



(12)发明专利申请

(10)申请公布号 CN 112309429 A

(43)申请公布日 2021.02.02

(21)申请号 201910695671.7

(22)申请日 2019.07.30

(71)申请人 上海流利说信息技术有限公司  
地址 200090 上海市杨浦区长阳路1687号  
西1168幢(C楼)一层C2101室

(72)发明人 刘晨晨 沈欣尧 顾怡炜 张蕾

(74)专利代理机构 北京信远达知识产权代理有限公司 11304

代理人 李兆轩

(51) Int. Cl.

G10L 25/51(2013.01)

G09B 19/06(2006.01)

G09B 5/04(2006.01)

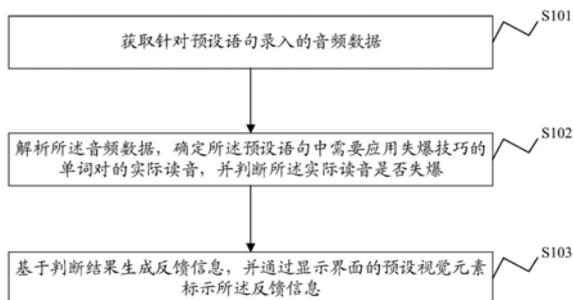
权利要求书2页 说明书10页 附图6页

(54)发明名称

一种失爆检测方法、装置、设备及计算机可读存储介质

(57)摘要

本申请公开了一种失爆检测方法、装置、设备和计算机可读存储介质,该方法包括:获取针对预设语句录入的音频数据;解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。本申请能够自动对录入的音频数据进行分析,以判断其中需要应用失爆技巧的单词对的实际读音是否失爆,生成相应的反馈信息,能够有效帮助学习者理解失爆的概念和掌握失爆技巧;并且避免了人工教学和纠错的过程,突破了学习时间和空间的限制,显著节约了学习成本。



1. 一种失爆检测方法,其特征在于,包括:  
获取针对预设语句录入的音频数据;  
解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;  
基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。
2. 根据权利要求1所述的失爆检测方法,其特征在于,所述获取针对预设语句录入的音频数据之前,还包括:  
利用所述预设语句的单词发音及预设失爆规则,筛选出需要应用失爆技巧的单词对;  
对所述单词对插入预设符号构建生成新的单词,并根据所述单词应用失爆技巧后的发音方式构建对应的音素序列,将所述音素序列添加至发音词典中。
3. 根据权利要求2所述的失爆检测方法,其特征在于,所述筛选出需要应用失爆技巧的单词对之后,还包括:  
通过所述显示界面的第一视觉元素对所述单词对中失爆部分的音标和字母进行标示,以提示所述失爆部分的发音特征。
4. 根据权利要求2所述的失爆检测方法,其特征在于,所述解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆,包括:  
解析所述音频数据,根据单词时间边界截取需要应用失爆技巧的单词对所对应的声学模型输出片段;  
将所述声学模型输出片段输入至预先基于所述发音词典构建的解码网络中,得到判断所述实际读音是否失爆的判断结果。
5. 根据权利要求1至4任一项所述的失爆检测方法,其特征在于,所述基于判断结果生成反馈信息,包括:  
若判定所述预设语句中需要应用失爆技巧的单词对的实际读音失爆,则判断所述实际读音是否符合预设条件;所述预设条件为所述实际读音的发音分数大于第一阈值,和/或所述实际读音的音素时长小于第二阈值;  
若判定所述实际读音符合所述预设条件,则生成所述实际读音正确失爆的反馈信息;  
若判定所述实际读音不符合所述预设条件,则生成所述实际读音未正确失爆的反馈信息。
6. 根据权利要求5所述的失爆检测方法,其特征在于,所述通过显示界面的预设视觉元素标示所述反馈信息之后,还包括:  
通过文字和/或语音的方式对需要应用失爆技巧的单词对的正确发音进行提示。
7. 一种失爆检测装置,其特征在于,包括:  
获取模块,用于获取针对预设语句录入的音频数据;  
判断模块,用于解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;  
生成模块,用于基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。
8. 一种失爆检测设备,其特征在于,应用于服务端,所述设备包括:

存储器,用于存储计算机程序;

处理器,用于执行所述计算机程序时实现如权利要求1至6任一项所述失爆检测方法的步骤。

9. 一种失爆检测设备,其特征在于,应用于客户端,所述设备包括:

音频采集装置,用于录入针对预设语句的音频数据;

通信装置,用于将所述音频数据发送至服务端,以便所述服务端解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息;

显示装置,用于将标示所述反馈信息的预设视觉元素在显示界面上进行显示。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质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如权利要求1至6任一项所述失爆检测方法的步骤。

## 一种失爆检测方法、装置、设备及计算机可读存储介质

### 技术领域

[0001] 本申请涉及语音技术领域,更具体地说,涉及一种失爆检测方法、装置、设备和一种计算机可读存储介质。

### 背景技术

[0002] 随着科学技术的发展,基于互联网的语言学习应用也得到了快速的发展。在一些语言学习应用中,应用提供商通过互联网将学习材料发送到客户端,用户经由客户端获取学习材料,进行对应的学习。对于语言学习,除了学习语法和词汇之外,发音能力是其中最重要的能力之一。通常情况下,用户会通过朗读、跟读等方式来提升自身的发音能力。然而,多数情况下用户无法得知自身发音是否准确。

