

(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 12 月 24 日 (24.12.2020)



(10) 国际公布号
WO 2020/253064 A1

(51) 国际专利分类号:
G10L 15/22 (2006.01)

(21) 国际申请号: PCT/CN2019/118415

(22) 国际申请日: 2019 年 11 月 14 日 (14.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910534912.X 2019年6月20日 (20.06.2019) 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 518000 (CN)。

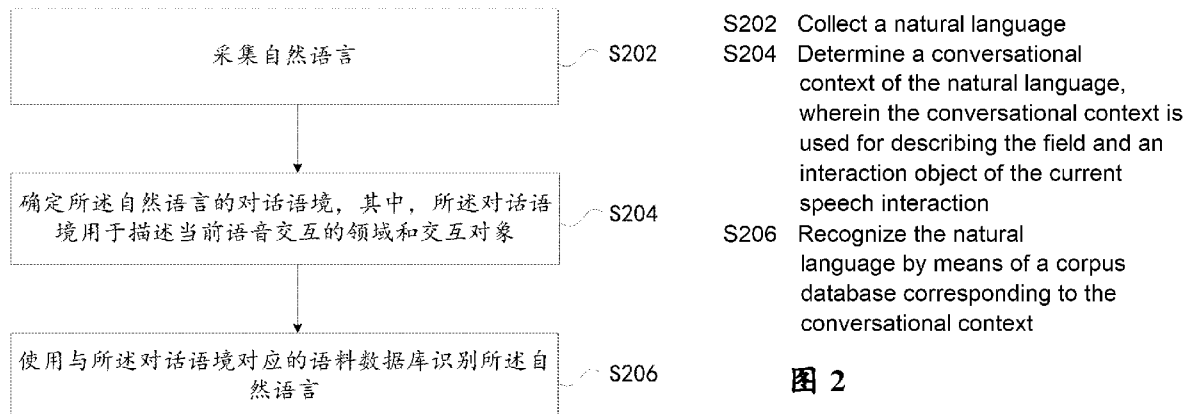
(72) 发明人: 张师玮(ZHANG, Shibei); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 北京中强智尚知识产权代理有限公司 (BEIJING STRONG WISDOM INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市朝阳区北苑东路 19 号院 2 号楼 2608 室, Beijing 100012 (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 AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 语音的识别方法及装置、计算机设备、存储介质



(57) Abstract: Provided are a speech recognition method and apparatus, and a computer device and a storage medium. The method comprises: collecting a natural language; determining a conversational context of the natural language, wherein the conversational context is used for describing the field and an interaction object of the current speech interaction; and recognizing the natural language by means of a corpus database corresponding to the conversational context. By means of the present invention, the technical problem in the prior art of a low speech efficiency is solved, speech recognition is more intelligent, and the speech recognition rate is improved.

(57) 摘要: 本发明实施例提供了一种语音的识别方法及装置、计算机设备、存储介质。一方面, 该方法包括: 采集自然语言; 确定所述自然语言的对话语境, 其中, 所述对话语境用于描述当前语音交互的领域和交互对象; 使用与所述对话语境对应的语料数据库识别所述自然语言。通过本发明, 解决了现有技术中语音效率低的技术问题, 语音识别更智能, 提高了语音识别率。

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)。

根据细则4.17的声明:

- 关于申请人有权申请并被授予专利 (细则
4.17(ii))
- 发明人资格 (细则4.17(iv))

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

语音的识别方法及装置、计算机设备、存储介质

【技术领域】

本申请涉及计算机领域，尤其涉及一种语音的识别方法及装置、计算机设备、存储介质。

【背景技术】

聊天机器人是人工智能中的重要领域，通过用户向聊天机器人发出自然语言，机器人可以完成一系列的操作。

相关技术中，聊天机器人只能做一些简单的闲聊对话，很少有能够分辨出人的真实意图，并且不能连续的针对某一个素材有限的语音或者结合上下文来理解对话意图，识别率低，需要获取大量的自然语言才能提高识别率，导致语音识别的应用范围一直不高。

针对相关技术中存在的上述问题，目前尚未发现有效的解决方案。

【发明内容】

有鉴于此，本申请实施例提供了一种语音的识别方法及装置、计算机设备、存储介质。

一方面，本申请实施例提供了一种语音的识别方法，所述方法包括：采集自然语言；确定所述自然语言的对话语境，其中，所述对话语境用于描述当前语音交互的领域和交互对象；使用与所述对话语境对应的语料数据库识别所述自然语言。

另一方面，本申请实施例提供了一种语音的识别装置，所述装置包括：采集模块，用于采集自然语言；确定模块，用于确定所述自然语言的对话语境，其中，所述对话语境用于描述当前语音交互的领域和交互对象；识别模

块，用于使用与所述对话语境对应的语料数据库识别所述自然语言。

根据本申请的又一个实施例，还提供了一种存储介质，所述存储介质中存储有计算机程序，其中，所述计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

根据本申请的又一个实施例，还提供了一种电子装置，包括存储器和处理器，所述存储器中存储有计算机程序，所述处理器被设置为运行所述计算机程序以执行上述任一项方法实施例中的步骤。

通过本申请，通过先确定对话语境和场景，进而在对应的语料数据库中识别自然语言，解决了现有技术中语音效率低的技术问题，语音识别更智能，提高了语音识别率。

【附图说明】

为了更清楚地说明本申请实施例的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其它的附图。

图1是本申请实施例的一种语音的识别移动终端的硬件结构框图；

图2是根据本申请实施例的语音的识别方法的流程图；

图3是本申请实施例中识别自然语言的流程图；

图4是本申请实施例交互示意图；

图5是根据本申请实施例的语音的识别装置的结构框图。

【具体实施方式】

下文中将参考附图并结合实施例来详细说明本申请。需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互组合。

