(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
2019 年 1 月 10 日 (10.01.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/007245 A1

(51) 国际专利分类号:
G10L 15/18 (2013.01)

(21) 国际申请号: PCT/CN2018/093216

(22) 国际申请日: 2018 年 6 月 28 日 (28.06.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710539394.1 2017年7月4日 (04.07.2017) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 吴栓 (WU, Shuan); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 刘广兴 (LIU, Guangxing); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 许毅 (XU, Yi); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) **Title:** PROCESSING METHOD, CONTROL METHOD AND RECOGNITION METHOD, AND APPARATUS AND ELECTRONIC DEVICE THEREFOR

(54) **发明名称**: 一种处理方法、控制方法、识别方法及其装置和电子设备

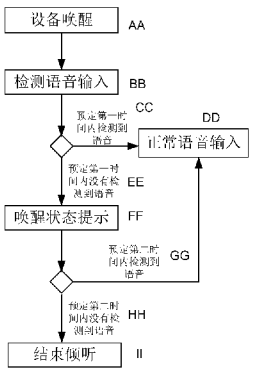


图 1

AA DEVICE WAKE-UP
BB DETECT VOICE INPUT
CC VOICE IS DETECTED IN A PRE-DETERMINED FIRST TIME
DD NORMAL VOICE INPUT
EE NO VOICE IS DETECTED IN THE PRE-DETERMINED FIRST TIME
FF WAKE-UP STATE PROMPT
GG VOICE IS DETECTED IN A PRE-DETERMINED SECOND TIME
HH NO VOICE IS DETECTED IN THE PRE-DETERMINED SECOND TIME
II END LISTENING

(57) Abstract: The embodiments of the present invention relate to a processing method, a control method and a recognition method, and an apparatus and electronic device therefor, in particular to a wake-up processing method, a wake-up sensitivity control method, a fast wake-up processing method and a voice control object recognition method, and corresponding apparatuses and electronic devices. The embodiments of the present invention provide an improved solution in the aspect of a prompt in device wake-up, the aspect of adjusting wakeup sensitivity based on different scenarios, the aspect of accurately recognizing and removing a fast wake-up word in quick wake-up, and the aspect of improving the determination of a control object for the execution of voice instructions, and improves the intelligence and accuracy of human-machine dialogue compared with the prior art.

(57) 摘要: 本发明实施例涉及一种处理方法、控制方法、识别方法及其装置和电子设备，尤其涉及唤醒处理方法、唤醒灵敏度控制方法、快捷唤醒处理方法以及语音控制对象识别方法以及相应的装置和电子设备。本发明实施例在设备唤醒的提示方面、根据不同的场景调整唤醒灵敏度方面、准确识别并去除快捷唤醒中的快捷唤醒词方面以及提高语音指令执行的控制对象的确定等方面，提出了改进方案，相比现有技术而言，提高了人机对话的智能化和准确度。

WO 2019/007245 A1

[见续页]

(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))。

一种处理方法、控制方法、识别方法及其装置和电子设备

本申请要求2017年07月04日递交的申请号为201710539394.1、发明名称为“一种处理方法、控制方法、识别方法及其装置和电子设备”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。

技术领域

本发明涉及唤醒处理方法、唤醒灵敏度控制方法以及语音控制对象识别方法以及相应的装置和电子设备。

背景技术

随着人工智能相关应用的越来越深入发展，语音识别技术作为智能化设备的基本的交互方式，扮演着越来越重要的角色。语音识别技术涉及到很多方面，其中包括通过语音指令来唤醒设备、对设备的操作进行控制、与设备进行人机对话以及针对多个设备的语音指令控制等。高效和准确的语音识别技术以及快捷便利的唤醒模式，是智能化设备的重要的发展方向。

发明内容

本发明提供了一种唤醒处理方法、装置及电子设备，能够在设备处于，唤醒状态，且没有语音输入的场景下，主动发出提示，从而让用户感知当前设备处于唤醒状态，方便进行语音输入。

为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供了一种唤醒处理方法，包括：

在设备被唤醒后，检测是否有语音输入；

如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

第二方面，提供了一种唤醒处理装置，包括：

语音检测模块，用于在设备被唤醒后，检测是否有语音输入；

唤醒提示模块，用于如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

第三方面，提供了一种电子设备，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

在设备被唤醒后，检测是否有语音输入；

如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

本发明提供的唤醒处理方法、装置及电子设备，在设备被唤醒后，如果用户在指定时间内没有发出语音指令，则设备可以输出表示设备处于唤醒状态的提示，让用户感知当前设备正处于唤醒状态；通过这样的机制，能够使得处于唤醒状态下的设备，在没有接收到任何用户输入的情况下，始终保持一个与用户进行沟通的状态，能够不断提示用户设备已经处于唤醒状态，并提示用户进行进一步操作，从而不需要用户再去判断设备是否唤醒。

本发明提供了一种唤醒灵敏度控制方法、装置及电子设备，能够根据设备所处的不同的应用场景，灵活调整设备的唤醒灵敏度，从而平衡了唤醒率和误唤醒率在使用方面存在的矛盾。

为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供了一种唤醒灵敏度控制方法，包括：

获取设备当前的应用场景信息；

根据所述应用场景信息，调整所述设备的唤醒灵敏度。

第二方面，提供了一种唤醒灵敏度控制装置，包括：

信息获取模块，用于获取设备当前的应用场景信息；

灵敏度调整模块，用于根据所述应用场景信息，调整所述设备的唤醒灵敏度。

第三方面，提供了一种电子设备，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

获取设备当前的应用场景信息；

根据所述应用场景信息，调整所述设备的唤醒灵敏度。

本发明提供的唤醒灵敏度控制方法、装置及电子设备，根据当前设备所处的应用场景，灵活调整设备的唤醒灵敏度。由于不再采用固定的唤醒灵敏度，从而能够根据应用场景来采用适合的唤醒灵敏度，平衡了唤醒率和误唤醒率在使用方面存在的矛盾。

本发明提供了一种快捷唤醒处理方法、装置及电子设备，能够在对语音指令进行语

义解析之前，先滤除掉唤醒词，从而使得语义解析的结果不会受到唤醒词的影响。

为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供了一种快捷唤醒处理方法，包括：

对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

对所述第二文本执行语义解析。

第二方面，提供了一种快捷唤醒处理装置，包括：

文本生成模块，用于对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

唤醒词过滤模块，用于针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

语义解析模块，用于对所述第二文本执行语义解析。

第三方面，提供了一种电子设备，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

对所述第二文本执行语义解析。

本实施例提供的快捷唤醒处理方法、装置及电子设备，在对识别后生成的文本进行语义解析之前识别并滤除掉唤醒词，从而使得语义解析的结果不会受到唤醒词的影响。

本发明提供了一种语音控制对象识别方法、装置及电子设备，能够在多设备同时运行的人机对话场景下，准确的识别出当前语音指令所要控制的对象设备。

为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供了一种语音控制对象识别方法，包括：

识别当前语音指令中体现控制意图的第一语义单元；

根据所述控制意图确定对应的一个或多个控制场景；

获取各个设备当前所处的运行状态；

根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当

前语音指令的控制对象。

第二方面，提供了另一种语音控制对象识别方法，包括：

获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

根据所述逻辑类确定所述当前语音指令的控制对象。

第三方面，提供了一种语音控制对象识别装置，包括：

语义识别模块，用于识别当前语音指令中体现控制意图的第一语义单元；

场景确定模块，用于根据所述控制意图确定对应的一个或多个控制场景；

状态获取模块，用于获取各个设备当前所处的运行状态；

第一对象确定模块，用于根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当前语音指令的控制对象。

第四方面，提供了另一种语音控制对象识别装置，包括：

逻辑类获取模块，用于获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

第二对象确定模块，用于根据所述逻辑类确定所述当前语音指令的控制对象。

第五方面，提供了一种电子设备，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

识别当前语音指令中体现控制意图的第一语义单元；

根据所述控制意图确定对应的一个或多个控制场景；

获取各个设备当前所处的运行状态；

根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当前语音指令的控制对象。

第六方面，提供了另一种电子设备，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

根据所述逻辑类确定所述当前语音指令的控制对象。

本实施例提供的语音控制对象的识别方法、装置及电子设备，根据不同设备的运行状态和当前语音指令体现的控制意图所确定的控制场景之间的匹配关系，确定当前语音指令的控制对象，或者，根据当前语音指令与前一次语音指令的上下文逻辑关系，确定当前语音指令的控制对象，从而提高了在从多个运行的设备中识别出语音控制对象的准确性。

上述说明仅是本发明技术方案的概述，为了能够更清楚了解本发明的技术手段，而可依照说明书的内容予以实施，并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂，以下特举本发明的具体实施方式。

附图说明

通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

图 1 为本发明实施例提供的唤醒处理的逻辑示意图；

图 2 为本发明实施例提供的唤醒处理方法的流程图；

图 3a 为本发明实施例提供的唤醒处理装置的结构示意图一；

图 3b 为本发明实施例提供的唤醒处理装置的结构示意图二；

图 4 为本发明实施例提供的电子设备的结构示意图；

图 5 为本发明实施例提供的唤醒灵敏度的控制逻辑示意图；

图 6 为本发明实施例提供的唤醒灵敏度的控制方法的流程图；

图 7 为本发明实施例提供的唤醒灵敏度的控制装置的结构示意图；

图 8 为本发明实施例提供的电子设备的结构示意图；

图 9 为本发明实施例提供的快捷唤醒的处理逻辑示意图；

图 10 为本发明实施例提供的快捷唤醒的处理方法的流程图；

图 11 为本发明实施例提供的快捷唤醒的处理装置的结构示意图；

图 12 为本发明实施例提供的电子设备的结构示意图；

图 13 为本发明实施例提供的语音控制对象识别方法的流程图一；

图 14 为本发明实施例提供的语音控制对象识别装置的结构示意图一；

图 15 为本发明实施例提供的语音控制对象识别装置的结构示意图二；

图 16 为本发明实施例提供的语音控制对象识别装置的结构示意图三；

图 17 为本发明实施例提供的电子设备的结构示意图一；

图 18 为本发明实施例提供的电子设备的结构示意图二。

具体实施方式

下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

实施例一

现有的人机会话场景中，设备从休眠状态下被唤醒到运行状态下，会通过灯光或者铃声来提示一下已经处于唤醒状态。但是，如果用户忽略了灯光或者铃声的提示时，就很难知道是否已经切换到唤醒状态或者是否还处于唤醒状态，这种提示对于用户来说并不方便。在唤醒之后，在用户不进行任何操作的情况下，设备也不会向用户提供任何反馈，以提示用户是否还处于唤醒状态。面对这种状态，用户不知道是应该输入唤醒指令还是直接输入指示设备的具体操作。

