(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 11 月 26 日 (26.11.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/233363 A1

(51) 国际专利分类号:
G10L 15/20 (2006.01) *G10L 15/26* (2006.01)

(21) 国际申请号: PCT/CN2020/087471

(22) 国际申请日: 2020 年 4 月 28 日 (28.04.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910430228.7 2019年5月22日 (22.05.2019) CN

(71) 申请人: 深圳壹账通智能科技有限公司 (ONE CONNECT SMART TECHNOLOGY CO., LTD (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(72) 发明人: 曹绪文 (CAO, Xuwen); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201室, Guangdong 518000 (CN)。

(74) 代理人: 深圳市隆天联鼎知识产权代理有限公司 (SHENZHEN LUNGTIN LIANDING INTELLECTUAL PROPERTY AGENT LTD.); 中国广东省深圳市福田区南园路上田大厦4A 刘抗美, Guangdong 518000 (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,

WO 2020/233363 A1

(54) **Title:** SPEECH RECOGNITION METHOD AND DEVICE, ELECTRONIC APPARATUS, AND STORAGE MEDIUM

(54) 发明名称: 语音识别的方法、装置、电子设备和存储介质

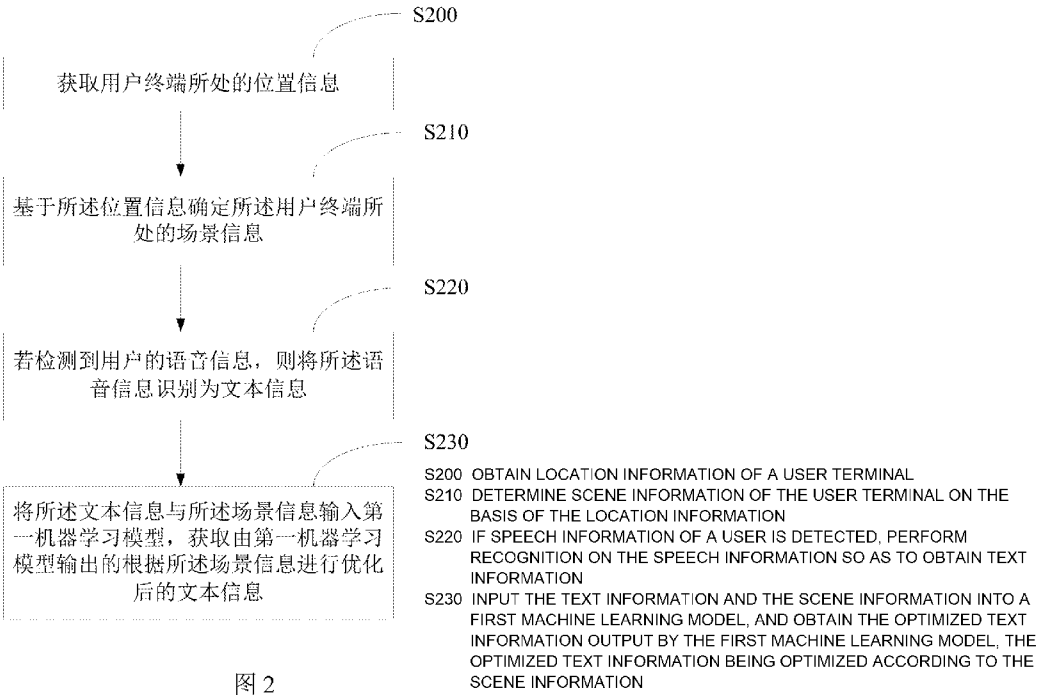


图 2

(57) **Abstract:** A speech recognition method and device (900), an electronic apparatus (1000), and a storage medium, relating to the field of biometrics. The method comprises: obtaining location information of a user terminal (110) (S200); determining scene information of the user terminal (110) on the basis of the location information (S210); if speech information of a user is detected, performing recognition on the speech information so as to obtain text information (S220); and inputting the text information and the scene information into a first machine learning model, and obtaining the optimized text information output by the first machine learning model, the optimized

[见续页]

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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))。

text information being optimized according to the scene information (S230). The invention improves the accuracy of speech recognition.

(57) 摘要：一种语音识别的方法、装置（900）、电子设备（1000）和存储介质，涉及生物识别领域，方法包括：获取用户终端（110）所处的位置信息（S200）；基于位置信息确定用户终端（110）所处的场景信息（S210）；若检测到用户的语音信息，则将语音信息识别为文本信息（S220）；将文本信息与场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据场景信息进行优化后的文本信息（S230）。提高了语音识别的准确率。

语音识别的方法、装置、电子设备和存储介质

技术领域

[0001] 本申请要求于2019年5月22日提交中国专利局、申请号为201910430228.7，发明名称为“语音识别的方法、装置、电子设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

[0002] 本申请涉及人工智能的生物识别领域，特别是涉及一种语音识别的方法、装置、电子设备和存储介质。

背景技术

[0003] 随着智能技术的发展，进行语音识别并根据识别的语音进行控制已成为智能技术应用中的一项重要内容，各种智能产品中也会应用语音识别技术实现智能化控制，而随着智能产品的增加以及对语音识别的准确度的要求越来越高，各种语音识别技术层出不穷。

[0004] 目前常用的语音识别方式是通过提取用户发出的待识别语音信息的特征，再根据识别算法对该用户发出的待识别语音信息进行识别。

发明概述

技术问题

[0005] 然而，发明人意识到在一些场景（如马路上）使用语音识别功能捕获到的待识别语音除了包含有人的语音之外，还会包含有诸如汽车鸣笛等杂音，其中人的语音信息是有效待识别的语音，而汽车鸣笛声等杂音是噪音，语音识别时会将汽车鸣笛等噪声也一同识别出来，进而导致语音识别准确性不足。

问题的解决方案

技术解决方案

[0006] 本申请实施例的目的在于提供一种语音识别方法、装置、计算机可读介质及电子设备，进而可以至少在一定程度上克服现有技术中语音识别准确率低的问题。

[0007] 根据本申请的第一方面，提供了一种语音识别的方法，包括：获取用户终端所

处的位置信息；基于所述位置信息确定所述用户终端所处的场景信息；若检测到用户的语音信息，则将所述语音信息识别为文本信息；将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[0008] 根据本申请的第二方面提供了一种语音识别装置，包括：第一获取模块：用于获取用户终端所处的位置信息；确定模块：用于基于所述位置信息确定所述用户终端所处的场景信息；识别模块：用于若检测到用户的语音信息，则将所述语音信息识别为文本信息；第二获取模块：用于将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[0009] 根据本申请的第三方面提供了一种语音识别的电子设备，包括：存储器，配置为存储可执行指令；处理器，配置为执行存储器中存储的可执行指令，以执行一种语音识别方法：

[0010] 其中，所述语音识别方法包括：

[0011] 获取用户终端所处的位置信息；

[0012] 基于所述位置信息确定所述用户终端所处的场景信息；

[0013] 若检测到用户的语音信息，则将所述语音信息识别为文本信息；

[0014] 将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[0015] 根据本申请的第四方面提供了一种计算机可读存储介质，其存储有计算机程序指令，当所述计算机指令被计算机执行时，使计算机执行语音识别方法：

[0016] 其中，所述语音识别方法包括：

[0017] 获取用户终端所处的位置信息；

[0018] 基于所述位置信息确定所述用户终端所处的场景信息；

[0019] 若检测到用户的语音信息，则将所述语音信息识别为文本信息；

[0020] 将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

发明的有益效果

有益效果

[0021] 在本申请的一些实施例中，通过获取用户终端所处的位置信息，基于所述位置信息确定所述用户终端所处的位置，将所述用户终端所处的位置与数据库中预存的位置与场景对应关系表确定所述用户终端设备所处的场景信息，当所述用户终端检测到有语音信息输入时将所述语音信息识别为文本信息，再将所述文本信息与所述场景信息一起输入预设的第一机器学习模型中，由第一机器学习模型将所述文本信息中包含的所述场景信息对应的噪声对应的文本信息滤除，获得优化后的文本信息。可见，本申请实施例，可快速且准确的滤除待识别语音对应的文本信息中包含的所述场景对应的噪声对应的文本信息；进而提高语音识别的准确性。

对附图的简要说明

附图说明

[0022] 图1示出根据本申请一示例实施方式的语音识别方法的使用环境的系统架构图。

[0023] 图2示出根据本公开一示例实施方式的语音识别方法的流程图。

[0024] 图3示出根据本公开一示例实施方式的基于所述位置信息确定所述用户终端所处的场景信息的详细流程图。

[0025] 图4示出根据本公开一示例实施方式的在若检测到用户的语音信息，则将所述语音信息识别为文本信息之前的流程图。

[0026] 图5示出根据本公开一示例实施方式的若检测到用户的语音信息，则将所述语音信息识别为文本信息的详细流程图。

[0027] 图6示出根据本公开一示例实施方式的在将所述文本信息和所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息之前的流程图。

