



(12)发明专利申请

(10)申请公布号 CN 106465074 A

(43)申请公布日 2017.02.22

(21)申请号 201580032959.6

(51)Int.Cl.

(22)申请日 2015.06.17

H04W 4/12(2009.01)

(85)PCT国际申请进入国家阶段日
2016.12.19

(86)PCT国际申请的申请数据
PCT/US2015/036112 2015.06.17

(87)PCT国际申请的公布数据
W02015/195725 EN 2015.12.23

(71)申请人 微软技术许可有限责任公司
地址 美国华盛顿州

(72)发明人 K.伍尔齐 L.金 P.哈尔沃森
S.乔里 R.豪金斯

(74)专利代理机构 中国专利代理(香港)有限公
司 72001

代理人 王兴秋 陈岚

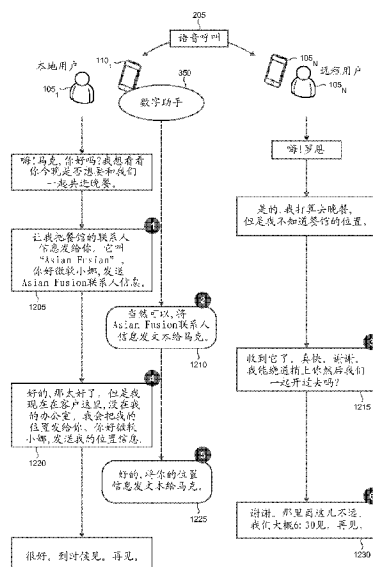
权利要求书1页 说明书17页 附图31页

(54)发明名称

数字助手在通信中的使用

(57)摘要

在设备(110)上操作的数字助手(350)被配置为通过监听语音和视频呼叫(205)并参与消息传送会话而作为主动参与者参加到在本地方(105-1)和远程方(105-N)之间的通信中。数字助手典型地可以通过使用关键词或短语(1205)的语音而被发起,并且然后使用语音或手势请求数字助手执行任务、提供信息和服务等。数字助手可以对请求做出响应并且采取适当的动作(1210)。在语音和视频呼叫中,与数字助手的交互(即,请求、响应和动作)可以被呼叫双方听到,就好像数字助手是呼叫中的第三方那样。在消息传送会话中,生成消息并且将其显示给每一个参与者,以使得他们可以看到与数字助手的交互,就好像它是参与者那样。



1. 一种用于在语音呼叫期间使用在设备上操作的数字助手的方法,该语音呼叫是在本地方和远程方之间通过通信网络承载的,该方法包括:

把数字助手监听配置为监听本地方和远程方之间的语音呼叫;

当本地方说出关键词或关键短语时发起数字助手;

当发起数字助手之后,对鉴于本地方在语音呼叫中所说的话而将要由数字助手执行的动作进行确定;

使得数字助手能够执行所确定的动作;以及

使数字助手返回到监听模式以便当本地方说出该关键词或关键短语时再发起数字助手。

2. 权利要求1所述的方法,还包括使得数字助手能够问问题,其中对问题的答复被使用于对将要执行的动作进行确定。

3. 权利要求1所述的方法,还包括配置数字助手以用于通过将音频注入到语音呼叫中而口头地向本地和远程方宣布所确定的动作。

4. 权利要求1所述的方法,还包括配置数字助手以用于视觉地向本地方宣布所确定的动作,视觉宣布包括在设备上显示的图形对象或文本对象的至少一个。

5. 权利要求4所述的方法,其中视觉宣布被配置成显示在与本地方相关联的设备或者与远程方相关联的设备中的一个上。

6. 权利要求4所述的方法,还包括使用电话应用来显示视觉宣布。

7. 权利要求4所述的方法,还包括使视觉宣布中的图形动画化。

8. 权利要求4所述的方法,还包括通过使图形对象或文本对象动画化而示出用于数字助手的状态。

9. 权利要求1所述的方法,还包括响应于本地用户的言语而配置数字助手以用于执行以下的至少一项:共享联系人信息、共享所存储的联系人、安排会议、查看用户的日程表、安排提醒、将另一方添加到当前呼叫、玩游戏、采购、记笔记、安排闹铃或唤醒提醒、将当前呼叫从语音呼叫改变成视频呼叫、将当前呼叫从视频呼叫改变成语音呼叫、重新连接掉线的呼叫、安排呼叫中断、与搜索服务交互、示出或共享电子邮件、共享或示出文件、发送到网站的链接或者发送到资源的链接。

10. 权利要求1所述的方法,还包括使得数字助手能够在执行动作之前从本地用户请求证实。

数字助手在通信中的使用

背景技术

[0001] 用户经常想要或需要进行多任务以使得可以执行各种任务而同时使用诸如智能电话或计算机之类的设备与其他人通信。然而,在通信中尝试多任务经常可能引起不连贯的、结结巴巴的或令人困惑的交互。一些常规的解决方案已经通过将附加控件(例如,按钮)添加到通信应用中来试图解决在尝试多任务时的不连贯通信的问题。不幸的是,这样的附加控件经常会使通信应用用户界面(GUI)对于用户而言是拥挤的、更难以操作的和令人困惑的。此外,这样的解决方案的结局是使得用户更容易离开通信应用且不做任何事情去使远程方保持专注于正发生的事情。其它常规解决方案使用某种形式的数字助手,它们在各种各样的计算平台上可获得,但是在与另一方通信期间以有用方式利用它们的能力非常有限。

[0002] 本背景被提供来介绍针对以下概要和详细描述的要旨上下文。本背景既不打算帮助确定所要求保护的题目的范围,也不打算被看作将所要求保护的题目限于用来解决以上提出的所有或任何缺点或问题的实现。

发明内容

[0003] 诸如智能电话、个人计算机或游戏操控台之类的设备上所支持的数字助手被配置成通过监听语音和视频呼叫,并参与消息传送会话,而作为主动参与者参加到在本地和远程方之间的通信中。数字助手典型地可以使用关键词或短语通过语音而被发起,然后使用语音命令、自然语言请求或者在一些情况下使用手势来请求数字助手执行任务、提供信息和服务等。数字助手可以对请求做出响应并且采取适当的动作。在语音和视频呼叫中,与数字助手的交互(即,请求、响应和动作)可以由呼叫的双方听到,就好像数字助手是呼叫中的第三方那样。在消息传送会话中,生成消息并且将其显示给每一个参与者,以使得他们可以看到与数字助手的交互,就好像其是参与者那样。

[0004] 在各种说明性示例中,数字助手使用来自各种源的输入去维持对呼叫状态和其它上下文的感知,使得它被使能去采取看起来自然的动作而同时减少用户的负担。因而,在约翰(用户)和汤姆、理查德以及哈里的四路会议呼叫上,用户可以说“为我们安排会议”,而数字助手理解:词语“我们”是指这四个参与者。数字助手可以确认该请求并且口头地陈述其正发送出会议请求。陈述被注入呼叫的音频中以使得所有会议呼叫参与者可以听到数字助手说话。在语音呼叫中,与数字助手的交互在由设备的电话应用所支持的UI上被视觉地证实。在视频呼叫中,数字助手交互被截取(capture)到画中画(PIP)覆盖中,画中画覆盖被本地显示且插入外出的视频流中以使得所有呼叫参与者可以看到它。

[0005] 由数字助手对上下文进行使用的其它示例包括:在执行给定动作或任务时考虑到时刻、用户的位置、用户的日程安排、其他方的属性等。因而,取决于用户是在国外度假胜地的假期中,还是用户是工作时段期间处于办公室中(正好在所安排的、审查针对所提出建筑的建筑规划的会议之前),数字助手可以采取不同的动作。例如,在度假情景中,数字助手可以在进行与亲戚的呼叫时通过在电子邮件或文本消息中将白天早前由用户拍摄的度假胜

地的照片发送给亲戚而响应于用户的请求“发送我拍摄的图片”。在办公室情景中,当用户在进行与助手的呼叫时,数字助手可以发送需要用来在会议时演示的针对所提出建筑物的当前工地的照片。

[0006] 在通信期间,可以由数字助手执行各种各样的任务并且采取动作。这些包括例如与用户交互(其可以涉及数字助手要求澄清问题和/或跟进(follow up with)用户)、执行各种任务、提供服务、搜集信息(例如,通过访问搜索服务)、操作设备、以及执行各种其它功能,以在不管是语音、视频还是消息传送的通信会话期间帮助用户。此外,在一些情况下,运行在设备上的应用可以与数字助手交互。例如,用户可能想要使用文本消息应用与数字助手交互(例如,使用通常称为“过顶(over-the-top)消息传送应用”的应用,而不是由移动网络运营商提供的第一方文本传送应用),其中数字助手可以表现得就像常规的消息传送参与者那样。这使能静默的交互以及提供便捷的方式来保持与数字助手的交互的记录。

[0007] 有利地,数字助手在通信中的当前使用使得能够以以下方式执行任务:使用流线型且简化的UI提供综合的特征集,而同时使得与数字助手的交互对于通信的其他方是可理解的。数字助手参与通信并且执行动作和任务,而不需要用户离开通信应用,以使得总体用户体验无缝地增强有附加功能性,而不会打断呼叫或通信或不得不致使呼叫或通信暂停。在真实世界场景中的成功的多任务也被使能,因为数字助手作为通信的受欢迎的第三方起作用,这样使得大幅地减少了对用户必须切换入和切换出各种应用和上下文的需要。

[0008] 提供本概要来以简化的形式介绍概念的选择,这些概念在下面的详细描述中被进一步描述。本概要既不打算标识所要求保护的的主题的关键特征或必要特征,也不打算用作确定所要求保护的的主题的范围时的帮助。此外,所要求保护的的主题不限于解决在本公开的任何部分中指出的任何或所有缺点的实现。将领会到,以上描述的主题可以被实现为计算机控制的装置、计算机进程、计算系统或者制品,诸如一个或多个计算机可读存储介质。这些和各种其它特征将从对以下详细描述的阅读以及对相关联的附图的审阅而显而易见。

附图说明

[0009] 图1示出了其中具有通信能力的设备通过网络交互的说明性环境;

图2示出了各种说明性通信类型,包括语音呼叫、消息传送通信和视频呼叫;

图3示出了包括数字助手组件的说明性分层架构;

图4和5示出了用户与数字助手之间的说明性界面;

图6示出了对数字助手的说明性输入以及可以由数字助手执行的一般功能的说明性分类;

图7、8和9示出了由数字助手显露的说明性用户界面(UI)的屏幕截图;

图10示出了其中由话筒捕获的音频被分开以使得其由数字助手和电话/视频呼叫应用接收的说明性布置;

图11示出了其中进入的音频被分开以使得其由数字助手和电话/视频呼叫应用接收的说明性布置;

图12示出了在数字助手与语音呼叫的各方之间的音频交互的说明性转录本(transcript);

图13-18示出了在语音呼叫通信期间在各种点处由设备上的电话应用示出的说明性UI

的屏幕截图；

图19和20示出了在消息传送会话期间在各种点处由消息传送应用示出的说明性UI的屏幕截图；

图21示出了在数字助手和视频呼叫的各方之间的音频交互的说明性转录本；

图22、23和24示出了在视频呼叫通信期间在各种点处由设备上的视频呼叫应用示出的说明性UI的屏幕截图；

图25示出了可以由数字助手在通信期间执行的功能的说明性分类；

图26示出了其中不同设备上的数字助手可以彼此交互的说明性布置；

图27示出了其中应用可以与数字助手交互的说明性布置；

图28和29示出了由消息传送应用在与数字助手交互时显露的各种UI的屏幕截图；

图30、31和32示出了在实现数字助手在通信中的当前使用时可以执行的说明性方法；

图33是说明性计算机系统的简化框图，诸如可以部分地被使用来实现数字助手在通信中的当前使用的个人计算机(PC)；

图34示出了可以部分地被使用来实现数字助手在通信中的当前使用的说明性设备的框图；

图35是说明性移动设备的框图；

图36是说明性多媒体操控台的框图。

[0010] 同样的参考标号在附图中指示同样的元件。除非以其它方式指示，否则元件不是按照比例绘制的。

具体实施方式

[0011] 图1示出了其中各种用户105采用通过通信网络115通信的相应设备110的说明性环境100。设备110提供语音电话能力并且典型地除各种其它特征之外还支持数据消费应用，诸如互联网浏览和多媒体(例如，音乐、视频等)消费。设备110可以包括例如用户设备、移动电话、蜂窝电话、功能型电话、平板计算机以及智能电话，用户经常利用它们来进行和接收语音和/或多媒体(即，视频)呼叫、参与消息传送(例如，文本传送)和电子邮件通信、使用采用数据的应用和访问服务、浏览万维网等。然而，还设想可在通信环境100内使用可替换类型的电子设备，只要它们配置有通信能力并且可以连接到通信网络115即可。这样的可替换的设备不同地包括手持式计算设备、PDA(个人数字助手)、便携式媒体播放器、使用头戴式受话器和耳机的设备(例如，蓝牙兼容设备)、手机平板(phablet)设备(即，组合智能电话/平板设备)、可穿戴式计算机、诸如GPS(全球定位系统)系统之类的导航设备、膝上型PC(个人计算机)、桌上型计算机、多媒体操控台、游戏系统等。在接下来的讨论中，术语“设备”的使用打算涵盖配置有通信能力并且能够连接到通信网络115的所有设备。

[0012] 环境100中的各种设备110可以支持不同特征、功能性和能力(此处一般地称为“特征”)。在给定设备上支持的一些特征可以类似于在其它设备上支持的那些特征，而其它特征可能是给定设备所独特的。在各种设备110上支持的特征之中的重叠程度和/或区别程度可以由于实现而变化。例如，一些设备110可以支持触摸控件、手势识别和语音命令，而其它设备可以使能更为有限的UI。一些设备可以支持视频消费和互联网浏览，而其它设备可以支持更为有限的媒体操控和网络接口特征。

[0013] 如所示出的,设备110可以访问通信网络115以便实现各种用户体验。通信网络可以包括以各种组合或子组合形式的各种各样网络类型和网络基础结构中的任何一种,包括蜂窝网络、卫星网络、诸如Wi-Fi和以太网网络之类的IP(互联网协议)网络、公共交换电话网络(PSTN)和/或诸如蓝牙网络之类的短程网络。网络基础结构可以例如由移动运营商、企业、互联网服务提供商(ISP)、电话服务提供商、数据服务提供商等支持。通信网络115典型地包括支持到互联网120的连接的接口,其典型地还被支持来使得移动设备110可以访问由一个或多个内容提供商125提供的内容并且在一些情况下访问服务提供商130。相应地,典型地使得通信网络115能够支持包括过顶通信的各种类型的设备到设备通信,以及不利用常规电话号码的通信,以便提供各方之间的连接性。

[0014] 设备110和通信网络115可以被配置为使能设备到设备的通信。如图2中所示,这样的设备到设备的通信200可以包括例如语音呼叫205、消息传送会话210和视频呼叫215。对于设备到设备的通信200的支持可以使用在设备110上运行的各种应用来提供。

[0015] 例如,图3示出了可以在支持这些应用的给定设备110上实例化的说明性分层架构300。架构300典型地被实现在软件中,然而在一些情况下也可以利用软件、固件和/或硬件的组合来实现。架构300被布置在层中,并且包括应用层305、OS(操作系统)层310和硬件层315。硬件层315向在其之上的层提供由设备100使用的各种硬件的抽象(例如输入和输出设备、联网和无线电硬件等)。在该说明性示例中,硬件层支持话筒320和音频端点325,其可以包括例如设备的内部扬声器、有线或无线头戴式受话器/听筒、外部扬声器/设备等。

[0016] 在该说明性示例中,应用层305支持各种应用(app)330(例如,web浏览器、地图app、电子邮件app等),以及电话app 335、消息传送app 340和视频呼叫app 345,诸如Skype™。App经常使用本地执行的代码实现。然而在一些情况下,这些app可以依赖由远程服务器或其它计算平台提供的服务和/或远程代码执行,其它计算平台是诸如由服务提供商130或其它基于云的资源支持的那些。尽管app 330、335、340和345在这里示出为被实例化在应用层305中的组件,但是将意识到,由给定app提供的功能性可以整体地或者部分地通过使用在OS或硬件层中被支持的组件来实现。

[0017] OS层310支持数字助手350和各种其它OS组件355。在一些情况下,数字助手350可以与服务提供商130交互,如由线360指示的。也就是说,数字助手在一些实现中可以部分地利用或者完全地利用在服务提供商130处支持的或者使用其它远程资源的远程代码执行。此外,其可以利用对于实现本文描述的各种特征和功能可能需要的其它OS组件355(和/或被实例化在架构300的其它层中的其它组件)和/或与其交互。尽管在本说明性示例中将数字助手350示出为被实例化在OS层310中,但是将意识到,由数字助手提供的功能性可以整体地或者部分地通过使用在应用或硬件层中被支持的组件来实现。

[0018] 如图4中所示,数字助手350可以采用自然语言用户界面(UI)405,其可以采取语音命令410作为来自用户105输入。语音命令410可以被使用来调用设备110上的各种动作、特征和功能、向系统和应用提供输入等。在一些情况下,语音命令410可以在特定用户体验的支持下独自地被利用,而在其它情况下,语音命令可以与其它非语音命令或输入组合地被利用,所述非语音命令或输入是诸如实现设备上的物理控件或者在UI上实现的虚拟控件的那些,或者使用手势的那些(如下文所述)。

[0019] 数字助手350还可以采用手势UI 505,如图5中所示。在这里,手势UI 505可以将由

用户105执行的手势510感测为输入,以调用设备110上的各种动作、特征和功能、向系统和应用提供输入等。用户手势510可以使用各种技术来感测,诸如光学感测、触摸感测、接近度感测等。在一些情况下,语音命令、手势以及真实或虚拟控件的物理操纵的各种组合可以被利用于与数字助手交互。在一些情景下,可以自动地调用数字助手。例如,因为数字助手典型地维持对呼叫状态和其它上下文的感知,所以可以通过诸如掉线的呼叫之类的特定上下文来调用数字助手。在该情况下,数字助手可以发起重拨以重新连接呼叫并且向用户告知该重新连接尝试。当数字助手检测到重新连接不可能(例如,没有蜂窝信号或者蜂窝信号较低)、线路繁忙时,或者在某数目的不成功的重新连接尝试之后,数字助手还可以向用户提供其它通信选项,诸如发送文本消息。

[0020] 图6示出了可以典型地由数字助手350支持的功能600的说明性分类。向数字助手350的输入典型地可以包括用户输入605(其中这样的用户输入可以包括从本地和远程方中的一方或两方到给定通信的输入)、来自内部源610的数据以及来自外部源615的数据。例如,来自内部源610的数据可以包括由设备110上的GPS(全球定位系统)组件或者某种其它位置感知组件报告的设备110的当前地理位置。来源于外部的数据615包括例如由外部系统、数据库、服务以及诸如服务提供商130等提供的数据(图1)。各种输入可以单独地或者以各种组合被使用,以使得数字助手350能够在其操作时利用上下文数据620。上下文数据可以包括例如时间/日期、用户的位置、语言、日程安排、安装在设备上的应用、用户的偏好、用户的行为(其中这样的行为在通知用户且用户准许的情况下被监控/追踪)、所存储的联系人(在一些情况下包括到本地用户或远程用户的社交图的链接,诸如由外部社交联网服务维持的那些)、呼叫历史、消息传送历史、浏览历史、设备类型、设备能力、通信网络类型和/或其中所提供的特征/功能性、移动数据计划约束/限制、与通信的其它方相关联的数据(例如,他们的日程安排、偏好等)等。在下文提供由数字助手对上下文的使用的附加说明性示例。

[0021] 如所示出的,功能600说明性地包括与用户交互625(例如通过自然语言UI和其它图形UI);执行任务630(例如,在用户的日程表中标记约会,发送消息和电子邮件等);提供服务635(例如,回答来自用户的问题,在地图上标出到目的地的方向等);搜集信息640(例如,寻找由用户请求的关于书籍或电影的信息,定位最近的意大利餐馆等);操作设备645(例如,设定偏好、调节屏幕明亮度、接通和关断诸如Wi-Fi和蓝牙之类的无线连接等);以及执行各种其他功能650。功能600的列表不打算穷举性的,并且可以由数字助手提供对于数字助手在通信中的当前使用的特定实现可能需要的其他功能。

[0022] 数字助手350可以通过显示在设备上的UI而显露给用户105。例如,图7、8和9示出了在数字助手在通信的当前使用中可以利用的UI的各种说明性屏幕截图。要强调的是,显示在附图中的特定UI可以根据特定实现的需要而从所显示的那样变化。图7中的UI 700示出了由图块(tile)705表示的数字助手(在该示例中命名为“微软小娜(Cortana)”),图块705与表示设备的开始屏幕上的其它app或用户体验的图块一起显示。数字助手还可以被配置为从设备上的任何UI内的任何位置启动,或者从任何当前用户体验内启动。例如,用户105可以处于电话呼叫中、在浏览web、观看视频或者听音乐,并且同时地从那些体验的任何一个内启动数字助手。在一些情况下,数字助手可以通过对物理或虚拟用户控件的操纵而启动,和/或在其它情况下通过语音命令和/或手势而启动。

[0023] 当用户例如通过触摸图块705或者通过调用语音命令或手势而调用数字助手时，在图8中示出的UI 800被显示在包括文本串805的设备110上，文本串805询问用户是否需要什么东西。在可替换实现中，可以采用文本到语音的翻译，使得可以播放音频消息以替代于文本串805或者作为文本串805的补充。如所示出的，UI包括被配置用于示出所接收的语音命令或其它用户输入的文本表示的框810。