需要说明的是，本申请的说明书和权利要求书及上述附图中的术语“第

一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。

实施例 1

本申请实施例一所提供的方法实施例可以在移动终端、中控设备、计算机终端或者类似的运算装置中执行。以运行在移动终端上为例，图 1 是本申请实施例的一种语音的识别移动终端的硬件结构框图。如图 1 所示，移动终端 10 可以包括一个或多个（图 1 中仅示出一个）处理器 102（处理器 102 可以包括但不限于微处理器 MCU 或可编程逻辑器件 FPGA 等的处理装置）和用于存储数据的存储器 104，可选地，上述移动终端还可以包括用于通信功能的传输设备 106 以及输入输出设备 108。本领域普通技术人员可以理解，图 1 所示的结构仅为示意，其并不对上述移动终端的结构造成限定。例如，移动终端 10 还可包括比图 1 中所示更多或者更少的组件，或者具有与图 1 所示不同的配置。

存储器 104 可用于存储计算机程序，例如，应用程序的软件程序以及模块，如本申请实施例中的语音的识别方法对应的计算机程序，处理器 102 通过运行存储在存储器 104 内的计算机程序，从而执行各种功能应用以及数据处理，即实现上述的方法。存储器 104 可包括高速随机存储器，还可包括非易失性存储器，如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中，存储器 104 可进一步包括相对于处理器 102 远程设置的存储器，这些远程存储器可以通过网络连接至移动终端 10。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

传输装置 106 用于经由一个网络接收或者发送数据。上述的网络具体实例可包括移动终端 10 的通信供应商提供的无线网络。在一个实例中，传输装置 106 包括一个网络适配器（Network Interface Controller，简称为 NIC），其可通过基站与其他网络设备相连从而可与互联网进行通讯。在一个实例中，

传输装置 106 可以为射频 (Radio Frequency, 简称为 RF) 模块, 其用于通过无线方式与互联网进行通讯。

在本实施例中提供了一种语音的识别方法, 图 2 是根据本申请实施例的语音的识别方法的流程图, 如图 2 所示, 该流程包括如下步骤:

步骤 S202, 采集自然语言;

步骤 S204, 确定所述自然语言的对话语境, 其中, 所述对话语境用于描述当前语音交互的领域和交互对象;

在本实施例中, 语境交互的领域为交互场景, 如天气查询, 路线导航, 对话聊天等, 交互对象为通过语音控制的目标对象, 如目标对象为系统的查询引擎, 通过查询引擎查询与自然语音对应的识别结果, 或者是根据自然语音中的指令信息去, 还包括调用引擎, 通过调用引擎来调用设备的软件或硬件, 以辅助识别, 如调用内置的 GPS 模块等, 还包括控制对象, 如通过中控台控制电视机等与中控台网络连接的设备时, 电视机为控制对象。

步骤 S206, 使用与所述对话语境对应的语料数据库识别所述自然语言。

通过本实施例的方案, 通过先确定对话语境和场景, 进而在对应的语料数据库中识别自然语言, 解决了现有技术中语音效率低的技术问题, 语音识别更智能, 提高了语音识别率。

本实施例的执行主体可以是聊天机器人, 语音控制台, 语音中控设备, 带有语音控制功能的电子设备等, 但不限于此。

本实施例中的对话语境包括对话场景, 机器交互参与的对象和设备等, 确定对话语境包括多种方案。

在一个示例中, 确定所述自然语言的对话语境包括:

S11, 提取所述自然语言中的关键字;

本实施例中的自然语言为自然人或者设备发出的语音, 在提取关键字时, 先将语音转换为文字, 然后使用预设的关键字列表, 在文字中查找是否包括

关键字列表中的任一关键字，如果包括则提取匹配的关键字。

S12，通过所述关键字匹配得到对话语境，其中，所述对话语境包括以下之一：知识搜索，互联网搜索，天气查询，路线导航，对话聊天，指令控制。

本实施例的关键字和对话语境通过预设映射表来匹配，如关键字集合{a,b,c}对应知识搜索的对话语境，关键字集合{b,c,d}对应互联网搜索的对话语境等。

在一些语境下，需要启动第三方软件（如地图，浏览器等），在一些语境下，需要用到本地知识图谱，在一些语境下，用到语料数据库。如路线导航的语境中，需要启动地图软件，如采集到的自然语言为“回家”，确定语境为路线导航，需要调用地图软件，搜索从当前位置到家的行车导航路线；或者在知识搜索的语境中，如采集到的自然语言为“张三的联系方式”，确定语境为知识搜索，在本地的通讯录中搜索出张三的联系方式等。

在另一个示例中，确定所述自然语言的对话语境包括：

S21，查找所述自然语言之前的历史对话记录；

本实施例的历史对话记录为当前自然语言之前的语音识别设备与语音识别对象之间的对话内容，该对话记录保存在语音设备的本地存储器，在查找历史对话记录时，可以根据时间作为参数，查找预设时间范围内的对话记录，例如，用户与语音机器人进行人机对话，在1分钟之内产生了三条对话记录，分别为：A,B,C，语音机器人当前接收到的待识别的自然语音为D，则可以根据A,B,C的内容来确定D的对话场景，如果D中的内容与A,B,C有关联性，则对话语境也是相同的，这是由于一次正常的人机对话都会包括多条对话记录。

