



(12) 发明专利

(10) 授权公告号 CN 109658916 B

(45) 授权公告日 2021.03.09

(21) 申请号 201811556330.3

G10L 15/26 (2006.01)

(22) 申请日 2018.12.19

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 108091321 A, 2018.05.29

申请公布号 CN 109658916 A

US 2018286383 A1, 2018.10.04

(43) 申请公布日 2019.04.19

KR 20170103209 A, 2017.09.13

(73) 专利权人 腾讯科技(深圳)有限公司

审查员 杨士林

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 樊邵婷

(74) 专利代理机构 广州华进联合专利商标代理

有限公司 44224

代理人 李文渊 何平

(51) Int.Cl.

G10L 13/02 (2013.01)

G10L 13/08 (2013.01)

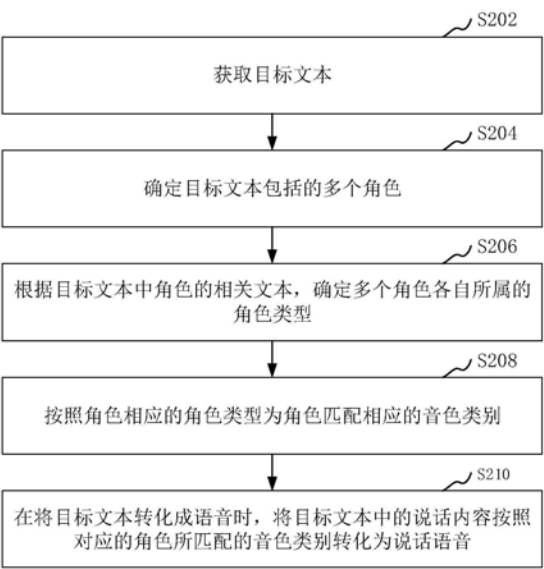
权利要求书3页 说明书14页 附图7页

(54) 发明名称

语音合成方法、装置、存储介质和计算机设备

(57) 摘要

本申请涉及一种语音合成方法、装置、计算机可读存储介质和计算机设备,所述方法包括:获取目标文本;确定所述目标文本包括的多个角色;根据所述目标文本中角色的相关文本,确定所述多个角色各自所属的角色类型;按照所述角色相应的角色类型为所述角色匹配相应的音色类别;在将所述目标文本转化成语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。本申请提供的方案可以大大扩展文字内容的表达。



1. 一种语音合成方法,包括:

获取目标文本;

确定所述目标文本包括的多个角色;所述角色包括主要角色和非主要角色;

根据所述目标文本中主要角色的相关文本,确定所述主要角色各自所属的角色类型;其中,所述相关文本是与所述主要角色相关的文本,包括角色指代词和角色描述文本中的至少一种;所述角色类型包括角色的性别类型、角色的性格类型和角色的体形类型中的至少一种;

按照所述主要角色相应的角色类型为所述主要角色匹配相应的音色类别;

依据所述目标文本中非主要角色的角色指代词,确定所述非主要角色的角色性别,并按照所述角色性别为所述非主要角色匹配相应的音色类别;

在将所述目标文本转化成语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。

2. 根据权利要求1所述的方法,其特征在于,所述获取目标文本包括:

接收语音合成请求;所述语音合成请求携带语音数据;

对所述语音数据进行语音识别,得到目标文本关键词;

依据所述目标文本关键词,查找与所述目标文本关键词相对应的目标文本。

3. 根据权利要求1所述的方法,其特征在于,所述根据所述目标文本中主要角色的相关文本,确定所述主要角色各自所属的角色类型,包括:

确定所述目标文本中分别与各所述主要角色相关的相关文本;

从预设角色类型库中,分别查找与各所述主要角色的相关文本相匹配的预设角色类型;

将查找到的预设角色类型,作为相应主要角色所属的角色类型。

4. 根据权利要求1所述的方法,其特征在于,所述在将所述目标文本转化成语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音,包括:

在将所述目标文本转化成语音时,确定所述目标文本中当前待转化的文本的内容;

在当前待转化的文本的内容为说话内容时,将所述说话内容按照对应的角色所匹配的音色类别转化为说话语音;

在当前待转化的文本的内容为旁白内容时,将所述旁白内容按预设音色类别转化为旁白语音。

5. 根据权利要求1所述的方法,其特征在于,所述在将所述目标文本转化成语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音,包括:

在将所述目标文本转化成语音时,确定所述目标文本中当前待转化的文本的内容;

在当前待转化的文本的内容为说话内容时,定位所述当前待转化的文本的邻近文本;

根据所述邻近文本确定与所述说话内容对应的角色;

将所述说话内容按照对应的角色所匹配的音色类别转化为说话语音。

6. 根据权利要求5所述的方法,其特征在于,所述在当前待转化的文本的内容为说话内容时,定位所述当前待转化的文本的邻近文本,包括:

在当前待转化的文本的内容为说话内容时,查找在所述当前待转化的文本之前、且与所述当前待转化的文本属于同个整句的第一文本;

当查找到所述第一文本、且所述第一文本包括对话词时,将所述第一文本作为所述当前待转化的文本的邻近文本;

否则,将在所述当前待转化的文本之后、且与所述当前待转化的文本属于同个整句的第二文本,作为所述当前待转化的文本的邻近文本。

7. 根据权利要求5所述的方法,其特征在于,所述根据所述邻近文本确定与所述说话内容对应的角色,包括:

当所述邻近文本中包括多个角色时,判断所述多个角色的角色指代词之间是否存在被动类型对话词;

当所述多个角色的角色指代词之间存在被动类型对话词时,将在后的角色指代词所对应的角色作为所述邻近文本中与所述说话内容对应的角色;

当所述多个角色的角色指代词之间不存在被动类型对话词时,将在前的角色指代词所对应的角色作为所述邻近文本中与所述说话内容对应的角色。

8. 一种语音合成装置,包括:

获取模块,用于获取目标文本;

确定模块,用于确定所述目标文本包括的多个角色;所述角色包括主要角色和非主要角色;

所述确定模块还用于根据所述目标文本中主要角色的相关文本,确定所述主要角色各自所属的角色类型;其中,所述相关文本是与所述主要角色相关的文本,包括角色指代词和角色描述文本中的至少一种;所述角色类型包括角色的性别类型、角色的性格类型和角色的体形类型中的至少一种;

匹配模块,用于按照所述主要角色相应的角色类型为所述主要角色匹配相应的音色类别;

所述匹配模块,还用于依据所述目标文本中非主要角色的角色指代词,确定所述非主要角色的角色性别,并按照所述角色性别为所述非主要角色匹配相应的音色类别;

转化模块,用于在将所述目标文本转化语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。

9. 根据权利要求8所述的装置,其特征在于,所述获取模块包括:

接收模块,用于接收语音合成请求;所述语音合成请求携带语音数据;

语音识别模块,用于对所述语音数据进行语音识别,得到目标文本关键词;

查找模块,用于依据所述目标文本关键词,查找与所述目标文本关键词相对应的目标文本。

10. 根据权利要求8所述的装置,其特征在于,所述确定模块还用于确定所述目标文本中分别与各所述主要角色相关的相关文本;从预设角色类型库中,分别查找与各所述主要角色的相关文本相匹配的预设角色类型;将查找到的预设角色类型,作为相应主要角色所属的角色类型。

11. 根据权利要求8所述的装置,其特征在于,所述转化模块还用于在将所述目标文本转化成语音时,确定所述目标文本中当前待转化的文本的内容;在当前待转化的文本的内容为说话内容时,将所述说话内容按照对应的角色所匹配的音色类别转化为说话语音;在当前待转化的文本的内容为旁白内容时,将所述旁白内容按预设音色类别转化为旁白语

音。

12. 根据权利要求8所述的装置,其特征在于,所述转化模块还用于在将所述目标文本转化成语音时,确定所述目标文本中当前待转化的文本的内容;在当前待转化的文本的内容为说话内容时,定位所述当前待转化的文本的邻近文本;根据所述邻近文本确定与所述说话内容对应的角色;将所述说话内容按照对应的角色所匹配的音色类别转化为说话语音。

13. 根据权利要求12所述的装置,其特征在于,所述转化模块还用于在当前待转化的文本的内容为说话内容时,查找在所述当前待转化的文本之前、且与所述当前待转化的文本属于同个整句的第一文本;当查找到所述第一文本、且所述第一文本包括对话词时,将所述第一文本作为所述当前待转化的文本的邻近文本;否则,将在所述当前待转化的文本之后、且与所述当前待转化的文本属于同个整句的第二文本,作为所述当前待转化的文本的邻近文本。

14. 根据权利要求12所述的装置,其特征在于,所述转化模块还用于当所述邻近文本中包括多个角色时,判断所述多个角色的角色指代词之间是否存在被动类型对话词;当所述多个角色的角色指代词之间存在被动类型对话词时,将在后的角色指代词所对应的角色作为所述邻近文本中与所述说话内容对应的角色;当所述多个角色的角色指代词之间不存在被动类型对话词时,将在前的角色指代词所对应的角色作为所述邻近文本中与所述说话内容对应的角色。

