

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 6 月 11 日 (11.06.2020)



(10) 国际公布号
WO 2020/114384 A1

(51) 国际专利分类号:
G10L 15/26 (2006.01) *G10L 17/04* (2013.01)
G10L 15/22 (2006.01) *G10L 17/22* (2013.01)

(21) 国际申请号: PCT/CN2019/122640

(22) 国际申请日: 2019 年 12 月 3 日 (03.12.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811488466.5 2018 年 12 月 6 日 (06.12.2018) 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: 孙尧 (SUN, Yao); 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,

(54) Title: VOICE INTERACTION METHOD AND DEVICE

(54) 发明名称: 一种语音交互方法和装置

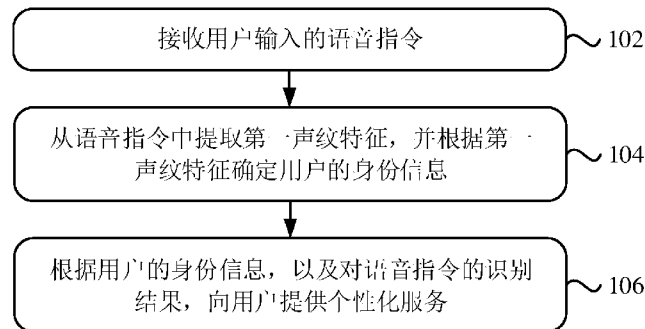


图 1

- 102 Receive a voice command inputted by a user
- 104 Acquire a first voice print feature from the voice command and determine user identity information according to the first voice print feature
- 106 Provide a personalized service to the user according to the user identity information and an identification result of the voice command

(57) Abstract: Provided are a voice interaction and device. The method comprises: receiving a voice command inputted by a user (102); acquiring a first voice print feature from the voice command and determining user identity information according to the first voice print feature (104); and providing a personalized service to the user according to the user identity information and an identification result of the voice command (106).

(57) 摘要: 一种语音交互方法和装置, 方法包括: 接收用户输入的语音指令 (102); 从语音指令中提取第一声纹特征, 并根据第一声纹特征确定用户的身份信息 (104); 根据用户的身份信息, 以及对语音指令的识别结果, 向用户提供个性化服务 (106)。

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))。

一种语音交互方法和装置

本申请要求 2018 年 12 月 06 日递交的申请号为 201811488466.5、发明名称为“一种语音交互方法和装置”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

本申请涉及计算机技术领域，尤其涉及一种语音交互方法和装置。

背景技术

随着科技的发展，智能设备的智能化程度越来越高。实际应用中，智能设备可以根据用户的历史使用数据、习惯、偏好等，为用户提供个性化服务。但是，针对于远场语音交互的共享型智能设备，由于面向的用户可能为多个，如何为每个用户都可以提供个性化服务，是当前需要考虑的问题。

因此，需要一种更加有效的语音交互方法。

15 发明内容

本说明书实施例提供一种语音交互方法和设备，用于实现共享型智能设备为不同用户提供符合用户需求的个性化服务。

第一方面，本说明书实施例提供了一种语音交互方法，包括：

接收用户输入的语音指令；

20 从所述语音指令中提取第一声纹特征，并根据所述第一声纹特征确定所述用户的身份信息；

根据所述用户的身份信息，以及对所述语音指令的识别结果，向所述用户提供个性化服务。

第二方面，本说明书实施例还提供了一种语音交互装置，用于执行如第一方面所述的语音交互方法，所述装置包括：

接收模块，接收用户输入的语音指令；

确定模块，从所述语音指令中提取第一声纹特征，并根据所述第一声纹特征确定所述用户的身份信息；

30 服务模块，根据所述用户的身份信息，以及对所述语音指令的识别结果，向所述用户提供个性化服务。

第三方面，本说明书实施例还提供了一种电子设备，包括：

存储器，存放程序；

处理器，执行所述存储器存储的程序，并具体执行如第一方面所述的语音交互方法。

第四方面，本说明书实施例还提供了一种计算机可读存储介质，所述计算机可读存储介质存储一个或多个程序，所述一个或多个程序当被包括多个应用程序的电子设备执行时，使得所述电子设备执行如第一方面所述的语音交互方法。

本说明书实施例采用的上述至少一个技术方案能够达到以下有益效果：

接收用户输入的语音指令，从语音指令中提取第一声纹特征，并根据第一声纹特征确定用户的身份信息，进而根据用户的身份信息，以及对语音指令的识别结果，向用户提供个性化服务，从而实现共享型智能设备为不同用户提供符合用户需求的个性化服务。

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图 1 为本说明书实施例提供的一种语音交互方法的流程示意图；

图 2 为本说明书实施例提供的 UP 模块的示意图；

图 3 为本说明书实施例提供的一种电子设备的结构示意图；

图 4 为本说明书实施例提供的一种语音交互装置的结构示意图。

具体实施方式

下面结合本说明书具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本说明书中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

以下结合附图，详细说明本说明书各实施例提供的技术方案。

图 1 为本说明书实施例提供的一种语音交互方法的流程示意图。所述方法可以如下所示。

步骤 102，接收用户输入的语音指令。

步骤 104，从语音指令中提取第一声纹特征，并根据第一声纹特征确定用户的身份信息。

步骤 106, 根据用户的身份信息, 以及对语音指令的识别结果, 向用户提供个性化服务。

针对于远场语音交互的共享型智能设备, 麦克风模块接收到用户输入的语音指令之后, 将语音指令传输至用户画像 (UP, User Profile) 模块, 其中, UP 模块部署在共享型智能设备中, 或者部署在共享型智能设备对应的私有云服务器上。

声纹 (VP, Voice Print) 是生物特征的一种, 用于描述携带言语信息的声波频谱。鉴于每个人的声纹特征的相对稳定性和强区分性, 可以将声纹特征作为识别一个人的重要特征。