[0024] 一个或多个图形对象815可以显示在UI 800上以向用户表示数字助手。图形对象815在本示例中是圆形形状，其可以被动画化(animate)以使得例如其在数字助手执行任务、提供信息、与用户交互等时改变它的形状、颜色、透明度、运动或外观。

[0025] 如图9的UI 900中所示，用户已经使用例如小键盘输入或语音命令而将串“打开邮件”905输入到框810中。响应于该输入，数字助手可以启动设备上的电子邮件应用。因而，例如当用户在呼叫中并且远程方询问“你是否已经阅读了我发送给你的电子邮件？”时，用户可以让数字助手打开来自远程方的电子邮件并且把它重读一遍。因为数字助手维持对呼叫状态和上下文的感知，所以其知道远程方的身份并且可以为用户定位期望的电子邮件。

[0026] 当该用户牵涉到与一个或多个远程方的语音或视频通信中时，数字助手可以被配置为通信的一部分并且按需要执行任务。如图10中所示，来自话筒320的音频在分离点1005被分开成两个流，使得电话及视频呼叫app 335和345以及数字助手350二者都可以从用户105接收音频信号。来自app的音频与由数字助手生成的音频被组合来创建组合的音频流1010，使得在通信的远端处的远程用户可以听到本地用户和数字助手二者在说什么。数字助手显露监听器1015，其监听来自用户的关键词或短语，该关键词或短语被使用来调用数字助手。同样地，如图11中所示，来自在远端处的远程方的进入的音频1110在分离点1105被分开成两个流，使得数字助手350和电话及视频呼叫app 335和345二者都可以接收进入的音频。要指出的是，术语“用户”和“方”可以在接下来的讨论中互换地使用。

[0027] 使数字助手350变成诸如电话呼叫、消息传送会话和音频呼叫之类的通信的一部分可以使能宽范围的丰富的用户体验。一个示范性体验在图12中示出，图12提供了在说明性语音呼叫205期间所谈的话的转录本。当各方在呼叫中交谈时，本地用户通过在块1205处在呼叫的点1处说出关键短语“你好微软小娜”而调用数字助手。要强调的是，这个特定的关键短语打算是说明性的，并且可以取决于数字助手在通信中的当前使用的特定实现的需要而利用其它关键词、短语或其它调用方法。

[0028] 例如，可以利用各种启发术、学习算法和/或策略，其考虑到可用的输入、上下文、用户的行为和与数字助手的交互。相应地，如果用户已经建立了例如在每周销售会议期间调用数字助手的历史，则可以为该会议的所有将来的实例自动地调用数字助手。或者，如果用户倾向于在驾车上下班的时候调用数字助手，则数字助手可以在下一次上下班行程期间自动地被调用(但是在用户驾车去购物商场时不必自动地调用)。在一些实现中，可以显露用户界面，这使得用户能够选择或调节数字助手将如何被自动调用。还可以使得用户能够在一些情况下推翻(override)自动调用，其中用户推翻的实例可以作为数字助手的已学习行为的部分而被包括。

[0029] 图13描绘了当用户在呼叫的点1处调用数字助手时由电话app在设备上示出的UI 1300的屏幕截图。要指出的是，在附图中示出的所有UI打算是说明性的，并且信息、所显露的特征和控件的呈现以及UI的总体外观和感觉可以通过实现而从所示出的那样变化。如图

13中所示,UI 1300示出了被叫方(即,远程方)的图片和姓名、所拨的号码和在UI底部的各种呼叫控件1305。与数字助手相关联的图形对象1310与文本串“正在听”(由参考标号1315指示)一起显示在UI上,以向用户图形化地示出数字助手成功启动、在听电话会谈并且能够致力于用户的任务、提供信息等。如上文所指出的,图形对象1310可以被动画化。

[0030] 在本地用户利用该示例中的关键短语而发起数字助手之后,用户请求数字助手向远程用户发送用于餐馆的联系人信息。在图12中的块1210,数字助手在呼叫的点2处通过说出将把联系人信息作为消息发送给远程用户来响应。在数字助手对用户请求的响应中的所生成的音频可以被本地方和远程方二者听到。数字助手还可以通过姓名查阅远程方。姓名的使用是数字助手可以如何应用其可获得的上下文数据的示例,这样使得其与各方的交互更为自然并且由数字助手支持的总体用户体验得以增强。也就是说,数字助手维持对呼叫上下文的感知,并且因而知道远程用户的身份以及其它呼叫参数。

[0031] 图14描绘了当数字助手在呼叫的点2处对用户的请求响应时由电话app在设备上示出的UI 1400的屏幕截图。文本串1415被显示在电话app的UI上,以便视觉地向本地用户证实数字助手在说的事情,即,餐馆联系人信息正被发送给在呼叫的远端处的远程用户。

[0032] 图15描绘了在图12中的块1215处在呼叫的点3处显示在远程用户的设备上的UI 1500的屏幕截图。此处,由数字助手发送的联系人信息作为新消息通知1505进入,其被显示在远程用户的设备上的UI的顶部。在本说明性示例中,通知示出发送者以及包含在消息中的内容的摘录。典型地,远程用户可以启动文本传送应用来查看消息的完整内容,其可以包括各种类型的联系人信息,诸如街道地址、到网站的链接、电话号码、地图等。

[0033] 返回到图12,随后在块1220处在呼叫的点4处,本地用户请求数字助手向远程用户发送用户的位置信息。用户通过使用关键短语(在本示例中是“你好微软小娜”)而发起数字助手。文本串“正在听”再次显示在电话app的UI 1600上,如由图16中的参考标号1615所指示的,以视觉地向本地用户证实数字助手正在呼叫中监听并且准备致力于任务、提供信息等。

[0034] 数字助手在图12中的块1225处在呼叫的点5处通过说出用户的位置信息将作为消息被发送给远程用户而响应。如之前那样,在数字助手对用户请求的响应中的所生成的音频可以被本地方和远程方二者都听到,并且数字助手可以通过姓名查阅远程用户。这提供了数字助手可以如何应用可获得的上下文数据来增强用户体验的另一个示例。此处,数字助手通过由设备的内部源供应的地理位置数据而维持对本地用户的位置的感知。

[0035] 图17描绘了当数字助手在呼叫的点5处对用户的请求响应时由电话app在设备上显示的UI 1700的屏幕截图。文本串1715被显示在电话app的UI上,以视觉地向本地用户证实数字助手在口头上说的事情,即,用户的位置信息正被发送给在语音呼叫的远端处的远程用户。

[0036] 图18描绘了在图12中的块1230处在呼叫的点6处显示于远程用户的设备上的UI 1800的屏幕截图。此处,由数字助手发送的本地用户的联系人信息作为新消息通知1805进入,新消息通知1805被显示在远程用户的设备上的UI的顶部。

[0037] 图19和20说明性地示出在本地方和远程方之间的消息传送会谈210的过程中可以利用数字助手。UI 1905和1910分别由本地和远程设备上的消息传送app显露。文本消息链被显示在每一个UI中,令外出的消息显示在右侧,而来自其他方的进入的消息显示在

左侧。尽管在这个特定说明性示例中使用了文本消息,但是要指出的是,数字助手在通信中的当前使用可以支持其他形式和类型的消息,包括多媒体消息、语音和视频消息。相应地,除非另外地声明,否则术语“消息传送”可以被视为是指所有这样的形式和类型的消息。此外,尽管目前的示例示出了消息链,但是要强调的是,数字助手也可以在每个消息的基础上与用户交互。

[0038] 在文本消息的交换期间的某一点处,本地用户通过说出关键短语“你好微软小娜”而启动数字助手,如由参考标号1915所指示的。本地用户然后口头地请求数字助手将联系人信息发送给远程用户。

[0039] 数字助手将其自身的文本消息注入本地设备上的消息传送app UI中,如由图19中的参考标号1920所指示的,以视觉地为本地用户示出其正如何对请求做出响应。此处,数字助手指示其正在文本消息1920中向远程用户发送联系人信息。在一些情况下,数字助手可以使用音频来响应请求。文本消息1920可以典型地以某种方式显著地标记,以使得本地用户清楚该消息是来自数字助手而不是来自远程用户。例如,消息可以标注有如所示的数字助手的名称,或者使用特殊文本、字体、颜色、图形等而高亮显示。在一些情况下也可以利用示出数字助手的存在的信息。例如,视情况而定,可以利用诸如“微软小娜已经进入会谈”或者“微软小娜已经离开会谈”之类的消息。

[0040] 在典型的实现中,来自数字助手的信息以全保真度持久存留在该用户与远程用户的通信历史中。这意味着,对于许多消息传送体验,诸如由丰富的过顶应用所支持的那些,通信的所有参与方均获得与数字助手交互的益处。因而,例如,远程消息传送应用用户在不必离开其当前上下文的情况下,可以看到对由数字助手提供的事件的指引。

[0041] 以与上文的电话呼叫示例相似的方式(其中使得远程用户在与本地设备处的本地用户交互时能够听到数字助手在说什么),此处远程用户被提供有进入的文本消息1925,其示出在本地用户和数字助手之间的交互。如图20中所示,数字助手在本地设备上的文本消息2005中以及在远程设备处的进入的文本消息2010中提供联系人信息。在本示例中,文本消息包括感兴趣的餐馆的名称和地址来作为可以被关注(follow)以得到更多信息的链接。在本示例中,还在文本消息中提供地图,其典型地可以被使用来启动相应设备上的地图app,地图app可以提供区域的地图、到餐馆的指引、感兴趣的附近点等。

[0042] 图21至24说明性地示出了可以在本地方和远程方之间的视频呼叫215的过程中如何利用数字助手。图21示出了视频呼叫的音频部分的转录本,其中本地用户在块2105处通过在视频呼叫的点1处说出“你好微软小娜”而启动数字助手。本地用户然后请求数字助手向远程用户发送餐馆联系人信息。图22示出了分别由本地和远程设备上的视频呼叫app在视频呼叫的点1处显露的UI 2205和2210。每一个UI示出了在视频呼叫的另一端处的参与方的大视频图像连同设备的用户的较小插入图像(也就是说,在典型的视频呼叫中,用户可以在视频呼叫app UI上的大视图中看到其他人,并且在小视图中看到他们自己)。

[0043] 以与上文讨论的语音呼叫示例类似的方式,PIP覆盖2215被插入在呈现于本地设备上的UI中的视频图像中。PIP覆盖在本示例中包括与数字助手相关联的图形对象和文本串。从本地设备外出的视频流还被配置成包括PIP覆盖2220,其可以显示在远程设备上的视频呼叫app UI上。相应地,在视频呼叫的上下文中本地用户和数字助手之间的交互通过使用视频PIP覆盖2220而指示给远程用户。

[0044] 图23示出了当数字助手在图21中的块2110处在视频呼叫的点2处响应用户的请求时,分别在本地和远程设备上的视频app UI 2305和2310。文本串2315和2320被显示在相应视频app的UI上的视频PIP覆盖中,以视觉地向本地和远程用户证实数字助手在视频呼叫中口头上说的事情,即,餐馆联系人信息正被发送给在视频呼叫的远端处的远程用户。

[0045] 图24示出了当数字助手在图21中的块2115处在视频呼叫的点3处响应用户的请求时,分别在本地和远程设备上的视频app UI 2405和2410。联系人信息通过视频流显示在相应视频app的UI上的窗口2415和2420中。在该示例中,窗口包括感兴趣的餐馆的名称和地址来作为可以被关注以得到更多信息的链接。在该示例中,还提供地图,其可以典型地被用来启动相应设备上的地图app。

[0046] 提供数字助手在通信中——在语音呼叫、消息传送和视频呼叫中——的使用的三个说明性示例,以强调通信中数字助手的当前使用的某些原理。然而,要强调的是,数字助手不限于如在以上示例中所示的任务的执行和信息的提供。图25提供了可以被提供的特征的各种说明性示例的分类2500,以及可以由数字助手在通信200中执行的功能。

[0047] 如所示,分类包括数字助手:共享当前位置(如由参考标号2505指示的);共享联系人信息2510;共享所存储的联系人2515;安排会议2520;查看用户的日程表2525;安排提醒2530(例如,呼叫中的远程用户可能要求本地用户在下班回家的路上取回牛奶;数字助手监听该呼叫并且可以安排提醒且然后在下班时提示用户);将另一方添加到当前的呼叫或群组消息传送线程2535;播放游戏2540,诸如益智问答(其中数字助手可以充当游戏主持人和/或提供答案);记笔记2550;安排闹钟或唤醒提醒2555;将当前呼叫从语音呼叫改变成视频呼叫2560;将当前呼叫从视频呼叫改变成语音呼叫2565;重新连接掉线的呼叫2570;安排呼叫中断2575(例如,让数字助手在特定时间向设备发出呼入以作为打断当前呼叫、会议或交互的方式——例如作为提供借口以结束当前呼叫、离开会议等的方式);与诸如Microsoft Bing™的搜索服务交互2580,以执行搜索并且显现涉及例如电影时间、餐馆、旅行、天气、交通、方向等的信息;示出或共享电子邮件和/或其它文件2585;发送到网站或其它资源的链接2590;采购2592(例如,从在线商店、网站等)或者提供其它特征和执行其它功能2595。

[0048] 在说明性示例中,数字助手可以被配置成维持对用户的安排、活动、行为和其它上下文的感知,以提供超出在呼叫内体验中所提供的那些服务的其它服务。例如,数字助手可以从用户的日程表和位置确定对于会议而言用户将迟到。用户可以偏向于不发送电子邮件(因为会议出席者可能不会查看他们的电子邮件和/或如果用户在驾车,则不可能靠边停车来发送电子邮件)。替代地,数字助手可以提出代表用户发出呼叫以便向其它会议出席者告知用户的迟到状态并且让他们知道用户在路上。在一些情况下,数字助手可以将UI显露给远程用户(即,会议出席者),这使得他们能够请求与用户的呼叫。在该情况下,数字助手可以对用户振铃并且在呼叫得到应答时连接各方。

[0049] 在另一个说明性示例中,当数字助手检测到用户对于会议迟到或者对于会议可能迟到时(例如,会议地点跨越校园,并且用户位于办公室而没有足够的时间到达那里),数字助手可以使用语音或视频设立会议桥并且利用适当的指令邀请会议参与者加入该桥。当安排会议开始时,数字助手可以代表用户向会议桥发出呼叫。

[0050] 图26示出了其中由本地和远程用户中的每一个使用的设备配置有相应的数字助

手350和2650的说明性示例。在该情况下,数字助手可以彼此交互,如由参考标号2605所指示的,以便在各方之间的通信200的过程中增强用户体验。例如,本地设备上的数字助手350可以从远程设备上的数字助手2650请求远程用户的日程表信息以便安排各方之间将来的会议。

[0051] 图27示出了其中设备110上的app 2725可以与数字助手350交互的说明性示例,所述交互是如由参考标号2705所指示的。例如,如在图28和29中的说明性消息传送app UI的屏幕截图中所示,用户105可以使用消息传送app发送文本消息来与数字助手交互,而不是如上文示例中那样使用语音命令。用户文本消息在UI中示于右边,而由数字助手生成的消息示于左边。文本传送使得用户和数字助手之间的交互能够相对静默地(并且异步地,这通常是优选的通信模式)发生,这例如在用户处于其中语音或手势的使用不可能或者不适当的位置处时可能是有利的。在一些情况下,静默操作可以在语音和视频通信的上下文中通过以下方式而被实现和/或进一步增强,即禁用扬声器电话操作以使得数字助手的语音仅在设备被拿到用户耳边时或者通过使用诸如头戴式受话器或耳机之类的音频端点设备才可以听到。在这样的情况下,数字助手可以私密地向用户提供信息,就像在用户耳边耳语那样。

[0052] 在另一个说明性示例中,数字助手可以被配置成在用户繁忙时应答进入的语音和视频呼叫。数字助手可以解释用户不能交谈(unavailable)并且提供UI,这使得远程用户能够使用音频或文本留下消息。UI可以显示对于该消息的文本,并且数字助手在一些情况下也可以向远程用户复述该消息。

[0053] 图28中的UI 2800示出了其中用户向数字助手写文本的一个说明性示例。此处,数字助手可以使用来自设备的电话呼叫状态的上下文,以在适当的时间触发给用户的提醒,如在块2805中示出的文本消息中所指示的。图29中的UI 2900示出了其中本地设备上的数字助手与远程设备上的数字助手交互以便安排各方之间的会议的另一个说明性示例,如在块2905中示出的文本消息中所指示的。

[0054] 图30示出了其中数字助手参与到在远程方和本地方之间的语音呼叫中的说明性方法3000的流程图。除非明确地陈述,否则在以下流程图中示出并且在随附文本中描述的方法或步骤不被限制于特定的次序或顺序。此外,方法或其步骤中的一些可以并发地发生或被执行,并且取决于给定的实现的要求而不必在这样的实现中执行所有的方法或步骤,并且一些方法或步骤可以可选地被利用。

[0055] 在步骤3005中,在由本地和远程方使用的设备之间建立语音呼叫。在步骤3010中,数字助手设立监听器以使得在呼叫期间本地用户可以通过说出关键词或短语而调用数字助手。典型地,如在步骤3015中所示,数字助手问候呼叫中的每一方。因为数字助手维持对呼叫上下文的感知,包括各方的身份,因而在一些情况下可以通过姓名而个性化问候。问候使每一个人知道数字助手是呼叫的一方并且准备执行任务以及提供服务。当用户说话时,数字助手监听,如在步骤3020中所示的。

[0056] 在步骤3025中,数字助手确定其可以响应于用户的言语而采取的动作。在典型的实现中,当进行确定时定位并且利用适用的上下文。也就是说,在一些情况下,取决于包括呼叫状态的上下文,数字助手可以采取不同的动作。此外,数字助手可以被配置为向用户问问题,例如以澄清该请求、或者执行与用户的一些后续交互,这些交互是在完成任务时可能

需要的。在步骤3030中,将音频注入到呼叫的流中以使得本地和远程用户可以听到数字助手确认用户的请求并且宣布其响应于该请求而正采取的动作(即,其是在共享联系人信息、记笔记,还是向呼叫添加某人等)。

[0057] 除数字助手口头地宣布动作之外,还在步骤3035中使用文本或图形把动作显示在本地设备上的电话app的UI上,并且在步骤3040中数字助手执行该动作。在一些实现中,数字助手可以在采取所宣布的动作之前要求用户证实该动作。在步骤3045中,数字助手返回到监听模式并且监听在呼叫中说话的用户,且方法3000可以酌情从步骤3020起重复。可替换地,数字助手可以离开语音呼叫并且对该效果做出宣布(例如,通过说“再见”),正如在步骤3050中所示的。

[0058] 图31示出了其中数字助手参与到在本地方和远程方之间的视频呼叫中的说明性方法3100的流程图。在步骤3105中,在由本地方和远程方使用的设备之间建立视频呼叫。在步骤3110中,数字助手设立监听器以使得在视频呼叫期间本地用户可以通过说出关键词或短语而调用数字助手。在步骤3115中,数字助手问候视频呼叫的每一方(其中问候可以通过姓名而个性化)以宣布它的存在。当用户说话时,用户助手监听,如在步骤3120中所示的。

[0059] 在步骤3125,数字助手确定其可以响应于用户的言语而采取的动作。在典型的实现中,当进行确定时定位并且利用适用的上下文,如同以上描述的语音呼叫的示例的那样。此外,数字助手可以被配置成向用户问问题,例如以澄清该请求、或者执行与用户的一些后续交互,这些交互是在完成任务时可能需要的。在步骤3130中,将音频注入到呼叫的流中以使得本地和远程用户可以听到数字助手确认用户的请求并且宣布其响应于该请求而正采取的动作。

[0060] 除数字助手口头地宣布动作之外,还在步骤3135中使用文本或图形把动作显示在本地设备上的电话app的UI上。示出本地用户和数字助手之间的交互的PIP覆盖在步骤3140中被注入到外出的视频流中,以使得视觉地向远程用户告知对于数字助手正发生什么。

[0061] 在步骤3145中,数字助手执行动作。在一些实现中,数字助手可以在采取所宣布的动作之前要求用户证实该动作。在步骤3150中,数字助手返回到监听模式并且监听在视频呼叫中说话的用户,并且方法3100可以酌情从步骤3120起重复。可替换地,数字助手可以离开视频呼叫并且对该效果做出宣布(例如,通过说出“再见”),如在步骤3155中所示的。

[0062] 图32示出了其中数字助手参与在本地方和远程方之间的消息传送会话的说明性方法3200的流程图。在步骤3205中,在由本地方和远程方使用的设备之间建立消息传送会话。在步骤3210中,数字助手设立监听器以使得在消息传送会话期间本地用户可以通过说出关键词或短语而调用数字助手。当用户说话时,数字助手监听,如在步骤3215中所示的。

[0063] 在步骤3220中,数字助手使用被发送给本地和远程用户二者的文本消息来宣布来自本地用户的请求,文本消息可以示出在消息传送app的UI上。在步骤3225中,数字助手确定其可以响应于用户的言语而采取的动作。在典型的实现中,当进行确定时定位并且利用适用的上下文,如同在以上描述的语音和视频呼叫的示例那样。在步骤3230中,数字助手确认用户的请求,并且使用被发送给本地和远程用户二者的文本消息宣布其正作为响应而采取的动作,文本消息可以示出在消息传送app的UI上。

[0064] 在步骤3235中,数字助手执行动作。在步骤3240中,数字助手返回到监听模式并且监听在视频呼叫中说话的用户,并且方法3200可以酌情从步骤3215起重复。

