



(12) 发明专利

(10) 授权公告号 CN 111710328 B

(45) 授权公告日 2024. 01. 12

(21) 申请号 202010547282.2

G10L 15/26 (2006.01)

(22) 申请日 2020.06.16

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 111710328 A

CN 107357778 A, 2017.11.17

CN 102999533 A, 2013.03.27

CN 106777073 A, 2017.05.31

(43) 申请公布日 2020.09.25

CN 108847241 A, 2018.11.20

(73) 专利权人 北京爱医声科技有限公司
地址 100095 北京市海淀区学清路10号院1
号楼A座19层1903

CN 110046350 A, 2019.07.23

CN 110162681 A, 2019.08.23

CN 110188353 A, 2019.08.30

(72) 发明人 陶焜

CN 110232114 A, 2019.09.13

CN 110767215 A, 2020.02.07

(74) 专利代理机构 北京鸿元知识产权代理有限公司 11327

CN 111275046 A, 2020.06.12

US 2017025117 A1, 2017.01.26

专利代理师 李玉琦 曹素云

US 2018286386 A1, 2018.10.04

(51) Int. Cl.

审查员 罗薇

G10L 15/06 (2013.01)

G10L 15/08 (2006.01)

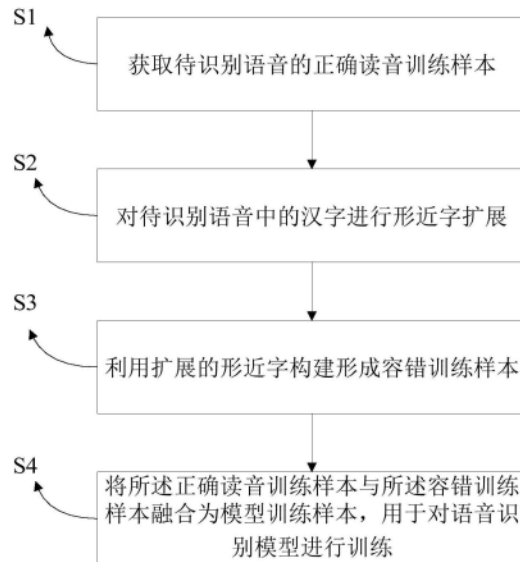
权利要求书2页 说明书7页 附图2页

(54) 发明名称

语音识别模型的训练样本选取方法、装置及介质

(57) 摘要

本发明公开了一种语音识别模型的训练样本选取方法、装置及介质,其中,方法包括:获取待识别语音的正确读音训练样本;对待识别语音中的汉字进行形近字扩展;利用扩展的形近字构建形成容错训练样本;将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,用于对语音识别模型进行训练。本发明对训练样本进行读音容错处理,提高了语音识别系统的容错性,达到即使在读音错误的情况下输入法也能提供正确候选词的目的,从而提升语音输入的准确率和实用性。



1. 一种语音识别模型的训练样本选取方法,其特征在于,包括:
 获取待识别语音的正确读音训练样本;
 对待识别语音中的汉字进行形近字扩展;
 利用扩展的形近字构建形成容错训练样本;
 将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,用于对语音识别模型进行训练,

通过去掉部首或替换部首的方式对待识别语音中的汉字进行形近字扩展,将潜在的错误读音也加入训练样本进行训练,对于与目标汉字完全同音的形近字,在分组时不予考虑,所述容错训练样本的权重通过下式获取:

$$W_{\text{near}}(a) = W_0 * \left(\sum_{i=1}^n P_i \right)$$

其中, $W_{\text{near}}(a)$ 表示容错训练样本的权重, a 表示容错训练样本中的形近字发音, W_0 表示预设的固定权重, n 表示容错训练样本中发音为 a 的形近字的数量, P_i 表示容错训练样本中发音为 a 的第 i 个形近字在汉语语料库中的单字出现概率。

2. 根据权利要求1所述的语音识别模型的训练样本选取方法,其特征在于,利用扩展的形近字构建形成容错训练样本的步骤包括:

根据发音对扩展的形近字进行分组,将相同发音的形近字分为一组;
 每组均对应生成一个容错训练样本。

3. 根据权利要求1所述的语音识别模型的训练样本选取方法,其特征在于,对语音识别模型进行训练时,设定容错训练样本的权重小于正确读音训练样本的权重。

4. 根据权利要求3所述的语音识别模型的训练样本选取方法,其特征在于,所述容错训练样本的权重由模型训练方法和容错训练样本中形近字在汉语语料库中的单字出现频率共同决定。

