



(12)发明专利申请

(10)申请公布号 CN 108304385 A

(43)申请公布日 2018.07.20

(21)申请号 201810135187.4

(22)申请日 2018.02.09

(71)申请人 叶伟

地址 200126 上海市浦东新区耀华路560号
204室

(72)发明人 叶伟

(74)专利代理机构 上海伯瑞杰知识产权代理有
限公司 31227

代理人 孟旭彤

(51) Int. Cl.

G06F 17/27(2006.01)

G06F 17/30(2006.01)

G06K 9/62(2006.01)

G10L 15/26(2006.01)

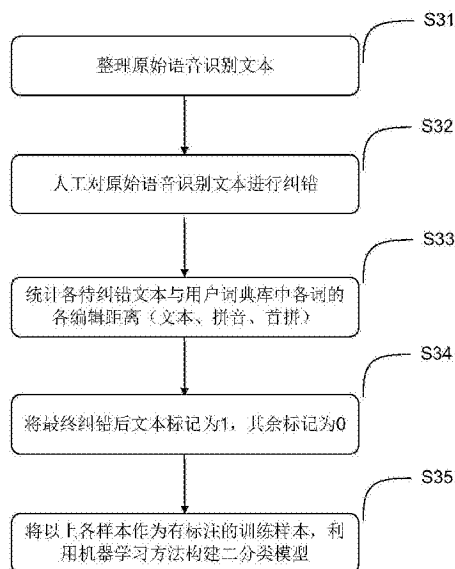
权利要求书2页 说明书3页 附图3页

(54)发明名称

一种语音识别文本纠错方法及装置

(57)摘要

一种语音识别文本纠错方法,该方法包括:生成用户词典库;生成候选待纠错文本集合;生成候选纠错后文本集合;以及对候选纠错后文本集合进行筛选,获得纠错后文本。本发明的实施例通过建立用户语料库的技术特征,构建出纠错后文本的判断模型,生成候选纠错后文本;并且建立用户语句预测模型,利用该模型,计算得到语音识别的纠错后文本,获得的有益效果之一是,在智能家居控制领域,为智能家电的语音识别控制的准确性提供了更进一步的实现方案。



1. 一种语音识别文本纠错方法,其特征在于,该方法包括:
生成用户词典库;
生成候选待纠错文本集合;
生成候选纠错后文本集合;以及
对候选纠错后文本集合进行筛选,获得纠错后文本。
2. 如权利要求1所述语音识别文本纠错方法,其特征在于,生成用户词典库的步骤包括:
收集用户文本语料,并进行整理;
对整理后的语料文本进行分词和去除停用词;
对分词结果进行词性标注、词频统计、拼音标注后构建用户词典库。
3. 如权利要求1所述语音识别文本纠错方法,其特征在于,生成候选待纠错文本集合的步骤包括:
对语音识别文本进行分词、去除停用词后,获得各分词结果;
检查各分词是否在用户词典库中,若在用户词典库中出现,则所述语音识别文本不作为候选待纠错文本,若不在用户词典库中,则所述语音识别文本作为候选待纠错文本。
4. 如权利要求1所述语音识别文本纠错方法,其特征在于,生成候选纠错后文本集合的步骤包括:
根据训练样本,生成判断用户词典库中各词是否是候选纠错后文本的判断模型,其中将编辑距离作为模型输入特征;
将所有候选纠错后文本替换掉待纠错文本,从而生成候选纠错后文本。
5. 如权利要求1所述语音识别文本纠错方法,其特征在于,对候选纠错后文本集合进行筛选后获得纠错后文本的步骤包括:
将统计获得的用户数据,作为原始输入参数,将用户语料意图及实体作为输出,构建用户语句预测模型;
利用用户语句预测模型,将语音识别文本的包括语音交互时间的原始参数输入,从而获得预测得到的用户意图及实体;
将预测所得到的用户意图及实体与候选纠错后文本进行相似度计算,得到相似度最高的候选纠错后文本作为唯一纠错后文本。
6. 如权利要求4所述语音识别文本纠错方法,其特征在于,候选纠错后文本判断模型的构建步骤包括:
整理原始语音识别文本;
人工对原始语音识别文本进行纠错,并记录下各纠错后文本;
统计各待纠错后文本与用户词典库各词编辑距离,此编辑距离包括文本编辑距离、首拼编辑距离、全拼音编辑距离,并将各编辑距离进行归一化,作为模型的输入;
标注用户词典库中各词是否是最终纠正后文本。若是,则标注为1,若不是则标注为0;
根据各标注样本,利用相关机器学习算法,生成判断是否是候选纠错后文本的模型。
7. 一种语音识别文本纠错方法装置,其特征在于,该装置包括存储器;以及
耦合到所述存储器的一个或多个处理器,处理器被配置为执行存储在所述存储器中的指令,所述处理器执行以下操作:

生成用户词典库；
生成候选待纠错文本集合；
生成候选纠错后文本集合；以及
对候选纠错后文本集合进行筛选，获得纠错后文本。

8. 如权利要求7所述语音识别文本纠错装置，其特征在于，生成用户词典库的步骤包括：

收集用户文本语料，并进行整理；
对整理后的语料文本进行分词和去除停用词；
对分词结果进行词性标注、词频统计、拼音标注后构建用户词典库。

9. 如权利要求7所述语音识别文本纠错装置，其特征在于，生成候选待纠错文本集合的步骤包括：

对语音识别文本进行分词、去除停用词后，获得各分词结果；
检查各分词是否在用户词典库中，若在用户词典库中出现，则所述语音识别文本不作为候选待纠错文本，若不在用户词典库中，则所述语音识别文本作为候选待纠错文本。

10. 如权利要求7所述语音识别文本纠错装置，其特征在于，生成候选纠错后文本集合的步骤包括：

根据训练样本，生成判断用户词典库中各词是否是候选纠错后文本的判断模型，其中将编辑距离作为模型输入特征；

将所有候选纠错后文本替换掉待纠错文本，从而生成候选纠错后文本。

11. 如权利要求7所述语音识别文本纠错装置，其特征在于，对候选纠错后文本集合进行筛选后获得纠错后文本的步骤包括：

将统计获得的用户数据，作为原始输入参数，将用户语料意图及实体作为输出，构建用户语句预测模型；

利用用户语句预测模型，将语音识别文本的包括语音交互时间的原始参数输入，从而获得预测得到的用户意图及实体；

将预测所得到的用户意图及实体与候选纠错后文本进行相似度计算，得到相似度最高的候选纠错后文本作为唯一纠错后文本。

12. 如权利要求7所述语音识别文本纠错装置，其特征在于，候选纠错后文本判断模型的构建步骤包括：

整理原始语音识别文本；
人工对原始语音识别文本进行纠错，并记录下各纠错后文本；
统计各待纠错后文本与用户词典库各词编辑距离，此编辑距离包括文本编辑距离、首拼编辑距离、全拼音编辑距离，并将各编辑距离进行归一化，作为模型的输入；
标注用户词典库中各词是否是最终纠正后文本。若是，则标注为1，若不是则标注为0；
根据各标注样本，利用相关机器学习算法，生成判断是否是候选纠错后文本的模型。

一种语音识别文本纠错方法及装置

技术领域

[0001] 本发明属于语音识别技术领域,特别涉及一种语音识别文本纠错方法及装置。

背景技术

[0002] 随着深度学习的普及,在计算机视觉、语音识别、自然语言处理等方面均取得重大突破。以语音识别为例,目前语音识别准确率已达到97%。以上技术的突破,使得语音识别的应用领域越来越广。由于相对于其他人机交互方式,语音交互更为符合人们的日常习惯,更为高效。可以预计,语音识别技术将广泛应用于智能家居、工业生产、通信、医疗、自动驾驶等各个领域。

[0003] 在实际语音交互过程中,由于用户发音不标准、噪音等各因素影响,音识别错误率较高。而现有技术都集中在提升语音识别准确率上,却缺乏对识别结果的纠错手段。以上原因,极大影响语音交互产品推广。

发明内容

[0004] 本发明的实施例提供了一种语音识别文本纠错方法及装置,目的在于解决语音识别技术中对于语音识别文本的纠错问题。

[0005] 本发明的实施例之一,一种语音识别文本纠错方法,包括,

[0006] 用户词典库生成方法:统计用户家庭常用语料,并进行整理。根据整理后的语料文本,进行文本分词、词性标注、词频统计、拼音标注。

[0007] 候选待纠错文本生成方法:首先,对语音识别文本进行分词、去除停用词后,检查各词是否在用户词典库中。若在用户词典库中,则不进行纠错。若不在用户词典库中,则作为候选待纠错文本。

[0008] 候选纠错后文本生成方法:首先根据训练样本,生成判断该词是否是纠错后文本的判断模型,其中将编辑距离作为模型输入特征。将所有候选纠错后文本替换掉待纠错文本,从而生成候选纠错后文本。