比如设备在从休眠切换到唤醒后，进行了一次响铃提示，如果用户没有听到这次响铃，便无法获知是否已经唤醒，而通过灯光方式来进行提示，用户在距离设备较远的时候也很难观察到灯光，因此，也无法清楚的获知是否处于唤醒状态。

本发明实施例改变了现有技术中，通过响铃和/或灯光方式无法使用户明确感知到设备是否处于唤醒状态，其核心思想在于，在唤醒状态下，通过增加检测是否有语音输入的功能，如果没有输入，则通过提示语来提示用户进行输入，从而让用户感知当前设备处于唤醒状态。

如图 1 所示，为本发明实施例提供的唤醒处理的逻辑示意图。在图 1 中，在设备被唤醒后，设备会先检测是否有语音输入，如果在预定的第一时间内检测到语音输入，则进行正常的语音输入流程以及后续操作；如果在预定的第一时间内没有检测到语音输入，则设备会发出语音提示，告知用户当前设备已处于唤醒状态，请用户输入语音指令。如果用户在设备发出处于唤醒状态的语音提示以后的预设第二时间内发出了语音指令，则设备采集该语音信号进行正常的语音输入流程以及后续操作，否则结束倾听，即关闭语音采集流程。用户如果想再次输入语音指令需要再次唤醒设备。

基于图 1 所示的唤醒处理的逻辑，如图 2 所示，为本发明实施例提供的唤醒处理方法流程图，该方法包括如下步骤：

S210，在设备被唤醒后，检测是否有语音输入。

具体地，用户可以通过直接唤醒和快捷唤醒的方式唤醒设备。所谓直接唤醒是指语音输入唤醒指令后等设备响应后，待设备被唤醒后再输入语音指令，等待设备反馈结果以实现人机会话。所谓快捷唤醒是指唤醒指令和语音指令一起进行语音设备，设备直接反馈结果实现人机会话。本方案中，关于设备的唤醒方式不局限于上述方式。

在设备被唤醒后，设备启动语音活动检测（Voice Activity Detection，VAD）流程，检测是否有语音输入。

当然，在设备被唤醒后，也可以进行响铃和/或灯光提醒，以告知用户设备以唤醒。

S220，如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

在设备被唤醒后，自动进行倒计时，倒计时的时长为预设的第一时间，如果在预定的第一时间内设备没有检测到语音输入，则输出表示设备处于唤醒状态的提示。如区别于设备唤醒初始状态的响铃内容和/或灯光闪烁方式，也可以直接通过语音的方式提示用户输入语音内容，例如语音输出“我在，请输入语音内容”。

用户听到或看到这些表示设备处于唤醒状态的提示后，可继续执行语音输入完成人机会话。

当然，如果当前设备没有连接到云端（云端负责对设备接收的语音指令进行识别，并将识别出的控制操作指令反馈给设备进行相应操作，本方案的改进点不涉及云端的处理环节），也可以通过语音提示的方式提醒用户网络没有连接好。

进一步地，上述方法还包括：如果在输出表示设备处于唤醒状态的提示之后的预定的第二时间内没有检测到语音输入，则关闭唤醒状态。

当设备输出表示设备处于唤醒状态的提示之后指定时间段内仍没有检测到语音输入，则表明用户可能不想进行人机会话，此时可控制设备关闭唤醒状态，进入到休眠模式。用户如果想再次输入语音指令需要再次唤醒设备。

本发明实施例提供的唤醒处理方法，在设备被唤醒后，如果用户在指定时间内没有发出语音指令，则设备可以输出表示设备处于唤醒状态的提示，让用户感知当前设备正处于唤醒状态；通过这样的机制，能够使得处于唤醒状态下的设备，在没有接收到任何用户输入的情况下，始终保持一个与用户进行沟通的状态，能够不断提示用户设备已经

处于唤醒状态，并提示用户进行进一步操作，从而不需要用户再去判断设备是否唤醒。

实施例二

如图 3a 所示，为本发明实施例的唤醒处理装置结构图，该唤醒处理装置可用于执行如图 2 所示的方法步骤，其包括：

语音检测模块 310，用于在设备被唤醒后，检测是否有语音输入；

唤醒提示模块 320，用于如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

进一步地，如图 3b 所示，上述唤醒处理装置还可包括唤醒关闭模块 330，用于如果在输出表示设备处于唤醒状态的提示之后的预定的第二时间内没有检测到语音输入，则关闭唤醒状态。

进一步地，上述表示设备处于唤醒状态的提示为语音提示。

本发明实施例提供的唤醒处理装置，在设备被唤醒后，如果用户在指定时间内没有发出语音指令，则设备可以输出表示设备处于唤醒状态的提示，让用户感知当前设备正处于唤醒状态；通过这样的机制，能够使得处于唤醒状态下的设备，在没有接收到任何用户输入的情况下，始终保持一个与用户进行沟通的状态，能够不断提示用户设备已经处于唤醒状态，并提示用户进行进一步操作，从而不需要用户再去判断设备是否唤醒。

实施例三

前面描述了唤醒处理装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 4 所示，其为本发明实施例的电子设备的结构示意图，具体包括：存储器 410 和处理器 420。

存储器 410，用于存储程序。

除上述程序之外，存储器 410 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 410 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 420，耦合至存储器 410，用于执行存储器 410 中的程序，以用于：

在设备被唤醒后，检测是否有语音输入；

如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 4 所示，电子设备还可以包括：通信组件 430、电源组件 440、音频组件 450、显示器 460 等其它组件。图 4 中仅示意性给出部分组件，并不意味着电子设备只包括图 4 所示组件。

通信组件 430 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 430 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 430 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

电源组件 440，为电子设备的各种组件提供电力。电源组件 440 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 450 被配置为输出和/或输入音频信号。例如，音频组件 450 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 410 或经由通信组件 430 发送。在一些实施例中，音频组件 450 还包括一个扬声器，用于输出音频信号。

显示器 460 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

实施例四

目前，在一些智能设备中，引入了语音唤醒机制，即检测环境的语音输入，并自动唤醒的机制。语音唤醒的两个基本的技术指标分别是唤醒率和误唤醒率。唤醒率和误唤醒率是相辅相成的。唤醒率过低影响用户唤醒设备的体验，误唤醒率过高会影响用户的

正常生活，造成不必要的干扰。

唤醒率和误唤醒率本质上主要都是由唤醒灵敏度决定的，决定唤醒灵敏度的因素可以是设备的传感器的灵敏度设置或者语音处理程序（唤醒引擎）对唤醒指令判定的灵敏度。

唤醒灵敏度越高，则越容易检测到周围环境的语音输入，在检测时，会更加容易导致设备唤醒，同时也会增加误唤醒的概率，而如果传感器的灵敏度较低，则虽然误唤醒的概率降低了，但同时设备也不容易被唤醒，唤醒率也就降低了，从而影响了用户的正常使用。

在现有技术中，采用的唤醒灵敏度固定的方式，无论是采用高唤醒灵敏度还是采用低唤醒灵敏度都无法解决对上述使用中存在的矛盾。

本发明实施例改变了现有技术中，对设备的唤醒灵敏度采用固定的方式，其核心思想在于，根据应用场景的不同，灵活调整设备的唤醒灵敏度。

如图 5 所示，为本发明实施例提供的唤醒灵敏度的控制逻辑示意图。在图 5 中，作为影响设备的唤醒灵敏度设置的一个重要因素，那就是应用场景，随着引用场景的不同，对唤醒灵敏度的要求也不同。例如，白天环境比较嘈杂，噪音比较大，则可以适当降低唤醒灵敏度，以减少误唤醒率；而夜晚环境较为安静，噪音比较小，则可以适当提高唤醒灵敏度，以提高唤醒率。在调整设备唤醒灵敏度时，可以从软件和硬件两个方面进行调整。

软件方面则可以调整设备中唤醒引擎的唤醒参数，例如通过设备的应用（APP）为唤醒引擎的唤醒灵敏度设置多个等级：灵敏、一般、稍弱、安静。用户可以用过手机 APP 来设置各时间段的灵敏等级，手机 APP 设置后，上传到相应云端记录该事件。云端在相应的时间点，推送相关的指令给设备，控制设备调整当前的唤醒引擎中的相关设置，从而调整唤醒灵敏度。

硬件方面，可以直接设置设备中声音传感器的声音采集参数，从而调整设备的唤醒灵敏度。

基于图 5 所示的唤醒处理的逻辑，如图 6 所示，为本发明实施例提供的唤醒灵敏度的控制方法流程图，该方法包括如下步骤：

S610，获取设备当前的应用场景信息。

其中，所述的应用场景可以为设备所处的时间段，可从设备的系统时钟获得，如白天、夜晚；也可以是周围环境，如设备是处在安静的居家室内，还是嘈杂的商场里等，

当然关于环境的应用场景信息可以由人工手动输入到设备中，如预先设置一个应用场景的安静等级：安静、轻度喧闹、嘈杂等几个等级。

S620，根据所述应用场景信息，调整设备的唤醒灵敏度。

根据检测到的设备当前的应用场景信息，调整设备的唤醒灵敏度，使得当应用场景较为安静时提高设备的唤醒灵敏度，当应用场景较为嘈杂时降低设备的唤醒灵敏度，从而根据应用场景的不同，灵活调整设备的唤醒灵敏度，兼顾降低误唤醒率和提高用户体验。

进一步地，上述方法还可包括：接收设备的应用的灵敏度设置信息，所谓的设备的应用，是指设备所对应的后台服务端，用户可以在手机 APP 上向服务端发送灵敏度设置信息，以通过服务端向设备发送设置指令，来调整设备的唤醒灵敏度。

进一步地，根据应用场景信息，调整设备的唤醒灵敏度可包括：根据应用场景信息和灵敏度设置信息，设置设备的唤醒引擎的唤醒参数和/或设置设备的声音传感器的声音采集参数。应用场景信息可包括设备所处的时间段信息。

例如，用户可以用过手机 APP 来设置各时间段的灵敏等级，并在手机 APP 上设置后，上传到相应云端记录该事件。云端在相应的时间点，推送相关的指令给设备，控制设备调整唤醒灵敏度；或者云端将针对不同时间段的唤醒灵敏度设置信息发送给设备，当设备检测到当前的时间段与云端推送的调整唤醒灵敏度的时间段相匹配时，控制设备调整当前的唤醒灵敏度。具体调整的对象可以是内置在设备中的唤醒引擎的相关设置参数，也可以是接收语音的声音传感器的声音采集参数。