[0028] 图7示出根据本公开一示例实施方式的在获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息之后的流程图。

[0029] 图8示出根据本公开一示例实施方式的在将所述优化后的文本信息和所述场景信息输入预设的意图识别模型之前的流程图。

[0030] 图9示出根据本公开一示例实施方式的语音识别装置的结构框图

[0031] 图10示出根据本公开一示例实施方式的语音识别的电子设备图。

[0032] 图11示出根据本公开一示例实施方式的语音识别的计算机可读存储介质图。

发明实施例

本发明的实施方式

[0033] 本申请涉及人工智能的语音处理技术领域，具体地，图1示出根据本公开一示例实施方式的语音识别方法的使用环境的构架图：

[0034] 该使用环境包括用户终端110、服务器120、数据库130。

[0035] 应该理解，图1中的用户终端、服务器和数据库的数目仅仅是示意性的。根据实现需要，可以具有任意数目的用户终端、服务器和数据库。比如服务器120可以是多个服务器组成的服务器集群等。

[0036] 在一实施例中服务器120通过用户终端110内置的全球定位系统（GPS）获取用户终端110所处的位置信息，服务器120基于用户终端110所处的位置信息确定用户终端110所处的位置，服务器120基于用户终端110所处位置与数据库130中预存的位置与场景信息对应关系确定用户终端110对应的场景信息，当用户终端110检测到有用户语音输入时，用户终端110将获取的用户语音信息发送至服务器120，服务器120将用户终端110发送的语音信息识别为文本信息，服务器120基于用户终端对应的场景信息和用户终端110发送的语音信息对应的文本信息，将用户终端发送的语音信息对应的文本信息中的用户终端110所处场景对应的噪声对应的文本信息滤除，输出优化后的文本信息，进而来提高语音识别的准确性。

[0037] 需要说明的是，本申请实施例所提供的数据处理方法一般由服务器120执行，相应地，数据处理装置一般设置于服务器120中。但是，在本申请的其它实施例中，终端也可以与服务器具有相似的功能，从而执行本申请实施例所提供的数据处理方案。

[0038] 图2示出根据本公开一示例实施方式的语音识别方法的流程图，可以包括如下步骤：

[0039] 步骤S200：获取用户终端所处的位置信息；

[0040] 步骤S210：基于所述位置信息确定所述用户终端所处的场景信息；

[0041] 步骤S220：若检测到用户的语音信息，则将所述语音信息识别为文本信息；

[0042] 步骤S230：将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[0043] 下面，将结合附图对本示例实施方式中上述数据字典生成的各步骤进行详细的解释以及说明。

[0044] 在步骤S200中：获取用户终端所处的位置信息。

[0045] 在一实施例中，所述位置信息是指表明所处的地方或地址的信息，通过获取用户终端所处的位置信息来确定用户终端所处的位置，进而确定用户终端对应的场景信息及场景信息对应的噪声，以便于后续的第一机器学习模型将所述文本信息中的噪声对应的文本信息滤除，调高语音识别的准确性。

[0046] 在一实施例中，位置信息可以是通过用户终端设备中内置的GPS模块获取的定位信息，也可以是用户终端设备的所有者在使用该用户终端设备时输入的表明位置的文本信息。

[0047] 继续参照图2所示，在步骤S210中：基于所述位置信息确定所述用户终端所处的场景信息。

[0048] 在一实施例中，所述场景信息是指表明所处情景的信息，通过用户所处的位置信息来确定该用户终端所处的位置，进而确定该用户终端所处的场景，再基于所述用户终端所处场景来确定通过用户终端设备获取的语音信息可能存在的噪声及噪声对应的文本信息。

[0049] 在一实施例中，如图3所示，步骤S210可以包括：

[0050] 步骤S2101：基于所述位置信息，确定所述用户终端所处位置；

[0051] 步骤S2102：基于所述用户终端所处位置，以及数据库中预存的位置与场景的对应关系，确定所述用户终端所处的场景信息。

[0052] 在一实施例中，用户终端所处的位置信是通过用户终端内置的GPS模块获取用户终端的定位信息，基于该GPS定位信息，确定用户终端所处位置为“学院路与创新路交叉口”，基于该用户终端所处的位置“学院路与创新路交叉口”，查询预存在数据库中标明各场景信息的地图，确定该用户终端所处的位置“学院路与创新路交叉口”对应的场景为“马路旁”。

[0053] 在一实施例中，用户终端所处的位置信息还可以是由该用户终端设备所有者在安装放置该用户终端设备时输入的。

[0054] 继续参照图2所示，在步骤S220中：若检测到用户的语音信息，则将所述语音信息识别为文本信息。

[0055] 在一实施例中，在检测到用户语音信息的时候，确定该用户语音信息对应的分贝值，基于该用户语音信息对应的分贝值的判断结果，来判断是否对该用户语音进行识别，以防止当用户终端处于较为嘈杂的场景时，将非目标用户的语音识别为文本信息，从而进一步降低对目标用户语音识别的准确率，同时也使用户终端在未检测到预设分贝值范围内的语音时处于待机状态，降低用户终端的耗电以节约资源。

[0056] 在一实施例中，可以通过第二机器学习模型来将语音信息识别为文本信息，在这种情况下需要事先对机器学习模型进行训练，具体训练过程如图4所示，可以包括如下步骤：

[0057] 步骤S410：接收用户自定义的语音片段，以及与所述语音片段对应的文本内容；

[0058] 步骤S420：识别所述语音片段的语素特征，根据所述语素特征和与所述语音片段对应的文本内容生成第二机器学习模型的训练样本；

[0059] 步骤S430：通过所述训练样本对所述第二机器学习模型进行训练生成语音识别模型，以基于该语音识别模型将所述语音信息识别为文本信息。

[0060] 通过将用户自定义的语音片段及所述自定义语音片段对应的文本信息作为训练样本生成机器学习模型的方式，来对所述用户输入的语音信息进行识别，既可提高对所述用于输入语音识别的准确率，又能满足所述用户的个性化需求。

[0061] 在一实施例中，获取的用户自定义语音片段是“you”，对应的文本信息是“无”等自定义语音片段与对应的文本信息， 将获取的自定义语音片段与对应的文本信息作为样本对第二机器学习模型进行训练生成语音识别模型，当获取的用户语音为“you”时，语音识别模型会将“you”识别为对应的文本信息“无”，以满足用户的个性化需求。

[0062] 在一时候例中，获取的用户自定义语音片段是“nongshalei”等带有浓重地方

方言的语音，对应的文本信息为“干什么呢”等地方方言对应的普通话文本信息，将获取的自定义语音片段与对应的文本信息作为训练样本对第二机器学习模型进行训练生成语音识别模型，当获取的用户语音为“nongshalei”等地方方言时，则会将该地方方言语音识别为对应的“干什么呢”等普通话文本信息，以提高对不同方言识别的准确率。

[0063] 在一实施例中，如图5所示，步骤S220还可以通过如下流程将语音信息识别为文本信息：

[0064] 步骤S2201：获取用户终端选定的自定义语音识别集；

[0065] 步骤S2202：将所述语音信息与所述用户终端选定的自定义语音识别集中包含的语音片段进行比对；

[0066] 步骤S2203：若所述语音信息与所述用户终端选定的自定义语音识别集中包含的目标语音片段相匹配，则将所述用户终端选定的自定义语音识别集中包含的所述目标语音片段对应的文本信息作为识别到的文本信息。

[0067] 通过将获取的用户语音与用户终端选定的自定义语音识别集中的语音片段进行匹配，来确定获取的用户语音对应的文本信息，可降低因语种的繁杂与同一语种间各地方言的差异化带来的语音识别的难度，，进而提高语音识别的准确率，同时也能提高对一些喜欢将不同语种语言或地方方言混杂使用的用户对应的语音识别的准确率。

[0068] 在一实施例中，获取的用户语音为“woverydiao”，该获取的用户语音“woverydiao”是一个混杂中英文及地方方言的语音信息，对于目前常见的语音识别模型很难识别该获取的用户语音，在本申请中，通过将该获取的用户语音与用户终端选定的自定义语音识别集中包含的语音片段进行匹配，发现存在与该获取的用户语音相同的目标语音片段，该目标语音片段对应的文本信息为“我很强”，则将该目标语音片段对应的文本信息“我很强”识别为该获取的用户语音“wovreydiao”的文本信息，这样即满足了不同用户的个性化需求又提高了对该用户语音识别的准确率。

[0069] 继续参照图2所示，在步骤S230中：将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后

的文本信息。

[0070] 在上述实施例中，通过机器学习模型对所述文本信息进行优化，滤除所述文本信息中包含的噪声，既可以提高语音识别的准确率又可短时间内快速处理大量的语音信息对应的数据。

