



(12)发明专利申请

(10)申请公布号 CN 110765763 A

(43)申请公布日 2020.02.07

(21)申请号 201910903618.1

(22)申请日 2019.09.24

(71)申请人 金蝶软件(中国)有限公司

地址 518000 广东省深圳市南山区科技园
科技南十二路2号金蝶软件园A座1-8
层

(72)发明人 宁义双 张良杰 闵刚

(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 于丽君

(51)Int.Cl.

G06F 40/232(2020.01)

G06F 40/289(2020.01)

G06F 40/242(2020.01)

G06F 16/332(2019.01)

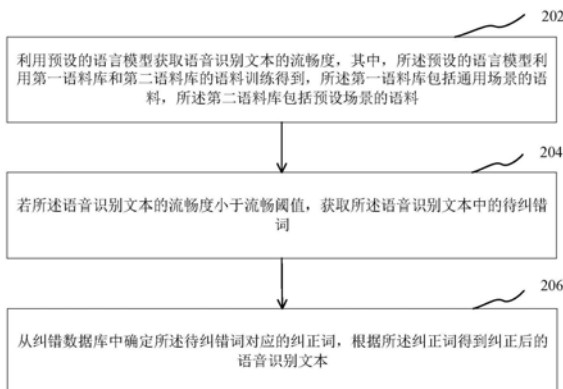
权利要求书2页 说明书12页 附图5页

(54)发明名称

语音识别文本的纠错方法、装置、计算机设备和存储介质

(57)摘要

本申请涉及一种语音识别文本的纠错方法、装置、计算机设备和存储介质。所述方法包括：利用预设的语言模型获取语音识别文本的流畅度，其中，所述预设的语言模型利用第一语料库和第二语料库的语料训练得到，所述第一语料库包括通用场景的语料，所述第二语料库包括预设场景的语料；若所述语音识别文本的流畅度小于流畅阈值，获取所述语音识别文本中的待纠错词；从纠错数据库中确定所述待纠错词对应的纠正词，根据所述纠正词得到纠正后的语音识别文本。本申请提高了对用户意图识别的准确率。



1. 一种语音识别文本的纠错方法,所述方法包括:

利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用场景的语料,所述第二语料库包括预设场景的语料;

若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词;

从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。

2. 根据权利要求1所述的方法,其特征在于,所述纠错数据库的构建方式包括:

获取所述第二语料库的语料;

利用分词词典对所述第二语料库的语料进行分词,得到候选词;

根据所述候选词以及所述候选词的拼音构建所述纠错数据库。

3. 根据权利要求2所述的方法,其特征在于,所述方法还包括:

获取所述候选词对应的混淆词;

将所述混淆词加入所述分词词典。

4. 根据权利要求3所述的方法,其特征在于,所述获取所述语音识别文本中的待纠错词,包括:

利用所述分词词典对所述语音识别文本进行分词,得到文本词;

计算各个所述文本词的平均绝对偏差值;

若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为所述待纠错词。

5. 根据权利要求2所述的方法,其特征在于,所述从纠错数据库中确定所述待纠错词对应的纠正词包括:

从所述纠错数据库中确定所述待纠错词对应的纠错候选词;

在所述纠错候选词中确定所述纠正词。

6. 根据权利要求5所述的方法,其特征在于,所述从所述纠错数据库中确定所述待纠错词对应的纠错候选词,包括:

获取所述待纠错词的拼音;

获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度;

将所述相似度大于相似阈值的候选词作为所述纠错候选词。

7. 根据权利要求6所述的方法,其特征在于,所述获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度,包括:

获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的编辑距离,利用所述编辑距离表征所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度。

8. 根据权利要求5所述的方法,其特征在于,所述在所述纠错候选词中确定所述纠正词,包括:

利用所述纠错候选词替换所述语音识别文本中的待纠错词,并利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度;

将所述流畅度满足预设条件的纠错候选词作为所述纠正词。

9. 根据权利要求8所述的方法,其特征在于,所述预设的语言模型为二元语言模型和三元语言模型;

所述利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度,包括:

将替换后的所述语音识别文本分别输入所述二元语言模型以及所述三元语言模型,得到所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度;

将所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度中的最大值,作为所述语音识别文本的流畅度。

10. 一种语音识别文本的纠错装置,其特征在于,所述装置包括:

获取模块,用于利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用场景的语料,所述第二语料库包括预设场景的语料;

所述获取模块,还用于若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词;

确定模块,用于从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。

11. 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9中任一项所述方法的步骤。

12. 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9中任一项所述的方法的步骤。

语音识别文本的纠错方法、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识别文本的纠错方法、装置、计算机设备和存储介质。

背景技术

[0002] 对于企业的应用而言,正确理解用户的意图是提升用户的满意度的关键。在语音交互系统中,对语音识别结果进行用户概念表征,得到用户的意图,其中,用户概念表征是指,通过对输入的信息进行加工,将所感知的事物的本质特点加以表达。

[0003] 但是,传统的语音识别技术仅从发音、语法的层面进行建模,导致语音识别结果存在不准确的问题,从而影响对用户意图识别的准确率。

发明内容

[0004] 基于此,有必要针对上述技术问题,提供一种能够提高对用户意图的识别的准确率的语音识别文本的纠错方法、装置、计算机设备和存储介质。

[0005] 一种语音识别文本的纠错方法,所述方法包括:

[0006] 利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用场景的语料,所述第二语料库包括预设场景的语料;

[0007] 若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词;

[0008] 从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。

[0009] 在其中一个实施例中,所述纠错数据库的构建方式包括:

[0010] 获取所述第二语料库的语料;

[0011] 利用分词词典对所述第二语料库的语料进行分词,得到候选词;

[0012] 根据所述候选词以及所述候选词的拼音构建所述纠错数据库。

[0013] 在其中一个实施例中,所述方法还包括:

[0014] 获取所述候选词对应的混淆词;

[0015] 将所述混淆词加入所述分词词典。

[0016] 在其中一个实施例中,所述获取所述语音识别文本中的待纠错词,包括:

[0017] 利用所述分词词典对所述语音识别文本进行分词,得到文本词;

[0018] 计算各个所述文本词的平均绝对偏差值;

[0019] 若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为所述待纠错词。

[0020] 在其中一个实施例中,所述从纠错数据库中确定所述待纠错词对应的纠正词包括:

- [0021] 从所述纠错数据库中确定所述待纠错词对应的纠错候选词；
- [0022] 在所述纠错候选词中确定所述纠正词。
- [0023] 在其中一个实施例中，所述从所述纠错数据库中确定所述待纠错词对应的纠错候选词，包括：
- [0024] 获取所述待纠错词的拼音；
- [0025] 获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度；
- [0026] 将所述相似度大于相似阈值的候选词作为所述纠错候选词。
- [0027] 在其中一个实施例中，所述获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度，包括：
- [0028] 获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的编辑距离，利用所述编辑距离表征所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度。
- [0029] 在其中一个实施例中，所述在所述纠错候选词中确定所述纠正词，包括：
- [0030] 利用所述纠错候选词替换所述语音识别文本中的待纠错词，并利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度；
- [0031] 将所述流畅度满足预设条件的纠错候选词作为所述纠正词。
- [0032] 在其中一个实施例中，所述预设的语言模型为二元语言模型和三元语言模型；
- [0033] 所述利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度，包括：
- [0034] 将替换后的所述语音识别文本分别输入所述二元语言模型以及所述三元语言模型，得到所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度；
- [0035] 将所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度中的最大值，作为所述语音识别文本的流畅度。
- [0036] 一种语音识别文本的纠错装置，所述装置包括：
- [0037] 获取模块，用于利用预设的语言模型获取语音识别文本的流畅度，其中，所述预设的语言模型利用第一语料库和第二语料库的语料训练得到，所述第一语料库包括通用场景的语料，所述第二语料库包括预设场景的语料；
- [0038] 所述获取模块，还用于若所述语音识别文本的流畅度小于流畅阈值，获取所述语音识别文本中的待纠错词；
- [0039] 确定模块，用于从纠错数据库中确定所述待纠错词对应的纠正词，根据所述纠正词得到纠正后的语音识别文本。
- [0040] 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：
- [0041] 利用预设的语言模型获取语音识别文本的流畅度，其中，所述预设的语言模型利用第一语料库和第二语料库的语料训练得到，所述第一语料库包括通用场景的语料，所述第二语料库包括预设场景的语料；
- [0042] 若所述语音识别文本的流畅度小于流畅阈值，获取所述语音识别文本中的待纠错词；
- [0043] 从纠错数据库中确定所述待纠错词对应的纠正词，根据所述纠正词得到纠正后的语音识别文本。
- [0044] 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：

[0045] 利用预设的语言模型获取语音识别文本的流畅度，其中，所述预设的语言模型利用第一语料库和第二语料库的语料训练得到，所述第一语料库包括通用场景的语料，所述第二语料库包括预设场景的语料；

[0046] 若所述语音识别文本的流畅度小于流畅阈值，获取所述语音识别文本中的待纠错词；

[0047] 从纠错数据库中确定所述待纠错词对应的纠正词，根据所述纠正词得到纠正后的语音识别文本。

[0048] 上述语音识别文本的纠错方法、装置、计算机设备和存储介质，利用预设的语言模型获取语音识别文本的流畅度，若语音识别文本的流畅度小于流畅阈值，获取语音识别文本中的待纠错词，从纠错数据库中确定待纠错词对应的纠正词，根据纠正词得到纠正后的语音识别文本，这样，对语音识别文本中的错误词进行检测和纠正，提高了对语音识别文本的识别的准确性，并且，利用第二语料库训练预设的语言模型和构建纠错数据库，对用户概念表征进行区分和增强，进而提高了对用户意图识别的准确率。

附图说明

[0049] 图1为一个实施例中语音识别文本的纠错方法的应用环境图；

[0050] 图2为一个实施例中语音识别文本的纠错方法的流程示意图；

[0051] 图3为一个实施例中语音识别文本的纠错方法的作用示意图；

[0052] 图4为一个实施例中纠错数据库的示意图；

[0053] 图5为另一个实施例中语音识别文本的纠错方法的流程示意图；

[0054] 图6为一个实施例中语音识别文本的纠错装置的结构框图；

[0055] 图7为另一个实施例中语音识别文本的纠错装置的结构框图；

[0056] 图8为一个实施例中计算机设备的内部结构图。

具体实施方式

[0057] 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。

[0058] 本申请提供的语音识别文本的纠错方法，可以应用于如图1所示的应用环境中。终端102或者服务器104利用预设的语言模型获取语音识别文本的流畅度，其中，所述预设的语言模型利用第一语料库和第二语料库的语料训练得到，所述第一语料库包括通用场景的语料，所述第二语料库包括预设场景的语料；若所述语音识别文本的流畅度小于流畅阈值，获取所述语音识别文本中的待纠错词；从纠错数据库中确定所述待纠错词对应的纠正词，根据所述纠正词得到纠正后的语音识别文本。

[0059] 其中，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。

[0060] 在一个实施例中，如图2所示，提供了一种语音识别文本的纠错方法，以该方法应

用于图1中的终端为例进行说明,包括以下步骤:

[0061] 步骤202,利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用场景的语料,所述第二语料库包括预设场景的语料。

[0062] 其中,通用场景的语料是指各个领域里普遍应用的语料,第一语料库包括通用场景的语料,在一个实施例中,第一语料库可选用维基百科数据集,维基百科数据集包括5000万条符合通用场景的正确表达。

[0063] 预设场景的语料是指各个领域里特定场景应用的语料,该特定场景可以是工作场景,比如财务(财务指标查询、费用报销、企业经营数据查询)、审批(出差审批、请假审批)、采购(商品采购)、管理(人力资源管理)等。第二语料库包括预设场景的语料,在一个实施例中,第二语料库可选用各个领域里工作场景的交互语料。由于每个领域对应有专业知识,这些专业知识对用户概念表征起到重要作用,因此通过第二语料库可加强对用户概念的解读。

