



(12)发明专利申请

(10)申请公布号 CN 107077503 A

(43)申请公布日 2017.08.18

(21)申请号 201580060292.0

(22)申请日 2015.11.06

(30)优先权数据

62/076,281 2014.11.06 US

62/085,852 2014.12.01 US

14/700,319 2015.04.30 US

(85)PCT国际申请进入国家阶段日

2017.05.05

(86)PCT国际申请的申请数据

PCT/US2015/059355 2015.11.06

(87)PCT国际申请的公布数据

W02016/073803 EN 2016.05.12

(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 M·E·沃格尔 O·克里希纳

J·于 V·R·托塔 J·C·塞托

C·杨 A·Y·白

D·J·斯特潘内齐 E·基金吉尔

(74)专利代理机构 永新专利商标代理有限公司

72002

代理人 孟杰雄 王英

(51)Int.Cl.

G06F 17/30(2006.01)

G06F 9/44(2006.01)

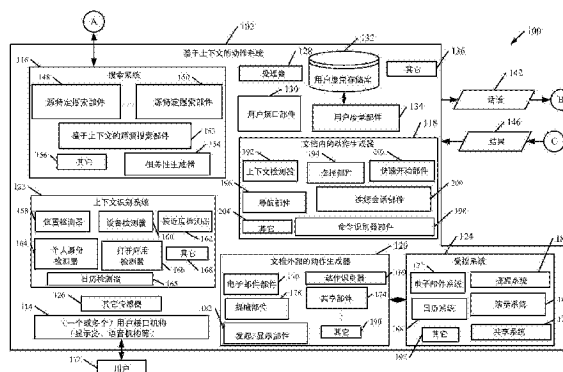
G10L 15/22(2006.01)

(54)发明名称

在语音用户接口中基于上下文的动作

(57)摘要

计算设备从用户处接收语音命令。所述设备基于语音命令输入来获得语言处理结果。所述结果包括意图和一组变元。所述设备还获得多种不同类型的上下文信息。基于所述意图、所述变元和所述上下文信息来识别动作，并且所述设备然后通过显示能够由所述用户启动以执行所述动作的用户可选输入机构来建议所述动作。所述设备还能够自动执行所述动作。



权利要求书2页 说明书19页 附图58页

1. 一种计算系统,包括:

用户接口机构,其接收话语,将所述话语发送到语言处理系统,并接收指示在所述话语中辨识出的意图和一组变元的语言处理结果;

上下文识别系统,其识别所述计算系统的当前上下文;

动作识别器,其基于所述当前上下文并基于在所述语言处理结果中的所述意图来识别动作;

跨源搜索部件,其基于所述当前上下文并基于在所述语言处理结果中的所述一组变元来搜索多个不同的信息源,以识别要对其执行所述动作的内容;以及

动作部件,其控制受控系统以对识别出的内容采取所述动作。

2. 根据权利要求1所述的计算系统,其中,所述话语隐含地识别所述内容,并且其中,所述跨源搜索部件使用所述变元作为搜索项来搜索所述多个不同的信息源,以识别多个不同的搜索结果,每个搜索结果识别不同的内容的项目。

3. 根据权利要求1所述的计算系统,还包括:

相关性生成器,其针对所述不同的搜索结果中的每个生成相关性量度以获得给定的搜索结果。

4. 根据权利要求3所述的计算系统,其中,所述相关性生成器基于所述当前上下文和所述变元,使用所述变元和所述当前上下文来另外消除模糊的搜索结果以获得所述给定的搜索结果。

5. 根据权利要求3所述的计算系统,其中,所述语言处理系统包括:

第一语言处理系统,其被设置在所述计算系统上,并包括言语辨识系统和自然语言理解系统,所述言语辨识系统生成所述话语的文本表示,并且所述自然语言理解系统识别在第一语言处理结果中的所述意图和所述变元,并且还基于所述当前上下文来另外消除模糊的变元。

6. 根据权利要求5所述的计算系统,其中,所述用户接口机构还将所述话语发送到远离所述计算系统的第二语言处理系统,所述第二语言处理系统生成第二语言处理结果,所述计算系统组合所述第一语言处理结果与所述第二语言处理结果以获得最终的语言处理结果。

7. 根据权利要求3所述的计算系统,其中,所述上下文识别系统包括:

个人身份检测器,其识别用户的个人身份;

接近度检测器,其识别在到所述计算系统的给定接近度内的其他人或事物;以及

打开应用检测器,其识别在所述计算系统上的打开应用。

8. 根据权利要求3所述的计算系统,其中,所述上下文识别系统包括:

设备检测器,其识别其上设置有所述计算系统的特定设备。

9. 根据权利要求3所述的计算系统,其中,所述上下文识别系统包括:

日历检测器,其识别用户的日历上的特定时间和项目。

10. 根据权利要求8所述的计算系统,其中,所述特定设备包括移动设备。

11. 根据权利要求8所述的计算系统,其中,所述上下文识别系统包括:

位置检测器,其识别在其上设置有所述计算系统的所述特定设备的位置。

12. 根据权利要求3所述的计算系统,其中,所述动作识别器将所述动作识别为以下中

的一个：发送、打开、安排会议、创建、附接、电子邮件发送、设置提醒、共享、呈现、发起和显示。

13. 一种计算机实施的方法，包括：

接收话语；

将所述话语发送到语言处理系统；

接收指示在所述话语中辨识出的意图和一组变元的语言处理结果；

检测所述计算系统的上下文；

分析所述上下文和在所述语言处理结果中的所述意图来识别要采取的动作；

基于所述上下文并基于在所述语言处理结果中的所述一组变元来搜索多个不同的信息源，以识别要对其执行所述动作的内容；并且

自动控制受控系统以对识别出的内容执行所述动作。

14. 一种移动设备，包括：

用户接口机构，其接收话语，将所述话语发送到语言处理系统，并接收指示在所述话语中辨识出的意图和一组变元的语言处理结果；

上下文识别系统，其识别所述计算系统的上下文；

动作识别器，其基于当前上下文并基于在所述语言处理结果中的所述意图来识别动作；

跨源搜索部件，其基于所述上下文并基于在所述语言处理结果中的所述一组变元来搜索多个不同的信息源，以识别要对其执行所述动作的内容；以及

动作部件，其控制受控系统以对识别出的内容采取所述动作。

15. 根据权利要求14所述的移动设备，其中，所述话语隐含地识别所述内容，并且其中，所述跨源搜索部件使用所述变元作为搜索项来搜索所述多个不同的信息源，以识别多个不同的搜索结果，每个搜索结果识别不同的内容的项目，并且还包括：

相关性生成器，其针对所述不同的搜索结果中的每个生成相关性量度以获得给定的搜索结果；并且所述相关性生成器基于所述上下文和所述变元，使用所述变元和所述当前上下文来另外消除模糊的搜索结果以获得所述给定的搜索结果。

在语音用户接口中基于上下文的动作

背景技术

[0001] 计算机系统被广泛使用。一些这种计算机系统被部署在设备中,所述设备包括台式计算机、膝上型计算机、平板计算机、智能电话、具有大屏的智能电话(例如,平板手机)以及多种其它移动设备和其它计算设备。

[0002] 对用户而言操作多个不同的计算设备是常见的。例如,用户可能在工作时具有台式计算机,而在家中具有另一计算机(例如,膝上型计算机)。用户还可以具有可以用于工作和个人使用两者的移动设备(例如,智能电话)。用户也可以具有用于专业工作和个人工作两者的其它设备。

[0003] 另外,可能存在用户可以生成并存储相关内容的多个不同位置。例如,用户可以在本地生成内容并将内容存储到设备(例如,智能电话、膝上型计算机、台式计算机等)。在客户端/服务器环境中,用户可以在服务器上生成信息并存储信息。用户还可以在存在多个不同服务器的环境中操作,并且用户可以在这些服务器中的任一个或全部上生成信息并存储信息。类似地,用户可以使用一个或多个基于云的服务,例如,基于云的存储、或其它基于云的服务。如果使用了基于云的服务,则用户也可以在云上生成内容并存储内容。

[0004] 为了定位相关内容,常常发现用户浏览许多不同的文件,以便定位其希望定位的特定文档或文件。类似地,当用户正在开会并希望与会议中的他人共享文档时,这通常需要用户自己定位文档,修改会议中所有的不同个人,识别所有这些个人的电子邮箱别名,编写具有所有上述信息的电子邮箱消息,然后附上文档并将其发送给这些个人。

[0005] 用于定位内容并与内容交互的这些当前过程是繁琐且耗时的。另外,尤其是在用诸如智能电话的相对小屏设备时,它们容易产生误差。

[0006] 上述讨论仅用于一般的背景信息,并不旨在用作辅助确定所要求保护的的主题的范围。

发明内容

[0007] 计算设备从用户处接收语音命令输入。所述设备基于语音命令输入来获得语言处理结果。所述结果包括意图和一组变元(argument)。所述设备还获得多种不同类型的上下文信息。基于所述意图、所述变元和所述上下文信息来识别动作,并且所述设备然后通过显示能够由所述用户启动以执行所述动作的用户可选输入机构来建议所述动作。所述设备也能够自动执行所述动作。

[0008] 提供该“发明内容”是为了引入在下文“具体实施方式”中进一步描述的简化形式的概念选择。该发明内容并不旨在识别所要求保护的的主题的关键特征或必要特征,也不旨在用作辅助确定所要求保护的的主题的范围。所要求保护的的主题不限于解决在背景中提及的任意或所有缺点的实施方式。

附图说明

[0009] 图1A和图1B(在本文统称为图1)示出了基于上下文的命令架构的一个范例的框

图。

[0010] 图2是言语处理结果的一个范例的框图。

[0011] 图3A-图3B(在本文统称为图3)示出了图示图1中示出的架构的基于上下文信息识别相关内容的操作的一个范例的流程图。

[0012] 图3C-图3H示出了用户接口显示的范例。

[0013] 图4是图示图1中示出的架构的对言语输入信号执行言语处理的操作的一个范例的流程图。

[0014] 图4A-图4H示出了用户接口显示的范例。

[0015] 图5是图示在图1中示出的架构在已经打开的文档内处理命令输入的操作的一个范例的流程图。

[0016] 图5A-图5V示出了用户接口显示的范例。

[0017] 图6A-图6B(在本文统称为图6)示出了在图1中示出的架构的执行快速开始过程的操作的一个范例的流程图。

[0018] 图6C-图6H示出了用户接口显示的各种范例。

[0019] 图7是示出在图1中示出的架构的在远程服务器环境中的一些部分的范例的框图。

[0020] 图8是示出在图1中图示的架构的在客户端/服务器环境中的一些部分的框图。

[0021] 图9是示出在图1中图示的架构的被设置在云计算架构中的部分的一个范例的框图。

[0022] 图10-图12示出了移动设备的各种范例。

[0023] 图13是能够被使用在先前附图中的各种架构中的计算系统的一个范例的框图。

具体实施方式

[0024] 图1A和图1B(在本文统称为图1)示出了基于上下文的命令架构100的一个范例的框图。架构100说明性地包括基于上下文的动作系统102,所述动作系统102与语言处理系统104和一组源106交互,所述源106说明性地包括多个不同的信息源108-110。信息源108-110本身能够是诸如最近使用文档的存储库、电子邮件消息、联系方式或人、日程信息(例如,日历信息)、在协作或演示系统中的文档或其它信息、设备信息、位置信息、以及多种其它信息的信息源。

[0025] 应当注意,在图1中示出的所有项目能够被定位在单个设备上,或者它们能够被分布在多个不同设备或位置之中。下面结合图7-图9来描述这个的多个不同范例。在图1中示出的范例中,基于上下文的动作系统102说明性地接收用户112输入的命令,并鉴于多个不同的上下文信息来解读这些命令。然后,其能够生成显示(例如,用户接口机构114)以用于与用户112交互。用户112说明性地与用户接口机构114交互,以便控制和操纵系统102。在一个范例中,机构114能够包括用户接口显示机构、由语音输入、按钮、按键等启动的语音机构、或多种其它机构。

[0026] 在图1中示出的范例中,系统102说明性地包括:搜索系统116,其用于搜索源108-110中的信息;动作生成器118,其在用户已经在文档内部时生成动作;动作生成器120,其在用户未在文档内交互时生成动作;上下文识别系统122;受控系统124;多个其它传感器126;处理器128;用户接口部件130;用户度量存储库132;用户度量部件134;并且系统102还能够

包括其它项目136。

[0027] 在图1中示出的架构100的范例中,语言处理系统104说明性地包括言语辨识系统138和自然语言理解系统140。当然,还能够预想到这些系统138、140中的任一个或两个也能够是基于上下文的动作系统104的部分,并且仅出于示例的原因而能够将其单独示出。

[0028] 言语辨识系统138能够是多种不同类型的言语辨识系统中的任一种,其采用语音输入(例如,响应于用户112向麦克风讲话而由麦克风生成的话语142)并生成话语142的上下文表示。例如,系统138能够使用声学模型和语言模型在隐马尔可夫过程中进行操作。其还能够使用基于神经网络的辨识或其它类型的辨识来进行操作。

[0029] 自然语言理解系统140说明性地接收话语142的文本表示,并生成对所述文本表示的解读。在一个范例中,系统140包括多个不同的语法143-144,所述语法143-144具有基于话语142中识别出的特定词语启动的规则。自然语言理解系统140通过识别话语142内的意图和各种变元来生成解读。意图能够是用户112期望的命令或者用户112试图执行的另一类型的动作或任务。变元包括能够识别将对其执行动作的特定内容的描述符信息。语言处理系统104说明性地返回文本表示、根据话语142辨识出的意图和变元作为语言处理结果146。将结合图2在下文更详细地描述该结果。

[0030] 在更详细描述架构100的总体操作之前,将首先提供架构100中的项目的一些的简略描述及其对应操作。搜索系统116说明性地包括多个不同的源特定搜索部件148-150。在一个范例中,每个部件148-150被配置为搜索特定信息源108-110。例如,当信息源108是文字处理文档的源时,搜索部件148然后可以被配置为搜索源108中的文字处理文档。当信息源110是包含电子邮件消息和接收者等的电子邮件系统时,搜索部件150然后可以被配置为搜索信息110的源,其包括电子邮件系统。当源108-110是日历或联系人目录时,能够使用对应的搜索部件148-250。

[0031] 在另一范例中,基于上下文的跨源搜索部件152能够自己或通过控制部件148-150(基于由系统104返回的结果中的意图和变元)识别将针对相关内容对其进行搜索的多个不同的信息源108-110,并能够搜索这些源108-110(或控制对应的部件148-160以对其进行搜索)以生成一组跨源搜索结果。相关性生成器154说明性地获得多种不同的上下文信息,并针对搜索系统116返回的内容的项目中的每个生成相关性分数。相关性分数能够用于生成内容的排序列表(或链接到所述内容)等。搜索系统116还能够包括其它项目156。

[0032] 上下文识别系统122说明性地包括多个不同的上下文检测器,所述多个不同的上下文检测器中的每个检测不同类型的上下文信息。例如,位置检测器158能够说明性地是定位系统(例如,GPS传感器或其它位置识别系统),其生成指示包括基于上下文的动作系统102(并由用户使用)的设备的当前位置的输出。设备检测器160获得关于其被定位的设备(或用户112正使用的)的上下文信息,例如,其是否是移动设备、台式计算机等。接近度检测器162检测设备(例如,用户112携带的设备)到其它项目(例如,其他个人、用户的工作位置、用户的家庭位置等)的接近度。个人识别检测器164说明性地检测用户的个人身份(假设用户选择进入或以其它方式同意该类型的检测)。日历检测器165能够检测天、周、月、会计期、年等的当前时间。其还能够说明性地识别用户日历上的过去、当前或将来的项目。这可以例如包括用户何时参加过会议、会议中的参与者、在会议中呈现的文档等。该信息能够被提供给相关性生成器154或者其能够被单独生成。打开应用检测器166检测例如以下的信息:在

用户112使用的设备或用户使用的其它设备上当前(或最近)打开的应用。上下文识别系统122能够包括多种其它检测器168,其还检测多种其它上下文信息。

[0033] 动作生成器120说明性地包括动作识别器169,所述动作识别器169基于上下文信息并基于意图和变元来识别将采取的动作。动作生成器120还包括能够基于用户112在结果146中表达的上下文信息、意图和变元而自己或通过受控系统124执行动作的一组部件。例如,话语142可以是“向我参加会议中的参与者发送当前规范”。在这种情况下,在话语142中表达的意图将包括多个步骤,例如识别“当前规范”,将其附到电子邮件(或者附到文档或以包括链接的其他方式附到文档),写电子邮件给用户112参加的“当前会议”的参与者,或许调整接收者访问控制权利或许可,然后发送所述电子邮件给参与者。因此,动作生成器120包括诸如电子邮件部件170的项目,所述电子邮件部件170能够在受控系统124中控制电子邮件系统172。共享部件174能够说明性地执行动作来与受控系统124中的各种共享系统176共享内容。例如,共享部件174可以将内容的项目邮件发送给另一用户的社交网站或工作网站,将内容的项目与合作网站上的工作组共享,或者将内容的项目邮件发送到博客上等。动作生成器120还能够包括这种项目作为提醒部件178,所述提醒部件178在受控系统124中控制提醒系统180。例如,话语142可以指示用户112想要在特定环境下被发送提醒。在这种情况下,意图是发送提醒,并且变元将识别发送提醒的特定环境。提醒部件178因此对此进行确定,并使用提醒系统180执行动作以发送期望的提醒。发起/显示部件182能够控制演示系统184或多种其它应用,使得能够基于在结果146中接收到的意图发起期望的应用并生成一个或多个期望的显示。在生成器120中的各种部件还能够控制诸如日历系统188的事物。例如,在动作生成器120中的项目能够插入用户112的约定,或向日历系统188中的其他会议被邀者发送会议请求。另外,当已知会议参与者时(例如,当用户112想要发送电子邮件给会议的所有参与者时),在动作生成器120中的项目能够与日历系统188交互以识别在给定会议中的参与者或被邀者,或者识别其它信息(例如,在给定会议中呈现的演示等),并使用所述信息执行动作。当然,动作生成器120能够包括多种其它部件190,所述多种其它部件190控制多种其它受控系统192,以便执行期望的动作。

[0034] 动作生成器118说明性地包括一组部件和检测器,所述一组部件和检测器能够基于上下文信息并基于各种用户输入来检测在已经打开的文档内的上下文并执行动作或者建议动作。例如,上下文检测器192能够检测打开的文档的上下文。当文档是文字处理文档时,能够检测当前被显示的特定页面。还可以检测多种其它元数据,例如,文档中的不同章节、文档上的不同作者或评论者、文档的各种修订日期、文档格式、访问控制许可等。选择部件194控制用户112的交互,所述交互指示用户希望选择在打开文档中的各种项目。导航部件196控制与文档的导航交互。命令识别器部件198识别具有显示给用户的对应用户输入机构的命令的子集,因此用户能够快速执行任意命令的子集中的任何项目。连续会话部件200说明性地在已经识别出的先前命令的上下文内处理各种结果146或来自用户112的其他输入。快速开始部件202控制在启动或初始授权文档期间显示文档的应用。当然,动作生成器118能够包括多种其它项目204,所述多种其它项目204也能够用于执行多种其它动作。

[0035] 图2示出了通过语言处理系统104生成的处理结果146的一个范例的框图。能够看出,结果146说明性地包括意图识别器206,所述意图识别器206识别在话语142中发现的意图。其还包括在话语142中辨识出的一组变元208-210。另外,结果146能够包括由言语辨识

系统138根据话语142生成的文本表示212。当前其也能够包括其它项目214。

[0036] 图3A和图3B(在本文统称为图3)图示了系统102(以及相关性生成器154)识别来自源106的内容的项目并为用户112显示它们的操作的一个范例。能够在任何期望时间确定内容的相关性,例如在用户第一次打开或发起系统102时,或者在用户请求系统102刷新被显示的文档的相关性时,或者例如自动基于上下文信息的改变。系统102因此首先确定是时候确定要向用户显现哪个相关内容。这由框216指示。当是时候识别相关内容时,相关性生成器154能够访问可以在相关性计算中使用的任何用户偏好或简档(profile)信息。这由框218指示。例如,用户可能偏好首先示出特定内容或者其它。其也访问上下文识别系统122并识别用户上下文。这由框220指示。例如,个人身份检测器164能够基于登录信息222来提供用户112的个人身份(再次,假设用户同意)。位置检测器158和设备检测器160能够用于确定用户112是在家224还是在工作226,或者在他处,以及用户正在使用何种设备。例如,用户可以在每次他或她在健身房时访问锻炼日志。即使文档可能在24小时中没有被访问,系统也能够感知用户物理存在于健身房中并在相关性计算中使用该信息。

[0037] 还能够识别其它用户上下文信息228(例如,使用模式等)。例如,其可以是靠近纳税时间。因此,单独基于新近信息或其它信息,不会将用户的先前纳税申报显现为相关的。然而,系统辨识每年的模式,此时,用户访问旧的纳税申报。该模式能够在相关性计算中使用。这只是一个范例。

[0038] 相关性生成器154还能够识别用户新近交互的内容的项目。这由框230指示。例如,信息源106中的一个是最近使用文档或内容的项目的列表。相关性生成器154因此能够从该源识别项目。能够识别前n个(或阈值数量)的内容的项目,如框232所指示的。能够按照新近次序识别内容的项目,如框234所指示的。能够识别用户最频繁访问的内容的项目,如框236所指示的;或者,能够识别由与用户相关的个人最频繁访问的内容的项目,如框238所指示的。例如,文档能够具有定义谁访问文档以及何时进行访问的元数据。当用户访问来自工作的信息并且用户属于一个或多个工作组时,或者当用户在一个或多个团队时或具有监督人时,该信息都能够从各种信息源108-110获得。相关性生成器154然后还能够识别其他用户各种最近访问的(或最频繁访问的)文档。这只是一个范例,并且还能够以其它方式识别最近交互的内容的项目,这由框240指示。

[0039] 相关性生成器154还能够识别交互类型。这由框242指示。例如,可以将一些类型的交互视为比其它的更重要。因此,生成器154能够识别在何处交互要编辑内容的项目244、要简单地打开内容的项目246、要对内容的项目进行评论248、要共享内容的项目250,或者交互是否是另一类型的交互252。

[0040] 生成器154还能够识别用户112或其他用户的不同类型的交互的频率。这由框254指示。

[0041] 另外,生成器154能够识别关于日历项目的交互是什么。这由框256指示。例如,如果用户总是访问在给定再发生会议中特定的内容的项目,并且该会议将要再次发生,则相关性生成器154可以确定在此时该特定的内容的项目尤其相关。因此,生成器154能够确定内容的项目是否在会议之前258、在会议期间260、在会议之后262交互,关于各种期限(例如,在特定类型期限之前立即)264、在用户的自由时间期间266、或相对其它日历项目268。

[0042] 生成器154还能够识别与用户112涉及的各个人,如框270所指示的。例如,识别系

统122能够识别用户112频繁发送消息的各个人。这由框272指示。能够识别用户112频繁在会议中的各个人,如框274所指示的。能够识别与用户112在组织关系表中密切相关的个人。这由框276指示。能够识别与用户112处于密切物理接近度的其他用户,如框278所指示的。能够识别与用户112个人相关的各个人(例如,在源106中来自社交网站等),如框280所指示的。能够识别用户112合著的各个人,如框282所指示的。能够识别用户是其成员的各种工作组或社交组或其它组284,并且还能够以其它方式识别其他人,这由框286指示。

[0043] 生成器154能够使用设备检测器160来识别用户112当前使用的特定设备。这由框288指示。例如,能够确定用户正在使用电话290、平板计算机292、膝上型计算机294、台式计算机296还是其它类型的设备298。还能够确定或检测设备使用的模式。例如,工作文档可能在台式计算机上更相关,这是因为用户通常在台式计算机上访问工作文档。存在在电话上更相关的其它文档(例如,一加仑汽油所行驶的里程跟踪器电子表格),因为所述文档主要在电话上使用。

[0044] 在评价各种内容的项目的相关性时,生成器154能够使用各种搜索部件148-152来识别存储内容的项目的内容存储位置。这由框300指示。例如,它们可以在本地被存储在本地数据存储库中,如框302所指示的。它们可能被存储在服务器上或以其它方式被远程地存储,如框304所指示的。它们可以被存储在特定的云服务306上或在其它位置308中。

[0045] 生成器154还能够使用其它传感器126来识别其它传感器输入310。例如,当其它传感器126包括加速计时,生成器154能够确定用户112当前是否在运动。这由框312指示。还能够确定用户112位于何处,如框314所指示的;或者其能够包括多种其它传感器输入316。

[0046] 相关性生成器154然后基于已经获得的所有信息来计算已经识别出的各种内容的项目的相关性。这由框318指示。例如,能够为关于内容的项目获得的每个上下文的内容或其它信息分配值。相关性计算可以简单地聚集所有这些值。这由框320指示。

[0047] 不同的信息项目也可以被认为比其它的更重要或更不重要。因此,能够基于其重要性对信息项目进行加权,如框322所指示的。加权还能够基于上下文信息动态改变。例如,如果用户正在工作,则可以用一种方式设置权重。如果用户在家中,则能够不同地设置权重。但这只是一个范例。存在能够基于所获得的相关性和上下文信息执行的多种其它类型的计算。这由框324指示。

[0048] 相关性生成器154然后能够生成相关内容的排序列表。这由框326指示。还能够存在用于每个不同类型的内容的多个不同类型的列表。这由框328指示。例如,特定类型的内容可以是演示系统中的“演示”(例如,幻灯片演示系统)。因此,对用户最相关的幻灯片演示(给出各种上下文信息)可以按相关性次序排序,并在显示器的一个部分中被显现给用户。然而,其它类型的内容可以包括电子邮件消息或电子表格文档。这些内容的项目可以基于其相关性进行排序,并且邮件消息的排序列表和电子表格的排序列表中的每个能够在用户接口显示器的另一单独可滚动的部分被显现给用户。这些仅是范例,并且还能够以其它方式生成相关内容的排序列表,这由框330指示。

