

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 1 月 10 日 (10.01.2019)



(10) 国际公布号
WO 2019/007249 A1

(51) 国际专利分类号:
G10L 17/22 (2013.01) *H04L 9/32* (2006.01)
G06F 21/32 (2013.01) *G10L 15/22* (2006.01)

(21) 国际申请号: PCT/CN2018/093231

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

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710542032.8 2017 年 7 月 5 日 (05.07.2017) CN

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

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

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(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,

(54) Title: INTERACTION METHOD, ELECTRONIC DEVICE, AND SERVER

(54) 发明名称: 一种交互方法、电子设备及服务器

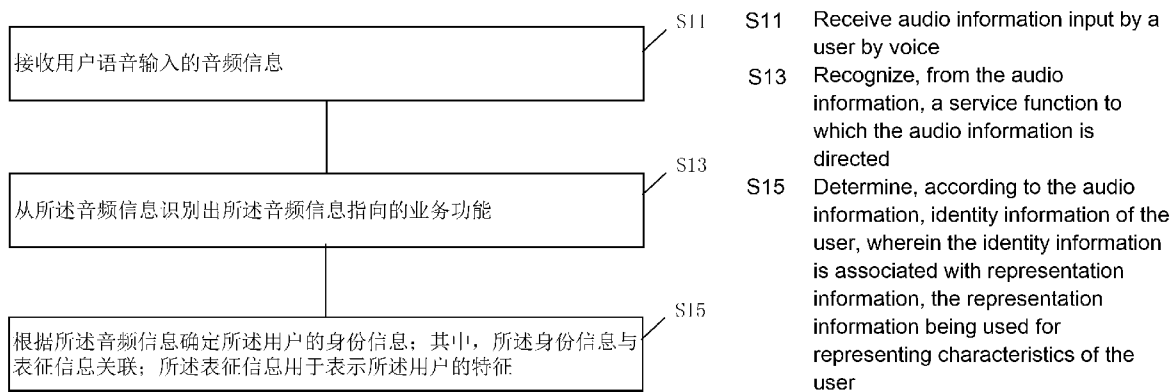


图 1

(57) Abstract: An interaction method, an electronic device, and a server. The method comprises: receiving audio information input by a user by voice (S11); recognizing, from the audio information, a service function to which the audio information is directed (S13); determining, according to the audio information, identity information of the user, wherein the identity information is associated with representation information, the representation information being used for representing characteristics of the user (S15); and using the representation information as an input of the service function to execute the service function.

(57) 摘要: 一种交互方法、电子设备及服务器, 其中, 该方法包括: 接收用户语音输入的音频信息 (S11); 从音频信息识别出音频信息指向的业务功能 (S13); 根据音频信息确定用户的身份信息, 其中, 该身份信息与表征信息关联, 该表征信息用于表示用户的特征 (S15); 将该表征信息作为业务功能的输入, 执行该业务功能。

AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

一种交互方法、电子设备及服务器

本申请要求 2017 年 07 月 05 日递交的申请号为 201710542032.8、发明名称为“一种交互方法、电子设备及服务器”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。

5

技术领域

本说明书中一个或者多个实施方式涉及计算机技术领域，特别涉及一种交互方法、电子设备及服务器。

10 背景技术

随着大数据时代的来临，越来越多的平台都会对用户的历史数据进行分析，从而总结出用户的行为特征，并且可以根据用户的行为特征，向其推送相应的信息。例如，用户在网购时，网购平台可以对用户的历史订单进行分析，从而得出用户比较感兴趣的商
品类型以及用户的消费能力。当用户在浏览网购平台的页面时，网购平台可以向用户推
荐其感兴趣的商品。

目前，用户通常会在平台中注册自身的账号信息。当用户在平台中登陆了账号信息后，用户在平台中产生的数据便可以与登陆的账号信息相关联。这样，平台可以针对账号信息关联的历史数据进行分析，从而得到该账号信息对应的用户的行为特征。

由上可见，现有技术中的这种数据分析的方式，通常是针对账号信息进行的。如果
不同的用户使用同一个账号信息，那么基于账号信息推送的业务信息则会不太准确。

发明内容

本说明书中一个或者多个实施方式的目的是提供一种交互方法、电子设备及服务器，能够为用户提供准确的业务信息。

为实现上述目的，本说明书中一个实施方式提供一种交互方法，包括：接收用户语音输入的音频信息；从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种电子设备，所述电子设备

包括语音录入单元、存储器以及处理器，其中：所述语音录入单元，用于接收用户语音输入的音频信息；所述存储器，用于存储用户的身份信息；所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将

5 所述表征信息作为所述业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收用户语音输入的音频信息；从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述业务功能信息和所述表征信息发送给服务器，

10 以用于所述服务器将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种电子设备，所述电子设备包括语音录入单元、存储器、网络通信端口以及处理器，其中：所述语音录入单元，用于接收用户语音输入的音频信息；所述存储器，用于存储用户的身份信息；所述网络通信端口，用于与服务器进行数据交互；所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述业务功能信息和所述表征信息通过所述网络通信端口发送给所述服务器，以用于所述服务器将所述

15 表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收用户语音输入的音频信息；从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；将所述业务功能信息和所述身份信息发送给服务器，以用于所述服务器确定与所述身份信息相关联的表征信息，以将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

20

为实现上述目的，本说明书中另一个实施方式还提供一种电子设备，所述电子设备包括语音录入单元、存储器、网络通信端口以及处理器，其中：所述语音录入单元，用于接收用户语音输入的音频信息；所述存储器，用于存储用户的身份信息；所述网络通信端口，用于与服务器进行数据交互；所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；将所述业务功能信息和所述身份信息通过所述网络通信端口发送给服务器，以用于所述服务器确定

25

30

与所述身份信息相关联的表征信息，以将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收用户语音输入的音频信息；从所述音频信息识别出所述音频信息指向的业务功能信息；将
5 所述业务功能信息和所述用户的音频信息发送给服务器，以用于所述服务器根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种电子设备，所述电子设备包括语音录入单元、网络通信端口以及处理器，其中：所述语音录入单元，用于接收用
10 户语音输入的音频信息；所述网络通信端口，用于与服务器进行数据交互；所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；将所述业务功能信息和所述用户的音频信息通过所述网络通信端口发送给所述服务器，以用于所述服务器根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

15 为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收用户语音输入的音频信息；将所述用户的音频信息发送给服务器，以用于所述服务器从所述音频信息识别出所述音频信息指向的业务功能信息，根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

20 为实现上述目的，本说明书中另一个实施方式还提供一种电子设备，所述电子设备包括语音录入单元、网络通信端口以及处理器，其中：所述语音录入单元，用于接收用户语音输入的音频信息；所述网络通信端口，用于与服务器进行数据交互；所述处理器，用于将所述用户的音频信息通过所述网络通信端口发送给服务器，以用于所述服务器从所述音频信息识别出所述音频信息指向的业务功能信息，根据所述音频信息确定所述用
25 户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收客户端发出的音频信息；从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。
30

为实现上述目的，本说明书中另一个实施方式还提供一种服务器，所述服务器包括网络通信端口、存储器以及处理器，其中：所述网络通信端口，用于接收客户端发出的音频信息；所述存储器，用于存储用户的身份信息；所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定用户的身份信息；其中，

5 所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收客户端发出的音频信息以及业务功能信息；根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息

10 息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种服务器，所述服务器包括网络通信端口、存储器以及处理器，其中：所述网络通信端口，用于接收客户端发出的音频信息以及业务功能信息；所述存储器，用于存储用户的身份信息；所述处理器，用于根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述

15 表征信息用于表示所述用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收客户端发出的身份信息以及业务功能信息；确定与所述身份信息相关联的表征信息；所述表征信息用于表示用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务

20 功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种服务器，所述服务器包括网络通信端口、存储器以及处理器，其中：所述网络通信端口，用于接收客户端发出的身份信息以及业务功能信息；所述存储器，用于存储表征信息；所述处理器，用于确定与所述身份信息相关联的表征信息；所述表征信息用于表示用户的特征；将所述表征信息

25 息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种交互方法，包括：接收客户端发出的业务功能信息和表征信息；其中，所述表征信息用于表示用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

为实现上述目的，本说明书中另一个实施方式还提供一种服务器，所述服务器包括

30 网络通信端口和处理器，其中：所述网络通信端口，用于接收客户端发出的业务功能信

息和表征信息；其中，所述表征信息用于表示用户的特征；所述处理器，用于将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

由上可见，本说明书一个或者多个实施方式提供的技术方案，可以将用户的身份信息与用户的音频信息进行关联。由于不同用户的声音特质也往往不同，因此在对用户的音频信息进行分析时，能够准确地识别出该用户的身份信息。该身份信息可以具备对应的表征信息，该表征信息便可以表示用户的行为特征。这样，当用户通过语音下达针对业务功能的执行指令时，便可以执行与该用户的表征信息相符的业务功能，从而能够准确地向用户提供相应的业务信息。

10 附图说明

为了更清楚地说明本说明中一个或者多个实施方式或现有技术中的技术方案，下面将对实施方式或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本说明书中记载的一些实施方式，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

15 图 1 为本说明书一个实施方式提供的交互方法流程图；

图 2 为本说明书一个实施方式中身份信息的确定方法流程图；

图 3 为本说明书一个实施方式中智能音箱的交互示意图；

图 4 为本说明书另一个实施方式中智能音箱的交互示意图；

图 5 为本说明书一个实施方式中智能洗衣机的交互示意图；

20 图 6 为本说明书一个实施方式中自动服务机的交互示意图；

图 7 为本说明书一个实施方式中电子设备的结构示意图；

图 8 为本说明书一个实施方式中交互方法的流程图。

具体实施方式

25 为了使本技术领域的人员更好地理解本说明书中一个或者多个实施方式的技术方案，下面将结合本说明书中一个或者多个实施方式的附图，对本说明书中一个或者多个实施方式中的技术方案进行清楚、完整地描述，显然，所描述的实施方式仅仅是本说明书中的一部分实施方式，而不是全部的实施方式。基于本说明书中的一个或者多个实施方式，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施方式，
30 都应当属于本说明书中一个或者多个实施方式保护的范围。

本说明书一个实施方式提供一种交互方法，所述交互方法可以应用于客户端和服务器的系统架构中。其中，客户端可以是具有录音功能的电子设备。具体的，例如，客户端可以是台式电脑、平板电脑、笔记本电脑、智能手机、数字助理、智能可穿戴设备、

5 导购终端、电视机、智能音箱、麦克风等。其中，智能可穿戴设备包括但不限于智能手环、智能手表、智能眼镜、智能头盔、智能项链等。或者，客户端也可以为能够运行于上述电子设备中的软件。例如，电子设备中提供录音功能，软件可以通过调用该录音功能录制音频信息。

在本说明书一个实施方式中，客户端可以是具有录音功能的电子设备。根据客户端

10 数据处理能力的不同，可以被划分成以下类别。

表 1

类别编号	类别名称	示例性设备	数据处理能力
1	初级网络设备	物联网设备、边缘节点设备	数据采集和发送
2	简单网络设备	智能可穿戴设备、POS (point of sale) 机	数据采集、发送、简单数据预处理
3	中级网络设备	智能家电、智能音箱	数据采集、发送、中等功耗数据预处理
4	智能网络设备	性能较好的智能手机、平板电脑	数据采集、发送、较高功耗的数据预处理、数据查询匹配
5	高性能设备	工作站、智能电话亭	数据采集、发送、数据预处理、数据查询匹配、数据存储