UP 模块从语音指令中提取第一声纹特征, 并根据第一声纹特征来确定用户的身份信息, 进而可以实现根据用户的身份信息, 为不同用户提供符合用户需求的个性化服务。

本说明书实施例中, 根据第一声纹特征确定用户的身份信息, 包括:

确定第一声纹特征数据库, 第一声纹特征数据库中包括多个用户身份标识, 以及与多个用户身份标识中任一用户身份标识对应的第二声纹特征;

将第一声纹特征与多个用户身份标识中任一用户身份标识对应的第二声纹特征进行匹配;

根据匹配结果, 确定用户对应的目标用户身份标识, 目标用户身份标识用于表示用户的身份信息;

第一声纹特征与第二声纹特征为相同或不同的声纹特征。

其中, 通过以下方式确定得到第一声纹特征数据库:

获取历史语音指令集合, 历史语音指令集合中包括多个历史语音指令;

提取多个历史语音指令中任一历史语音指令对应的第三声纹特征;

根据多个历史语音指令中任一历史语音指令对应的第三声纹特征, 对多个历史语音指令进行声纹聚类, 得到多个用户身份标识, 以及与多个用户身份标识中任一用户身份标识对应的历史语音指令;

根据多个用户身份标识中任一用户身份标识对应的历史语音指令, 确定该用户身份标识对应的第二声纹特征;

第二声纹特征和第三声纹特征为相同或不同的声纹特征。

其中, 声纹特征包括下述至少一种:

语言、音色、音调、语速、口音, 和频谱。

为了可以根据用户的声纹特征来确定用户的身份, UP 模块首先根据共享型智能设备

中的历史语音指令集合，通过声纹聚类的方式，构建包括多个用户身份标识，以及任一用户身份标识对应的第二声纹特征的第一声纹特征数据库（也可以称为，声纹池）。

图 2 为本说明书实施例提供的 UP 模块的示意图。

如图 2 所示，UP 模块 200 至少包括：声纹提取单元 201、声纹聚类模型 202、第一声纹特征数据库 203、UP 数据单元 204、自动语言识别（ASR，Automatic Speech Recognition）单元 205、获得标识单元 206、新增标识单元 207。

UP 模块 200 构建第一声纹特征数据库 203 的过程：

首先，UP 模块 200 获取同一共享型智能设备中的历史语音指令集合，历史语音指令集合中包括该共享型智能设备上所有的历史语音指令；

其次，UP 模块 200 中的声纹提取单元 201，从历史语音指令集合中的多个历史语音指令中，提取任一历史语音指令对应的第三声纹特征，第三声纹特征包括但不限于：语言、音色、音调、语速、口音、频谱，等。

然后，UP 模块 200 中的声纹聚类模型 202，根据任一历史语音指令对应的第三声纹特征进行声纹聚类，将声纹特征相似度很高的历史语音指令归为一类，确定为来自同一个用户的历史语音指令，并为每一类历史语音指令打上一个用户标签，即用户身份标识。

例如，设备 abc-用户 1 的历史语音指令，设备 abc-用户 2 的历史语音指令，等。

最后，UP 模块 200 中的声纹提取单元 201，根据一个用户身份标识对应的历史语音指令，确定该用户身份标识对应的第二声纹特征。进而将任一用户身份标识，以及该用户身份标识对应的第二声纹特征存储在第一声纹特征数据库 203 中。

例如，设备 abc-用户 1 对应的第二声纹特征（语言、音色、音调、语速、口音、频谱，等），设备 abc-用户 2 对应的第二声纹特征（语言、音色、音调、语速、口音、频谱，等），等。

在一实施例中，对于很久没有使用共享型智能设备的用户，为了节约空间，UP 模块 200 会在预设时长后删除该用户的数据，例如，该用户对应的历史语音指令，该用户对应的用户身份标识，以及该用户身份标识对应的第二声纹特征，等。

在另一实施例中，随着时间的变化，一个人年龄的增长或生理发育的影响，都可能会导致声音的变化。因此，UP 模块 200 中的声纹聚类模型 202 会定期更新第一声纹数据库 203 中存储的用户身份标识对应的第二声纹特征。

本说明书实施例中，还包括：

根据用户身份标识对应的第二声纹特征，确定用户身份标识对应的第一用户画像；

用户身份标识对应的第一用户画像，用于反映与用户身份标识对应的用户的年龄和/或性别。

具体地，根据用户身份标识对应的第二声纹特征，确定用户身份标识对应的第一用户画像，包括：

- 5 确定第二声纹特征数据库，第二声纹特征数据库中包括不同用户画像对应的第四声纹特征；

将用户身份标识对应的第二声纹特征，与不同用户画像对应的第四声纹特征进行匹配；

根据匹配结果，确定用户身份标识对应的第一用户画像；

- 10 其中，第二声纹特征与第四声纹特征为相同或不同的声纹特征。

第二声纹特征数据库，也即共享型智能设备对应的公有云服务器中的公共声纹特征数据库，其中包括不同用户画像对应的第四声纹特征，例如，不同性别用户对应的第四声纹特征、不同年龄的用户对应的第四声纹特征，等。

- 15 仍以上述图 2 为例，针对 UP 模块 200 中第一声纹特征数据库 203 中存储的多个用户身份标识以及任一用户身份标识对应的第二声纹特征，UP 模块 200 中的声纹聚类模型 202，将任一用户身份标识对应的第二声纹特征与第二声纹特征数据库中不同用户画像对应的第四声纹特征进行对比，从而判断该用户身份标识对应的第一用户画像（性别和/或年龄），并将任一用户身份标识对应的第一用户画像存储在 UP 模块 200 中的 UP 数据单元 204 中。

