

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 10 月 29 日 (29.10.2020)



(10) 国际公布号
WO 2020/215554 A1

- (51) 国际专利分类号:
G10L 15/22 (2006.01) **G10L 15/26** (2006.01)
- (21) 国际申请号: PCT/CN2019/102437
- (22) 国际申请日: 2019 年 8 月 26 日 (26.08.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910342353.2 2019 年 4 月 26 日 (26.04.2019) CN
- (71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国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, Guangdong 518033 (CN)。
- (72) 发明人: 魏韬(WEI, Tao); 中国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, Guangdong 518033 (CN)。 马骏(MA, Jun);

中国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, Guangdong 518033 (CN)。 王少军(WANG, Shaojun); 中国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, Guangdong 518033 (CN)。

(74) 代理人: 北京市京大律师事务所(BEIJING JINGDA LAW FIRM); 中国北京市海淀区海淀中街 16 号 7 层 2 单元 703, Beijing 100080 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: SPEECH RECOGNITION METHOD, DEVICE, AND APPARATUS, AND COMPUTER-READABLE STORAGE MEDIUM

(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

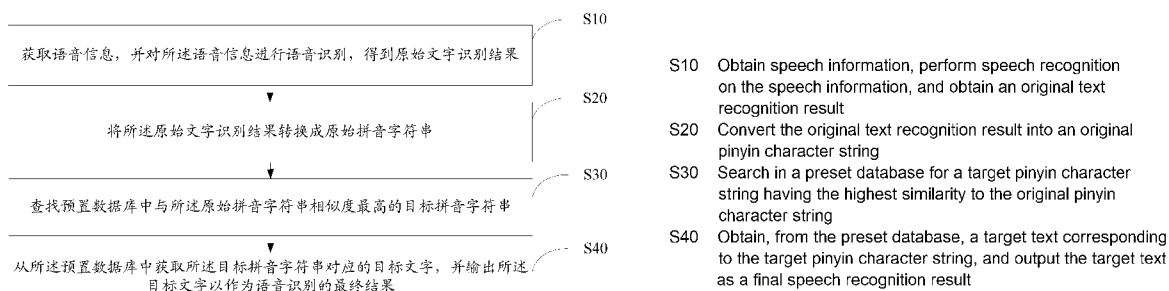


图 2

(57) Abstract: A speech recognition method, device, and apparatus, and a computer-readable storage medium. The method comprises: obtaining speech information, performing speech recognition on the speech information, and obtaining an original text recognition result (S10); converting the original text recognition result into an original pinyin character string (S20); searching in a preset database for a target pinyin character string having the highest similarity to the original pinyin character string (S30); and obtaining, from the preset database, a target text corresponding to the target pinyin character string, and outputting the target text as a final speech recognition result (S40). The method is used to obtain a speech recognition result and to further perform error correction on said result, thereby preventing errors from occurring in the speech recognition result due to pitch recognition errors, and improving speech recognition accuracy.

(57) 摘要: 一种语音识别方法、装置、设备及计算机可读存储介质, 该方法包括: 获取语音信息, 并对语音信息进行语音识别, 得到原始文字识别结果 (S10); 将原始文字识别结果转换成原始拼音字符串 (S20); 查找预置数据库中与原始拼音字符串相似度最高的目标拼音字符串 (S30); 从预置数据库中获取目标拼音字符串对应的目标文字, 并输出目标文字以作为语音识别的最终结果 (S40)。通过该方法, 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。

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国(除另有指明，要求每一种可提供的地区
保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

- 包括国际检索报告(条约第21条(3))。

语音识别方法、装置、设备及计算机可读存储介质

本申请要求于2019年4月26日提交中国专利局、申请号为201910342353.2、发明名称为“语音识别方法、装置、设备及计算机可读存储介质”的中国专利申请5 的优先权，其全部内容通过引用结合在申请中。

技术领域

本申请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。

10 背景技术

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。发明人意识到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15 处理出错，影响用户体验。