5. 一种语音识别模型的训练样本选取装置,其特征在于,包括:
 第一样本获取模块,用于获取待识别语音的正确读音训练样本;
 扩展模块,用于对待识别语音中的汉字进行形近字扩展;
 第二样本获取模块,用于利用扩展的形近字构建形成容错训练样本;
 样本融合模块,用于将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,所述模型训练样本用于对语音识别模型进行训练,

通过去掉部首或替换部首的方式对待识别语音中的汉字进行形近字扩展,将潜在的错误读音也加入训练样本进行训练,对于与目标汉字完全同音的形近字,在分组时不予考虑,所述容错训练样本的权重通过下式获取:

$$W_{\text{near}}(a) = W_0 * \left(\sum_{i=1}^n P_i \right)$$

其中, $W_{\text{near}}(a)$ 表示容错训练样本的权重, a 表示容错训练样本中的形近字发音, W_0 表示预设的固定权重, n 表示容错训练样本中发音为 a 的形近字的数量, P_i 表示容错训练样本中发音为 a 的第 i 个形近字在汉语语料库中的单字出现概率。

6. 一种电子装置,其特征在于,该电子装置包括:

至少一个处理器;

与至少一个处理器通信连接的存储器,所述存储器中包括训练样本选取程序,所述训练样本选取程序被处理器执行时实现如权利要求1至4中任一项所述的训练样本选取方法。

7. 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中包括训练样本选取程序,所述训练样本选取程序被处理器执行时,实现如权利要求1至4中任一项所述的训练样本选取方法。

语音识别模型的训练样本选取方法、装置及介质

技术领域

[0001] 本发明涉及人工智能技术领域,具体地,涉及一种语音识别模型的训练样本选取方法、装置及介质。

背景技术

[0002] 随着语音识别技术的飞速发展,智能语音输入法正逐渐成为文字录入的常见选择,被越来越多地应用于不同行业的众多场景中。智能语音输入法以语音识别技术为核心,主要包含特征提取、声学模型、语言模型以及字典与解码等部分,通过提取待识别语音数据的声学特征,基于声学模型将其解码为音素阵列,利用字典和语言模型输出相对应文字。这种策略以汉字的正确读音为基础和前提,如果读音错误,则很难得到正确的文字输出。而在实际使用场景中,汉字读音错误的情况十分普遍。对于不熟悉的汉字,人们通常会采取“读半边”或“读形近字”的方法,如“妊娠”中的“娠”、“齟齬”中的“齟”、“头孢克肟”中的“肟”等等。错误的读音导致语音输入法中的语音识别模型无法正确识别用户真正想要输入的汉字,使语音录入的效果和体验都大打折扣。

发明内容

[0003] 鉴于以上问题,本发明的目的是提供一种语音识别模型的训练样本选取方法、装置及介质,以解决当前智能语音输入法中的语音识别模型由于读音错误而无法正确识别输入文字的问题。

[0004] 为了实现上述目的,本发明的第一个方面是提供一种语音识别模型的训练样本选取方法,包括:

[0005] 获取待识别语音的正确读音训练样本;

[0006] 对待识别语音中的汉字进行形近字扩展;

[0007] 利用扩展的形近字构建形成容错训练样本;

[0008] 将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,用于对语音识别模型进行训练。

[0009] 优选地,通过去掉部首或替换部首的方式对待识别语音中的汉字进行形近字扩展。

[0010] 优选地,利用扩展的形近字构建形成容错训练样本的步骤包括:

[0011] 根据发音对扩展的形近字进行分组,将相同发音的形近字分为一组;

[0012] 每一个发音分组均对应生成一个容错训练样本。

[0013] 优选地,对语音识别模型进行训练时,设定容错训练样本的权重小于正确读音训练样本的权重。

[0014] 优选地,所述容错训练样本的权重由模型训练方法和容错训练样本中形近字在汉语语料库中的单字出现频率共同决定。

[0015] 优选地,所述容错训练样本的权重通过下式获取:

$$[0016] \quad W_{\text{near}}(a) = W_0 * \left(\sum_{i=1}^n P_i \right)$$

[0017] 其中, $W_{\text{near}}(a)$ 表示容错训练样本的权重, a 表示容错训练样本中的形近字发音, W_0 表示预设的固定权重, n 表示容错训练样本中发音为 a 的形近字的数量, P_i 表示容错训练样本中发音为 a 的第 i 个形近字在汉语语料库中的单字出现概率。

