(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119007721 A
(43)申请公布日 2024.11.22

(21)申请号 202411021194.3

(22)申请日 2024.07.29

(71)申请人 岚图汽车科技有限公司
地址 430000 湖北省武汉市武汉经济技术开发区人工智能科技园N栋研发楼3层N3010号

(72)发明人 肖永君 周冰 邵鹏 周鹏飞

(74)专利代理机构 深圳市世纪恒程知识产权代理事务所 44287
专利代理师 翁美铃

(51)Int.Cl.
G10L 15/22(2006.01)
G06F 3/16(2006.01)
G06F 16/33(2019.01)
G06F 8/61(2018.01)
G10L 15/28(2013.01)
G10L 15/18(2013.01)

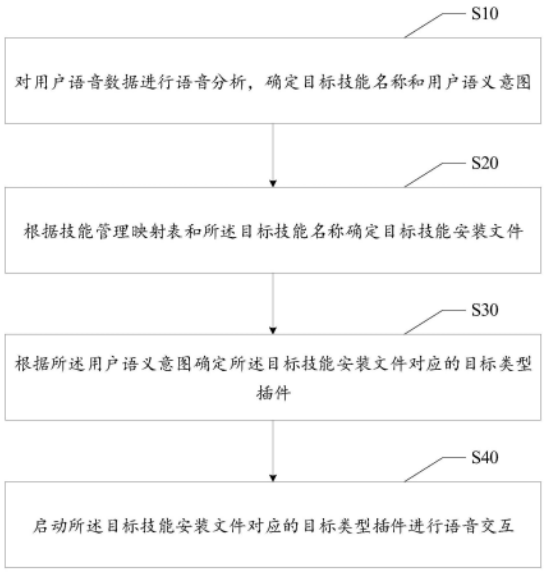
权利要求书2页 说明书12页 附图4页

(54)发明名称

语音交互方法、装置、设备及存储介质

(57)摘要

本申请公开了一种语音交互方法、装置、设备及存储介质，涉及车载语音技术领域，所述语音交互方法包括：对用户语音数据进行语音分析，确定目标技能名称和用户语义意图；根据技能管理映射表和所述目标技能名称确定目标技能安装文件；根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件；启动所述目标技能安装文件对应的目标类型插件进行语音交互。通过分析用户语音数据确定对应的技能安装文件，再根据用户语义意图启动对应的插件进行语音交互，解决了在不进行语义归一化的前提下实现平台化的问题，实现了语音技能实时更新。



CN 119007721 A

1.一种语音交互方法,其特征在于,所述语音交互方法包括:

对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;

根据技能管理映射表和所述目标技能名称确定目标技能安装文件;

根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;

启动所述目标技能安装文件对应的目标类型插件进行语音交互。

2.如权利要求1所述的方法,其特征在于,所述对用户语音数据进行语音分析,确定目标技能名称和用户语义意图之前,还包括:

分别获取历史技能安装文件的版本信息和云端技能安装文件的版本信息;

根据所述历史技能安装文件的版本信息和所述云端技能安装文件的版本信息确定技能版本状态;

根据所述技能版本状态启动语音服务,并采集用户语音数据。

3.如权利要求2所述的方法,其特征在于,所述根据所述技能版本状态启动语音服务,包括:

在所述技能版本状态为版本更新状态时,根据所述云端技能安装文件更新所述历史技能安装文件,得到更新技能安装文件;

解析所述更新技能安装文件得到对应的插件配置信息;

根据所述插件配置信息建立技能管理映射表,并启动语音服务。

4.如权利要求1所述的方法,其特征在于,所述对用户语音数据进行语音分析,确定目标技能名称和用户语义意图,包括:

对用户语音数据进行语音分析,得到对应的用户语音文本;

对所述用户语音文本进行关键字识别,得到语音识别结果;

根据所述语音识别结果确定目标技能名称和用户语义意图。

5.如权利要求1所述的方法,其特征在于,所述根据技能管理映射表和所述目标技能名称确定目标技能安装文件,包括:

根据技能管理映射表中技能插件信息映射表和所述目标技能名称确定目标技能插件信息;

根据所述目标技能插件信息和所述技能管理映射表中插件路径映射表确定目标插件存储路径;

根据所述目标技能插件信息和所述目标插件存储路径确定目标技能安装文件。

6.如权利要求1所述的方法,其特征在于,所述根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件,包括:

在所述用户语义意图为服务意图时,确定所述目标技能安装文件对应的目标类型插件为技能服务插件;

在所述用户语义意图为展示意图时,确定所述目标技能安装文件对应的目标类型插件为技能展示插件。

7.如权利要求1所述的方法,其特征在于,所述启动所述目标技能安装文件对应的目标类型插件进行语音交互,包括:

在所述目标技能安装文件对应的目标类型插件为技能展示插件时,获取待展示数据;

启动所述技能展示插件并展示所述待展示数据,以进行语音交互。

8.一种语音交互装置,其特征在于,所述装置包括:

分析模块,用于对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;

处理模块,用于根据技能管理映射表和所述目标技能名称确定目标技能安装文件;

识别模块,用于根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;

启动模块,用于启动所述目标技能安装文件对应的目标类型插件进行语音交互。

9.一种语音交互设备,其特征在于,所述设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述计算机程序配置为实现如权利要求1至7中任一项所述的语音交互方法的步骤。

10.一种存储介质,其特征在于,所述存储介质为计算机可读存储介质,所述存储介质上存储有计算机程序,所述计算机程序被处理器执行时实现如权利要求1至7中任一项所述的语音交互方法的步骤。

语音交互方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及车载语音技术领域,尤其涉及语音交互方法、装置、设备及存储介质。

背景技术

[0002] 车载语音是当前智能座舱中的一个重要组成部分,每个车企都投入大量成本打造自己的车载语音助手。在项目中常遇到一个问题,在新车型或项目常因为定点的语音供应商不同(如讯飞,思必驰,百度,自研等),旧项目中语音技能对接逻辑无法复用,无法满足平台化的需求。因此,如何在不进行语义归一化的前提下实现平台化,成为了亟待解决的问题。

发明内容

[0003] 本申请的主要目的在于提供一种语音交互方法、装置、设备及存储介质,旨在解决如何在不进行语义归一化的前提下实现平台化的技术问题。

[0004] 为实现上述目的,本申请提出一种语音交互方法,所述语音交互方法包括:

[0005] 对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;

[0006] 根据技能管理映射表和所述目标技能名称确定目标技能安装文件;

[0007] 根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;

[0008] 启动所述目标技能安装文件对应的目标类型插件进行语音交互。

[0009] 在一实施例中,所述对用户语音数据进行语音分析,确定目标技能名称和用户语义意图之前,还包括:

[0010] 分别获取历史技能安装文件的版本信息和云端技能安装文件的版本信息;

[0011] 根据所述历史技能安装文件的版本信息和所述云端技能安装文件的版本信息确定技能版本状态;