[0064] 语音识别文本是指,基于输入的语音,识别得到的文本数据。由于自然语言的多样性、复杂性,以及方言习惯不同,不同的用户对同一事物的表达可能不同,由此识别得到的文本数据也可能不同。比如,输入的语音为“仓库的库存剩下多少”,语音识别文本可能为“仓库的库存剩下多少”,也可能为“仓库的库存省下多少”。

[0065] 预设的语言模型是指为句子中每个词之间的前后关系建立的数学模型,其考虑至少两个词之间的前后关系,即下一个词的出现仅依赖于它前面的一个或几个词。预设的语言模型包括二元语言模型、三元语言模型、...、n元语言模型中的至少一个。

[0066] 如图3所示,预设的语言模型利用第一语料库和第二语料库中的语料训练得到。具体地,通过语言模型训练工具,利用第一语料库和第二语料库中的语料训练得到预设的语言模型。其中,语言模型训练工具可以是SRILM、IRSTLM、BerkeleyLM和KenLM等。

[0067] 以训练二元语言模型为例,统计第一语料库和第二语料库中相邻两个词一起出现的概率,并存储统计结果。为了简化计算,概率可取10为底的对数值,例如,“我们公司”可存储为“我们公司-1.25”。为了提升存储效率,存储文件可被转换为二进制文件。

[0068] 具体地,首先采用预设的语言模型对语音识别文本进行检错。将语音识别文本输入预设的语言模型,得到语音识别文本的流畅度,通过流畅度判断语音识别文本是否存在错误,其中,若流畅度小于流畅阈值,判定该语音识别文本存在错误,需要对语音识别文本进行纠正。

[0069] 在一个实施例中,预设的语言模型为二元语言模型和三元语言模型,二元语言模型和三元语言模型均通过第一语料库和第二语料库中的语料训练得到。将语音识别文本输入预设的语言模型,得到两个流畅度,若两个流畅度中的最大值小于流畅阈值,判定该语音识别文本存在错误。

[0070] 步骤204,若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词。

[0071] 其中,流畅度阈值用于判定语音识别文本是否存在错误,流畅度阈值可根据实际应用进行设定。若语音识别文本的流畅度大于或等于流畅阈值,判定该语音识别文本正确;若流畅度小于流畅阈值,判定该语音识别文本存在错误,需要对语音识别文本进行纠正。

[0072] 待纠错词是指语音识别文本中错误的文本词。在一个实施例中,对语音识别文本进行分词(利用分词工具,比如结巴分词工具),得到文本词。计算各个文本词的平均绝对偏差值,若一个文本词的平均绝对偏差大于偏差阈值,则判定该文本词错误,将该文本词作为待纠错词;若一个文本词的平均绝对偏差小于或等于偏差阈值,则判定该文本词正确。

[0073] 步骤206,从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。

[0074] 其中,纠错数据库用于对语音识别文本中的待纠错词进行替换。

[0075] 在一个实施例中,如图3所示,纠错数据库可通过第二语料库构建得到。获取第二语料库的语料,利用分词词典对第二语料库的语料进行分词,得到候选词,根据候选词以及候选词的拼音构建纠错数据库。

[0076] 在另一个实施例中,纠错数据库可通过第一语料库和第二语料库构建得到。获取第一语料库和第二语料库的语料,利用分词词典对第一语料库和第二语料库的语料进行分词,得到候选词,根据候选词以及候选词的拼音构建纠错数据库。

[0077] 其中,分词词典中存储了大量的词,其用于分词操作。当得到候选词时,获取候选词对应的混淆词,并将混淆词加入分词词典,以丰富分词词典的资源。

[0078] 具体地,从纠错数据库中确定待纠错词对应的纠错候选词,进一步的,在纠错候选词中确定纠正词。

[0079] 从纠错数据库中确定待纠错词对应的纠错候选词的方式可以是:获取待纠错词的拼音,获取待纠错词的拼音与纠错数据库中的候选词的拼音的相似度,将相似度大于相似阈值的候选词作为纠错候选词。

[0080] 在纠错候选词中确定纠正词的方式可以是:利用纠错候选词替换语音识别文本中的待纠错词,并利用预设的语言模型计算替换后的语音识别文本的流畅度,将流畅度满足预设条件的纠错候选词作为纠正词。在一个实施例中,将流畅度中的最大值对应的纠错候选词作为纠正词。

[0081] 具体地,利用纠正词替换语音识别文本中的待纠错词,得到纠正后的语音识别文本。

[0082] 上述语音识别文本的纠错方法中,利用预设的语言模型获取语音识别文本的流畅度,若语音识别文本的流畅度小于流畅阈值,获取语音识别文本中的待纠错词,从纠错数据库中确定待纠错词对应的纠正词,根据纠正词得到纠正后的语音识别文本,这样,对语音识别文本中的错误词进行检测和纠正,提高了对语音识别文本的识别的准确性,并且,利用第二语料库训练预设的语言模型和构建纠错数据库,对用户概念表征进行区分和增强,进而提高了对用户意图识别的准确率。

[0083] 在一个实施例中,所述纠错数据库的构建方式包括:获取所述第二语料库的语料;利用分词词典对所述第二语料库的语料进行分词,得到候选词;根据所述候选词以及所述候选词的拼音构建所述纠错数据库。

[0084] 其中,候选词是指第二语料库的语料中包含的词。

[0085] 具体地,利用分词工具(比如结巴分词工具)和分词词典对第二语料库的语料进行分词,得到候选词。比如,“我们公司的账户可以提取多少现金”,通过分词工具分词得到“我们”、“公司”、“的”、“账户”、“可以”、“提取”、“多少”、“现金”。