[0018] 为了实现上述目的, 本发明的第二个方面是提供一种语音识别模型的训练样本选取装置, 包括:

[0019] 第一样本获取模块, 用于获取待识别语音的正确读音训练样本;

[0020] 扩展模块, 用于对待识别语音中的汉字进行形近字扩展;

[0021] 第二样本获取模块, 用于利用扩展的形近字构建形成容错训练样本;

[0022] 样本融合模块, 用于将所述正确读音训练样本与所述容错训练样本融合为模型训练样本, 所述模型训练样本用于对语音识别模型进行训练。

[0023] 为了实现上述目的, 本发明的第三个方面是提供一种电子装置, 该电子装置包括: 至少一个处理器; 与至少一个处理器通信连接的存储器, 所述存储器中包括训练样本选取程序, 所述训练样本选取程序被处理器执行时实现如上所述的训练样本选取方法。

[0024] 为了实现上述目的, 本发明的第四个方面是提供一种计算机可读存储介质, 所述计算机可读存储介质中包括训练样本选取程序, 所述训练样本选取程序被处理器执行时, 实现如上所述的训练样本选取方法。

[0025] 与现有技术相比, 本发明具有以下优点和有益效果:

[0026] 本发明在对语音识别模型的训练样本选取时, 通过对待识别语音中的汉字进行形近字扩展, 实现读音容错处理, 提高了语音识别系统的容错性, 解决当前智能语音输入法中由于读音错误而无法正确识别的问题, 达到即使在读音错误的情况下输入法也能提供正确候选词的目的, 从而提升语音输入的准确率和实用性。

附图说明

[0027] 图1是本发明所述语音识别模型的训练样本选取方法的流程示意图;

[0028] 图2是本发明所述语音识别模型的训练样本选取装置的模块示意图。

具体实施方式

[0029] 下面将参考附图来描述本发明所述的实施例。本领域的普通技术人员可以认识到, 在不偏离本发明的精神和范围的情况下, 可以用各种不同的方式或其组合对所描述的实施例进行修正。因此, 附图和描述在本质上是说明性的, 而不是用于限制权利要求的保护范围。此外, 在本说明书中, 附图未按比例画出, 并且相同的附图标记表示相同的部分。

[0030] 图1是本发明所述语音识别模型的训练样本选取方法的流程示意图, 如图1所示, 本发明所述语音识别模型的训练样本选取方法, 包括:

[0031] 步骤S1, 获取待识别语音的正确读音训练样本, 其中, 正确读音训练样本为现有对语音识别模型进行训练时所确定的训练样本;

[0032] 步骤S2, 对待识别语音中的汉字进行形近字扩展;

[0033] 步骤S3,利用扩展的形近字构建形成容错训练样本;

[0034] 步骤S4,将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,用于对语音识别模型进行训练。

[0035] 本发明通过对训练样本进行容错处理,提高了语音识别模型的识别正确率。尤其是,在应用于智能语音输入法中进行文字录入时,可以在语音发音错误的情况下,录入正确的文字,提升语音输入的准确率。

[0036] 在步骤S2中,其主要是对冷僻汉字及易读错汉字进行形近字扩展,优选地,通过去掉部首或替换部首的方式对待识别语音中的汉字进行形近字扩展。当然也可以是其他扩展形式,只要是用户可能读错的发音均可。例如,常见抗生素“头孢克肟”中的“肟”(发音为“wo4”,后面的4表示声调是四声),可以基于去掉/替换部首的方式扩展其形近字为“亏”和“污”。对于一般的语音输入法而言,仅会将正确的读音样本{“头孢克肟”:[‘tou2’, ‘baol’, ‘ke4’, ‘wo4’]}作为训练样本进行训练,而本发明通过形近字扩展将潜在的错误读音也加入训练样本进行训练。例如,用户容易误将“肟”读成“亏”的发音‘kui1’或“污”的发音‘wu1’,利用扩展的形近字构建形成容错训练样本为{“头孢克肟”:[‘tou2’, ‘baol’, ‘ke4’, ‘kui1’]}和{“头孢克肟”:[‘tou2’, ‘baol’, ‘ke4’, ‘wu1’]}。再如,常见口腔疾病“龋齿”中的“龋”(念“qu3”)容易误读成“禹”的发音‘yu3’,利用本发明对训练样本进行选取时,扩展加入{“龋齿”:[‘yu3’, ‘chi3’]}的容错训练样本进行训练。同理,药品名“二甲双胍”的“胍”(念“gual”),容易误将“胍”读成“孤”的发音‘gu1’,因此,扩展加入{“二甲双胍”:[‘er4’, ‘jia3’, ‘shuang1’, ‘gu1’]}的容错训练样本进行训练。

