

## (19) 国家知识产权局



## (12) 发明专利



(10) 授权公告号 CN 112136102 B

(45) 授权公告日 2024.04.02

(21) 申请号 201980033449.9

(22) 申请日 2019.04.11

(65) 同一申请的已公布的文献号

申请公布号 CN 112136102 A

(43) 申请公布日 2020.12.25

(30) 优先权数据

2018-100418 2018.05.25 JP

(85) PCT国际申请进入国家阶段日

2020.11.18

(86) PCT国际申请的申请数据

PCT/JP2019/015875 2019.04.11

(87) PCT国际申请的公布数据

W02019/225201 JA 2019.11.28

(73) 专利权人 索尼公司

地址 日本东京

(72) 发明人 原昌宏 龟冈慎平

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

专利代理人 余刚

(51) Int.CI.

G06F 3/16 (2006.01)

G06F 3/01 (2006.01)

G10L 13/00 (2006.01)

G10L 15/00 (2013.01)

G10L 15/10 (2006.01)

G10L 15/30 (2013.01)

G10L 17/00 (2013.01)

(56) 对比文件

CN 104050966 A, 2014.09.17

CN 106469040 A, 2017.03.01

US 2017264451 A1, 2017.09.14

CN 105161106 A, 2015.12.16

US 9848082 B1, 2017.12.19

CN 103546788 A, 2014.01.29

CN 103631853 A, 2014.03.12

CN 106462617 A, 2017.02.22

US 2004044516 A1, 2004.03.04

US 2014129942 A1, 2014.05.08

US 2017358302 A1, 2017.12.14

审查员 李雪驹

权利要求书2页 说明书18页 附图10页

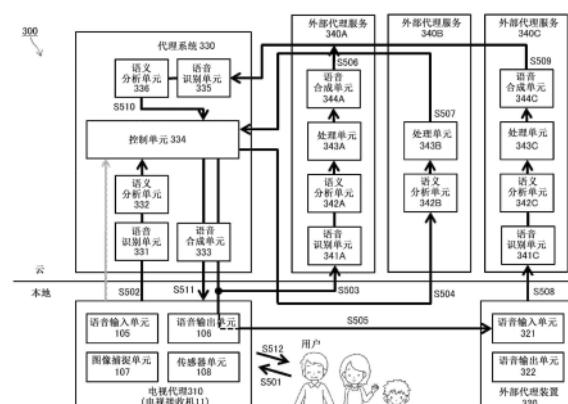
### (54) 发明名称

信息处理装置、信息处理方法以及信息处理系统

### (57) 摘要

提供了一种信息处理装置、一种信息处理方法和一种信息处理系统,通过其处理与代理的基于语音的对话。所述信息处理装置包括:通信单元,通过驻留在第一设备中的代理接收与用户对话有关的信息;以及控制单元,控制外部代理服务。控制单元收集以下至少一个:用户图像或语音;与用户对第一设备的操作有关的信息;或者由安装在第一设备中的传感器检测的传感器信息。控制单元控制外部代理服务的调用。

CN 112136102 B



1. 一种信息处理装置,包括:

语音识别单元,从驻留在第一设备中的代理接收与用户的查询对应的用户的话语,并转化为文本信息;以及

控制单元,基于所述文本信息调用多个外部代理服务中的一个或多个,

其中,所述控制单元在执行格式转换之后执行调用,以使所述调用符合每个外部代理服务的调用规范,

其中,当第一外部代理服务支持语音调用时,所述控制单元将对第一外部代理服务的所述文本信息的内容语音合成为语音数据,并将所述语音数据发送至所述第一外部代理服务,

当第二外部代理服务通过语音识别结果支持调用时,所述控制单元发送对第二外部代理服务的所述文本信息的内容,而不对所述内容进行语音合成。

2. 根据权利要求1所述的信息处理装置,其中,

所述控制单元收集信息,所收集的信息包括用户的图像或声音、与用户对所述第一设备的操作相关的信息、以及配备有所述第一设备的传感器检测到的传感器信息中的至少一个。

3. 根据权利要求1所述的信息处理装置,其中:

所述第一设备是电视接收机;以及

所述控制单元收集信息,所收集的信息包括用户的电视节目的观看历史和记录历史以及用户的遥控操作历史。

4. 根据权利要求2所述的信息处理装置,其中,

所述控制单元基于从所述第一设备收集的信息来管理用户的账户。

5. 根据权利要求1所述的信息处理装置,其中,

所述控制单元语音合成第二外部代理服务的调用内容,并在来自所述第一设备的所述第二外部代理服务的控制下执行外部代理装置的语音调用。

6. 根据权利要求1所述的信息处理装置,其中,

所述语音识别单元,识别来自外部代理服务的语音响应;所述信息处理装置进一步包括语义分析单元,所述语义分析单元对语音识别的结果进行语义分析。

7. 根据权利要求1所述的信息处理装置,其中,

通过API调用从外部代理服务接收响应。

8. 根据权利要求1所述的信息处理装置,其中,

所述控制单元通过从多个外部代理服务的响应中选择或合成,生成对用户对代理的查询的响应。

9. 根据权利要求1所述的信息处理装置,其中,

所述控制单元基于已经调用所述第一设备的用户的简档信息来控制外部代理服务的调用。

10. 根据权利要求9所述的信息处理装置,其中,

所述控制单元基于已经调用所述第一设备的用户的简档识别结果,执行具有简档切换功能的外部代理服务的调用。

11. 根据权利要求1所述的信息处理装置,其中,

所述控制单元基于已经调用所述第一设备的用户的敏感信息,控制外部代理服务的调用。

12.根据权利要求11所述的信息处理装置,其中,

所述控制单元用其他信息替换在用户调用中包括的敏感信息,然后执行外部代理服务的调用。

13.根据权利要求1所述的信息处理装置,其中,

所述控制单元基于用户的状态控制外部代理服务。

14.一种信息处理方法,包括:

从驻留在第一设备中的代理接收与用户的查询对应的用户的话语,并转化为文本信息;以及

基于所述文本信息调用多个外部代理服务中的一个或多个,

其中,在执行格式转换之后执行对多个外部代理服务中的一个或多个的调用,以使所述调用符合每个外部代理服务的调用规范,

其中,当第一外部代理服务支持语音调用时,所述方法包括将对第一外部代理服务的所述文本信息的内容语音合成为语音数据,并将所述语音数据发送至所述第一外部代理服务,

当第二外部代理服务通过语音识别结果支持调用时,所述方法包括发送对第二外部代理服务的所述文本信息的内容,而不对所述内容进行语音合成。

## 信息处理装置、信息处理方法以及信息处理系统

### 技术领域

[0001] 本说明书中公开的技术涉及一种处理基于语音的代理交互的信息处理装置和一种信息处理方法，并且涉及一种信息处理系统。

### 背景技术

[0002] 最近，在通过使用语音等与用户交互的同时，响应于使用和情况向用户呈现各种信息的代理开始传播。例如，传统上已知一种代理，其不仅对家用电器（例如，照明装置和空调）执行开/关和调节操作，而且例如当接收到关于天气预报、股票和交易所信息或新闻的查询时，通过语音响应于该查询，接收商品订单，并读出所购买书籍的内容。

[0003] 通常，通过安装在家庭等用户周围的代理装置和构建在云上的代理服务之间的合作来提供代理功能（例如，参考专利文献1）。例如，代理装置主要提供包括接受用户发出的语音的语音输入和响应于来自用户的查询的语音输出的用户界面。同时，在代理服务侧，执行高负载处理，高负载处理包括输入到代理装置的语音的语音识别和语义分析、诸如对应于用户查询的信息检索等处理、基于处理结果的语音合成等。

[0004] 此外，直接与用户交互的代理装置不仅是专用装置，而且是安装在室内的各种CE设备（例如，电视接收机、空调、录像机和洗衣机）以及各种信息设备，每个信息设备都具有驻留在其中的代理应用，各种信息设备包括物联网（IoT）装置、便携式信息终端（例如，智能电话和平板电脑）、交互式机器人、安装在汽车内部的汽车导航系统等（参考例如专利文献1）。

[0005] 现有技术

[0006] 专利文献

[0007] 专利文献1：日本未审查专利申请公开号2017-527844

[0008] 专利文献2：W02014/203495

[0009] 专利文献3：日本专利申请公开号2007-143010

[0010] 专利文献4：日本专利申请公开号2016-100613

### 发明内容

[0011] 本发明要解决的问题

[0012] 本说明书中公开的技术的目的是提供一种处理基于语音的代理交互的信息处理装置和一种信息处理方法，并提供一种信息处理系统。

[0013] 问题的解决方案

[0014] 本说明书中公开的技术的第一方面是一种信息处理装置，包括：

[0015] 通信单元，通过驻留在第一设备中的代理接收与用户交互相关的信息；以及

[0016] 控制单元，控制外部代理服务。

[0017] 所述控制单元收集信息，所收集的信息包括用户的图像或声音、与用户对第一设备的操作相关的信息、以及配备有第一设备的传感器检测到的传感器信息中的至少一个。

[0018] 此外,所述控制单元在执行格式转换之后执行调用,以使调用符合每个外部代理服务的调用规范。具体地,所述控制单元语音合成第一外部代理服务的调用内容,并执行第一外部代理服务的调用,或者语音合成第二外部代理服务的调用内容,并在来自第一设备的第二外部代理服务的控制下执行外部代理设备的语音调用。

[0019] 此外,本说明书中公开的技术的第二方面是一种信息处理装置,包括:

[0020] 通信单元,接收对用户的响应,所述响应通过处理与用户的交互相关的信息而生成;以及

[0021] 输出单元,输出响应。

[0022] 此外,本说明书中公开的技术的第三方面是一种信息处理方法,包括:

[0023] 通信步骤,用于通过驻留在第一设备中的代理接收与用户交互相关的信息;以及

[0024] 控制步骤,用于控制外部代理服务。

[0025] 此外,本说明书中公开的技术的第四方面是一种信息处理方法,包括:

[0026] 通信步骤,用于接收对用户的响应,所述响应通过处理与用户的交互相关的信息而生成;以及

[0027] 输出步骤,用于输出响应。

[0028] 此外,本说明书中公开的技术的第五方面是一种信息处理系统,包括:

[0029] 第一设备,与用户交互的代理驻留在该第一设备中;

[0030] 通信单元,通过代理接收与用户的交互相关的信息;以及

[0031] 控制单元,控制外部代理服务。

[0032] 然而,此处提到的“系统”是指逻辑上组装了多个装置(或分别实现特定功能的功能模块)的产品,因此装置或功能模块是否容纳在单个外壳中并不特别重要。

[0033] 本发明的效果

[0034] 根据本说明书中公开的技术,可以提供一种处理基于语音的代理交互的信息处理装置和一种信息处理方法、以及一种信息处理系统。

[0035] 应当注意,在本说明书中描述的效果应被解释为仅仅是说明性的,因此本发明的效果不限于在本说明书中描述的那些。此外,除了上述效果之外,本发明可以产生进一步的额外效果。

[0036] 基于以下描述的实施例和附图,根据以下详细描述,本说明书中公开的技术的其他目的、特征和优点将变得显而易见。

## 附图说明

[0037] 图1是示出本说明书中公开的技术的适用环境的示例的示图;

[0038] 图2是示意性示出能够作为代理装置操作的信息设备100的配置示例的示图;

[0039] 图3是示意性示出使用云的代理云平台的配置示例的示图;

[0040] 图4是示出代理系统330通过电视代理310在本地侧收集信息的状态的示图;

[0041] 图5是示出代理系统330调用外部代理服务340A、340B和340C的状态的示图;

[0042] 图6是示出代理系统330调用外部代理服务340的状态的示图;

[0043] 图7是示出当代理系统330通过使用用户简档信息不同地调用外部代理服务340时执行的处理步骤的流程图;

[0044] 图8是示出当代理系统330在保护用户敏感信息的同时不同地调用外部代理服务340时执行的处理步骤的流程图；

[0045] 图9是示出代理云平台300中的账户管理的示例的示图；

[0046] 图10是示出代理系统330响应于用户的状态来控制外部代理服务340A、340B和340C的状态的示图。

## 具体实施方式

[0047] 下面将参考附图详细描述本说明书中公开的技术的实施例。

[0048] 图1示出了客厅1，作为本说明书中公开的技术的适用环境的示例。电视接收机11和代理装置12安装在侧板13上。如下所述，电视接收机11具有驻留在其上的代理应用，并且还能够作为代理装置来操作。

