



(12)发明专利申请

(10)申请公布号 CN 110379413 A

(43)申请公布日 2019.10.25

(21)申请号 201910580572.4

(22)申请日 2019.06.28

(71)申请人 联想(北京)有限公司

地址 100085 北京市海淀区上地西路6号2  
幢2层201-H2-6

(72)发明人 赵泽清 汪俊杰

(74)专利代理机构 北京派特恩知识产权代理有  
限公司 11270

代理人 徐升升 张颖玲

(51)Int.Cl.

G10L 15/04(2013.01)

G10L 15/18(2013.01)

G10L 15/26(2006.01)

H04L 12/58(2006.01)

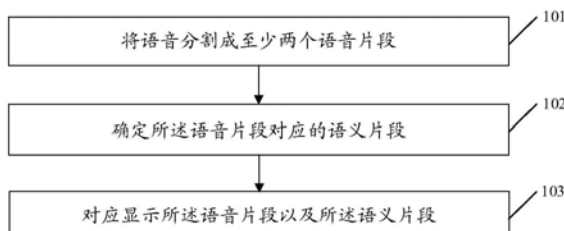
权利要求书1页 说明书11页 附图6页

(54)发明名称

一种语音处理方法、装置、设备及存储介质

(57)摘要

本申请实施例公开了一种语音处理方法,所述方法包括:将语音分割成至少两个语音片段;确定所述语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段。本申请实施例还公开了一种语音处理装置、设备及存储介质。



1. 一种语音处理方法,所述方法包括:  
将语音分割成至少两个语音片段;  
确定所述语音片段对应的语义片段;  
对应显示所述语音片段以及所述语义片段。
2. 根据权利要求1所述的方法,所述将语音分割成至少两个语音片段之前,包括:  
确定所述语音对应的时长;  
在所述时长大于预设指定时长的情况下,将语音分割成至少两个语音片段。
3. 根据权利要求1所述的方法,所述方法还包括:  
基于接收的针对所述语音的第一操作,显示操作界面;所述操作界面包括:语音分割按键;  
基于接收的针对所述语音分割按键的第二操作,生成语音分割指令;  
相应地,所述将语音分割成至少两个语音片段,包括:  
基于所述语音分割指令,将所述语音分割成至少两个语音片段。
4. 根据权利要求1所述的方法,所述将语音分割成至少两个语音片段,包括:  
根据所述语音中的间隔,确定所述语音的第一分割边界;  
根据所述第一分割边界,对所述语音进行分割,得到至少两个语音片段。
5. 根据权利要求1所述的方法,所述确定所述语音片段对应的语义片段,包括:  
确定所述语音片段对应的音素;  
将所述音素与设定的音素进行匹配;  
如果所述音素与设定的音素匹配,将所述设定的音素对应的语义信息确定为所述语音片段对应的语义片段。
6. 根据权利要求1所述的方法,所述确定所述语音片段对应的语义片段,包括:  
确定所述语音对应的语义信息;  
根据所述语音中的间隔,对所述语义信息添加标记;  
根据所述标记,确定所述语义信息的第二分割边界;  
根据所述第二分割边界,对所述语义信息进行分割,得到至少两个语义片段。
7. 根据权利要求1所述的方法,所述方法还包括:  
接收针对所述语音片段的第三操作;  
基于所述第三操作,对所述语音片段进行播放。
8. 一种语音处理装置,所述装置包括:分割模块、确定模块和显示模块;其中,  
所述分割模块,用于将语音分割成至少两个语音片段;  
所述确定模块,用于确定所述语音片段对应的语义片段;  
所述显示模块,用于对应显示所述语音片段以及所述语义片段。
9. 一种语音处理设备,包括处理器和用于存储能够在处理器上运行的计算机程序的存储器;其中,所述处理器用于运行所述计算机程序时,执行权利要求1至7任一项所述语音处理方法中的步骤。
10. 一种存储介质,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权利要求1至7任一项所述语音处理方法中的步骤。

## 一种语音处理方法、装置、设备及存储介质

### 技术领域

[0001] 本申请实施例涉及计算机技术领域,涉及但不限于一种语音处理方法、装置、设备及存储介质。

### 背景技术

[0002] 用户在聊天软件中有时会收到对方发来的一段较长的语音,如果想重新听某段语音,则需要从头开始听,比较麻烦。在相关技术中,目前的解决方案是将语音转成文字,当使用手指触碰对应文字的位置时,则从相应位置开始播放语音;但是,使用该解决方案在手机端进行操作时,由于手机端显示的文字比较小,很容易发生误触。

### 发明内容

[0003] 本申请实施例提供了一种语音处理方法、装置、设备及存储介质。

[0004] 本申请实施例的技术方案是这样实现的:

[0005] 第一方面,本申请实施例提供一种语音处理方法,所述方法包括:

[0006] 将语音分割成至少两个语音片段;

[0007] 确定所述语音片段对应的语义片段;

[0008] 对应显示所述语音片段以及所述语义片段。

[0009] 第二方面,本申请实施例提供一种语音处理装置,所述装置包括:分割模块、确定模块和显示模块;其中,

[0010] 所述分割模块,用于将语音分割成至少两个语音片段;

[0011] 所述确定模块,用于确定所述语音片段对应的语义片段;

[0012] 所述显示模块,用于对应显示所述语音片段以及所述语义片段。

[0013] 第三方面,本申请实施例还提供了一种语音处理设备,包括:处理器和用于存储能够在处理器上运行的计算机程序的存储器;其中,所述处理器用于运行所述计算机程序时,执行上述方案中任一项所述语音处理方法中的步骤。

[0014] 第四方面,本申请实施例还提供了一种存储介质,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方案中任一项所述语音处理方法中的步骤。

[0015] 本申请实施例中,将语音分割成至少两个语音片段;确定所述语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段;如此,用户可以根据显示的语义片段直接选择想要重新听的语音片段,更加方便快捷,提高了用户体验。

### 附图说明

[0016] 在附图(其不一定是按比例绘制的)中,相似的附图标记可在不同的视图中描述相似的部件。具有不同字母后缀的相似附图标记可表示相似部件的不同示例。附图以示例而非限制的方式大体示出了本文中所讨论的各个实施例。

[0017] 图1A为本申请实施例所提供的语音处理方法的实现流程示意图一;

- [0018] 图1B为本申请实施例所提供的语音处理方法的效果示意图一；
- [0019] 图2为本申请实施例所提供的语音处理方法的实现流程示意图二；
- [0020] 图3为本申请实施例所提供的语音处理方法的实现流程示意图三；
- [0021] 图4为本申请实施例所提供的语音处理方法的实现流程示意图四；
- [0022] 图5为本申请实施例所提供的语音处理方法的实现流程示意图五；
- [0023] 图6为本申请实施例所提供的语音处理方法的实现流程示意图六；
- [0024] 图7为本申请实施例所提供的语音处理方法的效果示意图二；
- [0025] 图8为本申请实施例所提供的语音处理方法的效果示意图三；
- [0026] 图9为本申请实施例所提供的语音处理方法的实现流程示意图七；
- [0027] 图10为本申请实施例所提供的语音处理装置的组成结构示意图；
- [0028] 图11为本申请实施例所提供的语音处理设备的硬件结构示意图。