在本实施方式中，初级网络设备的硬件设备较为简单，可以进行通过麦克风进行录音，生成音频信息。并将生成的音频信息通过网路通信模块发送给服务器。初级网络设备

15 备可以包括麦克风、网络通信单元、传感器和扬声器。初级网络设备可以基本上不需要对数据进行加工处理。初级网络设备还可以设置有其它的传感器，用于采集初级网络设备的工作参数。具体的，例如，初级网络设备可以是物联网设备、边缘节点设备等。

在本实施方式中，简单网络设备可以主要包括：麦克风、网络通信单元、处理器、存储器、扬声器等。简单网络设备相较于初级网络设备增强了数据处理的能力。简单网络设备

20 网络设备可以具有一个能处理简单逻辑运算的处理器，使得简单网络设备在采集到数据之后，可以对数据进行初步的预处理，比如可以根据音频信息生成特征矩阵。简单网络设备可以具有一个具有简单显示功能的显示模块，可以用于向用户反馈信息。具体的，例

如，简单网络设备可以是智能可穿戴设备、POS（point of sale）机等。例如，智能手环、较初级的智能手表、智能眼镜，或是线下购物场所内的结算设备（例如，POS 机）、移动式结算设备（例如，手持式 POS 机、附加在手持设备上的结算模块）等。

在本实施方式中，中级网络设备可以主要包括麦克风、网络通信单元、处理器、存储器显示器、扬声器等。中级网络设备的处理器的主频通常小于 2.0GHz，内存容量通常少于 2GB，存储器的容量通常少于 128GB。中级网络设备可以对录制的音频信息进行一定程度的处理，比如生成特征矩阵，对特征矩阵进行端点检测处理、降噪处理、语音识别等。具体的，例如，中级网络设备可以包括：智能家居中的智能家用电器、智能家庭终端、智能音箱、较高级的智能手表、较为初级的智能手机（比如，价位在 1000 元左右）、车载智能终端。

在本实施方式中，智能网络设备可以主要包括麦克风、网络通信单元、处理器、存储器、显示器、扬声器等硬件。智能网络设备可以具有较强的数据处理能力。智能网络设备的处理器的主频通常大于 2.0GHz，内存的容量通常小于 12GB，存储器的容量通常小于 1TB。可以对音频信息生成特征矩阵之后，可以进行端点检测处理、降噪处理、语音识别等。进一步的，智能网络设备还可以根据音频信息生成语音特征向量。在一些情况下，可以将语音特征向量与用户特征向量进行匹配，识别用户的身份。但这种匹配限于有限个数的用户特征向量，比如一个家庭中的各个家庭成员的用户特征向量。具体的，例如，智能网络设备可以包括：性能较好的智能手机、平板电脑、台式电脑、笔记本电脑等。

在本实施方式中，高性能设备可以主要包括麦克风、网络通信单元、处理器、存储器、显示器、扬声器等硬件。高性能设备可以具有大规模的数据运算处理能力，还可以提供强大的数据存储能力。高性能设备的处理器主频通常在 3.0GHz 以上，内存的容量通常大于 12GB，存储器容量可以在 1TB 以上。高性能设备可以对音频信息生成特征矩阵、端点检测处理、降噪处理、语音识别、生成语音特征向量，以及将语音特征向量与存储的大量的用户特征向量进行匹配。具体的，例如，高性能设备可以是工作站、配置很高的台式电脑、Kiosk 智能电话亭、自助服务机等。

当然，上述只是示例的方式列举了一些客户端。随着科学技术进步，硬件设备的性能可能会有提升，使得上述目前数据处理能力较弱的电子设备，也可能具备较强的处理能力。所以下文中实施方式引用上述表 1 中的内容，也仅作为示例参考，并不构成限定。

在本实施方式中，服务器可以是具有一定运算处理能力的电子设备。其可以具有网

络通信模块、处理器和存储器等。当然，上述服务器也可以是指运行于所述电子设备中的软体。上述服务器还可以为分布式服务器，可以是具有多个处理器、存储器、网络通信模块等协同运作的系统。或者，服务器还可以为若干服务器形成的服务器集群。

请参阅图 1，本说明书一个实施方式提供一种交互方法，该方法的执行主体可以是

5 客户端，所述方法可以包括以下步骤。

S11：接收用户语音输入的音频信息。

在本实施方式中，涉及用户的音频信息采集方法，具体地，可以通过客户端采集用户的音频信息。客户端可以在用户操作启动录音功能时，开始启动进行录取用户的语音，生成音频信息。客户端也可以自动启动录音功能，比如给客户端设置指定条件，当该条件达成时，启动录音功能。具体的，例如，指定一个时间，到达该时间时启动录音；或者，指定一个地点，到达该地点时启动录音；或者，设置一个环境音量，当环境音量符合设定的条件时，开始录音。在本实施方式中，生成音频信息的数量，可以是一个，也可以是多个。具体的，客户端在一次录音过程中，可以持续的将该次录音过程中的全部内容作为一个音频信息。也可以为，客户端在一次录音过程中，划分多个音频信息。比如，按照录音的时长划分音频信息。例如，每录制五分钟，形成一个音频信息。或者，按照数据量进行划分音频信息。例如，每个音频信息最多 5MB。

S13：从所述音频信息识别出所述音频信息指向的业务功能。

在本实施方式中，可以预先定义有业务指令集，在业务指令集中包括至少一个业务指令。业务指令可以指向一个特定的功能。当从用户的音频信息中识别出业务指令时，可以表示要执行该业务指令指向的业务功能。也可以为，仅仅通过预设识别算法得出用户说话表达的内容。具体的，预设识别算法可以采用隐式马尔科夫算法或神经网络算法等，对音频信息进行语音识别。

在本实施方式中，所述客户端可以作为模块集成于各种设备中。例如，所述客户端可以集成于咖啡机中，这样，用户对咖啡机说“我要喝咖啡”，那么该音频信息指向的业务功能便可以制作咖啡。又例如，所述客户端可以是智能手机中的语音识别模块，当用户说“我要付款”时，该音频信息指向的业务功能便可以付款。

在本说明书一个实施方式中，在识别业务功能时，可以通过语音转文字的技术来实现。具体地，首先可以将所述音频信息转换为对应的文字信息。具体地，可以识别所述音频信息中的音素，然后根据音素组合表，将各个音素组合为对应的文字信息。在得到文字信息后，可以将所述文字信息与业务指令库中的业务指令进行匹配。所述业务指令

库中可以存储业务指令，所述业务指令可以是通过文字或者词向量的形式进行存储的。这样，将识别出的文字信息与业务指令库中的各个业务指令进行匹配，从而可以将匹配得到的业务指令作为所述音频信息指向的业务功能。具体地，在进行业务指令匹配时，可以通过计算词向量之间的距离来实现。首先，可以将识别出的文字信息转换为对应的词向量，然后可以将转换得到的词向量与业务指令库中的各个业务指令的词向量求取欧氏距离。欧式距离越小，表明文字信息与业务指令越相似。这样，可以将欧氏距离最小的业务指令作为所述音频信息指向的业务功能。

S15: 根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征。

10 在本实施方式中，可以将用户的语音特征与所述用户的身份信息相关联。

在本实施方式中，可以采用音频信息采集方法获得用户的音频信息。其中，音频信息中的音频数据可以是用户的说话声音的录音。如此，根据音频信息生成的语音表征向量，可以对应表征的音频信息，也可以表征用户的一部分声音特质。由于每个用户生长发育过程，是各部不相同的，使得用户说话的声音，都具有一定的声音特质。进而，可以通过每个用户的声音特质分不同的用户。如此，语音表征向量可以通过表征用户的一部分声音特质，而可以用于标识用户。

在本实施方式中，针对用户采集的音频信息可以是一个或多个，可以对应每个音频信息采用音频信息处理方法生成对应的语音特征向量。当然，在一些情况下，也可以将一个以上音频信息同时进行按照音频信息处理方法进行运算处理，得到语音特征向量。此时，该语音特征向量可以对应该一个以上音频信息。

在本实施方式中，根据得到的语音特征向量，确定可以用于标识用户的用户特征向量。具体的，例如，若仅生成了一个语音特征向量，则可以将该语音特征向量作为用户的用户特征向量；若生成了多个语音特征向量，可以将在该多个语音特征向量中，选择一个相对表达用户的声音特质较多的语音特征向量，作为用户的用户特征向量；若生成了多个语音特征向量，还可以为将该多个语音特征向量中的部分或全部，进行进一步运算处理输出用户的用户特征向量。该运算处理可以包括但不限于针对该多个语音特征向量进行相应维度求和之后，再进一步计算均值。当然，还可以有其它算法，比如，运算处理时对多个语音特征向量的加权求和。

在本实施方式中，将用户的用户特征向量与用户的身份信息进行关联。用户的身份信息可以用于表示一个用户。用户的身份信息可以包括但不限于：用户名、昵称、真实

姓名、身份证、性别、联系电话、通信地址等等。将用户特征向量与用户的身份信息相关联，可以实现采集一个用户说话的音频信息后，可以通过音频信息的语音特征向量与用户特征向量的关联性，确定该用户的身份信息。

请参阅图 2，在本说明书一个实施方式中，针对身份信息的确定方法，可以包括以

5 下步骤。

步骤 S101：生成所述音频信息的语音特征向量。

在本实施方式中，可以约束用户语音输入的表达内容。比如，提供给用户一段文字让用户朗读。或者，询问用户问题，让用户回答。当然，也可以不约束用户表达的内容，而随机录制用户说的一句话或者一段话。

10 在本实施方式中，客户端可以在接收到用户的音频信息之后，生成语音特征向量。具体的，例如，用户可以根据其意志控制客户端，了解商品或者服务的信息。用户可以通过与客户端对话的方式，了解不同商品或者服务的细节参数等。在用户想要购买时，才启动生成用户的语音特征向量，进而确定用户的身份信息。进一步示例，客户端是智能音箱的场景下，可以为用户对智能音箱说“我要喝咖啡”，智能音箱通过语音识别出

15 用户表达的内容之后，才进一步的生成用户的语音特征向量。

在本实施方式中，生成语音特征向量时，可以先根据所述音频信息生成特征矩阵。具体地，可以根据预设算法，从音频信息中采集数据，输出包括所述音频信息的音频数据的特征的特征矩阵。用户的声会有用户自身的特征，比如音色、语调、语速等等。录制成音频信息时，可以从音频数据中的频率、振幅等角度，体现每个用户自身的声音

20 特征。使得将音频信息按照预设算法生成的特征矩阵，会包括音频信息中音频数据的特征。进而，基于特征矩阵生成的语音特征向量，可以用于表征该音频信息和音频数据。所述预设算法可以是 MFCC (Mel Frequency Cepstrum Coefficient)、MFSC (Mel Frequency Spectral Coefficient)、FMFCC (Fractional Mel Frequency Cepstrum Coefficient)、DMFCC (Discriminative)、LPCC (Linear Prediction Cepstrum Coefficient) 等。当然，所属领域

25 技术人员在本申请技术精髓启示下，还可能采用其它算法实现生成音频信息的特征矩阵，但只要其实现的功能和效果与本申请方式相同或相似，均应涵盖于本申请保护范围内。

在本实施方式中，在生成了特征矩阵后，可以将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的维度值，所述多个维度值形成所述语音特征向量。具体地，可以对所述特征矩阵按照不同的特征维度进行降维处理，得到可以表征