[0065] 图33是利用其可以实现数字助手在通信中的当前使用的说明性计算机系统3300的简化框图,所述计算机系统是诸如PC、客户端机器或服务器。计算机系统3300包括处理器3305、系统存储器3311和系统总线3314,系统总线3314将包括系统存储器3311的各种系统组件耦合到处理器3305。系统总线3314可以是若干类型的总线结构的任何一种,包括存储器总线或存储器控制器、外围总线或者使用任何各种总线架构的本地总线。系统存储器3311包括只读存储器(ROM)3317和随机存取存储器(RAM)3321。在ROM 3317中存储有基本输入/输出系统(BIOS)3325,其包含诸如在启动期间帮助在计算机系统3300内的元件之间转移信息的基本例程。计算机系统3300还可以包括:硬盘驱动器3328,用于从内部安放的硬盘(没有示出)读取并且向其写入;磁盘驱动器3330,用于从可拆卸磁盘3333(例如,软盘)读取或者向其写入;以及光盘驱动器3338,用于从可拆卸光盘3343读取或者向其写入,可拆卸光盘3343是诸如CD(压缩盘)、DVD(数字多功能盘)或其它光学介质。硬盘驱动器3328、磁盘驱动器3330和光盘驱动器3338分别通过硬盘驱动器接口3346、磁盘驱动器接口3349和光学驱动器接口3352连接到系统总线3314。驱动器及其相关联的计算机可读存储介质提供对用于计算机系统3300的计算机可读指令、数据结构、程序模块和其它数据的非易失性存储。尽管本说明性示例包括硬盘、可拆卸磁盘3333和可拆卸光盘3343,但是在数字助手在通信中的当前使用的一些应用中也可以使用其它类型的能存储可由计算机访问的数据的计算机可读存储介质,诸如盒式磁带、闪存存储卡、数字视频盘、数据盒式磁带、随机存取存储器(RAM)、只读存储器(ROM)等。此外,当在本文中使用时,术语计算机可读存储介质包括介质类型的一个或多个实例(例如,一个或多个磁盘、一个或多个CD等)。出于本说明书和权利要求的目的,短语“计算机可读存储介质”及其变形不包括波、信号和/或其它暂时性和/或非有形通信介质。

[0066] 许多程序模块可以被存储在硬盘、磁盘3333、光盘3343、ROM 3317或RAM 3321上,包括操作系统3355、一个或多个应用程序3357、其它程序模块3360和程序数据3363。用户可以通过输入设备将命令和信息输入到计算机系统3300中,输入设备是诸如键盘3366和指向设备3368,比如鼠标。其它输入设备(没有示出)可以包括话筒、操纵杆、游戏手柄、圆盘式卫星天线、扫描仪、追踪球、触摸板、触摸屏、触敏设备、语音命令模块或设备、用户运动或用户手势捕获设备等。这些和其它输入设备经常通过耦合到系统总线3314的串行端口接口3371而被连接到处理器3305,但是也可以通过其它接口连接,诸如并行端口、游戏端口或通用串行总线(USB)。监控器3373或其它类型的显示设备也经由诸如视频适配器3375这样的接口连接到系统总线3314。除监控器3373之外,个人计算机典型地还包括其它外围输出设备(没有示出),诸如扬声器和打印机。在图33中所示的说明性示例还包括主机适配器3378、小型计算机系统接口(SCSI)总线3383以及连接到SCSI总线3383的外部存储设备3376。

[0067] 计算机系统3300通过使用到一个或多个远程计算机(诸如远程计算机3388)的逻辑连接而可操作在联网环境中。远程计算机3388可以被选择为另一个个人计算机、服务器、路由器、网络PC、对等设备或者其它公共网络节点,并且典型地包括以上相对于计算机系统3300描述的元件中的许多或全部,然而仅在图33中示出单个代表性远程存储器/存储设备3390。在图33中描绘的逻辑连接包括局域网(LAN)3393和广域网(WAN)3395。这样的联网环境经常部署在例如办公室、企业范围的计算机网络、内联网和互联网中。

[0068] 当在LAN联网环境中使用时,计算机系统3300通过网络接口或适配器3396被连接

到局域网3393。当使用在WAN联网环境中时,计算机系统3300典型地包括宽带调制解调器3398、网络网关或者用于通过诸如互联网之类的广域网3395而建立通信的其它装置。可以在内部或外部的宽带调制解调器3398经由串行端口接口3371而连接到系统总线3314。在联网环境中,涉及计算机系统3300的程序模块或其一些部分可以被存储在远程存储器存储设备3390中。要指出,在图33中所示的网络连接是说明性的,并且取决于数字助手在通信中的当前使用的应用的特定要求,可以使用在计算机之间建立通信链接的其它手段。

[0069] 图34示出了用于能够执行本文描述的各种组件以便提供数字助手在通信中的当前使用的设备的说明性架构3400。因而,在图34中图示的架构3400示出了可以适配于服务器计算机、移动电话、PDA、智能电话、桌上型计算机、上网本计算机、平板计算机、GPS设备、游戏操控台和/或膝上型计算机的架构。架构3400可被用来执行本文提出的组件的任何方面。

[0070] 在图34中图示的架构3400包括CPU(中央处理单元)3402、系统存储器3404,包括RAM 3406和ROM 3408,以及将存储器3404耦合到CPU 3402的系统总线3410。在ROM 3408中存储有基本输入/输出系统,其包含诸如在启动期间帮助在架构3400内的元件之间转移信息的基本例程。架构3400还包括大容量存储设备3412以用于存储软件代码或其它计算机执行的代码,其被用来实现应用、文件系统和操作系统。

[0071] 大容量存储设备3412通过连接到总线3410的大容量存储控制器(没有示出)而被连接到CPU 3402。大容量存储设备3412及其相关联的计算机可读存储介质为架构3400提供非易失性存储。

[0072] 尽管本文中所包含的对计算机可读存储介质的描述是指大容量存储设备,诸如硬盘或CD-ROM驱动器,但是本领域技术人员应当意识到,计算机可读存储介质可以是能被架构3400访问的任何可获得的存储介质。

[0073] 作为示例而非限制,计算机可读存储介质可以包括易失性和非易失性、可拆卸和不可拆卸介质,其以任何方法或技术实现来用于存储诸如计算机可读指令、数据结构、程序模块或其它数据之类的信息。例如,计算机可读介质包括但不限于RAM、ROM、EPROM(可擦除可编程只读存储器)、EEPROM(电可擦除可编程只读存储器)、闪存存储器或其它固态存储器技术、CD-ROM、DVD、HD-DVD(高清晰度DVD)、蓝光或其它光学存储装置、盒式磁带、磁带、磁盘存储或其它磁性存储设备、或者可以被用来存储期望的信息并且可以被架构3400访问的任何其它介质。

[0074] 根据各种实施例,架构3400可以使用通过网络到远程计算机的逻辑连接而在联网环境中操作。架构3400可以通过连接到总线3410的网络接口单元3416而连接到网络。应当意识到,网络接口单元3416还可以被用来连接到其它类型的网络和远程计算机系统。架构3400还可以包括输入/输出控制器3418以用于接收和处理来自多个其它设备的输入,其它设备包括键盘、鼠标或电子触笔(没有在图34中示出)。类似地,输入/输出控制器3418可以向显示屏、打印机或者其它类型的输出设备(也没有在图34中示出)提供输出。

[0075] 应当意识到,本文描述的软件组件在被加载到CPU 3402上且被执行时,可以将CPU 3402和总体架构3400从通用计算系统变换成被定制来便利于本文提出的功能性的专用计算系统。CPU 3402可以用任何数目的晶体管或其它分立电路元件来构造,所述晶体管或其它分立电路单元可以单独地或者集体地呈现任何数目的状态。更明确地,CPU 3402可以响

应于包含在本文公开的软件模块内的可执行指令而作为有限状态机来操作。这些计算机可执行指令可以通过如下方式变换CPU 3402,即:指定CPU 3402如何在状态之间转移,从而变换构成CPU 3402的晶体管或者其它分立硬件元件。

[0076] 对本文提出的软件模块进行编码也可以变换本文提出的计算机可读存储介质的物理结构。在本描述的不同实现中,物理结构的特定变换可以取决于各种因素。这样的因素的示例可以包括但不限于:被使用来实现计算机可读存储介质的技术,计算机可读存储介质是被表征为主要还是次要存储装置等等。例如,如果计算机可读存储介质被实现为基于半导体的存储器,则本文公开的软件可以通过变换半导体存储器的物理状态而被编码在计算机可读存储介质上。例如,软件可以变换构成半导体存储器的晶体管、电容器或其它分立电路元件的状态。软件还可以变换这样的组件的物理状态以便在其上存储数据。

[0077] 作为另一示例,本文公开的计算机可读存储介质可以使用磁性或光学技术来实现。在这样的实现中,当在其中编码软件时,本文提出的软件可以变换磁性或光学介质的物理状态。这些变换可以包括更改给定磁性介质内的特定位置的磁性特性。这些变换还可以包括更改给定光学介质内的特定位置的物理特征或特性以改变那些位置的光学特性。在不脱离本描述的范围和精神的情况下,物理介质的其它变换是可能的,所提供的前面的示例仅是为了促进该讨论。

[0078] 鉴于上文所述,应当意识到,许多类型的物理变换发生在架构3400中以便存储和执行本文所提出的软件组件。还应当意识到,架构3400可以包括其它类型的计算设备,包括手持式计算机、嵌入式计算机系统、智能电话、PDA以及本领域技术人员所已知的其它类型的计算设备。还设想到,架构3400可以不包括在图34中示出的所有组件,可以包括未在图34中明确示出的其它组件,或者可以利用与图34中示出的架构完全不同的架构。

[0079] 图35是说明性设备110的功能框图,所述设备110诸如是移动电话或智能电话,其包括总体地在3502处示出的各种各样可选的硬件和软件组件。移动设备中的任何组件3502可以与任何其它组件通信,然而为了图示的方便而没有示出所有连接。移动设备可以是各种各样计算设备(例如,手机、智能电话、手持式计算机、PDA等)中的任何一种,并且可以允许与一个或多个移动通信网络3504——诸如蜂窝或卫星网络——的双向通信。

[0080] 所图示的设备110可以包括控制器或处理器3510(例如,信号处理器、微处理器、微控制器、ASIC(专用集成电路)或者其它控制和处理逻辑电路),以用于执行诸如信号编码、数据处理、输入/输出处理、功率控制和/或其它功能之类的任务。操作系统3512可以控制组件3502的分配和使用,包括电源状态、上锁(above-lock)状态和下锁(below-lock)状态,并且提供对于一个或多个应用程序3514的支持。应用程序可以包括常见的移动计算应用(例如,图像捕获应用、电子邮件应用、日程表、联系人管理器、web浏览器、消息传送应用)或者任何其它计算应用。

[0081] 所图示的设备110可以包括存储器3520。存储器3520可以包括不可拆卸存储器3522和/或可拆卸存储器3524。不可拆卸存储器3522可以包括RAM、ROM、闪速存储器、硬盘或其它公知的存储器存储技术。可拆卸存储器3524可以包括闪速存储器或订户身份模块(SIM)卡,其在GSM(全球移动通信系统)系统中是公知的,或者其它公知的存储器存储技术,诸如“智能卡”。存储器3520可被用于存储数据和/或代码以用于运行操作系统3512和应用程序3514。示例性数据可以包括将经由一个或多个有线或无线网络而发送到和/或接收

自一个或多个网络服务器或其它设备的网页、文本、图像、声音文件、视频数据或其它数据集。

[0082] 存储器3520还可以被布置作为或者包括一个或多个计算机可读存储介质,其以任何方法或技术被实现来用于存储诸如计算机可读指令、数据结构、程序模块或其它数据之类的信息。例如,计算机可读介质包括但不限于RAM、ROM、EPROM、EEPROM、闪存存储器或其它固态存储器技术、CD-ROM(压缩盘ROM)、DVD(数字多功能盘)、HD-DVD(高清晰度DVD)、蓝光或其它光学存储、盒式磁带、磁带、磁盘存储或其它磁性存储设备、或者可被用来存储期望的信息并可被设备110访问的任何其它介质。

[0083] 存储器3520可以被使用来存储订户标识符,诸如国际移动订户身份(IMSI),以及设备标识符,诸如国际移动设备标识符(IMEI)。这样的标识符可以被传送给网络服务器以识别用户和设备。设备110可以支持一个或多个输入设备3530;诸如:触摸屏3532;用于实现语音输入以用于语音识别、语音命令等的话筒3534;相机3536;物理键盘3538;追踪球3540;和/或接近度传感器3542;以及一个或多个输出设备3550,诸如扬声器3552和一个或多个显示器3554。使用手势识别的其它输入设备(没有示出)也可以在一些情况下被利用。其它可能的输出设备(没有示出)可以包括压电或触觉输出设备。一些设备可以服务于一个以上的输入/输出功能。例如,触摸屏3532和显示器3554可以被组合成单个输入/输出设备。

[0084] 无线调制解调器3560可以耦合到天线(没有示出)并且可以支持处理器3510和外部设备之间的双向通信,如本领域中很好理解的。调制解调器3560被一般性地示出,并且它可以包括用于与移动通信网络3504通信的蜂窝调制解调器和/或其它基于无线电的调制解调器(例如,蓝牙3564或Wi-Fi 3562)。无线调制解调器3560典型地被配置用于与一个或多个蜂窝网络(诸如GSM网络)通信,以用于在单个蜂窝网络内、在蜂窝网络之间、或者在设备与公共交换电话网络(PSTN)之间的数据和语音通信。

[0085] 设备还可以包括至少一个输入/输出端口3580、电源3582、诸如GPS接收器之类的卫星导航系统接收器3584、加速度计3586、陀螺仪(没有示出)和/或物理连接器3590,其可以是USB端口、IEEE 1394(火线)端口和/或RS-232端口。所图示的组件3502并非必需的或者全部包括的,因为可以删除任何组件并且可以添加其它组件。

[0086] 图36是多媒体操控台1104的说明性功能框图。多媒体操控台1104具有中央处理单元(CPU)3601,其具有1级高速缓存3602、2级高速缓存3604和闪存ROM(只读存储器)3606。1级高速缓存3602和2级高速缓存3604临时存储数据并且因而减少存储器访问周期的数目,由此改进处理速度和吞吐量。CPU 3601可以配置有多于一个核,以及因而有附加的1级和2级高速缓存3602和3604。闪存ROM 3606可以存储可执行代码,其在多媒体操控台1104通电时的自举过程的初始阶段期间加载。

[0087] 图形处理单元(GPU)3608和视频编码器/视频编解码器(编码器/译码器)3614形成用于高速和高分辨率图形处理的视频处理流水线。数据经由总线从GPU 3608被载送到视频编码器/视频编解码器3614。视频处理流水线将数据输出到A/V(音频/视频)端口3640以用于传送给电视或其它显示器。存储器控制器3610连接到GPU 3608以便便利处理器访问各种类型的存储器3612,诸如但不限于RAM。

[0088] 多媒体操控台1104包括I/O控制器3620、系统管理控制器3622、音频处理单元3623、网络接口控制器3624、第一USB(通用串行总线)主机控制器3626、第二USB控制器3628

和前面板I/O子配件3630,其优选地实现在模块3618上。USB控制器3626和3628充当用于外围控制器3642(1)和3642(2)、无线适配器3648和外部存储器设备3646(例如,闪存存储器、外部CD/DVD ROM驱动器、可拆卸介质等)的主机。网络接口控制器3624和/或无线适配器3648提供对网络(例如,互联网、家用网络等)的访问,且可以是多种多样的不同的有线或无线适配器组件的任何一种,包括以太网卡、调制解调器、蓝牙模块、线缆调制解调器等。

[0089] 系统存储器3643被提供来存储在自举过程期间加载的应用数据。提供了介质驱动器3644,并且可以包括DVD/CD驱动器、硬驱动器或其它可拆卸介质驱动器等。介质驱动器3644可以在多媒体操控台1104内部或外部。应用数据可以经由介质驱动器3644而被访问,以用于通过多媒体操控台1104来执行、回放等。介质驱动器3644经由总线连接到I/O控制器3620,所述总线诸如是串行ATA总线或其它高速连接(例如,IEEE 1394)。

[0090] 系统管理控制器3622提供涉及确保多媒体操控台1104的可用性的各种各样的服务功能。音频处理单元3623和音频编解码器3632形成具有高保真度和立体处理的对应音频处理流水线。音频数据经由通信链接被载送于音频处理单元3623和音频编解码器3632之间。音频处理流水线向A/V端口3640输出数据以用于由具有音频能力的外部音频播放器或设备再现。

[0091] 前面板I/O子配件3630支持电源按钮3650和弹出按钮3652、以及任何LED(发光二极管)或显露在多媒体操控台1104的外表面上的其它指示器的功能性。系统电源模块3636向多媒体操控台1104的组件提供电力。风扇3638冷却多媒体操控台1104内的电路。

[0092] 多媒体操控台1104内的CPU 3601、GPU 3608、存储器控制器3610以及各种其它组件经由一个或多个总线互连,所述总线包括串行和并行总线、存储器总线、外围总线、以及使用各种各样总线架构的任一种的处理器或本地总线。作为示例,这样的架构可以包括外围组件互连(PCI)总线、PCI-Express总线等。

[0093] 当多媒体操控台1104通电时,应用数据可以从系统存储器3634加载到存储器3612和/或高速缓存3602和3604中并且在CPU 3601上执行。应用可以呈现图形用户界面,其在导航到在多媒体操控台1104上可用的不同媒体类型时提供一致的用户体验。在操作中,包含在介质驱动器3644内的应用和/或其它介质可以从介质驱动器3644启动或播放以向多媒体操控台1104提供附加的功能性。

[0094] 多媒体操控台1104可以通过简单地将系统连接到电视或其它显示器而作为独立系统进行操作。在这种独立模式中,多媒体操控台1104允许一个或多个用户与系统交互、观看电影或听音乐。然而,在集成有通过网络接口控制器3624或无线适配器3648而可用的宽带连接性的情况下,多媒体操控台1104还可以作为更大网络社区中的参与者进行操作。

[0095] 当多媒体操控台1104通电时,设定量的硬件资源被预留用于由多媒体操控台操作系统进行的系统使用。这些资源可以包括存储器(例如,16MB)、CPU和GPU周期(例如,5%)、联网带宽(例如,8kps)等的预留。因为这些资源在系统自举时间预留,所以预留的资源从应用的视角来看不存在。

[0096] 特别地,存储器预留优选地大到足以包含启动内核、并发的系统应用和驱动器。CPU预留优选地是恒定的,使得如果预留的CPU使用没有被系统应用所用,则空闲线程将消耗任何未用的周期。

[0097] 关于GPU预留,由系统应用(例如,弹出窗口)生成的轻量消息被通过使用GPU中断

来安排代码以将弹出窗口渲染到覆盖中而显示。覆盖所需的存储器的量取决于覆盖区域大小,并且覆盖优选地按屏幕分辨率缩放。在由并发的系统应用使用全部用户界面的情况下,优选的是使用与应用分辨率无关的分辨率。可以使用缩放器来设定该分辨率,以使得消除对改变频率并使TV重新同步的需要。

[0098] 在多媒体控制台1104自举并且系统资源被预留之后,并发的系统应用执行以提供系统功能性。系统功能性被封装于在以上描述的预留系统资源内执行的一组系统应用中。操作系统内核相对于游戏应用线程标识是系统应用线程的线程。系统应用优选地被安排成以预定的时间和间隔在CPU 3601上运行以便向应用提供一致的系统资源视图。所述安排是最小化针对在操控台上运行的游戏应用的高速缓存中断。

[0099] 当并发的系统应用要求音频时,由于时间敏感性而与游戏应用异步地安排音频处理。多媒体控制台应用管理器(下文描述)在系统应用活动时控制游戏应用音频等级(例如,静音、衰减)。

[0100] 输入设备(例如,控制器3642(1)和3642(2))由游戏应用和系统应用共享。输入设备不是预留的资源,而是要在系统应用和游戏应用之间切换,使得每一个应用都将得到该设备的关注。应用管理器优选地控制输入流的切换,而无需知道游戏应用的知识,并且驱动器维持关于关注切换的状态信息。

[0101] 基于前文,应当意识到,已经在本文中公开了用于数字助手在通信中的使用的技术。尽管本文所提出的主题是以特定于计算机结构特征、方法和变形的动作、具体的计算机器和计算机可读存储介质的语言进行描述的,但是要理解,在所附权利要求中限定的本发明不一定限于在本文中描述的特定特征、动作或介质。而是,该特定特征、动作和介质是作为实现本权利要求的示例性形式而公开的。

[0102] 以上描述的主题仅作为说明而提供,其不应当被解释为限制性的。可以在不遵循所图示和描述的示例性实施例和应用、且不脱离以下权利要求中阐述的本发明的真实精神和范围的情况下,对本文描述的主题做出各种修改和改变。