[0012] 根据所述技能版本状态启动语音服务,并采集用户语音数据。

[0013] 在一实施例中,所述根据所述技能版本状态启动语音服务,包括:

[0014] 在所述技能版本状态为版本更新状态时,根据所述云端技能安装文件更新所述历史技能安装文件,得到更新技能安装文件;

[0015] 解析所述更新技能安装文件得到对应的插件配置信息;

[0016] 根据所述插件配置信息建立技能管理映射表,并启动语音服务。

[0017] 在一实施例中,所述对用户语音数据进行语音分析,确定目标技能名称和用户语义意图,包括:

[0018] 对用户语音数据进行语音分析,得到对应的用户语音文本;

[0019] 对所述用户语音文本进行关键字识别,得到语音识别结果;

[0020] 根据所述语音识别结果确定目标技能名称和用户语义意图。

[0021] 在一实施例中,所述根据技能管理映射表和所述目标技能名称确定目标技能安装文件,包括:

[0022] 根据技能管理映射表中技能插件信息映射表和所述目标技能名称确定目标技能插件信息；

[0023] 根据所述目标技能插件信息和所述技能管理映射表中插件路径映射表确定目标插件存储路径；

[0024] 根据所述目标技能插件信息和所述目标插件存储路径确定目标技能安装文件。

[0025] 在一实施例中，所述根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件，包括：

[0026] 在所述用户语义意图为服务意图时，确定所述目标技能安装文件对应的目标类型插件为技能服务插件；

[0027] 在所述用户语义意图为展示意图时，确定所述目标技能安装文件对应的目标类型插件为技能展示插件。

[0028] 在一实施例中，所述启动所述目标技能安装文件对应的目标类型插件进行语音交互，包括：

[0029] 在所述目标技能安装文件对应的目标类型插件为技能展示插件时，获取待展示数据；

[0030] 启动所述技能展示插件并展示所述待展示数据，以进行语音交互。

[0031] 此外，为实现上述目的，本申请还提出一种语音交互装置，所述语音交互装置包括：

[0032] 分析模块，用于对用户语音数据进行语音分析，确定目标技能名称和用户语义意图；

[0033] 处理模块，用于根据技能管理映射表和所述目标技能名称确定目标技能安装文件；

[0034] 识别模块，用于根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件；

[0035] 启动模块，用于启动所述目标技能安装文件对应的目标类型插件进行语音交互。

[0036] 此外，为实现上述目的，本申请还提出一种语音交互设备，所述设备包括：存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述计算机程序配置为实现如上文所述的语音交互方法的步骤。

[0037] 此外，为实现上述目的，本申请还提出一种存储介质，所述存储介质为计算机可读存储介质，所述存储介质上存储有计算机程序，所述计算机程序被处理器执行时实现如上文所述的语音交互方法的步骤。

[0038] 此外，为实现上述目的，本申请还提供一种计算机程序产品，所述计算机程序产品包括计算机程序，所述计算机程序被处理器执行时实现如上文所述的语音交互方法的步骤。

[0039] 本申请通过对用户语音数据进行语音分析，确定目标技能名称和用户语义意图；根据技能管理映射表和所述目标技能名称确定目标技能安装文件；根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件；启动所述目标技能安装文件对应的目标类型插件进行语音交互。通过分析用户语音数据确定对应的技能安装文件，再根据用户语义意图启动对应的插件进行语音交互，解决了在不进行语义归一化的前提下实现平台化的

问题,实现了语音技能实时更新。

附图说明

[0040] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本申请的实施例,并与说明书一起用于解释本申请的原理。

[0041] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例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,对于本领域普通技术人员而言,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。

[0042] 图1为本申请语音交互方法实施例一提供的流程示意图;

[0043] 图2为本申请语音交互方法实施例一提供的语音工程组件化示意图;

[0044] 图3为本申请语音交互方法实施例一提供的技能APK设计流程示意图;

[0045] 图4为本申请语音交互方法实施例二提供的流程示意图;

[0046] 图5为本申请实施例二提供的语音交互方法的简要流程示意图;

[0047] 图6为本申请实施例语音交互装置的模块结构示意图;

[0048] 图7为本申请实施例中语音交互方法涉及的硬件运行环境的设备结构示意图。

[0049] 本申请目的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0050] 应当理解,此处所描述的具体实施例仅仅用以解释本申请的技术方案,并不用于限定本申请。

[0051] 为了更好的理解本申请的技术方案,下面将结合说明书附图以及具体的实施方式进行详细的说明。

[0052] 本申请实施例的主要解决方案是:对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;根据技能管理映射表和所述目标技能名称确定目标技能安装文件;根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;启动所述目标技能安装文件对应的目标类型插件进行语音交互。

[0053] 车载语音是当前智能座舱中的一个重要组成部分,每个车企都投入大量成本打造自己的车载语音助手。在项目中常遇到一个问题,在新车型或项目常因为定点的语音供应商不同(如讯飞,思必驰,百度,自研等),旧项目中语音技能对接逻辑无法复用,无法满足平台化的需求。因此,如何在不进行语义归一化的前提下实现平台化,成为了亟待解决的问题。

[0054] 本申请通过对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;根据技能管理映射表和所述目标技能名称确定目标技能安装文件;根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;启动所述目标技能安装文件对应的目标类型插件进行语音交互。通过分析用户语音数据确定对应的技能安装文件,再根据用户语义意图启动对应的插件进行语音交互,解决了在不进行语义归一化的前提下实现平台化的问题,实现了语音技能实时更新。

[0055] 需要说明的是,本实施例的执行主体可以是一种具有数据处理、网络通信以及程序运行功能的计算服务设备,例如平板电脑、个人电脑、手机等,或者是一种能够实现上述

功能的语音交互设备等。以下以语音交互设备为执行主体为例,对本实施例及下述各实施例进行说明。

[0056] 基于此,本申请实施例提供了一种语音交互方法,参照图1,图1为本申请语音交互方法第一实施例的流程示意图。

[0057] 本实施例中,所述语音交互方法包括步骤S10～S40:

[0058] 步骤S10,对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;

[0059] 需要说明的是,因为不同语音供应商的语义(NLU)是不同的。为了实现平台化车企采用的通常办法是将供应商的语义进行归一化操作,即将不同家语音供应商的语义统一转换成一套车企内部制定的标准语义格式,然后再使用标准语义格式进行技能对接。这种做法也存在较多缺陷:a)工作量大,完整转换一套语义工作量大,当每引入一家新语义都要提前先进行语义归一化;b)语义本身具有可定制性和灵活性,供应商的语义平台一般都可以对技能进行增删修改查操作,就这意味着语义并不是固定的,语音供应商的每一次修改,都会影响到语义归一化模块正常运行;c)技能退化,因为归一化的语义标准格式具有通用、兼容等特点。这样在语义归一化过程中,会存在或多或少的数据丢失,从而使原本的一些优秀语义技能在转换后退化;d)不支持热更新,每一次技能对接的修改都要整个语音技能安装文件(APK,Android Package)进行升级。

