



(12)发明专利

(10)授权公告号 CN 104685494 B

(45)授权公告日 2018.08.14

(21)申请号 201380038146.9

(22)申请日 2013.06.11

(65)同一申请的已公布的文献号
申请公布号 CN 104685494 A

(43)申请公布日 2015.06.03

(30)优先权数据
13/493,714 2012.06.11 US

(85)PCT国际申请进入国家阶段日
2015.01.16

(86)PCT国际申请的申请数据
PCT/US2013/045265 2013.06.11

(87)PCT国际申请的公布数据
W02013/188463 EN 2013.12.19

(73)专利权人 亚马逊技术公司
地址 美国内华达州

(72)发明人 理查德-霍华德·苏普利
杰弗里-斯科特·海勒
丹尼尔·毕比里塔

(74)专利代理机构 广州华进联合专利商标代理
有限公司 44224
代理人 王程 何冲

(51)Int.Cl.
G06F 17/28(2006.01)

(56)对比文件
CN 102474902 A,2012.05.23,
CN 101431559 A,2009.05.13,

审查员 石爽

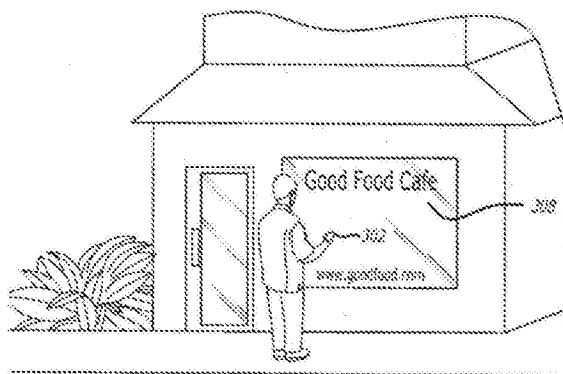
权利要求书2页 说明书12页 附图9页

(54)发明名称

文本辨识驱动的功能

(57)摘要

本发明公开用于将文本信息提供给应用程序、系统或服务的各种方法。具体来说,各个实施方案允许用户利用便携式计算设备的照相机来捕获图像。所述计算设备能够拍摄所述图像并处理所述图像以辨识、识别和/或隔离文本,以便将所述文本转发到应用程序或功能。所述应用程序或功能接着可利用所述文本来大致上实时地执行动作。所述文本可以包括电子邮件、电话号码、URL、地址等等,并且应用程序或功能可以拨打所述电话号码、导航到所述URL、打开通讯簿来保存联系人信息、显示地图来展示所述地址等等。



1. 一种计算机实施方法,其包括:
在配置有可执行指令的一个或多个计算机系统的控制下,
识别与图像中的文本相关联的图像信息;
根据与文本相关联的所述图像信息确定文本类型;
从多个功能中确定与所述文本类型相关联的至少一个功能;
使用全球定位系统接收便携式计算设备的位置数据;
至少基于为所述至少一个功能分配的优先级顺序和所述位置数据,确定最高优先级,
其中,所述分配的优先级顺序可通过代表用户在一段时间内的活动的历史用户输入数据调整;
提示与所述最高优先级相对应的第一功能;以及
允许计算设备执行所述第一功能。
2. 如权利要求1所述的计算机实施方法,其还包括:
提示用户确认要执行所述至少一个功能的意图。
3. 如权利要求1所述的计算机实施方法,其还包括:
提示用户在与所述文本类型相关联的第一动作与至少一个第二动作之间选择。
4. 如权利要求1所述的计算机实施方法,其还包括:
获得所述图像,所述图像是从所述计算设备的照相机获得。
5. 如权利要求1所述的计算机实施方法,其中所述图像是以下之一:当所述计算设备的照相机聚焦于对象上时自动获得的图像,或响应于用户输入而获得的图像。
6. 如权利要求1所述的计算机实施方法,其中所述至少一个功能包括以下操作中的至少一个:当所述文本类型指示电话号码时拨号;当所述文本类型指示电子邮件时打开电子邮件应用程序来撰写电子邮件;当所述文本类型指示网址时导航到URL;当所述文本类型指示至少一个电话号码和地址的存在时打开通讯簿来保存联系人信息;当所述文本类型指示物理地址时显示地图来展示地址;执行语言翻译;或求解数学方程式。
7. 如权利要求6所述的计算机实施方法,其还包括:
将所述图像加以阈值处理以标准化与文本相关联的所述图像信息,以便顾及所述图像的模糊、噪声或照明变化中的至少一个。
8. 如权利要求1所述的计算机实施方法,其中使用计算机视觉算法或光学字符识别算法中的至少一个来从所述图像至少识别与文本相关联的所述图像信息。
9. 如权利要求1所述的计算机实施方法,还包括:
根据辨识的语音确定动作;以及
至少基于所述动作允许所述计算设备执行所述第一功能,其中由与所述计算设备通信的远程服务器执行以下操作中的至少一个:识别与文本相关联的所述图像信息、确定所述文本类型,或确定所述至少一个功能。
10. 一种便携式计算设备,其包括:
处理器;以及
存储设备,其包括在由所述处理器执行时使所述处理器进行以下操作的指令:
识别图像中的文本;
从所述识别的文本识别文本模式;

从多个功能中确定与所述文本模式相关联的至少一个功能；
使用全球定位系统接收所述便携式计算设备的位置数据；
至少基于为所述至少一个功能分配的优先级顺序和所述位置数据，确定最高优先级，其中，所述分配的优先级顺序可通过代表用户在一段时间内的活动的历史用户输入数据调整；

提示与所述最高优先级相对应的第一功能；以及
执行所述第一功能。

11. 如权利要求10所述的便携式计算设备，其中在所述便携式计算设备上本地执行以下操作：识别所述文本、识别所述文本模式、确定所述至少一个功能以及执行所述至少一个功能。

12. 如权利要求10所述的便携式计算设备，其中由与所述便携式计算设备通信的远程服务器执行以下操作中的至少一个：识别所述文本、识别所述文本模式或确定所述至少一个功能。

13. 如权利要求10所述的便携式计算设备，其还包括：

照相机，所述照相机用于在其聚焦于对象上之后自动获得所述图像或响应于来自用户的输入而获得所述图像。

14. 如权利要求10所述的便携式计算设备，其还包括在由所述处理器执行时使所述处理器进行以下操作的指令：

使用光学字符辨识算法来分析所述图像的至少一个区域以辨识所述至少一个区域中的所述文本；其中：

用来确定至少一个功能的所述指令还包括在由所述处理器执行时使所述处理器确定与所述文本模式相关联的应用程序的指令；以及

用来执行至少一个功能的所述指令还包括在由所述处理器执行时使所述处理器将所述辨识的文本自动提供给所述应用程序的指令。

15. 如权利要求14所述的便携式计算设备，其还包括在由所述处理器执行时使所述处理器进行以下操作的指令：

使所述应用程序执行以下操作中的至少一个：当所述文本模式指示电话号码时拨号；当所述文本模式指示电子邮件时打开电子邮件应用程序来撰写电子邮件；当所述文本模式指示网址时导航到URL；当所述文本模式指示至少一个电话号码和地址的存在时打开通讯簿来保存联系人信息；当所述文本模式指示物理地址时显示地图来展示地址；执行语言翻译；或求解数学方程式。

文本辨识驱动的功能

[0001] 发明背景

[0002] 随着人们越来越多地利用各种计算设备(包括如平板计算机和智能电话的便携式设备),适应人们与这些设备交互的方式可为有利的。用户习惯于手动输入电话号码以进行电话呼叫,手动键入电子邮件地址以发送电子邮件,将网址手动键入网络浏览器中以查看网页等等。这些任务常常是乏味并且耗时的。在各种设备上已提供节省用户时间的各种方法,如将电话号码分配到收藏夹列表和对网址加书签。随着技术发展并且随着便携式计算设备上所提供的特征和服务发展并扩充,提供快捷方式和其它节省时间的方法的方式正在改变。

[0003] 附图简述

[0004] 将参照附图描述根据本公开的各个实施方案,在附图中:

[0005] 图1A、图1B和图1C示出根据一个实施方案的示例性实现方式,其中拿着便携式计算设备的用户将文本提供给地图应用程序;

[0006] 图2A、图2B和图2C示出根据一个实施方案的示例性实现方式,其中拿着便携式计算设备的用户将电话号码提供给电话应用程序;

[0007] 图3A和图3B示出根据一个实施方案的示例性实现方式,其中拿着便携式计算设备的用户将文本提供给浏览器;

[0008] 图4示出根据各个实施方案的用于将文本提供给便携式计算设备上的应用程序的示例性过程;

[0009] 图5A和图5B示出根据各个实施方案可以利用的计算设备的前视图和后视图实例;

