

(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
2021年11月25日 (25.11.2021)



(10) 国际公布号
WO 2021/232913 A1

- (51) 国际专利分类号:
G06F 9/4401 (2018.01) *G10L 15/22* (2006.01)
- (21) 国际申请号: PCT/CN2021/081332
- (22) 国际申请日: 2021年3月17日 (17.03.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010419583.7 2020年5月18日 (18.05.2020) CN
- (71) 申请人: **OPPO 广东移动通信有限公司 (GUANGDONG OPPO MOBILE TELECOMMUNICATIONS CORP., LTD.)** [CN/CN]; 中国广东省东莞市长安镇乌沙海滨路18号, Guangdong 523860 (CN)。
- (72) 发明人: 蒋燚 (JIANG, Yi); 中国广东省东莞市长安镇乌沙海滨路18号由OPPO广东移动通信有限公司转交, Guangdong 523860 (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, 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, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** VOICE INFORMATION PROCESSING METHOD AND APPARATUS, AND STORAGE MEDIUM AND ELECTRONIC DEVICE

(54) 发明名称: 语音信息处理方法及装置、存储介质和电子设备

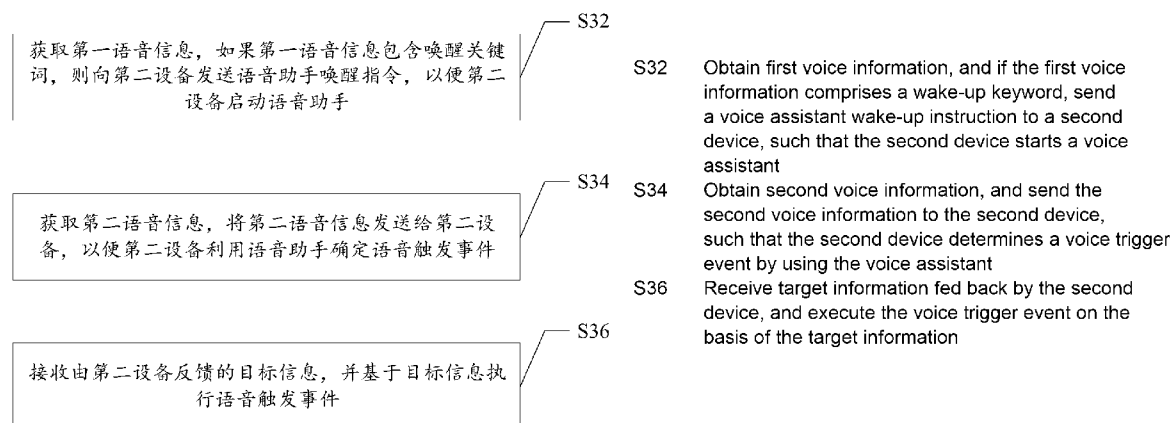


图 3

(57) **Abstract:** A voice information processing method, a voice information processing apparatus, a computer readable storage medium, and an electronic device, which relate to the technical field of voice processing. The voice information processing method comprises: a first device (1100) obtains first voice information, and in the case that the first voice information comprises a wake-up keyword, the first device (1100) sends a voice assistant wake-up instruction to a second device (1200), so that the second device (1200) starts a voice assistant; and then, the first device (1100) obtains second voice information and sends the second voice information to the second device (1200), and the second device (1200) determines a voice trigger event corresponding to the second voice information by using the voice assistant and feeds back target information associated with the execution of the voice trigger event to the first device (1100), so that the first device (1100) executes the voice trigger event on the basis of the target information. The method can reduce the operation pressure of the first device (1100).



WO 2021/232913 A1

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))。

(57) 摘要: 一种语音信息处理方法、语音信息处理装置、计算机可读存储介质和电子设备, 涉及语音处理技术领域。该语音信息处理方法包括: 第一设备(1100)获取第一语音信息, 在第一语音信息包含唤醒关键词的情况下, 第一设备(1100)向第二设备(1200)发送语音助手唤醒指令, 以便第二设备(1200)启动语音助手。之后, 第一设备(1100)获取第二语音信息, 并将第二语音信息发送给第二设备(1200), 第二设备(1200)利用语音助手确定与第二语音信息对应的语音触发事件, 并将与执行语音触发事件相关联的目标信息反馈给第一设备(1100), 以便第一设备(1100)基于该目标信息执行语音触发事件。上述方法可以减轻第一设备(1100)的运算压力。

语音信息处理方法及装置、存储介质和电子设备

相关申请的交叉引用

本申请要求于 2020 年 05 月 18 日提交的申请号为 202010419583.7、名称为“语音信息处理方法及装置、存储介质和电子设备”的中国专利申请的优先权，该中国专利申请的全部内容通过引用全部并入本文。

技术领域

本公开涉及语音处理技术领域，具体而言，涉及一种语音信息处理方法、语音信息处理装置、计算机可读存储介质和电子设备。

背景技术

智能眼镜作为一种可穿戴设备，将计算机技术与传统眼镜相结合，以实现丰富的功能。这些功能包括通过语音实现人机交互，语音处理过程的引入大大提高了智能眼镜的使用便捷性。

目前，语音处理过程依赖于智能眼镜配备的处理器。然而，受限于智能眼镜的运算能力，智能眼镜在进行语音处理时，常会遇到识别不佳、交互效果差的问题。如果智能眼镜配备高性能的处理器，则会增加智能眼镜的造价，普通用户无法接受。

发明内容

根据本公开的第一方面，提供了一种语音信息处理方法，应用于第一设备，该语音信息处理方法包括：获取第一语音信息，如果第一语音信息包含唤醒关键词，则向第二设备发送语音助手唤醒指令，以便第二设备启动语音助手；获取第二语音信息，将第二语音信息发送给第二设备，以便第二设备利用语音助手确定语音触发事件，语音触发事件与第二语音信息对应；接收由第二设备反馈的目标信息，并基于目标信息执行语音触发事件。

根据本公开的第二方面，提供了一种语音信息处理方法，应用于第二设备，该语音信息处理方法包括：响应语音助手唤醒指令，启动语音助手；其中，语音助手唤醒指令由第一设备在确定出第一语音信息包含唤醒关键词的情况下发送给第二设备；获取第一设备发送的第二语音信息，利用语音助手确定与第二语音信息对应的语音触发事件；将与执行语音触发事件相关联的目标信息反馈给第一设备，以便第一设备基于目标信息执行语音触发事件。

根据本公开的第三方面，提供了一种语音信息处理装置，应用于第一设备，包括：唤醒触发模块，被配置为获取第一语音信息，如果第一语音信息包含唤醒关键词，则向第二设备发送语音助手唤醒指令，以便第二设备启动语音助手；语音发送模块，被配置为获取第二语音信息，将第二语音信息发送给第二设备，以便第二设备利用语音助手确定语音触发事件，语音触发事件与第二语音信息对应；事件执行模块，被配置为接收由第二设备反馈的目标信息，并基于目标信息执行语音触发事件。

根据本公开的第四方面，提供了一种语音信息处理装置，应用于第二设备，包括：语音助手启动模块，被配置为响应语音助手唤醒指令，启动语音助手；其中，语音助手唤醒指令由第一设备在确定出第一语音信息包含唤醒关键词的情况下发送给第二设备；事件确定模块，被配置为获取第一设备发送的第二语音信息，利用语音助手确定与第二语音信息对应的语音触发事件；信息反馈模块，被配置为将与执行语音触发事件相关联的目标信息反馈给第一设备，以便第一设备基于目标信息执行语音触发事件。

根据本公开的第五方面，提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现上述的语音信息处理方法。

根据本公开的第六方面，提供了一种电子设备，包括处理器；存储器，被配置为存储

一个或多个程序，当一个或多个程序被处理器执行时，使得所述处理器实现上述的语音信息处理方法。

附图说明

5 图 1 示出了应用本公开实施例的语音信息处理方案的示例性系统架构的示意图；

图 2 示出了适于用来实现本公开实施例的电子设备的结构示意图；

图 3 示意性示出了根据本公开的示例性实施方式的应用于第一设备的语音信息处理方法的流程图；

10 图 4 示意性示出了根据本公开的示例性实施方式的应用于第二设备的语音信息处理方法的流程图；

图 5 示意性示出了根据本公开的示例性实施方式的实现语音信息处理过程的设备交互图；

图 6 示意性示出了本公开实施例的语音信息处理方案的整体软硬件架构图；

15 图 7 示意性示出了根据本公开的示例性实施方式的应用于第一设备的语音信息处理装置的方框图；

图 8 示意性示出了根据本公开的另一示例性实施方式的应用于第一设备的语音信息处理装置的方框图；

图 9 示意性示出了根据本公开的又一示例性实施方式的应用于第一设备的语音信息处理装置的方框图；

20 图 10 示意性示出了根据本公开的示例性实施方式的应用于第二设备的语音信息处理装置的方框图；

图 11 示意性示出了根据本公开的另一示例性实施方式的应用于第二设备的语音信息处理装置的方框图。

25 具体实施方式

现在将参考附图更全面地描述示例实施方式。然而，示例实施方式能够以多种形式实施，且不应被理解为限于在此阐述的范例；相反，提供这些实施方式使得本公开将更加全面和完整，并将示例实施方式的构思全面地传达给本领域的技术人员。所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多实施方式中。在下面的描述中，提供许多具体细节从而给出对本公开的实施方式的充分理解。然而，本领域技术人员将意识到，可以实践本公开的技术方案而省略所述特定细节中的一个或更多，或者可以采用其它的方法、组元、装置、步骤等。在其它情况下，不详细示出或描述公知技术方案以避免喧宾夺主而使得本公开的各方面变得模糊。

