

(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 3 日 (03.12.2020)



(10) 国际公布号
WO 2020/239013 A1

- (51) 国际专利分类号:
G10L 15/22 (2006.01) *G10L 17/22* (2013.01)
G10L 15/26 (2006.01) *H04M 1/725* (2006.01)
- (21) 国际申请号: PCT/CN2020/092888
- (22) 国际申请日: 2020 年 5 月 28 日 (28.05.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910472665.5 2019年5月31日 (31.05.2019) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 胡鹏 (HU, Peng); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 黄德才 (HUANG, Decai); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (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, 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)。

(54) Title: INTERACTION METHOD AND TERMINAL DEVICE

(54) 发明名称: 一种交互方法和终端设备

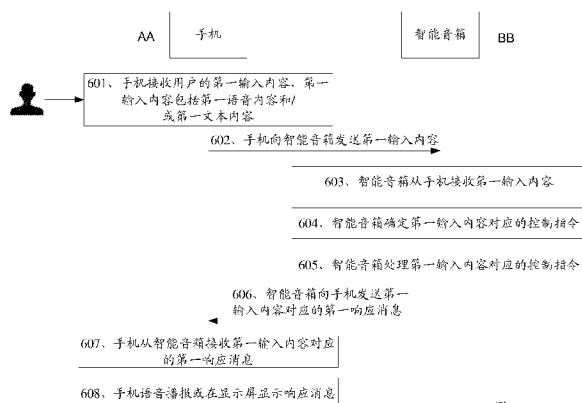


图 6

601 THE CELL PHONE RECEIVING FIRST INPUT CONTENT OF A USER, THE FIRST INPUT CONTENT COMPRISING FIRST VOICE CONTENT AND/OR FIRST TEXT CONTENT
602 THE CELL PHONE SENDING THE FIRST INPUT CONTENT TO THE SMART SPEAKER
603 THE SMART SPEAKER RECEIVING THE FIRST INPUT CONTENT FROM THE CELL PHONE
604 THE SMART SPEAKER DETERMINING A CONTROL INSTRUCTION CORRESPONDING TO THE FIRST INPUT CONTENT
605 THE SMART SPEAKER PROCESSING THE CONTROL INSTRUCTION CORRESPONDING TO THE FIRST INPUT CONTENT
606 THE SMART SPEAKER SENDING A FIRST RESPONSE MESSAGE CORRESPONDING TO THE FIRST INPUT CONTENT TO THE CELL PHONE
607 THE CELL PHONE RECEIVING THE FIRST RESPONSE MESSAGE CORRESPONDING TO THE FIRST INPUT CONTENT FROM THE SMART SPEAKER
608 THE CELL PHONE BROADCASTING THE RESPONSE MESSAGE BY MEANS OF VOICE OR DISPLAYING THE RESPONSE MESSAGE ON A DISPLAY SCREEN
AA CELL PHONE
BB SMART SPEAKER

(57) Abstract: Embodiments of the present application relate to the field of terminals, and provided thereby are an interaction method and a terminal device, which may enhance the flexibility of a user remotely operating an intelligent voice interaction device, and fully develop the capabilities of the intelligent voice interaction device. The method is as follows: a first terminal device receiving first input content from a second terminal device, the first input content comprising first voice content and/or first text content; the first terminal device determining a control instruction corresponding to the first input content; and the first terminal device processing the control instruction corresponding to the first input content. Embodiments of the present application are applied in remote interaction scenarios.

(57) 摘要: 本申请实施例提供一种交互方法和终端设备, 涉及终端领域, 能够增强用户远程操作智能语音交互设备的灵活性, 充分发挥智能语音交互设备的能力。其方法为: 第一终端设备从第二终端设备接收第一输入内容, 第一输入内容包括第一语音内容和/或第一文本内容; 第一终端设备确定第一输入内容对应的控制指令; 第一终端设备处理第一输入内容对应的控制指令。本申请实施例应用于远程交互场景中。

本国际公布：

- 包括国际检索报告(条约第21条(3))。

一种交互方法和终端设备

本申请要求在 2019 年 5 月 31 日提交中国国家知识产权局、申请号为 201910472665.5 的中国专利申请的优先权，发明名称为“一种交互方法和终端设备”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

本申请涉及终端领域，尤其涉及一种交互方法和终端设备。

背景技术

当前的智能语音交互设备（例如，智能音箱）是通过语音将用户与设备之间联系起来的。用户可以在智能音箱旁直接通过语音控制智能音箱，让智能音箱播放音乐、讲故事等等。但当用户远离智能音箱时（例如大于 5 米时），与智能音箱设备便无法取得联系。也就是说，用户在与智能音箱距离较近才能使用智能音箱，用户远离智能音箱时无法使用智能音箱，此时智能音箱处于闲置状态，造成资源浪费。

为解决上述问题，如图 1 所示，智能音箱与手机可以通过无线保真（wireless fidelity，WiFi）或者互联网（Internet）进行通信，用户可以通过操作手机应用（application，APP）上的指定菜单，发送指令给智能音箱，控制智能音箱进行音乐播放等操作。

但是，上述方案中，用户通常只能进行手机 APP 上已经预制好的有限的操作，例如音乐播放、暂停等，限制了智能音箱的功能。

发明内容

本申请实施例提供一种交互方法和终端设备，能够增强用户远程操作智能语音交互设备的灵活性，充分发挥智能语音交互设备的能力。

第一方面，本申请实施例提供一种交互方法，包括：第一终端设备从第二终端设备接收第一输入内容，第一输入内容包括第一语音内容和/或第一文本内容；第一终端设备确定第一输入内容对应的控制指令；第一终端设备处理第一输入内容对应的控制指令。

基于本申请实施例提供的方法，第一终端设备可以从第二终端设备接收第一输入内容，确定第一输入内容对应的控制指令，并处理该控制指令。这样一来，可以由第一终端设备基于第一输入内容（用户通过第二终端设备远程输入的内容）确定控制指令，而非用户根据手机 APP 上预制好的有限的指令进行操作，增强了用户与第一终端设备远程交互的灵活性，充分发挥第一终端设备的能力。

在一种可能的实现方式中，第一终端设备确定第一输入内容对应的控制指令包括：第一终端设备向服务器发送第一输入内容；第一终端设备从服务器接收第一输入内容对应的控制指令。服务器可以包括自动语音识别（automatic speech recognition，ASR）引擎和自然语言处理（natural language processing，NLP）引擎，ASR 引擎可以将第一语音内容转化为第一文本信息，NLP 引擎可以根据第一文本信息的语义得到控制指令。

在一种可能的实现方式中，第一终端设备处理第一输入内容对应的控制指令包括：第一终端设备执行第一输入内容对应的控制指令；或者，第一终端设备向第三终端设备发送第一输入内容对应的控制指令。第一终端设备可以接收服务器下发的控制指令，并转发给第三终端设备，从而实现对第三终端设备的控制，充分发挥了第一终端设备的能力，提高了用户体验。

在一种可能的实现方式中，该方法还包括：第一终端设备向第二终端设备发送第一输入内容对应的第一响应消息。第一响应消息可以包括服务器转写的控制指令以及智能音箱执行控制指令后的执行结果。以使用户通过第二终端设备输入第一输入内容后，可以获知根据第一输入内容得到的控制指令和针对该控制指令的执行结果，提高用户体验。

在一种可能的实现方式中，第一终端设备处理第一输入内容对应的控制指令包括：第一终端设备接收第二输入内容，第二输入内容包括第二语音内容和/或第二文本内容；若第一终端设备接收第二输入内容的时刻早于接收第一输入内容的时刻，第一终端设备处理第二输入内容对应的控制指令后，处理第一输入内容对应的控制指令。也就是说，智能音箱在处理不同的控制指令时，可以遵循先来先执行（先获取的控制指令先执行）的策略。

第二方面，本申请实施例提供一种交互方法，包括：第二终端设备从用户接收第一输入内容，第一输入内容包括第一语音内容和/或第一文本内容；第二终端设备向第一终端设备发送第一输入内容；第二终端设备从第一终端设备接收第一输入内容对应的第一响应消息；第二终端设备语音播报或在显示屏显示第一响应消息。

基于本申请实施例提供的方法，第二终端设备用户接收第一输入内容后，向第一终端设备发送第一输入内容，而后，从第一终端设备接收第一输入内容对应的第一响应消息，并语音播报或在显示屏显示第一响应消息。这样一来，用户通过第二终端设备输入第一输入内容后，可以获知根据第一输入内容得到的控制指令和针对该控制指令的执行结果，提高用户体验。

在一种可能的实现方式中，方法还包括：第二终端设备从第一终端设备接收第二响应消息，第二响应消息对应第二输入内容；第二终端设备语音播报或在显示屏显示第二响应消息。

这样一来，若其他用户（与输入第一输入内容的用户不同）在音箱旁输入语音内容（第二语音内容），或者其他用户通过第四终端设备（例如手机或智能穿戴设备等）输入第二输入内容，第一终端设备的用户可以获知其他用户对第二终端设备的操作（包括第二输入内容对应的控制指令和针对该控制指令的执行结果），提高了用户对第二终端设备的掌控能力，从而提高了用户体验。