[0060] 应当说明的是,如图2所示,将整个语音工程实现组件化,即每个语音技能对接的模块是一个组件,存在形式是一个独立的APK。那么A语音供应商具有20个技能,则对应20个技能组件APK。B语音供应商具有18个技能,则对应18个技能组件APK。语音开发人员尊重不同的供应商的语义上的差异,为A和B 2家语音供应商开发出38个技能组件APK。

[0061] 可以理解的是,用户语音数据指的是用户提出的语音问题或语音请求信息,目标技能名称指的是用户语音数据中包含的语音服务安装包的名称,用户语义意图指的是在进行自然语言交流时,所表达的目的或意图。

[0062] 在具体实施中,用户与车载语音进行交互时,通过用户提出的语音问题或语音请求信息进行语音分析,进而根据分析结果确定用户语音数据中包含的语音服务安装包的名称和用户所表达的目的或意图。

[0063] 在一种可行的实施方式中,步骤S10可以包括步骤A11～A13:

[0064] 步骤A11,对用户语音数据进行语音分析,得到对应的用户语音文本;

[0065] 可以理解的是,用户语音文本指的是用户语音数据对应文本数据。

[0066] 在具体实施中,为对用户的语音信息进行分析,将用户提出的语音问题或语音请求信息转换为文本数据,进而得到用户语音数据对应文本数据。

[0067] 步骤A12,对所述用户语音文本进行关键字识别,得到语音识别结果;

[0068] 可以理解的是,语音识别结果指的是语音文本中关键字的识别结果。

[0069] 在具体实施中,通过对用户语音数据对应文本数据中包含的关键字进行识别,得到语音文本中关键字的识别结果,例如,用户语音文本为“今天天气怎么样”,通过对关键字进行识别,得到关键字“天气”,进而确定关键字“天气”为语音识别结果。

[0070] 步骤A13,根据所述语音识别结果确定目标技能名称和用户语义意图。

[0071] 可以理解的是,例如,用户语音文本为“打开天气”,语音识别结果为“打开”和“天气”,进而确定目标技能名称为天气技能APK,用户语义意图为打开天气技能APK。

[0072] 步骤S20,根据技能管理映射表和所述目标技能名称确定目标技能安装文件;

[0073] 可以理解的是,技能管理映射表包括技能与插件APK路径的映射表和语义意图与插件信息的映射表,目标技能安装文件指的是需要安装的语音技能安装文件。

[0074] 在具体实施中,根据需求的语音技能APK名称通过语义意图与插件信息的映射表进行筛选,得到对应的技能APK,再通过技能与插件APK路径的映射表进行筛选,得到需求的语音技能APK的存储路径,进而得到需要安装的语音技能安装文件。

[0075] 在一种可行的实施方式中,步骤S20可以包括步骤A21～A23:

[0076] 步骤A21,根据技能管理映射表中技能插件信息映射表和所述目标技能名称确定目标技能插件信息;

[0077] 可以理解的是,插件信息映射表指的是语义意图与插件信息的映射表,目标技能插件信息指的是语音交互系统中可实现特定功能的组件或服务。

[0078] 在具体实施中,根据用户语音数据中包含的语音服务安装包的名称对语义意图与插件信息的映射表进行筛选,确定语音交互系统中可实现特定功能的组件或服务。

[0079] 步骤A22,根据所述目标技能插件信息和所述技能管理映射表中插件路径映射表确定目标插件存储路径;

[0080] 可以理解的是,技能管理映射表指的是技能与插件APK路径的映射表,目标插件存储路径指的是实现对应语音功能的组件存储地址。

[0081] 在具体实施中,根据语音交互系统中可实现特定功能的组件或服务通过特定技能(或功能)相关联的插件的APK(Android Application Package,Android操作系统用来分发和安装应用程序的文件格式)路径映射关系,确定实现对应语音功能的组件存储地址。

[0082] 步骤A23,根据所述目标技能插件信息和所述目标插件存储路径确定目标技能安装文件。

[0083] 可以理解的是,通过语音交互系统中可实现特定功能的组件或服务和实现对应语音功能的组件存储地址,进而得到需要安装的语音技能安装文件。

[0084] 步骤S30,根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;

[0085] 可以理解的是,目标类型插件包括服务插件(Service)和展示插件(View),对于一个语音技能它包含不同的意图,不同的意图对接方式也不一样,通常情况可分为2种:1种仅做播报不做技能卡片展示,如控制天气意图。另1种做播报以及带技能卡片展示,如查询天气意图。具体来说:对于不带技能卡片的意图对接通过Service插件完成,带技能卡片的意图对接通过view插件完成。在一个技能APK用于完成意图对接的Service和View本实施例称之为插件,语音服务作为宿主通过插件化技术加载技能APK中的插件到语音进程中技能卡片显示和播报。Service和View的插件表现形式不一样,其插件化技术实现方式也不一样。

[0086] 在具体实施中,例如,用户语义意图为“打开天气”,表明需要展示天气技能卡片,即做播报以及带技能卡片展示,确定目标类型插件为View插件。

[0087] 在一种可行的实施方式中,步骤S30可以包括步骤A31～A32:

[0088] 步骤A31,在所述用户语义意图为服务意图时,确定所述目标技能安装文件对应的目标类型插件为技能服务插件;

[0089] 可以理解的是,例如,用户语义意图为“今天天气怎么样”,即仅做播报不做技能卡

片展示,为服务意图,进而确定目标类型插件为技能服务插件(Service插件)。

[0090] 步骤A32,在所述用户语义意图为展示意图时,确定所述目标技能安装文件对应的目标类型插件为技能展示插件。

[0091] 可以理解的是,例如,用户语义意图为“打开天气”,即做播报以及带技能卡片展示,为展示(或控制)意图,进而确定目标类型插件为技能服务插件(View插件)。

[0092] 应当说明的是,Service插件采用主流的Hook技术实现,使用Hook Instrumentation实现,AMS是在SystemServer进程中,无法直接进行修改,只能在应用程序进程中修改。可以采用预先占坑的方式来解决没有在AndroidManifest.xml中显示声明的问题,具体做法是使用一个在AndroidManifest.xml中注册的Service来进行占坑,用来通过AMS的校验,Hook Instrumentation实现需要用到占坑Service如:(".S$1",".S$2"...),用占坑Service替换插件Service以及还原插件Service,达到最终启动的目的。

[0093] 需要说明的是,本实施例中View插件的做法是获取插件apk的类加载器classloader,然后利用插件的classloader去反射获取类的信息。对于高安卓版本获取类加载器可能会失败,会涉及到系统源码的一些定制修改。View的插件化会随安卓版本的安全特性的不同会有适配成本。也可以使用其他的插件框架实现。

[0094] 步骤S40,启动所述目标技能安装文件对应的目标类型插件进行语音交互。

[0095] 可以理解的是,启动需求的语音技能安装文件对应的插件,以实现与用户进行语音交互,例如,用户语义意图为“打开天气”,进而启动天气技能APK对应的View插件,以展示天气技能APK的技能卡片,最后完成与用户的语音交互。

