



(12) 发明专利

(10) 授权公告号 CN 113689854 B

(45) 授权公告日 2024. 01. 23

(21) 申请号 202110923097.3

G10L 15/26 (2006.01)

(22) 申请日 2021.08.12

G10L 15/30 (2013.01)

(65) 同一申请的已公布的文献号

G10L 15/04 (2013.01)

申请公布号 CN 113689854 A

H04L 67/146 (2022.01)

G06Q 10/0639 (2023.01)

(43) 申请公布日 2021.11.23

(56) 对比文件

(73) 专利权人 深圳追一科技有限公司

CN 106952646 A, 2017.07.14

地址 518057 广东省深圳市南山区粤海街

CN 112863526 A, 2021.05.28

道科技园社区科苑路8号讯美科技广

CN 111369990 A, 2020.07.03

场3号楼23A、23B

CN 104246750 A, 2014.12.24

(72) 发明人 吕子谋 常向月 杨晓璐

CN 104616652 A, 2015.05.13

(74) 专利代理机构 华进联合专利商标代理有限

CN 109819116 A, 2019.05.28

公司 44224

CN 110798327 A, 2020.02.14

专利代理师 陈小娜

CN 112201222 A, 2021.01.08

CN 112312064 A, 2021.02.02

(51) Int. Cl.

审查员 张鹤

G10L 15/22 (2006.01)

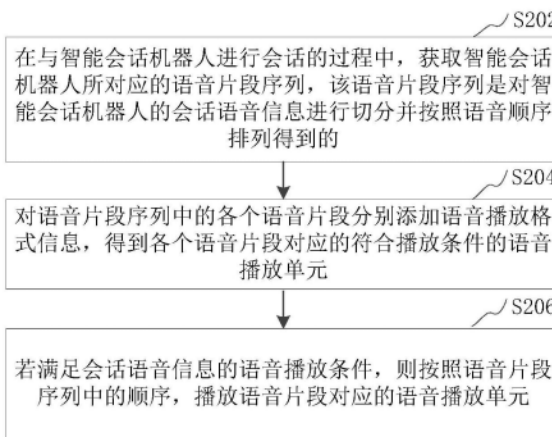
权利要求书3页 说明书19页 附图6页

(54) 发明名称

语音会话方法、装置、计算机设备和存储介质

(57) 摘要

本申请涉及一种语音会话方法、装置、计算机设备和存储介质。所述方法包括：在与智能会话机器人进行会话的过程中，获取所述智能会话机器人所对应的语音片段序列，所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的；对所述语音片段序列中的各个语音片段分别添加语音播放格式信息，得到各个所述语音片段对应的符合播放条件的语音播放单元；当满足所述会话语音信息的语音播放条件时，按照所述语音片段序列中的顺序，播放所述语音片段对应的语音播放单元。采用本方法能够提高语音会话流畅度。



1. 一种语音会话方法,其特征在于,所述方法包括:

在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列,所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;切分所述会话语音信息的步骤包括:获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段;

对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;

若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;所述语音播放条件是获取到针对语音提示信息的触发操作;

若所述当前语音播放单元播放失败,则获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序,从服务器获取所述目标片段顺序对应的语音片段,作为目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息,所述语音补充提示信息用于提示语音已经进行了更新。

2. 根据权利要求1所述的方法,其特征在于,所述从服务器获取所述目标片段顺序对应的语音片段,作为目标语音片段,包括:

向服务器发送语音片段获取请求,所述语音片段获取请求携带所述目标片段顺序以及所述会话语音信息的会话消息标识,以使得所述服务器响应于所述语音片段获取请求,根据所述会话消息标识从存储的所述语音片段序列中提取所述目标片段顺序对应的语音片段,作为目标语音片段;

接收所述服务器返回的所述目标语音片段;

所述方法还包括:

若接收到针对所述会话语音信息的播放操作,则按照所述语音片段序列中的顺序,播放所述语音片段对应的更新后的语音播放单元。

3. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

响应于针对所述会话语音信息的播放中断操作,停止播放所述语音播放单元并记录停止位置;

获取所述会话语音信息对应的文本信息;

在会话界面上展示所述文本信息,在所述文本信息中标记所述停止位置所对应的文本。

4. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

展示培训会话对应的培训会话界面;

获取所述培训会话对应的会话标识;

基于所述培训会话的会话标识建立所述培训会话界面与长连接网关之间的连接。

5. 根据权利要求4所述的方法,其特征在于,所述培训会话中的会话节点包括智能会话

机器人;所述在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列包括:

在与智能会话机器人进行会话的过程中,通过所述培训会话界面获取所述长连接网关通过实时对话语音通道传输的所述智能会话机器人的语音片段序列。

6. 根据权利要求4所述的方法,其特征在于,所述培训会话中的会话节点还包括培训对象,所述方法还包括:

通过所述培训会话界面采集所述培训对象的实时语音信息,将所述实时语音信息通过语音传输通道传输至长连接网关;其中,所述长连接网关通过所述会话标识将所述实时语音信息传输至对话控制平台,所述对话控制平台对所述实时语音信息进行语音识别,得到语音文本信息,将所述语音文本信息传输至培训对话服务,所述培训对话服务调用语义分析平台对所述语音文本信息进行会话质检,得到质检结果;

通过所述培训会话界面接收所述长连接网关通过语义分析通道传输的所述语音文本信息以及所述质检结果。

7. 一种语音会话方法,其特征在于,所述方法包括:

在智能会话机器人进行会话的过程中,获取所述智能会话机器人对应的会话语音信息;

获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;

根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;

基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段;

将所述语音片段序列发送给所述智能会话机器人对应的会话终端,以使得所述会话终端对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,所述会话终端按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;所述语音播放条件是获取到针对语音提示信息的触发操作;若所述当前语音播放单元播放失败,则所述会话终端获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序,从服务器获取所述目标片段顺序对应的语音片段,作为目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息,所述语音补充提示信息用于提示语音已经进行了更新。

8. 一种语音会话装置,其特征在于,所述装置包括:

语音片段序列获取模块,用于在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列,所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;切分所述会话语音信息的步骤包括:获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段;

语音播放单元得到模块,用于对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;

语音播放单元播放模块,用于若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;所述语音播放条件是获取到针对语音提示信息的触发操作;若所述当前语音播放单元播放失败,获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序,从服务器获取所述目标片段顺序对应的语音片段,作为目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息,所述语音补充提示信息用于提示语音已经进行了更新。

9. 一种语音会话装置,其特征在于,所述装置包括:

会话语音信息得到模块,用于在智能会话机器人进行会话的过程中,获取所述智能会话机器人对应的会话语音信息;语音片段序列得到模块,用于获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段;

语音片段序列发送模块,用于将所述语音片段序列发送给所述智能会话机器人对应的会话终端,以使得所述会话终端对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,所述会话终端按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;所述语音播放条件是获取到针对语音提示信息的触发操作;若所述当前语音播放单元播放失败,则所述会话终端获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序,从服务器获取所述目标片段顺序对应的语音片段,作为目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息,所述语音补充提示信息用于提示语音已经进行了更新。

10. 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求1至7中任一项所述的方法的步骤。

11. 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,所述计算机程序被处理器执行时实现权利要求1至7中任一项所述的方法的步骤。

## 语音会话方法、装置、计算机设备和存储介质

### 技术领域

[0001] 本申请涉及计算机技术领域,特别是涉及一种语音会话方法、装置、计算机设备和存储介质。

### 背景技术

[0002] 随着计算机技术的发展,智能培训机器人也迎来了蓬勃发展,智能培训机器人能够通过语音分析或者自然语言处理等技术,辅助学员通过典型录音观摩或者实战模拟训练等方式学习业务技能,在训练过程中得到实时辅助提醒、质量监督和得分评价,完成对业务流程、应答话术或者沟通技巧等技能的掌握。

[0003] 传统技术中,智能培训机器人在语音会话过程中往往是将机器人回复的语音直接传输到语音终端进行播放。

[0004] 然而,目前的语音会话方法,存在语音会话流畅度低的问题。

### 发明内容

[0005] 基于此,有必要针对上述技术问题,提供一种能够提高语音会话流畅度的语音会话方法、装置、计算机设备和存储介质。

[0006] 一种语音会话方法,所述方法包括:在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列,所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0007] 在其中一个实施例中,所述若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元包括:若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;若所述当前语音播放单元播放失败,则跳过所述当前语音播放单元,继续播放所述当前语音播放单元的下一语音播放单元。

[0008] 在其中一个实施例中,所述方法还包括:若所述当前语音播放单元播放失败,则获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序;向服务器发送语音片段获取请求,所述语音片段获取请求携带所述目标片段顺序以及所述会话语音信息的会话消息标识,以使得所述服务器响应于所述语音片段获取请求,根据所述会话消息标识从存储的所述语音片段序列中提取所述目标片段顺序对应的语音片段,作为目标语音片段;接收所述服务器返回的所述目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息;若接收到针对所述会话语音信息的播放操作,则按照所述语音片段序列中的顺序,播放所述语音片段对应的更新后的语音播放单元。