第二方面及其各种可能的实现方式的技术效果可以参见第一方面及其各种可能的实现方式的技术效果，此处不再赘述。

第三方面，本申请实施例提供一种第一终端设备，包括：接收单元，用于从第二终端设备接收第一输入内容，第一输入内容包括第一语音内容和/或第一文本内容；确定单元，用于确定第一输入内容对应的控制指令；处理单元，用于处理第一输入内容对应的控制指令。

在一种可能的实现方式中，确定单元用于：通过发送单元向服务器发送第一输入内容；通过接收单元从服务器接收第一输入内容对应的控制指令。

在一种可能的实现方式中，处理单元用于：执行第一输入内容对应的控制指令；或者，通过发送单元向第三终端设备发送第一输入内容对应的控制指令。

在一种可能的实现方式中，发送单元还用于：向第二终端设备发送第一输入内容对应的第一响应消息。

在一种可能的实现方式中，处理单元用于：通过接收单元接收第二输入内容，第二输入内容包括第二语音内容和/或第二文本内容；若第一终端设备接收第二输入内容的时刻早于接收第一输入内容的时刻，处理第二输入内容对应的控制指令后，处理第一输入内容对应的控制指令。

第四方面，本申请实施例提供一种第二终端设备，包括：接收单元，用于从用户接收第

一输入内容，第一输入内容包括第一语音内容和/或第一文本内容；发送单元，用于向第一终端设备发送第一输入内容；接收单元，还用于从第一终端设备接收第一输入内容对应的第一响应消息；处理单元，用于语音播报或在显示屏显示第一响应消息。

在一种可能的实现方式中，接收单元还用于：从第一终端设备接收第二响应消息，第二响应消息对应第二输入内容；处理单元，还用于语音播报或在显示屏显示第二响应消息。

第五方面，本申请实施例还提供了一种装置，该装置可以是第一终端设备或芯片。该装置包括处理器，用于实现上述第一方面提供的任意一种交互方法。该装置还可以包括存储器，用于存储程序指令和数据，存储器可以是集成在该装置内的存储器，或设置在该装置外的片外存储器。该存储器与该处理器耦合，该处理器可以调用并执行该存储器中存储的程序指令，用于实现上述第一方面提供的任意一种交互方法。该装置还可以包括通信接口，该通信接口用于该装置与其它设备（例如，第二终端设备）进行通信。

第六方面，本申请实施例还提供了一种装置，该装置可以是第二终端设备或芯片。该装置包括处理器，用于实现上述第二方面提供的任意一种交互方法。该装置还可以包括存储器，用于存储程序指令和数据，存储器可以是集成在该装置内的存储器，或设置在该装置外的片外存储器。该存储器与该处理器耦合，该处理器可以调用并执行该存储器中存储的程序指令，用于实现上述第二方面提供的任意一种交互方法。该装置还可以包括通信接口，该通信接口用于该装置与其它设备（例如，第一终端设备）进行通信。

第七方面，本申请实施例提供一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行上述第一方面或第二方面提供的任意一种交互方法。

第八方面，本申请实施例提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得计算机执行上述第一方面或第二方面提供的任意一种交互方法。

第九方面，本申请实施例提供了一种芯片系统，该芯片系统包括处理器，还可以包括存储器，用于实现上述第一方面或第二方面提供的任意一种交互方法。该芯片系统可以由芯片构成，也可以包含芯片和其他分立器件。

第十方面，本申请实施例提供了一种交互系统，所述系统包括第三方面中的第一终端设备和第四方面中的第二终端设备。

附图说明

图 1 为现有技术中的一种智能音箱与手机的通信架构示意图；

图 2 为一种适用于本申请实施例提供的交互方法的架构示意图；

图 3 为又一种适用于本申请实施例提供的交互方法的架构示意图；

图 4 为本申请实施例提供的一种第二终端设备的内部结构示意图；

图 5 为本申请实施例提供的一种第一终端设备的内部结构示意图；

图 6 为一种适用于本申请实施例提供的交互方法的信号交互示意图；

图 7 为本申请实施例提供的一种手机的桌面示意图；

图 8 为本申请实施例提供的一种智能音箱 APP 的输入界面示意图；

图 9 为本申请实施例提供的又一种智能音箱 APP 的输入界面示意图；

图 10 为本申请实施例提供的又一种第一终端设备的内部结构示意图；

图 11 为本申请实施例提供的又一种第二终端设备的内部结构示意图。

具体实施方式

本申请实施例提供一种交互方法和终端设备，应用于第一终端设备和第二终端设备组成的交互系统中，用户可以通过第二终端设备与第一终端设备远程交互。例如，应用于手机与

(配对的)智能家居产品(例如,智能音箱)组成的交互系统中。第一终端设备和第二终端设备之间可以通过新无线接入(new radio access technical, New RAT)、长期演进(long term evolution, LTE)、蓝牙(bluetooth, BT)、Wifi或其它协议进行通信,本申请不做限定。

如图2所示,为本申请实施例提供的一种交互系统的架构示意图,该系统可以包括第一终端设备(例如,手机10a)、第二终端设备(例如,智能音箱10b)、第一网络设备(例如,internet服务器11)和第二网络设备(例如,云服务器12)。第一终端设备可以通过internet服务器11接收第二终端设备发送的第一输入内容,并可以通过internet服务器11向第二终端设备发送第一输入内容的第一响应消息等。云服务器12可以用于解析语音内容和/或文本内容。例如,云服务器12可以通过ASR引擎将语音内容转换成文本信息,通过NLP引擎将文本信息转换成控制指令,以便第二终端设备根据控制指令做出响应。云服务器12可以是手机10a和智能音箱10b上安装的智能音箱APP对应的服务器,或者可以为集成在其他APP内的智能音箱程序对应的第三方服务器,本申请不做限定。

第一终端设备和internet服务器11之间,第二终端设备和internet服务器11,第二终端设备和云服务器12之间,可以通过无线的通信方式进行通信,无线的通信方式例如可以通过无线接入网设备(例如,基站)进行通信。在LTE网络中,基站可以为演进型基站(evolved node base station, eNB)。在第五代移动通信技术(5-Generation, 5G)网络中,基站可以为下一代基站(next generation node base station, gNB)、新型无线电基站(new radio eNB)、宏基站、微基站、高频基站或发送和接收点(transmission and reception point, TRP)等。

其中,本申请实施例提供的第一终端设备可以是用户设备(user equipment, UE),例如可以为手机、平板电脑、桌面型、膝上型笔记本电脑、超级移动个人计算机(ultra-mobile personal computer, UMPC)、手持计算机、上网本、个人数字助理(personal digital assistant, PDA)、车载终端等设备。第二终端设备可以为各种智能家居设备或UE,智能家居设备例如可以为智能音箱、智能电视、智能冰箱、智能洗衣机、智能电饭煲、智能洗碗机、智能扫地机器人等等。

在一种可能的设计中,如图3所示,交互系统还可以包括第三终端设备(例如扫地机器人10c),第三终端设备与第二终端设备连接。第三终端设备可以为各种智能家居设备或UE等等。

如图4所示,上述通信系统架构中的第二终端设备具体可以为手机100。手机100可以包括处理器110,内部存储器120,摄像头130,显示屏140,射频模块150,通信模块160,天线1,天线2,音频模块170,扬声器170A,受话器170B,麦克风170C,耳机接口170D等。

本发明实施例示意的结构并不构成对手机100的限定。可以包括比图示更多或更少的部件,或者组合某些部件,或者拆分某些部件,或者不同的部件布置。图示的部件可以以硬件,软件或软件和硬件的组合实现。

处理器110可以包括一个或多个处理单元,例如:处理器110可以包括应用处理器(application processor, AP),调制解调处理器,图形处理器(graphics processing unit, GPU),图像信号处理器(image signal processor, ISP),控制器,存储器,视频编解码器,数字信号处理器(digital signal processor, DSP),基带处理器,和/或神经网络处理器(Neural-network Processing Unit, NPU)等。其中,不同的处理单元可以是独立的器件,也可以是集成在同一个处理器中。

控制器可以是指手机100的各个部件按照指令协调工作的决策者。是手机100的神经

中枢和指挥中心。控制器根据指令操作码和时序信号，产生操作控制信号，完成取指令和执行指令的控制。

处理器 110 中还可以设置存储器，用于存储指令和数据。在一些实施例中，处理器中的存储器为高速缓冲存储器。可以保存处理器刚用过或循环使用的指令或数据。如果处理器需要再次使用该指令或数据，可从所述存储器中直接调用。避免了重复存取，减少了处理器的等待时间，因而提高了系统的效率。

在一些实施例中，处理器 110 可以包括接口。其中接口可以包括集成电路(inter-integrated circuit, I2C)接口，集成电路内置音频(inter-integrated circuit sound, I2S)接口，脉冲编码调制(pulse code modulation, PCM)接口，通用异步收发传输器(universal asynchronous receiver/transmitter, UART)接口，移动产业处理器接口(mobile industry processor interface, MIPI)，通用输入输出(general-purpose input/output, GPIO)接口，用户标识模块(subscriber identity module, SIM)接口，和/或通用串行总线(universal serial bus, USB)接口等。

