

(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
2017 年 5 月 4 日 (04.05.2017)



(10) 国际公布号
WO 2017/071070 A1

- (51) 国际专利分类号:
G05B 15/02 (2006.01)
- (21) 国际申请号: PCT/CN2015/099361
- (22) 国际申请日: 2015 年 12 月 29 日 (29.12.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510712870.6 2015 年 10 月 28 日 (28.10.2015) CN
- (71) 申请人: 小米科技有限责任公司 (XIAOMI INC.)
[CN/CN]; 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层, Beijing 100085 (CN)。
- (72) 发明人: 高斯太 (GAO, Sitai); 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层由小米科技有限责任公司转交, Beijing 100085 (CN)。 丁一 (DING, Yi); 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层由小米科技有限责任公司转交, Beijing 100085 (CN)。 侯恩星 (HOU, Enxing); 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层由小米科技有限责任公司转交, Beijing 100085 (CN)。
- (74) 代理人: 北京律智知识产权代理有限公司 (BEIJING INTELLEGAL INTELLECTUAL PROP-

ERTY AGENT LTD.); 中国北京市朝阳区慧忠路 5 号远大中心 B 座 1802, 1803, 1805, Beijing 100101 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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。

(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))。

(54) Title: SPEECH CONTROL METHOD AND APPARATUS FOR SMART DEVICE, CONTROL DEVICE AND SMART DEVICE

(54) 发明名称: 智能设备的语音控制方法、装置、控制设备及智能设备

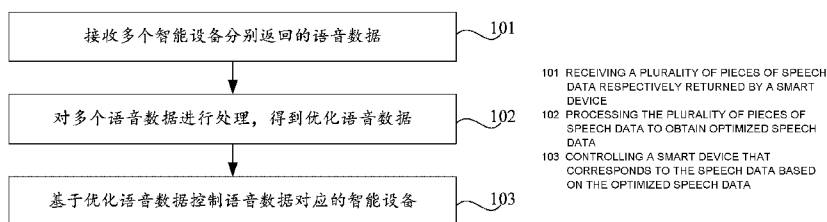


图 1

(57) Abstract: A speech control method and apparatus for a smart device, a control device and a smart device. The method comprises: receiving a plurality of pieces of speech data respectively returned by a smart device (101); processing the plurality of pieces of speech data to obtain optimized speech data (102); and controlling a smart device that corresponds to the speech data based on the optimized speech data (103). By means of the method, the control device can obtain optimized speech data by processing speech data from different positions and control a smart device corresponding to the optimized speech data based on the optimized speech data, thereby realizing speech control of the smart device, providing convenience for users in controlling the smart device and optimizing the user experience.

(57) 摘要: 一种智能设备的语音控制方法、装置、控制设备及智能设备, 方法包括: 接收多个智能设备分别返回的语音数据 (101); 对多个语音数据进行处理, 得到优化语音数据 (102); 基于优化语音数据控制语音数据对应的智能设备 (103)。应用本方法, 控制设备可以通过对来自不同位置处的语音数据进行处理, 得到优化语音数据, 并基于优化语音数据控制优化语音数据对应的智能设备, 从而实现了智能设备的语音控制, 为用户控制智能设备提供了方便, 优化了用户体验。



WO 2017/071070 A1

智能设备的语音控制方法、装置、控制设备及智能设备

本申请基于申请号为 201510712870.6、申请日为 2015 年 10 月 28 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此引入本申请作为参考。

技术领域

本公开涉及智能家居技术领域，尤其涉及一种智能设备的语音控制方法、装置、控制设备及智能设备。

背景技术

目前的语音识别技术，采用性能最好的处理芯片配合全向麦克风，在比较理想的情况下能够达到三米的识别距离。通常大型会议室中通过在会议室的不同位置设置多个麦克风，并对各麦克风收集的声音进行处理，来实现较好的语音识别效果。

相关技术中，对于家庭环境中的语音识别，如果要想实现上述效果需要根据家具及电器的摆设，在不同位置设置多个麦克风，成本较高。

发明内容

为克服相关技术中存在的问题，本公开提供了一种智能设备的语音控制方法、装置、控制设备及智能设备。

根据本公开实施例的第一方面，提供一种智能设备的语音控制方法，包括：
接收多个智能设备分别返回的语音数据；
对多个所述语音数据进行处理，得到优化语音数据；
基于所述优化语音数据控制所述语音数据对应的智能设备。

可选的，所述接收多个智能设备分别返回的语音数据，包括：
接收位于不同位置的多个智能设备分别返回的语音数据。

可选的，所述接收位于不同位置的多个智能设备返回的语音数据之前，所述方法还包括：

读取应用绑定的智能设备的基本信息；

基于所述基本信息确定具有录音功能的智能设备；

从具有录音功能的智能设备中确定待启动的智能设备；

向所述确定为待启动的智能设备发送开启指令，以开启所述智能设备。

可选的，所述从具有录音功能的智能设备中确定待启动的智能设备，包括：

显示具有录音功能的智能设备的列表；

基于用户针对所述列表的选择操作，将用户选中的智能设备确定为待启动录音功能的

智能设备。

可选的，所述从具有录音功能的智能设备中确定待启动的智能设备，包括：

基于定位技术确定用户的位置；

查找预先存储的各所述具有录音功能的智能设备的位置；

- 5 将位于包括所述用户的位置在内的设定范围内的具有录音功能的智能设备，确定为待启动的智能设备。

可选的，所述方法还包括：

在位于所述设定范围内的具有录音功能的智能设备的个数超过设定阈值时，读取位于所述设定范围内的具有录音功能的智能设备的历史使用数据；

- 10 基于所述历史使用数据确定待启动的智能设备。

可选的，所述历史使用数据包括：使用频率、最近一次使用时间、使用总时长中的任一项或多项。

可选的，所述对多个所述语音数据进行处理，得到优化语音数据，包括：

基于波束成形技术对多个所述语音数据进行处理，得到优化语音数据。

- 15 可选的，所述基于所述优化语音数据控制所述语音数据对应的智能设备，包括：

将所述优化语音数据发送给红外遥控设备，以使所述红外遥控设备基于所述优化语音数据中的语音信息查找对应的控制指令，基于所述优化语音数据中的设备名称查找对应的红外编码，并向所述红外编码发送所述控制指令。

可选的，所述基于所述优化语音数据控制所述语音数据对应的智能设备，包括：

- 20 将所述优化语音数据发送给服务器，以使所述服务器基于所述优化语音数据中的语音信息查找对应的控制指令，并将所述控制指令和所述优化语音数据中的设备名称发送给红外遥控设备，进一步使所述红外遥控设备将所述控制指令发送给所述设备名称对应的红外编码。

根据本公开实施例的第二方面，提供一种智能设备的语音控制方法，包括：

- 25 采集语音数据；

将所述语音数据发送给控制设备，以使所述控制设备基于所述语音数据及多个位于其他位置的智能设备采集的所述语音数据，控制所述语音数据对应的智能设备。

可选的，所述采集语音数据之前，所述方法包括：

基于所述控制设备发送的开启指令开启。

- 30 根据本公开实施例的第三方面，提供一种智能设备的语音控制装置，包括：接收模块、处理模块和控制模块；

所述接收模块，被配置为接收多个智能设备分别返回的语音数据；

所述处理模块，被配置为对所述接收模块接收到的多个所述语音数据进行处理，得到优化语音数据；

- 35 所述控制模块，被配置为基于所述处理模块得到的所述优化语音数据控制所述语音数

据对应的智能设备。

可选的，所述接收模块包括：接收子模块；

所述接收子模块，被配置为接收位于不同位置的多个智能设备分别返回的语音数据。

可选的，所述装置还包括：读取模块、第一确定模块、第二确定模块和开启模块；

5 所述读取模块，被配置为读取应用绑定的智能设备的基本信息；

所述第一确定模块，被配置为基于所述读取模块读取的基本信息确定具有录音功能的智能设备；

所述第二确定模块，被配置为从所述第一确定模块确定的具有录音功能的智能设备中确定待启动的智能设备；

10 所述开启模块，被配置为向所述第二确定模块确定为待启动的智能设备发送开启指令，以开启所述智能设备。

可选的，所述第二确定模块包括：显示子模块和第一确定子模块；

所述显示子模块，被配置为显示具有录音功能的智能设备的列表；

15 所述显示子模块，被配置为基于用户针对所述显示子模块所显示的列表中的选择操作，将用户选中的智能设备确定为待启动录音功能的智能设备。

可选的，所述第二确定模块包括：定位子模块、查找子模块和第二确定子模块；

所述定位子模块，被配置为基于定位技术确定用户的位置；

所述查找子模块，被配置为查找预先存储的各所述具有录音功能的智能设备的位置；

20 所述第二确定子模块，被配置为将位于包括所述定位子模块定位到的用户的位置在内的设定范围内的具有录音功能的智能设备，确定为待启动的智能设备。

可选的，所述第二确定模块还包括：读取子模块和第三确定子模块；

所述读取子模块，被配置为在所述第二确定子模块所确定的所述智能设备的个数超过设定阈值时，读取位于所述设定范围内的具有录音功能的智能设备的历史使用数据；

25 所述第三确定子模块，被配置为基于所述读取子模块读取的历史使用数据确定待启动的智能设备。

可选的，所述读取子模块读取的所述历史使用数据包括：使用频率、最近一次使用时间、使用总时长中的任一项或多项。

可选的，所述处理模块包括：处理子模块；

30 所述处理子模块，被配置为基于波束成形技术对所述接收模块接收的多个所述语音数据进行优化处理，得到优化语音数据。

可选的，所述控制模块包括：第一发送子模块；

35 所述第一发送子模块，被配置为将所述处理模块得到的所述优化语音数据发送给红外遥控设备，以使所述红外遥控设备基于所述优化语音数据中的语音信息查找对应的控制指令，基于所述优化语音数据中的设备名称查找对应的红外编码，并向所述红外编码发送所述控制指令。