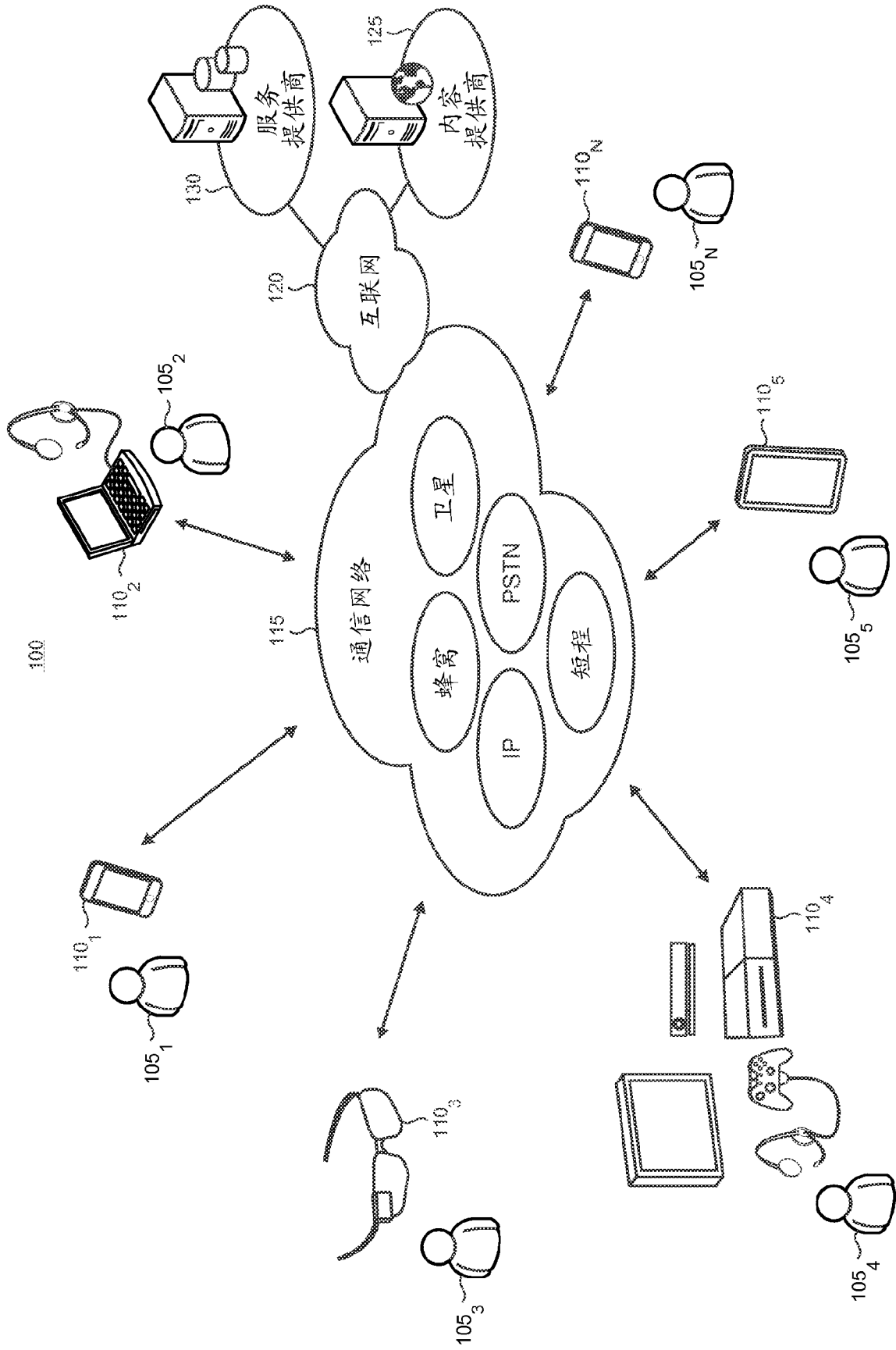


图 1

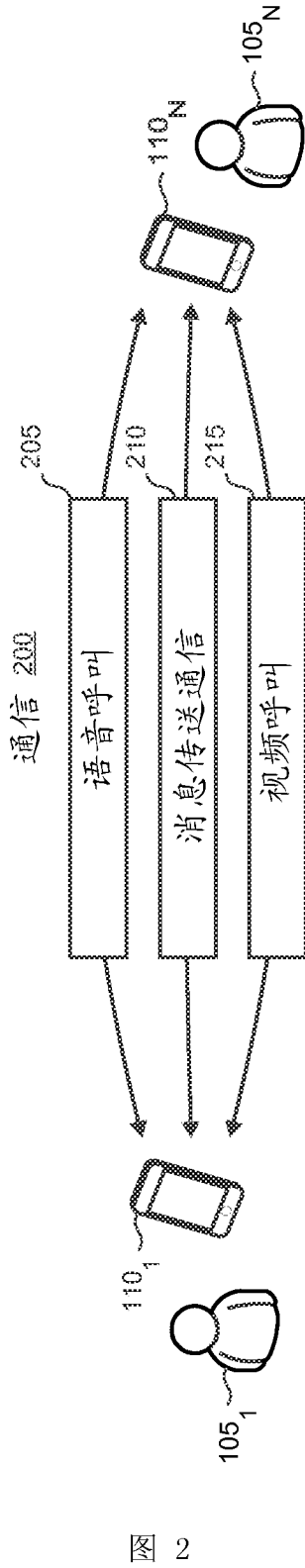


图 2

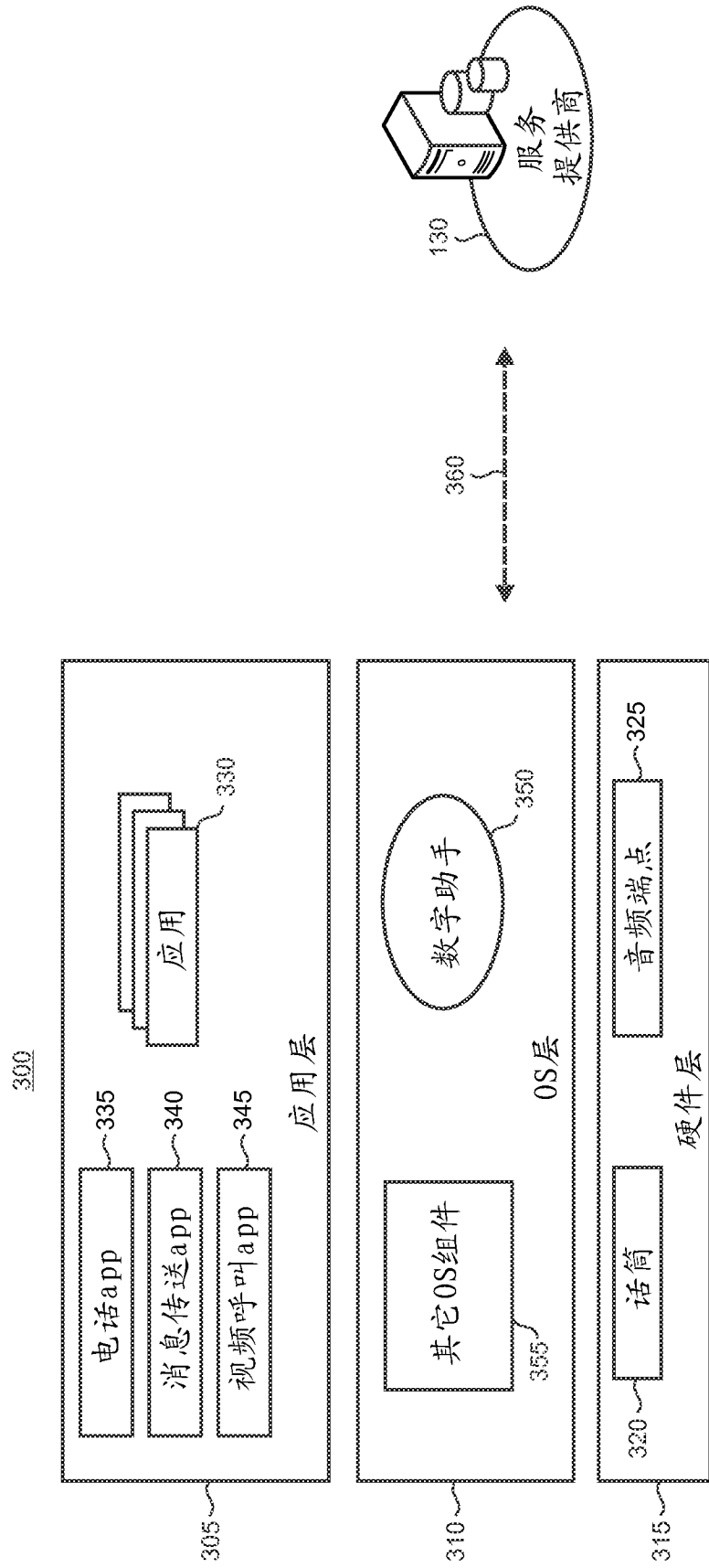


图 3

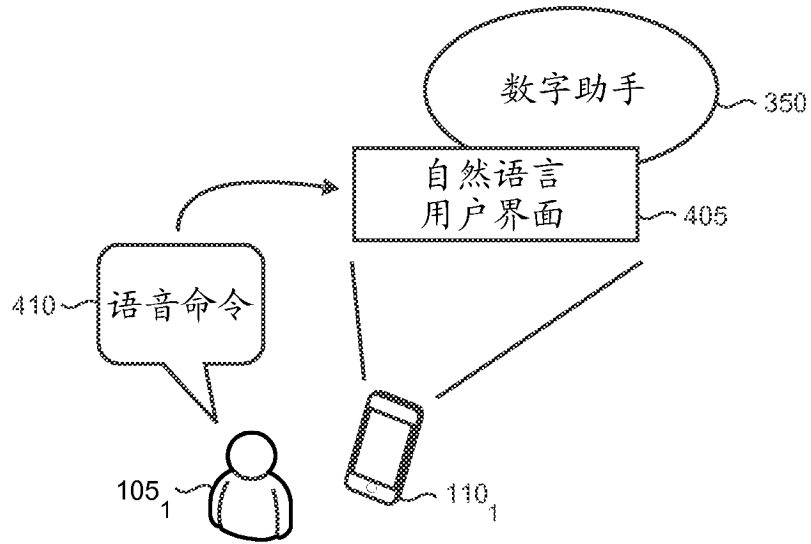


图 4

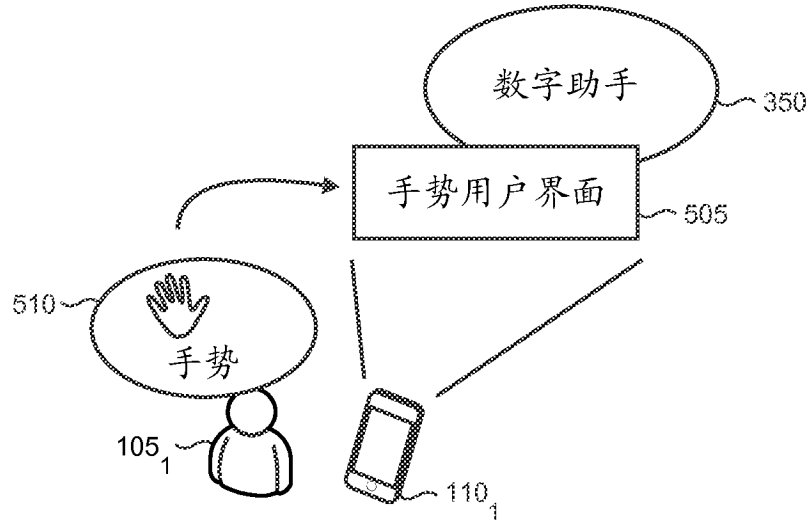


图 5

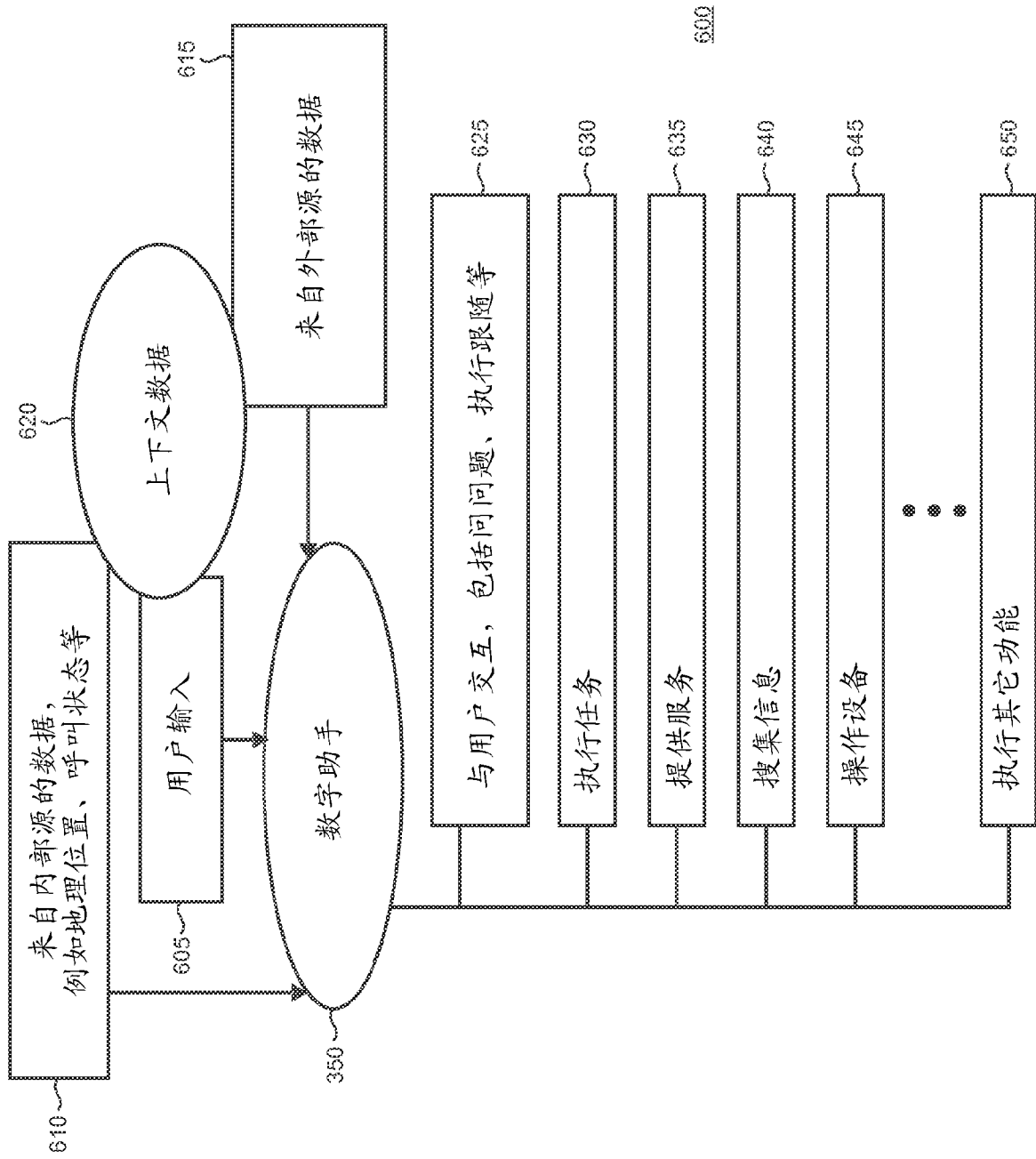


图 6

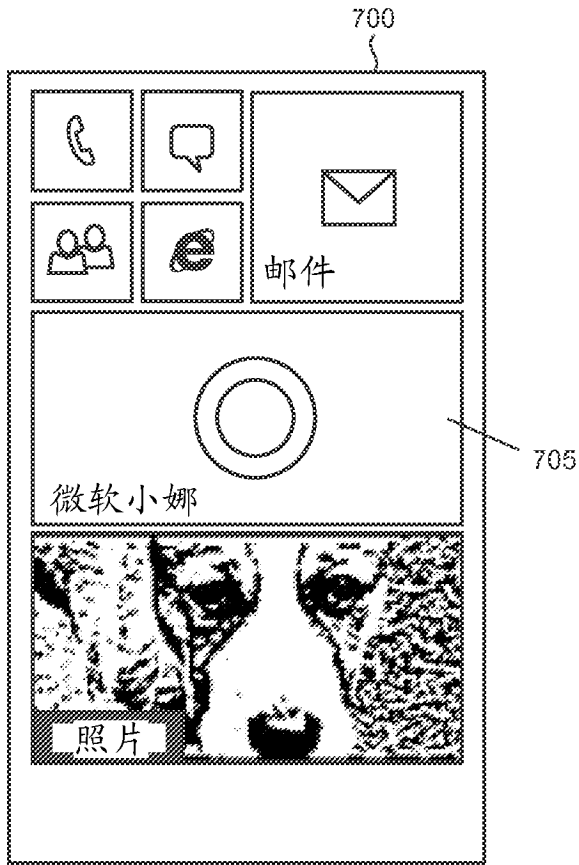


图 7

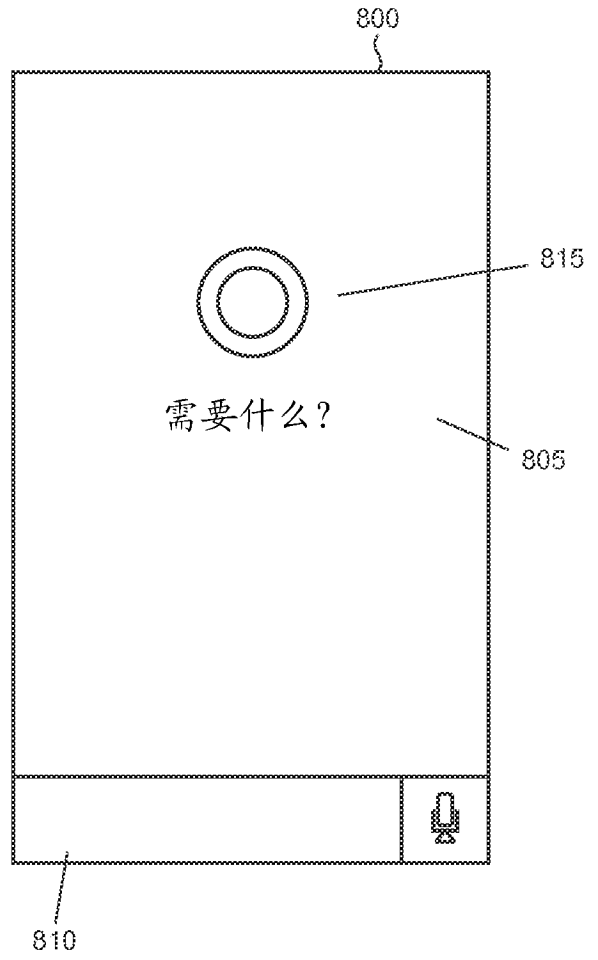