[0096] 在一种可行的实施方式中,步骤S40可以包括步骤A41～A42:

[0097] 步骤A41,在所述目标技能安装文件对应的目标类型插件为技能展示插件时,获取待展示数据;

[0098] 可以理解的是,待展示数据指的是展示或处理具体的数据,例如,天气数据和路线数据等。

[0099] 在具体实施中,在目标技能安装文件为天气技能APK,且对应的插件为展示插件时,需获取需要展示天气相关的数据,以进行展示。

[0100] 步骤A42,启动所述技能展示插件并展示所述待展示数据,以进行语音交互。

[0101] 在具体实施中,例如,当用户说“今天天气怎么样”,云端下发讯飞天气查询语义结果,语音服务APK根据语义中的信息找到要加载的插件(view),然后通过插件化技术创建全类名为com.voyah.skill.WeatherView的实例对象,添加到语音服务的技能卡片容器中,接着将语义通过view的tag属性传递给WeatherView实例对象,最后WeatherView实例对象在监听到view被显示时,获取从tag属性传递过来的数据解析并刷新view。

[0102] 应当说明的是,上面过程将语音工程通过技能分类形式实现组件化,整个语音工程包括一个宿主APK和多个插件APK。在安卓系统中每个APK都是一个独立运行的实体,应用之间的隔离的,所以需要使用插件化技术将技能APK运行在语音服务APK所在的进程中,具体使用哪一种插件化技术这里不做详细展开。首先,对于一个语音技能它包含不同的意图,不同的意图对接方式也不一样,通常情况可分为2种:1种仅做播报不做技能卡片展示,如控制天气意图。另1种做播报以及带技能卡片展示,如查询天气意图。具体来说:对于不带技能卡片的意图对接通过Service插件完成,带技能卡片的意图对接通过view插件完成。在一个

技能APK用于完成意图对接的Service和View本发明称之为插件,语音服务作为宿主通过插件化技术加载技能APK中的插件到语音进程中技能卡片显示和播报。Service和View的插件表现形式不一样,其插件化技术实现方式也不一样,后面再叙述。

[0103] 需要说明的是,如图3所示,a)对于一个技能APK,在APK的配置文件中申明进行技能对接的插件信息:如一个已存在的讯飞天气技能APK中申明了2个插件:WeatherView用来对接天气查询意图,WeatherService用来对接天气控制意图;b)在WeatherView代码逻辑中对接讯飞的天气查询意图,并完成对应的UI开发以及播报具体天气情况,在WeatherService代码逻辑中对接讯飞的天气操作意图,并完成打开和关闭天气应用动作以及播报执行结果;c)技能APK要做到足够小以保证下载和加载耗时非常快,所以技能APK不建议使用第三方库。同时因为技能APK运行在语音服务进程中,所以一些公共的SDK可以通过语音服务依赖,技能APK直接使用即可。

[0104] 本实施例通过对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;根据技能管理映射表和所述目标技能名称确定目标技能安装文件;根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;启动所述目标技能安装文件对应的目标类型插件进行语音交互。通过分析用户语音数据确定对应的技能安装文件,再根据用户语义意图启动对应的插件进行语音交互,解决了在不进行语义归一化的前提下实现平台化的问题,实现了语音技能实时更新。

[0105] 基于本申请第一实施例,在本申请第二种实施例中,与上述实施例一相同或相似的内容,可以参考上文介绍,后续不再赘述。在此基础上,请参照图4,步骤S10之前,所述语音交互方法还包括步骤S01～S03:

[0106] 步骤S01,分别获取历史技能安装文件的版本信息和云端技能安装文件的版本信息;

[0107] 可以理解的是,历史技能安装文件指的是本地的技能安装文件,云端技能安装文件指的是存储在云端的技能安装文件。

[0108] 在具体实施中,当语音技能APK开发完成后,将其发布到云端技能商店里。当语音服务启动后,分别获取本地的技能安装文件的版本信息和存储在云端的技能安装文件的版本信息。

[0109] 步骤S02,根据所述历史技能安装文件的版本信息和所述云端技能安装文件的版本信息确定技能版本状态;

[0110] 可以理解的是,技能版本状态包括版本需要更新状态和版本不需要更新状态。

[0111] 在具体实施中,将本地技能安装文件的版本信息和云端存储的技能安装文件的版本信息进行比较,在本地技能安装文件的版本信息和云端存储的技能安装文件的版本信息相同时,确定技能版本状态为不需要更新状态;在本地技能安装文件的版本信息和云端存储的技能安装文件的版本信息不相同时,确定技能版本状态为需要更新状态。

[0112] 步骤S03,根据所述技能版本状态启动语音服务,并采集用户语音数据。

[0113] 在具体实施中,在需要不更新技能版本时,直接启动语音服务;在需要更新技能版本时,先进行本地技能安装文件的版本更新,再启动语音服务,最后,在启动语音服务后,采集用户的语音数据。

[0114] 在一种可行的实施方式中,步骤S03可以包括步骤A031～A033:

[0115] 步骤A031,在所述技能版本状态为版本更新状态时,根据所述云端技能安装文件更新所述历史技能安装文件,得到更新技能安装文件;

[0116] 可以理解的是,更新技能安装文件指的是版本更新后的技能安装文件。

[0117] 在具体实施中,在需要更新本地技能文件的版本时,通过存储在云端对应的新技能安装文件对本地技能安装文件进行更新,得到版本更新后的技能安装文件。

[0118] 步骤A032,解析所述更新技能安装文件得到对应的插件配置信息;

[0119] 可以理解的是,插件配置信息包括插件类型、插件类名、处理意图以及供应商等。

[0120] 在具体实施中,将版本更新后的技能安装文件进行解析,得到插件类型、插件类名、处理意图以及供应商等插件配置信息。

[0121] 步骤A033,根据所述插件配置信息建立技能管理映射表,并启动语音服务。

[0122] 在具体实施中,根据插件类型、插件类名、处理意图以及供应商等插件配置信息分别建立技能与插件APK路径的映射、语义意图与插件信息的映射关系。

[0123] 应当说明的是,当讯飞天气技能APK开发完成后,将其发布到云端技能商店里。当语音服务启动后,语音服务主动检查本地的讯飞天气技能APK版本和云端的讯飞天气技能APK版本,如果云端的版本大于本地,静默地从云端将新APK下载并替换本地的版本。某些技能APK可作为收费项目运营,用户在技能商店中购买后提供下载和使用服务。