[0049] 生成器154然后使用用户接口部件130来显现相关内容。这由框332指示。例如,其能够显现内容的前N个项目,如框334所指示的。其能够显现相关内容,并基于内容类型在显示器设备上将其划出。这由框336指示。能够显示内容本身,如框338所指示的;或者链接(例如,超链接)到内容,如框340所指示的。其还能够采用总结部件来生成相关内容的总结并显

示这些总结。这由框342指示。器还能够以其它方式显现相关内容,这由框344指示。

[0050] 系统102然后能够接收与内容的项目的用户交互。这由框346指示。然后,基于用户交互执行动作。这由框348指示,然后在下文更详细地描述用户交互的多个范例。

[0051] 然而,首先,将描述用户接口显示的多个范例。图3C和图3D示出了用户接口显示在电话350(或其它移动设备)上以及在平板计算机352(或膝上型计算机)上看起来如何的范例。能够看出,显示中的每个分别说明性地包括内容显示部分354和356,其显示被视为相关的内容或到内容的链接。

[0052] 显示还包括过滤部分358和360,其具有能够被启动以过滤所显示内容的一组用户可选择过滤机构。例如,当前显示的内容被过滤为“相关的”。然而,用户还能够启动“新近”用户输入机构362以基于对其访问有多近来过滤所显示的内容。用户能够启动“频繁”用户输入机构364以基于与其交互有多频繁来过滤所显示的内容。用户能够启动“喜好”用户输入机构366以基于什么被用户视为喜好的来过滤所显示的内容,并且用户能够启动“更新”用户输入机构368来更新相关内容。

[0053] 图3E和图3F类似于图3C和图3D,除了用户现在正在启动“更新”用户输入机构368。图3G和图3H示出了相关性生成器154已经接收到指示用户希望重新计算或更新各种内容的项目的相关性的输入。例如,这可以是用户已经切换了位置或现在处于不同的会议(这不同于用户上次具有的计算出的相关性)或其它。图3G和图3H示出了正在显示的相关内容已经被改变,这是因为用户已经改变了物理位置。

[0054] 例如,在一个范例中,相关性生成器154已经使用了位置检测器158来检测用户112现在在工作而不是在家中。因此,各种内容的项目的相关性已经改变,使得工作相关的内容现在更相关,并且在列表中比个人或家庭相关的内容显示得更高。

[0055] 图4是图示架构100的对在系统102处由用户112接收到的言语输入(或话语142)的处理的流程的一个范例的流程图。出于本讨论的目的,假设用户的设备当前正在显示相关内容(例如,以上结合图3A-图3H示出的)。

[0056] 然后,用户说明性地提供指示用户将要提供语音命令的启动输入。备选地,设备可能总是收听并被特定语音命令激活。作为一个范例,用户能够触摸触敏显示屏或以其他方式发起系统的言语辨识功能。然后,系统102说明性地接收自然语言用户语音输入(或话语142)。这由图4中的框380指示。例如,这能够在用户的设备上的麦克风处接收,或者以其它方式进行接收。系统102然后将话语142提供给语言处理系统104。应当注意,系统104或系统104的部分138和140能够在本地被定位在用户的设备上、在服务器处、在云中、在多个不同位置中等。例如,用户设备可能具有相对小且简单的言语辨识系统和自然语言理解系统,因此简单的话语在本地被处理。云或服务器能够具有相对较大和复杂的言语辨识系统和自然语言处理系统,因此复杂的话语被发送到云。在另一范例中,话语能够在两处进行处理,并且能够使用具有最高置信度分数的结果。

[0057] 言语辨识系统138生成话语的文本表示,如框382所指示的。一旦生成文本表示,自然语言理解系统140就能够基于文本表示来识别话语142中的意图386。其还识别来自文本表示的变元388。其还能够识别其它信息,这由框390指示。

[0058] 意图说明性地对应于用户希望执行的动作。例如,用户可以发出例如“与Joe共享该文档”的短语。在该情况下,自然语言理解系统140将识别词语“共享”作为用户希望系统

执行的动作或命令。作为范例，词语“共享”可以激活各种语法143-144中的一个或多个规则。这些规则可以激活其它规则，所述规则中的每个具有在规则开始执行之前匹配的一组变元。当文本表示匹配给定规则时，该规则的意图和该规则的各种变元被连同文本表示一起输出作为结果146的部分。

[0059] 用户接口部件130然后向用户显示文本表示，如框392所指示的。动作生成器120中的动作识别器169基于意图来识别要采取的动作，并且其还能够基于上下文信息和变元来识别要采取的动作。这由框394指示。例如，其还能够访问一组规则143-144。规则说明性地将意图映射到给定的动作或一组动作。这能够由框396指示。

[0060] 能够以层级或依赖结构嵌套或布置规则，以便完成要求多个不同命令或动作的任务。或者，它们能够以其它方式彼此参考或交互。这由框398指示。例如，将文档发送到会议的一组参与者要求多个不同的动作。首先，其要求定位文档。其还要求定位一组参与者。其然后要求将这些参与者作为接收者来生成电子邮件。然后，其要求将定位的文档附到电子邮件并发送所述电子邮件。能够通过基于话语中表达的意图激活的规则序列来识别所有这些动作或命令。还能够以其它方式识别动作，这由框400指示。

[0061] 搜索系统116然后识别将基于变元和上下文信息对其执行动作的内容的项目。这由框402指示。例如，当意图是要发送具有附件的电子邮件时，变元将说明性地识别电子邮件的附件和电子邮件的接收者。变元可以明确地识别附件和参与者，或者它们可以以暗示的方式识别附件和参与者（例如，将附件识别为“我今早修订的规范”或将接收者识别为“本会议的参与者”）。在该情况下，搜索系统116说明性地搜索各种信息源106，以识别要被发送的附件和电子邮件的接收者。当其搜索各种信息源106时，将说明性地识别最相关的搜索结果（例如，最可能的接收者、最可能的附件等）。这样，能够使用相关性生成器来生成针对每个识别出的项目的相关性分数，或者还能够以其它方式生成相关性或估计最可能的接收者和附件。考虑到相关性由框404指示。还能够以其它方式406识别对其执行动作的内容的项目。

[0062] 一旦动作生成器120识别出将采取的动作，并已经使用搜索系统116来识别需要执行该动作的内容的项目，则说明性地使用动作生成器120中的一个部件来执行该动作，如框408所指示的。这能够通过受控系统124中的一个中执行一个或多个动作来进行，例如通过在电子邮件系统中将文档附到电子邮件，向组或社交网站发送电子邮件并同时共享文档等。执行多个动作由框410指示。还能够以其它方式执行动作，这由框412指示。

[0063] 图4A-图4H示出了在基于语音输入执行命令时生成的用户接口显示的各种范例。例如，图4A和图4B示出了分别在电话350和平板计算机352上用户接口显示可以是什么样的范例。在图示的范例中，用户提供了“请帮我找上周我发送给Rob Yong的Contoso销售平台，我认为我上周从家里进行了更新”的语音输入。这样，用户首先说明性地启动用户输入机构（例如，机构414），以向系统102指示用户将要讲出话语。话语被捕捉并被发送给语言处理系统104，在此言语辨识系统138生成其文本表示。在对应的用户接口显示上进行显示，例如分别在框416或418处显示的。

[0064] 图4C和图4D类似于图4A和图4B，除了能够看出搜索系统116具有识别出的多个不同的相关项目，其与用户的话语相关。为此，动作生成器120首先接收在话语中的意图是“发现”文档的指示。然后使用搜索系统116来定位与用户的查询相关的文档。这样，其

使用变元,例如“Contoso销售平台”、“上周”、“Rob Young”和“家里”来识别相关文档。在图4C和图4D中示出的显示指示所显示的内容的项目是响应于用户查询发现的。它们说明性地以由相关性生成器154确定的或以其它方式确定的与用户查询的相关性次序进行显示。显示部件182然后生成在电话350和/或平板计算机352上的用户接口显示(无论用户是否正在使用),并显示相关内容。

[0065] 图4E和图4F类似于图4C和图4D,除了它们示出了用户现在正在启动“共享”用户输入机构420来共享文档。这例如能够通过用户在触敏屏上触摸用户输入机构来进行。共享部件174然后使用搜索系统116来识别关于该文档应当与哪些人或组共享的推荐。这能够基于上述用户的上下文信息、文档、其他用户或其它上下文信息来进行。共享部件174然后使用用户接口部件130来显示对用户的共享推荐。这由图4G-图4H图示。

[0066] 在图4G和图4H中示出的范例中,共享推荐422被显示给用户进行选择。在所示的范例中,基于用户112与被分析的各个人的接近度以及基于其它上下文信息来识别共享推荐。这只是一个范例。如果共享推荐422是准确的,则用户简单地启动发送用户输入机构424,并且与在共享推荐中识别出的人共享所选的内容的项目。

[0067] 因此,描述的话语中的一些“意图”表示在任何给定文档的上下文外部的命令或采取的动作。例如,命令能够包括诸如打开、下载、打印、共享、创建、发送、安排约定、删除、邮件发送、发现等。这些动作能够在文档上或跨多个不同数据源(例如,文字处理文档、演示、电子邮件等)且跨不同设备(例如,命令“在我的台式计算机上打开我昨天从Nicole接收到的文档”可以由用户在移动设备上输入,但是该文档在用户的台式计算机上打开)的其它目标上执行。

[0068] 还应当注意,当用户打开文档时,系统102能够在文档上执行动作。例如,系统能够在文档内执行动作(例如,导航动作、格式化动作、选择动作等)。图5是图示系统102进行这样的操作的流程图。首先假设用户打开了文档。这由框440指示。例如,文档可以是文字处理文档442、电子表格444、演示文档(例如,幻灯片演示)446、电子邮件消息448、图像450、或其它文档452。

[0069] 上下文检测器192说明性地检测文档上下文。这由框454指示。例如,可以检测文档滚动到的页面、文档的格式、或多种其它上下文信息。

[0070] 动作生成器118然后访问用户度量部件134来识别对应于用户112的学习度量。这由框456指示。通过范例,用户度量部件134能够说明性地生成指示用户112在使用用户正在使用的特定应用时、在使用用户正在使用的特定设备时、在使用基于上下文的动作系统102时、或在使用其它项目时发展或体验如何的用户度量。作为一个范例,用户度量部件可以跟踪用户已经使用了给定应用中格式化特征多少次。其可以跟踪用户已经使用了何种类型的特征、用户已经使用系统104的频率、或者多种其它事物。这能够充当指示用户体验的度量(或能够基于该信息计算度量)。由于用户在使用给定应用或系统102时获得经验,因此系统可能不需要为用户生成许多帮助提示。然而,如果用户是新用户,则系统可能希望生成更多的帮助提示或更频繁地显示帮助提示。在任何情况下,生成器118访问针对当前用户112的用户学习度量,作为用户在与文档交互时如何体验或获知的指示。

[0071] 然后,命令识别器部件198基于上下文和用户学习度量来识别命令的上下文特定子集和显示频率。这由框458指示。例如,用户可以简单地查看文档。在该情况下,命令识别

器部件198可以预测用户可能希望执行命令或动作的第一子集中的一个。在另一范例中,用户可能已经选择了文档中的词语、短语、段落、对象或图像。在该情况下,命令识别器部件198说明性地预测用户可能希望执行命令的不同子集中的任一个。

[0072] 在识别命令的子集时,部件198可以访问一组命令预测规则,或者预测用户可能希望使用的各种命令的命令预测部件。在任一情况下,部件198基于应用的上下文、用户的上下文以及用户的体验水平(由学习度量表示)来识别用户可能希望采取的命令的子集。部件198还可以考虑用户特定使用或偏好数据。这由框460指示。部件198还将修改基于度量、上下文数据等的改变识别出的命令的子集。这由框462指示。能够以其它方式识别命令子集,如框464所指示的。

[0073] 命令识别器部件198然后使用用户接口部件130来显示用于识别出的命令子集的机构。它们以基于用户学习度量的频率显示。用户然后能够与显示机构交互来执行命令子集。显示用于用户交互的命令的识别出的子集由框466指示。范例可以是有用的。

[0074] 图5A-图5D示出了用户接口显示和命令显示机构的各种范例。在图5A中,例如,能够看出,用户在显示器468中简单地查看文档。因此,假定用户正在简单地查看文档,命令识别器部件198识别用户希望执行的命令的第一集合。命令子集的一个范例包括“转到”命令、“发现”命令、格式化“主题”命令等。然后生成一组显示元件470,其中,一个显示元件对应于识别出的命令子集中的每个。显示元件470说明性地是用户可启动显示元件,其能够由用户启动(例如,通过轻敲它们)来执行对应的动作或命令。因此,例如,如果用户启动“转到”显示元件,则导航部件196将用户导航到文档的识别出的部分。如果用户启动“主题”显示元件,则动作生成器118将用户导航通过格式化用户体验,这允许用户根据各种主题格式化文档。这些仅是范例。

[0075] 图5B示出了用户已经选择了词语“Redmond, Washington”。因此,命令识别器部件198已经识别出命令的第二子集并显示对应于这些命令的一组显示元件472。由于用户已经选择了一组词语,因此命令识别器部件198已经预测出很可能想要执行与当用户没有选择任何词语(如图5A所示)时不同的命令。因此,显示元件472对应于与显示元件470不同的命令。

[0076] 在图5C中能够看出,用户现在已经选择了整个段落。因此,命令识别器部件198预测出用户可能希望执行命令的又一子集中的任一个。部件198因此显示对应于这些命令的显示元件474的另一集合。

[0077] 图5D示出了用户现在已经选择了文档中的图像。基于该事实,命令识别器部件198预测用户很可能要执行命令的又一不同子集中的一个,并且显示对应于这些命令的显示元件476的不同集合。用户能够与显示元件中的一个交互,并且动作生成器118说明性地使用其部件中的一个来执行对应的动作或命令。这由图5中的框478和480指示。

[0078] 当用户这样做时,用户然后可以提供进一步的交互输入。在该情况下,在一个范例中,动作生成器118中的连续会话部件200解读在先前命令的上下文中的后续输入。这由框482指示。当然,命令能够是语音命令484或其它命令486。图5E-图5K图示了系统如何鉴于先前的命令来解读命令的一个范例。

[0079] 在图5E中能够看出用户已经选择了被显示文档中所有的文本。然后,用户启动用户输入机构484,指示用户将要提供语音输入。然后,用户讲出输入“使行距变宽”。图5F示出

了在488处显示的话语的文本表示。作为响应,语言处理系统104将意图识别为改变行距,并将其与结果146一起提供到连续会话部件200。命令识别器部件198用此作为文本信息,并识别涉及改变行距的命令的子集,且显示对应于这些命令的显示元件486的集合。

[0080] 用户敲击显示元件486,指示用户希望使得行距变宽。然后,动作生成器118使得行距变宽,如图5G所示。然后用户提供另一语音输入“更宽”。这被提供到系统104,在此将词语返回为结果146。部件200基于先前增加行距的指令识别这是连续指令以进一步增加行距。因此,其控制动作生成器118将行距增加得更宽。这在图5H中示出。图5I示出了命令识别器部件198维持与行距相关的两个显示元件490。这是因为已经解读为用户仍试图修改行距。因此,如图5J所示,用户能够简单地敲击这些用户输入机构中的一个,并且动作生成器118通过导航用户通过允许用户将行距手动设置为给定水平的另一用户体验来执行对应的动作,如图5K所指示的。

[0081] 返回图5,用户可以提供更多的交互,如框492所指示的。如果情况如此,则处理返回框454,在此再次检测文档上下文,访问用户学习度量等。

[0082] 在框478处,用户可能不曾启动命令识别器部件198所显示的显示元件中的一个(如图5A中的470)。在该情况下,系统检测用户是否提供在文档内的任何其它交互输入。这由框494指示。如果否,则对应于识别出的命令的显示元件最终从显示移除(例如,它们在特定时间量之后消失)。这由框496指示。例如,如果用户简单地查看文档,如图5A所示,并且用户没有启动任何显示元件470,则在预定时间段之后,它们被移除。预定时间段可以基于该特定用户的用户学习度量而变化。例如,如果用户是系统的非常有经验的用户,则系统将确定不需要长时间显示显示元件,这是因为用户已经知道如何使用它们。然而,如果用户是相对经验较少的用户或新用户,则可以在较长时间段内显示显示元件,使得用户能够复查它们、尝试它们、考虑各种选项等。另外,应当注意,在一个范例中,用户能够在任何时间访问命令的整个集合。例如,用户能够通过导航到命令菜单这样做,因此命令的全部集合可以被隐藏或隐匿,但是不会不可用。

[0083] 如果用户执行另一类型的交互(除了启动显示元件中的一个之外),则动作生成器118执行任何对应动作。这由框498指示。然后处理再次返回框454。这样,用户交互将是文档上下文的部分,其被命令识别器部件198在预测用户将可能使用的命令并显示对应的显示元件时进行考虑。图5L-图5O图示了这个的一个范例。

[0084] 图5L类似于图5A,除了能够看出用户现在提供触摸手势以滚动正被显示的文档。假定用户交互是上下文信息,那么命令识别器部件198预测用户很可能正在试图发现文档中的特定位置。因此,命令识别器部件198显示对应于允许用户这样做的命令的显示元件501的集合。例如,显示元件中的一个“转到页面”显示元件。当用户启动这个时,如图5N所指示的,动作生成器118显示菜单503,其允许用户选择用户想要滚动到文档的特定页面。导航部件196然后将用户导航到文档的该部分。再次,这仅是一个范例,并且存在多种其它用户动作或交互,其能够用作部件198的上下文信息以预测用户希望使用的可能命令。

[0085] 图5P-图5V示出了用户选择文档内不同项目的用户交互的一个范例。在相对小屏的设备上选择项目是非常繁琐的。用户通常不注意选择了错误的项目,或者在只想选一个时不注意地选择了多个项目。另外,用户通常需要滚动通过长的文档来选择同一类型的各种项目。图5P示出了用户简单地查看内有图片的文档。用户启动用户输入机构535,其向系

统指示用户将要提供语音命令。图5Q示出了用户已经提供了语音命令“我想格式化所有我的图片”。图5R示出了系统现在正在显示该输入的文本表示。图5S示出了选择部件194已经接收到指示意图是用户希望选择打开文档中所有图片并对其格式化的结果146。因此,选择部件194自动定位并选择所显示文档中的所有图片。命令识别器部件198还识别涉及格式化图片的命令的子集,并显示对应于所述命令的显示元件。图5S示出了用户正在敲击这些显示元件中的一个。图5T示出了同样的情形。图5U示出了用户已经选择了要替换所选图片的不同图片,并且图5V示出了已经插入了新选的图片。

[0086] 图6A和图6B(在本文统称为图6)示出了图示快速开始部件202允许用户快速开始编写文档的操作的一个范例的流程图。人们相信,一些用户难以开始编写文档,这是因为他们不确定如何建立格式等。因此,部件202帮助用户这样做。首先,接收打开新文档的用户输入。这由框505指示。再次,文档能够是文字处理文档507、电子表格文档509、演示文档511、或多种其它文档513。

[0087] 动作生成器120使用发起/显示部件182来发起相关应用并显示文档的空白画布。这由框515指示。动作生成器118然后接收请求类似于先前文档的格式化的用户输入。这由框517指示。再次,这能够是语音命令519或另一自然语言命令521。然后,系统对输入执行自然语言理解以识别文档、识别其格式化,并将所述格式化应用于本文档。

[0088] 对输入执行自然语言理解由框519指示。在一个范例中,快速开始部件202显示多个不同的相关文档,并允许用户从选择将从其识别格式化并将其应用到新文档的特定文档。确定用户是否具有多于一个选项由框521指示。能够基于用户偏好、基于管理建立、或以其它方式进行该确定。

[0089] 当没有向用户呈现多个不同选项时,取回在用户的自然语言输入中由用户识别出的文档,并且识别出格式设置。这由框523指示。处理然后跳到框525,在此将识别出的格式化应用给所显示的新文档。

[0090] 然而,如果在框521处将提供多个不同选项(或者如果自然语言处理结果是模糊的且识别出多个文档),则系统识别先前的文档和其它类似文档,如框527所指示的。文档能够是类似的,这是因为它们是新近的文档529,它们能够是类似的文档类型531,或者它们能够以其它方式类似533。

[0091] 对于每个识别出的文档,快速开始部件202生成并显示对应的显示元件。该显示元件说明性地识别文档并显示文档的一些内容,使得用户能够知道应用于基础文档的格式化。生成和显示对应于识别出的文档的可选显示元件由框535指示。

[0092] 然后,部件202接收这些显示元件中的一个的用户选择,如框537所指示的。识别对应文档的格式化,如框539所指示的。这能够通过如框541所指示的解析对应文档、通过如框543所指示的访问对应于文档的格式化元数据、或以其它方式545来进行。然后,在框525处,将识别出的格式化应用于所显示的新文档。

[0093] 快速开始部件202还能够显示用户输入机构,所述用户输入机构能够由用户启动以允许用户看到已经应用的特定格式化。当这个发生时,部件202在新的显示文档上公开格式化信息,因此用户能够看到其是如何格式化的。这由框527指示。然后,系统继续处理其它用户输入,例如进一步格式化输入、编写输入或其它输入。这由框529指示。

[0094] 图6C-图6H示出了这个的各种范例。在图6C中能够看出用户已经打开了文档,并且

对应应用已经显示了空白画布。然后,用户启动用户输入机构,使得用户能够提供语音命令。在图6D中,用户已经提供了语音命令“将这个建立为像我上学期的论文”。语言处理系统104生成针对该话语的文本表示和自然语言理解。意图是发现用户上学期的论文,从该论文提取格式化并将其应用于新打开的文档。

[0095] 图6E示出了动作生成器118已经使用搜索系统116识别对应于用户查询的多个相关文档。其显示了对应于相关文档中的每个的显示元件(例如缩略图或其它显示元件)。图6E示出了用户正在选择它们中的一个。图6F示出了用户正在提供用户输入以将该格式化应用于新打开的文档。图6G示出了用户输入机构531,其能够被用户启动以查看文档的不同章节的格式化。当用户启动用户输入机构531时,显示格式化信息。图6H示出了能够单独的显示多个不同章节及其自身的格式信息。

[0096] 图7-图9示出了指示图1中示出的架构100中的部件或项目能够被定位在不同位置处的各种范例。图7示出了多种不同的项目被定位在移动设备601上,所述移动设备601通过网络603连接到实施搜索系统116的搜索服务。网络603能够是局域网、广域网、蜂窝网、近场通信网络、或各种其它网络。

[0097] 图8示出了项目被设置在客户端/服务器架构中,其中,基于上下文的动作系统102连同语言处理系统104一起被部署在服务器604上。用户112通过生成用于与用户112交互的用户接口机构114的客户端设备606访问服务器604上的系统102。

[0098] 本讨论已经提及了处理器和服务器。在一个实施例中,处理器和服务器包括具有相关联的存储器和定时电路的计算机处理器,其未单独示出。它们是该系统或设备的功能部分(它们属于所述功能部分并被所述功能部分激活)并促进这些系统中的其它部件或项目的功能。

[0099] 而且,已经讨论了多个用户接口显示。它们能够采取多种不同形式,并且能够具有被布置在其上的多种不同的用户可启动输入机构。例如,用户可启动输入机构能够是文本框、复选框、图标、链接、下拉菜单、搜索框等。它们还能够以多种方式被启动。例如,能够使用点击设备(例如,轨迹球或鼠标)启动它们。能够使用硬件按钮、开关、控制杆或键盘、拇指开关或拇指垫等来启动它们。还能够使用虚拟键盘或其它虚拟致动器来启动它们。另外,当显示它们的屏幕是触敏屏幕时,能够使用触摸手势来启动它们。而且,当显示它们的设备具有言语辨识部件时,能够使用言语命令来启动它们。

[0100] 还讨论了多个数据存储库。应当注意,它们每个能够都能够被分解成多个数据存储库。所有的都能够是对于访问它们的系统为本地的,所有的都能够是远程的,或者一些能够是本地的而其它的是远程的。在本文中预想到所有这些配置。

[0101] 而且,附图示出了具有归因于每个框的功能的多个框。应当注意,能够使用较少的框,因此通过较少的部件执行功能。而且,能够使用更多的框,其中,功能被分布在更多的部件之中。

[0102] 图9是图1中示出的架构100的框图,除了其元件被设置在云计算架构500中。云计算提供计算、软件、数据访问和存储访问,这些不要求终端用户知道输送这些服务的系统的物理位置或配置。在各种实施例中,云计算通过广域网(例如互联网)使用适当的协议输送服务。例如,云计算提供商通过该广域网输送应用,并且能够通过网络浏览器或任何其它计算部件对其进行访问。架构100的软件或部件以及对应数据能够被存储于在远处位置处的

服务器上。云计算环境中的计算资源能够被固定在远程数据中心位置处或者它们能够是分散的。云计算基础设施能够通过共享数据中心输送服务,即使它们看起来是用于用户的单个接入点。因此,在本文中描述的部件和功能能够从在远处位置处的服务提供商使用云计算架构来提供。备选地,能够从传统的服务器提供它们,或者能够直接或以其它方式将它们安装在客户端设备上。

[0103] 描述旨在包括公共云计算和私有云计算两者。(公共的和私有的)云计算提供对资源的基本无缝池化,并减少管理和配置基础硬件基础设施的需要。

[0104] 公共云由供应商管理,并且通常支持使用相同的基础设施的多个消费者。而且,公共云与私有云相反,其能够将终端用户从管理硬件中释放出来。私有云由组织本身管理,并且基础设施通常不与其它组织进行共享。组织仍在一定程度上维护硬件,例如安装和维修等。