30 每个特征维度的维度值。进一步的，将维度值按照指定顺序排列便可以形成音频信息的

语音表征向量。具体的，可以通过卷积或者映射的算法对特征矩阵进行降维处理。在一个具体的示例中，可以采用 DNN（Deep Neural Network）、CNN（Convolutional Neural Network）和 RNN（Recurrent Neural Network）、深度学习或者上述算法的结合等，从特征矩阵中按照不同维度进行降维。

- 5 在一个实施方式中，为了进一步区分出音频信息中用户语音的音频数据和非用户语音的音频数据。语音特征向量的生成方法中，还可以包括端点检测处理。进而，可以在特征矩阵中减少将非用户语音的音频数据对应的数据，如此，可以在一定程度上提升生成的语音特征向量与用户之间的关联程度。端点检测处理的方法可以包括但不限于基于能量的端点检测、基于倒谱特征的端点检测、基于信息熵的端点检测、基于自身相关相似距离的端点检测等，在此不再列举。

步骤 S103：将所述语音特征向量与用户特征向量进行匹配。

步骤 S105：将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

在本实施方式中，客户端中可以预先按照用户注册方法设置有用户的用户特征向量。如此，可以通过将语音特征向量与用户特征向量进行匹配，实现对用户的身份进行验证。

- 15 在本实施方式中，将语音特征向量与用户特征向量进行匹配的方式，可以为根据二者进行运算，在二者之间符合某种关系时，可以认为匹配成功。具体的，例如，将二者做差后求和，将得到的数值作为匹配值，将该匹配值与一个设定阈值比较，在所述匹配值小于或等于设定阈值的情况下认为所述语音特征向量与所述用户特征向量匹配成功。或者，也可以将所述语音特征向量与所述用户特征向量求和，将得到的数值作为匹配值，
- 20 在所述匹配值大于或等于设定阈值的情况下认为所述语音特征向量与所述用户特征向量匹配成功。

- 在本实施方式中，客户端中可以具备多个用户的特征向量，这样，在获取到当前用户的语音特征向量之后，可以将该语音特征向量与多个用户的特征向量依次进行对比。在与各个用户的特征向量进行对比时，均可以得到对应的匹配值。这样，可以将匹配值
- 25 最大的用户的特征向量作为匹配得到的用户特征向量。在本实施方式中，在确定了所述语音特征向量对应的用户特征向量之后，便可以将该用户特征向量关联的身份信息作为所述用户的身份信息。

- 在本实施方式中，在确定了用户的身份信息后，所述身份信息通常与表征信息相关联。所述表征信息可以用于表示所述用户的特征。具体地，所述表征信息是针对用户的历史数据分析得到的。所述特征可以体现用户的偏好。例如，所述表征信息表明该用户
- 30

比较喜欢喝拿铁、在线下支付时经常使用微信支付、导航时通常选择规避拥堵的路线等。

S17: 将所述表征信息作为所述业务功能的输入, 执行所述业务功能。

在本实施方式中, 当客户端确定了用户的音频信息指向的业务功能, 并且确定了用户的表征信息后, 便可以将该表征信息作为执行业务功能的输入数据, 确定出符合用户表征信息的实现方式, 并通过该实现方式来执行所述业务功能。例如, 用户当前的音频信息是“我要喝咖啡”, 客户端根据用户的音频信息识别出该用户的表征信息后, 确定该表征信息表明用户偏爱喝拿铁, 那么客户端在执行制作咖啡这个业务功能时, 便可以制作拿铁咖啡, 从而符合用户的偏好。

在一个实施方式中, 在从所述音频信息识别出所述音频信息指向的业务功能时, 可以判断在所述音频信息中是否存在启动关键词。所述启动关键词可以是激活客户端的指定词汇或者短语。在本实施方式中, 客户端可以在判定存在所述启动关键词的情况下, 才从所述音频信息识别出所述音频信息指向的业务功能。如此, 在用户正常工作生活过程中, 不会受到客户端的干扰。而在用户需要使用客户端时, 可以通过给客户端发出启动指令, 开始与客户端进行交互。

在本实施方式中, 客户端可以处于一种待机的状态, 录制用户语音的音频信息, 以识别用户是否向客户端做出启动指令。具体的, 例如, 启动指令可以为识别到用户说“你好, XXX”。当客户端识别到用户说“你好, XXX”时, 便启动交互过程。

在一个实施方式中, 生成所述语音特征向量的方式包括以下至少一种: 任意选择一个所述用户的音频信息生成所述语音特征向量; 或者向所述用户提供指定信息, 并根据所述用户依照所述指定信息输入的语音信息生成所述语音特征向量。

在本实施方式中, 可以针对不同的场景, 生成语音特征向量所基于的音频信息不同。针对业务功能简单的场景, 可以根据任意一个录制的用户语音的音频信息生成语音特征向量, 以达到可以快速执行业务功能的目的。例如, 该业务功能简单的场景可以是点播歌曲的场景。

在本实施方式中, 对于一些相对复杂的或者安全系数需要较高的场景时, 可以要求录制用户设定的指定信息。所述指定信息可以是预先录制的一段文字、一句话或者模拟发声而无具体的含义。例如, 用户将自己的生日作为所述指定信息, 或者学一段鸟叫的口技作为指定信息。这样, 在用户下达用于表征业务功能的语音信息后, 客户端还可以

播放语音提示信息，该语音提示信息可以表明要求用户依照所述指定信息生成相应的语音信息。例如，客户端可以播放“请用正常语速阅读以下内容：XXXX”。这样，客户端在接收到依照所述指定信息输入的语音信息后，便可以基于该语音信息生成语义特征向量。这样处理的目的在于能够准确地识别出用户的身份信息。具体地，这种复杂的场景例如可以是涉及线下支付的场景。

在一个实施方式中，有时候用户提供的音频信息可能不太准确，或者音频信息中不包含确切的指令，在这种情况下，客户端可能会识别出多个业务功能。例如，用户的音频信息可以为“打开”，但是用户并没有指示具体打开什么。此时，客户端便可以针对该音频信息，向用户播放包含多个业务功能的提示语音信息。例如，所述提示语音信息可以为“您想打开什么？1. 窗户；2. 门；3. 空调；4. 电视，请您选择”。此时，用户可以针对该提示语音信息下达语音选择指令。例如，用户可以说“窗户”或者“1”。此时，客户端接收到所述用户下达的语音选择指令时，可以将所述语音选择指令指向的业务功能作为所述音频信息中识别出的业务功能。

在一个实施方式中，所述表征信息中可以包括针对所述业务功能的用户偏好数据。例如，所述业务功能为“支付”，那么针对该业务功能的用户偏好数据可以是“微信支付”，这样，客户端便可以从多个支付方式中选择微信支付的方式来实现支付的过程。此外，为了确保执行业务功能的准确性，客户端还可以向用户播放符合所述用户偏好数据的语音确认信息。例如，在确定用户偏好的支付方式为“微信支付”后，客户端可以播放“请问是否采用微信支付的方式？”。这样，在接收到用户下达的语音确认指令后，便可以执行符合所述用户偏好数据的业务功能。例如，用户回答“是的”，那么客户端便可以采用微信支付的方式完成支付的过程。

在一个实施方式中，用户如果没有在客户端中进行注册，或者用户在客户端中不存在身份信息，那么根据所述音频信息便无法确定所述用户的身份信息。此时，客户端便无法获取用户的表征信息，也就无法针对用户的偏好来执行业务功能。此时，客户端可以调用与所述业务功能相关联的备选信息集。所述备选信息集中包含的备选信息均可以表征所述业务功能的一种实现方式。例如，所述业务功能为“制作咖啡”，而与该业务功能相关的实现方式有“制作拿铁”、“制作卡布奇诺”、“制作焦糖咖啡”等。这样，

客户端可以向用户播放所述备选信息集中的备选信息，以供用户选择。在用户下达语音选择指令后，客户端可以接收用户下达的语音选择指令，所述语音选择指令可以指向所述备选信息集中的目标备选信息。例如，用户选择的可以是“制作焦糖咖啡”。这样，客户端可以根据所述目标备选信息表征的实现方式，执行所述业务功能。也就是说，客户端可以开始制作焦糖咖啡。

请参阅图 3。在一个场景示例中，客户端可以是具有一定程度上的运算能力的家用智能设备。例如，可以是上表 1 中的类别 3 型设备。在一个家用场景下，客户端可以被制造为智能音箱。智能音箱可以具有麦克风、扬声器、Wifi 模块、存储器、处理器等等。智能音箱可以实现普通的音频播放功能，并且配备有处理设备和网络设备以通过与用户对话以及与服务服务器数据交互，实现购物功能。

在本场景示例中，智能音箱可以通过识别唤醒词启动进一步的功能，在智能音箱识别到用户说出唤醒词之前，可以处于一种待机的状态。用户需要使用智能音箱时，可以说“你好，音箱”。智能音箱会录制该用户说的语音，并识别得出用户说话的内容为唤醒词。此时，智能音箱可以通过扬声器发生回答用户，“您好，您需要帮忙么？”。

在本场景示例中，用户想要购买空气净化器。用户可能会说：“我想买一台空气净化器，你有什么推荐么？”。智能音箱通过麦克风录音生成音频信息后，可以识别出该音频信息对应的身份信息。该身份信息可以与当前用户的偏好数据相关联。此时，智能音箱可以确定该用户对美的和小米的产品比较感兴趣。因此，智能音箱可以通过扬声器播放语音：“我有两个推荐，第一个是小米空气净化器 2 代，第二个是美的 KJ210G-C46 空气净化器”。用户可能会说：“我要购买小米空气净化器 2 代”。智能音箱便可以执行小米空气净化器 2 代的购买流程。

需要说明的是，智能音箱还可以与网购平台的服务器相连。智能音箱可以根据自身的配置，实现不同的功能。例如，智能音箱可以仅仅具备语音录入的功能，在获取到用户的音频信息后，可以将音频信息交由服务器处理。当然，智能音箱本身还可以实现一部分处理功能。例如，智能音箱可以对用户的音频信息进行识别，得到业务功能和身份信息之后，再将业务功能和身份信息发送给服务器，进行后续处理。当然，在实际应用过程中，该交互过程中的一个或者多个步骤可以分别在智能音箱和服务服务器中完成，本说明书对此并不作限定。

请参阅图 4。在一个场景示例中，智能音箱可以主动预测用户关注的商品或服务。

在本场景示例中，智能音箱可以位于用户家庭的客厅。用户在看电视过程中，智能音箱可以通过对电视声音的录音，得到音频信息，进而识别得到频道编码。或者，智能音箱可以通过识别电视节目的内容，与服务器进行交互，由服务器反馈相应的电视频道。

- 5 在本场景示例中，智能音箱可以存储有电视频道的节目清单。或者，也可以从服务器拉去电视频道的节目清单。在用户观看电视的过程中，可能对某一个广告的商品感兴趣，要进行购买。此时，用户可以说：“你好，音箱”，“我要购买这个广告中的商品”。智能音箱可以根据用户说话的时间点，和节目清单中给出的时间，确定用户要购买的商品信息。进一步的，智能音箱可以根据用户的语音的音频信息生成语音特征文件进行匹
- 10 配，验证用户的身份。在验证成功时，将用户的个人信息和商品信息发送给服务器，以使服务器从用户的金融账户中支付商品信息中的价款给该商品的卖家。