[0124] 需要说明的是,语音服务APK在启动过程中读取该技能APK中的配置信息,建立技能与插件APK路径的映射、语义意图与插件信息的映射关系。当用户说“今天武汉天气怎么样”,云端下发讯飞天气查询语义结果,语音服务APK根据语义中的信息找到要加载的插件(view),然后通过插件化技术创建全类名为com.voyah.skill.WeatherView的实例对象,添加到语音服务的技能卡片容器中,接着将语义通过view的tag属性传递给WeatherView实例对象,最后WeatherView实例对象在监听到view被显示时,获取从tag属性传递过来的数据解析并刷新view;当用户说“打开天气”,云端下发讯飞天气操作语义结果,语音服务APK根据语义中的信息找到要加载的插件(service),然后通过插件化技术启动全类名为com.voyah.skill.WeatherServcie的服务,接着将语义通过intent传递给WeatherService服务,最后WeatherService服务被启动后通过intent获取传递过来的语义,解析并调用接口执行响应。

[0125] 本实施例通过在启动语音服务时,分别获取历史技能安装文件的版本信息和云端技能安装文件的版本信息;根据所述历史技能安装文件的版本信息和所述云端技能安装文件的版本信息确定技能版本状态;根据所述技能版本状态启动语音服务,并采集用户语音数据。通过主动检查云端技能安装文件的版本信息和本地技能安装文件的版本信息,如果云端的版本大于本地,静默地从云端将新技能安装文件下载并替换本地的版本,实现了语音技能的实时更新。

[0126] 示例性地,为了助于理解本实施例结合上述实施例一后所得到的语音交互方法的实现流程,请参照图5,图5提供了一种语音交互方法的简要流程示意图,具体地:在技能APK开发完成后,将其发布到云端技能商店里。当语音服务启动后,语音服务主动检查本地的技能APK版本和云端的技能APK版本,如果云端的版本大于本地,静默地从云端将新APK下载并替换本地的版本,进而完成启动流程,再采集用户交互语音进行分析,根据语义中的供应商和技能名称从技能管理映射表找到对应的技能APK,再根据语义找到技能APK对应的插件,

即view插件或service插件,以传递语义数据。

[0127] 需要说明的是,上述示例仅用于理解本申请,并不构成对本申请语音交互方法的限定,基于此技术构思进行更多形式的简单变换,均在本申请的保护范围内。

[0128] 本申请还提供一种语音交互装置,请参照图6,所述语音交互装置包括:

[0129] 分析模块10,用于对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;

[0130] 处理模块20,用于根据技能管理映射表和所述目标技能名称确定目标技能安装文件;

[0131] 识别模块30,用于根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;

[0132] 启动模块40,用于启动所述目标技能安装文件对应的目标类型插件进行语音交互。

[0133] 可选地,所述分析模块10,还用于:

[0134] 分别获取历史技能安装文件的版本信息和云端技能安装文件的版本信息;

[0135] 根据所述历史技能安装文件的版本信息和所述云端技能安装文件的版本信息确定技能版本状态;

[0136] 根据所述技能版本状态启动语音服务,并采集用户语音数据。

[0137] 可选地,所述分析模块10,还用于:

[0138] 在所述技能版本状态为版本更新状态时,根据所述云端技能安装文件更新所述历史技能安装文件,得到更新技能安装文件;

[0139] 解析所述更新技能安装文件得到对应的插件配置信息;

[0140] 根据所述插件配置信息建立技能管理映射表,并启动语音服务。

[0141] 可选地,所述分析模块10,还用于:

[0142] 对用户语音数据进行语音分析,得到对应的用户语音文本;

[0143] 对所述用户语音文本进行关键字识别,得到语音识别结果;

[0144] 根据所述语音识别结果确定目标技能名称和用户语义意图。

[0145] 可选地,所述处理模块20,还用于:

[0146] 根据技能管理映射表中技能插件信息映射表和所述目标技能名称确定目标技能插件信息;

[0147] 根据所述目标技能插件信息和所述技能管理映射表中插件路径映射表确定目标插件存储路径;

[0148] 根据所述目标技能插件信息和所述目标插件存储路径确定目标技能安装文件。

[0149] 可选地,所述识别模块30,还用于:

[0150] 在所述用户语义意图为服务意图时,确定所述目标技能安装文件对应的目标类型插件为技能服务插件;

[0151] 在所述用户语义意图为展示意图时,确定所述目标技能安装文件对应的目标类型插件为技能展示插件。

[0152] 可选地,所述启动模块40,还用于:

[0153] 在所述目标技能安装文件对应的目标类型插件为技能展示插件时,获取待展示数

据;

[0154] 启动所述技能展示插件并展示所述待展示数据,以进行语音交互。

[0155] 本申请提供的语音交互装置,采用上述实施例中的语音交互方法,能够解决如何在不进行语义归一化的前提下实现平台化的技术问题。与现有技术相比,本申请提供的语音交互装置的有益效果与上述实施例提供的语音交互方法的有益效果相同,且所述语音交互装置中的其他技术特征与上述实施例方法公开的特征相同,在此不做赘述。

[0156] 本申请提供一种语音交互设备,语音交互设备包括:至少一个处理器;以及,与至少一个处理器通信连接的存储器;其中,存储器存储有可被至少一个处理器执行的指令,指令被至少一个处理器执行,以使至少一个处理器能够执行上述实施例一中的语音交互方法。

[0157] 下面参考图7,其示出了适于用来实现本申请实施例的语音交互设备的结构示意图。本申请实施例中的语音交互设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA(Personal Digital Assistant:个人数字助理)、PAD(Portable Application Description:平板电脑)、PMP(Portable Media Player:便携式多媒体播放器)、车载终端(例如车载导航终端)等等的移动终端以及诸如数字TV、台式计算机等等的固定终端。图7示出的语音交互设备仅仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0158] 如图7所示,语音交互设备可以包括处理装置1001(例如中央处理器、图形处理器等),其可以根据存储在只读存储器(ROM:Read Only Memory)1002中的程序或者从存储装置1003加载到随机访问存储器(RAM:Random Access Memory)1004中的程序而执行各种适当的动作和处理。在RAM1004中,还存储有语音交互设备操作所需的各种程序和数据。处理装置1001、ROM1002以及RAM1004通过总线1005彼此相连。输入/输出(I/O)接口1006也连接至总线。通常,以下系统可以连接至I/O接口1006:包括例如触摸屏、触摸板、键盘、鼠标、图像传感器、麦克风、加速度计、陀螺仪等的输入装置1007;包括例如液晶显示器(LCD:Liquid Crystal Display)、扬声器、振动器等的输出装置1008;包括例如磁带、硬盘等的存储装置1003;以及通信装置1009。通信装置1009可以允许语音交互设备与其他设备进行无线或有线通信以交换数据。虽然图中示出了具有各种系统的语音交互设备,但是应理解的是,并不要求实施或具备所有示出的系统。可以替代地实施或具备更多或更少的系统。

[0159] 特别地,根据本申请公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本申请公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信装置从网络上被下载和安装,或者从存储装置1003被安装,或者从ROM1002被安装。在该计算机程序被处理装置1001执行时,执行本申请公开实施例的方法中限定的上述功能。

[0160] 本申请提供的语音交互设备,采用上述实施例中的语音交互方法,能解决如何在不进行语义归一化的前提下实现平台化的技术问题。与现有技术相比,本申请提供的语音交互设备的有益效果与上述实施例提供的语音交互方法的有益效果相同,且该语音交互设备中的其他技术特征与上一实施例方法公开的特征相同,在此不做赘述。