[0071] 在一实施例中，获取优化后的文本信息还可通过，基于用户终端对应的场景信息，确定该场景信息对应的噪声文本信息，将获取的语音信息对应的文本信息中包含的该场景信息对应的噪声文本信息滤除，进而提高语音识别的准确率。

[0072] 在一实施例中，可以基于所述文本信息和所述场景信息通过第一机器学习模型的方式获取优化后的文本信息，在这种情况下需要事先对机器学习模型进行训练，具体训练过程如图6所示， 包括如下步骤：

[0073] 步骤S610：获取预先设置的语音信息样本集合中每个语音信息样本对应的文本信息和场景信息；

[0074] 步骤S620：确定所述语音信息样本集合中每个语音信息样本对应的优化后的文本信息；

[0075] 步骤S630：将获取的所述语音信息样本对应的文本信息和场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息，将所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息进行比对，如不一致，则调整所述第一机器学习模型的参数，直至所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息一致。

[0076] 在一实施例中，如图7所示，在图2所示的各个步骤之后，本申请实施例提供的语音识别的方法还可以包括如下步骤。

[0077] 步骤S240：将所述优化后的文本信息和所述场景信息输入预设的意图识别模型，获取由所述意图识别模型输出的所述语音信息所包含的意图信息。

[0078] 在一实施例中，意图信息是指语音表达的需求及目的，图7所示的实施例的技术方案可通过获取用户终端所处的场景信息与用户终端获取语音对应的文本信息，确定用户终端获取的语音信息对应的意图信息，进而按照用户终端获取的语音信息对应的意图信息执行相应指令，如当用户A晚上回家后，在房间B说“

开灯”，通常屋内的智能家居管理系统会将屋内的照明灯全部打开，但通过本申请的技术方案，基于接收到用户A“开灯”语音的用户终端确定用户A所处的场景为“房间B”，确定用户A“开灯”语音对应的意图信息为“打开房间B”的灯，则会将该意图信息发送至屋内的智能家居管理系统来打开房间B的灯，从而提高对用户语音指令的判断准确性，提高用户体验度。

[0079] 在一实施例中，可以基于所述优化后的文本信息和所述场景信息通过意图识别模型的方式获取所述意图信息，这种情况下需要事先对意图识别模型进行训练，具体训练过程如图8所示，可以包括如下步骤：：

[0080] 步骤S810：获取预先设置的语音信息样本集合中每个语音信息样本对应的优化后的文本信息和场景信息；

[0081] 步骤S820：确定所述语音信息样本对应的意图信息；

[0082] 步骤S830：将获取的所述语音信息样本对应的优化后的文本信息和场景信息输入意图识别模型，获取由所述意图识别模型输出的意图信息，将所述意图识别模型输出的意图信息与确定的所述语音信息样本对应的意图信息进行比对，如不一致则调整所述意图识别模型的参数，直至所述意图识别模型输出的意图信息与确定的所述语音样本对应的意图信息相一致。

[0083] 本申请还提供了一种语音识别的装置。参考图9所示，所述语音识别的装置900包括：第一获取模块910，确定模块920，识别模块930，第二获取模块940，其中：

[0084] 第一获取模块910，用于获取用户终端所处的位置信息；

[0085] 确定模块920，用于基于所述位置信息确定所述用户终端所处的场景信息；

[0086] 识别模块930，用于若检测到用户的语音信息，则将所述语音信息识别为文本信息；

[0087] 第二获取模块940，用于将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[0088] 在一实施例中，所述确定模块920还可配置为：基于所述位置信息，确定所述用户终端所处位置，基于所述用户终端所处位置，以及数据库中预存的位置与

场景的对应关系，确定所述用户终端所处的场景信息。

[0089] 在一实施例中，所述语音识别装置还包括：第二机器学习模型训练模块，用于接收用户自定义的语音片段，以及与所述语音片段对应的文本内容，识别所述语音片段的语素特征，根据所述语素特征和与所述语音片段对应的文本内容生成第二机器学习模型的训练样本，通过所述训练样本对所述第二机器学习模型进行训练生成语音识别模型，以基于该语音识别模型将所述语音信息识别为文本信息。

[0090] 在一实施例中，所述识别模块930还可配置为：用于获取用户终端选定的自定义语音识别集，将所述语音信息与所述用户终端选定的自定义语音识别集中包含的语音片段进行比对，若所述语音信息与所述用户终端选定的自定义语音识别集中包含的目标语音片段相匹配，则将所述用户终端选定的自定义语音识别集中包含的所述目标语音片段对应的文本信息作为识别到的文本信息。

[0091] 在一实施例中，所述语音识别装置还包括：第一机器学习模型训练模块，用于获取预先设置的语音信息样本集合中每个语音信息样本对应的文本信息和场景信息，确定所述语音信息样本集合中每个语音信息样本对应的优化后的文本信息，将获取的所述语音信息样本对应的文本信息和场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息，将所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息进行比对，如不一致，则调整所述第一机器学习模型的参数，直至所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息一致。

[0092] 在一实施例中，所述语音识别装置还包括：意图识别模块，用于获取预先设置的语音信息样本集合中每个语音信息样本对应的优化后的文本信息和场景信息。

[0093] 在一实施例中所述语音识别模块还包括：意图识别模型训练模块，用于获取预先设置的语音信息样本集合中每个语音信息样本对应的优化后的文本信息和场景信息，确定所述语音信息样本对应的意图信息，将获取的所述语音信息样本对应的优化后的文本信息和场景信息输入意图识别模型，获取由所述意图识别

模型输出的意图信息，将所述意图识别模型输出的意图信息与确定的所述语音信息样本对应的意图信息进行比对，如不一致则调整所述意图识别模型的参数，直至所述意图识别模型输出的意图信息与确定的所述语音样本对应的意图信息相一致。

[0094] 上述语音识别的装置中各模块的具体细节已经在对应的方法中进行了详细的描述，因此此处不再赘述。

[0095] 应当注意，尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元，但是这种划分并非强制性的。实际上，根据本公开的实施方式，上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之，上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

[0096] 此外，尽管在附图中以特定顺序描述了本公开中方法的各个步骤，但是，这并非要求或者暗示必须按照所述特定顺序来执行这些步骤，或是必须执行全部所示的步骤才能实现期望的结果。附加的或备选的，可以省略某些步骤，将多个步骤合并为一个步骤执行，以及/或者将一个步骤分解为多个步骤执行等。

[0097] 通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本公开实施方式的技术方案可以以软件产品的形式体现出来，所述软件产品可以存储在一个非易失性存储介质（可以是CD-ROM，U盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、移动终端、或者网络设备等）执行根据本公开实施方式的方法。

[0098] 在本申请的示例性实施例中，还提供了一种能够实现上述方法的电子设备。

[0099] 所属技术领域的技术人员能够理解，本申请的各个方面可以实现为系统、方法或程序产品。因此，本申请的各个方面可以具体实现为以下形式，即：完全的硬件实施方式、完全的软件实施方式（包括固件、微代码等），或硬件和软件方面结合的实施方式，这里可以统称为“电路”、“模块”或“系统”。

[0100] 下面参照图10来描述根据本申请的这种实施方式的电子设备1000。图0显示的电子设备1000仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任

何限制。

[0101] 如图10所示，电子设备1000以通用计算设备的形式表现。电子设备1000的组件可以包括但不限于：上述至少一个处理单元1010、上述至少一个存储单元1020、连接不同系统组件（包括存储单元1020和处理单元1010）的总线1030。

[0102] 其中，所述存储单元存储有程序代码，所述程序代码可以被所述处理单元1010执行，使得所述处理单元1010执行本说明书上述“示例性方法”部分中描述的根据本申请各种示例性实施方式的步骤。例如，所述处理单元1010可以执行如图2中所示步骤S200：获取用户终端所处的位置信息；步骤S210：基于所述位置信息确定所述用户终端所处的场景信息；步骤S220：若检测到用户的语音信息，则将所述语音信息识别为文本信息；步骤S230：将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息；

[0103] 存储单元1020可以包括易失性存储单元形式的可读介质，例如随机存取存储单元（RAM）10201和/或高速缓存存储单元10202，还可以进一步包括只读存储单元（ROM）10203。

[0104] 存储单元1020还可以包括具有一组（至少一个）程序模块10205的程序/实用工具10204，这样的程序模块10205包括但不限于：操作系统、一个或者多个应用程序、其它程序模块以及程序数据，这些示例中的每一个或某种组合中可能包括网络环境的实现。

[0105] 总线1030可以为表示几类总线结构中的一种或多种，包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

[0106] 电子设备1000也可以与一个或多个外部设备500（例如键盘、指向设备、蓝牙设备等）通信，还可与一个或者多个使得用户能与该电子设备1000交互的设备通信，和/或与使得该电子设备1000能与一个或多个其它计算设备进行通信的任何设备（例如路由器、调制解调器等等）通信。这种通信可以通过输入/输出（I/O）接口1050进行。并且，电子设备1000还可以通过网络适配器1060与一个或者多个网络（例如局域网（LAN），广域网（WAN）和/或公共网络，例如因特网

