



(12) 发明专利

(10) 授权公告号 CN 110164435 B

(45) 授权公告日 2024. 06. 25

(21) 申请号 201910342353.2

G10L 15/26 (2006.01)

(22) 申请日 2019.04.26

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 102254557 A, 2011.11.23

申请公布号 CN 110164435 A

CN 105895090 A, 2016.08.24

(43) 申请公布日 2019.08.23

审查员 谭宇玲

(73) 专利权人 平安科技(深圳)有限公司

地址 518033 广东省深圳市福田区福田街  
道福安社区益田路5033号平安金融中  
心23楼

(72) 发明人 魏韬 马骏 王少军

(74) 专利代理机构 北京市京大律师事务所

11321

专利代理师 居梦琪

(51) Int. Cl.

G10L 15/22 (2006.01)

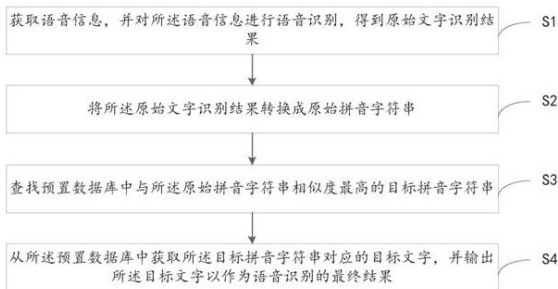
权利要求书2页 说明书9页 附图2页

(54) 发明名称

语音识别方法、装置、设备及计算机可读存储介质

(57) 摘要

本发明涉及语音交互技术领域,公开了一种语音识别方法、装置、设备及计算机可读存储介质,所述语音识别方法包括:获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;将所述原始文字识别结果转换成原始拼音字符串;查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。通过本发明,在得到语音识别结果后,进一步对语音识别结果进行纠错处理,避免了因音调识别错误而造成语音识别结果出错,提高了语音识别的准确率。



1. 一种语音识别方法,其特征在于,所述语音识别方法包括以下步骤:

获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;

将所述原始文字识别结果转换成原始拼音字符串;

创建业务主题对应的数据表;获取业务主题对应的历史业务提问记录,所述历史业务提问记录中包含若干个业务提问;获取每个业务提问对应的拼音字符串;将业务提问及对应的拼音字符串关联存储至所述数据表中;基于所述数据表,构建预置数据库;

显示业务主题选择界面,接收选择指令;基于所述选择指令,确定目标业务主题;根据所述目标业务主题,查找所述预置数据库中的对应目标数据表;计算所述原始拼音字符串与所述目标数据表中存储的各个拼音字符串的编辑距离;基于所述编辑距离,得到所述原始拼音字符串与所述各个拼音字符串的相似度;确定最高相似度对应的拼音字符串为目标字符串;

检测所述最高相似度是否大于或等于预设阈值;

若是,则从所述预置数据库中获取所述目标字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果;

若否,则显示文字输入界面,并接收基于所述文字输入界面输入的文字信息,并输出所述文字信息。

2. 如权利要求1所述的语音识别方法,其特征在于,在所述接收基于所述文字输入界面输入的文字信息,并输出所述文字信息之后,还包括:

获取所述文字信息对应的拼音字符串,并将所述文字信息以及所述文字信息对应的拼音字符串关联保存至所述目标数据表。

3. 一种语音识别装置,其特征在于,所述语音识别装置包括:

识别模块,用于获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;

转换模块,用于将所述原始文字识别结果转换成原始拼音字符串;

创建模块,用于创建业务主题对应的数据表;

获取模块,用于获取业务主题对应的历史业务提问记录,所述历史业务提问记录中包含若干个业务提问;获取每个业务提问对应的拼音字符串;

存储模块,用于将业务提问及对应的拼音字符串关联存储至所述数据表中;

构建模块,用于基于所述数据表,构建预置数据库;

查找模块,用于显示业务主题选择界面,接收选择指令;基于所述选择指令,确定目标业务主题;根据所述目标业务主题,查找所述预置数据库中的对应目标数据表;计算所述原始拼音字符串与所述目标数据表中存储的各个拼音字符串的编辑距离;基于所述编辑距离,得到所述原始拼音字符串与所述各个拼音字符串的相似度;确定最高相似度对应的拼音字符串为目标字符串;

