



(12) 发明专利

(10) 授权公告号 CN 109690480 B

(45) 授权公告日 2021.12.17

(21) 申请号 201780054824.9  
 (22) 申请日 2017.08.23  
 (65) 同一申请的已公布的文献号  
 申请公布号 CN 109690480 A  
 (43) 申请公布日 2019.04.26  
 (30) 优先权数据  
 62/384,500 2016.09.07 US  
 15/297,790 2016.10.19 US  
 (85) PCT国际申请进入国家阶段日  
 2019.03.06  
 (86) PCT国际申请的申请数据  
 PCT/US2017/048106 2017.08.23  
 (87) PCT国际申请的公布数据  
 W02018/048620 EN 2018.03.15  
 (73) 专利权人 微软技术许可有限责任公司  
 地址 美国华盛顿州  
 (72) 发明人 O·Z·克汉 R·萨里卡亚  
 D·杰特利  
 (74) 专利代理机构 北京市金杜律师事务所  
 11256  
 代理人 王茂华 彭梦晔

(51) Int.Cl.  
 G06F 9/451 (2018.01)  
 G06F 16/332 (2019.01)  
 G06F 40/30 (2020.01)  
 G10L 15/19 (2013.01)  
 G10L 15/06 (2013.01)  
 G10L 15/22 (2006.01)  
 (56) 对比文件  
 CN 102750270 A, 2012.10.24  
 CN 105702248 A, 2016.06.22  
 US 2016/0042735 A1, 2016.02.11  
 CN 104769584 A, 2015.07.08  
 CN 101493812 A, 2009.07.29  
 US 2004/0044517 A1, 2004.03.04  
 US 9009046 B1, 2015.04.14

谢炯坤.“面向人机互动的自然语言理解的研究”.《中国博士学位论文全文数据库 信息科技辑》.2015, I138-71.

J. Chai.etc.“Context-based multimodal input understanding in conversational systems”.《Proceedings. Fourth IEEE International Conference on Multimodal Interfaces》.2003,全文.

审查员 葛运滨

权利要求书3页 说明书14页 附图19页

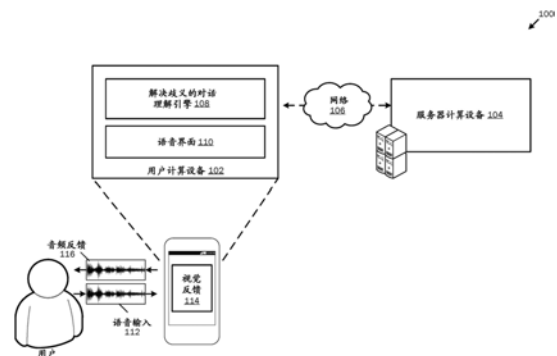
(54) 发明名称  
 解决歧义的对话理解系统

(57) 摘要

提供了解决对话理解系统中的歧义的系统、部件、设备、以及方法。一个非限制性示例是一种用于解决对话理解系统中的歧义的系统或方法。方法包括接收自然语言输入并基于自然语言输入来标识代理动作的步骤。方法还包括确定与代理动作相关联的歧义值并针对歧义条件评估歧义值的步骤。方法包括当确定歧义值满足歧义条件时的以下步骤：基于与所标识的代理动作相关联的歧义来选择提示动作；执行提示动作；响应

于提示动作而接收额外的输入；以及基于额外的输入来更新代理动作以解决歧义。方法还包括执行代理动作的步骤。

CN 109690480 B



1. 一种基于网络的解决歧义的对话理解系统,所述系统包括:  
至少一个处理器;以及  
存储器,其操作地连接到所述至少一个处理器并且存储指令,所述指令当由所述至少一个处理器运行时使所述至少一个处理器:  
经由网络接收来自用户计算设备的自然语言输入;  
至少部分基于所述自然语言输入来标识代理动作;  
确定所述代理动作中的歧义;  
确定与所述代理动作中的所述歧义相关联的歧义值;  
针对与所述代理动作相关联的危害性水平评估所述歧义值,所述危害性水平对应于执行不正确的代理动作的后果严重性,其中当所述代理动作可逆时,所述危害性水平降低;  
当确定所述歧义值满足所述危害性水平时:  
至少部分基于所述代理动作中的所述歧义来选择提示动作,其中所述提示动作包括由所述用户计算设备的音频组件输出的音频反馈和在所述用户计算设备的显示器上提供的视觉反馈中的一个或多个;  
通过所述用户计算设备来启动对所述提示动作的执行;  
响应于所述提示动作而接收额外的输入;以及  
至少部分基于所述额外的输入来更新所述代理动作以解决所述歧义;以及  
启动所述代理动作的执行。
2. 根据权利要求1所述的系统,其中所述歧义值至少部分基于至少两个歧义源。
3. 根据权利要求1所述的系统,其中所述自然语言输入是语音输入。
4. 根据权利要求3所述的系统,还包括用于以下的指令:  
确定与所述语音输入的自动语音识别相关联的歧义值。
5. 根据权利要求1所述的系统,还包括用于以下的指令:  
至少部分基于所述自然语言输入来执行语言理解以将含义归属给所述自然语言输入;  
至少部分基于被归属给所述自然语言输入的所述含义来从数据源中检索信息;以及  
至少部分基于被归属给所述自然语言输入的所述含义和所检索的所述信息来确定所述代理动作。
6. 根据权利要求5所述的系统,还包括用于以下的指令:  
访问与当前任务相关联的任务状态信息;以及  
至少部分基于所述任务状态信息来确定所述代理动作。
7. 根据权利要求1所述的系统,还包括用于以下的指令:  
将执行所述代理动作的结果发送到所述用户计算设备。
8. 根据权利要求1所述的系统,还包括用于以下的指令:  
将所述代理动作发送到所述用户计算设备以用于由所述用户计算设备执行。
9. 根据权利要求1所述的系统,还包括用于以下的指令:  
将所述提示动作发送到所述用户计算设备。
10. 根据权利要求1所述的系统,其中当所述视觉反馈和所述音频反馈两者都被所述用户计算设备输出时,所述视觉反馈被用于澄清所述音频反馈。
11. 根据权利要求1所述的系统,其中至少部分基于所述自然语言输入来标识所述代理

动作还包括访问与当前任务相关联的任务状态信息并至少部分基于所述任务状态信息来确定所述代理动作。

12. 根据权利要求1所述的系统,其中接收所述自然语言输入包括接收文本输入。

13. 一种用于解决对话理解系统中的歧义的方法,包括:

在计算设备处接收来自用户的语音输入;

对所述语音输入执行自动语音识别;

至少部分基于所述语音输入来标识代理动作;

确定所述代理动中是否存在歧义;

当确定所述代理动中存在歧义时,确定所述代理动作的危害性水平,所述危害性水平至少部分基于采取所述用户不意图的代理动作的一个或多个不意图的后果的严重性,其中当所述代理动作可逆时,所述危害性水平降低;

当确定所述危害性水平高于阈值时:

至少部分基于所述代理动作中的所述歧义来提供提示动作,所述提示动作是由所述计算设备的音频组件输出的音频提示和在所述计算设备的显示器上输出的视觉提示中的一个或多个;

响应于所述提示动作而接收额外的输入;以及

至少部分基于所述额外的输入来更新所述代理动作以解决所述歧义;以及

执行所述代理动作。

14. 根据权利要求13所述的方法,其中确定所述代理动作中是否存在歧义包括基于至少两个歧义源来确定歧义值。

15. 根据权利要求14所述的方法,其中确定所述代理动作中是否存在歧义包括确定与对所述语音输入执行自动语音识别相关联的歧义值。

16. 根据权利要求13所述的方法,其中标识所述代理动作包括:

至少部分基于所述语音输入来执行语言理解以将含义归属给所述语音输入;

至少部分基于被归属给所述语音输入的所述含义来从数据源中检索信息;以及

至少部分基于被归属给所述语音输入的所述含义和所检索的所述信息来确定所述代理动作。

17. 一种计算设备,包括:

至少一个处理器;以及

存储器,其被耦合到所述至少一个处理器并且存储指令,所述指令当由所述至少一个处理器运行时,执行方法,所述方法包括:

由所述计算设备接收初始自然语言输入;

至少部分基于所述初始自然语言输入来标识代理动作;

确定所述代理动中是否存在歧义;

当确定所述代理动中存在歧义时,确定所述代理动作的危害性水平,所述危害性水平至少部分基于执行不意图的代理动作的后果,其中当所述代理动作可逆时,所述危害性水平降低;

当确定所述危害性水平高于阈值时:

至少部分基于所述代理动作中的所述歧义来提供提示动作,所述提示动作是由所述计

算设备的音频组件输出的音频提示和在所述计算设备的显示器上输出的视觉提示中的一个或多个；

接收额外的输入；以及

使用所述额外的输入来更新所述代理动作以解决所述歧义；以及

执行所述代理动作。

18. 根据权利要求17所述的计算设备，其中确定所述代理动作中是否存在歧义包括基于至少两个歧义源来确定歧义值。

19. 根据权利要求17所述的计算设备，其中接收自然语言输入包括接收语音输入。

20. 根据权利要求17所述的计算设备，其中标识所述代理动作包括：

至少部分基于所述自然语言输入来执行语言理解以将含义归属给所述自然语言输入；

至少部分基于被归属给所述自然语言输入的所述含义来从数据源中检索信息；以及

至少部分基于被归属给所述自然语言输入的所述含义和所检索的所述信息来确定所述代理动作。

## 解决歧义的对话理解系统

### 背景技术

[0001] 对话理解系统接收通常以自然语言的形式用户输入并且之后解释输入以确定合适的响应或要执行的代理动作。通常,对话理解系统接收来自用户的口头输入,但是备选方案是可能的。

[0002] 对话理解系统被并入在诸如(例如,来自华盛顿州雷蒙德的微软公司的CORTANA®.软件和服务、来自加利福尼亚州山景城的谷歌公司的GOOGLE NOW®.软件和服务、以及来自加利福尼亚州库比蒂诺的苹果公司的SIRI®.软件和服务)个人数字助理的许多其他系统或搜索引擎中。对话理解系统由这些和其他类型的系统使用以接收用于执行各种任务(诸如执行搜索、发送消息、安排约会、执行预订、等等)的输入和指令。

[0003] 重要的是对话单元准确地确定用户想要什么。但是存在导致对话理解单元提供不正确结果或采取不期望动作的许多潜在错误源。例如,用户可能未清楚地说出,可能读错单词,可能拼错单词,可能使用错误的单词,或者可能使用具有多个含义的单词。一种解决这些错误的技术是在采取任何动作之前请求来自用户的反馈或确认。但是用户常常发现对这些请求做出响应是乏味且恼人的。

[0004] 正是关于这些和其他总体考虑做出了实施例。此外,尽管已经讨论了相对具体的问题,但是应当理解,这些实施例不应当限于解决在背景中标识的具体问题。

### 发明内容

[0005] 提供本发明内容从而以简化的形式介绍下面在具体实施方式部分中进一步描述的一系列概念。本发明内容不旨在标识要求保护的的主题的关键或必要特征,也不旨在作为确定要求保护的的主题的范围的辅助。

[0006] 一个非限制性示例是一种用于解决对话理解系统中的歧义的系统或方法。方法包括接收自然语言输入并基于自然语言输入来标识代理动作的步骤。方法还包括确定与代理动作相关联的歧义值并针对歧义条件评估歧义值的步骤。方法包括当确定歧义值满足歧义条件时的以下步骤:基于与所标识的代理动作相关联的歧义来选择提示动作;执行提示动作;响应于提示动作而接收额外的输入;以及基于额外的输入来更新代理动作以解决歧义。方法还包括执行代理动作的步骤。

[0007] 在附图和下面的描述中阐述一个或多个方面的细节。其他特征和优点将从以下详细描述的阅读和相关联的附图的审查变得显而易见。要理解,以下详细描述仅仅是说明性的并且不限制权利要求。

### 附图说明

[0008] 被并入本公开中并且构成本公开的一部分的附图示出各个方面。

[0009] 图1是用于解决对话理解系统中的歧义的示例系统的示意图。

- [0010] 图2是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0011] 图3是图1的示例解决歧义的对话理解引擎的示意性框图。
- [0012] 图4是图3的歧义解决引擎的示意性框图。
- [0013] 图5示出了由图1的系统的一些方面执行的解决对话理解系统中的歧义的示例方法。
- [0014] 图6示出了由图1的系统的一些方面执行的解决对话理解系统的语音输入中的歧义的示例方法。
- [0015] 图7是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0016] 图8是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0017] 图9是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0018] 图10是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0019] 图11是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0020] 图12是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0021] 图13是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0022] 图14是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0023] 图15是用户与图1的示例用户计算设备之间的示例交互的示意图。
- [0024] 图16是示出计算设备的示例物理部件的框图。
- [0025] 图17A和图17B是移动计算设备的框图。
- [0026] 图18是分布式计算系统的框图。