S22，根据所述历史对话记录确定所述对话语境。

在本示例的方案中，通过当前自然语言的上下文（主要是上文）来确定对话场景，匹配该场景领域的语料，进而识别：在AIML机制下，记录对话

中的上下文信息，基于上下文信息，通过配置好的话题库来匹配识别某一领域类的语料对话，进而使用该领域的语料进行识别。如：接收到用户的自然语言“太暗了”，通过之前的聊天记录，确定对话场景为“电视机控制”，进而在电视机领域的语料中识别出“背景太暗了”，进而调亮背景亮度，或者确定对话场景为“灯光控制”，进而在灯光照明领域的语料中识别出“照明设备太暗了”，进而打开照明设备。

在本实施例中，在使用与所述对话语境对应的语料数据库识别所述自然语言之前，还包括：设置所述语料数据库，其中，所述语料数据库包括：多轮对话语料，多场景语料，通过 solr (Search On Lucene Replication) 搜索引擎得到的网络语料。本实施例中的上述三种语料库是三个维度的数据库，可以单独使用，也可以相互配合使用，其中，多轮对话语料是指包括多组对话的语料库，可用于聊天，本地查询等，如饮食、电视、电影、音乐、工作等多个主题（每个类别包括多组逻辑对话），多场景语料为在多个特定场景下的语料库，可用于聊天，本地查询，机器控制等，如在饮食相关的领域和在音乐相关的领域，对于同样的问题：“请问人气高的地方”，则应答会有所差异，反馈的是分别是餐厅和音乐厅，网络语料则是以网络查询的结果作为语料。

可选的，采用 AIML 模型进行配置语料数据库，AIML 人工智能标记语言，AIML 描述一种称为 AIML 对象的数据对象的类 (class)，并部分描述了处理他们的计算机程序的行为 (behavior)。AIML 对象是由所谓的主题 (topics) 和类型 (categories) 的单元组成，其中包含任何解析或未解析的数据。其中，解析数据 (Parsed data) 是字符解析得到的字符数据 (character data)、和 AIML 元素 (AIML elements)。AIML 元素封装在包含在文档中的刺激-反应知识 (stimulus-response knowledge)，在这些元素中的字符数据 (Character data) 可以由 AIML 编译器 (AIML interpreter) 解析，也可以由响应器 (Responder)

解析。本实施例的 Solr 搜索引擎输入的是自然语言，输出对应的回答也是自然语言。在一个实施方式中，通过配置相应的问题模板来配置语料数据库，模板形式采用 AIML 模板形式进行编写，多组预设问题的回答语句通过网上爬取获得，如百度百科等，并根据问题的类型与相应的网址相链接。

图 3 是本申请实施例中识别自然语言的流程图，如图 3 所示，使用与所述对话语境对应的语料数据库识别所述自然语言包括：

S302，查询与所述对话语境对应的本地语料数据库，其中，所述本地语料数据库包括：多轮对话语料，多场景语料；

本实施例中的对话语境包括知识搜索，互联网搜索，天气查询，路线导航，对话聊天，指令控制等，其中，对话聊天为多轮对话语料，知识搜索，互联网搜索，天气查询，路线导航，指令控制为多场景语料，在多场景对话语料中，每一个对话语境的场景都对应一个专用的语料数据库，但是在多场景语料识别失败时，会转为多轮对话语料的语料库，进一步询问自然语言的实际含义。在一个示例中，采集到的自然语言为：“导航到前面写字楼”，确定为路线导航的对话语境，在本地语料数据库中查询“前面写字楼”没有成功，进而转到多轮对话语料，触发问题：“请问是哪一栋写字楼？”，用户回答：“A 写字楼”。

S304，使用所述本地语料数据库在本地识别所述自然语言；

可选的，在使用与所述对话语境对应的语料数据库识别所述自然语言之后，还包括：输出与识别结果对应的应答内容，或，生成与识别结果对应的控制指令。

将自然语言转换为语料数据库可以识别的机器语言后，在本地语料数据库查找对应的语义，在查找失败时，使用 solr 搜索引擎进行搜索，识别出语义后，进一步搜索出该语义下的应答内容，或者是针对目标对象的控制指令，如采集到自然语言为“今天的天气怎样”，可以搜索出天气状况，并输出应

答内容“今天天气晴，24℃，空气指数良好”，或者采集到的语言为“电视声音太大了”，匹配出用于调低电视音量的控制指令，并发送给电视的控制台。如果是在语料数据库识别出来的，则在预料数据库查找，如果是通过搜索引擎识别出来的，则通过搜索引擎查找。当然，也可以使用独立的数据库（该数据库保存了与每个语义对应的应答内容或控制指令）。图4是本申请实施例交互示意图，本地内置了互联网搜索，天气查询，路线导航三个语境的语料数据库。

S306，在使用所述本地语料数据库在本地识别所述自然语言失败时，调用所述 solr 搜索引擎识别所述自然语言，并在所述本地语料数据库中添加与所述自然语言对应的识别结果。

在使用 solr 搜索引擎也识别失败时，在采用所有语料数据库识别用户的自然语言都失败时，反馈错误，由用户来协助完成对话或操作，并进一步完善语料数据库，如将结果反馈到语料数据库，在本地语料库中修改或添加当前问答的识别内容，学习识别失败的自然语言。

在本实施例的实施方式中，采集自然语言包括：通过麦克风采集背景噪声的语音信号；解析所述背景噪声和所述语音信号的声线特性；根据历史语音的声线特性去掉所述背景噪声，将所述语音信号转换为文字，得到所述自然语言，其中，所述历史语音为在无噪音条件下采集的样本语音。