可选的，所述控制模块包括：第二发送子模块；

所述第二发送子模块，被配置为将所述处理模块得到的优化语音数据发送给服务器，以使所述服务器基于所述优化语音数据中的语音信息查找对应的控制指令，并将所述控制指令和所述优化语音数据中的设备名称发送给红外遥控设备，进一步使所述红外遥控设备将所述控制指令发送给所述设备名称对应的红外编码。

根据本公开实施例的第四方面，提供一种智能设备的语音控制装置，包括：采集模块和发送模块；

所述采集模块，被配置为采集语音数据；

所述发送模块，被配置为将所述采集模块采集的语音数据发送给控制设备，以使所述控制设备基于所述语音数据控制所述语音数据对应的智能设备。

可选的，所述装置还包括：开启模块；

所述开启模块，被配置为基于所述控制设备发送的开启指令开启。

根据本公开实施例的第五方面，提供一种控制设备，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

接收多个智能设备分别返回的语音数据；

对多个所述语音数据进行处理，得到优化语音数据；

基于所述优化语音数据控制所述语音数据对应的智能设备。

根据本公开实施例的第六方面，提供一种智能设备，其特征在于，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

采集语音数据；

将所述语音数据发送给控制设备，以使所述控制设备基于所述语音数据及多个位于其他位置的智能设备采集的所述语音数据，控制所述语音数据对应的智能设备。

本公开的实施例提供的技术方案可以包括以下有益效果：

本公开中控制设备可以通过对来自不同位置处的语音数据进行处理，得到优化语音数据，并基于该优化语音数据控制该优化语音数据对应的智能设备，从而实现了智能设备的语音控制，为用户控制智能设备提供了方便，优化了用户体验。

本公开中控制设备可以接收位于多个位置的智能设备采集的语音数据，对来自不同位置处的语音数据进行处理得到优化语音数据，因而保证了优化语音数据的质量，保证了语音识别的准确度，实现了对智能设备的语音控制。

本公开中控制设备可以先确定具有录音功能的智能设备，然后从具有录音功能的智能设备中确定待启动的智能设备，从而保证启动的智能设备能够录音。

本公开中控制设备可以基于用户的选择来确定待启动的智能设备，这种方式更加符合用户的需求，提高了用户体验。

本公开中控制设备可以对用户进行定位，然后基于用户的位置确定待启动的智能设备，这种方式能够保证待启动的智能设备距离用户的位置较近，从而保证采集到更为清晰

的语音数据，提高后续优化语音数据的辨识度，保证对智能设备的精确控制。

本公开中控制设备还可以结合智能设备的历史使用数据来确定待启动智能设备，由于历史使用数据能够一定程度上反映智能设备的性能，因而能够保证所录制的语音数据的质量。

5 本公开中控制设备可以将优化语音数据发送给红外遥控中心，由红外遥控中心实现对智能设备的精确语音控制。

本公开中智能设备可以将采集到的语音数据发送给控制设备，以使控制设备能够基于语音数据，及多个位于其他位置的智能设备采集的语音数据来控制语音数据对应的智能设备，从而实现了智能设备的精确语音控制，优化了用户体验。

10 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本公开。

附图说明

15 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

图 1 是本公开根据一示例性实施例示出的一种智能设备的语音控制方法流程图。

图 2 是本公开根据一示例性实施例示出的另一种智能设备的语音控制方法流程图。

图 3 是本公开根据一示例性实施例示出的另一种智能设备的语音控制方法流程图。

图 4 是本公开根据一示例性实施例示出的一种智能设备的语音控制应用场景示意图。

20 图 5 是本公开根据一示例性实施例示出的一种智能设备的语音控制装置框图。

图 6 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 7 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 8 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 9 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

25 图 10 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 11 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 12 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 13 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 14 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

30 图 15 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图。

图 16 是本公开根据一示例性实施例示出的一种用于智能设备的语音控制装置的一结构示意图。

图 17 是本公开根据一示例性实施例示出的一种用于智能设备的语音控制装置的一结构示意图。

35

具体实施方式

这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反，它们仅是与如所附权利要求要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

在本公开使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本公开。在本公开和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

应当理解，尽管在本公开可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本公开范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

如图1所示，图1是根据一示例性实施例示出的一种智能设备的语音控制方法流程图，该方法可以用于终端等控制设备中，包括以下步骤：

步骤101、接收多个智能设备分别返回的语音数据。

本公开中的终端可以是任何具有上网功能的智能终端，例如，可以具体为手机、平板电脑、PDA（Personal Digital Assistant，个人数字助理）等。其中，终端可以通过无线局域网接入路由器，并通过路由器访问公网上的服务器。本公开中的终端可以通过 App（Application，应用程序）来接收语音数据，例如智能家庭 App。

本公开中的智能设备包括智能家电、可穿戴设备等，智能设备都具有通信模块，例如 wifi（Wireless Fidelity，无线保真）模块，用于通过家中设置的路由器与终端、控制中心进行通信。其中控制中心可以为红外遥控中心，用于对各智能设备进行控制。

步骤102、对多个语音数据进行处理，得到优化语音数据。

本公开步骤中，采用波束成形技术对多个语音数据进行处理来得到优化语音数据。

步骤103、基于优化语音数据控制语音数据对应的智能设备。

上述实施例中，控制设备可以通过对来自不同位置处的语音数据进行处理，得到优化语音数据，并基于该优化语音数据控制该优化语音数据对应的智能设备，从而实现了智能设备的语音控制，而且由于优化语音数据是对来自不同位置处的语音数据进行处理而得到的，因而保证了优化语音数据的质量，保证了语音识别的准确度，实现了对智能设备的语音控制。

如图2所示，图2是根据一示例性实施例示出的另一种智能设备的语音控制方法流程图，该方法可以用于终端等控制设备中，包括以下步骤：

步骤201、读取应用绑定的智能设备的基本信息。

步骤 202、基于基本信息确定具有录音功能的智能设备。

本公开实施例中，终端可以基于智能家庭 App 来接收语音数据和控制智能设备，智能家庭 App 中关联有多个智能设备，且存储有各个智能设备的基本信息，终端通过读取基本信息能够确定哪些智能设备具有录音功能。

5 步骤 203、从具有录音功能的智能设备中确定待启动的智能设备。

本公开实施例中，终端需要确定待启动的智能设备，即采集语音数据的智能设备。

在一种公开方式中，终端的智能家庭 App 可以显示所绑定的智能设备中全部具有录音功能的智能设备，供用户选择，用户可以选择部分也可以选择全部作为待启动的智能设备，然后终端基于用户的选择操作，将用户选中的智能设备确定为待启动录音功能的智能设备。

10 在另一种公开方式中，由终端来确定要加入开启列表的智能设备。终端可以基于定位技术确定用户的位置；然后查找预先存储的各具有录音功能的智能设备的位置；然后终端将位于包括用户的位置在内的设定范围内的具有录音功能的智能设备，确定为待启动的智能设备。

15 例如终端定位到用户位于客厅中，基于设定范围，以用户的位置为中心半径为 2m 作圆的范围，则将该范围之内的具有录音功能的智能设备确定为待启动的智能设备。

或者，在终端的 App 中，存储有各个智能设备的放置位置，例如设备 1-4 放置在客厅，设备 5 和 6 放置在主卧，设备 7 和 8 放置在次卧。当终端确定用户位于客厅中，则确定设备 1-4 为待启动的智能设备。

20 并且，在本公开实施例的方式中，如果所确定的位于设定范围内的具有录音功能的智能设备的个数超过设定阈值（例如：6 个）时，读取位于设定范围内的具有录音功能的智能设备的历史使用数据；然后基于历史使用数据确定待启动的智能设备。该历史使用数据可以包括：使用频率、最近一次使用时间、使用总时长中的任一项或多项。例如，终端可以基于使用频率对智能设备进行排序，将使用频率最高的前 4 个智能设备确定为待启动的智能设备。

