各段音频数据进行依次播放,可在提高交流效率的同时,有效避免多个用户同时发言导致的发言冲突问题。

[0032] (2) 本申请提供的语音交互方法和相关设备中,参与在线会话的参与用户(下称第二用户)可采用第二客户端实现语言交互;具体地,当在在线多方会话的界面显示第一提示信息时,表征用户可在在线多方会话中进行语音交互,此时,第二用户可基于个人需求触发语音交互指令进行自由发言,使得第二客户端采集音频数据并上传该音频数据,以使该音频数据与在线多方会话中的其他待播放的音频数据一并进行依次播放。本申请的实施通过将采集的音频数据进行上传,以控制采集的音频数据与在线多方会话中的其他待播放的音

频数据一并进行依次播放,可有效提高会话期间各用户通过语音进行交流的效率,同时避免了多个用户同时发言导致的发言冲突问题。

[0033] 本申请附加的方面和优点将在下面的描述中部分给出,这些将从下面的描述中变得明显,或通过本申请的实践了解到。

附图说明

[0034] 为了更清楚地说明本申请实施例中的技术方案,下面将对本申请实施例描述中所需要使用的附图作简单地介绍。

[0035] 图1为本申请实施例提供了一种应用于第一客户端的语音交互方法的流程示意图;

[0036] 图2为本申请实施例提供了一种应用于第一客户端的语音交互方法的流程示意图;

[0037] 图3为本申请实施例提供了一种应用于第一客户端的语音交互方法中在线多方会话的界面示意图;

[0038] 图4为本申请实施例提供了一种应用于第一客户端的语音交互方法中目标窗口的示意图;

[0039] 图5为本申请实施例提供了一种应用于第二客户端的语音交互方法的流程示意图;

[0040] 图6为本申请实施例提供了一种应用于第二客户端的语音交互方法中在线多方会话的界面示意图;

[0041] 图7为本申请实施例提供了一种应用于第二客户端的语音交互方法中提示窗口的示意图;

[0042] 图8为本申请实施例提供的语音交互方法中第一客户端开启语音自由发言模式的流程图;

[0043] 图9为本申请实施例提供的语音交互方法中音频数据的采集与传送的流程图;

[0044] 图10为本申请实施例提供了一种语音交互方法的应用场景示意图;

[0045] 图11为本申请实施例提供了一种应用于第一客户端的语音交互装置的结构示意图;

[0046] 图12为本申请实施例提供了一种应用于第二客户端的语音交互装置的结构示意图;

[0047] 图13为本申请实施例提供了一种电子设备的结构示意图。

具体实施方式

[0048] 下面详细描述本申请的实施例,实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,仅用于解释本申请,而不能解释为对本申请的限制。

[0049] 本技术领域技术人员可以理解,除非特意声明,这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是,本申请的说明书中使用的措辞“包括”是指存在特征、整数、步骤、操作、元件和/或组件,但是并不排除存在或添加一个

或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。应该理解,当我们称元件被“连接”或“耦接”到另一元件时,它可以直接连接或耦接到其他元件,或者也可以存在中间元件。此外,这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和/或”包括一个或多个相关联的列出项的全部或任一单元和全部组合。

[0050] 为使本申请的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0051] 云计算(cloud computing)指IT基础设施的交付和使用模式,指通过网络以按需、易扩展的方式获得所需资源;广义云计算指服务的交付和使用模式,指通过网络以按需、易扩展的方式获得所需服务。这种服务可以是IT和软件、互联网相关,也可是其他服务。云计算是网格计算(Grid Computing)、分布式计算(Distributed Computing)、并行计算(Parallel Computing)、效用计算(Utility Computing)、网络存储(Network Storage Technologies)、虚拟化(Virtualization)、负载均衡(Load Balance)等传统计算机和网络技术发展融合的产物。

[0052] 随着互联网、实时数据流、连接设备多样化的发展,以及搜索服务、社会网络、移动商务和开放协作等需求的推动,云计算迅速发展起来。不同于以往的并行分布式计算,云计算的产生从理念上将推动整个互联网模式、企业管理模式发生革命性的变革。

[0053] 云会议是基于云计算技术的一种高效、便捷、低成本的会议形式。使用者只需要通过互联网界面,进行简单易用的操作,便可快速高效地与全球各地团队及客户同步分享语音、数据文件及视频,而会议中数据的传输、处理等复杂技术由云会议服务商帮助使用者进行操作。

[0054] 目前国内云会议主要集中在以SaaS(Software as a Service,软件即服务)模式为主体的服务内容,包括电话、网络、视频等服务形式,基于云计算的视频会议就叫云会议。

[0055] 在云会议时代,数据的传输、处理、存储全部由视频会议厂家的计算机资源处理,用户完全无需再购置昂贵的硬件和安装繁琐的软件,只需打开浏览器,登录相应界面,就能进行高效的远程会议。

[0056] 云会议系统支持多服务器动态集群部署,并提供多台高性能服务器,大大提升了会议稳定性、安全性、可用性。近年来,视频会议因能大幅提高沟通效率,持续降低沟通成本,带来内部管理水平升级,而获得众多用户欢迎,已广泛应用在政府、军队、交通、运输、金融、运营商、教育、企业等各个领域。毫无疑问,视频会议运用云计算以后,在方便性、快捷性、易用性上具有更强的吸引力,必将激发视频会议应用新高潮的到来。

[0057] 本申请所提及的在线教育(e-learning),又称远程教育、在线学习,是一种基于网络进行授课的行为,属于应用云计算技术的一种。

[0058] 下面以具体地实施例对本申请的技术方案以及本申请的技术方案如何解决上述技术问题进行详细说明。下面这几个具体的实施例可以相互结合,对于相同或相似的概念或过程可能在某些实施例中不再赘述。下面将结合附图,对本申请的实施例进行描述。

[0059] 在本申请实施例的语音交互方法中,该方法具体可以由本申请实施例的电子设备执行,具体地,该电子设备可以是客户端。客户端可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表等,但并不局限于此。在本申请实施时客户端与服务器可以通过有线或无线通信方式进行直接或间接地连接,本申请在此不做限制。

[0060] 如图1所示,本申请实施例提供的一种语音交互方法,应用于第一客户端时,包括以下步骤S101-S103:

[0061] 步骤S101:基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口。

[0062] 可选地,在线多方会话可以包括在线教育、在线会议等任何基于互联网进行远程交流的场景,本申请对此不作限定。具体地,在线多方会话区别于一对一的在线会话,在线多方会话可以包括两个用户的在线会话、也可以包括两个以上用户的在线会话,本申请实施例主要针对用户在两人以上的在线多方会话进行说明,但相关实施例也能在仅包括两个用户的在线会话中实施。

[0063] 具体地,在进行在线多方会话时,第一用户作为语音自由发言功能的控制方(可以是在在线多方会话的发起人、主持人等角色,该角色可进行转移),可以通过第一客户端在在线多方会话的界面中发起语音交互请求,该请求是针对在线多方会话的所有参与者提出,即针对除第一用户以外,当前在线多方会话中包括的所有用户(下称第二用户)。当第一用户通过第一客户端发起语音交互请求后,经相关处理,当前在线会话将进入语音自由发言模式,在该模式下,第一用户使用的第一客户端将显示目标窗口,且第二用户可自由发言以与第一用户或其他第二用户进行语音交流。

[0064] 步骤S102:接收由在线多方会话中包括的第二客户端采集的音频数据。

[0065] 可选地,在语音自由发言模式下,任一使用第二客户端的第二用户均可自由发现,在此期间,任一第二客户端均可采集一段或多段音频数据,本申请对任一第二用户的发言次数不作限制;但考虑到让更多的第二用户积极参与自由发言,可以限定任一第二客户端仅可采集一段音频数据,若第一段采集的音频数据无效,则视为未采集音频数据,直至采集到一段有效的音频数据。

[0066] 在一实施例中,考虑到发言的延时问题,可以设置在语音自由发言模式下,第二客户端在采集音频数据的同时,同步向第一客户端传送音频数据;即第一客户端接收音频数据的时间与第二客户端采集音频数据的时间同步。可选地,音频数据可以经由服务器完成中转,如由第二客户端采集音频数据并同步向服务器传送,由服务器向第一客户端传送。