[0009] 在其中一个实施例中,所述方法还包括:响应于针对所述会话语音信息的播放中断操作,停止播放所述语音播放单元并记录停止位置;获取所述会话语音信息对应的文本信息;在所述会话界面上展示所述文本信息,在所述文本信息中标记所述停止位置所对应的文本。

[0010] 在其中一个实施例中,所述方法还包括:展示培训会话对应的培训会话界面;获取所述培训会话对应的会话标识;基于所述培训会话的会话标识建立所述培训会话界面与长连接网关之间的连接。

[0011] 在其中一个实施例中,所述培训会话中的会话节点包括智能会话机器人;所述在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列包括:在与智能会话机器人进行会话的过程中,通过所述培训会话界面获取所述长连接网关通过实时对话语音通道传输的所述智能会话机器人的语音片段序列。

[0012] 在其中一个实施例中,所述培训会话中的会话节点还包括培训对象,所述方法还包括:通过所述培训会话界面采集所述培训对象的实时语音信息,将所述实时语音信息通过语音传输通道传输至长连接网关;其中,所述长连接网关通过所述会话标识将所述实时语音信息传输至对话控制平台,所述对话控制平台对所述实时语音信息进行语音识别,得到语音文本信息,将所述语音文本信息传输至培训对话服务,所述培训对话服务调用语义分析平台对所述语音文本信息进行会话质检,得到质检结果;通过所述培训会话界面接收所述长连接网关通过语义分析通道传输的所述语音文本信息以及所述质检结果。

[0013] 一种语音会话装置,所述装置包括:语音片段序列获取模块,用于在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列,所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;语音播放单元得到模块,用于对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;语音播放单元播放模块,用于若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0014] 在其中一个实施例中,语音播放单元播放模块,用于若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;若所述当前语音播放单元播放失败,则跳过所述当前语音播放单元,继续播放所述当前语音播放单元的下一语音播放单元。

[0015] 在其中一个实施例中,语音播放单元播放模块,用于若所述当前语音播放单元播放失败,则获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序;向服务器发送语音片段获取请求,所述语音片段获取请求携带所述目标片段顺序以及所述会话语音信息的会话消息标识,以使得所述服务器响应于所述语音片段获取请求,根据所述会话消息标识从存储的所述语音片段序列中提取所述目标片段顺序对应的语音片段,作为目标语音片段;接收所述服务器返回的所述目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息;若接收到针对所述会话语音信息的播放操作,则按照所述语音片段序列中的顺序,播放所述语音片段对应的更新后的语音播放单元。

[0016] 在其中一个实施例中,所述装置包括:播放中断操作响应模块,用于响应于针对所

述会话语音信息的播放中断操作,停止播放所述语音播放单元并记录停止位置;文本信息获取模块,用于获取所述会话语音信息对应的文本信息;文本信息展示模块,用于在所述会话界面上展示所述文本信息,在所述文本信息中标记所述停止位置所对应的文本。

[0017] 在其中一个实施例中,所述装置还包括:培训会话界面展示模块,用于展示培训会话对应的培训会话界面;会话标识获取模块,用于获取所述培训会话对应的会话标识;连接建立模型,用于基于所述培训会话的会话标识建立所述培训会话界面与长连接网关之间的连接。

[0018] 在其中一个实施例中,所述培训会话中的会话节点包括智能会话机器人;所述语音片段序列获取模块,还用于在与智能会话机器人进行会话的过程中,通过所述培训会话界面获取所述长连接网关通过实时对话语音通道传输的所述智能会话机器人的语音片段序列。

[0019] 在其中一个实施例中,所述培训会话中的会话节点还包括培训对象,所述装置还包括:实时语音信息采集模块,用于通过所述培训会话界面采集所述培训对象的实时语音信息,将所述实时语音信息通过语音传输通道传输至长连接网关;其中,所述长连接网关通过所述会话标识将所述实时语音信息传输至对话控制平台,所述对话控制平台对所述实时语音信息进行语音识别,得到语音文本信息,将所述语音文本信息传输至培训对话服务,所述培训对话服务调用语义分析平台对所述语音文本信息进行会话质检,得到质检结果;实时语音信息接收模块,用于通过所述培训会话界面接收所述长连接网关通过语义分析通道传输的所述语音文本信息以及所述质检结果。一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现以下步骤:在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列,所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0020] 在其中一个实施例中,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现以下步骤:若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;若所述当前语音播放单元播放失败,则跳过所述当前语音播放单元,继续播放所述当前语音播放单元的下一语音播放单元。

[0021] 在其中一个实施例中,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现以下步骤:若所述当前语音播放单元播放失败,则获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序;向服务器发送语音片段获取请求,所述语音片段获取请求携带所述目标片段顺序以及所述会话语音信息的会话消息标识,以使得所述服务器响应于所述语音片段获取请求,根据所述会话消息标识从存储的所述语音片段序列中提取所述目标片段顺序对应的语音片段,作为目标语音片段;接收所述服务器返回的所述目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息;若接收到针对所述会话语音信息的播放操作,则按照所述语音片段序列中的顺序,播放所述语音片段对应

的更新后的语音播放单元。

[0022] 在其中一个实施例中,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现以下步骤:响应于针对所述会话语音信息的播放中断操作,停止播放所述语音播放单元并记录停止位置;获取所述会话语音信息对应的文本信息;在所述会话界面上展示所述文本信息,在所述文本信息中标记所述停止位置所对应的文本。

[0023] 一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现以下步骤:在与智能会话机器人进行会话的过程中,获取所述智能会话机器人所对应的语音片段序列,所述语音片段序列是对所述智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。在其中一个实施例中,所述计算机程序被处理器执行时实现以下步骤:

[0024] 在其中一个实施例中,所述计算机程序被处理器执行时实现以下步骤:若满足所述会话语音信息的语音播放条件,则按照所述语音片段序列中的顺序,播放所述语音片段序列中的当前语音播放单元;若所述当前语音播放单元播放失败,则跳过所述当前语音播放单元,继续播放所述当前语音播放单元的下一语音播放单元。

[0025] 在其中一个实施例中,所述计算机程序被处理器执行时实现以下步骤:若所述当前语音播放单元播放失败,则获取所述当前语音播放单元对应的在所述语音片段序列的目标片段顺序;向服务器发送语音片段获取请求,所述语音片段获取请求携带所述目标片段顺序以及所述会话语音信息的会话消息标识,以使得所述服务器响应于所述语音片段获取请求,根据所述会话消息标识从存储的所述语音片段序列中提取所述目标片段顺序对应的语音片段,作为目标语音片段;接收所述服务器返回的所述目标语音片段,基于所述目标语音片段对所述语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于所述会话语音信息展示语音补充提示信息;若接收到针对所述会话语音信息的播放操作,则按照所述语音片段序列中的顺序,播放所述语音片段对应的更新后的语音播放单元。

[0026] 在其中一个实施例中,所述计算机程序被处理器执行时实现以下步骤:响应于针对所述会话语音信息的播放中断操作,停止播放所述语音播放单元并记录停止位置;获取所述会话语音信息对应的文本信息;在所述会话界面上展示所述文本信息,在所述文本信息中标记所述停止位置所对应的文本。

[0027] 上述语音会话方法、装置、计算机设备和存储介质,通过在与智能会话机器人进行会话的过程中,获取智能会话机器人所对应的语音片段序列,该语音片段序列是对智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。通过获取对会话语音信息切分之后的语音片段序列,并且按照所述语音片段序列中的顺序,播放语音片段对应的语音播放单元,能够达到减少语音会话时的卡顿现象,提高语音会话流畅度的目的。

[0028] 一种语音会话方法,所述方法包括:在智能会话机器人进行会话的过程中,获取所述智能会话机器人对应的会话语音信息;对所述会话语音信息进行切分,并按照语音顺序



对切分得到的语音片段进行排列,得到语音片段序列;将所述语音片段序列发送给所述智能会话机器人对应的会话终端,以使得所述会话终端对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,所述会话终端按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0029] 在其中一个实施例中,所述对所述会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列包括:获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段。

[0030] 一种语音会话装置,所述装置包括:会话语音信息得到模块,用于在智能会话机器人进行会话的过程中,获取所述智能会话机器人对应的会话语音信息;语音片段序列得到模块,用于对所述会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列;语音片段序列发送模块,用于将所述语音片段序列发送给所述智能会话机器人对应的会话终端,以使得所述会话终端对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,所述会话终端按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0031] 在其中一个实施例中,语音片段序列得到模块,用于获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段。

[0032] 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现以下步骤:在智能会话机器人进行会话的过程中,获取所述智能会话机器人对应的会话语音信息;对所述会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列;将所述语音片段序列发送给所述智能会话机器人对应的会话终端,以使得所述会话终端对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,所述会话终端按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0033] 在其中一个实施例中,所述处理器执行所述计算机程序时实现以下步骤:获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段。

[0034] 一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现以下步骤:在智能会话机器人进行会话的过程中,获取所述智能会话机器人对应