- 20 例如，UP 数据单元 204 中存储有，设备 abc-用户 1：女性、年轻；设备 abc-用户 2：男性、老年。

本说明书实施例中，还包括：

根据用户身份标识对应的历史语音指令，确定用户身份标识对应的第二用户画像；

- 25 用户身份标识对应的第二用户画像，用于反映与用户身份标识对应的用户的个性化需求。

具体地，根据用户身份标识对应的历史语音指令，确定用户身份标识对应的第二用户画像，包括：

将用户身份标识对应的历史语音指令转换为文本指令；

根据预设规则对文本指令进行分析；

- 30 根据分析结果，确定用户身份标识对应的第二用户画像。

仍以图 2 为例，为了更好地为不同用户提供符合用户需求的个性化服务，UP 模块 200 确定与用户的个性化需求相关的第二用户画像。

具体地，首先，UP 模块 200 中的 ASR 单元 205，将同一用户身份标识对应的历史语音指令转换为文本指令。例如，设备 abc-用户 1 对应的文本指令为：20180618 播放周杰伦的青花瓷、20180619 播放毛不易，等；设备 abc-用户 2 对应的文本指令为：20180618 播放邓丽君、20180619 播放陈百强，等。

然后，UP 模块 200 中的 UP 数据单元 204，根据预设规则，或特定的模型，对任一用户身份标识对应的文本指令进行分析，确定该用户身份标识对应，用于反映与该用户身份标识对应的用户的个性化需求的第二用户画像，以及将任一用户身份标识对应的第二用户画像补充存储在 UP 数据单元 204 中。

例如，UP 数据单元 204 中存储有，设备 abc-用户 1：女性、年轻、喜欢流行歌曲、特别喜欢周杰伦和毛不易；设备 abc-用户 2：男性、老年、喜欢老歌、特别喜欢邓丽君和陈百强；等。

随着时间的变化，不同用户的喜好可能会发生变化，因此，UP 模块会根据每个用户身份标识对应的历史语音指令，定期更新用户身份标识对应的第二用户画像。

仍以图 2 为例，当共享型智能设备中的麦克风模块接收到当前用户输入的语音指令之后，麦克风模块将该语音指令传输至 UP 模块 200，UP 模块 200 中的声纹提取单元 201 从该语音指令中提取第一声纹特征，UP 模块中的获得标识单元 206，将第一声纹特征与 UP 模块 200 中第一声纹特征数据库 203 中存储的，多个用户身份标识中任一用户身份标识对应的第二声纹特征，进行匹配，以及根据匹配结果，确定当前用户对应的目标用户身份标识，进而为当前用户提供个性化服务。

根据匹配结果确定用户对应的目标用户身份标识，进而为当前用户提供个性化服务的具体过程包括下述至少两种方式。

第一种：

本说明书实施例中，根据匹配结果，确定用户对应的目标用户身份标识，包括：

当第一声纹特征与多个用户身份标识中的一个用户身份标识对应的第二声纹特征之间的匹配度大于预设阈值时，将用户身份标识确定为目标用户身份标识。

仍以图 2 为例，UP 模块 200 中的获得标识单元 206，将声纹提取单元 201 从当前用户的语音指令中提取的第一声纹特征，与第一声纹特征数据库 203 中存储的任一用户身份标识对应的第二声纹特征进行匹配，若第一声纹特征与设备 abc-用户 1 对应的第

二声纹特征之间的匹配度大于预设阈值，因此，获得标识单元 206 可以确定当前用户的目标用户身份标识为设备 abc-用户 1。

针对于单条语音指令，采用上述根据声纹特征来确定用户身份的方式，使得召回率可以达到 80%，准确率可以达到 85%。其中，若可以排除音量过小、有噪音等干扰情况，

5 召回率还会有所提高。

需要说明的是，若用户注册了声纹，则根据声纹特征确定用户身份信息的准确率将大大提高。

本说明书实施例中，根据用户的身份信息，以及对语音指令的识别结果，向用户提供个性化服务，包括：

10 确定目标用户身份标识对应的第一用户画像和/或第二用户画像；

根据目标用户身份标识对应的第一用户画像和第二用户画像，以及对语音指令的识别结果，向用户提供个性化服务。

仍以上述图 2 为例，获得标识单元 206 确定当前用户的目标用户身份标识之后，UP 模块 200 从 UP 数据单元 204 中，确定与目标用户身份标识对应的第一用户画像和/或第二用户画像。根据与目标用户身份标识对应的第一用户画像和/或第二用户画像，可以了
15 解到当前用户的年龄、性别、个性化需求，等。

若 UP 模块 200 为部署在共享型智能设备中的模块，则 UP 模块 200 将当前用户的语音指令、第一声纹特征、第一用户画像和/或第二用户画像等信息传输至共享型智能设备对应的公有云服务器中的用户画像决策（UPD，User Profile Decision）模块，使得 UPD
20 模块根据上述相关信息，为当前用户提供个性化服务。

需要说明的是，UPD 模块为当前用户提供个性化服务的过程中，当前用户的隐私信息不会泄露到公有云服务器中，可以保证用户使用共享型智能设备的隐私安全性。

针对于单条语音指令，采用上述根据确定当前用户的用户画像的方式，在当前用户对应的历史语音指令数据充足的情况下（例如，一个月内有超过 10 条的历史语音指令），
25 召回率可以达到 85%，准确率可以达到 90%。

本说明书实施例中，图 1 所示的语音交互方法应用于智能音箱，向用户提供个性化服务包括下述至少一种：

音乐推荐和聊天。

当共享型智能设备为智能音箱时，智能音箱可以为用户提供音乐推荐、聊天等个性
30 化服务。

当多个用户共用一个智能音箱时，每个人有自己的个性化需求和点播记录。例如，老人爱听邓丽君，年轻人爱听周杰伦，儿童爱听儿歌，等。因此，部署在智能音箱中的 UP 模块或部署在智能音箱对应的私有云服务器上的 UP 模块，确定使用智能音箱的每个用户对应的第二声纹特征、第一用户画像（年龄、性别）、第二用户画像（个人爱好），
5 等。