[0010] 图6示出根据各个实施方案可以利用的计算设备的组件级实例;以及

[0011] 图7示出可实现各个实施方案的示例性环境。

具体实施方式

[0012] 根据本公开的各个实施方案的系统和方法可以克服在用于将文本信息提供给应用程序、系统或服务的常规方法中经历的前述缺点和其它缺点中的一个或多个。具体来说,各个实施方案允许便携式计算设备执行如下任务,如获得包括文本的图像信息、在图像中定位文本、识别文本的类型或模式(例如,电子邮件、电话号码、URL等等)、确定与文本相关联的功能或应用程序(例如,呼叫号码、打开互联网浏览器等等),和/或执行与文本相关联的功能或应用程序。也可以使用在网络上可利用的至少一个资源来执行这些任务中的至少一些。

[0013] 在一个实施方案中,用户可将照相机对准包括文本的对象,以捕获所述对象的图像。照相机可与能够拍摄图像并处理图像的便携式计算设备集成以辨识、识别和/或隔离文本,以便将所述文本发送到应用程序或功能。应用程序或功能接着可利用文本来大致上实时地执行动作。例如,沿街行走的用户可看见包括电话号码的广告。用户可使用本文中所述的教义,通过以下操作来呼叫所述电话号码:例如将智能电话的照相机对准号码以将号码输入至电话应用程序中,而不需要手动输入号码。在一个实例中,设备可输入来自图像的号

码并且等待用户指示以开始呼叫,或在另一个实例中,可由设备自动进行呼叫。这个实施例可以扩展到在网络浏览器中输入网址或URL、用于打开新消息窗口的电子邮件地址、用于打开地图应用程序以显示路线的物理地址和各种其它可动作的数据类型。

[0014] 下文参照各个实施方案来呈现各种其它应用、过程和用途。

[0015] 图1示出示例性情况100,其展示用户将便携式计算设备102放在事件宣传单106上方。事件宣传单106含有事件位置的物理地址,并且在这个实施例中,用户正在寻找到所述位置的路线。为了获得路线,用户可将设备102的照相机对准宣传单含有地址的那一部分以捕获或获得所述地址的图像,如通过图1A的屏幕104可看见。在一个实施方案中,在获得图像和/或识别图像中具有指示文本的存在的性质的一个或多个部分后,设备102上的应用程序自动运行光学字符识别(OCR)算法来辨识宣传单的成像文本。在各个实施方案中,OCR算法可以包括机器视觉算法和利用各种技术的其它图像预处理算法。OCR算法运行以识别并分析图像中可包括文本串或字符串的一个或多个区域。分析任何已识别的串以进一步识别模式,模式将指示感兴趣的数据对象或类型的存在,所述数据对象或类型如电子邮件地址、URL/网址、电话号码等等。返回参照图1,已识别指示物理地址的存在的串。在这个实施例中,如图1B中所示,提示用户确认用户要执行与物理地址相关联的特定动作或运行预先确定的例程或工作流程的意图。在这个实施例中,预先确定的例程使得地图应用程序被打开并且到所述地址的路线被显示。在这个实施例中,用户首先可通过按压‘是’按钮或通过触摸屏上作出选择来确认意图,由此在便携式计算设备102的屏幕104上显示具有到所述事件的路线的地图,如图1C中所示。在另一个实施方案中,用户可通过经由对话界面对设备讲话(例如,通过说出“是”)而不是按压按钮或在触摸屏上作出选择来确认意图。在受到提示时确认意图的其它方式也是可能的,所述方式如通过使设备倾斜等等。

[0016] 图2示出示例性情况200,其展示用户将便携式计算设备202放在名片206上方,如图2A中所示。如上所述,设备202或与所述设备通信的服务在名片的所捕获图像中定位文本,识别文本的类型或模式(例如,电子邮件地址、电话号码、URL等等),确定与文本的类型相关联的功能或应用程序(例如,呼叫号码、打开互联网浏览器等等),并且给所述应用程序或功能发送所定位文本的至少一部分以执行与它相关联的操作。在这个实施例中,名片206含有多个感兴趣的数据对象,这些数据对象是电话号码、电子邮件和物理地址。在一个实例中,在识别出多种数据类型的情况下,用户可将优先级分配到各个功能或应用程序,以便一种数据类型比另一种数据类型居先。在另一实例中,设备202可利用各个可利用的功能或应用程序之间或之中的选择来提示用户。

[0017] 返回参照图2,识别出两个或更多个数据对象的设备可利用如图2B中所示的选项来提示用户启动通讯簿应用程序,以便将对象保存为联系人信息。或者,可经由显示器204利用如“按1拨号、按2查找路线、按3发电子邮件或按4保存联系人”的消息来提示用户。在一个实施例中,可将数据类型的文本自动填充到相应数据字段类型中。此外,在各个实施方案中,设备202可被编程、预加载或包括可下载的插件来辨识类似名片的对象并且使默认工作流程与所述对象关联,在一些实例中所述默认工作流程可以是将数据保存为联系人。在图2的实施例中,如图2B中所示,提示用户保存与名片206相关联的联系人。然而,在此状况下,用户想要呼叫名片上的电话号码并且选择‘否’。在一个实例中,设备202可利用各个选项再次提示用户,或者用户可能已经设置应用程序优先级来自动运行预先确定的例程或执行预

先确定的操作。现在参照图2C,用户已选择呼叫电话号码的选项,或者用户已将拨打电话号码分配为优先操作,并且展示设备202正在呼叫号码。

[0018] 图3示出示例性情况300,其展示用户将便携式计算设备302对准窗户308。在这个实施例中,窗户308上印刷有网址,如图3A中所示。根据各个实施方案,沿街行走的用户可将与设备302相关联的照相机对准或瞄准网址来作为导航到所述网址的快捷方式,而不是将所述地址手动键入到网络浏览器中。在这个实施例中,用户必须选择屏幕304上的图标306或提供另一个此类输入,以便手动捕获含有网址的窗户的图像。在先前实施例中,照相机可包括例如陀螺仪、加速计、光传感器或其组合,以便预测用户要捕获图像并开始例程或执行操作、与特定数据对象关联的意图。在这个实施例中,用户拿起设备302,将照相机瞄准窗户308,并且按压图标306以捕获网址的图像。接着,如上所述,设备或服务运行OCR算法以在图像中定位文本,识别文本的类型或模式(在这个实施例中是网址),启动浏览器应用程序并且将用户导航到饭店的网址,如图3B中所示。在一些实施方案中,可利用各种过程、算法或其它上下文线索来确认用户的意图。在饭店实施例中,如本文所述的系统可利用来自全球定位系统(GPS)传感器的位置数据来确认用户的位置。例如,所述系统可利用位置数据来执行对在用户100英尺(100ft)内的本地企业的网络搜索,以确认用户是否希望被导航到系统最初确定的网址。在这个实施例中,例如,如果用户站在大致上靠近Good Food咖啡馆(或在其前面)的位置,那么系统(或计算设备)可以相当确信地确定用户希望被导航到所述咖啡馆的菜单页或评论页。用于确认用户要执行功能或启动应用程序的意图的各种其它过程、方法或算法也是可能的。

[0019] 图4示出根据各个实施方案的用于将文本信息提供给计算设备上的应用程序或功能的示例性过程400。应当理解,在各个实施方案的范围内,对于本文所述的任何过程都可以存在按类似或替代次序执行或者并行地执行的另外步骤、更少步骤或替代步骤,除非另外说明。在这个实施例中,获得图像或图像信息(例如,视频流)402。在各个实施方案中,从便携式计算设备的照相机应用程序获得图像或图像信息。在一个实例中,将设备悬停在面对照相机的对象上将使照相机自动捕获至少一个图像或记录图像序列。在另一个实例中,为了捕获图像,照相机需要来自用户的输入,例如,如通过用户选择触摸屏上的图标。处理所获得的图像信息来定位具有文本串或字符串的性质的至少一个区域404。在一个实例中,在接收图像信息后自动开始所述处理。照相机也可处于连续图像捕获模式。然而,这可能消耗大量能量,因此用户可以选择使设备在图像捕获和处理模式下连续运行或何时将设备禁用的各个选项或情况。例如,可选择在电池耗尽到特定等级时自动关闭连续捕获和处理模式的设置。在这个实施例中,使用光学字符辨识算法分析文本串以辨识文本串中的文本406。OCR算法可包括机器视觉算法和其它图像预处理技术或算法。识别出对应于所辨识文本的文本模式(例如,电子邮件、电话号码、URL等等)408。在这个实施例中,确定与文本模式相关联的应用程序410,并且将所辨识文本自动提供给所述应用程序。

[0020] 在一个实施方案中,设备可能未在文本的一部分内辨识出感兴趣的数据对象。在此状况下,设备可被编程来:为用户提供各种其它选项,如执行网络搜索;或搜索插件,用来辨识未经辨识的模式。可提示用户选择文本的一部分以将文本复制/粘贴到电子邮件、记事本或各种其它应用程序中。