[0037] 需要说明的是,本发明中对待识别语音进行形近字扩展时,不仅仅限于对冷僻汉字或易读错汉字进行扩展,也可以根据语音识别需求,对其他汉字进行扩展,对扩展的形近字进行自由组合,以扩大智能语音输入的应用范围。例如,对于“头孢克肟”,不仅可以对“肟”进行形近字扩展,形成上述形式的容错训练样本,还可以对“孢”进行形近字扩展,并对“肟”和“孢”的扩展发音进行自由组合,形成多个容错训练样本,如{“头孢克肟”:[‘tou2’, ‘pao2’, ‘ke4’, ‘kui1’]}和{“头孢克肟”:[‘tou2’, ‘pao2’, ‘ke4’, ‘wo4’]}等。

[0038] 当待识别语音中的目标汉字有多个形近字可供扩展时,利用扩展的形近字构建形成容错训练样本的步骤包括:根据发音对扩展的形近字进行分组,将相同发音的形近字分为一组;每一个发音分组均对应生成一个容错训练样本。其中,对于与目标汉字完全同音的形近字,在分组时不予考虑。例如,“肟”字除了正确读音‘wo4’外,其常见的形近字包括“污”(wu1)、“圪”(wu1)、“亏”(kui1)和“夸”(kua1),这四个字具有三种发音,对应到三个发音分组(wu1)、(kui1)和(kua1),形成相应的三个容错训练样本。

[0039] 考虑到错误读音的出现概率低于正确读音的出现概率,对语音识别模型进行训练时,设定容错训练样本的样本采样权重小于正确读音训练样本的样本采样权重。假设目标词语的正确读音训练样本的权重为 $w_{right}=1$,则扩展后的易错发音的容错训练样本的权重 $w_{near}<1$ 。

[0040] 进一步地,所述容错训练样本的权重由模型训练方法和容错训练样本中形近字在汉语语料库中的单字出现频率共同决定。其中,本发明对模型训练方法并无具体限定,根据具体的语音识别模型确定。对于待识别语音中的某个目标汉字有多个形近字可供扩展的情况,形成一个或多个发音分组。当某一个发音分组中有多个形近字时,该发音分组的出现频

率为组内各字的出现频率之和。发音分组的出现频率越高，误读为该发音的概率越高，因此对应的容错训练样本的权重 W_{near} 越高。

[0041] 具体地，所述容错训练样本的权重通过下式获取：

$$[0042] \quad W_{near}(a) = W_0 * \left(\sum_{i=1}^n P_i \right)$$

[0043] 其中， $W_{near}(a)$ 表示容错训练样本的权重， a 表示容错训练样本中的形近字发音， W_0 表示预设的固定权重， n 表示容错训练样本中发音为 a 的形近字的数量， P_i 表示容错训练样本中发音为 a 的第 i 个形近字在汉语语料库中的单字出现概率。

[0044] 例如，“脔”字除了正确读音‘wo4’外，其常见的形近字包括“污”(wu1)、“圪”(wu1)、“亏”(kui1)和“夸”(kua1)四个字，对应到三个发音分组(wu1)、(kui1)和(kua1)。统计可得这四个字在汉语语料库中的单字出现概率分别为“污”(P₁)，“圪”(P₂)，“亏”(P₃)，“夸”(P₄)。由此可得三个错误发音分组的出现概率分别为(wu1)：(P₁+P₂)、(kui1)：(P₃)和(kua1)：(P₄)。对应的容错训练样本权重分别计算为 $W_{near}(wu1) = W_0 * (P_1 + P_2)$ ， $W_{near}(kui1) = W_0 * (P_3)$ 和 $W_{near}(kua1) = W_0 * (P_4)$ 。其中， W_0 为预设的固定权重，根据经验设定，用于保证 W_{near} 的值小于 W_{right} 但又能产生足够的训练效果。基于扩展形近字的发音分组和其对应的权重 W_{near} ，生成容错训练样本并用于语音识别模型的训练，即可得到具有读音容错能力的语音识别模型，提高语音输入的智能识别准确率。