[0003] 在现代英语的朗读和学习过程中,爆破音占有重要地位,爆破音失爆的现象也很普遍,失爆具体是指当爆破音,如/p/、/b/、/t/、/d/、/k/、/g/后面带有其他辅音时,爆破音不会释放气流。由于失爆时需要学习者在做到爆破音口型的基础上不让气流泄出,而在中文中不存在类似的发音现象,掌握起来较为困难。

[0004] 传统的学习方案是通过人工进行现场教学和纠错,然而口语的学习需要不断进行练习,人工教学和纠错不但成本高,也对学习者的学习时间和空间造成了一定的限制。

### 发明内容

[0005] 本申请的目的在于提供一种失爆检测方法、装置、设备和一种计算机可读存储介质,以解决传统方法学习效率较低且有效练习的时间和空间受到限制的问题。

[0006] 为实现上述目的,本申请提供了一种失爆检测方法,包括:

[0007] 获取针对预设语句录入的音频数据;

[0008] 解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;

[0009] 基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。

[0010] 可选的,所述获取针对预设语句录入的音频数据之前,还包括:

[0011] 利用所述预设语句的单词发音及预设失爆规则,筛选出需要应用失爆技巧的单词对;

[0012] 对所述单词对插入预设符号构建生成新的单词,并根据所述单词应用失爆技巧后的发音方式构建对应的音素序列,将所述音素序列添加至发音词典中。

[0013] 可选的,所述筛选出需要应用失爆技巧的单词对之后,还包括:

[0014] 通过所述显示界面的第一视觉元素对所述单词对中失爆部分的音标和字母进行标示,以提示所述失爆部分的发音特征。

[0015] 可选的,所述解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆,包括:

[0016] 解析所述音频数据,根据单词时间边界截取需要应用失爆技巧的单词对所对应的声学模型输出片段;

[0017] 将所述声学模型输出片段输入至预先基于所述发音词典构建的解码网络中,得到判断所述实际读音是否失爆的判断结果。

[0018] 可选的,所述基于判断结果生成反馈信息,包括:

[0019] 若判定所述预设语句中需要应用失爆技巧的单词对的实际读音失爆,则判断所述实际读音是否符合预设条件;所述预设条件为所述实际读音的发音分数大于第一阈值,和/或所述实际读音的音素时长小于第二阈值;

[0020] 若判定所述实际读音符合所述预设条件,则生成所述实际读音正确失爆的反馈信息;

[0021] 若判定所述实际读音不符合所述预设条件,则生成所述实际读音未正确失爆的反馈信息。

[0022] 可选的,所述通过显示界面的预设视觉元素标示所述反馈信息之后,还包括:

[0023] 通过文字和/或语音的方式对需要应用失爆技巧的单词对的正确发音进行提示。

[0024] 为实现上述目的,本申请提供了一种失爆检测装置,包括:

[0025] 获取模块,用于获取针对预设语句录入的音频数据;

[0026] 判断模块,用于解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;

[0027] 生成模块,用于基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。

[0028] 为实现上述目的,本申请提供了一种失爆检测设备,应用于服务端,所述设备包括:

[0029] 存储器,用于存储计算机程序;

[0030] 处理器,用于执行所述计算机程序时实现前述公开的任一种失爆检测方法的步骤。

[0031] 为实现上述目的,本申请提供了一种失爆检测设备,应用于客户端,所述设备包括:

[0032] 音频采集装置,用于录入针对预设语句的音频数据;

[0033] 通信装置,用于将所述音频数据发送至服务端,以便所述服务端解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息;

[0034] 显示装置,用于将标示所述反馈信息的预设视觉元素在显示界面上进行显示。

[0035] 为实现上述目的,本申请提供了一种计算机可读存储介质,所述计算机可读存储介质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前述公开的任一种失爆检测方法的步骤。

[0036] 通过以上方案可知,本申请提供了一种失爆检测方法,包括:获取针对预设语句录入的音频数据;解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。本申请能够自动对录入的音频数据进行分析,以判断其中

需要应用失爆技巧的单词对的实际读音是否失爆,生成相应的反馈信息,能够有效帮助学习者理解失爆的概念并掌握失爆技巧;并且避免了人工教学和纠错的过程,突破了学习时间和空间的限制,显著节约了学习成本。

[0037] 本申请还公开了一种失爆检测装置、设备和一种计算机可读存储介质,同样能实现上述技术效果。

[0038] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性的,并不能限制本申请。

## 附图说明

[0039] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,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。

[0040] 图1为本申请实施例公开的一种失爆检测方法的流程图;

[0041] 图2为本申请实施例公开的另一种失爆检测方法的流程图;

[0042] 图3为在显示界面对用户失爆练习进行提示的视觉化展现示意图;

[0043] 图4为本申请实施例公开的又一种失爆检测方法的流程图;

[0044] 图5为在显示界面对用户失爆正误进行反馈的视觉化展现示意图;

[0045] 图6为本发明实施例提供的失爆检测装置的结构框图;

[0046] 图7为本发明实施例提供的失爆检测设备应用于服务端的结构框图;

[0047] 图8为本发明实施例提供的失爆检测设备应用于客户端的结构框图;

[0048] 图9为本发明实施例提供的失爆检测系统的结构框图。

## 具体实施方式

[0049] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,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,都属于本申请保护的范围。

[0050] 本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等(如果存在)是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的实施例能够以除了在这里图示或描述的内容以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0051] 需要说明的是,在本发明中涉及“第一”、“第二”等的描述仅用于描述目的,而不能理解为指示或暗示其相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。另外,各个实施例之间的技术方案可以相互结合,但是必须是以本领域普通技术人员能够实现为基础,当技术方案的