[0067] 具体地,针对第一客户端接收音频数据的情况进行举例说明:假设当前在线多方会话中除第一用户外,还包括第二用户A、B和C,在语音自由发言模式下(以开启语音自由发言模式的时间为起始时间),第3秒,第二用户A首次触发第二客户端采集音频数据(持续有7秒);第6秒,第二用户C触发第二客户端采集音频数据(持续12秒);第23秒,第二用户A二次触发第二客户端采集音频数据(持续有5秒);针对上述情况,第一客户端将在第3秒开始接收到第二用户A发言对应的音频数据,并持续该接收状态至第10秒;第一客户端在接收第二用户A首次发言的音频数据的过程中,在第6秒时将同步接收第二用户C发言对应的音频数据,并持续该接收状态至第18秒;第一客户端将在第23秒接收到第二用户A二次发言对应的音频数据,并持续该接收状态至第28秒。在该例子可见,当在线多方会话进入语音自由发言模式时,各第二用户可在任意时间自由发言,也可不发言,同时各第二用户可在语音自由发言模式的时间内多次发言。

[0068] 在一实施例中,第一客户端接收到的音频数据还可以包括自身采集的音频数据,具体地,本申请实施例提供的语音交互方法还包括步骤:响应于语音交互指令,采集音频数据;将采集的音频数据作为接收的音频数据在所述在线多方会话中播放。

[0069] 可选地,第一用户触发语音交互指令表征第一用户进入发言状态,第一客户端将开始采集音频数据。具体地,通过第一客户端采集的音频数据在采集过程中无需上传服务器,可直接作为接收的音频数据,与所接收到的第二客户端采集的音频数据一并在在线多方会话中进行播放。其中,语音交互指令的触发操作可以包括:对预设的硬件按键的长按操作、在在线多方会话的界面上对发言功能控件的触控操作。

[0070] 针对第一客户端接收音频数据的情况进行举例说明:假设当前在线多方会话中除第一用户外,还包括第二用户D、E和F,在语音自由发言模式下(以开启语音自由发言模式的时间为起始时间),第2秒,第二用户E触发第二客户端采集音频数据(持续有8秒);第11秒,第一用户触发第一客户端采集音频数据(持续有3秒);第13秒,第二用户D触发第二客户端采集音频数据(持续有7秒);针对上述情况,第一客户端将在第2秒开始接收到第二用户E发言对应的音频数据,并持续该接收状态至第10秒;第一客户端在第11秒开始接收到自身采集的音频数据,并持续该接收状态至第14秒;在第一客户端接收自身采集的音频数据的过程中,在第13秒接收到第二用户D发言对应的音频数据,并持续该接收状态至第20秒。在该例子可见,当在线多方会话进入语音自由发言模式时,各用户可在任意时间自由发言,也可不发言(包括第一用户和第二用户);同时在第一客户端采集音频数据的同时可以接收第二客户端采集的音频数据。

[0071] 步骤S103:响应于对目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。

[0072] 具体地,第一客户端在接收到各段音频数据后,可以通过目标窗口提供的目标功能控件控制各段音频数据依次播放,各段音频数据的播放时间互不重叠,可有效解决现有技术中多个用户同时发言导致发言冲突的问题。可选地,依次播放可以理解为各段音频数据基于第一客户端的控制顺序进行逐一播放;还可以理解为各段音频数据基于第一客户端接收的时间进行逐一播放;还可以适应具体情况进行调整。

[0073] 其中,依次播放的效果可以体现为各个参与当前在线多方会话的用户均可在相同时间听到相同音频数据对应的发言内容。

[0074] 本申请提供的语音交互方法及相关设备中,应用于在线教育的场景时,老师可采用第一客户端实现语音交互;具体地,老师可通过第一客户端在在线多方会话的界面中发起语音交互请求,即在授课过程中,第一客户端可控制语言发言的时间节点,且通过一次请求即可开启参加当前在线多方会话所有学生的语言交互权限,简便了老师的操作;进而基于语音交互请求,显示目标窗口;在第一客户端接收到由在线多方会话中包括的至少两个第二客户端采集的音频数据时,老师可通过目标窗口提供的目标功能控件控制接收到的各段音频数据在在线多方会话中依次播放,可在提高交流效率的同时,有效避免多个学生同时发言导致的发言冲突问题。

[0075] 在一实施例中,如图2所示,步骤S101基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口,包括步骤S201-S202:

[0076] 步骤S201:响应于在在线多方会话的界面上对开启语音交互功能控件的触控操作,向服务器发起语音交互请求。

[0077] 具体地,如图3所示,图3示出了第一用户作为当前在线多方会话的主持者(控制者)对应的在线多方会话的界面,在该界面上包括用于开启语音交互功能的功能控件,当第

一用户触控该功能控件时,第一客户端将响应于该触控操作向服务器发起语音交互请求。在本申请实施例中,功能控件显示于界面的顶部位置,其还可以设置在界面的其他位置,本申请对此不作限定。

[0078] 步骤S202:显示用于控制音频数据进行播放的目标窗口。

[0079] 具体地,如图4所示,该目标窗口可在第一客户端向服务器发起语音交互请求后,经服务器进行相关处理后,进入语音自由发言模式时以弹窗的形式显示。

[0080] 在一实施例中,如图4所示,目标窗口包括以下至少一项:显示区域、设置区域;其中,显示区域用于显示第一客户端当前接收到的音频数据;设置区域用于设置每一段音频数据的时长,该时长对应于第一客户端预设的语音交互时长,且该时长为第二客户端所采集音频数据的最长时长。

[0081] 其中,目标窗口提供的目标功能控件至少包括跳过音频数据、清空音频数据、暂停播放、调整播放顺序和停止语音交互对应的一个功能控件,在下述实施例中将详细叙述各个功能控件的情况。

[0082] 在一实施例中,步骤S102接收由在线多方会话中包括的第二客户端采集的音频数据,包括以下步骤A1-A2:

[0083] 步骤A1:接收服务器发送的由在线多方会话中包括的第二客户端采集的音频数据。

[0084] 具体地,第一客户端基于在线多方会话向服务器发送语音交互请求,经服务器的相关处理后,当前在线多方会话进入语音交互模式,且由服务器向第二客户端下发语音交互权限后,由服务器对第二客户端采集的音频数据进行中转(第二客户端向服务器发送所采集的音频数据,服务器将该音频数据向第一客户端转达)。

[0085] 步骤A2:基于接收各段音频数据的时间,以时间轴的形式在目标窗口中显示各段音频数据;其中,音频数据的显示内容由以下至少一项组合而成:各段音频数据分别对应第二客户端的账户名称、语音波形略缩图和时长。

[0086] 具体地,以第一客户端接收各段音频数据的时间为准,在目标窗口的显示区域以时间轴的形式显示各段音频数据;可选地,各段音频数据可采用同一数据通道显示,也可以采用不同的数据通道显示(即,一条通道显示一段音频数据,如图4所示)。

[0087] 其中,音频数据在显示区域中显示第二客户端的账户名称时,有利于第一用户迅速了解该音频数据所属的第二用户;在显示区域中显示语音波形略缩图时,有利于第一用户了解正在播放或待播放音频数据的语音质量;在显示区域中显示时长时,有利于第一用户了解正在播放或待播放音频数据的时长,以判断是否跳过、暂停或清空该音频数据。

[0088] 在一实施例中,如图9所示,步骤A2基于接收各段音频数据的时间,以时间轴的形式在目标窗口显示各段音频数据,包括以下步骤B1-B3:

[0089] 步骤B1:若当前正在接收的音频数据的时长不小于预设时长,将该音频数据基于接收的时间写入播放队列,并继续接收该音频数据。

[0090] 具体地,本申请实施例考虑到若音频数据的时长过短,有可能第二用户在输入音频数据或第二客户端采集音频数据时发生错误,导致音频数据的质量较差,因此,设定预设时长为较短的时长(如,可以为2秒),以筛选出质量较差的音频数据。预设时长可以基于实际情况进行设定,本申请对此不作限定。