）通信。如图所示，网络适配器1060通过总线1030与电子设备1000的其它模块通信。应当明白，尽管图中未示出，可以结合电子设备1000使用其它硬件和/或软件模块，包括但不限于：微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0107] 通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本申请实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是CD-ROM，U盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、终端装置、或者网络设备等）执行根据本申请实施方式的方法。

[0108] 在本申请的示例性实施例中，还提供了一种计算机可读存储介质，所述计算机可读存储介质为易失性存储介质或非易失性存储介质，其上存储有能够实现本说明书上述方法的程序产品。在一些可能的实施方式中，本申请的各个方面还可以实现为一种程序产品的形式，其包括程序代码，当所述程序产品在终端设备上运行时，所述程序代码用于使所述终端设备执行本说明书上述“示例性方法”部分中描述的根据本申请各种示例性实施方式的步骤。

[0109] 参考图11所示，描述了根据本申请的实施方式的用于实现上述方法的程序产品1100，其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码，并可以在终端设备，例如个人电脑上运行。然而，本申请的程序产品不限于此，在本文件中，可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0110] 所述程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM或闪存）、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0111] 计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了可读程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0112] 可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于无线、有线、光缆、RF等等，或者上述的任意合适的组合。

[0113] 可以以一种或多种程序设计语言的任意组合来编写用于执行本申请操作的程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如Java、C++等，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中，远程计算设备可以通过任意种类的网络，包括局域网（LAN）或广域网（WAN），连接到用户计算设备，或者，可以连接到外部计算设备（例如利用因特网服务提供商来通过因特网连接）。

[0114] 此外，上述附图仅是根据本申请示例性实施例的方法所包括的处理的示意性说明，而不是限制目的。易于理解，上述附图所示的处理并不表明或限制这些处理的时间顺序。另外，也易于理解，这些处理可以是例如在多个模块中同步或异步执行的。

权利要求书

[权利要求 1] 一种语音识别方法，其中，所述方法包括：

获取用户终端所处的位置信息；

基于所述位置信息确定所述用户终端所处的场景信息；

若检测到用户的语音信息，则将所述语音信息识别为文本信息；

将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[权利要求 2] 根据权利要求1所述的方法，其中，所述基于所述位置信息确定所述用户终端所处的场景信息包括：

基于所述位置信息，确定所述用户终端所处位置；

基于所述用户终端所处位置，以及数据库中预存的位置与场景的对应关系，确定所述用户终端所处的场景信息。

[权利要求 3] 根据权利要求1所述的方法，其中，在所述将所述语音信息识别为文本信息之前，所述方法还包括：

接收用户自定义的语音片段，以及与所述语音片段对应的文本内容；

识别所述语音片段的语素特征，根据所述语素特征和与所述语音片段对应的文本内容生成第二机器学习模型的训练样本；

通过所述训练样本对所述第二机器学习模型进行训练生成语音识别模型，以基于该语音识别模型将所述语音信息识别为文本信息。

[权利要求 4] 根据权利要求1所述的方法，其中，所述将所述语音信息识别为文本信息包括：

获取用户终端选定的自定义语音识别集；

将所述语音信息与所述用户终端选定的自定义语音识别集中包含的语音片段进行比对；

若所述语音信息与所述用户终端选定的自定义语音识别集中包含的目标语音片段相匹配，则将所述用户终端选定的自定义语音识别集中包含的所述目标语音片段对应的文本信息作为识别到的文本信息。

[权利要求 5] 根据权利要求1所述的方法，其中，在所述将所述文本信息和所述场

景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息之前，所述方法还包括：

获取预先设置的语音信息样本集合中每个语音信息样本对应的文本信息和场景信息；

确定所述语音信息样本集合中每个语音信息样本对应的优化后的文本信息；

将获取的所述语音信息样本对应的文本信息和场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息，将所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息进行比对，如不一致，则调整所述第一机器学习模型的参数，直至所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息一致。

[权利要求 6] 根据权利要求1至5中任一项所述的方法，其中，在所述获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息之后，所述方法还包括：

将所述优化后的文本信息和所述场景信息输入预设的意图识别模型，获取由所述意图识别模型输出的所述语音信息所包含的意图信息。

[权利要求 7] 根据权利要求6所述的方法，其中，在所述将所述优化后的文本信息和所述场景信息输入预设的意图识别模型之前，所述方法还包括：

获取预先设置的语音信息样本集合中每个语音信息样本对应的优化后的文本信息和场景信息；

确定所述语音信息样本对应的意图信息；

将获取的所述语音信息样本对应的优化后的文本信息和场景信息输入意图识别模型，获取由所述意图识别模型输出的意图信息，将所述意图识别模型输出的意图信息与确定的所述语音信息样本对应的意图信息进行比对，如不一致则调整所述意图识别模型的参数，直至所述意图识别模型输出的意图信息与确定的所述语音样本对应的意图信息相一致。

[权利要求 8] 一种语音识别装置，其中，包括：

第一获取模块：用于获取用户终端所处的位置信息；

确定模块：用于基于所述位置信息确定所述用户终端所处的场景信息；

识别模块：用于若检测到用户的语音信息，则将所述语音信息识别为文本信息；

第二获取模块：用于将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[权利要求 9] 一种语音识别的电子设备，其中，包括：

存储器，配置为存储可执行指令；

处理器，配置为执行存储器中存储的可执行指令，以实现一种语音识别方法：

其中，所述语音识别方法包括：

获取用户终端所处的位置信息；

基于所述位置信息确定所述用户终端所处的场景信息；

若检测到用户的语音信息，则将所述语音信息识别为文本信息；

将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[权利要求 10] 根据权利要求9所述的电子设备，其中，所述基于所述位置信息确定所述用户终端所处的场景信息包括：

基于所述位置信息，确定所述用户终端所处位置；

基于所述用户终端所处位置，以及数据库中预存的位置与场景的对应关系，确定所述用户终端所处的场景信息。

[权利要求 11] 根据权利要求9所述的电子设备，其中，在所述将所述语音信息识别为文本信息之前，所述方法还包括：

接收用户自定义的语音片段，以及与所述语音片段对应的文本内容；

识别所述语音片段的语素特征，根据所述语素特征和与所述语音片段

对应的文本内容生成第二机器学习模型的训练样本；

通过所述训练样本对所述第二机器学习模型进行训练生成语音识别模型，以基于该语音识别模型将所述语音信息识别为文本信息。

[权利要求 12] 根据权利要求9所述的电子设备，其中，所述将所述语音信息识别为文本信息包括：

获取用户终端选定的自定义语音识别集；

将所述语音信息与所述用户终端选定的自定义语音识别集中包含的语音片段进行比对；

若所述语音信息与所述用户终端选定的自定义语音识别集中包含的目标语音片段相匹配，则将所述用户终端选定的自定义语音识别集中包含的所述目标语音片段对应的文本信息作为识别到的文本信息。

[权利要求 13] 根据权利要求9所述的电子设备，其中，在所述将所述文本信息和所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息之前，所述方法还包括：

获取预先设置的语音信息样本集合中每个语音信息样本对应的文本信息和场景信息；

确定所述语音信息样本集合中每个语音信息样本对应的优化后的文本信息；

将获取的所述语音信息样本对应的文本信息和场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息，将所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息进行比对，如不一致，则调整所述第一机器学习模型的参数，直至所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息一致。

[权利要求 14] 根据权利要求9至13中任一项所述的电子设备，其中，在所述获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息之后，所述方法还包括：

将所述优化后的文本信息和所述场景信息输入预设的意图识别模型，

获取由所述意图识别模型输出的所述语音信息所包含的意图信息。

[权利要求 15] 根据权利要求14所述的电子设备，其中，在所述将所述优化后的文本信息和所述场景信息输入预设的意图识别模型之前，所述方法还包括：

获取预先设置的语音信息样本集合中每个语音信息样本对应的优化后的文本信息和场景信息；

确定所述语音信息样本对应的意图信息；

将获取的所述语音信息样本对应的优化后的文本信息和场景信息输入意图识别模型，获取由所述意图识别模型输出的意图信息，将所述意图识别模型输出的意图信息与确定的所述语音信息样本对应的意图信息进行比对，如不一致则调整所述意图识别模型的参数，直至所述意图识别模型输出的意图信息与确定的所述语音样本对应的意图信息相一致。