- 请参阅图 5。在一个场景示例中，客户端可以是具有数据采集能力的家用物联网设备。例如，可以是上表 1 中的类别 1 型设备。在一个家用场景下，客户端可以被制造为
- 15 洗衣机。洗衣机除了自身实现洗衣服所具有的功能之外，还具有传感器、Wifi 模块、麦克风和扬声器等等。洗衣机的传感器可以是温度传感器，用于感测洗衣机的环境温度，或者转速传感器，用于感测洗衣机波轮或者滚筒的转速。麦克风可以录制音频数据，扬声器可以播放服务器提供的音频数据。

- 在本场景示例中，麦克风录制音频数据，通过 Wifi 模块与家庭网络互连，将音频数
- 20 据发送给服务器。服务器可以反馈音频数据给所述洗衣机，通过所述扬声器播放。

- 在本场景示例中，用户家中缺少洗衣液，需要进行购买。用户可以对洗衣机说：“你好，洗衣机”。麦克风会将该语音录制成音频信息通过 Wifi 发送给服务器。服务器接收到音频信息之后，通过语音识别发现是唤醒词。并向洗衣机发送音频信息通过扬声器播放：“您好，您需要我帮您做什么？”。用户可能说：“我需要购买洗衣液，你有品牌
- 25 推荐么？”，服务器接收到音频信息后，进行语音识别和语义分析，得到向用户推荐的
- 品牌，通过发送音频信息答复用户：“目前 XX 品牌洗衣液正在打折促销，2L 洗衣液仅花费 10 元，您需要购买么？”。用户可能回答：“好的，我购买这个打折促销的 2L 洗衣液”。服务器接收到音频信息之后，通过语音识别和语义分析得出用户同意购买该推荐的商品。

请参阅图 6。在一个场景示例中，客户端可以是具有很强数据处理能力的公共服务设备。例如，可以是上表 1 中的类别 5 型设备。在一个家用场景下，客户端可以被制造为自助服务机。

在本场景示例中，自助服务机可以提供缴纳的水电费、上网费、订阅报刊杂志等功能。用户可以在自助服务机进行注册。用户注册成功之后，自助服务机中存储用户的用户特征向量和用户录入的个人信息，以用户再次操作时根据用户特征向量验证用户的身份。

在本场景示例中，已经完成注册的用户操作自助服务机订阅报纸。用户可以对自助服务机说：“你好，我要订阅报纸。”。自助服务机录制音频信息，并进行语音识别和语音分析之后，确认用户表达的内容。并根据音频信息生成语音特征向量，将该语音特征向量与本地存储的用户特征向量进行匹配，得到与该音频特征向量对应的用户特征向量，如此确定了用户的个人信息。与该个人信息关联的表征信息表明，该用户之前一直是订阅姑苏晚报。此时，自助服务机可以回答：“你好 XX 先生，您是需要订阅姑苏晚报吗？”。用户说：“是的”。此时，自助服务机可以为该用户进行姑苏晚报的订阅流程。

请参阅图 7，本说明书一个实施方式还提供一种电子设备，所述电子设备包括语音录入单元 100、存储器 200 以及处理器 300。

其中，所述语音录入单元 100，用于接收用户语音输入的音频信息。

所述存储器 200，用于存储用户的身份信息。

所述处理器 300，用于从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。

在本实施方式中，所述电子设备可以是具备用户交互界面的电子设备。例如，所述电子设备可以是台式电脑、平板电脑、笔记本电脑、智能手机、数字助理、智能可穿戴设备、导购终端、智能电视机等。通过所述用户交互界面，用户可以向所述电子设备下达业务指令。

此外，在本实施方式中，所述电子设备还可以是不具备用户交互界面的电子设备。

例如，所述电子设备可以是智能音箱、智能麦克风、机顶盒、电视盒子等。在这种情况下，用户可以通过语音与所述电子设备进行交互。具体地，所述电子设备可以根据上述的音频信息采集方法，采集得到用户的音频信息。然后可以通过上述的识别业务功能的方法，识别出该音频信息指向的业务功能。然后，所述电子设备还可以根据该音频信息，

5 确定用户的身份信息，从而可以确定与该身份信息相关联的表征信息。所述表征信息可以包括用户偏好数据，这样，根据所述用户偏好数据，从而可以按照用户的偏好，执行该业务功能。

请参阅图 8，本说明书中还提供一种交互方法，所述方法的执行主体为客户端，所

10 述方法可以包括以下步骤。

S21: 接收用户语音输入的音频信息。

S23: 从所述音频信息识别出所述音频信息指向的业务功能信息。

S25: 根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征。

15 S27: 将所述业务功能信息和所述表征信息发送给服务器，以用于所述服务器将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

20 本说明书还提供一种电子设备，所述电子设备包括语音录入单元、存储器、网络通信端口以及处理器。

其中，所述语音录入单元，用于接收用户语音输入的音频信息。

所述存储器，用于存储用户的身份信息。

25 所述网络通信端口，用于与服务器进行数据交互。

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述业务功能信息和所述表征信息通过所述网络通信端口发送给所述服务器，以用于所述服务器将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

30

本说明书还提供一种交互方法，所述方法的执行主体可以是客户端，所述方法可以包括以下步骤。

S31：接收用户语音输入的音频信息。

5 S33：从所述音频信息识别出所述音频信息指向的业务功能信息。

S35：根据所述音频信息确定所述用户的身份信息。

S37：将所述业务功能信息和所述身份信息发送给服务器，以用于所述服务器确定与所述身份信息相关联的表征信息，以将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

10 由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

15 本说明书还提供一种电子设备，所述电子设备包括语音录入单元、存储器、网络通信端口以及处理器。

其中，所述语音录入单元，用于接收用户语音输入的音频信息。

所述存储器，用于存储用户的身份信息。

所述网络通信端口，用于与服务器进行数据交互。

20 所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；将所述业务功能信息和所述身份信息通过所述网络通信端口发送给服务器，以用于所述服务器确定与所述身份信息相关联的表征信息，以将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

25 本说明书还提供一种交互方法，所述方法的执行主体可以是客户端，所述方法包括以下步骤。

S41：接收用户语音输入的音频信息。

S43：从所述音频信息识别出所述音频信息指向的业务功能信息。

30 S45：将所述业务功能信息和所述用户的音频信息发送给服务器，以用于所述服务器根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执

行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

5

本说明书还提供一种电子设备，所述电子设备包括语音录入单元、网络通信端口以及处理器。

其中，所述语音录入单元，用于接收用户语音输入的音频信息。

所述网络通信端口，用于与服务器进行数据交互。

10

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；将所述业务功能信息和所述用户的音频信息通过所述网络通信端口发送给所述服务器，以用于所述服务器根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

15

本说明书还提供一种交互方法，所述方法的执行主体可以为客户端，所述方法包括以下步骤。

S51：接收用户语音输入的音频信息。

S53：将所述用户的音频信息发送给服务器，以用于所述服务器从所述音频信息识别出所述音频信息指向的业务功能信息，根据所述音频信息确定所述用户的身份信息，以

20

将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

25

本说明书还提供一种电子设备，所述电子设备包括语音录入单元、网络通信端口以及处理器。

其中，所述语音录入单元，用于接收用户语音输入的音频信息。

所述网络通信端口，用于与服务器进行数据交互。

30

所述处理器，用于将所述用户的音频信息通过所述网络通信端口发送给服务器，以

用于所述服务器从所述音频信息识别出所述音频信息指向的业务功能信息，根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

5 本说明书还提供一种交互方法，所述交互方法的执行主体可以是服务器，所述方法可以包括以下步骤。

S61: 接收客户端发出的音频信息。

S63: 从所述音频信息识别出所述音频信息指向的业务功能。

S65: 根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；

10 所述表征信息用于表示所述用户的特征。

S67: 将所述表征信息作为所述业务功能的输入，执行所述业务功能。

在本实施方式中，从所述音频信息识别出所述音频信息指向的业务功能包括：将所述音频信息转换为对应的文字信息；将所述文字信息与业务指令库中的业务指令进行匹配，并将匹配得到的业务指令作为所述音频信息指向的业务功能。

15 在本实施方式中，根据所述音频信息确定所述用户的身份信息包括：生成所述音频信息的语音特征向量；将所述语音特征向量与用户特征向量进行匹配；将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

在本实施方式中，在生成语音特征向量的步骤中包括：根据所述音频信息生成特征矩阵；将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的

20 维度值，所述多个维度值形成所述语音特征向量。

在本实施方式中，从所述音频信息识别出所述音频信息指向的业务功能包括：判断在所述音频信息中是否存在启动关键词；在存在所述启动关键词的情况下，才从所述音频信息识别出所述音频信息指向的业务功能。

在本实施方式中，当从所述音频信息中识别出至少两个业务功能时，所述方法还包括：向所述客户端反馈包含所述两个业务功能的提示语音信息；接收所述客户端发来的语音选择指令，并将所述语音选择指令指向的业务功能作为所述音频信息中识别出的业务功能。

25

在本实施方式中，所述表征信息中包括针对所述业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：向所述客户端反馈符合所述用户偏好数据的语音确认信息；在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好

30

数据的业务功能。

在本实施方式中，若根据所述音频信息无法确定所述用户的身份信息，所述方法还包括：调用与所述业务功能相关联的备选信息集，并将所述备选信息集反馈至所述客户端；其中，所述备选信息集中的备选信息表征所述业务功能的一种实现方式；接收所述
5 客户端发来的语音选择指令，所述语音选择指令指向所述备选信息集中的目标备选信息；根据所述目标备选信息表征的实现方式，执行所述业务功能。

由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

10

本说明书还提供一种服务器，所述服务器包括网络通信端口、存储器以及处理器。

其中，所述网络通信端口，用于接收客户端发出的音频信息。

所述存储器，用于存储用户的身份信息。

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能；根据所述
15 音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。

本说明书还提供一种交互方法，所述交互方法的执行主体可以是服务器，所述方法可以包括以下步骤。

20 S71：接收客户端发出的音频信息以及业务功能信息。

S73：根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征。

S75：将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

25 在本实施方式中，根据所述音频信息确定所述用户的身份信息包括：生成所述音频信息的语音特征向量；将所述语音特征向量与用户特征向量进行匹配；将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

在本实施方式中，在生成语音特征向量的步骤中包括：根据所述音频信息生成特征矩阵；将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的
30 维度值，所述多个维度值形成所述语音特征向量。

在本实施方式中，所述表征信息中包括针对所述业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：向所述客户端反馈符合所述用户偏好数据的语音确认信息；在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据的业务功能。

5 在本实施方式中，若根据所述音频信息无法确定所述用户的身份信息，所述方法还包括：调用与所述业务功能相关联的备选信息集，并将所述备选信息集反馈至所述客户端；其中，所述备选信息集中的备选信息表征所述业务功能的一种实现方式；接收所述客户端发来的语音选择指令，所述语音选择指令指向所述备选信息集中的目标备选信息；根据所述目标备选信息表征的实现方式，执行所述业务功能。

10 由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

本说明书还提供一种服务器，所述服务器包括网络通信端口、存储器以及处理器。

15 其中，所述网络通信端口，用于接收客户端发出的音频信息以及业务功能信息。
所述存储器，用于存储用户的身份信息。

所述处理器，用于根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

20

本说明书还提供一种交互方法，所述交互方法的执行主体可以是服务器，所述方法可以包括以下步骤。

S81：接收客户端发出的身份信息以及业务功能信息。

S83：确定与所述身份信息相关联的表征信息；所述表征信息用于表示用户的特征。

25 S85：将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

在本实施方式中，所述表征信息中包括针对业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：向所述客户端反馈符合所述用户偏好数据的语音确认信息；在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据
30 的业务功能。