### 具体实施方式

[0027] 以下详细描述参考附图。在可能的情况下,相同的附图标记被使用在附图中并且以下描述参考相同或相似的元件。尽管示例可以被描述,但是修改、适应和其他实现方式是可能的。例如,可以对附图中示出的元件进行替代、添加、或修改,并且本文中描述的方法可以通过对所公开的方法的步骤进行替代、重新排序或添加阶段来修改。因此,以下详细描述不是限制性的,而是相反,合适的范围由随附权利要求限定。示例可以采取硬件实现方式、或者完全软件实现方式或者组合软件和硬件方面的实现方式的形式。因此,以下详细描述不应在限制性的意义上来理解。

[0028] 图1是用于解决对话理解系统中的歧义的示例系统100的示意图。在该示例中,系统100包括通过网络106进行通信的用户计算设备102和服务器计算设备104。用户计算设备102可以是用于实现用于与人类用户交互的解决歧义的对话理解引擎108的任何合适的计算设备。例如,用户计算设备102可以是以下中的至少一个:移动电话;智能电话;平板计算机;平板电话;智能手表;可穿戴计算机;个人计算机;台式计算机;笔记本计算机;游戏设备/计算机(例如,来自华盛顿州雷蒙德的微软公司的XBOX®游戏系统);媒体播放设备(例如,BLUERAY®或DVD播放器或流媒体播放器)、电视;导航系统;车载远程通信系统;等等。该列表仅仅是示例性的并且不应当被认为是限制性的。

[0029] 在一些方面中,用户计算设备102包括解决歧义的对话理解引擎108和语音界面110。解决歧义的对话理解引擎108接收用户输入,诸如语音输入112。尽管备选方案是可能

的,但是解决歧义的对话理解引擎108经由语音界面110接收用户输入。额外地或备选地,解决歧义的对话理解引擎108接收文本用户输入。

[0030] 解决歧义的对话理解引擎108基于所接收的用户输入来确定要执行的代理动作。取决于用户输入,解决歧义的对话理解引擎108可以执行许多不同类型的代理动作。代理动作的非限制性示例包括设置提醒、存储标注、执行搜索、启动电话呼叫、发送消息(例如,SMS或电子邮件)、安排约会、执行预订、以及提供方向。

[0031] 解决歧义的对话理解引擎108还确定所接收的用户输入是否有歧义。如果是,则解决歧义的对话理解引擎108确定是否要向用户确认歧义并且是否要请求来自用户的额外输入来解决歧义。基于歧义的类型和程度,解决歧义的对话理解引擎108确定合适的提示动作来解决歧义。提示动作可以包括视觉反馈114、音频反馈116中的一个或两者。

[0032] 在一些方面中,解决歧义的对话理解引擎108出于各种目的通过网络106与服务器计算设备104进行交互。例如,解决歧义的对话理解引擎108可以请求来自服务器计算设备104的数据以确定代理动作。解决歧义的对话理解引擎108可以将信息发送到服务器计算设备104,该信息可以与来自其他设备和用户的信息进行组合以构建用于标识并解决用户输入中的歧义的模式。在一些方面中,解决歧义的对话理解引擎108从服务器计算设备104中检索模型以用于评估并解决歧义。通常,模型使用机器学习技术基于从许多设备捕获的来自许多用户的数据来生成。

[0033] 语音界面110提供用于用户以口头语言与用户计算设备102进行交互的界面。在一些方面中,语音界面110包括可以为用户执行任务或服务的智能个人助理。智能个人助理的非限制性示例包括来自华盛顿州雷蒙德的微软公司的CORTANA®软件和服务、来自加利福尼亚州山景城的谷歌公司的GOOGLE NOW®软件和服务、以及来自加利福尼亚州库比蒂诺的苹果公司的SIRI®软件和服务。在其他方面中,语音界面110包括其他类型的应用或服务,诸如屏幕读取器或听写系统。

[0034] 尽管在图1的示例中,解决歧义的对话理解引擎108和语音界面110两者都处于用户计算设备102中,但是在其他方面中,解决歧义的对话理解引擎108和语音界面110中的一个或两者在服务器计算设备104上。在示例中,用户计算设备102将接收的语音输入112发送到服务器计算设备104以用于识别。

[0035] 图2是用户与用户计算设备202之间的示例交互的示意图。该示意图示出由诸如参考至少图1说明和描述的解决歧义的对话理解引擎108的解决歧义的对话理解引擎的示例操作。

[0036] 在该示例中,在A处,用户将口头输入200提供到用户计算设备202。更具体地,用户向用户计算设备202说“提醒我买牛奶”。用户计算设备202是参考图1说明和描述的用户计算设备101的示例。

[0037] 用户计算设备202包括解决歧义的对话理解引擎204和语音界面206。在B处,语音界面206接收口头输入200,并且解决歧义的对话理解引擎204确定要执行的动作以及是否存在关于输入或要执行的动作的任何歧义。取决于检测到的歧义的类型和程度,用户计算设备202生成针对用户的反馈。所提供的反馈通常提示用户进行额外的输入,用户计算设备202可以使用该额外的输入来解决歧义。

[0038] 由用户计算设备202生成的反馈的示例在C-F处被提供。通常,用户计算设备202生成来自示例中的仅仅一个示例的反馈。示例从较高歧义到较低歧义来布置。

[0039] 如在C处所示出的,解决歧义的对话理解引擎204生成包括音频反馈208和视觉反馈210的反馈。音频反馈208说“对不起,我不明白。你能重复你说了什么吗?”。视觉反馈210在用户计算设备202的用户界面上显示“未识别到”。解决歧义的对话理解引擎204还继续提供语音界面206,使得解决歧义的对话理解引擎204可以接收来自用户的重复的口头输入。该示例表示当解决歧义的对话理解引擎204确定在口头输入200的全部的识别中存在大量歧义的情形。由解决歧义的对话理解引擎204生成的反馈提示用户重复口头输入200。

[0040] 如在D处所示出的,解决歧义的对话理解引擎204生成包括音频反馈212和视觉反馈214的反馈。音频反馈212说“你能重复你想要被提醒什么吗?”。视觉反馈214在用户计算设备202的用户界面上显示“提醒?”。解决歧义的对话理解引擎204还继续提供语音界面206,使得解决歧义的对话理解引擎204可以当用户重复提醒的主题时接收来自用户的口头输入。该示例表示当解决歧义的对话理解引擎204确定在口头输入200的一部分(即,“买牛奶”)的识别中存在大量歧义,但是其他部分不包括很多歧义(即,“提醒我”)的情形。在该示例中提供的反馈传达确定提醒应当被创建的置信度(即低歧义)。