[权利要求 16] 一种计算机可读存储介质，其中，其存储有计算机程序指令，当所述计算机指令被计算机执行时，使计算机执行语音识别方法：

其中，所述语音识别方法包括：

获取用户终端所处的位置信息；

基于所述位置信息确定所述用户终端所处的场景信息；

若检测到用户的语音信息，则将所述语音信息识别为文本信息；

将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息。

[权利要求 17] 根据权利要求16所述的计算机可读存储介质，其中，所述基于所述位置信息确定所述用户终端所处的场景信息包括：

基于所述位置信息，确定所述用户终端所处位置；

基于所述用户终端所处位置，以及数据库中预存的位置与场景的对应关系，确定所述用户终端所处的场景信息。

[权利要求 18] 根据权利要求16所述的计算机可读存储介质，其中，在所述将所述语音信息识别为文本信息之前，所述方法还包括：

接收用户自定义的语音片段，以及与所述语音片段对应的文本内容；

识别所述语音片段的语素特征，根据所述语素特征和与所述语音片段对应的文本内容生成第二机器学习模型的训练样本；

通过所述训练样本对所述第二机器学习模型进行训练生成语音识别模型，以基于该语音识别模型将所述语音信息识别为文本信息。

[权利要求 19] 根据权利要求16所述的计算机可读存储介质，其中，所述将所述语音信息识别为文本信息包括：

获取用户终端选定的自定义语音识别集；

将所述语音信息与所述用户终端选定的自定义语音识别集中包含的语音片段进行比对；

若所述语音信息与所述用户终端选定的自定义语音识别集中包含的目标语音片段相匹配，则将所述用户终端选定的自定义语音识别集中包含的所述目标语音片段对应的文本信息作为识别到的文本信息。

[权利要求 20] 根据权利要求16所述的计算机可读存储介质，其中，在所述将所述文本信息和所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息之前，所述方法还包括：

获取预先设置的语音信息样本集合中每个语音信息样本对应的文本信息和场景信息；

确定所述语音信息样本集合中每个语音信息样本对应的优化后的文本信息；

将获取的所述语音信息样本对应的文本信息和场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息，将所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息进行比对，如不一致，则调整所述第一机器学习模型的参数，直至所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息一致。

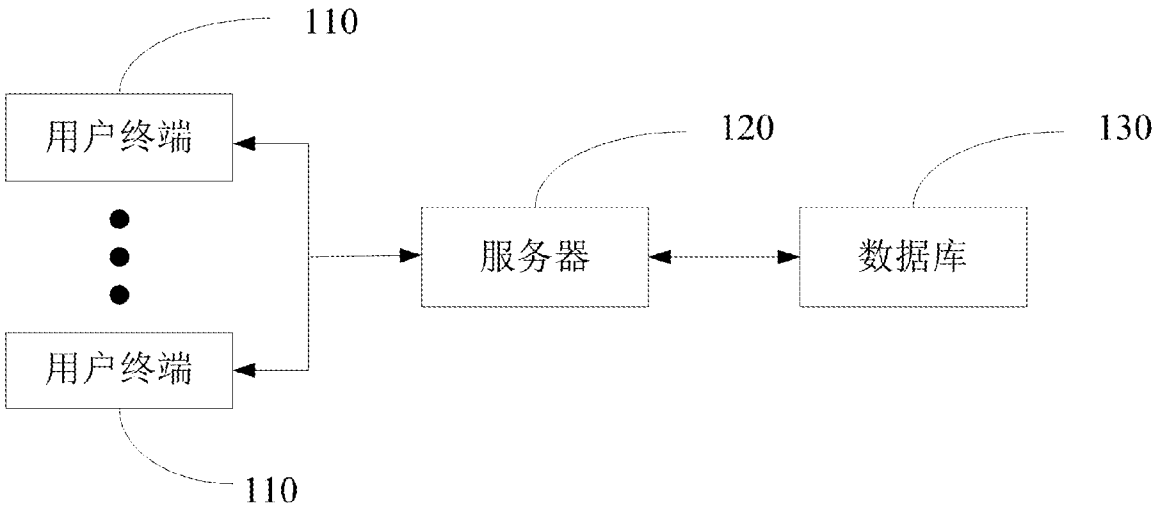

110
用户终端
120
130
服务器
数据库
用户终端
110

图 1

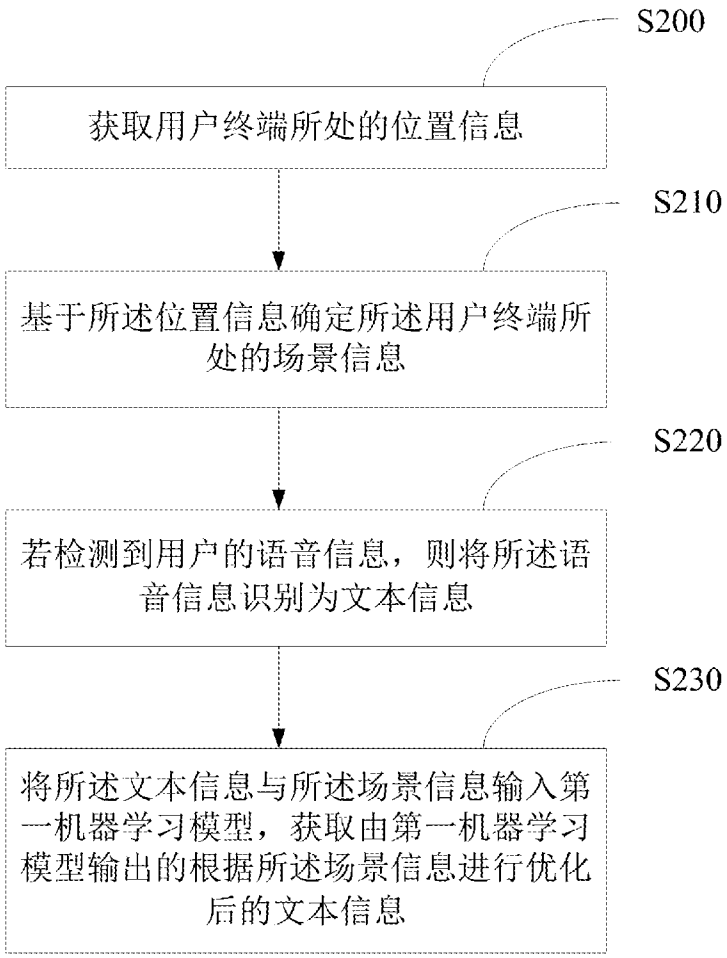

S200
获取用户终端所处的位置信息
S210
基于所述位置信息确定所述用户终端所处的场景信息
S220
若检测到用户的语音信息，则将所述语音信息识别为文本信息
S230
将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息

图 2

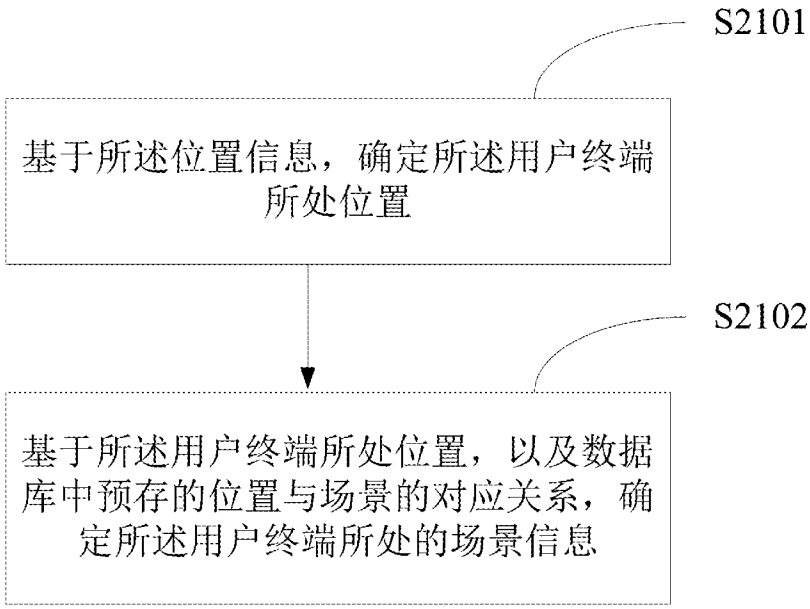

S2101
基于所述位置信息，确定所述用户终端所处位置
S2102
基于所述用户终端所处位置，以及数据库中预存的位置与场景的对应关系，确定所述用户终端所处的场景信息