发明内容

本申请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20 声调识别错误的技术问题。

为实现上述目的，本申请提供一种语音识别方法，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；

25 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。

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：

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；

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。

5 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：

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；

10 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，构建所述预置数据库。

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：

15 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。

20 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：

检测所述最高相似度是否大于或等于预设阈值；

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。

25 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：

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。

30 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：

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。

此外，为实现上述目的，本申请还提供一种语音识别装置，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；

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；

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。

此外，为实现上述目的，本申请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。

此外，为实现上述目的，本申请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。

本申请中，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。即在本申请中，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

图1为本申请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；

图2为本申请语音识别方法一实施例的流程示意图；

图3为本申请语音识别装置一实施例的功能模块示意图。

5 本申请目的的实现、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

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。

10 如图1所示，图1为本申请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。

本申请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。

15 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。

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。

25 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。

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。

30

参照图 2，图 2 为本申请语音识别方法一实施例的流程示意图。

本申请语音识别方法一实施例中，语音识别方法包括：

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；

5 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
10 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。

步骤 S20，将所述原始文字识别结果转换成原始拼音字符串；

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
15 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。

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；

20 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
25 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。

表 1

业务提问 1	拼音字符串 1
业务提问 2	拼音字符串 2

.....	
业务提问 n	拼音字符串 n

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。

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。

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：

- 1、sitten(k→s);
- 2、sittin(e→i);
- 3、sitting(→g)。

本实施例中，编辑距离与相似度的对应关系定义如下：

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。

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。

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。

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。

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。

进一步地，本申请语音识别方法一实施例中，上述步骤 S30 包括：

步骤 S301，显示业务主题选择界面，接收选择指令；

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。

步骤 S302，基于所述选择指令，确定目标业务主题；

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；

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 1，业务主题 2 对应的数据表 2。其中，数据表 1 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 2 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。

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 2 为目标数据表。

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。

本实施例中，若数据表 2 为目标数据表，则逐一计算原始拼音字符串与数据表 2 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 2 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 2 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 2 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。

本实施例中，由用户选择当前需要办理的业务的主题，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。

进一步地，本申请语音识别方法一实施例中，在上述步骤 S30 之前，还包括：

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。

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 1，业务主题 2 对应的数据表 2。其中，数据表 1 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 2 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。

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。

进一步地，本申请语音识别方法一实施例中，上述步骤 S304 包括：

- 5 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；

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。

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。

15 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：

- 1、sitten(k→s);
- 2、sittin(e→i);
- 3、sitting(→g)。

20 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；

本实施例中，编辑距离与相似度的对应关系定义如下：

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。

25 步骤 S3043，确定最高相似度对应的拼音字符串为目标字符串。

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。

30 进一步地，本申请语音识别方法一实施例中，在上述步骤 S3043 之后，

还包括:

检测所述最高相似度是否大于或等于预设阈值;

若大于或等于预设阈值, 则执行上述步骤 S40。

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, 5 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 10 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。

进一步地, 本申请语音识别方法一实施例中, 在上述检测所述最高相似 15 度是否大于或等于预设阈值的步骤之后, 还包括:

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。

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 20 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。

进一步地, 本申请语音识别方法一实施例中, 在上述接收基于所述文字 25 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:

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。

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 30 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，在数据库中沒有与之相似的文字表达。

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。

参照图 3，图 3 为本申请语音识别装置一实施例的功能模块示意图。

本申请语音识别装置一实施例中，语音识别装置包括：

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；

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；

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；

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。

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。

可选的，在一具体实施例中，所述查找模块包括：

指令接收单元，用于显示业务主题选择界面，接收选择指令；

主题确定单元，用于基于所述选择指令，确定目标业务主题；

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；

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。

可选的，在一具体实施例中，所述语音识别装置还包括：

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；

5 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。

可选的，在一具体实施例中，所述拼音字符串确定单元包括：

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；

10 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；

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。

可选的，在一具体实施例中，所述语音识别装置还包括：

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；

15 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。

可选的，在一具体实施例中，所述语音识别装置还包括：

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；

20 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。

可选的，在一具体实施例中，所述语音识别装置还包括：

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。

25 此外，本申请实施例还提出一种计算机可读存储介质，该计算机可读存储介质可以为非易失性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。

30 本申请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。

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。

上述本申请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，本申请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)执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请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，均同理包括在本申请的专利保护范围内。

权 利 要 求 书

1、一种语音识别方法，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
5 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。

