



(12) 发明专利申请

(10) 申请公布号 CN 112071287 A

(43) 申请公布日 2020. 12. 11

(21) 申请号 202010945536.6

(22) 申请日 2020.09.10

(71) 申请人 北京有竹居网络技术有限公司
地址 101200 北京市平谷区林荫北街13号
信息大厦802室

(72) 发明人 顾宇

(74) 专利代理机构 北京卫智畅科专利代理事务
所(普通合伙) 11557
代理人 陈佳

(51) Int.Cl.
G10G 3/04 (2006.01)

权利要求书2页 说明书10页 附图4页

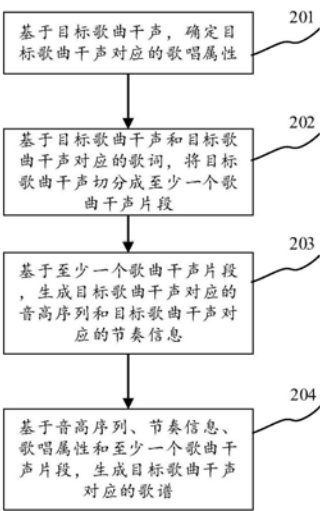
(54) 发明名称

用于生成歌谱的方法、装置、电子设备和计算机可读介质

(57) 摘要

本公开的实施例公开了用于生成歌谱的方法、装置、电子设备和计算机可读介质。该方法的一具体实施方式包括：基于目标歌曲干声，确定上述目标歌曲干声对应的歌唱属性；基于上述目标歌曲干声和上述目标歌曲干声对应的歌词，将上述目标歌曲干声切分成至少一个歌曲干声片段；基于上述至少一个歌曲干声片段，生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息；基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段，生成上述目标歌曲干声对应的歌谱。综合以上信息，可以快速的为上述目标歌曲干声生成歌谱。而且，生成歌谱的所需的成本将会大大降低。

200



1. 一种用于生成歌谱的方法,包括:

基于目标歌曲干声,确定所述目标歌曲干声对应的歌唱属性;

基于所述目标歌曲干声和所述目标歌曲干声对应的歌词,将所述目标歌曲干声切分成至少一个歌曲干声片段;

基于所述至少一个歌曲干声片段,生成所述目标歌曲干声对应的音高序列和所述目标歌曲干声对应的节奏信息;

基于所述音高序列、所述节奏信息、所述歌唱属性和所述至少一个歌曲干声片段,生成所述目标歌曲干声对应的歌谱。

2. 根据权利要求1所述的方法,其中,所述歌唱属性用于表征所述目标歌曲干声的发音特征。

3. 根据权利要求1所述的方法,其中,基于所述目标歌曲干声和所述目标歌曲干声对应的歌词,将所述目标歌曲干声切分成至少一个歌曲干声片段,还包括:

所述歌词的每个字对应至少一个歌曲干声片段,所述每个歌曲干声片段标有时间戳。

4. 根据权利要求1所述的方法,其中,在所述基于目标歌曲干声,确定所述目标歌曲干声对应的歌唱属性之前,还包括:

基于目标歌曲,生成所述目标歌曲对应的目标歌曲干声。

5. 根据权利要求4所述的方法,其中,所述基于目标歌曲,生成所述目标歌曲对应的目标歌曲干声,包括:

响应于所述目标歌曲中包括背景音乐,将所述目标歌曲输入预先训练的干声、背景音乐分离模型,生成所述目标歌曲干声。

6. 根据权利要求3所述的方法,其中,所述基于所述目标歌曲干声和所述目标歌曲干声对应的歌词,将所述干声切分成至少一个歌曲干声片段,包括:

将所述目标歌曲干声和所述歌词输入预先训练的歌曲干声切分模型,生成所述至少一个歌曲干声片段,其中,每个歌曲干声片段标识有时间戳。

7. 根据权利要求6所述的方法,其中,所述歌曲干声切分模型是以预先训练的语音识别模型为基础,使用预设的歌唱数据对所述语音识别模型训练调整得到的。

8. 根据权利要求3所述的方法,其中,所述基于所述至少一个歌曲干声片段,生成所述目标歌曲干声对应的音高序列和所述目标歌曲干声对应的节奏信息,包括:

对于所述歌词的每个字,将所述每个字对应的至少一个歌曲干声片段输入预先训练的音高检测模型,得到所述音高序列;

基于所述至少一个歌曲干声片段以及至少一个歌曲干声片段对应的时间戳,生成所述歌曲干声对应的节奏信息。

9. 根据权利要求1所述的方法,其中,所述目标歌曲干声对应的歌唱属性包括以下至少一项:颤音,换气。

10. 根据权利要求1所述的方法,其中,所述基于所述音高序列、所述节奏信息、所述歌唱属性和所述至少一个歌曲干声片段,生成所述目标歌曲干声对应的歌谱,包括:

将所述音高序列、所述节奏信息、所述歌唱属性和所述至少一个歌曲干声片段导入预设歌谱生成器中,生成所述目标歌曲干声对应的歌谱。

11. 一种用于生成歌谱的装置,包括:

确定单元,被配置成基于目标歌曲干声,确定所述目标歌曲干声对应的歌唱属性;

切分单元,被配置成基于所述目标歌曲干声和所述目标歌曲干声对应的歌词,将所述目标歌曲干声切分成至少一个歌曲干声片段;

第一生成单元,被配置成基于所述至少一个歌曲干声片段,生成所述目标歌曲干声对应的音高序列和所述目标歌曲干声对应的节奏信息;