[0091] 可选地,播放队列中可以存储有已播放、待播放以及正在播放的音频数据,当第一客户端将某一段音频数据写入播放队列时,该音频数据将出现在显示区域。

[0092] 步骤B2:若当前正在接收的音频数据的时长少于预设时长,删除该音频数据。

[0093] 结合步骤B1和B2,针对第一客户端如何处理所接收的音频数据进行叙述。假设第一客户端在开启语音自由发言模式的第7秒接收到一段音频数据A,第一客户端将对该音频数据A进行缓存;若在持续接收该音频数据A的过程中,接收该音频数据A的时长达预设时长,则将该段音频数据A写入播放列表,若当前播放列表为空,则可直接开始播放音频数据A;若当前播放列表存有其他音频数据,则基于写入播放列表的顺序依次播放音频数据。若在持续接收该音频数据A的过程中,接收该音频数据A的时长未达预设时长则删除缓存中的音频书A。

[0094] 步骤B3:基于播放队列以时间轴的形式在目标窗口中显示各段音频数据。

[0095] 可选地,所有写入播放队列的音频数据均将显示在目标窗口的显示区域中,第一用户可观看到各段音频数据的相关信息并对各段音频数据进行控制。

[0096] 在一实施例中,步骤S103响应于对目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放,包括以下C1-C4中的至少一项:

[0097] C1:响应于对跳过音频数据功能控件的触控操作,暂停播放当前正在播放的音频数据,并播放下一段音频数据。

[0098] 可选地,在各段音频数据基于第一客户端接收的时间依次进行播放时,当第一用户对跳过音频数的功能控件实施触控操作时,将跳过当前正在播放的音频数据,并播放播放队列中的下一段音频数据。

[0099] 可选地,由于可能存在同一时间接收到多段音频数据的情况,当第一用户对跳过音频数的功能控件(如图4所示,跳过当前对应的控件)实施触控操作时,将获取同一时间接收到的其他音频数据进行播放,若同一时间没有其他音频数据,则在下一接收时间接收到的多段音频数据中随机选择一段音频数据进行播放。

[0100] C2:响应于对清空音频数据功能控件的触控操作,删除以下至少一项:当前正在播放的音频数据、当前选定的至少一段音频数据和当前接收到的所有音频数据。

[0101] 具体地,当第一用户对清空音频数据的功能控件(如图4所示,清空发言对应的控件)实施触控操作时,可以在目标窗口该功能控件的相应位置显示二级功能入口,二级功能入口可以包括以下至少一项:用于删除当前正在播放的音频数据的功能控件、用于删除当前选定的至少一段音频数据的功能控件、以及用于删除当前接收到的所有音频数据的功能控件。

[0102] C3:响应于对暂停播放功能控件的触控操作,暂停播放音频数据。

[0103] 具体地,当第一用户对暂停播放音频数据的功能控件(如图4所示,暂停播放对应的控件)实施触控操作时,在线多方会话将整体暂停对音频数据的播放。在暂停播放音频数据的时间段内,第一用户可以输入音频数据以在在线多方会话中播放,在第一用户输入音频数据时第二用户无法输入音频数据;在暂停播放音频数据的时间段内,若第一用户无发言请求,则第二用户仍可自由发言,且第一客户端仍同步接收音频数据。

[0104] C4:响应于对调整播放顺序功能控件的触控操作,调整当前选定的至少一段音频数据的播放顺序。

[0105] 具体地,当第一用户对调整播放顺序的功能控件实施触控操作时,将调整第一用户选定的音频数据的播放顺序。可选地,第一用户可以通过选定任一音频数据后,拖动该音频数据往前或往后,触控调整播放顺序的功能控件以调整该音频数据的播放顺序。可选地,第一用户还可以通过选定任意两段音频数据后,触控调整播放顺序的功能控件以调换两端音频数据的播放顺序。

[0106] 在一实施例中,步骤S103控制接收到的各段音频数据依次播放之后,还包括步骤D1:

[0107] 步骤D1:响应于对停止语音交互功能控件的触控操作,向服务器发起停止语音交互请求,以停止在线多方会话中正在进行的语音交互行为;其中,停止的语音交互行为包括:停止播放音频数据、关闭目标窗口、删除当前接收的所有音频数据。

[0108] 具体地,当第一用户对停止语音交互的功能控件(如图4所示,停止发言对应的控件)实施触控操作时,第一客户端将向服务器发起停止语音交互请求,经服务器进行相关处理后,停止在线多方会话中正在进行的语音交互行为。

[0109] 基于同一发明构思,本申请实施例还提供一种语音交互方法,应用于第二客户端,如图5所示,包括以下步骤S501-S503:

[0110] 步骤S501:在在线多方会话的界面显示第一提示信息;第一提示信息用于提醒用户进行语音交互。

[0111] 可选地,如图6所示,为减少由于第一客户端开启语音自由发言模式而造成对第二客户端的干扰,将第一提示信息显示于在线多方会话的左下方;同时为了加强第一提示信息对第二用户的提醒作用,第一提示信息可基于预设频率闪烁显示(如在1秒中,前0.5秒显示第一提示信息,后0.5秒不显示第一提示信息)。

[0112] 在一可行的实施例中,还可以以弹窗的形式显示第一提示信息,也可以在显示第一提示信息的同时播放提示音。

[0113] 具体地,在显示第一提示信息之前,第二客户端可以接收根据第一客户端基于当前在线多方会话发起的语音交互请求下发的语音交互权限;语音交互权限由服务器向第二客户端下发,服务器向第二客户端下发语音交互权限可理解为服务器向第二客户端发送相关开启语音自由发言模式的信息,第二客户端在接收到该消息后,开启语音自由发言模式。在一实施例中,当服务器向第二客户端下发语音交互权限时,将在整个在线多方会话中同步开启语音自由发言模式,即第一客户端将正式进行语音自由发言模式,在此之前,第一客户端自发起语音交互请求起至正式进行语音自由发言模式期间处于准备阶段。

[0114] 步骤S502:响应于语音交互指令,采集音频数据并上传该音频数据,以使该音频数据与在线多方会话中的其他待播放的音频数据一并进行依次播放。

[0115] 可选地,第二用户触发语音交互指令表征第二用户进入发言状态,第二客户端将连麦采集音频数据。

[0116] 在一实施例中,上传音频数据可以理解为第二客户端将音频数据上传至服务器或向第一客户端传送。具体地,第二客户端采集音频数据的同时将音频数据向服务器发送,以使服务器将该音频数据同步转达至第一客户端中进行播放。

[0117] 具体地,其他待播放的音频数据可以包括其他第二客户端采集的音频数据、第一客户端采集的音频数据、第二客户端采集的其他音频数据中的至少一种。将第二客户端当

前采集的音频数据A上传后,该音频数据A将进入在线多方会话的待播放音频数据的播放队列中,与播放队列中待播放的音频数据一并进行依次播放。可选地,音频数据A进入播放队列的位置可以基于上传时间确定,也可以由接收端(可以理解为第一客户端)控制,按照实际需求控制音频数据A在播放队列中的位置,进而基于播放队列中各段音频数据的排序位置逐一播放各段音频数据。

[0118] 在一实施例中,语音交互方法还包括步骤S503:

[0119] 步骤S503:播放当前接收到的目标音频数据,目标音频数据包括第一客户端控制所接收到的音频数据依次播放的音频数据。

[0120] 可选地,在本申请实施例中并未限定步骤S503与步骤S502的先后执行顺序,可以先执行步骤S502,也可以先执行步骤S503,还可以两个步骤同时执行,即在采集音频数据的同时还可以播放目标音频数据(此时,正在发言的第二用户还可以适应正在播放的目标音频数据的内容调整当前发言的内容)。在一可行的实施例中,若执行步骤S502之前,正在播放目标音频数据,则可以在执行步骤S502采集音频数据时暂停播放目标音频数据,并在采集音频数据结束后,从结束的时间点起播放目标音频数据。举例说明:当在线多方会话进入语音自由发言模式时,第二客户端正在播放目标音频数据,当播放至第7秒时,执行步骤S502采集音频数据,此时将暂停目标音频数据的播放并持续采集音频数据的状态;此时采集音频数据的状态维持共6秒后结束,则此时将从目标音频数据第13秒所对应的位置继续播放该目标音频数据。