25 步骤 204、向确定为待启动的智能设备发送开启指令，以开启智能设备。

本公开步骤中，开启了智能设备可以理解为开启了该智能设备的录音功能。终端可以通过智能家庭 App 开启待启动的智能设备。

步骤 205、接收多个不同位置的智能设备采集的语音数据。

30 步骤 206、基于波束成形技术对多个语音数据进行处理，得到优化语音数据。

本公开步骤中，具体的处理可以包括：回声消除、信号处理以及加强处理等。

步骤 207、将优化语音数据发送给红外遥控设备，以使红外遥控设备基于优化语音数据中的语音信息查找对应的控制指令，基于优化语音数据中的设备名称查找对应的红外编码，并向红外编码发送控制指令。

35 本公开实施例中，红外遥控设备可以提取优化语音数据中的语音信息，例如“开启电

视”，并从预存储的数据中查找对应的控制指令；此外，优化语音数据中还携带有设备名称，例如“电视”，红外遥控设备可以查找对应的红外编码，将控制指令发送给该红外编码，从而实现了智能设备的语音控制。

在另一种公开方式中，终端还可以将该优化语音数据发送给服务器，以使服务器基于该优化语音数据中的语音信息查找对应的控制指令，并将控制指令和优化语音数据中的设备名称发送给红外遥控设备，进一步使红外遥控设备将控制指令发送给设备名称对应的红外编码。

上述实施例中，终端可以通过多种方式确定待启动的智能设备，例如，将用户选择的智能设备确定为智能设备，这种方式能够提高录音的满意度，优化用户体验。

终端还可以通过定位用户的位置来确定待启动的智能设备，这种方式能够确定出离用户最近的智能设备，从而提高语音录制的效果，保证优化语音数据的质量和识别度。

如图3所示，图3是根据一示例性实施例示出的另一种智能设备的语音控制方法流程图，该方法可以用于终端等控制设备中，包括以下步骤：

步骤301、采集语音数据。

本公开实施例中，由位于不同位置的多个具有录音功能的智能设备分别采集不同位置处的语音数据。智能设备可以基于控制设备发送的开启指令开启。

步骤302、将语音数据发送给控制设备，以使控制设备基于该语音数据及多个位于其他位置的智能设备采集的语音数据，控制语音数据对应的智能设备。

结合图1、2所示实施例，在一种公开方式中，智能设备可以将语音数据发送给控制设备，例如终端，终端将语音数据进行波束成形处理，得到优化语音数据，并将优化语音数据发送给红外遥控中心，红外遥控中心中存储有红外编码库，红外编码库中存储有各个智能设备的名称及对应的红外编码，还存储有语音信息与对应的控制指令。红外遥控中心提取优化语音数据中的语音信息，得到设备名称，查询红外编码库得到对应于设备名称的红外编码，并且基于语音信息查询到对应的控制指令，然后将该控制指令发送给智能设备的红外编码，以实现对该智能设备的语音控制。

在另一种公开方式中，智能设备将优化语音数据发送给终端之后，终端还可以将优化语音数据发送给服务器，服务器中存储有语音信息以及对应的控制指令，服务器基于优化语音数据提取到语音信息，查询得到对应的控制指令，服务器将控制指令以及设备名称发送给红外遥控器，红外遥控基于设备名称查找到对应的红外编码，并将控制指令发送给对应的红外编码，以实现对智能设备的语音控制。

上述实施例中，智能设备可以将采集的语音数据发送给终端，以便终端对该语音数据进行处理，得到优化语音数据，并基于该优化语音数据对智能设备进行控制，提高了优化语音数据的质量和识别度，优化了用户体验。

如图4所示，图4是本公开根据一示例性实施例示出的一种智能设备的语音控制应用场景示意图。在图4所示的场景中，包括：作为控制设备的智能手机，用于录音的智能设

备 1、智能设备 2 及智能设备 3，以及作为控制的对象的电视，智能手机中安装有智能家庭 App，智能手机通过智能家庭 App 控制所绑定的各个智能设备。

5 智能手机基于定位技术确定用户的位置，并查找预先存储的各具有录音功能的智能设备的位置；将位于包括用户的位置在内的设定范围（图 4 中所示的圆形范围）内的具有录音功能的智能设备 1、智能设备 2 和智能设备 3，确定为待启动的智能设备，然后开启待启动设备开始录音，接收位于不同位置的智能设备 1、智能设备 2 及智能设备 3 采集的语音数据，并对接收到的语音数据进行波束成形处理，得到优化语音数据“开启电视”，智能手机将该优化语音数据发送给红外遥控中心，以使红外遥控中心基于优化语音数据中的设备名称“电视”查找对应的红外编码，基于优化语音数据查找控制指令，并将控制指令
10 发送给红外编码，实现对电视的控制。

在图 4 所示应用场景中，实现智能设备的语音控制的具体过程可以参见前述对图 1-3 中的描述，在此不再赘述。

与前述智能设备的语音控制方法实施例相对应，本公开还提供了智能设备的语音控制装置及其所应用的控制设备和智能设备的实施例。

15 如图 5 所示，图 5 是本公开根据一示例性实施例示出的一种智能设备的语音控制装置框图，该装置可以包括：接收模块 510、处理模块 520 和控制模块 530。

其中，接收模块 510，被配置为接收多个智能设备分别返回的语音数据；

处理模块 520，被配置为对接收模块 510 接收到的多个语音数据进行处理，得到优化语音数据；

20 控制模块 530，被配置为基于处理模块 520 得到的优化语音数据控制语音数据对应的智能设备。

上述实施例中，控制设备可以通过对来自不同位置处的语音数据进行处理，得到优化语音数据，并基于该优化语音数据控制该优化语音数据对应的智能设备，从而实现了
对智能设备的语音控制，为用户控制智能设备提供了方便，优化了用户体验。

25 如图 6 所示，图 6 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该实施例在前述图 5 所示实施例的基础上，接收模块 510 可以包括：接收子模块 511。

接收子模块 511，被配置为接收位于不同位置的多个智能设备分别返回的语音数据。

上述实施例中，控制设备可以接收位于多个位置的智能设备采集的语音数据，对来自
30 不同位置处的语音数据进行处理得到优化语音数据，因而保证了优化语音数据的质量，保证了语音识别的准确度，实现了对智能设备的语音控制。

如图 7 所示，图 7 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该实施例在前述图 5 所示实施例的基础上，该装置还可以包括：读取模块 540、
第一确定模块 550、第二确定模块 560 和开启模块 570。

35 读取模块 540，被配置为读取应用绑定的智能设备的基本信息；

第一确定模块 550, 被配置为基于读取模块 540 读取的基本信息确定具有录音功能的智能设备;

第二确定模块 560, 被配置为从第一确定模块 550 确定的具有录音功能的智能设备中确定待启动的智能设备;

5 开启模块 570, 被配置为向第二确定模块 560 确定为待启动的智能设备发送开启指令, 以开启智能设备。

上述实施例中, 控制设备可以先确定具有录音功能的智能设备, 然后从具有录音功能的智能设备中确定待启动的智能设备, 从而保证启动的智能设备能够录音。

10 如图 8 所示, 图 8 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图, 该实施例在前述图 7 所示实施例的基础上, 该第二确定模块 560 可以包括: 显示子模块 561 和第一确定子模块 562。

其中, 显示子模块 561, 被配置为显示具有录音功能的智能设备的列表;

第一确定子模块 562, 被配置为基于用户在所述显示子模块 561 显示的列表中的选择操作, 将用户选中的智能设备确定为待启动录音功能的智能设备。

15 上述实施例中, 控制设备可以基于用户的选择来确定待启动的智能设备, 这种方式更加符合用户的需求, 提高了用户体验。

如图 9 所示, 图 9 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图, 该实施例在前述图 7 所示实施例的基础上, 第二确定模块 560 可以包括: 定位子模块 563、查找子模块 564 和第二确定子模块 565。

20 定位子模块 563, 被配置为基于定位技术确定用户的位置;

查找子模块 564, 被配置为查找预先存储的各具有录音功能的智能设备的位置;

第二确定子模块 565, 被配置为将位于包括定位子模块 563 定位到的用户的位置在内的设定范围内的具有录音功能的智能设备, 确定为待启动的智能设备。

25 上述实施例中, 控制设备可以对用户进行定位, 然后基于用户的位置确定待启动的智能设备, 这种方式能够保证待启动的智能设备距离用户的位置较近, 从而保证采集到更为清晰的语音数据, 提高后续优化语音数据的辨识度, 保证对智能设备的精确控制。

如图 10 所示, 图 10 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图, 该实施例在前述图 9 所示实施例的基础上, 第二确定模块 560 还可以包括: 读取子模块 566 和第三确定子模块 567。

30 读取子模块 566, 被配置为在第二确定子模块 565 所确定的智能设备的个数超过设定阈值时, 读取位于设定范围内的具有录音功能的智能设备的历史使用数据;