[0105] 在图9中示出的范例中,一些项目类似于在图1中示出的那些,并且它们被类似地编号。图9具体示出了在图1中能够被定位在云502(其能够是公共的、私有的、或部分公共的同时其它是私有的组合)中的各种项目。因此,用户112使用用户设备504通过云502访问这些项目。

[0106] 图9还描绘了云架构的另一实施例。图9示出了还预想到架构100的一些元件能够被设置在云502中而其它元件不被设置在云502中。通过范例的方式,数据存储库132和源106能够被设置在云502的外部,并通过云502进行访问。在另一范例中,系统102或搜索系统116或其它部分也能够设置在云502的外部。不管它们被定位在何处,它们能够直接由设备504通过网络(广域网或局域网)访问,它们能够由服务托管在远程站点处,或者它们能够作为服务通过云被提供或通过驻留在云中的连接服务访问。在本文中预想到所有这些架构。

[0107] 还应当注意,架构100或其部分能够被设置在多种不同的设备上。这些设备中的一些包括服务器、台式计算机、膝上型计算机、平板计算机、或其它移动设备,例如掌上计算机、蜂窝电话、智能电话、多媒体播放器、个人数字助理等。

[0108] 图10是能够被用作能够部署本系统(或其部分)的用户或客户端的手持设备16的手持或移动计算设备的一个图示性实施例的简化框图。图11-图12是手持或移动设备的范例。

[0109] 图10提供了客户端设备16的部件的一般框图,所述设备16能够运行架构100的部件或者能够运行与架构100交互的部件或运行两者。不是图1的所有元件都在设备16中示出,尽管预想到他们能够被设置在所图示的部件中的设备16上、或被添加到所图示的部件中的设备16上。在设备16中,通信链路13被提供为允许手持设备与其它计算设备通信,并且在一些范例中提供用于自动接收信息的通道,例如通过扫描。通信链路13的范例包括红外端口、串行/USB端口、电缆网络端口(例如,以太网端口)、以及允许通过一个或多个通信协议进行通信的无线网络端口,所述协议包括通用分组无线业务(GPRS)、LTE、HSPA、HSPA+和其它3G和4G无线协议、1Xrtt以及短消息服务,其是用于提供对网络的蜂窝接入的无线服务,例如,Wi-Fi协议和蓝牙协议,其提供对网络的本地无线连接。

[0110] 在其它范例中,在连接到SD卡接口15的可移除安全数字(SD)卡上接收应用或系统。SD卡接口15和通信链路13沿着总线19与处理器17(其还能够实施来自图1的处理器128或其它处理器)通信,所述总线还连接到存储器21和输入/输出(I/O)部件23,以及时钟25和

位置系统27。

[0111] 在一个实施例中，I/O部件23被提供为促进输入和输出操作。用于设备16的各种实施例的I/O部件23能够包括：输入部件，例如，按钮、触摸传感器、多点触摸传感器、光学或视频传感器、语音传感器、触摸屏、接近度传感器、麦克风、倾斜传感器、以及重力开关；输出部件，例如，显示设备、扬声器、以及打印机端口。还能够使用其它I/O部件23。

[0112] 时钟25说明性地包括实时时钟部件，其输出时间和日期。说明性地，其还能够提供用于处理器17的定时功能。

[0113] 位置系统27说明性地包括输出设备16的当前地理位置的部件。这能够包括例如全球定位系统(GPS)接收器、LORAN系统、航位推测系统、蜂窝三角测量系统、或其它定位系统。其还能够包括例如映射软件或导航软件，其生成期望的地图、导航路由和其它地理功能。

[0114] 存储器21存储操作系统29、网络设置31、应用33、应用配置设置35、数据存储37、通信驱动39、以及通信配置设置41。存储器21能够包括所有类型的有形的易失性和非易失性的计算机可读存储器设备。其还能够包括计算机存储媒介(如下所述)。存储器21存储计算机可读指令，所述计算机可读指令当被处理器17运行时使得处理器根据指令执行计算机实施的步骤或功能。应用154或数据存储156中的项目例如能够驻留在存储器21中。类似地，设备16能够具有客户端商业系统24，其能够运行各种商业应用或实施部分或所有意图104。处理器17也能够由其它部件激活以促进其功能。

[0115] 网络设置31的范例包括例如代理信息、互联网连接信息和映射的事物。应用配置设置35包括针对特定企业或用户裁剪应用的设置。通信配置设置41提供用于与其它计算机通信的参数，并包括例如GPRS参数、SMS参数、连接用户名称和密码的项目。

[0116] 应用33能够是先前存储于设备16上的应用，或者在使用期间安装的应用，但是这些能够是操作系统29的部分，或托管在设备16的外部。

[0117] 图11示出了设备16是平板计算机600的一个范例。在图6中，计算机600被示为具有用户接口显示屏602。屏幕602能够是触摸屏(因此来自用户手指的触摸手势能够用于与应用交互)，或者是从笔或触笔接收输入的笔支持接口。还能够使用屏幕上虚拟键盘。当然，还可以通过合适的附接机构(例如，无线链接或USB端口)附接到键盘或其它用户输入设备。计算机600还能够说明性地接收语音输入。

[0118] 还能够使用设备16的额外范例。设备16能够是特征电话、智能电话或移动电话。电话包括：一组按键，其用于拨打电话号码；显示器，其能够显示包括应用图像、图标、网页、照片和视频的图像；以及控制按钮，其用于选择在显示器上示出的项目。电话能够包括天线，所述天线用于接收蜂窝电话信号，例如，通用分组无线业务(GPRS)和1Xrtt以及短消息服务(SMS)信号。在一些范例中，电话还包括：安全数字(SD)卡槽，其用于容纳SD卡。

[0119] 移动设备还能够是个人数字助理或多媒体播放器或平板计算设备等(下文称作PDA)。PDA能够包括感应屏，其在触笔(或其它指针，例如用户的手指)被定位在屏幕上时感知触笔的部分。这允许用户选择、突出显示和移动屏幕上的项目以及绘制和书写。PDA还能够包括多个用户输入键或按钮，其允许用户滚动通过在显示器上显示的菜单选项或其它显示选项，并允许用户改变应用或选择用户输入功能，而不接触显示器。PDA能够包括：内部天线和红外发射器/接收器，其允许与其它计算机进行无线通信；以及连接端口，其允许硬件连接到其它计算设备。这种硬件连接通常由支架支撑，所述支架通过串行端口或USB端口连

接到其它计算机。这样,这些连接是非网络连接。

[0120] 图12还示出了电话是智能电话71。智能电话71具有触敏显示器73,其显示图标或瓦片(tile)或其它用户输入机构75。机构75能够由用户用于运行应用、打电话、执行数据转移操作等。一般而言,智能电话71被构建在移动操作系统上,并提供比特征电话更高级的计算能力和连接性。

[0121] 注意,设备16的其它形式是可能的。

[0122] 图13是(例如)能够部署架构100或其部分的计算环境的一个范例。参考图13,用于实施一些实施例的示例性系统包括计算机810形式的通用计算设备。计算机810的部件可以包括但不限于:处理单元820(其能够包括处理器128或设备504或其它设备中的处理器)、系统存储器830、以及系统总线821,所述系统总线821将包括系统存储器的各种系统部件耦合到处理单元820。系统总线821可以是多种类型的总线结构中的任一种,包括存储器总线或存储器控制器、外围总线、以及使用多个总线架构中的任一种的本地总线。通过范例而非限制的方式,这种架构包括工业标准架构(ISA)总线、微通道架构(MCA)总线、扩展ISA(EISA)总线、视频电子标准协会(VESA)局部总线、外围部件互联(PCI)总线(但也已知为夹层总线)。关于图1描述的存储器和程序能够被部署在图13的对应部分中。

[0123] 计算机810通常包括多个计算机可读媒介。计算机可读媒介能够是能够被计算机810访问的任何可用媒介,其包括易失性和非易失性媒介、可移除和不可移除媒介。通过范例而非限制的方式,计算机可读媒介可以包括计算机存储媒介和通信媒介。计算机存储媒介不同于且不包括已调数据信号或载波。其包括硬件存储媒介,包括易失性和非易失性的、可移除和不可移除的媒介,其以任何方法或技术实施以用于存储信息,例如,计算机可读指令、数据结构、程序模块或其它数据。计算机存储介质包括但不限于:RAM、ROM、EEPROM、闪存或其它存储器技术、CD-ROM、数字通用盘(DVD)或其它光盘存储设备、磁带盒、磁带、磁盘存储设备或其它磁存储设备,或者能够用于存储期望信息并能够被计算机810访问的任何其它媒介。通信媒介通常在传输机构中实施计算机可读指令、数据结构、程序模块或其它数据,并包括任何信息输送媒介。术语“已调数据信号”意指具有其特征集中的一个或多个的或以编码信号中的信息的方式改变的信号。通过范例而非限制的方式,通信媒介包括有线媒介,例如,有线网络或直连线连接;以及无线媒介,例如声学、RF、红外和其它无线媒介。上述任何组合也应当被包含在计算机可读媒介的范围内。

[0124] 系统存储器830包括易失性和/或非易失性存储器形式的计算机存储媒介,例如,只读存储器(ROM)831和随机存取存储器(RAM)832。基本输入/输出系统833(BIOS)包含通常被存储在ROM 831中的基本例程,其帮助例如在启动期间在计算机810内的元件之间转移信息。RaM 832通常包含数据和/或程序模块,其对处理单元820立即可用和/或当前被处理单元820操作。通过范例而非限制的方式,图13图示了操作系统834、应用程序835、其它程序模块836和程序数据837。

[0125] 计算机810还可以包括其它可移除/不可移除的易失性/非易失性计算机存储媒介。仅通过范例的方式,图13图示了硬盘驱动器841,其从不可移除的非易失性磁性媒介进行读取或写入不可移除的非易失性磁性媒介;以及光盘驱动器855,其从可移除的非易失性光盘856(例如,CD ROM或其它光学媒介)进行读取或写入可移除的非易失性光盘856(例如,CD ROM或其它光学媒介)。能够在示例性操作环境中使用的其它可移除/不可移除的易失

性/非易失性计算机存储媒介包括但不限于：磁带盒、闪存卡、数字通用盘、数字录像带、固态RAM、固态ROM等。硬盘驱动器841通常通过不可移除存储器接口（例如接口840）连接到系统总线821，并且光盘驱动器855通常通过可移除存储器接口（例如接口850）连接到系统821。

[0126] 备选地或额外地，在本文中描述的功能能够至少部分地由一个或多个硬件逻辑部件来执行。例如但不限于，能够使用的说明性类型的硬件逻辑部件包括：现场可编程门阵列（FPGA）、专用集成电路（ASIC）、程序专用标准产品（ASSP）、单片系统（SOC）、复杂可编程逻辑设备（CPLD）等。

[0127] 以上讨论的且在图13中图示的驱动器及其相关联的计算机存储媒介提供对计算机可读指令、数据结构、程序模块和计算机810的其它数据的存储。在图13中，例如，将硬盘驱动器841图示为存储操作系统844、应用程序845、其它程序模块846以及程序数据847。注意，这些部件能够与操作系统834、应用程序835、其它程序模块836以及程序数据837相同或不同。操作系统844、应用程序845、其它程序模块846以及程序数据847在这里给出了不同的编号以最低限度地图示它们是不同副本。

[0128] 用户可以通过输入设备将命令和信息输入到计算机810中，所述输入设备例如是键盘862、麦克风863以及指针设备861（例如，鼠标、轨迹球或触摸垫）。其它输入设备（未示出）可以包括控制杆、游戏垫、卫星盘、扫描仪等。这些输入设备和其它输入设备通常通过被耦合到系统总线的用户输入接口860连接到处理单元820，但是可以通过其它接口和总线结构（例如，并行端口、游戏端口或通用串行总线（USB））进行连接。视觉显示器891或其它类型的显示设备还经由接口（例如，视频接口890）连接到系统总线821。除了监视器外，计算机还包括其它外围输出设备，例如扬声器897和打印机896，其能够通过输出外围接口895而被连接。

[0129] 计算机810能够在网络环境中使用到一个或多个远程计算机（例如，远程计算机880）的逻辑连接进行操作。远程计算机880还可以是个人计算机、手持设备、服务器、路由器、网络PC、对等设备或其它公用网络节点，并且通常包括以上关于计算机810描述的许多元件或所有元件。在图13中描绘的逻辑连接包括局域网（LAN）871和广域网（WAN）873，但还可以包括其它网络。这种网络环境在办公室、企业宽带计算机网络、内联网和互联网中常见。

[0130] 当在LAN网络环境中使用时，计算机810通过网络接口或适配器870连接到LAN 871。当在WAN网络环境中使用时，计算机810通常包括调制解调器872或用于通过WAN 873（例如，互联网）建立通信的其它器件。在外部或内部的调制解调器872可以经由用户输入接口860或其它适当机构连接到系统总线821。在网络环境中，关于计算机810或其部分描绘的程序模块可以被存储在远程存储器存储设备中。通过范例而非限制的方式，图13图示了驻留在远程计算机880上的远程应用程序885。应当理解，所示出的网络连接是示例性的，并且可以使用建立计算机之间的通信连接的其它手段。

[0131] 还应当注意，在本文中描述的不同实施例能够以不同方式进行组合。也就是说，一个或多个实施例的部分能够与一个或多个其它实施例的部分进行组合。所有这些都在本文中被预想到。

[0132] 范例1是一种计算系统，包括：

[0133] 用户接口机构,其接收话语,将所述话语发送到语言处理系统,并接收指示在所述话语中辨识出的意图和一组变元的语言处理结果;

[0134] 上下文识别系统,其识别所述计算系统的当前上下文;

[0135] 动作识别器,其基于所述当前上下文并基于在所述语言处理结果中的所述意图来识别动作;以及

[0136] 跨源搜索部件,其基于所述当前上下文并基于在所述语言处理结果中的所述一组变元来搜索多个不同的信息源,以识别要对其执行所述动作的内容。

[0137] 范例2是任何或所有先前范例的计算系统,还包括:

[0138] 动作部件,其控制受控系统以对识别出的内容采取所述动作。

[0139] 范例3是任何或所有先前范例的计算系统,其中,所述话语隐含地识别所述内容,并且其中,所述跨源搜索部件使用所述变元作为搜索项来搜索所述多个不同的信息源,以识别多个不同的搜索结果,每个搜索结果识别不同的内容的项目。

[0140] 范例4是任何或所有先前范例的计算系统,还包括:

[0141] 相关性生成器,其针对所述不同的搜索结果中的每个生成相关性量度以获得给定的搜索结果。

[0142] 范例5是任何或所有先前范例的计算系统,其中,所述相关性生成器基于当前上下文和所述变元,使用所述变元和所述当前上下文来另外消除模糊的搜索结果以获得所述给定的搜索结果。

[0143] 范例6是任何或所有先前范例的计算系统,其中,所述语言处理系统包括:

[0144] 第一语言处理系统,其被设置在所述计算系统上,并包括言语辨识系统和自然语言理解系统,所述言语辨识系统生成所述话语的文本表示,并且所述自然语言理解系统识别在第一语言处理结果中的所述意图和所述变元,并且还基于所述当前上下文来另外消除模糊的变元。

[0145] 范例7是任何或所有先前范例的计算系统,其中,所述用户接口机构还将所述话语发送到远离所述计算系统的第二语言处理系统,所述第二语言处理系统生成第二语言处理结果,所述计算系统组合所述第一语言处理结果与所述第二语言处理结果以获得最终的语音处理结果。

[0146] 范例8是任何或所有先前范例的计算系统,其中,所述上下文识别系统包括:

[0147] 个人身份检测器,其识别用户的个人身份。

[0148] 范例9是任何或所有先前范例的计算系统,其中,所述上下文识别系统包括:

[0149] 接近度检测器,其识别在到所述计算系统的给定接近度内的其他人或事物。

[0150] 范例10是任何或所有先前范例的计算系统,其中,所述上下文识别系统包括:

[0151] 打开应用检测器,其识别在所述计算系统上的打开应用。

[0152] 范例11是任何或所有先前范例的计算系统,其中,所述上下文识别系统包括:

[0153] 设备检测器,其识别在其上设置有所述计算系统的特定设备。

[0154] 范例12是任何或所有先前范例的计算系统,其中,所述上下文识别系统包括:

[0155] 日历检测器,其识别用户的日历上的特定时间和项目。

[0156] 范例13是任何或所有先前范例的计算系统,其中,所述特定设备包括移动设备。

[0157] 范例14是任何或所有先前范例的计算系统,其中,所述上下文识别系统包括:

[0158] 位置检测器,其识别在其上设置有所述计算系统的所述特定设备的位置。

[0159] 范例15是任何或所有先前范例的计算系统,其中,所述动作识别器将所述动作识别为以下中的一个:发送、打开、安排会议、创建、附接、电子邮件发送、设置提醒、共享、呈现、发起和显示。

[0160] 范例16是一种计算机实施的方法,包括:

[0161] 接收话语;

[0162] 将所述话语发送到语言处理系统;

[0163] 接收指示在所述话语中辨识出的意图和一组变元的语言处理结果;

[0164] 检测所述计算系统的上下文;

[0165] 分析所述上下文和在所述语言处理结果中的所述意图来识别要采取的动作;

[0166] 基于所述上下文并基于在所述语言处理结果中的所述一组变元来搜索多个不同的信息源,以识别要对其执行所述动作的内容;并且

[0167] 自动控制受控系统以对识别出的内容执行所述动作。

[0168] 范例17是任何或所有先前范例的计算机实施的方法,还包括:

[0169] 生成对应于所述不同的搜索结果中的每个的相关性量度;并且

[0170] 基于所述对应的相关性量度来对所述不同的搜索结果进行排序。

[0171] 范例18是任何或所有先前范例的计算机实施的方法,还包括:

[0172] 基于当前上下文和所述变元,使用所述变元和所述上下文来另外消除模糊的搜索结果以获得给定的搜索结果。

[0173] 范例19是一种移动设备,包括:

[0174] 用户接口机构,其接收话语,将所述话语发送到语言处理系统,并接收指示在所述话语中辨识出的意图和一组变元的语言处理结果;

[0175] 上下文识别系统,其识别所述计算系统的上下文;

[0176] 动作识别器,其基于当前上下文并基于在所述语言处理结果中的所述意图来识别动作;

[0177] 跨源搜索部件,其基于所述上下文并基于在所述语言处理结果中的所述一组变元来搜索多个不同的信息源,以识别要对其执行所述动作的内容;以及

[0178] 动作部件,其控制受控系统以对识别出的内容采取所述动作。

[0179] 范例20是任何或所有先前范例的移动设备,其中,所述话语隐含地识别所述内容,并且其中,所述跨源搜索部件使用所述变元作为搜索项来搜索所述多个不同的信息源,以识别多个不同的搜索结果,每个搜索结果识别不同的内容的项目,并且还包括:

[0180] 相关性生成器,其针对所述不同的搜索结果中的每个生成相关性量度以获得给定的搜索结果;并且所述相关性生成器基于所述上下文和所述变元,使用所述变元和所述当前上下文来另外消除模糊的搜索结果以获得所述给定的搜索结果。

[0181] 尽管以专用于结构特征和/或方法动作的语言描述了主题,但是应当理解,在权利要求中定义的主题不必限于上述特定特征或动作。相反,上述特定特征和动作被公开为实施权利要求的范例形式。