第二生成单元,被配置成基于所述音高序列、所述节奏信息、所述歌唱属性和所述至少一个歌曲干声片段,生成所述目标歌曲干声对应的歌谱。

12.一种电子设备,包括:

一个或多个处理器;

存储装置,其上存储有一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-10中任一所述的方法。

13.一种计算机可读介质,其上存储有计算机程序,其中,所述程序被处理器执行时实现如权利要求1-10中任一所述的方法。

用于生成歌谱的方法、装置、电子设备和计算机可读介质

技术领域

[0001] 本公开的实施例涉及计算机技术领域，具体涉及用于生成歌谱的方法、装置、电子设备和计算机可读介质。

背景技术

[0002] 对歌曲而言，要想完美的演绎出一首歌曲。歌曲对应的歌谱至关重要。目前，相关合成歌谱的方法可以通过人工标注的方式生成歌谱。首先，对歌曲演奏者而言，往往对演奏者的演唱能力要求较高。而且，还需要专业的标注人员对演奏者所演奏的歌曲进行生成歌谱。这就需要标注人员不仅了解语言学，还需要拥有较高的音准。因此，生成歌谱的成本较高，而且，人工生成歌谱的速度较慢。

发明内容

[0003] 本公开的内容部分用于以简要的形式介绍构思，这些构思将在后面的具体实施方式部分被详细描述。本公开的内容部分并不旨在标识要求保护的技术方案的关键特征或必要特征，也不旨在用于限制所要求的保护的技术方案的范围。

[0004] 本公开的一些实施例提出了用于生成歌谱的方法、装置、设备和计算机可读介质，来解决以上背景技术部分提到的技术问题。

[0005] 第一方面，本公开的一些实施例提供了一种用于生成歌谱的方法，该方法包括：基于目标歌曲干声，确定上述目标歌曲干声对应的歌唱属性；基于上述目标歌曲干声和上述目标歌曲干声对应的歌词，将上述目标歌曲干声切分成至少一个歌曲干声片段，；基于上述至少一个歌曲干声片段，生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息；基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段，生成上述目标歌曲干声对应的歌谱。

[0006] 第二方面，本公开的一些实施例提供了一种用于生成歌谱的装置，装置包括：确定单元，被配置成基于目标歌曲干声，确定上述目标歌曲干声对应的歌唱属性；切分单元，被配置成基于上述目标歌曲干声和上述目标歌曲干声对应的歌词，将上述目标歌曲干声切分成至少一个歌曲干声片段；第一生成单元，被配置成基于上述至少一个歌曲干声片段，生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息；第二生成单元，被配置成基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段，生成上述目标歌曲干声对应的歌谱。

[0007] 第三方面，本公开的一些实施例提供了一种电子设备，包括：一个或多个处理器；存储装置，其上存储有一个或多个程序，当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如第一方面中任一的方法。

[0008] 第四方面，本公开的一些实施例提供了一种计算机可读介质，其上存储有计算机程序，其中，程序被处理器执行时实现如第一方面中任一的方法。

[0009] 本公开的上述各个实施例中的一个实施例具有如下有益效果：基于目标歌曲干

声,确定上述目标歌曲干声对应的歌唱属性,上述歌唱属性用于表征上述目标歌曲干声的发音特征。例如,颤音、换气等。基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段,其中,上述歌词的每个字对应至少一个歌曲干声片段,上述每个歌曲干声片段标有时间戳;基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息;这里,可以使用音高检测软件,快速且准确的确定出歌曲干声对应的音高序列。节奏信息是指歌曲演奏速度的快慢。作为示例,可以以每分钟60节拍的速度演唱歌曲。最终,基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。例如,五线谱。该实施方式可以自动生成目标歌曲干声生成歌谱所需要的歌唱属性、音高序列、节奏信息和至少一个歌曲干声片段。综合以上信息,可以快速的为上述目标歌曲干声生成歌谱。而且,生成歌谱的所需的成本将会大大降低。

附图说明

[0010] 结合附图并参考以下具体实施方式,本公开各实施例的上述和其他特征、优点及方面将变得更加明显。贯穿附图中,相同或相似的附图标记表示相同或相似的元素。应当理解附图是示意性的,原件和元素不一定按照比例绘制。

[0011] 图1是根据本公开的一些实施例的一种用于生成歌谱的方法的一个应用场景的示意图;

[0012] 图2是根据本公开的一种用于生成歌谱的方法的一些实施例的流程图;

[0013] 图3是根据本公开的一种用于生成歌谱的方法的一些实施例的流程图;

[0014] 图4是根据本公开的一种用于生成歌谱的装置的一些实施例的结构示意图;

[0015] 图5是适于用来实现本公开的一些实施例的电子设备的结构示意图。

具体实施方式

[0016] 下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例,然而应当理解的是,本公开可以通过各种形式来实现,而且不应该被解释为限于这里阐述的实施例。相反,提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是,本公开的附图及实施例仅用于示例性作用,并非用于限制本公开的保护范围。

[0017] 另外还需要说明的是,为了便于描述,附图中仅示出了与本公开相关的部分。在不冲突的情况下,本公开中的实施例及实施例中的特征可以相互组合。

[0018] 需要注意,本公开中提及的“第一”、“第二”等概念仅用于对不同的装置、模块或单元进行区分,并非用于限定这些装置、模块或单元所执行的功能的顺序或者相互依存关系。