35 此外，附图仅为本公开的示意性图解，并非一定是按比例绘制。图中相同的附图标记表示相同或类似的部分，因而将省略对它们的重复描述。附图中所示的一些方框图是功能实体，不一定必须与物理或逻辑上独立的实体相对应。可以采用软件形式来实现这些功能实体，或在一个或多个硬件模块或集成电路中实现这些功能实体，或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

40 附图中所示的流程图仅是示例性说明，不是必须包括所有的步骤。例如，有的步骤还可以分解，而有的步骤可以合并或部分合并，因此实际执行的顺序有可能根据实际情况改变。另外，下面所有的术语“第一”、“第二”仅是为了区分的目的，不应作为本公开内容的限制。

图 1 示出了应用本公开实施例的语音信息处理方案的示例性系统架构的示意图。

45 如图 1 所示，系统架构可以包括第一设备 1100 和第二设备 1200。其中，第一设备 1100 可以是接收语音信息并执行语音信息对应事件的设备，具体的，第一设备 1100 可以是 AR

(Augmented Reality, 增强现实) 眼镜、VR (Virtual Reality, 虚拟现实) 眼镜或 MR (Mixed Reality, 混合现实) 眼镜, 然而, 第一设备 1100 还可以是例如智能头盔等具有显示功能的其他可穿戴设备。第二设备 1200 可以是配置有语音助手的设备, 利用该语音助手对从第一设备 1100 接收到的语音信息进行分析, 确定出语音信息对应的事件并将相关信息反馈给第一设备 1100, 以便第一设备 1100 执行语音信息对应的事件。具体的, 第二设备 1200 可以是手机、平板、个人计算机等。

在本公开示例性实施方式的语音信息处理过程中, 首先, 第一设备 1100 获取第一语音信息, 对第一语音信息进行关键词识别, 确定第一语音信息是否包含唤醒关键词。在第一语音信息包含唤醒关键词的情况下, 第一设备 1100 向第二设备 1200 发送语音助手唤醒指令。第二设备 1200 响应该语音助手唤醒指令启动安装于第二设备 1200 上的语音助手。

接下来, 如果第一设备 1100 获取到第二语音信息, 则可以将第二语音信息发送给第二设备 1200, 第二设备 1200 可以利用语音助手确定与第二语音信息对应的语音触发事件, 并将与该语音触发事件相关联的目标信息发送给第一设备 1100, 第一设备 1100 可以基于该目标信息执行语音触发事件。

在一些实施例中, 第二设备 1200 确定语音触发事件的过程, 可以仅由第二设备 1200 来实现, 也就是说, 第二设备 1200 利用语音助手对第二语音信息进行分析 (包括但不限于语音识别、语义识别、语音合成等操作), 根据分析结果确定出语音触发事件。

在另一些实施例中, 实现本公开语音信息处理过程的架构还可以包括服务器 1300, 在这种情况下, 第二设备 1200 可以将第二语音信息发送给服务器 1300, 由服务器 1300 对第二语音信息进行分析, 并将分析结果反馈给第二设备 1200。

需要说明的是, 本公开对语音触发事件的类型不做限制, 以第一设备 1100 是 AR 眼镜为例, 语音触发事件可以包括调节 AR 眼镜的音量、查看天气、接通电话、设置日程、录屏、截屏、打开/关闭相册、打开/关闭指定应用程序、关机等。

另外, 第二设备 1200 可以将语音助手和/或语音触发事件的用户界面 (User Interface, UI) 发送给第一设备 1100 进行显示。

应当注意的是, 第二设备 1200 发送给第一设备 1100 供第一设备 1100 显示的用户界面, 与第二设备 1200 自身显示的用户界面可以不同。具体的, 第二设备 1200 发送给第一设备 1100 的用户界面可以是由第二设备 1200 或服务器 1300 渲染生成的三维界面图像, 而第二设备 1200 自身显示的用户界面通常为二维界面, 并且二者的界面布置方式及内容也存在差别。

在第一设备 1100 例如为 AR 眼镜的实例中, 可以呈现出三维立体效果, 以使用户查看。

图 2 示出了适于用来实现本公开示例性实施方式的电子设备的示意图。本公开示例性实施方式中的第一设备和/或第二设备可以配置为如图 2 所示的形式。需要说明的是, 图 2 示出的电子设备仅是一个示例, 不应对本公开实施例的功能和使用范围带来任何限制。

本公开的电子设备至少包括处理器和存储器, 存储器被配置为存储一个或多个程序, 当一个或多个程序被处理器执行时, 使得处理器可以实现本公开示例性实施方式的语音信息处理方法。

具体的, 如图 2 所示, 电子设备 200 可以包括: 处理器 210、内部存储器 221、外部存储器接口 222、通用串行总线(Universal Serial Bus, USB)接口 230、充电管理模块 240、电源管理模块 241、电池 242、天线 1、天线 2、移动通信模块 250、无线通信模块 260、音频模块 270、扬声器 271、受话器 272、麦克风 273、耳机接口 274、传感器模块 280、显示屏 290、摄像模组 291、指示器 292、马达 293、按键 294 以及用户标识模块(Subscriber Identification Module, SIM)卡接口 295 等。其中传感器模块 280 可以包括深度传感器、压力传感器、陀螺仪传感器、气压传感器、磁传感器、加速度传感器、距离传感器、接近光传感器、指纹传感器、温度传感器、触摸传感器、环境光传感器及骨传导传感器等。

可以理解的是，本申请实施例示意的结构并不构成对电子设备 200 的具体限定。在本申请另一些实施例中，电子设备 200 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。

5 处理器 210 可以包括一个或多个处理单元，例如：处理器 210 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、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，也可以集成在一个或多个处理
10 器中。另外，处理器 210 中还可以设置存储器，用于存储指令和数据。

USB 接口 230 是符合 USB 标准规范的接口，具体可以是 MiniUSB 接口，MicroUSB 接口，USBTypeC 接口等。USB 接口 230 可以用于连接充电器为电子设备 200 充电，也可以用于电子设备 200 与外围设备之间传输数据。也可以用于连接耳机，通过耳机播放音频。该接口还可以用于连接其他电子设备，例如 AR 设备等。

15 充电管理模块 240 用于从充电器接收充电输入。其中，充电器可以是无线充电器，也可以是有线充电器。电源管理模块 241 用于连接电池 242、充电管理模块 240 与处理器 210。电源管理模块 241 接收电池 242 和/或充电管理模块 240 的输入，为处理器 210、内部存储器 221、显示屏 290、摄像模组 291 和无线通信模块 260 等供电。

20 电子设备 200 的无线通信功能可以通过天线 1、天线 2、移动通信模块 250、无线通信模块 260、调制解调处理器以及基带处理器等实现。

移动通信模块 250 可以提供应用在电子设备 200 上的包括 2G/3G/4G/5G 等无线通信的解决方案。

25 无线通信模块 260 可以提供应用在电子设备 200 上的包括无线局域网(Wireless Local Area Networks, WLAN) (如无线保真(Wireless Fidelity, Wi-Fi)网络)、蓝牙(Bluetooth, BT)、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。

30 电子设备 200 通过 GPU、显示屏 290 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 290 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 210 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。

电子设备 200 可以通过 ISP、摄像模组 291、视频编解码器、GPU、显示屏 290 及应用处理器等实现拍摄功能。在一些实施例中，电子设备 200 可以包括 1 个或 N 个摄像模组 291，N 为大于 1 的正整数，若电子设备 200 包括 N 个摄像头，N 个摄像头中有一个是主摄像头。

35 内部存储器 221 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。内部存储器 221 可以包括存储程序区和存储数据区。外部存储器接口 222 可以用于连接外部存储卡，例如 Micro SD 卡，实现扩展电子设备 200 的存储能力。

电子设备 200 可以通过音频模块 270、扬声器 271、受话器 272、麦克风 273、耳机接口 274 及应用处理器等实现音频功能。例如音乐播放、录音等。

40 音频模块 270 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 270 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 270 可以设置于处理器 210 中，或将音频模块 270 的部分功能模块设置于处理器 210 中。

45 扬声器 271，也称“喇叭”，用于将音频电信号转换为声音信号。电子设备 200 可以通过扬声器 271 收听音乐，或收听免提通话。受话器 272，也称“听筒”，用于将音频电信号转换成声音信号。当电子设备 200 接听电话或语音信息时，可以通过将受话器 272 靠近人

耳接听语音。麦克风 273，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 273 发声，将声音信号输入到麦克风 273。电子设备 200 可以设置至少一个麦克风 273。耳机接口 274 用于连接有线耳机。