手机 100 的无线通信功能可以通过天线 1，天线 2，射频模块 150，通信模块 160，调制解调器以及基带处理器等实现。

天线 1 和天线 2 用于发射和接收电磁波信号。手机 100 中的每个天线可用于覆盖单个或多个通信频带。不同的天线还可以复用，以提高天线的利用率。例如：可以将蜂窝网天线复用为无线局域网分集天线。在一些实施例中，天线可以和调谐开关结合使用。

射频模块 150 可以提供应用在手机 100 上的包括第二代(2nd generation, 2G)/第三代(3rd generation, 3G)/第四代(4th generation, 4G)/第五代(5th generation, 5G)等无线通信的解决方案的通信处理模块。可以包括至少一个滤波器，开关，功率放大器，低噪声放大器(low noise amplifier, LNA)等。射频模块由天线 1 接收电磁波，并对接收的电磁波进行滤波，放大等处理，传送至调制解调器进行解调。射频模块 150 还可以对经调制解调器调制后的信号放大，经天线 1 转为电磁波辐射出去。在一些实施例中，射频模块 150 的至少部分功能模块可以被设置于处理器 110 中。在一些实施例中，射频模块 150 的至少部分功能模块可以与处理器 110 的至少部分模块被设置在同一个器件中。

调制解调器可以包括调制器和解调器。调制器用于将待发送的低频基带信号调制成中高频信号。解调器用于将接收的电磁波信号解调为低频基带信号。随后解调器将解调得到的低频基带信号传送至基带处理器处理。低频基带信号经基带处理器处理后，被传递给应用处理器。应用处理器通过音频设备(不限于扬声器，受话器等)输出声音信号，或通过显示屏显示图像或视频。在一些实施例中，调制解调器可以是独立的器件。在一些实施例中，调制解调器可以独立于处理器，与射频模块或其他功能模块设置在同一个器件中。

通信模块 160 可以提供应用在手机 100 上的包括无线局域网(wireless local area networks, WLAN)(例如，WiFi)、蓝牙，全球导航卫星系统(global navigation satellite system, GNSS)，调频(frequency modulation, FM)，近距离无线通信技术(near field communication, NFC)，红外技术(infrared, IR)等无线通信的解决方案的通信处理模块。通信模块 160 可以是集成至少一个通信处理模块的一个或多个器件。通信模块 160 经由天线 2 接收电磁波，将电磁波信号调频以及滤波处理，将处理后的信号发送到处理器。通信模块 160 还可以从处理器接收待发送的信号，对其进行调频，放大，经天线 2 转为电磁波辐射出去。

在一些实施例中，手机 100 的天线 1 和射频模块 150 耦合，天线 2 和通信模块 160 耦合。使得手机 100 可以通过无线通信技术与网络以及其他设备通信。所述无线通信技术可以包括全球移动通讯系统(global system for mobile communications, GSM)，通用分组无线服务(general

packet radio service, GPRS), 码分多址接入 (code division multiple access, CDMA), 宽带码分多址 (wideband code division multiple access, WCDMA), 时分码分多址 (time-division code division multiple access, TD-SCDMA), LTE, 5G 新无线通信 (New Radio, NR), BT, GNSS, WLAN, NFC, FM, 和/或 IR 技术等。所述 GNSS 可以包括全球卫星定位系统(global positioning system, GPS), 全球导航卫星系统 (global navigation satellite system, GLONASS), 北斗卫星导航系统 (beidou navigation satellite system, BDS), 准天顶卫星系统 (quasi-zenith satellite system, QZSS) 和/或星基增强系统 (satellite based augmentation systems, SBAS)。

手机 100 通过 GPU, 显示屏 140, 以及应用处理器等实现显示功能。GPU 为图像处理的微处理器, 连接显示屏和应用处理器。GPU 用于执行数学和几何计算, 用于图形渲染。处理器 110 可包括一个或多个 GPU, 其执行程序指令以生成或改变显示信息。

显示屏 140 用于显示图像, 视频等。显示屏包括显示面板。显示面板可以采用液晶显示屏(liquid crystal display, LCD), 有机发光二极管(organic light-emitting diode, OLED), 有源矩阵有机发光二极体或主动矩阵有机发光二极体(active-matrix organic light emitting diode 的, AMOLED), Miniled, MicroLed, Micro-oLed, 量子点发光二极管 (quantum dot light emitting diodes, QLED)等。在一些实施例中, 手机 100 可以包括 1 个或 N 个显示屏, N 为大于 1 的正整数。

手机 100 可以通过 ISP, 摄像头 130, 视频编解码器, GPU, 显示屏 140 以及应用处理器等实现拍摄功能。

ISP 用于处理摄像头反馈的数据。例如, 拍照时, 打开快门, 光线通过镜头被传递到摄像头感光元件上, 光信号转换为电信号, 摄像头感光元件将所述电信号传递给 ISP 处理, 转化为肉眼可见的图像。ISP 还可以对图像的噪点, 亮度, 肤色进行算法优化。ISP 还可以对拍摄场景的曝光, 色温等参数优化。在一些实施例中, ISP 可以设置在摄像头 130 中。

摄像头 130 用于捕获静态图像或视频。物体通过镜头生成光学图像投射到感光元件。感光元件可以是电荷耦合器件 (charge coupled device, CCD) 或互补金属氧化物半导体 (complementary metal-oxide-semiconductor, CMOS) 光电晶体管。感光元件把光信号转换成电信号, 之后将电信号传递给 ISP 转换成数字图像信号。ISP 将数字图像信号输出到 DSP 加工处理。DSP 将数字图像信号转换成标准的 RGB, YUV 等格式的图像信号。在一些实施例中, 手机 100 可以包括 1 个或 N 个摄像头, N 为大于 1 的正整数。

内部存储器 120 可以用于存储计算机可执行程序代码, 所述可执行程序代码包括指令。处理器 110 通过运行存储在内部存储器 120 的指令, 从而执行手机 100 的各种功能应用以及数据处理。存储器 120 可以包括存储程序区和存储数据区。其中, 存储程序区可存储操作系统, 至少一个功能所需的应用程序 (比如声音播放功能, 图像播放功能等) 等。存储数据区可存储手机 100 使用过程中所创建的数据 (比如音频数据, 电话本等) 等。此外, 存储器 120 可以包括高速随机存取存储器, 还可以包括非易失性存储器, 例如至少一个磁盘存储器件, 闪存器件, 其他易失性固态存储器件, 通用闪存存储器 (universal flash storage, UFS) 等。

手机 100 可以通过音频模块 170, 扬声器 170A, 受话器 170B, 麦克风 170C, 耳机接口 170D, 以及应用处理器等实现音频功能。例如音乐播放, 录音等。

音频模块 170 用于将数字音频信息转换成模拟音频信号输出, 也用于将模拟音频输入转换为数字音频信号。音频模块还可以用于对音频信号编码和解码。在一些实施例中, 音频模块 170 可以设置于处理器 110 中, 或将音频模块的部分功能模块设置于处理器 110 中。

扬声器 170A, 也称“喇叭”, 用于将音频电信号转换为声音信号。手机 100 可以通过扬声

器收听音乐，或收听免提通话。

受话器 170B，也称“听筒”，用于将音频电信号转换成声音信号。当手机 100 接听电话或语音信息时，可以通过将受话器靠近人耳接听语音。

麦克风 170C，也称“话筒”，“传声器”，用于将声音信号转换为电信号。当拨打电话或发送语音信息时，用户可以通过人嘴靠近麦克风发声，将声音信号输入到麦克风。手机 100 可以设置至少一个麦克风。在一些实施例中，手机 100 可以设置两个麦克风，除了采集声音信号，还可以实现降噪功能。在一些实施例中，手机 100 还可以设置三个，四个或更多麦克风，实现采集声音信号，降噪，还可以识别声音来源，实现定向录音功能等。

耳机接口 170D 用于连接有线耳机。耳机接口可以是 USB 接口，也可以是 3.5mm 的开放移动终端平台（open mobile terminal platform, OMTP）标准接口，美国蜂窝电信工业协会（cellular telecommunications industry association of the USA, CTIA）标准接口。

如图 5 所示，上述通信系统架构中的第一终端设备例如可以为智能家居设备 200。智能家居设备 200 中可以包括处理器 201、显示屏 202、存储模块 203、通信模块 204、射频模块 205、天线 01、天线 02、麦克风 206 以及扬声器 207 等部件。各部件的功能可以参考上文相关描述，在此不做赘述。

在一些实施例中，智能家居设备 200 的天线 01 和通信模块耦合，天线 02 和射频模块耦合。使得智能家居设备 200 可以通过无线通信技术与网络以及其他设备通信。所述无线通信技术可以包括 LTE，5G NR，WLAN 等。从而，智能家居设备 200 可以与 internet 服务器 11 以及云服务器 12（也可以称为云端）交互。

可以理解的是，上述智能家居设备 200 可以具有比图 5 中所示出的更多的或者更少的部件，可以组合两个或更多的部件，或者可以具有不同的部件配置。图 5 中所示出的各种部件可以在包括一个或多个信号处理或专用集成电路在内的硬件、软件、或硬件和软件的组合中实现。