15. 一种计算机可读存储介质,存储有计算机程序,所述计算机程序被处理器执行时,使得所述处理器执行如权利要求1至7中任一项所述方法的步骤。

16. 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,所述计算机程序被所述处理器执行时,使得所述处理器执行如权利要求1至7中任一项所述方法的步骤。

语音合成方法、装置、存储介质和计算机设备

技术领域

[0001] 本申请涉及文字语音转换技术领域,特别是涉及一种语音合成方法、装置、计算机可读存储介质和计算机设备。

背景技术

[0002] 随着计算机技术的发展,出现了文字语音转换技术。文字语音转换技术主要是将计算机内的文本转换成连续自然的语音。传统的将文字转化成语音的方式,通常是采用TTS(Text To Speech)技术,根据文本合成相应的语音。

[0003] 然而,传统的将文字转化成声音的方式,通常全篇只有一个嗓音,并且多为女性。当需转换的文字内容非常多,比如播报书籍或故事时,采用单一的嗓音会局限对文字内容的表达。

发明内容

[0004] 基于此,有必要针对传统的将文字转化成声音的方式局限了文字内容的表达的技术问题,提供一种语音合成方法、装置、计算机可读存储介质和计算机设备。

[0005] 一种语音合成方法,包括:

[0006] 获取目标文本;

[0007] 确定所述目标文本包括的多个角色;

[0008] 根据所述目标文本中角色的相关文本,确定所述多个角色各自所属的角色类型;

[0009] 按照所述角色相应的角色类型为所述角色匹配相应的音色类别;

[0010] 在将所述目标文本转化成语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0011] 一种语音合成装置,包括:

[0012] 获取模块,用于获取目标文本;

[0013] 确定模块,用于确定所述目标文本包括的多个角色;

[0014] 所述确定模块还用于根据所述目标文本中角色的相关文本,确定所述多个角色各自所属的角色类型;

[0015] 匹配模块,用于按照所述角色相应的角色类型为所述角色匹配相应的音色类别;

[0016] 转化模块,用于在将所述目标文本转化成语音时,将所述目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0017] 一种计算机可读存储介质,存储有计算机程序,所述计算机程序被处理器执行时,使得所述处理器执行所述语音合成方法的步骤。

[0018] 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,所述计算机程序被所述处理器执行时,使得所述处理器执行所述语音合成方法的步骤。

[0019] 上述语音合成方法、装置、计算机可读存储介质和计算机设备,确定目标文本所包括的多个角色,并根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。再

按照角色相应的角色类型为角色匹配相应的音色类别。这样,通过角色的相关文本来确定该角色对应的音色类别,使得该角色的音色类别可以准确地贴近该角色。在将目标文本转化成语音时,就可将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。这样,在进行对话内容播报时,不同的角色可采用与该角色贴近的音色类别进行播报,既能区分不同的角色,又能给听众一种身临其境的感受,大大扩展了文字内容的表达。

附图说明

- [0020] 图1为一个实施例中语音合成方法的应用环境图;
- [0021] 图2为一个实施例中语音合成方法的流程示意图;
- [0022] 图3为一个实施例中根据目标文本中各角色的相关文本,确定多个角色各自所属的角色类型的步骤的流程示意图;
- [0023] 图4为一个实施例中在将目标文本合成语音时,将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音步骤的流程示意图;
- [0024] 图5为一个实施例中在将目标文本合成语音时,将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音步骤的流程示意图;
- [0025] 图6为一个实施例中对未匹配角色类型的角色所对应的说话内容进行语音转化的步骤的流程图;
- [0026] 图7为一个实施例中语音合成方法的流程示意图;
- [0027] 图8为一个实施例中语音合成方法的应用场景图;
- [0028] 图9为一个实施例中语音合成装置的结构框图;
- [0029] 图10为另一个实施例中语音合成装置的结构框图;
- [0030] 图11为一个实施例中计算机设备的结构框图。

具体实施方式

[0031] 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,此处所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。

[0032] 图1为一个实施例中语音合成方法的应用环境图。参照图1,该语音合成方法应用于语音合成系统。该语音合成系统包括终端110和服务器120。终端110和服务器120通过网络连接。终端110具体可以是台式终端或移动终端,移动终端具体可以是音箱、手机、平板电脑、笔记本电脑等中的至少一种。服务器120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。

[0033] 其中,终端110和服务器120可分别单独用于执行上述语音合成方法,或者终端110和服务器120协同执行上述语音合成方法。比如,终端110或服务器120获取目标文本,并执行该语音合成方法,将目标文本转化成目标语音。或者,终端110向服务器120发送语音合成请求,服务器120根据语音合成请求获取目标文本,并执行该语音合成方法,将目标文本转化为目标语音。

[0034] 如图2所示,在一个实施例中,提供了一种语音合成方法。本实施例主要以该方法应用于上述图1中的计算机设备(如终端110或服务器120)来举例说明。参照图2,该语音合

成方法具体包括如下步骤:

[0035] S202,获取目标文本。

[0036] 其中,目标文本是待转化为语音的文本,具体可以是存储于存储介质中的、且以电子形式存在的文本。具体地,计算机设备可接收语音合成请求,将语音合成请求所指定的文本作为待播报的目标文本。计算机设备可根据该语音合成请求查找本地存储的目标文本,或者通过USB(Universal Serial Bus,通用串行总线)接口连接或网络连接等通信方式,接收其他计算机设备发送的目标文本。

[0037] 在一个实施例中,步骤S202,也就是获取目标文本包括以下步骤:接收语音合成请求;语音合成请求携带语音数据;对语音数据进行语音识别,得到目标文本关键词;依据目标文本关键词,查找与目标文本关键词相对应的目标文本。

[0038] 其中,语音识别是从语音数据中识别出文本的处理过程。在一个实施例中,终端可采集用户录制的语音数据,并根据该语音数据生成语音合成请求。终端可直接在本地对语音数据进行语音识别,得到与语音数据相对应的文本,并从文本中确定目标文本关键词。终端也可以将包括有语音数据的语音合成请求发送至服务器,由服务器对该语音数据进行语音识别,得到与语音数据相对应的文本,并从文本中确定目标文本关键词。进一步地,计算机设备(终端或服务器)可在本地或通过搜索引擎从网络资源中查找与该目标文本关键词相匹配的目标文本。

[0039] 在一个实施例中,计算机设备(终端或服务器)从文本中确定目标文本关键词的方式具体可采用如下方式:计算机设备可预先设定关键词集合,当计算机设备对语音数据识别得到文本后,对该文本进行分词处理,并删减掉停用词后,得到词序列。计算机设备可将词序列中的各词与该关键词集和中的各词进行比对,将匹配成功的词作为目标文本关键词。

[0040] 其中,停用词(Stop Words)是指在信息检索中,为节省存储空间和提高搜索效率,在处理自然语言数据(或文本)之前或之后会自动过滤掉的某些字或词,比如一些应用十分广泛的词、语气助词、客套词、介词或连接词等。

[0041] 以该语音合成方法应用于服务器来举例说明:当用户触发语音录制操作时,终端可通过录音装置录制用户发出的声音,得到语音数据。比如,用户启动录音按键,并发出“请播放《白雪公主和七个小矮人》的故事”的声音。终端将相应的语音数据发送至服务器。服务器对语音数据进行语音识别,得到“请播放白雪公主和七个小矮人的故事”的文本。进而可确定目标文本关键词为“白雪公主和七个小矮人”。

[0042] 上述实施例中,接收包含有语音数据的语音合成请求,将语音数据进行语音识别,得到目标文本关键词。依据目标文本关键词,查找与目标文本关键词相对应的目标文本。这样,用户可通过语音请求的方式触发对目标文本的语音播报,大大提高了用户和计算机设备之间的互动性和灵活性。

[0043] 在一个实施例中,语音合成请求中携带目标文本关键词,计算机设备可直接根据目标文本关键词,查找与目标文本关键词相对应的目标文本。

[0044] S204,确定目标文本包括的多个角色。

[0045] 其中,角色是小说或剧中具有特定行为特征的人或物,具体可以是故事中的人、或具有人物特性的其他生物等。比如《红楼梦》中的贾宝玉和林黛玉,又或比如《小红帽》中的

小红帽和大灰狼等。

[0046] 在一个实施例中,计算机设备可遍历目标文本,根据目标文本中包括的角色指代词确定该目标文本中涉及到的多个角色。其中,角色指代词是指代不同角色的词,具体可以是角色的称谓、名字、或人称代词等。其中,称谓是人们因亲属或其他关系而建立起来的称呼、名称,比如爸爸、妈妈、和叔叔等。名字是角色的名称,比如贾宝玉、林黛玉、张三、和王五等。人称代词是指代人物的代词,比如他们、她们、他和她等。

[0047] 在一个实施例中,可人为对不同文本中的角色打上相应的标签,也就是人为确定不同文本中的角色。比如《白雪公主和七个小矮人》的文本,很明显,《白雪公主和七个小矮人》中的角色包括“白雪公主”、“皇后”、“七个小矮人”、“王子”等,可预先确定《白雪公主和七个小矮人》中的角色。当《白雪公主和七个小矮人》为目标文本时,计算机设备则可直接将“白雪公主”、“皇后”、“七个小矮人”、“王子”等作为该目标文本所包括的多个角色。