[0041] 如在E处所示出的,解决歧义的对话理解引擎204生成包括音频反馈216和视觉反馈218的反馈。音频反馈216说“你能帮忙选择你的意思是什么吗?”。视觉反馈218包括在用户计算设备202的用户界面上显示“请选择你的意思是什么:”的提示220。视觉反馈218还包括选项222(“1.买牛奶的提醒。”)和选项224(“2.主题为“买牛奶”的标注。)。尽管未示出,但是在一些方面中,解决歧义的对话理解引擎204还继续提供语音界面206,使得解决歧义的对话理解引擎204可以接收来自用户的额外的口头输入以解决歧义。该示例表示当解决歧义的对话理解引擎204具有关于口头输入中的单词的高水平置信度,但是具有关于基于这些单词要执行的合适动作的歧义时的情形。各个方面提供解决歧义的各种选项。例如,在一些方面中,用户可以利用用户计算设备202的屏幕上的触摸输入来选择选项222或选项224。备选地,在一些方面中,用户可以说出选项的部分以选择或说出与选项相关联的数字(例如,选择选项222的1或“提醒”,选择选项224的2或“标注”)。

[0042] 如在F处所示出的,解决歧义的对话理解引擎204生成包括音频反馈226和视觉反馈228的反馈。音频反馈226说“好,我可以提醒你买牛奶。你想要什么时间被提醒?”。视觉反馈218包括确认语句230(“买牛奶的提醒”)和随后的问题232(“什么时间?”)。解决歧义的对话理解引擎204还继续提供语音界面206,使得解决歧义的对话理解引擎204可以当用户回答随后的问题232时接收来自用户的口头输入。该示例表示当解决歧义的对话理解引擎204具有关于口头输入中的单词和基于这些单词要执行的动作的高水平置信度时的情形。确认语句230告诉用户解决歧义的对话理解引擎204采取什么动作,并且因为需要更多信息来回答随后的问题232,所以用户具有在其不正确时停止/更改动作的机会。

[0043] 在C-F处示出的各种示例仅仅是示例并且存在可以由解决歧义的对话理解引擎204生成的许多其他类型的反馈。在全文中更详细地讨论了一些方面用于评估歧义并选择合适的用户反馈以解决歧义的技术。一些方面针对最小化提示用户提供信息的时间量的目标平衡消除歧义的目标。

[0044] 图3是解决歧义的对话理解引擎300的示意性框图。解决歧义的对话理解引擎300

是图1中示出的解决歧义的对话理解引擎108的示例。解决歧义的对话理解引擎300包括输入引擎302、自动语音识别引擎304、语言理解引擎306、知识检索引擎310、任务状态管理引擎312、代理动作确定引擎314、歧义解决引擎316、以及用户界面引擎318。

[0045] 输入引擎302接收来自各种源的输入。例如,输入引擎302可以从语音界面接收口头语音输入。输入引擎302还可以接收文本输入,文本输入诸如可以由用户键入。输入引擎302还可以诸如经由网络传输从另一计算设备接收语音或文本输入。

[0046] 自动语音识别引擎304识别口头输入中的单词(或语音的其他部分)。在一些方面中,自动语音识别引擎304确定针对所识别的单词(或语音的其他部分)的置信度分数。语言理解引擎306将语言理解模型应用到由自动语音识别引擎304识别的或已经以其他方式(例如,作为文本输入)接收的单词。语言理解引擎306使用语言理解模型来推测用户的意图。

[0047] 知识检索引擎310从一个或多个本地或远程数据源检索知识。例如,如果用户输入包括单词“呼叫亨利”,则知识检索引擎310可以访问地址簿数据源以查找名叫亨利的联系人。

[0048] 任务状态管理引擎312基于输入来访问与当前任务状态(或信念状态)相关联的信息/上下文和/或更新当前任务状态。例如,任务状态管理引擎312可以确定输入“买牛奶”与先前启动的任务相关以生成提醒并且将使用该输入来生成买牛奶的提醒而不是将“买牛奶”评估为独立语句。

[0049] 代理动作确定引擎314基于用户输入来确定要执行的代理动作。歧义解决引擎316评估与用户输入和所确定的代理动作相关联的歧义以确定是否以及如何执行代理动作之前寻求来自用户的额外输入。

[0050] 用户界面引擎318基于用户输入来生成用户界面。例如,如果代理动作被执行,则用户界面引擎318可以生成用户界面以指示代理动作已经被执行,并且在合适的情况下示出执行代理动作的结果。用户界面引擎318还可以生成询问用户的用户界面以在继续执行代理动作之前请求额外输入。

[0051] 图4是歧义解决引擎400的示意性框图。歧义解决引擎400是图3中示出的歧义解决引擎316的示例。歧义解决引擎400包括歧义评估引擎402、危害性评估引擎404、以及解决选择引擎406。

[0052] 歧义评估引擎402量化在对话理解过程期间的各个点处的歧义。歧义评估引擎402可以使用由参考图3更详细地讨论的输入引擎302的部件中的至少一些部件生成的各种统计置信度值。

[0053] 危害性评估引擎404评估所标识的代理动作的危害性。危害性是指当不期望特定代理动作时执行该特定代理动作的后果。具有高危害性的代理动作将在其在不期望时被执行的情况下具有更大后果。交易性代理动作常常具有高危害性。将具有较高危害性的交易性代理动作的示例包括将消息发送到另一用户/设备并启动银行交易。对比之下,具有低危害性的代理动作在其在不期望时被执行的情况下具有更小后果。信息性代理动作常常具有低危害性。例如,检索电影列表的信息性代理大动作将具有较低危害性。在一些方面中,危害性基于动作是否可逆(例如,发送消息)。

[0054] 解决选择引擎406确定是否需要来自用户的额外输入来解决与可以被执行的特定代理动作相关的歧义。一般地,如果不需要另外的输入来解决歧义,则将执行代理动作(或

所标识的代理动作中的至少一个步骤将被执行)。如果需要另外的输入来解决歧义,则解决选择引擎406将确定如何请求额外信息。例如,解决选择引擎406可以确定特定视觉或音频反馈应当被提供给用户以提示用户提供额外信息来解决歧义。反馈可以明确地或隐含地指示歧义已经被标识。