在本实施例的方案在应用在聊天机器人上时，基于国际机器人的通用语言 AIML 的模型进行优化，由于 AIML 的标签配置非常灵活，可以支持自定义场景的配置，多轮对话语料，应用于多种场景，可以让机器人拥有更智能化，更专业化的对话功能；后台增加 solr 搜索引擎可以对接更为丰富的语料资源，并且适应基于语义的快速搜索和对话功能；机器人具备学习功能，可以在与人的对话中学习之前不能理解的语句与对答，拥有自我成长功能。

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根据

上述实施例的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）执行本申请各个实施例所述的方法。

实施例 2

在本实施例中还提供了一种语音的识别装置，该装置用于实现上述实施例及优选实施方式，已经进行过说明的不再赘述。如以下所使用的，术语“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现，但是硬件，或者软件和硬件的组合的实现也是可能并被构想的。

图 5 是根据本申请实施例的语音的识别装置的结构框图，如图 5 所示，该装置包括：采集模块 50，确定模块 52，识别模块 54，其中，

采集模块 50，用于采集自然语言；

确定模块 52，用于确定所述自然语言的对话语境，其中，所述对话语境用于描述当前语音交互的领域和交互对象；

识别模块 54，用于使用与所述对话语境对应的语料数据库识别所述自然语言。

可选的，所述确定模块包括：提取单元，用于提取所述自然语言中的关键字；匹配单元，用于通过所述关键字匹配得到对话语境，其中，所述对话语境包括以下之一：知识搜索，互联网搜索，天气查询，路线导航，对话聊天，指令控制。

可选的，所述确定模块包括：查找单元，用于查找所述自然语言之前的历史对话记录；确定单元，用于根据所述历史对话记录确定所述对话语境。

可选的，所述装置还包括：设置模块，用于在所述识别模块使用与所述对话语境对应的语料数据库识别所述自然语言之前，设置所述语料数据库，其中，所述语料数据库包括：多轮对话语料，多场景语料，通过 solr 搜索引擎得到的网络语料。

可选的，所述识别模块包括：查询单元，用于查询与所述对话语境对应的本地语料数据库，其中，所述本地语料数据库包括：多轮对话语料，多场景语料；第一识别单元，用于使用所述本地语料数据库在本地识别所述自然语言；第二识别单元，用于在使用所述本地语料数据库在本地识别所述自然语言失败时，调用所述 solr 搜索引擎识别所述自然语言，并在所述本地语料数据库中添加与所述自然语言对应的识别结果。

可选的，所述采集模块包括：采集单元，用于通过麦克风采集背景噪声的语音信号；解析单元，用于解析所述背景噪声和所述语音信号的声线特性；降噪单元，用于根据历史语音的声线特性去掉所述背景噪声，将所述语音信号转换为文字，得到所述自然语言，其中，所述历史语音为在无噪音条件下采集的样本语音。

可选的，所述装置还包括：处理单元，用于在所述识别模块使用与所述对话语境对应的语料数据库识别所述自然语言之后，输出与识别结果对应的应答内容，或，生成与识别结果对应的控制指令。

需要说明的是，上述各个模块是可以通过软件或硬件来实现的，对于后者，可以通过以下方式实现，但不限于此：上述模块均位于同一处理器中；或者，上述各个模块以任意组合的形式分别位于不同的处理器中。

实施例 3

在本申请所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可

以有另外的划分方式，例如，多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。

上述以软件功能单元的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中，包括若干指令用以使得一台计算机装置（可以是个人计算机，服务器，或者网络装置等）或处理器（Processor）执行本申请各个实施例所述方法的部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory, ROM）、随机存取存储器（Random Access Memory, RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

本申请的实施例还提供了一种存储介质，该存储介质中存储有计算机程序，其中，该计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

可选地，在本实施例中，上述存储介质可以被设置为存储用于执行以下步骤的计算机程序：

S1，采集自然语言；

S2, 确定所述自然语言的对话语境, 其中, 所述对话语境用于描述当前语音交互的领域和交互对象;

S3, 使用与所述对话语境对应的语料数据库识别所述自然语言。

可选地, 在本实施例中, 上述存储介质可以包括但不限于: U 盘、只读存储器 (Read-Only Memory, 简称为 ROM)、随机存取存储器 (Random Access Memory, 简称为 RAM)、移动硬盘、磁碟或者光盘等各种可以存储计算机程序的介质。

本申请的实施例还提供了一种电子装置, 包括存储器和处理器, 该存储器中存储有计算机程序, 该处理器被设置为运行计算机程序以执行上述任一项方法实施例中的步骤。

可选地, 上述电子装置还可以包括传输设备以及输入输出设备, 其中, 该传输设备和上述处理器连接, 该输入输出设备和上述处理器连接。

可选地, 在本实施例中, 上述处理器可以被设置为通过计算机程序执行以下步骤:

S1, 采集自然语言;

S2, 确定所述自然语言的对话语境, 其中, 所述对话语境用于描述当前语音交互的领域和交互对象;

S3, 使用与所述对话语境对应的语料数据库识别所述自然语言。

以上所述仅为本申请的较佳实施例而已, 并不用以限制本申请, 凡在本申请的精神和原则之内, 所做的任何修改、等同替换、改进等, 均应包含在本申请保护的范围之内。

权利要求书

1. 一种语音的识别方法，其中，所述方法包括：

采集自然语言；

确定所述自然语言的对话语境，其中，所述对话语境用于描述当前语音交互的领域和交互对象；

使用与所述对话语境对应的语料数据库识别所述自然语言；

其中，在使用与所述对话语境对应的语料数据库识别所述自然语言之前，还包括：设置所述语料数据库，其中，所述语料数据库包括：多轮对话语料，多场景语料，通过 solr 搜索引擎得到的网络语料；