图 8

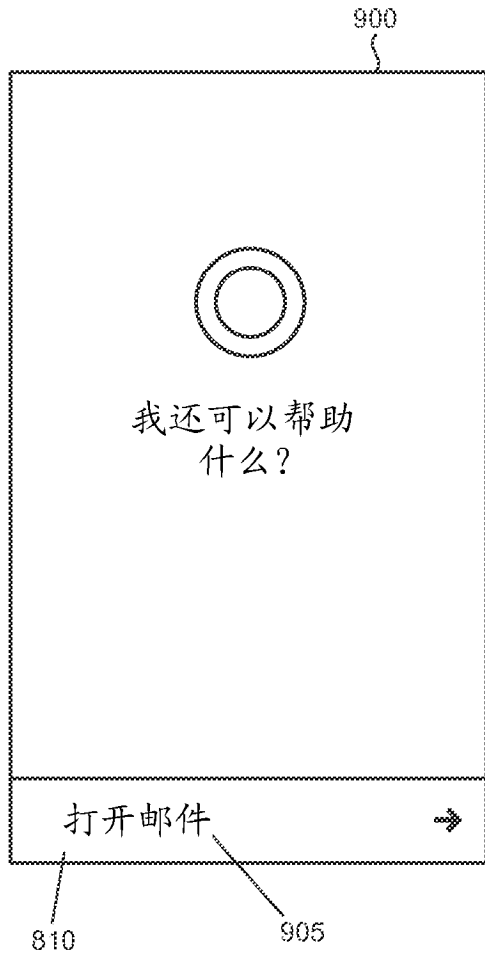


图 9

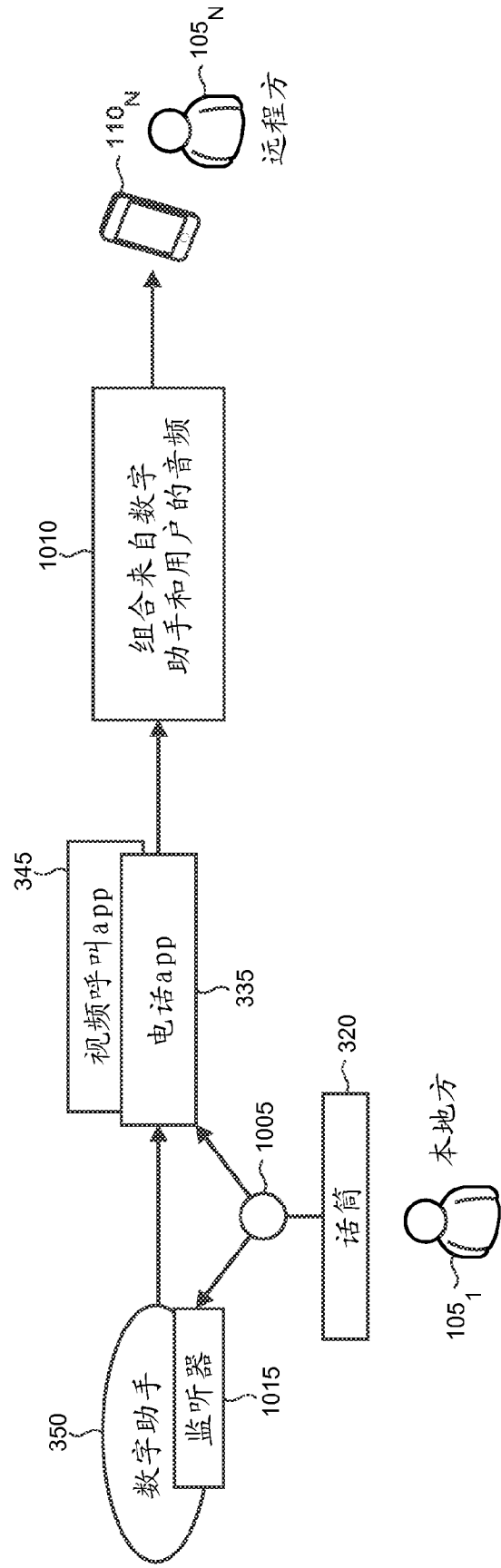


图 10

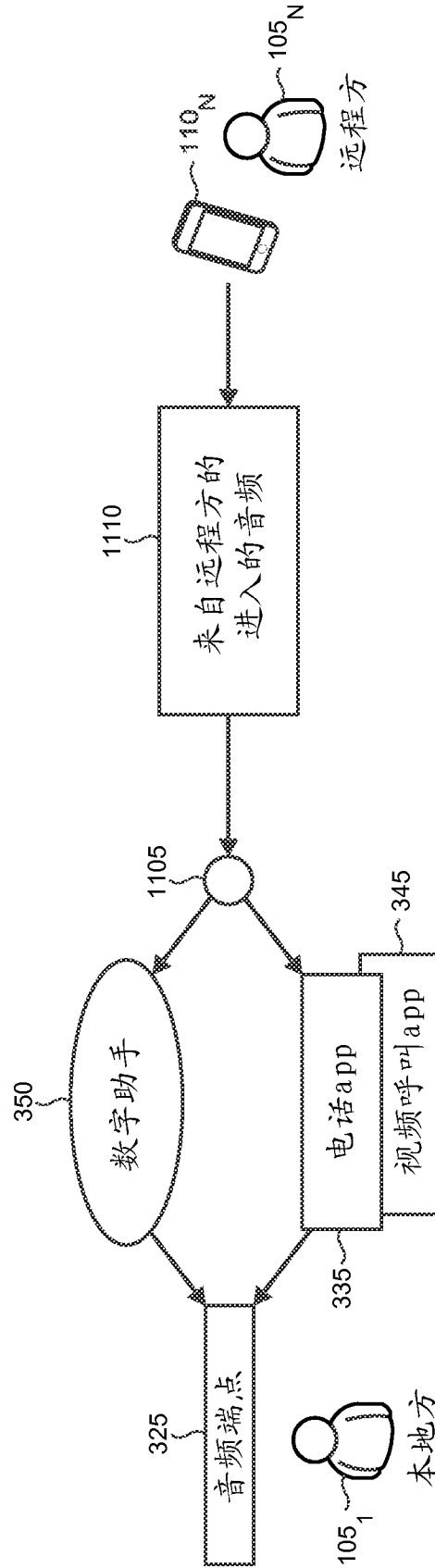


图 11

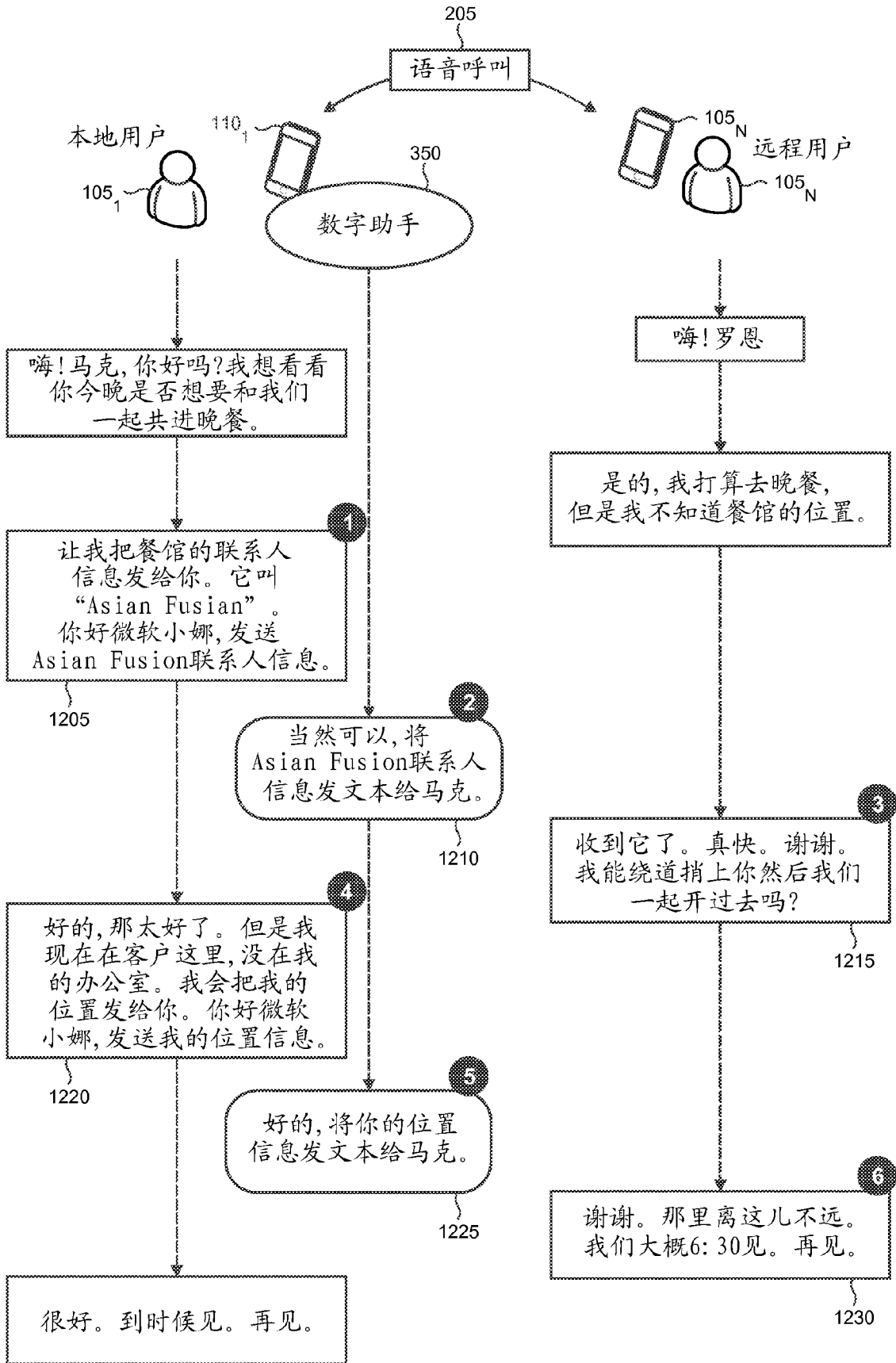


图 12

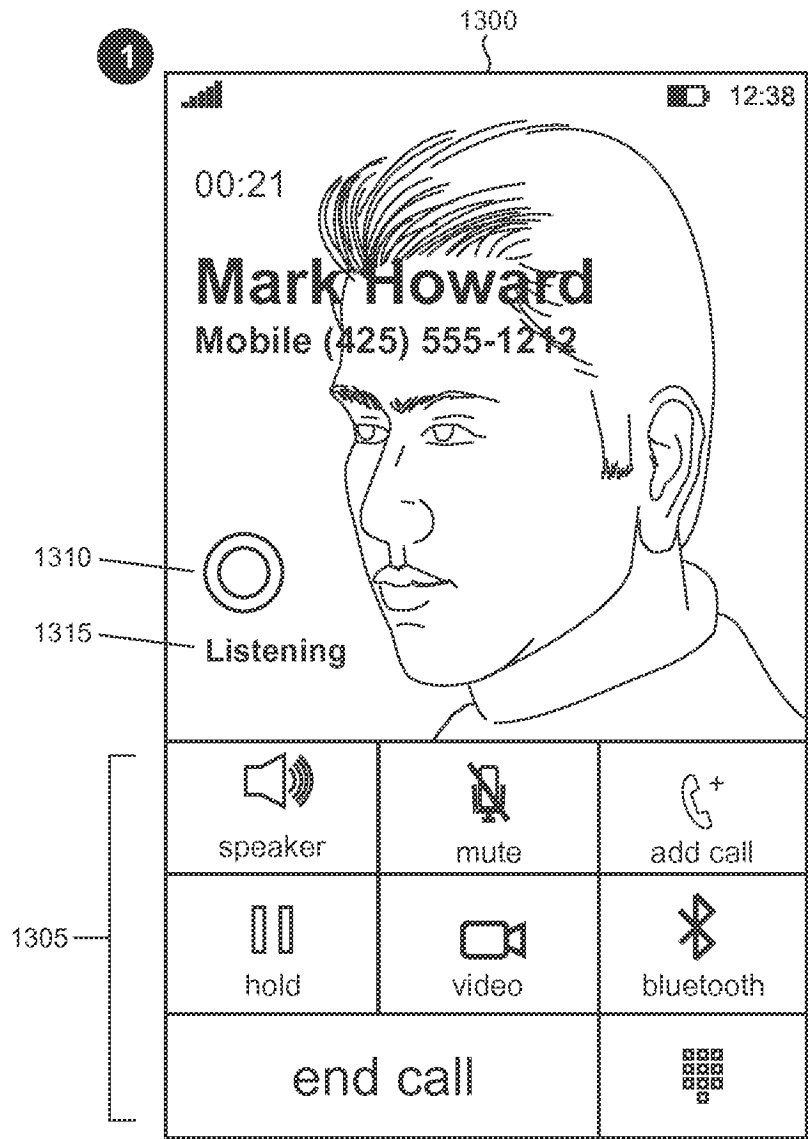


图 13

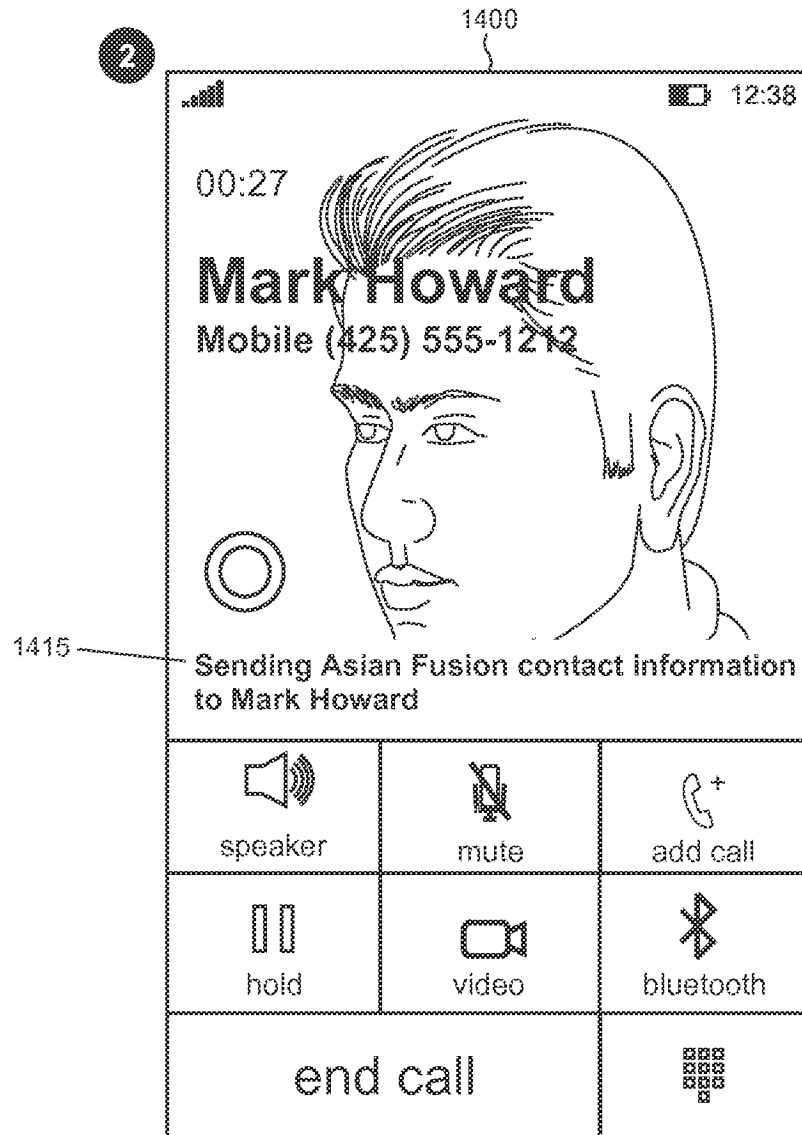


图 14

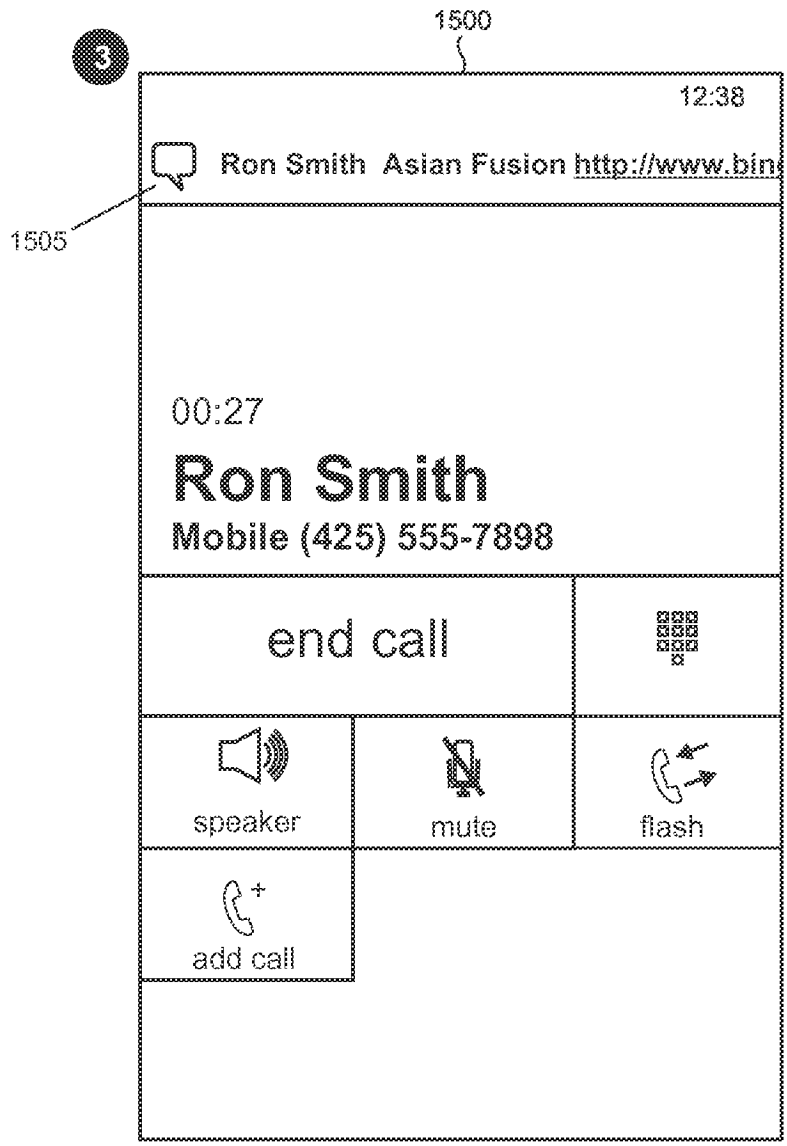


图 15