[0045] 需要说明的是，本发明仅仅是对训练样本进行选取，不会改变既有语音识别模型的模型结构或训练超参数，仅在语音识别模型的训练样本准备阶段起作用。在为语音识别模型准备训练样本时，在正常的正确读音训练样本之外，扩展生成一部分容错训练样本一并加入训练，权重 W_{near} 越高的样本有越大的概率被选中加入训练。由于容错训练样本的数量远低于正确读音训练样本的数量，可以保证在提高容错能力的同时不会影响正常发音文本的识别。

[0046] 在智能语音输入法中，当用户读完一句话后，输入法会通过发音声学模型识别到多组相应的候选词，再通过上下文的语言模型选出概率最大的候选词。对于现有的训练样本选取方式得到的语音识别模型而言，当用户的句子中包括“tou2”，‘baol’，‘ke4’，‘wu1’”的发音时，候选词只有“头孢克圪”，即使通过上下文的语言模型也无法实现正确的候选词识别。但利用本发明的训练样本选取方式训练得到的语音识别模型，候选词将包括“头孢克圪”、“头孢克脔”等更多选项，让用户实际想表达的(但是读错了的)正确的候选词能够出现在这个语言模型阶段。通过语言模型结合候选词前后的上下文信息，从多个候选词中选出概率最大的结果，从而可以以较大概率成功识别出用户真实意图想输入的词语“头孢克脔”。因此，通过本发明的训练样本选取方法，对训练样本进行容错处理，使得语音输入法在用户读音错误的情况下仍然可以给出正确的候选字词。

[0047] 图2是本发明所述语音识别模型的训练样本选取装置的模块示意图，如图2所示，本发明所述语音识别模型的训练样本选取装置，包括：

[0048] 第一样本获取模块1，用于获取待识别语音的正确读音训练样本；

[0049] 扩展模块2，用于对待识别语音中的汉字进行形近字扩展；

[0050] 第二样本获取模块3，用于利用扩展的形近字构建形成容错训练样本；

[0051] 样本融合模块4,用于将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,所述模型训练样本用于对语音识别模型进行训练。

[0052] 扩展模块2主要是对冷僻汉字及易读错汉字进行形近字扩展,优选地,通过去掉部首或替换部首的方式对待识别语音中的汉字进行形近字扩展。例如,常见抗生素“头孢克肟”中的“肟”(发音为“wo4”,后面的4表示声调是四声),可以基于去掉/替换部首的方式扩展其形近字为“亏”和“污”。对于一般的语音输入法而言,仅会将正确的读音样本{“头孢克肟”:[‘tou2’, ‘baol’, ‘ke4’, ‘wo4’]}作为训练样本进行训练,而本发明通过形近字扩展将潜在的错误读音也加入训练样本进行训练。例如,用户容易误将“肟”读成“亏”的发音‘kui1’或“污”的发音‘wu1’,利用扩展的形近字构建形成容错训练样本为{“头孢克肟”:[‘tou2’, ‘baol’, ‘ke4’, ‘kui1’]}和{“头孢克肟”:[‘tou2’, ‘baol’, ‘ke4’, ‘wu1’]}。再如,常见口腔疾病“龋齿”中的“龋”(念“qu3”)容易误读成“禹”的发音‘yu3’,利用本发明对训练样本进行选取时,扩展加入{“龋齿”:[‘yu3’, ‘chi3’]}的容错训练样本进行训练。同理,药品名“二甲双胍”的“胍”(念“gual”),容易误将“胍”读成“孤”的发音‘gu1’,因此,扩展加入{“二甲双胍”:[‘er4’, ‘jia3’, ‘shuang1’, ‘gu1’]}的容错训练样本进行训练。

[0053] 当待识别语音中的目标汉字有多个形近字可供扩展时,第二样本获取模块3通过下述方式构建形成容错训练样本:根据发音对扩展的形近字进行分组,将相同发音的形近字分为一组;每一个发音分组均对应生成一个容错训练样本。其中,对于与目标汉字完全同音的形近字,在分组时不予考虑。例如,“肟”字除了正确读音‘wo4’外,其常见的形近字包括“污”(wu1)、“圪”(wu1)、“亏”(kui1)和“夸”(kua1),这四个字具有三种发音,对应到三个发音分组(wu1)、(kui1)和(kua1),形成相应的三个容错训练样本。