本实施例提供的唤醒灵敏度的控制方法，根据当前设备所处的应用场景，灵活调整设备的唤醒灵敏度。由于不再采用固定的唤醒灵敏度，从而能够根据应用场景来采用适合的唤醒灵敏度，平衡了唤醒率和误唤醒率在使用方面存在的矛盾。

实施例五

如图 7 所示，为本发明实施例的唤醒灵敏度的控制装置结构图，该唤醒处理装置可内置在设备中，可用于执行如图 6 所示的方法步骤，其包括：

信息获取模块 710，用于获取设备当前的应用场景信息；

灵敏度调整模块 720，用于根据应用场景信息，调整设备的唤醒灵敏度。

进一步地，上述信息获取模块 710 还用于接收设备的应用的灵敏度设置信息，

相应的，灵敏度调整模块 720 具体用于，根据应用场景信息和灵敏度设置信息，设

置设备的唤醒引擎的唤醒参数和/或设置设备的声音传感器的声音采集参数。

进一步地，上述应用场景信息可包括设备所处的时间段信息。

本实施例提供的唤醒灵敏度的控制装置，根据当前设备所处的应用场景，灵活调整设备的唤醒灵敏度。由于不再采用固定的唤醒灵敏度，从而能够根据应用场景来采用适合的唤醒灵敏度，平衡了唤醒率和误唤醒率在使用方面存在的矛盾。

实施例六

前面描述了唤醒灵敏度的控制装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 8 所示，其为本发明实施例的电子设备的结构示意图，具体包括：存储器 810 和处理器 820。

存储器 810，用于存储程序。

除上述程序之外，存储器 810 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 810 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 820，耦合至存储器 810，用于执行存储器 810 中的程序，以用于：

获取设备当前的应用场景信息；

根据所述应用场景信息，调整所述设备的唤醒灵敏度。

上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 8 所示，电子设备还可以包括：通信组件 830、电源组件 840、音频组件 850、显示器 860 等其它组件。图 8 中仅示意性给出部分组件，并不意味着电子设备只包括图 8 所示组件。

通信组件 830 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 830 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 830 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）

技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

电源组件 840，为电子设备的各种组件提供电力。电源组件 840 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 850 被配置为输出和/或输入音频信号。例如，音频组件 850 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 810 或经由通信组件 830 发送。在一些实施例中，音频组件 850 还包括一个扬声器，用于输出音频信号。

显示器 860 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

实施例七

在目前的语音技术中，可以通过输入唤醒词来唤醒设备，然后设备会将后续接收到的语音内容送到云端进行识别，然后返回具体的针对设备的指令或者是回答内容等。

但是，在快捷唤醒的场景下，即唤醒词和语音指示内容一起作为一条语音指令输入的情况下， 设备会把唤醒词和语音指示内容一起送到云端做语义理解，而云端不具备识别唤醒词的能力，因此，这样会导致语义理解出现各种问题，造成如语音理解会进错领域，或者答非所问。并且，不同设备设定的唤醒词也会有差异，每个设备中的唤醒词识别引擎也是不同的，导致云端也很难统一进行处理。

本发明实施例改善了现有技术中，云端没有过滤唤醒词，而影响后续语义理解的问题，其核心思想在于，在云端也设置唤醒词识别引擎，从而在进行语义解析之前，先滤除掉唤醒词。

如图 9 所示，为本实施例提供的快捷唤醒的处理逻辑图。在图 9 中，设备首先将检测到的包含唤醒词的音频信息传输到云端的自动语音识别（ASR）进行处理，识别出包含唤醒词的文本，然后对唤醒词进行过滤，生成唤醒词被过滤后的文本再进行语义解析。其中，在经唤醒词识别引擎识别出唤醒词后，云端可以将包含唤醒词的文本返回至设备，由设备对文本中的唤醒词进行过滤，然后再将过滤后的文本返回云端继续进行语义解析处理。

基于图 9 所示的快捷唤醒的处理逻辑，如图 10 所示，为本发明实施例提供的快捷唤醒的处理方法流程图，该方法包括如下步骤：

S101，对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

人机对话场景中，针对快捷唤醒设备的应用场景，设备在接收到唤醒词+语音指令的第一音频信息后，将该音频信息传送至云端，云端对该第一音频信息进行识别处理后，生成与该第一音频信息对应的第一文本。

S102，针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

在实际应用场景中，云端并不知道哪条音频信息中包括唤醒词，并且如果在进行识别时，如果音频清晰，则可识准确别出语义正确的文本内容，且该内容在进行后续语义解析时也不会因发生误解析，而进入错误领域，产生错误指令。但是，如果音频不清晰，则可能语音识别出语义错误的文本内容，且该内容在进行后续语义解析时很可能因发生误解析而进入错误领域，产生错误指令，或者根本无法判断出执行指令，而使得对设备的控制或反馈失败。因此，在对识别处理后得到的第一文本进行语义解析之前，有必要过滤掉其中的唤醒词。

本步骤在对设备上传的第一音频信息进行识别处理后，对生成的第一文本先进行过滤唤醒词的处理，生成去掉唤醒词后的第二文本，以防止唤醒词在进行语义解析时出现错误解析。

S103，对第二文本执行语义解析。

进一步地，上述对来自设备的包含唤醒词的第一音频信息进行识别的处理包括：

使用云端的识别模型对第一音频信息进行识别，其中， 识别模型包括设备中唤醒词识别引擎所使用的唤醒词词库。

为了使云端能够过准确识别出第一语音信息中包含的唤醒词，在识别模型中可增加专门用于识别设备唤醒词的唤醒词词库，该唤醒词词库与设备中唤醒词识别引擎所使用的唤醒词词库相同。这样在对第一音频信息进行识别时，就可以准确识别出唤醒词。

例如，当输入至设备的第一音频信息为“你好，请开灯”，如果语音不太清晰，则可能会识别为“你要开灯”（识别模型也会根据上下文的情况进行一些智能判断）。在这样的情况下，当识别算法模型中，包含唤醒词的唤醒词词库时，就可以很容易识别出该唤醒词。这样，当针对“你好，请开灯”这样的音频进行识别时，会在识别句子的前部分时，更多的考虑是唤醒词，当识别到可能是“你好”还是“你要”时，根据唤醒词

的词库，会最终确定为“你好”，这样就不会出现上述的错误了。

进一步地，上述针对第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本可包括：

将第一文本发送给设备；

设备根据该设备的唤醒词识别引擎所使用的唤醒词词库，过滤掉第一文本中的唤醒词，生成第二文本，并将该第二文本发送给云端。

最后再由云端继续执行对第二文本的语义解析。

本实施例提供的快捷唤醒的处理方法，在对识别后生成的文本进行语义解析之前识别并滤除掉唤醒词，从而使得语义解析的结果不会受到唤醒词的影响。

实施例八

如图 11 所示，为本发明实施例的快捷唤醒的处理装置结构图，该快捷唤醒的处理装置可内置在云端中，可用于执行如图 10 所示的方法步骤，其包括：

文本生成模块 111，用于对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

唤醒词过滤模块 112，用于针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

语义解析模块 113，用于对第二文本执行语义解析。

进一步地，上述文本生成模块 111 具体用于，

使用云端的识别模型对第一音频信息进行识别，其中，识别模型包括设备中唤醒词识别引擎所使用的唤醒词词库。

进一步地，上述唤醒词过滤模块 112 具体用于，

将第一文本发送给设备；

接收设备返回的过滤掉第一文本中的唤醒词后的第二文本。

在实际应用场景中，云端可在识别出语音指令中的唤醒词后，将包含唤醒词的第一文本返回给设备，由设备对第一文本中的唤醒词进行过滤，然后再将唤醒词过滤之后的第二文本返回至云端，以是云端继续进行文本的语义解析。由此可见，云端与设备之间通过任务交互构成一个处理系统，共同完成对第一文本中的唤醒词的过滤操作。

本实施例提供的快捷唤醒的处理装置，在对识别后生成的文本进行语义解析之前识别并滤除掉唤醒词，从而使得语义解析的结果不会受到唤醒词的影响。

实施例九

前面描述了快捷唤醒的处理装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 12 所示，其为本发明实施例的电子设备的结构示意图，具体包括：存储器 121 和处理器 122。

存储器 121，用于存储程序。

除上述程序之外，存储器 121 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 121 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 122，耦合至存储器 121，用于执行存储器 121 中的程序，以用于：

获取设备当前的应用场景信息；

根据所述应用场景信息，调整所述设备的唤醒灵敏度。

上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 12 所示，电子设备还可以包括：通信组件 123、电源组件 124、音频组件 125、显示器 126 等其它组件。图 12 中仅示意性给出部分组件，并不意味着电子设备只包括图 12 所示组件。

通信组件 123 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 123 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 123 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

电源组件 124，为电子设备的各种组件提供电力。电源组件 124 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 125 被配置为输出和/或输入音频信号。例如，音频组件 125 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦

克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 121 或经由通信组件 123 发送。在一些实施例中，音频组件 125 还包括一个扬声器，用于输出音频信号。

显示器 126 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

实施例十

在智能语音控制领域中，用户发出的一条语音指令，该条指令反应用户明确唯一意图，但语句本身可能具有多个意图表现，此种情况是我们常说的“歧义”。在语音智能系统可以控制多个设备时，会出现对多个设备之间的控制冲突，比如用户发出一条“播放”的语音指令，可能是播放音乐播放设备中的音乐，也可能是播放视频播放机中的电影等。

针对这样的问题，本实施例提供了一种语音控制对象识别方法，如图 13 所示，其为本发明实施例提供的语音控制对象识别方法的流程图，包括：

S131：识别当前语音指令中体现控制意图的第一语义单元。具体地，用户的语音指令一般会包含多个语义单元，这里所说的语义单元可以是一个字或者一个词或者一个短语，也可以是构成语音指令中的一个句子，语义单元应当是能够完整表达一个含义的单元。在本步骤中，第一语义单元应当是能够体现控制意图的语义单元，其中，控制意图是指人机交互系统中让设备执行的某一项具体功能，例如播放、暂停、提高音量。相同的控制意图可能会由不同的语音指令的表达形式，也就是说会对应不同的第一语义单元。比如，语音控制指令为“把视频给我播放一下”、“给我放映一下视频”、“我要看一下电影”，在这这些语音控制指令中，“播放”、“放映”、“看一下”都是第一语义单元，其都指向“播放”这一控制意图，对于控制意图的确定可以使用已有的语义分析技术，在这里就不再赘述。