在一实施例中，当接收到当前用户的语音指令“我想听歌”时，UP 模块可以根据从当前用户的语音指令中提取的第一声纹特征，确定当前用户的身份信息，从而确定当前用户的第二用户画像（个人爱好）。使得 UPD 模块可以根据当前用户的第二用户画像（个人爱好），为当前用户播放符合用户个性化需求的歌曲。

10 例如，为喜欢周杰伦的用户播放周杰伦的歌曲，为喜欢邓丽君的用户播放邓丽君的歌曲，等。

在另一实施例中，当接收到当前用户的语音指令“我想听歌”时，UP 模块可以根据从当前用户的语音指令中提取的第一声纹特征，确定当前用户的身份信息，从而确定当前用户的第一用户画像（年龄）、第二用户画像（个人爱好）。使得 UPD 模块可以根据可以当前用户的第一用户画像（年龄）和第二用户画像（个人爱好），为当前用户播放符合用户个性化需求的歌曲，以及主动问候当前用户。
15

例如，当前用户为儿童时，UPD 模块为当前用户播放儿歌，并可以在播放儿歌之前主动问候当前用户“宝贝，下面给你播放儿歌呦”。

在另一实施例中，当接收到当前用户的语音指令“今天心情不好”时，UP 模块可以根据从当前用户的语音指令中提取的第一声纹特征，确定当前用户的身份信息，从而确定当前用户的第一用户画像（年龄、性别）、第二用户画像（个人爱好）。使得 UPD 模块可以根据当前用户的第一用户画像（年龄、性别）和第二用户画像（个人爱好），确定与当前用户聊一些符合用户个性化需求的话题。
20

例如，与老人聊养生话题，与年轻人聊世界杯话题，与儿童聊幼儿园话题，等。

25 针对同一智能音箱，可以根据声纹识别出当前用户，从而实现智能音箱为不同用户提供符合用户需求的个性化服务。

第二种：

本说明书实施例中，还包括：

30 当第一声纹特征与多个用户身份标识中任一用户身份标识对应的第二声纹特征之间的匹配度均不大于预设阈值时，在第一声纹特征数据库中增加一个新的用户身份标识。

仍以上述图 2 为例，UP 模块 200 中的获得标识单元 206，将声纹提取单元 201 从当前用户的语音指令中提取的第一声纹特征，与第一声纹特征数据库 203 中存储的任一用户身份标识对应的第二声纹特征进行匹配，若第一声纹特征与第一声纹特征数据库 203 中存储的任一用户身份标识对应的第二声纹特征之间的匹配度均不大于预设阈值，则获得标识单元 206 可以确定当前用户为新用户。此时，新增标识单元 207 生成一个新的用户身份标识，进而将该新的用户身份标识确定为当前用户的目标用户身份标识。

由于 UP 模块 200 中的 UP 数据单元 204 中，未存储有与该新的用户身份标识对应的第一用户画像和第二用户画像，因此，共享型智能设备无法为当前用户提供精准的个性化服务。但是，共享型智能设备可以根据历史服务数据，为当前用户提供粗略的个性化服务。

例如，针对智能音箱，若当前用户为新用户，且根据智能音箱中的历史播放数据可知，使用该智能音箱的用户经常点播周杰伦和毛不易的歌曲时，则智能音箱为新用户播放周杰伦或毛不易的歌曲，而不会播放该智能音箱从来没有播放过的摇滚歌曲，等。从而实现为当前新用户提供粗略的个性化服务。

仍以上述图 2 为例，针对于新增标识单元 207 生成的新的用户身份标识，作为临时用户身份标识添加到声纹聚类模型 202 中。

若之后预设时间段内未接收到与该临时用户身份标识对应的用户发送的语音指令，则 UP 模块 200 可以确定该与该临时用户身份标识对应的用户仅为临时出现的使用用户，可以将其对应的相关数据删除；

若之后预设时间段内接收到与该临时用户身份标识对应的用户发送的语音指令超过预设数量，则 UP 模块 200 可以确定共享型智能设备新增加了一个稳定的使用用户（例如，家庭中增加了保姆、配偶、子女等新成员），则将该临时用户身份标识升级为一个稳定的用户身份标识，并将该用户身份标识添加到第一声纹特征数据库 203 中，以及在 UP 数据单元 204 中确定与其对应的第一用户画像和/或第二用户画像，等。

本说明书实施例记载的技术方案，接收用户输入的语音指令，从语音指令中提取第一声纹特征，并根据第一声纹特征确定用户的身份信息，进而根据用户的身份信息，以及对语音指令的识别结果，向用户提供个性化服务，从而实现共享型智能设备为不同用户提供符合用户需求的个性化服务。

图 3 为本说明书实施例提供的一种电子设备的结构示意图。如图 3 所示，在硬件层

面，该电子设备包括处理器，可选地还包括内部总线、网络接口、存储器。其中，存储器可能包含内存，例如高速随机存取存储器(Random-Access Memory, RAM)，也可能还包括非易失性存储器(non-volatile memory)，例如至少1个磁盘存储器等。当然，该电子设备还可能包括其他业务所需要的硬件。

5 处理器、网络接口和存储器可以通过内部总线相互连接，该内部总线可以是ISA(Industry Standard Architecture, 工业标准体系结构)总线、PCI(Peripheral Component Interconnect, 外设部件互连标准)总线或EISA(Extended Industry Standard Architecture, 扩展工业标准结构)总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示，图3中仅用一个双向箭头表示，但并不表示仅有一根总线或
10 一种类型的总线。

存储器，存放程序。具体地，程序可以包括程序代码，所述程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器，并向处理器提供指令和数据。

处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行，在逻辑层面上形成语音交互装置。处理器，执行存储器所存放的程序，并具体执行图1所示的方法
15 实施例的步骤。