### 具体实施方式

[0029] 为使本申请实施例的目的、技术方案和优点更加清楚，下面将结合本申请实施例中的附图，对本申请的具体技术方案做进一步详细描述。以下实施例用于说明本申请，但不用来限制本申请的范围。

[0030] 在详述本申请实施例时，为便于说明，表示器件结构的剖面图会不依一般比例作局部放大，而且所述示意图只是示例，其在此不应限制本申请保护的范围。此外，在实际制作中应包含长度、宽度及深度的三维空间尺寸。

[0031] 本申请实施例提供的语音处理方法可应用于语音处理装置，语音处理装置可实施于语音处理设备上。语音处理装置将接收到的语音分割成至少两个语音片段；确定至少两个语音片段对应的语义片段；对应显示至少两个语音片段以及所述语义片段，根据显示的语音片段和该语音片段对应的语义片段，用户可以选择需要重新听的语音片段。

[0032] 本申请实施例提供一种语音处理方法，该方法应用于实施语音处理方法的语音处理设备，语音处理设备中的各功能模块可以由语音处理设备（如终端设备、服务器）的硬件资源，如处理器等计算资源、传感器等探测资源、通信资源协同实现。

[0033] 语音处理设备可以是任何具有信息处理能力的电子设备，在一种实施例中，电子设备可以是智能终端，例如可以是笔记本等具有无线通信能力的移动终端、AR/VR设备。在另一种实施例中，电子设备还可以是不便移动的具有计算功能的终端设备，比如台式计算机、桌面电脑、服务器等。

[0034] 当然，本申请实施例不局限于提供为方法和硬件，还可有多种实现方式，例如提供为存储介质（存储有用于执行本申请实施例提供的语音处理方法的指令）。

[0035] 图1A为本申请实施例中的语音处理方法的实现流程示意图一，如图1A所示，该方法包括以下步骤：

[0036] 步骤101：将语音分割成至少两个语音片段；

[0037] 这里，语音处理设备接收到一段语音，将接收到的语音分割成至少两个语音片段。其中，语音处理设备接收的语音可以是另一设备发送的，也可以是服务器发送的。

[0038] 当语音处理设备接收到一段语音后，按照指定的规则，对接收到的语音进行分割；这里，指定的规则可以为按照语音的时长平均分割，也可以为按照语音中的间隔，对语音进

行分割。

[0039] 如果指定的规则为:按照语音的时长平均分割,则接收一段语音后,确定语音的时长,并对该时长进行平均分割得到平均时长,按照平均时长对接收到的语音进行平均分割。比如:接收一段时间长为18秒的语音,将该语音平均分割成长为9秒的语音片段。

[0040] 如果指定的规则为:按照语音中的间隔,对语音进行分割,则接收一段语音后,确定该语音中的间隔,根据确定的间隔,对接收到的语音进行分割。比如:接收的一段语音为“A1…A2…A3”,根据确定的该语音中的间隔“…”将该语音分割成语音片段“A1”、“A2”、“A3”。

[0041] 在对语音分割时,将语音分割成至少两个的语音片段。比如:将语音A分割为两个语音片段A1、A2;又比如:将语音A分割为五个语音片段A1、A2、A3、A4、A5。

[0042] 这里,在对语音进行分割之前可以先判断下该语音的时长,当该语音的时长较长时,再对该语音进行分割。

[0043] 步骤102:确定所述语音片段对应的语义片段;

[0044] 这里,在确定语音片段对应的语义片段时,可以根据语义识别模型来确定语音片段对应的语义片段。语义识别模型需要根据多种语音样本和语义样本来用过训练来获得。

[0045] 这里,当将语音分割成至少两个语音片段,确定该语音片段对应的语义片段。比如:分割的语音片段为“A1”、“A2”、“A3”,确定语音片段“A1”对应的语义片段为“这道题”,语音片段“A2”对应的语义片段为“太难了”,语音片段“A3”对应的语义片段为“我不会做”;又比如:分割的语音片段为“A1”、“A2”、“A3”、“A4”,确定语音片段“A1”对应的语义片段为“这次旅游”,语音片段“A2”对应的语义片段为“一定是很棒的体验”,语音片段“A3”对应的语义片段为“太棒了”,语音片段“A4”对应的语义片段为“真的很期待”。

[0046] 这里,在确定语音片段对应的语义片段时,可以对分割得到的至少两个语音片段,分别确定至少两个语音片段对应的语义片段;也可以确定该语音对应的语义,再将一段语义分割成多个语义片段。

[0047] 步骤103:对应显示所述语音片段以及所述语义片段。

[0048] 这里,在获得了至少两个语音片段和该语音片段对应的语义片段后,语音识别设备显示至少两个语音片段和该语音片段对应的语义片段。

[0049] 其中,显示的语音片段和该语音片段对应的语义片段,可以按照接收语音的先后顺序在语音识别设备的显示屏上依次排列。

[0050] 在显示语音片段和该语音片段对应的语义片段时,可以是在语音识别设备的显示屏上显示的。比如,如图1B所示:在语音识别设备的显示屏分别显示语音片段“A1”及语音片段“A1”对应的语义片段11为“这次旅游”,语音片段“A2”及语音片段“A2”对应的语义片段12为“一定是很棒的体验”,语音片段“A3”及语音片段“A3”对应的语义片段13为“太棒了”,语音片段“A4”及语音片段“A4”对应的语义片段14为“真的很期待”。

[0051] 本申请实施例提供的语音处理方法,将语音分割成至少两个语音片段;确定所述语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段;如此,用户可以根据显示的语义片段直接选择想要重新听的语音片段,更加方便快捷,提高了用户体验。

[0052] 本申请实施例中提供一种语音处理方法,如图2所示,该方法包括以下步骤:

[0053] 步骤201:确定所述语音对应的时长;

[0054] 这里,在语音处理设备接收到一段语音后,确定接收到的语音对应的时长,比如:语音处理设备接收到一段语音A,根据接收语音的开始时间和结束时间,确定接收到的语音A对应的时长为30秒。

[0055] 步骤202:在所述时长大于预设指定时长的情况下,将语音分割成至少两个语音片段;