当一个语音指令中，主要只包含体现控制意图的第一语义单元时，就容易造成控制指令或者说是控制意图在多个设备之间的冲突。例如：“播放”、“暂停”、“停止”、“重播”、“退出”、“打开声音”、“关闭声音”、“提高音量”、“降低音量”，这些控制意图往往是很多设备都具备的，因此，很容易产生冲突。

S132：根据控制意图确定对应的一个或多个控制场景。在步骤 S131 中，已经识别出了控制意图，在步骤 S132 中，可以根据该控制意图将可能的控制场景全部列出。例如，在当前的环境中，具有如下三个设备，该三个设备均能接收语音指令：

智能电视：与网络连接，能够进行在线视频播放以及视频搜索；

智能音箱：具有 wifi 和蓝牙功能，能够播放在线音乐或者通过蓝牙连接的其他设备中的音乐；

智能电脑。

以控制意图为“播放”为例，其对应的控制场景可能为：

1）智能音箱通过 wifi 进行在线音乐播放。

2）智能音箱播放与其有蓝牙连接的其他设备中的音乐。

3）智能电视播放当前暂停的视频。

4）智能电视播放当前搜索结果中对应的视频。

5）运行智能电脑中音视频播放 APP 的播放功能。

上述的多个控制场景，也就对应了该控制意图可能的控制对象。

S133：获取各个设备当前所处的运行状态。这里所说的运行状态可以是各个设备的开机状态、运行着哪些应用、以及各个应用所处的控制状态等。例如，针对上述的智能电视、智能音箱以及智能电脑的运行状态进行获取，可以假设存在如下状态：

a）智能电视处于一个电影的暂停播放状态。

b）智能电视执行完一个电影名称的搜索，并显示了搜索结果。

c）智能音箱与手机蓝牙连接，并且处于音乐播放暂停状态，所播放音乐为手机中的音乐。

d）智能音箱处于 wifi 连接状态，并且处于在线音乐播放暂停状态。

f）智能电脑中运行着听书软件，并且处于播放暂停状态。

需要说明的是，以上各个状态并不是并存关系，例如，状态 a 和状态 b 对于同一个智能电视来说当前只会存在于一个状态中，状态 c 和状态 d 对于同一个智能音箱而言当前只会存在于一个状态中。

在实际应用中，虽然也会存在智能设备会同时运行多个应用情形，但是，可以将当前处于激活状态的应用或者处于平面显示的应用的状态作为当前的状态。

另外，需要说明的是，步骤 S131 和 S132 与步骤 S133 之间本身没有先后顺序，步骤 S133 可以与步骤 S131 和 S132 同时执行，也行在步骤 S131 之前执行，也可以在步骤

S132 之后执行，也可以在步骤，步骤 S131 和 S132 之间执行。

S134：根据各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定当前语音指令的控制对象。步骤 S132 和步骤 S133 已经获取到了各个设备的运行状态和控制场景，在步骤 S134 中，将会对各个设备的运行状态和一个或多个控制场景之间的匹配关系进行分析，看一下各个设备的运行状态与那种一控制场景匹配，然后就可以根据相应的控制场景确定相应的控制对象，并执行进一步的控制操作。

需要说明的是，这里所述的控制对象可以是设备也可以是设备中的具体某个应用或者进程。

仍然以上述的智能电视、智能音箱以及智能电脑为例，可能存在如下的情形：

A1）如果智能电视处于电影的暂停播放状态（上述状态 a），智能音箱处于虽然与手机连接但是并未执行播放进程，智能电脑中的听书软件没有运行。针对这种情形，可以确定该“播放”控制意图指向智能电视中的电影播放进程。

A2）如果智能电视处于执行完一个电影名称的搜索并显示了搜索结果的状态（上述状态 b），智能音箱处于虽然与手机连接但是并未执行播放进程，智能电脑中的听书软件没有运行。针对这种情形，可以确定该“播放”控制意图指向智能电视中的电影播放进程。

A3）如果智能音箱与手机蓝牙连接，并且处于音乐播放暂停状态，所播放音乐为手机中的音乐，智能电视只是显示主页界面，智能电脑中的听书软件没有运行。针对这种情形，可以确定该“播放”控制意图指向智能音箱中的音乐播放进程。

A4）智能音箱处于 wifi 连接状态，并且处于在线音乐播放暂停状态，智能电视只是显示主页界面，智能电脑中的听书软件没有运行。针对这种情形，可以确定该“播放”控制意图指向智能音箱中的音乐播放进程。

A5）如果智能电脑中运行着听书软件，并且处于播放暂停状态，智能音箱处于虽然与手机连接但是并未执行播放进程，智能电视只是显示主页界面。针对这种情形，可以确定该“播放”控制意图指向智能电脑中的听书软件应用。

通过上面的举例可以看出，通过综合分析各个设备的运行状态和当前语音指令中的控制意图对应的一个或多个控制场景，在一些情况下是可以确定出当前语音指令的合理的控制对象，从而能够更加准确地确定用户发出的当前语音指令的确定指向，从而方便用户的语音控制，减少实际控制对象和用户想要的控制对象之间的判断偏差，提高设备的智能化水平。

需要说明的是，基于上述的语音控制对象识别方法整个流程也可能存在无法唯一判定出用户的当前语音指令所指向的控制对象的情形，在这种情况下，可以通过后面介绍的其他处理方式来进一步判断，也可以直接向用户发出语音提示让用户进一步明确控制需求或者控制对象。

此外，在一些情形中，用户发出的当前语音指令中会包含能够体现控制对象的一些第二语义单元，如果存在能够第二语义单元，则可以基于此直接确定控制对象或筛选掉一部分控制对象。因此，在上述的步骤 S131 之前，还可以包括：

S130：识别当前语音指令中体现控制对象的第二语义单元，如果存在第二语义单元，则根据第二语义单元确定控制对象或者排除掉部分控制对象后，执行步骤 S131，如果不存在第二语音单元，则执行步骤 S131。例如，当前语音控制指令为“播放这个视频”，则如果环境中只有智能电视、智能音箱以及智能电脑三个设备，则可以直接把智能音箱排除掉，进而通过后续的设备的运行状态和可能场景的视频可以很容易确定出控制对象。如果环境中只有智能电视和智能音箱这两个设备，则可以直接确定出控制对象为智能电视中的视频播放进程。

如果通过上述的处理步骤仍然无法确定控制对象，例如在无法获取到各个设备当前所处的运行状态或者无法根据获取到运行状态确定所述当前语音指令的控制对象的情况下，则可以执行如下处理：

S135：获取当前语音指令的控制意图所对应的逻辑池中记录的最近一次历史语音指令所对应的逻辑类，所述逻辑池包括多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，根据所述逻辑类确定所述当前语音指令的控制对象。

该步骤主要是基于当前语音指令之前的历史语音指令并结合当前语音指令进行判断，属于基于语音指令的上下文的判断。不过，本实施例的特别之处在于，针对不同的控制意图构建了相应的逻辑池，逻辑池包括多个逻辑类，并且在设备开机后，对执行过的历史语音指令进行记录，记录方式为按照将历史语音指令按照逻辑类进行分别记录。在实际应用中，逻辑池的各个逻辑类中可以只保存属于该逻辑类的最近一次历史语音指令即可。

这里所说的逻辑类为控制意图所对应的具体控制领域或者具体控制场景等，通过逻辑类并结合环境中存在设备，可以确定具体的控制对象。

此外，需要说明的是，上述步骤 S135 中基于逻辑池来确定当前语音指令的控制对象的方法也可以作为一个独立的方案来执行，即不用在通过上述步骤 S130 到 S134 后无法

确定出控制对象的情况下才执行。

进一步地，如果在所述逻辑池中无记录或者无法根据所述逻辑类确定所述当前语音指令的控制对象的情况下，可以执行如下处理：

S136：获取从当前语音指令的控制意图对应的控制对象队列中获取优先级最高的控制对象作为当前语音指令的控制对象，其中，所述控制对象队列中记录有根据用户行为习惯统计获得的所述控制意图对应的控制对象，并按照统计次数的多少进行优先级的排序，统计次数越高优先级越高。例如，根据用户的行为习惯，在“播放”这个意图下，对音箱的控制次数较多，则可以在步骤 S130 到 S134 都无法确定控制对象的情况下，将控制对象确定为音箱的音乐播放进程。

本实施例的语音控制对象识别方法，通过对设备的运行状态和控制场景之间的匹配关系进行分析，从而确定语音指令的控制对象，减少实际控制对象和用户想要的控制对象之间的判断偏差，提高设备的智能化水平。此外，本实施例还引入了逻辑池机制和控制对象队列等机制来进行独立或者辅助的控制对象判断，从而进一步提高了确定语音指令的控制对象的准确度。

实施例十一

如图 14 所示，为本发明实施例的语音控制对象识别装置结构图，该语音控制对象识别装置可设置在云端中，可用于执行如图 13 所示的方法步骤，其包括：

语义识别模块 141，用于识别当前语音指令中体现控制意图的第一语义单元；

场景确定模块 142，用于根据控制意图确定对应的一个或多个控制场景；

状态获取模块 143，用于获取各个设备当前所处的运行状态；

第一对象确定模块 144，用于根据各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定当前语音指令的控制对象。

进一步地，如图 15 所示，上述装置中还包括：

第一处理模块 145，用于识别当前语音指令中体现控制对象的第二语义单元，如果存在第二语义单元，则根据第二语义单元确定控制对象或者排除掉部分控制对象后，执行获取当前语音指令中体现控制意图的第一语义单元的处理，否则，执行获取当前语音指令中体现控制意图的第一语义单元的处理。

进一步地，如图 15 所示，上述语音控制对象识别装置中还可包括：逻辑类获取模块 146 用于，

在无法获取到各个设备当前所处的运行状态或者无法根据获取到运行状态确定当前语音指令的控制对象的情况下，

获取当前语音指令的控制意图所对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池包括多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令；

第二对象确定模块 147，用于根据所述逻辑类确定所述当前语音指令的控制对象。

进一步地，如图 15 所示，上述语音控制对象识别装置中还可包括：第二处理模块 148 用于，

在逻辑池中无记录或者无法根据逻辑类确定当前语音指令的控制对象的情况下，

获取从当前语音指令的控制意图对应的控制对象队列中获取优先级最高的控制对象作为当前语音指令的控制对象，其中，控制对象队列中记录有根据用户行为习惯统计获得的控制意图对应的控制对象，并按照统计次数的多少进行优先级的排序，统计次数越高优先级越高。

需要说明的是，上述逻辑类获取模块 146 和第二对象确定模块 147 也可以单独构成一个语音控制对象识别装置，如图 16 所示，直接根据当前语音指令进行语音控制对象的识别处理。