[0054] 考虑到错误读音的出现概率低于正确读音的出现概率,对语音识别模型进行训练时,设定容错训练样本的样本采样权重小于正确读音训练样本的样本采样权重。假设目标词语的正确读音训练样本的权重为 $W_{\text{right}}=1$,则扩展后的易错发音的容错训练样本的权重 $W_{\text{near}} < 1$ 。

[0055] 进一步地,所述容错训练样本的权重由模型训练方法和容错训练样本中形近字在汉语语料库中的单字出现频率共同决定。其中,本发明对模型训练方法并无具体限定,根据具体的语音识别模型确定。对于待识别语音中的某个目标汉字有多个形近字可供扩展的情况,形成一个或多个发音分组。当某一个发音分组中有多个形近字时,该发音分组的出现频率为组内各字的出现频率之和。发音分组的出现频率越高,误读为该发音的概率越高,因此对应的容错训练样本的权重 W_{near} 越高。

[0056] 具体地,所述容错训练样本的权重通过下式获取:

$$[0057] \quad W_{\text{near}}(\mathbf{a}) = W_0 * \left(\sum_{i=1}^n P_i \right)$$

[0058] 其中, $W_{\text{near}}(\mathbf{a})$ 表示容错训练样本的权重, \mathbf{a} 表示容错训练样本中的形近字发音, W_0 表示预设的固定权重, n 表示容错训练样本中发音为 \mathbf{a} 的形近字的数量, P_i 表示容错训练样本中发音为 \mathbf{a} 的第 i 个形近字在汉语语料库中的单字出现概率。

[0059] 本发明所述语音识别模型的训练样本选取方法应用于电子装置,所述电子装置可以是电视机、智能手机、平板电脑、计算机等终端设备。

[0060] 所述电子装置包括：至少一个处理器，与至少一个处理器通信连接的存储器，所述存储器中包括训练样本选取程序，所述训练样本选取程序被处理器执行时实现如下所述的训练样本选取方法的步骤：

[0061] 获取待识别语音的正确读音训练样本；

[0062] 对待识别语音中的汉字进行形近字扩展；

[0063] 利用扩展的形近字构建形成容错训练样本；

[0064] 将所述正确读音训练样本与所述容错训练样本融合为模型训练样本，用于对语音识别模型进行训练。

[0065] 所述电子装置还可以包括网络接口和通信总线等。其中，网络接口可以包括标准的有线接口、无线接口，通信总线用于实现各个组件之间的连接通信。

[0066] 存储器包括至少一种类型的可读存储介质，可以是闪存、硬盘、光盘等非易失性存储介质，也可以是插接式硬盘等，且并不限于此，可以是非暂时性方式存储指令或软件以及任何相关联的数据文件并向处理器提供指令或软件程序以使该处理器能够执行指令或软件程序的任何装置。本发明中，存储器存储的软件程序包括训练样本选取程序，并可以向处理器提供该训练样本选取程序，以使得处理器可以执行该训练样本选取程序，实现训练样本选取方法。

[0067] 处理器可以是中央处理器、微处理器或其他数据处理芯片等，可以运行存储器中的存储程序，例如，本发明中训练样本选取程序。

[0068] 所述电子装置还可以包括显示器，显示器也可以称为显示屏或显示单元。在一些实施例中显示器可以是LED显示器、液晶显示器、触控式液晶显示器以及有机发光二极管(Organic Light-Emitting Diode, OLED)触摸器等。显示器用于显示在电子装置中处理的信息以及用于显示可视化的工作界面。

[0069] 所述电子装置还可以包括用户接口，用户接口可以包括输入单元(比如键盘)、语音输出装置(比如音响、耳机)等。

[0070] 需要说明的是，本发明之电子装置的具体实施方式与上述语音识别模型的训练样本选取方法、装置的具体实施方式大致相同，在此不再赘述。

[0071] 在其他实施例中，语音识别模型的训练样本选取程序还可以被分割为一个或者多个模块，一个或者多个模块被存储于存储器中，并由处理器执行，以完成本发明。本发明所称的模块是指能够完成特定功能的一系列计算机程序指令段。所述训练样本选取程序可以被分割为：第一样本获取模块1、扩展模块2、第二样本获取模块3和样本融合模块4。上述模块所实现的功能或操作步骤均与上文类似，此处不再详述，示例性地，例如其中：