结合出现相互矛盾或无法实现时应当认为这种技术方案的结合不存在,也不在本发明要求的保护范围之内。

[0052] 本发明实施例可以用于发音学习场景,尤其是语言学习中的发音学习场景或者发音校正场景,其中语言包括但不限于英语、法语、德语、日语等外语,以及粤语、四川话等汉语分支。本发明实施例涉及的语言学习场景例如可以是语言学习软件或语言学习终端中的发音测评场景、发音校正场景等场景,也可以是其他语言学习场景,本发明实施例中并不限定。

[0053] 下面对本申请实施例的应用场景进行详细阐述,用户可以通过客户端进行发音学习,客户端可以在显示界面上显示用户待学习的内容,并且还可以通过扬声器等音频播放装置向用户输出语音形式的音频内容。在用户进行语音的发音学习时,客户端可以通过音频采集装置采集用户发音时的音频数据,以便后续进行失爆检测操作。可以理解的是,执行失爆检测操作的主体可以为客户端,也可以为服务端,这均不影响本申请的实现。

[0054] 本发明实施例中客户端可以包括但不限于:智能手机、平板电脑、MP4、MP3、PC、PDA、可穿戴设备和头戴显示设备等;服务端可以包括但不限于:单个网络服务器、多个网络服务器组成的服务器组或基于云计算由大量计算机或网络服务器构成的云。

[0055] 结合上述应用场景,本申请所提供的失爆检测方法的一种具体实施方式的流程图如图1所示,该方法具体包括:

[0056] S101:获取针对预设语句录入的音频数据;

[0057] 本实施例中,预设语句可以为需要应用失爆技巧的语句,具体包括一个或多个语句,每个语句可以包括两个及以上单词。用户可以通过显示屏查看该预设语句并进行朗读,以通过客户端录入针对该预设语句的语音,经过音频采集装置采集后获得语音对应的音频数据。

[0058] S102:解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;

[0059] 在本步骤中,通过对音频数据进行解析,得到预设语句中需要应用失爆技巧的单词对的实际读音,并判断上述实际语音是否失爆。其中,该过程可以由客户端执行,也可以由服务端执行,本实施例对此不做限定。

[0060] S103:基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。

[0061] 可以理解的是,本实施例基于上述判断实际语音是否失爆的结果可以生成相应的反馈信息,并通过显示界面上的预设视觉元素向用户进行直观反馈。当然,还可以通过播放相应的音效来进行反馈,这均不影响本申请的实现。

[0062] 通过以上方案可知,本申请提供的一种失爆检测方法,包括:获取针对预设语句录入的音频数据;解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。本申请能够自动对录入的音频数据进行分析,以判断其中需要应用失爆技巧的单词对的实际读音是否失爆,生成相应的反馈信息,能够辅助学习者理解失爆的概念,有效帮助学习者掌握失爆技巧;并且避免了人工教学和纠错的过程,突破了学习时间和空间的限制,显著节约了学习成本。

[0063] 本申请实施例公开了另一种失爆检测方法,相对于上一实施例,本实施例对技术方案作了进一步的说明和优化。参见图2所示,具体的:

[0064] S201:利用预设语句的单词发音及预设失爆规则,筛选出需要应用失爆技巧的单词对;

[0065] 在具体实施中,可以通过遍历预设语句中的相邻单词对,基于预设失爆规则筛选出需要应用失爆技巧的单词对,其中,预设失爆规则为需要进行失爆的条件信息,例如,前一词以“持续音+/t//d/”结尾,后一词以辅音开头,则其中/t//d/需要应用失爆技巧。

[0066] S202:对所述单词对插入预设符号构建生成新的单词,并根据所述单词应用失爆技巧后的发音方式构建对应的音素序列,将所述音素序列添加至发音词典中;

[0067] 本步骤中,在筛选得到需要应用失爆技巧的单词对之后,在该单词对的中间插入预设符号构建生成新的单词,同时根据其在失爆时的发音方式构建对应的音素序列,例如,预设符号可以为@@,当检测到单词对为“stick to”时,对其插入预设符号,构建生成新的单词“stick@@to”,对应的发音为/stitu/,其中k失爆,将新构建的单词对应发音添加至语音识别系统的发音词典中。其中,发音词典是构建传统语音识别系统的必要组成部分,识别系统仅会输出存在于发音词典中的单词,发音词典可以预先构建得到,在失爆检测过程中会根据需求向发音词典中添加新构建的单词发音。

[0068] 可以理解的是,本实施例在筛选出需要应用失爆技巧的单词对之后,还可以通过显示界面的第一视觉元素对单词对中失爆部分的音标和字母进行标示,以提示失爆部分的发音特征。其中,第一视觉元素可以为虚线,在显示界面展示需要应用失爆技巧的短语或句子及其对应的音标,其中失爆部分的字母,利用虚线从中间划开,虚线右边部分利用预设颜色进行标记,以提示用户无需完整的发音;对于失爆部分的音标,利用虚线从中间划开,虚线左边利用特殊颜色标注,同时,失爆音素右边紧邻的整个音素用特殊颜色标注,以提醒用户注意是由于这两个特殊颜色的音素出现了组合情况时才发生了失爆现象,从视觉上直观的对用户进行提醒,以辅助用户理解失爆的概念及对应的发音特征。

[0069] 如图3所示,对于预设语句“stick to”,首先检测到二者为需要应用失爆技巧的单词对。对失爆部分的字母利用虚线从中间划开,虚线右边部分标为灰色,并且在音标部分直接标明失爆部分的发音特征。