实施例十二

前面图 14 描述了语音控制对象识别装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 17 所示，其为本发明实施例的电子设备的结构示意图，具体包括：存储器 171 和处理器 172。

存储器 171，用于存储程序。

除上述程序之外，存储器 171 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 171 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 172，耦合至存储器 171，用于执行存储器 171 中的程序，以用于：

识别当前语音指令中体现控制意图的第一语义单元；

根据所述控制意图确定对应的一个或多个控制场景；

获取各个设备当前所处的运行状态；

根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当前语音指令的控制对象。

上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 17 所示，电子设备还可以包括：通信组件 173、电源组件 174、音频组件 175、显示器 176 等其它组件。图 17 中仅示意性给出部分组件，并不意味着电子设备只包括图 17 所示组件。

通信组件 173 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 173 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 173 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

电源组件 174，为电子设备的各种组件提供电力。电源组件 174 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 175 被配置为输出和/或输入音频信号。例如，音频组件 175 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 171 或经由通信组件 173 发送。在一些实施例中，音频组件 175 还包括一个扬声器，用于输出音频信号。

显示器 176 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

实施例十三

前面图 16 描述了语音控制对象识别装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 18 所示，其为本发明实施例的电子设备的结构示意图，具体包括：

存储器 181 和处理器 182。

存储器 181，用于存储程序。

除上述程序之外，存储器 181 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 181 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 182，耦合至存储器 181，用于执行存储器 181 中的程序，以用于：

获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

根据所述逻辑类确定所述当前语音指令的控制对象。

上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 18 所示，电子设备还可以包括：通信组件 183、电源组件 184、音频组件 185、显示器 186 等其它组件。图 18 中仅示意性给出部分组件，并不意味着电子设备只包括图 18 所示组件。

通信组件 183 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 183 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 183 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

电源组件 184，为电子设备的各种组件提供电力。电源组件 184 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 185 被配置为输出和/或输入音频信号。例如，音频组件 185 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 181 或经由通信组件 183 发送。在一些实施例中，音频组件 185 还包括一个扬声器，用于输出音频信号。

显示器 186 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

本领域普通技术人员可以理解：实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时，执行包括上述各方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权 利 要 求 书

1.一种唤醒处理方法，其特征在于，包括：

在设备被唤醒后，检测是否有语音输入；

如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

2.根据权利要求 1 所述的方法，其特征在于，所述方法还包括：

如果在输出表示设备处于唤醒状态的提示之后的预定的第二时间内没有检测到语音输入，则关闭唤醒状态。

3.根据权利要求 1 或 2 所述的方法，其特征在于，所述表示设备处于唤醒状态的提示为语音提示。

4.一种唤醒处理装置，其特征在于，包括：

语音检测模块，用于在设备被唤醒后，检测是否有语音输入；

唤醒提示模块，用于如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

5.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

在设备被唤醒后，检测是否有语音输入；

如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示。

6.一种唤醒灵敏度控制方法，其特征在于，包括：

获取设备当前的应用场景信息；

根据所述应用场景信息，调整所述设备的唤醒灵敏度。

7.根据权利要求 6 所述的方法，其特征在于，所述方法还包括：接收设备的应用的灵敏度设置信息，

所述根据所述应用场景信息，调整所述设备的唤醒灵敏度包括：

根据所述应用场景信息和所述灵敏度设置信息，设置所述设备的唤醒引擎的唤醒参数和/或设置所述设备的声音传感器的声音采集参数。

8.根据权利要求 6 所述的方法，其特征在于，所述应用场景信息包括设备所处的时间段信息。

9.一种唤醒灵敏度控制装置，其特征在于，包括：

信息获取模块，用于获取设备当前的应用场景信息；

灵敏度调整模块，用于根据所述应用场景信息，调整所述设备的唤醒灵敏度。

10.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

获取设备当前的应用场景信息；

根据所述应用场景信息，调整所述设备的唤醒灵敏度。

11.一种快捷唤醒处理方法，其特征在于，包括：

对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

对所述第二文本执行语义解析。

12.根据权利要求 11 所述的处理方法，其特征在于，所述对来自设备的包含唤醒词的第一音频信息进行识别包括：

使用云端的识别模型对所述第一音频信息进行识别，其中，所述识别模型包括所述设备中唤醒词识别引擎所使用的唤醒词词库。

13.根据权利要求 11 所述的处理方法，其特征在于，所述针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本包括：

将所述第一文本发送给所述设备；

所述设备根据该设备的唤醒词识别引擎所使用的唤醒词词库，过滤掉所述第一文本中的唤醒词，生成所述第二文本，并将该第二文本发送给云端。

14.一种快捷唤醒处理装置，其特征在于，包括：

文本生成模块，用于对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

唤醒词过滤模块，用于针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

语义解析模块，用于对所述第二文本执行语义解析。

15.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本；

针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本；

对所述第二文本执行语义解析。

16.一种语音控制对象识别方法，其特征在于，包括：

识别当前语音指令中体现控制意图的第一语义单元；

根据所述控制意图确定对应的一个或多个控制场景；

获取各个设备当前所处的运行状态；

根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当前语音指令的控制对象。

17.根据权利要求 16 所述的方法，其特征在于，所述方法还包括：

识别当前语音指令中体现控制对象的第二语义单元，如果存在所述第二语义单元，则根据所述第二语义单元确定所述控制对象或者排除掉部分控制对象后，执行所述获取当前语音指令中体现控制意图的第一语义单元的处理，否则，执行所述获取当前语音指令中体现控制意图的第一语义单元的处理。

18.根据权利要求 16 所述的方法，其特征在于，所述方法还包括：

在无法获取到各个设备当前所处的运行状态或者无法根据获取到运行状态确定所述当前语音指令的控制对象的情况下，

获取当前语音指令的控制意图所对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池包括多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

根据所述逻辑类确定所述当前语音指令的控制对象。

19.根据权利要求 18 所述的方法，其特征在于，所述方法还包括：

在所述逻辑池中无记录或者无法根据所述逻辑类确定所述当前语音指令的控制对象的情况下，

获取从当前语音指令的控制意图对应的控制对象队列中获取优先级最高的控制对象作为当前语音指令的控制对象，其中，所述控制对象队列中记录有根据用户行为习惯统计获得的所述控制意图对应的控制对象，并按照统计次数的多少进行优先级的排序，统计次数越高优先级越高。

20.一种语音控制对象的识别方法，其特征在于，包括：

获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令；

根据所述逻辑类确定所述当前语音指令的控制对象。

21.一种语音控制对象的识别装置，其特征在于，包括：

语义识别模块，用于识别当前语音指令中体现控制意图的第一语义单元；

场景确定模块，用于根据所述控制意图确定对应的一个或多个控制场景；

状态获取模块，用于获取各个设备当前所处的运行状态；

第一对象确定模块，用于根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当前语音指令的控制对象。

22.一种语音控制对象的识别装置，其特征在于，包括：

逻辑类获取模块，用于获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

第二对象确定模块，用于根据所述逻辑类确定所述当前语音指令的控制对象。

23.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

识别当前语音指令中体现控制意图的第一语义单元；

根据所述控制意图确定对应的一个或多个控制场景；

获取各个设备当前所处的运行状态；

根据所述各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定所述当前语音指令的控制对象。

24.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

获取当前语音指令对应的逻辑池中记录的最近一次语音指令所对应的逻辑类，所述逻辑池多个逻辑类，每个逻辑类中记录有属于该逻辑类的历史语音指令，

根据所述逻辑类确定所述当前语音指令的控制对象。

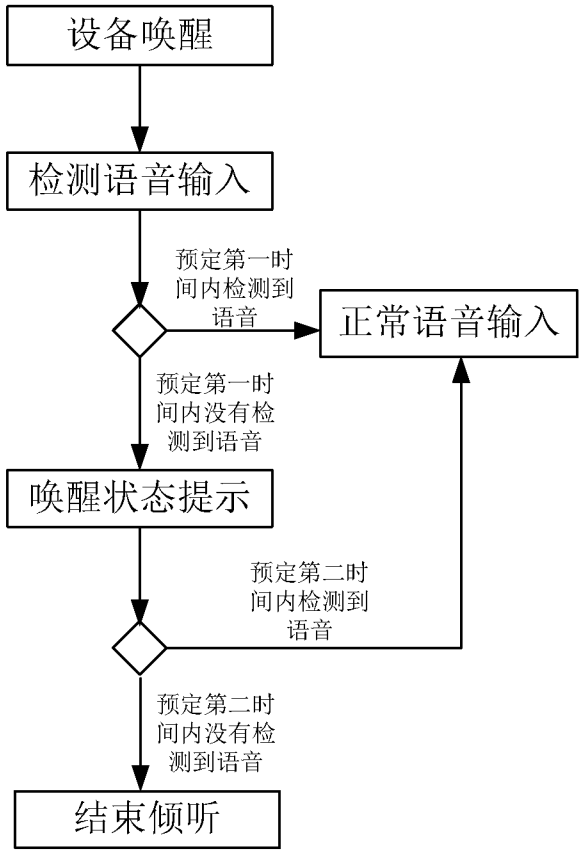

设备唤醒
检测语音输入
预定第一时间内检测到语音
正常语音输入
预定第一时间内没有检测到语音
唤醒状态提示
预定第二时间内检测到语音
预定第二时间内没有检测到语音
结束倾听