[0009] 在智能家居控制方面,候选纠错后文本筛选方法:将所有原先统计的用户数据,包括用户ID、语音交互文本时间、传感器数据、家电状态等参数作为原始输入参数。将整理后所得的用户语料意图及实体作为输出,构建用户语句预测模型。利用该模型,将语音交互时时间、传感器数据、家电状态等参数输入,从而获得预测得到的用户意图及实体。将预测所得到的用户意图及实体与候选纠错后文本进行相似度计算,得到相似度最高的候选纠错后文本作为唯一纠错后文本。

[0010] 本发明的实施例通过建立用户语料库的技术特征,构建出纠错后文本的判断模型,生成候选纠错后文本;并且建立用户语句预测模型,利用该模型,计算得到语音识别的纠错后文本,获得的有益效果之一是,在智能家居控制领域,为智能家电的语音识别控制的准确性提供了更进一步的实现方案。

附图说明

[0011] 通过参考附图阅读下文的详细描述,本发明示例性实施方式的上述以及其他目的、特征和优点将变得易于理解。在附图中,以示例性而非限制性的方式示出了本发明的若干实施方式,其中:

[0012] 图1是本发明一个实施例中的基于用户语料库的语音识别纠错方法的流程图。

[0013] 图2是本发明一个实施例中用户词典库构建流程图。

[0014] 图3是本发明一个实施例中构建候选纠错后文本模型流程图。

具体实施方式

[0015] 根据一个或多个实施例,如图1所示,是语音识别结果文本纠错方法的流程示意图。流程步骤包括,

[0016] S11:构建用户词典库。将收集到的各用户文本语料进行清洗整理,并进行分词、去除停用词。将分词结果根据词频、词性、拼音为一个完整索引,构建用户词典库,并且其中各拼音用分割符分割。

[0017] S12:准候选待纠错文本生成。将用户语音交互文本经过去除停用词后,获得各分词结果。检查各词是否在用户词典库中出现,若在各用户词典库中出现,则不作为候选待纠错文本,反之则作为候选待纠错文本。检查时,针对分完后的各词一一在用户词典库中检查,若词在用户词典库中则无需纠错;若词不在用户词典库中才作为待纠错文本。

[0018] S13:候选纠错后文本生成。首先根据之前收集到的语料数据,收集整理各待纠正文本与纠正后文本的文本编辑距离、首字母编辑距离、全拼音编辑距离,并将以上各编辑距离进行归一化作为输入特征,进行候选纠错后文本生成模型。计算出各待纠错文本与用户词典库中各词的文本编辑距离、首字母编辑距离、全拼音编辑距离,并将其作为模型输入,得出用户词典库中哪些词作为候选纠错文本,哪些不作为候选纠错文本。如果,只有一个候选纠错后文本,则将其作为文本纠错后结果;如果有多个候选纠错后文本,则继续进行候选纠错后文本筛选。

[0019] S14:从各候选纠错后文本中筛选,挑选出最有可能的纠错后文本。该步骤中涉及到利用机器学习技术,构建基于用户环境的用户语料预测模型。根据用户产生语音交互的时间、地点、传感器数据、家电数据等维度构建用户语料预测模型。其输出主要包括意图、实体目标、操控方式。如输出:家电控制意图、实体为空调、操控方式为制热。将各纠错后文本与模型输出结果进行相似度计算,选取相似度最高的候选纠错后文本作为唯一纠错后文本。

[0020] 根据一个或多个实施例,如图2所示,用户词典库构建流程图。用户词典库生成方法包括:统计用户家庭常用语料,并进行整理。根据整理后的语料文本,进行文本分词、词性标注、词频统计、拼音标注,构建用户词典库。

[0021] 根据一个或多个实施例,如图3所示,构建候选纠错后文本判断模型流程示意图。

[0022] S31,整理原始语音识别文本,并进行脱敏;

[0023] S32,人工对原始语音识别文本进行纠错,并记录下各纠错后文本;

[0024] S33,统计各待纠错后文本与用户词典库各词编辑距离,此编辑距离包括文本编辑

距离、首拼编辑距离、全拼音编辑距离,并将各编辑距离进行归一化,作为模型的输入;

[0025] S34,标注词典库中各词是否是最终纠正后文本。若是,则标注为1,若不是则标注为0;

[0026] S35,根据以上各标注样本,利用相关机器学习算法,生成判断是否是候选纠错后文本的模型。

[0027] 根据一个或多个实施例,一种语音识别文本纠错方法装置,该装置包括存储器;以及耦合到所述存储器的一个或多个处理器,处理器被配置为执行存储在所述存储器中的指令,所述处理器执行以下操作:

[0028] 生成用户词典库;