[0070] S203:获取针对所述预设语句录入的音频数据;

[0071] S204:解析所述音频数据,根据单词时间边界截取需要应用失爆技巧的单词对所对应的声学模型输出片段;

[0072] 需要指出的是,本实施例可以通过语音识别的强制切分对齐确定每个音素的位置,并可根据每个单词的音节找到每个音节和单词的时间边界,进一步根据时间边界截取当前需要应用失爆技巧的单词对所对应的声学模型输出片段。

[0073] S205:将所述声学模型输出片段输入至预先基于所述发音词典构建的解码网络中,得到判断所述实际读音是否失爆的判断结果;

[0074] 其中,解码网络为预先构建的只有两条支路的语法网络,其中一条由两个未进行失爆的单词构成,如“stick”和“to”,另一条则由对应新添加的应用失爆技巧的单词构成,如“stick@@to”,再与状态转移、上下文、发音规则一同构建解码网络,该解码网络有且仅有表示失爆或未失爆的两种输出。

[0075] 将上述声学模型输出片段输入至解码网络中进行解码,若解码结果为两个独立单词,如“stick”和“to”,则表示该音频数据在对应单词对上未进行失爆;若解码结果为上述新添加至发音词典中的应用失爆技巧的单词,如“stick@@to”,则表示该音频数据在对应单词对上进行了失爆。

[0076] S206:基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。

[0077] 进一步地,本申请所提供的失爆检测方法的又一种具体实施方式中,在解码网络确定音频数据在对应单词对上进行了失爆之后,还可以进行进一步判断,以提高判断的准确性。该实施例对应的流程图如图4所示,该方法具体包括:

[0078] S301:利用预设语句的单词发音及预设失爆规则,筛选出需要应用失爆技巧的单词对;

[0079] S302:对所述单词对插入预设符号构建生成新的单词,并根据所述单词应用失爆技巧后的发音方式构建对应的音素序列,将所述音素序列添加至发音词典中;

[0080] S303:获取针对所述预设语句录入的音频数据;

[0081] S304:解析所述音频数据,根据单词时间边界截取需要应用失爆技巧的单词对所对应的声学模型输出片段;

[0082] S305:将所述声学模型输出片段输入至预先基于所述发音词典构建的解码网络中,得到判断所述实际读音是否失爆的判断结果;

[0083] S306:若判定所述预设语句中需要应用失爆技巧的单词对的实际读音失爆,则判断所述实际读音是否符合预设条件;

[0084] 其中,上述预设条件为所述实际读音的发音分数大于第一阈值,和/或所述实际读音的音素时长小于第二阈值。

[0085] S307:若判定所述实际读音符合所述预设条件,则生成所述实际读音正确失爆的反馈信息;

[0086] S308:若判定所述实际读音不符合所述预设条件,则生成所述实际读音未正确失爆的反馈信息。

[0087] 通过预设发音分数计算方法,计算上述单词对的发音分数,将得分低于预设第一阈值的语音数据判定为未正确失爆。其中,预设第一阈值可以为对大量标注数据进行统计分析获得。

[0088] 一种预设发音分数计算方法可以为:计算失爆单词对应的实际发音数据的后验概率评价指标;计算失爆单词对应的实际发音数据的时长评价指标;将后验概率评价指标以及时长评价指标输入至预先建立的评价模型中,得到实际发音数据的发音分数。其中,计算失爆单词对应的实际发音数据的后验概率评价指标的过程可以具体包括:计算每一个音素的似然度得分;将失爆单词对应的实际发音数据的似然度得分除以所有音素的似然度得分,得到失爆单词对应的实际发音数据的后验概率评价指标。具体地,预先建立的评价模型可以为线性回归模型。

[0089] 计算失爆单词对应的实际发音数据的时长评价指标的过程可以具体包括:预先在标准发音数据上统计每一个音素的音素时长,通过高斯模型建立表征音素时长与时长评价指标的对应关系;确定失爆单词对应的实际发音数据的音素时长;根据高斯模型确定当前

音素时长所对应的时长评价指标。

[0090] 失爆单词的音素时长可以具体为：失爆单词对中前一个单词的末尾音素以及后一个单词的首个音节的时间长度，该时间长度由上文中的音素边界计算获得。将该音素时长低于预设第二阈值的语音数据判为失爆，否则判为未失爆。该预设第二阈值同样通过对大量标注数据的分析统计获得。

[0091] 在上述任一实施例的基础上，本申请所提供的失爆检测方法中，上述通过显示界面的预设视觉元素标示所述反馈信息可以具体包括：通过显示界面的第二视觉元素标示出实际读音是否正确失爆。例如，在显示界面可以失爆的单词对之间通过特殊标识表示失爆，在进行了正确失爆的情况下特殊标识呈现预设第一颜色，在未进行正确失爆的情况下该特殊标识呈现预设第二颜色。预设第一颜色与预设第二颜色为不同的颜色。

[0092] 此外，还可以通过文字和/或语音的方式对需要应用失爆技巧的单词对的正确发音进行提示。例如，在显示界面显示失爆对应的音标，或者显示失爆时的发音技巧。进一步地，还可以播放用户录入的音频及示范音频帮助用户进行对比并模仿正确发音。

[0093] 参照图5，该实施例中，预设语句为“stick to”，通过界面上方左边的大圆圈指示用户是否进行了失爆，该圆圈颜色变绿则表示进行了正确失爆，该圆圈颜色变红则表示未进行正确失爆。