[0048] 在一个实施例中,计算机设备可确定该目标文本所包括的多个主要角色。其中,主要角色是小说或剧中处于中心地位的角色。具体地,计算机设备可遍历目标文本,统计不同角色所对应的角色指代词出现的频次,将大于或等于预设频次的角色指代词所对应的角色作为主要角色;计算机设备还可以将各频次从大至小进行排名,将排名名词小于预设值的频次所对应的角色作为主要角色。也就是将出现频次最多的前几名的角色作为主要角色。或者,可预先人为确定不同文本中的主要角色,并打上相应的标签。计算机设备可根据目标文本中相应的标签确定目标文本包括的多个主要角色。

[0049] S206,根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。

[0050] 其中,角色的相关文本是与该角色相关的文本,具体可以是该角色的角色指代词,和/或描述该角色的描述文本。其中角色指代词,比如角色的称谓、名字、或人称代词等。角色的描述文本具体可包括角色性格描述、角色外貌体形描述等的文本,比如善良的、邪恶的或身高九尺、浓眉大眼等。

[0051] 角色类型是具有共同点的一类角色所对应的类别,比如角色的性别类型、角色的性格类型、角色的体形类型等。其中,角色的性别类型比如男和女;角色的性格类型比如温柔和凶狠;角色的体形类型比如纤细和粗壮等。

[0052] 具体地,计算机设备可根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。比如,当某个角色的角色指代词为称为词“哥哥”时,可判定该角色为男性;当该角色的描述信息包括“憨厚的”、“身高九尺”、“体形健硕”等描述词时,可确定该角色的性格类型属于“善良”类型,体形类型属于“粗壮”类型。

[0053] 在一个实施例中,步骤S206,也就是根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型的步骤具体包括以下步骤:

[0054] S302,确定目标文本中分别与各角色相关的相关文本。

[0055] 具体地,计算机设备可定位目标文本中各角色的角色名字,在角色名字出现的文本的上下文文本中查找与该角色相关的相关文本。或者,计算机设备可遍历目标文本,确定目标文本中所有的角色指代词和/或描述角色的描述文本。再根据目标文本的语义内容、或者相关文本与角色名字的邻近关系,分析各角色指代词、以及描述文本分别所对应的角色。将与某一角色相关的角色指代词和/或角色描述文本作为该角色的相关文本。

[0056] 在一个实施例中,确定目标文本中所有的角色指代词和/或描述角色的描述文本

的方式具体可以采用如下方式:计算机设备可预先建立角色的相关文本库,在相关文本库中收录有与角色相关的各相关信息。在分析目标文本时,可对目标文本中的整句分别进行分词处理,得到目标文本所包括的各词,将各词与相关文本库进行比对,将匹配成功的词作为目标文本中的角色指代词和/或描述角色的描述文本。

[0057] S304,从预设角色类型库中,分别查找与各角色的相关文本相匹配的预设角色类型。

[0058] 其中,预设角色类型库是包括多种角色类型、角色的相关文本、以及角色类型和角色的相关文本之间的对应关系的数据库。比如,当预设角色类型库包括性别类型时,相应的与该性别类型对应的相关文本也包括在内;当预设角色类型库包括角色的性格类型时,相应的与该性格类型对应的相关文本也包括在内;当预设角色类型库包括角色的体形类型时,相应的与该体形类型对应的相关文本也包括在内。

[0059] 举例说明,当角色的类型包括性别类型如男女时,则可根据角色的称谓、名字和人物指代词建立相应的角色类型库,比如,对常用的称谓建立相关的角色类型库:女性:姐姐、妹妹、妈妈、祖母、小姐、夫人、姑娘、仙女、宫女、王后等;男性:弟弟、哥哥、爸爸、先生、公子、壮士、男士、小伙子、侍卫等。对常用的名字也建立相关的角色类型库:女性:娜、婷、娟、琳等;男性:龙、锋、鹏、涛等。对角色的人称代词进行判断并建立相应的角色类型库:女性:她和她们等;男性:他和他们等。

[0060] 相应地,当角色的类型包括角色的性格类型时,则可根据常用的性格描述文本建立相关的角色类型库;当角色的类型包括角色的体形类型时,则可根据常用的体形描述用于建立相关的角色类型库。

[0061] 具体地,计算机设备可根据各角色相关的相关文本,从预设角色类型库中,分别查找与各角色的相关文本相匹配的预设角色类型。在一个实施例中,当角色的相关文本为预设角色类型库中的相关文本时,则可直接将该相关文本所对应的角色类型作为相匹配的角色类型。

[0062] S306,将查找到的预设角色类型,作为相应角色所属的角色类型。

[0063] 具体地,计算机设备可将查找到的所有的预设角色类型均作为该角色所属的角色类型。计算机设备也可根据相关文本与预设角色类型库中的预设角色类型的匹配程度,选择匹配程度最高的预设角色类型作为该角色所属的角色类型。其中,匹配程度可依据角色的相关文本与预设角色类型库中的预设相关文本的重合程度来确定。或者,计算机设备可按角色的不同角色类型的优先级顺序确定该角色所属的角色类型。比如,当角色类型包括角色的性别类型、角色的性格类型、角色的体形类型时,且角色的性别类型的优先级高于角色的性格类型的优先级,角色的性格类型的优先级高于角色的体形类型的优先级。当查找到的与某一角色相匹配的角色类型包括角色的性别类型和角色的性格类型时,则选择角色的性别类型相对应的预设角色类型作为相应角色所属的角色类型。

[0064] 上述实施例中,根据与各角色相关的相关文本,从预设角色类型库中,分别查找与各角色相匹配的预设角色类型,并将查找到的预设角色类型作为该角色所属的角色类型,可方便准确地根据预设角色类型库查找到与角色匹配的角色类型。

[0065] S208,按照角色相应的角色类型为角色匹配相应的音色类别。

[0066] 具体地,计算机设备可按照角色相应的角色类型为角色匹配相应的音色类别。其

中,不同的音色类别对应不同的声音。在一个实施例中,计算机设备可为各角色分配与各角色的角色类型全部匹配或部分匹配的音色类别。计算机设备可将相应的文本按照匹配的音色类别转化成相应的语音。

[0067] S210,在将目标文本转化成语音时,将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0068] 其中,说话内容是指角色说话的内容,说话内容具体可以是角色自己和自己说话的内容,比如内心独白、或自言自语等;说话内容还可以是角色和其他角色之间的对话内容,比如多人对话、或多方会谈等。

[0069] 下面以《白雪公主和七个小矮人》中的选段来说明什么是说话内容。《白雪公主和七个小矮人》中的选段内容如下:

[0070] 王后敲着门喊道:“卖杂货,多好的杂货呀!”白雪公主从窗户往外看去,说道:“老人家,你好!你卖的是啥?”王后回答道:“好东西,好漂亮的东西,有各种颜色的带子和线筒。”白雪公主暗想:“这老太婆,好像并不是那种坏人,就让她进来吧。”

[0071] 其中,“卖杂货,多好的杂货呀!”、“老人家,你好!你卖的是啥?”、以及“好东西,好漂亮的东西,有各种颜色的带子和线筒。”分别是王后和白雪公主的对话内容,可以分别理解成与“王后”对应的说话内容、以及与“白雪公主”对应的说话内容。而“这老太婆,好像并不是那种坏人,就让她进来吧。”是“白雪公主”心里想的内容,虽然并未说出口,但可以理解成是“白雪公主”在心里自己对自己说的话,因而也是说话内容。

[0072] 下面再以其他角色(比如,非“人”的角色)为例说明什么是说话内容。其中,其他角色是指特殊的具有人物特性的其他生物的角色,比如具有人物特性的外星人、动物、或植物等。

[0073] 《小红帽》故事选段中有如下内容:“你好,小红帽。”大灰狼说。“你好,狼先生。”小红帽说。“这么早,你要到哪里去啊?”大灰狼问。“我要到奶奶家去。”小红帽回答。

[0074] 很明显地,“大灰狼”这个角色所说的话,比如“你好,小红帽”、以及“这么早,你要到哪里去啊”都可称作说话内容。同样地,“小红帽”这个角色所说的话也可称作说话内容。

[0075] 还比如:外星人布鲁看到地球上高高的楼房,惊呆了,心想:“这里是哪里啊,这些大大的方盒子又是什么呢?”。这句话中“外星人布鲁”内心所想的内容也可称作说话内容。

[0076] 具体地,计算机设备在将目标文本转化成语音时,在当前待转化的文本的内容为说话内容时,将说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0077] 在一个实施例中,步骤S210,也就是在将目标文本转化成语音时,将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音的步骤,具体包括以下步骤:

[0078] S402,在将目标文本转化成语音时,确定目标文本中当前待转化的文本的内容。

[0079] 具体地,计算机设备在将目标文本转化成语音时,可以段落或整句为单位进行合成。计算机设备可对当前待转化的文本的内容进行对话判断。计算机设备可遍历当前待转化的文本,确定当前待转化的文本中的标点符号,当当前待转化的文本中包括说话符号,比如双引号(“)或单引号(‘)时,则判断说话符号所限定的文本中是否存在表示整句结束的结束符号,如句号(。)、感叹号(!)或问号(?)等。当说话符号所限定的文本中存在结束符号时,则判定该说话符号所确定的文本的内容为说话内容。否则,当前待转化的文本的内容为旁白内容。

[0080] S404,在当前待转化的文本的内容为说话内容时,将说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0081] 具体地,在当前待转化的文本的内容为说话内容时,计算机设备可将说话内容按照对应的角色所匹配的音色类别转化为说话语音。可以理解的是,不同角色所对应的说话内容,可采用不同音色类别的声音进行转化。

[0082] S406,在当前待转化的文本的内容为旁白内容时,将旁白内容按预设音色类别转化为旁白语音。

[0083] 具体地,在当前待转化的文本的内容不是说话内容时,可认为该内容为旁白内容。计算机设备可将旁白内容按预设音色类别转化为旁白语音。其中,预设音色类别是预先设置的某种音色类别,在在当前待转化的文本的内容为旁白内容时,则无需区分不同角色,直接将旁白内容按预设音色类别转化为旁白语音。这样,可以明确地区分出哪些是说话内容,哪些是旁白内容,可方便用户理解目标文本,进一步地扩展了文字内容的表达。

[0084] 上述语音合成方法,确定目标文本所包括的多个角色,并根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。再按照角色相应的角色类型为角色匹配相应的音色类别。这样,通过角色的相关文本来确定该角色对应的音色类别,使得该角色的音色类别可以准确地贴近该角色。在将目标文本转化成语音时,就可将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。这样,在进行对话内容播报时,不同的角色可采用与该角色贴近的音色类别进行播报,既能区分不同的角色,又能给听众一种身临其境的感受,大大扩展了文字内容的表达。

[0085] 在一个实施例中,步骤S210,也就是在将目标文本转化成语音时,将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音的步骤,具体包括以下步骤:

[0086] S502,在将目标文本转化成语音时,确定目标文本中当前待转化的文本的内容。

[0087] 具体地,计算机设备在将目标文本转化成语音时,可以段落或整句为单位进行合成。计算机设备可对当前待转化的文本的内容进行对话判断。计算机设备可遍历当前待转化的文本,确定当前待转化的文本中的标点符号,在当前待转化的文本中包括说话符号时,进一步判断说话符号所限定的文本中是否存在表示整句结束的结束符号。当说话符号所限定的文本中存在结束符号时,则判定该说话符号所确定的文本的内容为说话内容。否则,当前待转化的文本的内容为旁白内容。

[0088] S504,在当前待转化的文本的内容为说话内容时,定位当前待转化的文本的邻近文本。

[0089] 具体地,在当前待转化的文本的内容为说话内容时,计算机设备可定位当前待转化的文本的邻近文本。其中,邻近文本是当前待转化的文本的邻近的上下文文本,具体可以是当前待转化的文本的前面或后面的文本。

[0090] 在一个实施例中,步骤S504具体包括:在当前待转化的文本的内容为说话内容时,查找在当前待转化的文本之前、且与当前待转化的文本属于同个整句的第一文本;当查找第一文本、且第一文本包括对话词时,将第一文本作为当前待转化的文本的邻近文本;否则,将在当前待转化的文本之后、且与当前待转化的文本属于同个整句的第二文本,作为当前待转化的文本的邻近文本。

[0091] 其中,属于同个整句是指共同构成具有完整语义的句子,且该句子中包括有结束

符号。具体地,在当前待转化的文本的内容为说话内容时,计算机设备可查找在当前待转化的文本之前、且与当前待转化的文本属于同个整句的第一文本。也就是查找说话符号之前的、且与该说话符号中的文本共同构成一个整句的第一文本。

[0092] 当查找到第一文本时,判断第一文本中是否存在对话词。其中,对话词是提示说话的词,包括主动类型对话词和被动类型对话词。其中,主动类型对话词是指主动说话的词,比如,说、喊、回答、想、或道等。被动类型说话词是表示听到别人说话的词,比如,听见、听到、或听说等。当查找到第一文本、且第一文本包括对话词时,将第一文本作为当前待转化的文本的邻近文本。若未查找到第一文本、或者第一文本中不包含对话词时,则将在当前待转化的文本之后、且与当前待转化的文本属于同个整句的第二文本,作为当前待转化的文本的邻近文本。

[0093] 上述实施例中,考虑到文本内容中关于说话内容和对话词之间的先后关系,在当前待转化的文本的内容为说话内容时,优先将当前待转化的文本之前、且与当前待转化的文本属于同个整句的第一文本作为邻近文本。当第一文本中不存在对话词,或不存在第一文本时,则考虑将在后的文本作为邻近文本,可以更准确地定位当前待转化的文本的邻近文本,进而提高了后续根据邻近文本确定相应角色的准确性。

[0094] S506,根据邻近文本确定与说话内容对应的角色。

[0095] 具体地,计算机设备可根据邻近文本中出现的角色指代词,确定与说话内容对应的角色。在一个实施例中,邻近文本中仅出现唯一的角色指代词,那么毫无疑问,计算机设备可直接将该角色指代词所对应的角色作为与说话内容对应的角色。

[0096] 在一个实施例中,邻近文本中出现多个角色指代词,那么此时,计算机设备可根据多个角色指代词之间的文本确定与对话内容对应的角色。

[0097] 在一个实施例中,步骤S506具体包括:当邻近文本中包括多个角色时,判断多个角色的角色指代词之间是否存在被动类型对话词;当多个角色的角色指代词之间存在被动类型对话词时,将在后的角色指代词所对应的角色作为邻近文本中与说话内容对应的角色;当多个角色的角色指代词之间不存在被动类型对话词时,将在前的角色指代词所对应的角色作为邻近文本中与说话内容对应的角色。

[0098] 具体地,当邻近文本中包括多个角色时,计算机设备可判断多个角色的角色指代词之间是否存在被动类型对话词。其中,被动类型对话词是表示听到别人说话的词,比如,听见、听到、或听说等。当多个角色指代词之间存在被动类型对话词时,将在后的角色指代词所对应的角色作为该邻近文本中与说话内容对应的角色。当多个角色指代词之间不存在被动类型对话词时,将在前的角色指代词所对应的角色作为该邻近文本中与说话内容对应的角色。

[0099] 举例说明,当待合成文本以及邻近文本的结构为提示语+说话内容时,比如,小男孩高兴地喊道:“谢谢叔叔!”(其中“喊道”为主动类型对话词),计算机设备可将“小男孩”作为与该说话内容对应的角色。

[0100] 在待合成文本以及邻近文本的结构为说话内容+提示语时,计算机设备可将角色指代词所指代的角色作为与说话内容对应的角色。比如,“他们一定会回来的!”小明激动地说。其中“说”为主动类型对话词,计算机设备将“小明”作为与该说话内容对应的角色。

[0101] 在待合成文本以及邻近文本的结构为说话内容+提示语+说话内容时,计算机设备

可将角色指代词所指代的角色作为与前后说话内容分别对应的角色。比如，“是的！”小红解释道，“因为天气预报播报明天是晴天”。其中“解释道”为主动类型对话词，计算机设备将“小红”作为与前后说话内容分别对应的角色。

[0102] 当邻居文本中包括多个角色时，计算机设备可根据多个角色指代词之间是否存在被动类型对话词来确定相应的角色。比如，她看着他说：“…”，此时“看着”是一种与听觉无关的词，并不是被动类型对话词，因而采用在前的角色指代词“她”所对应的角色作为与说话内容对应的角色。还比如，她听见他说：“…”，此时“听见”是被动类型对话词，因而采用在后的角色指代词“他”所对应的角色作为与说话内容对应的角色。

[0103] 上述实施例中，当邻近文本中包括多个角色时，通过判断多个角色的角色指代词之间是否存在被动类型对话词，可以灵活准确地确定与说话内容对应的角色。

[0104] S508，将说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0105] 具体地，在当前待转化的文本的内容为说话内容时，计算机设备可将说话内容按照对应的角色所匹配的音色类别转化为说话语音。可以理解的是，不同角色所对应的说话内容，可采用不同音色类别的声音进行转化。

[0106] 上述实施例中，在当前待转化的文本的内容为说话内容时，根据当前待转化的文本的邻近文本，也就是提示语文本来确定与说话内容对应的角色，可以准确地确定与说话内容对应的角色，进而准确地通过不同的音色类别区分出不同的“说话者”，大大扩展了文字内容的表达。

[0107] 在一个实施例中，该语音合成方法还包括对未匹配角色类型的角色所对应的说话内容进行语音转化的步骤，该步骤具体包括：当目标文本中的说话内容所对应的角色未成功匹配角色类型时，依据角色的角色指代词确定角色的角色性别；在将目标文本转化成语音时，采用与角色性别相匹配的音色类别，将未成功匹配角色类型的角色所对应的说话内容转化为说话声音。