第三确定子模块 567, 被配置为基于读取子模块 566 读取的历史使用数据确定待启动的智能设备。

35 其中读取子模块 566 读取的历史使用数据包括: 使用频率、最近一次使用时间、使用总时长中的任一项或多项。

上述实施例中，控制设备还可以结合智能设备的历史使用数据来确定待启动智能设备，由于历史使用数据能够一定程度上反映智能设备的性能，因而能够保证所录制的语音数据的质量。

5 如图 11 所示，图 11 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该实施例在前述图 5 所示实施例的基础上，处理模块 520 可以包括：处理子模块 521。

处理子模块 521，被配置为基于波束成形技术对接收模块 510 接收的多个语音数据进行处理，得到优化语音数据。

10 上述实施例中，控制设备可以基于波束成形技术对多个语音数据进行处理，进一步提高语音识别成功率。

如图 12 所示，图 12 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该实施例在前述图 5 所示实施例的基础上，控制模块 530 可以包括：第一发送子模块 531。

15 第一发送子模块 531，被配置为将处理模块 520 得到的优化语音数据发送给红外遥控设备，以使红外遥控设备基于优化语音数据中的语音信息查找对应的控制指令，基于优化语音数据中的设备名称查找对应的红外编码，并向红外编码发送控制指令。

上述实施例中，控制设备可以将优化语音数据发送给红外遥控中心，由红外遥控中心实现对智能设备的精确语音控制。

20 如图 13 所示，图 13 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该实施例在前述图 5 所示实施例的基础上，控制模块 530 可以包括：第二发送子模块 532。

25 第二发送子模块 532，被配置为将处理模块得到的优化语音数据发送给服务器，以使服务器基于优化语音数据中的语音信息查找对应的控制指令，并将控制指令和优化语音数据中的设备名称发送给红外遥控设备，进一步使红外遥控设备将控制指令发送给设备名称对应的红外编码。

上述实施例中，控制设备可以将优化语音数据发送给服务器，由服务器及红外遥控中心实现对智能设备的精确语音控制。

上述图 5 至图 13 示出的智能设备的语音控制装置实施例可以应用在控制设备中。

30 如图 14 所示，图 14 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该装置可以应用于智能设备中，该装置可以包括：采集模块 610 和发送模块 620。

采集模块 610，被配置为采集语音数据；

发送模块 620，被配置为将采集模块 610 采集的语音数据发送给控制设备，以使控制设备基于语音数据控制语音数据对应的智能设备。

35 上述实施例中，智能设备可以将采集到的语音数据发送给控制设备，以使控制设备能够基于语音数据，及多个位于其他位置的智能设备采集的语音数据来控制语音数据对应的

智能设备，从而实现了智能设备的精确语音控制，优化了用户体验。

如图 15 所示，图 15 是本公开根据一示例性实施例示出的另一种智能设备的语音控制装置框图，该实施例在前述图 14 所示实施例的基础上，该装置还可以包括：开启模块 630。

开启模块 630，被配置为基于控制设备发送的开启指令开启。

5 上述图 14-15 示出的智能设备的语音控制装置实施例可以应用在用于采集语音数据的智能设备中。

上述装置中各个单元的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程，在此不再赘述。

10 对于装置实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本公开方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

15 相应的，本公开还提供一种控制设备，所述控制设备包括有处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

接收多个智能设备分别返回的语音数据；

对多个所述语音数据进行处理，得到优化语音数据；

基于所述优化语音数据控制所述语音数据对应的智能设备。

20 相应的，本公开还提供一种智能设备，所述智能设备包括有处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

采集语音数据；

将所述语音数据发送给控制设备，以使所述控制设备基于所述语音数据及多个位于其他位置的智能设备采集的所述语音数据，控制所述语音数据对应的智能设备。

25 如图 16 所示，图 16 是本公开根据一示例性实施例示出的一种用于智能设备的语音控制装置 1600 的一结构示意图（控制设备侧）。例如，装置 1600 可以是具有路由功能的移动电话，计算机，数字广播终端，消息收发设备，游戏控制台，平板设备，医疗设备，健身设备，个人数字助理等。

30 参照图 16，装置 1600 可以包括以下一个或多个组件：处理组件 1602，存储器 1604，电源组件 1606，多媒体组件 1608，音频组件 1610，输入/输出（I/O）的接口 1612，传感器组件 1614，以及通信组件 1616。

35 处理组件 1602 通常控制装置 1600 的整体操作，诸如与显示，电话呼叫，数据通信，相机操作和记录操作相关联的操作。处理组件 1602 可以包括一个或多个处理器 1620 来执行指令，以完成上述的方法的全部或部分步骤。此外，处理组件 1602 可以包括一个或多个模块，便于处理组件 1602 和其他组件之间的交互。例如，处理组件 1602 可以包括多媒

体模块，以方便多媒体组件 1608 和处理组件 1602 之间的交互。

存储器 1604 被配置为存储各种类型的数据以支持在装置 1600 的操作。这些数据的示例包括用于在装置 1600 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 1604 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。

电源组件 1606 为装置 1600 的各种组件提供电力。电源组件 1606 可以包括电源管理系统，一个或多个电源，及其他与为装置 1600 生成、管理和分配电力相关联的组件。

多媒体组件 1608 包括在所述装置 1600 和用户之间的提供一个输出接口的屏幕。在一些实施例中，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。所述触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中，多媒体组件 1608 包括一个前置摄像头和/或后置摄像头。当装置 1600 处于操作模式，如拍摄模式或视频模式时，前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

音频组件 1610 被配置为输出和/或输入音频信号。例如，音频组件 1610 包括一个麦克风(MIC)，当装置 1600 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 1604 或经由通信组件 1616 发送。在一些实施例中，音频组件 1610 还包括一个扬声器，用于输出音频信号。

I/O 接口 1612 为处理组件 1602 和外围接口模块之间提供接口，上述外围接口模块可以是键盘，点击轮，按钮等。这些按钮可包括但不限于：主页按钮、音量按钮、启动按钮和锁定按钮。

传感器组件 1614 包括一个或多个传感器，用于为装置 1600 提供各个方面的状态评估。例如，传感器组件 1614 可以检测到装置 1600 的打开/关闭状态，组件的相对定位，例如所述组件为装置 1600 的显示器和小键盘，传感器组件 1614 还可以检测装置 1600 或装置 1600 一个组件的位置改变，用户与装置 1600 接触的存在或不存在，装置 1600 方位或加速/减速和装置 1600 的温度变化。传感器组件 1614 可以包括接近传感器，被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件 1614 还可以包括光传感器，如 CMOS 或 CCD 图像传感器，用于在成像应用中使用。在一些实施例中，该传感器组件 1614 还可以包括加速度传感器，陀螺仪传感器，磁传感器，压力传感器，微波传感器或温度传感器。

通信组件 1616 被配置为便于装置 1600 和其他设备之间有线或无线方式的通信。装置

1600 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 1616 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，所述通信组件 1616 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。

在示例性实施例中，装置 1600 可以被一个或多个应用专用集成电路（ASIC）、数字信号处理器（DSP）、数字信号处理设备（DSPD）、可编程逻辑器件（PLD）、现场可编程门阵列（FPGA）、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

在示例性实施例中，还提供了一种包括指令的非临时性计算机可读存储介质，例如包括指令的存储器 1604，上述指令可由装置 1600 的处理器 1620 执行以完成上述方法。例如，所述非临时性计算机可读存储介质可以是 ROM、随机存取存储器（RAM）、CD-ROM、磁带、软盘和光数据存储设备等。

如图 17 所示，图 17 是本公开根据一示例性实施例示出的一种用于智能设备的语音控制装置 1700 的一结构示意图（智能设备侧）。例如，装置 1700 可以是具有路由功能的移动电话，计算机，数字广播终端，消息收发设备，游戏控制台，平板设备，医疗设备，健身设备，个人数字助理等。

参照图 17，装置 1700 可以包括以下一个或多个组件：处理组件 1702，存储器 1704，电源组件 1706，多媒体组件 1708，音频组件 1710，输入/输出（I/O）的接口 1712，传感器组件 1714，以及通信组件 1716。

处理组件 1702 通常控制装置 1700 的整体操作，诸如与显示，电话呼叫，数据通信，相机操作和记录操作相关联的操作。处理组件 1702 可以包括一个或多个处理器 1720 来执行指令，以完成上述的方法的全部或部分步骤。此外，处理组件 1702 可以包括一个或多个模块，便于处理组件 1702 和其他组件之间的交互。例如，处理组件 1702 可以包括多媒体模块，以方便多媒体组件 1708 和处理组件 1702 之间的交互。

存储器 1704 被配置为存储各种类型的数据以支持在装置 1700 的操作。这些数据的示例包括用于在装置 1700 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 1704 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。

电源组件 1706 为装置 1700 的各种组件提供电力。电源组件 1706 可以包括电源管理系统，一个或多个电源，及其他与为装置 1700 生成、管理和分配电力相关联的组件。