[0121] 具体地,当第一客户端接收到各段音频数据时,将控制各段音频数据依次播放,各段音频数据在第一客户端体现为依次播放的同时,将向服务器发送播放消息,由服务器将相关音频数据下发至各个第二客户端,以使第二客户端同步播放相关的音频数据,实现在在线多方会话中依据第一客户端的控制依次对各段音频数据进行播放。

[0122] 本申请实施例提供的语音交互方法中,应用于在线教育的场景时,学生可采用第二客户端实现语言交互;具体地,当老师通过第一客户端向服务器基于在线多方会话发起语音交互请求时,服务器将会根据该请求向第二客户端下发语音交互权限,此时,学生可基于个人需求触发语音交互指令进行自由发言,使得第二客户端采集音频数据并同步传送至服务器,由服务器将该音频数据同步向第一客户端传送;第二客户端在接收到目标音频数据后进行播放,该目标音频数据包括第一客户端控制所接收到的音频数据依次播放的音频数据。本申请的实施可有效提高授课期间学生通过语音与老师进行交流的效率。

[0123] 在一实施例中,步骤S502中的响应于语音交互指令包括以下F1-F2中的至少一项:

[0124] F1:响应于对预设的硬件按键的长按操作,开启音频数据的采集功能。

[0125] 具体地,通过预设的硬件按键可以是基于在线多方会话的应用程序进行设置,如图7所示,硬件按键可以是F2。其中,长按操作可理解为长按硬件按键,若松开则为指示停止采集音频数据。

[0126] F2:响应于在在线多方会话的界面上对发言功能控件的触控操作,开启音频数据的采集功能。

[0127] 具体地,可在在线多方会话的界面上设置第二用户用于发言的功能控件,当第二用户触控该功能控件时开启音频数据的采集功能;相应地,可以同时弹窗显示提示窗口(如图7所示),当第二用户触控提示窗口上设置的停止发言的功能控件时,退出提示窗口的显

示并停止采集音频数据。

[0128] 在一实施例中,在步骤S501在在线多方会话的界面显示第一提示西悉尼之前,还包括步骤接收根据第一客户端在在线多方会话的界面中发起的语音交互请求下发的语音交互权限。语音交互权限中携带有预设的语音交互时长;可选地,该语音交互时长可以由第一客户端通过目标窗口的设置区域设置。具体地,步骤S502中的采集音频数据并上传该音频数据,还包括以下步骤G1-G2中的至少一项:

[0129] 步骤G1:当正在采集的音频数据的时长达预设的语音交互时长时,中断对该音频数据的采集和/或关闭采集功能。

[0130] 具体地,步骤G1中的语音交互时长与上述实施例中的步骤S202中目标窗口中设置区域用于设置每一段音频数据的时长相应。其中,中断对音频数据的采集和/或关闭采集功能为强制性停止音频数据的采集功能。本申请实施例有利于控制每一位第二用户的发言时间,避免由单一第二用户长时间发言而导致其他第二用户无法积极参与到交流中。

[0131] 步骤G2:当正在采集音频数据时,响应于停止语音交互指令,中断对该音频数据的采集和/或关闭采集功能,若该音频数据的时长小于预设时长,显示语音交互失败的提示信息。

[0132] 具体地,步骤G2中的中断或关闭对音频数据的采集功能与上述实施例中的步骤B3相应。其中,显示语音交互失败的提示信息可采用与图6或图7所示的提示方式进行,以提醒第二用户当前的发言无效将不会在在线多方会话中进行播放。

[0133] 在一实施例中,如图7所示,步骤S502响应于语音交互指令,采集音频数据并上传该音频数据,还包括步骤H1:

[0134] 步骤H1:显示提示窗口,提示窗口包括用于显示所采集音频数据时长的提示框。

[0135] 具体地,提示框中显示有所采集音频数据的时长“正在录(音23秒)”,还可以显示有当前用户按下的硬件按键。

[0136] 在一实施例中,当开启语音自由发言模式时,无需用户触发语音交互指令则直接在在线多方会话的界面将显示有提示窗口,当第二用户未进行发言时,图7中“发言中”的字样将显示为“待发言”等字样,当第二用户触发语音交互指令时,该字样将改变为“发言中”,以有效提醒第二用户可进行自由发言。

[0137] 可选地,步骤S503中的播放当前接收到的目标音频数据,还包括:显示播放窗口。具体地,播放窗口包括用于显示当前正在播放的目标音频数据的区域,在显示该目标音频数据时可以同步显示采集该目标音频数据的第二客户端的账户信息(如账户名称)。基于此,在播放当前接收到的目标音频数据时,用户可以通过播放窗口了解当前正在发言的用户信息,以提高语音交互的效率。

[0138] 在一实施例中,步骤S503中的播放当前接收到的目标音频数据之后,还包括步骤I1和步骤I2:

[0139] 步骤I1:接收服务器根据第一客户端发起的停止语音交互请求对语音交互权限进行回收的第二提示信息。

[0140] 步骤I2:显示第二提示信息或基于第二提示信息停止显示第一提示信息。

[0141] 可选地,步骤I1中第一客户端发起停止语音交互请求与上述实施例中的步骤D1中的向服务器发起停止语音交互请求相应。

[0142] 具体地,当第一用户通过第一客户端发起停止语音交互请求后,经服务器回收第二客户端语音交互的权限后,将关闭当前的语音自由发言模式,第二客户端将不再显示第一提示信息。其中,在服务器回收语音交互权限时向第二客户端下发第二提示信息,以使第二客户端显示该第二提示信息或基于该第二提示信息关闭语音自由发言模式并停止显示第一提示信息。

[0143] 下面结合图8、图9和图10对上述实施例描述的语音交互方法进行叙述,具体以在线教育的场景进行叙述。

[0144] 在整个语音交互系统中,包括老师使用的第一客户端1001、学生使用的第二客户端1003以及服务器1002,三者可有线或无线通信连接。其中,服务器扮演着对信息进行处理和中转的角色;第一客户端为语音交互系统中的控制端,可有效控制语音自由发言模式的开启与关闭;第二客户端为语音交互系统中的参与者,可在语音自由发言模式下采集音频数据并通过服务器传送至第一客户端,以在在线多方会话中播放。

[0145] 其中,在线多方会话在在线教育的场景中体现为授课过程,老师在指定时间通过在线平台的某一课堂进行授课,学生在指定时间进入在线平台对应的课堂进行上课。过程中,涉及到需要学生自由发言讨论的环节时,老师可开启语音自由发言模式,加入该课堂的所有学生可在该模式下自由发言(无限定发言的起始时间)。当学生发言时,第二客户端将基于第二用户的相关触控操作(长按F2按键),采集音频数据并经服务器同步传送至第一客户端,由第一客户端控制所接收到的音频数据在在线多方会话中依次播放;当学生松开F2按键时,第二客户端关闭音频数据的采集功能,并向服务器发送发言结束信令,由服务器通知第一客户端第二用户发言结束。

[0146] 在上述实施例中,触控操作可以理解为接触式触控操作、也可以理解为非接触式触控操作;在一可行的实施例中,触控操作还包括通过鼠标、键盘、触控笔等输入设备对各项功能控件的触发操作(如通过鼠标进行的点击操作、通过键盘进行的长按操作)。

[0147] 在一实施例,如图11所示,提供了一种语音交互装置110,应用于第一客户端,包括:窗口显示模块111、接收模块112、控制模块113。其中:窗口显示模块111用于基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口;接收模块112,用于接收由在在线多方会话中包括的第二客户端采集的音频数据;控制模块113,用于响应于对目标窗口提供的目标功能控件的触控操作,控制接收到的各段音频数据依次播放。

[0148] 可选地,窗口显示模块111包括:请求单元,用于响应于在当前在线多方会话的界面上对开启语音交互功能控件的触控操作,向服务器发起语音交互请求;第一显示单元,用于显示用于控制音频数据进行播放的目标窗口。

[0149] 可选地,目标窗口包括以下至少一项:显示区域,用于显示当前接收到的音频数据;设置区域,用于设置每一段音频数据的语音交互时长;目标功能控件,用于控制音频数据进行播放;其中,目标功能控件至少包括跳过音频数据、清空音频数据、暂停播放、调整播放顺序和停止语音交互中对应的一个功能控件。