其中，使用与所述对话语境对应的语料数据库识别所述自然语言包括：查询与所述对话语境对应的本地语料数据库，其中，所述本地语料数据库包括：多轮对话语料，多场景语料；使用所述本地语料数据库在本地识别所述自然语言；在使用所述本地语料数据库在本地识别所述自然语言失败时，调用所述 solr 搜索引擎识别所述自然语言，并在所述本地语料数据库中添加与所述自然语言对应的识别结果。

2. 根据权利要求 1 所述的方法，确定所述自然语言的对话语境包括：

提取所述自然语言中的关键字；

通过所述关键字匹配得到对话语境，其中，所述对话语境包括以下之一：知识搜索，互联网搜索，天气查询，路线导航，对话聊天，指令控制。

3. 根据权利要求 1 所述的方法，确定所述自然语言的对话语境包括：

查找所述自然语言之前的历史对话记录；

根据所述历史对话记录确定所述对话语境。

4. 根据权利要求 1 所述的方法，采集自然语言包括：

通过麦克风采集背景噪声的语音信号；

解析所述背景噪声和所述语音信号的声线特性；

根据历史语音的声线特性去掉所述背景噪声，将所述语音信号转换为文字，得到所述自然语言，其中，所述历史语音为在无噪音条件下采集的样本语音。

5. 根据权利要求1所述的方法，在使用与所述对话语境对应的语料数据库识别所述自然语言之后，所述方法还包括：

输出与识别结果对应的应答内容，或，生成与识别结果对应的控制指令。

6. 根据权利要求5所述的方法，在生成与识别结果对应的控制指令之后，所述方法还包括：

建立与控制对象的通信链路；

通过所述通信链路向所述控制对象发送所述控制指令。

7. 根据权利要求1所述的方法，在所述solr搜索引擎识别所述自然语言失败时，所述方法还包括：

反馈识别错误的指示信息，并获取用户输入的协助识别结果；

将所述协助识别结果反馈到语料数据库，在本地语料库中修改或添加当前问答的识别内容。

8. 一种语音的识别装置，其中，所述装置包括：

采集模块，用于采集自然语言；

确定模块，用于确定所述自然语言的对话语境，其中，所述对话语境用于描述当前语音交互的领域和交互对象；

识别模块，用于使用与所述对话语境对应的语料数据库识别所述自然语言。

9. 根据权利要求8所述的装置，所述确定模块包括：

提取单元，用于提取所述自然语言中的关键字；

匹配单元，用于通过所述关键字匹配得到对话语境，其中，所述对话语境包括以下之一：知识搜索，互联网搜索，天气查询，路线导航，对话聊天，指令控制；

其中，所述装置还包括：设置模块，用于在所述识别模块使用与所述对

话语境对应的语料数据库识别所述自然语言之前，设置所述语料数据库，其中，所述语料数据库包括：多轮对话语料，多场景语料，通过 solr 搜索引擎得到的网络语料。

其中，所述识别模块包括：查询单元，用于查询与所述对话语境对应的本地语料数据库，其中，所述本地语料数据库包括：多轮对话语料，多场景语料；第一识别单元，用于使用所述本地语料数据库在本地识别所述自然语言；第二识别单元，用于在使用所述本地语料数据库在本地识别所述自然语言失败时，调用所述 solr 搜索引擎识别所述自然语言，并在所述本地语料数据库中添加与所述自然语言对应的识别结果。

10. 根据权利要求 8 所述的装置，所述确定模块包括：

查找单元，用于查找所述自然语言之前的历史对话记录；

确定单元，用于根据所述历史对话记录确定所述对话语境。

11. 根据权利要求 8 所述的装置，所述装置还包括：

建立模块，用于在所述识别模块生成与识别结果对应的控制指令之后，建立与控制对象的通信链路；

发送模块，用于通过所述通信链路向所述控制对象发送所述控制指令。

12. 根据权利要求 8 所述的装置，所述识别模块包括：

处理模块，用于在所述识别模块通过所述 solr 搜索引擎识别所述自然语言失败时，反馈识别错误的指示信息，并获取用户输入的协助识别结果；

学习模块，用于将所述协助识别结果反馈到语料数据库，在本地语料库中修改或添加当前问答的识别内容。

13. 根据权利要求 8 所述的装置，所述采集模块包括：

采集单元，用于通过麦克风采集背景噪声的语音信号；

解析单元，用于解析所述背景噪声和所述语音信号的声线特性；

降噪单元，用于根据历史语音的声线特性去掉所述背景噪声，将所述语

音信号转换为文字，得到所述自然语言，其中，所述历史语音为在无噪音条件下采集的样本语音。

14. 根据权利要求 8 所述的装置，所述装置还包括：

处理单元，用于在所述识别模块使用与所述对话语境对应的语料数据库识别所述自然语言之后，输出与识别结果对应的应答内容，或，生成与识别结果对应的控制指令。

15. 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一种语音的识别方法的步骤，包括：

采集自然语言；

确定所述自然语言的对话语境，其中，所述对话语境用于描述当前语音交互的领域和交互对象；

使用与所述对话语境对应的语料数据库识别所述自然语言；

其中，在使用与所述对话语境对应的语料数据库识别所述自然语言之前，还包括：设置所述语料数据库，其中，所述语料数据库包括：多轮对话语料，多场景语料，通过 solr 搜索引擎得到的网络语料；