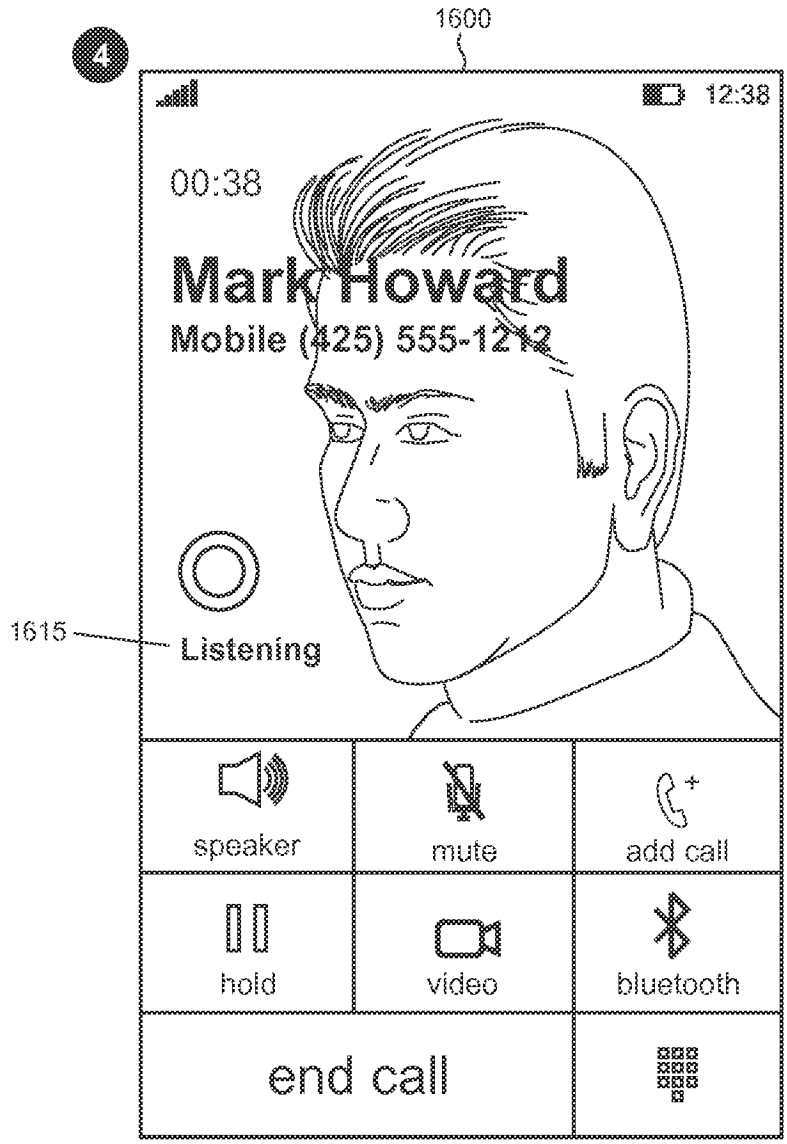


图 16

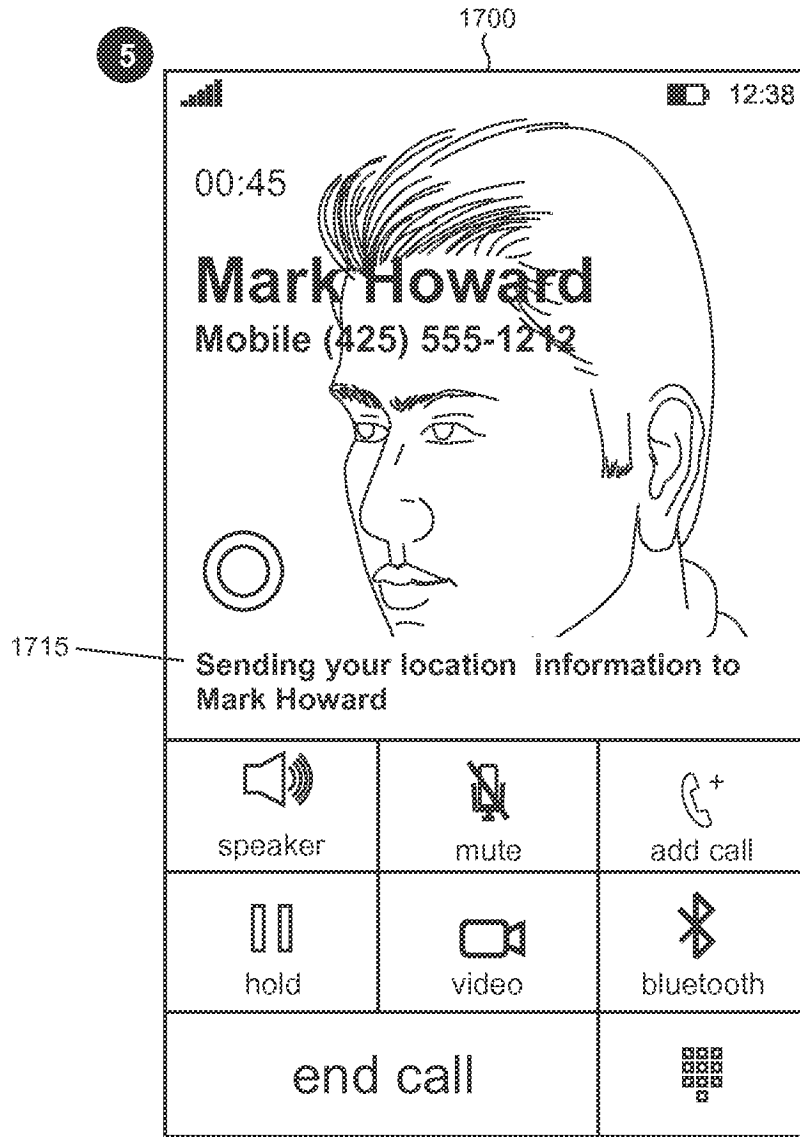


图 17

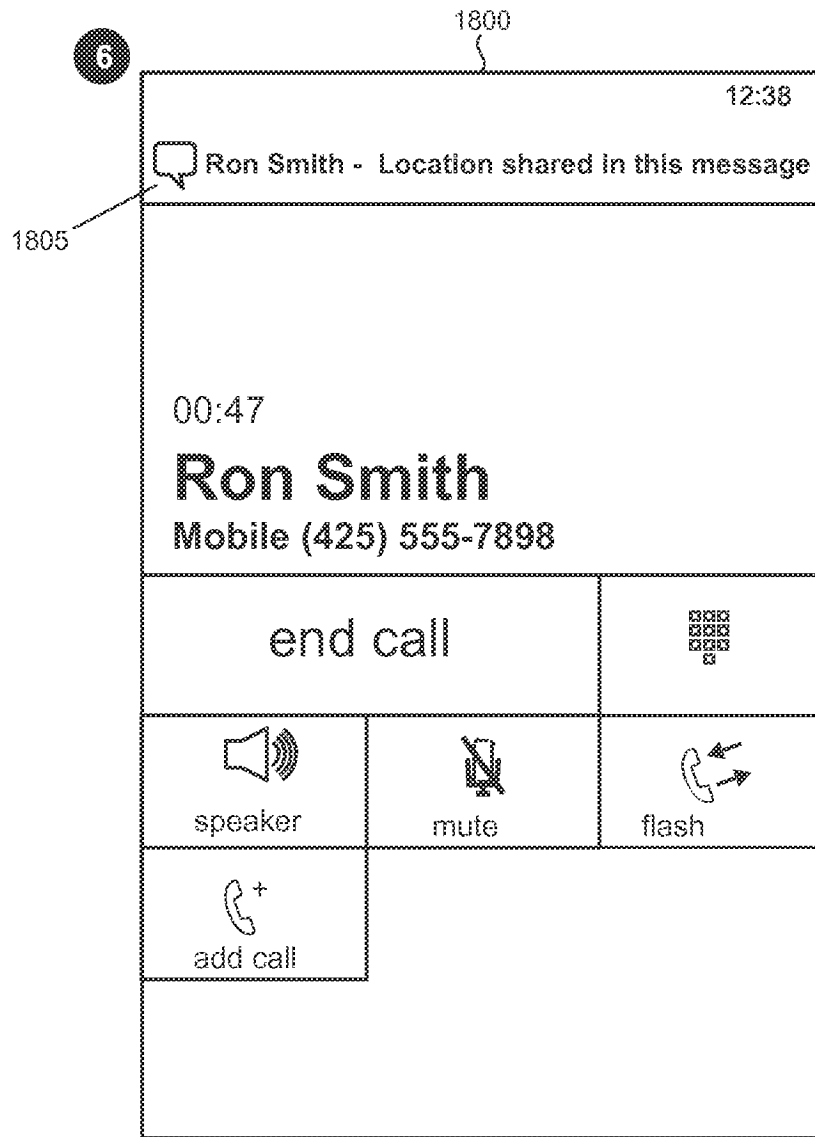


图 18

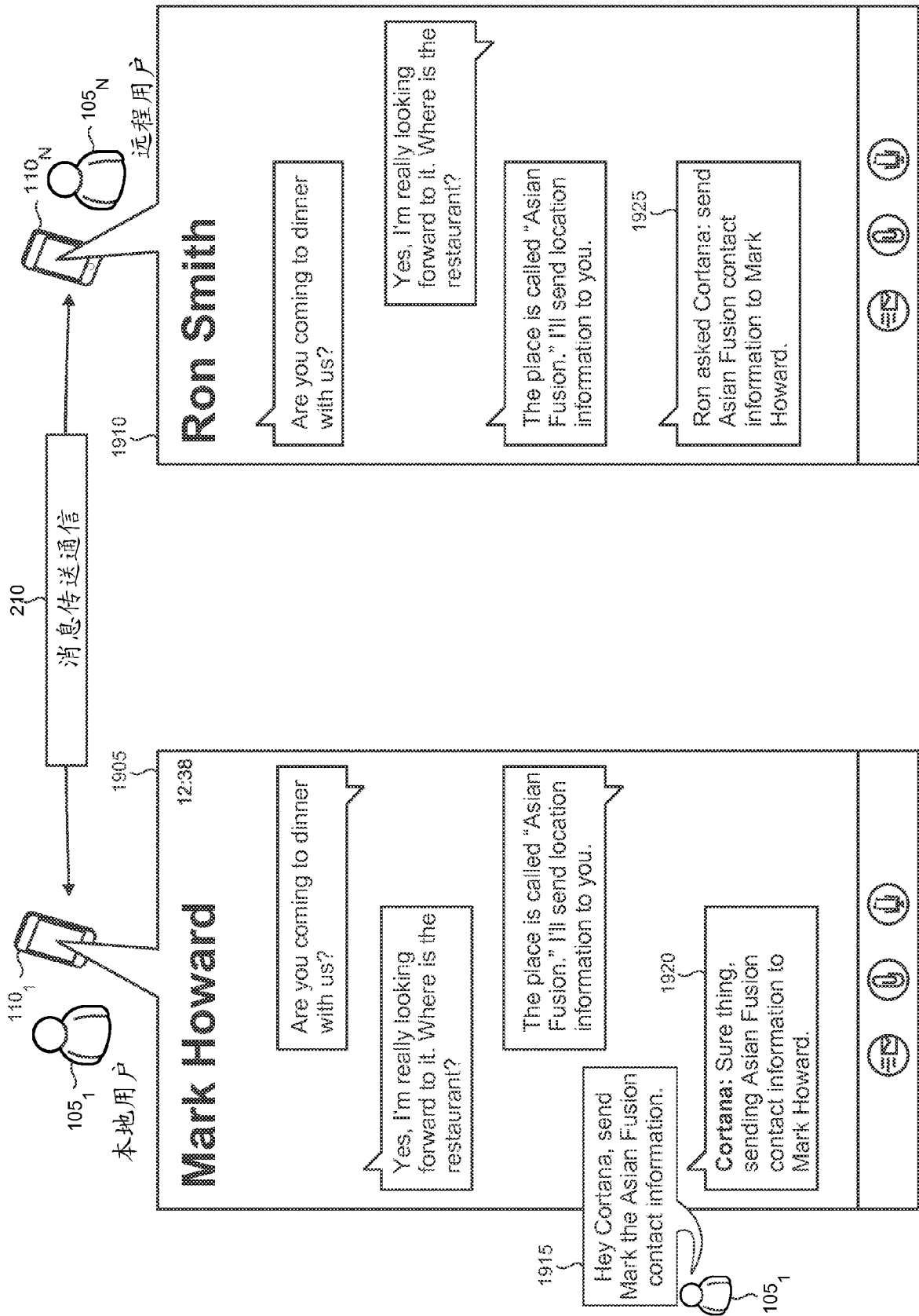


图 19

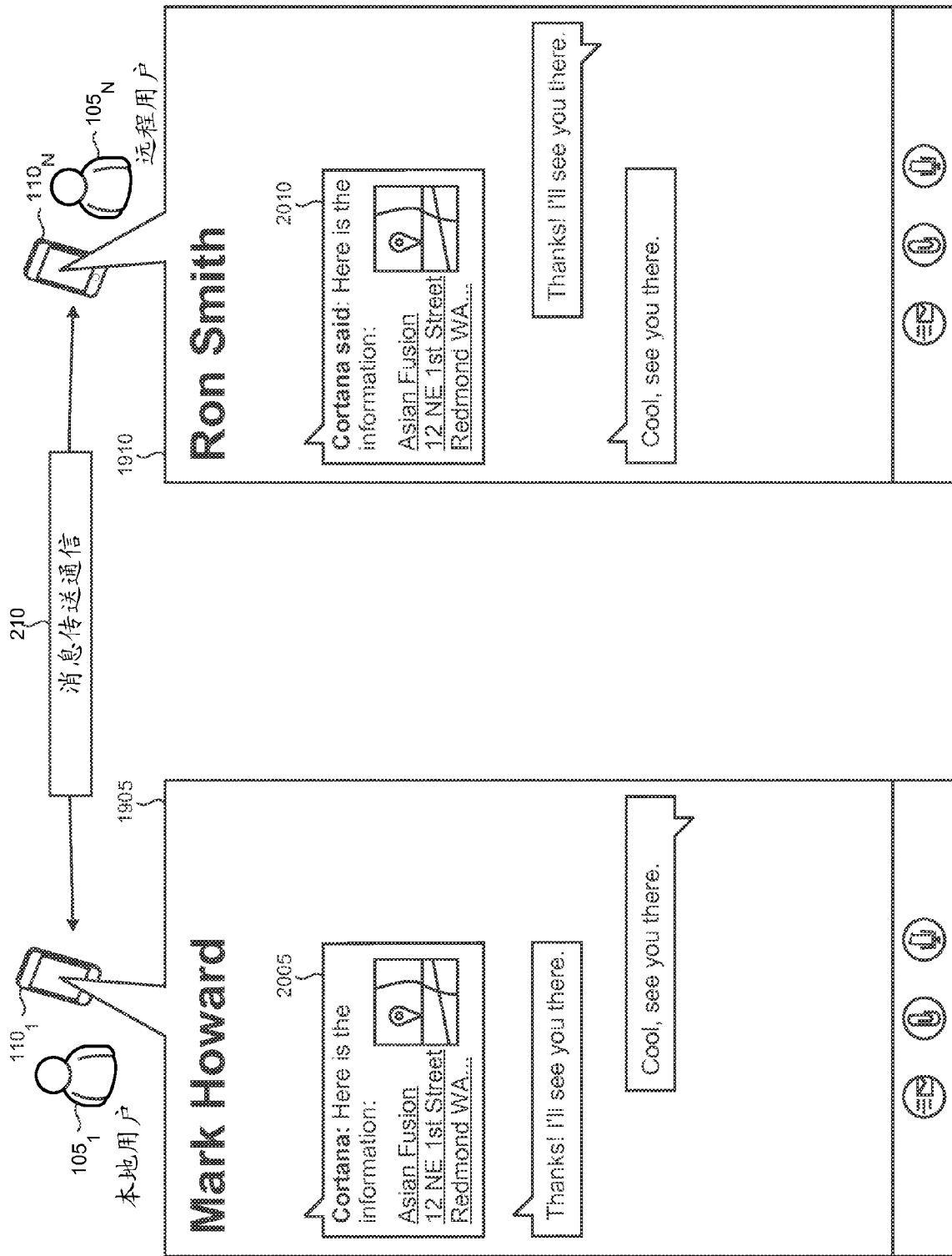


图 20

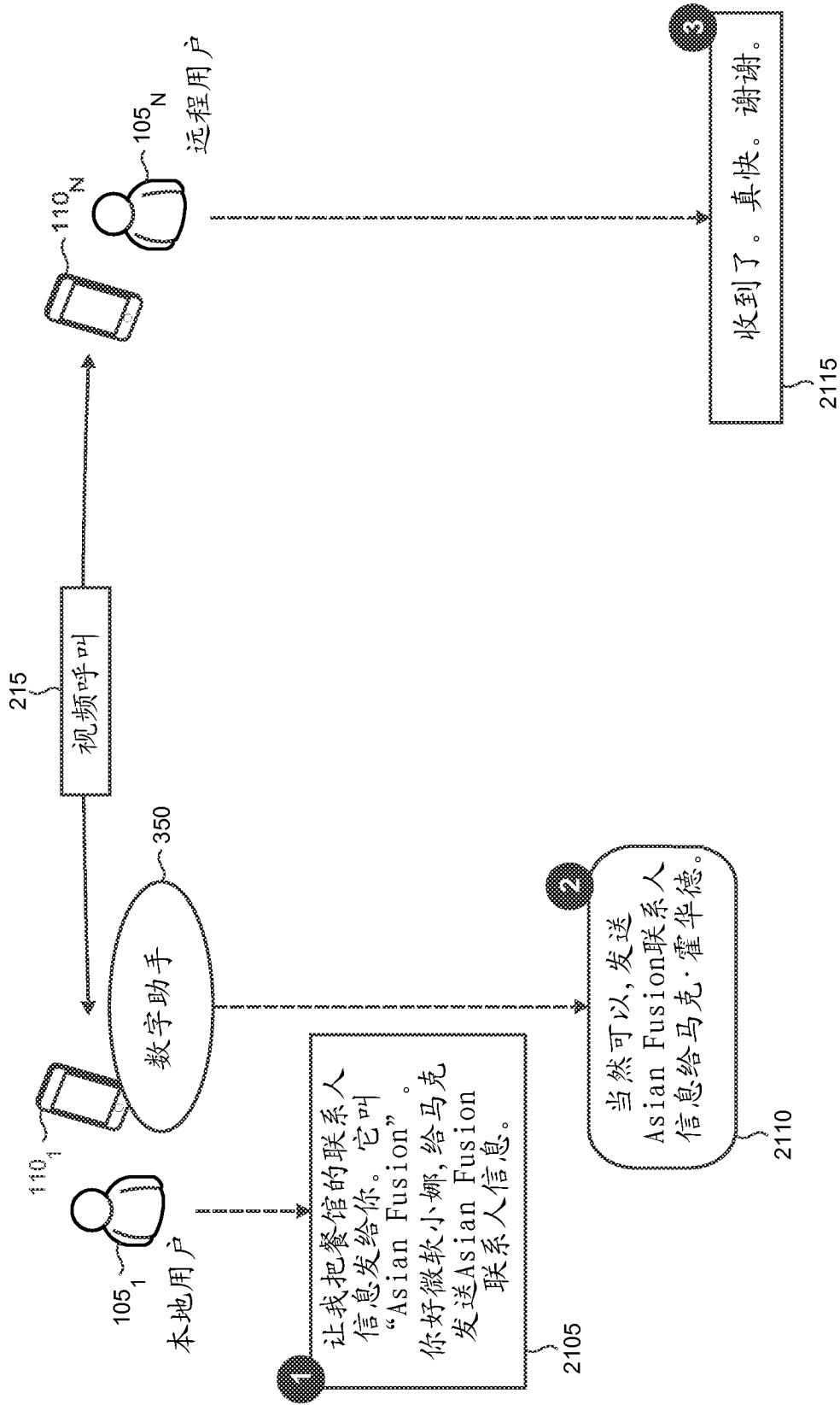


图 21

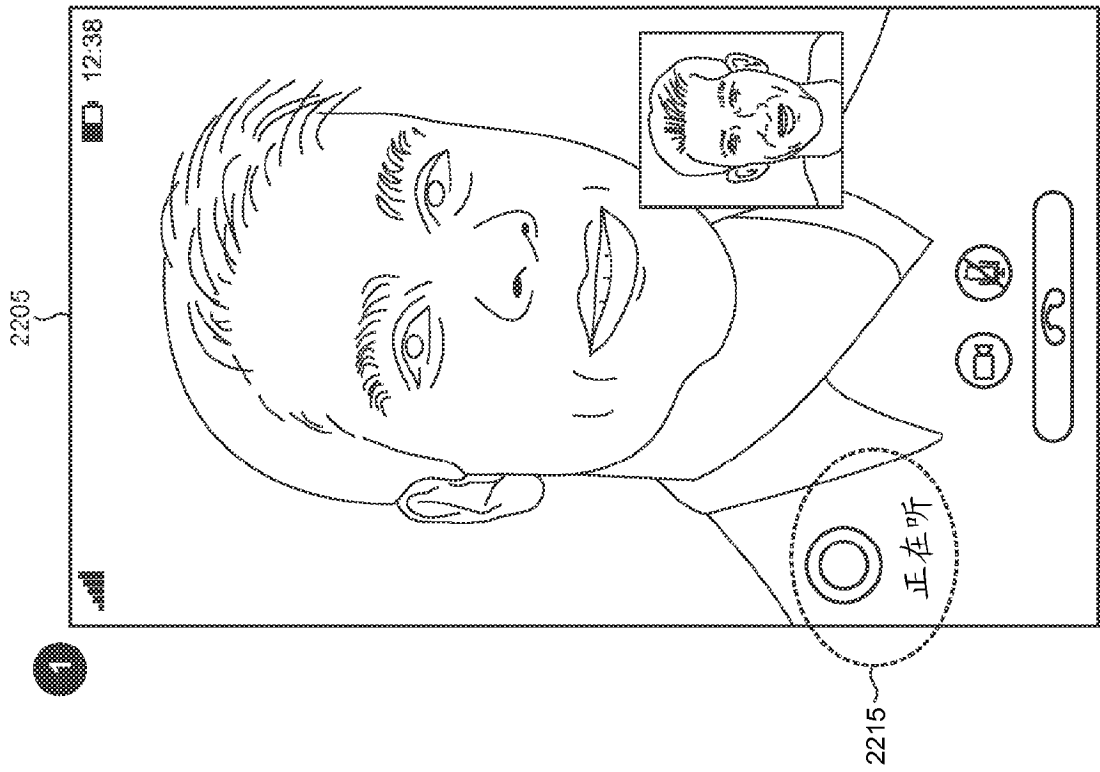
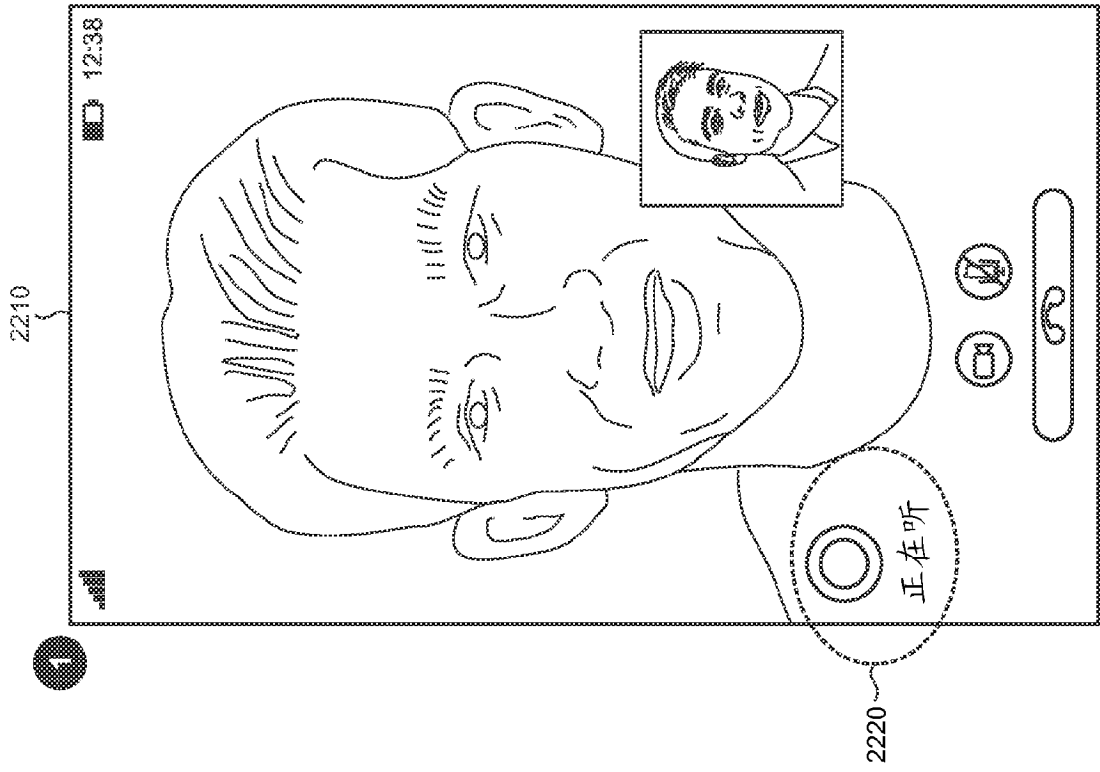