[0108] 在一个实施例中，当相关文本为角色的描述文本，且计算机设备根据目标文本中各角色的相关文本，无法准确地确定多个角色各自所属的角色类型时，计算机设备可依据角色的角色指代词确定角色的角色性别，并从预设角色类型库中随机选择与相应角色性别匹配的音色类型。

[0109] 在一个实施例中，计算机设备为提高语音合成的效率，可只为主要角色进行角色匹配。也就是，当计算机设备确定的目标文本包括的多个角色为主要角色时，未被确定的角色（非主要角色）就并未匹配相应的角色类型。此时，计算机设备在将目标文本转化成语音时，可根据非主要角色的角色指代词确定相应的角色性别，并从预设角色类型库中随机选择与相应角色性别匹配的音色类型。

[0110] 进一步地，计算机设备可采用与角色性别相匹配的音色类别，将未成功匹配角色类型的角色所对应的说话内容转化为说话声音。

[0111] 参考图6，图6为一个实施例中对未匹配角色类型的角色所对应的说话内容进行语音转化的步骤的流程图。当计算机设备仅为主要角色匹配相应的角色类型时，如图6所示，步骤S602，在当前待转化的文本的内容为说话内容时，定位当前待转化的文本的邻近文本。步骤S604，对邻居文本中的角色进行判断。当该角色为主要角色时，跳转至步骤S606，将说话内容按照与主要角色相匹配的音色类别转化为说话语音。当该角色不是主要角色时，则

跳转至步骤S608,判断该角色的角色性别。当该角色的角色性别为男性时,则跳转至步骤S610,采用男声播报该说话内容;当该角色的角色性别为女性时,则跳转至步骤S612,采用女声播报该说话内容。当说话内容合成完毕后,跳转至步骤S614,当说话内容播报结束后,切换回旁白语音。

[0112] 上述实施例中,当目标文本中的说话内容所对应的角色未成功匹配角色类型时,可依据角色的角色指代词确定角色的角色性别,在将目标文本转化成语音时,采用与角色性别相匹配的音色类别,将未成功匹配角色类型的角色所对应的说话内容转化为说话声音,大大提高了语音合成的效率。

[0113] 下面以《白雪公主和七个小矮人》的故事选段为例,详细说明本方案。同时,以与角色的相关文本为角色指代词、角色类型为角色的性别类型为例来进行说明。其中,目标文本为:

[0114] 王后听了大吃一惊,因为她知道这面镜子是从不说假话的,一定是那仆人蒙骗了她,她决不能容忍有任何比她更漂亮的人活在这个世上。所以,她把自己装扮成一个卖杂货的老太婆,翻山越岭来到了那七个小矮人的住处。她敲着门喊道:“卖杂货,多好的杂货呀!”白雪公主从窗户往外看去,说道:“老人家,你好!你卖的是什么啊?”她回答道:“好东西,好漂亮的东西,有各种颜色的带子和线筒。”白雪公主暗想:“这老太婆,好像并不是那种坏人,就让她进来吧。”想到这里,她跑过去打开门。老太婆进来后说道:“哎呀!看你的胸带多差呀,来吧,让我给你系上一根漂亮的新带子。”白雪公主做梦也没想到这会有危险,所以她走上前去站在了老太婆的面前。老太婆很熟练地将带子给她系在胸前,系着系着,突然,她猛地用力将带子拉紧,白雪公主便被勒得透不过气来,很快失去知觉倒在了地上,就像死去了一样。看到她的样子,恶毒的王后说道:“这下你的美丽该结束了吧!”说完放心地走了。

[0115] 晚上,七个小矮人回来了,当他们看到他们诚实可爱的白雪公主躺在地上一动不动,就像死了一样时,他们的心马上缩紧了,急忙上前将她抬了起来,他们马上剪断了带子。过了一会儿,白雪公主慢慢地开始呼吸了,不久她又活了过来。听她讲完事情的经过后,他们说道:“那个老太婆就是王后,下次你要当心,在我们离开后,千万不要让任何人进来。”

[0116] 首先,根据目标文本的内容,可知,在这两段文本中,出现最多次数的是“白雪公主”和“王后”,因此计算机设备可将两者定为主要角色。计算机设备可根据“白雪公主”和“王后”的称谓,识别为女性,因此给“白雪公主”和“王后”分别匹配两种不同的女性声音:女A,女B。

[0117] 在将目标文本转化成语音时,当出现冒号和坐引号时,意味着要出现对话了。计算机设备可对说话人进行识别,如果是白雪公主,则配与女A,如果是王后,则配与女B。如果是其它次要角色,则按照性别配与不同的男声和女声。如小矮人出现时,通过第三人称“他们”,判断为男性,并配与相应的男声。遇到右引号时,则恢复旁白声音。

[0118] 在一个具体实施例中,语音合成方法具体包括以下步骤:

[0119] S702,接收语音合成请求;语音合成请求携带语音数据。

[0120] S704,对语音数据进行语音识别,得到目标文本关键词。

[0121] S706,依据目标文本关键词,查找与目标文本关键词相对应的目标文本。

[0122] S708,确定目标文本包括的多个角色。

[0123] S710,确定目标文本中分别与各角色相关的相关文本。

[0124] S712,从预设角色类型库中,分别查找与各角色的相关文本相匹配的预设角色类型。

[0125] S714,将查找到的预设角色类型,作为相应角色所属的角色类型。

[0126] S716,按照各角色相应的角色类型为各角色匹配相应的音色类别。

[0127] S718,在将目标文本转化成语音时,确定目标文本中当前待转化的文本。

[0128] S720,在当前待转化的文本的内容为旁白内容时,将旁白内容按预设音色类别转化为旁白语音。

[0129] S722,在当前待转化的文本的内容为说话内容时,查找在当前待转化的文本之前、且与当前待转化的文本属于同个整句的第一文本。

[0130] S724,当查找到第一文本、且第一文本包括对话词时,将第一文本作为当前待转化的文本的邻近文本。

[0131] S726,否则,将在当前待转化的文本之后、且与当前待转化的文本属于同个整句的第二文本,作为当前待转化的文本的邻近文本。

[0132] S728,根据邻近文本确定与说话内容对应的角色。

[0133] S730,将说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0134] S732,当目标文本中的说话内容所对应的角色未成功匹配角色类型时,依据角色的角色指代词确定角色的角色性别。

[0135] S734,在将目标文本转化成语音时,采用与角色性别相匹配的音色类别,将未成功匹配角色类型的角色所对应的说话内容转化为说话声音。

[0136] 上述语音合成方法,确定目标文本所包括的多个角色,并根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。再按照角色相应的角色类型为角色匹配相应的音色类别。这样,通过角色的相关文本来确定该角色对应的音色类别,使得该角色的音色类别可以准确地贴近该角色。在将目标文本转化成语音时,就可将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。这样,在进行对话内容播报时,不同的角色可采用与该角色贴近的音色类别进行播报,既能区分不同的角色,又能给听众一种身临其境的感受,大大扩展了文字内容的表达。

[0137] 图7为一个实施例中语音合成方法的流程示意图。应该理解的是,虽然图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7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,这些子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的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0138] 在具体应用场景中,参考图8,图8示出了一个实施例中语音合成方法的应用场景图。参考图8,用户可启动音箱810,用户和音箱810进行语音交互,比如用户对音箱810说“请播放《白雪公主和七个小矮人》”。音箱810采集用户声音,录制用户说的话,并生成相应的语音数据。音箱810根据语音数据生成语音合成请求,并通过网络连接发送至服务器820。服务器820对录制的语音数据进行语音识别,得到“请播放《白雪公主和七个小矮人》”,并获取相应的《白雪公主和七个小矮人》的文本,再将《白雪公主和七个小矮人》中的说话内容按照对

应角色所匹配的音色类别转化成相应的说话语音,通过音箱进行语音播报。音箱在播报《白雪公主和七个小矮人》时,根据《白雪公主和七个小矮人》中不同的角色采用相应匹配的音色类别来进行播报,这样,用户听到的有声书就不会从头至尾都只有一种声音,而是像真实的人物对话那样,可以明显地区分出不同的角色,大大扩展了文字内容的表达。

[0139] 如图9所示,在一个实施例中,提供了语音合成装置900,包括获取模块901、确定模块902、匹配模块903和转化模块904。

[0140] 获取模块901,用于获取目标文本。

[0141] 确定模块902,用于确定目标文本包括的多个角色。

[0142] 确定模块902还用于根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。

[0143] 匹配模块903,用于按照角色相应的角色类型为角色匹配相应的音色类别。

[0144] 转化模块904,用于在将目标文本转化成语音时,将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0145] 如图10所示,在一个实施例中,获取模块901包括接收模块9011、语音识别模块9012和查找模块9013,其中:

[0146] 接收模块9011,用于接收语音合成请求;语音合成请求携带语音数据。

[0147] 语音识别模块9012,用于对语音数据进行语音识别,得到目标文本关键词。

[0148] 查找模块9013,用于依据目标文本关键词,查找与目标文本关键词相对应的目标文本。

[0149] 在一个实施例中,确定模块902还用于确定目标文本中分别与各角色相关的相关文本;从预设角色类型库中,分别查找与各角色的相关文本相匹配的预设角色类型;将查找到的预设角色类型,作为相应角色所属的角色类型。