图 3

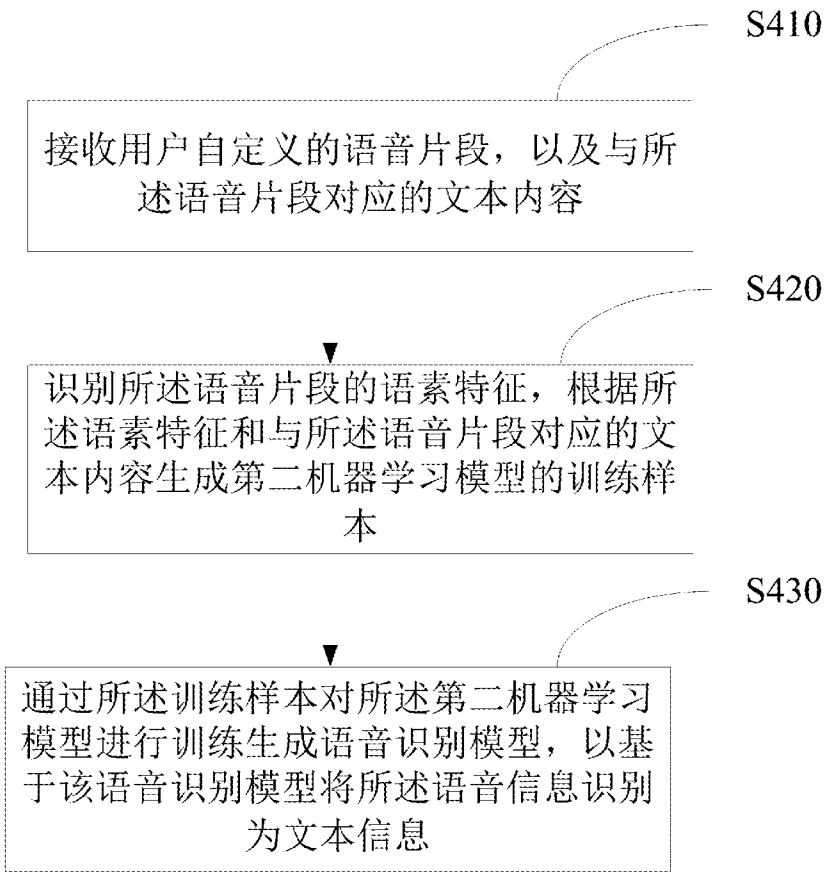

S410
接收用户自定义的语音片段，以及与所述语音片段对应的文本内容
S420
识别所述语音片段的语素特征，根据所述语素特征和与所述语音片段对应的文本内容生成第二机器学习模型的训练样本
S430
通过所述训练样本对所述第二机器学习模型进行训练生成语音识别模型，以基于该语音识别模型将所述语音信息识别为文本信息

图 4

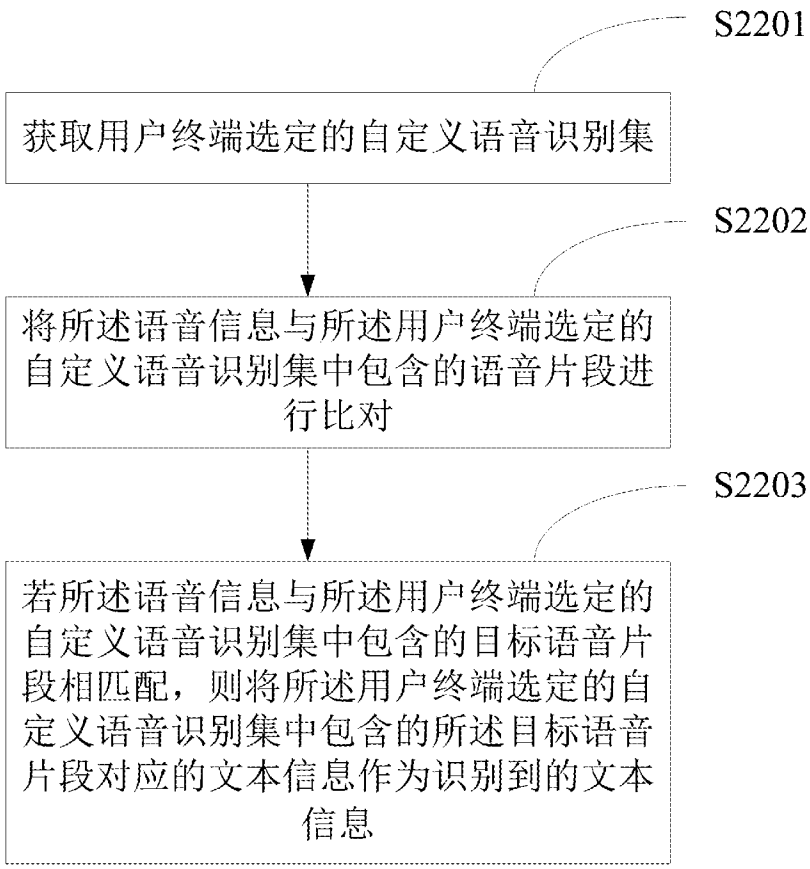

S2201
获取用户终端选定的自定义语音识别集
S2202
将所述语音信息与所述用户终端选定的自定义语音识别集中包含的语音片段进行比对
S2203
若所述语音信息与所述用户终端选定的自定义语音识别集中包含的目标语音片段相匹配，则将所述用户终端选定的自定义语音识别集中包含的所述目标语音片段对应的文本信息作为识别到的文本信息

图 5

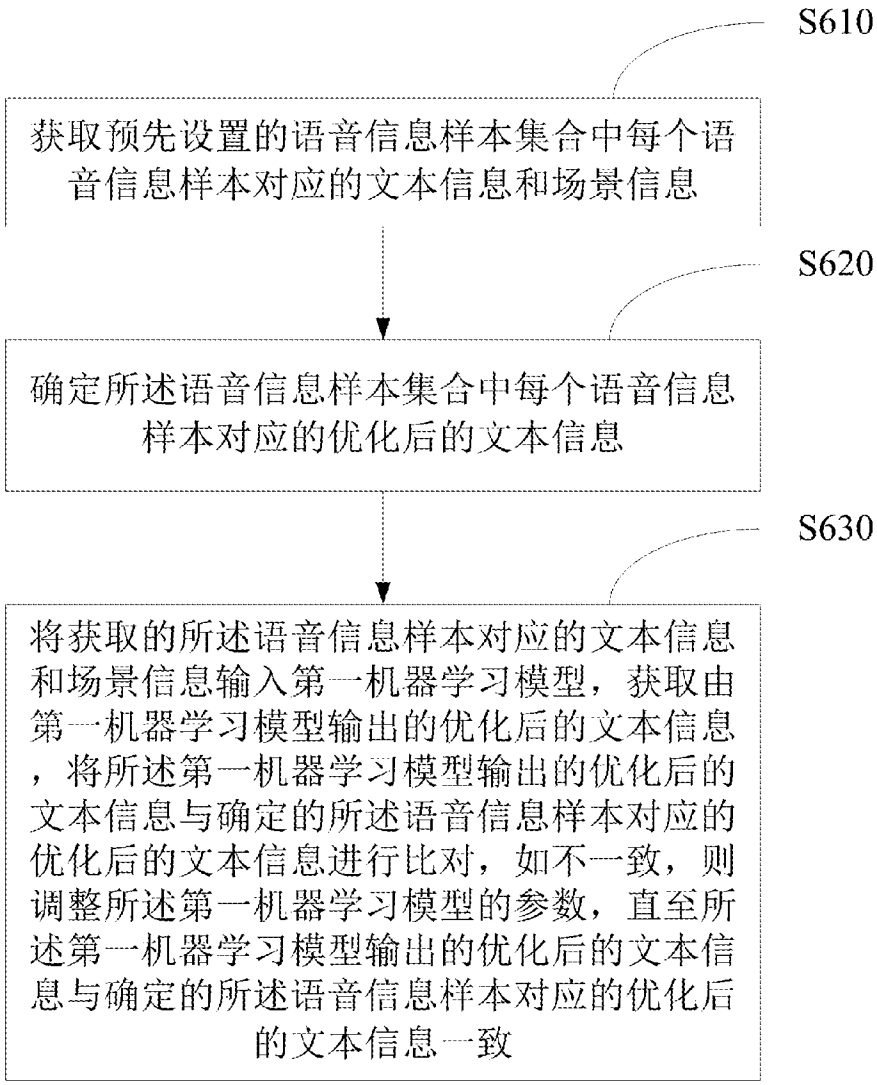

S610
获取预先设置的语音信息样本集合中每个语音信息样本对应的文本信息和场景信息
S620
确定所述语音信息样本集合中每个语音信息样本对应的优化后的文本信息
S630
将获取的所述语音信息样本对应的文本信息和场景信息输入第一机器学习模型，获取由第一机器学习模型输出的优化后的文本信息，将所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息进行比对，如不一致，则调整所述第一机器学习模型的参数，直至所述第一机器学习模型输出的优化后的文本信息与确定的所述语音信息样本对应的优化后的文本信息一致