由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

5 本说明书还提供一种服务器，所述服务器包括网络通信端口、存储器以及处理器。
其中，所述网络通信端口，用于接收客户端发出的身份信息以及业务功能信息。
所述存储器，用于存储表征信息。

所述处理器，用于确定与所述身份信息相关联的表征信息；所述表征信息用于表示用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行
10 所述业务功能。

本说明书还提供一种交互方法，所述交互方法的执行主体可以是服务器，所述方法可以包括以下步骤。

S91：接收客户端发出的业务功能信息和表征信息；其中，所述表征信息用于表示用
15 户的特征。

S93：将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实
20 现方式，可以参见前述实施方式的描述，这里便不再赘述。

在本实施方式中，所述表征信息中包括针对业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：向所述客户端反馈符合所述用户偏好数据的语音确认信息；在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据的业务功能。

25 由上可见，本实施方式中的各个步骤的执行过程与前述实施方式中的描述类似，只不过前述实施方式中的部分步骤的执行主体从客户端转变为服务器。各个步骤的具体实现方式，可以参见前述实施方式的描述，这里便不再赘述。

本说明书还提供一种服务器，所述服务器包括网络通信端口和处理器。

30 其中，所述网络通信端口，用于接收客户端发出的业务功能信息和表征信息；其中，

所述表征信息用于表示用户的特征。

所述处理器，用于将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

5 在本说明书中，所述语音录入单元可以将声音转换成电信号形成音频信息。所述语音录入单元可以采取电阻式麦克风、电感式麦克风、电容式麦克风、铝带式麦克风、动圈式麦克风或驻极体麦克风等形式。

所述存储器包括但不限于随机存取存储器（Random Access Memory, RAM）、只读存储器（Read-Only Memory, ROM）、缓存（Cache）、硬盘（Hard Disk Drive, HDD）
10 或者存储卡（Memory Card）。所述存储器可以用于存储计算机程序指令。网络通信单元可以是依照通信协议规定的标准设置的，用于进行网络连接通信的接口。

所述网络通信端口可以是与不同的通信协议进行绑定，从而可以发送或接收不同数据的虚拟端口。例如，所述网络通信端口可以是负责进行 web 数据通信的 80 号端口，也可以是负责进行 FTP 数据通信的 21 号端口，还可以是负责进行邮件数据通信的 25 号
15 端口。此外，所述网络通信端口还可以是实体的通信接口或者通信芯片。例如，其可以为无线移动网络通信芯片，如 GSM、CDMA 等；其还可以为 Wifi 芯片；其还可以为蓝牙芯片。

所述处理器可以按任何适当的方式实现。例如，所述处理器可以采取例如微处理器或处理器以及存储可由该（微）处理器执行的计算机可读程序代码（例如软件或固件）
20 的计算机可读介质、逻辑门、开关、专用集成电路（Application Specific Integrated Circuit, ASIC）、可编程逻辑控制器和嵌入微控制器的形式等等。

本说明书中的各个实施方式均采用递进的方式描述，各个实施方式之间相同相似的部分互相参见即可，每个实施方式重点说明的都是与其他实施方式的不同之处。

25

由上可见，本说明书一个或者多个实施方式提供的技术方案，可以将用户的身份信息
与用户的音频信息进行关联。由于不同用户的声音特质也往往不同，因此在对用户的
音频信息进行分析时，能够准确地识别出该用户的身份信息。该身份信息可以具备对应的
表征信息，该表征信息便可以表示用户的行为特征。这样，当用户通过语音下达针对
30 业务功能的执行指令时，便可以执行与该用户的表征信息相符的业务功能，从而能够准

确地向用户提供相应的业务信息。

在 20 世纪 90 年代, 对于一个技术的改进可以很明显地区分是硬件上的改进 (例如, 对二极管、晶体管、开关等电路结构的改进) 还是软件上的改进 (对于方法流程的改进)。

- 5 然而, 随着技术的发展, 当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此, 不能说一个方法流程的改进就不能用硬件实体模块来实现。例如, 可编程逻辑器件 (Programmable Logic Device, PLD) (例如现场可编程门阵列 (Field Programmable Gate Array, FPGA)) 就是这样一种集成电路, 其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统 “集成” 在一片 PLD 上, 而不需要
- 10 请芯片制造厂商来设计和制作专用的集成电路芯片。而且, 如今, 取代手工地制作集成电路芯片, 这种编程也多半改用 “逻辑编译器 (logic compiler)” 软件来实现, 它与程序开发撰写时所用的软件编译器相类似, 而要编译之前的原始代码也得用特定的编程语言来撰写, 此称之为硬件描述语言 (Hardware Description Language, HDL), 而 HDL 也
- 15 并非仅有一种, 而是有许多种, 如 ABEL (Advanced Boolean Expression Language)、AHDL (Altera Hardware Description Language)、Confluence、CUPL (Cornell University Programming Language)、HDCal、JHDL (Java Hardware Description Language)、Lava、Lola、MyHDL、PALASM、RHDL (Ruby Hardware Description Language) 等, 目前最普遍使用的是 VHDL (Very-High-Speed Integrated Circuit Hardware Description Language)
- 20 与 Verilog2。本领域技术人员也应该清楚, 只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中, 就可以很容易得到实现该逻辑方法流程的硬件电路。

- 本领域技术人员也知道, 除了以纯计算机可读程序代码方式实现控制器以外, 完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件, 而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。
- 25 或者甚至, 可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

- 通过以上的实施方式的描述可知, 本领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件平台的方式来实现。基于这样的理解, 本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来, 该计算机软件产
- 30

品可以存储在存储介质中，如 ROM/RAM、磁碟、光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施方式或者实施方式的某些部分所述的方法。

- 虽然通过实施方式描绘了本申请，本领域普通技术人员知道，本申请有许多变形和
- 5 变化而不脱离本申请的精神，希望所附的权利要求包括这些变形和变化而不脱离本申请的精神。

权利要求书

1. 一种交互方法，其特征在于，包括：

接收用户语音输入的音频信息；

从所述音频信息识别出所述音频信息指向的业务功能；

5 根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；
所述表征信息用于表示所述用户的特征；

将所述表征信息作为所述业务功能的输入，执行所述业务功能。

2. 根据权利要求 1 所述的方法，其特征在于，从所述音频信息识别出所述音频信息
10 指向的业务功能包括：

将所述音频信息转换为对应的文字信息；

将所述文字信息与业务指令库中的业务指令进行匹配，并将匹配得到的业务指令作
为所述音频信息指向的业务功能。

15 3. 根据权利要求 1 所述的方法，其特征在于，根据所述音频信息确定所述用户的身份
信息包括：

生成所述音频信息的语音特征向量；

将所述语音特征向量与用户特征向量进行匹配；

将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

20

4. 根据权利要求 3 所述的方法，其特征在于，在生成语音特征向量的步骤中包括：

根据所述音频信息生成特征矩阵；

将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的维
度值，所述多个维度值形成所述语音特征向量。

25

5. 根据权利要求 3 所述的方法，其特征在于，在生成语音特征向量的步骤中包括以
下至少之一：

任意选择一个所述用户的音频信息生成所述语音特征向量；或者，

向所述用户提供指定信息，并根据所述用户依照所述指定信息输入的语音信息生成
30 所述语音特征向量。

6. 根据权利要求 1 所述的方法, 其特征在于, 从所述音频信息识别出所述音频信息指向的业务功能包括:

判断在所述音频信息中是否存在启动关键词;

5 在存在所述启动关键词的情况下, 才从所述音频信息识别出所述音频信息指向的业务功能。

7. 根据权利要求 1 所述的方法, 其特征在于, 当从所述音频信息中识别出至少两个业务功能时, 所述方法还包括:

10 向所述用户播放包含所述两个业务功能的提示语音信息;

接收所述用户下达的语音选择指令, 并将所述语音选择指令指向的业务功能作为所述音频信息中识别出的业务功能。

8. 根据权利要求 1 所述的方法, 其特征在于, 所述表征信息中包括针对所述业务功能的用户偏好数据; 相应地, 在执行所述业务功能之前, 所述方法还包括:

15 向用户播放符合所述用户偏好数据的语音确认信息;

在接收到用户下达的语音确认指令后, 执行符合所述用户偏好数据的业务功能。

9. 根据权利要求 1 所述的方法, 其特征在于, 若根据所述音频信息无法确定所述用户

20 的身份信息, 所述方法还包括:

调用与所述业务功能相关联的备选信息集, 并向用户播放所述备选信息集中的备选信息; 其中, 所述备选信息集中的备选信息表征所述业务功能的一种实现方式;

接收用户下达的语音选择指令, 所述语音选择指令指向所述备选信息集中的目标备选信息;

25 根据所述目标备选信息表征的实现方式, 执行所述业务功能。

10. 一种电子设备, 其特征在于, 所述电子设备包括语音录入单元、存储器以及处理器, 其中:

所述语音录入单元, 用于接收用户语音输入的音频信息;

30 所述存储器, 用于存储用户的身份信息;

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。

5

11. 一种交互方法，其特征在于，包括：

接收用户语音输入的音频信息；

从所述音频信息识别出所述音频信息指向的业务功能信息；

根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；

10 所述表征信息用于表示所述用户的特征；

将所述业务功能信息和所述表征信息发送给服务器，以用于所述服务器将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

12. 根据权利要求 11 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能信息包括：

15

将所述音频信息转换为对应的文字信息；

将所述文字信息与业务指令库中的业务指令进行匹配，并将匹配得到的业务指令作为所述音频信息指向的业务功能信息。

20

13. 根据权利要求 11 所述的方法，其特征在于，根据所述音频信息确定所述用户的身份信息包括：

生成所述音频信息的语音特征向量；

将所述语音特征向量与用户特征向量进行匹配；

将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

25

14. 根据权利要求 13 所述的方法，其特征在于，在生成语音特征向量的步骤中包括：根据所述音频信息生成特征矩阵；

将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的维度值，所述多个维度值形成所述语音特征向量。

30

15. 根据权利要求 13 所述的方法，其特征在于，在生成语音特征向量的步骤中包括以下至少之一：

任意选择一个所述用户的音频信息生成所述语音特征向量；或者，

向所述用户提供指定信息，并根据所述用户依照所述指定信息输入的语音信息生成

5 所述语音特征向量。

16. 根据权利要求 11 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能信息包括：

判断在所述音频信息中是否存在启动关键词；

10 在存在所述启动关键词的情况下，才从所述音频信息识别出所述音频信息指向的业务功能信息。

17. 根据权利要求 11 所述的方法，其特征在于，当从所述音频信息中识别出至少两个业务功能信息时，所述方法还包括：

15 向所述用户播放包含所述两个业务功能信息的提示语音信息；

接收所述用户下达的语音选择指令，并将所述语音选择指令指向的业务功能信息作为所述音频信息中识别出的业务功能信息。

20 18. 一种电子设备，其特征在于，所述电子设备包括语音录入单元、存储器、网络通信端口以及处理器，其中：

所述语音录入单元，用于接收用户语音输入的音频信息；

所述存储器，用于存储用户的身份信息；

所述网络通信端口，用于与服务器进行数据交互；

25 所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述业务功能信息和所述表征信息通过所述网络通信端口发送给所述服务器，以用于所述服务器将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