[0150] 在一个实施例中,转化模块904还用于在将目标文本转化成语音时,确定目标文本中当前待转化的文本;在当前待转化的文本的内容为说话内容时,将说话内容按照对应的角色所匹配的音色类别转化为说话语音;在当前待转化的文本的内容为旁白内容时,将旁白内容按预设音色类别转化为旁白语音。

[0151] 在一个实施例中,转化模块904还用于在将目标文本转化成语音时,确定目标文本中当前待转化的文本;在当前待转化的文本的内容为说话内容时,定位当前待转化的文本的邻近文本;根据邻近文本确定与说话内容对应的角色;将说话内容按照对应的角色所匹配的音色类别转化为说话语音。

[0152] 在一个实施例中,转化模块904还用于在当前待转化的文本的内容为说话内容时,查找在当前待转化的文本之前、且与当前待转化的文本属于同个整句的第一文本;当查找到第一文本、且第一文本包括对话词时,将第一文本作为当前待转化的文本的邻近文本;否则,将在当前待转化的文本之后、且与当前待转化的文本属于同个整句的第二文本,作为当前待转化的文本的邻近文本。

[0153] 在一个实施例中,转化模块904还用于当邻近文本中包括多个角色时,判断多个角色的角色指代词之间是否存在被动类型对话词;当多个角色的角色指代词之间存在被动类型对话词时,将在后的角色指代词所对应的角色作为邻近文本中与说话内容对应的角色;当多个角色的角色指代词之间不存在被动类型对话词时,将在前的角色指代词所对应的角

色作为邻近文本中与说话内容对应的角色。

[0154] 在一个实施例中,确定模块902还用于当目标文本中的说话内容所对应的角色未成功匹配角色类型时,依据角色的角色指代词确定角色的角色性别。转化模块904还用于在将目标文本转化成语音时,采用与角色性别相匹配的音色类别,将未成功匹配角色类型的角色所对应的说话内容转化为说话声音。

[0155] 上述语音合成装置,确定目标文本所包括的多个角色,并根据目标文本中角色的相关文本,确定多个角色各自所属的角色类型。再按照角色相应的角色类型为角色匹配相应的音色类别。这样,通过角色的相关文本来确定该角色对应的音色类别,使得该角色的音色类别可以准确地贴近该角色。在将目标文本转化成语音时,就可将目标文本中的说话内容按照对应的角色所匹配的音色类别转化为说话语音。这样,在进行对话内容播报时,不同的角色可采用与该角色贴近的音色类别进行播报,既能区分不同的角色,又能给听众一种身临其境的感受,大大扩展了文字内容的表达。

[0156] 图11示出了一个实施例中计算机设备的内部结构图。该计算机设备具体可以是图1中的终端110或服务器120。如图11所示,该计算机设备包括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,存储器包括非易失性存储介质和内存存储器。该计算机设备的非易失性存储介质存储有操作系统,还可存储有计算机程序,该计算机程序被处理器执行时,可使得处理器实现语音合成方法。该内存存储器中也可储存有计算机程序,该计算机程序被处理器执行时,可使得处理器执行语音合成方法。

[0157] 本领域技术人员可以理解,图11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。

[0158] 在一个实施例中,本申请提供的语音合成装置可以实现为一种计算机程序的形式,计算机程序可在如图11所示的计算机设备上运行。计算机设备的存储器中可存储组成该语音合成装置的各个程序模块,比如,图8所示的获取模块、确定模块、匹配模块和转化模块。各个程序模块构成的计算机程序使得处理器执行本说明书中描述的本申请各个实施例的语音合成方法中的步骤。

[0159] 例如,图11所示的计算机设备可以通过如图8所示的语音合成装置中的获取模块执行步骤S202。计算机设备可通过确定模块执行步骤S204和S206。计算机设备可通过匹配模块执行步骤S208。计算机设备可通过转化模块执行步骤S210。

[0160] 在一个实施例中,提供了一种计算机设备,包括存储器和处理器,存储器存储有计算机程序,计算机程序被处理器执行时,使得处理器执行上述语音合成方法的步骤。此处语音合成方法的步骤可以是上述各个实施例的语音合成方法中的步骤。

[0161] 在一个实施例中,提供了一种计算机可读存储介质,存储有计算机程序,计算机程序被处理器执行时,使得处理器执行上述语音合成方法的步骤。此处语音合成方法的步骤可以是上述各个实施例的语音合成方法中的步骤。

[0162] 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,所述的程序可存储于一非易失性计算机可读取存储介质中,该程序在执行时,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/或易失性存储器。非易失性存储器可包括只读存储器 (ROM)、可编程ROM (PROM)、电可编程ROM (EPROM)、电可擦除可编程ROM (EEPROM) 或闪存。易失性存储器可包括随机存取存储器 (RAM) 或者外部高速缓冲存储器。作为说明而非局限, RAM以多种形式可得, 诸如静态RAM (SRAM)、动态RAM (DRAM)、同步DRAM (SDRAM)、双数据率SDRAM (DDRSDRAM)、增强型SDRAM (ESDRAM)、同步链路 (Synchlink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接RAM (RDRAM)、直接存储器总线动态RAM (DRDRAM)、以及存储器总线动态RAM (RDRAM) 等。

[0163] 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。

[0164] 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, 但并不能因此而理解为对本申请专利范围的限制。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。