[0161] 应当理解,本申请公开的各部分可以用硬件、软件、固件或它们的组合来实现。在

上述实施方式的描述中,具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0162] 以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以所述权利要求的保护范围为准。

[0163] 本申请提供一种计算机可读存储介质,具有存储在其上的计算机可读程序指令(即计算机程序),计算机可读程序指令用于执行上述实施例中的语音交互方法。

[0164] 本申请提供的计算机可读存储介质例如可以是U盘,但不限于电、磁、光、电磁、红外线、或半导体的系统、系统或器件,或者任意以上的组合。计算机可读存储介质的更具体地例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM:Random Access Memory)、只读存储器(ROM:Read Only Memory)、可擦式可编程只读存储器(EPROM:Erasable Programmable Read Only Memory或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM:CD-Read Only Memory)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本实施例中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、系统或者器件使用或者与其结合使用。计算机可读存储介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:电线、光缆、RF(Radio Frequency:射频)等等,或者上述的任意合适的组合。

[0165] 上述计算机可读存储介质可以是语音交互设备中所包含的;也可以是单独存在,而未装配入语音交互设备中。

[0166] 上述计算机可读存储介质承载有一个或者多个程序,当上述一个或者多个程序被语音交互设备执行时,使得语音交互设备:对用户语音数据进行语音分析,确定目标技能名称和用户语义意图;根据技能管理映射表和所述目标技能名称确定目标技能安装文件;根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件;启动所述目标技能安装文件对应的目标类型插件进行语音交互。

[0167] 可以以一种或多种程序设计语言或其组合来编写用于执行本申请的操作的计算机程序代码,上述程序设计语言包括面向对象的程序设计语言—诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN:Local Area Network)或广域网(WAN:Wide Area Network)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0168] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的是,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执

行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0169] 描述于本申请实施例中所涉及到的模块可以通过软件的方式实现,也可以通过硬件的方式来实现。其中,模块的名称在某种情况下并不构成对该单元本身的限定。

[0170] 本申请提供的可读存储介质为计算机可读存储介质,所述计算机可读存储介质存储有用于执行上述语音交互方法的计算机可读程序指令(即计算机程序),能够解决如何在不进行语义归一化的前提下实现平台化的技术问题。与现有技术相比,本申请提供的计算机可读存储介质的有益效果与上述实施例提供的语音交互方法的有益效果相同,在此不做赘述。

[0171] 本申请还提供一种计算机程序产品,包括计算机程序,所述计算机程序被处理器执行时实现如上述的语音交互方法的步骤。

[0172] 本申请提供的计算机程序产品能够解决如何在不进行语义归一化的前提下实现平台化的技术问题。与现有技术相比,本申请提供的计算机程序产品的有益效果与上述实施例提供的语音交互方法的有益效果相同,在此不做赘述。

[0173] 以上所述仅为本申请的部分实施例,并非因此限制本申请的专利范围,凡是在本申请的技术构思下,利用本申请说明书及附图内容所作的等效结构变换,或直接/间接运用在其他相关的技术领域均包括在本申请的专利保护范围内。

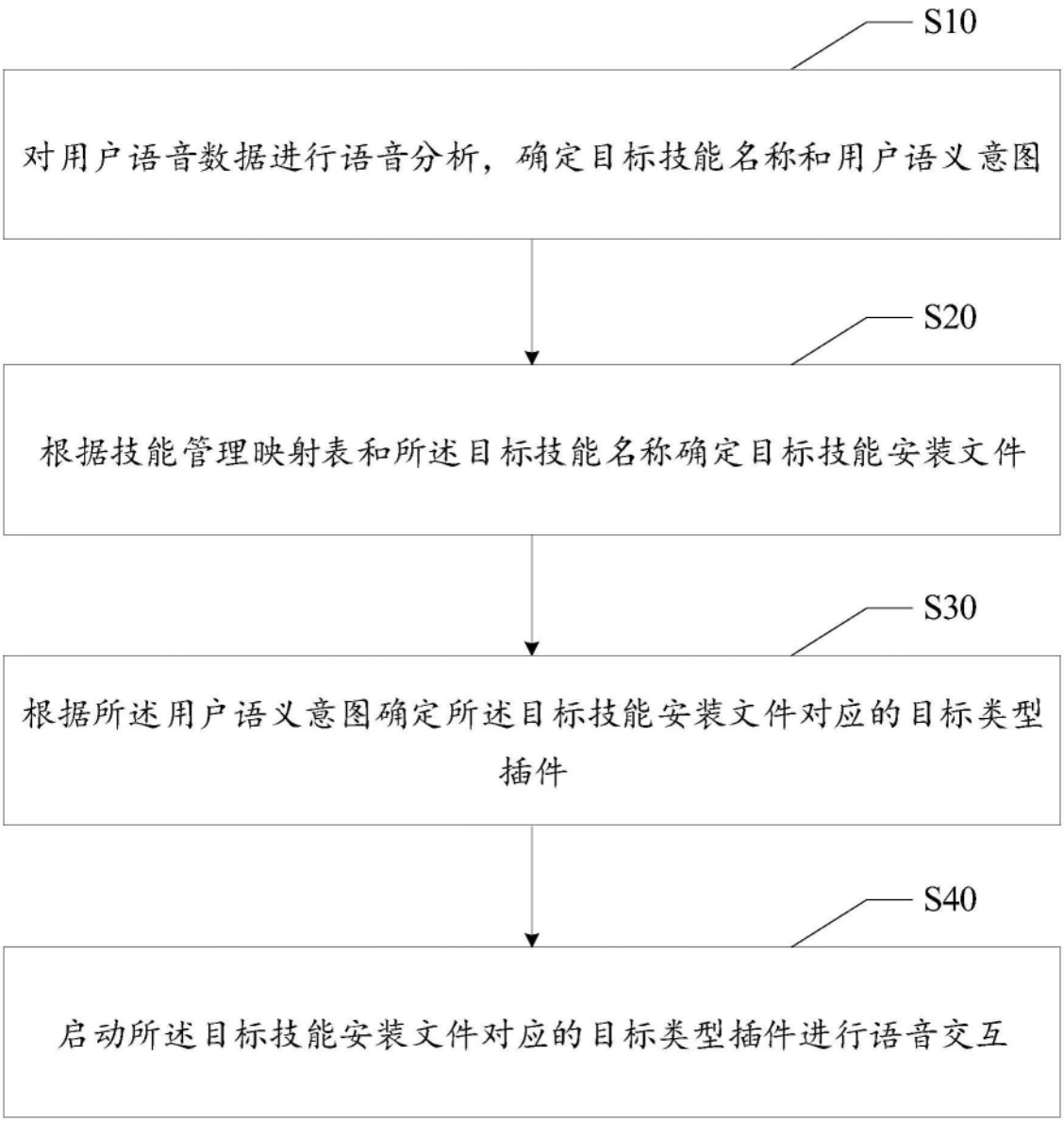

S10
对用户语音数据进行语音分析，确定目标技能名称和用户语义意图
S20
根据技能管理映射表和所述目标技能名称确定目标技能安装文件
S30
根据所述用户语义意图确定所述目标技能安装文件对应的目标类型插件
S40
启动所述目标技能安装文件对应的目标类型插件进行语音交互