[0086] 获取每个候选词的拼音,并将候选词与拼音关联存储。在一个实施例中,如图4所示,词与拼音以键值对的方式进行存储。

[0087] 上述语音识别文本的纠错方法中,根据第二语料库构建纠错数据库,实现对用户概念表征进行区分和增强。

[0088] 在一个实施例中,所述方法还包括:获取所述候选词对应的混淆词;将所述混淆词加入所述分词词典。

[0089] 其中,混淆词是指与候选词的发音接近或者相同的词。

[0090] 分词词典中存储了大量的词,其用于分词操作。当得到候选词时,获取候选词对应的混淆词,并将混淆词加入分词词典,以丰富分词词典的资源,提高对语音识别文本的分词的准确度。

[0091] 具体地,对每一个候选词中的每个字进行替换,得到候选词对应的混淆词。在一个实施例中,利用字级别混淆集对候选词中的每个字进行替换。比如,“现金”对应的混淆词可以是“先进”、“线进”、“现近”、“现进”等。

[0092] 上述语音识别文本的纠错方法中,将混淆词加入分词词典,丰富了分词词典的资源,提高对语音识别文本的分词的准确度。

[0093] 在一个实施例中,所述获取所述语音识别文本中的待纠错词,包括:利用所述分词词典对所述语音识别文本进行分词,得到文本词;计算各个所述文本词的平均绝对偏差值;若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为所述待纠错词。

[0094] 其中,待纠错词是指语音识别文本中错误的词;文本词是指语音识别文本中的词;偏差阈值用于判定文本词是否错误,偏差阈值可根据实际应用进行设定。

[0095] 具体地,利用分词工具(比如结巴分词工具)和分词词典对语音识别文本进行分词,得到文本词。计算各个文本词的平均绝对偏差值,若一个文本词的平均绝对偏差大于偏差阈值,则判定该文本词错误,将该文本词作为待纠错词;若一个文本词的平均绝对偏差小于或等于偏差阈值,则判定该文本词正确。

[0096] 上述语音识别文本的纠错方法中,通过文本词的平均绝对偏差值判断文本词是否存在错误,提高了纠错的准确度。

[0097] 在一个实施例中,所述从纠错数据库中确定所述待纠错词对应的纠正词包括:从所述纠错数据库中确定所述待纠错词对应的纠错候选词;在所述纠错候选词中确定所述纠正词。

[0098] 其中,纠错候选词是指,可能用于纠正所述待纠错词的词的集合。比如,语音识别文本为“我们公司的账户可以提取多少线进”,“线进”为待纠错词,那么纠错候选词可能为“先进”、“现金”等。

[0099] 具体地,从所述纠错数据库中确定所述待纠错词对应的纠错候选词的方式可以是:通过拼音相似度从纠错数据库中确定纠错候选词。其中,拼音相似度可通过拼音的编辑距离进行确定。

[0100] 具体地,在纠错候选词中确定纠正词的方式可以是:利用纠错候选词替换语音识别文本中的待纠错词,并利用预设的语言模型计算替换后的语音识别文本的流畅度,将流畅度满足预设条件的纠错候选词作为纠正词。

[0101] 上述语音识别文本的纠错方法中,可以结合拼音的编辑距离和预设的语言模型对

语音识别文本进行纠错,进一步提高了选取纠正词的准确率。

[0102] 在一个实施例中,所述从所述纠错数据库中确定所述待纠错词对应的纠错候选词,包括:获取所述待纠错词的拼音;获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度;将所述相似度大于相似阈值的候选词作为所述纠错候选词。

[0103] 由于纠错数据库中存储有候选词以及候选词的拼音,可通过比对待纠错词的拼音与纠错数据库中的候选词的拼音的相似度来确定纠错候选词。

[0104] 具体地,待纠错词的拼音与纠错数据库中的候选词的拼音的相似度,可通过计算待纠错词的拼音与候选词的拼音的编辑距离来确定,其中,编辑距离是用来度量两个序列相似程度的指标。以拼音为例,拼音的编辑距离是指在两个拼音之间,由其中一个拼音转换为另一个拼音所需要的最少字符编辑操作次数。待纠错词的拼音与候选词的拼音之间的编辑距离越小,待纠错词与候选词之间的相似度越大,因此将相似度大于相似阈值的候选词作为纠错候选词。

[0105] 在一个实施例中,拼音的编辑距离的计算方式如下:

$$[0106] \quad distance(t_0, t_i) = |len(t_0) - len(t_i)| * \left(\sum_{w \in t_0} len_p(w) + \sum_{w \in t_i} len_p(w) \right) / (len(t_0) + len(t_i))$$

[0107] 其中, t_0 和 t_i 分别为待纠错词和纠错数据库中的候选词, $len(x)$ 为词 x 中包含的字的个数, $len_p(x)$ 为词 x 的拼音中包含的字符的个数。

[0108] 上述语音识别文本的纠错方法中,利用拼音的编辑距离确定纠错候选词,提高了选取纠错候选词的准确率。

[0109] 在一个实施例中,所述在所述纠错候选词中确定所述纠正词,包括:利用所述纠错候选词替换所述语音识别文本中的待纠错词,并利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度;将所述流畅度满足预设条件的纠错候选词作为所述纠正词。

[0110] 其中,预设的语言模型用于计算一个句子流畅的概率。预设的语言模型可为 N 元语言模型, N 可为一、二、三、四等。 N 元语言模型是指在一个句子中,对于其中一个位置,根据该位置的前 $N-1$ 个词,计算每个待选词在该位置时,该句子流畅的概率。预设的语言模型也可为至少两个 N 元语言模型的组合,比如预设的语言模型可为二元语言模型和三元语言模型等。

[0111] 具体地,利用纠错候选词替换语音识别文本中的待纠错词,利用预设的语言模型计算替换后的语音识别文本的流畅度,通过流畅度确定纠正词。利用预设的语言模型计算替换后的语音识别文本的流畅度的方式为:将替换后的语音识别文本输入预设的语言模型,得到替换后的语音识别文本的流畅度。

[0112] 其中,预设条件用于在纠错候选词中筛选纠正词,可根据实际应用进行设定。在一个实施例中,将预设的语言模型输出的替换后的语音识别文本的流畅度中,最大的流畅度对应的纠错候选词作为纠正词。