图 1

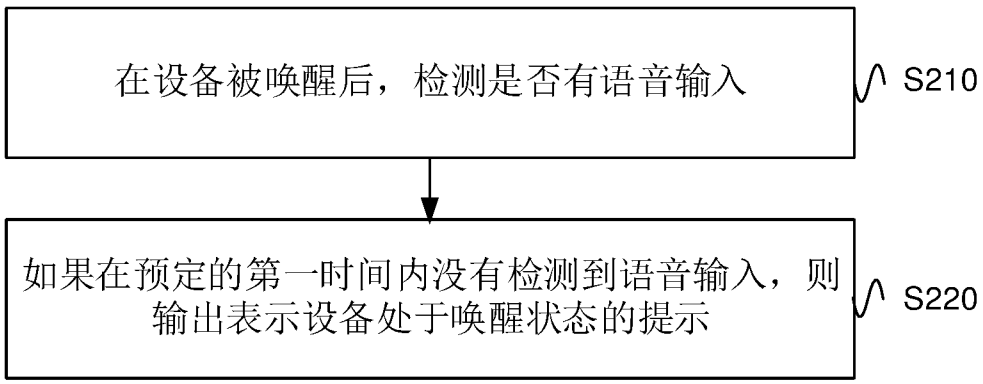

在设备被唤醒后，检测是否有语音输入
S210
如果在预定的第一时间内没有检测到语音输入，则输出表示设备处于唤醒状态的提示
S220

图 2

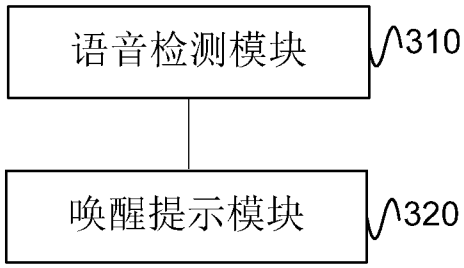


图 3a

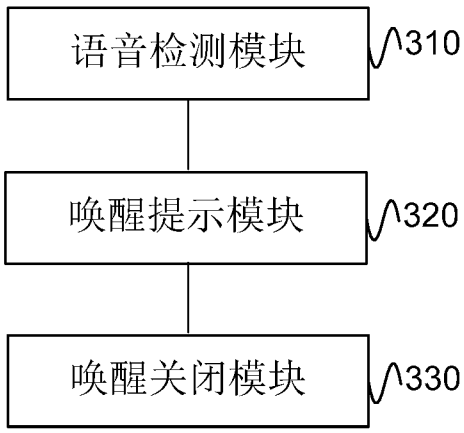


图 3b

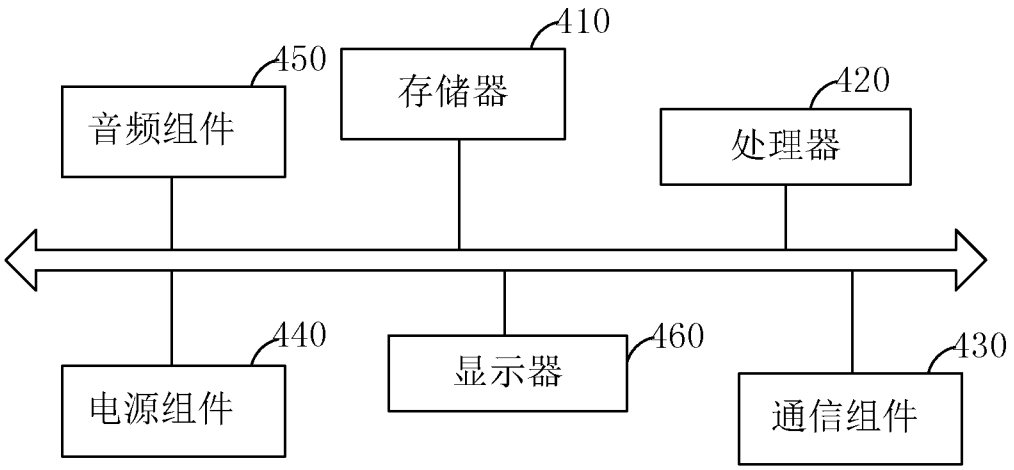

410
450
存储器
420
音频组件
处理器
440
460
430
显示器
电源组件
通信组件

图 4

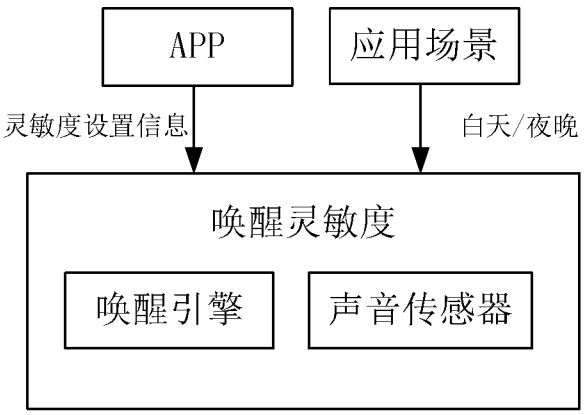

APP
应用场景
灵敏度设置信息
白天/夜晚
唤醒灵敏度
唤醒引擎
声音传感器

图 5

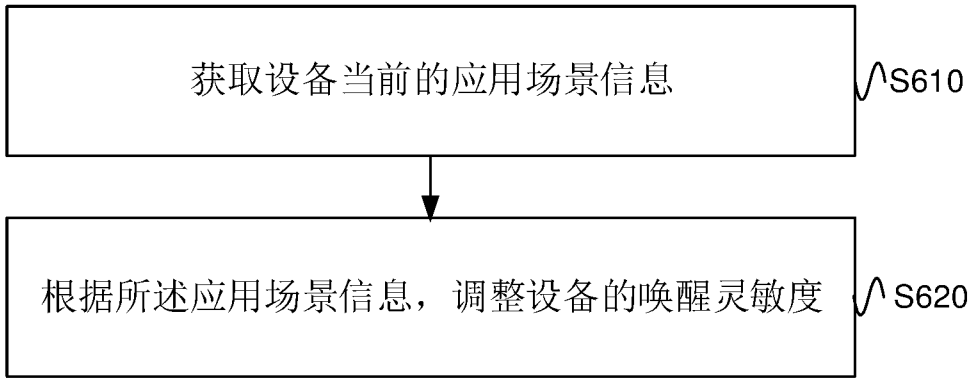

获取设备当前的应用场景信息
S610
根据所述应用场景信息，调整设备的唤醒灵敏度
S620

图 6

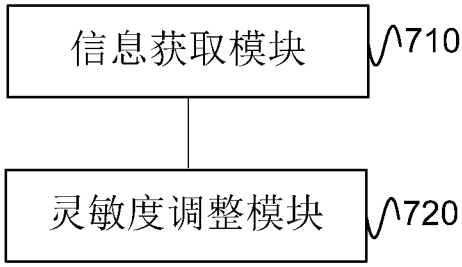

信息获取模块
710
灵敏度调整模块
720

图 7

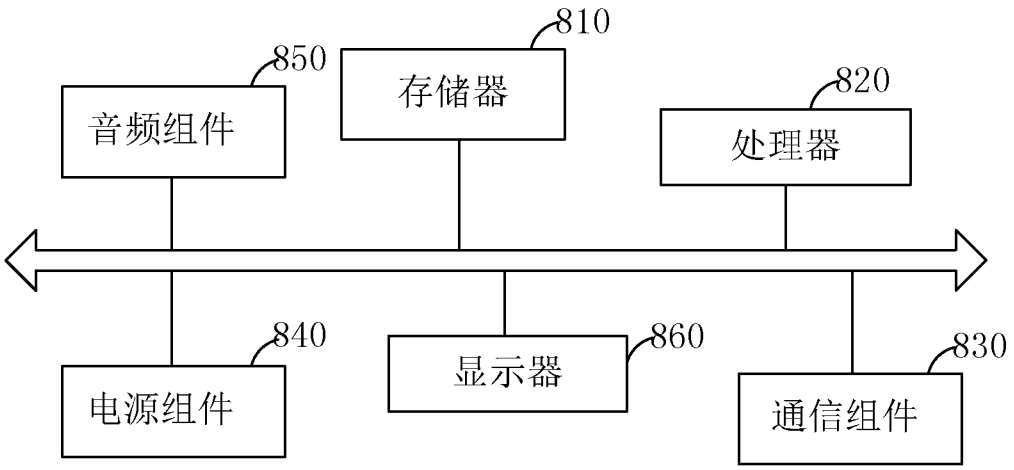

810
850
存储器
820
音频组件
处理器
840
860
830
显示器
电源组件
通信组件

图 8

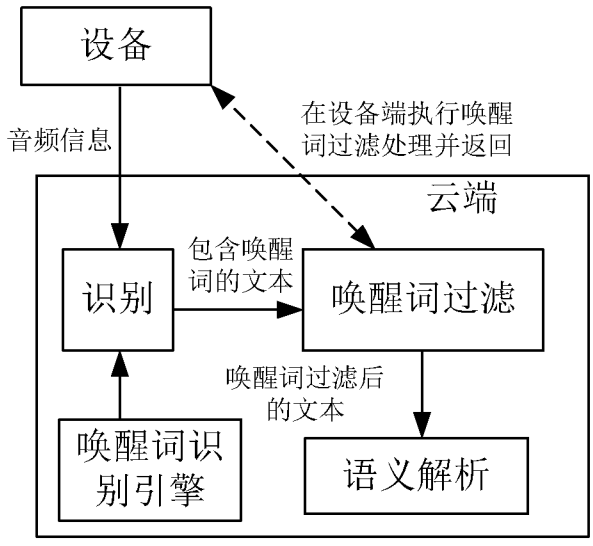

设备
音频信息
在设备端执行唤醒词过滤处理并返回
云端
包含唤醒词的文本
识别
唤醒词过滤
唤醒词过滤后的文本
唤醒词识别引擎
语义解析

图 9

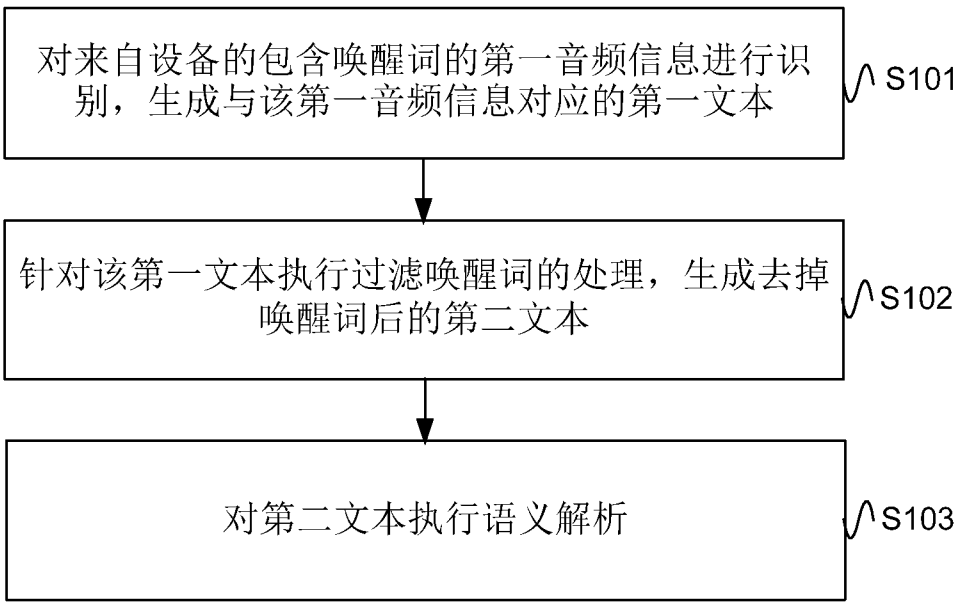

对来自设备的包含唤醒词的第一音频信息进行识别，生成与该第一音频信息对应的第一文本
S101
针对该第一文本执行过滤唤醒词的处理，生成去掉唤醒词后的第二文本
S102
对第二文本执行语义解析
S103