的会话语音信息;对所述会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列;将所述语音片段序列发送给所述智能会话机器人对应的会话终端,以使得所述会话终端对所述语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个所述语音片段对应的符合播放条件的语音播放单元;若满足所述会话语音信息的语音播放条件,所述会话终端按照所述语音片段序列中的顺序,播放所述语音片段对应的语音播放单元。

[0035] 在其中一个实施例中,所述计算机程序被处理器执行时实现以下步骤:获取所述会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据所述第一播放时长确定所述当前语音片段对应的下一语音片段的第二播放时长,所述第二播放时长所对应的语音传输时长小于所述第一播放时长;基于所述第二播放时长从所述会话语音信息中切分得到所述当前语音片段对应的下一语音片段。

[0036] 上述语音会话方法、装置、计算机设备和存储介质,通过在智能会话机器人进行会话的过程中,获取智能会话机器人对应的会话语音信息;对该会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列;将该语音片段序列发送给智能会话机器人对应的会话终端,以使得会话终端对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;会话终端若满足会话语音信息的语音播放条件,会话终端按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。在发送会话语音信息时,对该会话语音信息进行切分,将切分得到的语音片段序列发送给智能会话机器人对应的会话终端,由于切分后的语音片段序列可以是一段一段的进行传输,因此会话终端可以根据已接收到的语音片段播放语音,减少语音会话时的卡顿现象,提高了语音会话的流畅度。

## 附图说明

- [0037] 图1为一个实施例中语音会话方法的应用环境图;
- [0038] 图2为一个实施例中语音会话方法的流程示意图;
- [0039] 图3为一个实施例中播放语音片段对应的语音播放单元步骤的流程示意图;
- [0040] 图4为另一个实施例中语音会话方法的流程示意图;
- [0041] 图5为另一个实施例中语音会话方法的流程示意图;
- [0042] 图6为一个实施例中语音会话方法的流程示意图;
- [0043] 图7为一个实施例中语音片段序列步骤的流程示意图;
- [0044] 图8为一个实施例中智能机器人的软件架构图;
- [0045] 图9为一个实施例中语音会话装置的结构框图;
- [0046] 图10为一个实施例中语音会话装置的结构框图;
- [0047] 图11为一个实施例中计算机设备的内部结构图。
- [0048] 图12为另一个实施例中计算机设备的内部结构图。

## 具体实施方式

[0049] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本申请,并不

用于限定本申请。

[0050] 本申请提供的语音会话方法,可以应用于如图1所示的应用环境中,具体应用到一种语音会话系统中。该语音会话系统包括会话终端102和服务器104,其中,会话终端102通过网络与服务器104进行通信。会话终端102执行一种语音会话方法,具体的,在与智能会话机器人进行会话的过程中,会话终端102从服务器104获取智能会话机器人所对应的语音片段序列,该语音片段序列是对该智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;对该语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;在会话终端102的会话界面上展示上述会话语音信息对应的语音提示信息,若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。其中,会话终端102可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备,服务器104可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

[0051] 在一个实施例中,如图2所示,提供了一种语音会话方法,以该方法应用于图1中的会话终端为例进行说明,包括以下步骤:

[0052] 步骤S202,在与智能会话机器人进行会话的过程中,获取智能会话机器人所对应的语音片段序列,该语音片段序列是对智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的。

[0053] 其中,智能会话机器人是指能够提供语音会话交互的智能机器人。例如,培训机构使用的智能培训机器人或者银行使用的智能机器人等。语音片段序列是由多个语音片段排列成的序列。语音片段序列中的语音片段可以是通过智能会话机器人的会话语音信息进行切分得到的,切分得到的各个语音片段的大小可以相同也可以不同。智能会话机器人的会话语音信息是指由智能机器人产生的语音信息。语音片段在会话语音信息中的位置越靠前,则语音片段在语音片段序列中的排序越靠前。

[0054] 具体地,在与智能会话机器人进行会话的过程中,语音会话终端通过预设的API (Application Programming Interface) 接口,获取智能会话机器人所对应的语音片段序列,该语音片段序列是对智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的。

[0055] 在一个实施例中,会话终端通过实时对话语音通道获取智能会话机器人所对应的语音片段序列。其中的实时对话语音通道是通过实时对话语音对应的指令类型进行划分的通道。例如,实时对话语音对应的指令类型为用户指令类型,则实时对话语音通道为语音数据传输通道。

[0056] 步骤S204,对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元。

[0057] 其中,语音播放格式信息是指能够使语音片段具有播放功能的信息。例如,在语音片段后添加WAV头部,添加WAV头部后的语音片段,具备播放功能等。携带有WAV的声音文件,为一种标准数字音频文件。语音播放单元是能够独立播放的语音单元。

[0058] 具体地,会话终端获取智能会话机器人所对应的语音片段序列之后,对语音片段序列中的各个语音片段分别添加能够使各个语音片段进行播放的语音播放格式信息,并且通过会话终端上的浏览器上的音频编码函数编码为支持播放的语音播放单元。例如,将语

音片段序列中的各个语音片段添加WAV头部,然后调用浏览器的音频编码接口函数,编码为支持播放的新的音频文件,存放在另一接口函数中。例如,调用浏览器的音频编码接口函数AudioContext,编码为支持播放的新的音频文件,存放在另一接口函数AudioBuffer中。其中AudioContext和AudioBuffer分别为Web(World Wide Web)浏览器中的音频文件对应的应用接口函数。

[0059] 步骤S206,若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。

[0060] 其中,语音播放条件是指播放会话语音信息所需要满足的条件,语音播放条件例如可以是对语音提示信息的触发操作,例如可以对是语音提示信息的点击操作。

[0061] 具体地,会话终端得到各个语音片段对应的符合播放条件的语音播放单元后,当会话终端确定满足会话语音信息的语音播放条件时,则按照语音片段序列中的各个语音片段的顺序,播放语音片段对应的语音播放单元。

[0062] 在一些实施例中,会话终端可以展示会话语音信息对应的语音提示信息,其中,语音提示信息用于对会话语音信息进行提示,语音提示信息可以包括会话语音信息的时长等,语音提示信息例如可以是“6秒”,表示会话语音信息的时长为6秒。

[0063] 上述语音会话方法中,通过在与智能会话机器人进行会话的过程中,获取智能会话机器人所对应的语音片段序列,该语音片段序列是对智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。通过获取对会话语音信息切分之后的语音片段序列,并且按照语音片段序列中的顺序,播放语音片段对应的语音播放单元,能够达到减少语音会话时的卡顿现象,提高语音会话流畅度的目的。

[0064] 在一个实施例中,如图3所示,若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段对应的语音播放单元包括:

[0065] 步骤S302,若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段序列中的当前语音播放单元。

[0066] 其中,当前语音播放单元为当前语音片段对应的语音播放单元,当前语音片段是指语音片段序列中当前时间播放到的语音片段。当前语音播放单元随着播放时间的变化而变化,例如在第1秒到第3秒播放的是语音片段1,第4秒到第6秒播放的是语音片段2,则对于第1秒到第3秒,当前语音播放单元为语音片段1,对于第4秒到第6秒,当前语音播放单元为语音片段2。

[0067] 具体地,会话终端可以将会话语音信息展示在会话界面上,会话终端可以通过会话界面检测语音播放条件,若检测到会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段序列中的当前语音播放单元。其中,会话界面可以是任意的用于会话的界面,例如可以是培训会话对应的培训会话界面,培训会话是指用于进行培训所进行的会话。

[0068] 步骤S304,若当前语音播放单元播放失败,则跳过当前语音播放单元,继续播放当前语音播放单元的下一语音播放单元。

[0069] 其中,当前语音播放单元的下一语音播放单元是指语音片段序列中与当前语音片段相邻并且排列在当前语音片段之后的语音片段所对应的语音播放单元。当前语音片段为当前语音播放单元对应的语音片段。例如语音片段序列为[R1,R2,R3],其中,当前语音片段为R1,则当前语音播放单元的下一语音播放单元为语音片段R2对应的语音播放单元。

[0070] 具体地,语音播放单元播放时,会话终端按照语音片段序列中语音片段的排列顺序对语音片段对应的语音播放单元进行播放,若当前语音播放单元播放完毕,则播放当前语音播放单元的下一语音播放单元,若当前语音播放单元播放失败,则跳过当前语音播放单元,继续播放当前语音播放单元的下一语音播放单元。例如,语音片段序列[R1,R2,R3]中,R1对应的语音播放单元为V1,R2对应的语音播放单元为V2,R3对应的语音播放单元为V3,若当前语音播放单元为V1,若V1播放完毕,则播放当前语音播放单元的下一语音播放单元V2,若V1播放失败,则跳过当前语音播放单元为V1,继续播放当前语音播放单元的下一语音播放单元V2。