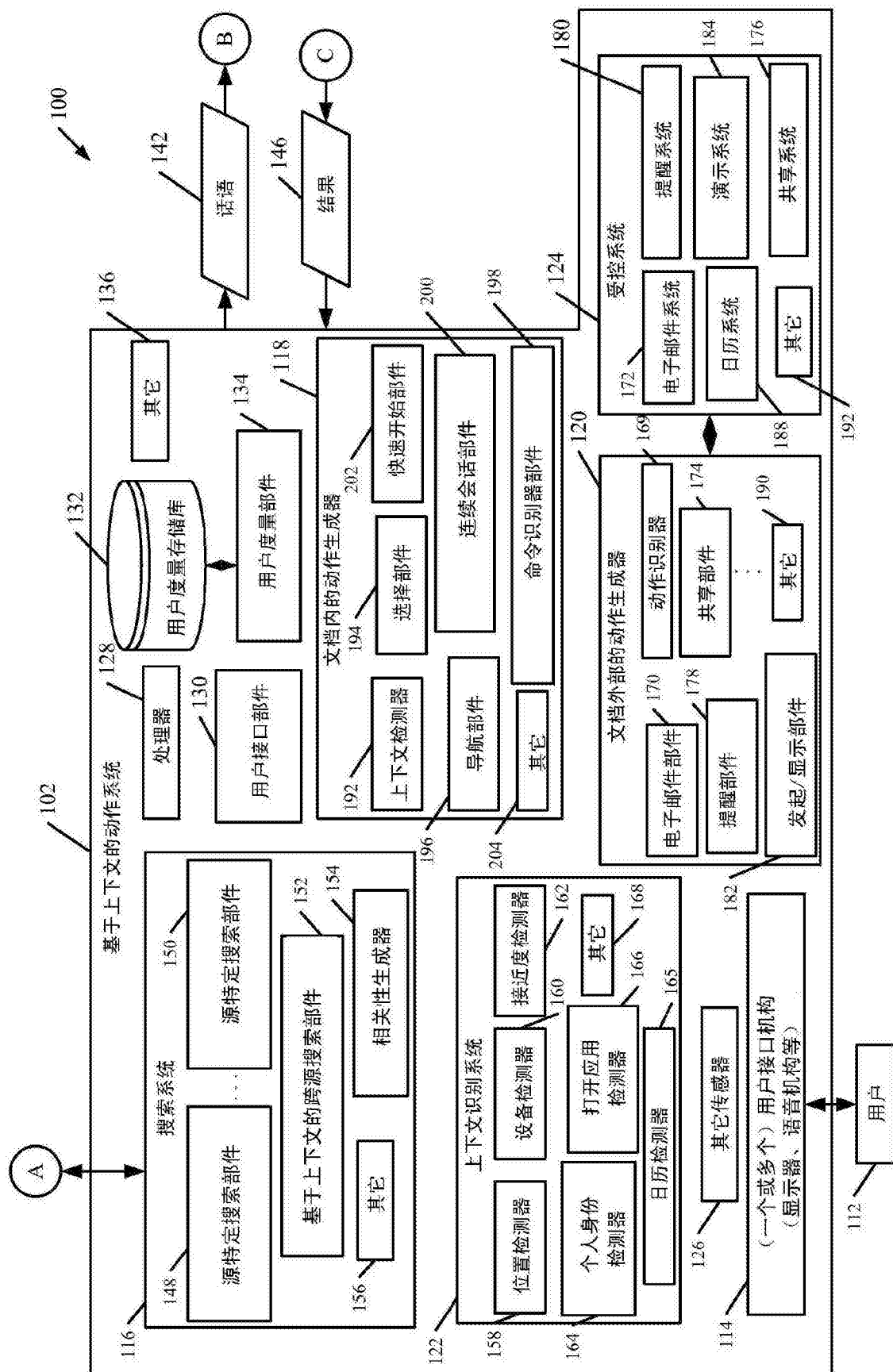


图1A

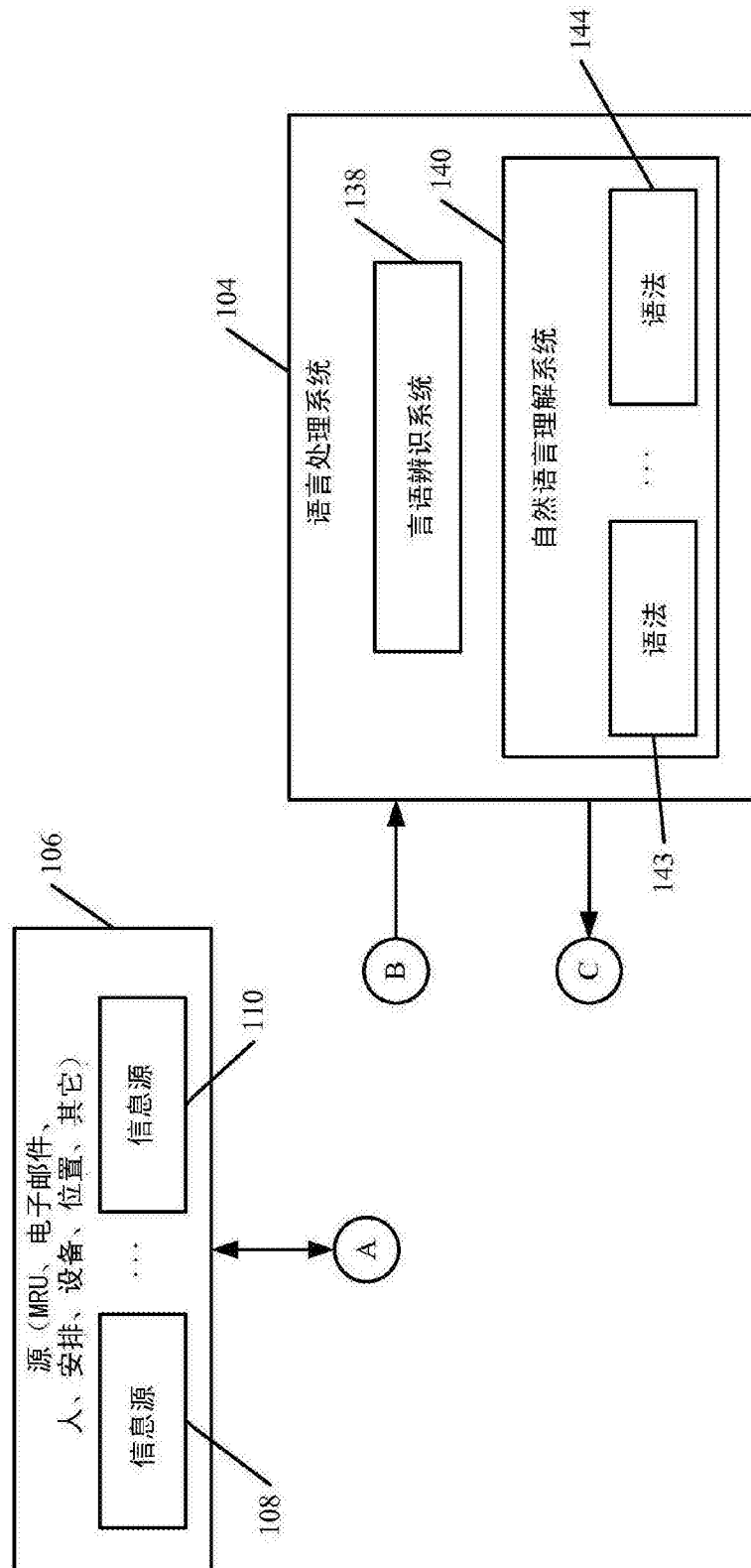


图1B

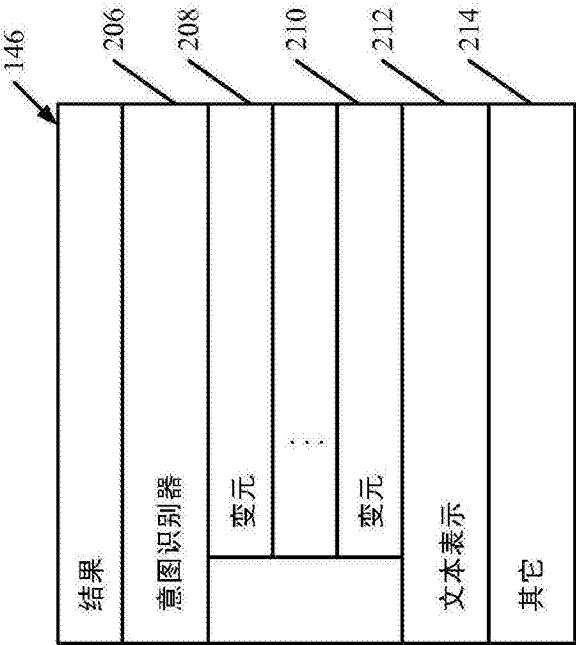


图2

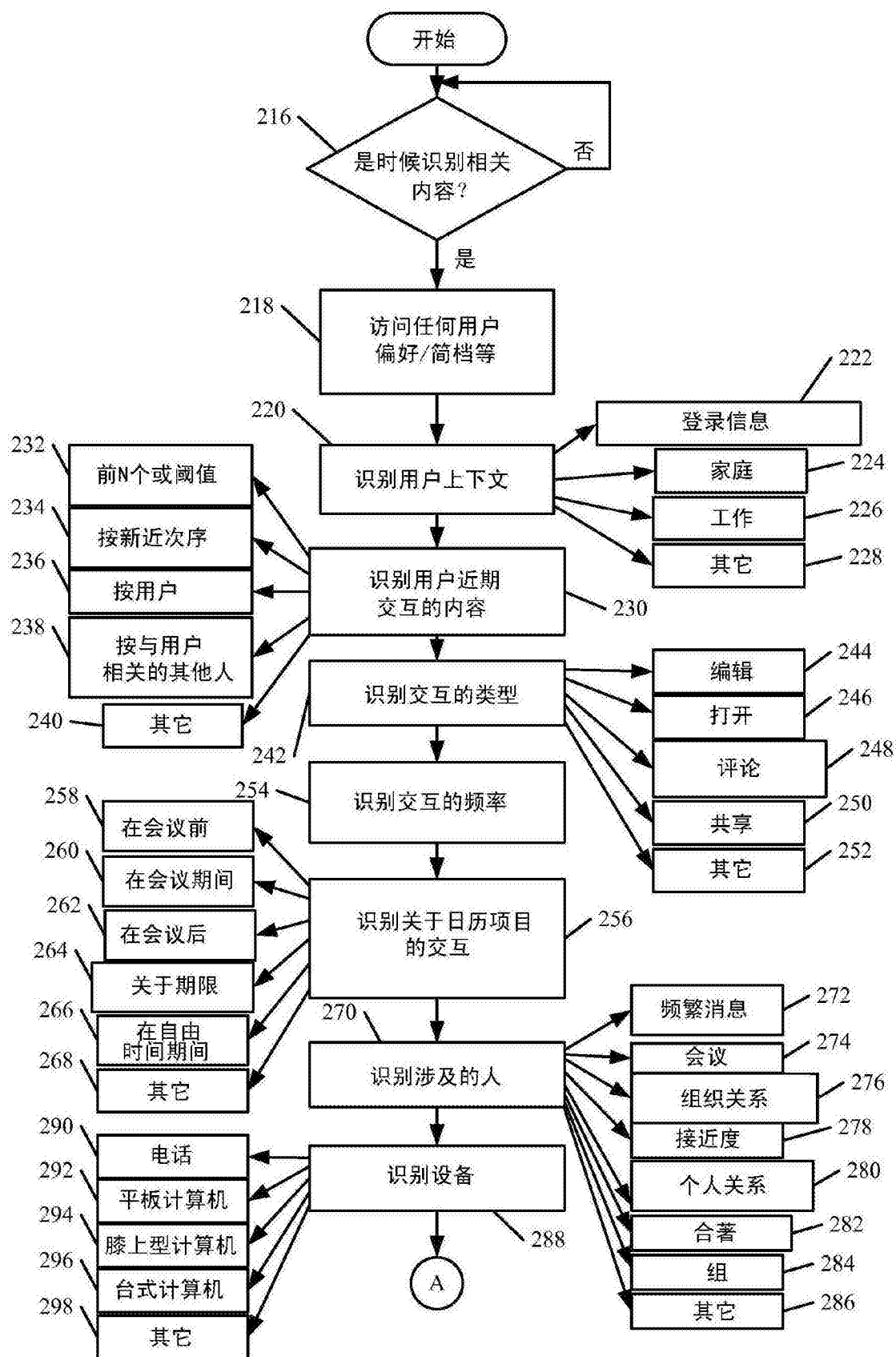


图3A

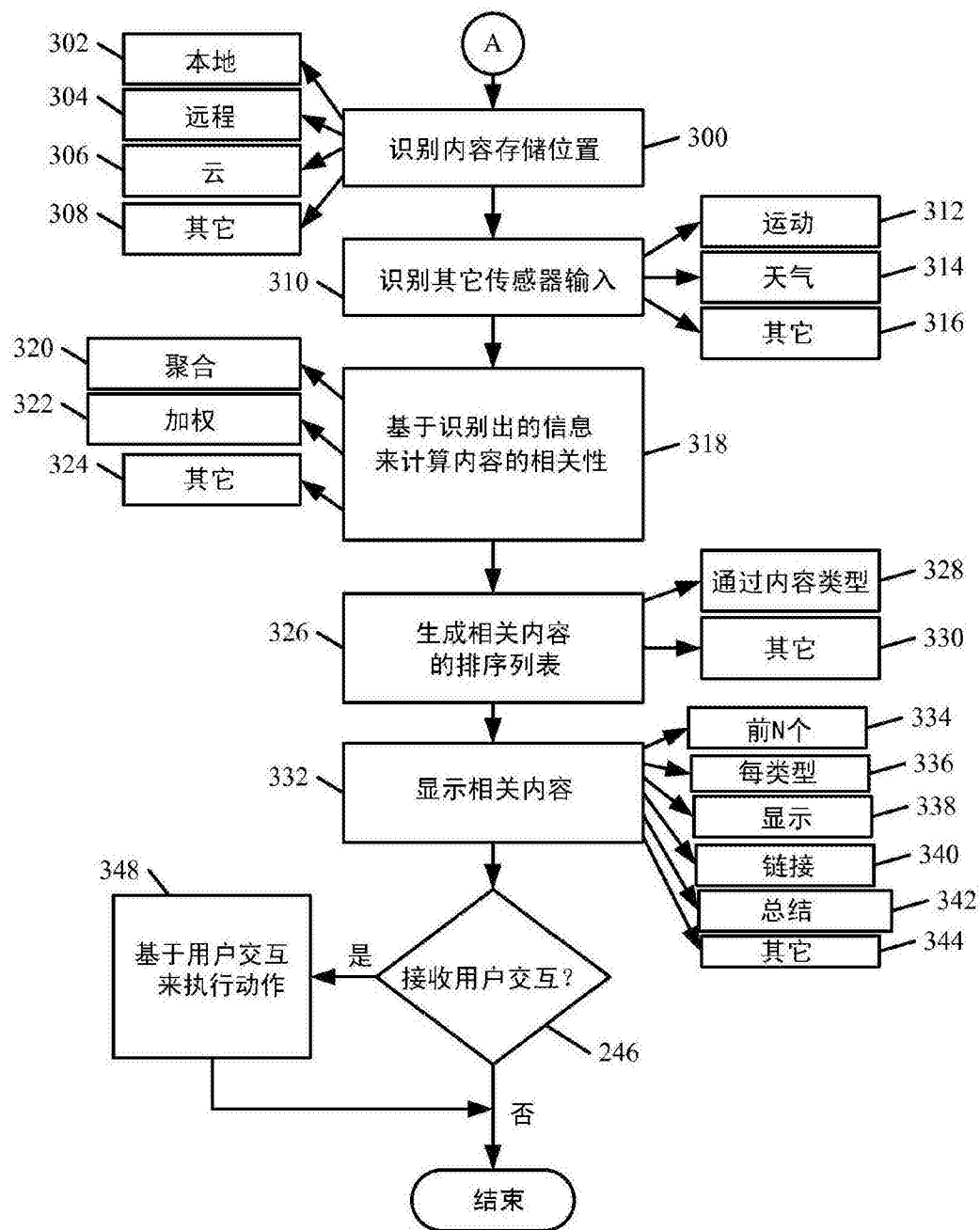


图3B

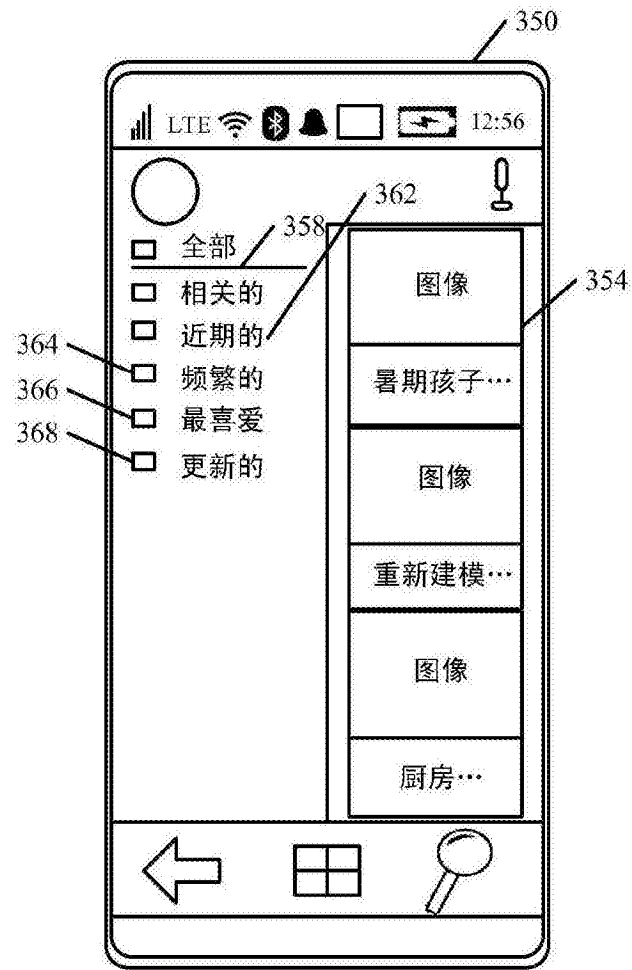


图3C

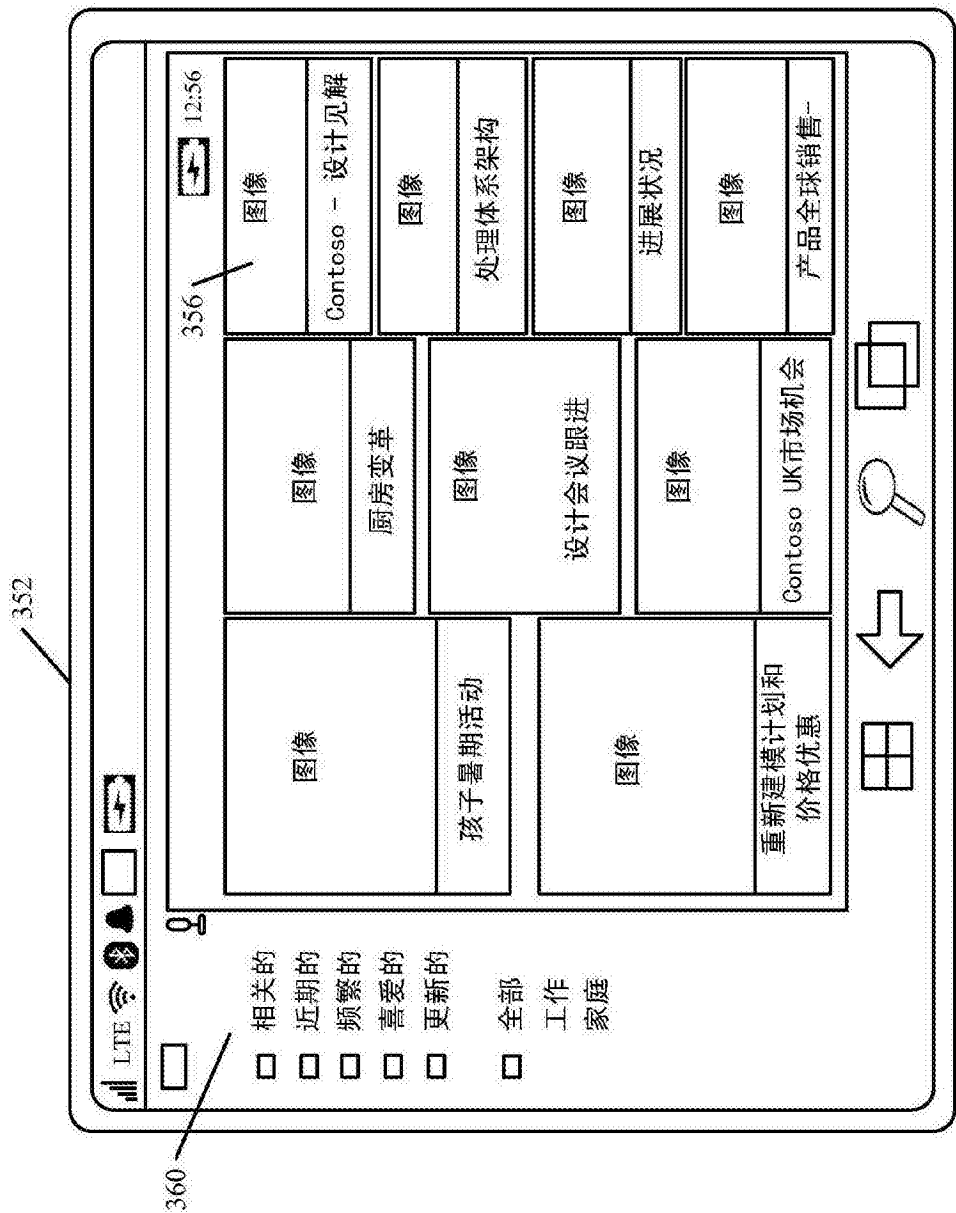


图3D

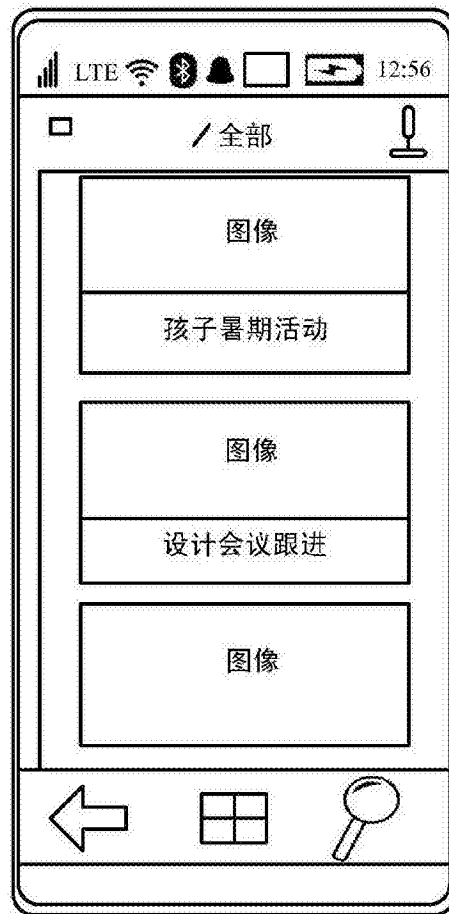


图3E

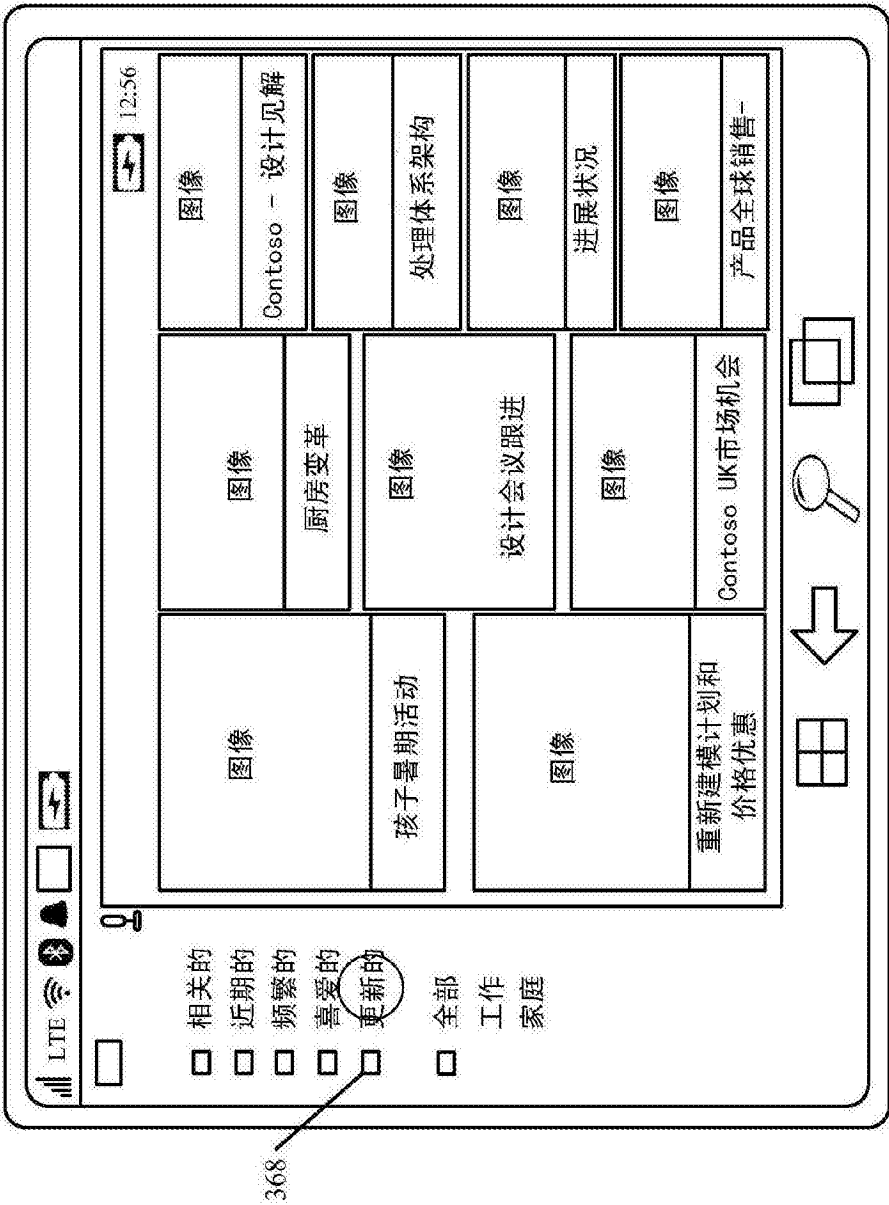


图3F



图3G

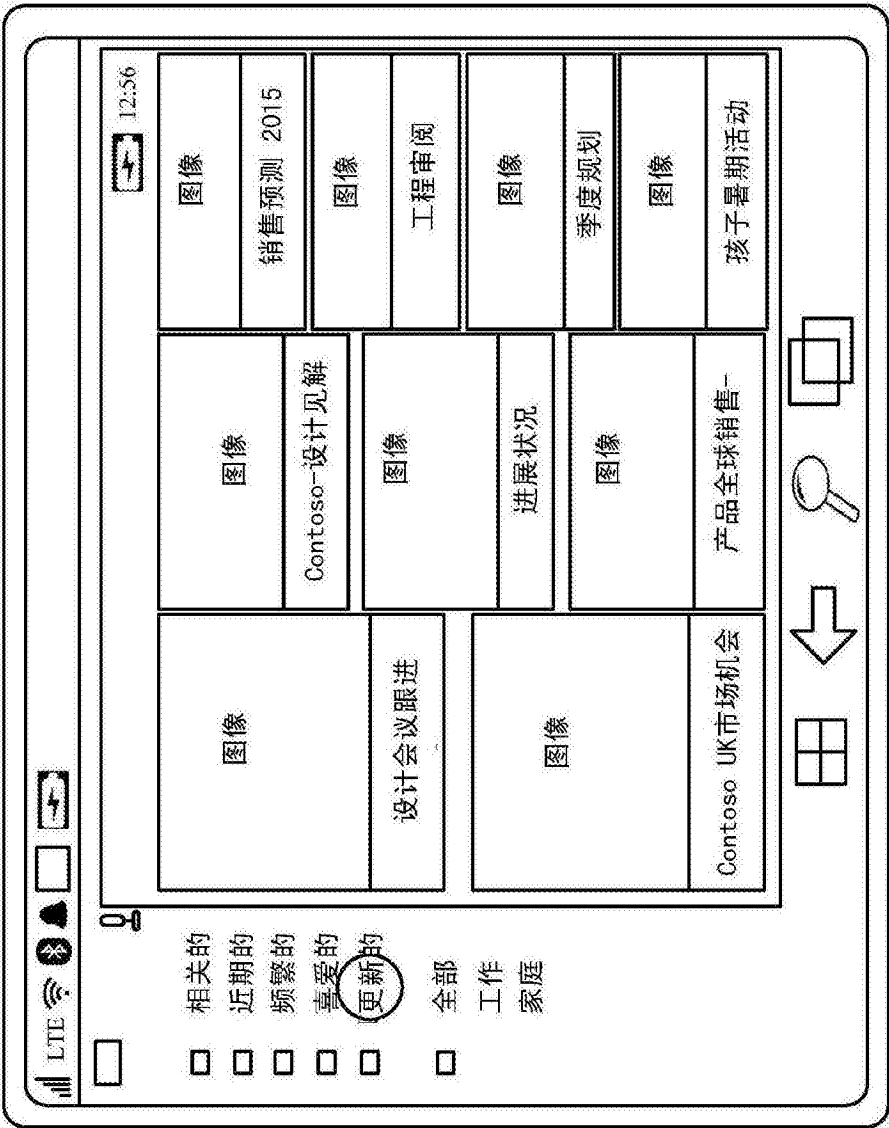


图3H

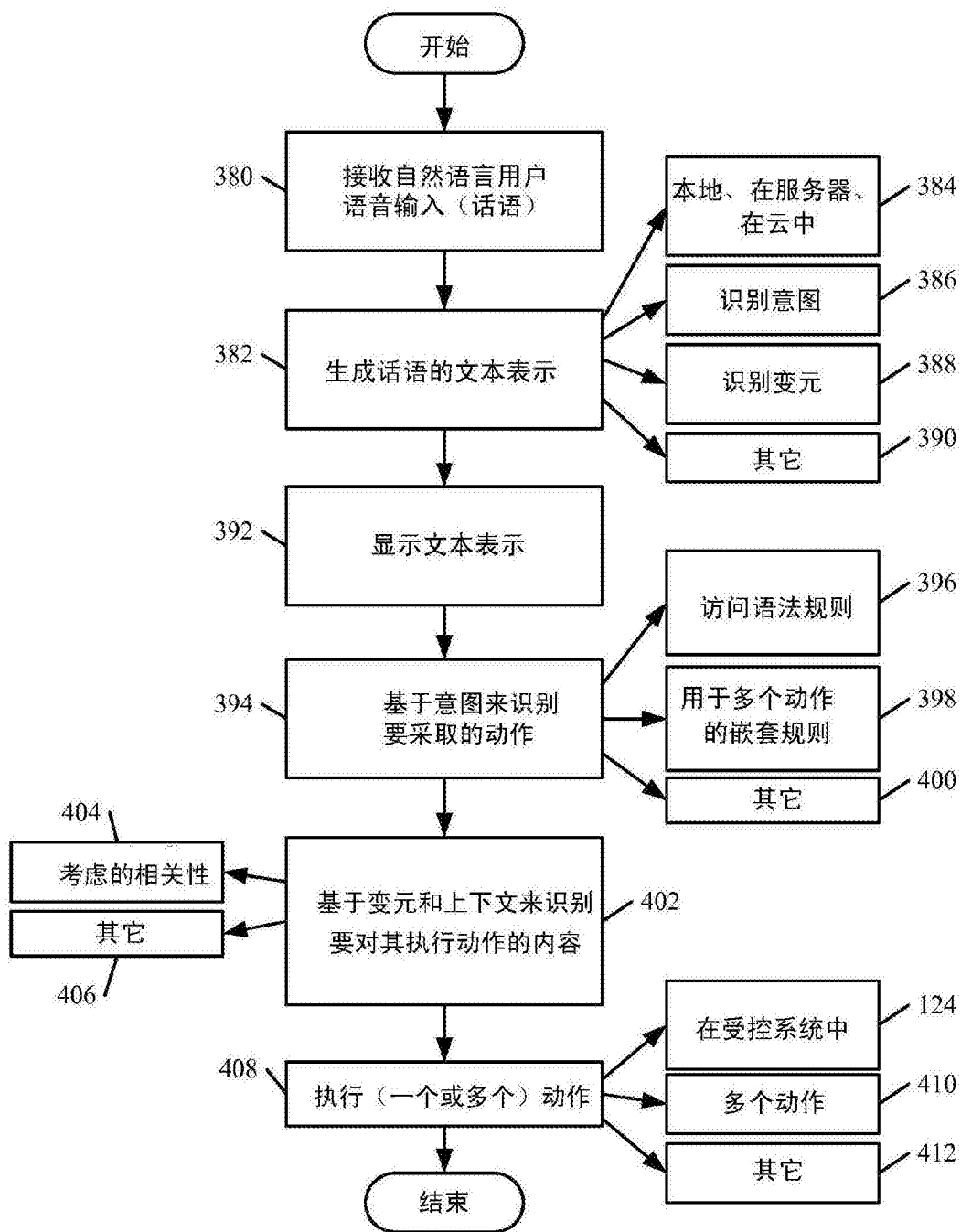