[0029] 生成候选待纠错文本集合;

[0030] 生成候选纠错后文本集合;以及

[0031] 对候选纠错后文本集合进行筛选,获得纠错后文本。

[0032] 所述存储器包括计算机可读记录/存储介质,如随机存取存储器(RAM)、只读存储器(ROM)、闪存存储器、光盘、磁盘、固态盘等等。

[0033] 值得说明的是,虽然前述内容已经参考若干具体实施方式描述了本发明创造的精神和原理,但是应该理解,本发明并不限于所公开的具体实施方式,对各方面的划分也不意味着这些方面中的特征不能组合,这种划分仅是为了表述的方便。本发明旨在涵盖所附权利要求的精神和范围内所包括的各种修改和等同布置。

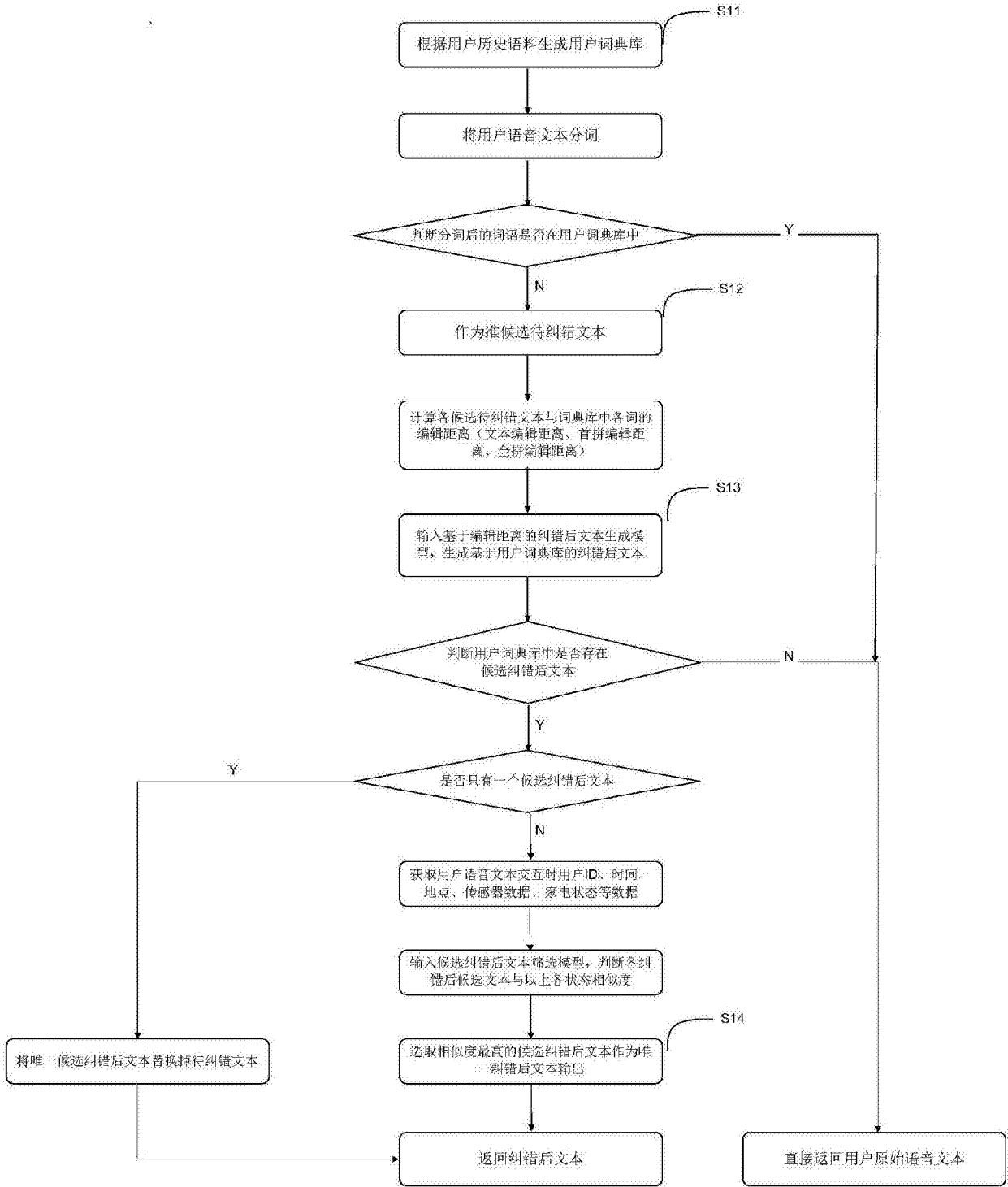


图1

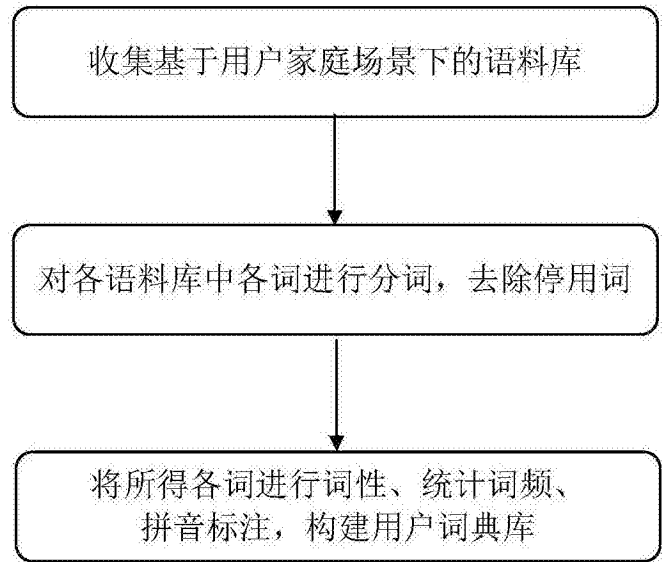


图2

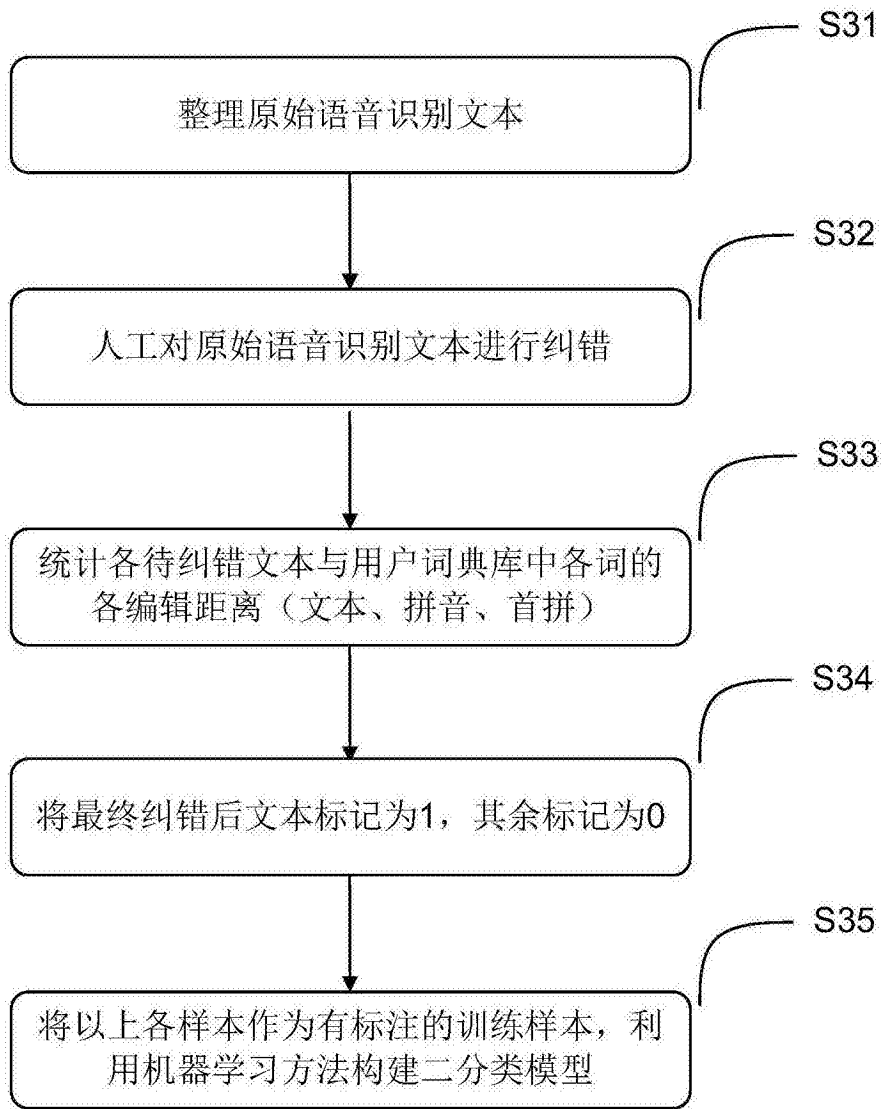


图3