[0071] 在一些实施例中,会话终端可以利用会话终端中设置的播放设备进行播放,在播放语音播放单元的过程中,播放设备可以在播放完当前语音播放单元时,再获取当前语音播放单元的下一语音播放单元进行播放,或者播放设备可以在播放完当前语音播放单元之前,获取当前语音播放单元的下一语音播放单元,或者播放可以在播放完当前语音播放单元之前,获取当前语音播放单元之后的多个语音播放单元,多个是指至少两个。当然,在播放之前,播放设备可以将语音片段序列中各个语音片段分别对应的语音播放单元缓存至播放设备中。本实施例中,通过若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段序列中的当前语音播放单元,若当前语音播放单元播放失败,则跳过当前语音播放单元,继续播放当前语音播放单元的下一语音播放单元,能够达到准确对语音播放单元进行播放,提高播放语音播放单元流畅度的目的。

[0072] 在一个实施例中,如图4所示,语音会话方法还包括:

[0073] 步骤S402,若当前语音播放单元播放失败,则获取当前语音播放单元对应的在语音片段序列的目标片段顺序。

[0074] 其中,片段顺序是指语音片段在语音片段序列中的排列顺序。例如,语音片段序列为{R1,R2,R3,R4,R5},语音片段R4的片段顺序为4。目标片段顺序是指当前语音播放单元对应的语音片段在语音片段序列中的排序。

[0075] 步骤S404,向服务器发送语音片段获取请求,该语音片段获取请求携带目标片段顺序以及会话语音信息的会话消息标识,以使得服务器响应于语音片段获取请求,根据会话消息标识从存储的语音片段序列中提取目标片段顺序对应的语音片段,作为目标语音片段。

[0076] 其中,语音片段获取请求用于请求获取语音片段,会话消息标识用于唯一标识会话消息,会话消息标识可以根据需要设置,例如可以根据会话语音信息出现的先后顺序进行设置,当然也可以采用其他方式设置,这里不做限定,会话消息标识例如可以为001。语音片段获取请求中可以包括会话语音信息的会话消息标识以及播放失败的语音片段在语音片段序列中的排序,例如携带目标片段顺序。例如,会话消息为开始某项培训时,对应的会话标识为该项培训对应的识别标志,例如该项培训对应的识别标志为数字001,则001即是会话消息标识。目标语音片段为语音片段序列中排序为目标片段顺序的语音片段。

[0077] 具体地,会话消息标识与语音片段序列之间存在一一对应关系,会话终端在确定下一语音播放单元对应的在语音片段序列的片段顺序后,向服务器发送语音片段获取请求,服务器可以响应于语音片段获取请求,根据会话消息标识从存储的语音片段序列中提取目标片段顺序对应的语音片段,作为目标语音片段。

[0078] 步骤S406,接收服务器返回的目标语音片段,基于目标语音片段对语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于会话语音信息展示语音补充提示信息。

[0079] 其中,语音补充提示信息用于提示语音已经进行了更新语音补充提示信息例如可以为“语音已更新,请查看更新语音”。

[0080] 具体地,服务器响应于语音片段获取请求,根据会话消息标识从存储的语音片段序列中提取目标片段顺序对应的语音片段,作为目标语音片段后,将目标语音片段返回至会话终端,会话终端可以基于目标语音片段对语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于会话语音信息展示语音补充提示信息。例如,语音补充提示信息为“语音已更新,请查看更新语音”等,并将该语音补充提示信息展示在会话界面上。其中,会话终端可以为目标语音片段生成语音播放单元,作为目标语音播放单元,利用目标语音播放单元替换掉语音片段序列中排序为目标片段顺序的语音片段对应的语音播放单元,得到更新后的语音播放单元。

[0081] 步骤S408,若接收到针对会话语音信息的播放操作,则按照语音片段序列中的顺序,播放语音片段对应的更新后的语音播放单元。

[0082] 具体地,会话终端接收到语音补充提示信息后,若接收到针对会话语音信息的播放操作,例如,接收到对会话界面上的语音提示信息的点击操作,则会话终端可以按照语音片段序列中的顺序,播放语音片段对应的更新后的语音播放单元。

[0083] 本实施例中,通过若当前语音播放单元播放失败,则获取当前语音播放单元对应的在语音片段序列的目标片段顺序,向服务器发送语音片段获取请求,该语音片段获取请求携带目标片段顺序以及会话语音信息的会话消息标识,以使得服务器响应于语音片段获取请求,根据会话消息标识从存储的语音片段序列中提取目标片段顺序对应的语音片段,作为目标语音片段,接收服务器返回的目标语音片段,基于目标语音片段对语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于会话语音信息展示语音补充提示信息,若接收到针对会话语音信息的播放操作,则按照语音片段序列中的顺序,播放语音片段对应的更新后的语音播放单元,能够达到及时获取缺失的语音片段,并对该语音片段对应的语音播放单元进行更新,对更新后的语音播放单元进行播放的目的。

[0084] 在一个实施例中,如图5所示,语音会话方法还包括:

[0085] 步骤S502,响应于针对会话语音信息的播放中断操作,停止播放语音播放单元并记录停止位置。

[0086] 其中,播放中断操作是指触发会话语音信息停止播放的操作。例如,在会话语音信息的播放过程中,若会话终端获取到对正在播放的语音的点击操作,则触发停止会话语音信息的播放。播放中断操作还可以包括语音信息输入操作,当播放会话语音信息的同时,若会话终端接收到语音信息输入操作,则确定获取到会话语音信息对应的播放中断操作,停止播放会话语音信息。语音信息输入操作用于触发向会话中输入语音信息。

[0087] 具体地,会话终端在接收到播放中断操作时,响应于针对会话语音信息的播放中断操作,停止播放正在播放的语音播放单元,并记录停止位置,例如,响应于针对会话语音信息的播放中断操作,假设会话语音信息总时长为30秒,在播放到会话语音信息的第15秒时获取到播放中断操作,则停止位置为15秒。

[0088] 步骤S504,获取会话语音信息对应的文本信息。

[0089] 其中,会话语音信息对应的文本信息可以是对会话语音信息进行语音识别所识别出的文本信息。会话语音信息对应的文本信息可以是会话终端从服务器获取的,也可以是会话终端对会话语音信息进行语音识别所得到的。

[0090] 具体地,当会话终端停止播放会话语音信息并记录停止位置时,会话终端可以获取该会话语音信息对应的文本信息,例如可以从服务器获取会话语音信息对应的文本信息,或者对会话语音信息进行语音识别得到文本信息。

[0091] 在一个实施例中,会话终端利用语音识别模块,对会话语音信息进行语音识别,该识别出的文本信息作为会话语音信息对应的文本信息。例如,在会话终端上的培训考试界面上,获取机器人读取的题目对应的会话语音信息,在答题用户无需机器人读取完毕已经得知答案的情况下,会话终端获取到对会话语音信息的点击操作,暂停播放会话语音信息,获取会话语音信息的文本信息,展示文本信息,在文本信息中标记停止位置对应的文本。步骤S506,在会话界面上展示文本信息,在文本信息中标记停止位置所对应的文本。

[0092] 具体地,会话终端获取会话语音信息对应的文本信息,在会话界面上展示文本信息,在文本信息中标记停止位置所对应的文本。例如,停止位置为会话语音信息播放到15秒处的位置,在会话语音信息对应的文本信息中,标记会话语音信息播放到15秒处的文本信息。

[0093] 在一些实施例中,会话终端在播放智能会话机器人的会话语音信息的同时,当接收到语音信息输入操作时,确定获取到会话语音信息的播放中断操作,响应于针对会话语音信息的播放中断操作,停止播放语音播放单元并记录停止位置,获取会话语音信息对应的文本信息,在会话界面上展示文本信息,在文本信息中标记停止位置所对应的文本,会话终端还可以获取语音信息输入操作所输入的用户语音信息,获取对用户语音信息进行识别所得到用户文本信息,在展示会话语音信息对应的文本信息的同时,展示用户文本信息。其中,用户文本信息可以是会话终端对用户语音信息进行识别得到的,也可以是服务器对用户语音信息进行识别得到的。

[0094] 在一些实施例中,会话终端响应于针对会话语音信息的播放中断操作,停止播放语音播放单元并记录停止位置,会话终端还可以向服务器发送语音片段传输停止请求,语音片段传输停止请求中可以携带被停止播放的会话语音信息的标识,语音片段传输停止请求用于请求服务器停止向会话终端发送被停止播放的会话语音信息对应的语音片段。

[0095] 本实施例中,通过响应于针对会话语音信息的播放中断操作,停止播放语音播放单元并记录停止位置,获取会话语音信息对应的文本信息,在会话界面上展示文本信息,在文本信息中标记停止位置所对应的文本,能够达到准确计算会话语音中每一次会话语音信息执行后的时间的目的,缩短会话时间,提高语音会话效率。

[0096] 在一个实施例中,该方法还包括:展示培训会话对应的培训会话界面,获取培训会话对应的会话标识,基于培训会话的会话标识建立培训会话界面与长连接网关之间的连

接。

[0097] 其中,培训会话是用于培训的会话,培训会话界面为培训会话对应的界面,用户展示培训会话中的会话信息,培训会话中的会话节点可以包括虚拟用户,也可以包括真实的用户,虚拟用户例如可以是智能会话机器人,真实的用户例如可以是需要培训的人员。培训会话界面也可以称为培训对话界面。会话标识用于唯一识别一个会话。会话节点是指参与会话的成员。会话也可以称为对话。