检测模块,用于检测所述最高相似度是否大于或等于预设阈值;

输出模块,用于在检测所述最高相似度大于或等于预设阈值时,从所述预置数据库中获取所述目标字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果,以及在检测所述最高相似度小于预设阈值时,显示文字输入界面,并接收基于所述文字输入界面输入的文字信息,并输出所述文字信息。

4. 一种语音识别设备,其特征在於,所述语音识别设备包括:存储器、处理器及存储在所述存储器上并在所述处理器上运行的语音识别程序,所述语音识别程序被所述处理器执行时实现如权利要求1或2所述的语音识别方法的步骤。

5. 一种计算机可读存储介质,其特征在於,所述计算机可读存储介质上存储有语音识别程序,所述语音识别程序被处理器执行时实现如权利要求1或2所述的语音识别方法的步骤。

## 语音识别方法、装置、设备及计算机可读存储介质

### 技术领域

[0001] 本发明涉及语音交互技术领域,尤其涉及一种语音识别方法、装置、设备及计算机可读存储介质。

### 背景技术

[0002] 目前,在语音交互业务中,一般是将用户语音识别为文字,然后根据识别到的文字进行业务处理。单对中文语音进行识别时,因为用户发音不标准或当前识别机制的缺陷,往往会将某个字的声调识别错误,例如,用户想要表达的是“补领驾驶证”,若识别为“不领驾驶证”,则导致业务处理出错,影响用户体验。

### 发明内容

[0003] 本发明的主要目的在于提供一种语音识别方法、装置、设备及计算机可读存储介质,旨在解决现有技术中由于识别机制的缺陷,往往会将某个字的声调识别错误的技术问题。

[0004] 为实现上述目的,本发明提供一种语音识别方法,所述语音识别方法包括以下步骤:

[0005] 获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;

[0006] 将所述原始文字识别结果转换成原始拼音字符串;

[0007] 查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;

[0008] 从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。

[0009] 可选地,所述查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串的步骤包括:

[0010] 显示业务主题选择界面,接收选择指令;

[0011] 基于所述选择指令,确定目标业务主题;

[0012] 根据所述目标业务主题,查找所述预置数据库中的对应目标数据表;

[0013] 确定所述目标数据表中与所述原始拼音字符串相似度最高的拼音字符串为目标拼音字符串。

[0014] 可选地,在所述查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串的步骤之前,还包括:

[0015] 创建业务主题对应的数据表;

[0016] 获取业务主题对应的历史业务提问记录,所述历史业务提问记录中包含若干个业务提问;

[0017] 获取每个业务提问对应的拼音字符串;

[0018] 将业务提问及对应的拼音字符串关联存储至所述数据表中;

[0019] 基于所述数据表,构建所述预置数据库。

[0020] 可选地,所述确定所述目标数据表中与所述原始拼音字符串相似度最高的拼音字符串为目标拼音字符串的步骤包括:

[0021] 计算所述原始拼音字符串与所述目标数据表中存储的各个拼音字符串的编辑距离;

[0022] 基于所述编辑距离,得到所述原始拼音字符串与所述各个拼音字符串的相似度;

[0023] 确定最高相似度对应的拼音字符串为目标字符串。

[0024] 可选地,在所述确定最高相似度对应的拼音字符串为目标字符串的步骤之后,还包括:

[0025] 检测所述最高相似度是否大于或等于预设阈值;

[0026] 若所述最高相似度大于或等于预设阈值,则执行获取所述目标拼音字符串对应的目标文字,并输出所述目标文字的步骤。

[0027] 可选地,在所述检测所述最高相似度是否大于或等于预设阈值的步骤之后,还包括:

[0028] 若所述最高相似度小于预设阈值,则显示文字输入界面;

[0029] 接收基于所述文字输入界面输入的文字信息,并输出所述文字信息。

[0030] 可选地,在所述接收基于所述文字输入界面输入的文字信息,并输出所述文字信息之后,还包括:

[0031] 获取所述文字信息对应的拼音字符串,并将所述文字信息以及所述文字信息对应的拼音字符串关联保存至所述目标数据表。

[0032] 此外,为实现上述目的,本发明还提供一种语音识别装置,所述语音识别装置包括:

[0033] 识别模块,用于获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;

[0034] 转换模块,用于将所述原始文字识别结果转换成原始拼音字符串;

[0035] 查找模块,用于查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;

[0036] 输出模块,用于从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。

[0037] 可选地,所述查找模块包括:

[0038] 指令接收单元,用于显示业务主题选择界面,接收选择指令;

[0039] 主题确定单元,用于基于所述选择指令,确定目标业务主题;

[0040] 数据表确定单元,用于根据所述目标业务主题,查找所述预置数据库中的对应目标数据表;

[0041] 拼音字符串确定单元,用于确定所述目标数据表中与所述原始拼音字符串相似度最高的拼音字符串为目标拼音字符串。

[0042] 可选地,所述语音识别装置还包括:

[0043] 创建模块,用于创建业务主题对应的数据表;

[0044] 获取模块,用于获取业务主题对应的历史业务提问记录,所述历史业务提问记录中包含若干个业务提问;获取每个业务提问对应的拼音字符串;

- [0045] 存储模块,用于将业务提问及对应的拼音字符串关联存储至所述数据表中;
- [0046] 构建模块,用于基于所述数据表,构建预置数据库。
- [0047] 可选地,所述拼音字符串确定单元包括:
- [0048] 计算子单元,用于计算所述原始拼音字符串与所述目标数据表中存储的各个拼音字符串的编辑距离;
- [0049] 换算子单元,用于基于所述编辑距离,得到所述原始拼音字符串与所述各个拼音字符串的相似度;
- [0050] 确定子单元,用于确定最高相似度对应的拼音字符串为目标字符串。
- [0051] 可选地,所述语音识别装置还包括:
- [0052] 检测模块,用于检测所述最高相似度是否大于或等于预设阈值;
- [0053] 第一判定模块,用于若所述最高相似度大于或等于预设阈值,则执行获取所述目标拼音字符串对应的目标文字,并输出所述目标文字的步骤。
- [0054] 可选地,所述语音识别装置还包括:
- [0055] 第一判定模块,用于若所述最高相似度小于预设阈值,则显示文字输入界面;
- [0056] 所述输出模块,还用于接收基于所述文字输入界面输入的文字信息,并输出所述文字信息。
- [0057] 可选地,所述语音识别装置还包括:
- [0058] 更新模块,用于获取所述文字信息对应的拼音字符串,并将所述文字信息以及所述文字信息对应的拼音字符串关联保存至所述目标数据表。
- [0059] 此外,为实现上述目的,本发明还提供一种语音识别设备,所述语音识别设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的语音识别程序,所述语音识别程序被所述处理器执行时实现如上所述的语音识别方法的步骤。
- [0060] 此外,为实现上述目的,本发明还提供一种计算机可读存储介质,所述计算机可读存储介质上存储有语音识别程序,所述语音识别程序被处理器执行时实现如上所述的语音识别方法的步骤。
- [0061] 本发明中,将接收到的语音识别为原始文字后,将原始文字转化成原始拼音,然后从数据库中查找与原始拼音最相似的目标拼音,最后输出目标拼音对应的目标文字。即在本发明中,在得到语音识别结果后,进一步对语音识别结果进行纠错处理,避免了因音调识别错误而造成语音识别结果出错,提高了语音识别的准确率。

## 附图说明

- [0062] 图1为本发明实施例方案涉及的硬件运行环境的语音识别设备结构示意图;
- [0063] 图2为本发明语音识别方法一实施例的流程示意图;
- [0064] 图3为本发明语音识别装置一实施例的功能模块示意图。
- [0065] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

## 具体实施方式

- [0066] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。
- [0067] 如图1所示,图1为本发明实施例方案涉及的硬件运行环境的语音识别设备结构示

意图。

[0068] 本发明实施例语音识别设备可以是PC,也可以是智能手机、平板电脑、便携计算机等终端设备。

[0069] 如图1所示,该语音识别设备可以包括:处理器1001,例如CPU,网络接口1004,用户接口1003,存储器1005,通信总线1002。其中,通信总线1002用于实现这些组件之间的连接通信。用户接口1003可以包括显示屏(Display)、输入单元比如键盘(Keyboard),可选用户接口1003还可以包括标准的有线接口、无线接口。网络接口1004可选地可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器1005可以是高速RAM存储器,也可以是稳定的存储器(non-volatile memory),例如磁盘存储器。存储器1005可选地还可以是独立于前述处理器1001的存储装置。