上述如图1所述的方法可以应用于处理器中，或者由处理器实现。处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器，包括中央处理器(Central Processing Unit, CPU)、网络处理器(Network Processor, NP)等；还可以是数字信号处理器(Digital Signal Processor, DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列(Field-Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本说明书实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本
25 说明书实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤。

30 该电子设备可执行图1所示方法实施例执行的方法，并实现上述图1所示方法实施

例的功能，本说明书实施例在此不再赘述。

本说明书实施例还提出了一种计算机可读存储介质，该计算机可读存储介质存储一个或多个程序，该一个或多个程序包括指令，该指令当被包括多个应用程序的电子设备执行时，能够使该电子设备执行图 1 所示实施例中的语音交互方法，并具体执行图 1 所示方法实施例的步骤。

图 4 为本说明书实施例提供的一种语音交互装置的结构示意图。图 4 所示的装置 400 可以用于执行上述图 1-图 2 所示实施例的方法，装置 400 包括：

接收模块 401，接收用户输入的语音指令；

10 确定模块 402，从语音指令中提取第一声纹特征，并根据第一声纹特征确定用户的身份信息；

服务模块 403，根据用户的身份信息，以及对语音指令的识别结果，向用户提供个性化服务。

可选地，确定模块 402，进一步包括：

15 第一确定单元，确定第一声纹特征数据库，第一声纹特征数据库中包括多个用户身份标识，以及与多个用户身份标识中任一用户身份标识对应的第二声纹特征；

匹配单元，将第一声纹特征与多个用户身份标识中任一用户身份标识对应的第二声纹特征进行匹配；

20 第二确定单元，根据匹配结果，确定用户对应的目标用户身份标识，目标用户身份标识用于表示用户的身份信息；

其中，第一声纹特征与第二声纹特征为相同或不同的声纹特征。

可选地，通过以下方式确定得到第一声纹特征数据库：

获取历史语音指令集合，历史语音指令集合中包括多个历史语音指令；

提取多个历史语音指令中任一历史语音指令对应的第三声纹特征；

25 根据多个历史语音指令中任一历史语音指令对应的第三声纹特征，对多个历史语音指令进行声纹聚类，得到多个用户身份标识，以及与多个用户身份标识中任一用户身份标识对应的历史语音指令；

根据多个用户身份标识中任一用户身份标识对应的历史语音指令，确定用户身份标识对应的第二声纹特征；

30 其中，第二声纹特征和第三声纹特征为相同或不同的声纹特征。

可选地，第二确定单元，具体用于：

当第一声纹特征与多个用户身份标识中的一个用户身份标识对应的第二声纹特征之间的匹配度大于预设阈值时，将该用户身份标识确定为目标用户身份标识。

可选地，第二确定单元，具体用于：

- 5 当第一声纹特征与多个用户身份标识中任一用户身份标识对应的第二声纹特征之间的匹配度均不大于预设阈值时，在第一声纹特征数据库中增加一个新的用户身份标识。

可选地，确定模块 402，还用于：

根据用户身份标识对应的第二声纹特征，确定用户身份标识对应的第一用户画像；

用户身份标识对应的第一用户画像，用于反映与用户身份标识对应的用户的年龄和/

- 10 或性别。

可选地，确定模块 402，具体用于：

确定第二声纹特征数据库，第二声纹特征数据库中包括不同用户画像对应的第四声纹特征；

- 15 将用户身份标识对应的第二声纹特征，与不同用户画像对应的第四声纹特征进行匹配；

根据匹配结果，确定用户身份标识对应的第一用户画像；

其中，第二声纹特征与第四声纹特征为相同或不同的声纹特征。

可选地，确定模块 402，还用于：

根据用户身份标识对应的历史语音指令，确定用户身份标识对应的第二用户画像；

- 20 用户身份标识对应的第二用户画像，用于反映与用户身份标识对应的用户的个性化需求。

可选地，确定模块 402，具体用于：

将用户身份标识对应的历史语音指令转换为文本指令；

根据预设规则对文本指令进行分析；

- 25 根据分析结果，确定用户身份标识对应的第二用户画像。

可选地，服务模块 403，具体用于：

确定目标用户身份标识对应的第一用户画像和第二用户画像；

根据目标用户身份标识对应的第一用户画像和第二用户画像，以及对语音指令的识别结果，向用户提供个性化服务。

- 30 可选地，装置 400 为智能音箱，向用户提供个性化服务包括下述至少一种：

音乐推荐和聊天。

可选地，声纹特征包括下述至少一种：

语言、音色、音调、语速、口音，和频谱。

- 根据语音交互装置，接收模块接收用户输入的语音指令；确定模块从语音指令中提
- 5 取第一声纹特征，并根据第一声纹特征确定用户的身份信息；服务模块根据用户的身份信息，以及对语音指令的识别结果，向用户提供个性化服务，从而实现共享型智能设备为不同用户提供符合用户需求的个性化服务。

- 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，
- 10 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）。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
- 15 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
- 20 言来撰写，此称之为硬件描述语言（Hardware Description Language, HDL），而 HDL 也并非仅有一种，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
- 25 使用的是 VHDL（Very-High-Speed Integrated Circuit Hardware Description Language）与 Verilog。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。

- 控制器可以按任何适当的方式实现，例如，控制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）的计算机
- 30 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,控制器的例子包括但不限于以下微控制器: ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20 以及 Silicone Labs C8051F320,存储器控制器还可以被实现为存储器的控制逻辑的一部分。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。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。

上述实施例阐明的系统、装置、模块或单元,具体可以由计算机芯片或实体实现,或者由具有某种功能的产品来实现。一种典型的实现设备为计算机。具体的,计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

为了描述的方便,描述以上装置时以功能分为各种单元分别描述。当然,在实施本申请时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

本领域内的技术人员应明白,本发明的实施例可提供为方法、系统或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或