[0113] 上述语音识别文本的纠错方法中,通过预设的语言模型确定纠正词,提高了选取纠正词的准确率。

[0114] 在一个实施例中,所述预设的语言模型为二元语言模型和三元语言模型;所述利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度,包括:将替换后的所述语音识别文本分别输入所述二元语言模型以及所述三元语言模型,得到所述二元语言模型

输出的流畅度以及所述三元语言模型输出的流畅度;将所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度中的最大值,作为所述语音识别文本的流畅度。

[0115] 其中,二元语言模型和三元语言模型均通过第一语料库和第二语料库中的语料训练得到。

[0116] 具体地,将替换后的语音识别文本输入预设的语言模型,得到二元语言模型输出的流畅度以及三元语言模型输出的流畅度,将两个流畅度中的最大值作为语音识别文本的流畅度。

[0117] 上述语音识别文本的纠错方法中,通过二元语言模型和三元语言模型确定纠正词,提高了选取纠正词的准确率。

[0118] 如图5所示,对一实施例中的语音识别文本的纠错方法进行详细介绍:

[0119] 步骤502,利用预设的语言模型获取语音识别文本的流畅度;

[0120] 步骤504,若所述语音识别文本的流畅度小于流畅阈值,利用分词词典对所述语音识别文本进行分词,得到文本词;

[0121] 步骤506,计算各个所述文本词的平均绝对偏差值;

[0122] 步骤508,若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为待纠错词;

[0123] 步骤510,获取所述待纠错词的拼音;

[0124] 步骤512,获取所述待纠错词的拼音与纠错数据库中的候选词的拼音的相似度;

[0125] 步骤514,将所述相似度大于相似阈值的候选词作为所述纠错候选词;

[0126] 步骤516,利用所述纠错候选词替换所述语音识别文本中的待纠错词,并利用预设的语言模型计算替换后的所述语音识别文本的流畅度;

[0127] 步骤518,将所述流畅度满足预设条件的纠错候选词作为所述待纠错词的纠正词。

[0128] 上述语音识别文本的纠错方法中,利用预设的语言模型获取语音识别文本的流畅度,若语音识别文本的流畅度小于流畅阈值,获取语音识别文本中的待纠错词,从纠错数据库中确定待纠错词对应的纠正词,根据纠正词得到纠正后的语音识别文本,这样,对语音识别文本中的错误词进行检测和纠正,提高了对语音识别文本的识别的准确性,并且,利用第二语料库训练预设的语言模型和构建纠错数据库,对用户概念表征进行区分和增强,进而提高了对用户意图识别的准确率。

[0129] 应该理解的是,虽然图2和图5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和图5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。

[0130] 在一个实施例中,如图6所示,提供了一种语音识别文本的纠错装置600,包括:获取模块602和确定模块604,其中:

[0131] 获取模块602,用于利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用

场景的语料,所述第二语料库包括预设场景的语料;

[0132] 所述获取模块602,还用于若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词;

[0133] 确定模块604,用于从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。

[0134] 上述语音识别文本的纠错装置600,利用预设的语言模型获取语音识别文本的流畅度,若语音识别文本的流畅度小于流畅阈值,获取语音识别文本中的待纠错词,从纠错数据库中确定待纠错词对应的纠正词,根据纠正词得到纠正后的语音识别文本,这样,对语音识别文本中的错误词进行检测和纠正,提高了对语音识别文本的识别的准确性,并且,利用第二语料库训练预设的语言模型和构建纠错数据库,对用户概念表征进行区分和增强,进而提高了对用户意图的识别的准确率。

[0135] 在一个实施例中,如图7所示,所述识别文本的纠错装置600还包括分词模块606和构建模块608,其中,所述获取模块602,还用于获取所述第二语料库的语料;所述分词模块606,用于利用分词词典对所述第二语料库的语料进行分词,得到候选词;所述构建模块608,用于根据所述候选词以及所述候选词的拼音构建所述纠错数据库。

[0136] 在一个实施例中,所述识别文本的纠错装置600还包括加入模块,所述获取模块602,还用于获取所述候选词对应的混淆词;所述加入模块,用于将所述混淆词加入所述分词词典。

[0137] 在一个实施例中,所述获取模块602,还用于利用所述分词词典对所述语音识别文本进行分词,得到文本词;计算各个所述文本词的平均绝对偏差值;若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为所述待纠错词。

[0138] 在一个实施例中,所述确定模块604,还用于从所述纠错数据库中确定所述待纠错词对应的纠错候选词;在所述纠错候选词中确定所述纠正词。

[0139] 在一个实施例中,所述确定模块604,还用于获取所述待纠错词的拼音;获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度;将所述相似度大于相似阈值的候选词作为所述纠错候选词。

[0140] 在一个实施例中,所述确定模块604,还用于获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的编辑距离,利用所述编辑距离表征所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度。在一个实施例中,所述确定模块604,还用于利用所述纠错候选词替换所述语音识别文本中的待纠错词,并利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度;将所述流畅度满足预设条件的纠错候选词作为所述纠正词。

[0141] 在一个实施例中,所述确定模块604,还用于将替换后的所述语音识别文本分别输入所述二元语言模型以及所述三元语言模型,得到所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度;将所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度中的最大值,作为所述语音识别文本的流畅度。关于语音识别文本的纠错装置的具体限定可以参见上文中对于语音识别文本的纠错方法的限定,在此不再赘述。上述语音识别文本的纠错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或者终端,其内部结构图可以如图8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、网络接口和数据库。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识别文本的纠错数据。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识别文本的纠错方法。

[0143] 本领域技术人员可以理解,图8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。

[0144] 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,存储器中存储有计算机程序,该处理器执行计算机程序时实现以下步骤:

[0145] 利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用场景的语料,所述第二语料库包括预设场景的语料;

[0146] 若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词;

[0147] 从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。

[0148] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:

[0149] 获取所述第二语料库的语料;

[0150] 利用分词词典对所述第二语料库的语料进行分词,得到候选词;

[0151] 根据所述候选词以及所述候选词的拼音构建所述纠错数据库。

[0152] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:

[0153] 获取所述候选词对应的混淆词;

[0154] 将所述混淆词加入所述分词词典。

[0155] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:

[0156] 利用所述分词词典对所述语音识别文本进行分词,得到文本词;

[0157] 计算各个所述文本词的平均绝对偏差值;

[0158] 若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为所述待纠错词。

[0159] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:

[0160] 从所述纠错数据库中确定所述待纠错词对应的纠错候选词;

[0161] 在所述纠错候选词中确定所述纠正词。

[0162] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:

[0163] 获取所述待纠错词的拼音;

[0164] 获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度;

- [0165] 将所述相似度大于相似阈值的候选词作为所述纠错候选词。
- [0166] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:
- [0167] 获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的编辑距离,利用所述编辑距离表征所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度。
- [0168] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:
- [0169] 利用所述纠错候选词替换所述语音识别文本中的待纠错词,并利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度;
- [0170] 将所述流畅度满足预设条件的纠错候选词作为所述纠正词。
- [0171] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:
- [0172] 所述利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度,包括:
- [0173] 将替换后的所述语音识别文本分别输入所述二元语言模型以及所述三元语言模型,得到所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度;
- [0174] 将所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度中的最大值,作为所述语音识别文本的流畅度。
- [0175] 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,计算机程序被处理器执行时实现以下步骤:
- [0176] 利用预设的语言模型获取语音识别文本的流畅度,其中,所述预设的语言模型利用第一语料库和第二语料库的语料训练得到,所述第一语料库包括通用场景的语料,所述第二语料库包括预设场景的语料;
- [0177] 若所述语音识别文本的流畅度小于流畅阈值,获取所述语音识别文本中的待纠错词;
- [0178] 从纠错数据库中确定所述待纠错词对应的纠正词,根据所述纠正词得到纠正后的语音识别文本。
- [0179] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:
- [0180] 获取所述第二语料库的语料;
- [0181] 利用分词词典对所述第二语料库的语料进行分词,得到候选词;
- [0182] 根据所述候选词以及所述候选词的拼音构建所述纠错数据库。
- [0183] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:
- [0184] 获取所述候选词对应的混淆词;
- [0185] 将所述混淆词加入所述分词词典。
- [0186] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:
- [0187] 利用所述分词词典对所述语音识别文本进行分词,得到文本词;
- [0188] 计算各个所述文本词的平均绝对偏差值;
- [0189] 若所述文本词的平均绝对偏差值大于偏差阈值,则判定所述文本词为所述待纠错词。
- [0190] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:
- [0191] 从所述纠错数据库中确定所述待纠错词对应的纠错候选词;
- [0192] 在所述纠错候选词中确定所述纠正词。
- [0193] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:

- [0194] 获取所述待纠错词的拼音；
- [0195] 获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度；
- [0196] 将所述相似度大于相似阈值的候选词作为所述纠错候选词。
- [0197] 在一个实施例中，计算机程序被处理器执行时还实现以下步骤：
- [0198] 获取所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的编辑距离，利用所述编辑距离表征所述待纠错词的拼音与所述纠错数据库中的候选词的拼音的相似度。
- [0199] 在一个实施例中，计算机程序被处理器执行时还实现以下步骤：
- [0200] 利用所述纠错候选词替换所述语音识别文本中的待纠错词，并利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度；
- [0201] 将所述流畅度满足预设条件的纠错候选词作为所述纠正词。
- [0202] 在一个实施例中，计算机程序被处理器执行时还实现以下步骤：
- [0203] 所述利用所述预设的语言模型计算替换后的所述语音识别文本的流畅度，包括：
- [0204] 将替换后的所述语音识别文本分别输入所述二元语言模型以及所述三元语言模型，得到所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度；
- [0205] 将所述二元语言模型输出的流畅度以及所述三元语言模型输出的流畅度中的最大值，作为所述语音识别文本的流畅度。
- [0206] 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取存储介质中，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/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。
- [0207] 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。
- [0208] 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。