[0056] 这里,预先设置指定时长,在接收到一段语音对应的时长大于该预设指定时长的情况下,可以认为接收到一段语音为长语音,并将该长语音分割成至少两个语音片段。

[0057] 比如:预设指定时长为20秒,接收到一段语音对应的时长为30秒,接收的一段语音对应的时长大于该预设指定时长,将30秒长的语音分割成至少两个语音片段。

[0058] 在接收到一段语音对应的时长小于该预设指定时长的情况下,可以认为接收到一段语音为短语音,不需要对其进行分割。

[0059] 需要说明的是,预设指定时长可以根据实际情况设置,本申请实施例对此不作限制。

[0060] 步骤203:确定所述语音片段对应的语义片段;

[0061] 步骤204:对应显示所述语音片段以及所述语义片段。

[0062] 其中,步骤203至步骤204分别参见上述实施例中的步骤102至步骤103。

[0063] 本申请实施例提供的语音处理方法,确定语音对应的时长;在时长大于预设指定时长的情况下,将语音分割成至少两个语音片段;确定语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段;如此,能够在确定语音为长语音时,对该语音进行分割,得到至少两个语音片段,提高了用户体验。

[0064] 本申请实施例中提供一种语音处理方法,如图3所示,该方法包括以下步骤:

[0065] 步骤301:基于接收的针对所述语音的第一操作,显示操作界面;

[0066] 这里,当语音处理设备接收到一段语音后,将接收到的语音显示到语音处理设备的显示屏上,用户对显示屏上语音进行操作,语音处理设备基于接收的针对语音的第一操作,显示操作界面;该操作界面包括:语音分割按键。

[0067] 比如:语音处理设备接收到一段时长为30秒的语音,并将该语音显示在显示屏上,用户对该语音进行右击操作,语音处理设备接收到针对语音的右击操作,显示操作界面;该操作界面包括:语音分割按键。

[0068] 需要说明的是,第一操作可以为点击、触摸等触控操作,本申请实施例对此不作限制。

[0069] 步骤302:基于接收的针对所述语音分割按键的第二操作,生成语音分割指令;

[0070] 这里,当语音处理设备的显示屏上显示操作界面后,用户针对显示操作界面上的语音分割按键进行操作,语音处理设备基于接收的针对语音分割按键的第二操作,生成语音分割指令。

[0071] 比如:用户对显示操作界面上的语音分割按键进行点击,语音处理设备接收到针对语音分割按键的点击操作,生成语音分割指令。

[0072] 需要说明的是,第二操作可以为点击、触摸等触控操作,本申请实施例对此不作限制。

[0073] 步骤303:基于所述语音分割指令,将所述语音分割成至少两个语音片段;

[0074] 这里,基于语音分割按键,生成语音分割指令后,基于语音分割指令,将语音分割成至少两个语音片段。

[0075] 其中,语音分割指令还可以携带语音分割的规则,根据语音分割指令和语音分割的规则,对语音进行分割。

[0076] 步骤304:确定所述语音片段对应的语义片段;

[0077] 步骤305:对应显示所述语音片段以及所述语义片段。

[0078] 其中,步骤304至步骤305分别参见上述实施例中的步骤102至步骤103。

[0079] 本申请实施例提供的语音处理方法,基于接收的针对语音的第一操作,显示操作界面;操作界面包括:语音分割按键;基于接收的针对语音分割按键的第二操作,生成语音分割指令;基于语音分割指令,将语音分割成至少两个语音片段;确定语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段;如此,能够基于接收的针对语音分割按键的第二操作,生成语音分割指令,对语音进行分割,提高了用户体验。

[0080] 本申请实施例中提供一种语音处理方法,如图4所示,该方法包括以下步骤:

[0081] 步骤401:根据所述语音中的间隔,确定所述语音的第一分割边界;

[0082] 这里,在将语音分割为至少两个语音片段时,根据语音中的间隔,确定语音的第一分割边界。其中,语音中的间隔可以为语音中的停顿。

[0083] 比如:接收到的语音为“A1…A2…A3”,将语音中的停顿“…”,确定为语音的第一分割边界。

[0084] 步骤402:根据所述第一分割边界,对所述语音进行分割,得到至少两个的语音片段;

[0085] 这里,根据确定的第一分割边界,对语音进行分割,得到至少两个语音片段。比如:接收到的语音为“A1…A2…A3”,根据语音中的停顿“…”,对语音进行分割,得到三个语音片段:语音片段“A1”、语音片段“A2”、语音片段“A3”。

[0086] 步骤403:确定所述语音片段对应的音素;

[0087] 这里,在确定语音片段对应的语义片段时,当获得了至少两个语音片段后,分别确定至少两个语音片段对应的语义片段。

[0088] 在确定至少两个语音片段对应的语义片段时,先确定该语音片段对应的音素。比如:至少两个语音片段为“A1”、“A2”,确定语音片段“A1”对应的音素为“zhedaoti”,确定语音片段“A2”对应的音素为“tainanle”。

[0089] 这里,可以根据语义识别模型,确定语音片段对应的音素,语义识别模型中预先存储语音片段中的每个文字的发音对应的音素。

[0090] 步骤404:将所述音素与设定的音素进行匹配;

[0091] 这里,将确定的语音片段对应的音素,与设定的音素进行匹配。比如:确定语音片段“A1”对应的音素为“zhedaoti”,将该音素与所有文字对应的音素进行匹配。

[0092] 这里,设定的音素可以为所有文字对应的音素,预先存储在语音处理设备中。

[0093] 步骤405:如果所述音素与设定的音素匹配,将所述设定的音素对应的语义信息确定为所述语音片段对应的语义片段;

[0094] 这里,在将音素与设定的音素进行匹配的过程中,如果语音片段对应的音素,与设定的音素匹配,将设定的音素对应的语义信息确定为语音片段对应的语义片段。

[0095] 比如:语音片段“A1”对应的音素为“zhedaoti”,与设定的音素“zhe”、“dao”、“ti”匹配,设定的音素“zhe”、“dao”、“ti”对应的语义信息分别为“这”、“道”、“题”,将语义信息“这”、“道”、“题”,确定为语音片段“A1”对应的语义片段“这道题”。

[0096] 步骤406:对应显示所述语音片段以及所述语义片段。

[0097] 其中,步骤406参见上述实施例中的步骤103。

[0098] 本申请实施例提供的语音处理方法,根据所述语音中的间隔,确定所述语音的第一分割边界;根据所述第一分割边界,对所述语音进行分割,得到至少两个语音片段;确定所述语音片段对应的音素;将所述音素与设定的音素进行匹配;如果所述音素与设定的音素匹配,将所述设定的音素对应的语义信息确定为所述语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段;如此,根据分割边界,对语音进行分割,并根据音素,确定语音片段对应的语义片段,提高了用户体验。

[0099] 本申请实施例中提供一种语音处理方法,如图5所示,该方法包括以下步骤:

[0100] 步骤501:将语音分割成至少两个语音片段;

[0101] 其中,步骤501参见上述实施例中的步骤101。

[0102] 步骤502:确定所述语音对应的语义信息;