[0070] 本领域技术人员可以理解,图1中示出的语音识别设备结构并不构成对语音识别设备的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

[0071] 如图1所示,作为一种计算机存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及语音识别程序。

[0072] 在图1所示的语音识别设备中,网络接口1004主要用于连接后台服务器,与后台服务器进行数据通信;用户接口1003主要用于连接客户端(用户端),与客户端进行数据通信;而处理器1001可以用于调用存储器1005中存储的语音识别程序,并执行以下语音识别方法各个实施例的操作。

[0073] 参照图2,图2为本发明语音识别方法一实施例的流程示意图。

[0074] 本发明语音识别方法一实施例中,语音识别方法包括:

[0075] 步骤S10,获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;

[0076] 本实施例中,在一些语音交互业务中,客户可通过语音的形式提出需要问的问题。在语音识别设备获取到用户发出的语音信息时,对语音信息进行语音识别,得到语音信息对应的原始文字识别结果。其中,对语音信息进行语音识别依靠现有的语音识别技术实现。语音识别技术,也被称为自动语音识别Automatic Speech Recognition, (ASR),其目标是将人类的语音中的词汇内容转换为计算机可读的输入,例如二进制编码或者字符序列,从而根据二进制编码或者字符序列将语音转换成文字。

[0077] 步骤S20,将所述原始文字识别结果转换成原始拼音字符串;

[0078] 本实施例中,可以借助编程技术,实现文字转拼音。例如pinyin4j,pinyin4j是sourceforge.net上的一个开源项目,用来处理将中文转换成拼音的场景。目前,文字转拼音技术较为成熟,可选地技术手段较多,可根据实际需要选择具体的技术手段实现文字(原始文字识别结果)转拼音(原始拼音字符串),在此不作限制。

[0079] 步骤S30,查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;

[0080] 本实施例中,步骤S30之前可收集历史处理过的业务中,用户提过的问题或经常提的问题。例如,以交通业务为例,收集处理交通业务时,业务处理人员遇到的业务提问。将业务提问以文字的形式整理出来,得到历史业务提问记录,历史业务提问记录中包括:业务提

问1、业务提问2……业务提问n。然后借助文字转拼音技术,得到业务提问1对应的拼音字符串、业务提问2对应的拼音字符串……业务提问n对应的拼音字符串,将每个业务提问与其对应的拼音字符串关联保存至预置数据库中。存储形式如表1所示,表1为预置数据库的数据存储结构示意图。

[0081] 表1

[0082]	业务提问 1	拼音字符串 1
	业务提问 2	拼音字符串 2
	.....	.....
[0083]	业务提问 n	拼音字符串 n

[0084] 按照上述方式预先构建预置数据库之后,再将原始文字识别结果转换成原始拼音字符串之后,逐一计算原始拼音字符串与预置数据库中的每个拼音字符串的相似度。具体的,可通过计算原始拼音字符串与预置数据库中的每个拼音字符串的编辑距离,确定原始拼音字符串与预置数据库中的每个拼音字符串的相似度,编辑距离越小,相似度越高。

[0085] 其中编辑距离(EditDistance)是指两个字串之间,由一个转成另一个所需的最少编辑操作次数。以莱文斯坦距离为例,莱文斯坦距离又称Levenshtein距离,是编辑距离的一种。指两个字串之间,由一个转成另一个所需的最少编辑操作次数。允许的编辑操作包括将一个字符替换成另一个字符,插入一个字符,删除一个字符。

[0086] 例如将kitten编辑转成sitting,需要的最小编辑距离为3,具体编辑处理过程如下:

[0087] 1、sitten(k→s);

[0088] 2、sittin(e→i);

[0089] 3、sitting(→g)。

[0090] 本实施例中,编辑距离与相似度的对应关系定义如下:

[0091] 编辑距离为0,相似度为100%;编辑距离为1,相似度为99%;编辑距离为2,相似度为98%……以此类推,直至编辑距离大于或等于100时,相似度为0。

[0092] 参照上述说明,即可得到原始拼音字符串与预置数据库中的每个拼音字符串的相似度。然后进一步从得到的相似度中选取最高相似度,并将最高相似度对应的拼音字符串作为目标拼音字符串。

[0093] 步骤S40,从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。

[0094] 本实施例中,确定目标拼音字符串后,进一步从数据库中获取目标拼音字符串对应的目标文字,然后输出目标文字,以供基于目标文字进行后续的业务处理。

[0095] 本实施例中,获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;将所述原始文字识别结果转换成原始拼音字符串;查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。通过本实施例,在得到语音识别结果后,进一步对语音识别结果进行纠错处理,避免了因音调识别错误而造成语音



识别结果出错,提高了语音识别的准确率。

[0096] 进一步地,本发明语音识别方法一实施例中,上述步骤S30包括:

[0097] 步骤S301,显示业务主题选择界面,接收选择指令;

[0098] 在本实施例中,显示业务主题选择界面,以供用户选择需要办理的业务的业务主题。对于某个交互业务来说,可分为若干个业务主题。例如交通业务,可以分为:业务主题1-查分业务(办理驾驶证分数查询、销分等)、业务主题1-证件业务(办理证件挂失、证件更新等)。可以是在业务主题选择界面上显示所有的业务主题,由用户点击需要的业务主题,触发选择指令,语音识别设备接收到该选择指令。

[0099] 步骤S302,基于所述选择指令,确定目标业务主题;

[0100] 步骤S303,根据所述目标业务主题,查找所述预置数据库中的对应目标数据表;

[0101] 本实施例中,预置数据库中存储了不同业务主题对应的数据表。例如,在创建预置数据库时,可创建业务主题1对应的数据表A,业务主题2对应的数据表B。其中,数据表A中存储有:业务提问A1以及对应的拼音字符串、业务提问A2以及对应的拼音字符串……业务提问An以及对应的拼音字符串。业务提问A1至An为历史处理过的查分业务中,用户提过的问题或经常提的问题。数据表B中存储有:业务提问B1以及对应的拼音字符串、业务提问B2以及对应的拼音字符串……业务提问Bn以及对应的拼音字符串。业务提问B1至Bn为历史处理过的证件业务中,用户提过的问题或经常提的问题。

[0102] 一实施例中,若用户当前想要办理证件业务,则选择业务主题选择界面上的证件业务,接收基于用户操作触发的选择指令,根据该选择指令即可确定证件业务为目标业务主题,然后进一步确定数据表B为目标数据表。

[0103] 步骤S304,确定所述目标数据表中与所述原始拼音字符串相似度最高的拼音字符串为目标拼音字符串。

[0104] 本实施例中,若数据表B为目标数据表,则逐一计算原始拼音字符串与数据表B中的每个拼音字符串的相似度。具体的,可通过计算原始拼音字符串与数据表B中的每个拼音字符串的编辑距离,确定原始拼音字符串与数据表B中的每个拼音字符串的相似度,编辑距离越小,相似度越高。即可得到原始拼音字符串与数据表B中的每个拼音字符串的相似度。然后进一步从得到的相似度中选取最高相似度,并将最高相似度对应的拼音字符串作为目标拼音字符串。

[0105] 本实施例中,由用户选择当前需要办理的业务的业务主题,然后从用户选择的业务主题对应的数据表中确定与原始拼音字符串相似度最高的目标拼音字符串,缩小了对比范围,减少了计算量,加快了查找结果,从而加快了整体处理速度。

[0106] 进一步地,本发明语音识别方法一实施例中,在上述步骤S30之前,还包括:

[0107] 创建业务主题对应的数据表;获取业务主题对应的历史业务提问记录,所述历史业务提问记录中包含若干个业务提问;获取每个业务提问对应的拼音字符串;将业务提问及对应的拼音字符串关联存储至所述数据表中;基于所述数据表,构建预置数据库。