多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中，计算设备包括一个或多个处理器（CPU）、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器（RAM）和/或非易失性内存等形式，如只读存储器（ROM）或闪存（flash RAM）。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存（PRAM）、静态随机存取存储器（SRAM）、动态随机存取存储器（DRAM）、其他类型的随机存取存储器（RAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、快闪记忆体或其他内存技术、只读光盘只读存储器（CD-ROM）、数字多功能光盘（DVD）或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体（transitory media），如调制的数据信号和载波。

还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

本申请可以在由计算机执行的计算机可执行指令的一般上下文中描述，例如程序模块。一般地，程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本申请，在这些分布式计算环境中，由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中，程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分

互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于系统实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

- 5 以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围之内。

权利要求书

1、一种语音交互方法，包括：

接收用户输入的语音指令；

从所述语音指令中提取第一声纹特征，并根据所述第一声纹特征确定所述用户的身份信息；

根据所述用户的身份信息，以及对所述语音指令的识别结果，向所述用户提供个性化服务。

2、如权利要求 1 所述的方法，根据所述第一声纹特征确定所述用户的身份信息，包括：

确定第一声纹特征数据库，所述第一声纹特征数据库中包括多个用户身份标识，以及与所述多个用户身份标识中任一用户身份标识对应的第二声纹特征；

将所述第一声纹特征与所述多个用户身份标识中任一用户身份标识对应的所述第二声纹特征进行匹配；

根据匹配结果，确定所述用户对应的目标用户身份标识，所述目标用户身份标识用于表示所述用户的身份信息；

其中，所述第一声纹特征与所述第二声纹特征为相同或不同的声纹特征。

3、如权利要求 2 所述的方法，通过以下方式确定得到所述第一声纹特征数据库：

获取历史语音指令集合，所述历史语音指令集合中包括多个历史语音指令；

提取所述多个历史语音指令中任一历史语音指令对应的第三声纹特征；

根据所述多个历史语音指令中任一历史语音指令对应的所述第三声纹特征，对所述多个历史语音指令进行声纹聚类，得到所述多个用户身份标识，以及与所述多个用户身份标识中任一用户身份标识对应的历史语音指令；

根据所述多个用户身份标识中任一用户身份标识对应的历史语音指令，确定所述用户身份标识对应的所述第二声纹特征；

其中，所述第二声纹特征和所述第三声纹特征为相同或不同的声纹特征。

4、如权利要求 3 所述的方法，根据匹配结果，确定所述用户对应的目标用户身份标识，包括：

当所述第一声纹特征与所述多个用户身份标识中的一个用户身份标识对应的所述第二声纹特征之间的匹配度大于预设阈值时，将所述用户身份标识确定为所述目标用户身份标识。

5、如权利要求 4 所述的方法，还包括：

当所述第一声纹特征与所述多个用户身份标识中任一用户身份标识对应的所述第二声纹特征之间的匹配度均不大于所述预设阈值时，在所述第一声纹特征数据库中增加一个新的用户身份标识。

5 6、如权利要求 3 所述的方法，还包括：

根据所述用户身份标识对应的所述第二声纹特征，确定所述用户身份标识对应的第一用户画像；

所述用户身份标识对应的第一用户画像，用于反映与所述用户身份标识对应的用户的年龄和/或性别。

10 7、如权利要求 6 所述的方法，根据所述用户身份标识对应的所述第二声纹特征，确定所述用户身份标识对应的第一用户画像，包括：

确定第二声纹特征数据库，所述第二声纹特征数据库中包括不同用户画像对应的第四声纹特征；

15 将所述用户身份标识对应的所述第二声纹特征，与所述不同用户画像对应的所述第四声纹特征进行匹配；

根据匹配结果，确定所述用户身份标识对应的第一用户画像；

其中，所述第二声纹特征与所述第四声纹特征为相同或不同的声纹特征。

8、如权利要求 7 所述的方法，还包括：

20 根据所述用户身份标识对应的历史语音指令，确定所述用户身份标识对应的第二用户画像；

所述用户身份标识对应的第二用户画像，用于反映与所述用户身份标识对应的用户的个性化需求。

9、如权利要求 8 所述的方法，根据所述用户身份标识对应的历史语音指令，确定所述用户身份标识对应的第二用户画像，包括：

25 将所述用户身份标识对应的历史语音指令转换为文本指令；

根据预设规则对所述文本指令进行分析；

根据分析结果，确定所述用户身份标识对应的第二用户画像。

10、如权利要求 9 所述的方法，根据所述用户的身份信息，以及对所述语音指令的识别结果，向所述用户提供个性化服务，包括：

30 确定所述目标用户身份标识对应的第一用户画像和第二用户画像；

根据所述目标用户身份标识对应的第一用户画像和第二用户画像，以及对所述语音指令的识别结果，向所述用户提供个性化服务。

11、如权利要求 10 所述的方法，其特征在于，所述方法应用于智能音箱，向所述用户提供个性化服务包括下述至少一种：

5 音乐推荐和聊天。

12、如权利要求 2、3、7 任一项所述的方法，其特征在于，所述声纹特征包括下述至少一种：

语言、音色、音调、语速、口音，和频谱。

13、一种语音交互装置，用于执行如权利要求 1-12 任一项所述的语音交互方法，
10 所述装置包括：

接收模块，接收用户输入的语音指令；

确定模块，从所述语音指令中提取第一声纹特征，并根据所述第一声纹特征确定所述用户的身份信息；

15 服务模块，根据所述用户的身份信息，以及对所述语音指令的识别结果，向所述用户提供个性化服务。

14、一种电子设备，包括：

存储器，存放程序；

处理器，执行所述存储器存储的程序，并具体执行如权利要求 1-12 任一项所述的语音交互方法。

20 15、一种计算机可读存储介质，所述计算机可读存储介质存储一个或多个程序，所述一个或多个程序当被包括多个应用程序的电子设备执行时，使得所述电子设备执行如权利要求 1-12 任一项所述的语音交互方法。