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。其中，在本申请的描述中，除非另有说明，“至少一个”是指一个或多个，“多个”是指两个或两个以上。另外，为了便于清楚描述本申请实施例的技术方案，在本申请的实施例中，采用了“第一”、“第二”等字样对功能和作用基本相同的相同项或相似项进行区分。本领域技术人员可以理解“第一”、“第二”等字样并不对数量和执行次序进行限定，并且“第一”、“第二”等字样也并不限定一定不同。

为了便于理解，以下结合附图对本申请实施例提供的交互方法进行具体介绍。

如图 6 所示，本申请实施例提供一种交互方法，以第一终端设备为智能音箱，第二终端设备为手机为例进行说明，包括：

601、手机接收用户的第一输入内容，第一输入内容包括第一语音内容和/或第一文本内容。

用户的手机可以安装智能音箱 APP 客户端，用户可以通过手机上的 APP 客户端发送第一语音内容和/或第一文本内容给智能音箱。

举例来说，当用户希望远程操作家中的智能音箱时，如图 7 所示，用户可以在手机的桌面 701 上点击智能音箱 APP 的图标 702。当手机检测到用户点击桌面 701 上的智能音箱 APP 的图标 702 的操作后，可以启动智能音箱 APP，显示如图 8 所示的图形用户界面（graphical user interface, GUI），该 GUI 可以称为输入界面 801。该输入界面 801 可以包括语音输入图标 802 和输入语音的提示信息 803 “请输入语音”，或者语音的提示信息可以是“请说出您希望执行

的操作”等（图8中未示出），输入语音的提示信息可以是用户点击语音输入图标802后显示的；文本输入部分可以包括文本框804和输入文本的提示信息805“请输入文字”，或者该提示信息可以是“请编辑出您希望执行的操作”（图8中未示出）等。

可选的，用户可以通过手机的其他APP（例如，微信APP）中的智能音箱小程序或公众号来操作智能音箱，或者用户可以通过浏览器登录智能音箱的网页来操作智能音箱，本申请不做限定。

若第一输入内容包括第一语音内容，手机可以通过麦克风拾取用户发出的语音信号（第一语音内容）。示例性的，用户在进行语音输入时，可以点击语音输入图标，此时手机启动麦克风检测用户发出的语音信号（用户说的话），第一语音内容可以为“请打开家里的扫地机器人清扫客厅”或“请清扫客厅”。用户停止说话后预设时间间隔后，手机可以认为语音输入完毕。可选的，手机可以通过提示音或提示信息提示用户语音输入完毕。可选的，手机可以预设用户可以输入的语音长度，例如规定用户输入的语音长度大于1s小于或等于60s。可选的，若用户发出的语音内容的音量过小（例如，小于20dB），可以通过提示音或提示信息提示用户调整音量大小，以便麦克风可以更好的拾取用户的语音内容。

若第一输入内容包括第一文本内容，手机可以通过输入装置（例如触摸屏或键盘）接收用户输入的文本内容（第一文本内容）。例如，第一文本内容可以是用户自己编辑的一段心得或感悟等。或者，第一文本内容可以是用户从浏览器或微信等应用复制的一篇小短文或一首小诗。示例性的，用户在进行文本输入时，可以点击文本框803，文本框中的光标闪烁提示用户输入文本内容，可选的，文本框周围（例如文本框的下方或上方）可以显示输入法（例如拼音输入法或手写输入法）或快捷方式（复制、粘贴等）等，方便用户的操作。

若第一输入内容包括第一语音内容和第一文本内容，手机可以通过麦克风拾取用户发出的语音内容（第一语音内容），并通过输入装置接收用户输入的文本内容（第一文本内容）。例如，第一文本内容可以是用户编辑的一段文字，第一语音内容可以是“请用声情并茂/高昂/低沉的语气朗读这段文字（即第一文本内容）”。用户可以先输入第一语音内容，再输入第一文本内容，或者用户可以先输入第一文本内容再输入第一语音内容，或者用户可以同时输入第一语音内容和第一文本内容，本申请不做限定。

可选的，第一输入内容可以包括图像信息，例如网页截图、火车票、飞机票的订票信息的截图等。举例来说，用户可以通过智能音箱APP输入第一语音内容或第一文本内容：“预定图片中航次的机票”，并输入机票信息的截图。

这样，当用户不在智能音箱附近的时候，可以利用互联网环境，通过手机的麦克风拾取用户发出的语音内容，和/或，通过输入装置接收用户输入的文本内容，使手机与智能音箱实现远程交互，能够使用户远程无障碍操作智能音箱，并通过智能音箱与家中其他成员进行沟通，充分发挥了家中智能音箱的能力。

602、手机向智能音箱发送第一输入内容。

手机可以直接将用户输入的语音内容（第一语音内容）和/或用户输入的文本内容（第一文本内容）发送给智能音箱，无需对第一输入内容进行相应处理，例如进行语义提取或操作指令匹配等操作。

举例来说，当用户对手机说：“请打开家里的扫地机器人清扫客厅”，手机可以直接将“请打开家里的扫地机器人清扫客厅”这句话的音频信息发送给智能音箱，而无需根据“请打开家里的扫地机器人清扫客厅”这句话的语义匹配相应的操作指令。

603、智能音箱从手机接收第一输入内容。

604、智能音箱确定第一输入内容对应的控制指令。

智能音箱可以将第一输入内容上传服务器（例如，云服务器）。若第一输入内容为第一语音内容，云服务器的 ASR 引擎可以将第一语音内容转化为第一文本信息，云服务器的 NLP 引擎可以根据第一文本信息的语义得到控制指令。该控制指令可以包括第一文本信息的垂类和槽位。示例性的，若用户说“请帮我订一张后天早上去北京的机票”或者“我想订后天早上的机票去北京”，NLP 引擎可以根据关键字做意图识别确定第一文本信息的垂类为“订机票”，并确定该垂类的槽位包括：出发时间为“后天早上”，目的地为“北京”。或者，若用户说“讲一个小笑话”，NLP 引擎可以根据关键字做意图识别确定垂类为“讲故事”。并确定该垂类的槽位包括：故事类型为“笑话”，故事长短为“短”。云服务器向智能音箱下发控制指令，智能音箱接收云服务器下发的控制指令。

若第一输入内容为文本内容，云服务器的 NLP 引擎可以根据文本信息的语义得到控制指令，而后云服务器向智能音箱下发该控制指令，智能音箱接收云服务器下发的控制指令。

605、智能音箱处理第一输入内容对应的控制指令。

第一输入内容对应的控制指令可以是针对智能音箱的，即需要智能音箱根据控制指令做出响应；第一输入内容对应的控制指令也可以是针对第三终端设备（例如，家用的扫地机器人），即智能音箱转发控制指令给第三终端设备，第三终端设备根据控制指令做出响应。

若控制指令是针对智能音箱的，智能音箱执行第一输入内容对应的控制指令。例如，若控制指令为播放（PLAY），智能音箱播放音乐。

若控制指令是针对第三终端设备的，智能音箱可以向第三终端设备发送第一输入内容对应的控制指令，以便第三终端设备执行第一输入内容对应的控制指令，从而控制第三终端设备进行相应操作，例如开机、关机等等。其中，第三终端设备可以是与智能音箱匹配的终端设备，例如，可以是与智能音箱匹配的扫地机器人，空调、冰箱、洗衣机、智能窗帘等等。

举例来说，当用户需要远程操作家中的扫地机器人而手机没有与家中的扫地机器人匹配时，手机没有控制扫地机器人指令，无法直接控制扫地机器人，此时手机可以通过智能音箱控制扫地机器人。例如，用户可以通过手机的智能语音 APP 客户端输入如下语音内容：“请打开家里的扫地机器人清扫客厅”。智能音箱接收到这条语音内容后，可以将该语音内容上传云服务器，云服务器的 ASR 引擎可以将该语音内容转化为文本信息，云服务器的 NLP 引擎可以根据该文本信息的语义得到控制指令“清扫客厅”，云服务器向智能音箱下发控制指令，智能音箱接收云服务器下发的控制指令，并转发给扫地机器人，从而实现对扫地机器人的控制，充分发挥了智能音箱的能力，提高了用户体验。

606、智能音箱向手机发送第一输入内容对应的第一响应消息。

若控制指令是针对智能音箱的，智能音箱执行第一输入内容对应的控制指令后，向手机发送第一响应消息，第一响应消息可以包括云服务转写的控制指令以及智能音箱执行控制指令后的执行结果，以使用户通过第二终端设备输入第一输入内容后，可以获知根据第一输入内容得到的控制指令和针对该控制指令的执行结果，提高了用户体验。举例来说，用户 A 可以通过手机的 APP 端发送“播放荷塘月色”的语音内容，智能音箱接收到该语音内容后，通过云服务器确定该语音内容对应的控制指令，根据该控制指令播放相应歌曲，并将播放状态（即第一输入内容对应的第一响应消息）发送到手机 APP 端。

若控制指令是针对第三终端设备的，智能音箱从第三终端设备接收第一输入内容对应的执行结果后，可以向手机发送第一响应消息，第一响应消息可以包括云服务转写的控制指令以及第三终端设备执行控制指令后的执行结果。