[0055] 各种因素被考虑以确定是否寻求另外的用户输入来解决歧义。例如,各种歧义可以针对一个或多个条件被量化并评估。如果条件被满足,则解决选择引擎406可以确定没有必要寻求额外用户输入来解决歧义。条件可以基于各种因素。例如,如果所确定的代理动作具有高危害性,则条件可以需要很低的歧义分数(或高置信度分数)以在没有额外用户输入的情况下前进。作为另一示例,如果最近已经向用户询问了额外的信息,则条件可以被调整使得更高的歧义分数将仍然满足条件(例如,以便不以对信息的太多请求惹恼用户)。在一些方面中,在用户在冗长或复杂的任务中间的情况下,条件将要求更低的歧义分数,以便最小化在系统不正确地退出需要用户重做许多步骤的长任务的情况下将发生的用户挫折感。

[0056] 解决选择引擎406考虑各种因素来确定如何解决歧义。例如,如果两个备选的代理动作被标识为可能,则解决选择引擎406可以提示用户选择选项之一。备选地,如果存在关于单个标识的代理动作的歧义,则解决选择引擎406可以提示用户确认动作。作为另一备选方案,如果歧义防止完全定义代理动作,则解决选择引擎406可以提示用户提供(或澄清)定义代理动作必需的缺失信息。

[0057] 可以被选择以解决歧义的一些示例技术(或提示动作)包括隐式口头反馈、隐式视觉反馈、具有自然语言理解/策略提示的显式口头反馈、要求重复、再问一次并且呈现备选方案的显式口头反馈。参考至少图7-15来说明和描述这些技术的示例。

[0058] 图5示出了由系统100的一些方面执行的解决对话理解系统中的歧义的示例方法500。在示例中,方法500可以在包括被配置为存储并且执行操作、程序或指令的至少一个处理器的设备上执行。

[0059] 在操作502处,接收输入。如以上所描述的,输入可以是口头输入、文本输入、或另一类型的输入。输入可以直接从用户接收或者可以由另一计算设备发送。

[0060] 在操作504处,基于输入来标识代理动作。标识代理动作可以包括执行自动语音识别、语言理解、知识检索、以及任务状态管理。

[0061] 在操作506处,确定是否存在与输入或所标识的代理动作相关的歧义。在各个方面中,在确定是否存在与所标识的代理动作相关的歧义中考虑各种因素。例如,自动语音识别可以生成置信度分数,并且语言理解引擎可以生成语言建模似然分数(例如,用户多久陈述相似的事物)以及语言理解语义分数(例如,该表达是否对特定领域/意图有意义?)、后知识检索上下文更新(例如,从包括列表、词典、数据库或其他web服务和源的额外数据源检索信息的影响)、以及与(在信念状态更新中计算的)相同过程中的先前表达的关系。另外,在对话理解过程中的任何点处的可行备选方案的存在可以指示关于结果的潜在歧义。这些因素中的一些因素或全部因素可以用于计算歧义分数。

[0062] 确定是否存在歧义可以包括将歧义分数(或置信度分数)与特定阈值进行比较。在一些方面中,存在与对话理解过程的各个部分相关的多个歧义分数。在一些方面中,这些多个分数与一个或多个阈值进行单独地比较以确定是否存在歧义(例如,如果分数中的任何分数超过阈值,则确定存在歧义)。备选地,在一些方面中,多个分数被一起考虑以确定是否

存在歧义(例如,如果第一歧义分数超过第一阈值并且第二歧义分数超过第二阈值,则确定存在歧义)。各种组合和备选方案也是可能的。

[0063] 如果确定存在歧义,则方法前进到操作508。如果相反,确定不存在歧义,则方法前进到操作516,在那里执行代理动作。在操作516之后,该方法结束。

[0064] 在操作508处,所标识的动作的后果被评估以确定所标识的动作的危害性。危害性可以被评估为定量或定性分数。如以上所描述的,危害性涉及不正确地执行所标识的代理动作的后果的严重性。

[0065] 在操作510处,标识备选的代理动作。然而,取决于输入,可以不存在要标识的任何合适的备选的代理动作。

[0066] 在操作512处,确定是否应当寻求额外的输入。如果是,则方法前进到操作514。如果不是,则方法前进到操作516,在那里代理动作被执行。如先前所描述的,在确定是否寻求额外的输入中考虑各种因素。额外地,在一些方面中,机器学习技术用于构建何时寻求输入的模式。

[0067] 在操作514处,选择并执行解决歧义的提示动作。提示动作可以包括音频或视觉提示中的一个或两者以重复先前输入中的一些或全部、以选择选项、或者以提供额外的信息。在一些方面中,选择提示动作包括确定是否要向用户呈现由操作510标识的备选的代理动作中的任何。在提示动作被执行之后,方法返回到操作502,在那里过程使用响应于提示动作接收的新输入来重复。

[0068] 图6示出了由系统100的一些方面执行的解决对话理解系统的语音输入中的歧义的示例方法600。在示例中,方法600可以在包括被配置为存储并且执行操作、程序或指令的至少一个处理器的设备上执行。

[0069] 在操作602处,接收输入。操作602可以类似于先前描述的操作502。

[0070] 在操作604处,对语音输入执行自动语音识别以识别语音输入中的单词。在操作606处,确定语音识别歧义。语音识别歧义可以例如基于针对语音识别结果中的一些或全部的置信度分数来确定。在一些方面中,如果语音识别歧义超过阈值,则方法600前进到操作630,在那里选择提示动作以解决歧义。以这种方法,方法600避免基于高度歧义的自动语音识别结果来执行对话理解中的许多步骤。

[0071] 在操作608处,对在语音输入中识别的单词执行语言理解。在操作610处,确定语言理解歧义。语言理解歧义可以例如基于针对语言理解结果中的一些或全部的置信度分数来确定。在一些方面中,如果语言识别歧义超过阈值,则方法600前进到操作630,在那里选择提示动作以解决歧义。以这种方法,方法600避免在语言理解结果高度歧义时执行对话理解中的一些步骤。

[0072] 在操作612处,检索知识以帮助基于语音输入来确定要执行的代理动作。基于来自操作608的语音输入的语言理解来检索知识。如以上所描述的,知识检索可以用于从内部或外部数据源检索信息。例如,如果语言理解确定用户想要启动对“劳伦”的电话呼叫,则知识检索可以在联系人数据存储中检查名叫劳伦的联系人。