[0098] 具体地,终端可以基于培训会话的会话标识,在培训会话界面与长连接网络之间建立连接,从而培训会话界面可以从长连接网关获取数据。培训会话界面与长连接网关之间建立的连接(传输通道)可以包括实时对话语音通道或信令语义分析通道。实时对话语音通道用于传输语音信息,信令语义分析通道用于传输控制命令、质检结果等信息。相比于传统的技术中只有一条数据传输通道,能够让语音和文本数据传输分离,互不影响。

[0099] 在一个实施例中,培训会话中的会话节点包括智能会话机器人,在与智能会话机器人进行会话的过程中,获取智能会话机器人所对应的语音片段序列包括:在与智能会话机器人进行会话的过程中,通过培训会话界面获取长连接网关通过实时对话语音通道传输的智能会话机器人的语音片段序列。

[0100] 具体地,智能会话机器人的语音片段序列可以是培训会话界面从长连接网关得到的。长连接网关可以通过与培训会话界面之间建立的实时对话语音通道,将智能会话机器人的语音片段序列传输至培训会话界面。

[0101] 在一个实施例中,培训会话中的会话节点还包括培训对象,该方法还包括:通过培训会话界面采集培训对象的实时语音信息,将实时语音信息通过语音传输通道传输至长连接网关;其中,长连接网关通过会话标识将实时语音信息传输至对话控制平台,对话控制平台对实时语音信息进行语音识别,得到语音文本信息,将语音文本信息传输至培训对话服务,培训对话服务调用语义分析平台对语音文本信息进行会话质检,得到质检结果;通过培训会话界面接收长连接网关通过语义分析通道传输的语音文本信息以及质检结果。

[0102] 其中,培训对象是指需要被培训的对象,例如可以是培训学员,培训对象为真实的用户。语义分析通道是指培训会话界面与长连接网关之间建立的信令语义分析通道。

[0103] 具体地,当终端确定培训会话中包括培训对象时,可以响应于培训对象在培训会话界面中触发的语音输入操作,对培训对象进行语音采集,得到培训对象对应的实时语音信息,将实时语音信息通过语音传输通道传输至长连接网关。

[0104] 在一些实施例中,长连接网关可以通过会话标识,在对话控制平台与培训会话界面之间建立长连接。对话控制平台可以根据会话标识从培训对话服务获取配置信息,并获取培训会话中的培训节点的信息。培训对话服务可以获取培训会话的缓存的培训节点的信息,将缓存的培训节点的信息发送至对话控制平台。培训节点的信息包括节点类型,节点类型可以包括机器人或真实用户中的至少一个。对话控制平台可以根据节点类型确定会话节点为机器人还是培训对象。

[0105] 在一个实施例中,如图6所示,提供了一种语音会话方法,以该方法应用于图1中的服务器为例进行说明,包括以下步骤:

[0106] 步骤S602,在智能会话机器人进行会话的过程中,获取智能会话机器人对应的会话语音信息。



[0107] 具体地,在智能会话机器人进行会话的过程中,当服务器接收到会话请求时,通过会话请求中的会话标识,获取存储的智能会话机器人对应的会话语音信息,例如服务器中可以存储有各个会话标识分别对应的智能会话机器人,服务器可以根据会话标识确定智能会话机器人,从而获取该智能会话机器人的会话语音信息。会话请求中可以包括会话标识。会话请求可以是会话终端发送至服务器的。

[0108] 步骤S604,对会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列。

[0109] 其中,语音顺序是指会话语音信息中的语音信息在会话语音信息中的位置。语音片段在会话语音信息中的位置越靠前,则语音片段在语音片段序列中的位置越靠前。

[0110] 具体地,服务器在获取到智能会话机器人对应的会话语音信息后,可以对会话语音信息进行切分,并且根据语音的时序顺序对切分得到的语音片段进行排列,得到语音片段序列,各个语音片段的大小可以相同也可以不同,例如可以对按照预测语音信息大小对会话语音信息进行切分,得到大小相同的各个语音片段,排列成语音片段序列。

[0111] 在一个实施例中,服务器可以根据预设语音信息大小对会话语音信息进行切分,其中,预设语音信息大小可以小于最大传输单元MTU(Maximum Transfer Unit)。假设,若服务器获取到的会话语音信息的大小小于最大传输单元MTU,则可以将会话语音信息直接发送给会话终端,不进行切分。

[0112] 步骤S606,将语音片段序列发送给智能会话机器人对应的会话终端,以使得会话终端对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;若满足会话语音信息的语音播放条件,会话终端按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。

[0113] 具体地,服务器获取智能会话机器人对应的会话语音信息,对会话语音信息进行切分得到语音片段序列,通过长连接网关将语音片段序列中的语音片段按照先后顺序发送给智能会话机器人对应的会话终端,以使得会话终端对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元,若满足会话语音信息的语音播放条件,会话终端按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。

[0114] 上述语音会话方法中,通过在智能会话机器人进行会话的过程中,获取智能会话机器人对应的会话语音信息;对该会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列;将该语音片段序列发送给智能会话机器人对应的会话终端,以使得会话终端对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;若满足会话语音信息的语音播放条件,则会话终端按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。在发送会话语音信息时,对该会话语音信息进行切分,将切分得到的语音片段序列发送给智能会话机器人对应的会话终端,由于切分后的语音片段序列可以是一段一段的进行传输,因此会话终端可以根据已接收到的语音片段播放语音,减少语音会话时的卡顿现象,提高了语音会话的流畅度。

[0115] 在一个实施例中,如图7所示,对会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列包括:

[0116] 步骤S702,获取会话语音信息对应的切分得到的当前语音片段的第一播放时长。

[0117] 其中,当前语音片段可以为会话语音序列中的任意一个语音片段。播放时长是指语音片段对应的时长。第一播放时长是指当前语音片段对应的播放时长,例如为3秒。

[0118] 具体地,服务器可以获取语音片段序列中各个语音片段分别对应的播放时长,例如,服务器可以通过音频监听插件获取到各个语音片段的播放时长,服务器可以存储各个语音片段分别对应的播放时长。

[0119] 步骤S704,根据第一播放时长确定当前语音片段对应的下一语音片段的第二播放时长,第二播放时长所对应的语音传输时长小于第一播放时长。

[0120] 其中,语音传输时长是指语音在传输线路中传输的时间长度。例如,语音片段从服务器传输到会话终端所使用的时间为30毫秒,则该语音片段的传输时长为30毫秒。第二播放时长是中当前语音片段的下一语音片段的播放时长。第二播放时长所对应的语音传输时长小于第一播放时长,例如第二播放时长所对应的语音传输时长为30毫秒,第一播放时长为3秒。

[0121] 具体地,各个语音片段在通信链路中传输,都存在有语音传输时长,服务器可以获取当前语音片段的第一播放时长,通过当前语音片段的第一播放时长确定下一语音片段的第二播放时长,例如服务器根据第一播放时长确定当前语音片段的下一语音片段的语音传输时长,基于语音传输时长确定第二播放时长。其中,第二播放时长所对应的语音传输时长小于第一播放时长。

[0122] 步骤S706,基于第二播放时长从会话语音信息中切分得到当前语音片段对应的下一语音片段。

[0123] 具体地,服务器得到下一语音片段的播放时长(即第二播放时长),利用第二播放时长从会话语音信息中切分得到当前语音片段对应的下一语音片段,该下一语音片段的播放时长为第二播放时长。

[0124] 在一个实施例中,服务器可以根据语音片段时长以及对会话语音信息的采样率、采样位深和声道数等得到语音片段的大小,例如,语音片段的大小表示为A,片段时长表示为T,采样率表示为V,采样位深表示为S,声道数表示为B,则语音片段的大小A表示为公式:

[0125]  $A=T*V*S*B$ 。

[0126] 本实施例中,通过获取会话语音信息对应的切分得到的当前语音片段的第一播放时长,根据第一播放时长确定当前语音片段对应的下一语音片段的第二播放时长,基于第二播放时长从会话语音信息中切分得到当前语音片段对应的下一语音片段,能够达到准确得到当前语音片段对应的下一语音片段的目的,能够使得在当前语音片段播放完毕之前,下一语音片段已传输完成,从而使语音会话更加顺畅,提高了语音播放的流畅性。

[0127] 在一个实施例中,以智能培训机器人为例进行阐述。具体地,在对话节点为机器人节点时,语音平台获取机器人要回复的语音信息,语音平台将回复的语音信息按照预定大小切成多个语音分片,根据配置文件设置的大小决定语音切分大小;如果合成语音信息上游模块发来的语音信息大小小于配置文件设置的语音分片的大小,则直接发送语音信息到客户端,对语音信息不做切分。当前语音分片在传输时间上需小于上一片语音分片的实际播放时间,确保客户端能提前缓存语音分片;语音平台将多个分片按照一定的次序依此传输给统一长连接网关,再由统一长连接网关通过语音传输通道直接传输给客户端,即前端;