多媒体组件 1708 包括在所述装置 1700 和用户之间的提供一个输出接口的屏幕。在一些实施例中，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。所述触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中，多媒体组件 1708 包括一个前置摄像头和/或后置摄像头。当装置 1700 处于操作模式，
5 如拍摄模式或视频模式时，前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

音频组件 1710 被配置为输出和/或输入音频信号。例如，音频组件 1710 包括一个麦克风（MIC），当装置 1700 处于操作模式，如呼叫模式、记录模式和语音识别模式时，
10 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 1704 或经由通信组件 1716 发送。在一些实施例中，音频组件 1710 还包括一个扬声器，用于输出音频信号。

I/O 接口 1712 为处理组件 1702 和外围接口模块之间提供接口，上述外围接口模块可以是键盘，点击轮，按钮等。这些按钮可包括但不限于：主页按钮、音量按钮、启动按钮和锁定按钮。

15 传感器组件 1714 包括一个或多个传感器，用于为装置 1700 提供各个方面的状态评估。例如，传感器组件 1714 可以检测到装置 1700 的打开/关闭状态，组件的相对定位，例如所述组件为装置 1700 的显示器和小键盘，传感器组件 1714 还可以检测装置 1700 或装置 1700 一个组件的位置改变，用户与装置 1700 接触的存在或不存在，装置 1700 方位或加速/减速和装置 1700 的温度变化。传感器组件 1714 可以包括接近传感器，被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件 1714 还可以包括光传感器，如
20 CMOS 或 CCD 图像传感器，用于在成像应用中使用。在一些实施例中，该传感器组件 1714 还可以包括加速度传感器，陀螺仪传感器，磁传感器，压力传感器，微波传感器或温度传感器。

通信组件 1716 被配置为便于装置 1700 和其他设备之间有线或无线方式的通信。装置
25 1700 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 1716 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，所述通信组件 1716 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。

30 在示例性实施例中，装置 1700 可以被一个或多个应用专用集成电路（ASIC）、数字信号处理器（DSP）、数字信号处理设备（DSPD）、可编程逻辑器件（PLD）、现场可编程门阵列（FPGA）、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

在示例性实施例中，还提供了一种包括指令的非临时性计算机可读存储介质，例如包
35 括指令的存储器 1704，上述指令可由装置 1700 的处理器 1720 执行以完成上述方法。例如，

所述非临时性计算机可读存储介质可以是 ROM、随机存取存储器（RAM）、CD-ROM、磁带、软盘和光数据存储设备等。

5 本领域技术人员在考虑说明书及实践这里公开的发明后，将容易想到本公开的其它实施方案。本公开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下面的权利要求指出。

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

权利要求

1、一种智能设备的语音控制方法，其特征在于，包括：

接收多个智能设备分别返回的语音数据；

对多个所述语音数据进行处理，得到优化语音数据；

5 基于所述优化语音数据控制所述语音数据对应的智能设备。

2、根据权利要求1所述的方法，其特征在于，所述接收多个智能设备分别返回的语音数据，包括：

接收位于不同位置的多个智能设备分别返回的语音数据。

10 3、根据权利要求1所述的方法，其特征在于，所述接收位于不同位置的多个智能设备返回的语音数据之前，所述方法还包括：

读取应用绑定的智能设备的基本信息；

基于所述基本信息确定具有录音功能的智能设备；

从具有录音功能的智能设备中确定待启动的智能设备；

向所述确定为待启动的智能设备发送开启指令，以开启所述智能设备。

15 4、根据权利要求3所述的方法，其特征在于，所述从具有录音功能的智能设备中确定待启动的智能设备，包括：

显示具有录音功能的智能设备的列表；

基于用户针对所述列表的选择操作，将用户选中的智能设备确定为待启动录音功能的智能设备。

20 5、根据权利要求3所述的方法，其特征在于，所述从具有录音功能的智能设备中确定待启动的智能设备，包括：

基于定位技术确定用户的位置；

查找预先存储的各所述具有录音功能的智能设备的位置；

25 将位于包括所述用户的位置在内的设定范围内的具有录音功能的智能设备，确定为待启动的智能设备。

6、根据权利要求5所述的方法，其特征在于，所述方法还包括：

在位于所述设定范围内的具有录音功能的智能设备的个数超过设定阈值时，读取位于所述设定范围内的具有录音功能的智能设备的历史使用数据；

基于所述历史使用数据确定待启动的智能设备。

30 7、根据权利要求6所述的方法，其特征在于，所述历史使用数据包括：使用频率、最近一次使用时间、使用总时长中的任一项或多项。

8、根据权利要求1所述的方法，其特征在于，所述对多个所述语音数据进行处理，得到优化语音数据，包括：

基于波束成形技术对多个所述语音数据进行处理，得到优化语音数据。

35 9、根据权利要求1所述的方法，其特征在于，所述基于所述优化语音数据控制所述

语音数据对应的智能设备，包括：

将所述优化语音数据发送给红外遥控设备，以使所述红外遥控设备基于所述优化语音数据中的语音信息查找对应的控制指令，基于所述优化语音数据中的设备名称查找对应的红外编码，并向所述红外编码发送所述控制指令。

- 5 10、根据权利要求 1 所述的方法，其特征在于，所述基于所述优化语音数据控制所述语音数据对应的智能设备，包括：

将所述优化语音数据发送给服务器，以使所述服务器基于所述优化语音数据中的语音信息查找对应的控制指令，并将所述控制指令和所述优化语音数据中的设备名称发送给红外遥控设备，进一步使所述红外遥控设备将所述控制指令发送给所述设备名称对应的红外编码。

- 11、一种智能设备的语音控制方法，其特征在于，包括：

采集语音数据；

将所述语音数据发送给控制设备，以使所述控制设备基于所述语音数据及多个位于其他位置的智能设备采集的所述语音数据，控制所述语音数据对应的智能设备。

- 15 12、根据权利要求 11 所述的方法，其特征在于，所述采集语音数据之前，所述方法包括：

基于所述控制设备发送的开启指令开启。

13、一种智能设备的语音控制装置，其特征在于，包括：接收模块、处理模块和控制模块；

- 20 所述接收模块，被配置为接收多个智能设备分别返回的语音数据；

所述处理模块，被配置为对所述接收模块接收到的多个所述语音数据进行处理，得到优化语音数据；

所述控制模块，被配置为基于所述处理模块得到的所述优化语音数据控制所述语音数据对应的智能设备。

- 25 14、根据权利要求 13 所述的装置，其特征在于，所述接收模块包括：接收子模块；所述接收子模块，被配置为接收位于不同位置的多个智能设备分别返回的语音数据。

15、根据权利要求 13 所述的装置，其特征在于，所述装置还包括：读取模块、第一确定模块、第二确定模块和开启模块；

所述读取模块，被配置为读取应用绑定的智能设备的基本信息；

- 30 所述第一确定模块，被配置为基于所述读取模块读取的基本信息确定具有录音功能的智能设备；

所述第二确定模块，被配置为从所述第一确定模块确定的具有录音功能的智能设备中确定待启动的智能设备；

- 35 所述开启模块，被配置为向所述第二确定模块确定为待启动的智能设备发送开启指令，以开启所述智能设备。

16、根据权利要求 15 所述的装置，其特征在于，所述第二确定模块包括：显示子模块和第一确定子模块；

所述显示子模块，被配置为显示具有录音功能的智能设备的列表；

5 所述第一确定子模块，被配置为基于用户针对所述显示子模块所显示的列表中的选择操作，将用户选中的智能设备确定为待启动录音功能的智能设备。

17、根据权利要求 15 所述的装置，其特征在于，所述第二确定模块包括：定位子模块、查找子模块和第二确定子模块；

所述定位子模块，被配置为基于定位技术确定用户的位置；

所述查找子模块，被配置为查找预先存储的各所述具有录音功能的智能设备的位置；

10 所述第二确定子模块，被配置为将位于包括所述定位子模块定位到的用户的位置在内的设定范围内的具有录音功能的智能设备，确定为待启动的智能设备。

18、根据权利要求 17 所述的装置，其特征在于，所述第二确定模块还包括：读取子模块和第三确定子模块；

15 所述读取子模块，被配置为在所述第二确定子模块所确定的所述智能设备的个数超过设定阈值时，读取位于所述设定范围内的具有录音功能的智能设备的历史使用数据；

所述第三确定子模块，被配置为基于所述读取子模块读取的历史使用数据确定待启动的智能设备。

19、根据权利要求 18 所述的装置，其特征在于，所述读取子模块读取的所述历史使用数据包括：使用频率、最近一次使用时间、使用总时长中的任一项或多项。

20 20、根据权利要求 13 所述的装置，其特征在于，所述处理模块包括：处理子模块；

所述处理子模块，被配置为基于波束成形技术对所述接收模块接收的多个所述语音数据进行处理，得到优化语音数据。

21、根据权利要求 13 所述的装置，其特征在于，所述控制模块包括：第一发送子模块；

25 所述第一发送子模块，被配置为将所述处理模块得到的所述优化语音数据发送给红外遥控设备，以使所述红外遥控设备基于所述优化语音数据中的语音信息查找对应的控制指令，基于所述优化语音数据中的设备名称查找对应的红外编码，并向所述红外编码发送所述控制指令。