[0049] 另外，在客厅1中，沙发14被安装成面向电视接收机11，并且沙发桌15设置在沙发14的前面。

[0050] 在图1所示的示例中，三个父母和一个孩子坐在沙发14上。这三个人是电视接收机11上显示的电视节目的观众，并且是在电视接收机11上操作的代理或代理装置12的用户。这三个人向代理查询，并等待答复。另外，假设每个用户可以使用遥控器(未示出)来执行电视接收机11的各种远程操作，包括频道选择、音量调节、输入输出装置的切换等。或者，每个用户还可以通过语音指示驻留在电视接收机11上的代理执行电视接收机11的各种操作，包括频道选择、音量调节、输入输出装置的切换等。

[0051] 从电视接收机11发出的声音的声波作为直接波到达坐在沙发14上的三个人，并且还作为从墙面、天花板和地板表面反射的波到达三个人。从电视接收机11发出的声音包括在频道选择之后正在接收的电视节目的声音以及从代理返回的合成声音。类似地，从代理装置12发出的声音作为直接波到达坐在沙发14上的三个人，并且也作为来自墙面、天花板和地板表面的反射波到达三个人。

[0052] 应当注意，本说明书中公开的技术的适用环境不限于如图1所示的客厅，客厅由普通家庭中的家庭成员共享。本说明书中公开的技术也可以应用于安装了用于提供代理功能的电视接收机的各种房间，这些房间包括作为特定用户的私人房间的书房、卧室等。此外，本说明书中公开的技术不仅可以应用于家庭，还可以应用于公司的办公室。

[0053] 图2示意性地示出了能够作为代理装置操作的信息设备100的配置示例。此处，图1中的电视接收机11被假设为信息设备100。信息设备100设置有处理单元101、调谐器102、通信单元103、显示单元104、语音输入单元105、语音输出单元106、图像捕捉单元107、传感器单元108、遥控接收单元109和记录单元110。

[0054] 调谐器102选择并接收地面广播的广播信号和卫星广播的广播信号。此外，通信单元103通过使用诸如以太网(注册商标)等有线通信或诸如Wi-Fi(注册商标)等无线通信连接到外部网络，例如，因特网。此外，通信单元103可以根据包括例如数字生活网络联盟(DLNA)(注册商标)的标准通过家庭网络互连到每个家庭内CE设备，或者可以进一步设置有与IoT装置的接口功能。

[0055] 显示单元104由例如液晶、有机EL元件等构成，设置有水平与垂直比为16:9的屏幕，并用于显示已经由调谐器102选择和接收的节目内容的视频和电子节目指南(EPG)以及

数据广播内容，并显示屏幕显示 (OSD)。应当注意，触摸传感器可以叠加在显示单元104的屏幕的部分或整个区域上。

[0056] 语音输入单元105由声音收集元件(例如，麦克风)配置，并且用于输入在安装有信息设备100的房间中出现的语音。电视节目的观众的话语或者使用语音代理的用户的话语可以被称为房间中出现的语音。语音输入单元105可以设置有多个麦克风组合成的麦克风阵列(一部分或全部麦克风可以从外部连接到信息设备100)。或者，语音输入单元105可以包括安装在电视接收机11的遥控器中的麦克风，或者可以包括安装在电视接收机11的外部装置中的麦克风，该外部装置包括智能电话、可佩戴装置等)。在语音输入单元105设置有多个麦克风的情况下，波束形成处理使得来自期望声源位置的声音(例如，与语音代理说话的用户的语音)增加，并且相反，使得抑制来自不必要声源位置的声音(例如，其他用户的语音、从房间中的其他AV设备输出的声音等)。

[0057] 语音输出单元106由声音生成元件(例如，扬声器)构成。语音输出单元106用于已经由调谐器102选择和接收的节目内容和数据广播内容的语音输出、从语音代理返回的合成语音的输出等。除了锥形扬声器之外，平板型扬声器(例如，参考专利文献3)可以用于语音输出单元106。此外，语音输出单元106可以设置有多个扬声器组合成的扬声器阵列(多声道扬声器或超级多声道扬声器)(部分或整个扬声器可以从外部连接到信息设备100)。在语音输出单元106设置有多个扬声器的情况下，声像可以被定位在期望的位置(例如，语音代理的语音的声像被定位在成为相对于已经进行查询的用户的预定相对位置的位置，或者被定位在声像没有与电视节目的声音混合的位置)，或者可以使声音难以从期望位置之外的位置听到。此外，还可以使用诸如“波前合成”等声场再现技术，该技术通过使用多个麦克风来吸收声场中的声音的波前，并且基于所获得的声音获取信号来再现声场(例如，参考专利文献4)。

[0058] 图像捕捉单元107由包括例如诸如互补金属氧化物半导体(CMOS)和电荷耦合器件(CCD)等图像传感器的相机构成，并且主要对位于信息设备100的前方(例如，显示单元104的屏幕前方)的用户进行图像捕捉。图像捕捉单元107可以是例如使用两个相机的立体相机或者使用三个或更多个相机的多眼相机。此外，一部分或全部相机可以从外部连接到信息设备100。

[0059] 传感器单元108主要感测安装信息设备100的房间的环境信息。可以任意确定传感器单元108的配置(换言之，要包括的一种传感器元件)。例如，传感器单元108可以包括对象检测传感器和深度传感器。基于对象检测传感器和深度传感器的检测结果(如果需要，结合由图像捕捉单元107捕捉的图像的图像识别结果)，可以检测安装在房间中的家具的布局。此外，传感器单元108可以设置有检测环境信息的环境传感器，该环境传感器包括照度传感器、温度传感器、湿度传感器等。此外，传感器单元108可以设置有红外传感器和人体检测传感器。此外，传感器单元108可以设置有检测用户的脉动、出汗、脑电波、肌电位、呼气等的活体传感器。形成传感器单元108的一部分或整个传感器单元可以从外部连接到信息设备100。

[0060] 遥控接收单元109通过使用红外通信、短距离无线通信等接收从遥控器(未示出)发送的遥控命令。遥控器的操作者例如是在图1所示的起居室中的电视接收机11的观看用户。

[0061] 记录单元110由例如诸如硬盘驱动器(HDD)等大容量记录装置配置,并且主要用于记录由调谐器102接收的节目内容。尽管记录单元110设置在信息设备100内部,但是也存在记录单元110通过诸如高清晰度多媒体接口(HDMI)(注册商标)等接口和通用串行总线(USB)从外部连接到信息设备100的情况。

[0062] 处理单元101包括处理器和存储器,执行加载到存储器中的程序,以执行各种处理,并且全面控制信息设备100中的操作。在处理单元101中,基本上,在由操作系统(OS)提供的执行环境下执行各种应用。例如,在可以使用多处理器的情况下,或者在OS支持多线程执行的情况下,可以并行执行的所有处理单元被读入可以并行执行处理单元的存储器中。

[0063] 可以称为在处理单元101中执行的处理的处理包括:基于由遥控接收单元109接收的遥控命令的设备100中的操作,该操作包括频道选择、音量调节、记录、图像质量调节等;输出由调谐器102选择和接收的节目内容的视频和声音的处理;记录在记录单元110上的节目内容的再现输出处理;执行通过调谐器102或通信单元103接收的数据广播应用;以及执行各种应用,例如,语音代理。此外,通过使处理单元101并行执行多个代理应用,信息设备100还能够用作多个代理装置。

[0064] 通常,通过安装在用户周围的代理装置和构建在云上的代理服务之间的合作来提供代理功能。图3示意性地示出了使用云的代理云平台300的配置示例。代理云平台300分为本地端和云端。

[0065] 代理云平台300的本地侧对应于例如图1所示的客厅1,并且包括电视代理310和外部代理装置320。

[0066] 电视代理310是驻留在信息设备100(对应于图1中的电视接收机11)中的代理,并且能够利用信息设备100配备的每个功能模块,该功能模块包括语音输入单元105、语音输出单元106、图像捕捉单元107、传感器单元108等。

[0067] 外部代理装置320对应于图1中的代理装置12。外部代理装置320是专用于语音代理的装置,并且设置有用于与用户交互的语音输入单元321和语音输出单元322。然而,外部代理装置320不限于专门用于语音代理的装置,其也被称为“智能扬声器”、“AI扬声器”、“AI助手”等。外部代理装置320可以是除电视接收机11之外的CE设备,或者可以是驻留在IoT装置、信息终端等中的语音代理。

[0068] 从用户的角度来看,电视代理310与作为执行基于语音的交互的代理的外部代理装置320没有区别。然而,由于前者通过电视接收机11与用户交互(或提供与电视具有高亲和力的代理服务),所以前者被称为“电视代理”;并且由于后者在电视接收机11之外与用户交互(或者提供与电视广播服务关系不大的代理服务),后者被称为“外部代理装置”。因此,为了方便起见,两者都有所区别。

[0069] 同时,代理云平台300的云侧包括代理系统330和多个外部代理服务340A、340B、340C...。从用户的角度来看,从提供响应本地用户查询的代理服务的角度来看,代理系统330与外部代理服务340A、340B、340C...没有不同。然而,由于前者通过直接与电视代理310合作来提供代理服务,所以前者被称为“代理系统”;并且因为后者不直接与电视代理310合作,而是从电视接收机11的外部提供代理服务,所以后者被称为“外部代理服务”。因此,为了方便起见,这两者都有所区别。代理系统330可以构建在由例如电视接收机11的产品制造商或销售代理管理的服务器装置上。然而,代理系统330不受特定管理形式的限制。此外,外

部代理服务340A、340B、340C、... (在图3所示的示例中,外部代理服务340C) 直接与安装在本地的外部代理装置320合作。

[0070] 代理系统330设置有语音识别单元331、语义分析单元332、语音合成单元333和控制单元334。

[0071] 语音识别单元331语音识别由电视代理310侧的语音输入单元105收集的用户话语,并将用户话语转换成文本信息。语义分析单元332基于语音识别的文本信息对用户的话语数据进行语义分析。另外,语音合成单元333将作为对用户的响应的文本信息转换成语音数据。语音数据传输到电视代理310,并且从语音输出单元106被语音输出到作为查询源的用户。应当注意,电视代理310可以适于通过在显示单元104的屏幕上不仅使用语音输出而且使用视频输出来输出对来自用户的查询的响应。此外,例如,电视代理310可以适于通过一起使用经由家庭网络连接到电视接收机11的其他CE设备或其他IoT装置的输出功能来输出对用户查询的响应。

[0072] 控制单元334执行响应来自电视代理310的用户的查询的处理。例如,为了使控制单元334响应来自用户的查询,该查询包括搜索用户引用的信息、获取用户请求的内容以及订购商品,控制单元334可以设置有学习功能或人工智能(AI)功能。

[0073] 此外,为了响应来自用户的查询,控制单元334能够进一步调用(call)外部代理服务340A、340B、340C、...。控制单元334可以调用所有可请求的外部代理服务340A、340B、340C、...或者可以选择并调用一些外部代理装置。例如,在外部代理服务340A、340B、340C、...分别具有特性的特定区域的情况下,控制单元334可以根据来自用户的查询内容、进行查询的用户的简档信息等,接受或拒绝请求目的地外部代理服务。另外,在用户进行指定特定外部代理服务的“启动字”的语音输入的情况下,控制单元334可以原样请求对应于输入的启动字的外部代理服务,或者可以在考虑启动字的同时接受或拒绝请求目的地外部代理服务。

[0074] 此外,还考虑了一些从代理系统330调用外部代理服务340A、340B、340C、...的方法。例如,通过语义分析单元332对来自用户的查询进行语义分析而获得的文本信息可以传输到外部代理服务340B,或者通过将通过语音合成单元333对文本信息进行语音合成而获得的语音数据传输到外部代理服务340A,可以以类似于来自用户的语音输入的模式进行查询。或者,在代理系统330向外部代理服务340C发出请求的情况下,通过将由语音合成单元333语音合成的语音数据传输到电视代理310,从电视代理310的语音输出单元106输出听起来像在房间中进行查询的用户的语音的声音,从而可以通过外部代理装置320进行查询。

[0075] 此外,还考虑了一些从请求目的地外部代理服务340A、340B、340C、...响应于代理系统100的方法。例如,与外部代理服务340B一样,存在包括文本信息等(或在语音合成之前)的处理结果传输到代理服务330的情况。另外,与外部代理服务340A和340C一样,也存在通过语音合成处理结果获得的语音数据传输到代理服务330的情况。代理服务330还设置有语音识别单元335和语义分析单元336,语音识别单元335对从例如外部代理服务340A和340C传输的语音数据进行语音识别,语义分析单元336对语音识别结果进行语义分析。