其中，使用与所述对话语境对应的语料数据库识别所述自然语言包括：查询与所述对话语境对应的本地语料数据库，其中，所述本地语料数据库包括：多轮对话语料，多场景语料；使用所述本地语料数据库在本地识别所述自然语言；在使用所述本地语料数据库在本地识别所述自然语言失败时，调用所述 solr 搜索引擎识别所述自然语言，并在所述本地语料数据库中添加与所述自然语言对应的识别结果。

16. 根据权利要求 15 所述的计算机设备，确定所述自然语言的对话语境包括：

提取所述自然语言中的关键字；

通过所述关键字匹配得到对话语境，其中，所述对话语境包括以下之一：知识搜索，互联网搜索，天气查询，路线导航，对话聊天，指令控制。

17. 根据权利要求 15 所述的计算机设备, 确定所述自然语言的对话语境包括:
查找所述自然语言之前的历史对话记录;

根据所述历史对话记录确定所述对话语境。

18. 一种计算机存储介质, 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一种语音的识别方法的步骤, 包括:

采集自然语言;

确定所述自然语言的对话语境, 其中, 所述对话语境用于描述当前语音交互的领域和交互对象;

使用与所述对话语境对应的语料数据库识别所述自然语言;

其中, 在使用与所述对话语境对应的语料数据库识别所述自然语言之前, 还包括: 设置所述语料数据库, 其中, 所述语料数据库包括: 多轮对话语料, 多场景语料, 通过 solr 搜索引擎得到的网络语料;

其中, 使用与所述对话语境对应的语料数据库识别所述自然语言包括: 查询与所述对话语境对应的本地语料数据库, 其中, 所述本地语料数据库包括: 多轮对话语料, 多场景语料; 使用所述本地语料数据库在本地识别所述自然语言; 在使用所述本地语料数据库在本地识别所述自然语言失败时, 调用所述 solr 搜索引擎识别所述自然语言, 并在所述本地语料数据库中添加与所述自然语言对应的识别结果。

19. 根据权利要求 18 所述的计算机存储介质, 确定所述自然语言的对话语境包括:

提取所述自然语言中的关键字;

通过所述关键字匹配得到对话语境, 其中, 所述对话语境包括以下之一: 知识搜索, 互联网搜索, 天气查询, 路线导航, 对话聊天, 指令控制。

20. 根据权利要求 18 所述的计算机存储介质, 确定所述自然语言的对话语境包括:

查找所述自然语言之前的历史对话记录；
根据所述历史对话记录确定所述对话语境。

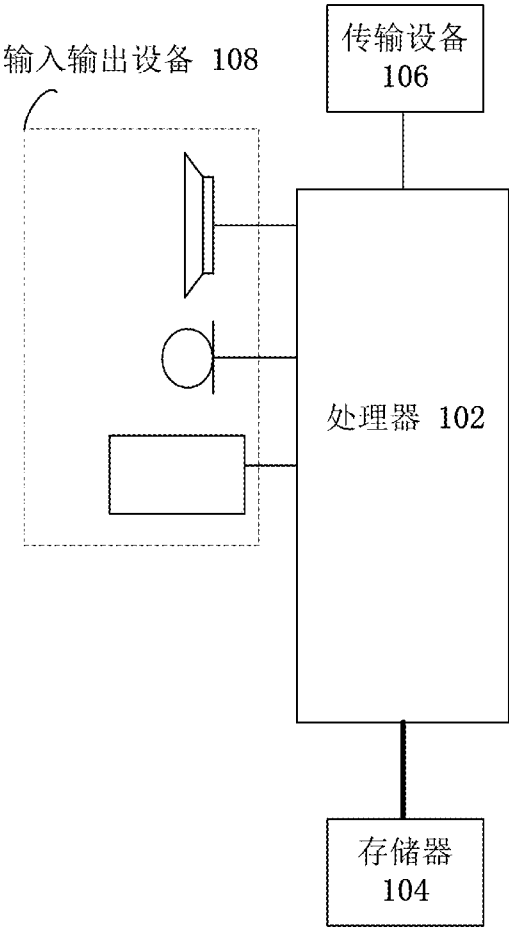


图 1

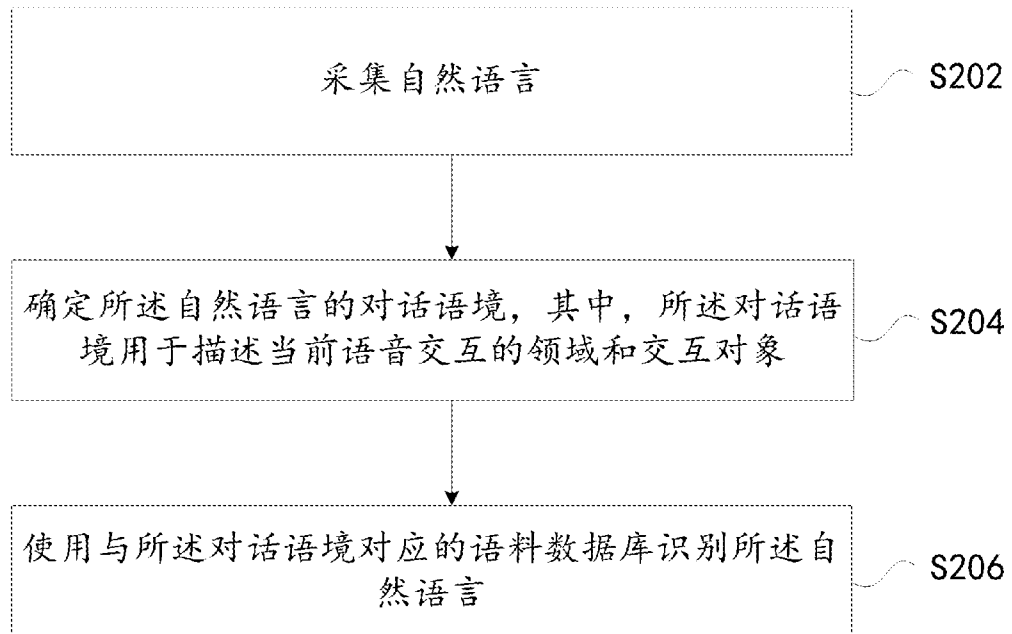


图 2

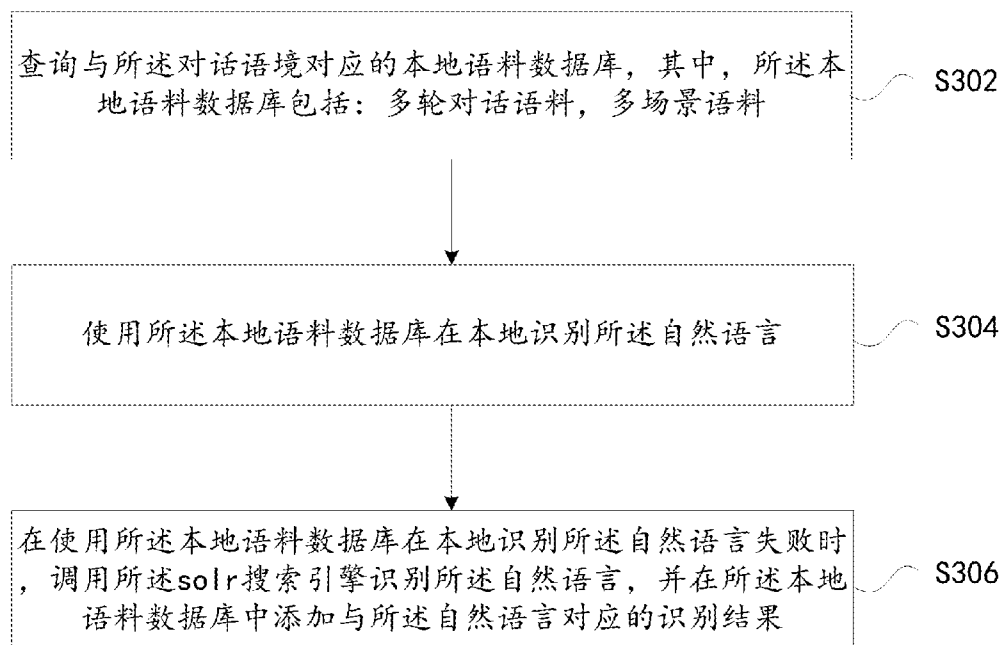


图 3

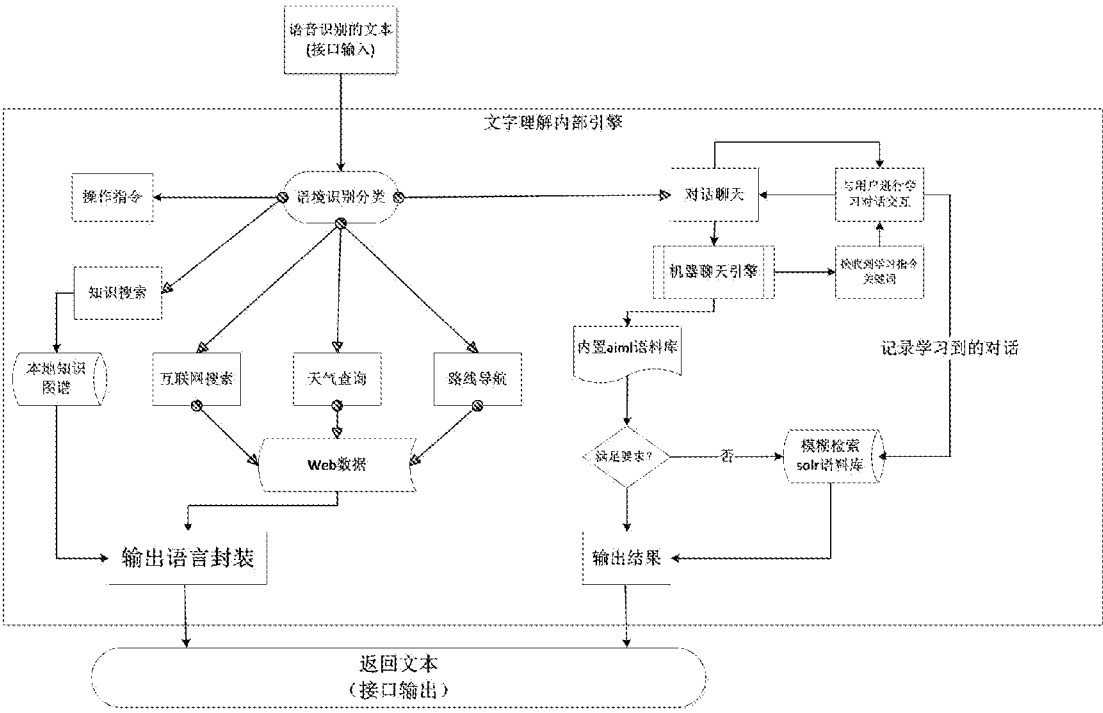


图 4

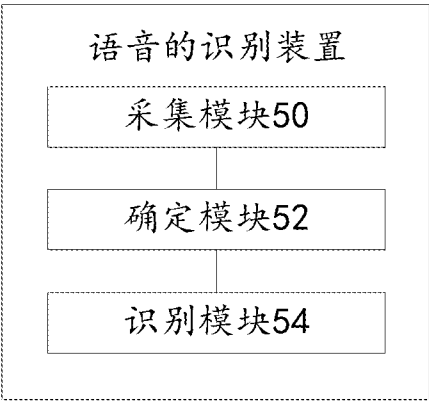


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/118415

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/22(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 G06F