针对电子设备 200 中传感器模块 280 可以包括的传感器，深度传感器用于获取景物的深度信息。压力传感器用于感受压力信号，可以将压力信号转换成电信号。陀螺仪传感器可以用于确定电子设备 200 的运动姿态。气压传感器用于测量气压。磁传感器包括霍尔传感器。电子设备 200 可以利用磁传感器检测翻盖皮套的开合。加速度传感器可检测电子设备 200 在各个方向上(一般为三轴)加速度的大小。距离传感器用于测量距离。接近光传感器可以包括例如发光二极管(LED)和光检测器，例如光电二极管。指纹传感器用于采集指纹。温度传感器用于检测温度。触摸传感器可以将检测到的触摸操作传递给应用处理器，以确定触摸事件类型。可以通过显示屏 290 提供与触摸操作相关的视觉输出。环境光传感器用于感知环境光亮度。骨传导传感器可以获取振动信号。

按键 294 包括开机键，音量键等。按键 294 可以是机械按键。也可以是触摸式按键。马达 293 可以产生振动提示。马达 293 可以用于来电振动提示，也可以用于触摸振动反馈。指示器 292 可以是指示灯，可以用于指示充电状态，电量变化，也可以用于指示消息，未接来电，通知等。SIM 卡接口 295 用于连接 SIM 卡。电子设备 200 通过 SIM 卡和网络交互，实现通话以及数据通信等功能。

本申请还提供了一种计算机可读存储介质，该计算机可读存储介质可以是上述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

计算机可读存储介质例如可以是一——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。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机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。

计算机可读存储介质可以发送、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。计算机可读存储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、电线、光缆、RF 等等，或者上述的任意合适的组合。

计算机可读存储介质承载有一个或者多个程序，当上述一个或者多个程序被一个该电子设备执行时，使得该电子设备实现如下述实施例中所述的方法。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现，所描述的单元也可以设置在处理器中。其中，这些单元的名称在某种情况下并不构成对该单元本身的限定。

图 3 示意性示出了本公开的示例性实施方式的应用于第一设备的语音信息处理方法的流程图。参考图 3，应用于第一设备的语音信息处理方法可以包括以下步骤：

S32. 获取第一语音信息, 如果第一语音信息包含唤醒关键词, 则向第二设备发送语音助手唤醒指令, 以便第二设备启动语音助手。

在本公开的示例性实施方式中, 第一设备配备有语音收录装置, 具体的, 第一设备可以配备有单个麦克风或麦克风阵列, 以获取第一设备周围的语音信息。

5 另外, 第一设备还配备有 DSP (Digital Signal Process) 芯片, 被配置为在获取第一语音信息的情况下, 对第一语音信息进行分析, 以确定第一语音信息是否包含唤醒关键词, 该唤醒关键词用于唤醒第二设备上的语音助手。

容易理解的是, 第一设备通常配备有声卡。在一些实施例中, 第一设备配备的语音收录装置和 DSP 芯片均可以与声卡连接, 通过声卡将语音收录装置获取到的第一语音信息发送给 DSP 芯片。然而, 在这种情况下, 需要启动声卡才能使 DSP 芯片获取到第一语音信息。

在另一些实施例中, 第一设备配备的语音收录装置可以直接与 DSP 芯片连接, 由此, DSP 芯片获取第一语音信息的过程无需依赖声卡的启动即可实现。

15 针对步骤 S32, 第一设备的语音收录装置可以获取第一语音信息, DSP 芯片可以对第一语音信息进行分析, 判断第一语音信息中是否包含唤醒关键词, 在第一语音信息包含唤醒关键词的情况下, 第一设备可以向第二设备发送语音助手唤醒指令, 第二设备可以响应该语音助手唤醒指令启动语音助手。

在本公开的一些实施例中, 第一设备与第二设备通过 USB 连接, 这种有线的连接方式可以避免信息传输丢帧的问题, 提高了语音处理的准确性。

20 在本公开的另一些实施例中, 第一设备还可以通过蓝牙或 WiFi 的方式与第二设备连接, 本公开对此不做限制。

可以理解的是, 在第一设备获取到包含唤醒关键词的第一语音信息之前, 第二设备的语音助手处于关闭状态, 整个系统处于半休眠状态, 这样有助于减少系统的功耗。

25 此外, 在第二设备启动语音助手后, 第二设备可以将语音助手的三维界面图像发送给第一设备, 以便在第一设备的显示端可以显示出语音助手的三维界面图像, 例如, 用户可以在 AR 眼镜的镜片上看到语音助手的三维界面图像, 以提示用户语音助手已开启, 可以进行语音命令的录入。

语义助手的三维界面图像可以是由第二设备渲染生成, 具体的, 可以是在响应语音助手唤醒指令后进行渲染而生成的图像。也可以是预先渲染生成并存储的图像, 响应语音助手唤醒指令而调取出的图像, 本公开对渲染的方式、时间等均不做限制。

30 S34. 获取第二语音信息, 将第二语音信息发送给第二设备, 以便第二设备利用语音助手确定语音触发事件。

在第二设备启动语音助手的情况下, 第一设备的麦克风可以获取第二语音信息, 并将第二语音信息发送给第二设备。应该理解的是, 在第二设备的语音助手开启的情况下, 第一设备可以控制用于进行唤醒关键词识别的 DSP 芯片处于休眠状态。

在第一设备与第二设备通过 USB 连接的情况下, 可以通过 USB 重定向的方式使第二设备获取到第二语音信息。也就是说, 第一设备的语音收录装置被重定向为第二设备的语音输入, 第二设备的语音助手可以监听第一设备录入的语音。

40 针对 USB 重定向过程, 具体的, 首先, 通过第二设备中与第一设备匹配的服务可以监听第一设备的 USB 插入事件, 应该理解的是, 该服务仅响应第一设备的 USB 插入事件, 而不响应其他设备的 USB 插入事件; 接下来, 第二设备可以通过回调函数将自身的语音收录装置重置为第一设备的语音收录装置。

45 在第一设备为 AR 眼镜且第二设备为手机的实例中, 首先, 手机的 AR service (AR 服务) 可以监听 AR 眼镜 USB 插入事件, 同时可以提供安全验证; 接下来, 手机可以将其语音收录装置重置为 AR 眼镜的麦克风; 然后, 用户直接通过 AR 眼镜的麦克风录入语音信

息，该语音信息会直接传送至手机的语音收录装置，并传递给手机上实现与 AR 眼镜交互的上层应用（即，语音助手）。

根据本公开的一些实施例，第二设备可以利用语音助手对第二语音信息进行分析（包括但不限于语音识别、语义识别、语音合成等操作），以确定出与第二语音信息对应的语音触发事件。例如，分析出的语音触发事件为“调低音量”；又如，分析出的语音触发事件为“关机”；再如，分析出的语音触发事件为“打开相册”，等等。

在这些实施例中，可以在第二设备不联网（即，不连接服务器）的情况下实现。

根据本公开的另一些实施例，第二设备可以利用语音助手将第二语音信息发送给服务器，由服务器对第二语音信息进行分析，确定出语音触发事件，并将语音触发事件反馈给第二设备。另外，服务器可以根据业务实际需求实现动态扩容服务。

在一个实施例中，服务器可以基于第一语音处理过程确定出语音触发事件。该第一语音处理过程可以包括：首先，服务器对第二语音信息进行语音识别，以将第二语音信息转换为文本信息；接下来，服务器可以对该文本信息进行语义识别，确定出文本信息对应的语义，并生成语义识别结果，其中，语义识别结果可以以文本的形式呈现；然后，服务器根据语义识别结果确定出与第二语音信息对应的语音触发事件。在这种情况下，语音触发事件可以例如为控制第一设备的事件。例如，控制第一设备关机。

在另一个实施例中，服务器可以基于第二语音处理过程确定出语音触发事件。该第二语音处理过程可以包括：首先，服务器对第二语音信息进行语音识别，以将第二语音信息转换为文本信息；接下来，服务器可以对该文本信息进行语义识别，得到语音识别结果；然后，服务器可以根据该语音识别结果确定与第二语音信息对应的响应文本，并将该响应文本转换为第三语音信息，作为语音触发事件包含的信息。在这种情况下，第二语音信息可以是询问天气的信息，对应的语音触发事件可以是语音播报天气情况，而上述响应文本指的是以文本形式确定出的实时天气信息，第三语音信息是对响应文本进行文本转换语音的过程而生成的语音信息，例如为“晴天”、“小雨”等与实时天气相关的语音信息，可以对应于下述向第二设备反馈的目标信息。

需要说明的是，一方面，上面语音分析的过程仅是示例，本公开对具体过程不做限制；另一方面，上述语音分析的过程也可以在第二设备上实现，在这种情况下，方案可以不需要服务器的参与。

S36. 接收由第二设备反馈的目标信息，并基于目标信息执行语音触发事件。

在第二设备确定出与第二语音信息对应的语音触发事件之后，第二设备可以确定出与执行该语音触发事件相关联的目标信息，并将该目标信息发送给第一设备，以便第一设备基于该目标信息执行语音触发事件。其中，目标信息可以是控制第一设备状态的指令，或者可以是与显示语音触发事件对应内容相关的信息。

以语音触发事件为“调低音量”为例，目标信息可以是调低第一设备音量的指令，第一设备在接收到该指令后，可以调低自身音量。

又如，以语音触发事件为“打开相册”为例，目标信息可以包括相册内各图像的信息，以便第一设备接收到目标信息后，可以在第一设备的显示端显示出相册中包含的图像。例如，可以在 AR 眼镜的镜片上显示出相册中包含的照片。