[0076] 此外,控制单元334在语义分析结果级别聚集来自外部代理服务340A、340B和340C中的每一个的响应,并生成响应来自用户的查询的响应语句。可以任意确定用于聚集多个响应的方法。可以通过与代理系统330自身获得的响应进行比较,来选择来自特定外部代理

服务的响应;可以基于包括多数决定等的预定选择标准,来选择来自特定外部代理服务的响应;或者可以通过合成来自多个外部代理服务的响应来生成一个响应语句。在任何情况下,语音合成单元333将控制单元334生成的响应语句转换成语音数据,并将语音数据传输到电视代理310。随后,在电视代理310中,从语音输出单元106向用户输出对作为查询源的用户的响应的语音。

[0077] 应当注意,在语音识别单元331、语义分析单元332、语音合成单元333、控制单元334、语音识别单元335和语义分析单元336中,代理系统330配备的一部分或全部功能模块可以包含到电视代理310中。或者,这些功能模块331至336可以构建在一个服务器装置上,或者可以以分布式方式设置在多个服务器装置中。

[0078] 相应外部代理服务340A、340B、340C、...的功能配置基本相似。下面将代表性地描述外部代理服务340C的功能配置。

[0079] 外部代理服务340C配备有语音识别单元341C、语义分析单元342C、处理单元343C和语音合成单元344C。

[0080] 语音识别单元341C语音识别由外部代理装置320的语音输入单元321收集的话语,并将该话语转换成文本信息。应当注意,还假设由语音输入单元321收集的语音不仅包括用户的话语,还包括从电视代理310的语音输出单元106输出的语音的情况。

[0081] 语义分析单元342C基于语音识别的文本信息对用户的话语数据进行语义分析。处理单元343C执行响应来自用户的查询的处理。例如,当语义分析结果显示已经从用户语音输入了指定外部代理服务340C的“启动字”时,处理单元343C开始执行处理。

[0082] 语音合成单元344C将作为对用户的响应的文本信息转换成语音数据。语音数据传输到外部代理装置320,并且从语音输出单元322被语音输出到作为查询源的用户。应当注意,尽管由语音合成单元333生成的语音数据传输到外部代理装置320并被语音输出,但是也假设语音数据传输到代理系统330并被输入到语音识别单元335的情况。

[0083] 顺便提及,为了简化附图,图3示出了云侧的一个代理系统330向一个电视代理310提供代理服务。然而,假设一个代理系统330同时向在多个电视接收机上运行的多个电视代理提供代理服务。换言之,实际上,一个代理系统330同时向安装在家庭中的大量电视接收机的电视代理提供代理服务。

[0084] 此外,由于页面空间的限制,图3示出了三个外部代理服务340A、340B和340C。然而,与代理系统330合作的外部代理服务的数量不限于三个,因此假设代理系统330与两个或更少或四个或更多外部代理服务合作。

[0085] 此外,尽管图3示出了一个外部代理装置320,但是当实现本说明书中公开的技术时,外部代理装置并不总是必要的组件。在其他情况下,应当理解,即使两个或更多外部代理装置在本地侧共存,也可以类似地实现本说明书中公开的技术。

[0086] 图3所示的多个代理服务中的外部代理服务340A、340B和340C也被称为例如“AI代理”或“AI扬声器”,并且在提交本申请时已经是广泛使用的服务。假设这种外部代理服务从属于无限数量的外部代理装置(对应于图3中的外部代理装置320),这些外部代理装置可以经由广域网(例如,互联网)连接。每个用户都可以通过自己拥有的外部代理装置登录到外部代理服务。每个外部代理服务为已经登录到其的每个用户设置账户,并且利用与该账户相关联的简档信息和历史信息来管理每个用户的简档信息和历史信息(以下也称为“账户

切换”功能)。此外,例如,通过连续更新简档信息和历史信息,或者通过采用学习功能,每当从用户接收到查询时,外部代理服务能够提供针对每个单独用户定制或个性化的精细服务。

[0087] 此外,在图3所示的多个代理服务中,假设代理系统330由例如电视接收机11的产品制造商或销售代理来管理,并且通过直接与电视代理310合作来提供代理服务。尽管代理系统330也是AI代理,但是假设电视代理310的用户基本上限于观看电视接收机11的用户。代理系统330识别已经由电视接收机11的图像捕捉单元107和语音输入单元105(即,在本地侧)捕捉的图像和语音,基于个人识别的结果来识别用户,并且基于用户来管理账户。随后,每当从用户接收到查询时,或者基于传感器单元108(可以包括图像捕捉单元107和语音输入单元105)的检测结果,通过更新每个用户的简档信息和历史信息,或者通过采用学习功能,代理系统330能够提供针对每个单独用户定制或个性化的精细服务。

[0088] 应当注意,在图3所示的代理云平台300中,尽管代理系统330被配置为与驻留在电视接收机11中的电视代理直接合作,但是代理系统330能够与之直接合作的设备类型不限于电视接收机。例如,还假设驻留在各种信息设备中的代理直接与代理系统330合作的修改示例,各种信息设备包括:可以由用户携带的信息终端,例如,智能电话和平板电脑;用户佩戴的可穿戴设备,或各种CE设备,例如,空调、录音机和洗衣机;IoT装置;交互式机器人;安装在汽车内部的汽车导航系统等。此外,在代理系统330与电视接收机之外的设备合作的情况下,外部代理装置320可以是驻留在电视接收机中的语音代理。

[0089] 随后,将描述代理云平台300的操作示例。

[0090] 图4示出了代理系统330通过电视代理310在本地侧收集信息的状态。

[0091] 电视代理310总是监视例如电视节目的观看历史、当前观看情况、电视接收机11本身的状态(开/关、音量、图像质量设置等)、记录历史和电视接收机11的遥控器的操作历史,并将这些信息传输到代理系统330。在代理系统330侧,与用户观看电视接收机11中的电视节目的历史相关的这些信息总是从电视代理310收集的,并且与例如用户简档信息和用户账户信息相关联地管理这些信息。此外,电视代理310可以向代理系统330传输从根据诸如DLNA(注册商标)等标准经由家庭网络与电视接收机11互连的每个CE设备获得的或者从安装在同一房间中的IoT装置获得的信息。

[0092] 另外,电视代理310向代理系统330传输由电视接收机11设置有的图像捕捉单元107捕捉的用户图像的面部图像和输入到语音输入单元105中的用户的语音数据。在代理系统330侧,从电视代理310接收的图像经受面部识别,并且语音数据被语音识别,从而执行个人识别处理。或者,电视代理310可以适于对由图像捕捉单元107捕捉的图像进行图像识别,或者对输入到语音输入单元105中的语音进行语音识别,然后将个人识别处理的结果传输到代理系统330。代理系统330可以适于在每次新识别一个人时设置新的用户账户。此外,代理系统330可以适于基于用户分发已经通过电视代理310从本地侧收集的信息片段,并且通过将每个片段与用户账户相关联来管理信息片段。此外,代理系统330可以适于基于已经从电视代理310接收的图像数据和语音数据来收集与用户状态相关的信息(例如,用户是否正在观看正由电视接收机11选择和接收的电视节目)。

[0093] 此外,电视代理310向代理系统330传输由信息设备100设置有的传感器单元108检测到的检测数据。在传感器单元108包括对象检测传感器、深度传感器和检测环境信息的环

境传感器(照度传感器、温度传感器、湿度传感器等)的情况下,代理系统330可以收集这些传感器信息来管理用户的环境信息。此外,在传感器单元108设置有检测用户的脉动、出汗、脑电波、肌电位、呼气等的活体传感器的情况下,代理系统330可以从电视代理310收集这些传感器信息,以基于用户管理生物信息。

[0094] 顺便提及,还存在电视代理310通过电视接收机11获得的信息包括用户的敏感信息的情况。因此,电视代理310可以适于在向代理系统330提供本地信息之前屏蔽用户的敏感信息。将在后面详细描述敏感信息的屏蔽。

[0095] 电视代理310和代理系统330可以适于通过按个人识别划分用户简档来基于个人执行账户管理。例如,通过对由图像捕捉单元107捕捉的用户图像的面部图像进行面部识别,或者通过对从语音输入单元105输入的用户话语进行语音识别来执行个人识别。另外,电视代理310和代理系统330可以基于电视节目的观看历史、遥控器的操作历史等来执行个人识别。例如,可以基于用户将观看偏好的倾向与观看电视节目的习惯(包括一周中的某一天、时区等)分开,因此,通过基于用户将操作遥控器和各种设备的操作习惯分开,操作习惯可以用于个人识别。

[0096] 图9示出了代理云平台300中的账户管理的示例。

[0097] 基于例如由图像捕捉单元107捕捉的用户图像的面部识别结果,电视代理310掌握由“父亲”、“母亲”和“孩子”组成的家庭结构,设置“父亲”、“母亲”和“孩子”各自的账户,并管理每个家庭成员的简档。

[0098] 此外,代理系统330与电视代理310共享用户的账户信息。换言之,在图9所示的示例中,代理系统330掌握由“父亲”、“母亲”和“孩子”组成的家庭结构,并设置“父亲”、“母亲”和“孩子”各自的账户。然而,代理系统330不需要亲自识别用户,并且代理系统330可以通过例如家庭关系等来管理账户。

[0099] 同时,例如,当从代理系统330调用外部代理服务340C时,并且当划分用户的简档,以便从外部代理装置320调用时,外部代理服务340C可以划分简档,以设置账户。此外,外部代理装置320可以适于通过识别当被用户调用时获得的语音来设置账户。在图9所示的示例中,外部代理服务340C和外部代理装置320共享“父亲”的账户信息。

[0100] 应当注意的是,在提交本申请时,对于如何由代理管理用户账户没有特别的安排。在本实施例中,假设代理系统330能够基于一个电视代理310中的个人识别来设置多个用户的账户(在图9所示的示例中,设置了“父亲”、“母亲”和“孩子”各自的账户)。

[0101] 此外,假设外部代理服务340A、340B和340C都能够任意执行账户管理。例如,在只为一个外部代理装置320设置代表用户的一个账户的情况下,当账户被分配给作为一个用户的一个外部代理装置320时,可以允许在用户账户下注册多个用户的简档,或者可以允许在一个用户的账户下注册多个子用户。另外,在也从代理系统330调用外部代理服务340A、340B和340C中的每一个的情况下,外部代理服务340A、340B和340C中的每一个可以将类似的账户管理应用于代理系统330,账户管理包括:仅设置代表性用户的一个账户;在一个用户账户下分配多个用户的简档;并且在一个用户账户下注册多个子用户。

[0102] 图5示出了代理系统330调用外部代理服务340A、340B和340C的状态。例如,当用户向电视代理310进行查询时,代理系统330调用外部代理服务340A、340B和340C。

[0103] 当用户通过语音向电视代理310进行查询时(S501),语音输入单元105收集用户的

语音,并将该语音传输到代理系统330。随后,在代理系统330中,语音识别单元331语音识别用户的语音,并且语义分析单元332执行语义分析(S502)。另外,为了响应于来自用户的查询,控制单元334调用外部代理服务340A、340B和340C。

[0104] 此时,控制单元334可以调用所有可调用的外部代理服务340A、340B和340C,或者可以选择性地调用任何外部代理服务。例如,根据来自用户的查询的内容、已经进行查询的用户的简档信息等,控制单元334可以接受或拒绝外部代理服务调用外部代理服务。另外,在用户进行指定特定外部代理服务的“启动字”的语音输入的情况下,控制单元334可以原样调用对应于输入的启动字的外部代理服务,或者可以接受或拒绝外部代理服务调用外部代理服务,同时考虑启动字。

[0105] 即使在调用任何外部代理服务的情况下,代理系统330也对来自用户的输入语音进行语音识别和语义分析,随后,在控制单元334执行格式转换,以符合每个外部代理服务的调用规范之后,代理系统330执行调用。

[0106] 在代理系统330调用外部代理服务340A的情况下,由于外部代理服务340A支持语音调用,代理系统330通过语音合成单元333将对外部代理服务340A的查询的内容转换成语音数据,已经由控制单元334生成该内容,然后将语音数据发送到外部代理服务340A(S503)。例如,由语音合成单元333转换的语音数据可以与从电视代理310传输的语音数据基本相同,或者可以是已经由控制单元334处理其内容的语音数据。