30 22、根据权利要求 13 所述的装置，其特征在于，所述控制模块包括：第二发送子模块；

所述第二发送子模块，被配置为将所述处理模块得到的优化语音数据发送给服务器，以使所述服务器基于所述优化语音数据中的语音信息查找对应的控制指令，并将所述控制指令和所述优化语音数据中的设备名称发送给红外遥控设备，进一步使所述红外遥控设备将所述控制指令发送给所述设备名称对应的红外编码。

35 23、一种智能设备的语音控制装置，其特征在于，包括：采集模块和发送模块；

所述采集模块，被配置为采集语音数据；

所述发送模块，被配置为将所述采集模块采集的语音数据发送给控制设备，以使所述控制设备基于所述语音数据及多个位于其他位置的智能设备采集的所述语音数据，控制所述语音数据对应的智能设备。

5 24、根据权利要求 23 所述的装置，其特征在于，所述装置还包括：开启模块；

所述开启模块，被配置为基于所述控制设备发送的开启指令开启。

25、一种控制设备，其特征在于，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

接收多个智能设备分别返回的语音数据；

10 对多个所述语音数据进行处理，得到优化语音数据；

基于所述优化语音数据控制所述语音数据对应的智能设备。

26、一种智能设备，其特征在于，包括：处理器；用于存储处理器可执行指令的存储器；其中，所述处理器被配置为：

采集语音数据；

15 将所述语音数据发送给控制设备，以使所述控制设备基于所述语音数据及多个位于其他位置的智能设备采集的所述语音数据，控制所述语音数据对应的智能设备。

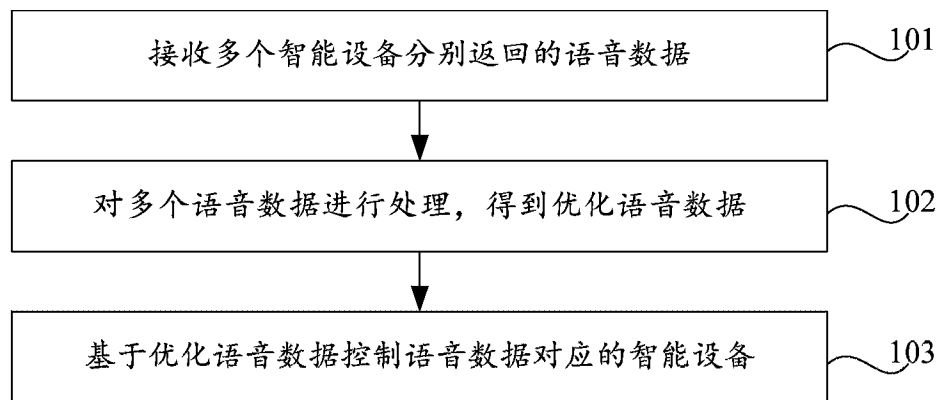


图1

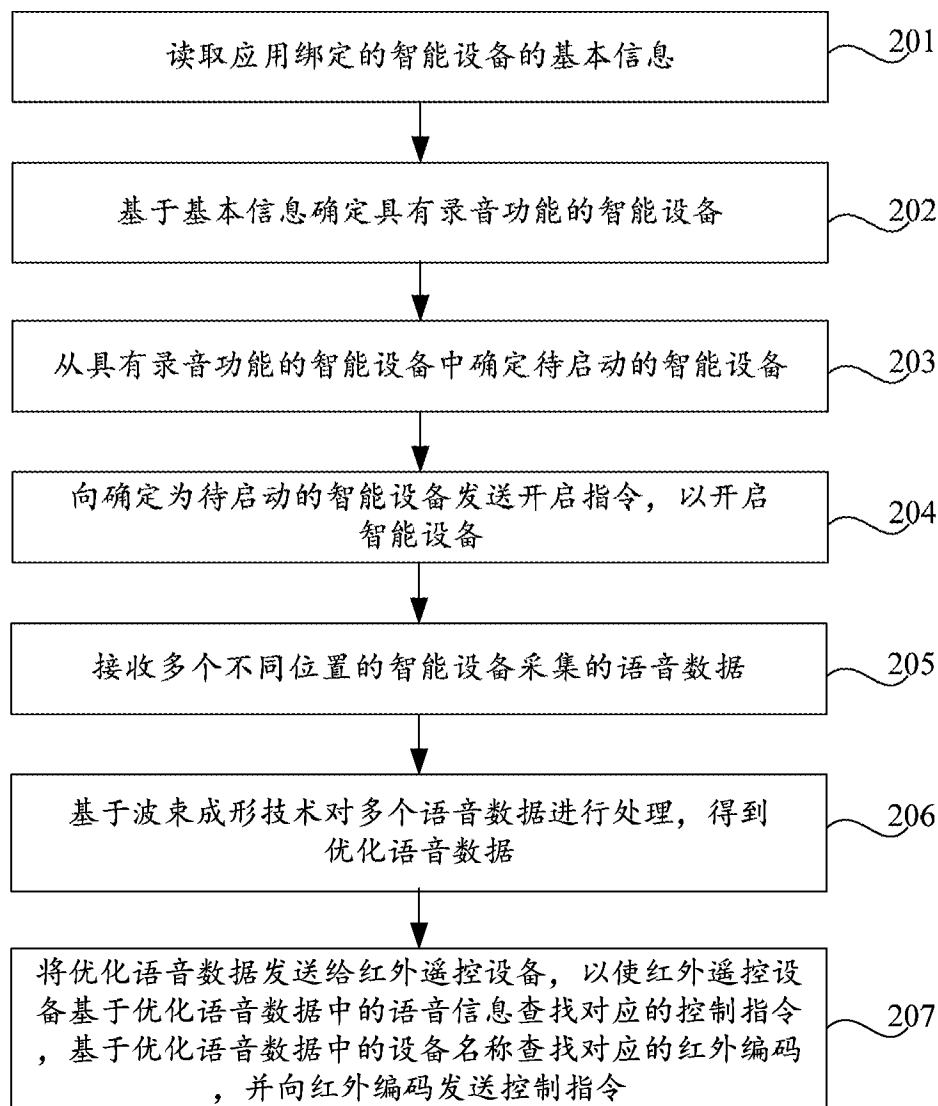


图2

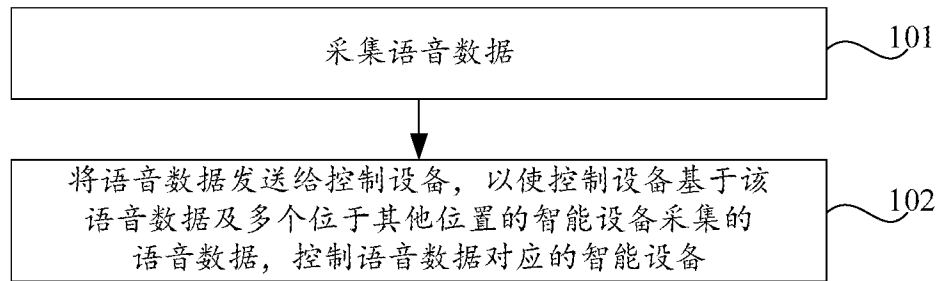


图3

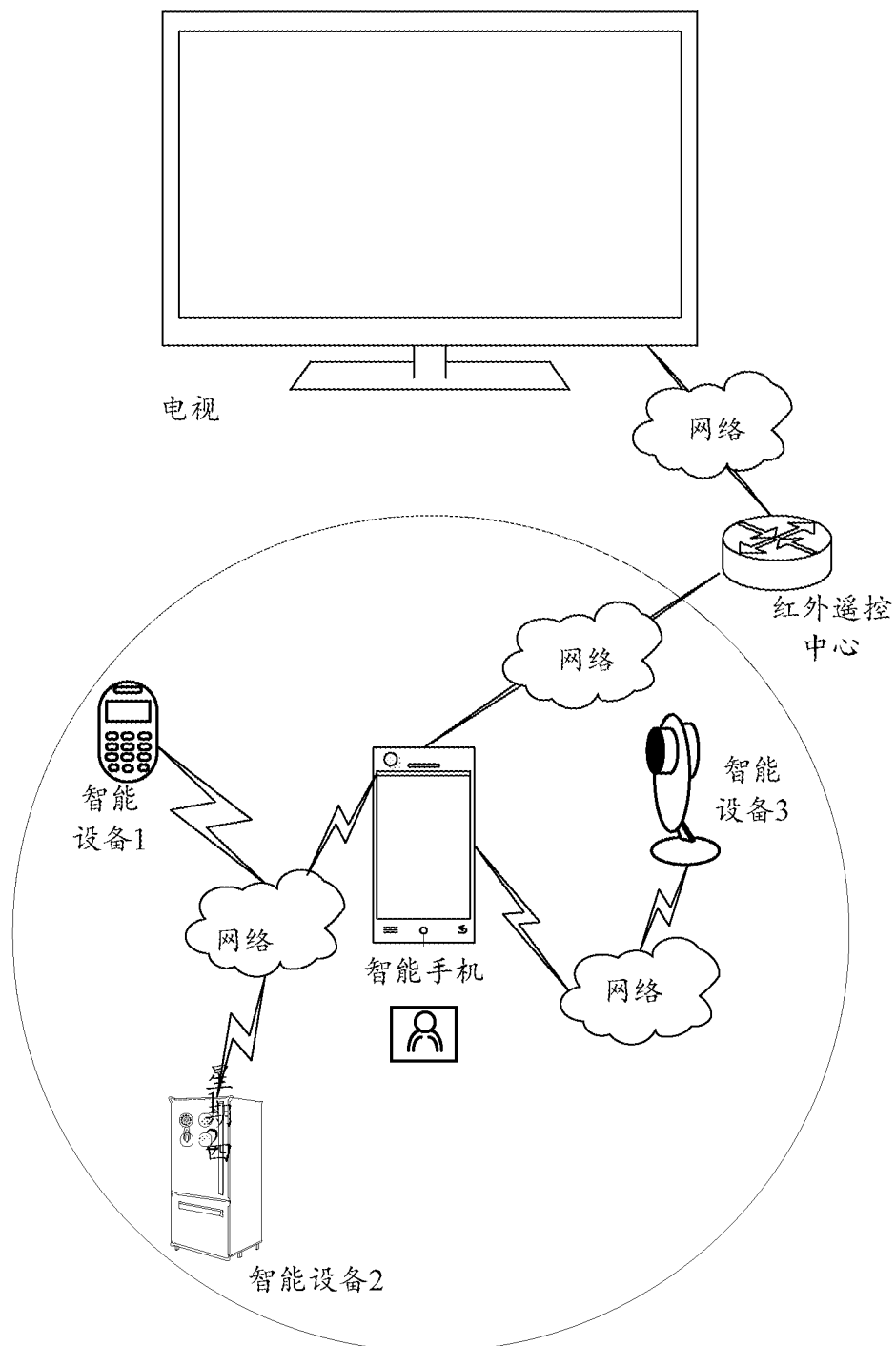


图4

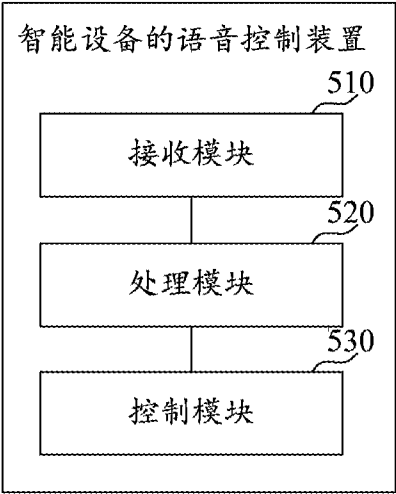


图5

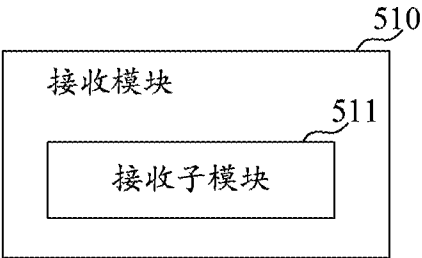


图6

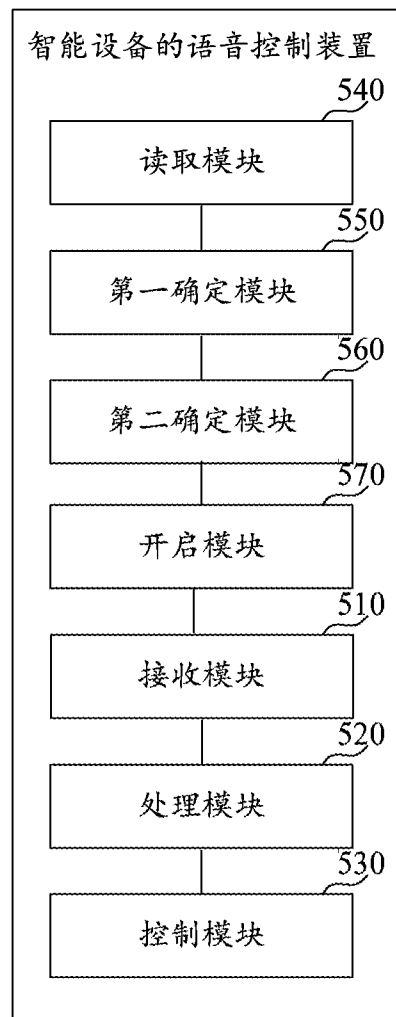


图7

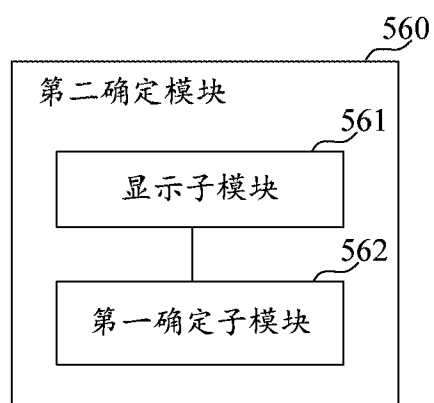


图8

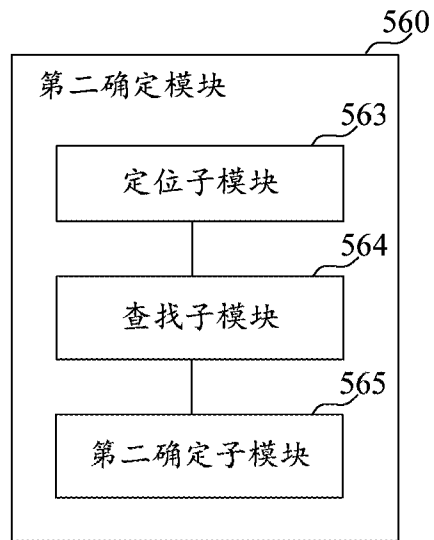


图9

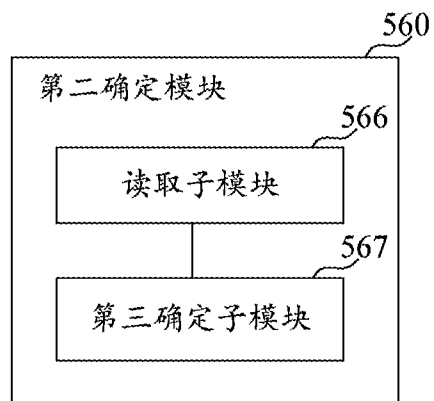


图10

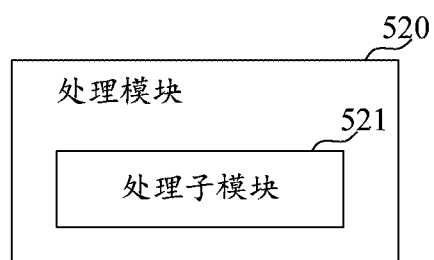


图11

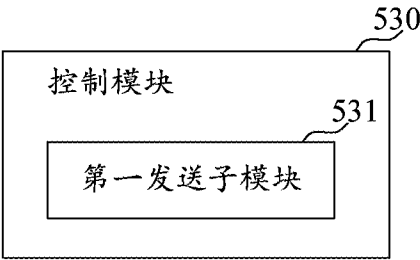


图12

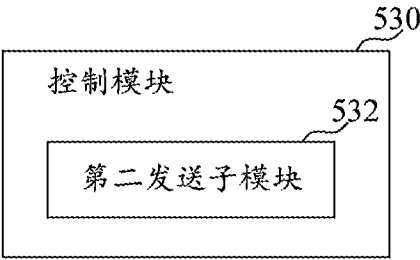


图13