607、手机从智能音箱接收第一输入内容对应的第一响应消息。

608、手机语音播报或在显示屏显示响应消息。

即手机可以语音播报和/或在显示屏上显示智能音箱通过云服务器转写后的控制指令和执行结果（即响应消息）。

如图 9 所示，假设用户 805 语音输入：“吃晚饭的时候播放半小时轻音乐，晚饭 18: 50 开始”。手机可以在 APP 上显示用户 805 输入了语音内容 806，并显示智能音箱 807 处理后的控制指令 808 为“18: 50 开始播放轻音乐 30 分钟”，以及执行结果 809 为“18: 50 开始播放轻音乐”和执行结果 810 “19: 20 结束播放轻音乐”。可选的，手机可以在 APP 上显示用户 805 输入语音内容的时间信息 811，以及手机接收到智能音箱发送的控制指令 808 的时间信息 812、手机接收到执行结果 809 的时间信息 813 和手机接收到执行结果 810 的时间信息 814。进一步的，手机还可以将控制指令 808、执行结果 809 和执行结果 810 语音播报给用户。或者，手机可以不显示控制指令 808、执行结果 809 和执行结果 810，仅通过语音播报给用户。例如，当手机检测到用户正在听音乐且手机屏幕为黑屏状态时，手机语音播报智能音箱反馈的响应消息，无需点亮屏幕，可以节省耗电。

另外，该交互方法还可以包括：

609、智能音箱接收第二输入内容，所述第二输入内容包括第二语音内容和/或第二文本内容。

第二输入内容可以是其他用户（与输入第一输入内容的用户不同）在音箱旁输入的语音内容（第二语音内容），或者第二输入内容可以是其他用户通过第四终端设备（例如用户的手机或智能穿戴设备等）发送的第二语音内容和/或第二文本内容。

若智能音箱接收第二输入内容的时刻早于接收第一输入内容的时刻，智能音箱处理第二输入内容对应的控制指令后，处理第一输入内容对应的控制指令；若智能音箱接收第二输入内容的时刻晚于接收第一输入内容的时刻，智能音箱处理第一输入内容对应的控制指令后，处理第二输入内容对应的控制指令。也就是说，智能音箱在处理不同的控制指令时，遵循先来先执行策略。

以下结合具体场景对智能音箱处理不同的控制指令的先后顺序进行说明：

场景 1、用户 A 通过手机的 APP 端发送“播放音乐”的语音内容，智能音箱接收到语音内容后，通过云服务器确定该语音内容对应的指令，根据该指令播放音乐，并可以将播放状态发送到手机 APP 端。在播放音乐的过程中，用户 B 在音箱旁通过语音指令控制暂停播放，预期结果为：智能音箱暂停音乐播放，并可以将暂停状态发送到手机 APP 端。

场景 2、智能音箱正在进行音乐播放，用户 B 在音箱旁通过语音指令控制暂停，而后用户 A 通过手机的 APP 端发送“播放音乐”的语音内容，智能音箱接收到语音内容后，通过云服务器确定该语音内容对应的指令，预期结果为：智能音箱暂停音乐播放后又继续播放音乐，并可以将播放状态发送到手机 APP 端。

场景 3、用户 A 通过手机的 APP 端发送“播放音乐”的语音内容，智能音箱接收到语音内容后，通过云服务器确定该语音内容对应的指令，根据该指令播放音乐，在播放音乐的过程中（例如，播放了一首歌的 1/3 后），用户 B 在音箱旁通过语音控制音乐播放（用户 B 请求播放的内容与用户 A 请求播放的内容正好一致，例如都要求播放“荷塘月色”），预期结果为：智能音箱继续播放“荷塘月色”，无需重头开始播放“荷塘月色”，能够满足用户需求且避免重复执行相同的指令。

需要说明的是，步骤 609 和步骤 603 之间没有必然的执行先后顺序，步骤 609 可以在

步骤 603 之前执行,也可以在步骤 603 之后执行,也可以和步骤 603 同时执行,本实施例对此不作具体限定。

可以理解的是,智能音箱处理第二输入内容对应的控制指令后,可以向手机发送第二输入内容对应的第二响应消息。手机从智能音箱接收第二响应消息,第二响应消息对应第二输入内容,手机语音播报或在显示屏显示第二响应消息。这样一来,若其他用户(与输入第一输入内容的用户不同)在音箱旁输入语音内容(第二语音内容),或者其他用户通过第四终端设备(例如用户的手机或智能穿戴设备等)输入第二输入内容,第一终端设备的用户可以获知其他用户对第二终端设备的操作(包括第二输入内容对应的控制指令和针对该控制指令的执行结果),提高了用户对第二终端设备的掌控能力,从而提高了用户体验。

基于本申请实施例提供的方法,第二终端设备(例如,手机)可以从用户接收第一输入内容,并向第一终端设备(例如,智能语音交互设备)发送第一输入内容。第一终端设备从第二终端设备接收第一输入内容,通过服务器确定第一输入内容对应的控制指令,并处理该控制指令。这样一来,可以由第一终端设备基于用户通过第二终端设备远程输入的内容(即第一输入内容)确定控制指令,而非用户根据手机 APP 上预制好的有限的指令进行操作,增强了用户与第一终端设备远程交互的灵活性,充分发挥第一终端设备的能力。

上述本申请提供的实施例中,分别从第一终端设备、第二终端设备以及第一终端设备和第二终端设备之间交互的角度对本申请实施例提供的方法进行了介绍。为了实现上述本申请实施例提供的方法中的各功能,第一终端设备和第二终端设备可以包括硬件结构和/或软件模块,以硬件结构、软件模块、或硬件结构加软件模块的形式来实现上述各功能。上述各功能中的某个功能以硬件结构、软件模块、还是硬件结构加软件模块的方式来执行,取决于技术方案的特定应用和设计约束条件。

在采用对应各个功能划分各个功能模块的情况下,图 10 示出了上述实施例中所涉及的装置 10 的一种可能的结构示意图,该装置可以为第一终端设备,该第一终端设备包括:接收单元 1001、确定单元 1002 和处理单元 1002。在本申请实施例中,接收单元 1001,用于从第二终端设备接收第一输入内容,第一输入内容包括第一语音内容和/或第一文本内容;确定单元 1002,用于确定第一输入内容对应的控制指令;处理单元 1003,用于处理第一输入内容对应的控制指令。可选的,第一终端设备还可以包括发送单元 1004(图 10 中未示出),用于向第二终端设备发送第一输入内容对应的第一响应消息。

在图 6 所示的方法实施例中,接收单元 1001 用于支持第一终端设备执行图 6 中的过程 603;确定单元 1002 用于支持第一终端设备执行图 6 中的过程 604;处理单元 1003 用于支持第一终端设备执行图 6 中的过程 605;发送单元 1004,用于支持第一终端设备执行图 6 中的过程 606。其中,上述方法实施例涉及的各步骤的所有相关内容均可以援引到对应功能模块的功能描述,在此不再赘述。

在采用对应各个功能划分各个功能模块的情况下,图 11 示出了上述实施例中所涉及的装置 11 的一种可能的结构示意图,该装置可以为第二终端设备,该第二终端设备包括:接收单元 1101、发送单元 1102 和处理单元 1103。在本申请实施例中,接收单元 1101,用于从用户接收第一输入内容,第一输入内容包括第一语音内容和/或第一文本内容;发送单元 1102,用于向第一终端设备发送第一输入内容;接收单元 1101,还用于从第一终端设备接收第一输入内容对应的第一响应消息;处理单元 1103,用于语音播报或在显示屏显示第一响应消息。

在图 6 所示的方法实施例中,接收单元 1101 用于支持第二终端设备执行图 6 中的过程 601 和 607;发送单元 1102 用于支持第二终端设备执行图 6 中的过程 602;处理单元 1103 用

于支持第二终端设备执行图 6 中的过程 608。其中，上述方法实施例涉及的各步骤的所有相关内容均可以援引到对应功能模块的功能描述，在此不再赘述。

本申请实施例中对模块的划分是示意性的，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，另外，在本申请各个实施例中的各功能模块可以集成在一个处理器中，也可以是单独物理存在，也可以两个或两个以上模块集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。示例性地，在本申请实施例中，接收单元和发送单元可以集成至收发单元中。

本申请实施例提供的方法中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本发明实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、网络设备、用户设备或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（digital subscriber line, DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机可以存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质（例如，软盘、硬盘、磁带）、光介质（例如，数字视频光盘（digital video disc, DVD））、或者半导体介质（例如，固态硬盘（solid state drives, SSD））等。

显然，本领域的技术人员可以对本申请实施例进行各种改动和变型而不脱离本申请的精神和范围。这样，倘若本申请实施例的这些修改和变型属于本申请权利要求及其等同技术的范围之内，则本申请也意图包含这些改动和变型在内。