[0107] 在这种情况下,在外部代理服务340A中,从代理系统330传输的语音数据由语音识别单元341A进行语音识别,并进一步由语义分析单元342A进行语义分析。然后处理单元343A执行响应来自代理系统330的查询的处理。随后,由语音合成单元344A将作为来自代理系统330的查询的处理结果的文本信息转换成语音数据,然后将语音数据返回给代理系统330(S506)。

[0108] 此外,在代理系统330调用外部代理服务340B的情况下,由于外部代理服务340B通过语音识别结果支持调用,所以代理系统330将对外部代理服务的查询内容原样传输到外部代理服务340B(即,没有语音合成内容),已经由控制单元334生成该内容(S504)。对外部代理服务的查询的内容(已经由控制单元334生成该内容)可以与从电视代理310传输的语音数据的语音识别结果基本相同,或者可以是语音数据,已经由例如控制单元334处理其内容。

[0109] 在这种情况下,在外部代理服务340B中,来自代理系统330的查询由语义分析单元343B进行语义分析,并且处理单元343B执行响应来自代理系统330的查询的处理,并且将处理结果原样返回给代理系统330(即,不语音合成处理结果)(S507)。代理系统330侧的控制单元334准备用于从外部代理装置320接收响应的应用编程接口(API)。

[0110] 另外,在代理系统330调用外部代理服务340C的情况下,由于外部代理服务340C支持从与其合作的外部代理装置320进行调用,所以代理系统330通过语音合成单元333将对外部代理服务的查询的内容转换成语音数据,已经由控制单元334生成该查询,然后将语音数据传输到电视代理310,这使得通过语音从电视代理310调用外部代理装置320(S505)。例如,从电视代理310调用外部代理装置320的语音可以与用户在S501中向电视代理310进行查询时使用的语音基本相同,或者可以是语音,已经由控制单元334处理其内容。此外,语音可以包括用于启动外部代理装置320的“启动字”。

[0111] 在这种情况下,由外部代理装置320的语音输入单元321收集的语音传输到外部代理服务340C (S508)。此外,在外部代理服务340C中,从外部代理装置320传输的语音数据由语音识别单元341C进行语音识别,并进一步由语义分析单元342C进行语义分析。处理单元343C然后执行响应来自代理系统330的查询的处理。随后,作为来自代理系统330的查询的处理结果的文本信息被语音合成单元344转换成语音数据,然后该语音数据被返回给代理系统330 (S509)。

[0112] 在代理系统330中,从外部代理服务340A和外部代理服务340C中的每一个返回的语音数据处理结果由语音识别单元335进行语音识别,并且语音识别的结果进一步由语义分析单元336进行语义分析,然后提供给控制单元334。此外,控制单元334通过从外部代理服务340B调用的API来接收响应。

[0113] 控制单元334在语义分析结果级别聚集来自外部代理服务340A、340B和340C中的每一个的响应,并生成响应来自用户的查询的响应语句。可以任意确定用于聚集多个响应的方法。可以通过与代理系统330自身获得的响应进行比较来选择来自特定外部代理服务的响应,或者可以基于包括多数决定等的预定选择标准来选择来自特定外部代理服务的响应。或者,可以通过合成来自多个外部代理服务的响应来生成一个响应语句。

[0114] 语音合成单元333将控制单元334生成的响应语句转换成语音数据,并将语音数据传输到电视代理310 (S511)。随后,在电视代理310中,从语音输出单元106向用户输出对作为查询源的用户的响应的语音 (S512)。

[0115] 简而言之,通过使用安装在用户所在房间的电视接收机11上的电视代理310,代理系统330主要使用语音交互来提供代理服务。此外,除了用户向用户自身返回查询处理的结果之外,通过调用外部代理服务340A、340B和340C,代理系统330还能够通过利用外部资源来响应用户。代理系统330能够通过转换成符合每个外部代理服务340A、340B和340C的调用规范的格式来调用每个外部代理服务340A、340B和340C。期望代理系统330通过使用外部代理服务340A、340B和340C的调用向用户返回更高质量的响应。

[0116] 在此处,作为当代理系统330调用外部代理服务340A、340B和340C时产生的一个问题,可以提到多条用户信息混合的问题。

[0117] 假设当家庭成员正在观看如图6所示的客厅中的电视接收机11正在选择和接收的电视节目时,一个家庭成员已经向电视代理310进行了查询。当代理系统330调用外部代理服务340时,即使在家庭成员中的任何人进行查询的情况下,该查询也是由作为“代理系统330”的同一用户的账户进行的,并且相应地,所有家庭成员的多条信息(简档信息、历史信息等)混合。例如,即使家庭成员中的任何人向电视代理310进行查询,如果代理系统330总是合成相同的语音,来调用外部代理服务340A,则外部代理服务340A不能够掌握是谁进行了查询。结果,即使外部代理服务340配备有账户切换功能,也难以提供对应于作为查询源的每个用户的精细服务。

[0118] 代理系统330能够对由图像捕捉单元107捕捉的用户的图像进行面部识别,对由语音输入单元105输入的语音进行语音识别,基于已经通过电视接收机11收集的观看历史和遥控操作历史或传感器单元108的检测结果来识别作为查询源的用户的简档等。

[0119] 因此,代理系统330识别已经通过语音进行查询的用户的简档,并且根据外部代理服务340的账户切换功能来映射用户的简档,并且通过使用该简档来不同地执行调用。

[0120] 例如,在代理系统330调用外部代理服务340A或340C的情况下,在查询来自父亲的情况下,合成成人男性音调和音色以执行调用,在查询来自母亲的情况下,合成成人女性音调和音色以执行调用,并且在查询来自儿童的情况下,合成儿童音调和音色以执行调用。此外,在代理系统330调用外部代理服务340B的情况下,在查询来自父亲的情况下,通过具有成人男声的调用语句来执行调用,并且在查询来自儿童的情况下,通过具有童声的调用语句来执行调用。不用说,代理系统330可以添加用户简档信息,作为元数据,来调用外部代理服务340。

[0121] 外部代理服务340A、340B和340C中的每一个基于如何调用代理系统330来识别用户的简档,以适当地执行账户切换,并且向代理系统330返回对查询的响应。

[0122] 在外部代理服务340A、340B和340C的每一个中,有各种用于账户切换的方法。例如,当从代理系统330接收到查询时,每次切换简档时,可以分配单独的用户账户。此外,当账户作为一个用户被分配给代理系统330时,可以在用户账户下注册相应的家庭成员的多个简档。或者,当一个账户作为一个用户被分配给代理系统330时,相应的家庭成员可以在用户账户下注册为子用户。

[0123] 图7以流程图的形式示出了当代理系统330通过使用用户的简档信息不同地调用外部代理服务340时执行的处理步骤。

[0124] 当代理系统330通过电视代理310输入来自用户的语音时(步骤S701),代理系统330基于图像捕捉单元107捕捉的用户图像的面部识别结果、输入语音的语音识别结果和传感器单元108的检测结果来识别用户的简档(步骤S702)。

[0125] 接下来,代理系统330对已经在步骤S701中输入的语音进行语义分析,并选择应该被调用的外部代理服务(步骤S703)。

[0126] 接下来,代理系统330检查所选择的外部代理服务是否具有账户切换功能(步骤S704)。

[0127] 在此处,在被调用的外部代理服务具有账户切换功能的情况下(步骤S704中的是),通过使用在步骤S702中识别的用户的简档,代理系统330生成符合外部代理服务的调用,并执行该调用(步骤S705)。例如,代理系统330用适合于包括用户的年龄、性别等的简档的音调和音色调用外部代理服务。

[0128] 同时,在被调用的外部代理服务不具有账户切换功能的情况下(步骤S704中为否),则在不使用在步骤S702中识别的用户的简档的情况下,代理系统330通过代表用户执行外部代理服务的调用(步骤S706)。

[0129] 另外,作为当代理系统330调用外部代理服务340A、340B和340C时产生的另一个问题,会涉及将泄露用户的敏感信息的问题。

[0130] 在语音输入单元105在安装在客厅中的电视接收机11中输入语音的使用模式中,家庭成员倾向于无意识地进行包括与敏感信息相关的单词的对话,因此从用户到代理系统330的查询可以包括敏感信息。在其他情况下,也存在这样的可能性,尽管用户不打算进行查询,但是将在代理系统330侧自动确定已经进行了查询,并且因此将启动服务。

[0131] 在JIS Q 15001:2006的个人信息保护和管理标准中,提到了敏感信息“与思想、信仰和宗教有关的问题”、“导致社会歧视的问题,例如,种族、民族、家庭出身和注册住所(不包括仅关于注册住所所在地区的信息的情况)、身心障碍、犯罪记录等”、“与集体活动行为

有关的问题,例如,劳工的组织权、集体谈判权等”、“与行使政治权利有关的问题,例如,参加群众示威、行使请愿权等”、以及“与保健和性生活有关的问题”。此外,除了上述标准中规定的问题之外,具有商业、金融或个人特征的各种信息(例如,密码和账号)也是敏感的个人信息。如果此类敏感信息被非法人员获取、使用和传播,会带来各种风险,例如,侵犯个人的人格和财产。另外,通常的情况是,通过图像捕捉单元107对客厅等中的家庭状态进行图像捕捉而获得的视频包括应当被保护的各种敏感信息,该敏感信息包括面部图像。

[0132] 一些保护敏感信息的技术是已知的。例如,存储敏感信息的装置以具有高安全级别的模式运行。然而,在通过使用安装在客厅中的电视接收机11中(或外部连接到电视接收机11)内置的语音输入单元105、图像捕捉单元107、传感器单元108等来收集信息的情况下,不可能防止混合敏感信息。

[0133] 如果代理系统330调用包含敏感信息的外部代理服务340A、340B和340C,则敏感信息有可能从任何外部代理服务开始传播,并且将被非法使用。此外,通过使用这样的方法,其中,在从电视代理310收集的多条信息中,敏感信息被代理系统330简单地用黑色着色,以隐藏敏感信息,不能读取查询的内容,因此外部代理服务340A、340B和340C不能提供正常的代理服务。

[0134] 因此,通过用其他信息替换来自用户的查询中包括的敏感信息,代理系统330调用外部代理服务340A、340B和340C,同时保持内容本身处于可读状态。

[0135] 例如,在来自用户的查询包括疾病名称的情况下,疾病名称被替换为单词“疾病”,在来自用户的查询包括个人名称的情况下,个人名称被替换为不能识别个人的单词,例如,“人”、“父亲”、“他”(代词)和“像xxx这样的人”,并且在来自用户的查询包括“地址”的情况下,删除街道地址、城镇和村庄名称,以保持一种状态,在这种状态下,虽然无法完全识别个人,但可以读出查询的意图。此外,在包括用户视频和用户语音的查询的情况下,用户的面部图像和语音被平均面部图像和平均语音代替,其年龄和性别与用户相同,从而防止在保持用户简档的同时识别个人。

[0136] 此外,代理系统330可以具有用于替换包括如上所述的敏感信息的文字、图像和声音的数据库,或者能够使用这种外部数据库。

[0137] 图8以流程图的形式示出了当不同地调用外部代理服务340时执行的处理步骤,同时保护用户的敏感信息。

[0138] 当代理系统330通过电视代理310输入来自用户的查询时(步骤S801),代理系统330对查询中包括的视频和语音进行识别处理,并分析查询的意图(步骤S802)。

[0139] 接下来,代理系统330基于步骤S802中的识别结果和分析结果,检查在视频、语音或查询中是否包括用户的敏感信息(步骤S803)。

[0140] 随后,在已经通过电视代理310输入的信息中包括用户的敏感信息的情况下(步骤S803中的是),代理系统330用其他信息替换敏感信息,同时查询的意图保持在可读状态(步骤S804)。

[0141] 在步骤S804中,例如,在来自用户的查询包括疾病名称的情况下,疾病名称被替换为单词“疾病”,在来自用户的查询包括个人名称的情况下,个人名称被替换为不能识别个人的单词,例如,“人”、“父亲”、“他”(代词)和“像xxx这样的人”,并且在来自用户的查询包括“地址”的情况下,移除街道地址和城镇及村庄名称。此外,用户的面部图像和语音被替换

为平均面部图像和平均语音,其年龄和性别与用户的相同。

[0142] 简而言之,当代理系统330调用外部代理服务340A、340B和340C来响应来自用户的查询时,代理系统330可以说变成了网关,因此可以执行屏蔽处理,以防止敏感信息流出到外部代理服务340A、340B和340C。