[0021] 在各个实施方案中,如以上参照图2所述,功能或应用程序可涉及:当文本模式指

示电话号码时拨打号码,或当文本模式指示电话号码和物理地址的存在时打开通讯簿来保存联系人信息。当文本模式指示网址时,可将应用程序导航到URL,如图3中所述。当文本模式指示物理地址时,应用程序可显示地图来展示地址,如图1中所述。功能或应用程序也可以执行各种语言翻译。应用程序也可以是具有求解数学方程的能力或与插件相关联的计算器,所述插件用于识别账单的特征以便在和朋友在饭店吃饭时以多种方式计算小费或拆分账单。在一个实施方案中,可提示用户在与特定数据对象相关联的第一动作与至少一个第二动作之间选择。返回参照名片实施例,与名片辨识相关联的预先确定的第一动作可以是在通讯簿应用程序中将名片上的信息保存为新联系人。第二动作可以是呼叫名片上的电话号码,并且第三动作可以是查找到所述地址的路线。

[0022] 在一个实施方案中,与许多实施方案的便携式计算设备通信的系统或服务可监测用户在一段时间内的活动,以调整应用程序所执行的动作。例如,如果用户总是选择呼叫名片上的号码并且从不选择保存联系人,那么设备可调整算法来自动呼叫或首先提示用户呼叫号码。这个实施例也可扩展到其它数据类型。在其它实施方案中,设备可预加载有各种行为预测算法或意图软件,用来在各种情况下预测用户的意图。例如,设备可包括面对用户的照相机、陀螺仪、加速计、光传感器或其组合。面对用户的照相机可使用面部辨识和手势跟踪算法来确定用户的注视方向。光传感器可确定处于用户手提包或口袋中的可能性。在这个实施例中,设备可进入节能模式并且禁止捕获或处理图像,直到传感器检测到足够量的光。陀螺仪或加速计可用来检测用户如何拿着设备并移动设备。在一些实例中,用户进行可预测的运动或以指示拍摄或捕获图像的特定方式拿着设备。在一个实施例中,记录移动的陀螺仪和加速计可以再次使设备处于节能模式。在这个实施例中,将很难捕获相当好的图像,并且设备可被编程来合理推断这些不是正常的图像捕获条件。在本公开的范围内,也可利用各种其它方法或预测用户意图。

[0023] 各个实施方案利用图像预处理算法和/或技术来使文本标准化以便进行识别和/或辨识。现实世界中的文本表现为许多颜色、形状、字体类型,所述字体类型可为倾斜的、模糊的、具有不同大小等等。可利用机器视觉图像处理技术,并且所述技术可包括阈值处理(将灰度图像转换成黑色和白色,或使用基于灰度值的分离)、分割、斑点提取、模式辨识、条形码和数据矩阵码读取、计量(测量对象尺寸)、定位、边缘检测、颜色分析、过滤(例如,形态学过滤)和模板匹配(对特定模式进行查找、匹配和/或计数)。在本公开的范围内,也可利用用于将文本处理成更容易由OCR辨识的文本的各种其它技术。在一些实施方案中,分析多个图像或视频流以减小眩光或噪声以及其它图像缺陷。

[0024] 图5示出可根据各个实施方案来使用的示例性便携式计算设备500的前视图和后视图。尽管展示了一种类型的便携式计算设备(例如,智能手机、电子书阅读器或平板计算机),但应理解,能够确定、处理并提供输入的各种其它类型电子设备也可根据本文所论述的各个实施方案加以使用。设备可包括例如笔记本计算机、个人数据助理、蜂窝电话、视频游戏控制台或控制器以及便携式媒体播放器等等。

[0025] 在这个实施例中,便携式计算设备500具有显示屏502(例如,液晶显示器(LCD)元件),所述显示屏可操作来将图像内容显示给设备的一个或多个用户或观看者。在至少一些实施方案中,显示屏使用例如电容式或电阻式触摸技术来提供基于触摸或轻扫的输入。此类显示元件可用来例如允许用户通过按压对应于按钮(如右边或左边的鼠标按钮、触摸点

等等)的图像的显示区域来提供输入。同样,设备在设备的其它区域上(如在设备的侧面或背面上)也可具有触敏或压敏材料510。尽管在至少一些实施方案中,用户可通过触摸或挤压此类材料来提供输入,但是在其它实施方案中,所述材料可用来通过图案化的表面相对于材料的移动来检测设备的运动。

[0026] 示例性便携式计算设备可以包括用于如常规图像和/或视频捕获等目的的一个或多个图像捕获元件。如本文其它地方所论述,图像捕获元件也可用于如确定运动并接收手势输入的目的。尽管在这个实施例中,便携式计算设备包括在设备“前面”的一个图像捕获元件504和在设备“后面”的一个图像捕获元件512,但是应当理解,图像捕获元件也(或者)可放在设备的侧面或拐角上,并且可能存在任何适合数量的类似或不同类型的捕获元件。每个图像捕获元件例如可以是照相机、电荷耦合设备(CCD)、运动检测传感器或红外传感器,或可利用另一种图像捕获技术。

[0027] 便携式计算设备也可包括至少一个麦克风506或能够捕获音频数据的其它音频捕获元件,如在某些实施方案中可用来确定位置变化或接收用户输入。在一些设备中,可能仅有一个麦克风,而在其它设备中,可能在设备的每个侧面和/或拐角上或在其它适合位置存在至少一个麦克风。

[0028] 在这个实施例中,设备500也可包括至少一个运动或位置确定元件508,这个元件可操作来提供如设备位置、方向、运动或取向的信息。这些元件可包括例如加速计、惯性传感器、电子陀螺仪、电子罗盘和GPS元件。各种类型的运动或取向变化可用来向设备提供可触发用于另一设备的至少一个控制信号的输入。示例性设备也包括至少一个通信机构514,如可以包括可操作来与一个或多个便携式计算设备通信的至少一个有线或无线组件。设备也包括电源系统516,如可以包括可操作来通过常规插上插头的方法或通过其它方法来再充电的电池,所述其它方式如通过接近充电板或其它此类设备来进行电容性充电。在各个实施方案的范围内,各种其它元件和/或组合是可能的。

[0029] 为了提供如参照图5所述的功能,图6示出便携式计算设备600的一组示例性基础组件,所述便携式计算设备如参照图5所述的设备500。在这个实施例中,设备包括至少一个处理器602,用来执行可存储在至少一个存储器设备或元件604中的指令。如本领域普通技术人员将显而易见,设备可包括许多类型的存储器、数据存储器或计算机可读存储介质,如用于由处理器602执行的程序指令的第一数据存储器,相同的或独立的存储器可用于图像或数据,可移动式存储器可用来与其它设备共享信息,等等。

[0030] 设备通常将会包括某种类型的显示元件606,如触摸屏、电子墨水(e-墨水)、有机发光二极管(OLED)或液晶显示器(LCD),但是如便携式媒体播放器的设备可能经由其它手段(如通过音频扬声器)来传送信息。如所论述的,在许多实施方案中,设备将会包括至少一个图像捕获元件608,如能够对设备附近的用户、人或对象进行成像的一个或多个照相机。在至少一些实施方案中,设备可使用图像信息来确定用户的手势或运动,这将允许用户在不实际接触和/或移动便携式设备的情况下通过便携式设备提供输入。图像捕获元件也可用来确定设备的周围环境,如本文所论述。图像捕获元件可包括任何适合的技术以在用户操作设备时捕获用户的图像,所述图像捕获元件如具有足够的分辨率、聚焦范围和可视区域的CCD图像捕获元件。

[0031] 在许多实施方案中,设备将包括至少一个音频元件610,如一个或多个音频扬声器

和/或麦克风。麦克风可用来促进语音支持功能,如语音辨识、数字记录等等。音频扬声器可执行音频输出。在一些实施方案中,音频扬声器可与设备分开存在。如上所述与许多实施方案相关的设备也可包括至少一个光学字符辨识器(OCR)元件612和与其相关联的各种图像预处理算法。

[0032] 设备可包括至少一个额外输入设备,所述至少一个额外输入设备能够从用户接收常规输入。这种常规输入可包括例如按钮、触摸板、触摸屏、方向盘、操纵杆、键盘、鼠标、轨迹球、小键盘或用户可用来向设备输入命令的任何其它此类设备或元件。在一些实施方案中,这些I/O设备甚至同样可以通过无线红外或蓝牙或其它链路来连接。然而,在一些实施方案中,此类设备可能根本不包括任何按钮,且可能仅通过视觉命令和音频命令的组合加以控制,使得用户可在不必与设备接触的情况下控制设备。

[0033] 示例性设备也包括一个或多个无线组件614,所述无线组件可操作来与特定无线信道的通信范围内的一个或多个便携式计算设备通信。无线信道可以是用来允许设备进行无线通信的任何适合信道,如蓝牙信道、蜂窝信道或Wi-Fi信道。应当理解,设备可具有本领域已知的一个或多个常规有线通信连接。示例性设备包括本领域已知的各种电源组件616,用来将电力提供给便携式计算设备,所述电源组件可包括与电源板或类似设备一起使用的电容性充电元件,如本文其它地方所论述。示例性设备也可包括至少一个触敏和/或压敏元件618,如围绕设备外壳的触敏材料、能够将基于挤压的输入提供给设备的至少一个区域等等。例如,在一些实施方案中,这种材料可用来确定运动,如设备或用户手指的运动,而在其它实施方案中,所述材料将用来提供特定输入或命令。