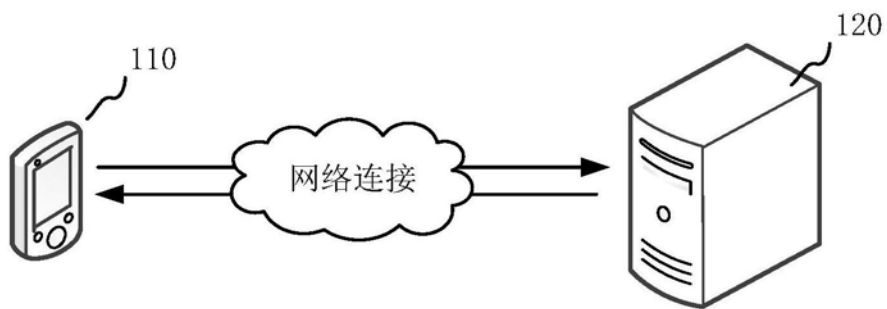


图1

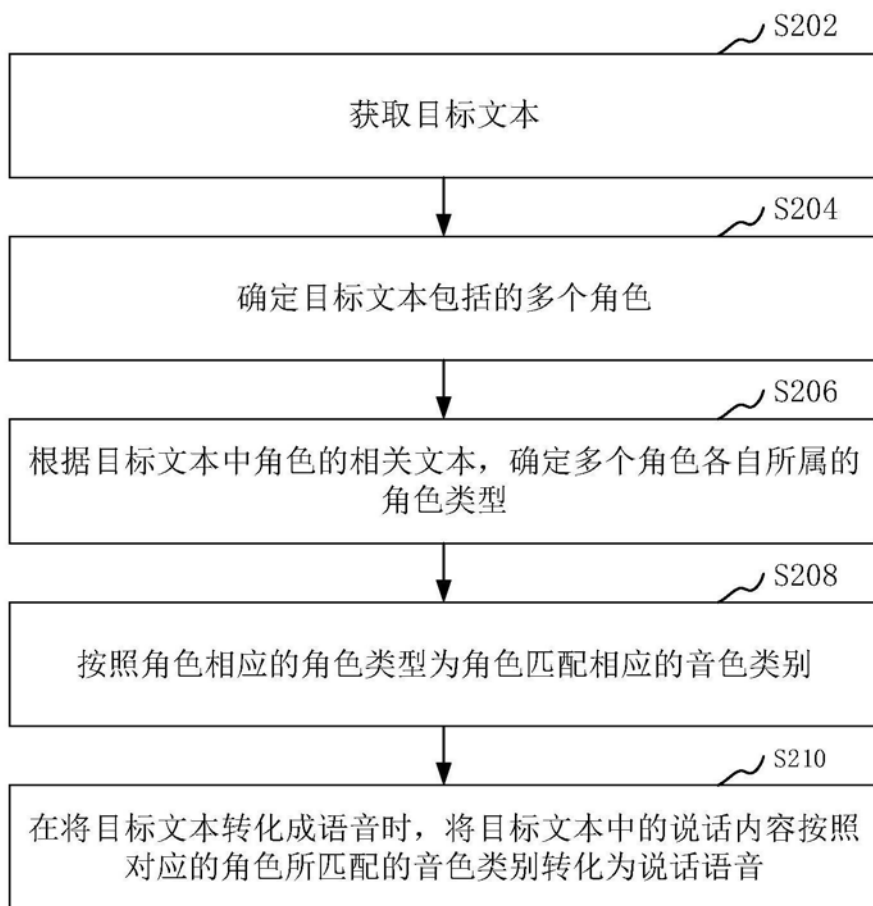


图2

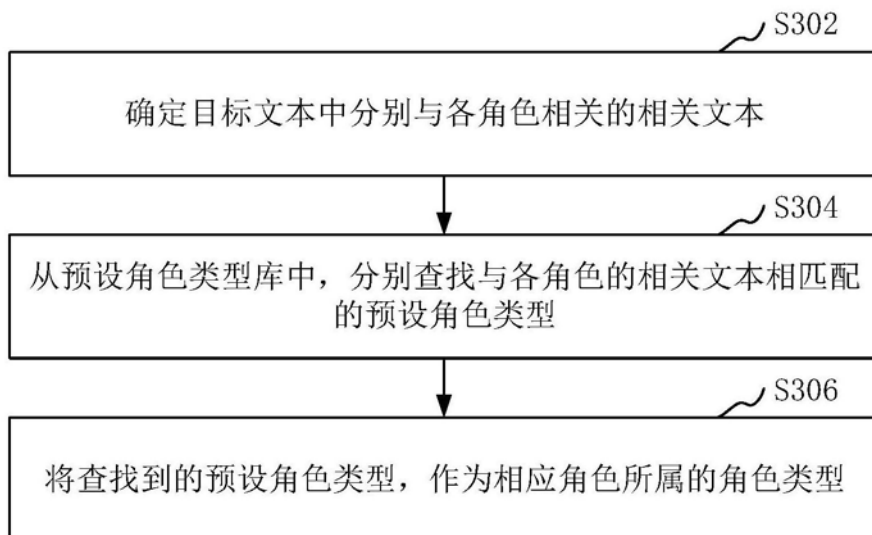


图3

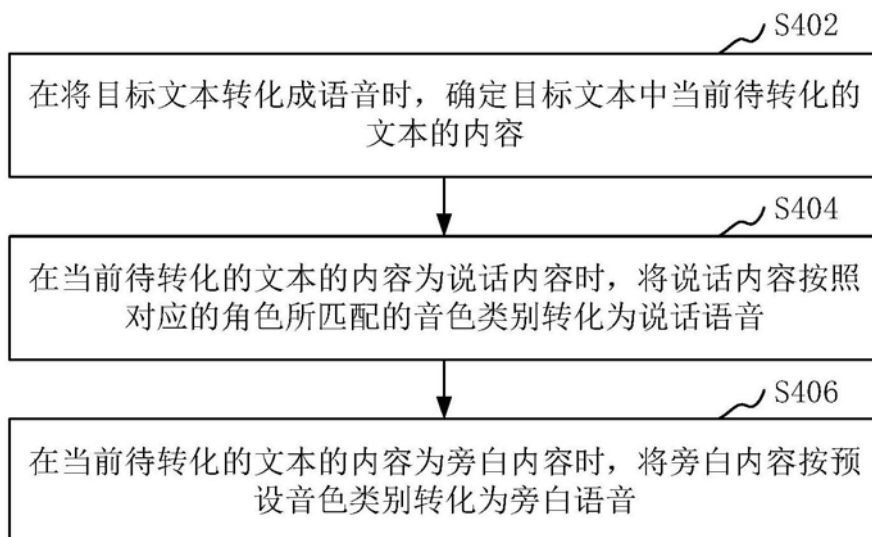


图4

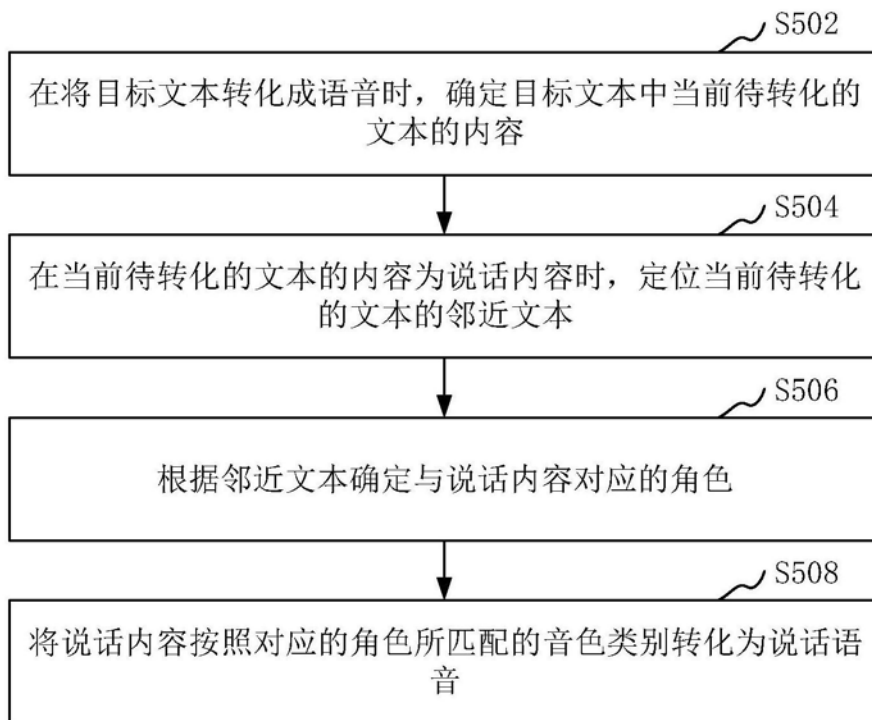


图5

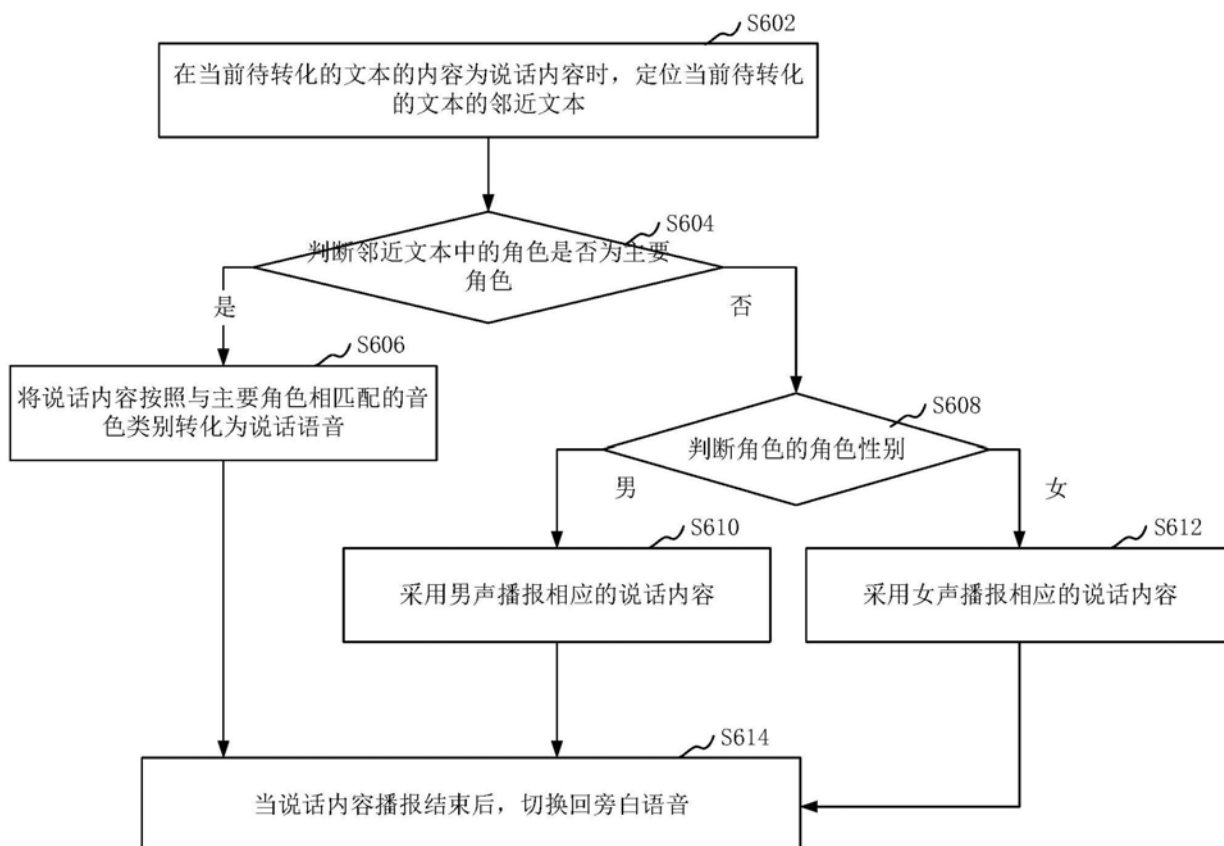