[0143] 应当注意,尽管图8示出了当代理系统330调用外部代理服务340A、340B和340C时,代理系统330屏蔽敏感信息的示例,但是当代理系统330在本地侧收集信息时,电视代理310可以屏蔽用户的敏感信息。

[0144] 上面已经集中解释由代理系统330或外部代理服务340A、340B和340C响应用户查询的操作。代理系统330和外部代理服务340A、340B和340C也能够主动与用户交谈。然而,如果当用户正在观看电视节目时,从电视代理310或外部代理装置320与用户通话,则阻止用户观看电视节目。另外,在用户正在观看记录的节目的情况下,或者当用户正在使用电视接收机11之外的设备时(例如,当用户干扰音乐时,当用户正在打电话时,当用户正在玩游戏时),类似的情况也适用于此。

[0145] 代理系统330能够通过驻留在电视接收机11中的电视代理310来确定用户的观看状态,并且能够基于确定的结果来适当地控制与用户通话的时间。相反,外部代理服务340C仅可以通过输入到外部代理装置320的语音输入单元321中的语音来掌握用户的状态。即使在用户不发声的无声状态下,也不可能确定是否允许与用户交谈或者用户是否正在安静地观看电视节目。

[0146] 因此,代理系统330被配置为基于通过电视代理310收集的本地侧的信息来确定用户的状态,并根据用户的状态来控制外部代理的操作。

[0147] 具体地,当用户正在观看电视节目时,代理系统330禁止外部代理服务340A、340B和340C。相反,当用户结束观看电视节目时,代理系统330启用外部代理服务340A、340B和340C。

[0148] 另外,在外部代理服务340A、340B和340C可以向代理系统330进行查询的情况下,外部代理服务340A、340B和340C可以适于在执行独立动作之前向代理系统330进行关于用户的状态或者关于是否可以执行独立动作的查询。当用户正在观看电视节目时,代理系统330通过指示不能执行独立动作来响应于查询,并抑制来自电视代理310或外部代理装置320的用户调用。相反,如果用户结束观看电视节目,则代理系统330响应查询指示可以执行主动行动。应当注意,外部代理服务340A、340B和340C能够通过例如语音或API调用向代理系统330进行查询。

[0149] 图10示出了代理系统330响应于用户的状态来控制外部代理服务340A、340B和340C的状态。

[0150] 代理系统330通过电视代理310获得本地侧的情况,即,电视节目的观看历史、当前观看情况、电视接收机11本身的状态(开/关、音量、图像质量设置等)、记录历史、电视接收机11的遥控器的操作历史等(S1001),并由控制单元334确定用户是否正在观看电视节目。

[0151] 当用户正在观看电视节目时,控制单元334禁止外部代理服务340A、340B和340C。控制单元334以符合每个外部代理服务的调用规范的格式执行禁止处理。具体地,当控制单元334生成禁止外部代理服务340A的指令时,该指令被语音合成单元333转换成语音数据,然后传输到外部代理服务340A(S1002)。另外,控制单元334将禁止外部代理服务340B的指

令原样(换言之,没有语音合成该指令)传输到外部代理服务340B(S1003)。此外,当控制单元334生成禁止外部代理服务340C的指令时,该指令被语音合成单元333转换成语音数据,然后传输到电视代理310,并且随后电视代理310通过语音指示外部代理装置320禁止外部代理服务340C(S1004)。在代理系统330如上所述禁止外部代理服务340A、340B和340C的时间段期间,外部代理服务340A、340B和340C抑制独立用户调用。

[0152] 应当注意,除了代理系统330同时禁止所有外部代理服务340A、340B和340C的情况之外,还存在仅禁止一部分外部代理服务的情况。

[0153] 相反,当用户结束观看电视节目时,控制单元334启用外部代理服务340A、340B和340C。控制单元334以符合每个外部代理服务的调用规范的格式执行启用处理。具体地,当控制单元334生成启用外部代理服务340A的指令时,该指令被语音合成单元333转换成语音数据,然后传输到外部代理服务340A(S1002)。另外,控制单元334向外部代理服务340B直接传输启用外部代理服务340B的指令(换言之,没有语音合成该指令)(S1003)。此外,当控制单元334生成启用外部代理服务340C的指令时,该指令被语音合成单元333转换成语音数据,然后传输到电视代理310,并且随后电视代理310指示外部代理装置320通过语音启用外部代理服务340C(S1004)。当代理系统330如上所述启用已经处于禁止状态的外部代理服务340A、340B和340C时,允许外部代理服务340A、340B和340C重新开始独立用户调用。

[0154] 应当注意,除了代理系统330同时启用所有外部代理服务340A、340B和340C的情况之外,还存在仅启用一部分外部代理服务的情况。

[0155] 此外,在执行独立用户调用之前,外部代理服务340A、340B和340C中的每一个都能够向代理系统330查询用户的状态或者是否能够执行独立动作。

[0156] 在外部代理服务340A中,当处理单元343A生成对代理系统330的查询时,文本信息被语音合成单元344A转换成语音数据,然后传输到代理系统330(S1005)。当用户正在观看电视节目时,这意味着应该抑制外部代理服务340A的用户调用,当控制单元334生成禁止外部代理服务340A的指令时,该指令被语音合成单元333转换成语音数据,然后传输到外部代理服务340A(S1002)。同时,在用户已经结束观看电视节目的情况下,这意味着允许外部代理服务340A的用户调用,当控制单元334生成启用外部代理服务340A的指令时,该指令被语音合成单元333转换成语音数据,然后传输到外部代理服务340A(S1002)。

[0157] 另外,在外部代理服务340B中,处理单元343B生成对代理系统330的查询,并通过API调用向代理系统330进行查询(S1006)。当用户正在观看电视节目时,这意味着应该抑制外部代理服务340B的用户调用,控制单元334将禁止外部代理服务340B的指令原样传输到外部代理服务340B(即,没有语音合成该指令)(S1003)。同时,在用户已经结束观看电视节目的情况下,这意味着允许外部代理服务340B的用户调用,控制单元334将启用外部代理服务340B的指令原样传输到外部代理服务340B(即,没有语音合成该指令)(S1003)。

[0158] 另外,在外部代理服务340C中,当处理单元343C生成对代理系统330的查询时,文本信息被语音合成单元344C转换成语音数据,然后传输到代理系统330(S1007)。当用户正在观看电视节目时,这意味着应该抑制外部代理服务340C的用户调用,当控制单元334生成禁止外部代理服务340C的指令时,该指令被语音合成单元333转换成语音数据,然后传输到电视代理310,并且电视代理310通过语音指示外部代理装置320禁止外部代理服务340C(S1004)。同时,在用户已经结束观看电视节目的情况下,这意味着允许外部代理服务340C

的用户调用,当控制单元334生成启用外部代理服务340C的指令时,该指令被语音合成单元333转换成语音数据,然后传输到电视代理310,并且电视代理310指示外部代理装置320通过语音启用外部代理服务340C(S1004)。

[0159] 顺便提及,可以在任意时间进行从外部代理服务340A、340B和340C中的每一个到代理系统330的查询。然而,不用说,两个或更多外部代理服务可以适于同步地向代理系统330进行查询。

[0160] 工业适用性

[0161] 至此,已经参考具体实施例详细描述了本说明书中公开的技术。然而,显而易见的是,本领域技术人员可以在不脱离本说明书中公开的技术要点的情况下修正或替换前述实施例。

[0162] 已经针对实施例解释了本说明书,在该实施例中,本说明书中公开的技术被应用于驻留在电视接收机中的电视代理。然而,本说明书中公开的技术要点不限于此。本说明书中公开的技术也可以类似地应用于各种代理系统,每个代理系统使用驻留在诸如空调、记录器和洗衣机等各种CE设备中的代理,并且驻留在各种信息设备中,包括:IoT装置;用户携带的便携式信息终端,例如,智能电话和平板电脑;用户佩戴的可穿戴设备;交互式机器人;安装在汽车内部的汽车导航系统等。因此,可以根据需要使用外部代理服务。

[0163] 简而言之,本说明书中公开的技术已经以说明的形式进行了描述,并且不应被解释为限制本说明书的内容。为了确定本说明书中公开的技术要点,应该考虑权利要求。

[0164] 应当注意,也可以如下配置本说明书中公开的技术:

[0165] (1)一种信息处理装置,包括:

[0166] 通信单元,通过驻留在第一设备中的代理接收与用户交互相关的信息;以及

[0167] 控制单元,控制外部代理服务。

[0168] (1-1)根据前述(1)所述的信息处理装置,还包括

[0169] 语音识别单元,识别来自用户的语音输入;语义分析单元,对语音识别的结果进行语义分析;以及语音合成单元,对用户的响应进行语音合成,

[0170] 其中,从第一设备语音输出所述语音合成单元的语音合成结果。

[0171] (2)根据前述(1)所述的信息处理装置,其中,

[0172] 所述控制单元收集信息,所收集的信息包括用户的图像或声音、与用户对第一设备的操作相关的信息、以及配备有第一设备的传感器检测到的传感器信息中的至少一个。

[0173] (3)根据前述(1)或(2)中任一项所述的信息处理装置,其中:

[0174] 所述第一设备是电视接收机;以及

[0175] 所述控制单元收集信息,所收集的信息包括用户的电视节目的观看历史和记录历史以及用户的遥控操作历史。

[0176] (4)根据前述(2)或(3)中任一项所述的信息处理装置,其中,

[0177] 所述控制单元基于从第一设备收集的信息来管理用户的账户。

[0178] (5)根据前述(1)至(4)中任一项所述的信息处理装置,其中,

[0179] 所述控制单元控制外部代理服务的调用。

[0180] (6)根据前述(5)所述的信息处理装置,其中,

[0181] 所述控制单元在执行格式转换之后执行调用,以使调用符合每个外部代理服务的

调用规范。

- [0182] (7) 根据前述(6)所述的信息处理装置,其中,
- [0183] 所述控制单元语音合成第一外部代理服务的调用内容,并执行第一外部代理服务的调用。
- [0184] (8) 根据前述(6)所述的信息处理装置,其中,
- [0185] 所述控制单元语音合成第二外部代理服务的调用内容,并在来自第一设备的第二外部代理服务的控制下执行外部代理设备的语音调用。
- [0186] (9) 根据前述(5)至(8)中任一项所述的信息处理装置,还包括:
- [0187] 语音识别单元,识别来自外部代理服务的语音响应;以及
- [0188] 语义分析单元,对语音识别的结果进行语义分析。
- [0189] (10) 根据前述(5)至(8)中任一项所述的信息处理装置,其中,
- [0190] 通过API调用从外部代理服务接收响应。
- [0191] (11) 根据前述(5)所述的信息处理装置,其中,
- [0192] 所述控制单元通过从多个外部代理服务的响应中选择或合成,生成对用户对代理的查询的响应。
- [0193] (12) 根据前述(1)至(11)中任一项所述的信息处理装置,其中,
- [0194] 所述控制单元基于已经调用第一设备的用户的简档信息来控制外部代理服务的调用。
- [0195] (13) 根据前述(12)所述的信息处理装置,其中,
- [0196] 所述控制单元基于已经调用第一设备的用户的简档识别结果,执行具有简档切换功能的外部代理服务的调用。
- [0197] (14) 根据前述(1)至(13)中任一项所述的信息处理装置,其中,
- [0198] 所述控制单元基于已经调用第一设备的用户的敏感信息,控制外部代理服务的调用。
- [0199] (15) 根据前述(14)所述的信息处理装置,其中,
- [0200] 所述控制单元用其他信息替换在用户调用中包括的敏感信息,然后执行外部代理服务的调用。
- [0201] (16) 根据前述(1)至(15)中任一项所述的信息处理装置,其中,
- [0202] 所述控制单元基于用户的状态控制外部代理服务。
- [0203] (16-1) 根据前述(16)所述的信息处理装置,其中,
- [0204] 所述控制单元基于用户对第一设备的使用状态来控制外部代理服务。
- [0205] (16-2) 根据前述(16)所述的信息处理装置,其中,
- [0206] 所述控制单元基于用户的状态禁止和启用外部代理服务。
- [0207] (16-3) 根据前述(16)所述的信息处理装置,其中,
- [0208] 所述控制单元基于用户的状态响应关于从外部代理服务调用代理的查询。
- [0209] (17) 一种信息处理装置,包括:
- [0210] 通信单元,接收对用户的响应,所述响应通过处理与用户的交互相关的信息而生成;以及
- [0211] 输出单元,输出响应。