[0072] 第一样本获取模块1，用于获取待识别语音的正确读音训练样本；

[0073] 扩展模块2，用于对待识别语音中的汉字进行形近字扩展；

[0074] 第二样本获取模块3，用于利用扩展的形近字构建形成容错训练样本；

[0075] 样本融合模块4，用于将所述正确读音训练样本与所述容错训练样本融合为模型训练样本，所述模型训练样本用于对语音识别模型进行训练。

[0076] 本发明的一个实施例中，计算机可读存储介质可以是任何包含或存储程序或指令的有形介质，其中的程序可以被执行，通过存储的程序指令相关的硬件实现相应的功能。例如，计算机可读存储介质可以是计算机磁盘、硬盘、随机存取存储器、只读存储器等。本发明

并不限于此,可以是以非暂时性方式存储指令或软件以及任何相关数据文件或数据结构并且可提供给处理器以使处理器执行其中的程序或指令的任何装置。所述计算机可读存储介质中包括训练样本选取程序,所述训练样本选取程序被处理器执行时,实现如下的训练样本选取方法:

[0077] 获取待识别语音的正确读音训练样本;

[0078] 对待识别语音中的汉字进行形近字扩展;

[0079] 利用扩展的形近字构建形成容错训练样本;

[0080] 将所述正确读音训练样本与所述容错训练样本融合为模型训练样本,用于对语音识别模型进行训练。

[0081] 需要说明的是,本发明之计算机可读存储介质的具体实施方式与上述语音识别模型的训练样本选取方法、装置的具体实施方式大致相同,在此不再赘述。

[0082] 需要说明的是,在本文中,术语“包括”、“包含”或者其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、装置、物品或者方法不仅包括那些要素,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、装置、物品或者方法所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

[0083] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在如上所述的一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等等)执行本发明各个实施例所述的方法。

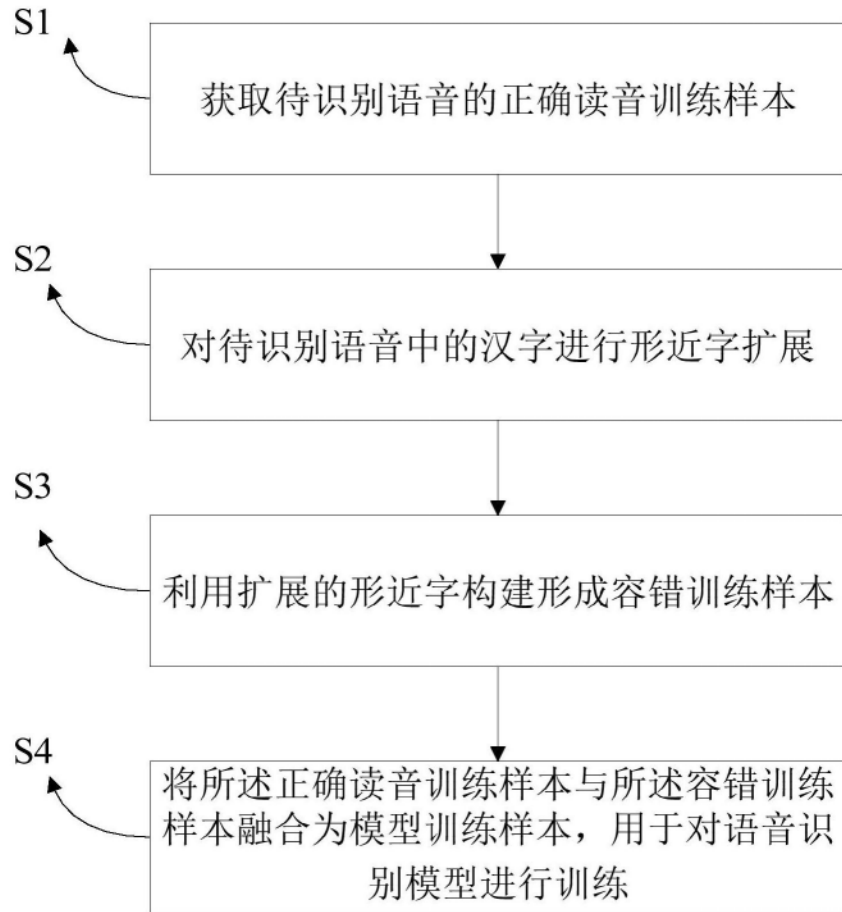


图1



图2