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)

CNPAT, CNKI, WPI, EPODOC: 语音 交互 对话 场景 数据库 搜索引擎 本地 服务器 voice interaction search database local sever solr

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110265013 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 20 September 2019 (2019-09-20) description, paragraphs 31-96	1-20
X	CN 109299320 A (SHANGHAI ZHIZHEN INTELLIGENT NETWORK TECHNOLOGY CO., LTD.) 01 February 2019 (2019-02-01) description, paragraphs 30-50	1-20
X	CN 106201424 A (BEIJING GAN WEI ROBOT TECHNOLOGY CO., LTD.) 07 December 2016 (2016-12-07) description, paragraphs 84-115	1-20
Y	CN 109902158 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 18 June 2019 (2019-06-18) description, paragraphs 4-57	1-20
Y	CN 106297782 A (BEIJING INTELLIGENT STEWARD CO., LTD.) 04 January 2017 (2017-01-04) description, paragraphs 62-105	1-20
A	US 2015019221 A1 (CHUNGHWA PICTURE TUBES, LTD.) 15 January 2015 (2015-01-15) 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

19 February 2020

Date of mailing of the international search report

26 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
Information on patent family members

International application No.

PCT/CN2019/118415

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
CN	110265013	A	20 September 2019	None		
CN	109299320	A	01 February 2019	None		
CN	106201424	A	07 December 2016	None		
CN	109902158	A	18 June 2019	None		
CN	106297782	A	04 January 2017	None		
US	2015019221	A1	15 January 2015	TW	201503105	A 16 January 2015

国际检索报告

国际申请号

PCT/CN2019/118415

A. 主题的分类 G10L 15/22 (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 G06F 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)) CNPAT, CNKI, WPI, EPDOC: 语音 交互 对话 场景 数据库 搜索引擎 本地 服务器 voice interaction search database local sever solr																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 110265013 A (平安科技深圳有限公司) 2019年 9月 20日 (2019 - 09 - 20) 说明书第31-96段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 109299320 A (上海智臻智能网络科技股份有限公司) 2019年 2月 1日 (2019 - 02 - 01) 说明书第30-50段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 106201424 A (北京甘为乐博科技有限公司) 2016年 12月 7日 (2016 - 12 - 07) 说明书第84-115段</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 109902158 A (平安科技深圳有限公司) 2019年 6月 18日 (2019 - 06 - 18) 说明书第4-57段</td> <td>1-20</td> </tr> <tr> <td>Y</td> <td>CN 106297782 A (北京智能管家科技有限公司) 2017年 1月 4日 (2017 - 01 - 04) 说明书第62-105段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2015019221 A1 (CHUNGHWA PICTURE TUBES, LTD.) 2015年 1月 15日 (2015 - 01 - 15) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 110265013 A (平安科技深圳有限公司) 2019年 9月 20日 (2019 - 09 - 20) 说明书第31-96段	1-20	X	CN 109299320 A (上海智臻智能网络科技股份有限公司) 2019年 2月 1日 (2019 - 02 - 01) 说明书第30-50段	1-20	X	CN 106201424 A (北京甘为乐博科技有限公司) 2016年 12月 7日 (2016 - 12 - 07) 说明书第84-115段	1-20	Y	CN 109902158 A (平安科技深圳有限公司) 2019年 6月 18日 (2019 - 06 - 18) 说明书第4-57段	1-20	Y	CN 106297782 A (北京智能管家科技有限公司) 2017年 1月 4日 (2017 - 01 - 04) 说明书第62-105段	1-20	A	US 2015019221 A1 (CHUNGHWA PICTURE TUBES, LTD.) 2015年 1月 15日 (2015 - 01 - 15) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 110265013 A (平安科技深圳有限公司) 2019年 9月 20日 (2019 - 09 - 20) 说明书第31-96段	1-20																					
X	CN 109299320 A (上海智臻智能网络科技股份有限公司) 2019年 2月 1日 (2019 - 02 - 01) 说明书第30-50段	1-20																					
X	CN 106201424 A (北京甘为乐博科技有限公司) 2016年 12月 7日 (2016 - 12 - 07) 说明书第84-115段	1-20																					
Y	CN 109902158 A (平安科技深圳有限公司) 2019年 6月 18日 (2019 - 06 - 18) 说明书第4-57段	1-20																					
Y	CN 106297782 A (北京智能管家科技有限公司) 2017年 1月 4日 (2017 - 01 - 04) 说明书第62-105段	1-20																					
A	US 2015019221 A1 (CHUNGHWA PICTURE TUBES, LTD.) 2015年 1月 15日 (2015 - 01 - 15) 全文	1-20																					
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
<table border="0"> <tr> <td> * 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的在先申请或专利 “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 </td> <td> “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 </td> </tr> </table>			* 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的在先申请或专利 “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																			
* 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的在先申请或专利 “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 2月 19日		国际检索报告邮寄日期 2020年 2月 26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		受权官员 方婷 电话号码 86-010-53961654																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/118415

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	110265013	A	2019年 9月 20日	无	
CN	109299320	A	2019年 2月 1日	无	
CN	106201424	A	2016年 12月 7日	无	
CN	109902158	A	2019年 6月 18日	无	
CN	106297782	A	2017年 1月 4日	无	
US	2015019221	A1	2015年 1月 15日	TW 201503105 A	2015年 1月 16日