图4

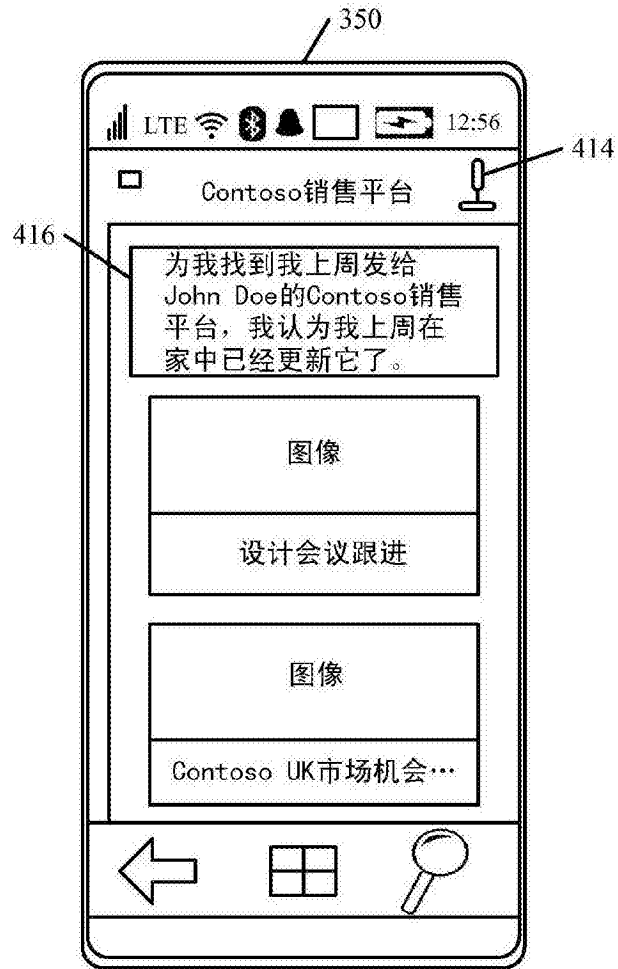


图4A

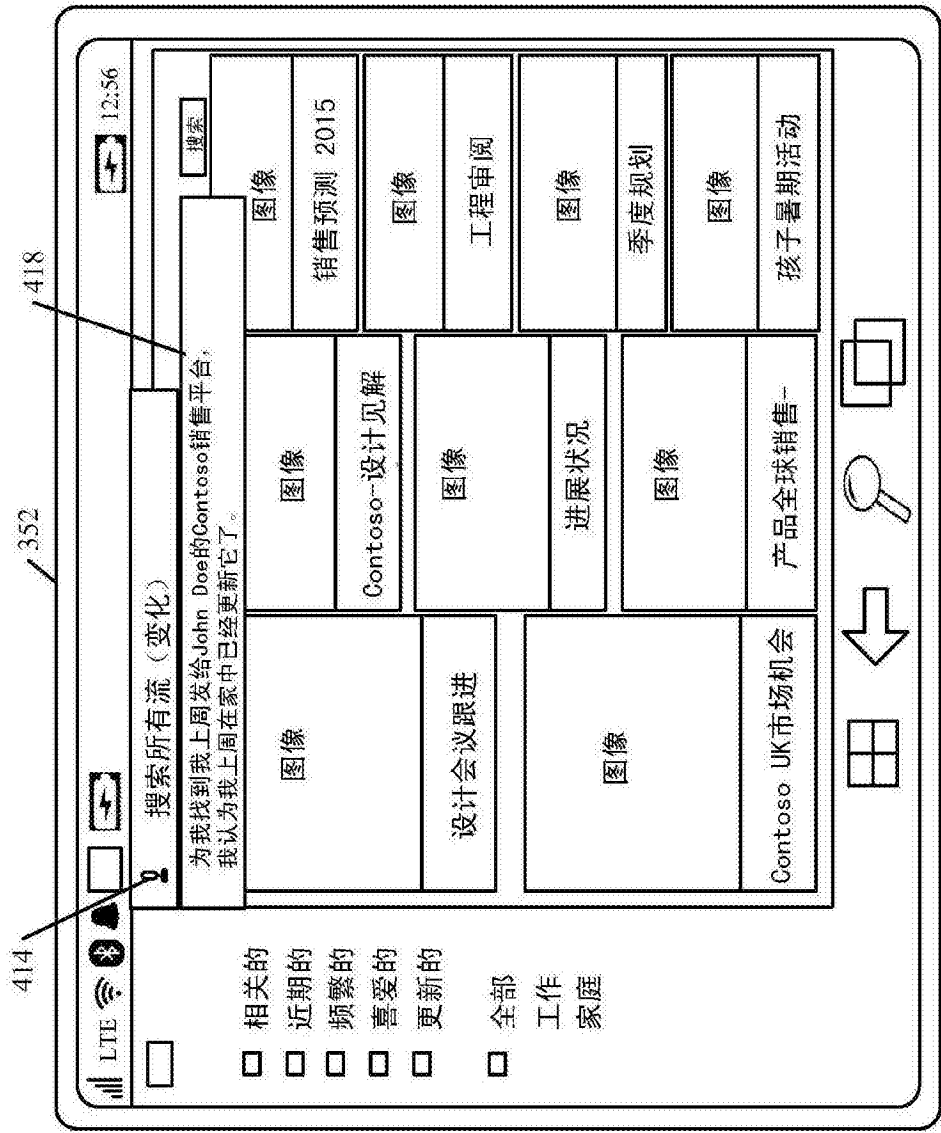


图4B

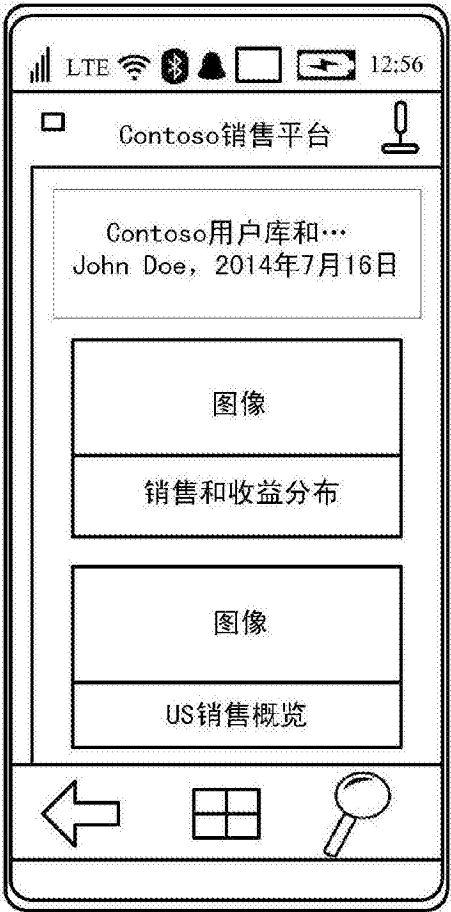


图4C

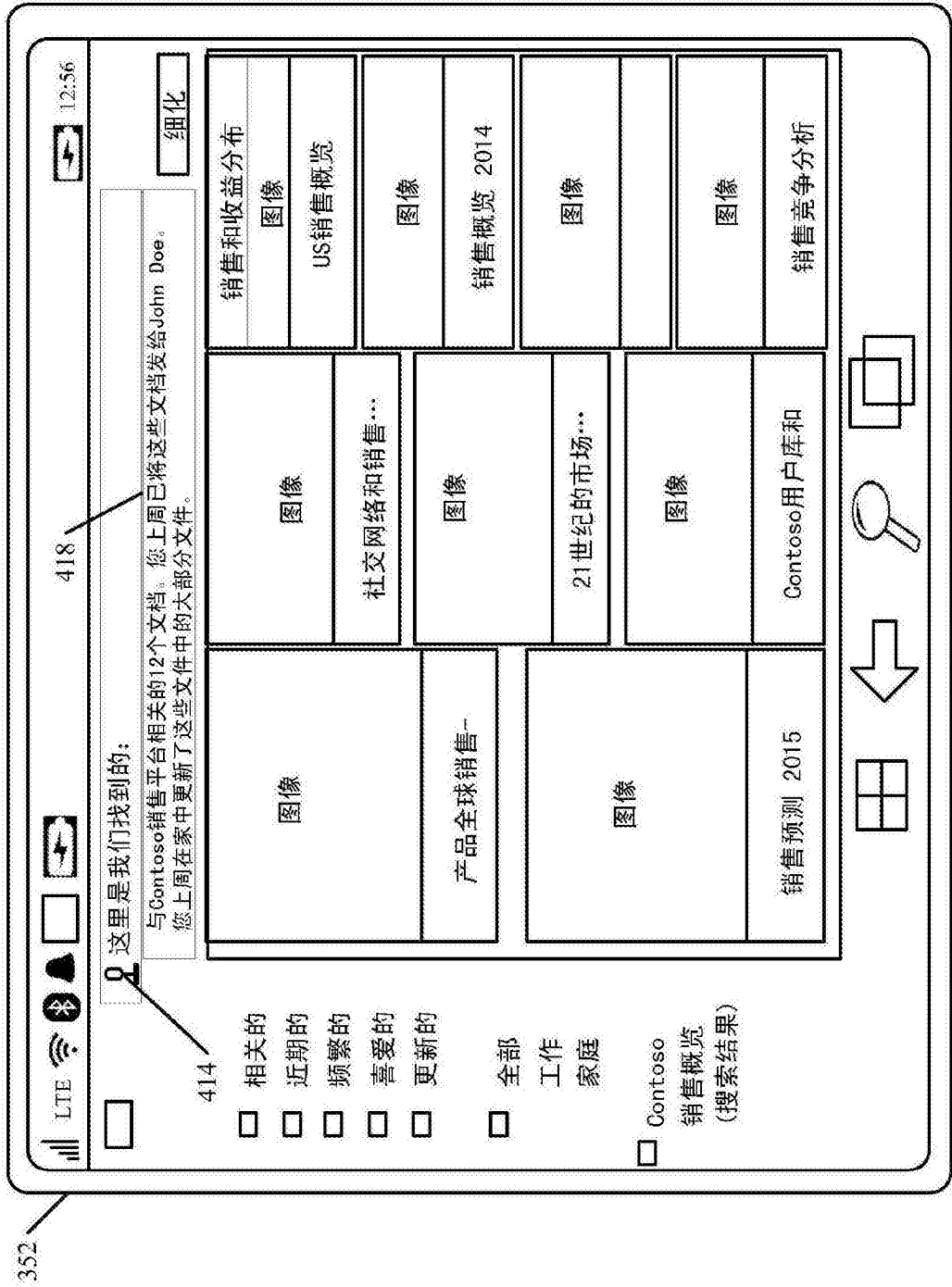


图4D

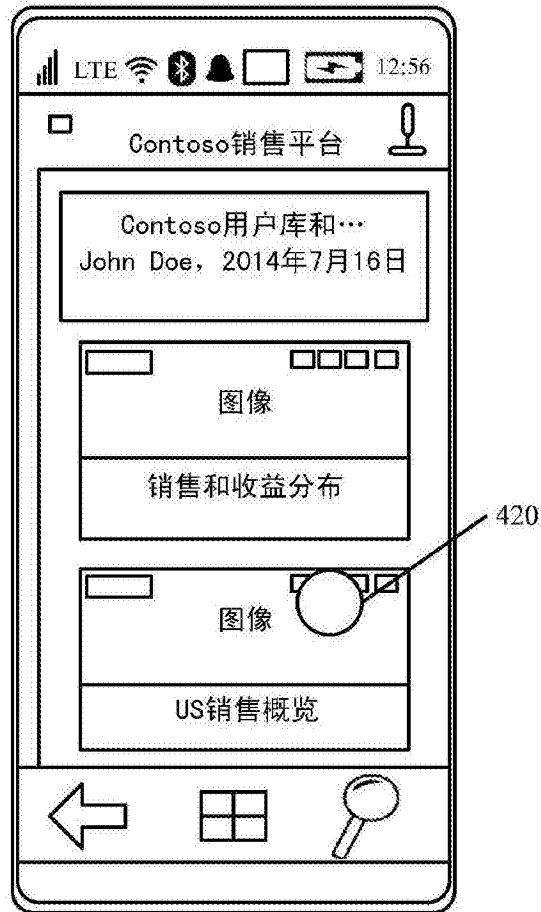


图4E

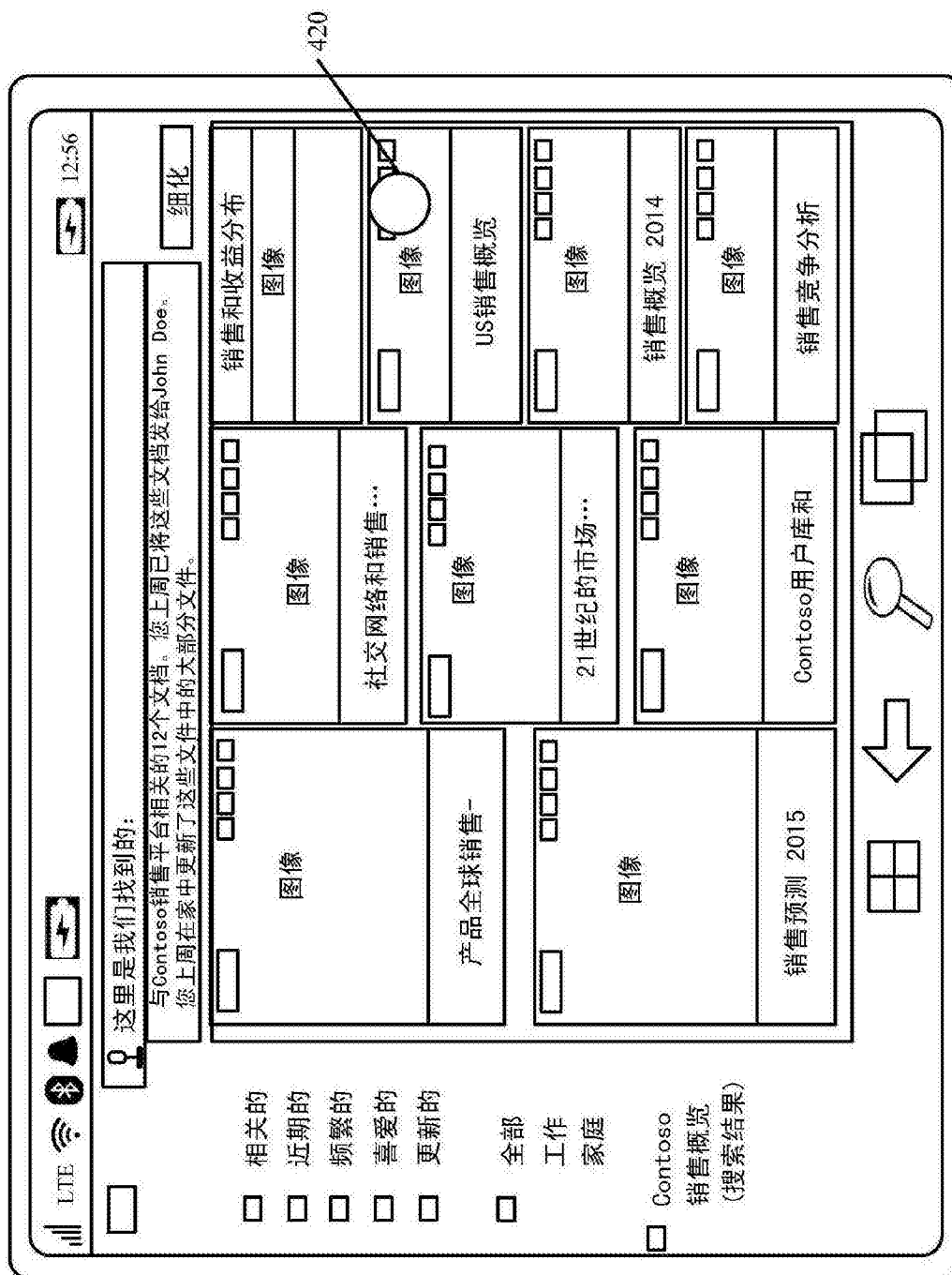


图4F

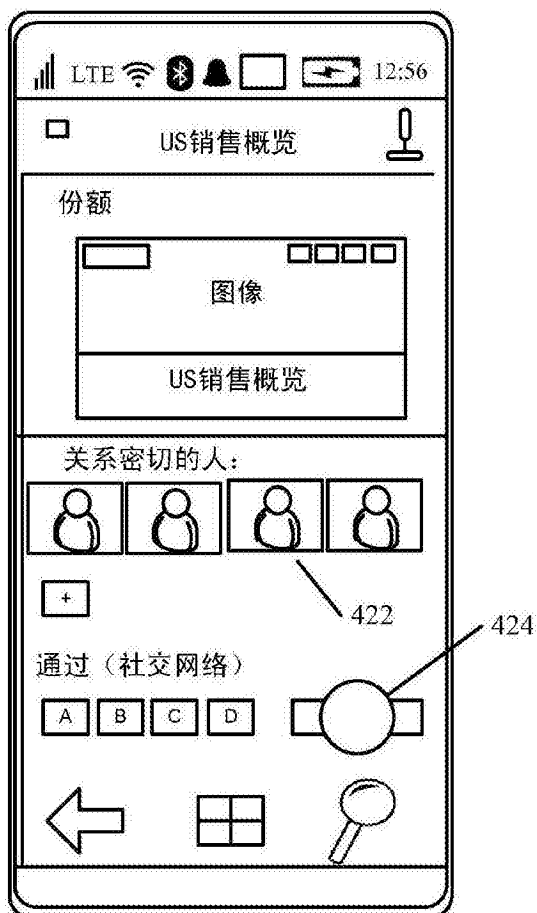


图4G

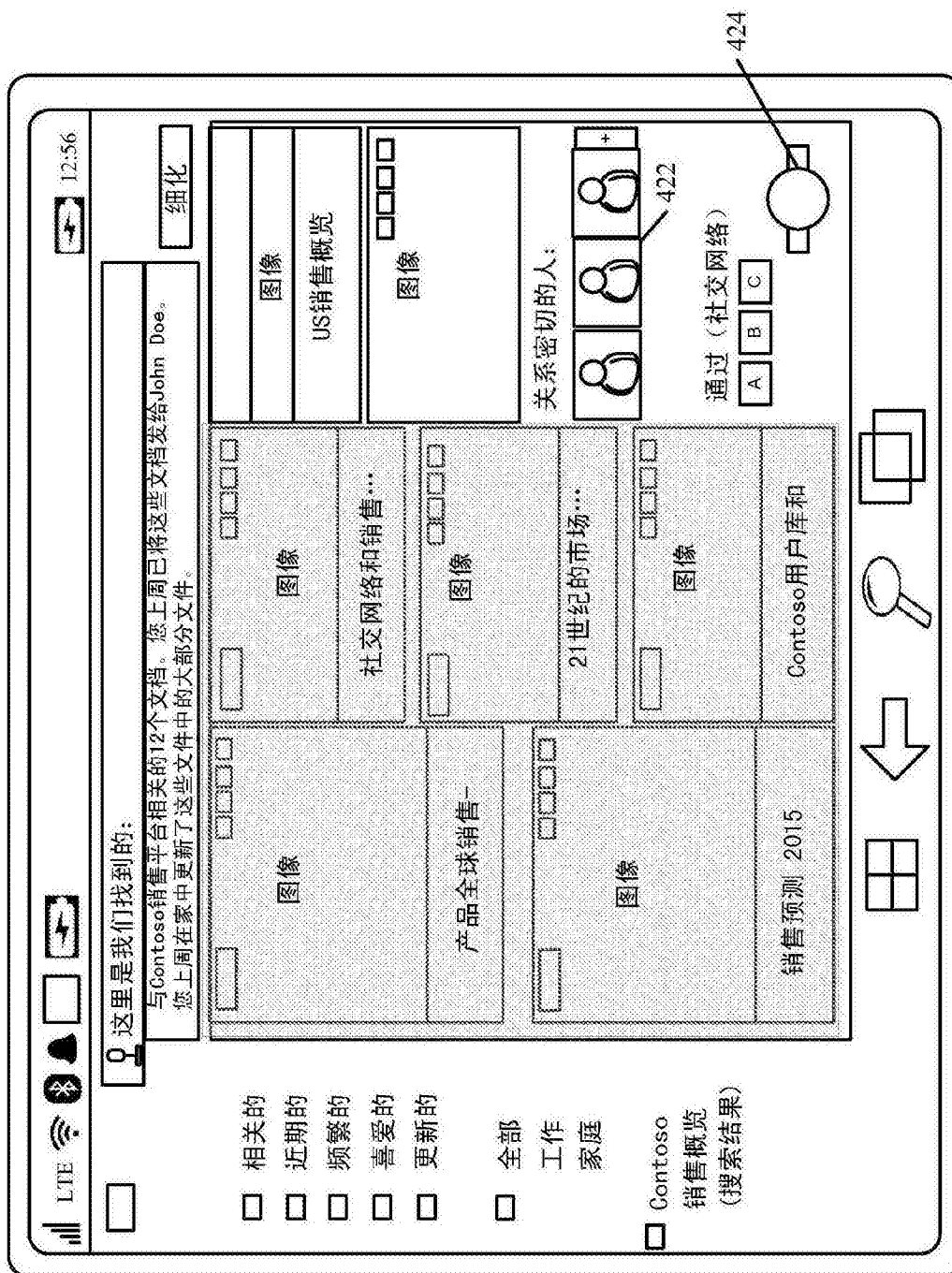


图4H

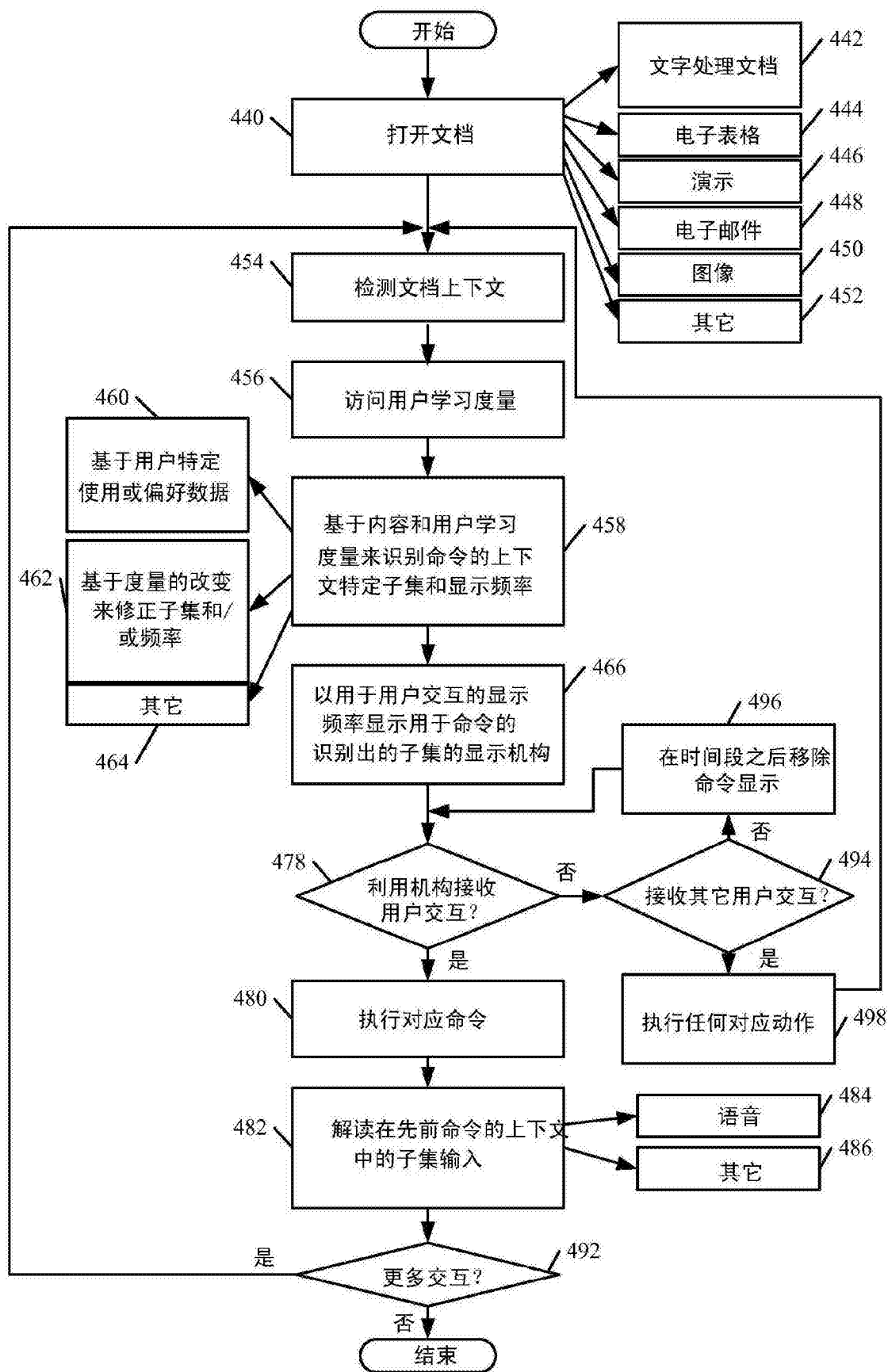


图5

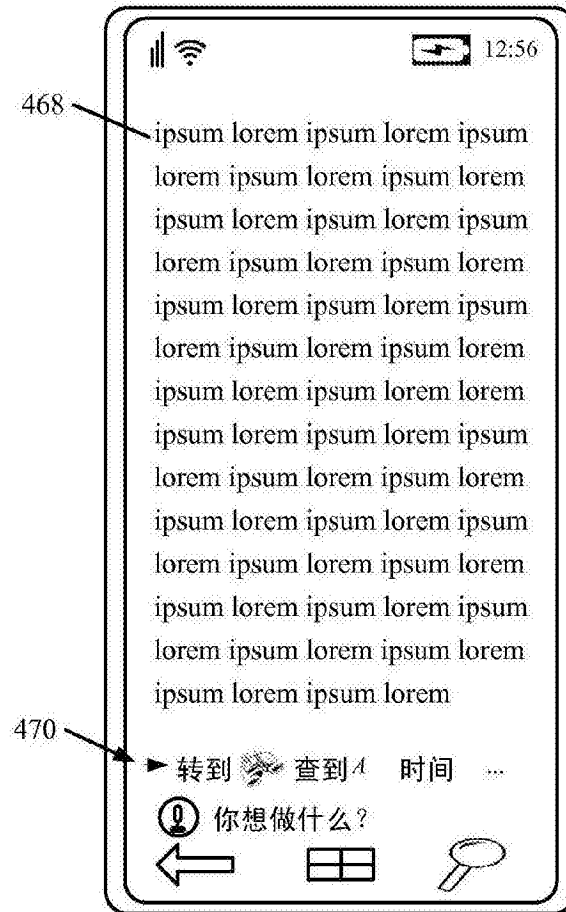


图5A

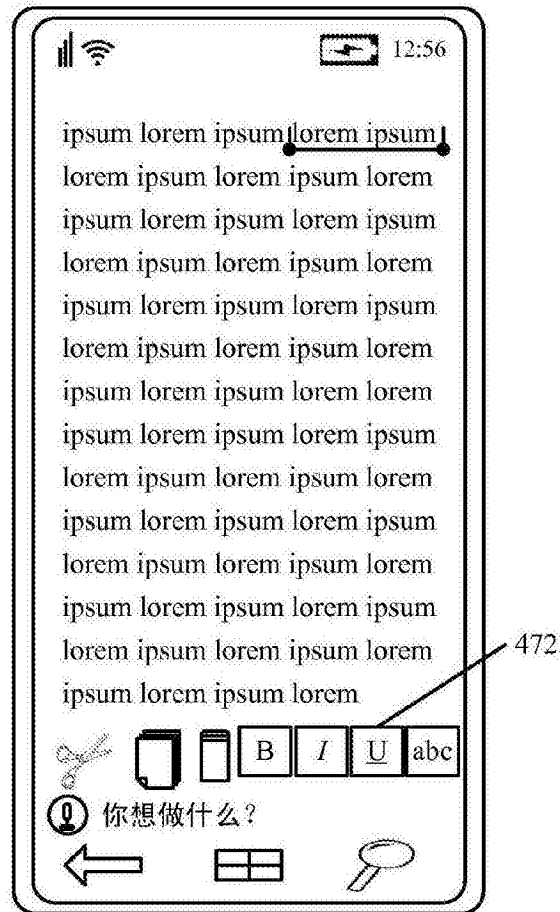