[0094] 本实施例中通过显示界面视觉元素辅助学习用户练习失爆技巧，在练习过程中加强对概念的理解，并能快速定位自己在练习时的问题所在。具体可以通过形象化的视觉元素图形，例如，利用虚线对字母进行切分，并采用将字母的一半设为灰色的方式来标识发音时口型到位但是气流不能释放的发音方式，辅助用户直观理解单词的失爆情况。

[0095] 下面对本发明实施例提供的一种失爆检测装置进行介绍，下文描述的失爆检测装置与上文描述的失爆检测方法可相互对应参照。

[0096] 图6为本发明实施例提供的失爆检测装置的结构框图，参照图6失爆检测装置可以包括：

[0097] 获取模块100，用于获取针对预设语句录入的音频数据；

[0098] 判断模块200，用于解析所述音频数据，确定所述预设语句中需要应用失爆技巧的单词对的实际读音，并判断所述实际读音是否失爆；

[0099] 生成模块300，用于基于判断结果生成反馈信息，并通过显示界面的预设视觉元素标示所述反馈信息。

[0100] 作为一种具体实施方式，本申请实施例还可以进一步包括：

[0101] 筛选模块，用于在获取针对预设语句录入的音频数据之前，利用所述预设语句的单词发音及预设失爆规则，筛选出需要应用失爆技巧的单词对；

[0102] 构建模块，用于对所述单词对插入预设符号构建生成新的单词，并根据所述单词应用失爆技巧后的发音方式构建对应的音素序列，将所述音素序列添加至发音词典中。

[0103] 作为一种具体实施方式，本申请实施例还可以进一步包括：

[0104] 标示模块，用于在筛选出需要应用失爆技巧的单词对之后，通过所述显示界面的第一视觉元素对所述单词对中失爆部分的音标和字母进行标示，以提示所述失爆部分的发音特征。

[0105] 作为一种具体实施方式，本申请实施例中判断模块200可以具体包括：

[0106] 截取单元,用于解析所述音频数据,根据单词时间边界截取需要应用失爆技巧的单词对所对应的声学模型输出片段;

[0107] 输入单元,用于将所述声学模型输出片段输入至预先基于所述发音词典构建的解码网络中,得到判断所述实际读音是否失爆的判断结果。

[0108] 作为一种具体实施方式,本申请实施例中生成模块300可以具体包括:

[0109] 判断单元,用于若判定所述预设语句中需要应用失爆技巧的单词对的实际读音失爆,则判断所述实际读音是否符合预设条件;所述预设条件为所述实际读音的发音分数大于第一阈值,和/或所述实际读音的音素时长小于第二阈值;

[0110] 第一反馈单元,用于若判定所述实际读音符合所述预设条件,则生成所述实际读音正确失爆的反馈信息;

[0111] 第二反馈单元,用于若判定所述实际读音不符合所述预设条件,则生成所述实际读音未正确失爆的反馈信息。

[0112] 作为一种具体实施方式,本申请实施例还可以进一步包括:

[0113] 提示模块,用于在通过显示界面的预设视觉元素标示所述反馈信息之后,通过文字和/或语音的方式对需要应用失爆技巧的单词对的正确发音进行提示。

[0114] 本实施例的失爆检测装置用于实现前述的失爆检测方法,因此失爆检测装置中的具体实施方式可见前文中的失爆检测方法的实施例部分,例如,获取模块100,判断模块200,生成模块300,分别用于实现上述失爆检测方法中的步骤S101,S102,S103,所以,其具体实施方式可以参照相应的各个部分实施例的描述,在此不再赘述。

[0115] 本申请能够自动对录入的音频数据进行分析,以判断其中需要应用失爆技巧的单词对的实际读音是否失爆,生成相应的反馈信息,能够辅助学习者理解失爆的概念,有效帮助学习者掌握失爆技巧;并且避免了人工教学和纠错的过程,突破了学习时间和空间的限制,显著节约了学习成本。

[0116] 此外,本申请还提供了一种失爆检测设备,应用于服务端1,如图7所示,所述设备包括:

[0117] 存储器11,用于存储计算机程序;

[0118] 处理器12,用于执行所述计算机程序时实现如下步骤:获取针对预设语句录入的音频数据;解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息。

[0119] 其中,存储器11至少包括一种类型的可读存储介质,所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器(例如,SD或DX存储器等)、磁性存储器、磁盘、光盘等。存储器11在一些实施例中可以是失爆检测设备的内部存储单元,例如硬盘。存储器11在另一些实施例中也可以是失爆检测设备的外部存储设备,例如插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。进一步地,存储器11还可以既包括失爆检测设备的内部存储单元也包括可以外部存储设备。存储器11不仅可以用于存储安装于失爆检测设备的应用软件及各类数据,例如失爆检测的程序01的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0120] 处理器12在一些实施例中可以是一中央处理器(Central Processing Unit,

CPU)、控制器、微控制器、微处理器或其他数据处理芯片,用于运行存储器11中存储的程序代码或处理数据,例如执行失爆检测的程序01等。

[0121] 可选地,所述处理器12用于执行所述计算机程序时具体实现如下步骤:在获取针对预设语句录入的音频数据之前,利用所述预设语句的单词发音及预设失爆规则,筛选出需要应用失爆技巧的单词对;对所述单词对插入预设符号构建生成新的单词,并根据所述单词应用失爆技巧后的发音方式构建对应的音素序列,将所述音素序列添加至发音词典中。