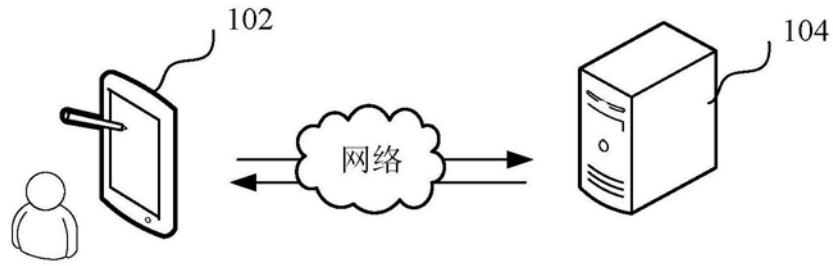


图1

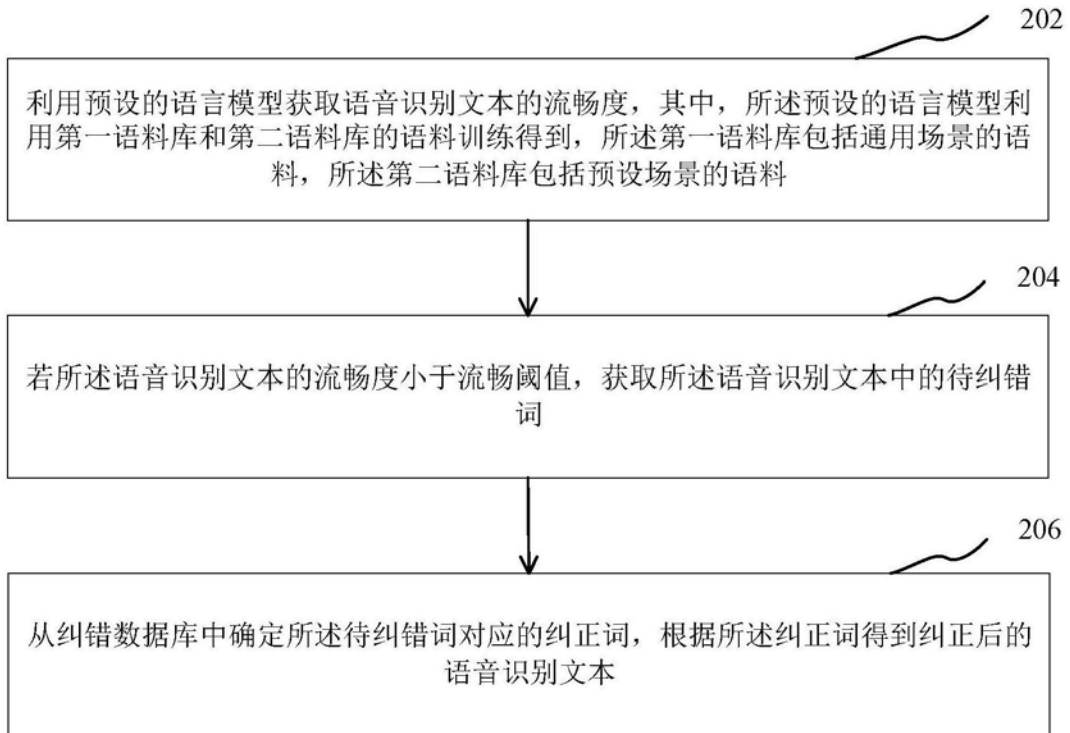


图2

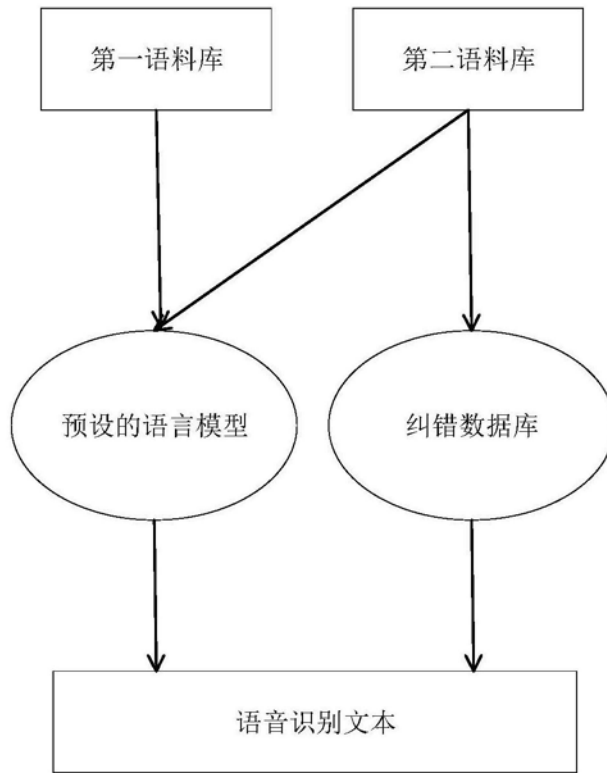


图3

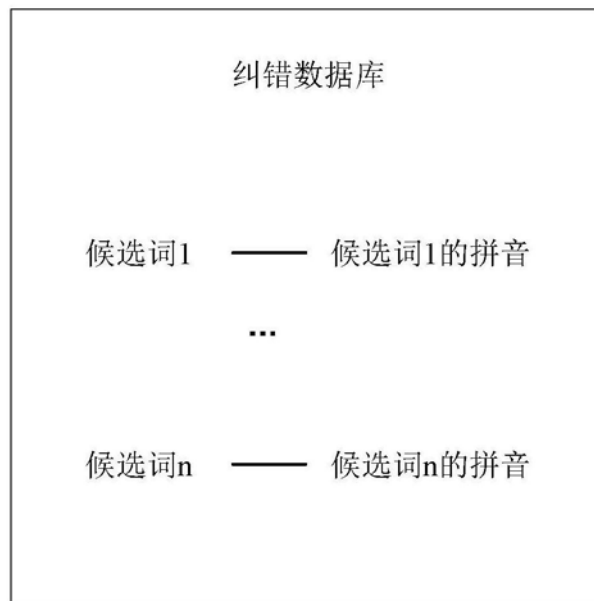


图4