图 22

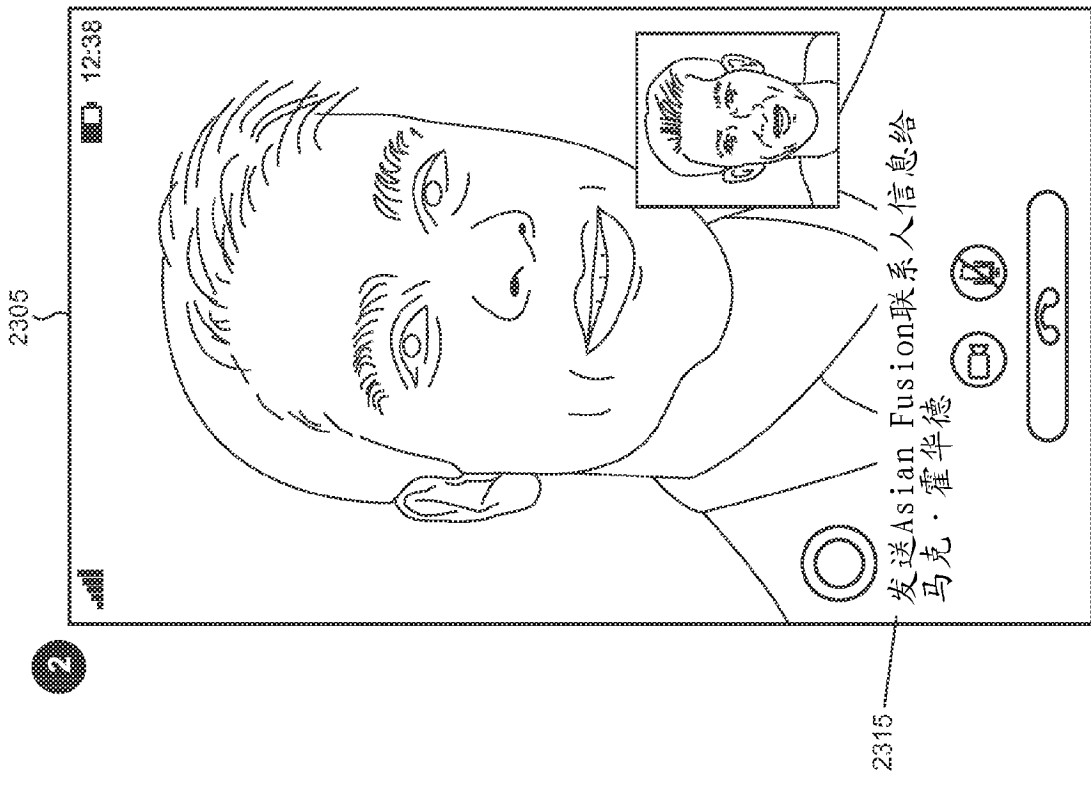
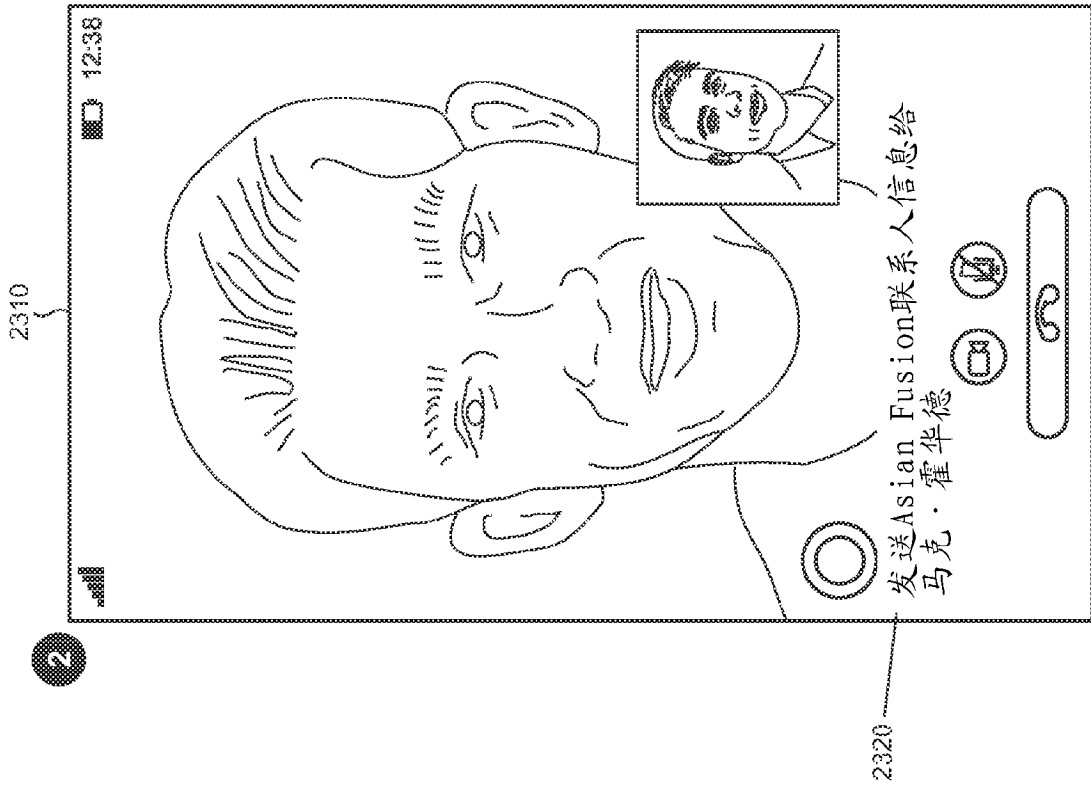


图 23

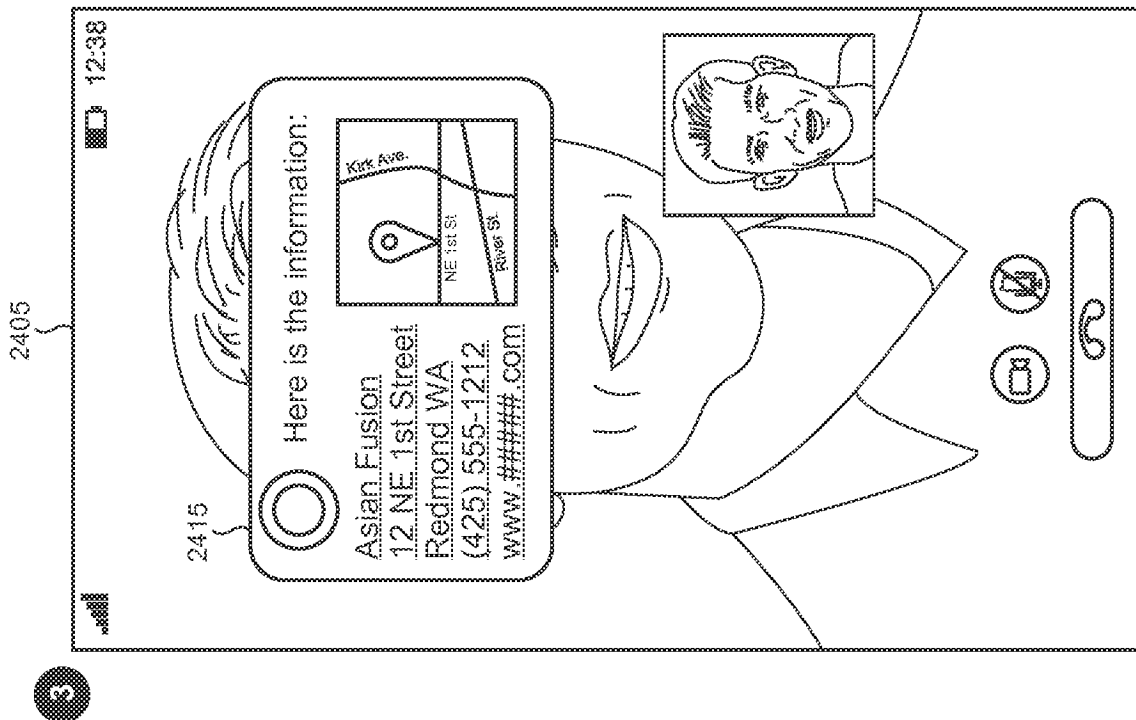
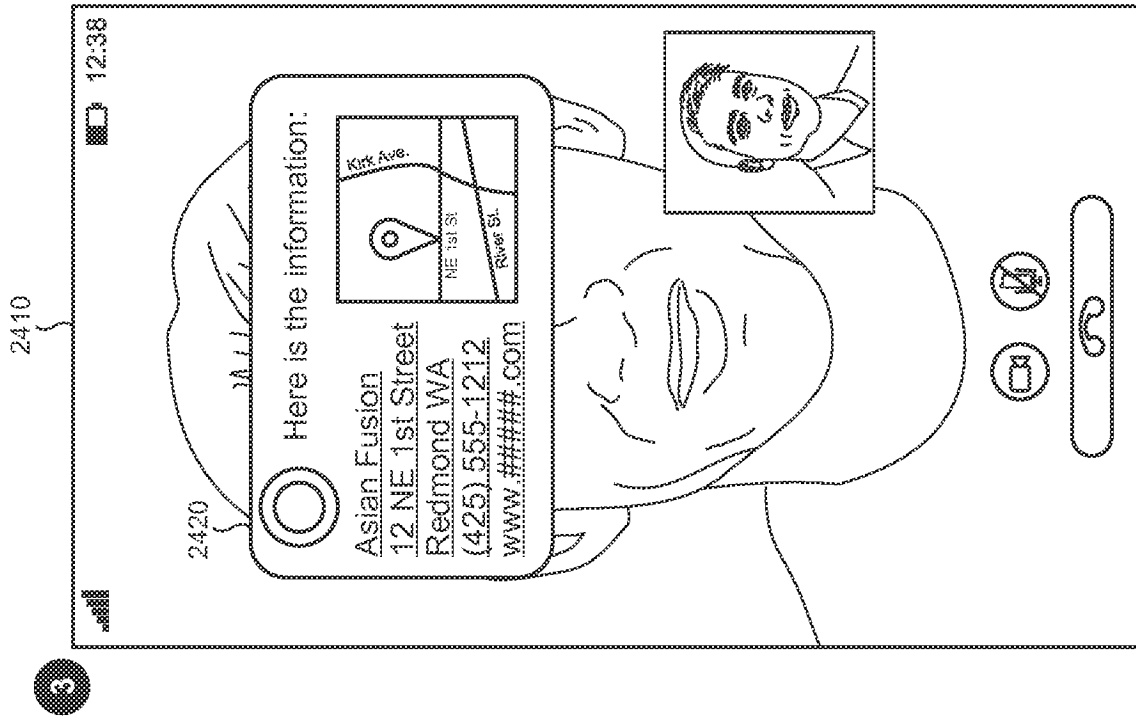