图1

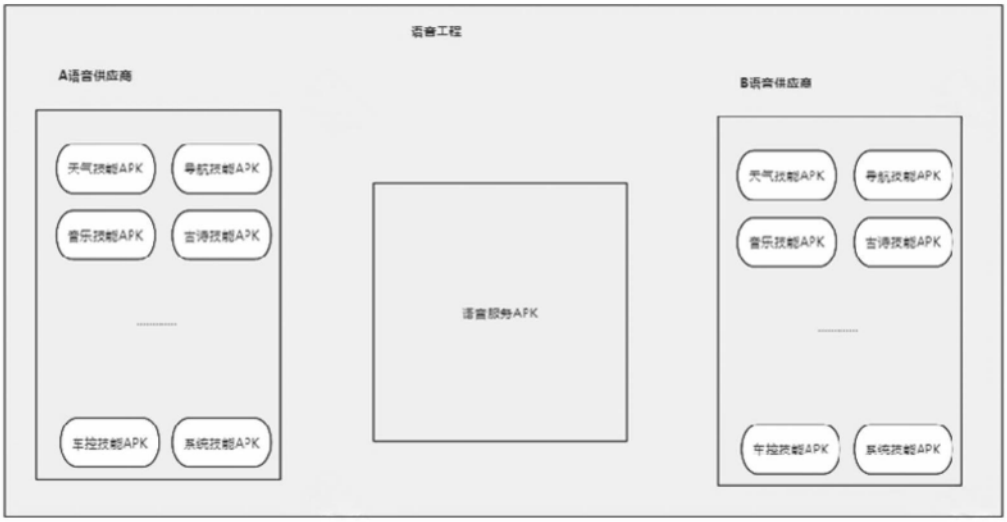


图2

```
{
    "list": [
        {
            "clazz": "com.voyah.skill.WeahterView",  //插件类名
            "intent": "Query", //处理的意图
            "type": "view"     // //插件类型
        },
        {
            "clazz": "com.voyah.skill.WeatherService",  //插件类名
            "intent": "INSTRUCTION", //处理的意图
            "type": "service"        // //插件类型
        }
    ],
    "skillId": "weather", // 对应技能
    "vendor": "iflytek" // 供应商
}
```

图3

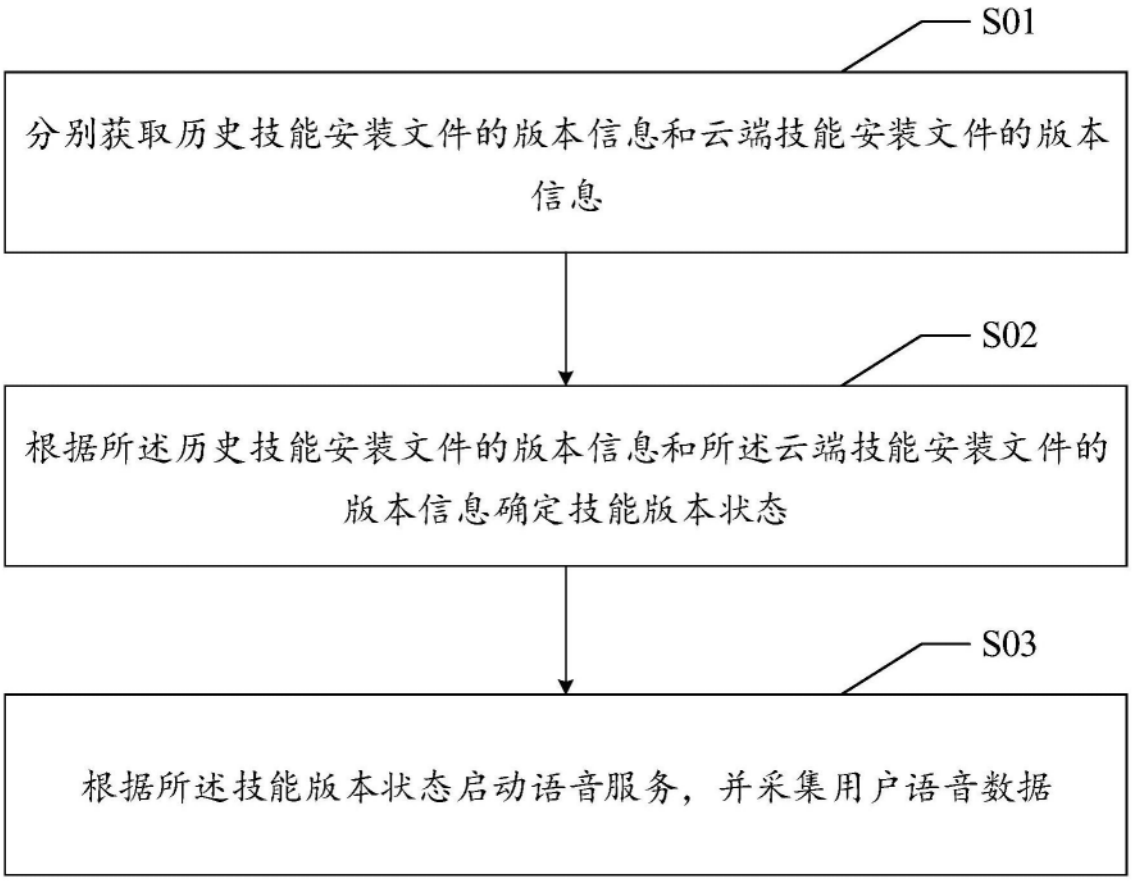

S01
分别获取历史技能安装文件的版本信息和云端技能安装文件的版本信息
S02
根据所述历史技能安装文件的版本信息和所述云端技能安装文件的版本信息确定技能版本状态
S03
根据所述技能版本状态启动语音服务，并采集用户语音数据