10 2、如权利要求 1 所述的语音识别方法，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：

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；

15 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；

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。

20 3、如权利要求 2 所述的语音识别方法，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：

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；

25 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，构建所述预置数据库。

4、如权利要求 2 所述的语音识别方法，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：

30 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。

5

5、如权利要求4所述的语音识别方法，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：

检测所述最高相似度是否大于或等于预设阈值；

10 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。

6、如权利要求5所述的语音识别方法，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：

若所述最高相似度小于预设阈值，则显示文字输入界面；

15 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。

7、如权利要求6所述的语音识别方法，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：

20 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。

8、一种语音识别装置，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；

25 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；

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；

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。

30

9、如权利要求 8 所述的语音识别装置，所述查找模块包括：

指令接收单元，用于显示业务主题选择界面，接收选择指令；

主题确定单元，用于基于所述选择指令，确定目标业务主题；

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

5 的对应目标数据表；

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。

10、如权利要求 9 所述的语音识别装置，所述语音识别装置还包括：

10 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；

15 构建模块，用于基于所述数据表，构建预置数据库。

11、如权利要求 9 所述的语音识别装置，所述拼音字符串确定单元包括：

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；

20 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；

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。

12、如权利要求 11 所述的语音识别装置，所述语音识别装置还包括：

25 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；

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。

13、如权利要求 12 所述的语音识别装置，所述语音识别装置还包括：

30 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;

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。

- 5 14、如权利要求 13 所述的语音识别装置,所述语音识别装置还包括:
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。

- 10 15、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下语音识别方法的步骤:

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;

- 15 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。

- 20 16、如权利要求 15 所述的语音识别设备,所述语音识别程序被所述处理
器执行时还实现如下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:

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;

- 25 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。

- 30 17、如权利要求 16 所述的语音识别设备,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,所述语音识别
程序被所述处理器执行时还实现如下语音识别方法的步骤:

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；

5 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，构建所述预置数据库。

18、如权利要求 16 所述的语音识别设备，所述语音识别程序被所述处理器执行时还实现如下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：

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；

15 确定最高相似度对应的拼音字符串为目标字符串。

19、如权利要求 18 所述的语音识别设备，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，所述语音识别程序被所述处理器执行时还实现如下语音识别方法的步骤：