- [0212] (18) 一种信息处理方法,包括:
  - [0213] 通信步骤,用于通过驻留在第一设备中的代理接收与用户交互相关的信息;以及
  - [0214] 控制步骤,用于控制外部代理服务。
- [0215] (19) 一种信息处理方法,包括:
  - [0216] 通信步骤,用于接收对用户的响应,所述响应通过处理与用户的交互相关的信息而生成;以及
  - [0217] 输出步骤,用于输出响应。
- [0218] (20) 一种信息处理系统,包括:
  - [0219] 第一设备,与用户交互的代理驻留在所述第一设备中;
  - [0220] 通信单元,接收与通过代理与用户的交互相关的信息;以及
  - [0221] 控制单元,控制外部代理服务。
- [0222] (20-1) 根据前述(20)所述的信息处理系统,其中,
- [0223] 所述第一设备接收并向用户输出响应,所述响应通过处理与用户交互相关的信息而生成。
- [0224] 附图标记列表
- [0225] 100信息设备(电视接收机) 101处理单元
- [0226] 102调谐器 103通信单元 104显示单元 105语音输入单元
- [0227] 106语音输出单元 107图像捕捉单元 108传感器单元
- [0228] 109遥控接收单元 110记录单元 300代理云平台
- [0229] 310电视代理 320外部代理装置 321语音输入单元
- [0230] 322语音输出单元 330代理系统 331语音识别单元
- [0231] 332语义分析装置 333语音合成单元 334控制单元
- [0232] 335语音识别单元 336语义分析单元。

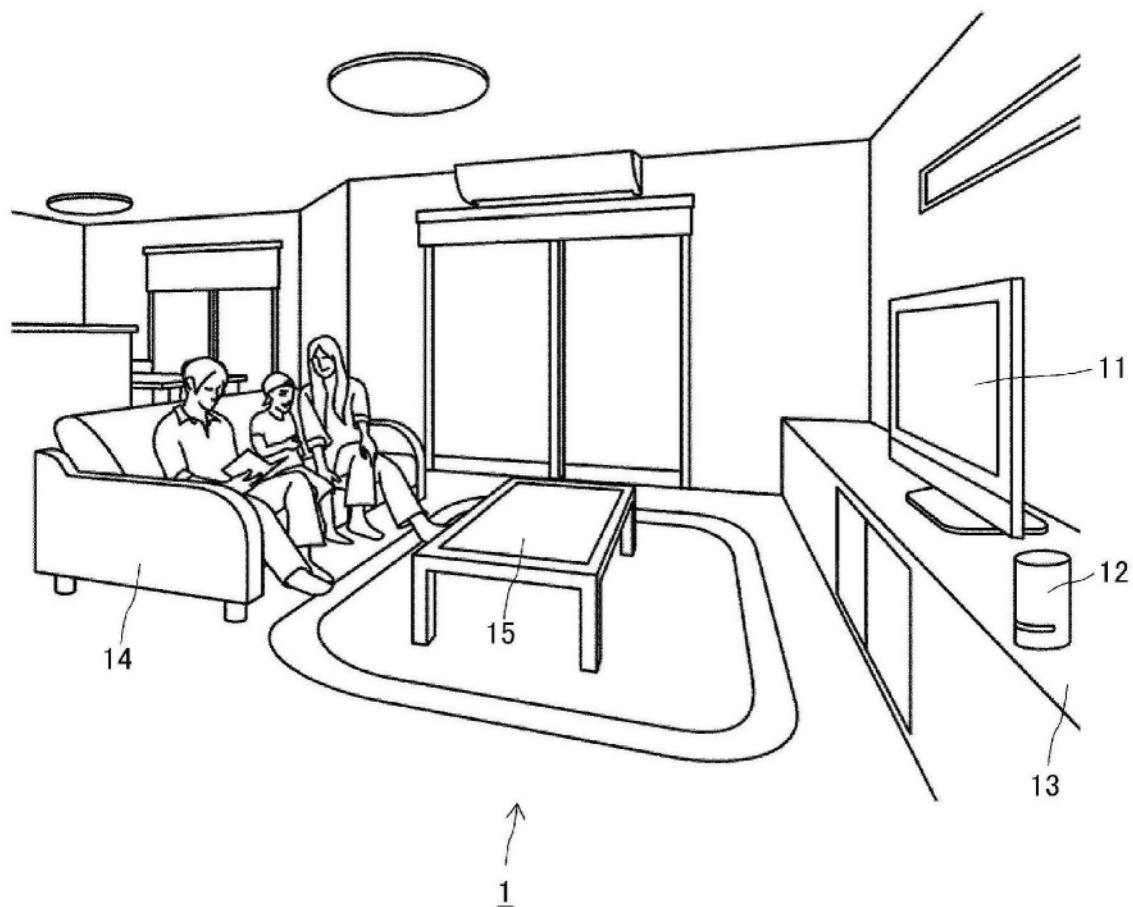


图1

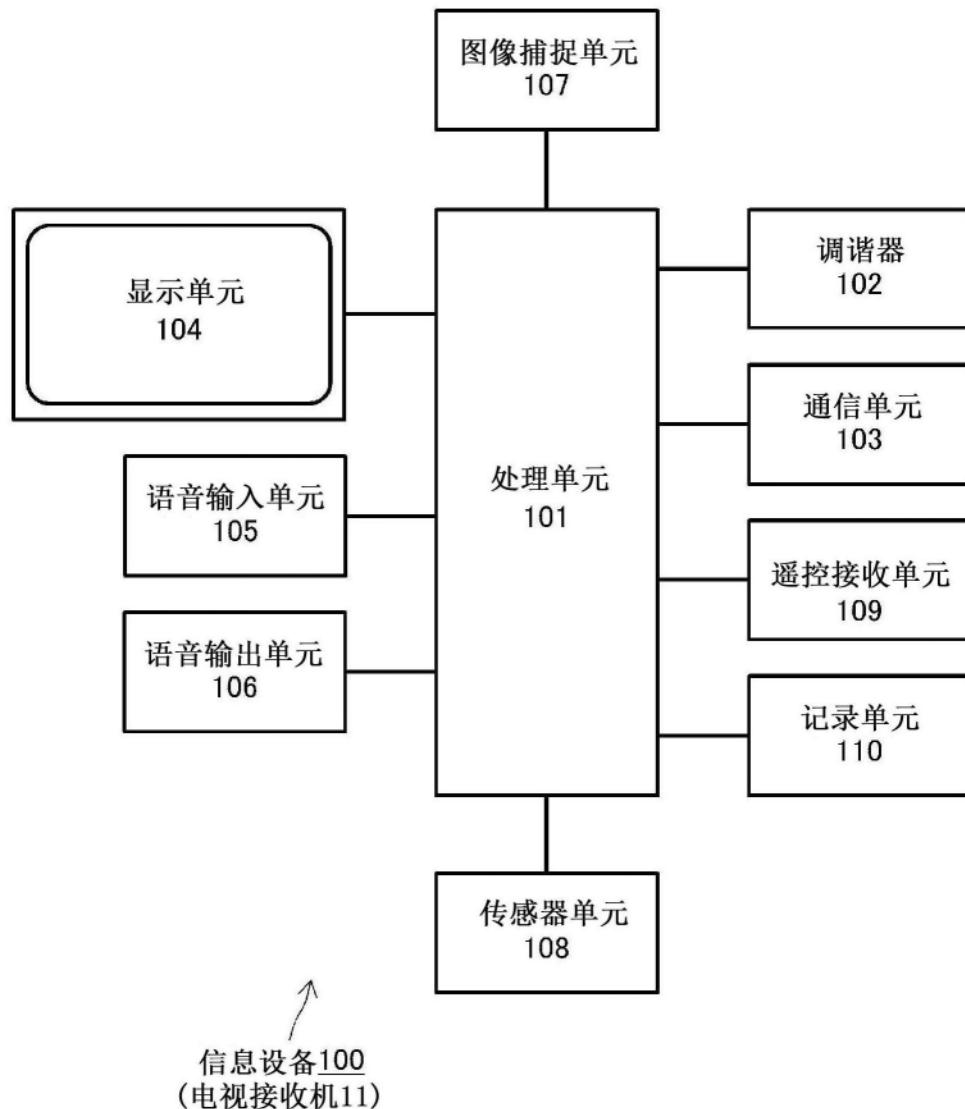
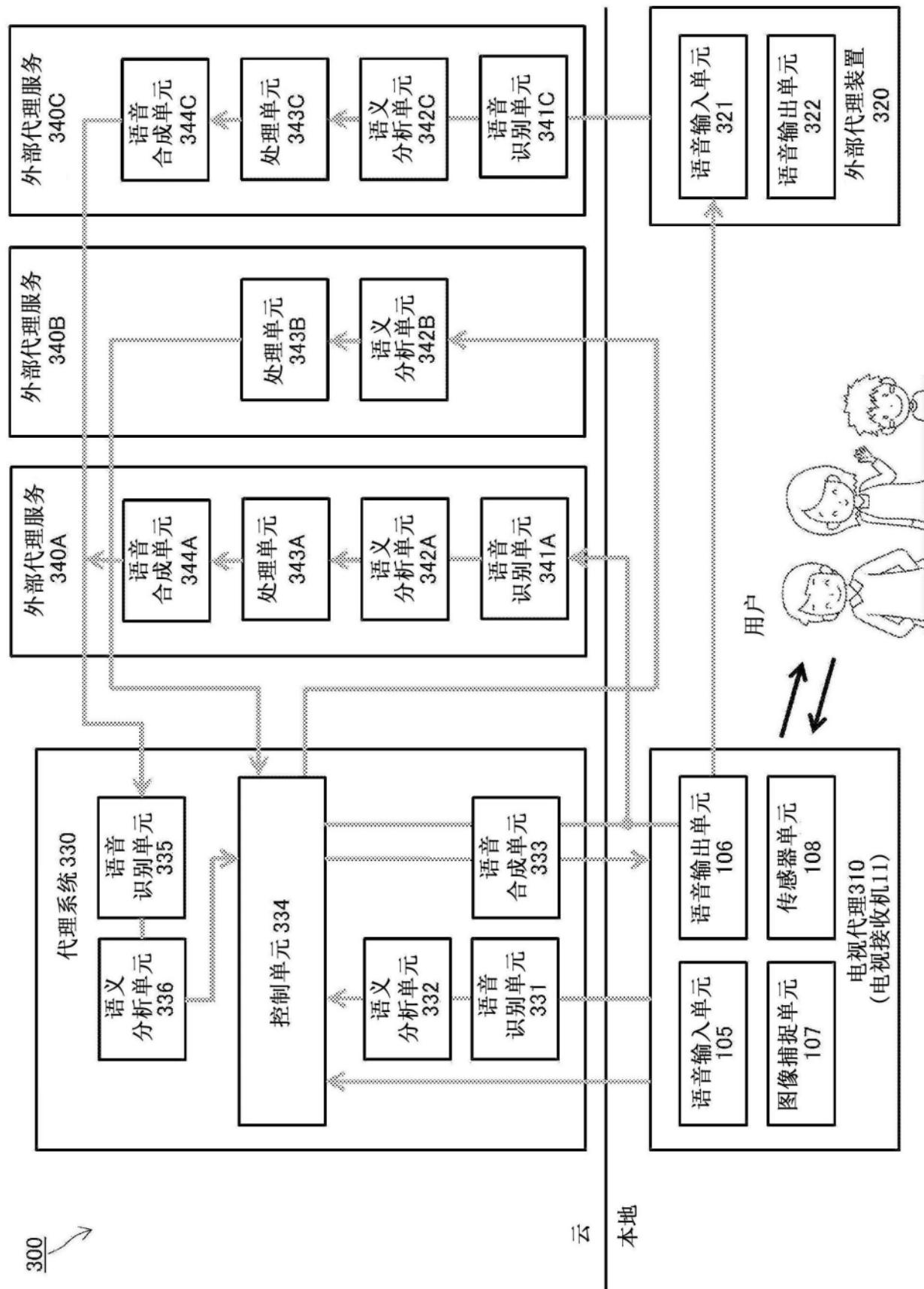