再如，以语音触发事件为“关机”为例，目标信息可以是控制第一设备关机的指令，第一设备在接收到该指令后，可以关机。

可以理解的是，第二设备的语音助手可以集成语音触发事件与本地系统命令的映射关系，以便在确定出语音触发事件的情况下，可以利用这些系统命令对第一设备进行控制。

此外，在第二设备利用语音助手确定出语音触发事件后，本公开示例性实施方式还包括向第一设备发送与语音触发事件对应的三维界面图像的方案。

具体的，第二设备可以确定出语音触发事件对应的三维界面图像，并将该三维界面图

像发送给第一设备，以便第一设备在显示端显示出该三维界面图像，例如，用户可以在 AR 眼镜的镜片上看到“调低音量”对应的三维界面图像。

与上述语音助手的三维界面图像类似地，与语音触发事件对应的三维界面图像由第二设备渲染生成，且本公开对渲染的方式、时间等均不做限制。可以理解的是，本公开所述的三维界面图像可以包含文字、符号、静态图片、动态图片、视频中的一种或多种的组合。

考虑到用户可能较长时间未发出语音而导致功耗增加的问题，在本公开的一些实施例中，在第一设备向第二设备发送语音助手唤醒指令后，第一设备开始计时，经历预定时间段（例如，20 秒）后，如果未获取到第二语音信息，则第一设备向第二设备发送语音助手关闭指令，第二设备可以响应该语音助手关闭指令，关闭语音助手。

另外，也可以第二设备自行计时，经历预定时间段后，未获取到第二语音信息，则第二设备自行关闭语音助手。

本公开示例性实施方式还提供了一种应用于第二设备的语音信息处理方法的流程图。参考图 4，应用于第二设备的语音信息处理方法可以包括以下步骤：

S42. 响应语音助手唤醒指令，启动语音助手；其中，语音助手唤醒指令由第一设备在确定出第一语音信息包含唤醒关键词的情况下发送给第二设备；

S44. 获取第一设备发送的第二语音信息，利用语音助手确定与第二语音信息对应的语音触发事件；

S46. 将与执行语音触发事件相关联的目标信息反馈给第一设备，以便第一设备基于目标信息执行语音触发事件。

步骤 S42 至步骤 S46 的具体过程已在上述步骤 S32 至步骤 S36 中进行了详细说明，在此不再赘述。

下面将参考图 5 对本公开一个实施例的实现语音处理过程的设备交互图进行说明。

在步骤 S502 中，第一设备获取第一语音信息，并通过配备的 DSP 芯片判断第一语音信息中是否包含唤醒关键词，其中，唤醒关键词可以是用户自定义的关键词，例如，“小布”。在步骤 S504 中，如果第一语音信息包含唤醒关键词，则第一设备可以向第二设备发送语音助手唤醒指令。

在步骤 S506 中，第二设备启动语音助手。另外，在步骤 S508 中，第二设备可以向第一设备发送语音助手的三维界面图像。在步骤 S510 中，第一设备可以在其显示端显示出语音助手的三维界面图像。例如，可以在 AR 眼镜的镜片上显示出语音助手的三维界面图像。

步骤 S502 至步骤 S510 示例性描述了通过第一设备的唤醒服务来启动第二设备的语音助手的过程，下面将参考步骤 S512 至步骤 S528 来说明通过语音控制第一设备的过程。

在步骤 S512 中，第一设备获取第二语音信息，并将第二语音信息发送给第二设备。

在步骤 S514 中，第二设备将第二语音信息发送给服务器。在步骤 S516 中，服务器对第二语音信息进行分析，得到分析结果。在步骤 S518 中，服务器将第二语音信息的分析结果反馈给第二设备。在步骤 S520 中，第二设备根据服务器的分析结果确定语音触发事件。

在步骤 S522 中，第二设备向第一设备发送目标信息，该目标信息与执行语音触发事件相关联，也就是说，目标信息是第一设备能够执行语音触发事件必须的信息。在步骤 S524 中，第一设备执行语音触发事件。

另外，在步骤 S526 中，第二设备向第一设备发送与语音触发事件对应的三维界面图像。在步骤 S528 中，第一设备显示与语音触发事件对应的三维界面图像。

下面将参考图 6 对本公开实施例的语音信息处理方案的整体软硬件架构进行说明。在该实施例中，第一设备为 AR 眼镜，第二设备为手机。其中，AR 眼镜与手机通过 USB 连接，手机与服务器可以通过 3G、4G、5G、WiFi 等方式连接。

针对 AR 眼镜，其上的操作系统可以是 RTOS（实时操作系统），AR 眼镜本身配备有 DSP 芯片，用于提供唤醒服务。在 AR 眼镜的麦克风获取到上述第一语音信息时，利用关键词识别引擎来确定第一语音信息是否包含唤醒关键词。另外，在唤醒手机的语音助手的情况下，AR 眼镜可以呈现出由手机渲染的语音助手的三维用户界面。

5 针对手机，操作系统中可以配置有 AR 软件平台（例如，ARCore 或 ARKit）。在 AR 软件平台的基础上，可以配置有语音助手的应用程序（APP），可以响应 AR 眼镜的唤醒服务生成的指令启动语音助手 APP。通过语音助手用户界面的 UI 交互，可以对 AR 眼镜的麦克风进行复位，另外，还可以对应 AR 眼镜上显示出的语音助手的用户界面。

此外，利用手机中语音助手 APP 提供的语音软件开发工具包，可以实现与服务器的语音语义分析引擎的交互，以将上述第二语音信息发送给服务器，由服务器的语音语义分析引擎对第二语音信息进行分析，并反馈分析结果。

需要说明的是，图 6 的架构仅是示例，不应作为本公开内容的限制。

应当注意，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该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。

进一步的，本示例实施方式中还提供了一种应用于第一设备的语音信息处理装置。

图 7 示意性示出了本公开的示例性实施方式的应用于第一设备的语音信息处理装置的方框图。参考图 7，根据本公开的示例性实施方式的应用于第一设备的语音信息处理装置 7 可以包括唤醒触发模块 71、语音发送模块 73 和事件执行模块 75。

具体的，唤醒触发模块 71 可以被配置为获取第一语音信息，如果第一语音信息包含唤醒关键词，则向第二设备发送语音助手唤醒指令，以便第二设备启动语音助手；语音发送模块 73 可以被配置为获取第二语音信息，将第二语音信息发送给第二设备，以便第二设备利用语音助手确定与第二语音信息对应的语音触发事件；事件执行模块 75 可以被配置为接收由第二设备反馈的与执行语音触发事件相关联的目标信息，并基于目标信息执行语音触发事件。

根据本公开的示例性实施例，参考图 8，相比于语音信息处理装置 7，语音信息处理装置 8 还可以包括图像显示模块 81。

具体的，图像显示模块 81 可以被配置为执行：在第二设备启动语音助手后，接收语音助手的三维界面图像；其中，语音助手的三维界面图像由第二设备渲染生成；在第一设备的显示端显示出语音助手的三维界面图像。

根据本公开的示例性实施例，第一设备与第二设备通过 USB 连接，在这种情况下，语音发送模块 73 可以被配置为执行：获取第二语音信息；通过 USB 重定向的方式将第二语音信息发送给第二设备。

35 根据本公开的示例性实施例，图像显示模块 81 还可以被配置为执行：接收与语音触发事件对应的三维界面图像；其中，与语音触发事件对应的三维界面图像由第二设备渲染生成；在第一设备的显示端显示出与语音触发事件对应的三维界面图像。

根据本公开的示例性实施例，参考图 9，相比于语音信息处理装置 7，语音信息处理装置 9 还可以包括关闭触发模块 91。

40 具体的，关闭触发模块 91 可以被配置为执行：在向第二设备发送语音助手唤醒指令后，开始计时；经历预定时间段后，如果未获取到第二语音信息，则向第二设备发送语音助手关闭指令，以便第二设备关闭语音助手。

进一步的，本示例实施方式中还提供了一种应用于第二设备的语音信息处理装置。

45 图 10 示意性示出了本公开的示例性实施方式的应用于第二设备的语音信息处理装置的方框图。参考图 10，根据本公开的示例性实施方式的应用于第二设备的语音信息处理装

置 10 可以包括语音助手启动模块 101、事件确定模块 103 和信息反馈模块 105。

具体的,语音助手启动模块 101 可以被配置为响应语音助手唤醒指令,启动语音助手;其中,语音助手唤醒指令由第一设备在确定出第一语音信息包含唤醒关键词的情况下发送给第二设备;事件确定模块 103 可以被配置为获取第一设备发送的第二语音信息,利用语音助手确定与第二语音信息对应的语音触发事件;信息反馈模块 105 可以被配置为将与执行语音触发事件相关联的目标信息反馈给第一设备,以便第一设备基于目标信息执行语音触发事件。

根据本公开的示例性实施例,信息反馈模块 105 还可以被配置为执行:在启动语音助手后,向第一设备发送语音助手的三维界面图像,以便第一设备在显示端显示出语音助手的三维界面图像;其中,语音助手的三维界面图像由第二设备渲染生成。

根据本公开的示例性实施例,第一设备与第二设备通过 USB 连接,在这种情况下,事件确定模块 103 获取第一设备发送的第二语音信息的过程可以被配置为执行:通过 USB 重定向的方式从第一设备获取第二语音信息。

根据本公开的示例性实施例,事件确定模块 103 还可以被配置为执行:利用语音助手将第二语音信息发送给服务器,以便服务器确定与第二语音信息对应的语音触发事件;获取由服务器确定出的与第二语音信息对应的语音触发事件。