[0019] 需要注意,本公开中提及的“一个”、“多个”的修饰是示意性而非限制性的,本领域技术人员应当理解,除非在上下文另有明确指出,否则应该理解为“一个或多个”。

[0020] 本公开实施方式中的多个装置之间所交互的消息或者信息的名称仅用于说明性的目的,而并不是用于对这些消息或信息的范围进行限制。

[0021] 下面将参考附图并结合实施例来详细说明本公开。

[0022] 图1是根据本公开一些实施例的用于生成歌谱的方法的一个应用场景的示意图。

[0023] 如图1所示,电子设备101首先通过目标歌曲干声102确定上述目标歌曲干声102的

歌唱属性103,然后上述执行主体通过目标歌曲干声102和目标歌曲干声对应的歌词104,将上述目标歌曲干声102切分成至少一个目标歌曲干声片段105,基于上述至少一个目标歌曲干声片段105,生成目标歌曲干声102对应的音高序列106和目标歌曲干声102对应的节奏信息107。最后通过目标歌曲干声102的歌唱属性103,至少一个目标歌曲干声片段105,目标歌曲干声102对应的音高序列106和目标歌曲干声102对应的节奏信息107生成目标歌曲干声对应的歌谱108。

[0024] 可以理解的是,用于生成歌谱的方法可以是由电子设备101来执行,或者也可以是由服务器来执行,上述方法的执行主体还可以包括上述电子设备101与上述服务器通过网络相集成所构成的设备,或者还可以是各种软件程序来执行。其中,电子设备101可以是具有信息处理能力的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、膝上型便携计算机和台式计算机等等。此外,执行主体也可以体现为服务器、软件等。当执行主体为软件时,可以安装在上述所列举的电子设备中。其可以实现成例如用来提供分布式服务的多个软件或软件模块,也可以实现成单个软件或软件模块。在此不做具体限定。

[0025] 应该理解,图1中的终端设备数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备。

[0026] 继续参考图2,示出了根据本公开的用于生成歌谱的方法的一些实施例的流程200。该用于生成歌谱的方法,包括以下步骤:

[0027] 步骤201,基于目标歌曲干声,确定目标歌曲干声对应的歌唱属性。

[0028] 在一些实施例中,根据目标歌曲干声,用于生成歌谱的方法的执行主体(例如图1所示的电子设备)可以通过各种方法确定上述目标歌曲干声对应的歌唱属性。这里,干声是指无背景音乐的纯人声。目标歌曲干声可以是任意干声。

[0029] 在一些实施例的一些可选的实现方式中,在上述基于目标歌曲干声,确定上述目标歌曲干声对应的歌唱属性之前,还包括:根据目标歌曲,可以通过各种方式生成目标歌曲对应的目标歌曲干声。这里,目标歌曲可以是任意歌曲。实践中,可以从网络上下下载的歌曲,还可以是用户自己演唱的歌曲。其中,上述目标歌曲可以包括干声和背景音乐。此外,目标歌曲还可以只存在干声。当目标歌曲只存在干声时,直接可以得到所需要的干声。当目标歌曲包括干声和背景音乐时,可以使用相关的音频分割软件来对干声和背景音乐分割。从而得到上述目标歌曲干声。

[0030] 可选的,响应于目标歌曲中包括背景音乐,还可以将目标歌曲输入预先训练的干声、背景音乐分离模型,生成上述干声。这里,上述干声、背景音乐分离模型是以神经网络为基础结构训练得到的模型。

[0031] 在一些实施例的一些可选的实现方式中,上述歌唱属性用于表征上述目标歌曲干声的发音特征。例如,颤音、换气等。

[0032] 作为示例,可以使用一些相关的歌唱属性检测软件(例如,检测颤音的软件;检测换气的软件等)来检测上述目标歌曲干声对应的歌唱属性。

[0033] 步骤202,基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段。

[0034] 在一些实施例中,基于目标歌曲干声和目标歌曲干声对应的歌词,上述执行主体可以将目标歌曲干声切分成至少一个歌曲干声片段。其中,上述歌词的每个字对应至少一

个歌曲干声片段。上述歌曲干声片段可以是表征某一个音素的发音片段。实践中,可以是歌曲干声片段的格式可以是wav,上述wav是使用用于存储波形数据的标准数字音频文件格式。例如,表示“a”的发音片段。在汉语中,音素分为元音与辅音两大类。

[0035] 在一些实施例的一些可选的实现方式中,基于目标歌曲干声和目标歌曲干声对应的歌词,将目标歌曲干声切分成至少一个歌曲干声片段,还包括:歌词的每个字对应至少一个歌曲干声片段,每个歌曲干声片段标有时间戳。作为示例,可以使用音频切分软件将上述目标歌曲干声切分为至少一个歌曲干声片段。需要说明的是,上述切分可以是按照音素切分。也可以按照音节来切分。例如,现有歌词“我爱祖国”对应的干声,对于字“我”而言,按照音素切分,可以切分为两个歌曲干声片段。上述两个歌曲干声片段分别表示“w”和“o”的发音。类似的,对于其余三个字,均可以得到对应切分的歌曲干声片段。步骤203,基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息。

[0036] 在一些实施例中,基于至少一个歌曲干声片段,上述执行主体可以使用各种方法来生成上述干声对应的音高序列和上述目标歌曲干声对应的节奏信息。这里,节奏信息是指歌曲演奏速度的快慢。上述节奏信息可以用节拍来表示。作为示例,可以以每分钟60节拍的速度演唱歌曲。