图 24

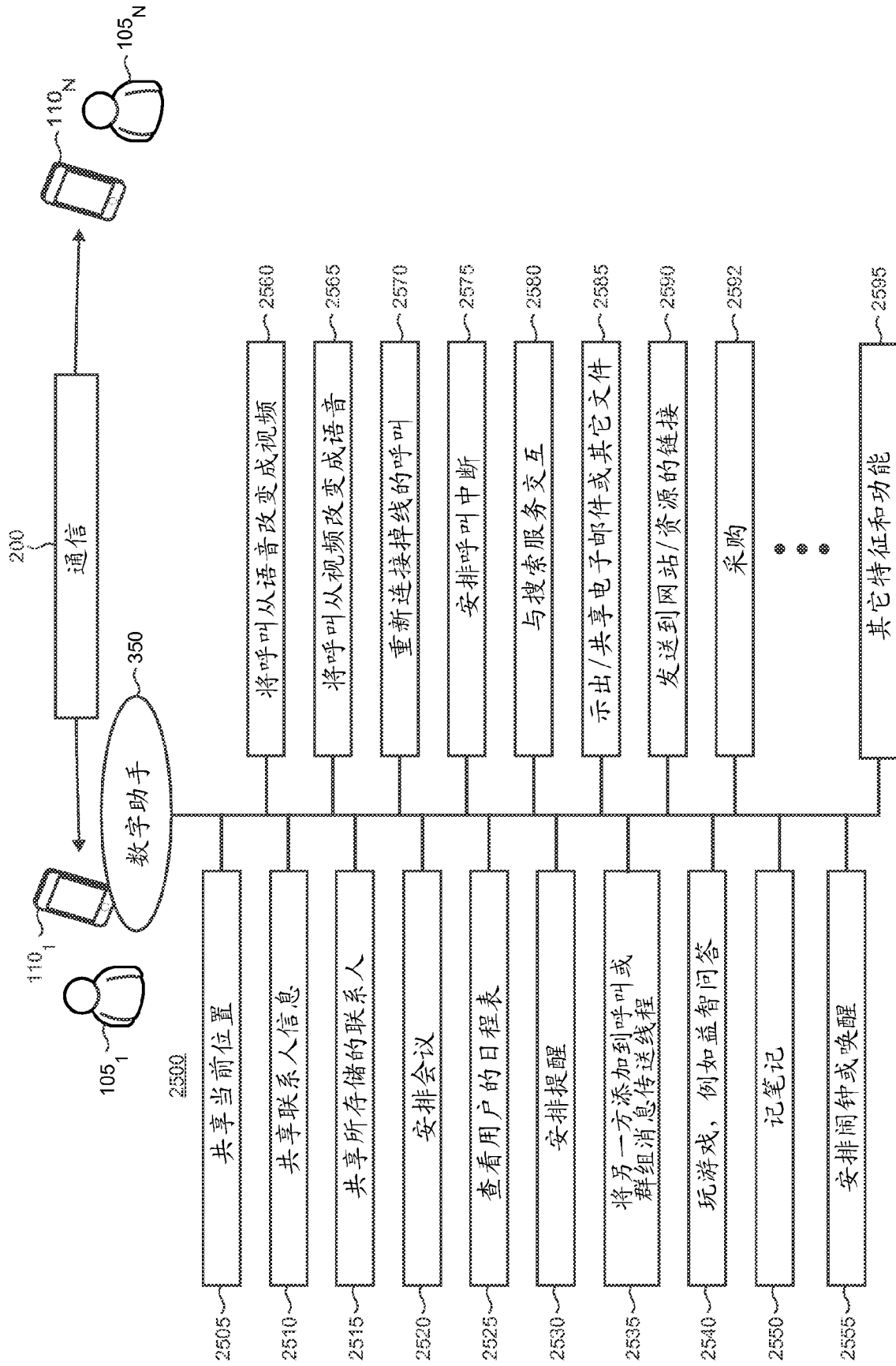


图 25

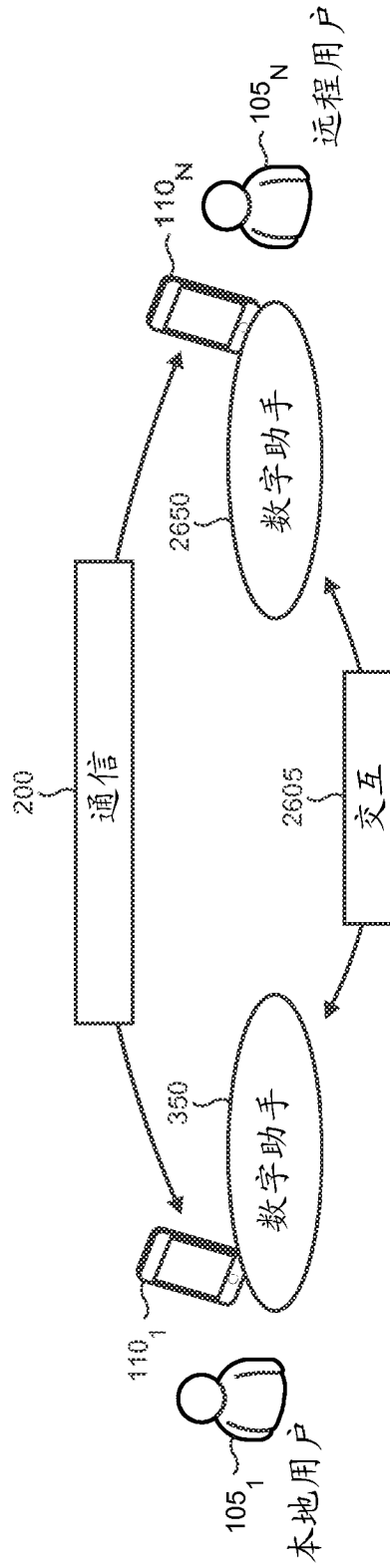


图 26

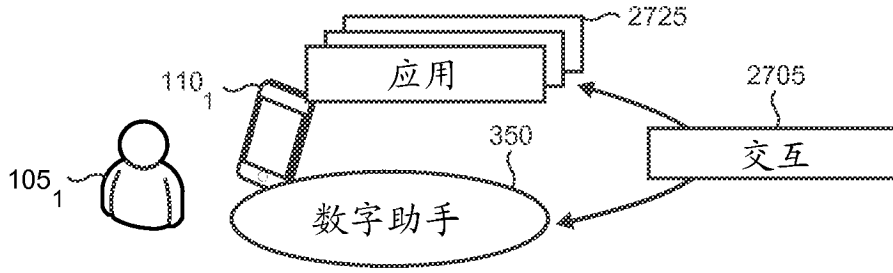


图 27

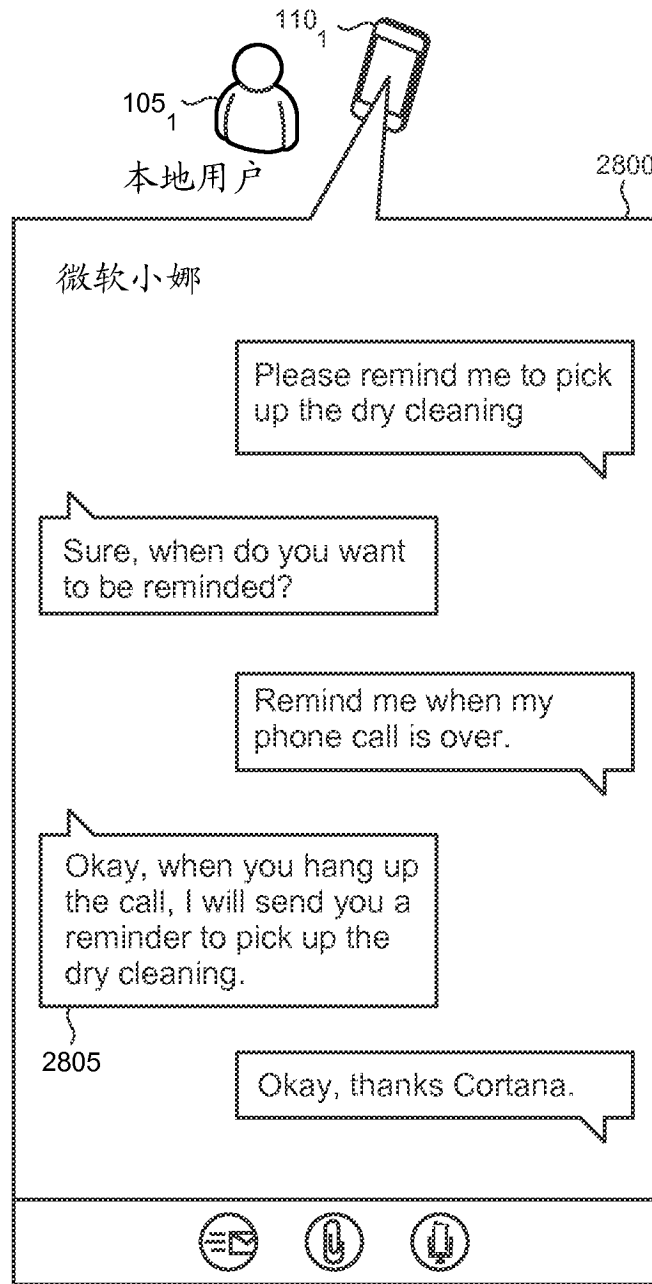


图 28



图 29

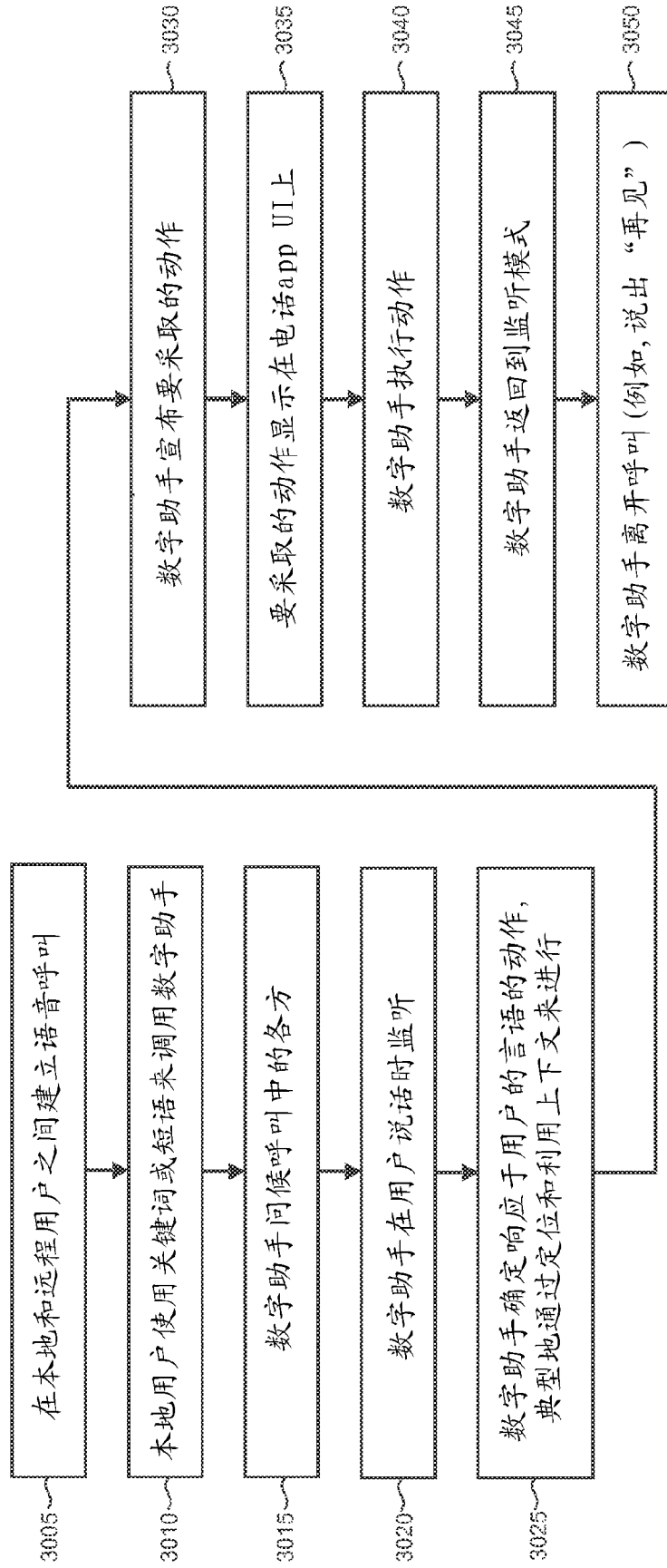


图 30

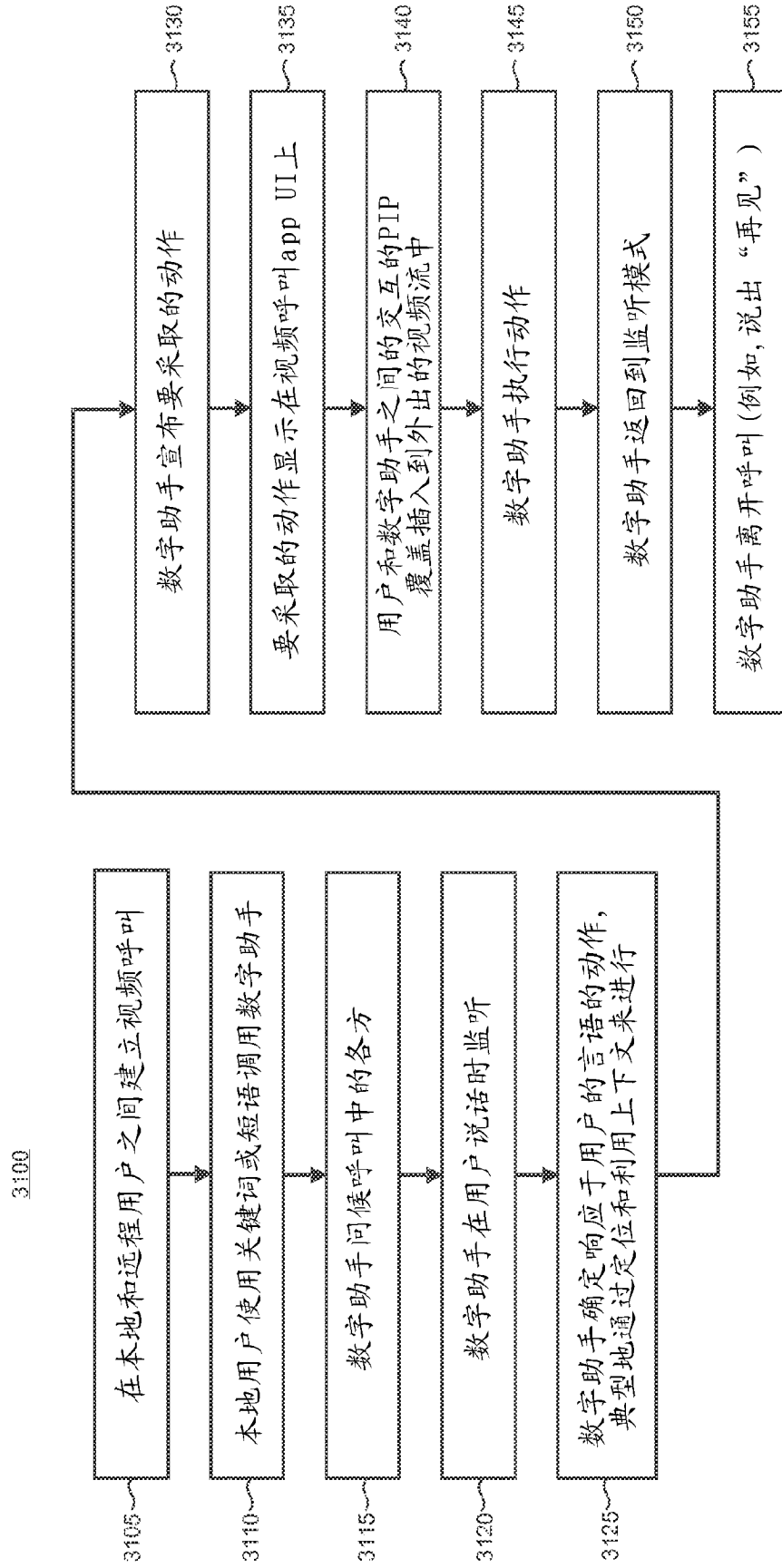


图 31

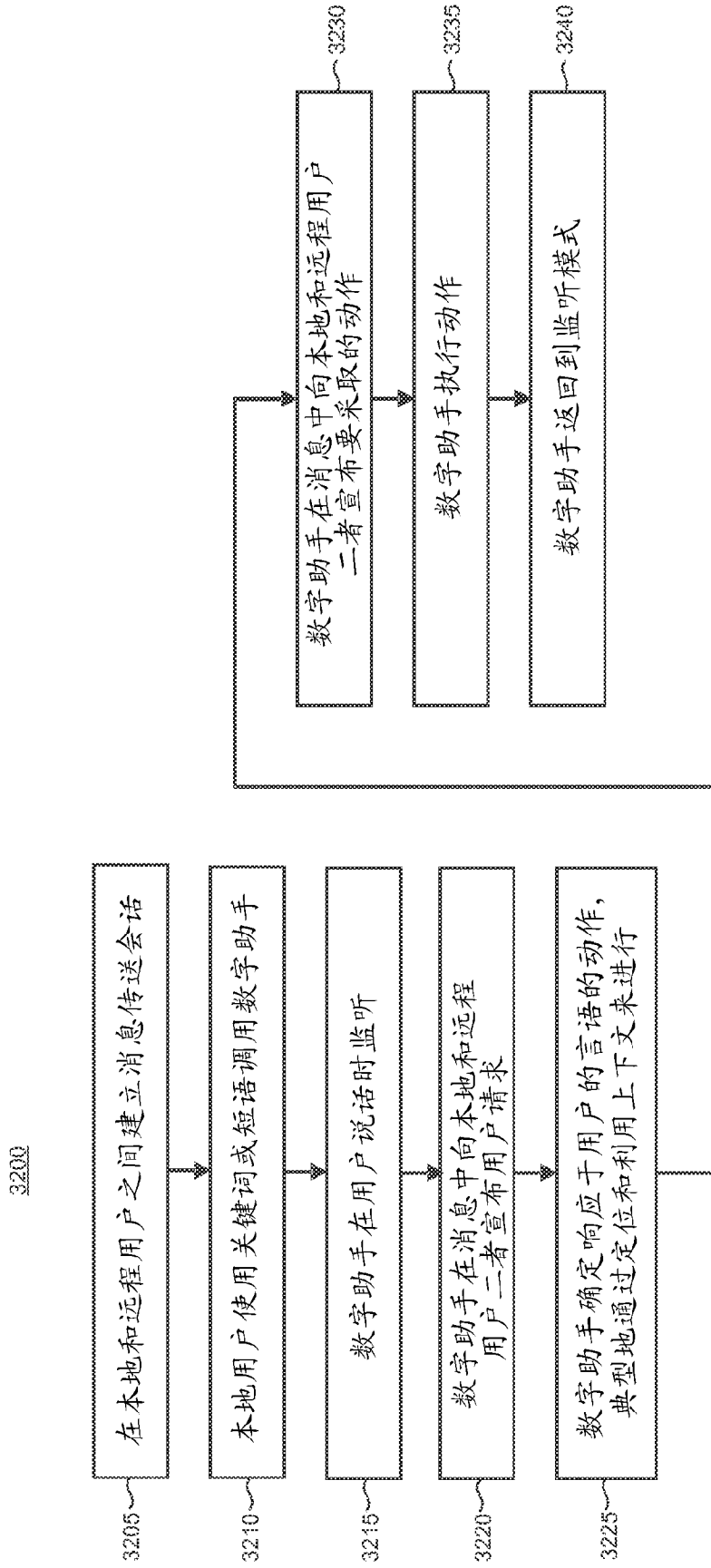


图 32

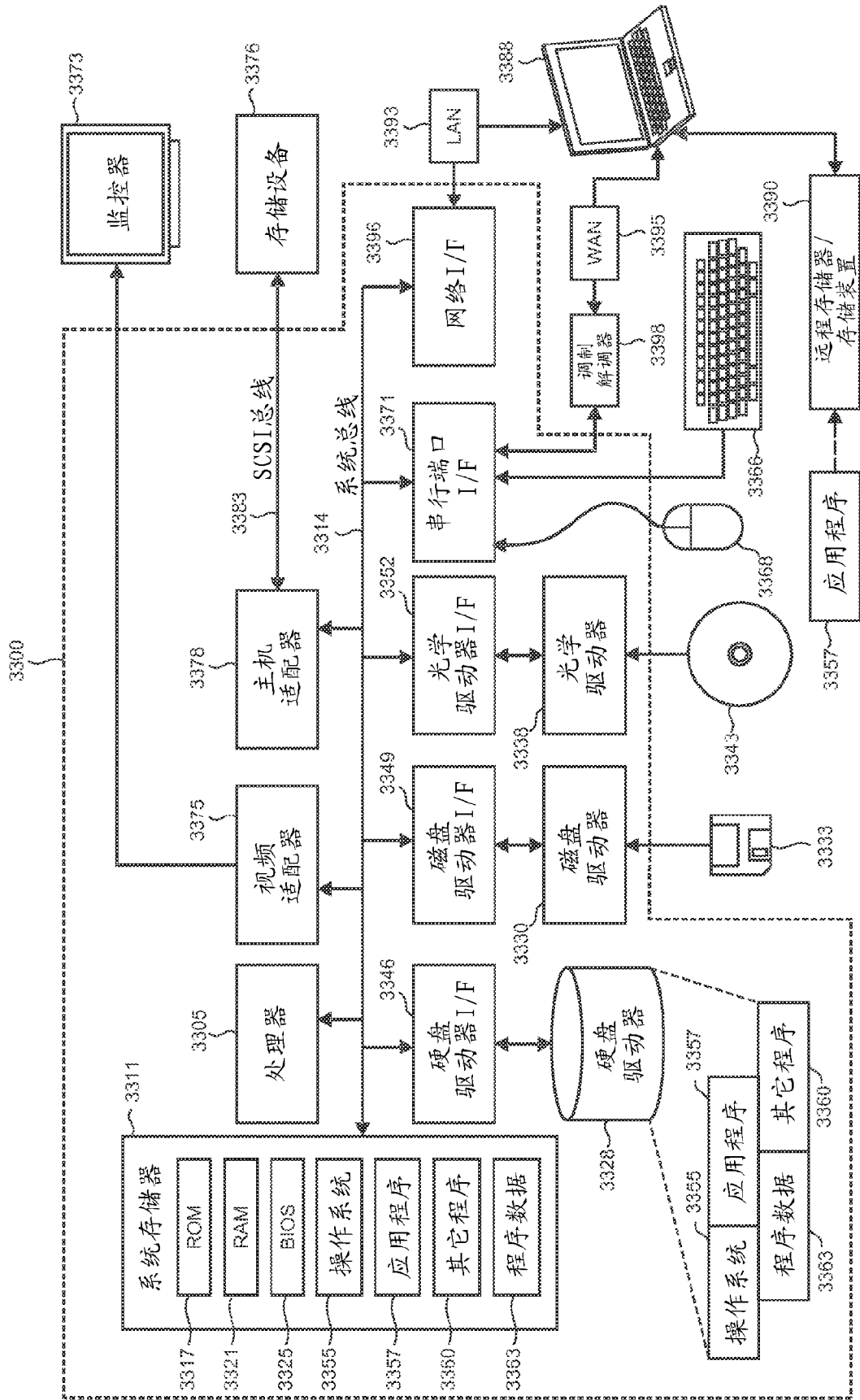


图 33

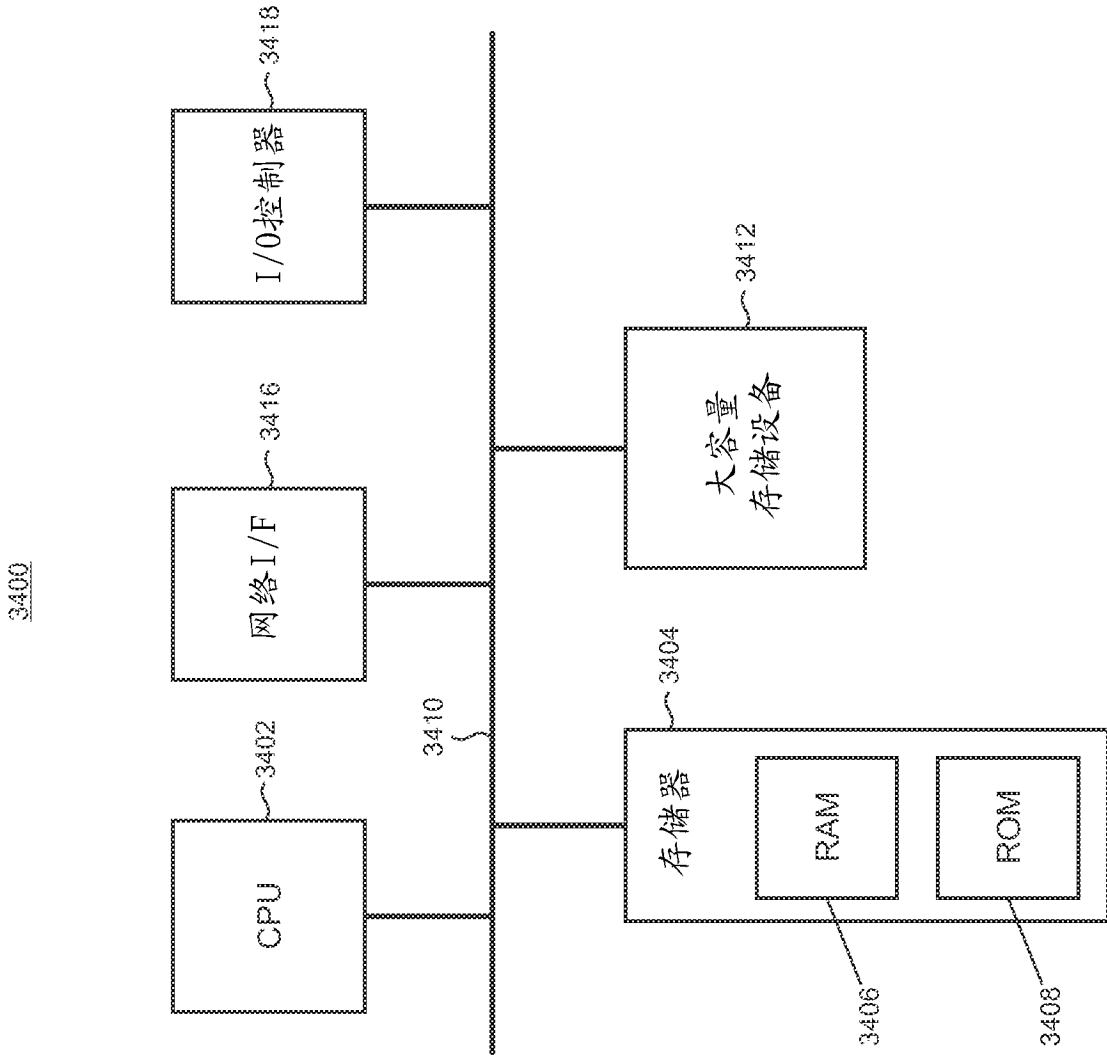


图 34

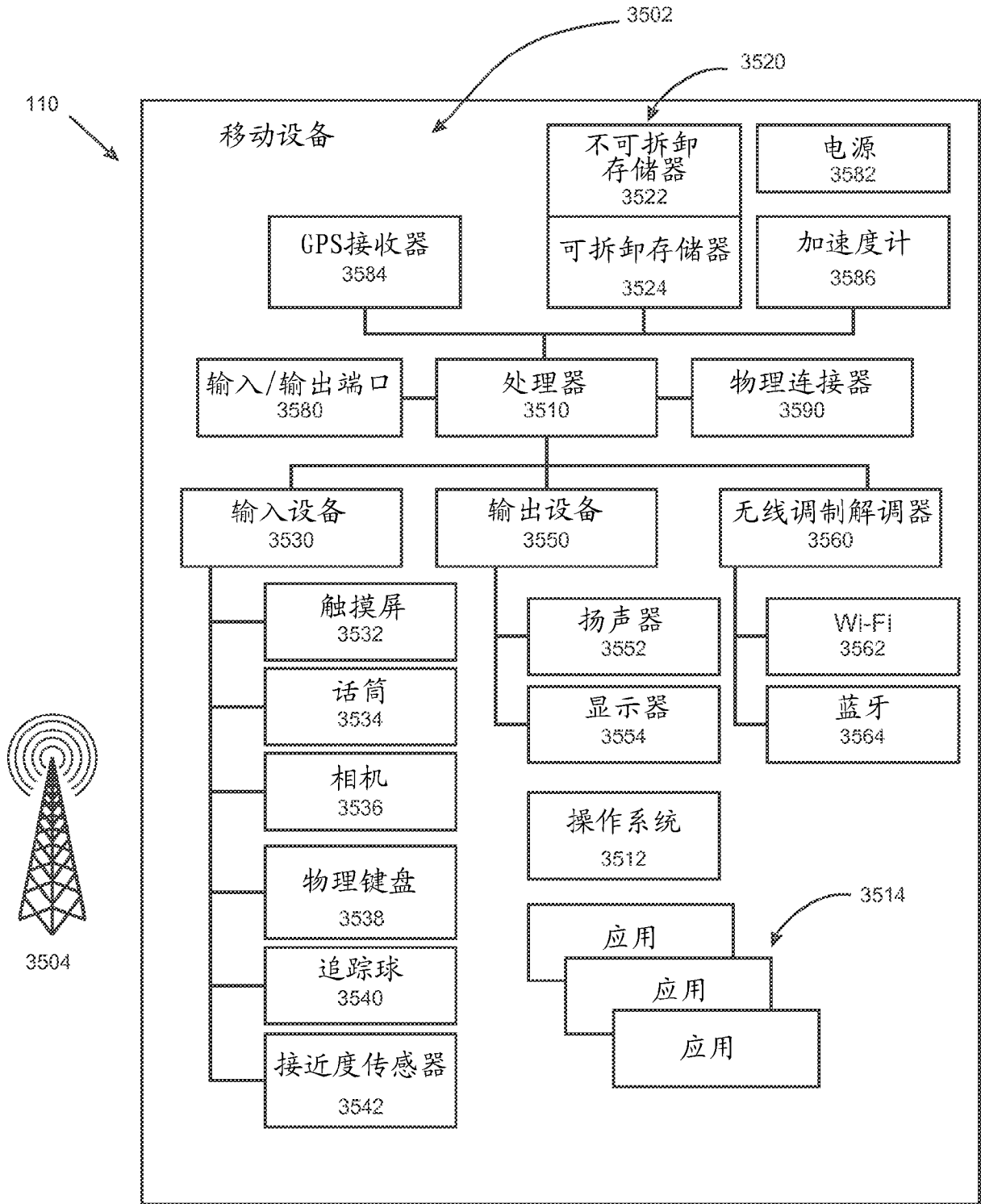


图 35

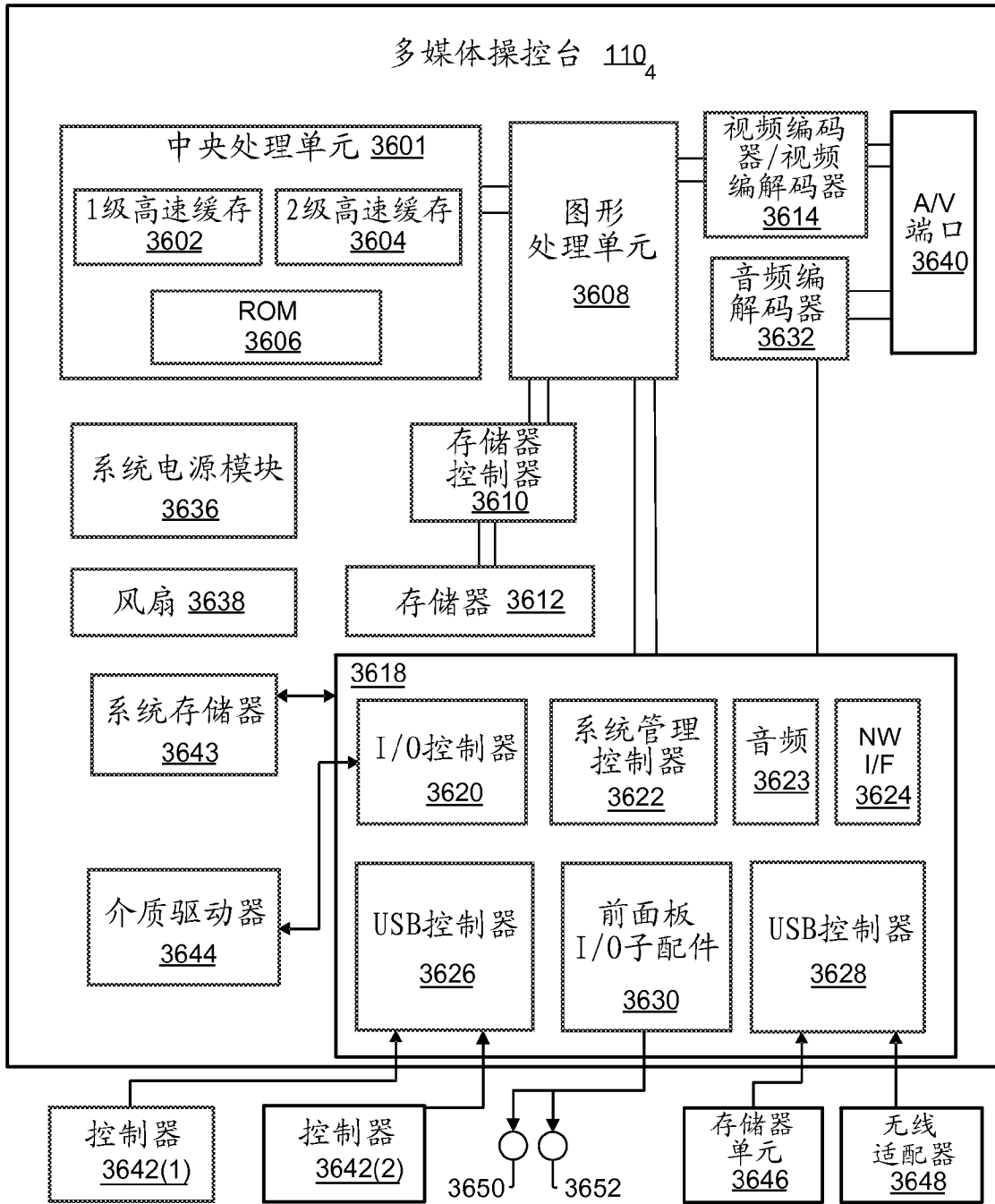


图 36