图 6

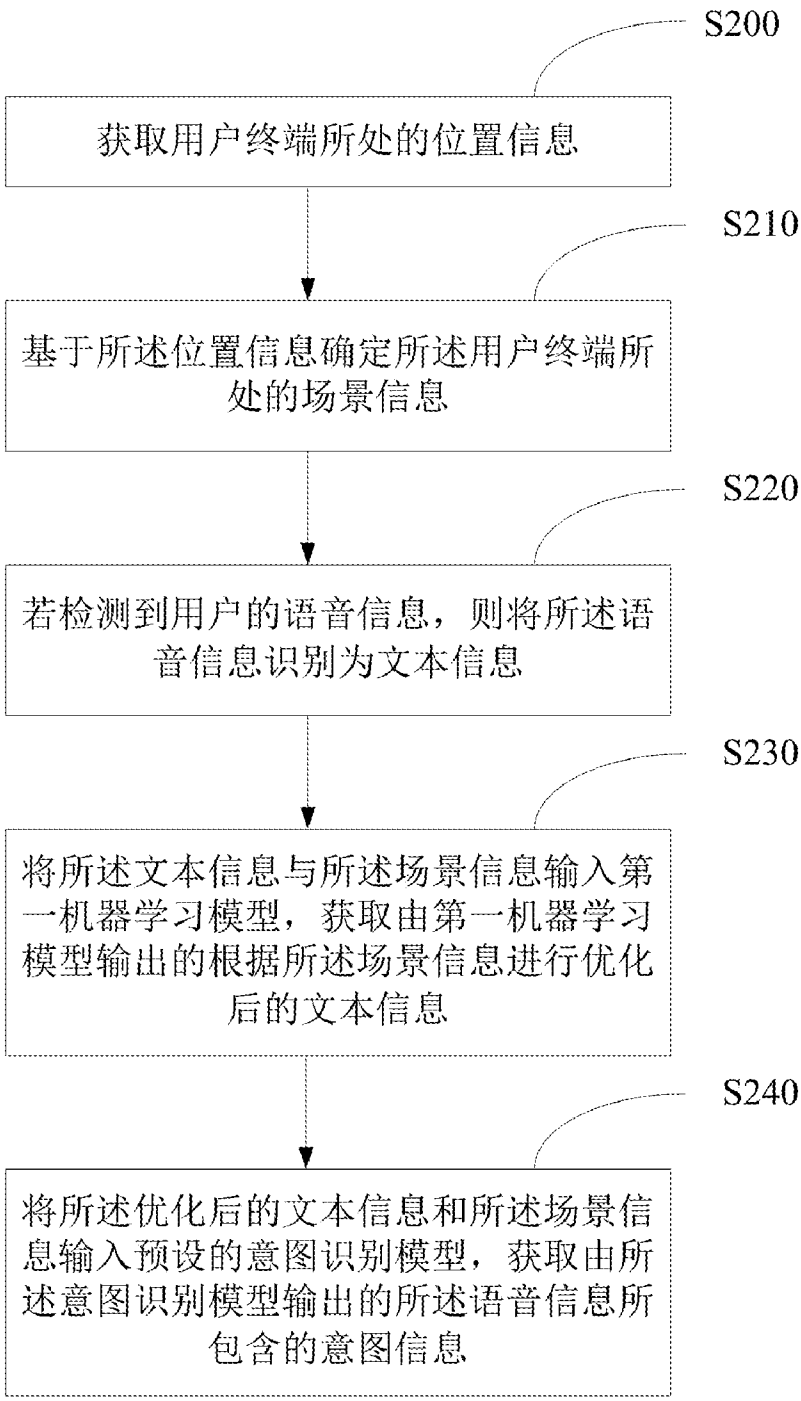

S200
获取用户终端所处的位置信息
S210
基于所述位置信息确定所述用户终端所处的场景信息
S220
若检测到用户的语音信息，则将所述语音信息识别为文本信息
S230
将所述文本信息与所述场景信息输入第一机器学习模型，获取由第一机器学习模型输出的根据所述场景信息进行优化后的文本信息
S240
将所述优化后的文本信息和所述场景信息输入预设的意图识别模型，获取由所述意图识别模型输出的所述语音信息所包含的意图信息

图 7

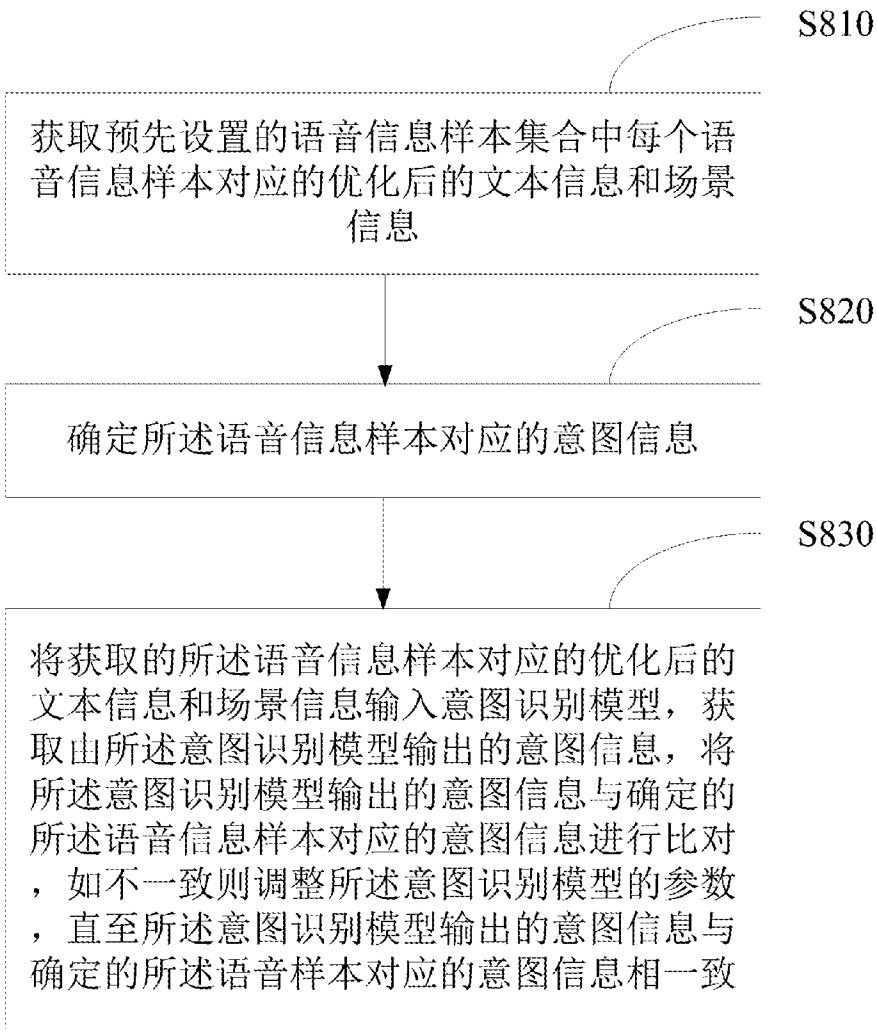

S810
获取预先设置的语音信息样本集合中每个语音信息样本对应的优化后的文本信息和场景信息
S820
确定所述语音信息样本对应的意图信息
S830
将获取的所述语音信息样本对应的优化后的文本信息和场景信息输入意图识别模型，获取由所述意图识别模型输出的意图信息，将所述意图识别模型输出的意图信息与确定的所述语音信息样本对应的意图信息进行比对，如不一致则调整所述意图识别模型的参数，直至所述意图识别模型输出的意图信息与确定的所述语音样本对应的意图信息相一致