[0037] 作为示例,上述执行主体可以使用音高检测软件来监测上述至少一个歌曲干声片段的音高,从而生成上述干声对应的音高序列。实践中,对于“我爱祖国”这段干声而言,经步骤202会得到“我爱祖国”对应的歌曲干声片段,即“w、o,”、“a、i”、“z、u”、“g、u、o”。然后,将每个唱字对应的歌曲干声片段使用音高检测软件来检测这个歌唱字对应的音高。类似的,可以得到其余唱字对应的音高。组合这些唱字对应的音高,即可得到上述“我爱祖国”这段干声对应的音高序列。上述执行主体可以使用相关的节奏检测软件来对上述干声进行检测,进而生成上述干声所对应的节奏信息。作为示例,假设有10个歌曲干声片段,依次为歌曲干声片段1、歌曲干声片段2、歌曲干声片段3、歌曲干声片段4、歌曲干声片段5、歌曲干声片段6、歌曲干声片段7、歌曲干声片段8、歌曲干声片段9和歌曲干声片段10。对这10个歌曲干声片段而言,每个歌曲干声片段对应的时间戳可以是:1秒,2秒,3秒,4秒,5秒,6秒,7秒,8秒,9秒和10秒。根据上述每个片段的时间戳,可以确定出上述干声对应的节拍是每秒一拍,即每分钟60节拍。

[0038] 在一些实施例的一些可选的实现方式中,对于上述歌词的每个字,将上述每个字对应的至少一个歌曲干声片段输入预先训练的音高检测模型,得到上述音高序列。基于上述至少一个歌曲干声片段以及至少一个歌曲干声片段对应的时间戳,生成上述歌曲干声对应的节奏信息。

[0039] 对于上述歌词的每个字,上述执行主体可以将上述每个字对应的至少一个歌曲干声片段输入预先训练的音高检测模型,进而得到音高序列。这里,上述音高检测模型可以是基于数据信号处理器的音高检测器。

[0040] 步骤204,基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。

[0041] 在一些实施例中,上述方法的执行主体可以将上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段输入至预设的歌谱生成器中生成目标歌曲干声对

应的歌谱。实践中,上述歌谱可以是五线谱。这里,最终生成的歌谱可以以文件的形式存储。例如以xml文件形式存储,上述xml是一种文件格式。

[0042] 作为示例,可以将上述信息输入相关的生成歌谱软件,从而生成上述歌谱。此外,经以上步骤得到上述信息后,还可以通过人工标注的形式生成歌谱。需要说明的是,传统的生成歌谱,往往将大量的时间花费在上述信息的收集和确定过程中,因此速度较慢,生成歌谱难度较大。而这里,是在生成歌谱信息已经收集好的基础上,根据一些预先设定的规则生成歌谱,生成歌谱的速度较快。

[0043] 本公开的上述各个实施例中的一个实施例具有如下有益效果:

[0044] 基于目标歌曲干声,确定上述目标歌曲干声对应的歌唱属性,上述歌唱属性用于表征上述目标歌曲干声的发音特征。例如,颤音、换气等。基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段,其中,上述歌词的每个字对应至少一个歌曲干声片段,上述每个歌曲干声片段标有时间戳;基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息;这里,可以使用音高检测软件,快速且准确的确定出歌曲干声对应的音高序列。节奏信息是指歌曲演奏速度的快慢。作为示例,可以以每分钟60节拍的速度演唱歌曲。最终,基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。例如,五线谱。该实施方式可以自动生成目标歌曲干声生成歌谱所需要的歌唱属性、音高序列、节奏信息和至少一个歌曲干声片段。综合以上信息,可以快速的为上述目标歌曲干声生成歌谱。而且,生成歌谱的所需的成本将会大大降低。

[0045] 进一步参考图3,其示出了用于生成歌谱的方法的另一些实施例的流程300。该用于生成歌谱的方法的流程300,包括以下步骤:

[0046] 步骤301,基于目标歌曲干声,确定目标歌曲干声对应的歌唱属性。

[0047] 在一些实施例中,步骤301的具体实现及所带来的技术效果可以参考图2对应的那些实施例中的步骤201,在此不再赘述。

[0048] 步骤302,将上述目标歌曲干声和上述目标歌曲干声对应的歌词输入预先训练的歌曲干声切分模型,生成至少一个歌曲干声片段。

[0049] 在一些实施例中,上述执行主体可以将上述目标歌曲干声和上述目标歌曲干声对应的歌词输入预先训练的歌曲干声切分模型,生成至少一个歌曲干声片段。其中,每个歌曲干声片段标识有时间戳。

[0050] 这里,上述歌曲干声切分模型是以一些端到端模型为基础,改进得到的模型。上述端到端模型可以是以下之一:LAS(Listen Attend and Spell,听注意和拼写),CTC(Connectionist Temporal Classification,连接主义时间分类)和RNNT(Recurrent Neural Network Transducer,循环神经网络转化器)。

[0051] 在一些实施例的一些可选的实现方式中,上述歌曲干声切分模型是以预先训练的语音识别模型为基础,使用预设的歌唱数据对上述语音识别模型训练调整得到的。

[0052] 步骤303,基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息。

[0053] 步骤304,基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。在一些实施例中,步骤303至步骤304的具体