[0108] 本实施例中,在创建预置数据库时,可创建业务主题1对应的数据表A,业务主题2对应的数据表B。其中,数据表A中存储有:业务提问A1以及对应的拼音字符串、业务提问A2以及对应的拼音字符串……业务提问An以及对应的拼音字符串。业务提问A1至An为历史处理过的查分业务中,用户提过的问题或经常提的问题。数据表B中存储有:业务提问B1以及

对应的拼音字符串、业务提问B2以及对应的拼音字符串……业务提问Bn以及对应的拼音字符串。业务提问B1至Bn为历史处理过的证件业务中,用户提过的问题或经常提的问题。

[0109] 通过本实施例,将业务提问及其对应的拼音,存储至业务提问所属的业务主题对应的数据表中。后续由用户选择当前需要办理的业务的业务主题,然后从用户选择的业务主题对应的数据表中确定与原始拼音字符串相似度最高的目标拼音字符串,缩小了对比范围,减少了计算量,加快了查找结果,从而加快了整体处理速度。

[0110] 进一步地,本发明语音识别方法一实施例中,上述步骤S304包括:

[0111] 步骤S3041,计算所述原始拼音字符串与所述目标数据表中存储的各个拼音字符串的编辑距离;

[0112] 本实施例中,首先计算原始拼音字符串与目标数据表中的每个拼音字符串的编辑距离,编辑距离越小,相似度越高。

[0113] 其中编辑距离(EditDistance)是指两个字串之间,由一个转成另一个所需的最少编辑操作次数。以莱文斯坦距离为例,莱文斯坦距离又称Levenshtein距离,是编辑距离的一种。指两个字串之间,由一个转成另一个所需的最少编辑操作次数。允许的编辑操作包括将一个字符替换成另一个字符,插入一个字符,删除一个字符。

[0114] 例如将kitten编辑转成sitting,需要的最小编辑距离为3,具体编辑处理过程如下:

[0115] 1、sitten(k→s);

[0116] 2、sittin(e→i);

[0117] 3、sitting(→g)。

[0118] 步骤S3042,基于所述编辑距离,得到所述原始拼音字符串与所述各个拼音字符串的相似度;

[0119] 本实施例中,编辑距离与相似度的对应关系定义如下:

[0120] 编辑距离为0,相似度为100%;编辑距离为1,相似度为99%;编辑距离为2,相似度为98%……以此类推,直至编辑距离大于或等于100时,相似度为0。

[0121] 步骤S3043,确定最高相似度对应的拼音字符串为目标字符串。

[0122] 本实施例中,通上述步骤S3041至步骤S3042,即可得到原始拼音字符串与目标数据表中的每个拼音字符串的相似度。然后进一步从得到的相似度中选取最高相似度,并将最高相似度对应的拼音字符串作为目标拼音字符串。

[0123] 进一步地,本发明语音识别方法一实施例中,在上述步骤S3043之后,还包括:

[0124] 检测所述最高相似度是否大于或等于预设阈值;

[0125] 若大于或等于预设阈值,则执行上述步骤S40。

[0126] 本实施例中,有可能存在最高相似度的值较小的情况。若存在这种情况,则表明原始拼音字符串与目标拼音字符串差别较大,相应的,表明目标文字与原始文字识别结果的差别较大,即目标文字并不是用户想要表达的意思。因此,设置一预设阈值,例如设置为90%(该值可根据实际情况进行设置),只有相似度大于或等于90%时(说明原始拼音字符串与目标拼音字符串差别较小,相应的,表明目标文字与原始文字识别结果的差别较小,目标文字很大概率即为用户想要表达的意思),才执行步骤S40。即只有在原始拼音字符串与目标拼音字符串差别较小时,才输出目标拼音字符串对应的目标文字。提高了语音识别的

准确度。

[0127] 进一步地,本发明语音识别方法一实施例中,在上述检测所述最高相似度是否大于或等于预设阈值的步骤之后,还包括:

[0128] 若所述最高相似度小于预设阈值,则显示文字输入界面;接收基于所述文字输入界面输入的文字信息,并输出所述文字信息。

[0129] 本实施例中,若最高相似度小于预设阈值,则说明原始拼音字符串与目标拼音字符串差别较大,相应的,表明目标文字与原始文字识别结果的差别较大,即目标文字并不是用户想要表达的意思。因此,需要更换交互方式。例如,显示文字输入界面,以供用户输入文字信息。在接收到用户输入的文字信息后,输出文字信息,以供基于该文字信息进行后续的业务处理。