[0103] 这里,在确定语义信息时,可以将接收到的一段语音作为一个整体,确定一段语音对应的语义信息。

[0104] 在确定一段语音对应的语义信息时,先确定该语音对应的音素,将确定的语音对应的音素,与设定的音素进行匹配,如果语音对应的音素,与设定的音素匹配,将设定的音素对应的语义信息确定为语音对应的语义信息。

[0105] 比如:一段语音为“A1…A2”,确定语音“A1…A2”对应的音素为“zhedaoti…tainanle”,将该音素与所有文字对应的音素进行匹配,语音“A1…A2”对应的音素为“zhedaoti…tainanle”,与设定的音素“zhe”、“dao”、“ti”、“tai”、“nan”、“le”匹配,设定的音素对应的语义信息分别为“这”、“道”、“题”、“太”、“难”、“了”,将语义信息“这”、“道”、“题”、“太”、“难”、“了”,确定为语音片段“A1…A2”对应的语义信息“这道题太难了”。

[0106] 这里,可以根据语义识别模型,确定语音片段对应的音素,语义识别模型中预先存储语音片段中的每个文字的发音对应的音素。设定的音素可以为所有文字对应的音素,预先存储在语音处理设备中。

[0107] 步骤503:根据所述语音中的间隔,对所述语义信息添加标记;

[0108] 这里,在确定了语义信息后,根据语音中的间隔,对语义信息添加标记;其中,语音中的间隔可以为语音中的停顿,对语义信息添加标记可以为对语义信息添加标点符号。

[0109] 比如:接收到的语音为“A1…A2”,对应的语义信息为“这道题太难了”,根据语音中的停顿“…”,对语义信息“这道题太难了”添加标点,得到“这道题,太难了”。

[0110] 步骤504:根据所述标记,确定所述语义信息的第二分割边界;

[0111] 这里,将对语义信息添加的标记,作为语义信息的第二分割边界。比如:语义信息为“这道题,太难了”,将语义信息中标点符号“,”,确定为该语义信息的第二分割边界。

[0112] 步骤505:根据所述第二分割边界,对所述语义信息进行分割,得到至少两个语义片段;

[0113] 这里,根据确定的第二分割边界,对语义信息进行分割,得到至少两个语义片段。



比如:接收到的语义信息为“这道题,太难了”,根据语义信息中的标点符号“,”,对语义信息进行分割,得到两个语义片段:语义片段“这道题”、语义片段“太难了”。

[0114] 步骤506:对应显示所述语音片段以及所述语义片段。

[0115] 其中,步骤506参见上述实施例中的步骤103。

[0116] 本申请实施例提供的语音处理方法,将语音分割成至少两个语音片段;确定所述语音对应的语义信息;根据所述语音中的间隔,对所述语义信息添加标记;根据所述标记,确定所述语义信息的第二分割边界;根据所述第二分割边界,对所述语义信息进行分割,得到至少两个语义片段;对应显示所述语音片段以及所述语义片段;如此,根据第二分割边界,对语义信息进行分割,得到至少两个语义片段,提高了用户的体验。

[0117] 本申请实施例中提供一种语音处理方法,如图6所示,该方法包括以下步骤:

[0118] 步骤601:将语音分割成至少两个语音片段;

[0119] 步骤602:确定所述语音片段对应的语义片段;

[0120] 步骤603:对应显示所述语音片段以及所述语义片段;

[0121] 其中,步骤601至步骤603分别参见上述实施例中的步骤101至步骤103。

[0122] 步骤604:接收针对所述语音片段的第三操作;

[0123] 这里,显示了语音片段和语音片段对应的语义片段后,接收针对语音片段的第三操作。比如:在语音识别设备的显示屏分别显示语音片段“A1”及语音片段“A1”对应的语义片段为“这次旅游”,语音片段“A2”及语音片段“A2”对应的语义片段为“一定是很棒的体验”,语音片段“A3”及语音片段“A3”对应的语义片段为“太棒了”,语音片段“A4”及语音片段“A4”对应的语义片段为“真的很期待”,用户点击语音片段“A1”,语音识别设备接收到针对语音片段“A1”的点击操作。

[0124] 需要说明的是,第三操作可以为点击、触摸等触控操作,本申请实施例对此不作限制。

[0125] 步骤605:基于所述第三操作,对所述语音片段进行播放。

[0126] 这里,基于语音识别设备接收到第三操作,对语音片段进行播放。比如:用户点击语音片段“A1”,语音识别设备接收到针对语音片段“A1”的点击操作,对语音片段“A1”进行播放。

[0127] 本申请实施例提供的语音处理方法,将语音分割成至少两个语音片段;确定所述语音片段对应的语义片段;对应显示所述语音片段以及所述语义片段;接收针对所述语音片段的第三操作;基于所述第三操作,对所述语音片段进行播放,如此,用户可以根据显示的语义片段直接选择想要重新听的语音片段,更加方便快捷,提高了用户体验。

[0128] 本申请实施例中通过具体场景对本申请实施例提供的语音处理方法进行说明。

[0129] 在本申请实施例中,在用户收到对方发送的长语音后,选择“长语音分割”按键,根据语音中的停顿将长语音及其识别文本分割成片段,展示给用户。

[0130] 在一示例中,如图7所示,当语音处理设备,接收到长语音71时,显示屏上显示功能选择框72,显示功能选择框72中包括:使用听筒模式按键73、收藏按键74、长语音分割按键75,基于用户在功能选择框72上的操作,长语音分割按键75被选中。语音处理设备执行长语音分割按键75对应的分割功能,对接收到长语音71进行分割,得到该长语音71对应的三个语音片段及对应的文字并对应显示,如图8所示:语音片段A1及对应的文字74“这次旅游是

一次很棒的经验”、语音片段A2及对应的文字75“太棒啦”、语音片段A3及对应的文字76“快快准备东西吧”。

[0131] 本申请实施例的语音处理方法的实现流程示意图,如图9所示:

[0132] 步骤901:将长语音识别成文本,并加标点。

[0133] 步骤902:根据标点将长语音和文本分割成语音片段及其对应的文本片段。

[0134] 步骤903:将语音片段及其文本片段展示给用户。

[0135] 本申请实施例的技术效果为:用户可以根据识别出的文本直接选择想要重新听的语音片段,更加方便快捷,提高了用户体验。

[0136] 本申请实施例还提供一种语音处理装置,该装置所包括的各模块、各模块所包括的各单元,可以通过语音处理装置的处理器来实现;当然也可通过具体的逻辑电路实现;在实施的过程中,处理器可以为中央处理器(CPU)、微处理器(MPU)、数字信号处理器(DSP)或现场可编程门阵列(FPGA)等。

[0137] 如图10所示,语音处理装置100包括:

[0138] 分割模块1001,用于将语音分割成至少两个语音片段;

[0139] 确定模块1002,用于确定所述语音片段对应的语义片段;

[0140] 显示模块1003,用于对应显示所述语音片段以及所述语义片段。

[0141] 在一些实施例中,分割模块1001包括:第一确定单元和分割单元;其中,

[0142] 第一确定单元,用于确定所述语音对应的时长;

[0143] 分割单元,用于在所述时长大于预设指定时长的情况下,将语音分割成至少两个语音片段。

[0144] 在一些实施例中,语音处理装置100还包括:显示模块和生成模块;其中,

[0145] 显示模块,用于基于接收的针对所述语音的第一操作,显示操作界面;所述操作界面包括:语音分割按键;

[0146] 生成模块,用于基于接收的针对所述语音分割按键的第二操作,生成语音分割指令;

[0147] 相应地,分割模块1001,用于基于所述语音分割指令,将所述语音分割成至少两个语音片段。

[0148] 在一些实施例中,分割模块1001还包括:第二确定单元和第三确定单元;其中,

[0149] 第二确定单元,用于根据所述语音中的间隔,确定所述语音的第一分割边界;

[0150] 第三确定单元,用于根据所述第一分割边界,对所述语音进行分割,得到至少两个语音片段。

[0151] 在一些实施例中,确定模块1002还包括:第四确定单元、匹配单元和第五确定单元;其中,

[0152] 第四确定单元,用于确定所述语音片段对应的音素;

[0153] 匹配单元,用于将所述音素与设定的音素进行匹配;

[0154] 第五确定单元,用于如果所述音素与设定的音素匹配,将所述设定的音素对应的语义信息确定为所述语音片段对应的语义片段。

[0155] 在一些实施例中,确定模块1002还包括:第六确定单元、标记单元、第七确定单元和第八确定单元;其中,

- [0156] 第六确定单元,用于确定所述语音对应的语义信息;
- [0157] 标记单元,用于根据所述语音中的间隔,对所述语义信息添加标记;
- [0158] 第七确定单元,用于根据所述标记,确定所述语义信息的第二分割边界;
- [0159] 第八确定单元,用于根据所述第二分割边界,对所述语义信息进行分割,得到至少两个语义片段。
- [0160] 在一些实施例中,语音处理装置100还包括:接收模块和播放模块;其中,
- [0161] 接收模块,用于接收针对所述语音片段的第三操作;
- [0162] 播放模块,用于基于所述第三操作,对所述语音片段进行播放。
- [0163] 需要说明的是:上述实施例提供的语音处理装置在语音处理时,仅以上述各程序模块的划分进行举例说明,实际应用中,可以根据需要而将上述处理分配由不同的程序模块完成,即将装置的内部结构划分成不同的程序模块,以完成以上描述的全部或者部分处理。另外,上述实施例提供的语音处理装置与语音处理方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。
- [0164] 图11所示的语音处理设备110包括:至少一个处理器1110、存储器1140、至少一个网络接口1120和用户接口1130。语音处理设备110中的各个组件通过总线系统1150耦合在一起。可理解,总线系统1150用于实现这些组件之间的连接通信。总线系统1150除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图11中将各种总线都标为总线系统1150。
- [0165] 用户接口1130可以包括显示器、键盘、鼠标、轨迹球、点击轮、按键、按钮、触感板或者触摸屏等。
- [0166] 存储器1140可以是易失性存储器或非易失性存储器,也可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(ROM,Read Only Memory)。易失性存储器可以是随机存取存储器(RAM,Random Access Memory)。本发明实施例描述的存储器1140旨在包括任意适合类型的存储器。
- [0167] 本发明实施例中的存储器1140能够存储数据以支持语音处理设备110的操作。这些数据的示例包括:用于在语音处理设备110上操作的任何计算机程序,如操作系统和应用程序。其中,操作系统包含各种系统程序,例如框架层、核心库层、驱动层等,用于实现各种基础业务以及处理基于硬件的任务。应用程序可以包含各种应用程序。
- [0168] 其中,处理器1110用于运行所述计算机程序时,以实现上述实施例中提供的语音处理方法中的步骤。
- [0169] 作为本发明实施例提供的方法采用软硬件结合实施的示例,本发明实施例所提供的方法可以直接体现为由处理器1110执行的软件模块组合,例如本本发明实施例提供的语音处理装置,语音处理装置的软件模块可以存储于存储器1140,处理器1110读取存储器1140中软件模块包括的可执行指令,结合必要的硬件(例如,包括处理器1110以及连接到总线1150的其他组件)完成本发明实施例提供的语音处理方法。
- [0170] 作为示例,处理器1110可以是一种集成电路芯片,具有信号的处理能力,例如通用处理器、数字信号处理器(DSP,Digital Signal Processor),或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等,其中,通用处理器可以是微处理器或者任何常规的处理器等。

[0171] 这里需要指出的是:以上语音处理设备实施例的描述,与上述方法描述是类似的,具有同方法实施例相同的有益效果,因此不做赘述。对于本申请语音处理设备实施例中未披露的技术细节,本领域的技术人员请参照本申请方法实施例的描述而理解,为节约篇幅,这里不再赘述。

[0172] 在示例性实施例中,本申请实施例还提供了一种存储介质,可以为计算机可读存储介质,例如包括存储计算机程序的存储器,上述计算机程序可由处理器处理,以完成前述方法所述步骤。计算机可读存储介质可以是FRAM、ROM、PROM、EPROM、EEPROM、Flash Memory、磁表面存储器、光盘、或CD-ROM等存储器。

[0173] 本申请实施例还提供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处理时实现上述实施例中提供的语音处理方法中的步骤。

[0174] 这里需要指出的是:以上计算机介质实施例的描述,与上述方法描述是类似的,具有同方法实施例相同的有益效果,因此不做赘述。对于本申请存储介质实施例中未披露的技术细节,本领域的技术人员请参照本申请方法实施例的描述而理解,为节约篇幅,这里不再赘述。

[0175] 上述本申请实施例揭示的方法可以应用于所述处理器中,或者由所述处理器实现。所述处理器可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过所述处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的所述处理器可以是通用处理器、DSP,或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。所述处理器可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者任何常规的处理器等。结合本申请实施例所公开的方法的步骤,可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于存储介质中,该存储介质位于存储器,所述处理器读取存储器中的信息,结合其硬件完成前述方法的步骤。