实现及所带来的技术效果可以参考图2对应的那些实施例中的步骤203至步骤204,在此不再赘述。

[0054] 从图3中可以看出,与图2对应的一些实施例的描述相比,图3对应的一些实施例中的用于生成歌谱的装置300体现了使用歌曲干声切分模型来对干声进行切分。因为歌曲干声切分模型是经过大量歌唱数据训练得到的。同时,歌曲干声切分模型是在预先训练好的语音识别模型的基础上,通过歌唱数据对语音识别模型进行微调得到的。所以经模型切分出的歌曲干声片段是相对准确的,每个字可以准确的匹配到对应的歌曲干声片段。为接下来准确生成歌谱提供了基础。

[0055] 进一步参考图4,作为对上述各图所示方法的实现,本公开提供了一种用于生成歌谱的装置的一些实施例,这些装置实施例与图2所示的那些方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0056] 如图4所示,一些实施例的用于生成歌谱的装置400包括:确定单元401,切分单元402,第一生成单元403、第二生成单元404。确定单元401被配置成基于目标歌曲干声,确定上述目标歌曲干声对应的歌唱属性;切分单元402被配置成基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段;第一生成单元403被配置成基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息;第二生成单元404被配置成基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。

[0057] 在一些实施例的一些可选的实现方式中,其中,歌唱属性用于表征目标歌曲干声的发音特征。

[0058] 在一些实施例的一些可选的实现方式中,上述切分单元402还可以包括对应子单元,其中,对应单元被配置成歌词的每个字对应至少一个歌曲干声片段,每个歌曲干声片段标有时间戳。

[0059] 在一些实施例的一些可选的实现方式中,上述装置400还包括第三生成单元,其中,第三生成单元被配置成基于目标歌曲,生成上述目标歌曲对应的目标歌曲干声。

[0060] 在一些实施例的一些可选的实现方式中,上述第三生成单元可以进一步被配置成响应于上述目标歌曲中包括背景音乐,将上述目标歌曲输入预先训练的干声、背景音乐分离模型,生成上述目标歌曲干声。

[0061] 在一些实施例的一些可选的实现方式中,切分单元402可以进一步被配置成:将上述目标歌曲干声和上述目标歌曲干声对应的歌词输入预先训练的歌曲干声切分模型,生成上述至少一个歌曲干声片段,其中,上述每个歌曲干声片段标识有时间戳。

[0062] 在一些实施例的一些可选的实现方式中,上述歌曲干声切分模型是以预先训练的语音识别模型为基础,使用预设的歌唱数据对上述语音识别模型训练调整得到的。

[0063] 在一些实施例的一些可选的实现方式中,第一生成单元403可以进一步被配置成对于上述歌词的每个字,将上述每个字对应的至少一个歌曲干声片段输入预先训练的音高检测模型,得到上述音高序列。基于上述至少一个歌曲干声片段以及至少一个歌曲干声片段对应的时间戳,生成上述歌曲干声对应的节奏信息。

[0064] 在一些实施例的一些可选的实现方式中,上述目标歌曲干声对应的歌唱属性包括

以下至少一项:颤音,换气。

[0065] 在一些实施例的一些可选的实现方式中,上述装置400的第二生成单元404进一步被配置成将上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段导入预设歌谱生成器中,生成上述目标歌曲干声对应的歌谱。

[0066] 可以理解的是,该装置400中记载的诸单元与参考图2描述的方法中的各个步骤相对应。由此,上文针对方法描述的操作、特征以及产生的有益效果同样适用于装置400及其中包含的单元,在此不再赘述。

[0067] 下面参考图5,其示出了适于用来实现本公开的一些实施例的电子设备(例如图1中的电子设备)500的结构示意图。图5示出的电子设备仅仅是一个示例,不应对本公开的实施例的功能和使用范围带来任何限制。

[0068] 如图5所示,电子设备500可以包括处理装置(例如中央处理器、图形处理器等)501,其可以根据存储在只读存储器(ROM)502中的程序或者从存储装置508加载到随机访问存储器(RAM)503中的程序而执行各种适当的动作和处理。在RAM 503中,还存储有电子设备500操作所需的各种程序和数据。处理装置501、ROM 502以及RAM503通过总线504彼此相连。输入/输出(I/O)接口505也连接至总线504。

[0069] 通常,以下装置可以连接至I/O接口505:包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置506;包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置507;包括例如磁带、硬盘等的存储装置508;以及通信装置509。通信装置509可以允许电子设备500与其他设备进行无线或有线通信以交换数据。虽然图5示出了具有各种装置的电子设备500,但是应理解的是,并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。图5中示出的每个方框可以代表一个装置,也可以根据需要代表多个装置。

[0070] 特别地,根据本公开的一些实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本公开的一些实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的一些实施例中,该计算机程序可以通过通信装置509从网络上被下载和安装,或者从存储装置508被安装,或者从ROM 502被安装。在该计算机程序被处理装置501执行时,执行本公开的一些实施例的方法中限定的上述功能。

[0071] 需要说明的是,本公开的一些实施例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——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。在本公开的一些实施例中,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。而在本公开的一些实施例中,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。

[0072] 在一些实施方式中,客户端、服务器可以利用诸如HTTP(HyperText Transfer Protocol,超文本传输协议)之类的任何当前已知或未来研发的网络协议进行通信,并且可以与任意形式或介质的数字数据通信(例如,通信网络)互连。通信网络的示例包括局域网(“LAN”),广域网(“WAN”),网际网(例如,互联网)以及端对端网络(例如,ad hoc端对端网络),以及任何当前已知或未来研发的网络。