[0150] 可选地,接收模块112,包括:接收单元,用于接收服务器发送的由所述在线多方会话中包括的第二客户端采集的音频数据;第二显示单元,用于基于接收各段音频数据的时间,以时间轴的形式在目标窗口中显示各段音频数据;音频数据的显示内容由以下至少一项组合而成:各段音频数据分别对应第二客户端的账户名称、语音波形略缩图和时长。

[0151] 可选地第二显示单元包括：第一接收子单元，用于若当前正在接收的音频数据的时长不小于预设时长，将该音频数据基于接收的时间写入播放队列，并继续接收该音频数据；第二接收子单元，用于若当前正在接收的音频数据的时长小于预设时长，删除该音频数据；显示子单元，用于基于播放队列在显示区域以时间轴的形式在目标窗口显示各段音频数据。

[0152] 可选地，控制模块113至少包括以下一个单元：第一暂停单元，用于响应于对跳过音频数据功能控件的触控操作，暂停播放当前正在播放的音频数据，并播放下一段音频数据；删除单元，用于响应于对清空音频数据功能控件的触控操作，删除以下至少一项：当前正在播放的音频数据、当前选定的至少一段音频数据和当前接收到的所有音频数据；第二暂停单元，用于响应于对暂停播放功能控件的触控操作，暂停播放音频数据；调整单元，用于响应于对调整播放顺序功能控件的触控操作，调整当前选定的至少一段音频数据的播放顺序。

[0153] 可选地，语音交互装置110还包括停止模块，用于响应于对停止语音交互功能控件的触控操作，向服务器发起停止语音交互请求，以停止在线多方会话中正在进行的语音交互行为；其中，停止的语音交互行为包括以下至少一项：停止播放音频数据、关闭目标窗口、删除当前接收的所有音频数据。

[0154] 在一实施例，如图12所示，提供了一种语音交互装置120，应用于第二客户端，包括：显示模块121、采集模块122。其中：显示模块121，用于在在线多方会话的界面显示第一提示信息；所述第一提示信息用于提醒用户进行语音交互；采集模块122，用于响应于语音交互指令，采集音频数据并上传该音频数据，以使该音频数据与所述在线多方会话中的与其他待播放的音频数据一并进行依次播放。

[0155] 可选地，装置120还包括播放模块，用于播放当前接收到的目标音频数据，所述目标音频数据包括第一客户端控制所接收到的音频数据依次播放的音频数据。

[0156] 可选地，采集模块122至少包括以下一个单元：第一开启单元，用于响应于对预设的硬件按键的长按操作，开启音频数据的采集功能；第二开启单元，用于响应于在在线多方会话的界面上对发言功能控件的触控操作，开启音频数据的采集功能。

[0157] 可选地，显示模块121还用于接收根据第一客户端在在线多方会话的界面中发起的语音交互请求下发的语音交互权限；语音交互权限中携带有预设的语音交互时长；采集模块122还包括以下至少一个单元：第一处理单元，用于当正在采集的音频数据的时长达预设的语音交互时长时，中断对该音频数据的采集和/或关闭采集功能；第二处理单元，用于当正在采集音频数据时，响应于停止语音交互指令，中断对该音频数据的采集和/或关闭采集功能，若该音频数据的时长小于预设时长，显示语音交互失败的提示信息。

[0158] 可选地，采集模块122还用于显示提示窗口，提示窗口包括用于显示所采集音频数据时长的提示框。

[0159] 可选地，播放模块还用于显示播放窗口，播放窗口包括用于显示当前正在播放的目标音频数据的区域。

[0160] 本申请实施例的语音交互装置可执行本申请的实施例所提供的一种语音交互方法，其实现原理相类似，本申请各实施例中的语音交互装置中的各模块所执行的动作是与本申请各实施例中的语音交互方法中的步骤相对应的，对于语音交互装置的各模块的详细

功能描述具体可以参见前文中所示的对应的语音交互方法中的描述,此处不再赘述。

[0161] 基于与本申请的实施例中所示的方法相同的原理,本申请的实施例中还提供了一种电子设备,该电子设备可以包括但不限于:处理器和存储器;存储器,用于存储计算机操作指令;处理器,用于通过调用计算机操作指令执行实施例所示的语音交互方法。

[0162] 在一个可选实施例中提供了一种电子设备,如图13所示,图13所示的电子设备1300包括:处理器1301和存储器1303。其中,处理器1301和存储器1303相连,如通过总线1302相连。可选地,电子设备1300还可以包括收发器1304。需要说明的是,实际应用中收发器1304不限于一个,该电子设备1300的结构并不构成对本申请实施例的限定。

[0163] 处理器1301可以是CPU(Central Processing Unit,中央处理器),通用处理器,DSP(Digital Signal Processor,数据信号处理器),ASIC(Application Specific Integrated Circuit,专用集成电路),FPGA(Field Programmable Gate Array,现场可编程门阵列)或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请内容所描述的各种示例性的逻辑方框,模块和电路。处理器1301也可以是实现计算功能的组合,例如包含一个或多个微处理器组合,DSP和微处理器的组合等。处理器1301通过调用计算机操作指令执行上述实施例所示的语音交互方法,该语音交互方法包括:基于当前在线多方会话发起语音交互请求;接收由在线多方会话中包括的至少两个第二客户端采集的音频数据;控制接收到的各段音频数据在在线多方会话中依次播放。

[0164] 总线1302可包括一通路,在上述组件之间传送信息。总线1302可以是PCI(Peripheral Component Interconnect,外设部件互连标准)总线或EISA(Extended Industry Standard Architecture,扩展工业标准结构)总线等。总线1302可以分为地址总线、数据总线、控制总线等。为便于表示,图13中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0165] 存储器1303可以是ROM(Read Only Memory,只读存储器)或可存储静态信息和指令的其他类型的静态存储设备,RAM(Random Access Memory,随机存取存储器)或者可存储信息和指令的其他类型的动态存储设备,也可以是EEPROM(Electrically Erasable Programmable Read Only Memory,电可擦可编程只读存储器)、CD-ROM(Compact Disc Read Only Memory,只读光盘)或其他光盘存储、光碟存储(包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等)、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质,但不限于此。

[0166] 存储器1303用于存储执行本申请方案的应用程序代码,并由处理器1301来控制执行。处理器1301用于执行存储器1303中存储的应用程序代码,以实现前述方法实施例所示的内容。

[0167] 其中,电子设备包括但不限于:移动电话、笔记本电脑、数字广播接收器、PDA(个人数字助理)、PAD(平板电脑)、PMP(便携式多媒体播放器)、车载终端(例如车载导航终端)等等的移动终端以及诸如数字TV、台式计算机等等的固定终端。图13示出的电子设备仅仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0168] 根据本申请的一个方面,提供了一种计算机程序产品或计算机程序,该计算机程

序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述各种可选实现方式中提供的语音交互方法。

[0169] 本申请实施例提供了一种计算机可读存储介质,该计算机可读存储介质上存储有计算机程序,当其在计算机上运行时,使得计算机可以执行前述方法实施例中相应内容。

[0170] 应该理解的是,虽然附图的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,其可以以其他的顺序执行。而且,附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,其执行顺序也不必然是依次进行,而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0171] 需要说明的是,本申请上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本申请中,计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:电线、光缆、RF(射频)等等,或者上述的任意合适的组合。

[0172] 上述计算机可读介质可以是上述电子设备中所包含的;也可以是单独存在,而未装配入该电子设备中。

[0173] 上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被该电子设备执行时,使得该电子设备执行上述实施例所示的方法。

[0174] 可以以一种或多种程序设计语言或其组合来编写用于执行本申请的操作的计算机程序代码,上述程序设计语言包括面向对象的程序设计语言——诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言——诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0175] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、方法和计算机程

序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的是,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0176] 描述于本申请实施例中所涉及到的模块可以通过软件的方式实现,也可以通过硬件的方式来实现。其中,模块的名称在某种情况下并不构成对该模块本身的限定,例如,窗口显示模块还可以被描述为“用于基于在在线多方会话的界面中发起的语音交互请求,显示目标窗口的模块”。