权 利 要 求 书

1、一种交互方法，其特征在于，包括：

第一终端设备从第二终端设备接收第一输入内容，所述第一输入内容包括第一语音内容和/或第一文本内容；

所述第一终端设备确定所述第一输入内容对应的控制指令；

所述第一终端设备处理所述第一输入内容对应的所述控制指令。

2、根据权利要求1所述的交互方法，其特征在于，所述第一终端设备确定所述第一输入内容对应的控制指令包括：

所述第一终端设备向服务器发送所述第一输入内容；

所述第一终端设备从所述服务器接收所述第一输入内容对应的所述控制指令。

3、根据权利要求1或2所述的交互方法，其特征在于，所述第一终端设备处理所述第一输入内容对应的所述控制指令包括：

所述第一终端设备执行所述第一输入内容对应的所述控制指令；或者

所述第一终端设备向第三终端设备发送所述第一输入内容对应的所述控制指令。

4、根据权利要求1-3任一项所述的交互方法，其特征在于，所述方法还包括：

所述第一终端设备向所述第二终端设备发送所述第一输入内容对应的第一响应消息。

5、根据权利要求1-4任一项所述的交互方法，其特征在于，所述第一终端设备处理所述第一输入内容对应的所述控制指令包括：

所述第一终端设备接收第二输入内容，所述第二输入内容包括第二语音内容和/或第二文本内容；

若所述第一终端设备接收所述第二输入内容的时刻早于接收所述第一输入内容的时刻，所述第一终端设备处理所述第二输入内容对应的控制指令后，处理所述第一输入内容对应的控制指令。

6、一种交互方法，其特征在于，包括：

第二终端设备从用户接收第一输入内容，所述第一输入内容包括第一语音内容和/或第一文本内容；

所述第二终端设备向第一终端设备发送所述第一输入内容；

所述第二终端设备从所述第一终端设备接收所述第一输入内容对应的第一响应消息；

所述第二终端设备语音播报或在显示屏显示所述第一响应消息。

7、根据权利要求6所述的交互方法，其特征在于，所述方法还包括：

所述第二终端设备从所述第一终端设备接收第二响应消息，所述第二响应消息对应第二输入内容；

所述第二终端设备语音播报或在显示屏显示所述第二响应消息。

8、一种第一终端设备，其特征在于，包括：

接收单元，用于从第二终端设备接收第一输入内容，所述第一输入内容包括第一语音内容和/或第一文本内容；

确定单元，用于确定所述第一输入内容对应的控制指令；

处理单元，用于处理所述第一输入内容对应的所述控制指令。

9、根据权利要求8所述的第一终端设备，其特征在于，所述确定单元用于：

通过发送单元向服务器发送所述第一输入内容；

通过所述接收单元从所述服务器接收所述第一输入内容对应的所述控制指令。

10、根据权利要求 8 或 9 所述的第一终端设备，其特征在于，所述处理单元用于：
执行所述第一输入内容对应的所述控制指令；或者
通过所述发送单元向第三终端设备发送所述第一输入内容对应的所述控制指令。

11、根据权利要求 8-10 任一项所述的第一终端设备，其特征在于，所述发送单元还用于：
向所述第二终端设备发送所述第一输入内容对应的第一响应消息。

12、根据权利要求 8-11 任一项所述的第一终端设备，其特征在于，所述处理单元用于：
通过所述接收单元接收第二输入内容，所述第二输入内容包括第二语音内容和/或第二文本内容；

若所述第一终端设备接收所述第二输入内容的时刻早于接收所述第一输入内容的时刻，
处理所述第二输入内容对应的控制指令后，处理所述第一输入内容对应的控制指令。

13、一种第二终端设备，其特征在于，包括：

接收单元，用于从用户接收第一输入内容，所述第一输入内容包括第一语音内容和/或第一文本内容；

发送单元，用于向第一终端设备发送所述第一输入内容；

所述接收单元，还用于从所述第一终端设备接收所述第一输入内容对应的第一响应消息；
处理单元，用于语音播报或在显示屏显示所述第一响应消息。

14、根据权利要求 13 所述的第二终端设备，其特征在于，所述接收单元还用于：
从所述第一终端设备接收第二响应消息，所述第二响应消息对应第二输入内容；
所述处理单元，还用于语音播报或在显示屏显示所述第二响应消息。

15、一种第一终端设备，其特征在于，所述第一终端设备包括处理器和存储器；

所述存储器用于存储计算机执行指令，当所述第一终端设备运行时，所述处理器执行所述存储器存储的所述计算机执行指令，以使所述第一终端设备执行如权利要求 1-5 中任一项所述的交互方法。

16、一种第二终端设备，其特征在于，所述第二终端设备包括处理器和存储器；

所述存储器用于存储计算机执行指令，当所述第二终端设备运行时，所述处理器执行所述存储器存储的所述计算机执行指令，以使所述第二终端设备执行如权利要求 6 或 7 所述的交互方法。