[0176] 可以理解,本申请实施例的存储器(存储器)可以是易失性存储器或者非易失性存储器,也可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是ROM、可编程只读存储器(PROM,Programmable Read-Only Memory)、可擦除可编程只读存储器(EPROM,Erasable Programmable Read-Only Memory)、电可擦除可编程只读存储器(EEPROM,Electrically Erasable Programmable Read-Only Memory)、磁性随机存取存储器(FRAM,ferromagnetic random access memory)、快闪存储器(Flash Memory)、磁表面存储器、光盘、或只读光盘(CD-ROM,Compact Disc Read-Only Memory);磁表面存储器可以是磁盘存储器或磁带存储器。易失性存储器可以是随机存取存储器(RAM,Random Access Memory),其用作外部高速缓存。通过示例性但不是限制性说明,许多形式的RAM可用,例如静态随机存取存储器(SRAM,Static Random Access Memory)、同步静态随机存取存储器(SSRAM,Synchronous Static Random Access Memory)、动态随机存取存储器(DRAM,Dynamic Random Access Memory)、同步动态随机存取存储器(SDRAM,Synchronous Dynamic Random Access Memory)、双倍数据速率同步动态随机存取存储器(DDRSDRAM,Double Data Rate Synchronous Dynamic Random Access Memory)、增强型同步动态随机存取存储器(ESDRAM,Enhanced Synchronous Dynamic Random Access Memory)、同步连接动态随机存取存储器(SLDRAM,SyncLink Dynamic Random Access Memory)、直接内存总线随机存取存

储器(DRRAM, Direct Rambus Random Access Memory)。本申请实施例描述的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

[0177] 本领域内的技术人员应明白,本申请实施例的语音处理方法的其他构成以及作用,对于本领域的技术人员而言都是已知的,为了减少冗余,本申请实施例不做赘述。

[0178] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”“具体示例”或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本申请的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0179] 尽管已经示出和描述了本申请的实施例,本领域的普通技术人员可以理解:在不脱离本申请的原理和宗旨的情况下可以对这些实施例进行多种变化、修改、替换和变型,本申请的范围由权利要求及其等同限定。