[0177] 以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解,本申请中所涉及的公开范围,并不限于上述技术特征的特定组合而成的技术方案,同时也应涵盖在不脱离上述公开构思的情况下,由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

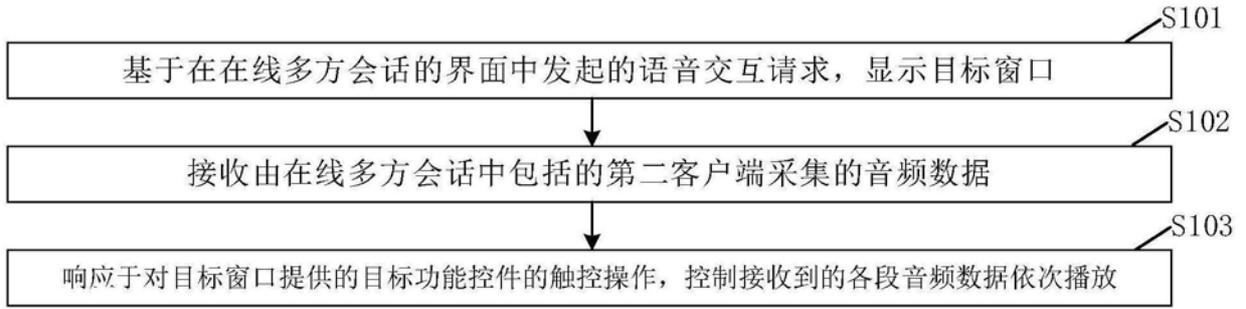


图1

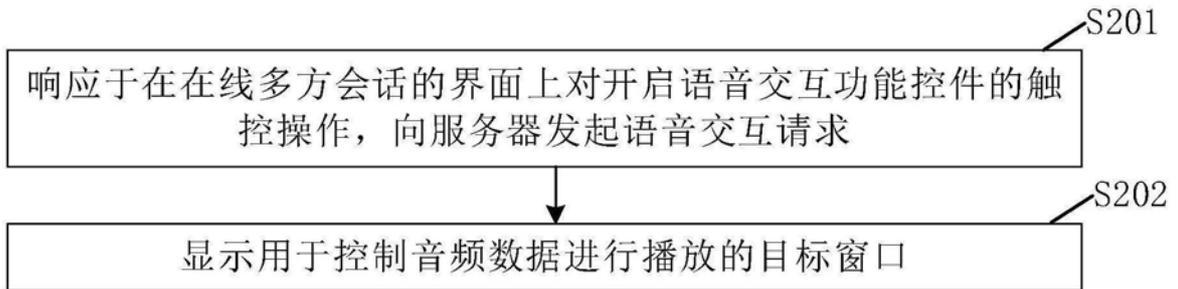