20 检测所述最高相似度是否大于或等于预设阈值；

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；

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。

25

20、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下语音识别方法的步骤：

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；

30 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。

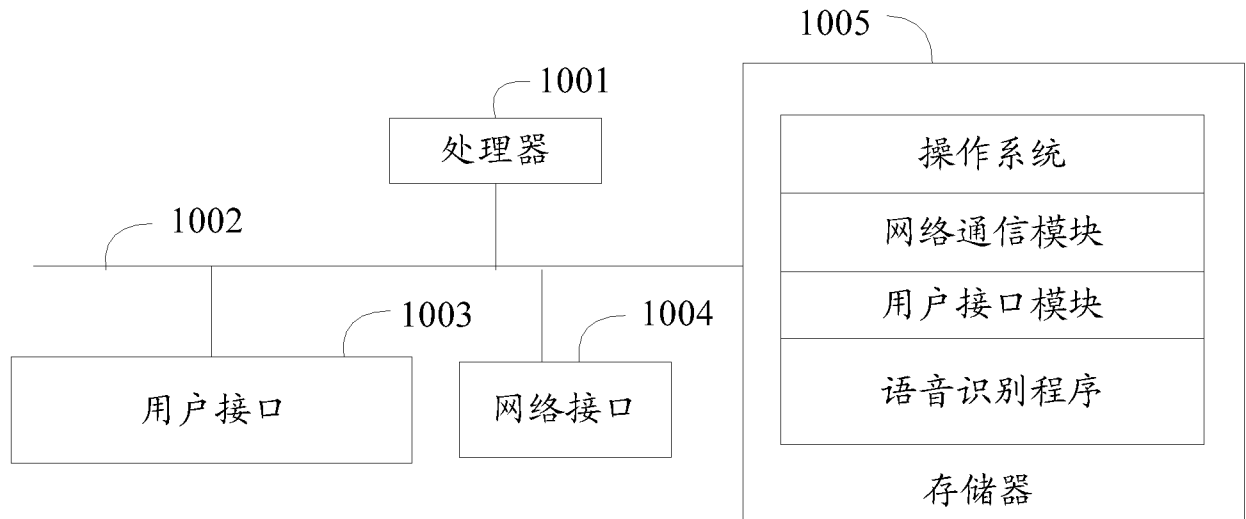


图 1

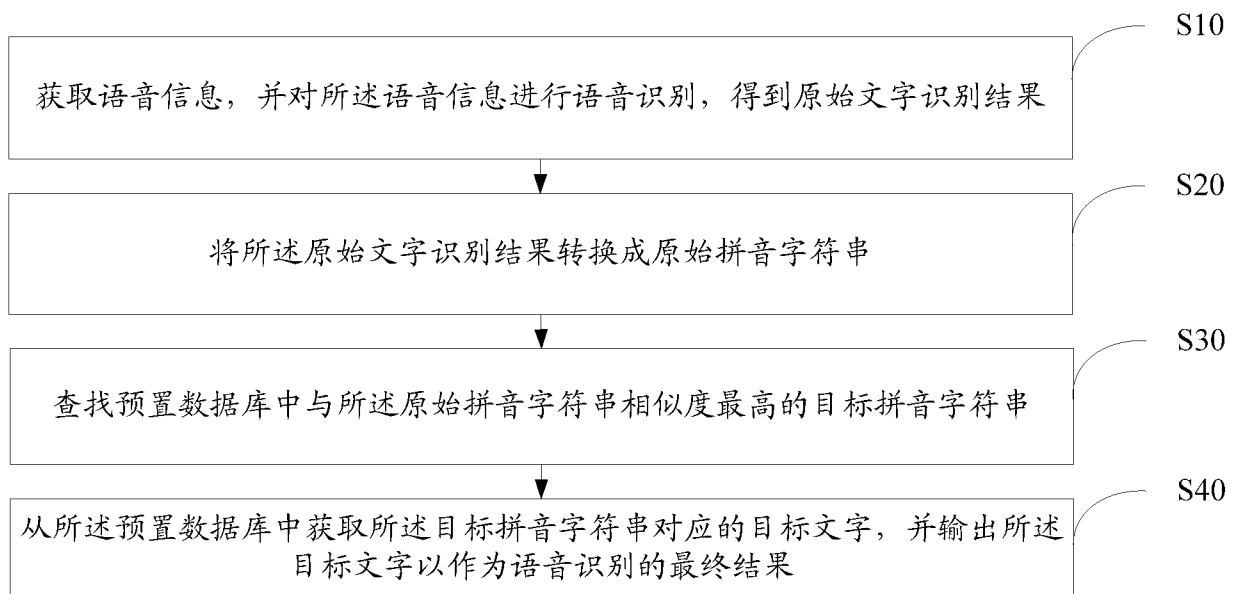


图 2



图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/102437

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/22(2006.01)i; G10L 15/26(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, VEN: 语音, 识别, 拼音, 目标, 文本, 文字, 汉字, 匹配, 相似, 阈值, 对话框, 界面, 窗口, audio, speech, voice, Chinese, character?, target, match+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110164435 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 23 August 2019 (2019-08-23) claims 1-10, and description, full text	1-20
Y	CN 105895090 A (LETV HOLDING (BEIJING) CO., LTD. et al.) 24 August 2016 (2016-08-24) description, paragraphs [0032]-[0100], and figures 1-4	1-20
Y	CN 102254557 A (SHENZHEN ZIDONG TECHNOLOGY CO., LTD.) 23 November 2011 (2011-11-23) description, paragraphs [0008]-[0101], and figures 1-6	1-20
Y	CN 108268637 A (GUANGDONG SWORD-DATA TECHNOLOGY CO., LTD.) 10 July 2018 (2018-07-10) description, paragraphs [0042]-[0096], and figures 1-7	1-20
A	CN 106297799 A (LETV HOLDING (BEIJING) CO., LTD. et al.) 04 January 2017 (2017-01-04) entire document	1-20
A	CN 109036419 A (NUBIA TECHNOLOGY CO., LTD.) 18 December 2018 (2018-12-18) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