[0073] 上述计算机可读介质可以是上述装置中所包含的;也可以是单独存在,而未装配入该电子设备中。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:

[0074] 基于目标歌曲干声,确定上述目标歌曲干声对应的歌唱属性;基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段;基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息;基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。

[0075] 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公开的一些实施例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)。

[0076] 附图中的流程图和框图,图示了按照本公开各种实施例的系统、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,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。

[0077] 描述于本公开的一些实施例中的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中,例如,可以描述为:一种处理器包括确定单元、切分单元、第一生成单元和第二生成单元。其中,这些单元的名称在某种情况下并不构成对该单元本身的限定,例如,第一生成单元还可以被描述为“确定上述目标歌曲干声对应的歌唱属性的单元”。

[0078] 本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如,非限制性地,可以使用的示范类型的硬件逻辑部件包括:现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等等。

[0079] 根据本公开的一个或多个实施例,提供了一种用于生成歌谱的方法方法,包括:基于目标歌曲干声,确定上述目标歌曲干声对应的歌唱属性;基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段;基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息;基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。

[0080] 根据本公开的一个或多个实施例,其中,歌唱属性用于表征目标歌曲干声的发音特征。

[0081] 根据本公开的一个或多个实施例,基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段,还包括:上述歌词的每个字对应至少一个歌曲干声片段,上述每个歌曲干声片段标有时间戳。根据本公开的一个或多个实施例,在上述基于目标歌曲干声,确定上述目标歌曲干声对应的歌唱属性之前,还包括:基于目标歌曲,生成上述目标歌曲对应的目标歌曲干声。

[0082] 根据本公开的一个或多个实施例,上述基于目标歌曲,生成上述目标歌曲对应的目标歌曲干声,包括:响应于上述目标歌曲中包括背景音乐,将上述目标歌曲输入预先训练的干声、背景音乐分离模型,生成上述目标歌曲干声。

[0083] 根据本公开的一个或多个实施例,上述基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述干声切分成至少一个歌曲干声片段,包括:将上述目标歌曲干声和上述目标歌曲干声对应的歌词输入预先训练的歌曲干声切分模型,生成上述至少一个歌曲干声片段,其中,上述每个歌曲干声片段标识有时间戳。

[0084] 根据本公开的一个或多个实施例,上述歌曲干声切分模型是以预先训练的语音识别模型为基础,使用预设的歌唱数据对上述语音识别模型训练调整得到的。

[0085] 根据本公开的一个或多个实施例,上述基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息,包括:对于上述歌词的每个字,将上述每个字对应的至少一个歌曲干声片段输入预先训练的音高检测模型,得到上述音高序列。基于上述至少一个歌曲干声片段以及至少一个歌曲干声片段对应的时间戳,生成上述歌曲干声对应的节奏信息。

[0086] 根据本公开的一个或多个实施例,上述目标歌曲干声对应的歌唱属性包括以下至少一项:颤音,换气。

[0087] 根据本公开的一个或多个实施例,上述基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱,包括:将上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段导入预设歌谱生成器中,生成上述目标歌曲干声对应的歌谱。

[0088] 根据本公开的一个或多个实施例,提供了一种用于生成歌谱的方法装置,包括:确定单元,切分单元,第一生成单元、和第二生成单元。确定单元,被配置成基于目标歌曲干

声,确定上述目标歌曲干声对应的歌唱属性;切分单元,被配置成基于上述目标歌曲干声和上述目标歌曲干声对应的歌词,将上述目标歌曲干声切分成至少一个歌曲干声片段;第一生成单元,被配置成基于上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的音高序列和上述目标歌曲干声对应的节奏信息;第二生成单元,被配置成基于上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段,生成上述目标歌曲干声对应的歌谱。

[0089] 根据本公开的一个或多个实施例,其中,歌唱属性用于表征目标歌曲干声的发音特征。

[0090] 根据本公开的一个或多个实施例,上述切分单元进一步被配置成:上述歌词的每个字对应至少一个歌曲干声片段,上述每个歌曲干声片段标有时间戳。

[0091] 根据本公开的一个或多个实施例,上述装置还包括第三生成单元,其中,第三生成单元被配置成基于目标歌曲,生成上述目标歌曲对应的目标歌曲干声。

[0092] 根据本公开的一个或多个实施例,上述第三生成单元可以进一步被配置成响应于上述目标歌曲中包括背景音乐,将上述目标歌曲输入预先训练的干声、背景音乐分离模型,生成上述目标歌曲干声。

[0093] 根据本公开的一个或多个实施例,上述切分单元可以进一步被配置成将上述目标歌曲干声和上述目标歌曲干声对应的歌词输入预先训练的歌曲干声切分模型,生成上述至少一个歌曲干声片段,其中,上述每个歌曲干声片段标识有时间戳。

[0094] 根据本公开的一个或多个实施例,上述歌曲干声切分模型是以预先训练的语音识别模型为基础,使用预设的歌唱数据对上述语音识别模型训练调整得到的。

[0095] 根据本公开的一个或多个实施例,上述第一生成单元可以进一步被配置成对于上述歌词的每个字,将上述每个字对应的至少一个歌曲干声片段输入预先训练的音高检测模型,得到上述音高序列。基于上述至少一个歌曲干声片段以及至少一个歌曲干声片段对应的时间戳,生成上述歌曲干声对应的节奏信息。