[0122] 可选地,所述处理器12用于执行所述计算机程序时具体实现如下步骤:在筛选出需要应用失爆技巧的单词对之后,通过所述显示界面的第一视觉元素对所述单词对中失爆部分的音标和字母进行标示,以提示所述失爆部分的发音特征。

[0123] 可选地,所述处理器12用于执行所述计算机程序时具体实现如下步骤:解析所述音频数据,根据单词时间边界截取需要应用失爆技巧的单词对所对应的声学模型输出片段;将所述声学模型输出片段输入至预先基于所述发音词典构建的解码网络中,得到判断所述实际读音是否失爆的判断结果。

[0124] 可选地,所述处理器12用于执行所述计算机程序时具体实现如下步骤:若判定所述预设语句中需要应用失爆技巧的单词对的实际读音失爆,则判断所述实际读音是否符合预设条件;所述预设条件为所述实际读音的发音分数大于第一阈值,和/或所述实际读音的音素时长小于第二阈值;若判定所述实际读音符合所述预设条件,则生成所述实际读音正确失爆的反馈信息;若判定所述实际读音不符合所述预设条件,则生成所述实际读音未正确失爆的反馈信息。

[0125] 可选地,所述处理器12用于执行所述计算机程序时具体实现如下步骤:在通过显示界面的预设视觉元素标示所述反馈信息之后,通过文字和/或语音的方式对需要应用失爆技巧的单词对的正确发音进行提示。

[0126] 可以理解的是,本申请实施例中服务端可以包括但不限于:单个网络服务器、多个网络服务器组成的服务器组或基于云计算由大量计算机或网络服务器构成的云。

[0127] 此外,本申请还提供了一种失爆检测设备,应用于客户端2,如图8所示,所述设备包括:

[0128] 音频采集装置21,用于录入针对预设语句的音频数据;

[0129] 通信装置22,用于将所述音频数据发送至服务端,以便所述服务端解析所述音频数据,确定所述预设语句中需要应用失爆技巧的单词对的实际读音,并判断所述实际读音是否失爆;基于判断结果生成反馈信息,并通过显示界面的预设视觉元素标示所述反馈信息;

[0130] 显示装置23,用于将标示所述反馈信息的预设视觉元素在显示界面上进行显示。

[0131] 可选地,本申请实施例所提供的失爆检测设备中所述显示装置23还可以用于:在筛选出需要应用失爆技巧的单词对之后,通过所述显示界面的第一视觉元素对所述单词对中失爆部分的音标和字母进行标示,以提示所述失爆部分的发音特征。

[0132] 可以理解的是,本申请实施例中客户端可以包括但不限于:智能手机、平板电脑、MP4、MP3、PC、PDA、可穿戴设备和头戴显示设备等。

[0133] 进一步地,本申请还提供了一种失爆检测系统,如图9所示,该系统包括上述任一

种服务端1以及上述任一种客户端2。用户可以通过客户端进行发音学习,客户端可以在显示界面上显示用户待学习的内容,并且还可以通过扬声器等音频播放装置向用户输出语音形式的音频内容,在用户进行语音的发音学习时,客户端可以通过音频采集装置采集用户发音时的音频数据,并将音频数据发送至服务端,由服务端进行失爆检测的过程。在服务端对音频数据进行分析并得到反馈信息之后,将该反馈信息发送至客户端。通过客户端的显示装置对反馈信息进行显示,向用户提供视觉辅助信息。

[0134] 此外,本申请还提供了一种计算机可读存储介质,所述计算机可读存储介质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前述实施例公开的任一种失爆检测方法的步骤。

[0135] 本申请所提供的失爆检测设备、失爆检测系统、计算机可读存储介质与前述失爆检测方法相对应。所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0136] 综上,本申请能够自动对录入的音频数据进行分析,以判断其中需要应用失爆技巧的单词对的实际读音是否失爆,生成相应的反馈信息,能够帮助学习者理解失爆的概念,有效帮助学习者掌握失爆技巧;并且避免了人工教学和纠错的过程,突破了学习时间和空间的限制,显著节约了学习成本。

[0137] 说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的系统而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以对本申请进行若干改进和修饰,这些改进和修饰也落入本申请权利要求的保护范围内。

[0138] 还需要说明的是,在本说明书中,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”或者任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,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“包括一个……”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。