[0034] 在一些实施方案中,设备可包括激活和/或停用检测和/或命令模式的能力,如在从用户或应用程序接收到命令时或重试以确定音频输入或视频输入时等等。在一些实施方案中,设备可包括例如可用来激活一个或多个检测模式的红外检测器或运动传感器。例如,当没有用户在房间里时,设备可能不试图检测设备或与设备通信。例如,如果红外检测器(即,具有一个像素分辨率的检测器,所述检测器检测状态的变化)检测到用户进入房间,那么设备可激活检测或控制模式,使得设备在用户需要时可以就绪,但在用户不在附近时节省电力和资源。

[0035] 根据各个实施方案的计算设备可以包括光检测元件,所述光检测元件能够确定设备是否暴露于环境光或是否处于相对或完全黑暗中。此类元件可在许多方面有益。在某些常规设备中,使用光检测元件来确定用户何时将手机放到用户面部(使光检测元件大致上被屏蔽掉环境光),这可以触发一动作,如暂时关闭电话的显示元件(因为在将设备放到用户耳朵上时,用户看不见显示元件)。光检测元件可与来自其它元件的信息联合使用,以调整设备的功能。例如,如果设备不能够检测用户的观看位置并且用户没有拿着设备,但设备暴露于环境光,那么设备可能确定的是,它可能已被用户放下并且可能关闭显示元件并且禁用某个功能。如果设备不能够检测用户的观看位置,用户没有拿着设备并且设备不再暴露于环境光,那么设备可能确定的是,设备已被放在用户可能无法进入的袋子或其它隔室中并且因此可能关闭或禁用在其它情况下可能可利用的额外特征。在一些实施例中,为了激活设备的某个功能,用户必须看着设备、拿着设备或拿出设备放在光下。在其它实施方案中,设备可以包括可在不同模式下工作的显示元件,所述模式如反射模式(用于明亮情况)和发射模式(用于黑暗情况)。基于检测到的光,设备可以改变模式。在使用麦克风的情况

下,设备可禁用大致上与省电无关的其它特征。例如,设备可使用语音辨识来确定靠近设备的人(如儿童),并且可基于所述确定来禁用或启用如互联网访问或父母控制的特征。此外,设备可分析所记录的噪声以试图确定环境,如设备是否在车内或飞机上,并且所述确定可帮助决定启用/禁用哪些特征或基于其它输入采取哪些动作。如果使用语音辨识,那么单词可用作输入,无论是直接对设备说的单词还是通过对话间接拾取的单词。例如,如果设备确定它在车内、面对用户并且检测到如“饿”或“吃”的单词,那么设备可能打开显示元件并显示关于附近饭店的信息等等。出于隐私和其它此类目的,用户可以选择关闭语音记录和对话监测。

[0036] 在以上实施例的一些中,设备所采取的动作涉及停用某个功能以实现减少电力消耗的目的。然而,应当理解,动作可对应于可使用设备来调整类似或其它潜在问题的其它功能。例如,某些功能(如请求网页内容、在硬盘驱动器上搜索内容和打开各种应用程序)可能需要一定量的时间来完成。对于资源有限的设备或被大量使用的设备来说,同时发生的许多此类操作可使设备变慢或甚至锁定,这可以导致低效率,使用户体验降级并且可能使用更多电力。

[0037] 为了解决这些和其它此类问题中的至少一些,根据各个实施方案的方法也可利用如用户注视方向的信息来激活可能使用的资源,以便展开对处理能力、存储空间和其它此类资源的需要。

[0038] 在一些实施方案中,设备可具有足够的处理能力,并且成像元件和相关联的分析算法可以足够灵敏以仅仅基于所捕获图像来区分设备的运动、用户头部的运动、用户眼睛的运动和其它此类运动。在其它实施方案中,如在可能希望过程利用相当简单的成像元件和分析方法的情况下,可能希望包括能够确定设备的当前取向的至少一个取向确定元件。在一个实施例中,至少一个取向确定元件是至少一个单轴或多轴加速计,所述加速计能够检测如设备的三维位置和设备移动的幅度和方向以及振动、震动等等的因素。使用如加速计的元件来确定设备的取向或移动的方法在本领域中也是已知的并且本文将不进行详细论述。在各个实施方案的范围内,同样可使用用于检测取向和/或移动的其它元件来用作取向确定元件。当来自加速计或类似元件的输入与来自照相机的输入一起使用时,可更准确地解释相对移动,从而允许更精确的输入和/或复杂性更低的图像分析算法。

[0039] 例如,当使用计算设备的成像元件来检测设备和/或用户的运动时,计算设备可使用图像中的背景来确定移动。例如,如果用户以固定取向(例如,距离、角度等等)将设备拿向用户,并且用户改变对周围环境的取向,那么仅仅分析用户的图像将不能检测到设备取向的变化。相反,在一些实施方案中,计算设备仍然可通过辨识用户背后的背景影像的变化来检测设备的移动。因此,例如,如果对象(例如,窗户、图片、树、灌木、建筑、车等等)在图像中向左或向右移动,那么设备可确定的是,所述设备已改变取向,即使设备相对于用户的取向并未改变。在其它实施方案中,设备可检测到用户已相对于设备移动并且相应地调整。例如,如果用户相对于设备向左或向右倾斜他们的头部,那么显示元件上所呈现的内容可能同样地倾斜以使内容与用户保持相同取向。

[0040] 如所论述,可以根据所述的实施方案在各种环境中实现不同方法。例如,图7示出用于实现根据各个实施方案的各方面的环境700的实施例。如将了解,尽管出于解释目的使用基于网络的环境,但是可视情况使用不同环境来实现各个实施方案。系统包括电子客户

端设备702,所述电子客户端设备可包括可操作来在适合网络704上发送和接收请求、消息或信息并且将信息传送给设备用户的任何适合设备。此类客户端设备的实例包括个人计算机、手机、手持式消息传递设备、膝上计算机、机顶盒、个人数据助理、电子书阅读器等。网络可包括任何适合网络,其包括内部网、互联网、蜂窝网、局域网或任何其它此类网络或上述网络的组合。网络可以是“推送式”网络、“拉动式”网络或其组合。在“推送式”网络中,服务器中的一个或多个向客户端设备推送出数据。在“拉动式”网络中,服务器中的一个或多个在客户端设备请求数据之后将数据发送到客户端设备。此类系统所用的组件可以至少部分取决于所选网络和/或环境的类型。用于经由此类网络通信的协议和组件是众所周知的,因而本文不再详细论述。网络上的通信可以经由有线或无线连接及其组合来实现。在这个实施例中,网络包括互联网,因为环境包括用于接收请求并且响应于所述请求而提供内容的网络服务器706,然而对于其它网络来说,可以使用满足类似目的的替代设备,如本领域普通技术人员将显而易见。

[0041] 所示出的环境包括至少一个应用程序服务器708和数据存储器710。应当理解,可以存在可以链接起来或以其它方式来配置的若干应用程序服务器、层或其它元件、过程或组件,这些应用程序服务器、层或其它元件、过程或组件可交互来执行如从适合的数据存储器获取数据的任务。如本文所使用,术语“数据存储器”指能够存储、访问和检索数据的任何设备或设备组合,所述设备或设备组合可以包括在任何标准环境、分布式环境或群集式环境中任何组合和任何数量的数据服务器、数据库、数据存储设备和数据存储介质。应用程序服务器708可包括任何适合的硬件和软件,用来按需要与数据存储器810集成来执行用于客户端设备的一个或多个应用程序的各方面并且处置应用程序的大部分数据访问和业务逻辑。应用程序服务器与数据存储器协同提供访问控制服务,并且能够生成将要传送给用户的内容,如文本、图片、音频和/或视频,在这个实施例中,所述内容可由网络服务器706以HTML、XML或另一适合的结构化语言的形式提供给用户。所有请求和响应的处置以及客户端设备702与应用程序服务器708之间的内容递送可由网络服务器706来处置。应当理解,网络服务器和应用程序服务器不是必要的,且仅仅是示例性组件,因为本文所论述的结构化代码可在如本文其它地方所论述的任何适合的设备或主机上执行。

[0042] 数据存储器710可包括若干独立的数据表、数据库或其它数据存储机构和介质,用来存储与特定方面相关的数据。例如,所示出的数据存储器包括用于存储内容(例如,生成数据)712和用户信息716的机构,所述机构可用来为生成侧提供内容。还展示数据存储器包括用于存储日志或会话数据714的机构。应当理解,可能存在可能需要存储在数据存储器中的许多其它方面,如页面图像信息和访问权信息,所述方面可视情况存储在上文列出的机构中的任何机构中或存储在数据存储器710的中额外机构中。数据存储器710可通过与它相关联的逻辑来操作,以便从应用程序服务器708接收指令,并且响应于所述指令而获取数据、更新数据或以其它方式处理数据。在一个实施例中,用户可以针对某种类型的项目提交搜索请求。在此状况下,数据存储器可能访问用户信息来验证用户的身份,并且可访问目录详细信息以获取有关所述类型的项目的信息。接着可将信息如以网页上的结果列表的形式返回给用户,用户能够经由用户设备702上的浏览器来查看所述列表。可在浏览器的专用页面或窗口中查看感兴趣的特定项目的信息。