前端接收到语音分片后对分片进行处理,生成语音分片的自增长索引标记,由前端的字典映射储存自增长索引,保证了播放时的顺序正确;前端接收到的语音分片后先在该语音分片上添加WAV头部,然后通过客户端的AudioContext函数编码为支持播放的AudioBuffer函数,并对上述语音分片进行噪声处理,储存在字典对应自增长索引的位置,以及,接收对该语音分片的播放触发操作,当上述语音分片播放结束后,继续查询字典映射中是否存在可播放的语音分片,直到字典映射清空为止;在播放语音分片过程中,若接收到用户录入的语音,则语音平台生成用户语音对应的文本信息,通过统一长连接网关发送给前端,以使前端中断传输语音分片,同时停止播放正在播放的语音分片,并清空字典映射,同时展示用户语音对应的文本信息。清空字典可以保证在有用户录入语音时,不再播放机器人的语音。本实施例能够解决智能培训机器人在播放机器人语音信息过程存在播放卡顿的问题。可以理解的,本实施例中的客户端可以是Chrome (Google Chrome) 浏览器或者Safari (Apple Safari) 浏览器等主流客户端。

[0128] 在一个实施例中,调用AudioContext可以通过接收智能培训机器人上的会话开始操作,响应于该会话开始操作,调用AudioContext的resume函数,将接收到的语音分片信息,转换为能够支持播放的AudioBuffer,兼容多种浏览器的可播放效果。

[0129] 在一个实施例中,当前端接收到语音分片,对该语音分片进行添加WAV头部失败或者调用浏览器AudioContext失败或者其他能够导致该语音分片转换为能够满足播放条件播放音频失败时,前端删除该分片以及该分片在字典中对应的存储位置。

[0130] 在一个实施例中,上述智能机器人的软件架构,如图8所示,包括交互展示层、业务逻辑层和技术平台层。具体地,交互展示层用于界面展示与用户交互逻辑处理,该界面可以为PC(Personal Computer)网页、手机APP(Application)或者平板电脑APP等。业务逻辑层处理复杂的业务逻辑,保障系统的可用性以及扩展性等,技术平台层提供更底层的语音,语义等技术支撑。交互展示层包括培训对话界面和培训管理界面。培训管理界面用于管理员配置产品并进行发布,例如,配置试题或者配置规则等。培训对话界面包括SDK(Software Development Kit)和适配不同终端的UI(User Interface)界面。SDK包含有展示层各种处理数据的通用逻辑。业务逻辑层包括长连接网关、培训对话服务、培训管理后台等。其中,培训管理后台用于配置培训对话流程的各个节点、题目、任务等。配置的数据存储在数据库中。长连接网关用于维持与培训对话用户和服务器之间长链接,统一接收客服人员培训时录入的语音信息,并将客服人员录入的语音信息传输给技术平台层的语言平台进行识别、合成、编解码等,并可以接收语音平台发送的机器人说的语音信息。长连接网关根据会话ID(Identity document)记录与培训对话界面和对话控制平台的连接信息,通过心跳机制检查连接的有效性,以保证会话过程中的连接不易断掉。长连接网关与培训对话界面的SDK之间建立两条数据传输通道,即实时对话语音通道和信令语义分析通道。实时对话语音通道用于传输语音信息,信令语义分析通道用于传输控制命令或者质检结果等信息。相比于传统的技术中只有一条数据传输通道,能够让语音和文本数据传输分离,互不影响。培训对话服务用来接收考试请求,通过雪花算法产生会话ID,获取配置数据,如配置的试题及对话节点等信息以及控制对话流程。其中的语义分析平台用来对文本信息进行质检。

[0131] 在一个实施例中,以培训学员考试为例,描述上述智能机器人的软件架构中各个模块之间的交互过程:客服人员利用平板电脑进行考试,登录账号,进入培训试题考试界

面,培训对话界面获取考试界面上的考试开始操作,培训对话界面根据考试按钮的开始操作,向培训对话服务发起考试请求;培训对话服务根据发起考试请求的时间戳以及服务器的IP(Internet Protocol)地址,通过雪花算法会生成一个全局的会话ID,并缓存本次会话数据,将会话ID返回给培训对话界面;培训对话界面根据会话ID与长连接网关建立两条WS(Websocket)连接,长连接网关根据会话ID与对话控制平台建立连接,连接建立后,长连接网关通过会话ID的唯一性维持培训对话界面与对话控制平台之间的长连接;当对话控制平台与长连接网关建立连接成功之后,对话控制平台根据会话ID向培训对话服务获取配置信息,并获取培训对话中的节点信息。培训对话服务根据获取到的会话ID获取缓存的节点信息并返回给对话控制平台。当对话控制平台获取到节点信息之后,如果是机器人节点,则将配置的机器人文本信息转义成语音信息推送给长连接网关,长连接网关通过实时对话语音通道传给培训对话界面,培训对话界面播放机器人语音。若对话控制平台获取到是培训学员节点,则对话控制平台通过发送开始录音指令给长连接网关,长连接网关通过信令通道通知培训对话界面进行录音。培训对话界面采集用户的实时语音信息,并将实时语音信息通过语音传输通道传输给长连接网关;长连接网关通过会话ID将客服人员录入的语音信息传输对话控制平台;对话控制平台对用户录入的语音信息进行语音识别转换为文本信息,然后将文本信息传输给培训对话服务。培训对话服务获取到对话控制平台传过来的信息时,如果是机器人节点,则直接缓存信息;如果是培训学员节点,则调用语义分析平台对学员文本信息进行会话质检,得到质检结果,质检结果一般是正确或者错误,通过质检结果对本次节点信息进行节点分数的计算,以及判断是否下一次节点需要进行跳转等逻辑操作。并将文本信息和质检结果一起传输给统一网关;会话质检是可配置的,可以不需要;长连接网关会将文本信息和质检结果一起通过语义分析通道传输给培训对话界面。应该理解的是,虽然图2-7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,这些步骤可以以其它的顺序执行。而且,图2-7中的至少一部分步骤可以包括多个步骤或者多个阶段,这些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,这些步骤或者阶段的执行顺序也不必然是依次进行,而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

[0132] 在一个实施例中,如图9所示,提供了一种语音会话装置900,语音片段序列获取模块902,用于在与智能会话机器人进行会话的过程中,获取智能会话机器人所对应的语音片段序列,语音片段序列是对智能会话机器人的会话语音信息进行切分并按照语音顺序排列得到的;语音播放单元得到模块904,用于对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;语音播放单元播放模块906,用于若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。

[0133] 在一个实施例中,语音播放单元播放模块904,用于若满足会话语音信息的语音播放条件,则按照语音片段序列中的顺序,播放语音片段序列中的当前语音播放单元;若当前语音播放单元播放失败,则跳过当前语音播放单元,继续播放当前语音播放单元的下一语音播放单元。

[0134] 在一个实施例中,语音播放单元播放模块904,用于若当前语音播放单元播放失

败,则获取当前语音播放单元对应的在语音片段序列的目标片段顺序;向服务器发送语音片段获取请求,语音片段获取请求携带目标片段顺序以及会话语音信息的会话消息标识,以使得服务器响应于语音片段获取请求,根据会话消息标识从存储的语音片段序列中提取目标片段顺序对应的语音片段,作为目标语音片段;接收服务器返回的目标语音片段,基于目标语音片段对语音片段序列对应的语音播放单元进行更新,并在会话界面上,对应于会话语音信息展示语音补充提示信息;若接收到针对会话语音信息的播放操作,则按照语音片段序列中的顺序,播放语音片段对应的更新后的语音播放单元。

[0135] 在一个实施例中,装置包括:播放中断操作响应模块,用于响应于针对会话语音信息的播放中断操作,停止播放语音播放单元并记录停止位置;文本信息获取模块,用于获取会话语音信息对应的文本信息;文本信息展示模块,用于在会话界面上展示文本信息,在文本信息中标记停止位置所对应的文本。

[0136] 在一个实施例中,装置还包括:培训会话界面展示模块,用于展示培训会话对应的培训会话界面;会话标识获取模块,用于获取培训会话对应的会话标识;连接建立模型,用于基于培训会话的会话标识建立培训会话界面与长连接网关之间的连接。

[0137] 在一个实施例中,培训会话中的会话节点包括智能会话机器人;语音片段序列获取模块,还用于在与智能会话机器人进行会话的过程中,通过培训会话界面获取长连接网关通过实时对话语音通道传输的智能会话机器人的语音片段序列。

[0138] 在一个实施例中,培训会话中的会话节点还包括培训对象,装置还包括:实时语音信息采集模块,用于通过培训会话界面采集培训对象的实时语音信息,将实时语音信息通过语音传输通道传输至长连接网关;其中,长连接网关通过会话标识将实时语音信息传输至对话控制平台,对话控制平台对实时语音信息进行语音识别,得到语音文本信息,将语音文本信息传输至培训对话服务,培训对话服务调用语义分析平台对语音文本信息进行会话质检,得到质检结果;实时语音信息接收模块,用于通过培训会话界面接收长连接网关通过语义分析通道传输的语音文本信息以及质检结果。