图6

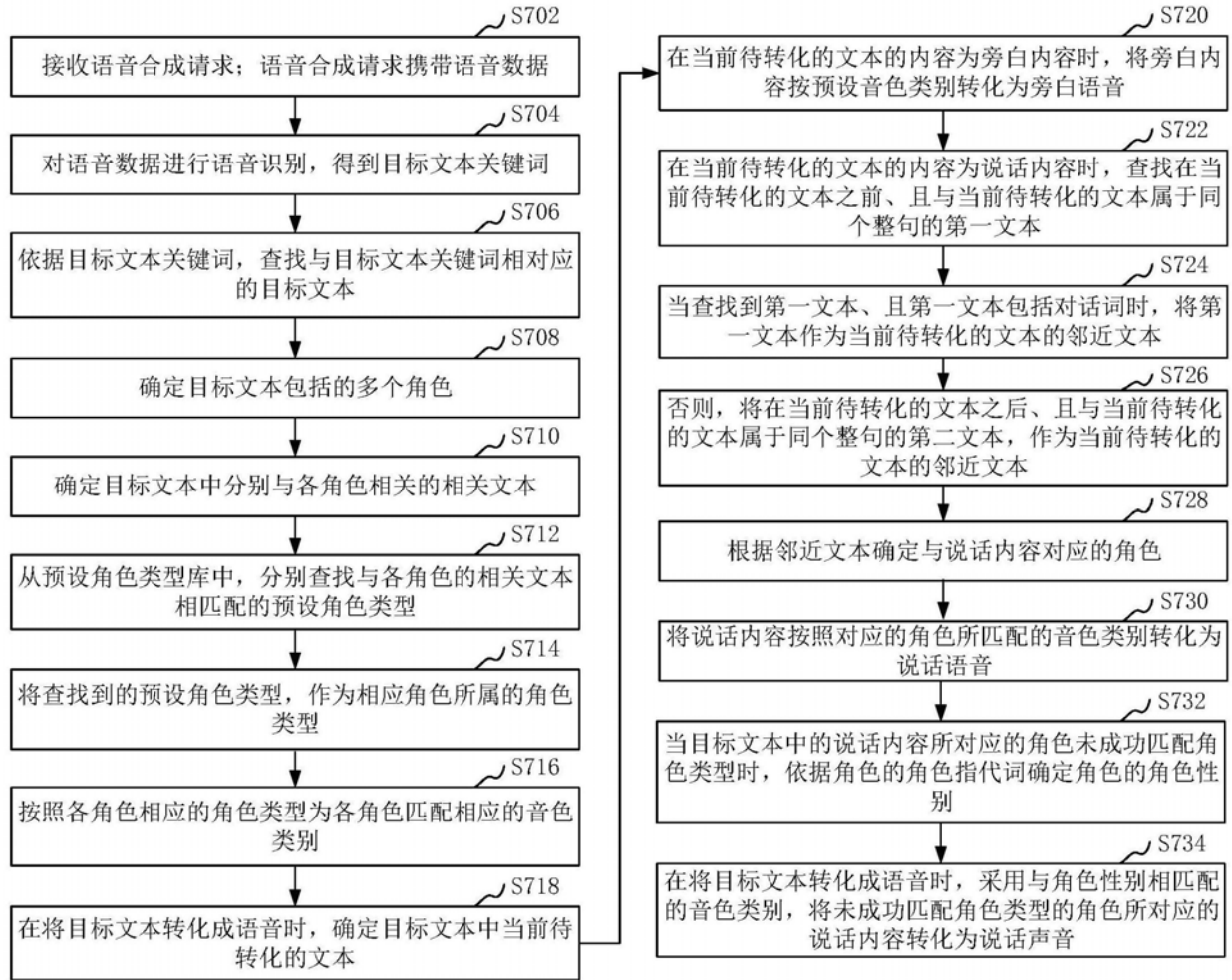


图7

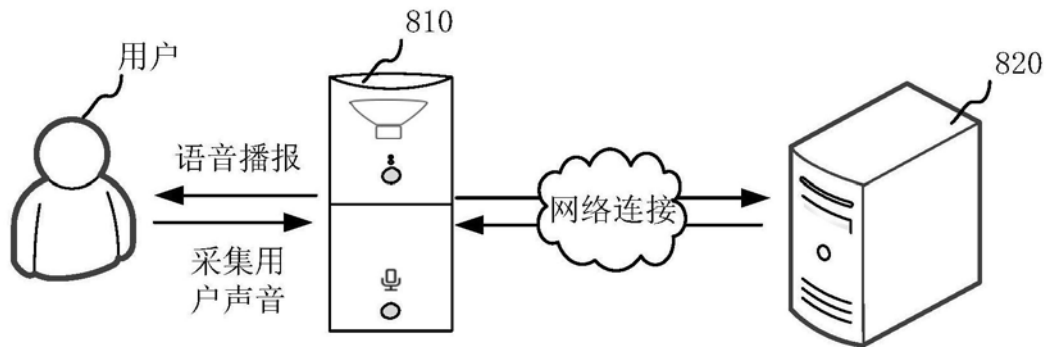


图8

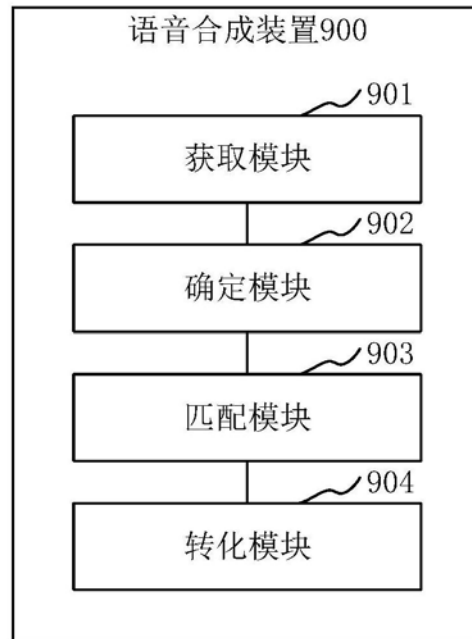


图9

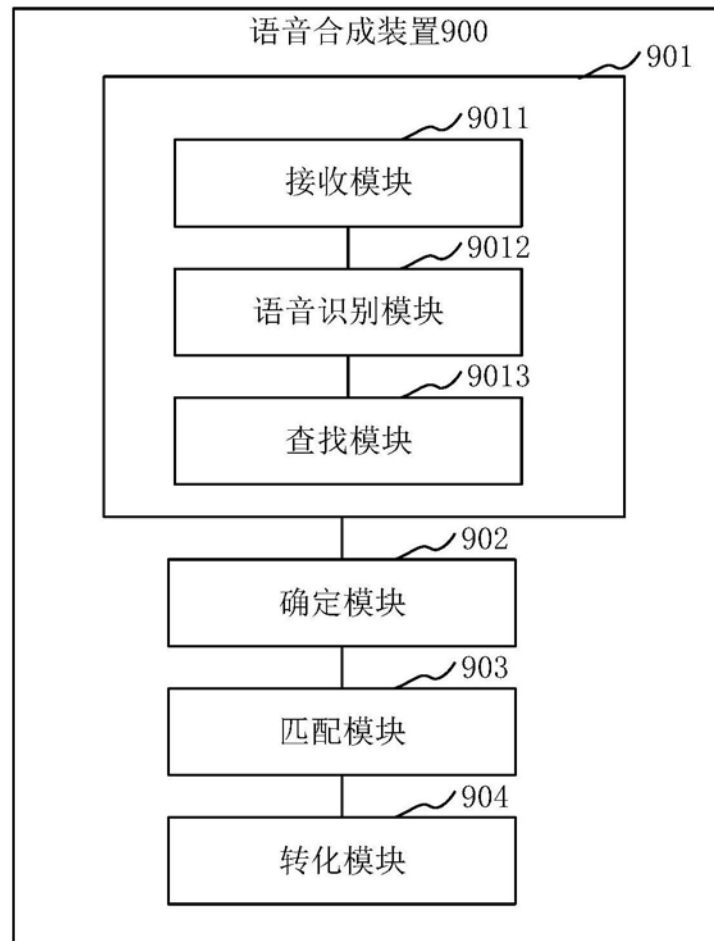


图10

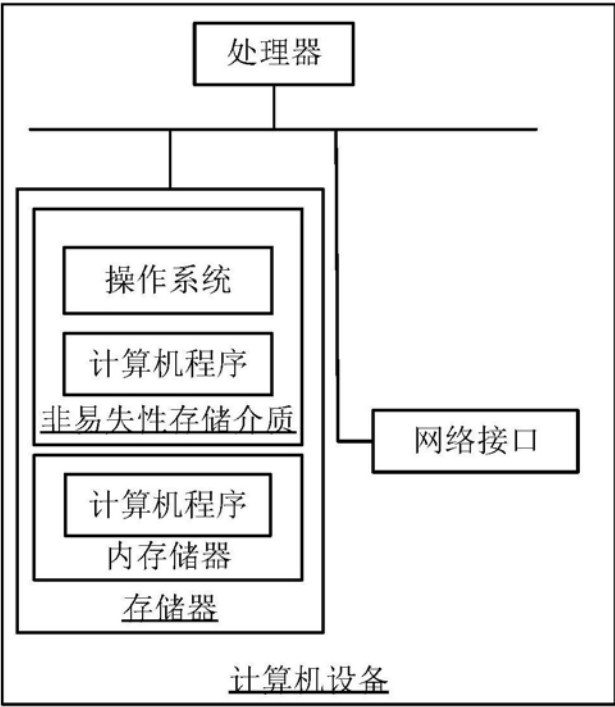


图11