17、一种计算机可读存储介质，其特征在于，包括指令，当其在计算机上运行时，使得计算机执行权利要求 1-5 中任一项所述的交互方法。

18、一种计算机可读存储介质，其特征在于，包括指令，当其在计算机上运行时，使得计算机执行权利要求 6 或 7 所述的交互方法。

19、一种芯片系统，其特征在于，包括处理器和存储器，所述处理器执行所述存储器存储的计算机执行指令，以实现如权利要求 1-5 中任一项所述的交互方法。

20、一种芯片系统，其特征在于，包括处理器和存储器，所述处理器执行所述存储器存储的计算机执行指令，以实现如权利要求 6 或 7 所述的交互方法。

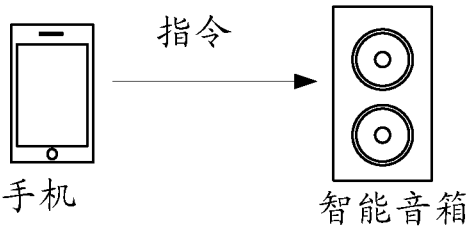


图 1

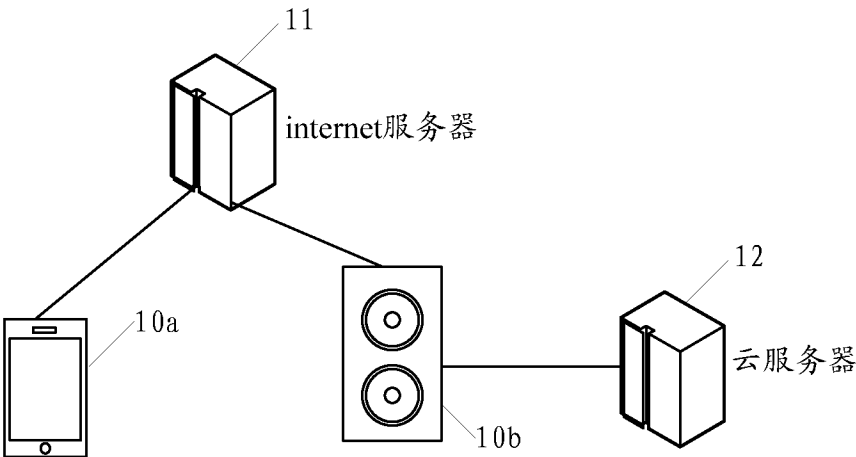


图 2

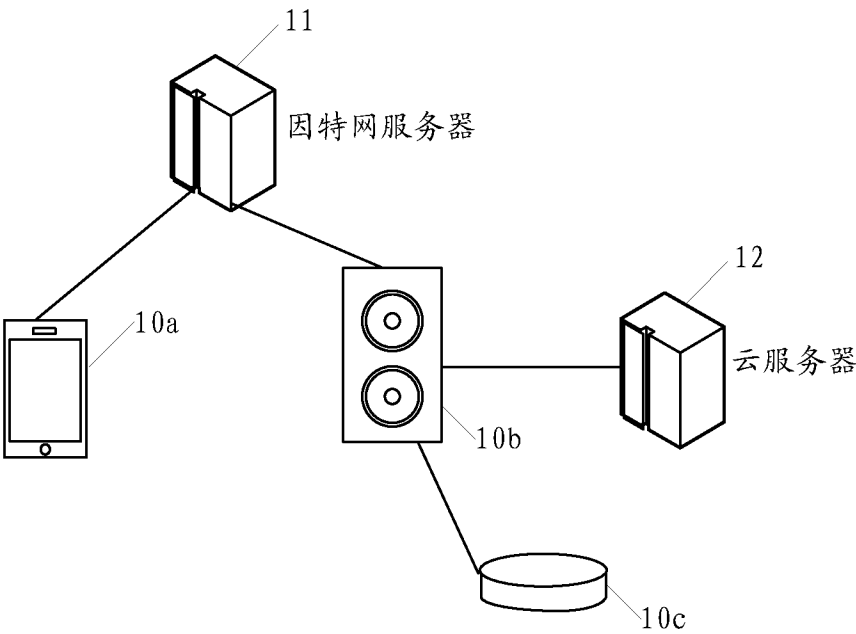


图 3

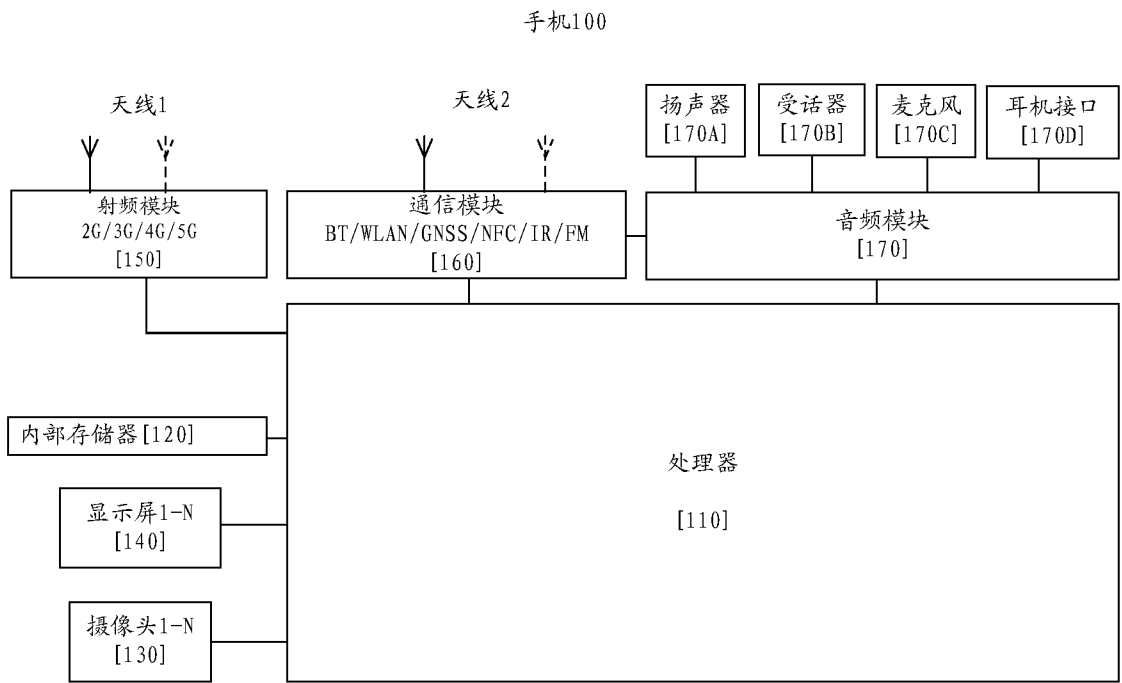


图 4

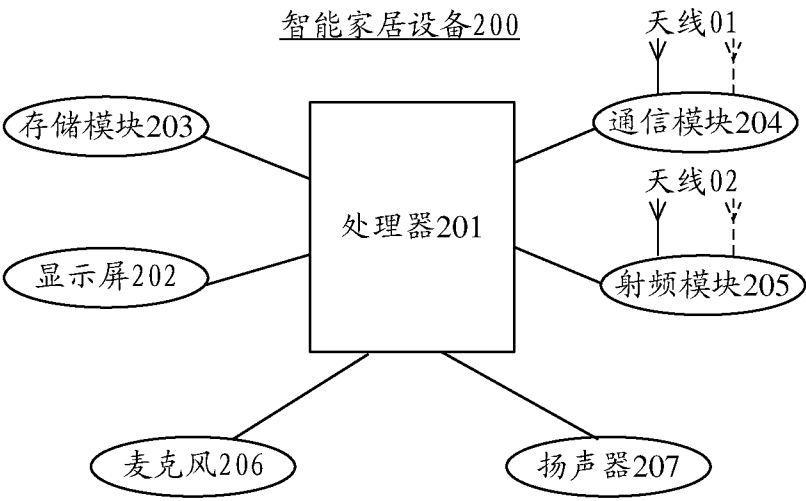


图 5

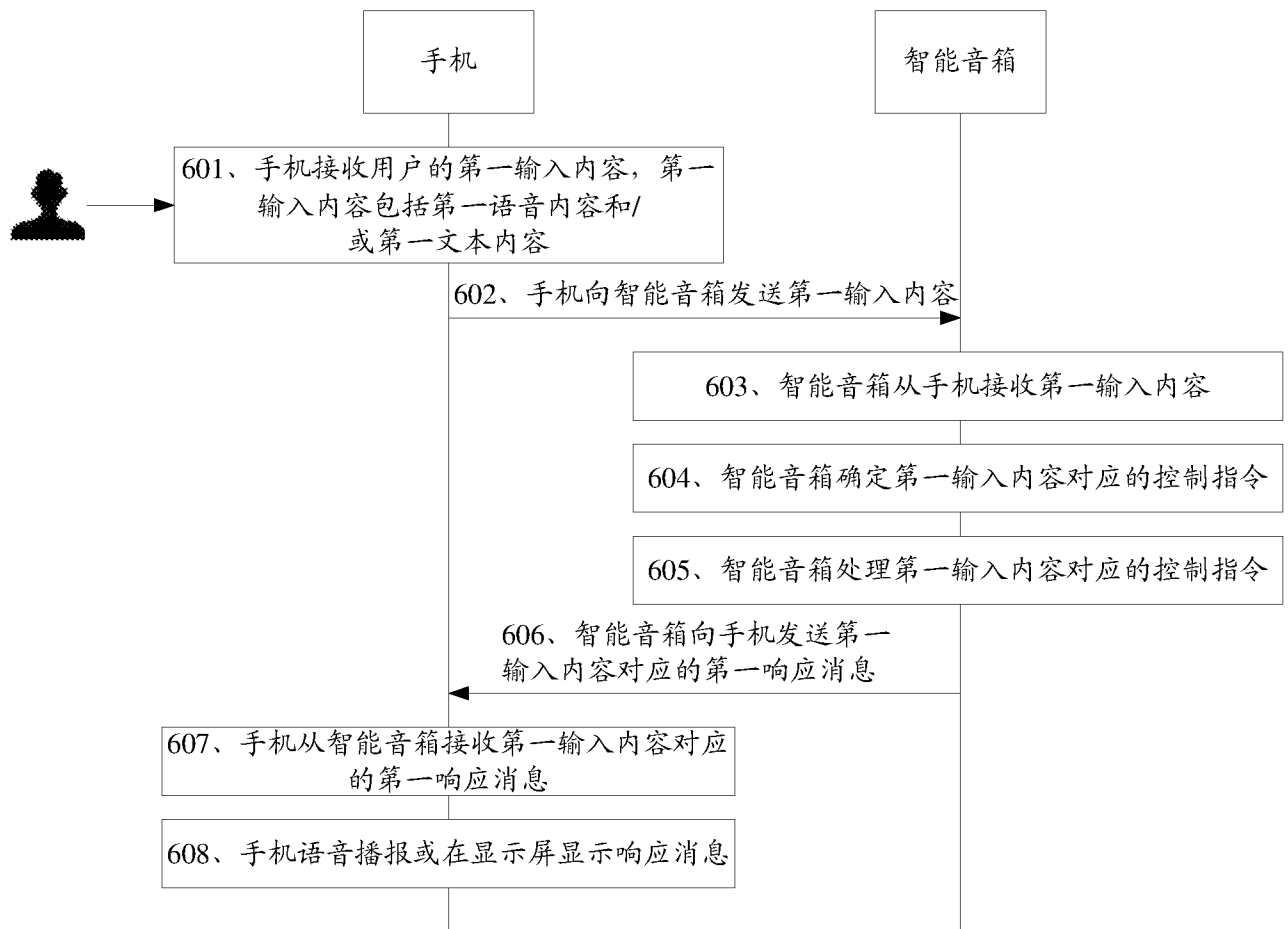


图 6



图 7

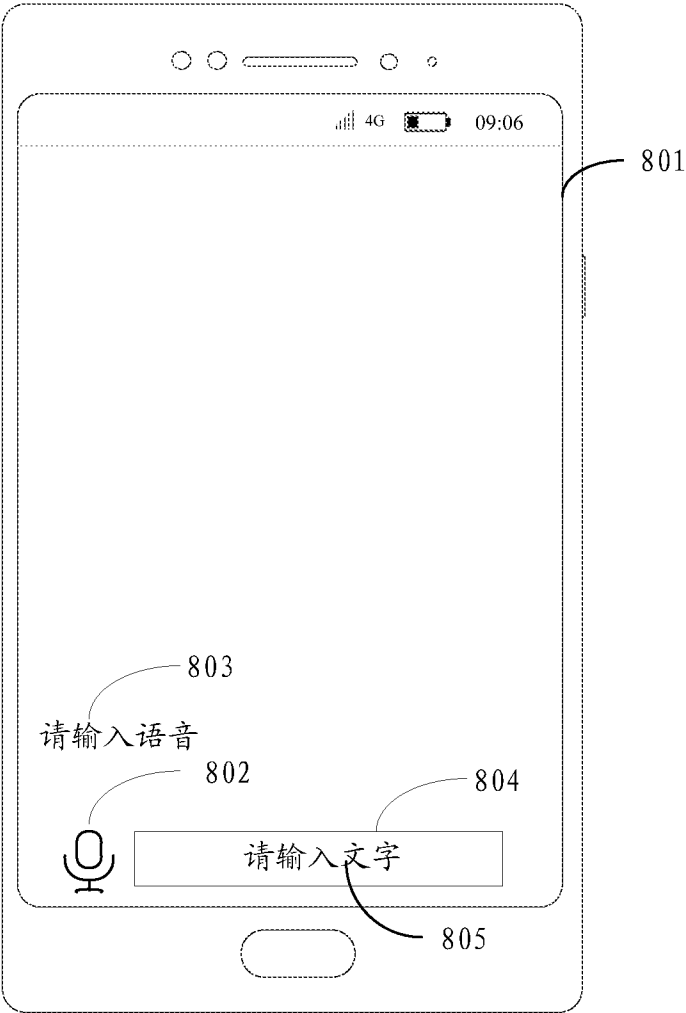


图 8

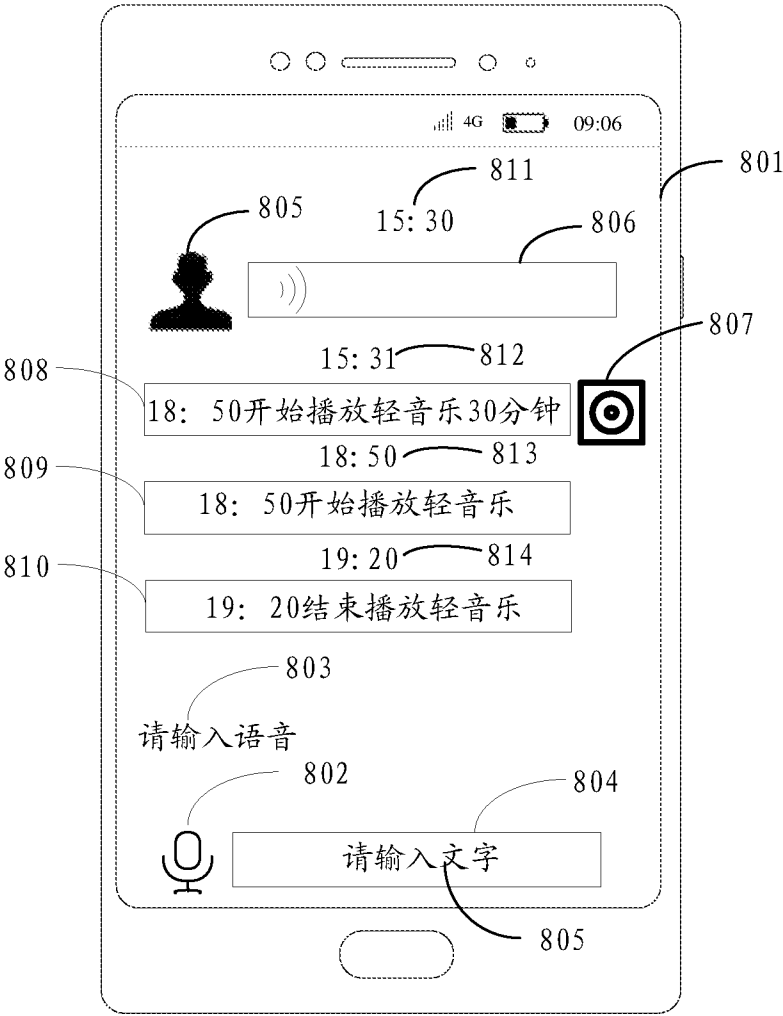


图 9

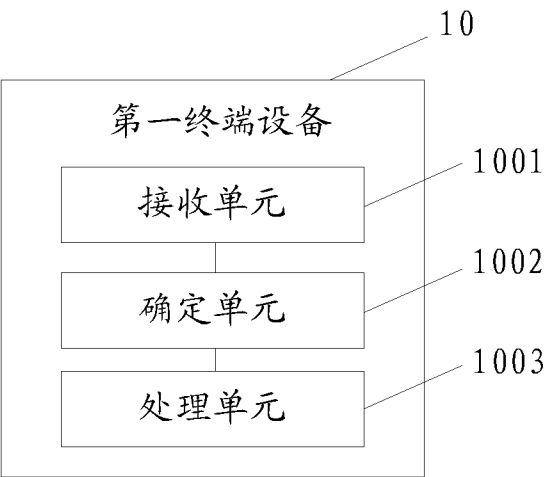


图 10

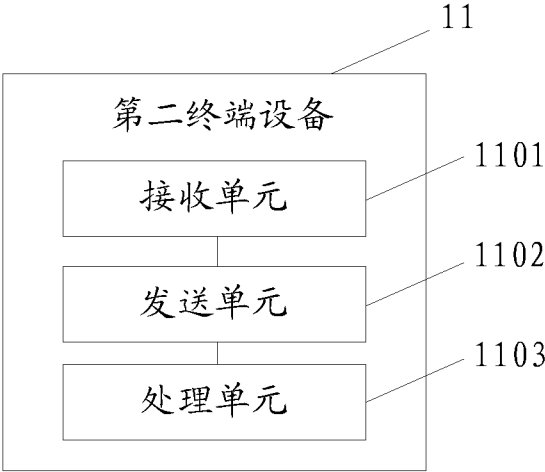


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/092888

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/22(2006.01)i; G10L 15/26(2006.01)i; G10L 17/22(2013.01)i; H04M 1/725(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; H04M

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; 万方; 超星; VEN; USTXT; EPTXT; WOTXT; IEEE: 文本, 终端, 家电, 智能, 云端, 音箱, 控制, 语音, 命令, 指令, 识别, 辨识, voice, speech, terminal, mobile, control+, telephone, command

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101272418 A (YULONG COMPUTER TELECOMMUNICATION SCIENTIFIC (SHENZHEN) CO., LTD.) 24 September 2008 (2008-09-24) description page 6 line 6 to page 12 bottom line, figures 1-3	1-20
X	CN 105991825 A (ZTE CORPORATION) 05 October 2016 (2016-10-05) description paragraphs [0074]-[0143], figures 1-7	1-20
X	CN 102427418 A (FUZHOU HAIJING SCIENCE & TECH DEVELOPMENT CO., LTD.) 25 April 2012 (2012-04-25) description, paragraphs [0019]-[0022], and figures 1 and 2	1-20
X	CN 102170617 A (ZTE CORPORATION) 31 August 2011 (2011-08-31) description, paragraphs [0022]-[0070], and figures 1-4	1-20
X	CN 1655206 A (DENG, Qinghui) 17 August 2005 (2005-08-17) description page 4 line 8- page 7 bottom line, figures 1-3	1-20
A	JP 2014164241 A (NIPPON TELEGRAPH & TELEPHONE) 08 September 2014 (2014-09-08) 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

20 July 2020