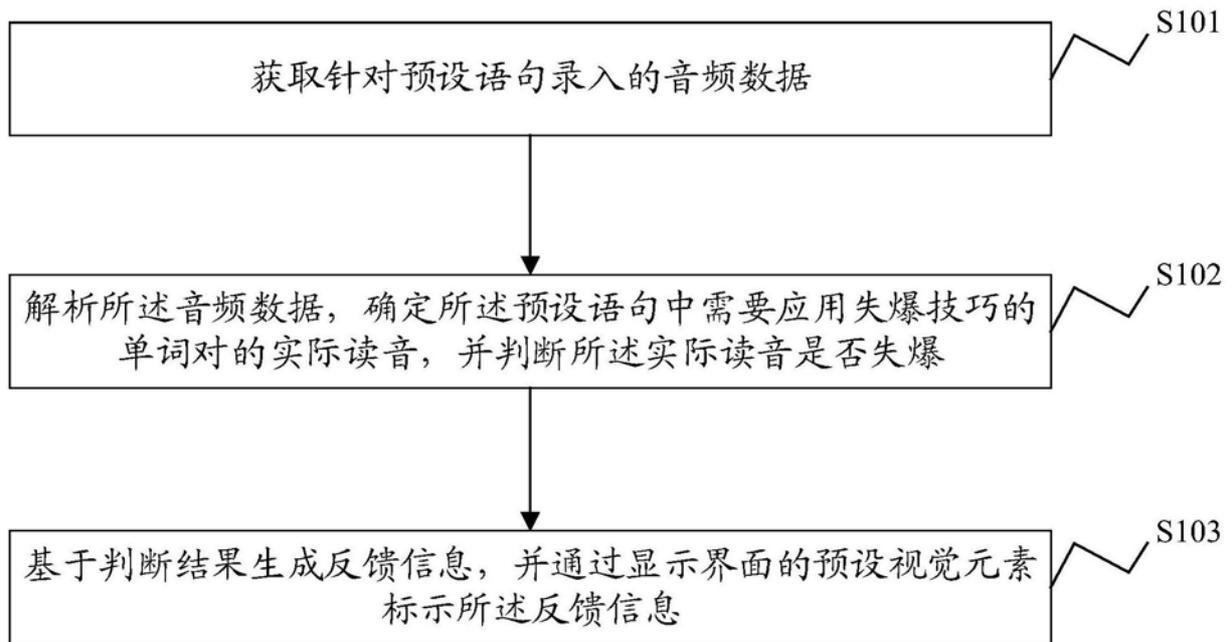


图1

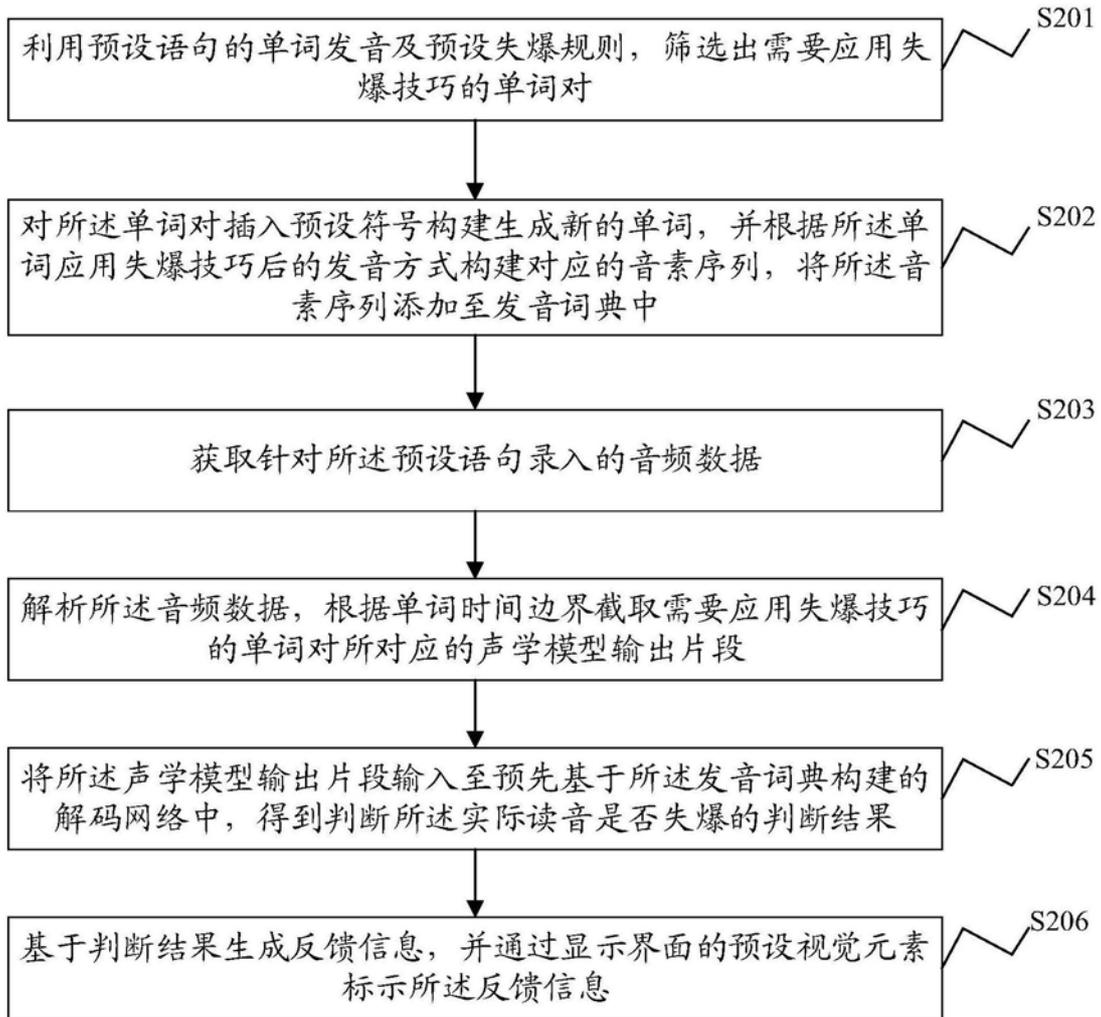


图2



图3

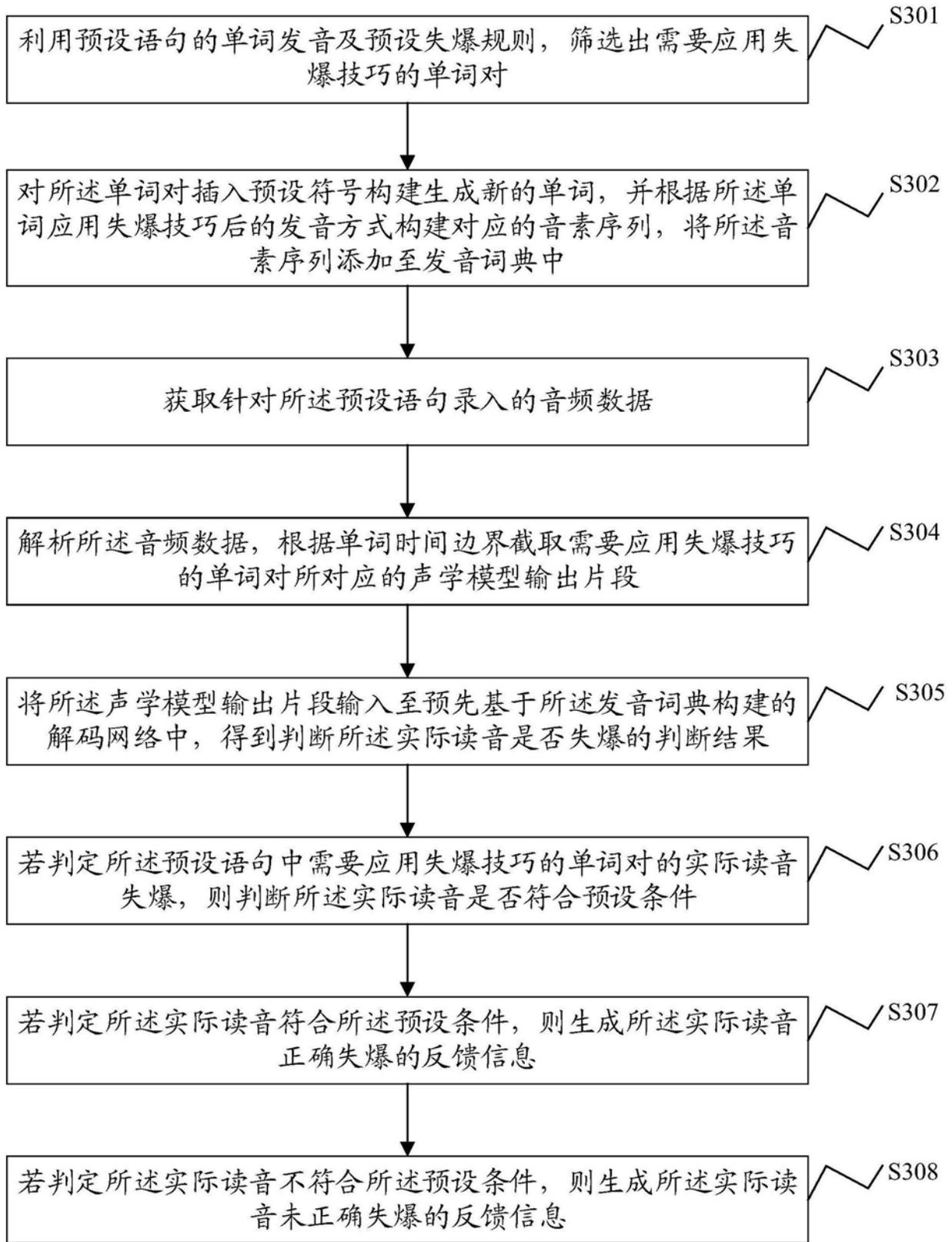


图4



图5