[0139] 在一个实施例中,如图10所示,提供了一种语音会话装置1000,包括:会话语音信息得到模块1002、语音片段序列得到模块1004和语音片段序列发送模块1006,其中:会话语音信息得到模块1002,用于在智能会话机器人进行会话的过程中,获取智能会话机器人对应的会话语音信息;语音片段序列得到模块1004,用于对会话语音信息进行切分,并按照语音顺序对切分得到的语音片段进行排列,得到语音片段序列;语音片段序列发送模块1006,用于将语音片段序列发送给智能会话机器人对应的会话终端,以使得会话终端对语音片段序列中的各个语音片段分别添加语音播放格式信息,得到各个语音片段对应的符合播放条件的语音播放单元;若满足会话语音信息的语音播放条件,则会话终端按照语音片段序列中的顺序,播放语音片段对应的语音播放单元。

[0140] 在一个实施例中,语音片段序列得到模块1004,用于获取会话语音信息对应的切分得到的当前语音片段的第一播放时长;根据第一播放时长确定当前语音片段对应的下一语音片段的第二播放时长,第二播放时长所对应的语音传输时长小于第一播放时长;基于第二播放时长从会话语音信息中切分得到当前语音片段对应的下一语音片段。

[0141] 关于语音会话装置的具体限定可以参见上文中对于语音会话方法的限定,在此不再赘述。上述语音会话装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上

述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中,也可以以软件形式存储于计算机设备中的存储器中,以便于处理器调用执行以上各个模块对应的操作。

[0142] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是服务器,其内部结构图可以如图11所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。其中,该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的数据库用于存储语音会话数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种语音会话方法。

[0143] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是终端,其内部结构图可以如图12所示。该计算机设备包括通过系统总线连接的处理器、存储器、通信接口、显示屏和输入装置。其中,该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统和计算机程序。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的通信接口用于与外部的终端进行有线或无线方式的通信,无线方式可通过WIFI、运营商网络、NFC(近场通信)或其他技术实现。该计算机程序被处理器执行时以实现一种语音会话方法。该计算机设备的显示屏可以是液晶显示屏或者电子墨水显示屏,该计算机设备的输入装置可以是显示屏上覆盖的触摸层,也可以是计算机设备外壳上设置的按键、轨迹球或触控板,还可以是外接的键盘、触控板或鼠标等。

[0144] 本领域技术人员可以理解,图11或者图12中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0145] 在一个实施例中,还提供了一种计算机设备,包括存储器和处理器,存储器中存储有计算机程序,该处理器执行计算机程序时实现上述各方法实施例中的步骤。

[0146] 在一个实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现上述各方法实施例中的步骤。

[0147] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和易失性存储器中的至少一种。非易失性存储器可包括只读存储器(Read-Only Memory,ROM)、磁带、软盘、闪存或光存储器等。易失性存储器可包括随机存取存储器(Random Access Memory,RAM)或外部高速缓冲存储器。作为说明而非局限,RAM可以是多种形式,比如静态随机存取存储器(Static Random Access Memory,SRAM)或动态随机存取存储器(Dynamic Random Access Memory,DRAM)等。

[0148] 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0149] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但不能因此而理解为对发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