图4

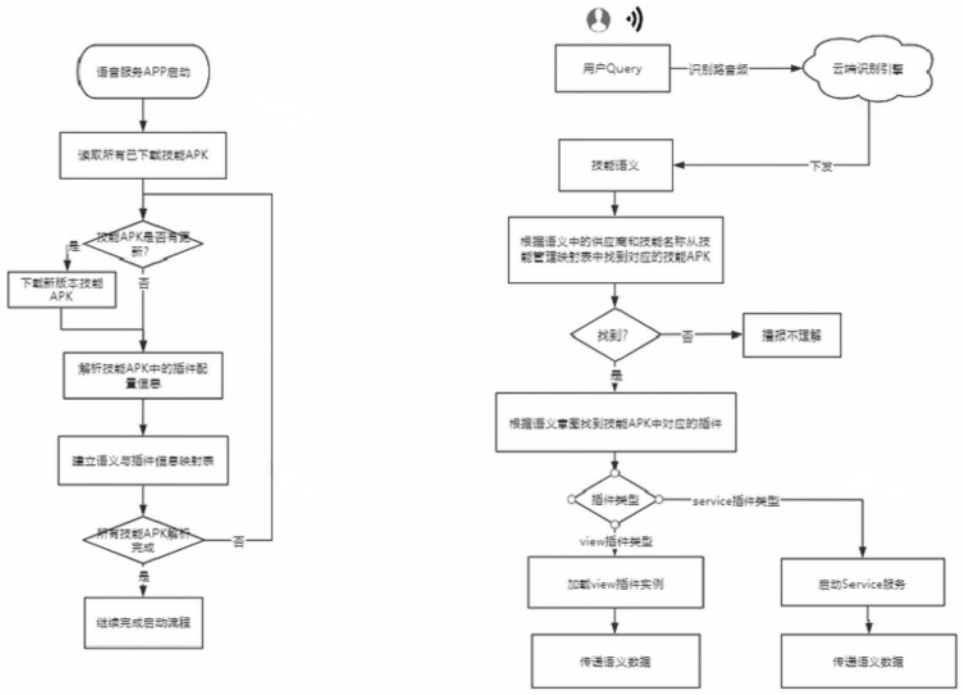

语音服务APP启动
读取所有已下载技能APK
技能APK是否有更新?
是
否
下载新版本技能APK
解析技能APK中的插件配置信息
建立语义与插件信息映射表
所有技能APK解析完成
否
是
继续完成启动流程
用户Query
识别语音流
云端识别引擎
技能语义
下发
根据语义中的供应商和技能名称从技能管理映射表中找到对应的技能APK
找到?
否
播报不理解
是
根据语义查图找到技能APK中对应的插件
插件类型
service插件类型
view插件类型
加载view插件实例
启动Service服务
传递语义数据
传递语义数据

图5

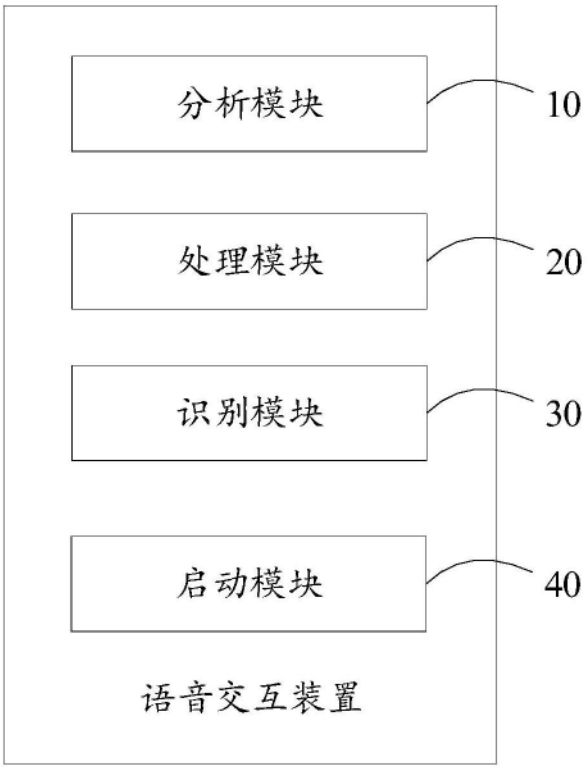

分析模块
10
处理模块
20
识别模块
30
启动模块
40
语音交互装置

图6

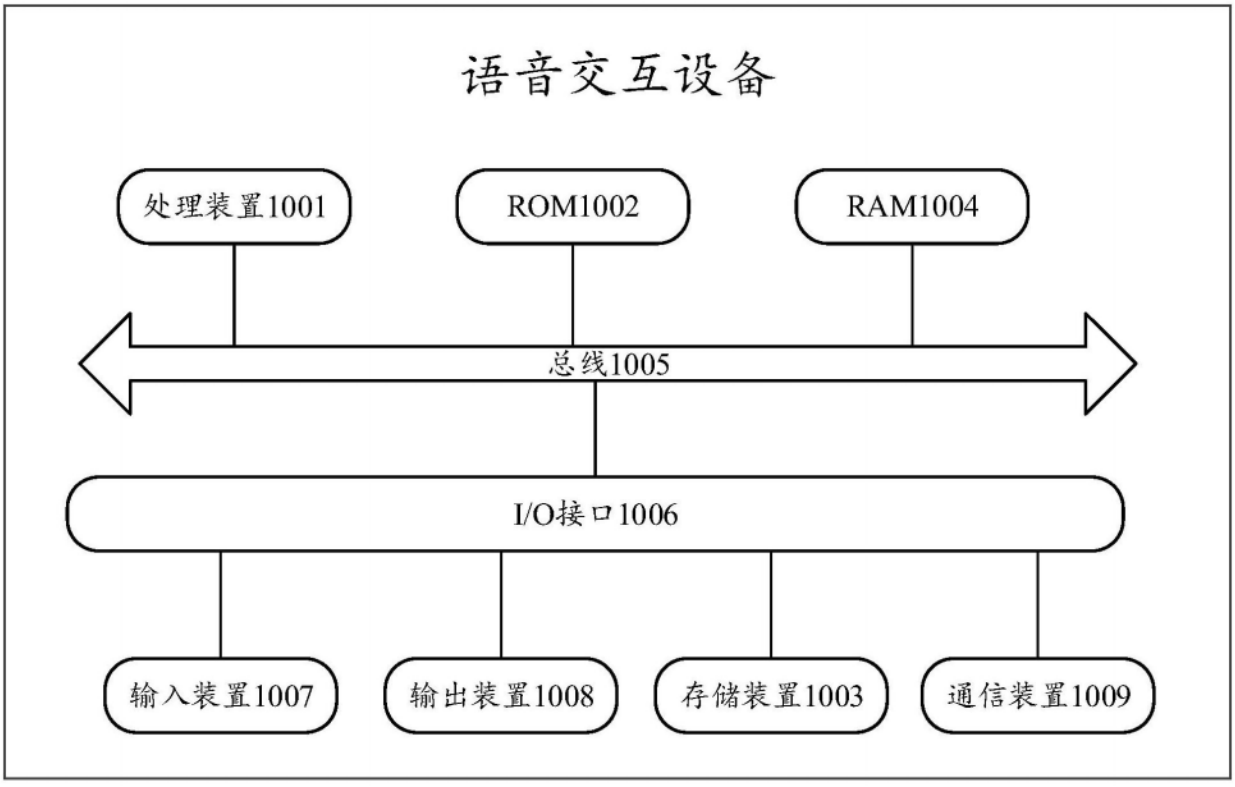

语音交互设备
处理装置1001
ROM1002
RAM1004
总线1005
I/O接口1006
输入装置1007
输出装置1008
存储装置1003
通信装置1009

图7