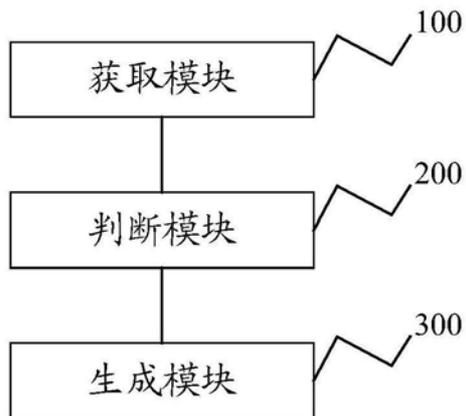


图6

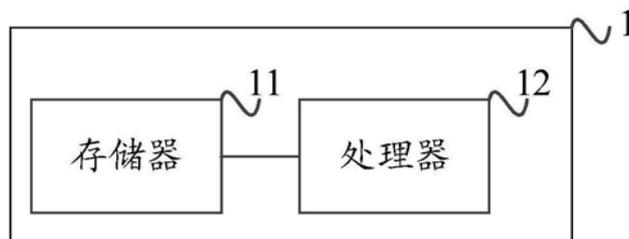


图7

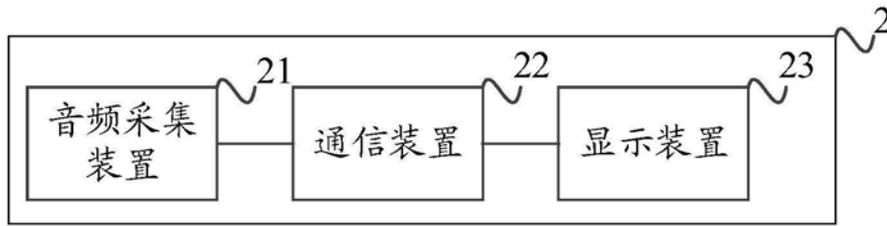


图8

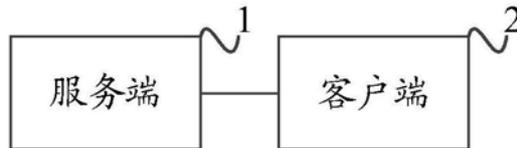


图9