30 19. 一种交互方法，其特征在于，包括：

接收用户语音输入的音频信息；

从所述音频信息识别出所述音频信息指向的业务功能信息；

根据所述音频信息确定所述用户的身份信息；

- 5 将所述业务功能信息和所述身份信息发送给服务器，以用于所述服务器确定与所述身份信息相关联的表征信息，以将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

20. 根据权利要求 19 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能信息包括：

- 10 将所述音频信息转换为对应的文字信息；

将所述文字信息与业务指令库中的业务指令进行匹配，并将匹配得到的业务指令作为所述音频信息指向的业务功能信息。

- 15 21. 根据权利要求 19 所述的方法，其特征在于，根据所述音频信息确定所述用户的身份信息包括：

生成所述音频信息的语音特征向量；

将所述语音特征向量与用户特征向量进行匹配；

将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

- 20 22. 根据权利要求 21 所述的方法，其特征在于，在生成语音特征向量的步骤中包括：根据所述音频信息生成特征矩阵；

将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的维度值，所述多个维度值形成所述语音特征向量。

- 25 23. 根据权利要求 21 所述的方法，其特征在于，在生成语音特征向量的步骤中包括以下至少之一：

任意选择一个所述用户的音频信息生成所述语音特征向量；或者，

向所述用户提供指定信息，并根据所述用户依照所述指定信息输入的语音信息生成所述语音特征向量。

24. 根据权利要求 19 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能信息包括：

判断在所述音频信息中是否存在启动关键词；

在存在所述启动关键词的情况下，才从所述音频信息识别出所述音频信息指向的业

5 务功能信息。

25. 根据权利要求 19 所述的方法，其特征在于，当从所述音频信息中识别出至少两个业务功能信息时，所述方法还包括：

向所述用户播放包含所述两个业务功能信息的提示语音信息；

10 接收所述用户下达的语音选择指令，并将所述语音选择指令指向的业务功能信息作为所述音频信息中识别出的业务功能信息。

26. 根据权利要求 19 所述的方法，其特征在于，若根据所述音频信息无法确定所述用户的身份信息，所述方法还包括：

15 调用与所述业务功能相关联的备选信息集，并向用户播放所述备选信息集中的备选信息；其中，所述备选信息集中的备选信息表征所述业务功能信息的一种实现方式；

接收用户下达的语音选择指令，所述语音选择指令指向所述备选信息集中的目标备选信息；

20 将所述业务功能信息和所述目标备选信息发送至所述服务器，以使得所述服务器根据所述目标备选信息表征的实现方式，执行所述业务功能信息表示的业务功能。

27. 一种电子设备，其特征在于，所述电子设备包括语音录入单元、存储器、网络通信端口以及处理器，其中：

所述语音录入单元，用于接收用户语音输入的音频信息；

25 所述存储器，用于存储用户的身份信息；

所述网络通信端口，用于与服务器进行数据交互；

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；根据所述音频信息确定所述用户的身份信息；将所述业务功能信息和所述身份信息通过所述网络通信端口发送给服务器，以用于所述服务器确定与所述身份信息相关联的表征信息，以将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业

30

务功能。

28. 一种交互方法，其特征在于，包括：

接收用户语音输入的音频信息；

5 从所述音频信息识别出所述音频信息指向的业务功能信息；

将所述业务功能信息和所述用户的音频信息发送给服务器，以用于所述服务器根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

10 29. 根据权利要求 28 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能信息包括：

将所述音频信息转换为对应的文字信息；

将所述文字信息与业务指令库中的业务指令进行匹配，并将匹配得到的业务指令作为所述音频信息指向的业务功能信息。

15

30. 根据权利要求 28 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能信息包括：

判断在所述音频信息中是否存在启动关键词；

在存在所述启动关键词的情况下，才从所述音频信息识别出所述音频信息指向的业务功能信息。

20

31. 根据权利要求 28 所述的方法，其特征在于，当从所述音频信息中识别出至少两个业务功能信息时，所述方法还包括：

向所述用户播放包含所述两个业务功能信息的提示语音信息；

25 接收所述用户下达的语音选择指令，并将所述语音选择指令指向的业务功能信息作为所述音频信息中识别出的业务功能信息。

32. 一种电子设备，其特征在于，所述电子设备包括语音录入单元、网络通信端口以及处理器，其中：

30 所述语音录入单元，用于接收用户语音输入的音频信息；

所述网络通信端口，用于与服务器进行数据交互；

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能信息；将所述业务功能信息和所述用户的音频信息通过所述网络通信端口发送给所述服务器，以用于所述服务器根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

33. 一种交互方法，其特征在于，包括：

接收用户语音输入的音频信息；

将所述用户的音频信息发送给服务器，以用于所述服务器从所述音频信息识别出所述音频信息指向的业务功能信息，根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

34. 一种电子设备，其特征在于，所述电子设备包括语音录入单元、网络通信端口以及处理器，其中：

所述语音录入单元，用于接收用户语音输入的音频信息；

所述网络通信端口，用于与服务器进行数据交互；

所述处理器，用于将所述用户的音频信息通过所述网络通信端口发送给服务器，以用于所述服务器从所述音频信息识别出所述音频信息指向的业务功能信息，根据所述音频信息确定所述用户的身份信息，以将所述身份信息关联的表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

35. 一种交互方法，其特征在于，包括：

接收客户端发出的音频信息；

从所述音频信息识别出所述音频信息指向的业务功能；

根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；

将所述表征信息作为所述业务功能的输入，执行所述业务功能。

36. 根据权利要求 35 所述的方法，其特征在于，从所述音频信息识别出所述音频信

息指向的业务功能包括：

将所述音频信息转换为对应的文字信息；

将所述文字信息与业务指令库中的业务指令进行匹配，并将匹配得到的业务指令作为所述音频信息指向的业务功能。

5

37. 根据权利要求 35 所述的方法，其特征在于，根据所述音频信息确定所述用户的身份信息包括：

生成所述音频信息的语音特征向量；

将所述语音特征向量与用户特征向量进行匹配；

10

将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

38. 根据权利要求 37 所述的方法，其特征在于，在生成语音特征向量的步骤中包括：根据所述音频信息生成特征矩阵；

15 将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的维度值，所述多个维度值形成所述语音特征向量。

39. 根据权利要求 35 所述的方法，其特征在于，从所述音频信息识别出所述音频信息指向的业务功能包括：

判断在所述音频信息中是否存在启动关键词；

20

在存在所述启动关键词的情况下，才从所述音频信息识别出所述音频信息指向的业务功能。

40. 根据权利要求 35 所述的方法，其特征在于，当从所述音频信息中识别出至少两个业务功能时，所述方法还包括：

25

向所述客户端反馈包含所述两个业务功能的提示语音信息；

接收所述客户端发来的语音选择指令，并将所述语音选择指令指向的业务功能作为所述音频信息中识别出的业务功能。

30 41. 根据权利要求 35 所述的方法，其特征在于，所述表征信息中包括针对所述业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：

向所述客户端反馈符合所述用户偏好数据的语音确认信息；

在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据的业务功能。

5 42. 根据权利要求 35 所述的方法，其特征在于，若根据所述音频信息无法确定所述用户的身份信息，所述方法还包括：

调用与所述业务功能相关联的备选信息集，并将所述备选信息集反馈至所述客户端；其中，所述备选信息集中的备选信息表征所述业务功能的一种实现方式；

10 接收所述客户端发来的语音选择指令，所述语音选择指令指向所述备选信息集中的目标备选信息；

根据所述目标备选信息表征的实现方式，执行所述业务功能。

43. 一种服务器，其特征在于，所述服务器包括网络通信端口、存储器以及处理器，其中：

15 所述网络通信端口，用于接收客户端发出的音频信息；

所述存储器，用于存储用户的身份信息；

所述处理器，用于从所述音频信息识别出所述音频信息指向的业务功能；根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为所述业务功能的输入，执行所述业务功能。

20

44. 一种交互方法，其特征在于，包括：

接收客户端发出的音频信息以及业务功能信息；

根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；

25 将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

45. 根据权利要求 44 所述的方法，其特征在于，根据所述音频信息确定所述用户的身份信息包括：

30 生成所述音频信息的语音特征向量；

将所述语音特征向量与用户特征向量进行匹配；

将匹配得到的用户特征向量关联的身份信息作为所述用户的身份信息。

46. 根据权利要求 45 所述的方法，其特征在于，在生成语音特征向量的步骤中包括：

5 根据所述音频信息生成特征矩阵；

将所述特征矩阵按照多个特征维度进行降维处理，得到多个用于表征特征维度的维度值，所述多个维度值形成所述语音特征向量。

47. 根据权利要求 44 所述的方法，其特征在于，所述表征信息中包括针对所述业务

10 功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：

向所述客户端反馈符合所述用户偏好数据的语音确认信息；

在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据的业务功能。

15 48. 根据权利要求 44 所述的方法，其特征在于，若根据所述音频信息无法确定所述用户的身份信息，所述方法还包括：

调用与所述业务功能相关联的备选信息集，并将所述备选信息集反馈至所述客户端；其中，所述备选信息集中的备选信息表征所述业务功能的一种实现方式；

20 接收所述客户端发来的语音选择指令，所述语音选择指令指向所述备选信息集中的目标备选信息；

根据所述目标备选信息表征的实现方式，执行所述业务功能。

49. 一种服务器，其特征在于，所述服务器包括网络通信端口、存储器以及处理器，其中：

25 所述网络通信端口，用于接收客户端发出的音频信息以及业务功能信息；

所述存储器，用于存储用户的身份信息；

所述处理器，用于根据所述音频信息确定用户的身份信息；其中，所述身份信息与表征信息关联；所述表征信息用于表示所述用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

50. 一种交互方法，其特征在于，包括：

接收客户端发出的身份信息以及业务功能信息；

确定与所述身份信息相关联的表征信息；所述表征信息用于表示用户的特征；

将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务
5 功能。

51. 根据权利要求 50 所述的方法，其特征在于，所述表征信息中包括针对业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：

向所述客户端反馈符合所述用户偏好数据的语音确认信息；

10 在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据的业务功能。

52. 一种服务器，其特征在于，所述服务器包括网络通信端口、存储器以及处理器，其中：

15 所述网络通信端口，用于接收客户端发出的身份信息以及业务功能信息；

所述存储器，用于存储表征信息；

所述处理器，用于确定与所述身份信息相关联的表征信息；所述表征信息用于表示用户的特征；将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行
20 所述业务功能。