图 8

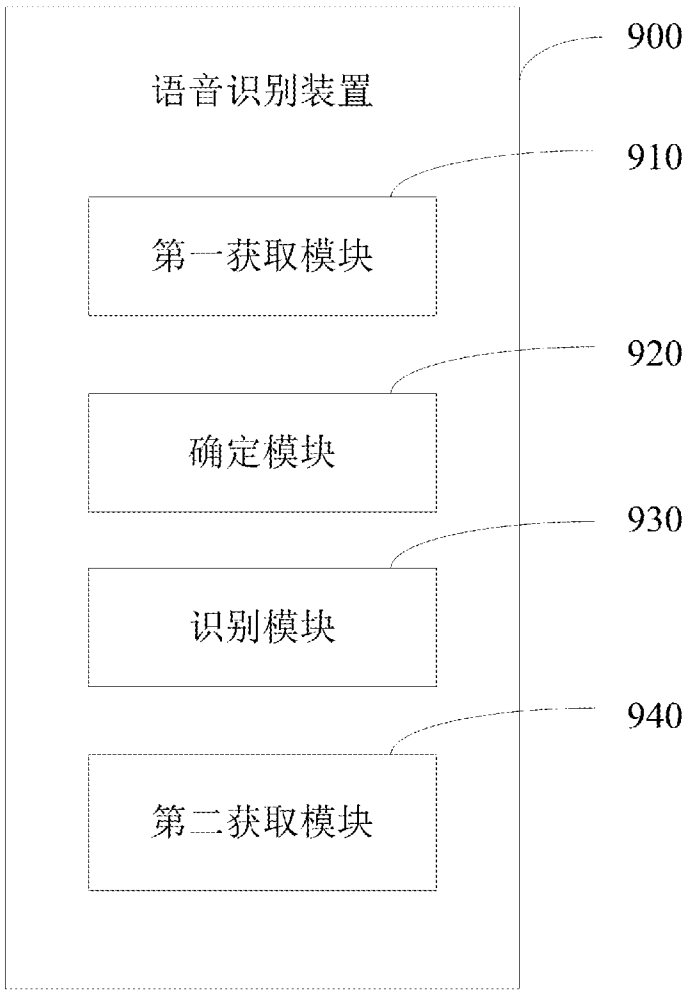


图 9

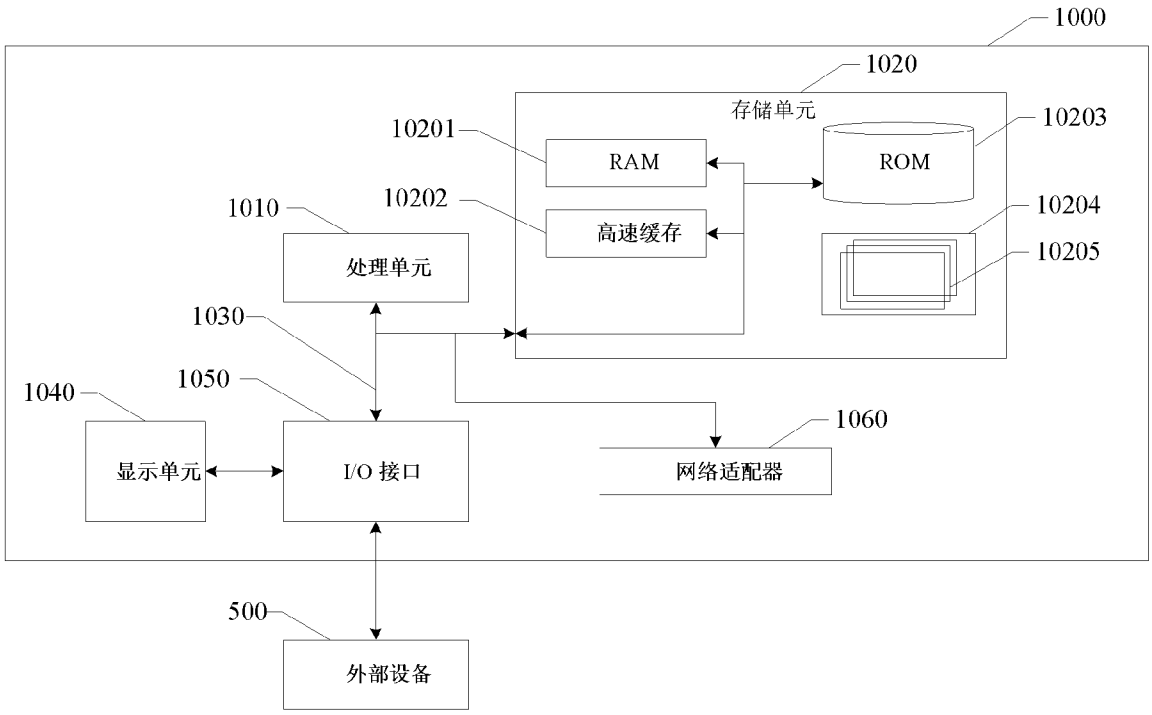

1000
1020
存储单元
10201
RAM
10203
ROM
10202
高速缓存
10204
10205
1010
处理单元
1030
1050
1040
1060
显示单元
I/O 接口
网络适配器
500
外部设备

图 10

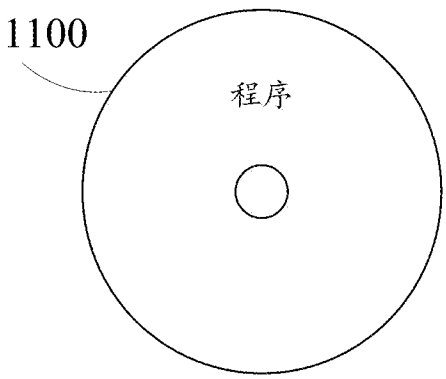

1100
程序

图 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2020/087471

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/20(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)

G10L15

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)

CNKI, CNPAT, WPI, EPODOC: 地理, 坐标, 位置, 场景, 训练, 调整, 参数, 语音识别, 机器学习, 深度学习, 神经网络, 文本, 语义, 语素, 意图, geography, coordinates, location, scene, training, adjustment, parameters, speech recognition, machine learning, deep learning, neural network, text, semantic, morpheme, intention.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110349575 A (SHENZHEN ONECONNECT TECHNOLOGY CO., LTD.) 18 October 2019 (2019-10-18) claims 1-10, description paragraphs [0032]-[0110]	1-20
X	CN 105719649 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 29 June 2016 (2016-06-29) description, paragraphs [0036]-[0094]	1-4, 6-12, 14-19
Y	CN 105719649 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 29 June 2016 (2016-06-29) description, paragraphs [0036]-[0094]	5, 13, 20
Y	CN 106683662 A (CHINA TELECOM CORPORATION LIMITED) 17 May 2017 (2017-05-17) description, paragraphs [0093]-[0101]	5, 13, 20
X	CN 105448292 A (BEIJING YUSHANZHI INFORMATION TECHNOLOGY CO., LTD.) 30 March 2016 (2016-03-30) description, paragraphs [0005]-[0047]	1-4, 6-12, 14-19

☑ 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	Date of mailing of the international search report
28 June 2020	**24 July 2020**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 109509473 A (VIVO COMMUNICATION TECHNOLOGY CO., LTD.) 22 March 2019 (2019-03-22) description, paragraphs [0031]-[0051]	1-4, 6-12, 14-19
X	CN 104240698 A (SHANGHAI BOSHI INFORMATION SCIENCE & TECHNOLOGY CO., LTD.) 24 December 2014 (2014-12-24) description, paragraphs [0003]-[0014]	1, 8, 9, 16
A	US 2015120618 A1 (SAMSUNG ELECTRONICS CO., LTD.) 30 April 2015 (2015-04-30) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2020/087471

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 110349575 A	18 October 2019	None	
CN 105719649 A	29 June 2016	CN 105719649 B	05 July 2019
CN 106683662 A	17 May 2017	None	
CN 105448292 A	30 March 2016	CN 105448292 B	12 March 2019
CN 109509473 A	22 March 2019	None	
CN 104240698 A	24 December 2014	None	
US 2015120618 A1	30 April 2015	US 9948764 B2	17 April 2018
		US 2018227410 A1	09 August 2018
		KR 20150047803 A	06 May 2015

国际检索报告

国际申请号
PCT/CN2020/087471

A. 主题的分类

G10L 15/20(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)

G10L15

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））

CNKI，CNPAT，WPI，EPODOC：地理，坐标，位置，场景，训练，调整，参数，语音识别，机器学习，深度学习，神经网络，文本，语义，语素，意图，geography， coordinates， location， scene， training， adjustment， parameters， speech recognition， machine learning， deep learning， neural network， text， semantic， morpheme， intention。

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 110349575 A (深圳壹账通智能科技有限公司) 2019年 10月 18日 (2019-10-18) 权利要求1-10，说明书第[0032]-[0110]段	1-20
X	CN 105719649 A (百度在线网络技术北京有限公司) 2016年 6月 29日 (2016-06-29) 说明书第[0036]-[0094]段	1-4，6-12，14-19
Y	CN 105719649 A (百度在线网络技术北京有限公司) 2016年 6月 29日 (2016-06-29) 说明书第[0036]-[0094]段	5，13，20
Y	CN 106683662 A (中国电信股份有限公司) 2017年 5月 17日 (2017-05-17) 说明书第[0093]-[0101]段	5，13，20
X	CN 105448292 A (北京羽扇智信息科技有限公司) 2016年 3月 30日 (2016-03-30) 说明书第[0005]-[0047]段	1-4，6-12，14-19
X	CN 109509473 A (维沃移动通信有限公司) 2019年 3月 22日 (2019-03-22) 说明书第[0031]-[0051]段	1-4，6-12，14-19
X	CN 104240698 A (上海伯释信息科技有限公司) 2014年 12月 24日 (2014-12-24) 说明书第[0003]-[0014]段	1、8、9、16

☑ 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	国际检索报告邮寄日期
2020年 6月 28日	2020年 7月 24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 王玥 电话号码 86-（10）- 53962545

国际检索报告

国际申请号
PCT/CN2020/087471

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	US 2015120618 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2015年 4月 30日 (2015-04-30) 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/CN2020/087471

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	110349575	A	2019年 10月 18日		无		
CN	105719649	A	2016年 6月 29日	CN	105719649	B	2019年 7月 5日
CN	106683662	A	2017年 5月 17日		无		
CN	105448292	A	2016年 3月 30日	CN	105448292	B	2019年 3月 12日
CN	109509473	A	2019年 3月 22日		无		
CN	104240698	A	2014年 12月 24日		无		
US	2015120618	A1	2015年 4月 30日	US	9948764	B2	2018年 4月 17日
				US	2018227410	A1	2018年 8月 9日
				KR	20150047803	A	2015年 5月 6日