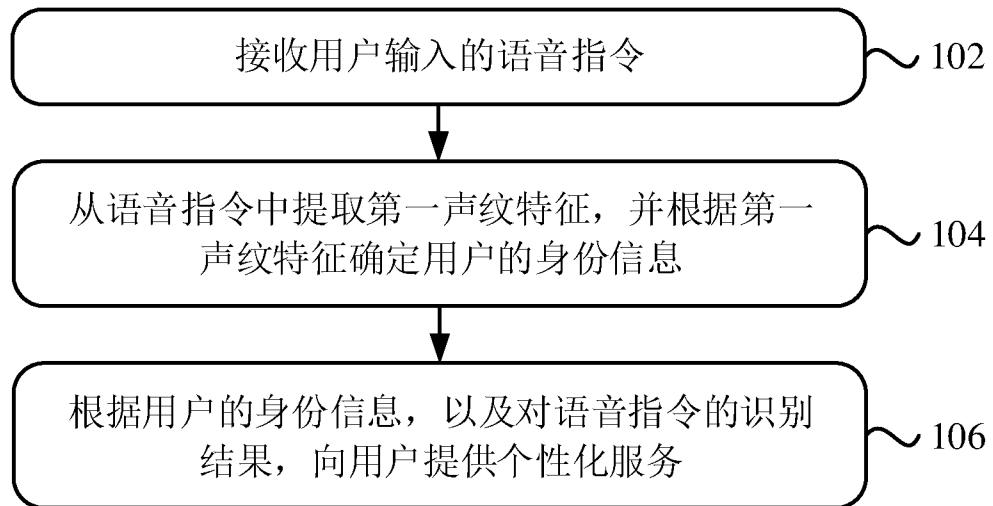


图 1

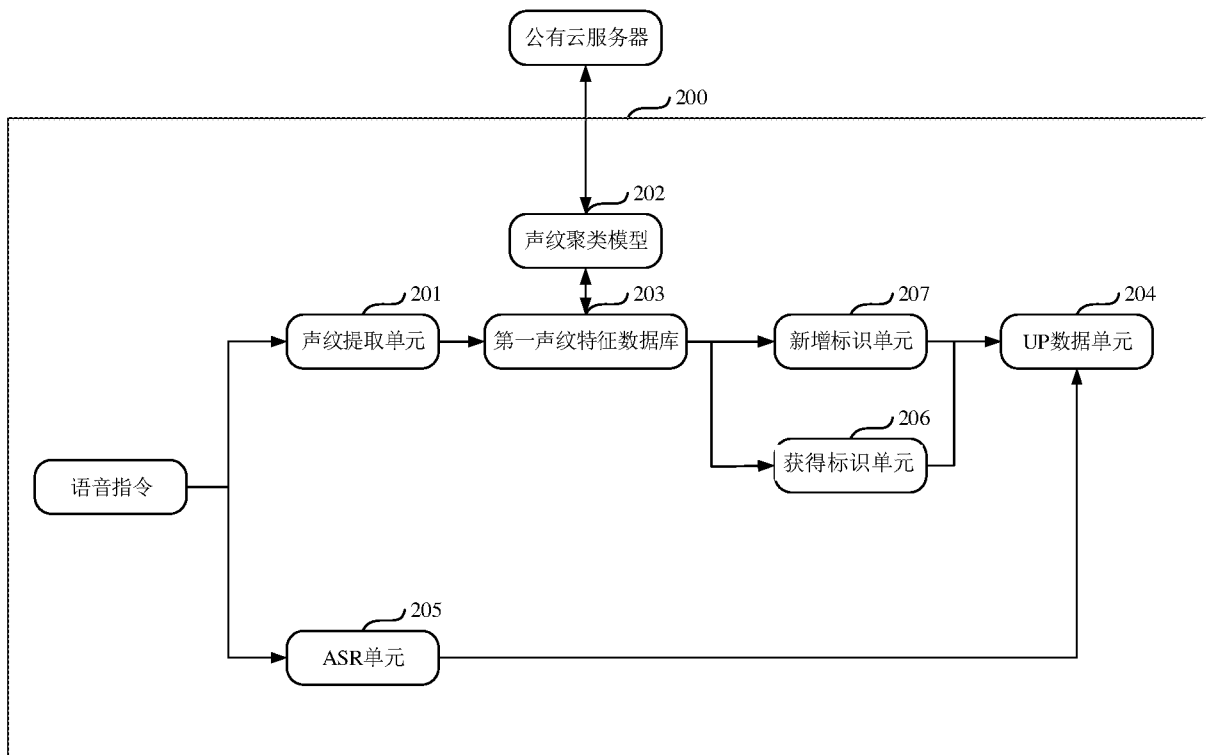


图 2

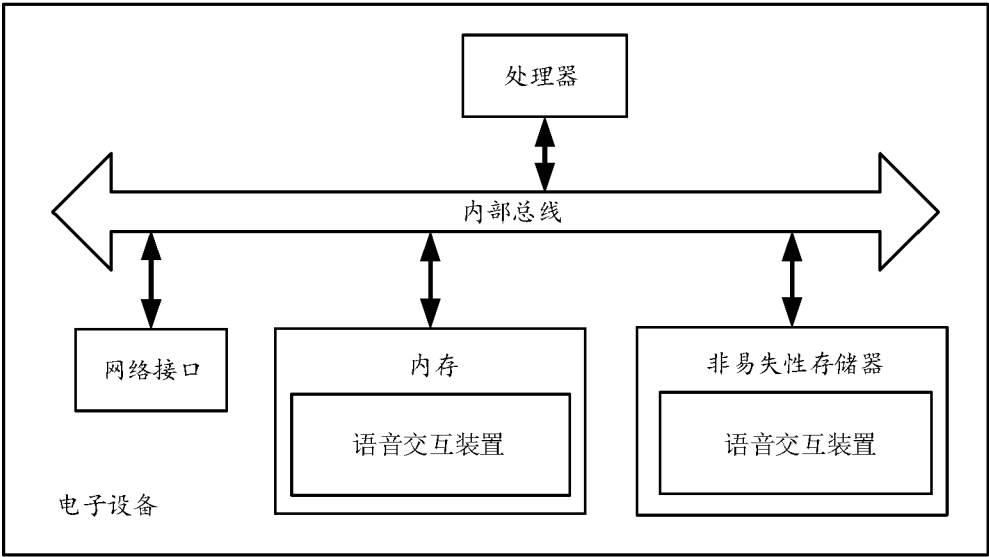


图 3

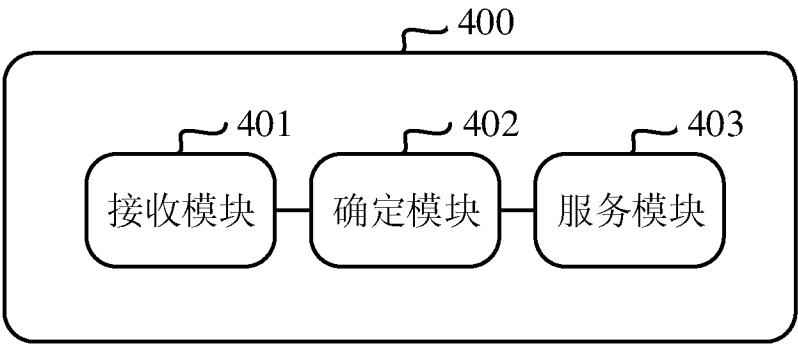


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/122640

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/26(2006.01)i; G10L 15/22(2006.01)i; G10L 17/04(2013.01)i; G10L 17/22(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L15/-; G10L17/-

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)

WPI, EPODOC, CNPAT, CNKI: 阿里巴巴, 孙尧, 语音, 声纹, 说话人, 说话者, 特征, 指令, 命令; speech, voice, recogni+, identif+, voiceprint, speaker, feature, character, order, command

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 107481720 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 15 December 2017 (2017-12-15) claims 1-16, description, paragraphs [0089]-[0249], and figures 1-9	1-15
X	CN 105810200 A (SHENZHEN YONGYIDA ROBOT CO., LTD.) 27 July 2016 (2016-07-27) description, paragraphs [0029]-[0048], and figures 1-2	1-15
A	CN 105609103 A (HEFEI VRVIEW INFORMATION TECHNOLOGY CO., LTD.) 25 May 2016 (2016-05-25) entire document	1-15
A	CN 105979376 A (LESHI ZHIXIN ELECTRONIC SCIENCE AND TECHNOLOGY (TIANJIN) CO., LTD.) 28 September 2016 (2016-09-28) entire document	1-15
A	CN 103730120 A (SHENZHEN ARATEK BIOMETRICS TECHNOLOGY CO., LTD.) 16 April 2014 (2014-04-16) entire document	1-15
A	US 2010158207 A1 (DHAWAN, V. et al.) 24 June 2010 (2010-06-24) entire document	1-15



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

22 February 2020

Date of mailing of the international search report

03 March 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
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/CN2019/122640

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2017244700 A1 (YAP, K.R.) 24 August 2017 (2017-08-24) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/122640