[0043] 每个服务器通常将包括操作系统,所述操作系统提供用于所述服务器的一般管理

和操作的可执行程序指令,且每个服务器通常将包括存储指令的计算机可读介质,所述指令在由服务器的处理器执行时可使服务器执行它的预期功能。操作系统的适合实现方式和服务器的一般功能是众所周知的或可商购的,并且易于由本领域普通技术人员实现,尤其是根据本文中的公开来实现。

[0044] 在一个实施方案中,环境是分布式计算环境,所述环境利用经由通信链路、使用一个或多个计算机网络或直接连接来互联的若干计算机系统和组件。然而,本领域普通技术人员应理解,这种系统可在具有比图7所示出的组件更少或更多个组件的系统中同样顺利地操作。因此,图7中的系统700的描绘本质上应视为说明性的,并且不限制本公开的范围。

[0045] 可进一步在广泛范围的操作环境中实现各个实施方案,所述环境在一些状况下可包括一个或多个用户计算机或可用来操作多个应用程序中的任何一个的计算设备。用户或客户端设备可包括多个通用个人计算机中的任何一个,如运行标准操作系统的台式计算机或膝上计算机,以及运行移动软件并且能够支持多个网络连接协议和消息传递协议的蜂窝设备、无线设备和手持设备。这种系统也可包括多个工作站,所述工作站运行各种可商购的操作系统和用于如开发和数据库管理的目的的其它已知应用程序中的任何一个。这些设备也可包括其它电子设备,如虚拟终端、精简型客户端、游戏系统和能够经由网络通信的其它设备。

[0046] 大多数实施方案利用本领域技术人员所熟悉的至少一个网络来支持使用各种可商购的协议(如TCP/IP、OSI、FTP、UPnP、NFS、CIFS和AppleTalk)中的任何一个的通信。例如,网络例如可为局域网、广域网、虚拟专用网、互联网、内部网、外联网、公共交换电话网、红外网络、无线网络以及上述网络的任何组合。

[0047] 在利用网络服务器的实施方案中,网络服务器可以运行各种服务器或中间层应用程序中的任何一个,所述服务器包括HTTP服务器、FTP服务器、CGI服务器、数据服务器、Java服务器和业务应用程序服务器。所述服务器也可以能够响应于来自用户设备的请求而执行程序或脚本,如通过执行一个或多个网络应用程序,所述网络应用程序可以实现为用任何编程语言(如Java[®]、C、C#或C++)或任何脚本语言(如Perl、Python或TCL以及上述脚本语言的组合)编写的一个或多个脚本或程序。服务器也可包括数据库服务器,其包括但不限于可商购自Oracle[®]、Microsoft[®]、Sybase[®]和IBM[®]的数据库服务器。

[0048] 环境可包括如上文所论述的各种数据存储介质以及其它存储器和存储介质。这些介质可驻留在各种位置,如在一个或多个计算机本地(和/或驻留在一个或多个计算机中)的存储介质上,或远离网络上的计算机中的任何一个或所有计算机。在一组特定实施方案中,信息可以驻留在本领域技术人员所熟悉的存储区域网络(SAN)中。类似地,可视情况在本地存储和/或远程存储用于执行归属于计算机、服务器或其它网络设备的功能的任何必要文件。在系统包括计算机化设备的情况下,每个此类设备可包括可经由总线来电耦合的硬件元件,所述元件包括例如至少一个中央处理器(CPU)、至少一个输入设备(例如,鼠标、键盘、控制器、触敏显示元件或小键盘)和至少一个输出设备(例如,显示设备、打印机或扬声器)。此类系统也可以包括一个或多个存储设备,如磁盘驱动器、光学存储设备和固态存储设备(如随机存取存储器(RAM)或只读存储器(ROM)),以及可移动式介质设备、存储卡、闪存卡等。

[0049] 此类设备也可包括计算机可读存储介质读取器、通信设备(例如调制解调器、网卡

(无线或有线)、红外通信设备等)和如上所述的工作存储器。计算机可读存储介质读取器可与计算机可读存储介质连接或被配置来接收计算机可读存储介质,计算机可读存储介质表示远程存储设备、本地存储设备、固定式存储设备和/或可移除式存储设备以及用于临时和/或更永久地含有、存储、传输和检索计算机可读信息的存储介质。系统和各种设备通常也将包括位于至少一个工作存储器设备内的多个软件应用程序、模块、服务或其它元件,其包括操作系统和应用程序,如客户端应用程序或网络浏览器。应当了解,替代性实施方案相比上文所描述的实施方案可以有众多变化。例如,也可使用定制硬件,和/或特定元件可以在硬件、软件(包括可移植的软件,如小程序)或两者中实现。此外,可以采用与如网络输入/输出设备的其它计算设备的连接。

[0050] 含有代码或部分代码的存储介质和计算机可读介质可包括本领域已知或已使用的任何适合介质,其包括存储介质和通信介质,如但不限于用于存储和/或传输信息(如计算机可读指令、数据结构、程序模块或其它数据)的任何方法或技术中所实现的易失性介质和非易失性介质、可移动式介质和不可移动式介质,包括RAM、ROM、EEPROM、闪存或其它存储器技术、CD-ROM、数字通用光盘(DVD)或其它光学存储器、磁盒、磁带、磁盘存储器或其它磁性存储设备,或可用来存储所需信息且可供系统设备访问的任何其它介质。基于本文所提供的公开和教义,本技术领域普通技术人员将了解实现各个实施方案的其它方式和/或方法。

[0051] 因此,应在说明性意义而不是限制性意义上理解说明书和附图。然而,将显而易见的是:在不脱离如在权利要求中阐述的本发明的更宽广精神和范围的情况下,可以对其做出各种修改和改变。

[0052] 条款

[0053] 1.一种计算机实施方法,其包括:

[0054] 在配置有可执行指令的一个或多个计算机系统的控制下,

[0055] 获得由便携式计算设备的照相机捕获的图像信息;

[0056] 处理所述图像信息以定位所述图像的具有文本串性质的至少一个区域;

[0057] 使用光学字符辨识算法来分析所述至少一个区域以辨识所述至少一个区域中的文本;

[0058] 识别对应于所述辨识的文本的文本模式;

[0059] 确定与所述文本模式相关联的应用程序;以及

[0060] 将所述辨识的文本自动提供给所述应用程序。

[0061] 2.如条款1所述的计算机实施方法,其中所述获得、处理、分析、识别、确定和提供在获得图像信息之后自动发生。

[0062] 3.如条款1所述的计算机实施式方法,其还包括:

[0063] 使所述应用程序执行以下操作中的至少一个:当所述文本模式指示电话号码时拨号;当所述文本模式指示电子邮件时打开电子邮件应用程序来撰写电子邮件;当所述文本模式指示网址时导航到URL;当所述文本模式指示至少一个电话号码和地址的存在时打开通讯簿来保存联系人信息;当所述文本模式指示物理地址时显示地图来展示地址;执行语言翻译;或求解数学方程式。

[0064] 4.如条款3所述的计算机实施方法,其还包括:

- [0065] 监测一段时间内的用户活动以调整由所述应用程序执行的动作。
- [0066] 5.如条款1所述的计算机实施方法,其还包括:
- [0067] 通过使用加速计或陀螺仪中的至少一个感测所述设备的移动或所述设备的取向中的至少一个来确认用户意图,以将所述辨识的文本提供给所述应用程序。
- [0068] 6.一种计算机实施方法,其包括:
- [0069] 在配置有可执行指令的一个或多个计算机系统的控制下,
- [0070] 识别与图像中的文本相关联的图像信息;
- [0071] 根据与文本相关联的所述图像信息来确定文本类型;
- [0072] 确定与所述文本类型相关联的至少一个功能;以及
- [0073] 允许计算设备使用所述文本来执行所述至少一个功能。
- [0074] 7.如条款6所述的计算机实施方法,其还包括:
- [0075] 提示用户确认要执行所述至少一个功能的意图。
- [0076] 8.如条款6所述的计算机实施方法,其还包括:
- [0077] 提示用户在与所述文本类型相关联的第一动作与至少一个第二动作之间选择。
- [0078] 9.如条款6所述的计算机实施方法,其还包括:
- [0079] 获得所述图像,所述图像是从所述计算设备的照相机获得。
- [0080] 10.如条款6所述的计算机实施方法,其中所述图像是以下之一:当所述计算设备的照相机聚焦在对象上时自动获得的图像,或响应于用户输入而获得的图像。
- [0081] 11.如条款6所述的计算机实施方法,其中所述至少一个功能包括以下操作中的至少一个:当所述文本类型指示电话号码时拨号;当所述文本类型指示电子邮件时打开电子邮件应用程序来撰写电子邮件;当所述文本类型指示网址时导航到URL;当所述文本类型指示至少一个电话号码和地址的存在时打开通讯簿来保存联系人信息;当所述文本类型指示物理地址时显示地图来展示地址;执行语言翻译;或求解数学方程式。
- [0082] 12.如条款9所述的计算机实施方法,其还包括:
- [0083] 将所述图像加以阈值处理以标准化与文本相关联的所述图像信息,以便顾及所述图像的模糊、噪声或照明变化中的至少一个。
- [0084] 13.如条款6所述的计算机实施方法,其中使用计算机视觉算法或光学字符辨识算法中的至少一个来从所述图像至少识别与文本相关联的所述图像信息。
- [0085] 14.如条款6所述的计算机实施方法,其中由与所述计算设备通信的远程服务器执行以下操作中的至少一个:识别与文本相关联的所述图像信息、确定所述文本类型或确定所述至少一个功能。
- [0086] 15.一种计算机实施方法,其包括:
- [0087] 在配置有可执行指令的一个或多个计算机系统的控制下,
- [0088] 从由便携式计算设备的照相机捕获的数字图像辨识字符串;
- [0089] 从所述字符串识别字符模式;以及
- [0090] 确定与所述字符模式相关联的应用程序;以及
- [0091] 将所述字符模式提供给所述应用程序。
- [0092] 16.如条款15所述的计算机实施方法,其中所述应用程序在被执行时执行以下操作中的至少一个:当所述字符模式指示电话号码时拨号;当所述字符模式指示网址时导航