图2



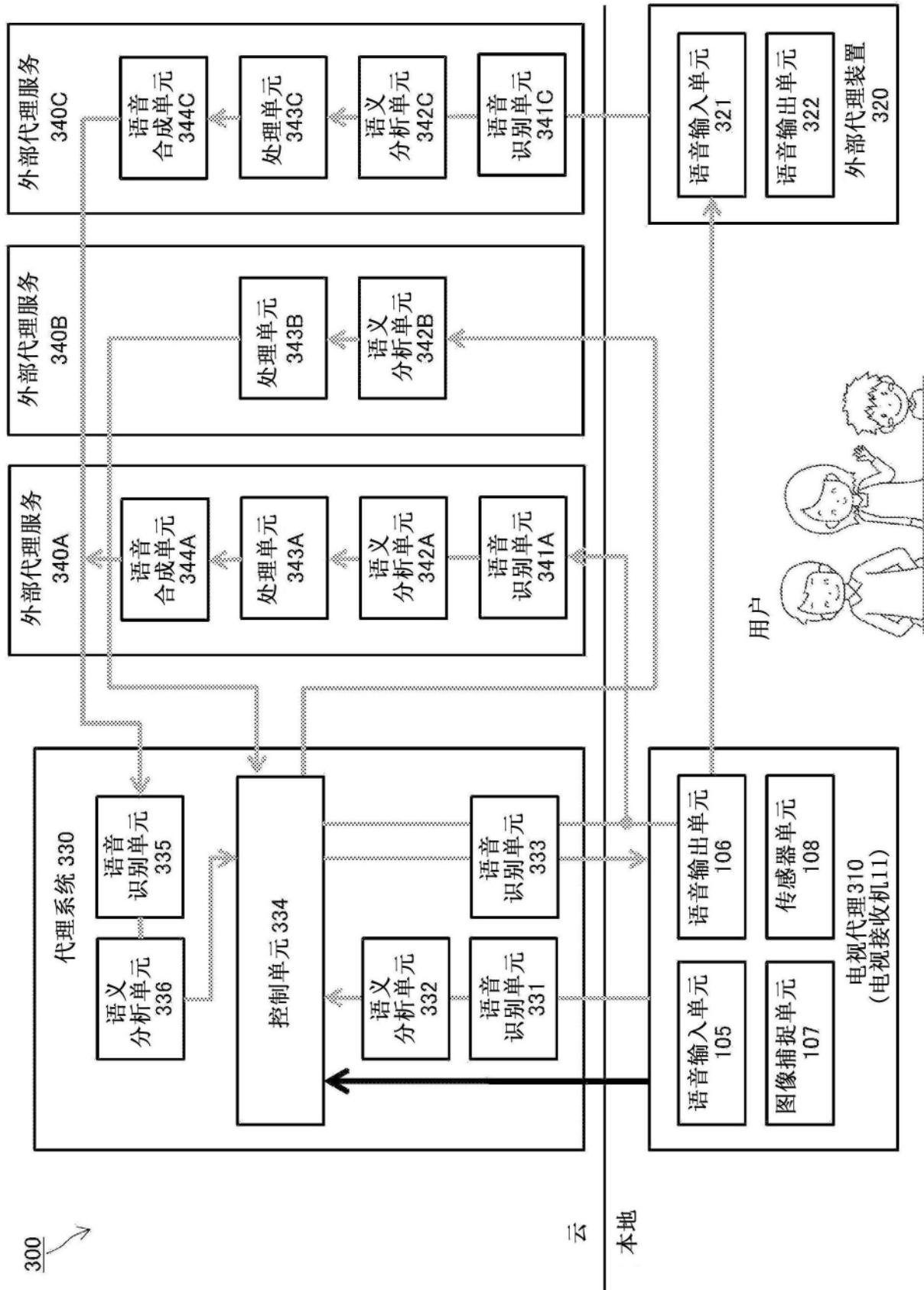


图4

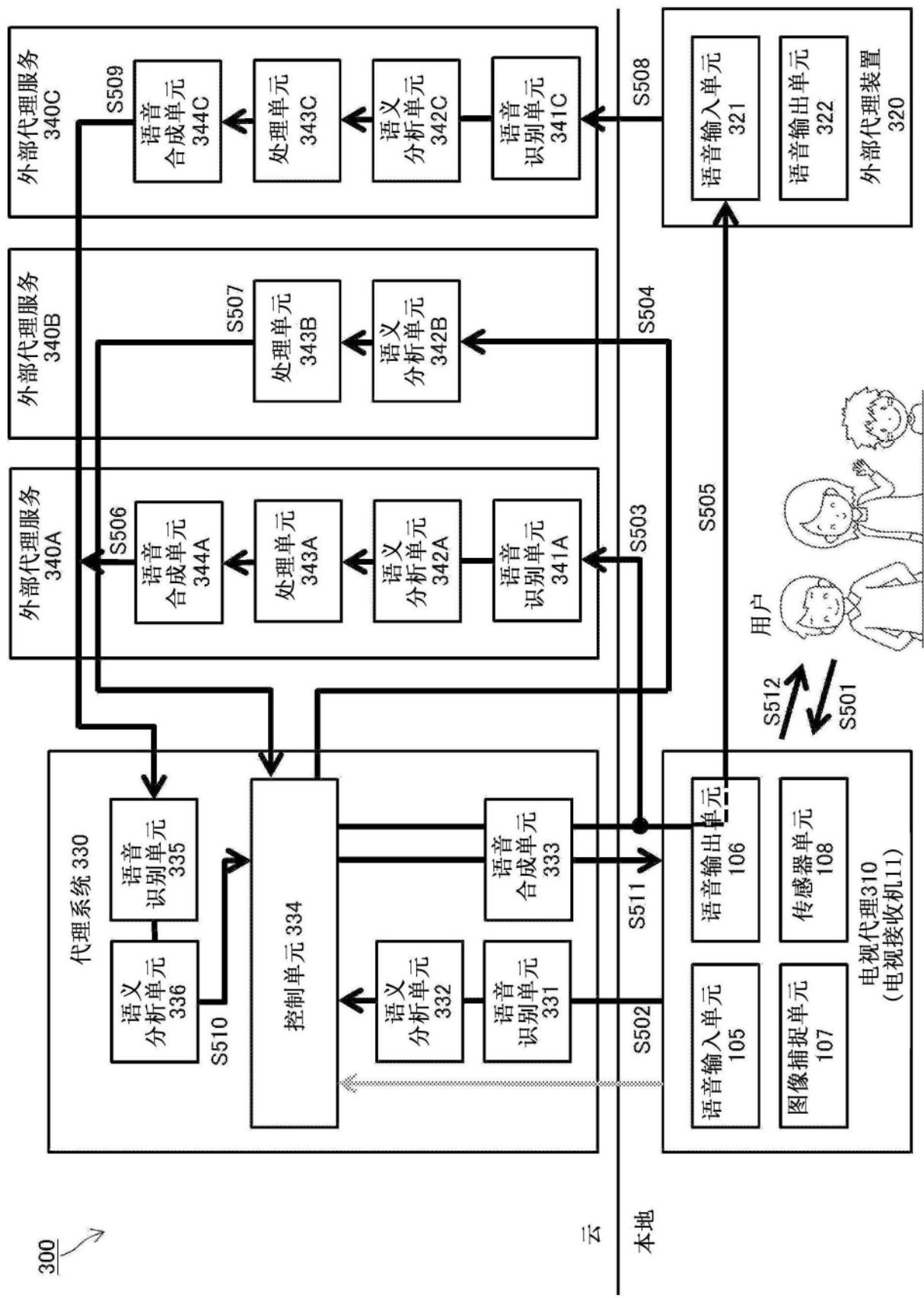


图5

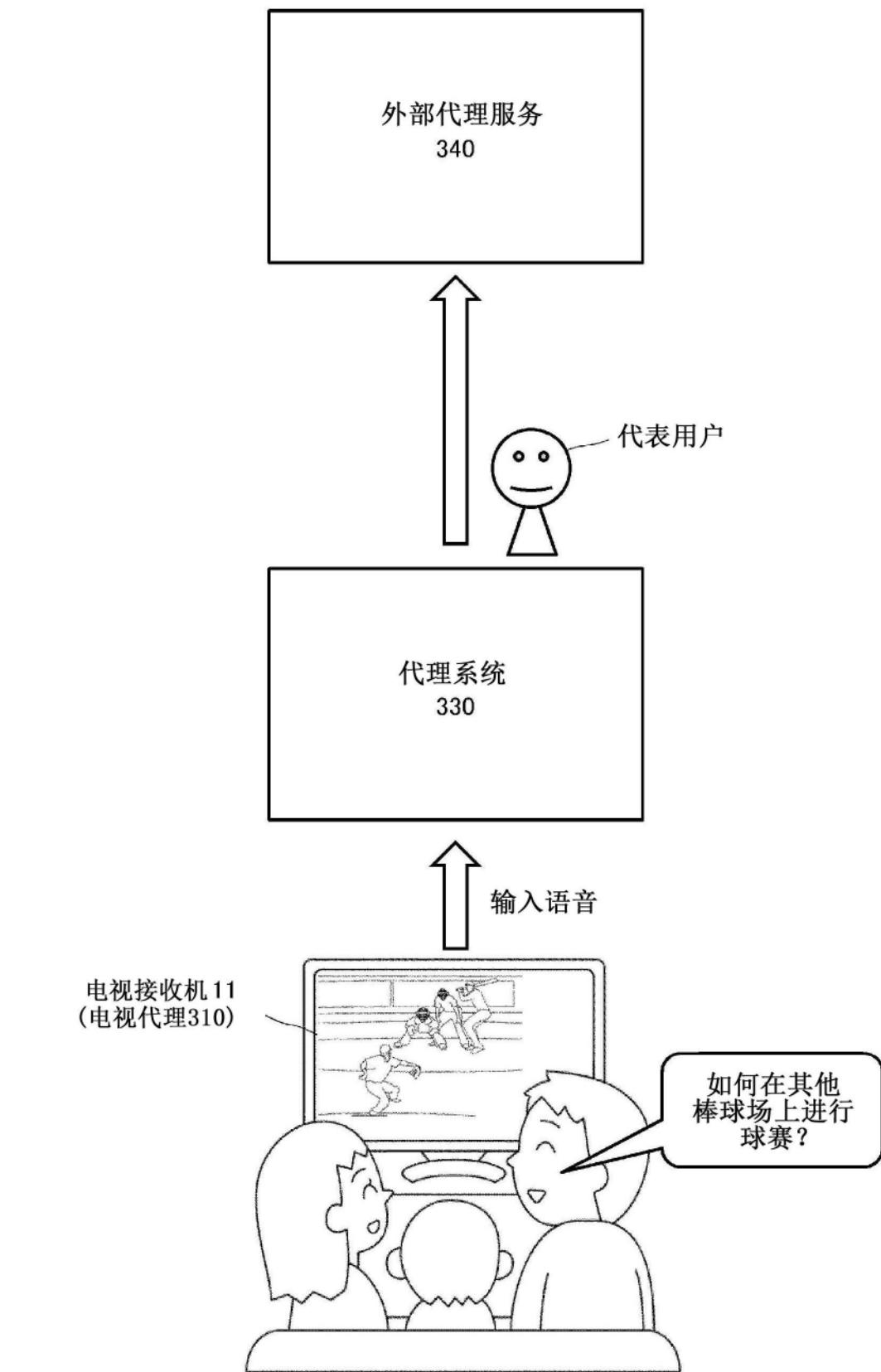


图6

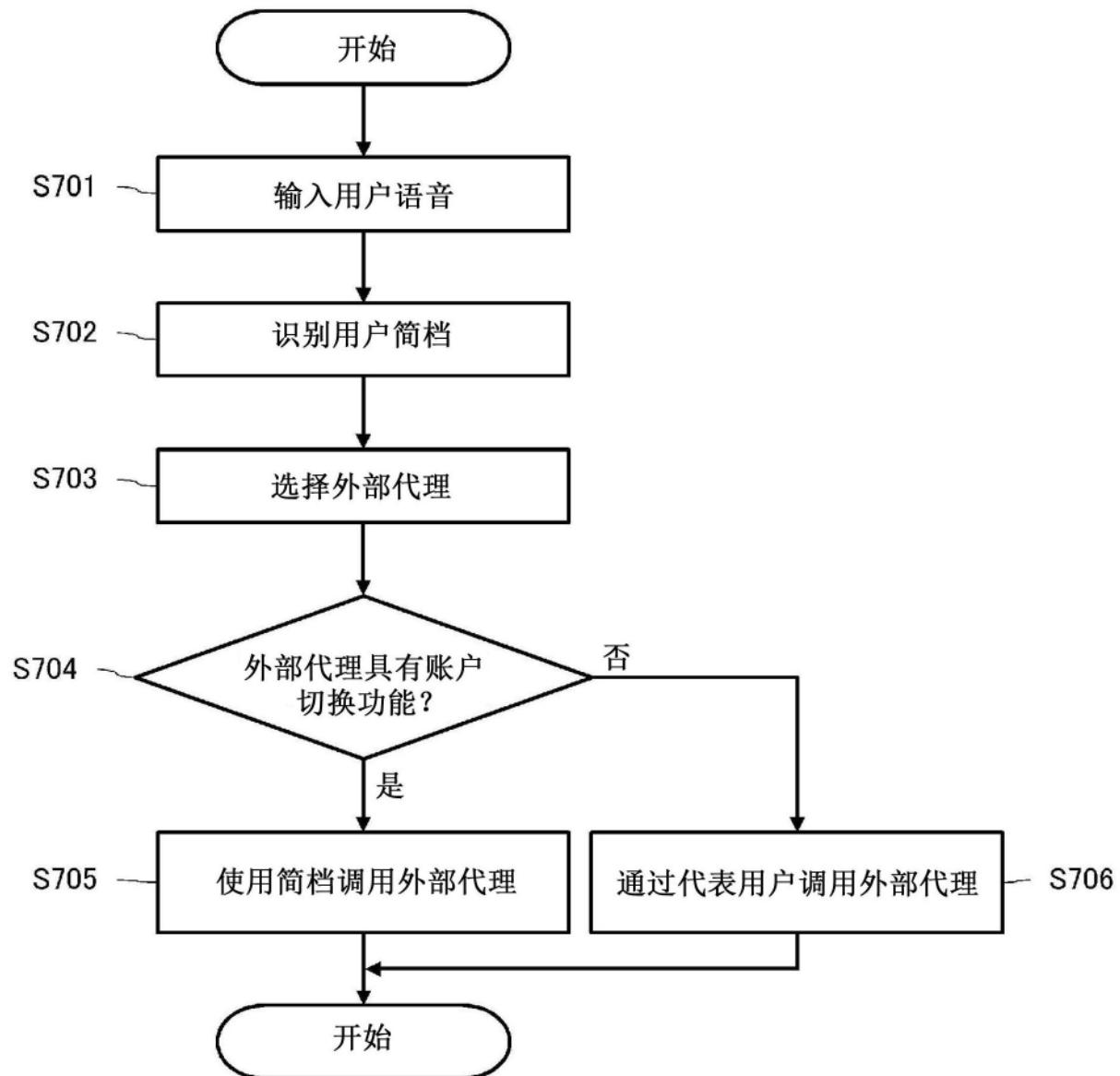