53. 一种交互方法，其特征在于，包括：

接收客户端发出的业务功能信息和表征信息；其中，所述表征信息用于表示用户的特征；

将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务
25 功能。

54. 根据权利要求 53 所述的方法，其特征在于，所述表征信息中包括针对业务功能的用户偏好数据；相应地，在执行所述业务功能之前，所述方法还包括：

向所述客户端反馈符合所述用户偏好数据的语音确认信息；

30 在接收到所述客户端发来的语音确认指令后，执行符合所述用户偏好数据的业务功

能。

55. 一种服务器，其特征在于，所述服务器包括网络通信端口和处理器，其中：

所述网络通信端口，用于接收客户端发出的业务功能信息和表征信息；其中，所述

5 表征信息用于表示用户的特征；

所述处理器，用于将所述表征信息作为执行所述业务功能信息表示的业务功能的输入，执行所述业务功能。

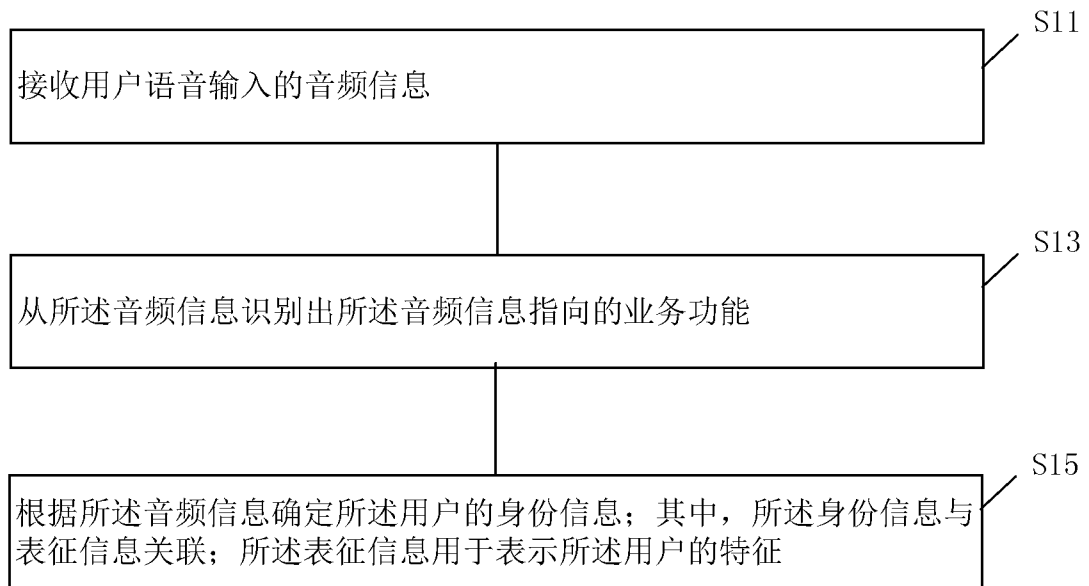


图 1

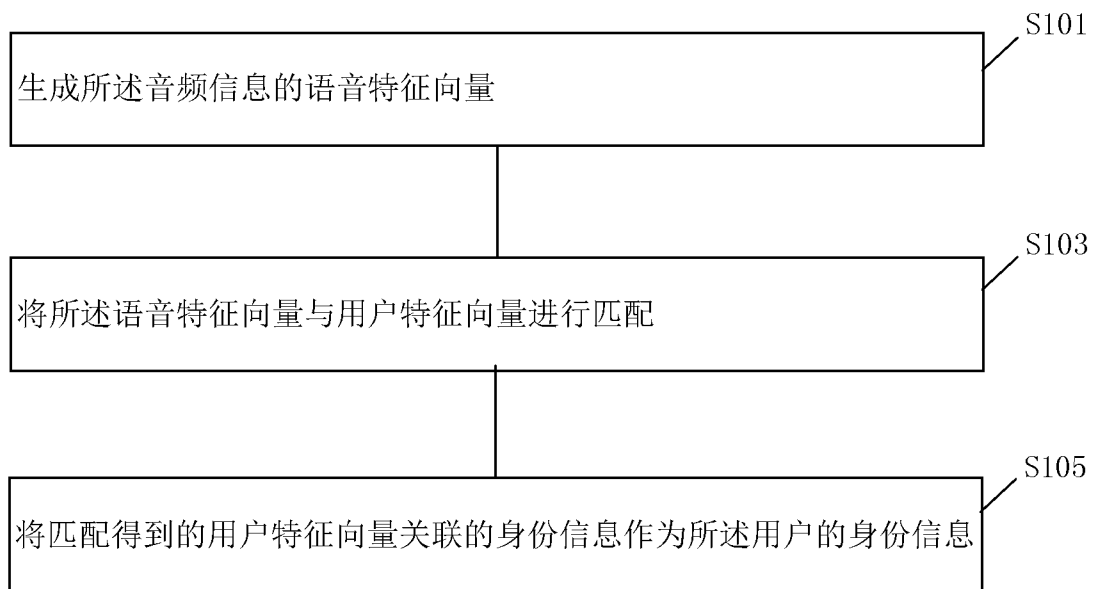


图 2

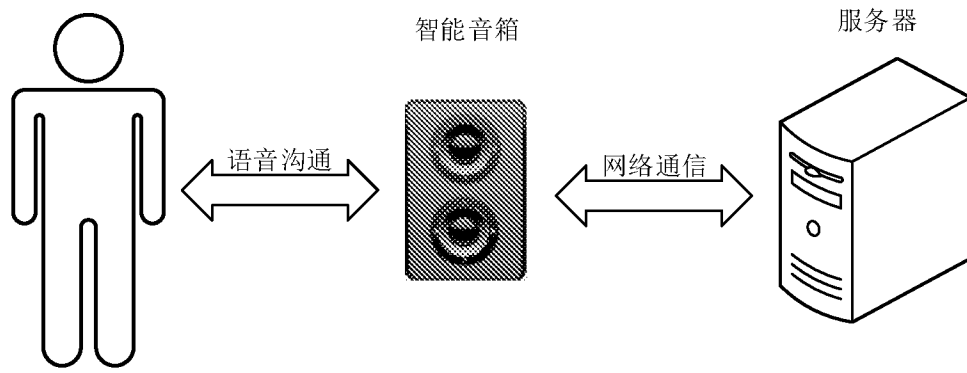


图 3

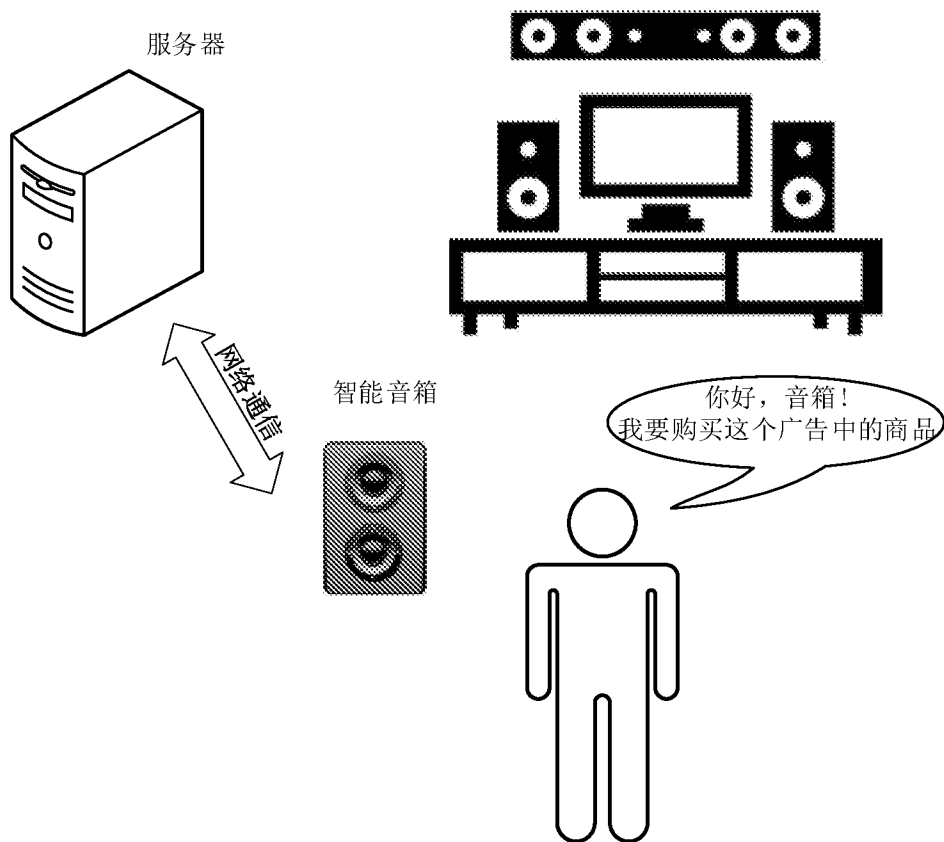


图 4

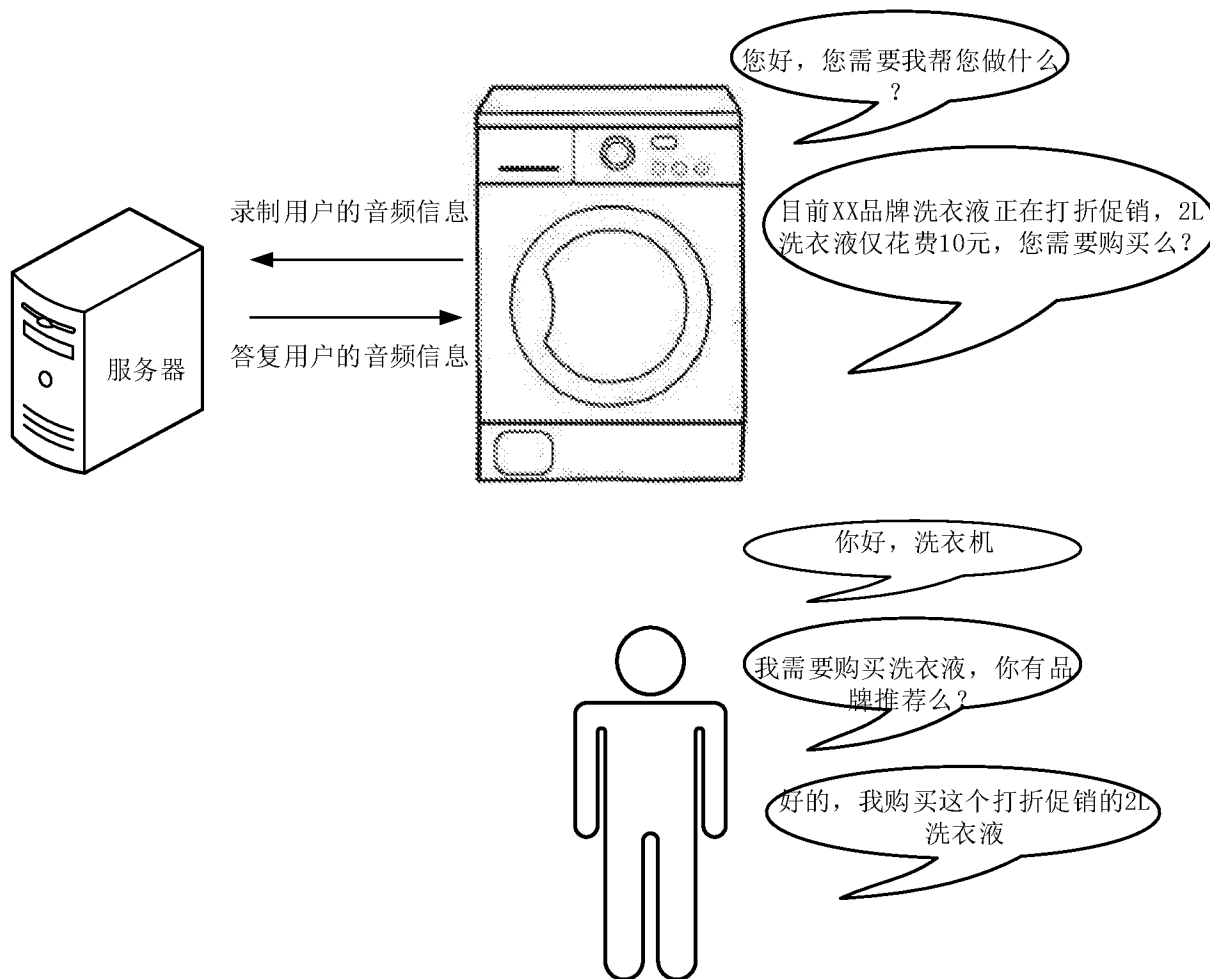


图 5

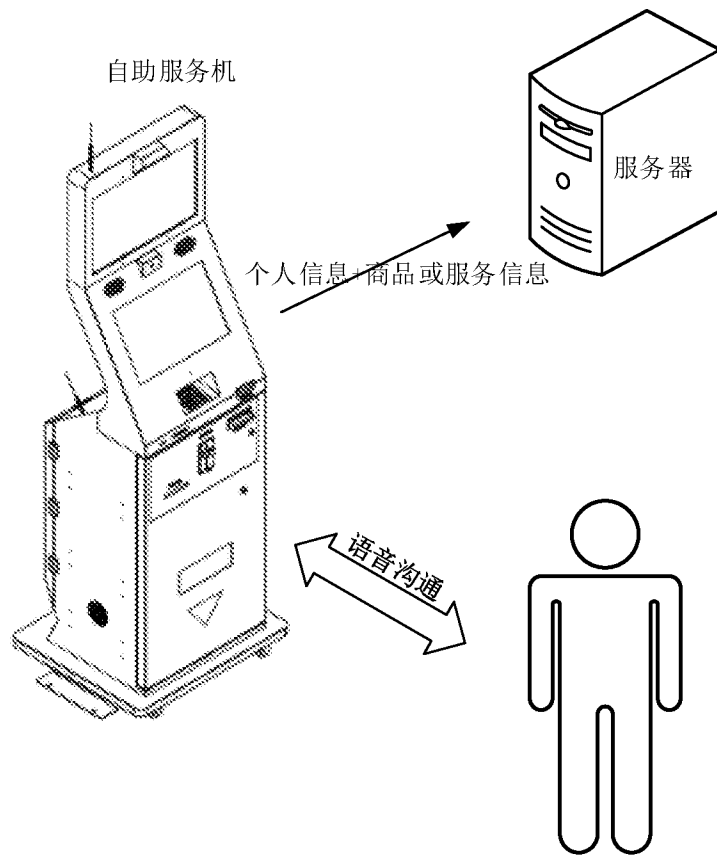


图 6

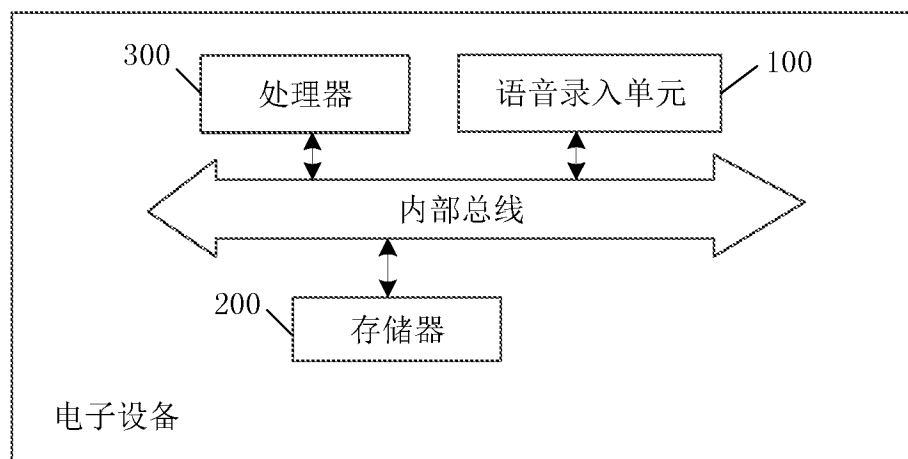


图 7

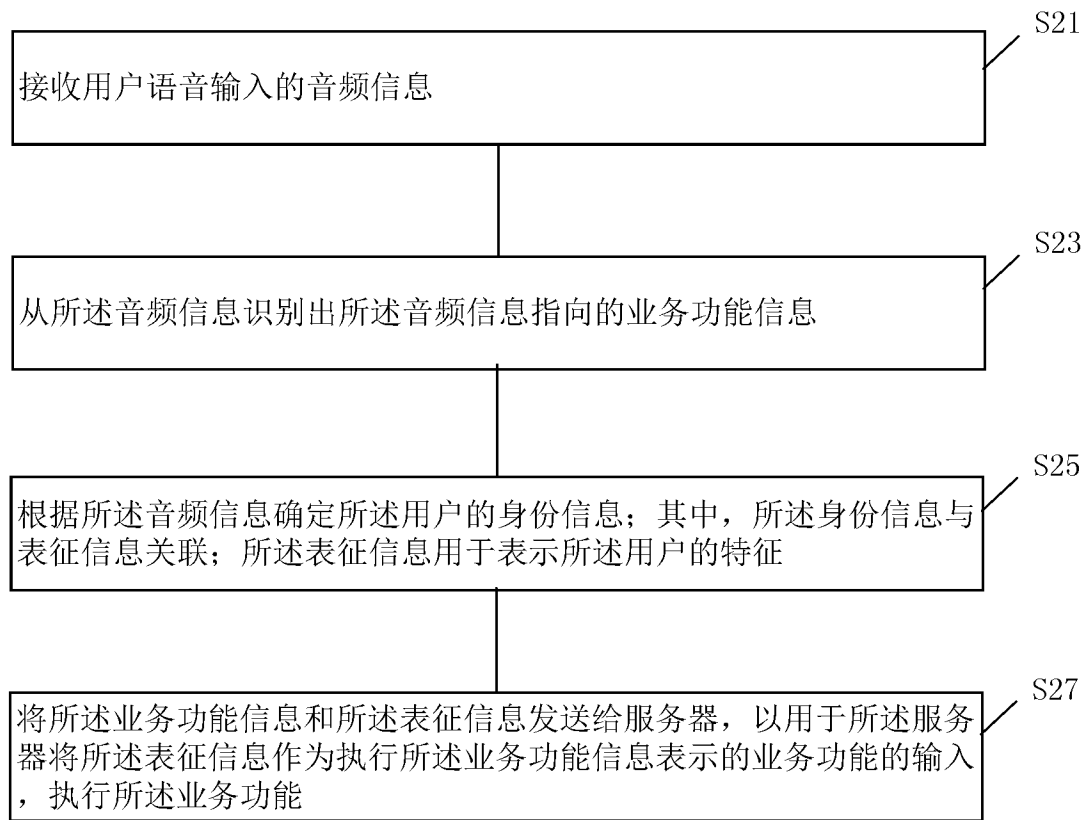


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/093231

A. CLASSIFICATION OF SUBJECT MATTER

G10L 17/22(2013.01)i; G06F 21/32(2013.01)i; H04L 9/32(2006.01)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)

G10L; G06F; H04L

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, WANFANG, VEN: 阿里巴巴, 谷歌, 交互, 智能, 语音, 音频, 声音, 声纹, 用户, 账户, 身份, 画像, 偏好, 喜好, 习惯, 历史, 个性, 命令, 指令, 口令, interaction, intelligence, intelligent, speech, voice, sound, voice print, user, account, identity, representation, preference, fancy, custom, habit, history, individuality, personality, selfhood, instruction, command, order

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107481720 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 15 December 2017 (2017-12-15) description, paragraphs [0089]-[0159], and figure 1	1-10
X	CN 105374355 A (KOMOO INTELLIGENT TECHNOLOGY CO., LTD.) 02 March 2016 (2016-03-02) description, paragraphs [0051]-[0092], and figures 1-2	1-55
X	CN 105810200 A (SHENZHEN YONGYIDA ROBOT CO., LTD.) 27 July 2016 (2016-07-27) description, paragraphs [0036]-[0047], and figure 2	1-10
X	CN 105719646 A (SHI, WEITING) 29 June 2016 (2016-06-29) description, paragraphs [0003] and [0037]-[0052], and figure 2	1-10
A	CN 103761462 A (ANHUI USTC IFLYTEK CO., LTD.) 30 April 2014 (2014-04-30) entire document	1-55
A	CN 106875944 A (SHENZHEN FIGIGANTIC ELECTRONIC CO., LTD.) 20 June 2017 (2017-06-20) entire document	1-55

☒ 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

28 September 2018

Date of mailing of the international search report

09 October 2018