到URL;当所述字符模式指示至少一个电话号码和地址的存在时打开通讯簿来保存联系人信息;当所述字符模式指示物理地址时显示地图来展示地址;执行语言翻译;或求解数学方程式。

[0093] 17.如条款15所述的计算机实施方法,其还包括:

[0094] 提示用户在与所述字符模式相关联的第一动作与至少一个第二动作之间选择。

[0095] 18.一种便携式计算设备,其包括:

[0096] 处理器;以及

[0097] 存储设备,其包括在由所述处理器执行时使所述处理器进行以下操作的指令:

[0098] 识别图像中的文本;

[0099] 从所述识别的文本识别文本模式;

[0100] 确定与所述文本模式相关联的至少一个功能;以及

[0101] 执行与所述文本模式相关联的所述至少一个功能。

[0102] 19.如条款18所述的便携式计算设备,其中在所述便携式计算设备上本地执行以下操作:识别所述文本、识别所述文本模式、确定所述至少一个功能以及执行所述至少一个功能。

[0103] 20.如条款18所述的便携式计算设备,其中由与所述便携式计算设备通信的远程服务器来执行以下操作中的至少一个:识别所述文本、识别所述文本模式或确定所述至少一个功能。

[0104] 21.如条款18所述的便携式计算设备,其还包括:

[0105] 照相机,所述照相机用于在其聚焦于对象上之后自动获得所述图像或响应于来自用户的输入而获得所述图像。

[0106] 22.一种非暂时性计算机可读存储介质,其存储指令,所述指令在由处理器执行时使所述处理器进行以下操作:

[0107] 识别图像中的文本;

[0108] 从所述识别的文本识别文本模式;

[0109] 确定与所述文本模式相关联的至少一个功能;以及

[0110] 使用所述文本模式执行所述至少一个功能。

[0111] 23.如条款22所述的非暂时性计算机可读存储介质,其中在所述便携式计算设备上本地执行以下操作中的至少一个:识别所述文本、识别所述文本模式、确定所述至少一个功能或执行所述至少一个功能。

[0112] 24.如条款22所述的非暂时性计算机可读存储介质,其存储指令,所述指令在由处理器执行时使所述处理器进一步进行以下操作:

[0113] 监测一段时间内的用户活动以调整由所述至少一个功能执行的动作。

[0114] 25.如条款22所述的非暂时性计算机可读存储介质,其中所述文本模式包括以下中至少一个:指示电话号码的模式、指示电子邮件地址的模式、指示URL的模式或指示街道地址的模式。