图2

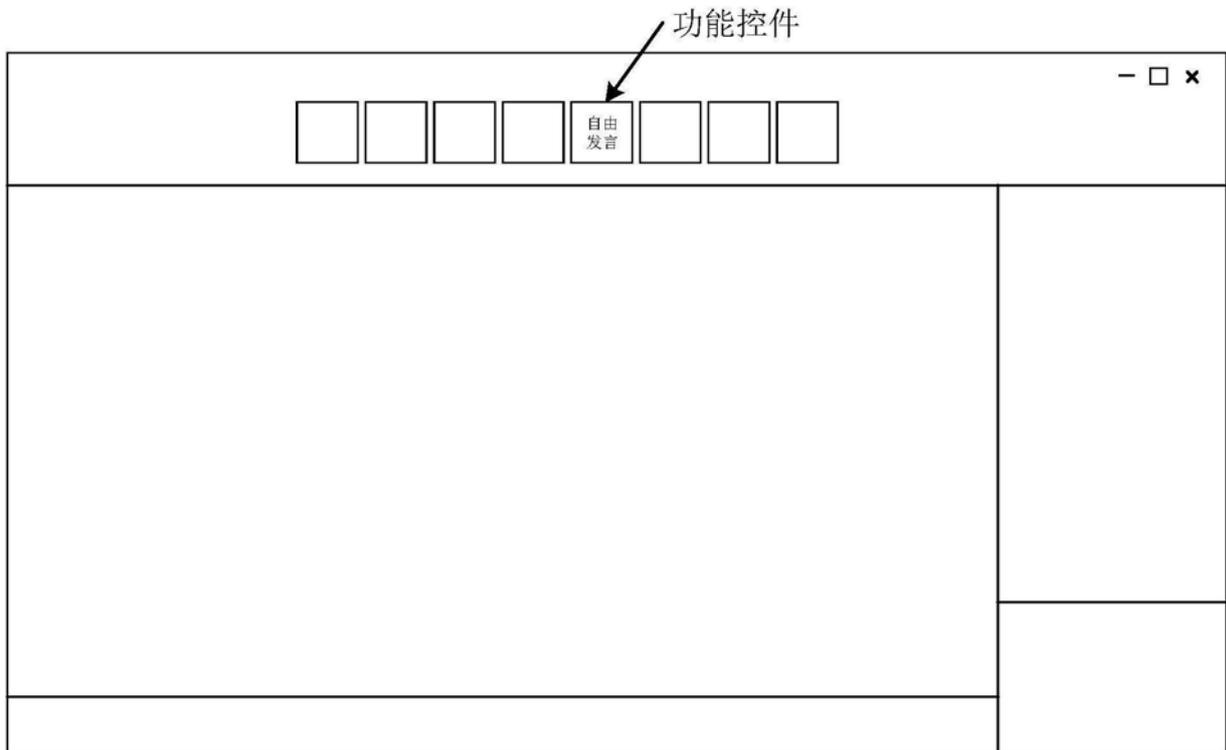


图3

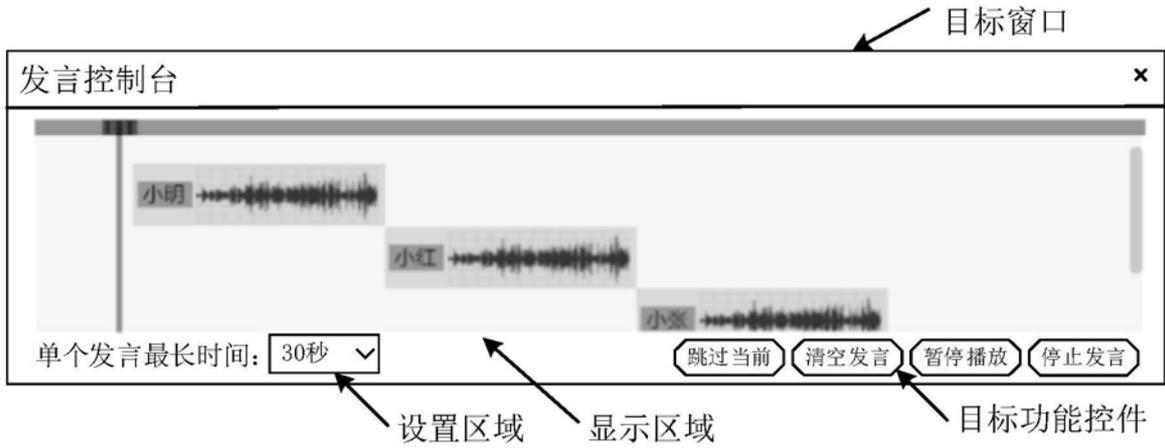


图4

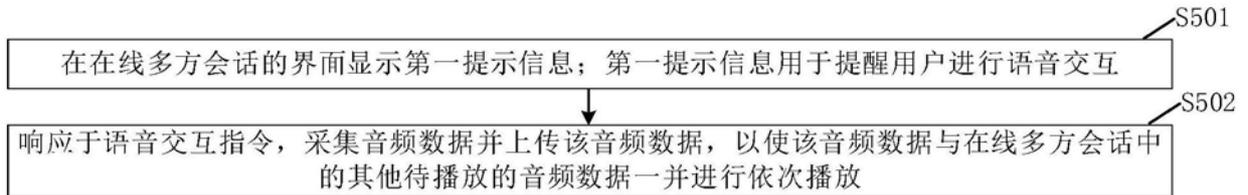


图5

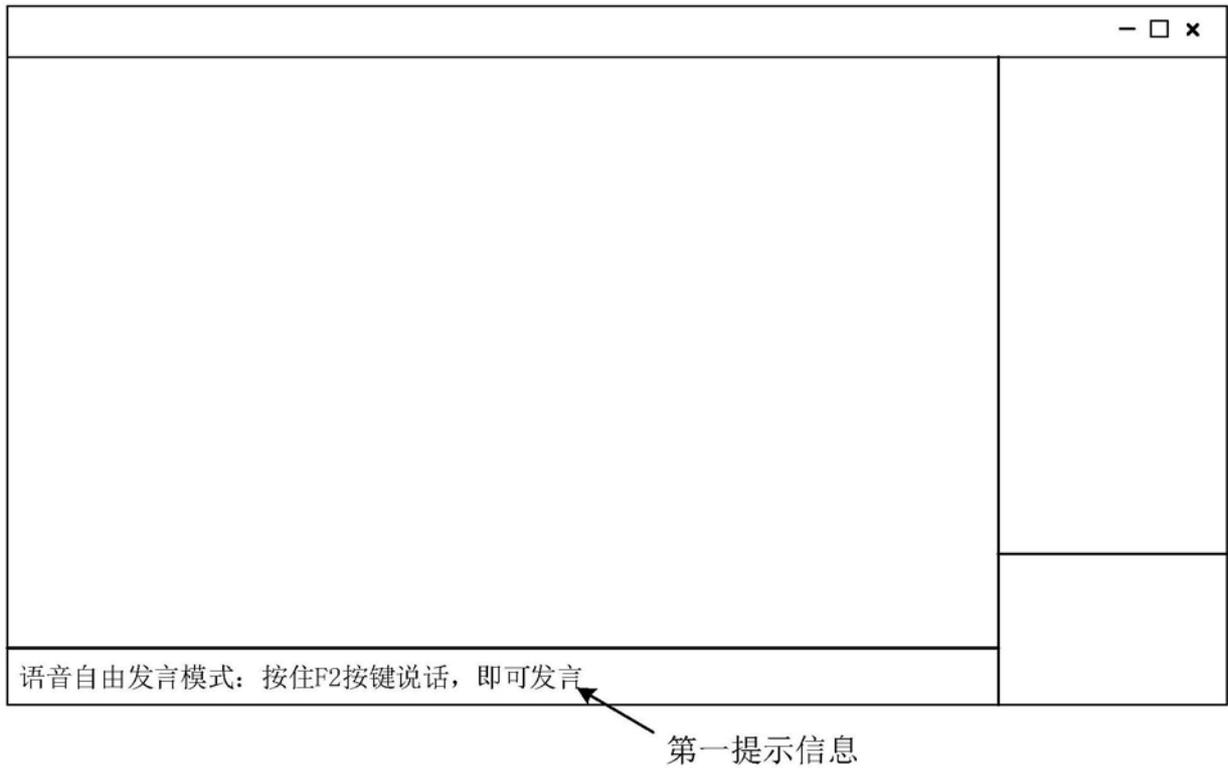


图6



图7

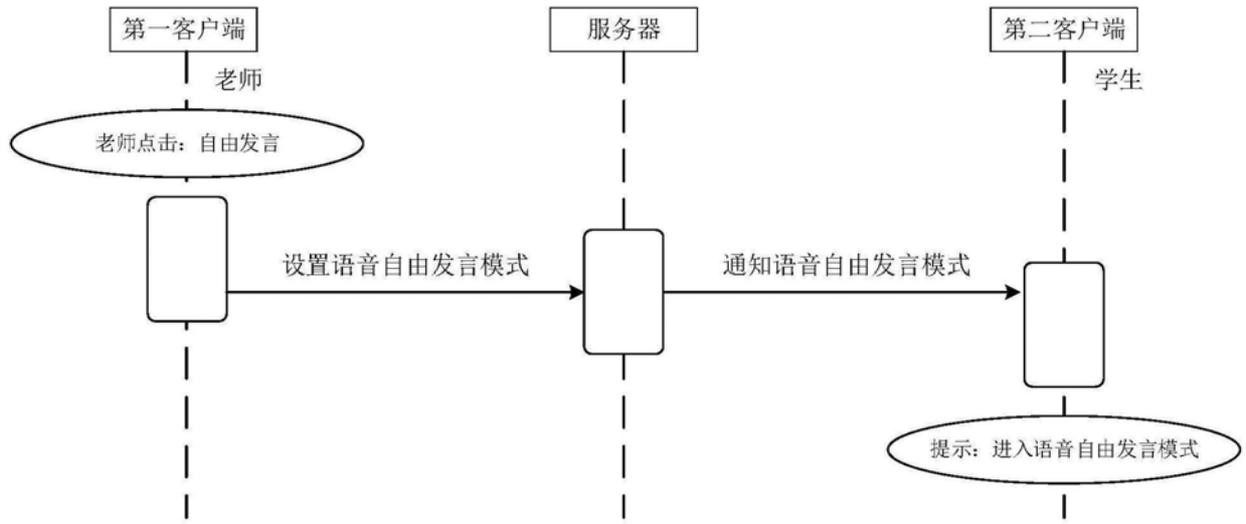