[0073] 在操作614处,确定知识检索歧义。语言理解歧义可以例如基于针对检索的记录中的一些或全部的置信度分数或者基于匹配记录的数量来确定。继续启动对劳伦的电话呼叫的示例语音输入,如果在用户的地址簿中找到多个名叫劳伦的联系人,则可以确定存在与

检索的知识相关的大量歧义。在一些方面中,如果知识检索歧义超过阈值,则方法600前进到操作630,在那里选择提示动作以解决歧义。以这种方法,方法600避免在知识检索结果高度歧义时执行对话理解中的一些步骤。

[0074] 在操作616处,执行任务状态管理。任务状态管理可以包括确定当前任务状态并基于语言理解和基于在操作602中接收的语音输入检索的知识来更新任务状态。在操作618处,确定任务状态歧义。任务状态歧义可以例如基于针对任务状态的置信度分数来确定(例如,语言理解结果和检索的知识是否在任务状态方面有意义)。在一些方面中,如果任务状态歧义超过阈值,则方法600前进到操作630,在那里选择提示动作以解决歧义。以这种方式,方法600避免在任务状态管理结果高度歧义时确定代理动作。

[0075] 在操作620处,基于语音输入来确定代理动作。如先前所描述的,各种各样的代理动作可以取决于语音输入来执行。在操作622处,确定代理动作歧义。代理动作歧义可以基于系统多么确信所确定的代理动作适合语音输入。因为语音输入被提供为自然语言输入并且不一定包括任何特定命令语言或关键字,所以多个代理动作可以看上去是可能合适的。在这种情形下,所确定的代理动作可以被确定为具有歧义。

[0076] 在操作624处,确定代理动作危害性。如以上所描述的,代理动作的危害性是指当用户不想要动作被执行时(即,当其不是用户针对语音输入意图的结果时)执行该动作的后果的严重性。

[0077] 在操作626处,确定用于执行所标识的代理动作的条件是否被满足。例如,条件可以包括针对对话理解过程中的先前步骤中的每个步骤的最大歧义水平。另一条件可以与跨对话理解过程中的先前步骤中的全部步骤的总歧义相关。在一些方面中,数值分数针对所确定的歧义中的每个被确定并且那些分数被组合并且与最大阈值进行比较。在一些方面中,数值分数在被组合之前被加权。在一些方面中,阈值基于代理动作的危害性来调整(例如,如果代理动作具有较高的危害性,则阈值被降低以限制总歧义)。

[0078] 如果确定用于执行代理动作的条件被满足,则方法600前进到操作628,在那里执行代理动作。如果不是,则方法前进到操作630。

[0079] 在操作630处,选择提示动作以解决歧义。提示动作向用户提供反馈,该反馈可以是音频的、视觉的、或两者,以提供可以用于解决所标识的歧义的额外输入。在操作632处,执行所选择的提示动作。其后,方法返回到操作602,在那里使用新输入来重复方法。

[0080] 图7是用户与用户计算设备702之间的示例交互的示意图。用户计算设备702是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中,提示动作提供隐式音频反馈。

[0081] 在该示例中,用户提供说“提醒我买牛奶”的口头输入700。作为响应,用户计算设备702提供音频反馈704和视觉反馈706。音频反馈704说“好的,我可以提醒你买牛奶。你想要什么时间被提醒?”。视觉反馈706包括提醒708。提醒包括:主题字段710,其显示“买牛奶”;以及时间字段712,其显示“?”。

[0082] 图8是用户与用户计算设备802之间的示例交互的示意图。用户计算设备802是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中,提示动作提供隐式音频反馈。

[0083] 在该示例中,用户提供说“提醒我买牛奶”的口头输入800。作为响应,用户计算设

备802提供音频反馈804和视觉反馈806。音频反馈804说“好的,你想要什么时间被提醒?”。视觉反馈806包括提醒808。提醒包括:主题字段810,其显示“买牛奶”;以及时间字段812,其显示“?”。

[0084] 隐式视觉反馈允许系统将其对口头输入800的理解与实际对话流分离显示并且指示其当前知识,而不要求用户确认或拒绝任何事情。以这种方式,用户意识到系统的当前信念,并且可以决定在必要时校正系统,但是对话流不被中断。

[0085] 图9是用户与用户计算设备902之间的示例交互的示意图。用户计算设备902是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中,提示动作提供关于自然语言理解的显式音频反馈。

[0086] 在该示例中,用户提供说“提醒我买牛奶”的口头输入900。作为响应,用户计算设备902提供音频反馈904和视觉反馈906。音频反馈904说“你想要被提醒买牛奶吗?”。视觉反馈906包括提醒908。提醒包括:主题字段810,其显示“买牛奶”;以及时间字段812,其显示“?”。

[0087] 在该示例中,系统显式地重复用户说了什么并且要求用户确认。额外地,系统提供什么代理动作将被执行的指示。该显式音频反馈在系统确定存在高水平的歧义(例如,系统缺乏其理解的置信度),并且系统已经标识了仅仅一个潜在代理动作的情况下是有用的。代理动作的指示警告用户在用户肯定回答的情况下系统将做什么。

[0088] 图10是用户与用户计算设备1002之间的示例交互的示意图。用户计算设备1002是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中,提示动作提供显式音频反馈。

[0089] 在该示例中,用户提供说“提醒我买牛奶”的口头输入1000。作为响应,用户计算设备1002提供音频反馈1004和视觉反馈1006。音频反馈1004说“你是说“提醒我买牛奶”吗?”。视觉反馈1006包括显示消息“你是说“提醒我买牛奶”吗?”的消息面板1008。

[0090] 在该示例中,系统显式地重复用户说了什么并且向用户要求来自用户的确认。该显式口头反馈在系统确定存在高水平的歧义(例如,系统缺乏该理解的置信度),并且系统已经标识了仅仅一个潜在代理动作,或者甚至尚未基于歧义的程度来标识代理动作的情况下是有用的。

[0091] 图11是用户与用户计算设备1102之间的示例交互的示意图。用户计算设备1102是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中,提示动作简单地要求用户重复口头输入1200。

[0092] 在该示例中,用户提供说“提醒我买牛奶”的口头输入1100。作为响应,用户计算设备1102提供音频反馈1104和视觉反馈1106。音频反馈1104说“对不起,我不明白。你能重复你说了什么吗?”。视觉反馈1106包括显示消息“未识别”的消息面板1108。