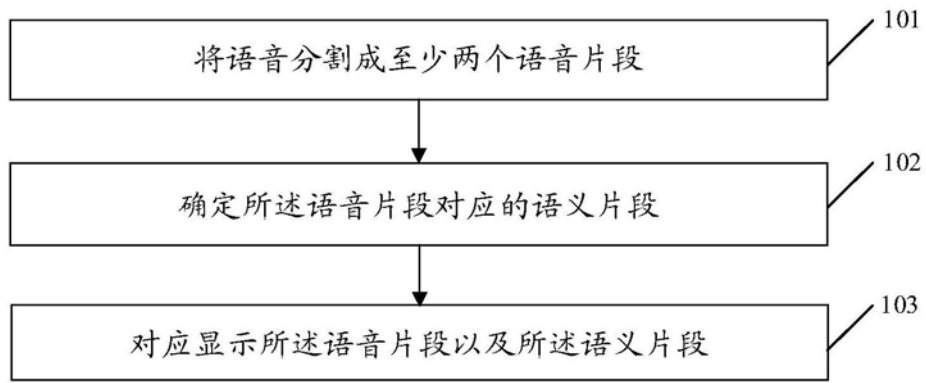


图1A

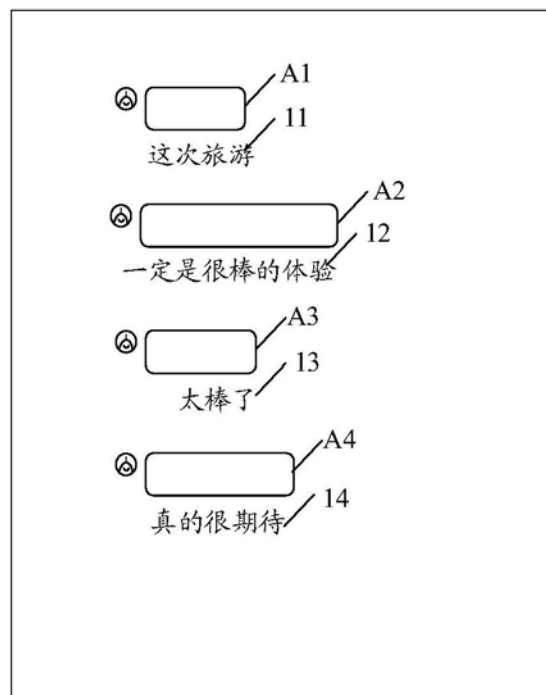


图1B

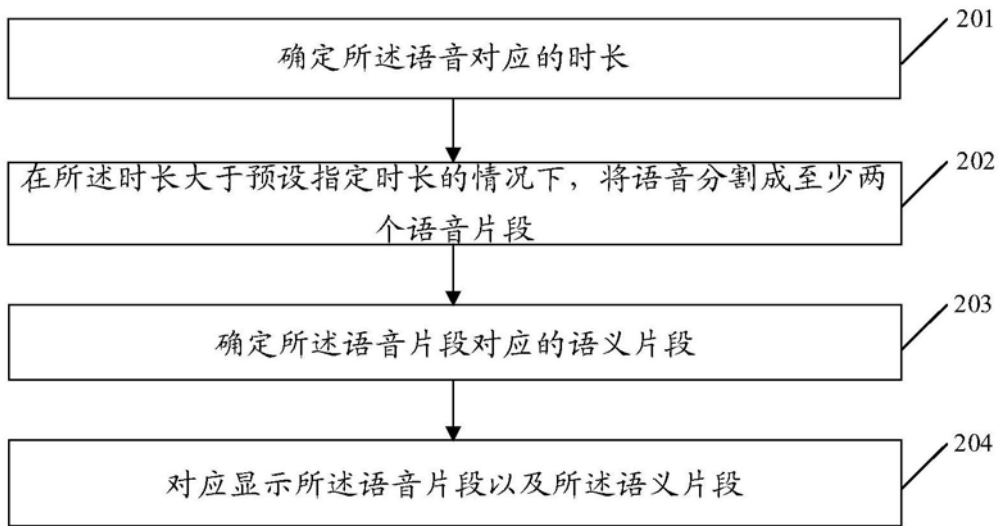


图2

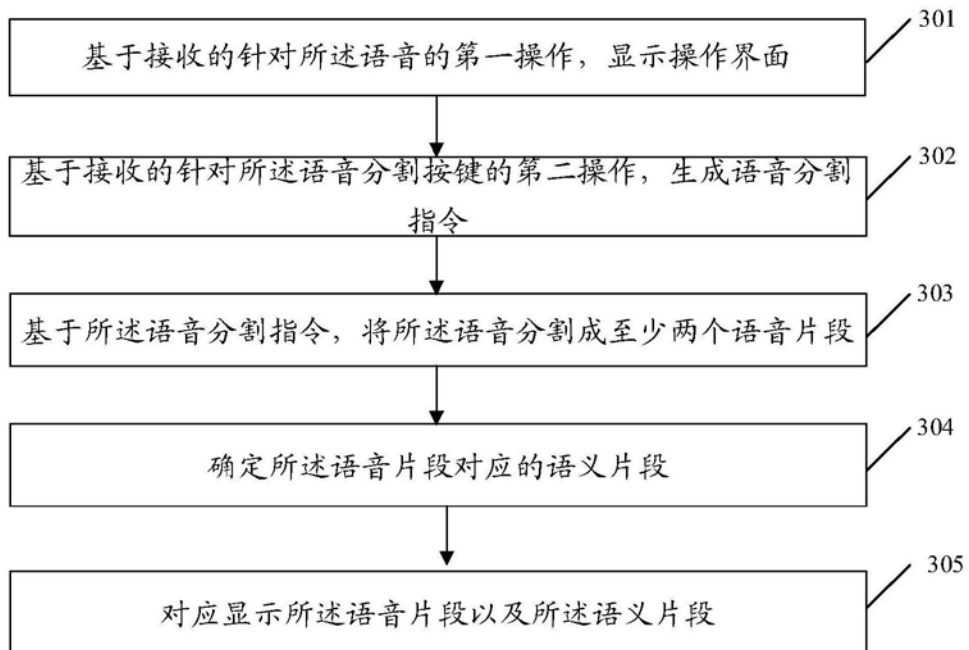


图3

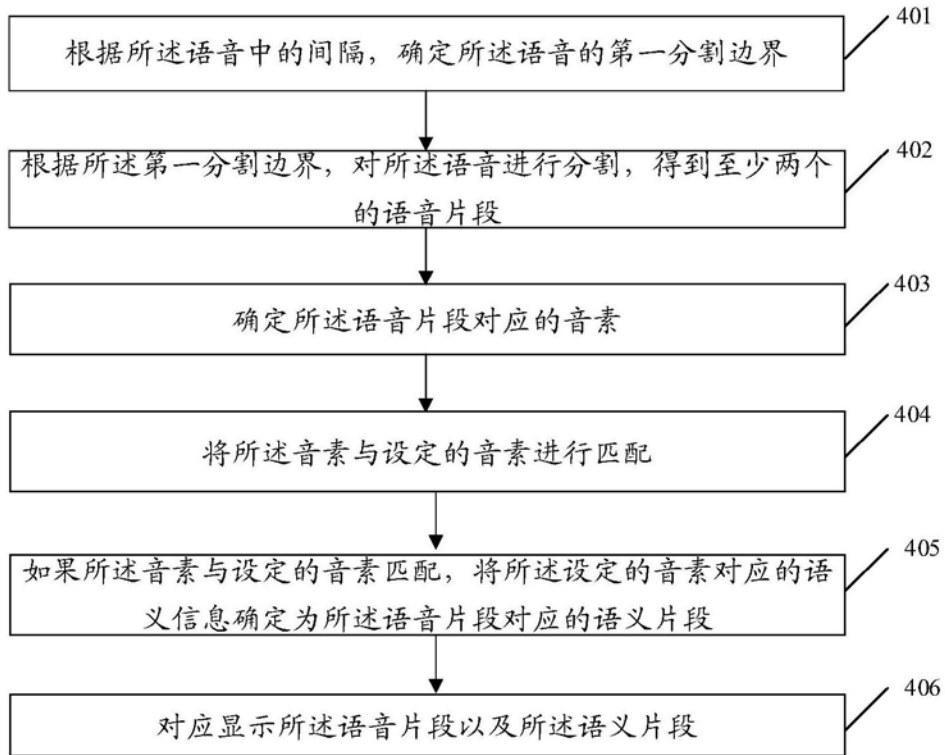


图4

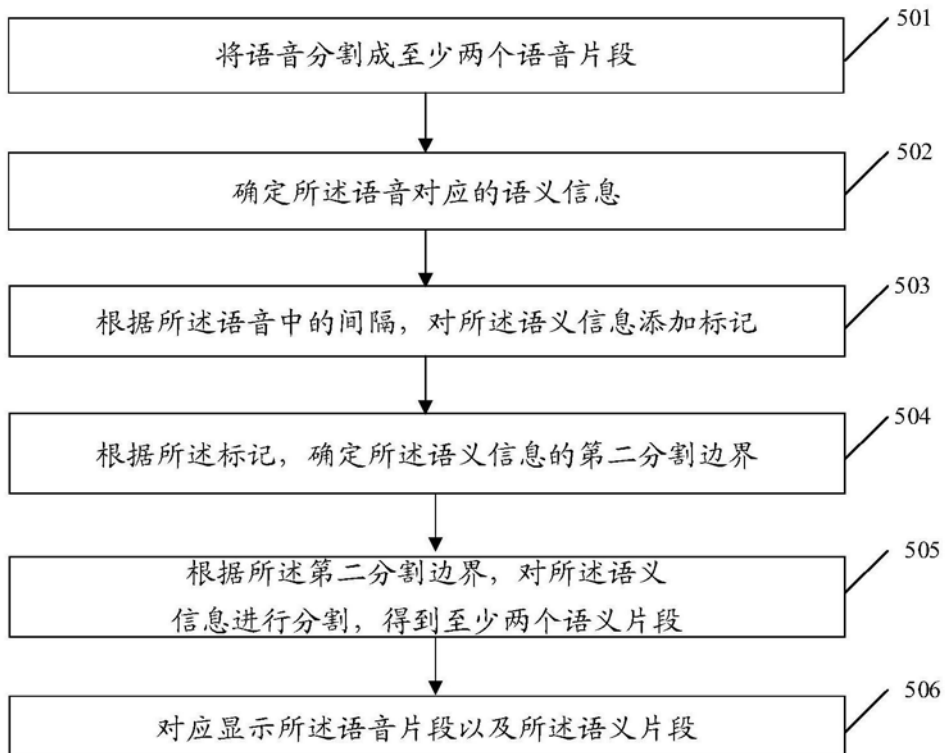


图5