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

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/093231

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 106682090 A (SHANGHAI ZHIZHEN INTELLIGENT NETWORK TECHNOLOGY CO., LTD.) 17 May 2017 (2017-05-17) entire document	1-55

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/093231

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	107481720	A	15 December 2017	None	
CN	105374355	A	02 March 2016	None	
CN	105810200	A	27 July 2016	None	
CN	105719646	A	29 June 2016	None	
CN	103761462	A	30 April 2014	CN 103761462	B 12 October 2016
CN	106875944	A	20 June 2017	None	
CN	106682090	A	17 May 2017	None	

A. 主题的分类 G10L 17/22(2013.01)i; G06F 21/32(2013.01)i; H04L 9/32(2006.01)i; G10L 15/22(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; G06F; H04L 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, VEN:阿里巴巴, 许毅, 交互, 智能, 语音, 音频, 声音, 声纹, 用户, 账户, 身份, 画像, 偏好, 喜好, 习惯, 历史, 个性, 命令, 指令, 口令, interaction, intelligence, intelligent, speech, voice, sound, voice print, user, account, identity, representation, preference, fancy, custom, habit, history, individuality, personality, selfhood, instruction, command, order		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 107481720 A (百度在线网络技术北京有限公司) 2017年 12月 15日 (2017 - 12 - 15) 说明书[0089]-[0159]段及图1	1-10
X	CN 105374355 A (厦门科牧智能技术有限公司) 2016年 3月 2日 (2016 - 03 - 02) 说明书第[0051]-[0092]段及图1-2	1-55
X	CN 105810200 A (深圳前海勇艺达机器人有限公司) 2016年 7月 27日 (2016 - 07 - 27) 说明书第[0036]-[0047]段及图2	1-10
X	CN 105719646 A (史唯廷) 2016年 6月 29日 (2016 - 06 - 29) 说明书第[0003]、[0037]-[0052]段及图2	1-10
A	CN 103761462 A (安徽科大讯飞信息科技股份有限公司) 2014年 4月 30日 (2014 - 04 - 30) 全文	1-55
A	CN 106875944 A (深圳市丰巨泰科电子有限公司) 2017年 6月 20日 (2017 - 06 - 20) 全文	1-55
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2018年 9月 28日		2018年 10月 9日
ISA/CN的名称和邮寄地址		受权官员
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		喻新 电话号码 86-010-62085168

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 106682090 A（上海智臻智能网络科技股份有限公司）2017年 5月 17日（2017 - 05 - 17） 全文	1-55

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/093231

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	107481720	A	2017年 12月 15日	无	
CN	105374355	A	2016年 3月 2日	无	
CN	105810200	A	2016年 7月 27日	无	
CN	105719646	A	2016年 6月 29日	无	
CN	103761462	A	2014年 4月 30日	CN 103761462	B 2016年 10月 12日
CN	106875944	A	2017年 6月 20日	无	
CN	106682090	A	2017年 5月 17日	无	