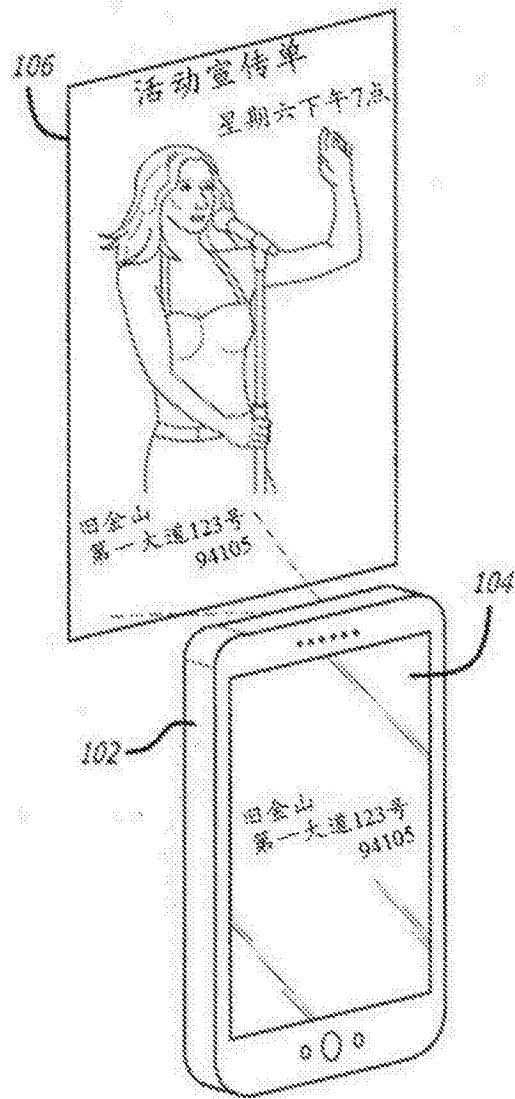


图1A

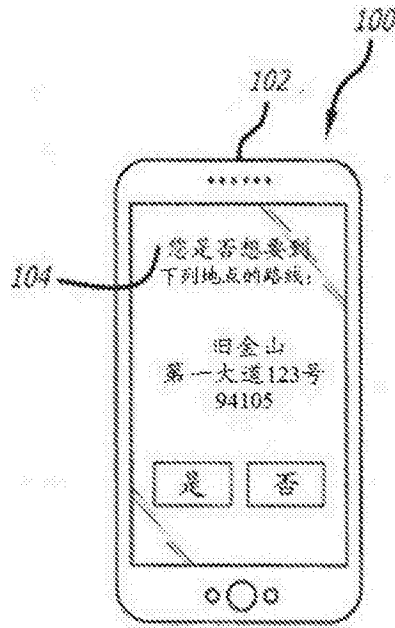


图1B

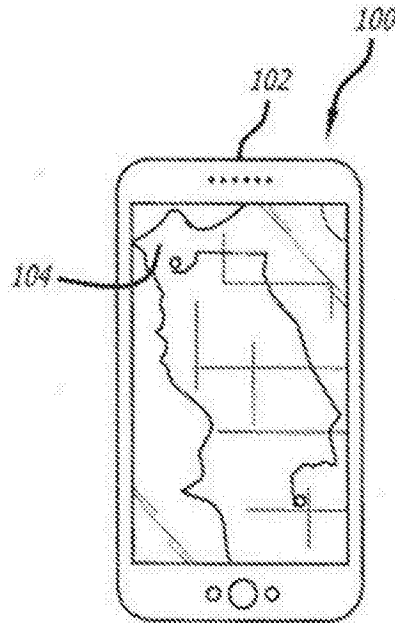


图1C

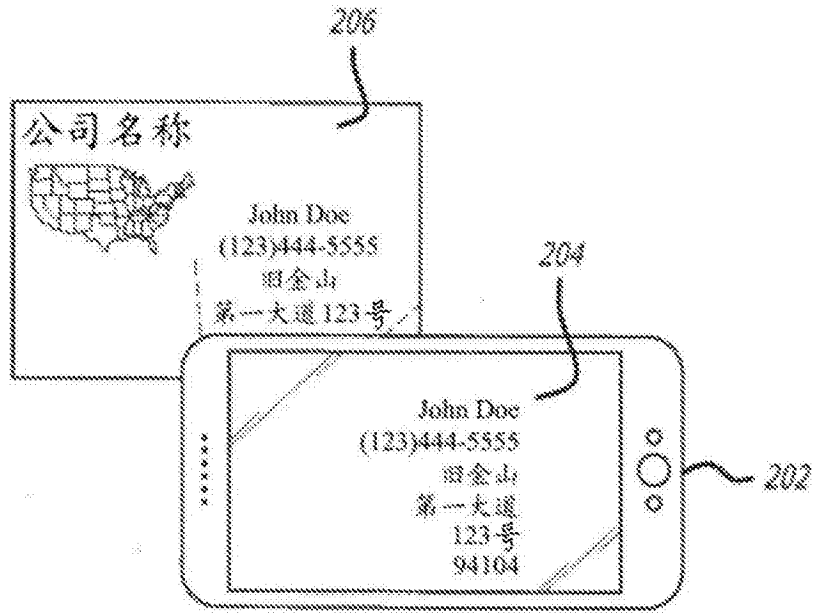


图2A

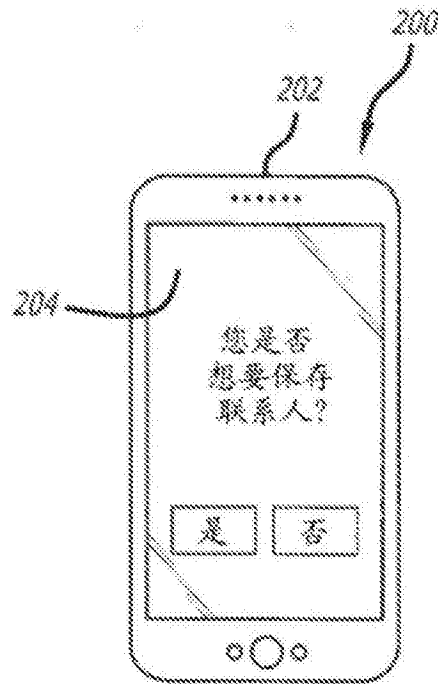


图2B

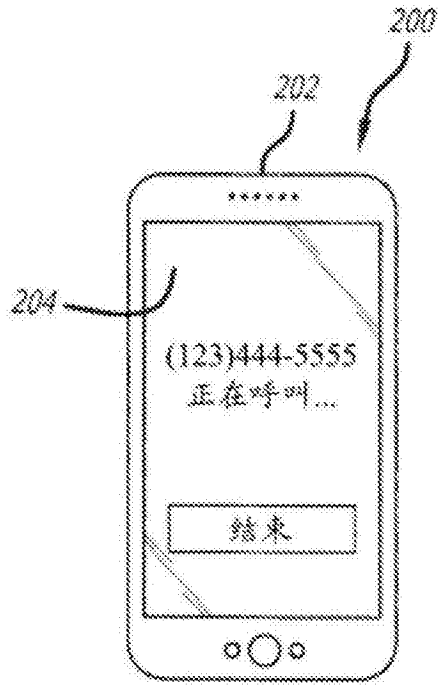


图2C

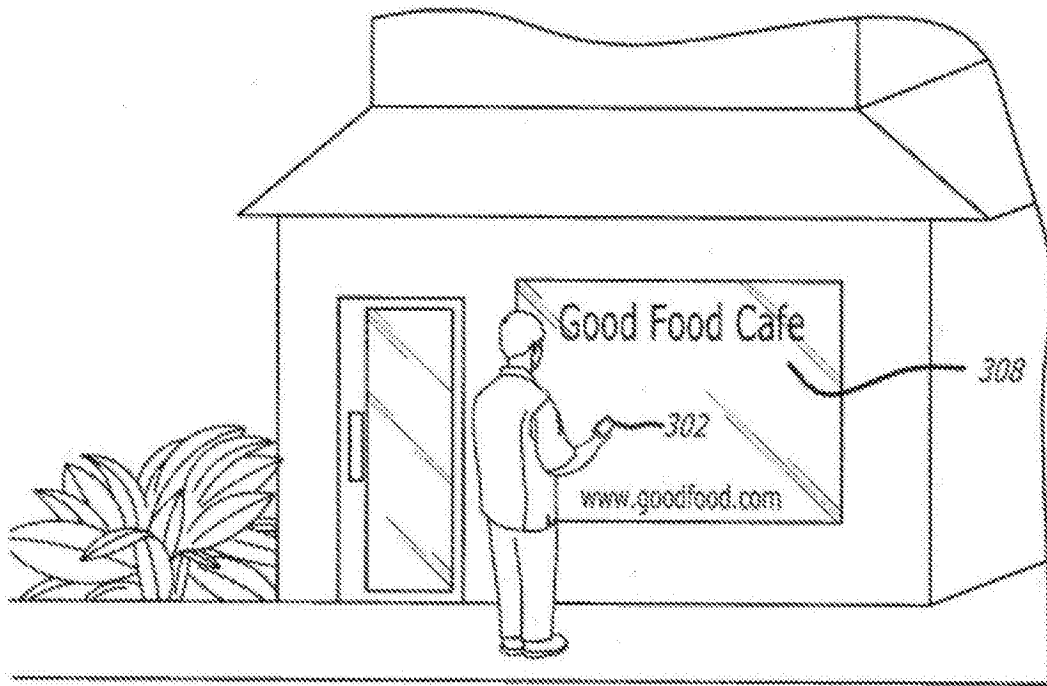