[0096] 根据本公开的一个或多个实施例,上述目标歌曲干声对应的歌唱属性包括以下至少一项:颤音,换气。

[0097] 根据本公开的一个或多个实施例,上述装置的第二生成单元进一步被配置成将上述音高序列、上述节奏信息、上述歌唱属性和上述至少一个歌曲干声片段导入预设歌谱生成器中,生成上述目标歌曲干声对应的歌谱。

[0098] 根据本公开的一个或多个实施例,提供了一种电子设备,包括:一个或多个处理器;存储装置,其上存储有一个或多个程序,当一个或多个程序被一个或多个处理器执行,使得一个或多个处理器实现如上述任一实施例描述的方法。

[0099] 根据本公开的一个或多个实施例,提供了一种计算机可读介质,其上存储有计算机程序,其中,程序被处理器执行时实现如上述任一实施例描述的方法。

[0100] 以上描述仅为本公开的一些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,本公开的实施例中所涉及的范围,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本公开构思的情况下,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。例如上述特征与本公开的实施例中公开的(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。

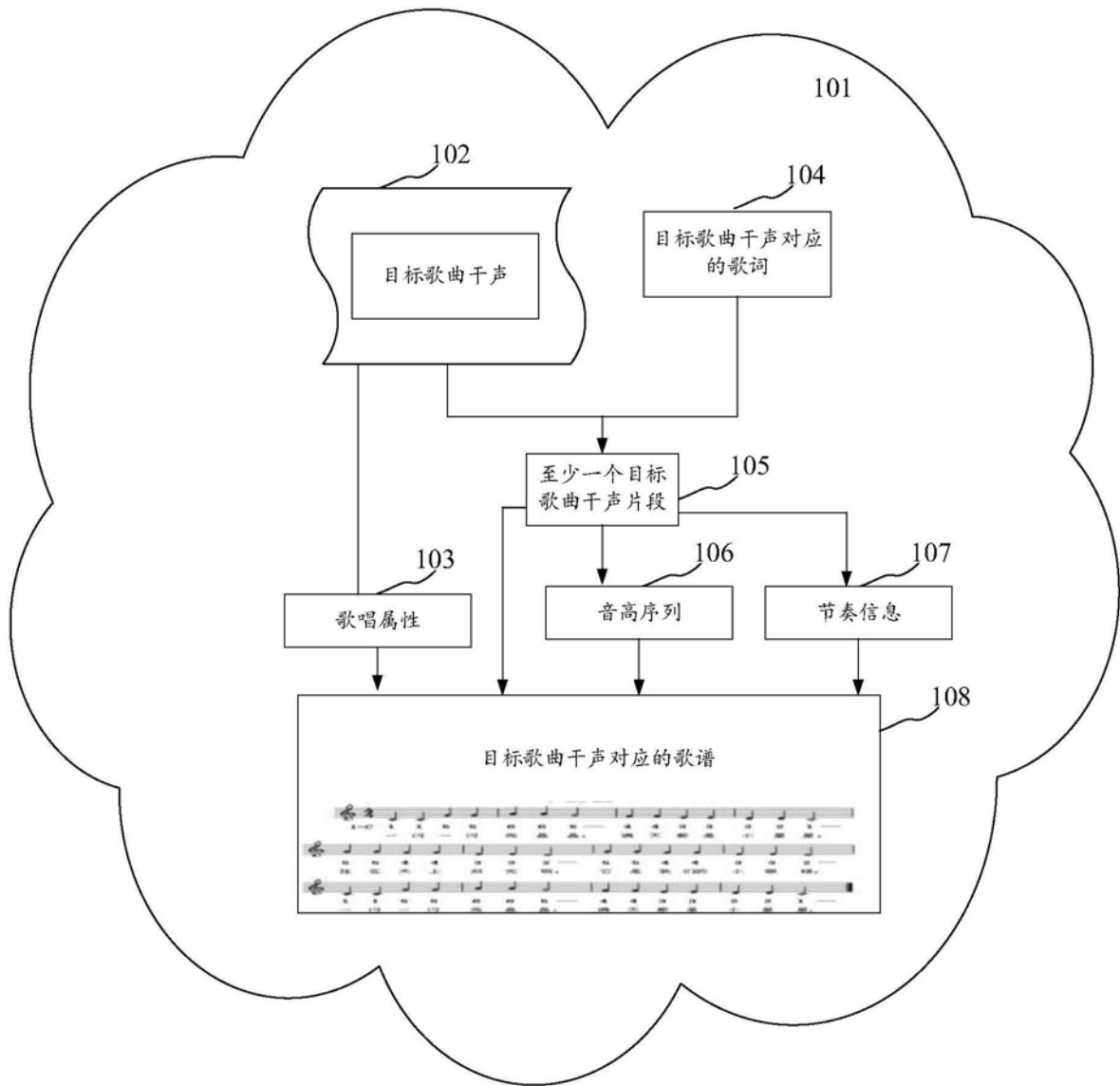


图1

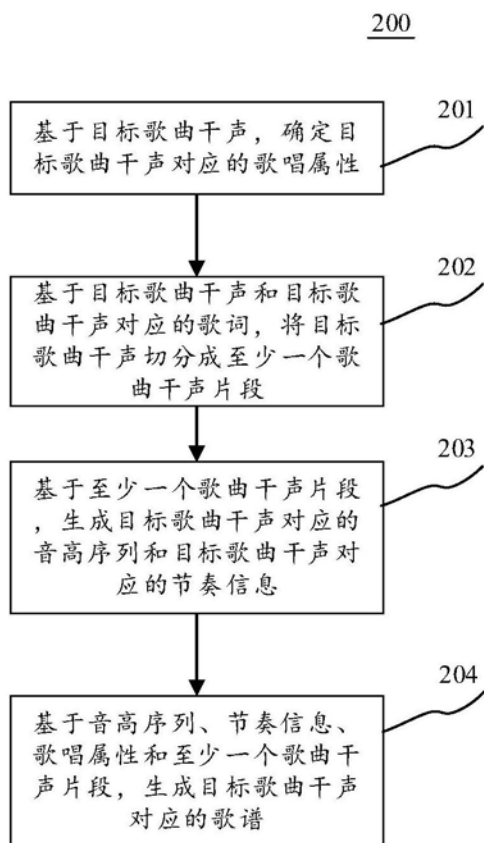


图2

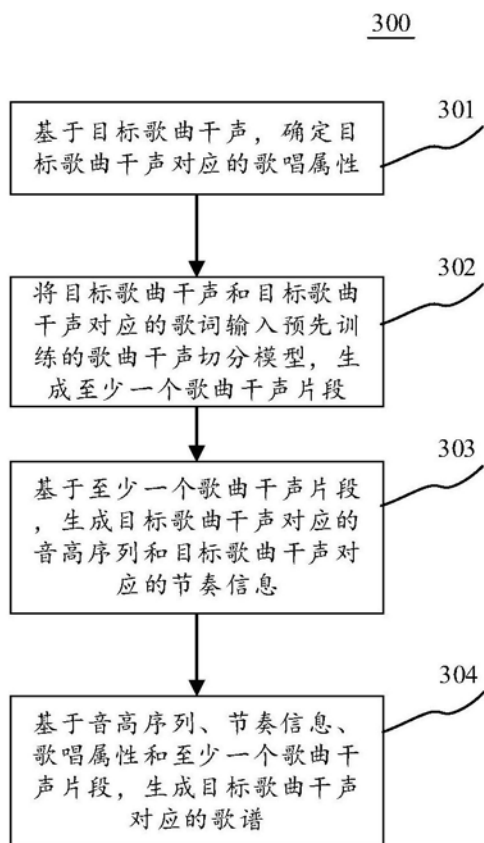


图3

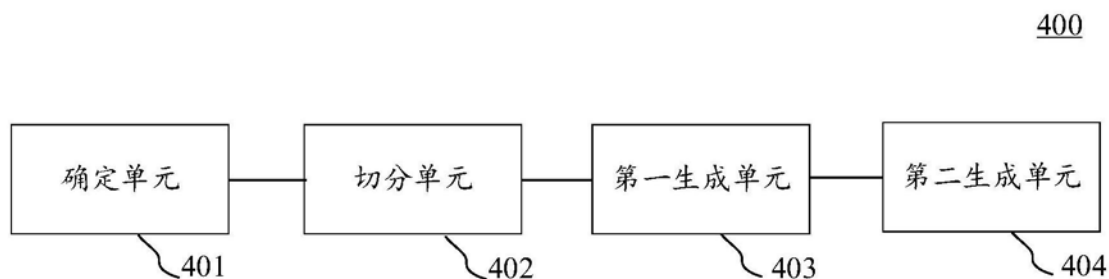


图4

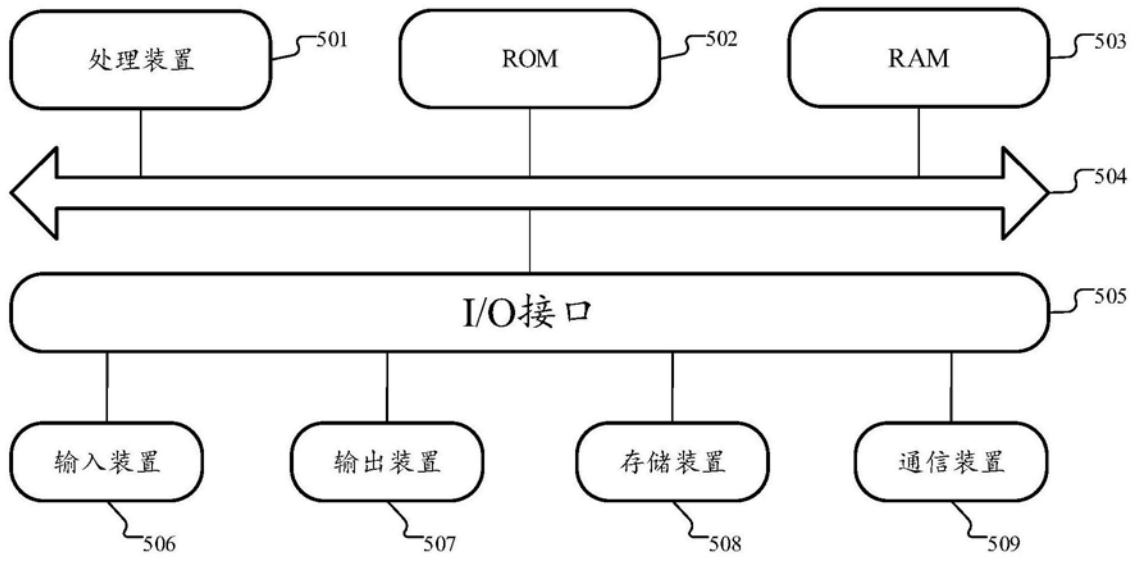
500

图5