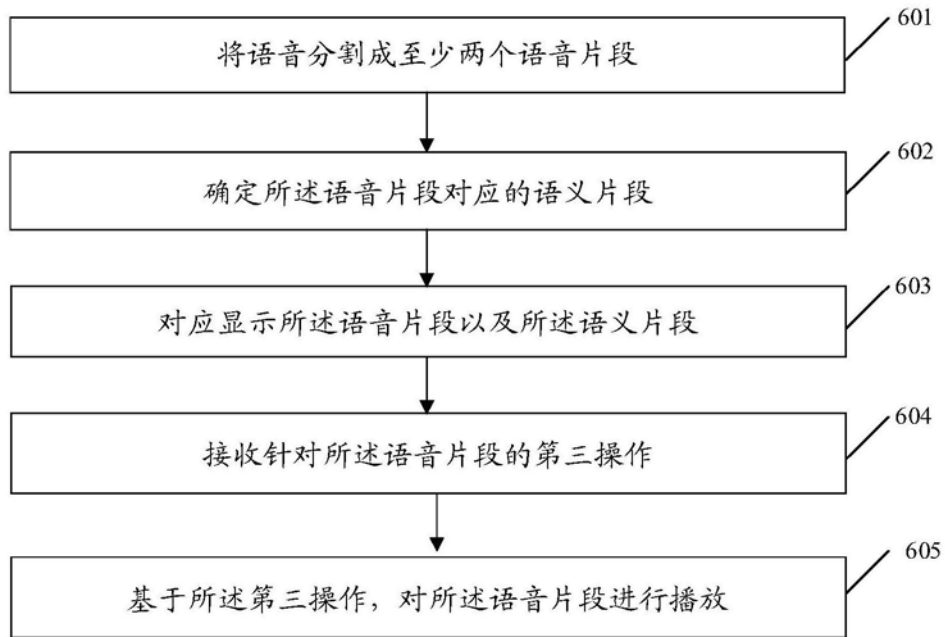


图6

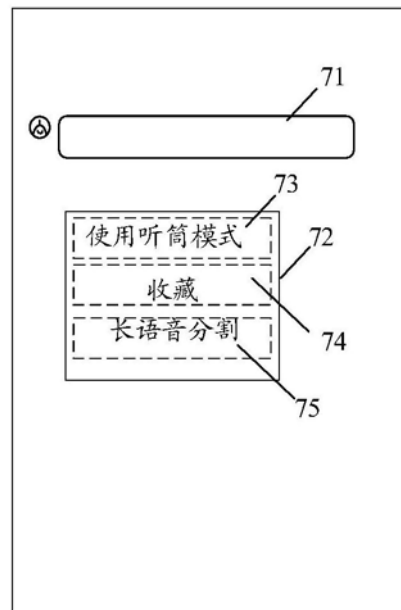


图7

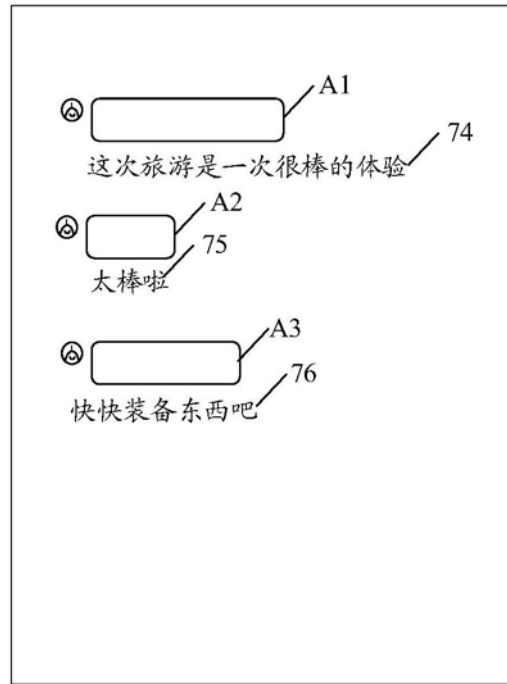


图8

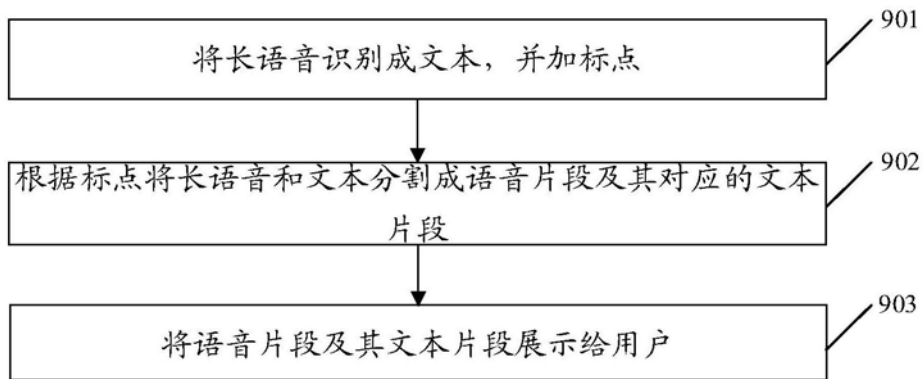


图9

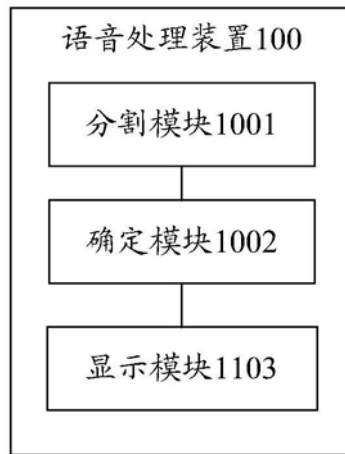


图10

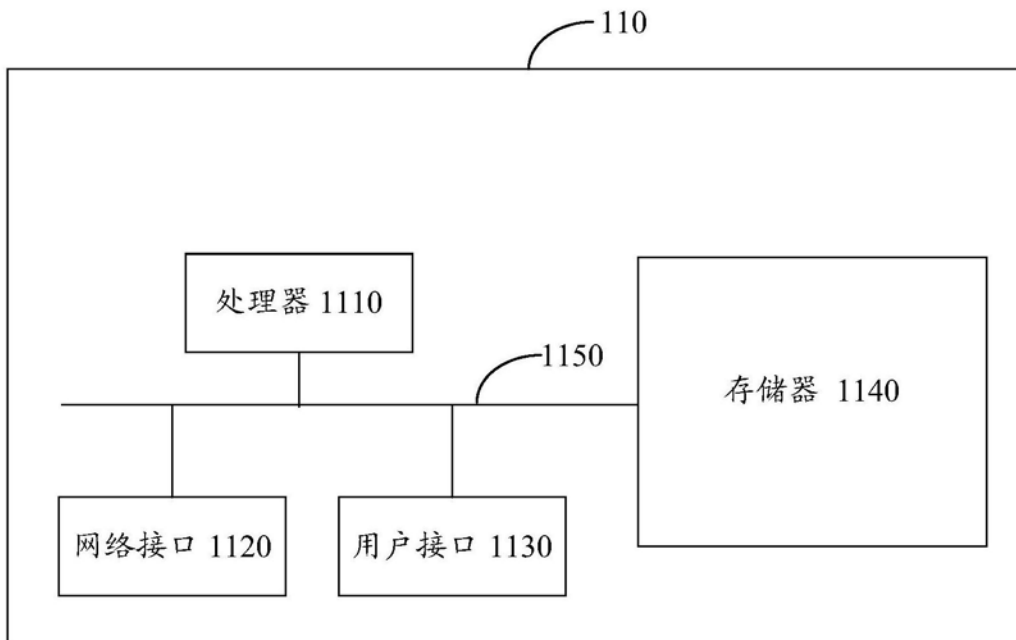


图11