[0130] 进一步地,本发明语音识别方法一实施例中,在上述接收基于所述文字输入界面输入的文字信息,并输出所述文字信息的步骤之后,还包括:

[0131] 获取所述文字信息对应的拼音字符串,并将所述文字信息以及所述文字信息对应的拼音字符串关联保存至所述目标数据表。

[0132] 本实施例中,若最高相似度小于预设阈值,则说明原始拼音字符串与目标拼音字符串差别较大,相应的,表明目标文字与原始文字识别结果的差别较大,即目标文字并不是用户想要表达的意思。也说明针对当前用户输入的语音,在数据库中没有与之相似的文字表达。

[0133] 因此,将信息输入方式更换为文字输入方式。并对此时用户输入的文字信息进行文字转拼音,然后将当前输入的文字与对应的拼音字符串关联保存至数据库中。即对数据库的数据进行了增加,使数据库的数据更加丰富,提升了方案的适用性。

[0134] 参照图3,图3为本发明语音识别装置一实施例的功能模块示意图。

[0135] 本发明语音识别装置一实施例中,语音识别装置包括:

[0136] 识别模块10,用于获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;

[0137] 转换模块20,用于将所述原始文字识别结果转换成原始拼音字符串;

[0138] 查找模块30,用于查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;

[0139] 输出模块40,用于从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。

[0140] 本实施例中,获取语音信息,并对所述语音信息进行语音识别,得到原始文字识别结果;将所述原始文字识别结果转换成原始拼音字符串;查找预置数据库中与所述原始拼音字符串相似度最高的目标拼音字符串;从所述预置数据库中获取所述目标拼音字符串对应的目标文字,并输出所述目标文字以作为语音识别的最终结果。通过本实施例,在得到语音识别结果后,进一步对语音识别结果进行纠错处理,避免了因音调识别错误而造成语音识别结果出错,提高了语音识别的准确率。

[0141] 此外,本发明实施例还提出一种计算机可读存储介质,所述计算机可读存储介质上存储有语音识别程序,所述语音识别程序被处理器执行时实现如上语音识别方法各个实施例的操作。

[0142] 本发明计算机可读存储介质的具体实施例与上述语音识别方法的各个实施例基本相同,在此不做赘述。

[0143] 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、方法、物品或者系统不仅包括那些要素,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、方法、物品或者系统所固有的要素。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、方法、物品或者系统中还存在另外的相同要素。