[0093] 在该示例中,系统要求用户重复口头输入。再问一次的这种策略可以在当多个潜在识别值被标识时的高度歧义的自动语音识别结果(例如,低ASR置信度)的情况下是有用的。该系统可以使用从处理重复的口头输入得到的信息以及从处理口头输入1100得到的自动语音识别结果来确定正确的识别结果。

[0094] 图12是用户与用户计算设备1202之间的示例交互的示意图。用户计算设备1202是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统

生成的示例提示动作。在该示例中，提示动作要求用户重复口头输入1200的部分。

[0095] 在该示例中，用户提供说“提醒我买牛奶”的口头输入1200。作为响应，用户计算设备1202提供音频反馈1204和视觉反馈1206。音频反馈1204说“你能重复你想要被提醒什么吗？”。视觉反馈1206包括具有显示“？”的主题字段的提醒1208。

[0096] 在该示例中，系统向用户指示提醒将被创建并且要求用户重复对应于提醒的主题的口头输入的部分。要求部分重复的这种策略可以在针对口头输入1200的部分并且当多个（或没有）潜在识别值被标识时的高度歧义的自动语音识别结果（例如，低ASR置信度）的情况下是有用的。该系统可以使用从处理重复的口头输入部分得到的信息以及从处理口头输入1200得到的自动语音识别结果来确定正确的识别结果。

[0097] 图13-15包括在要求用户从多个备选中进行选择的情况下的交互的示意性框图。呈现备选的这种策略在以充分的歧义标识到多个良好候选时是有用的，使得用户可以之后快速地标识意图的输入或动作。

[0098] 图13是用户与用户计算设备1302之间的示例交互的示意图。用户计算设备1302是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中，提示动作要求用户在口头输入1300的备选转录之间进行选择。

[0099] 在该示例中，用户提供说“提醒我买牛奶”的口头输入1300。作为响应，用户计算设备1302提供音频反馈1304和视觉反馈1306。音频反馈1304说“你能选择你说了什么吗？”。视觉反馈1306包括说“你能选择你说了什么吗？”的提示文本1308。视觉反馈1305还包括第一选项1310（“1.提醒我买牛奶”）和第二选项1312（“2.提醒我买丝绸”）。

[0100] 图14是用户与用户计算设备1402之间的示例交互的示意图。用户计算设备1402是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中，提示动作要求用户在基于口头输入1400的备选识别的潜在提醒之间进行选择。

[0101] 在该示例中，用户提供说“提醒我买牛奶”的口头输入1400。作为响应，用户计算设备1402提供音频反馈1404和视觉反馈1406。音频反馈1404说“你能选择你说了什么吗？”。视觉反馈1406包括说“你能选择你说了什么吗？”的提示文本1408。视觉反馈1406还包括：第一提醒1410，其具有“买牛奶”的主题；以及第二提醒1412，其具有“买丝绸”的主题。

[0102] 图15是用户与用户计算设备1502之间的示例交互的示意图。用户计算设备1502是用户计算设备102的示例。该示意图示出由诸如参考至少图1说明和描述的系统100的系统生成的示例提示动作。在该示例中，提示动作要求用户在基于口头输入1500的备选潜在动作之间进行选择。

[0103] 在该示例中，用户提供说“提醒我买牛奶”的口头输入1500。作为响应，用户计算设备1502提供音频反馈1504和视觉反馈1506。音频反馈1504说“你能选择你的意思是什么吗？”。视觉反馈1506包括说“你能选择你的意思是什么吗？”的提示文本1508。视觉反馈1506还包括：提醒1510，其具有“买牛奶”的主题；以及标注1512，其具有“买丝绸”的主题。用户可以做出响应以指示是创建标注还是提醒。

[0104] 当向用户呈现备选时，系统还可以考虑来自对话理解过程中的各个点中的多种类型的备选（例如，自动语音识别、动作确定、等等）。例如，系统可以示出不同的自动语音识别

备选,或者语言理解假设,或者不同的后知识取回结果,或者甚至来自可以向用户提供不同选项的来自任务(或信念)状态更新的备选选项。在一些方面中,来自对话理解过程的稍后步骤的备选被选择用于显示(例如,语言理解而不是自动语音识别),除非在较早的步骤处存在明显歧义(例如,自动语音识别具有低置信度)。有益地,从过程中的稍后步骤生成的备选将通常向用户提供系统的当前解释和意图的更好理解。

[0105] 使用哪个方法以及向用户呈现哪些备选(以及哪个抽象)的选择可以通过各种方式做出。在一些方面中,基于规则的系统使用各种置信度阈值与规则来确定针对给定情形的最好选择。一些方面使用可以确定要使用的正确方法的统计机器学习模型(判别式或生成式)。备选地或额外地,一些方面使用对提供不同类型的反馈的成本(例如,用户时间/烦恼度)进行建模的决策理论模型来选择如何提供反馈。一些方面跨不同领域的宽泛用户集合共享预测模型。该数据可以帮助系统获知关于系统认为存在歧义是否正确,以及用户是否接受或拒绝提供反馈,以及用户在不同类型的情形下偏好哪些类型的方法来要求澄清。

[0106] 图16-18和相关联的描述提供示例被实践于其中的各种操作环境的讨论。然而,参考图16-18说明和讨论的设备和系统是出于示例和说明的目的并且不限制被用于实践本文中描述的方面的大量计算设备配置。

[0107] 图16是示出可以利用其来实践本公开的示例的计算设备1600的物理部件(即,硬件)的框图。在一个基本配置中,计算设备1600包括至少一个处理单元1602和系统存储器1604。根据一方面,取决于计算设备的配置和类型,系统存储器1604包括但不限于易失性存储设备(例如,随机访问存储器)、非易失性存储设备(例如,只读存储器)、闪存、或这样的存储器的任何组合。根据一方面,系统存储器1604包括操作系统1605和适合于运行软件应用1650的一个或多个程序模块1606。根据一方面,系统存储器1604包括解决歧义的对话理解引擎1352和语音界面1354。操作系统1605例如适合于控制计算设备1600的操作。另外,各方面结合图形库、其他操作系统或者任何其他应用程序被实践并且不限于任何具体应用或系统。该基本配置在图16中由虚线1608之内的那些部件示出。根据一方面,计算设备1600具有额外的特征或功能。例如,根据一方面,计算设备1600包括额外数据存储设备(可移除的和/或不可移除的),诸如,例如,磁盘、光盘、或磁带。这样的额外存储设备在图16中由可移除存储设备1609和不可移除存储设备1610示出。