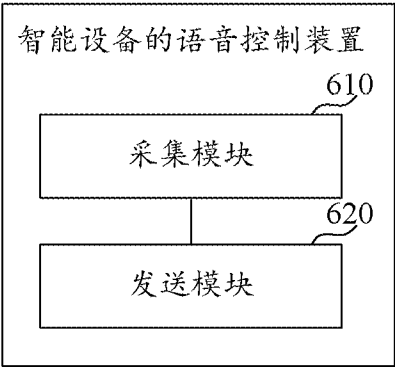


图14

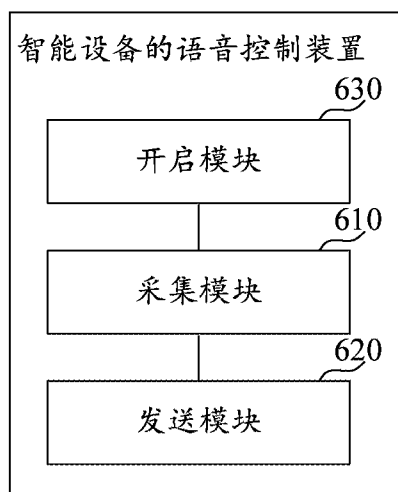


图15

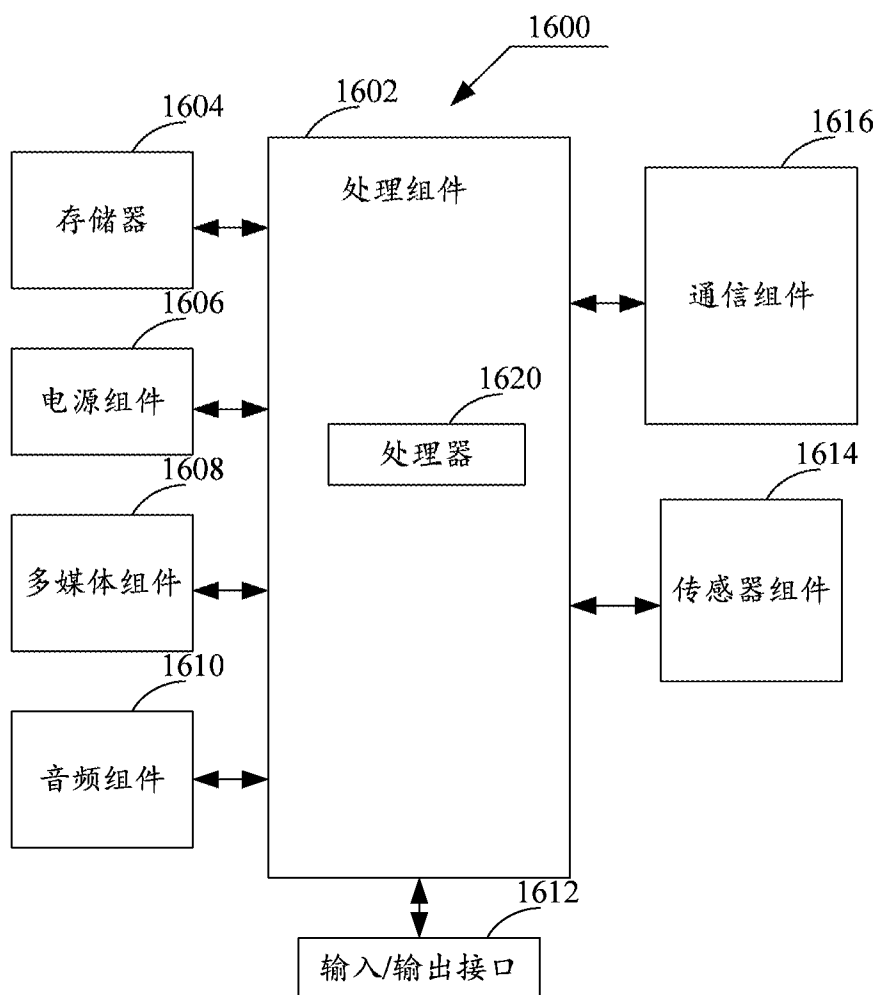


图16

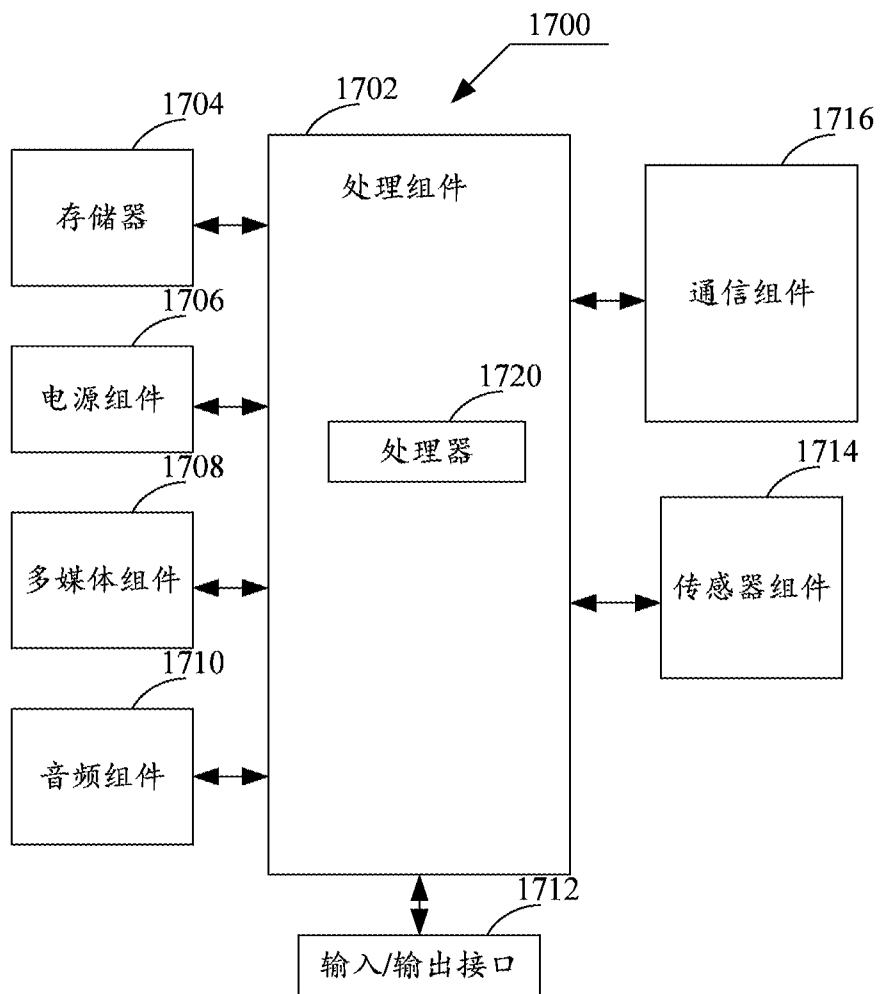


图17

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/099361**A. CLASSIFICATION OF SUBJECT MATTER**

G05B 15/02 (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)

G05B; 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)

CNTXT; CNABS; VEN: intelligent, location, record, several, phoneme, sound, pronounc+, vox, voice, speech, multi+, lots, many, plural+, microphone?, control+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN204390737U (IFLYTEK CO., LTD.), 10 June 2015 (10.06.2015), description, paragraphs 0015-0033, and figures 1-4	1-2, 8-14, 20-26
Y	CN 104935615 A (CHONGQING SHENSHU TECHNOLOGY CO., LTD.), 23 September 2015 (23.09.2015), description, paragraphs 0033-0051, and figures 1-4	1-2, 8-14, 20-26
A	US 6230138 B1 (VISTEON GLOBAL TECH INC.), 08 May 2001 (08.05.2001), the whole document	1-26
A	CN 102750946 A (PARAGON AG), 24 October 2012 (24.10.2012), the whole document	1-26

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 26 July 2016 (26.07.2016)	Date of mailing of the international search report 02 August 2016 (02.08.2016)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer SUN, Jiejun Telephone No.: (86-10) 62084166