根据本公开的示例性实施例,信息反馈模块 105 还可以被配置为执行:在利用语音助手确定与第二语音信息对应的语音触发事件后,向第一设备发送与语音触发事件对应的三维界面图像,以便第一设备在显示端显示出与语音触发事件对应的三维界面图像;其中,与语音触发事件对应的三维界面图像由第二设备渲染生成。

根据本公开的示例性实施例,参考图 11,相比于语音信息处理装置 10,语音信息处理装置 11 还可以包括语音助手关闭模块 111。

具体的,语音助手关闭模块 111 可以被配置为执行:响应语音助手关闭指令,关闭语音助手。

由于本公开实施方式的语音信息处理装置的各个功能模块与上述方法实施方式中相同,因此在此不再赘述。

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,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)执行根据本公开实施方式的方法。

此外,上述附图仅是根据本公开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。

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。

本领域技术人员在考虑说明书及实践这里公开的内容后,将容易想到本公开的其他实施例。本申请旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由权利要求指出。

应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可

以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限。

权利要求

1. 一种语音信息处理方法，应用于第一设备，包括：

获取第一语音信息，如果所述第一语音信息包含唤醒关键词，则向第二设备发送语音助手唤醒指令，以便所述第二设备启动语音助手；

获取第二语音信息，将所述第二语音信息发送给所述第二设备，以便所述第二设备利用所述语音助手确定语音触发事件，所述语音触发事件与所述第二语音信息对应；

接收由所述第二设备反馈的目标信息，并基于所述目标信息执行所述语音触发事件。
2. 根据权利要求 1 所述的语音信息处理方法，其中，在所述第二设备启动语音助手后，所述语音信息处理方法还包括：

接收所述语音助手的三维界面图像；其中，所述语音助手的三维界面图像由所述第二设备渲染生成；

在所述第一设备的显示端显示所述语音助手的三维界面图像。
3. 根据权利要求 1 所述的语音信息处理方法，其中，所述第一设备与所述第二设备通过 USB 连接；其中，获取第二语音信息，将所述第二语音信息发送给所述第二设备，包括：

获取所述第二语音信息；

通过 USB 重定向的方式将所述第二语音信息发送给所述第二设备。
4. 根据权利要求 1 或 2 所述的语音信息处理方法，其中，所述语音信息处理方法还包括：

接收与所述语音触发事件对应的三维界面图像；其中，与所述语音触发事件对应的三维界面图像由所述第二设备渲染生成；

在所述第一设备的显示端显示与所述语音触发事件对应的三维界面图像。
5. 根据权利要求 1 至 3 中任一项所述的语音信息处理方法，其中，所述语音信息处理方法还包括：

在向所述第二设备发送所述语音助手唤醒指令后，开始计时；

经历预定时间段后，如果未获取到所述第二语音信息，则向所述第二设备发送语音助手关闭指令，以便所述第二设备关闭所述语音助手。
6. 一种语音信息处理方法，应用于第二设备，包括：

响应语音助手唤醒指令，启动语音助手；其中，所述语音助手唤醒指令由第一设备在确定出第一语音信息包含唤醒关键词的情况下发送给所述第二设备；

获取所述第一设备发送的第二语音信息，利用所述语音助手确定与所述第二语音信息对应的语音触发事件；

将与执行所述语音触发事件相关联的目标信息反馈给所述第一设备，以便所述第一设备基于所述目标信息执行所述语音触发事件。
7. 根据权利要求 6 所述的语音信息处理方法，其中，在启动所述语音助手后，所述语音信息处理方法还包括：

向所述第一设备发送所述语音助手的三维界面图像，以便所述第一设备在显示端显示出所述语音助手的三维界面图像；

其中，所述语音助手的三维界面图像由所述第二设备渲染生成。
8. 根据权利要求 6 所述的语音信息处理方法，其中，所述第一设备与所述第二设备通过 USB 连接；其中，获取所述第一设备发送的第二语音信息包括：

通过 USB 重定向的方式从所述第一设备获取所述第二语音信息。
9. 根据权利要求 6 所述的语音信息处理方法，其中，利用所述语音助手确定与所述第二语音信息对应的语音触发事件包括：

利用所述语音助手将所述第二语音信息发送给服务器,以便所述服务器确定与所述第二语音信息对应的语音触发事件;

获取由所述服务器确定出的语音触发事件,所述语音触发事件与所述第二语音信息对应。

5 10. 根据权利要求 9 所述的语音信息处理方法,其中,所述语音触发事件基于所述服务器执行第一语音处理过程而确定出;其中,所述第一语音处理过程包括:

将所述第二语音信息转换为文本信息;

对所述文本信息进行语义识别,根据语义识别结果确定与所述第二语音信息对应的语音触发事件。

10 11. 根据权利要求 9 所述的语音信息处理方法,其中,所述语音触发事件基于所述服务器执行第二语音处理过程而确定出,所述第二语音处理过程包括:

将所述第二语音信息转换为文本信息;

对所述文本信息进行语义识别,得到语义识别结果;

根据所述语义识别结果确定与所述第二语音信息对应的响应文本;

15 将所述响应文本转换为第三语音信息,作为所述语音触发事件包含的信息。

12. 根据权利要求 6 或 7 所述的语音信息处理方法,其中,在利用所述语音助手确定与所述第二语音信息对应的语音触发事件后,所述语音信息处理方法还包括:

向所述第一设备发送与所述语音触发事件对应的三维界面图像,以便所述第一设备在显示端显示出与所述语音触发事件对应的三维界面图像;

20 其中,与所述语音触发事件对应的三维界面图像由所述第二设备渲染生成。

13. 根据权利要求 6 至 11 中任一项所述的语音信息处理方法,其中,所述语音信息处理方法还包括:

响应语音助手关闭指令,关闭所述语音助手。

14. 一种语音信息处理装置,应用于第一设备,包括:

25 唤醒触发模块,被配置为获取第一语音信息,如果所述第一语音信息包含唤醒关键词,则向第二设备发送语音助手唤醒指令,以便所述第二设备启动语音助手;

语音发送模块,被配置为获取第二语音信息,将所述第二语音信息发送给所述第二设备,以便所述第二设备利用所述语音助手确定语音触发事件,所述语音触发事件与所述第二语音信息对应;

30 事件执行模块,被配置为接收由所述第二设备反馈的目标信息,并基于所述目标信息执行所述语音触发事件。

15. 根据权利要求 14 所述的语音信息处理装置,其中,所述语音信息处理装置还包括:

35 图像显示模块,被配置为在所述第二设备启动所述语音助手后,接收所述语音助手的三维界面图像;其中,所述语音助手的三维界面图像由所述第二设备渲染生成;以及在所述第一设备的显示端显示出所述语音助手的三维界面图像。

16. 根据权利要求 14 所述的语音信息处理装置,其中,所述第一设备与所述第二设备通过 USB 连接;语音发送模块还被配置为获取所述第二语音信息,通过 USB 重定向的方式将所述第二语音信息发送给所述第二设备。

40 17. 一种语音信息处理装置,应用于第二设备,包括:

语音助手启动模块,被配置为响应语音助手唤醒指令,启动语音助手;其中,所述语音助手唤醒指令由第一设备在确定出第一语音信息包含唤醒关键词的情况下发送给所述第二设备;

45 事件确定模块,被配置为获取所述第一设备发送的第二语音信息,利用所述语音助手确定与所述第二语音信息对应的语音触发事件;

信息反馈模块,被配置为将与执行所述语音触发事件相关联的目标信息反馈给所述第一设备,以便所述第一设备基于所述目标信息执行所述语音触发事件。

5 18. 根据权利要求 17 所述的语音信息处理装置,其中,所述信息反馈模块还被配置为在启动所述语音助手后,向所述第一设备发送所述语音助手的三维界面图像,以便所述第一设备在显示端显示出所述语音助手的三维界面图像,所述语音助手的三维界面图像由所述第二设备渲染生成。

19. 一种计算机可读存储介质,其上存储有计算机程序,所述程序被处理器执行时实现如权利要求 1 至 13 中任一项所述的语音信息处理方法。

10 20. 一种电子设备,包括:
处理器;