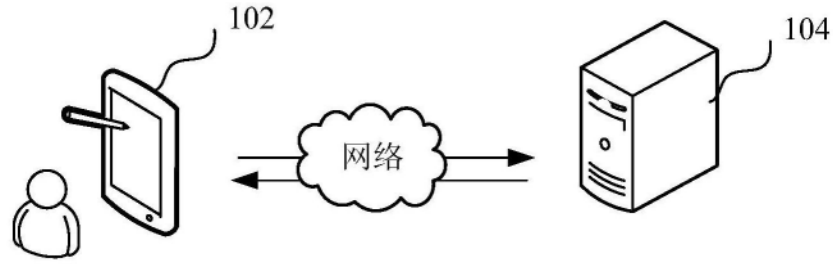


图1

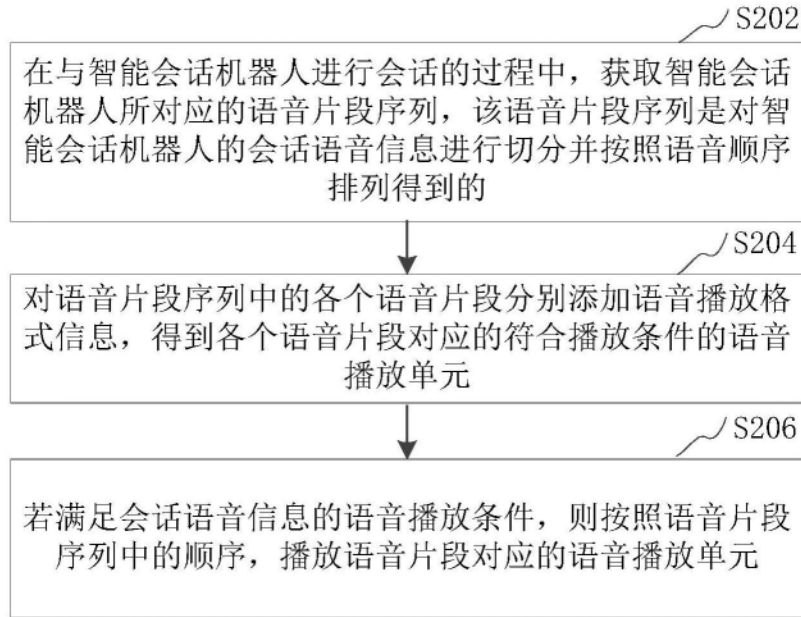


图2

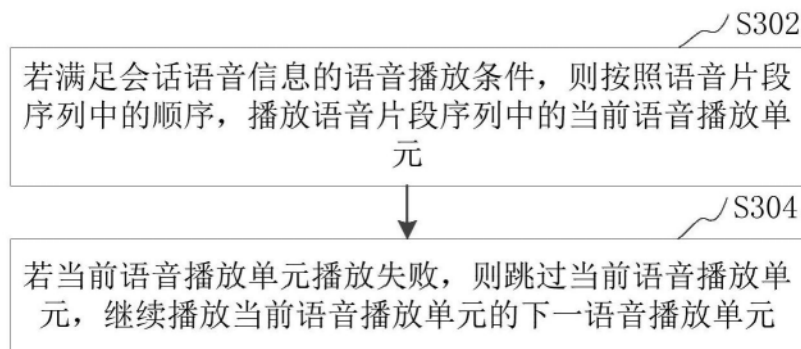


图3



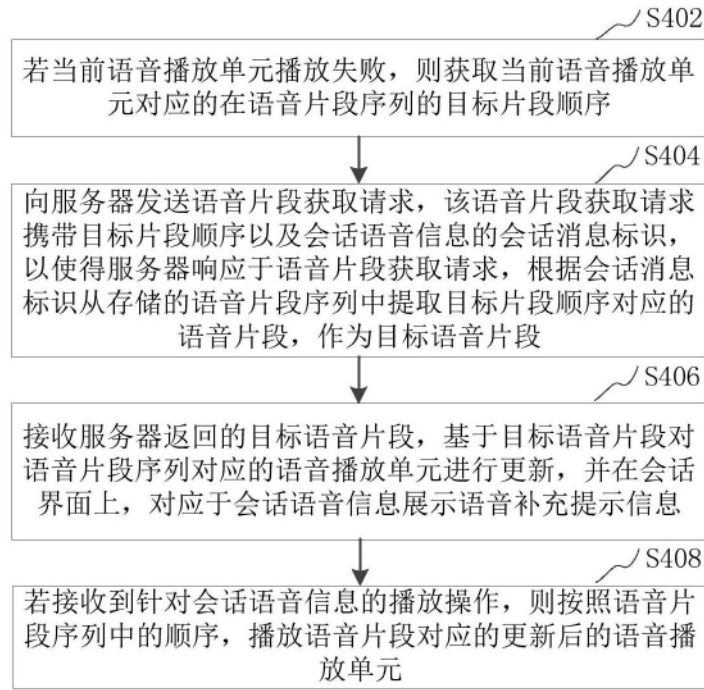


图4

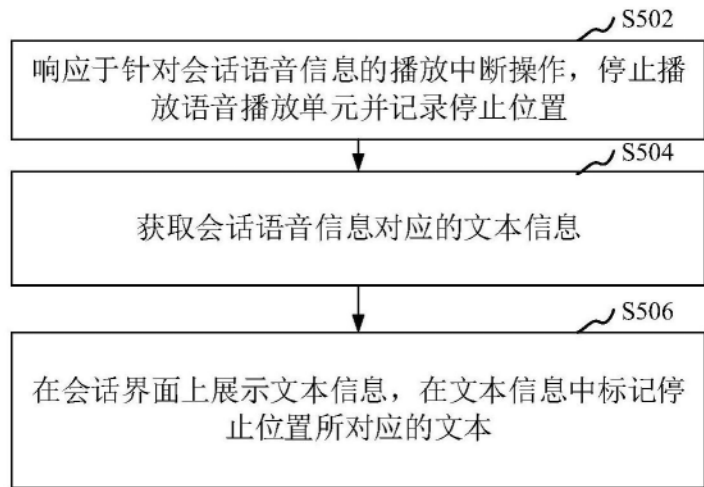


图5

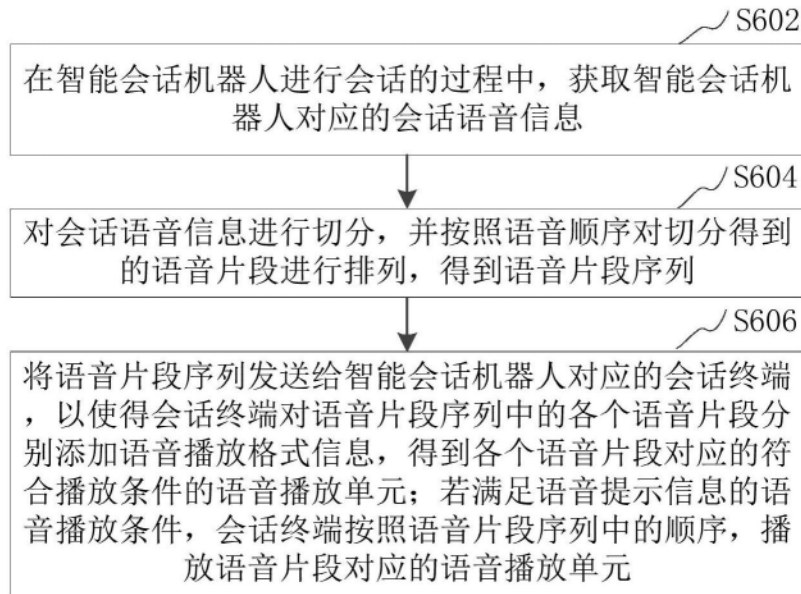


图6

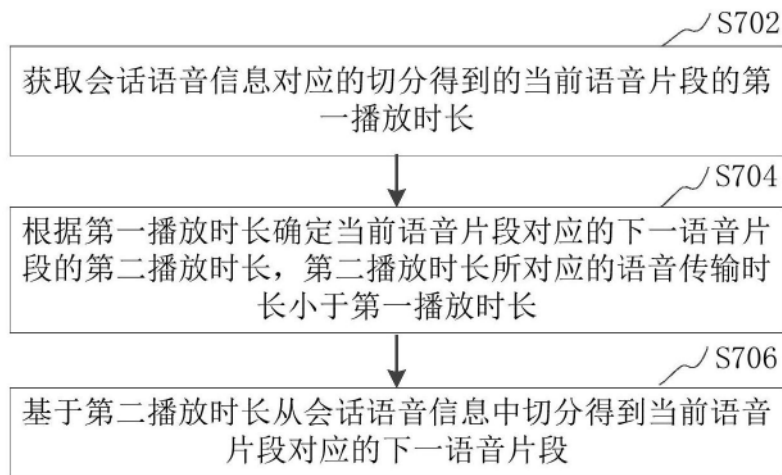


图7

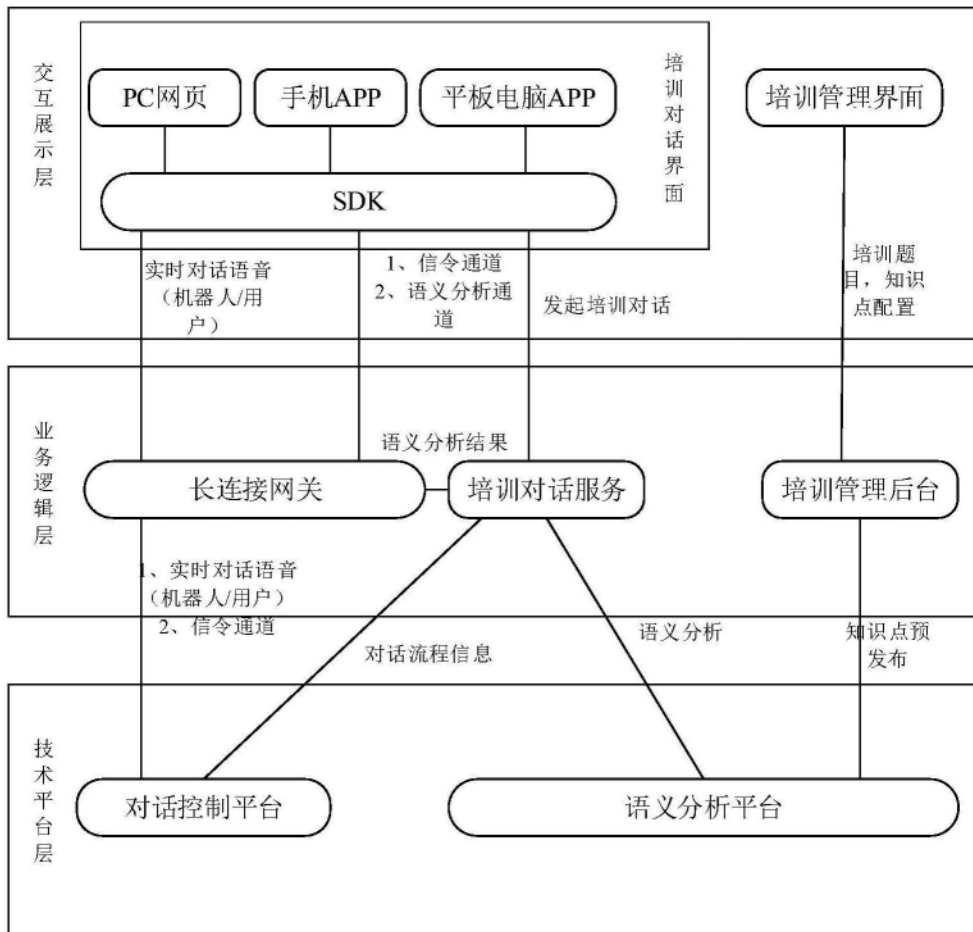


图8

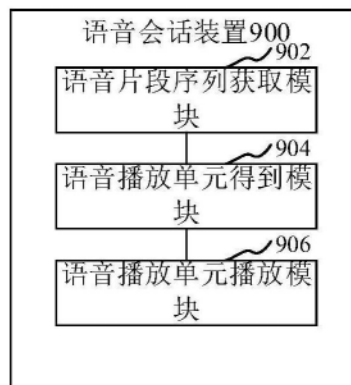


图9

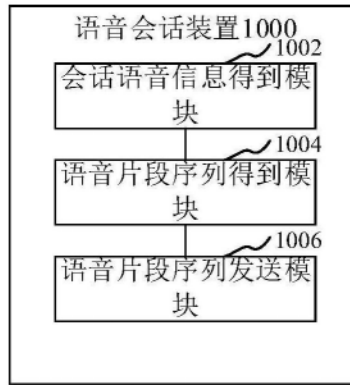


图10

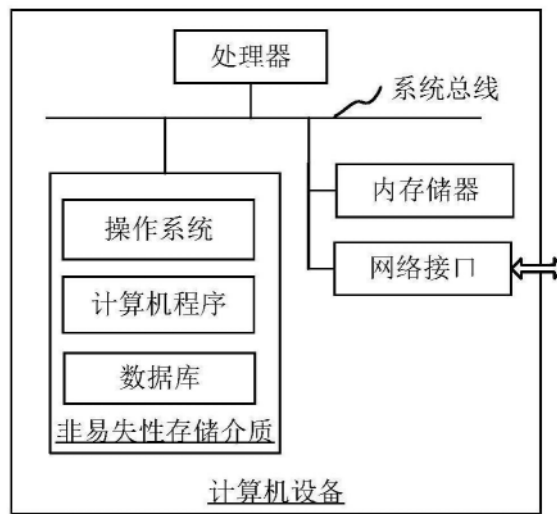


图11

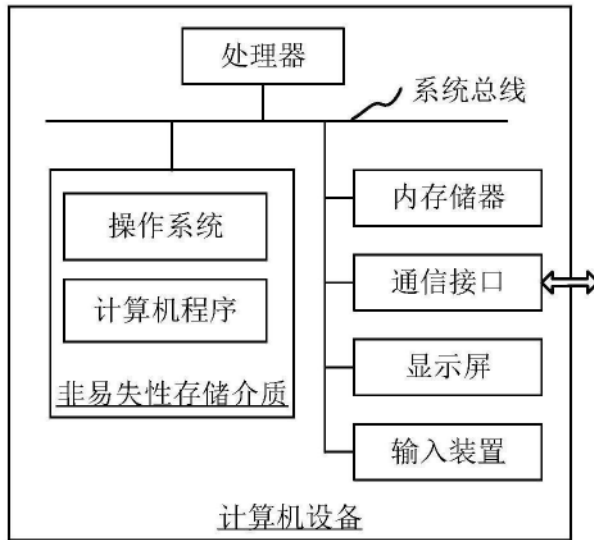


图12