图7

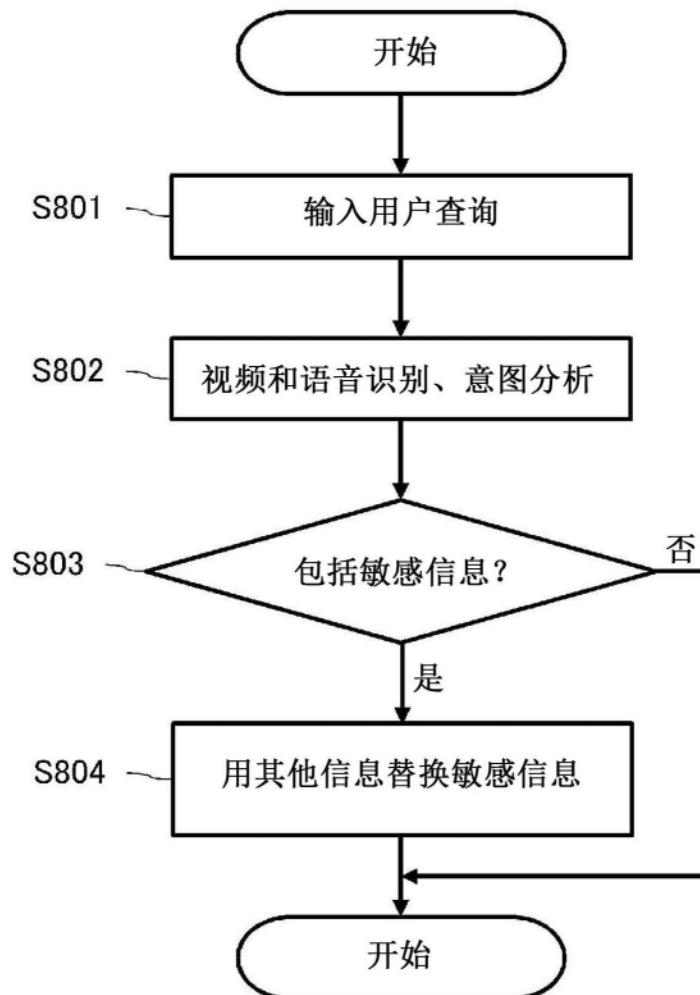


图8

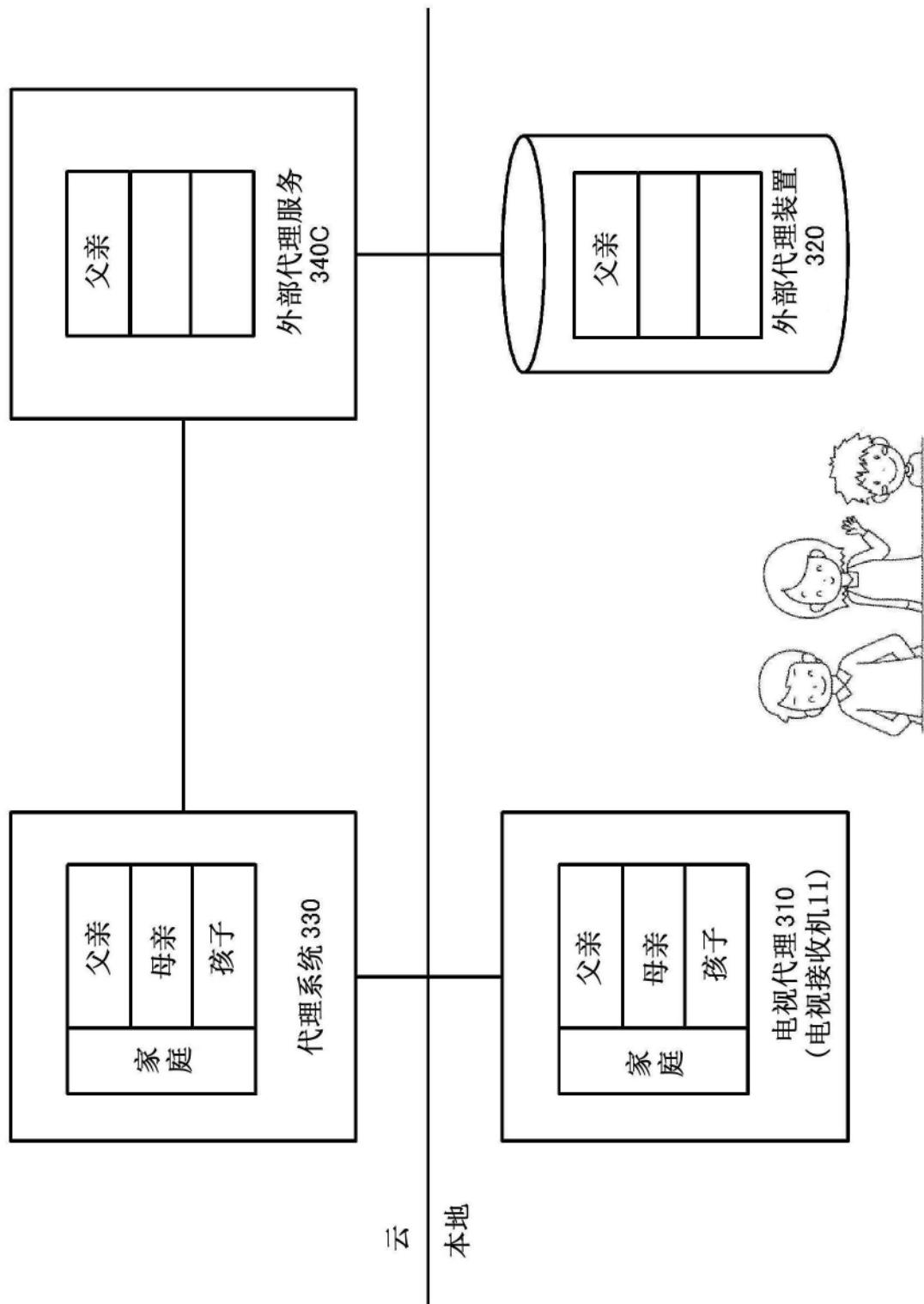
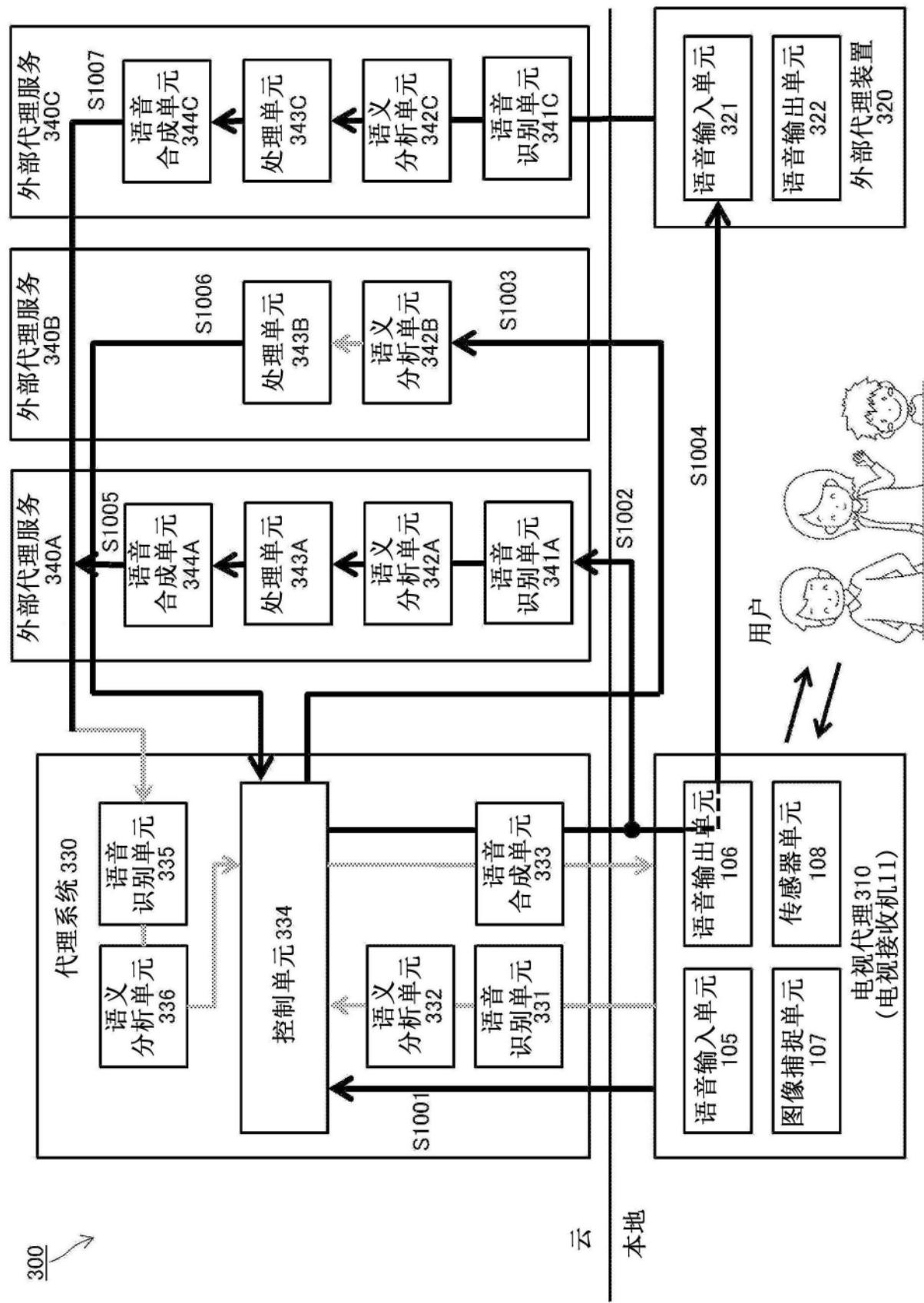


图9



300 ↗

图 10