Date of mailing of the international search report

20 August 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/CN2020/092888

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	101272418	A	24 September 2008	CN	101272418	B	18 July 2012
CN	105991825	A	05 October 2016	WO	2016124008	A1	11 August 2016
CN	102427418	A	25 April 2012	None			
CN	102170617	A	31 August 2011	WO	2012136024	A1	11 October 2012
CN	1655206	A	17 August 2005	None			
JP	2014164241	A	08 September 2014	JP	6028289	B2	16 November 2016

国际检索报告

国际申请号

PCT/CN2020/092888

A. 主题的分类 G10L 15/22 (2006.01) i; G10L 15/26 (2006.01) i; G10L 17/22 (2013.01) i; H04M 1/725 (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; H04M 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; 万方; 超星; VEN; USTXT; EPTXT; WOTXT; IEEE: 文本, 终端, 家电, 智能, 云端, 音箱, 控制, 语音, 命令, 指令, 识别, 辨识, voice, speech, terminal, mobile, control+, telephone, command		
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 101272418 A (宇龙计算机通信科技深圳有限公司) 2008年 9月 24日 (2008 - 09 - 24) 说明书第6页第6行-第12页最后一行, 附图1-3	1-20
X	CN 105991825 A (中兴通讯股份有限公司) 2016年 10月 5日 (2016 - 10 - 05) 说明书第[0074]-[0143]段, 附图1-7	1-20
X	CN 102427418 A (福州海景科技开发有限公司) 2012年 4月 25日 (2012 - 04 - 25) 说明书第[0019]-[0022]段, 附图1、2	1-20
X	CN 102170617 A (中兴通讯股份有限公司) 2011年 8月 31日 (2011 - 08 - 31) 说明书第[0022]-[0070]段, 附图1-4	1-20
X	CN 1655206 A (邓清辉) 2005年 8月 17日 (2005 - 08 - 17) 说明书第4页第8行-第7页最后一行, 附图1-3	1-20
A	JP 2014164241 A (NIPPON TELEGRAPH & TELEPHONE) 2014年 9月 8日 (2014 - 09 - 08) 全文	1-20
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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 7月 20日		国际检索报告邮寄日期 2020年 8月 20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 周群 电话号码 86- (0512) -88997268

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/092888

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	101272418	A	2008年 9月 24日	CN	101272418	B	2012年 7月 18日
CN	105991825	A	2016年 10月 5日	WO	2016124008	A1	2016年 8月 11日
CN	102427418	A	2012年 4月 25日	无			
CN	102170617	A	2011年 8月 31日	WO	2012136024	A1	2012年 10月 11日
CN	1655206	A	2005年 8月 17日	无			
JP	2014164241	A	2014年 9月 8日	JP	6028289	B2	2016年 11月 16日