存储器,被配置为存储一个或多个程序,当所述一个或多个程序被所述处理器执行时,使得所述处理器实现如权利要求 1 至 13 中任一项所述的语音信息处理方法。

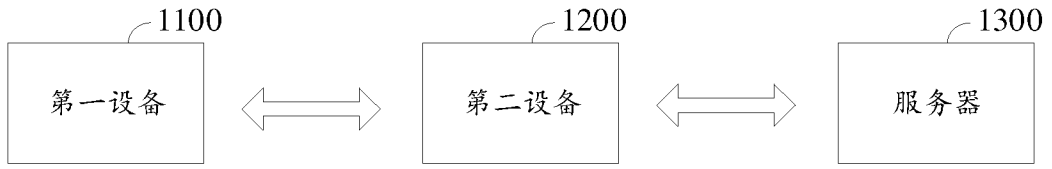


图 1

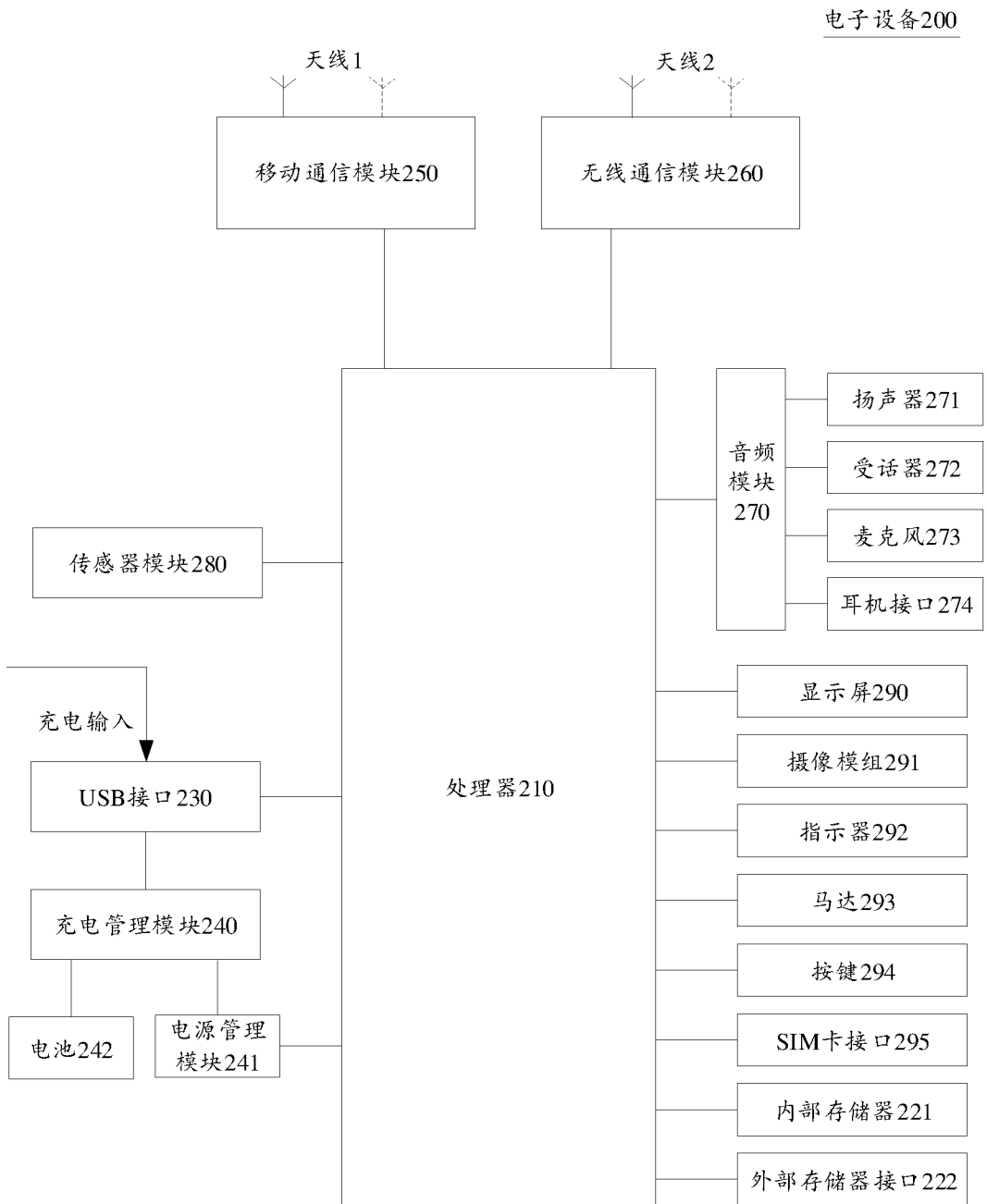


图 2

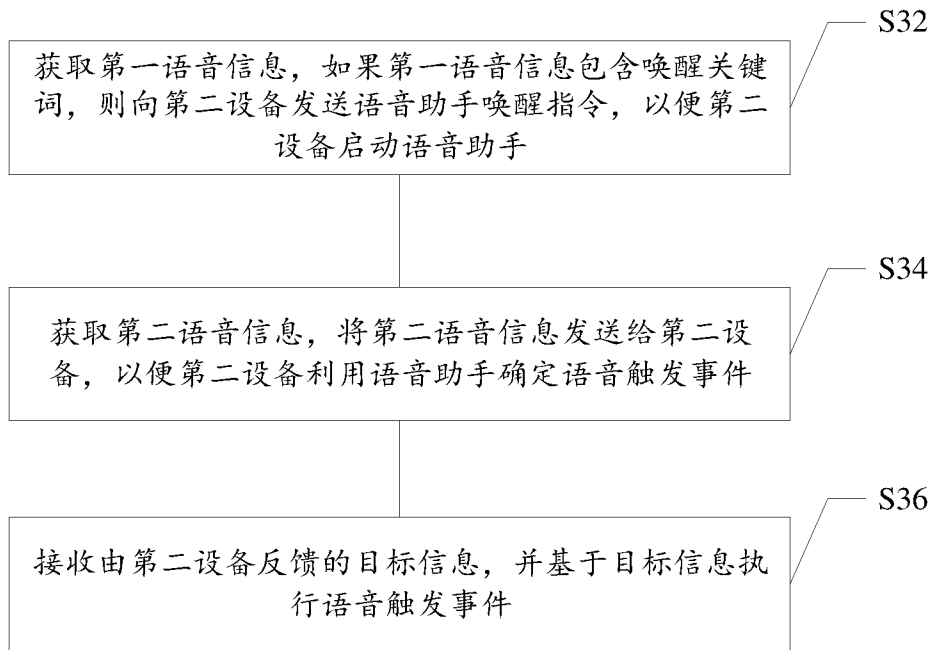


图 3

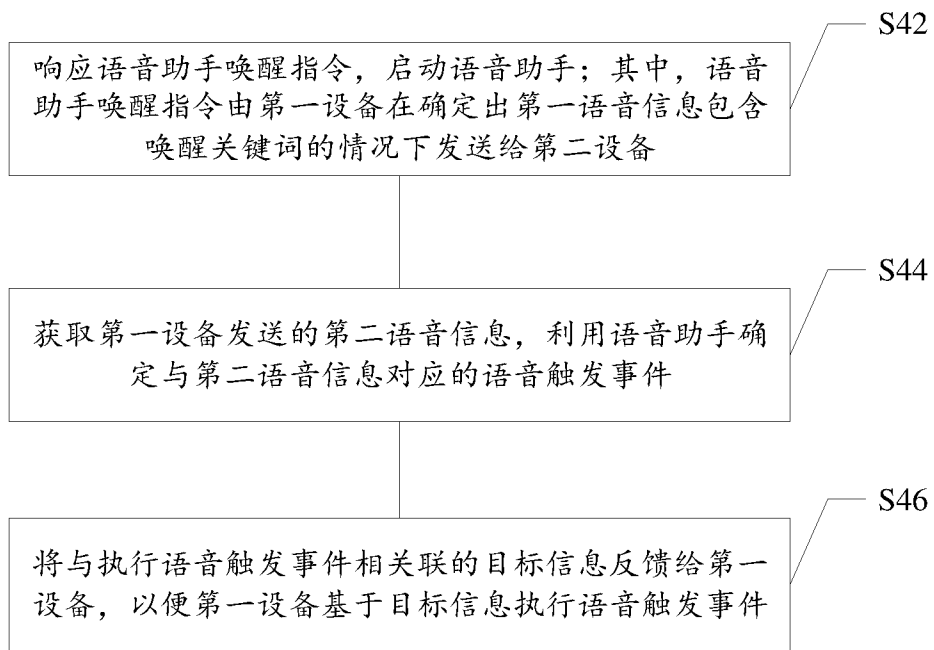


图 4

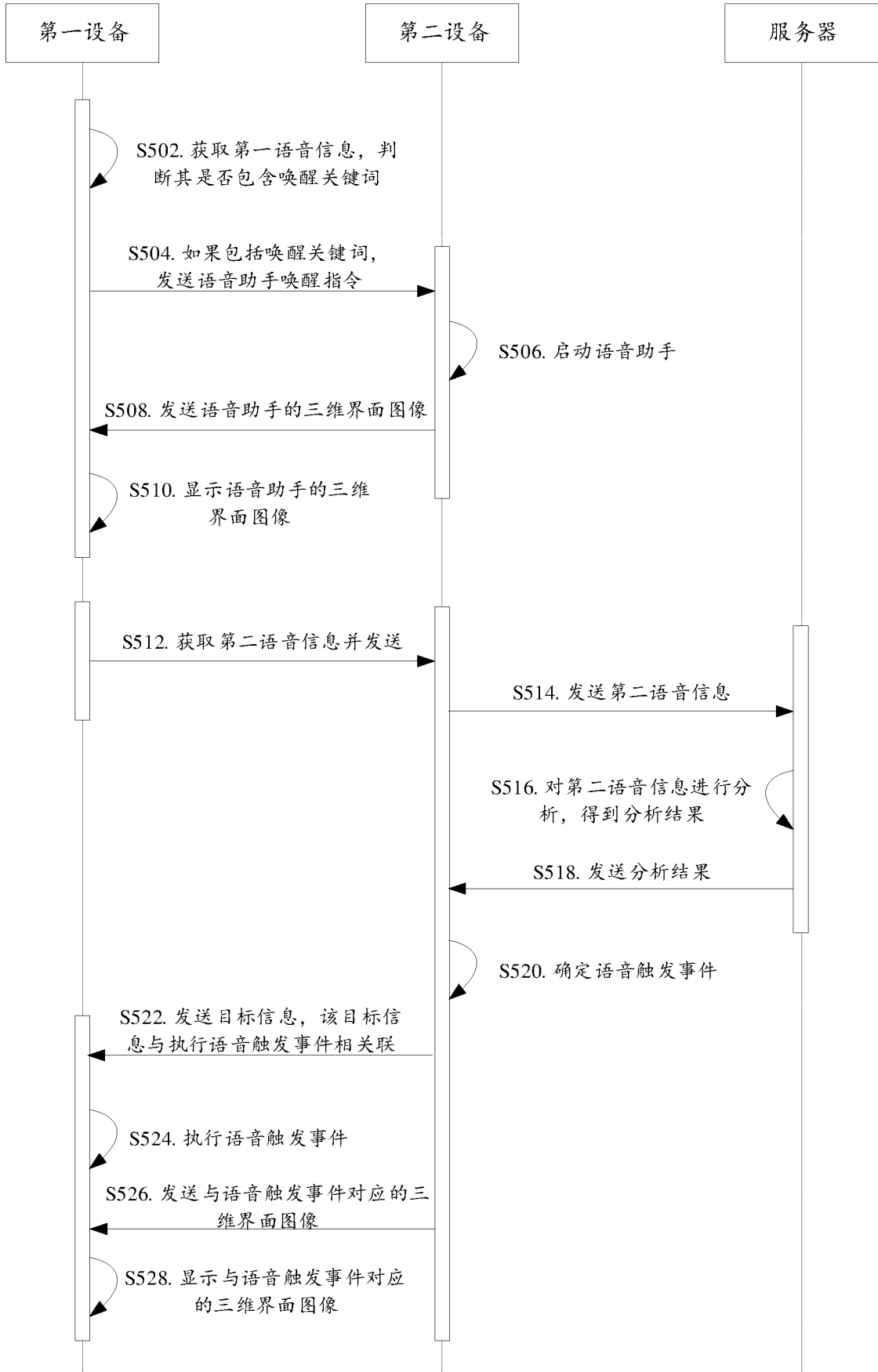


图 5

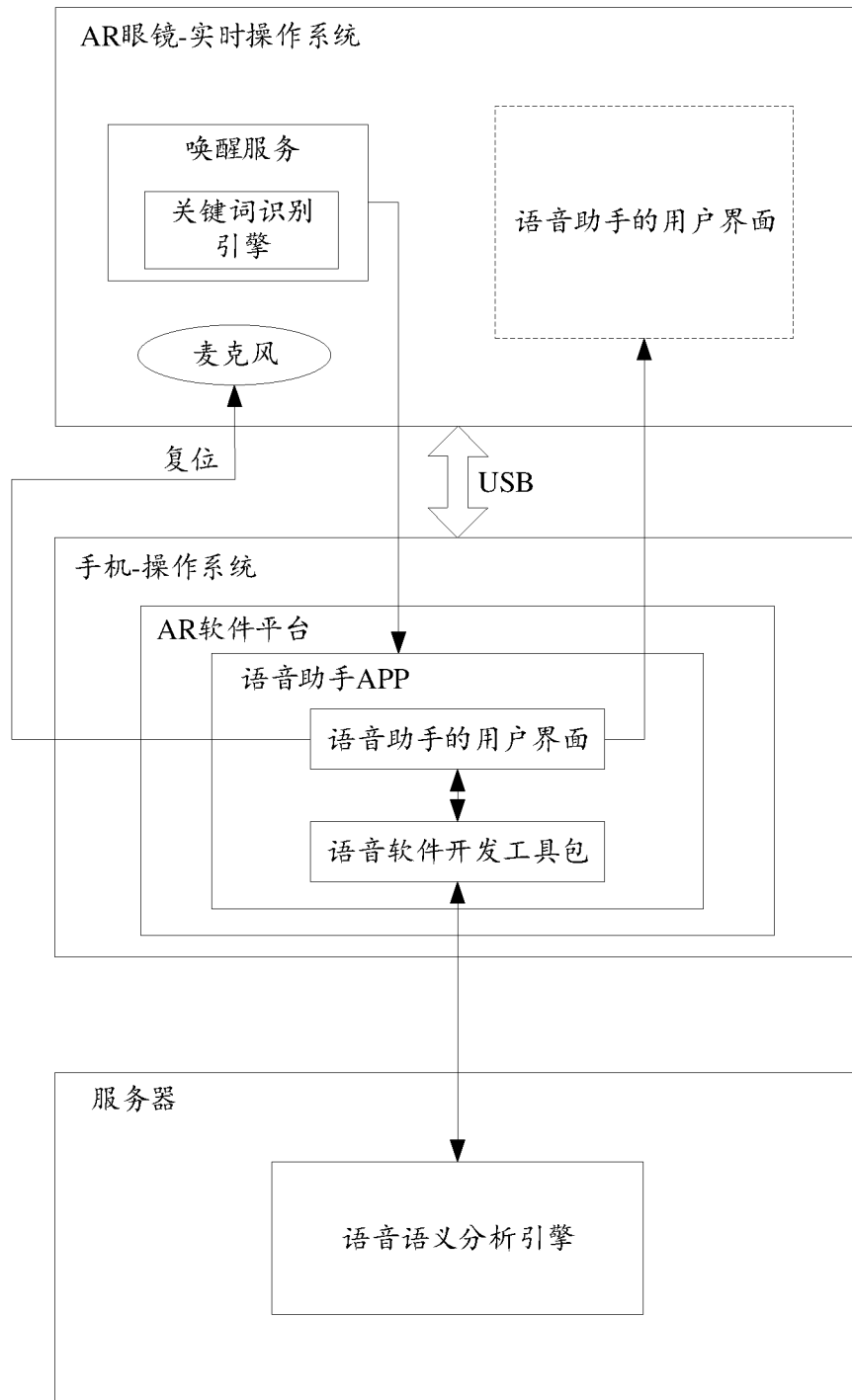


图 6

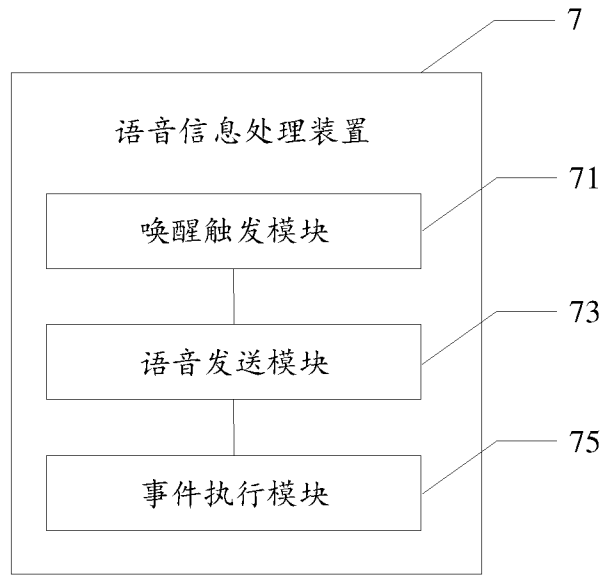


图 7

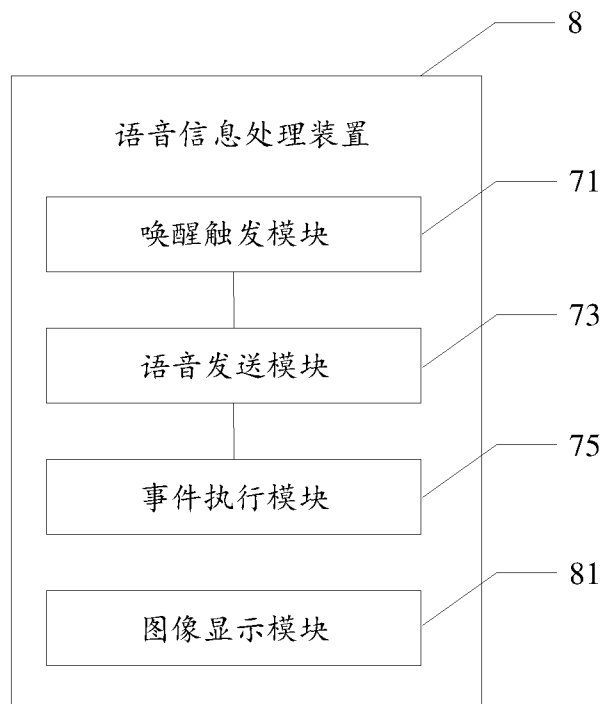


图 8

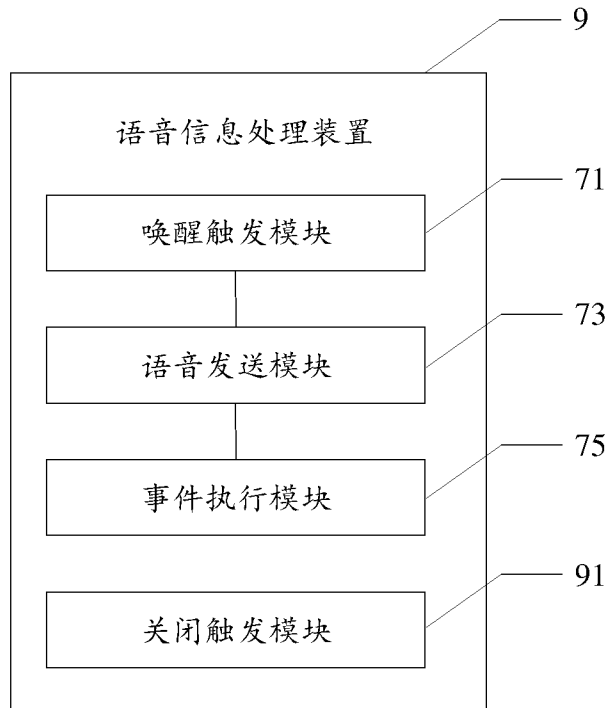


图 9

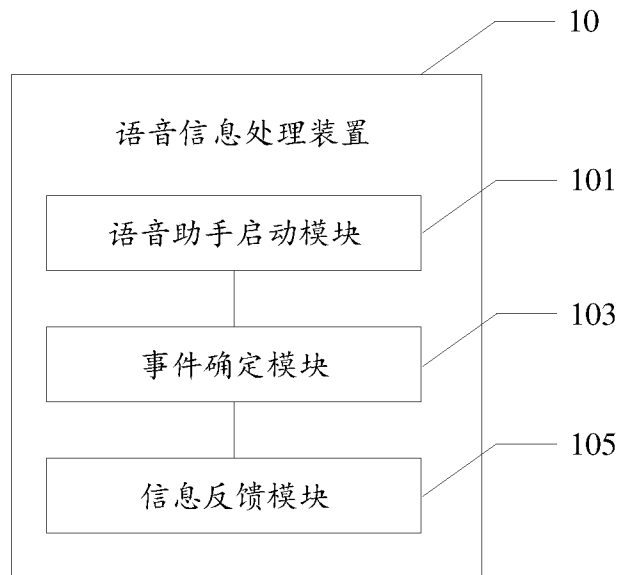


图 10

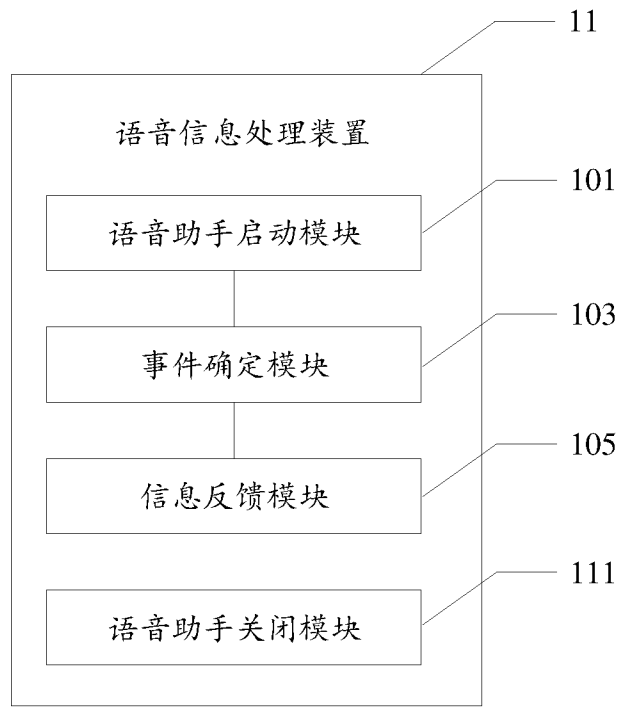


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/081332

A. CLASSIFICATION OF SUBJECT MATTER		
G06F 9/4401(2018.01)i; G10L 15/22(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G06F; 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) CNPAT, WPI, EPODOC, CNKI, GOOGLE, ISI: 可穿戴, 耳机, 智能眼镜, 语音助手, 唤醒, 启动, 开启, 传送, 发送, 传输, 手机, 电子设备, 服务器, 交互; wearable device, headset, intelligent glasses, smart glasses, wake up, voice, speech, assistant, transmit, phone, interaction		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111694605 A (GUANGDONG OPPO MOBILE TELECOMMUNICATIONS CORP., LTD.) 22 September 2020 (2020-09-22) claims 1-17	1-20
X	CN 103558916 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 05 February 2014 (2014-02-05) description paragraphs [0050]-[0130], [0228]-[0231], figure 1	1-20
X	CN 107978316 A (XI'AN FENGYU INFORMATION TECHNOLOGY CO., LTD.) 01 May 2018 (2018-05-01) description, paragraphs [0105]-[0182], and figure 1	1-20
X	CN 111096680 A (GUANGDONG MIDEA KITCHEN APPLIANCES MANUFACTURING CO., LTD. et al.) 05 May 2020 (2020-05-05) description, paragraphs [0063]-[0082]	1-20