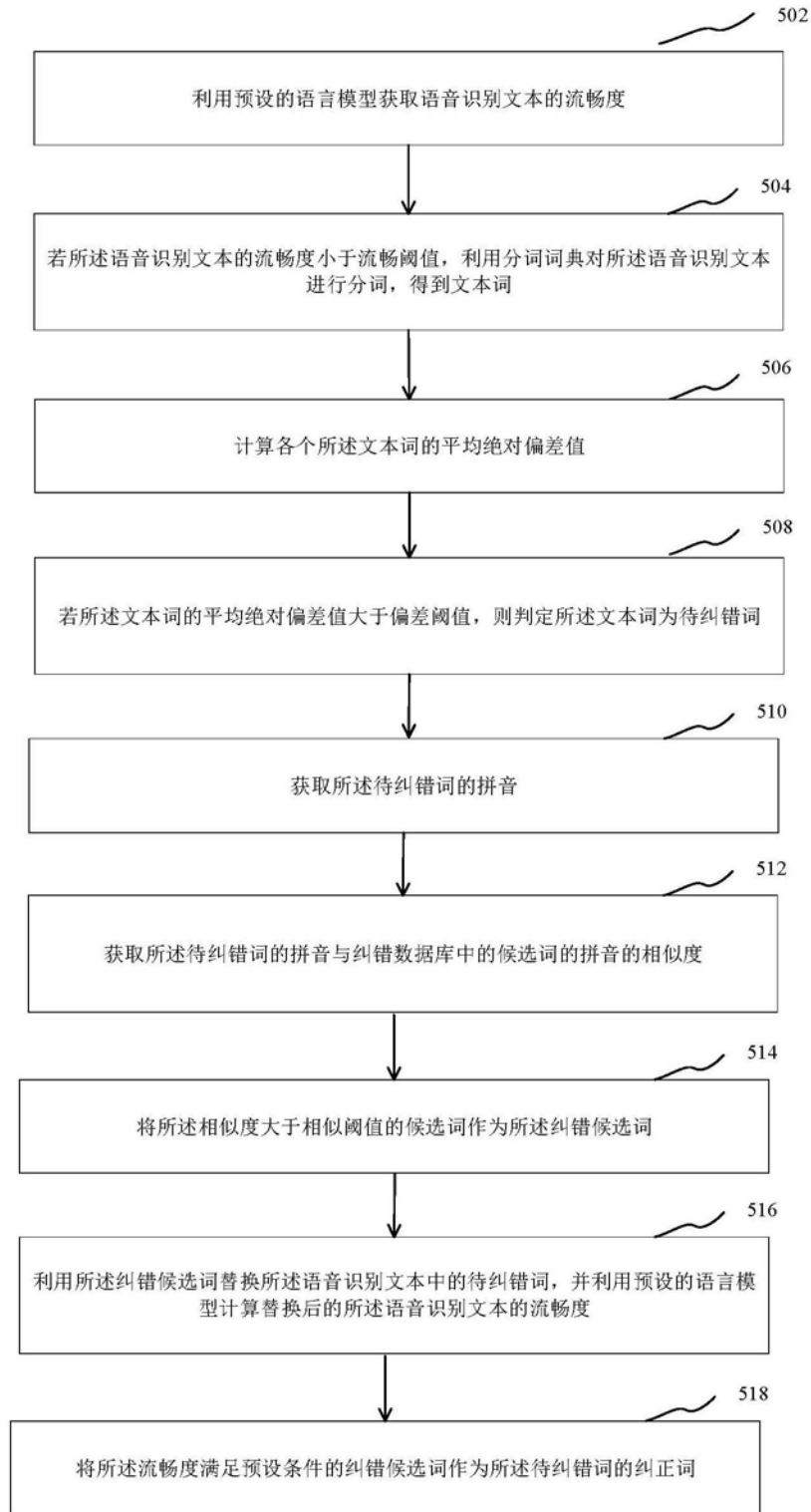


图5

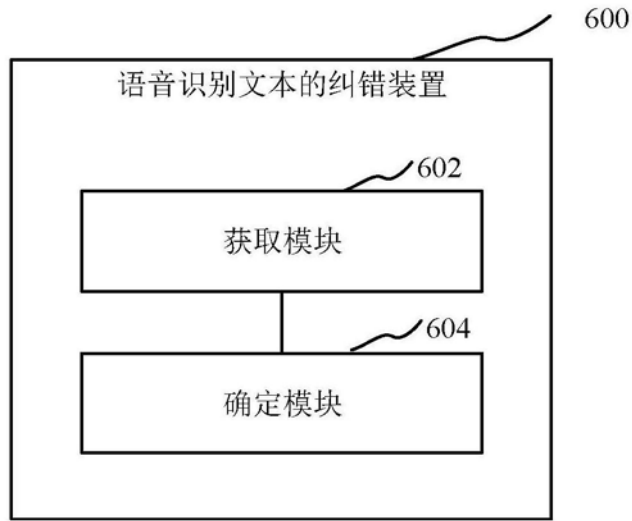


图6

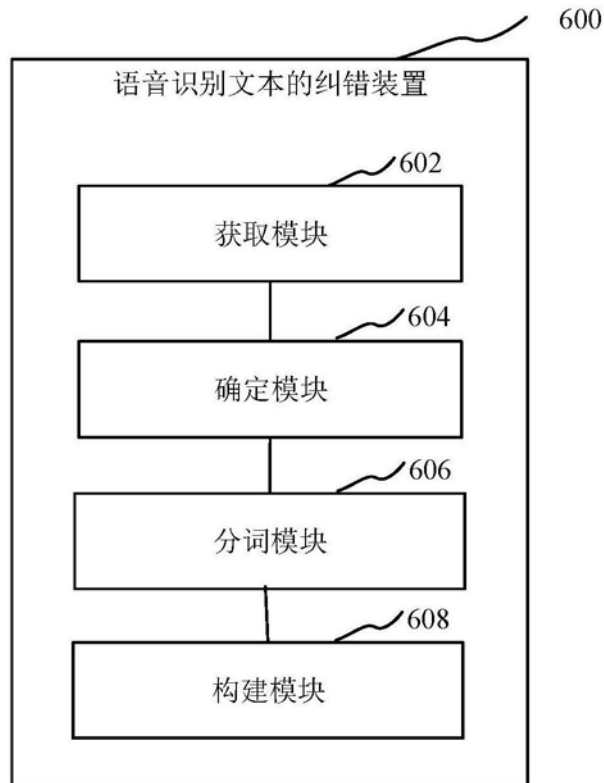


图7

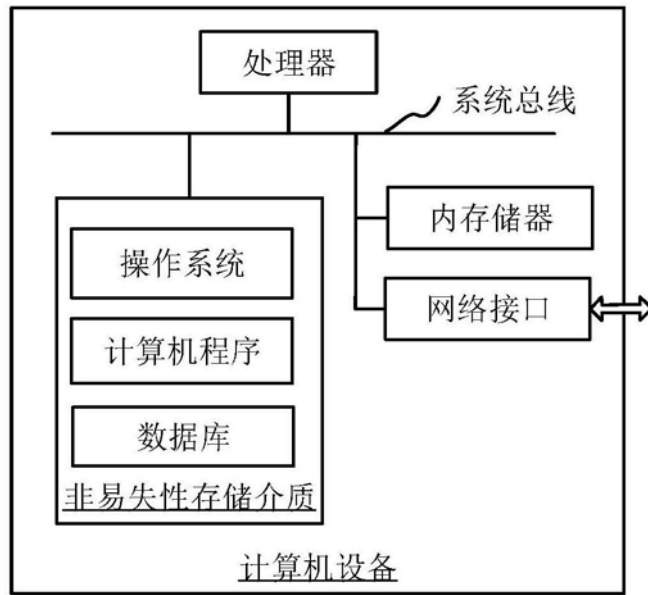


图8