[0144] 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。

[0145] 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)执行本发明各个实施例所述的方法。

[0146] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

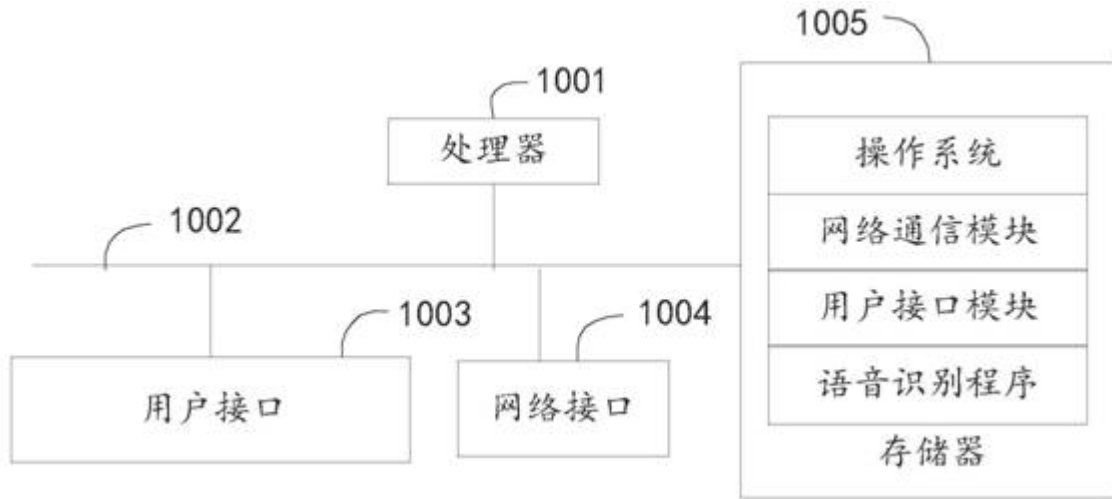


图1

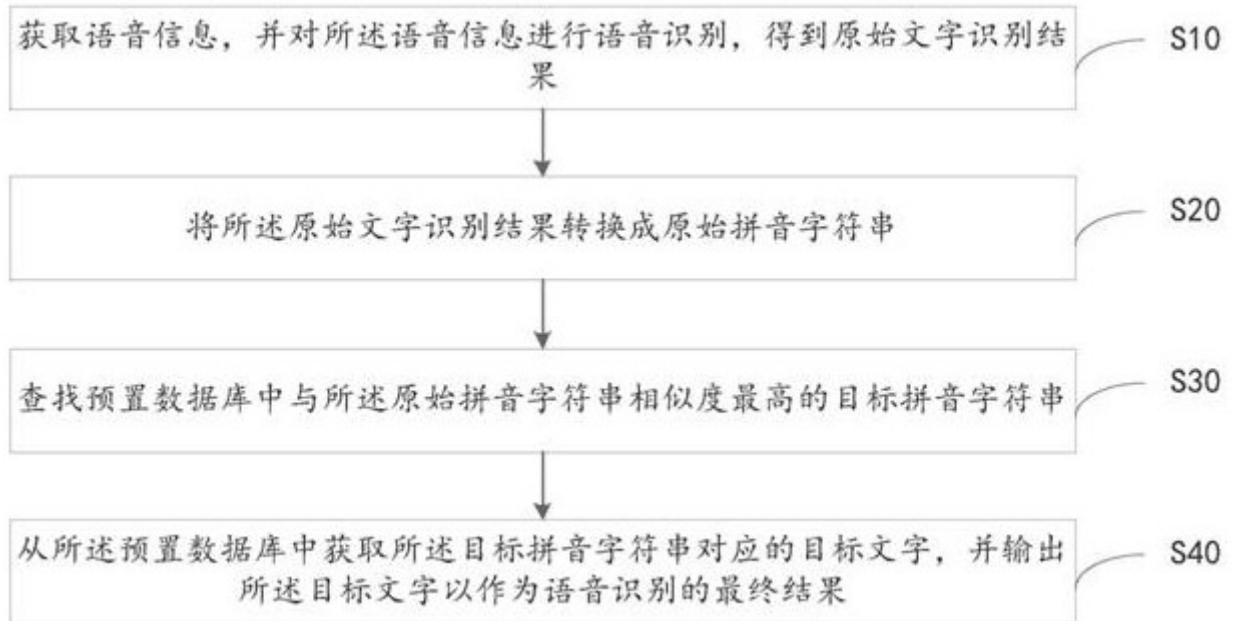


图2

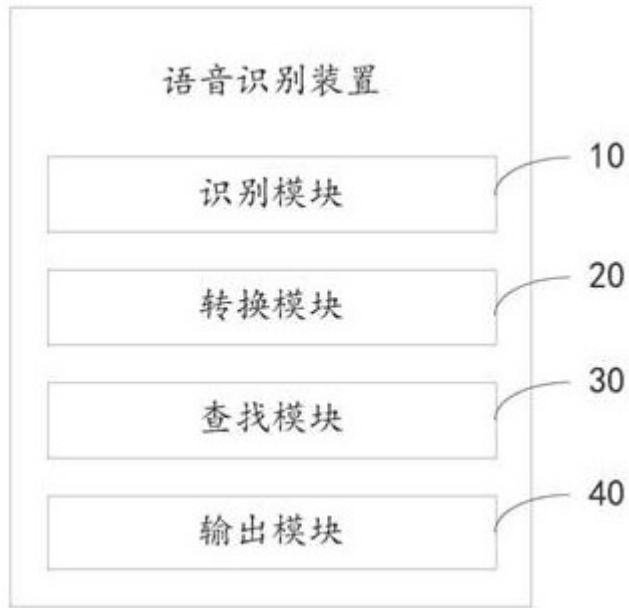


图3