International application No.
PCT/CN2015/099361

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2015/099361

A. 主题的分类

G05B 15/02(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)

G05B; 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))

CNTXT;CNABS;VEN:智能, 控制, 语音, 定位, 录音, 多, 几, 麦克风, phoneme, sound, pronounc+, vox, voice, speech, multi+, lots, many, plural+, microphone?, control+

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
Y	CN 204390737 U (科大讯飞股份有限公司) 2015年 6月 10日 (2015-06-10) 说明书0015-0033段, 图1-4	1-2, 8-14, 20-26
Y	CN 104935615 A (重庆深蜀科技有限公司) 2015年 9月 23日 (2015-09-23) 说明书0033-0051段, 图1-4	1-2, 8-14, 20-26
A	US 6230138 B1 (VISTEON GLOBAL TECH INC) 2001年 5月 8日 (2001-05-08) 全文	1-26
A	CN 102750946 A (帕拉贡股份公司) 2012年 10月 24日 (2012-10-24) 全文	1-26

☐ 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

2016年 7月 26日

国际检索报告邮寄日期

2016年 8月 2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

授权官员

孙洁君

电话号码 (86-10)62084166

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2015/099361

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	204390737	U	2015年 6月 10日	无	
CN	104935615	A	2015年 9月 23日	无	
US	6230138	B1	2001年 5月 8日	无	
CN	102750946	A	2012年 10月 24日	US 2012221341 A1	2012年 8月 30日
				DE 102011012573 A1	2012年 8月 30日
				US 8996383 B2	2015年 3月 31日

表 PCT/ISA/210 (同族专利附件) (2009年7月)