图5B

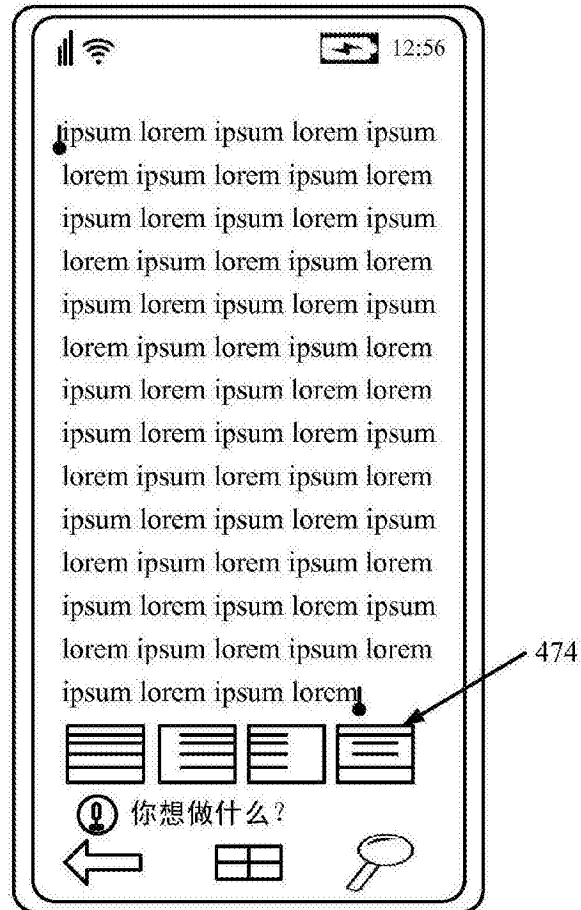


图5C

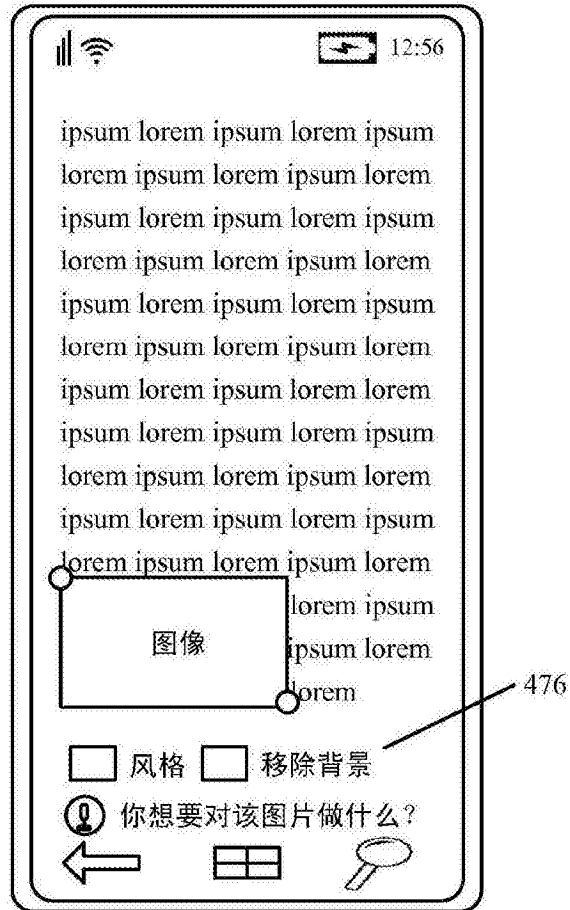


图5D

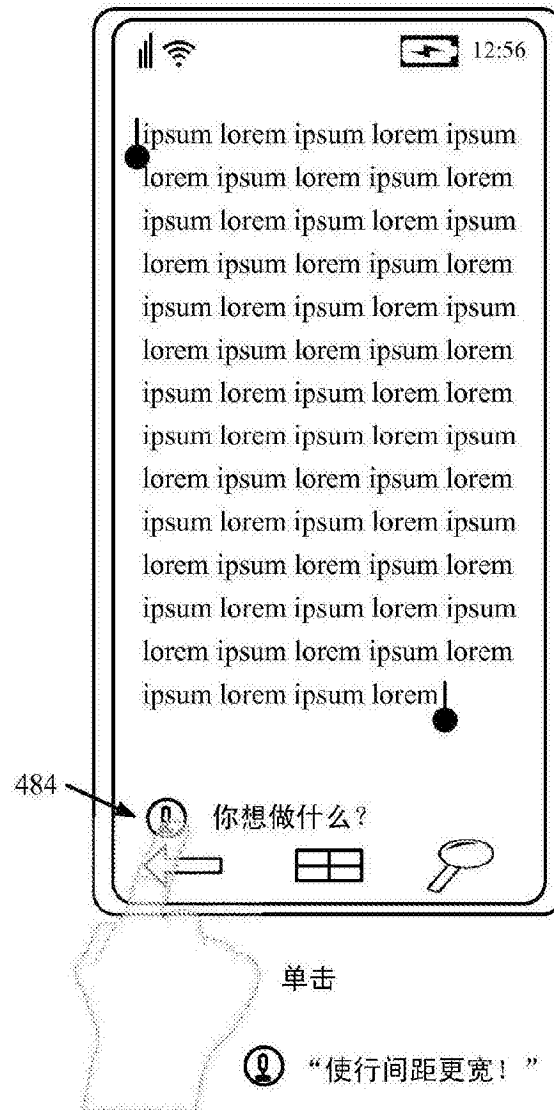


图5E

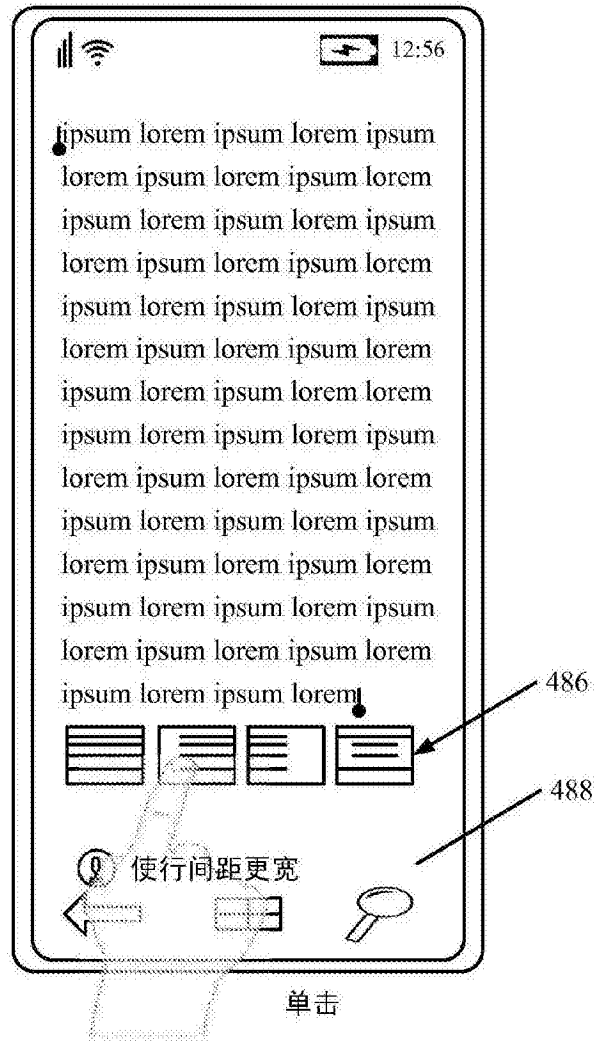


图5F

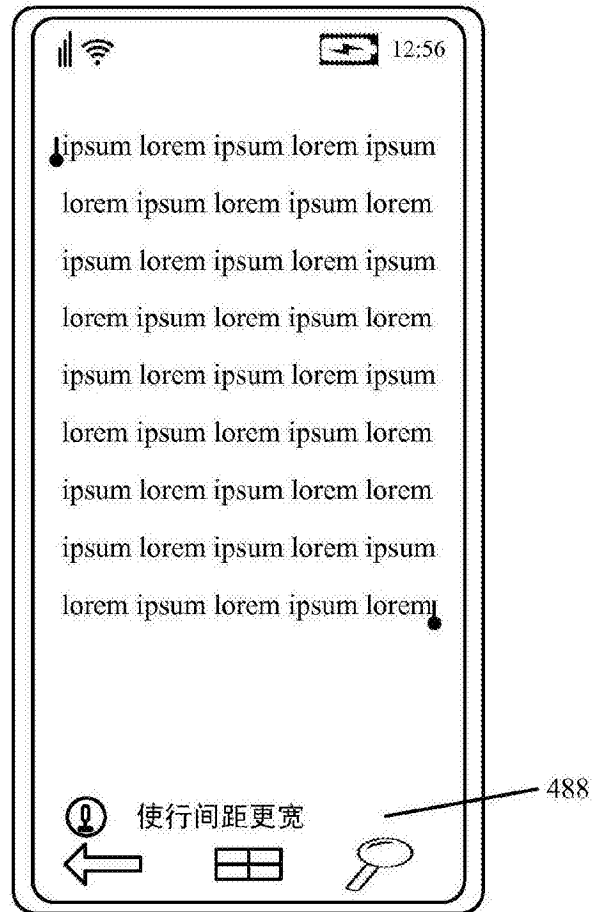
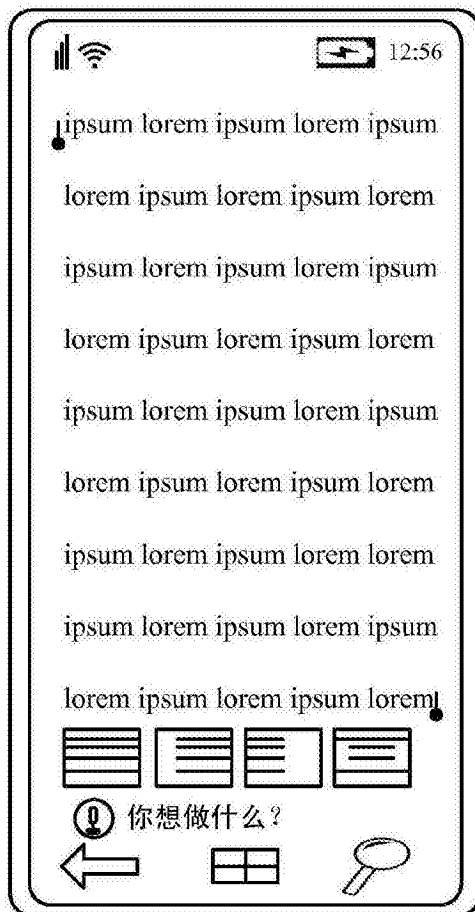


图5G



“更多”

图5H

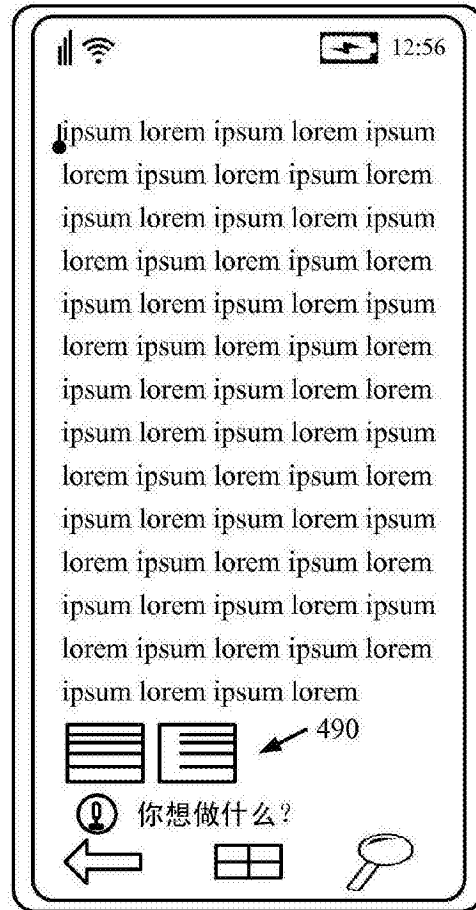


图5I

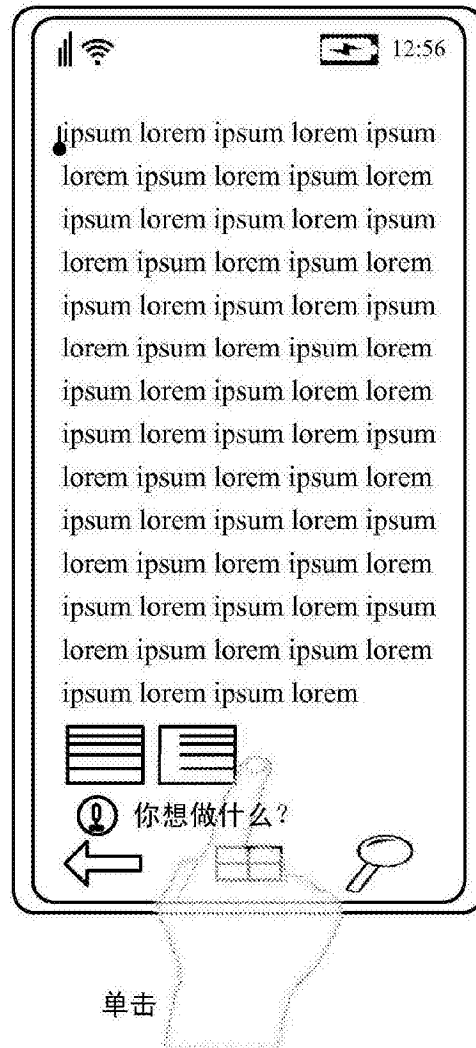


图5J

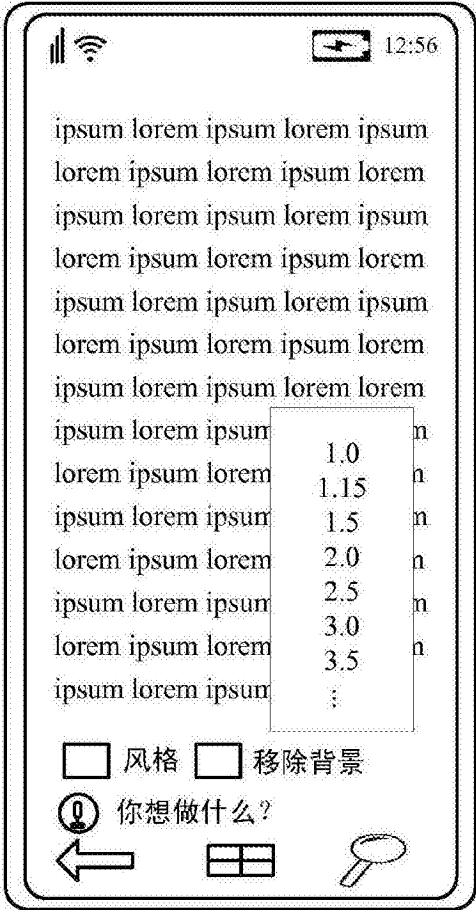
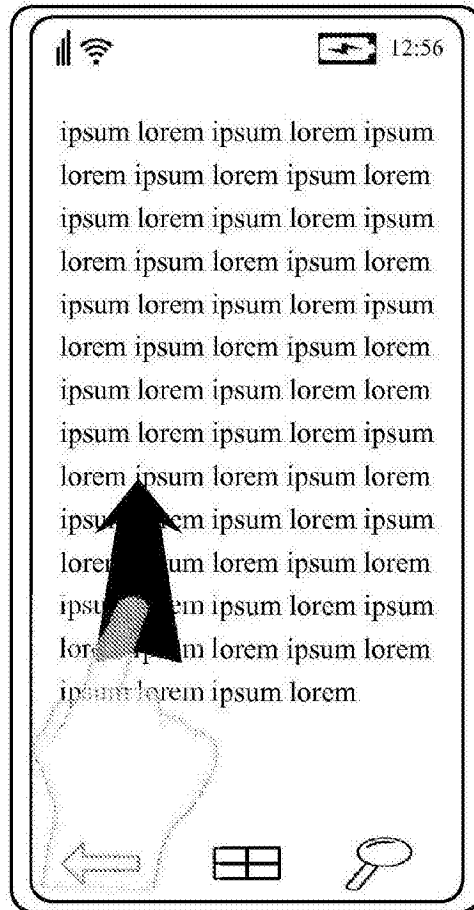
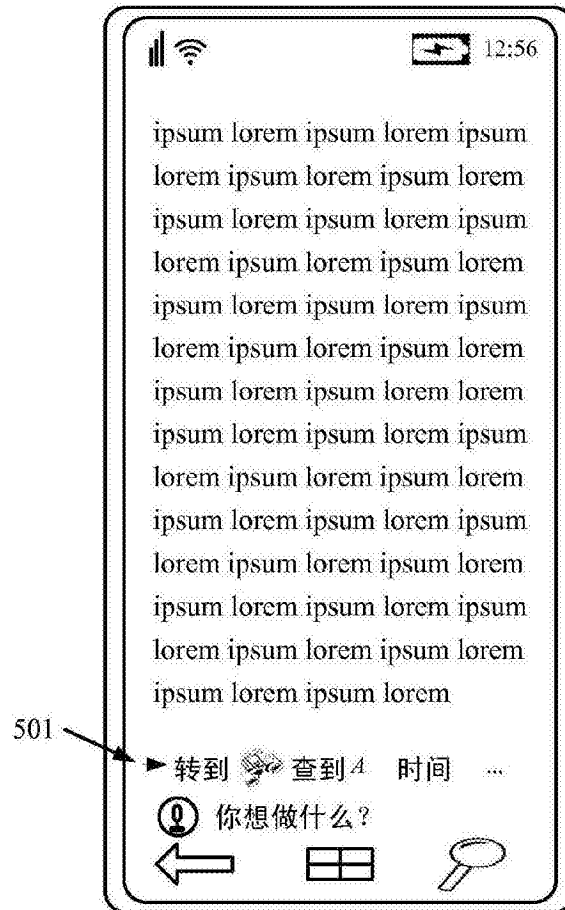