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	105810200	A	27 July 2016	None			
CN	105609103	A	25 May 2016	None			
CN	105979376	A	28 September 2016	WO	2017092342	A1	08 June 2017
				US	2017164049	A1	08 June 2017
CN	103730120	A	16 April 2014	None			
US	2010158207	A1	24 June 2010	US	9313307	B2	12 April 2016
US	2017244700	A1	24 August 2017	US	10044710	B2	07 August 2018

国际检索报告

国际申请号

PCT/CN2019/122640

A. 主题的分类

G10L 15/26(2006.01)i; G10L 15/22(2006.01)i; G10L 17/04(2013.01)i; G10L 17/22(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L15/-; G10L17/-

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))

WPI, EPDOC, CNPAT, CNKI: 阿里巴巴, 孙尧, 语音, 声纹, 说话人, 说话者, 特征, 指令, 命令; speech, voice, recogni+, identif+, voiceprint, speaker, feature, character, order, command

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 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) 权利要求1-16、说明书第[0089]-[0249]段、附图1-9	1-15
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[0029]-[0048]段、图1-2	1-15
A	CN 105609103 A (合肥寰景信息技术有限公司) 2016年 5月 25日 (2016 - 05 - 25) 全文	1-15
A	CN 105979376 A (乐视致新电子科技天津有限公司) 2016年 9月 28日 (2016 - 09 - 28) 全文	1-15
A	CN 103730120 A (深圳市亚略特生物识别科技有限公司) 2014年 4月 16日 (2014 - 04 - 16) 全文	1-15
A	US 2010158207 A1 (DHAWAN, Vishal 等) 2010年 6月 24日 (2010 - 06 - 24) 全文	1-15
A	US 2017244700 A1 (YAP, Kurt Ransom) 2017年 8月 24日 (2017 - 08 - 24) 全文	1-15

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2020年 2月 22日

国际检索报告邮寄日期

2020年 3月 3日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

陈红红

电话号码 86-(10)-53962642

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/122640

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	105810200	A	2016年 7月 27日	无			
CN	105609103	A	2016年 5月 25日	无			
CN	105979376	A	2016年 9月 28日	WO	2017092342	A1	2017年 6月 8日
				US	2017164049	A1	2017年 6月 8日
CN	103730120	A	2014年 4月 16日	无			
US	2010158207	A1	2010年 6月 24日	US	9313307	B2	2016年 4月 12日
US	2017244700	A1	2017年 8月 24日	US	10044710	B2	2018年 8月 7日