A	US 2018293982 A1 (YUTOU TECHNOLOGY HANGZHOU CO., LTD.) 11 October 2018 (2018-10-11) entire document	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 15 April 2021		Date of mailing of the international search report 27 April 2021
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/CN2021/081332

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111694605	A	22 September 2020	None			
CN	103558916	A	05 February 2014	WO	2015066949	A1	14 May 2015
CN	107978316	A	01 May 2018	None			
CN	111096680	A	05 May 2020	None			
US	2018293982	A1	11 October 2018	JP	2018533139	A	08 November 2018
				TW	201714168	A	16 April 2017
				CN	106572418	A	19 April 2017
				EP	3361712	A1	15 August 2018
				WO	2017059812	A1	13 April 2017

国际检索报告

国际申请号

PCT/CN2021/081332

<p>A. 主题的分类</p> <p>G06F 9/4401(2018.01)i; G10L 15/22(2006.01)i</p> <p>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</p>																				
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06F; G10L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT, WPI, EPDOC, CNKI, GOOGLE, ISI: 可穿戴, 耳机, 智能眼镜, 语音助手, 唤醒, 启动, 开启, 传送, 发送, 传输, 手机, 电子设备, 服务器, 交互; wearable device, headset, intelligent glasses, smart glasses, wake up, voice, speech, assistant, transmit, phone, interaction</p>																				
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 111694605 A (OPPO广东移动通信有限公司) 2020年 9月 22日 (2020 - 09 - 22) 权利要求1-17</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 103558916 A (百度在线网络技术北京有限公司) 2014年 2月 5日 (2014 - 02 - 05) 说明书第[0050]-[0130]、[0228]-[0231]段、图1</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 107978316 A (西安蜂语信息科技有限公司) 2018年 5月 1日 (2018 - 05 - 01) 说明书第[0105]-[0182]段、图1</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 111096680 A (广东美的厨房电器制造有限公司 等) 2020年 5月 5日 (2020 - 05 - 05) 说明书第[0063]-[0082]段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2018293982 A1 (YUTOU TECHNOLOGY HANGZHOU CO., LTD.) 