图5K



滚动，滚动，滚动

图5L



预测UI出现

图5M

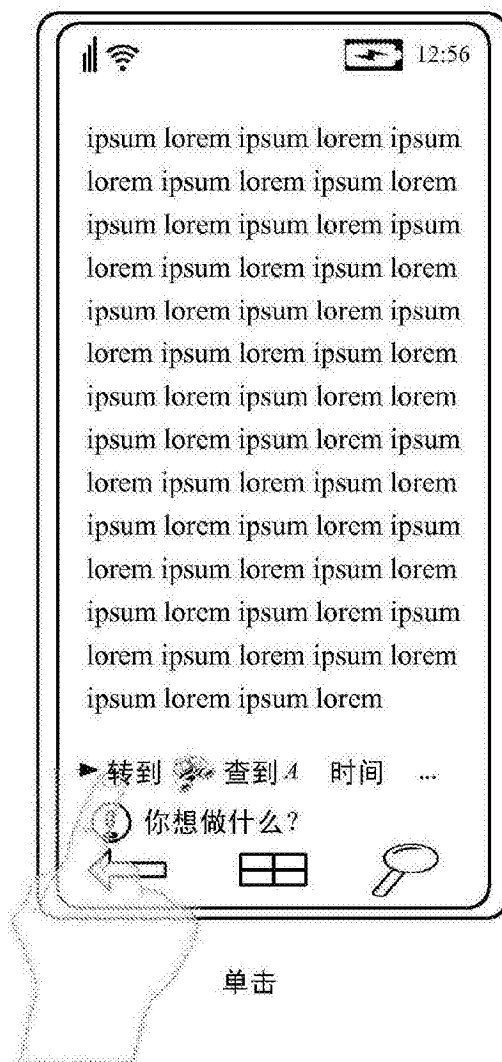


图5N

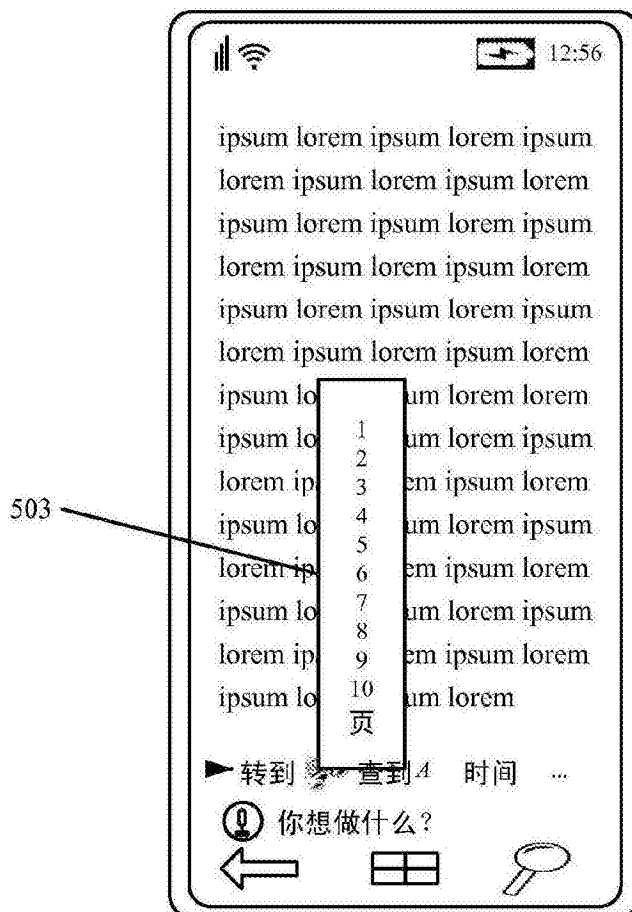


图50

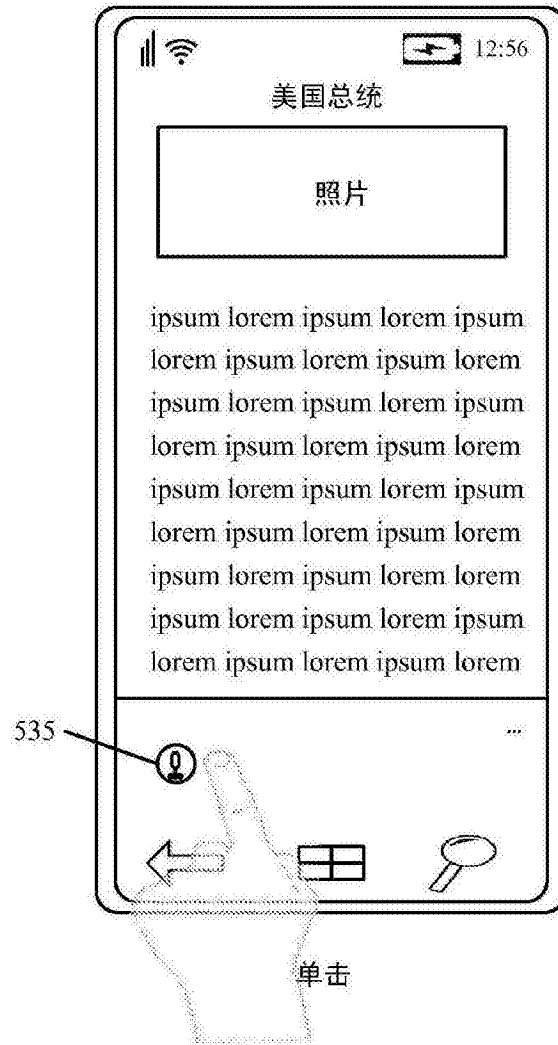
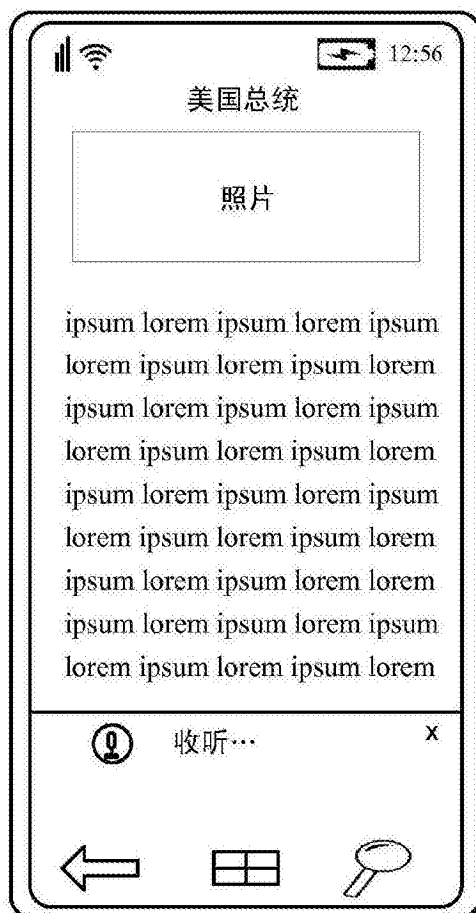


图5P



“我想要格式化我的全部图片”

图5Q

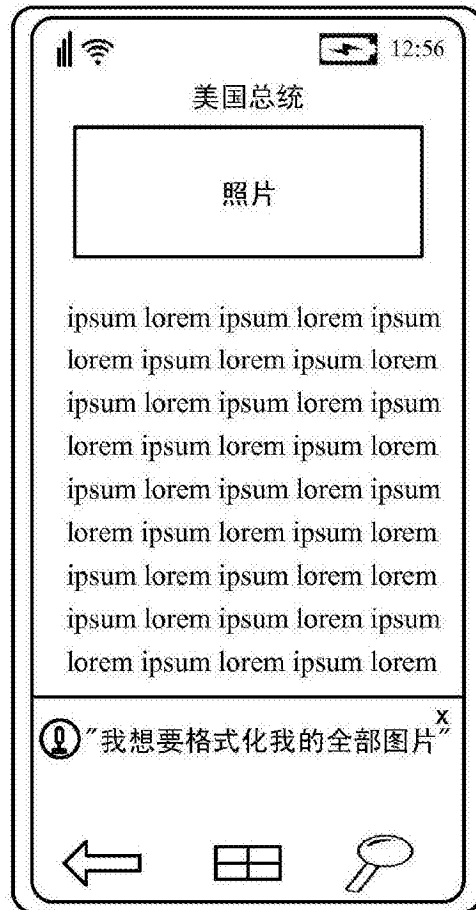


图5R

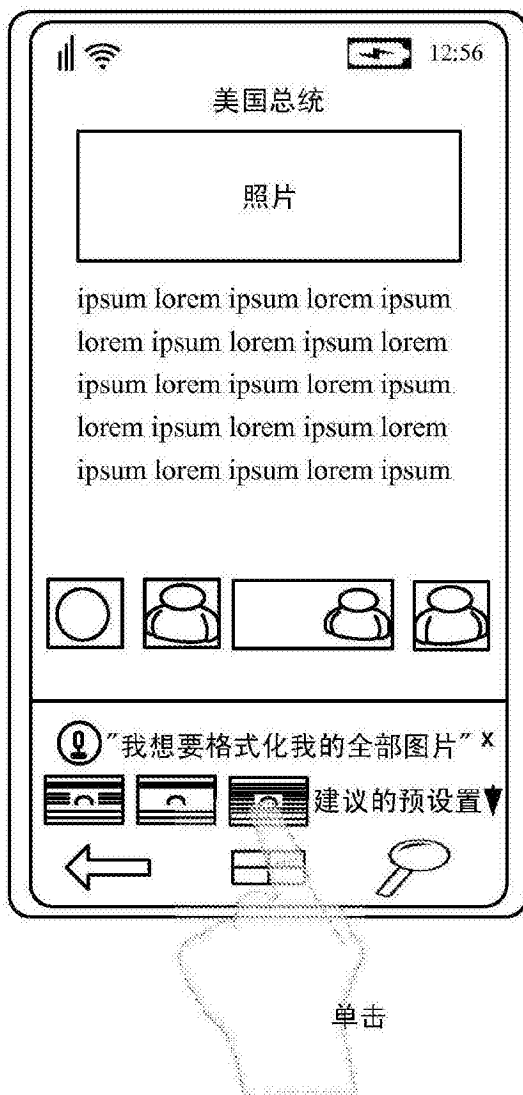


图5S

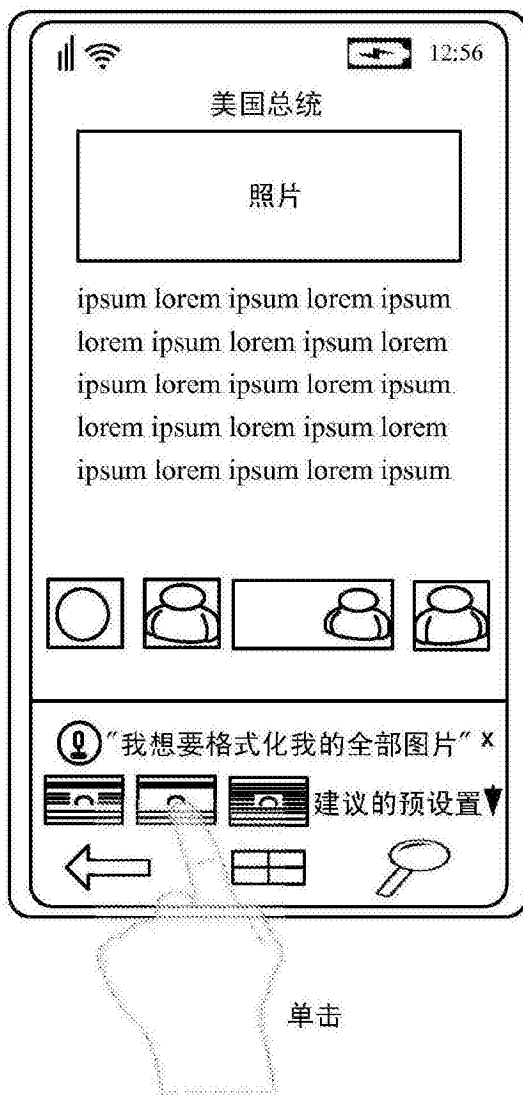
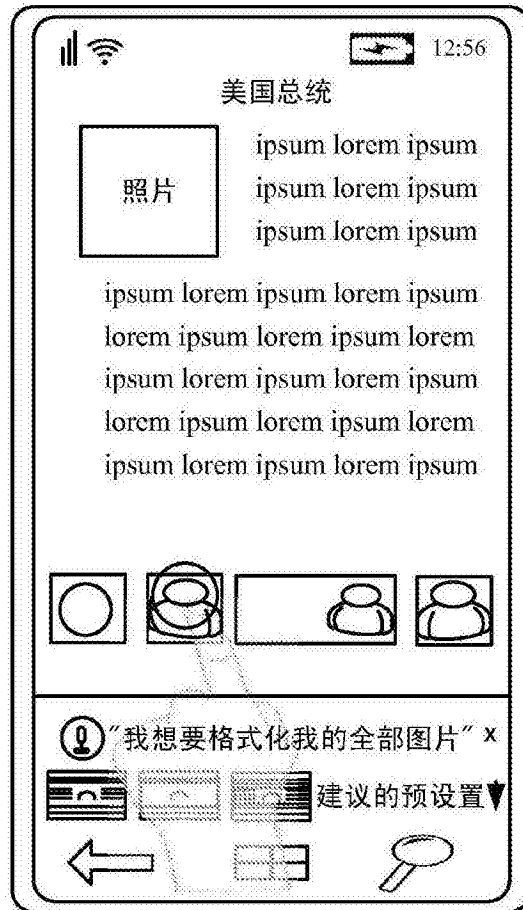


图5T



点击且保持

图5U

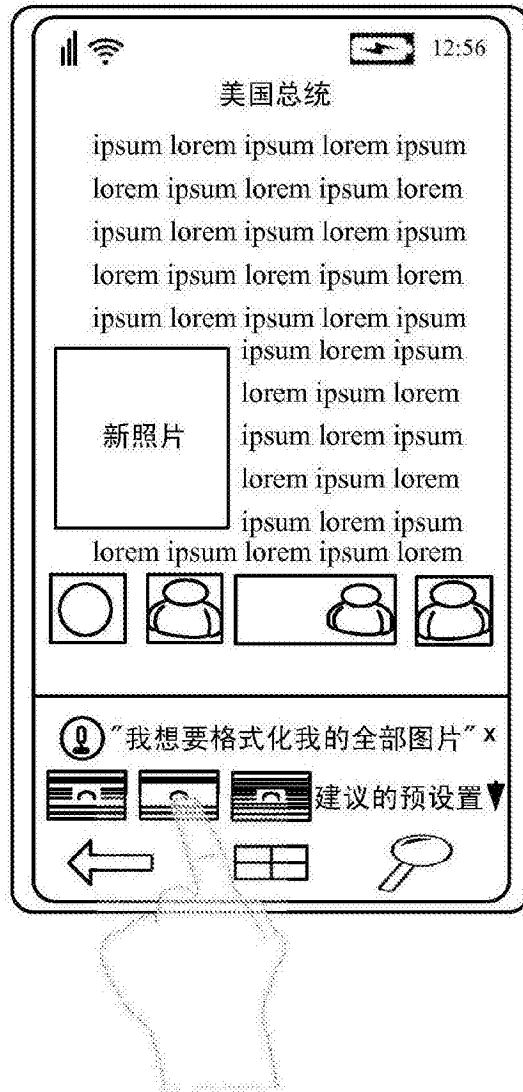


图5V

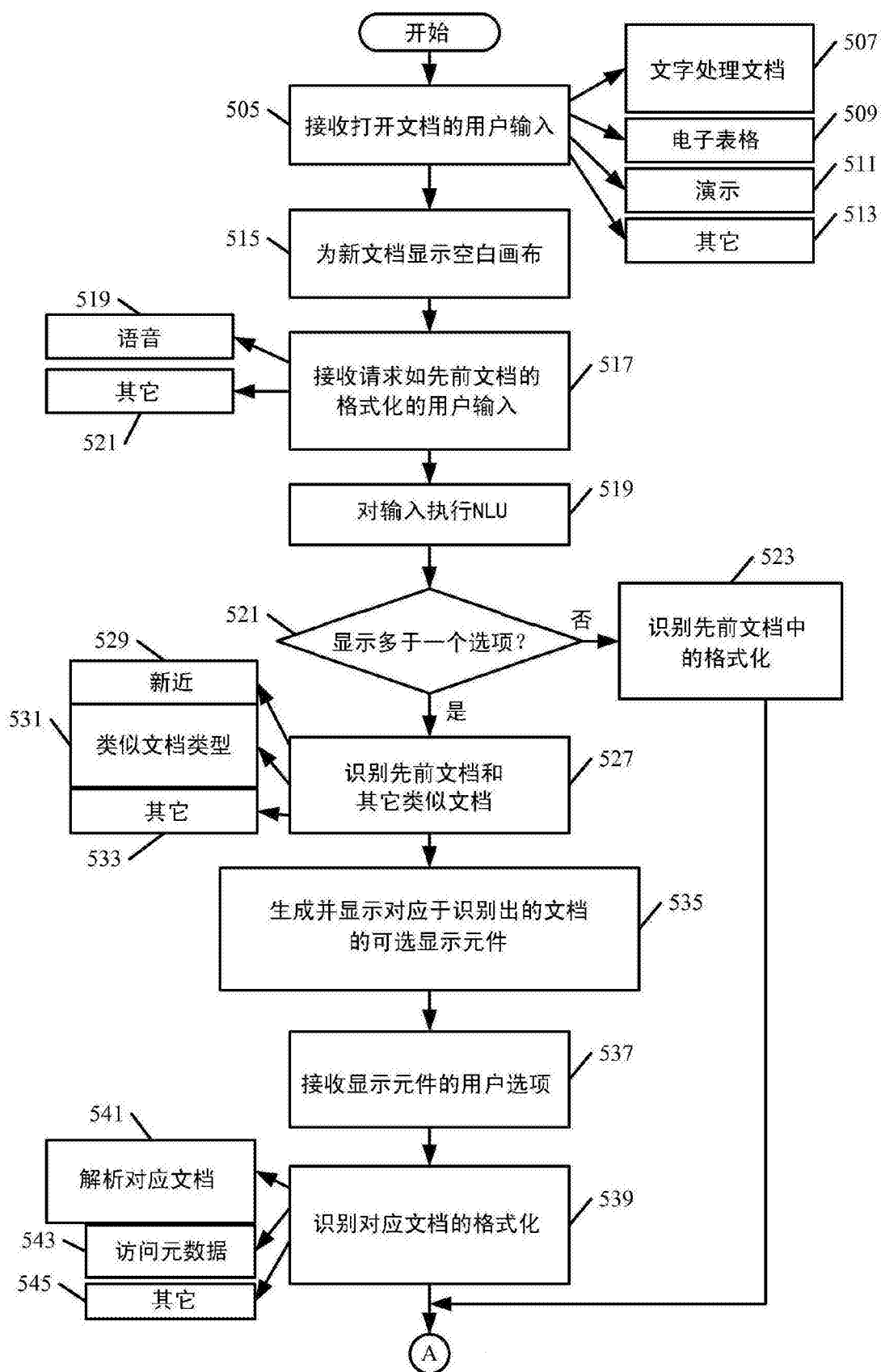


图6A

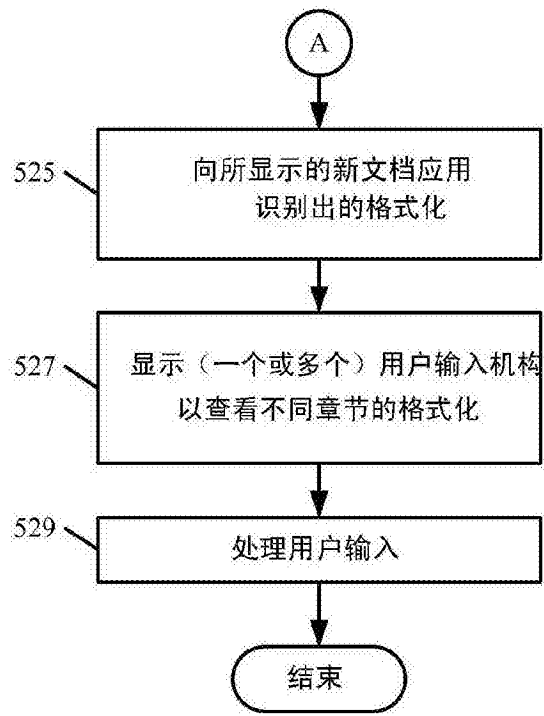


图6B

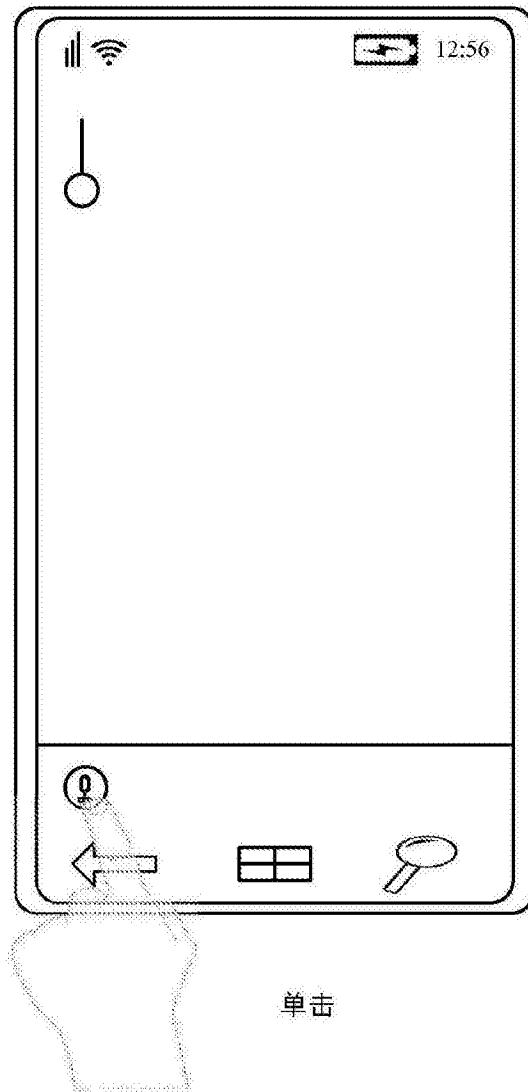
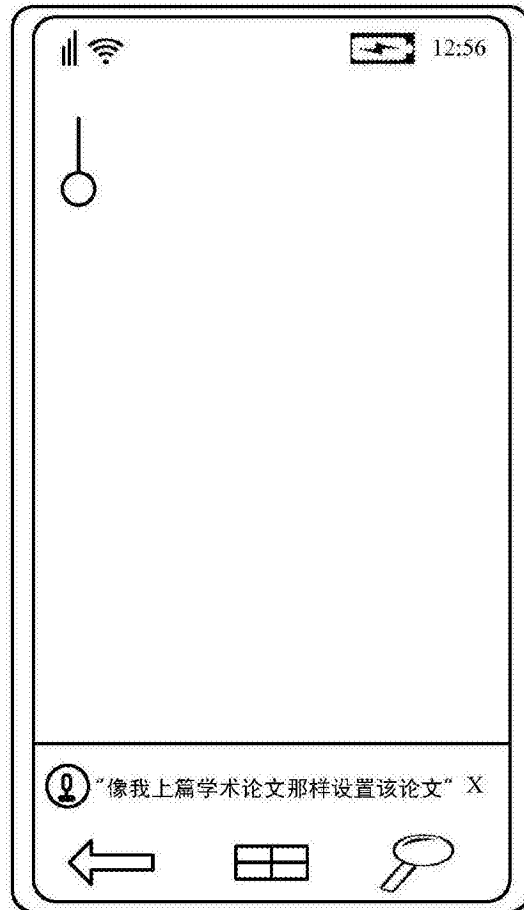
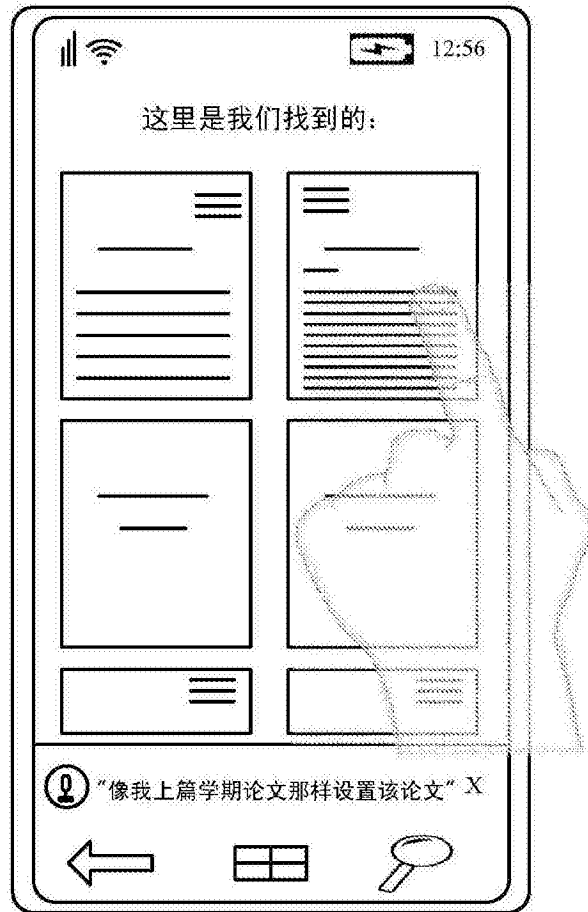


图6C



① “像我上篇学术论文
那样设置该论文”

图6D

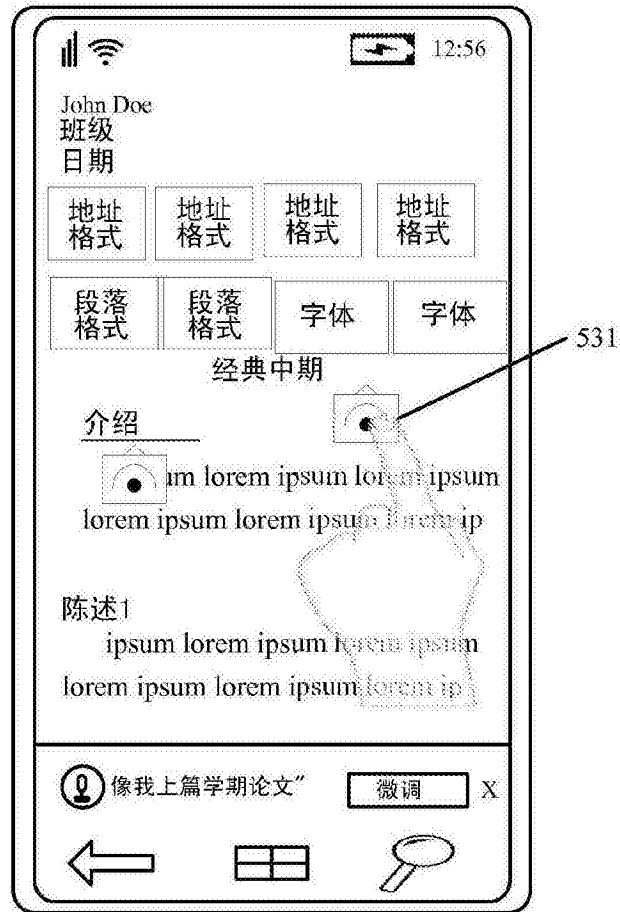


① “像我上篇学期论文那样设置该论文”