22 January 2020

Date of mailing of the international search report

03 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/102437

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107657947 A (BEIJING Baidu NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 02 February 2018 (2018-02-02) entire document	1-20
A	CN 109545190 A (UNION MOBILE PAY CO., LTD.) 29 March 2019 (2019-03-29) entire document	1-20
A	CN 109256125 A (BEIJING Baidu NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 22 January 2019 (2019-01-22) entire document	1-20
A	CN 109036420 A (NUBIA TECHNOLOGY CO., LTD.) 18 December 2018 (2018-12-18) entire document	1-20
A	CN 101206859 A (TSINGHUA UNIVERSITY et al.) 25 June 2008 (2008-06-25) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/102437

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110164435	A	23 August 2019	None			
CN	105895090	A	24 August 2016	WO	2017166649	A1	05 October 2017
CN	102254557	A	23 November 2011	CN	102254557	B	24 July 2013
CN	108268637	A	10 July 2018	None			
CN	106297799	A	04 January 2017	None			
CN	109036419	A	18 December 2018	None			
CN	107657947	A	02 February 2018	None			
CN	109545190	A	29 March 2019	None			
CN	109256125	A	22 January 2019	None			
CN	109036420	A	18 December 2018	None			
CN	101206859	A	25 June 2008	CN	100470633	C	18 March 2009

国际检索报告

国际申请号

PCT/CN2019/102437

A. 主题的分类

G10L 15/22(2006.01)i; G10L 15/26(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, VEN: 语音, 识别, 拼音, 目标, 文本, 文字, 汉字, 匹配, 相似, 阈值, 对话框, 界面, 窗口, audio, speech, voice, Chinese, character?, target, match+

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 110164435 A (平安科技深圳有限公司) 2019年 8月 23日 (2019 - 08 - 23) 权利要求1-10, 以及说明书正文	1-20
Y	CN 105895090 A (乐视控股北京有限公司 等) 2016年 8月 24日 (2016 - 08 - 24) 说明书第[0032]-[0100]段, 图1-4	1-20
Y	CN 102254557 A (深圳市子栋科技有限公司) 2011年 11月 23日 (2011 - 11 - 23) 说明书第[0008]-[0101]段, 图1-6	1-20
Y	CN 108268637 A (广东蔚海数问大数据科技有限公司) 2018年 7月 10日 (2018 - 07 - 10) 说明书第[0042]-[0096]段, 图1-7	1-20
A	CN 106297799 A (乐视控股北京有限公司 等) 2017年 1月 4日 (2017 - 01 - 04) 全文	1-20
A	CN 109036419 A (努比亚技术有限公司) 2018年 12月 18日 (2018 - 12 - 18) 全文	1-20
A	CN 107657947 A (百度在线网络技术北京有限公司) 2018年 2月 2日 (2018 - 02 - 02) 全文	1-20

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布, 在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2020年 1月 22日

国际检索报告邮寄日期

2020年 2月 3日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

张亚玲

电话号码 86-(010)-62085722

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 109545190 A (联动优势科技有限公司) 2019年 3月 29日 (2019 - 03 - 29) 全文	1-20
A	CN 109256125 A (百度在线网络技术北京有限公司) 2019年 1月 22日 (2019 - 01 - 22) 全文	1-20
A	CN 109036420 A (努比亚技术有限公司) 2018年 12月 18日 (2018 - 12 - 18) 全文	1-20
A	CN 101206859 A (清华大学 等) 2008年 6月 25日 (2008 - 06 - 25) 全文	1-20

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/102437

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	110164435	A	2019年 8月 23日	无			
CN	105895090	A	2016年 8月 24日	W0	2017166649	A1	2017年 10月 5日
CN	102254557	A	2011年 11月 23日	CN	102254557	B	2013年 7月 24日
CN	108268637	A	2018年 7月 10日	无			
CN	106297799	A	2017年 1月 4日	无			
CN	109036419	A	2018年 12月 18日	无			
CN	107657947	A	2018年 2月 2日	无			
CN	109545190	A	2019年 3月 29日	无			
CN	109256125	A	2019年 1月 22日	无			
CN	109036420	A	2018年 12月 18日	无			
CN	101206859	A	2008年 6月 25日	CN	100470633	C	2009年 3月 18日