2018年 10月 11日 (2018 - 10 - 11) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 111694605 A (OPPO广东移动通信有限公司) 2020年 9月 22日 (2020 - 09 - 22) 权利要求1-17	1-20	X	CN 103558916 A (百度在线网络技术北京有限公司) 2014年 2月 5日 (2014 - 02 - 05) 说明书第[0050]-[0130]、[0228]-[0231]段、图1	1-20	X	CN 107978316 A (西安蜂语信息科技有限公司) 2018年 5月 1日 (2018 - 05 - 01) 说明书第[0105]-[0182]段、图1	1-20	X	CN 111096680 A (广东美的厨房电器制造有限公司 等) 2020年 5月 5日 (2020 - 05 - 05) 说明书第[0063]-[0082]段	1-20	A	US 2018293982 A1 (YUTOU TECHNOLOGY HANGZHOU CO., LTD.) 2018年 10月 11日 (2018 - 10 - 11) 全文	1-20
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 111694605 A (OPPO广东移动通信有限公司) 2020年 9月 22日 (2020 - 09 - 22) 权利要求1-17	1-20																		
X	CN 103558916 A (百度在线网络技术北京有限公司) 2014年 2月 5日 (2014 - 02 - 05) 说明书第[0050]-[0130]、[0228]-[0231]段、图1	1-20																		
X	CN 107978316 A (西安蜂语信息科技有限公司) 2018年 5月 1日 (2018 - 05 - 01) 说明书第[0105]-[0182]段、图1	1-20																		
X	CN 111096680 A (广东美的厨房电器制造有限公司 等) 2020年 5月 5日 (2020 - 05 - 05) 说明书第[0063]-[0082]段	1-20																		
A	US 2018293982 A1 (YUTOU TECHNOLOGY HANGZHOU CO., LTD.) 2018年 10月 11日 (2018 - 10 - 11) 全文	1-20																		
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2021年 4月 15日</p>		<p>国际检索报告邮寄日期</p> <p>2021年 4月 27日</p>																		
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>张琳琳</p> <p>电话号码 86-(10)-53961404</p>																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/081332

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	111694605	A	2020年 9月 22日	无			
CN	103558916	A	2014年 2月 5日	WO	2015066949	A1	2015年 5月 14日
CN	107978316	A	2018年 5月 1日	无			
CN	111096680	A	2020年 5月 5日	无			
US	2018293982	A1	2018年 10月 11日	JP	2018533139	A	2018年 11月 8日
				TW	201714168	A	2017年 4月 16日
				CN	106572418	A	2017年 4月 19日
				EP	3361712	A1	2018年 8月 15日
				WO	2017059812	A1	2017年 4月 13日