图8

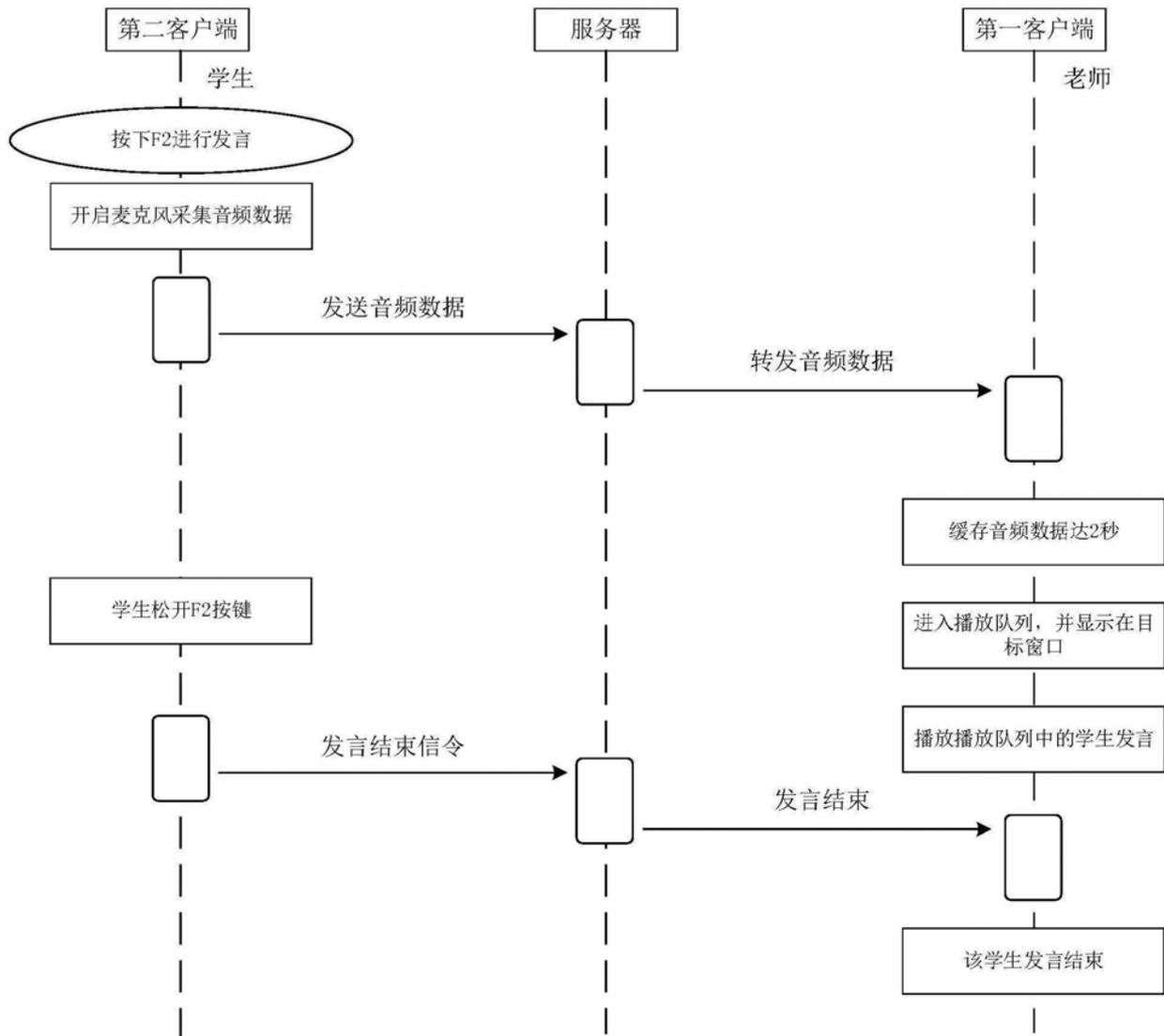


图9

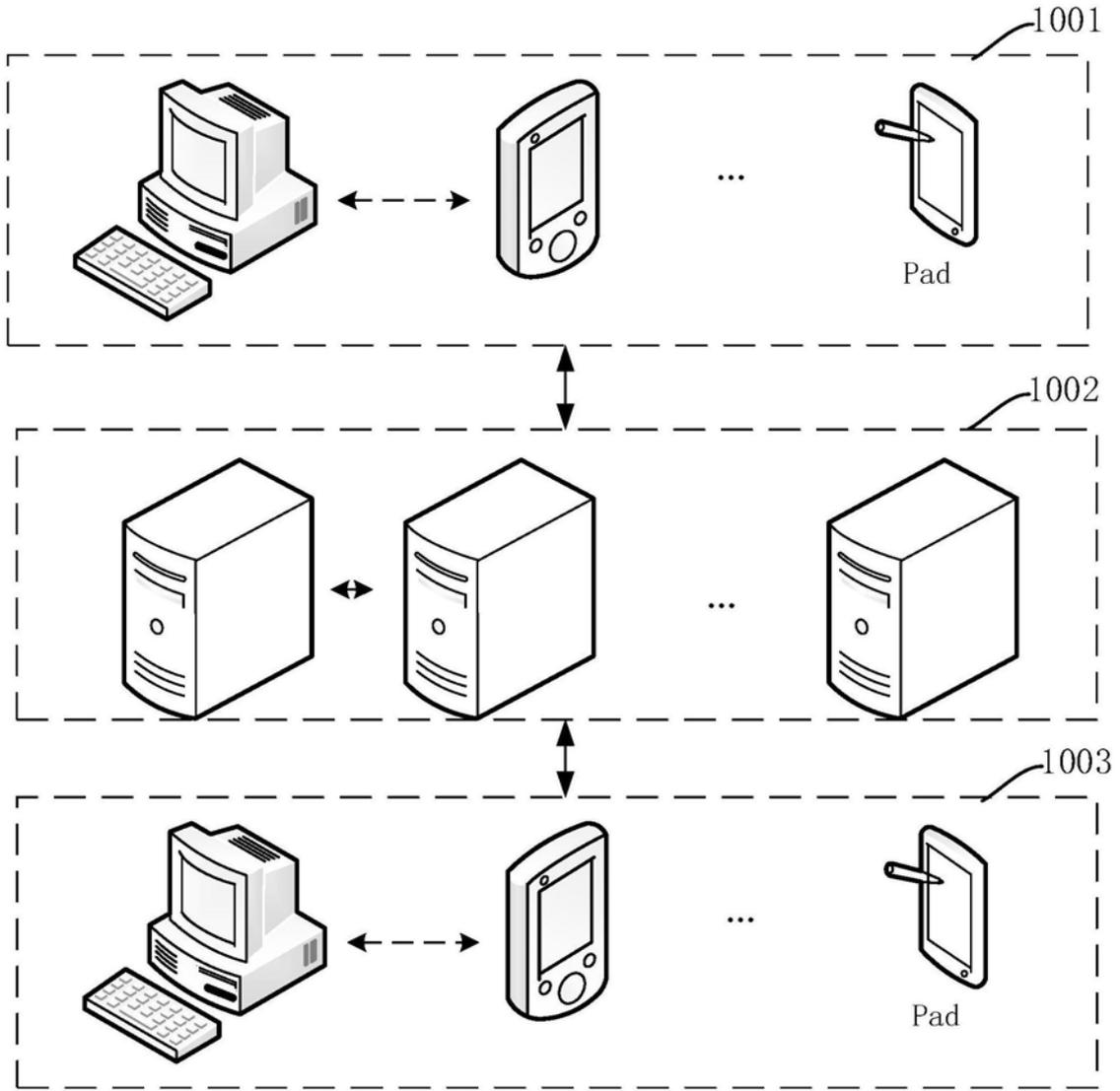


图10

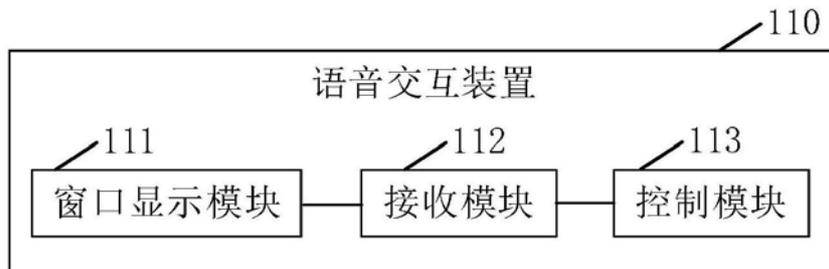


图11

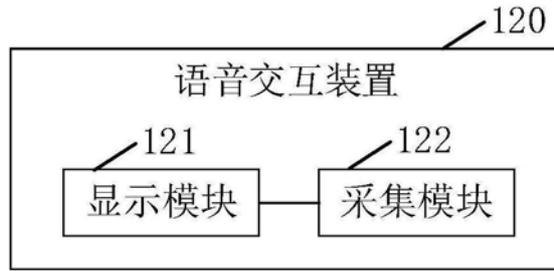


图12

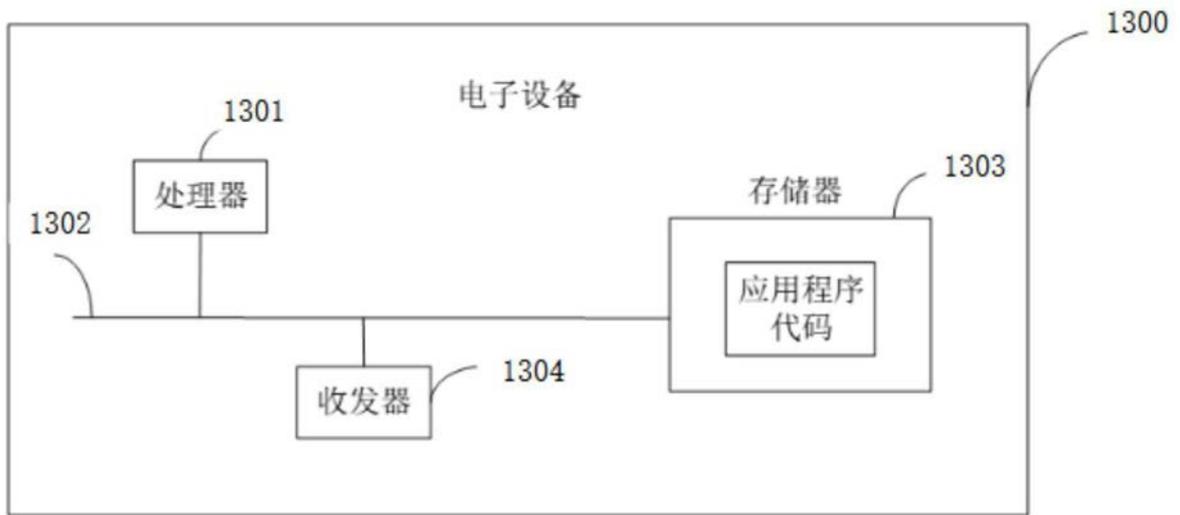


图13