图3A

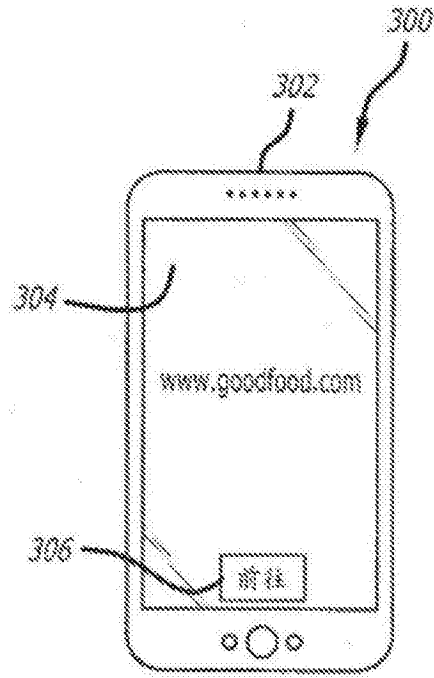


图3B

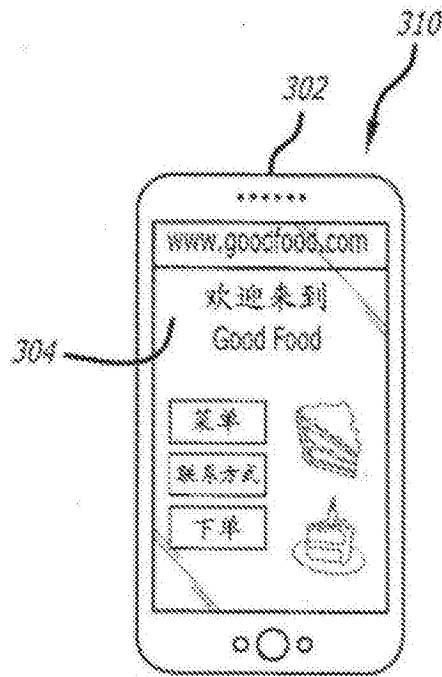


图3C

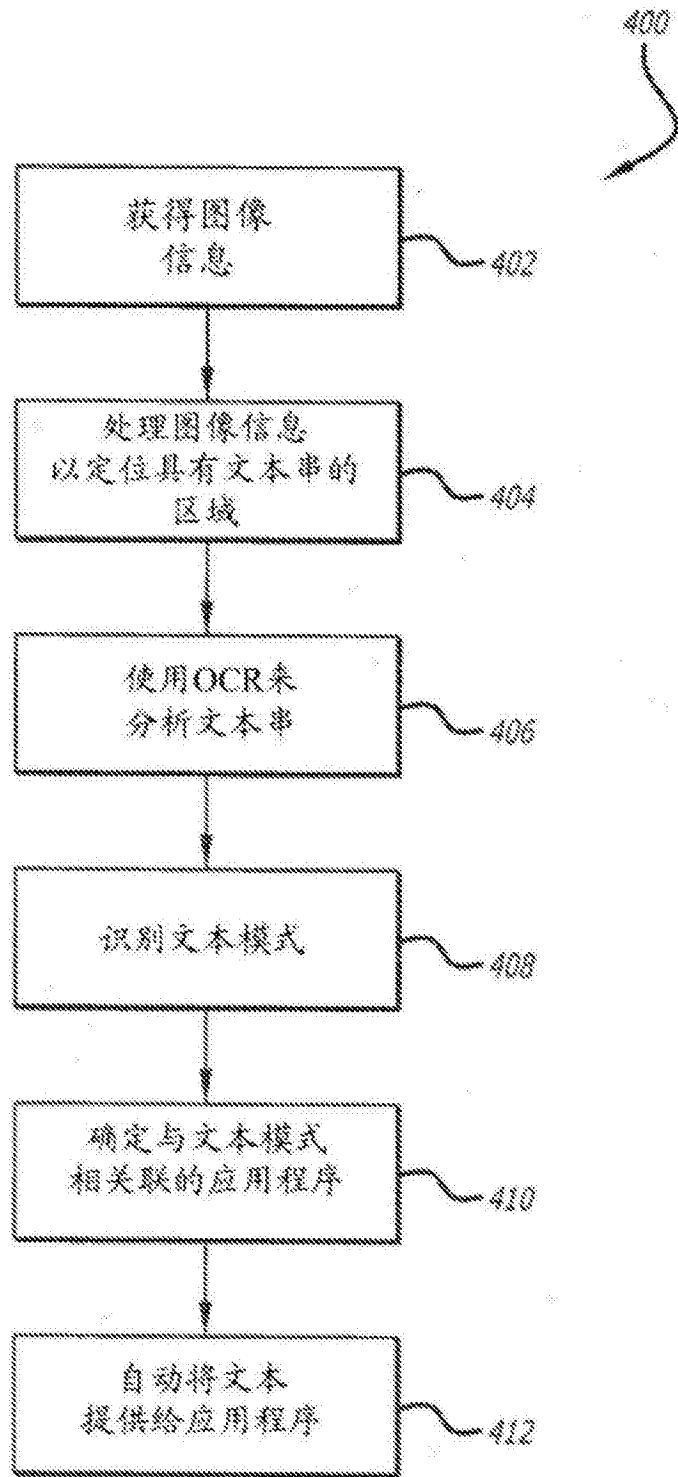


图4

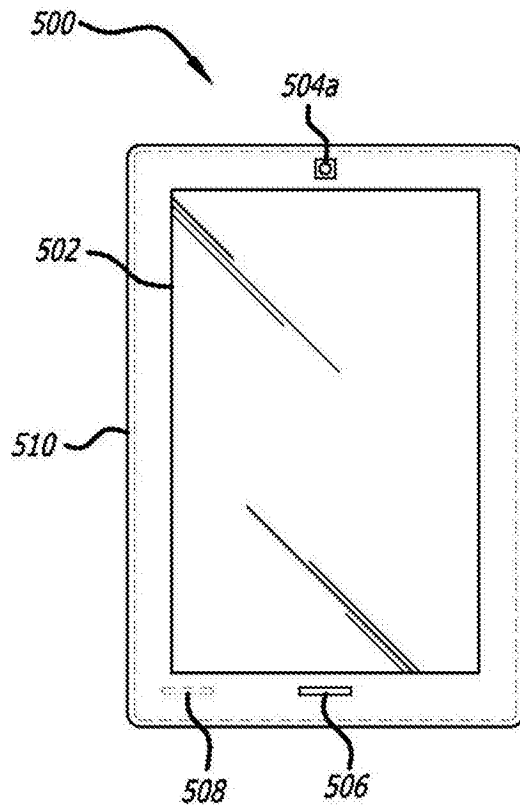


图5A

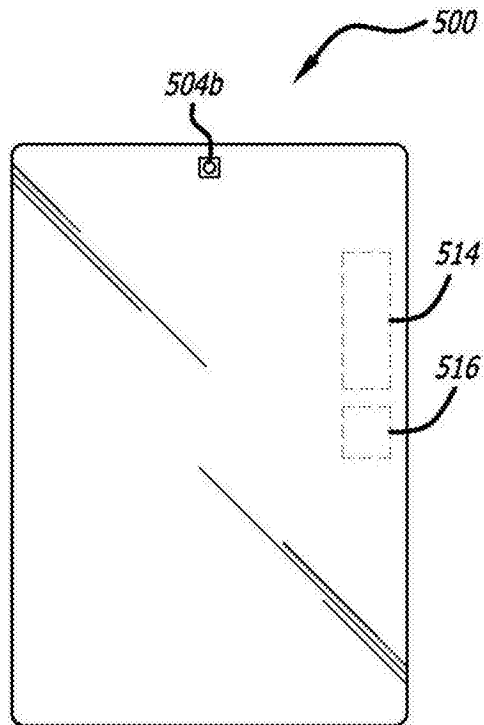


图5B

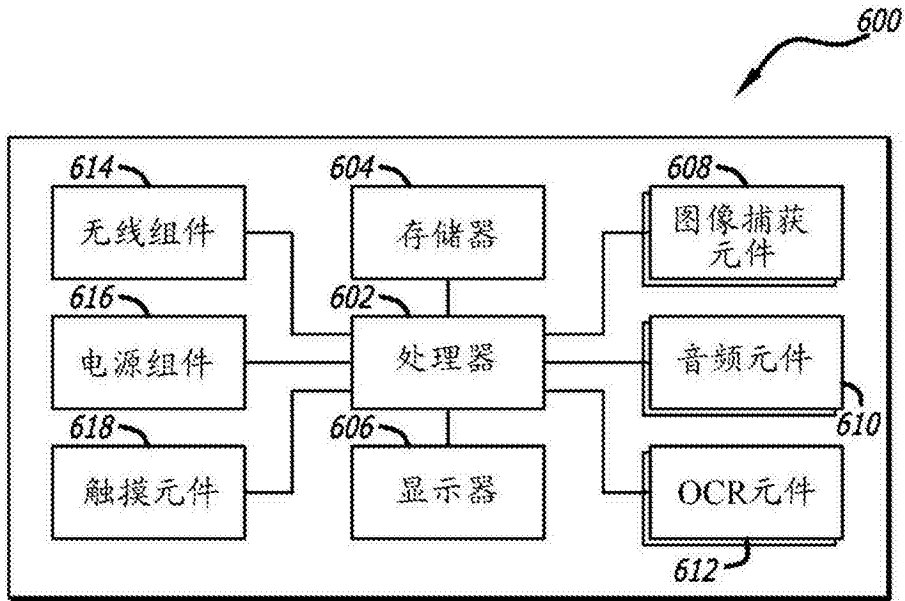


图6

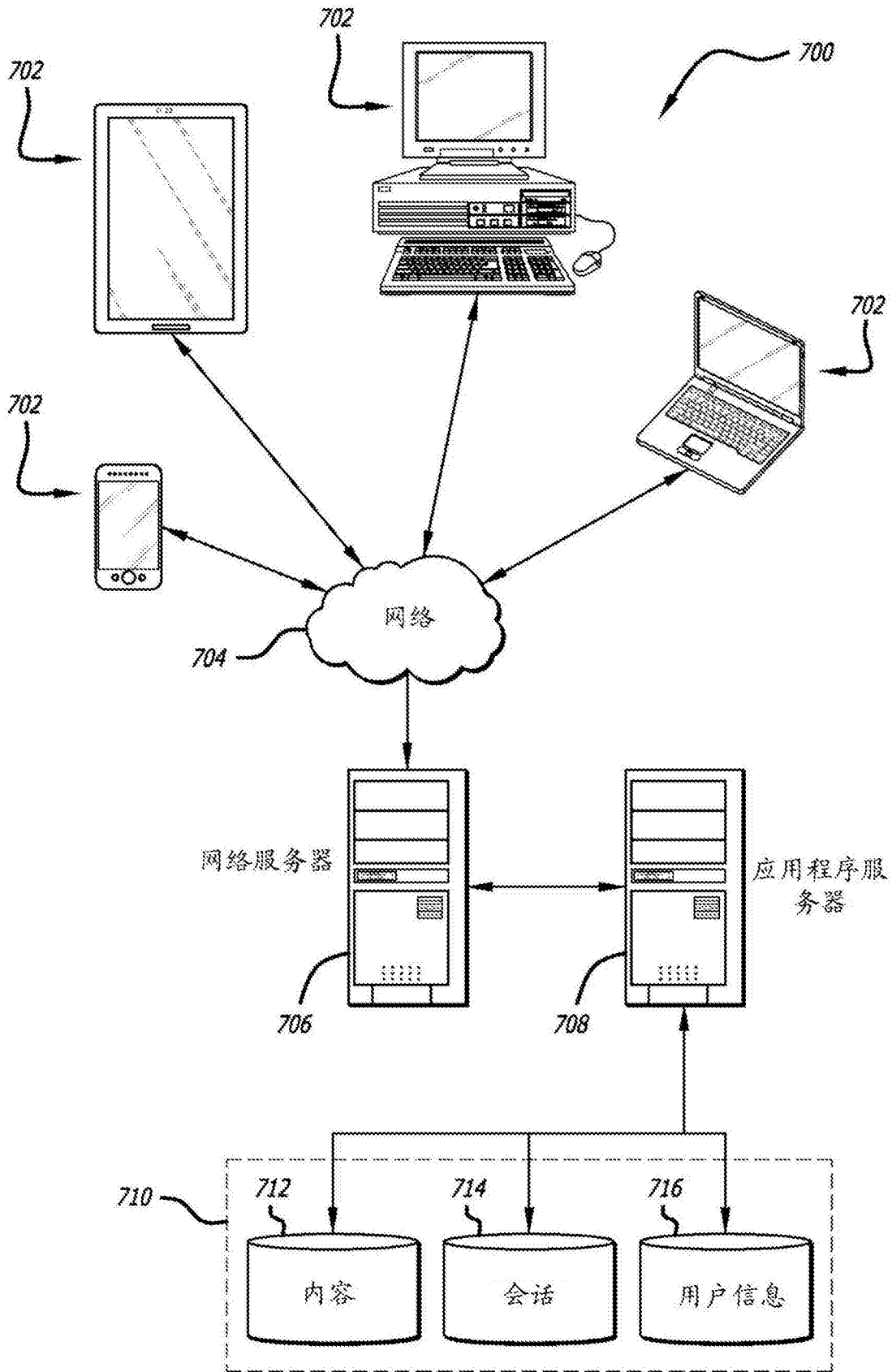


图7