图6E

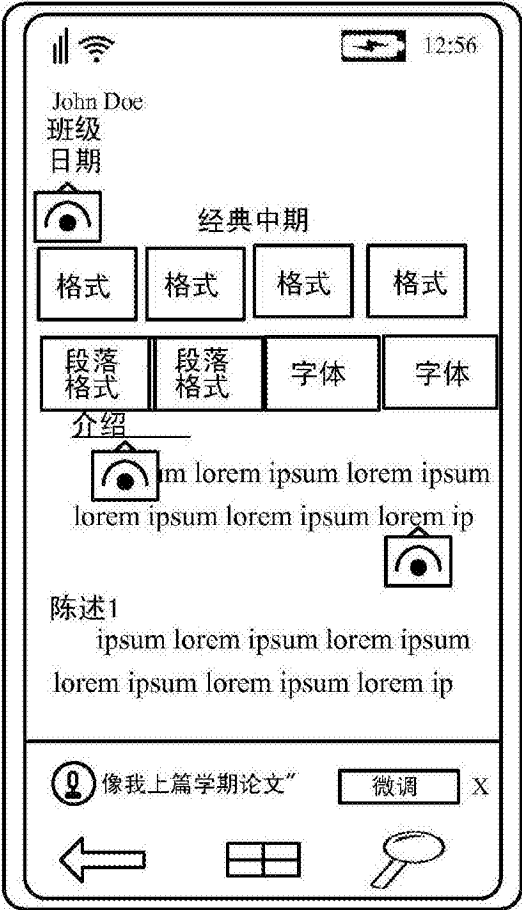


图6F



单击

图6G



单击

图6H

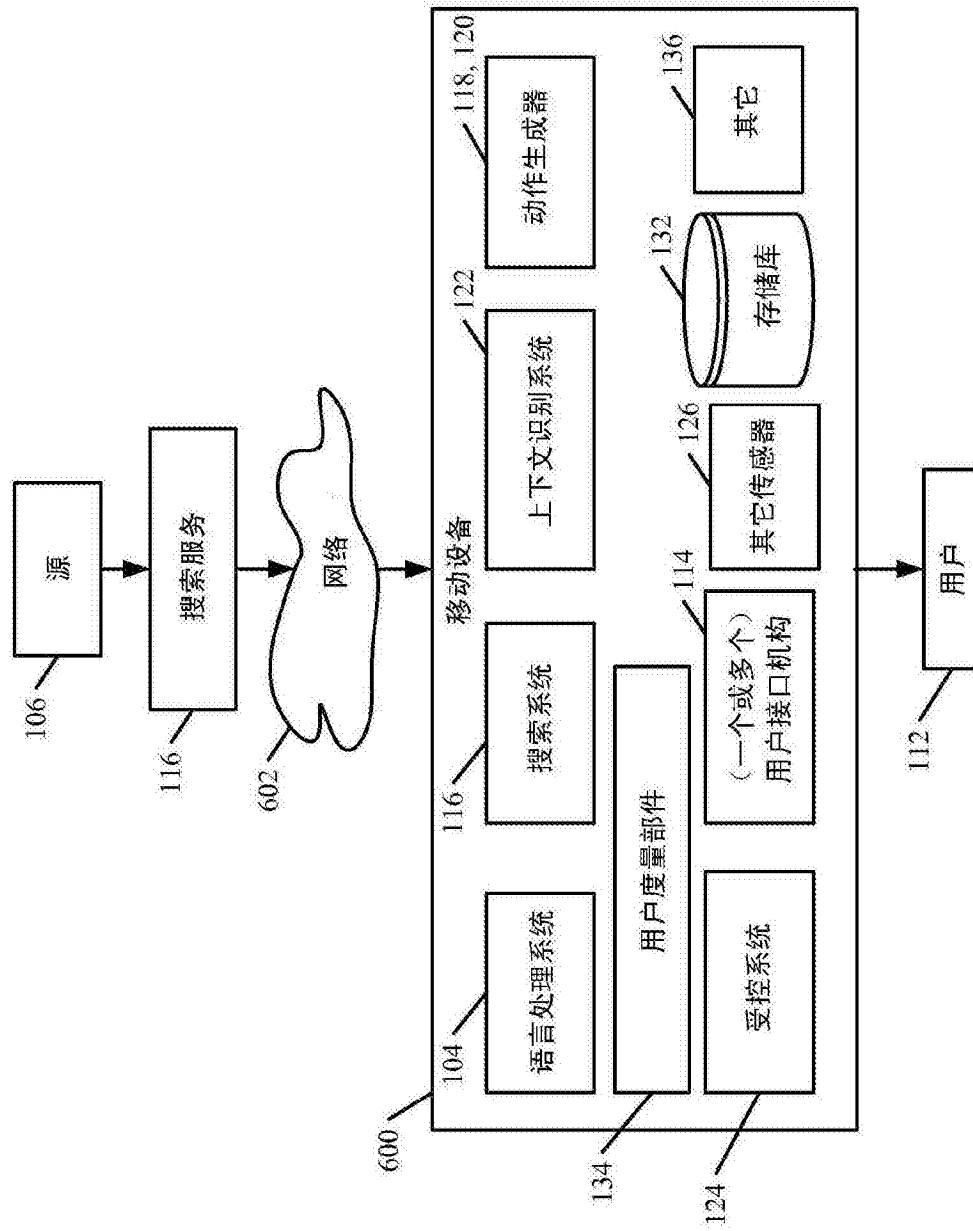


图7

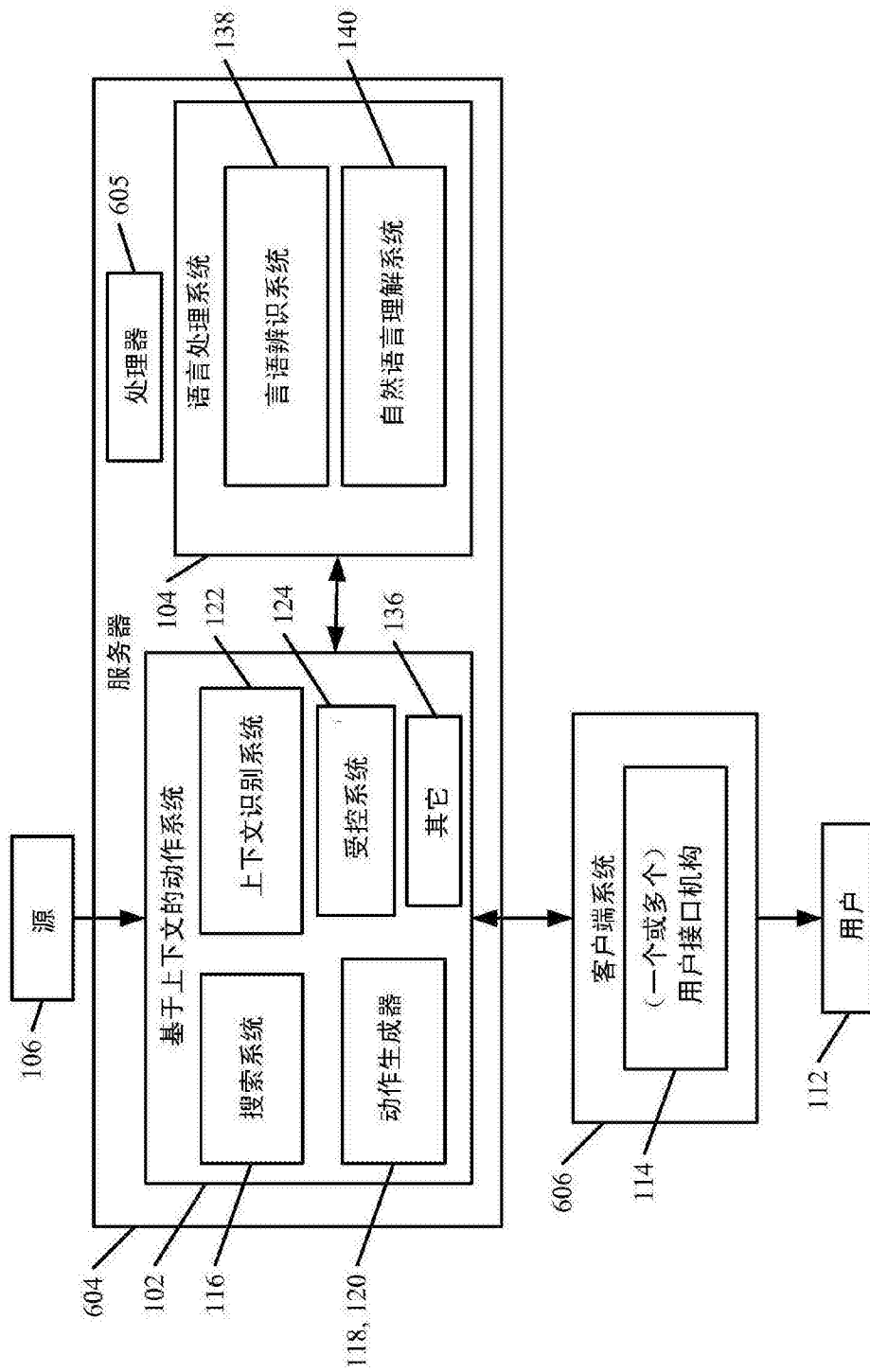


图8

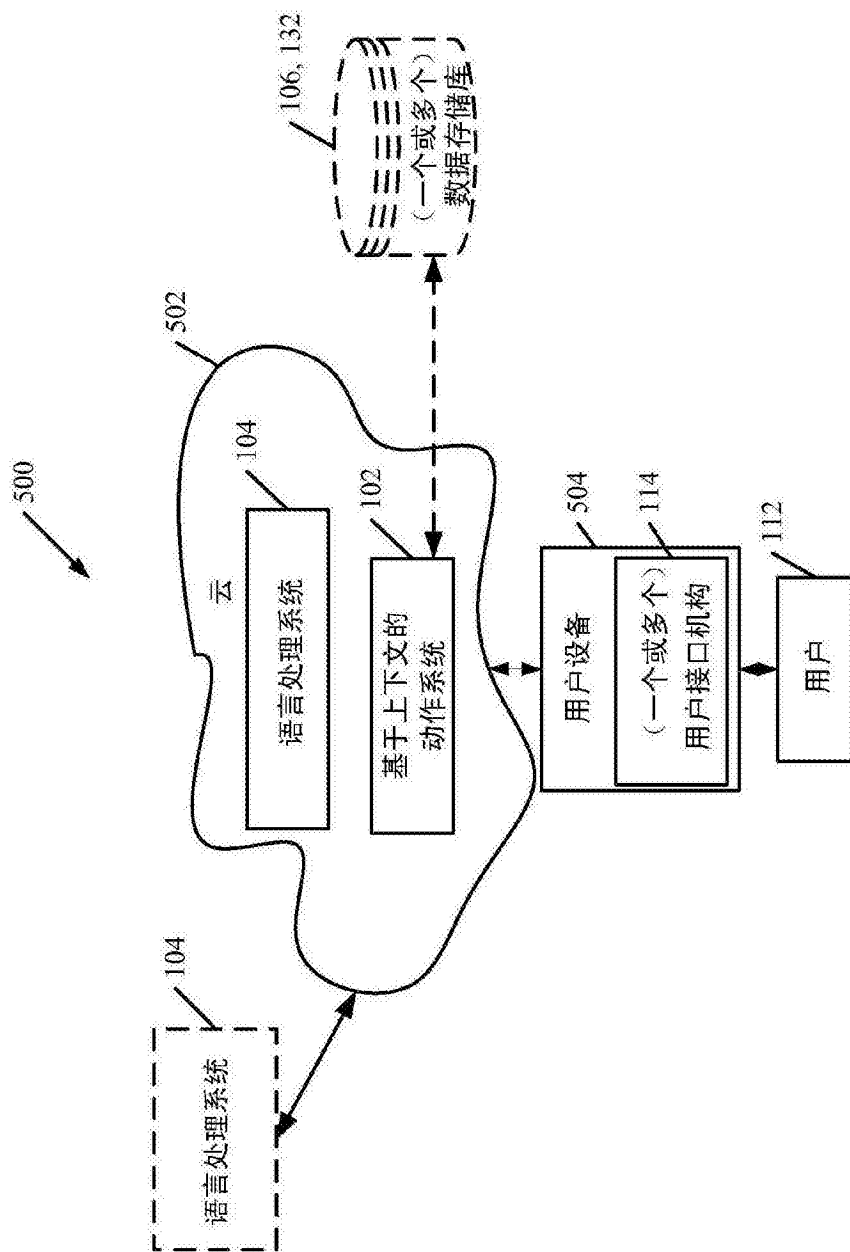


图9

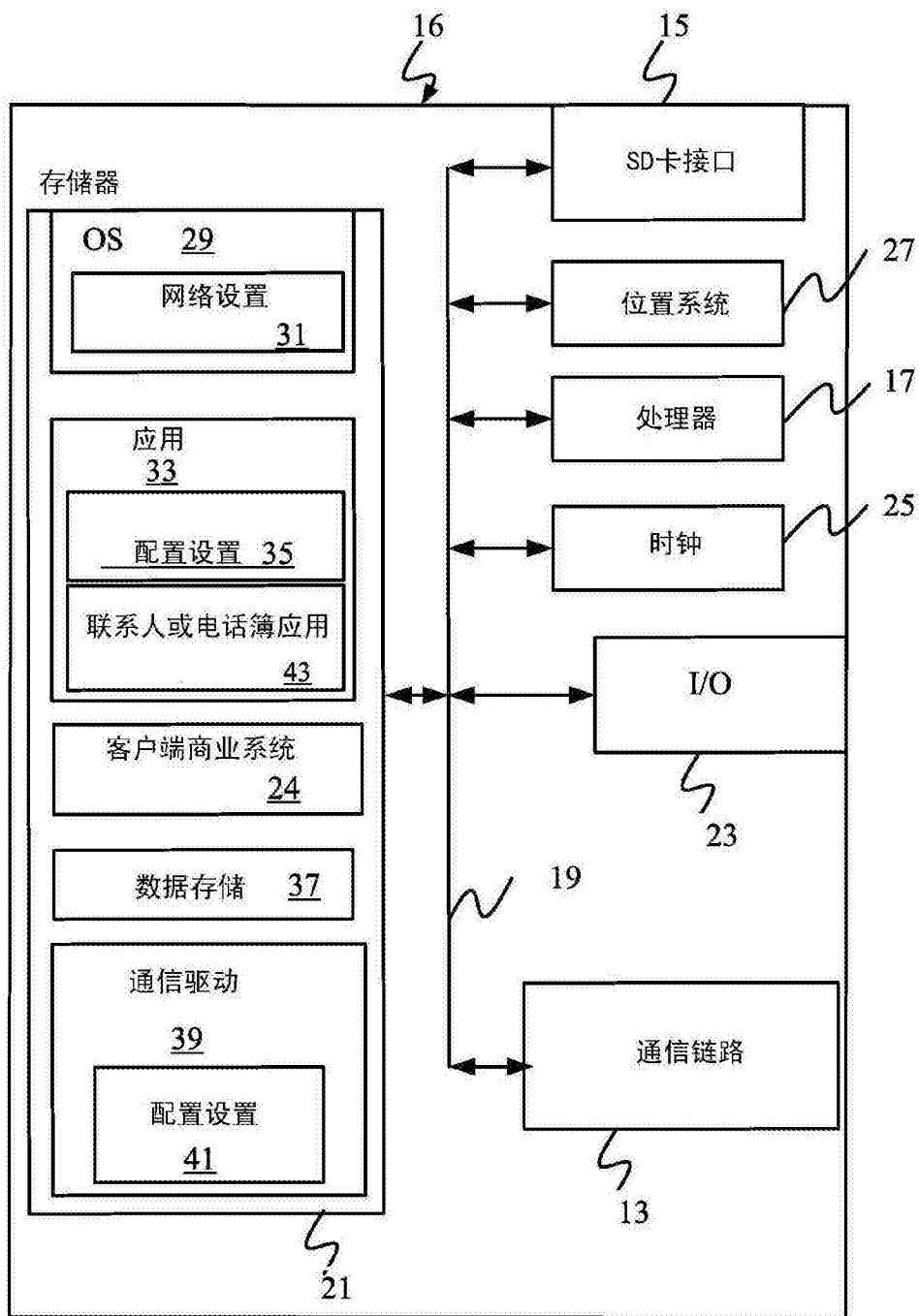


图10

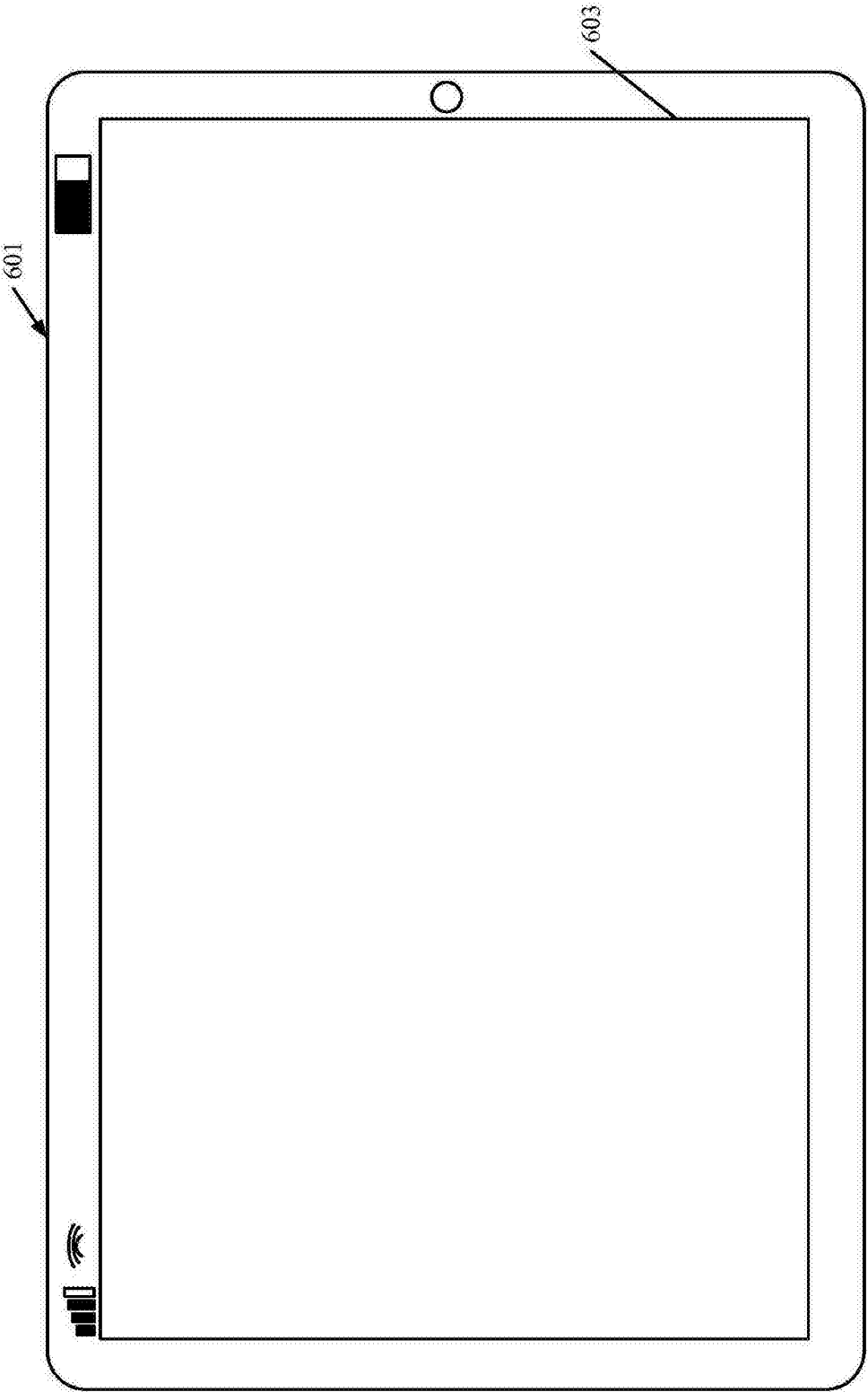


图11

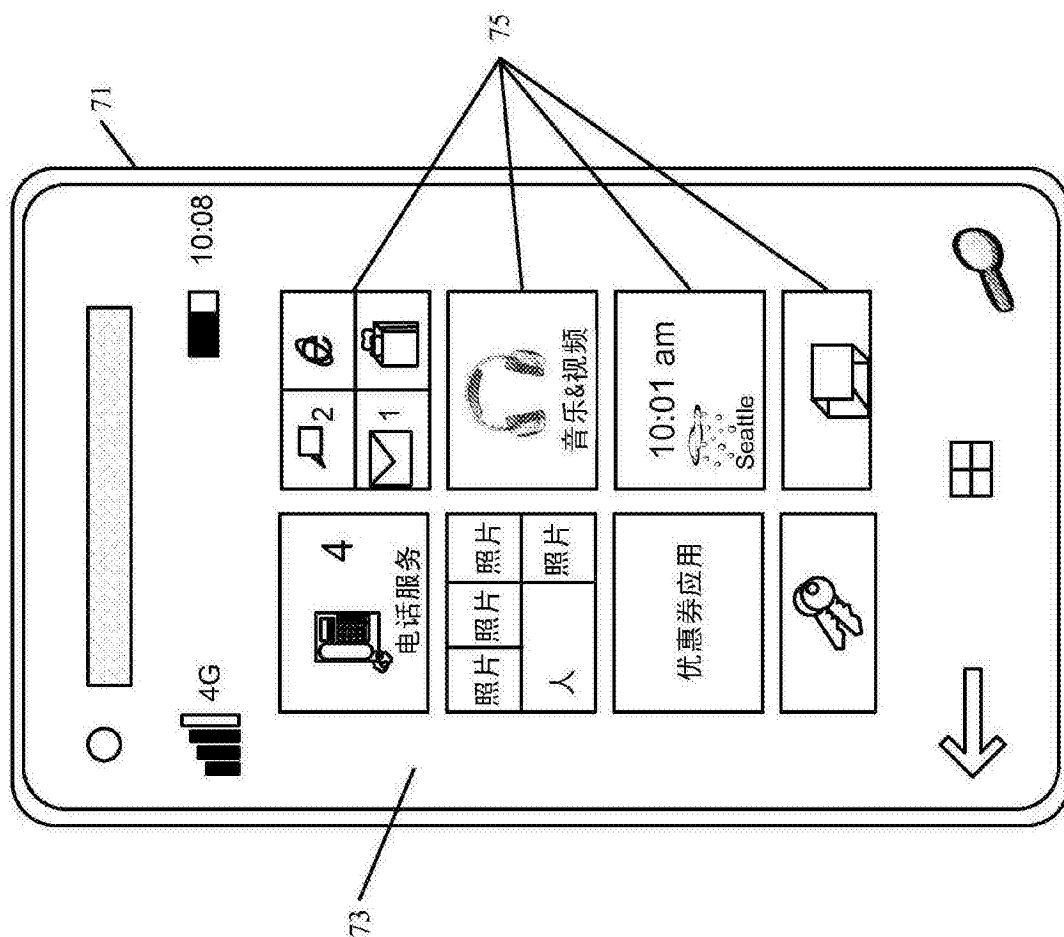


图12

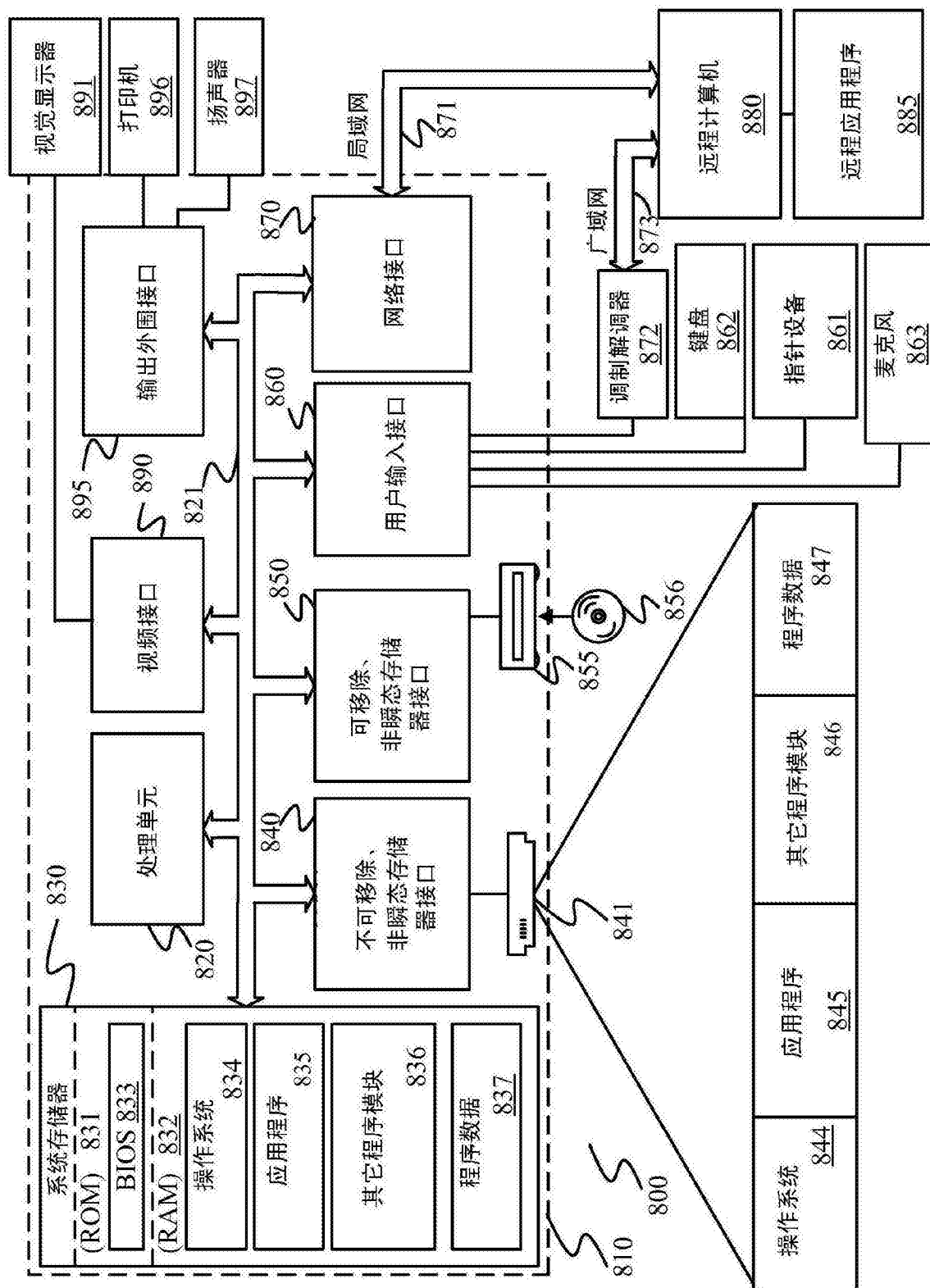


图13