图 10

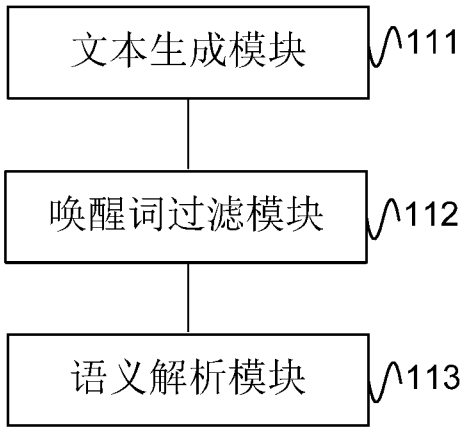

文本生成模块
111
唤醒词过滤模块
112
语义解析模块
113

图 11

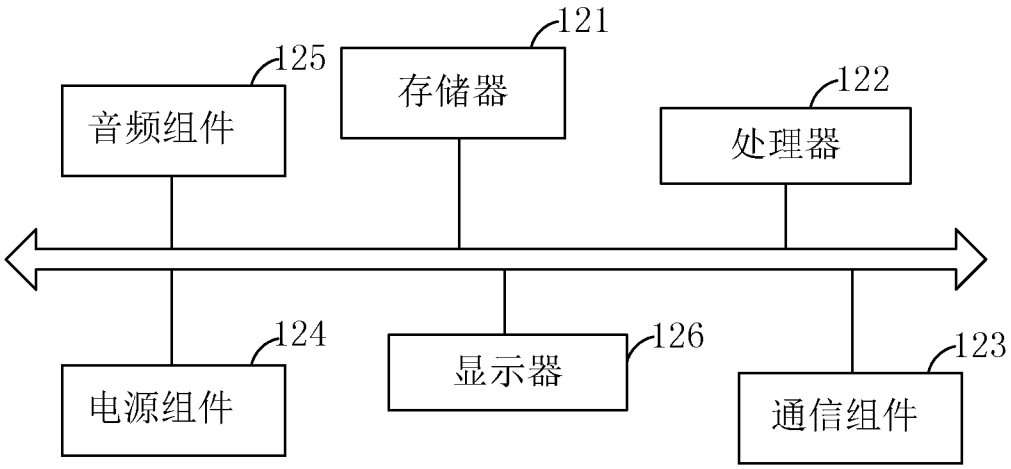


图 12

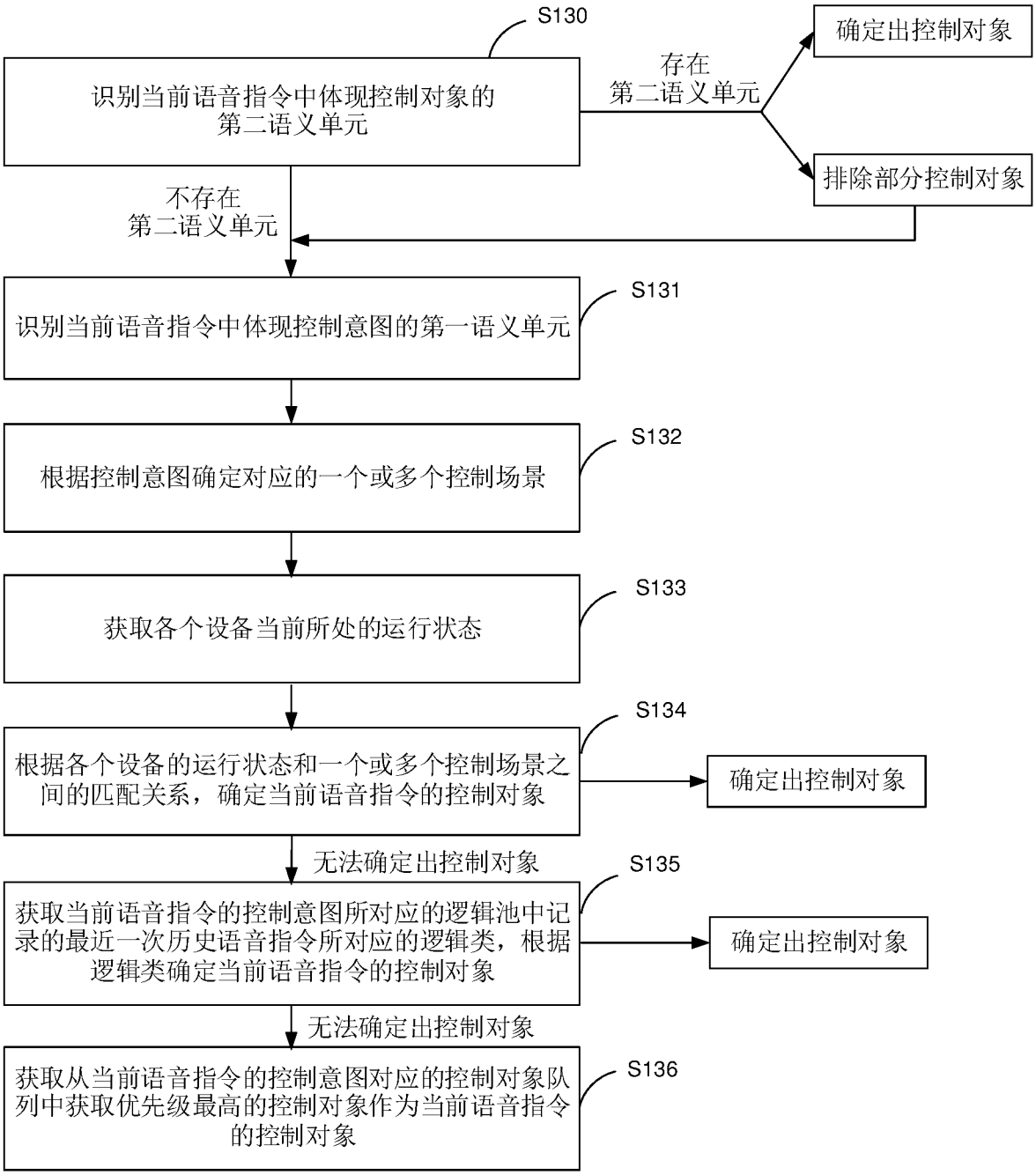

S130
识别当前语音指令中体现控制对象的第二语义单元
存在第二语义单元
确定出控制对象
排除部分控制对象
不存在第二语义单元
S131
识别当前语音指令中体现控制意图的第一语义单元
S132
根据控制意图确定对应的一个或多个控制场景
S133
获取各个设备当前所处的运行状态
S134
根据各个设备的运行状态和一个或多个控制场景之间的匹配关系，确定当前语音指令的控制对象
确定出控制对象
无法确定出控制对象
S135
获取当前语音指令的控制意图所对应的逻辑池中记录的最近一次历史语音指令所对应的逻辑类，根据逻辑类确定当前语音指令的控制对象
确定出控制对象
无法确定出控制对象
S136
获取从当前语音指令的控制意图对应的控制对象队列中获取优先级最高的控制对象作为当前语音指令的控制对象

图 13

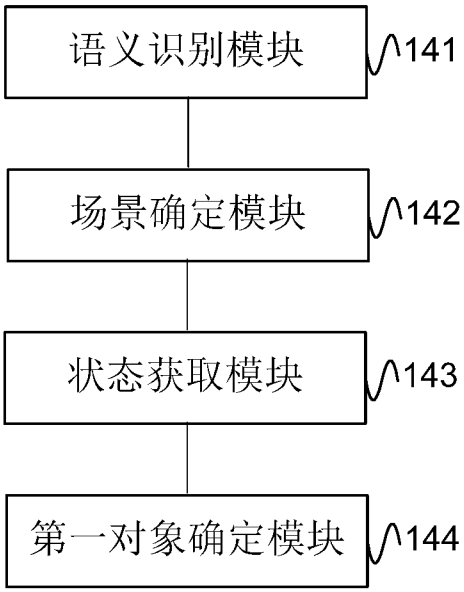


图 14

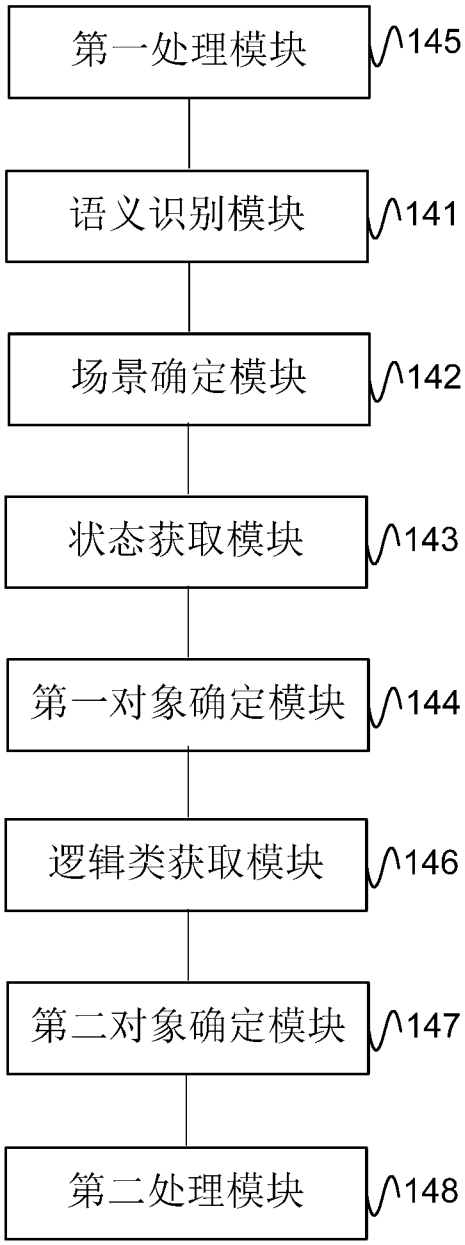

第一处理模块
145
语义识别模块
141
场景确定模块
142
状态获取模块
143
第一对象确定模块
144
逻辑类获取模块
146
第二对象确定模块
147
第二处理模块
148

图 15

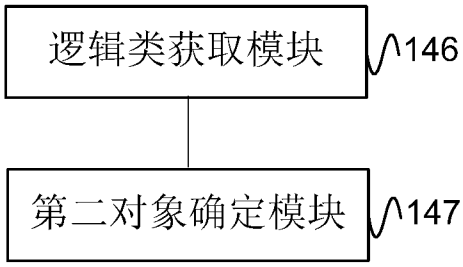

逻辑类获取模块
146
第二对象确定模块
147

图 16

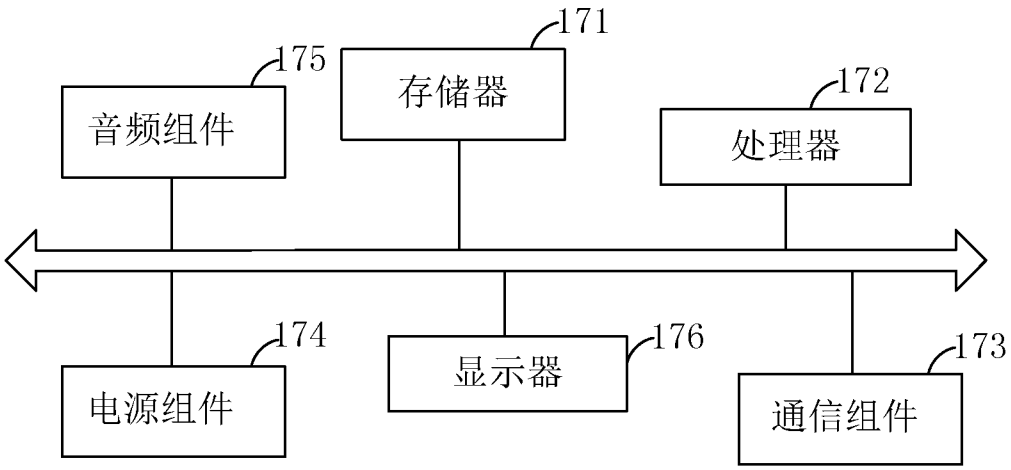

171
175
存储器
音频组件
172
处理器
174
176
173
电源组件
显示器
通信组件

图 17

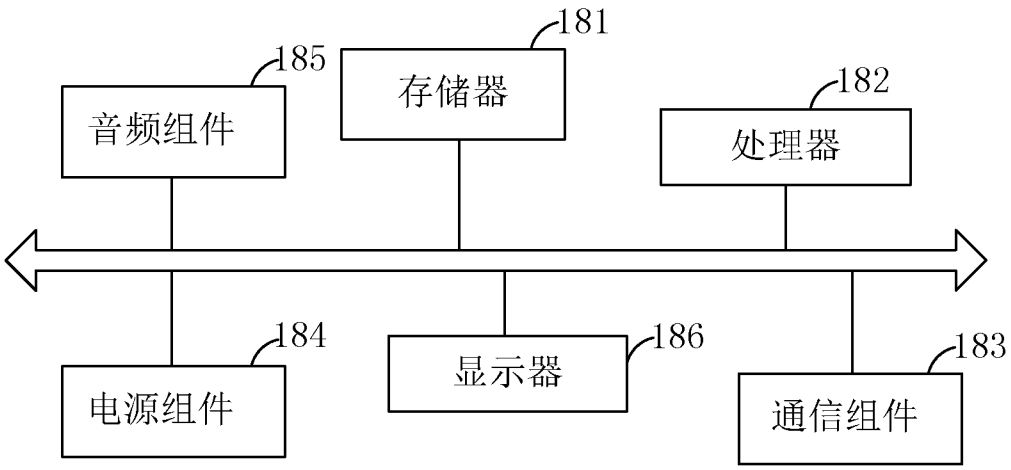


图 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/093216

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/18(2013.01)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; CNTXT; CNKI; USTXT; EPTXT; WOTXT; VEN: 阿里巴巴, 吴栓, 语音识别, 唤醒, 激活, 场景, 灵敏度, 语音提醒, 智能家居, 语义, 阈值, 唤醒词, 语音终端, 冲突, 逻辑类, 天猫精灵, voice activity detect+, logic+, siri, recognit+, prior+, AI, ASR, rouse, voice, activit+, artificial intelligen+, awak+, speech, wake, arouse, smart home, home automat+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104347072 A (GUANGDONG MIDEA AIR-CONDITIONING EQUIPMENT CO., LTD.) 11 February 2015 (2015-02-11) description, paragraphs [0091]-[0153], and figures 5-11	1-5
X	CN 103971680 A (HUAWEI DEVICE CO., LTD.) 06 August 2014 (2014-08-06) description, paragraphs [0040]-[0180], and figures 1-11	6-10
X	CN 106782554 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 31 May 2017 (2017-05-31) description, paragraphs [0040]-[0115], and figures 2-4	11-15
A	CN 105355201 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 24 February 2016 (2016-02-24) description, paragraphs [0023]-[0126], and figures 1-6	1-24
A	JP 2017058507 A (NIPPON TELEGR & TELEPH CORP. <NTT>) 23 March 2017 (2017-03-23) entire document	1-24

☐ 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
09 August 2018	**31 August 2018**

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

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/093216

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

[1] Independent claims 1, 4 and 5 are a first group of claims, and relate to a wakeup processing method and apparatus, and an electronic device; independent claims 6, 9 and 10 are a second group of claims and relate to a wakeup sensitivity control method and apparatus and an electronic device; independent claims 11, 14 and 15 are a third group of claims and relate to a fast wakeup processing method and apparatus, and an electronic device; independent claims 16, 21 and 23 are a fourth group of claims and relate to a method and apparatus for recognizing a voice control object, and an electronic device; and independent claims 20, 22 and 24 are a fifth group of claims and relate to a method and apparatus for recognizing a voice control object, and an electronic device. The above-mentioned first to fifth groups of claims do not have a same or corresponding special technical feature which presents the contribution of the invention over the prior art, there is no technical relationship therebetween, and they do not fall within a single general inventive concept, and therefore do not satisfy the requirement of unity and do not comply with PCT Rule 13.1.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☑ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/093216

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	104347072	A	11 February 2015	None			
CN	103971680	A	06 August 2014	JP	2017058691	A	23 March 2017
				US	9666186	B2	30 May 2017
				WO	2014114048	A1	31 July 2014
				US	2014207447	A1	24 July 2014
				EP	2763134	A1	06 August 2014
				CN	103971680	B	05 June 2018
				JP	6101196	B2	22 March 2017
				EP	2763134	B1	04 January 2017
				JP	2014142626	A	07 August 2014
CN	106782554	A	31 May 2017	None			
CN	105355201	A	24 February 2016	None			
JP	2017058507	A	23 March 2017	None			

国际检索报告

国际申请号 PCT/CN2018/093216

A. 主题的分类

G10L 15/18(2013.01)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;CNTXT;CNKI;USTXT;EPTXT;WOTXT;VEN:阿里巴巴，吴栓，语音识别，唤醒，激活，场景，灵敏度，语音提醒，智能家居，语义，阈值，唤醒词，语音终端，冲突，逻辑类，天猫精灵，voice activity detect+，logic+，siri，recognit+，prior+，AI，ASR，rouse，voice，activit+，artificial intelligen+，awak+，speech，wake，arouse，smart home，home automat+

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
X	CN 104347072 A (广东美的制冷设备有限公司) 2015年 2月 11日 (2015-02-11) 说明书第[0091]-[0153]段，图5-11	1-5
X	CN 103971680 A (华为终端有限公司) 2014年 8月 6日 (2014-08-06) 说明书第[0040]-[0180]段，图1-11	6-10
X	CN 106782554 A (百度在线网络技术北京有限公司) 2017年 5月 31日 (2017-05-31) 说明书第[0040]-[0115]段，图2-4	11-15
A	CN 105355201 A (百度在线网络技术北京有限公司) 2016年 2月 24日 (2016-02-24) 说明书第[0023]-[0126]段，图1-6	1-24
A	JP 2017058507 A (NIPPON TELEGR & TELEPH CORP <NTT>) 2017年 3月 23日 (2017-03-23) 全文	1-24

☐ 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
2018年 8月 9日	2018年 8月 31日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(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 陈小军 电话号码 86-(0512)-88997347

国际检索报告

国际申请号
PCT/CN2018/093216

第III栏　缺乏发明单一性的意见(续第1页第3项)

本国际检索单位在该国际申请中发现多项发明，即：

[1] 独立权利要求1、4、5为第一组权利要求，涉及一种唤醒处理方法、装置以及电子设备；独立权利要求6、9、10为第二组权利要求，涉及一种唤醒灵敏度控制方法、装置以及电子设备；独立权利要求11、14、15为第三组权利要求，涉及一种快捷唤醒处理方法、装置以及电子设备；独立权利要求16、21、23为第四组权利要求，涉及一种语音控制对象的识别方法、装置以及电子设备；独立权利要求20、22、24为第五组权利要求，涉及一种语音控制对象的识别方法、装置以及电子设备；上述第一至五组权利要求之间，两两不具有相同或相应的体现发明对现有技术作出贡献的特定技术特征，不存在技术关联，不属于一个总的发明构思，因此不满足单一性的要求，不符合PCT实施细则13.1的规定。

1. ☐ 由于申请人按时缴纳了被要求缴纳的全部附加检索费，本国际检索报告涉及全部可作检索的权利要求。

2. ☑ 由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求进行检索，本单位未通知缴纳任何加费。

3. ☐ 由于申请人仅按时缴纳了部分被要求缴纳的附加检索费，本国际检索报告仅涉及已缴费的那些权利要求，具体地说，是权利要求：

4. ☐ 申请人未按时缴纳被要求缴纳的附加检索费。因此，本国际检索报告仅涉及权利要求书中首先提及的发明；包含该发明的权利要求是：

对异议的意见

☐ 申请人缴纳了附加检索费，同时提交了异议书，适用时，缴纳了异议费。

☐ 申请人缴纳了附加检索费，同时提交了异议书，但未在通知书规定的时间期限内缴纳异议费。

☐ 缴纳附加检索费时未提交异议书。

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/093216

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	104347072	A	2015年 2月 11日	无			
CN	103971680	A	2014年 8月 6日	JP	2017058691	A	2017年 3月 23日
				US	9666186	B2	2017年 5月 30日
				WO	2014114048	A1	2014年 7月 31日
				US	2014207447	A1	2014年 7月 24日
				EP	2763134	A1	2014年 8月 6日
				CN	103971680	B	2018年 6月 5日
				JP	6101196	B2	2017年 3月 22日
				EP	2763134	B1	2017年 1月 4日
				JP	2014142626	A	2014年 8月 7日
CN	106782554	A	2017年 5月 31日	无			
CN	105355201	A	2016年 2月 24日	无			
JP	2017058507	A	2017年 3月 23日	无			