[0108] 如以上所陈述的,根据一方面,许多程序模块和数据文件被存储于系统存储器1604中。当在处理单元1602上运行时,程序模块1606(例如,应用1350、解决歧义的对话理解引擎1352以及语音界面1354)执行包括但不限于图5和图6中示出的方法500和600的步骤中的一个或多个步骤的过程。根据一方面,其他程序模块根据示例来使用并且包括诸如电子邮件和联系人应用、文字处理应用、电子表格应用、数据库应用、幻灯片演示应用、绘图或计算机辅助应用程序、等等的应用。

[0109] 根据一方面,各方面被实践于包括分立电子元件的电子电路中、包含逻辑门的封装或集成电子芯片中、利用微处理器的电路中、或者包含电子元件或微处理器的单个芯片上。例如,各方面经由片上系统(SOC)实践,其中图16中示出的部件中的每个或许多被集成到单个集成电路上。根据一方面,这样的SOC设备包括一个或多个处理单元、图形单元、通信单元、系统虚拟化单元以及其全部功能被集成(或“烧写”)到芯片衬底上作为单个集成电路的各种应用。当经由SOC操作时,本文中描述的功能经由与单个集成电路(芯片)上的计算设

备1600的其他部件集成的专用逻辑来操作。根据一方面,本公开的各方面使用能够执行逻辑操作的其他技术来实践,这些逻辑操作诸如例如为AND、OR和NOT,包括但不限于机械的、光学的、流体的和量子的技术。另外,各方面被实践于通用计算机内或任何其他电路或系统中。

[0110] 根据一方面,计算设备1600具有一个或多个输入设备1612,诸如键盘、鼠标、笔、声音输入设备、触摸输入设备、等等。根据一方面,还包括诸如显示器、扬声器、打印机、等等的(一个或多个)输出设备1614。前述设备是示例并且其他可以被使用。根据一方面,计算设备1600包括允许与其他计算设备1618进行通信的一个或多个通信连接1616。合适的通信连接1616的示例包括但不限于射频(RF)发射器、接收器、和/或收发器电路;通用串行总线(USB)、并行端口和/或串行端口。

[0111] 如本文中所使用的术语计算机可读介质包括计算机存储介质。计算机存储介质包括以任何方法或技术实施的用于存储诸如计算机可读指令、数据结构、或程序模块的信息的易失性介质和非易失性介质、可移除介质和不可移除介质。系统存储器1604、可移除存储设备1609、以及不可移除存储设备1610全部是计算机存储介质示例(即,存储器存储设备)。根据一方面,计算机存储介质包括RAM、ROM、电可擦可编程只读存储器(EEPROM)、闪存或其他存储器技术、CD-ROM、数字多用盘(DVD)或其他光学存储设备、磁带盒、磁带、磁盘存储设备或其他磁性存储设备、或者能够被用于存储信息并且能够由计算设备1600访问的任何其他制品。根据一方面,任何这样的计算机存储介质是计算设备1600的部分。计算机存储介质不包括载波或其他传播的数据信号。

[0112] 根据一方面,通信介质由计算机可读指令、数据结构、程序模块或诸如载波或其他传输介质的经调制的数据信号中的其他数据来实现并且包括任何信息递送介质。根据一方面,术语“经调制的数据信号”描述具有以使得在信号中编码信息的方式设置或改变的一个或多个特性的信号。通过举例而非限制的方式,通信介质包括诸如有线网络或直接有线连接的有线介质和诸如声学、射频(RF)、红外和其他无线介质的无线介质。

[0113] 图17A和17B示出了可以利用其来实践这些方面的移动计算设备1700,例如,移动电话、智能电话、平板个人计算机、膝上型计算机、等等。参考图17A,示出了用于实施这些方面的移动计算设备1700的示例。在基本配置中,移动计算设备1700是具有输入元件和输出元件两者的手持计算机。移动计算设备1700通常包括显示器1705以及允许用户将信息输入到移动计算设备1700中的一个或多个输入按钮1710。根据一方面,移动计算设备1700的显示器1705用作输入设备(例如,触摸屏显示器)。如果被包括,则可选的副输入元件1715允许另外的用户输入。根据一方面,副输入元件1715是旋转开关、按钮或任何其他类型的手动输入元件。在备选示例中,移动计算设备1700包含更多或更少的输入元件。例如,显示器1705在一些示例中可以不是触摸屏。在备选示例中,移动计算设备1700是便携式电话系统,诸如蜂窝电话。根据一方面,移动计算设备1700包括可选的小键盘1735。根据一方面,可选的小键盘1735是物理小键盘。根据另一方面,可选的小键盘1735是在触摸屏显示器上生成的“软”小键盘。在各种方面中,输出元件包括用于示出图形用户界面(GUI)的显示器1705、视觉指示器1720(例如,发光二极管)、和/或音频换能器1725(例如,扬声器)。在一些示例中,移动计算设备1700包含用于向用户提供触觉反馈的振动换能器。在又一示例中,移动计算设备1700包含输入和/或输出端口,诸如音频输入(例如,麦克风插孔)、音频输出(例如,耳

机插孔)、以及视频输出(例如,HDMI端口)以将信号发送到外部设备或者从外部设置接收信号。在又一示例中,移动计算设备1700包含外围设备端口1740,诸如音频输入(例如,麦克风插孔)、音频输出(例如,耳机插孔)、以及视频输出(例如,HDMI端口)以将信号发送到外部设备或者从外部设置接收信号。

[0114] 图17B是示出了移动计算设备的一个示例的体系结构的框图。即,移动计算设备1700包含系统(即,体系结构)1702以实施一些示例。在一个示例中,系统1702被实施为能够运行一个或多个应用(例如,浏览器、电子邮件、日历、联系人管理器、消息传送客户端、游戏以及媒体客户端/播放器)的“智能电话”。在一些示例中,系统1702被集成为计算设备,诸如集成的个人数字助理(PDA)和无线电话。

[0115] 根据一方面,一个或多个应用程序1750被加载到存储器1762中并且在操作系统1764上运行或者结合该操作系统运行。应用程序的示例包括电话拨号程序、电子邮件程序、个人信息管理(PIM)程序、文字处理程序、电子表格程序、因特网浏览器程序、消息传送程序、等等。根据一方面,解决歧义的对话理解引擎1752和语音界面1754被加载到存储器1762中。系统1702还包括存储器1762内的非易失性存储区域1768。非易失性存储区域1768用于存储永久信息,该永久信息在系统1702关机的情况下不应当丢失。应用程序1750可以使用并存储非易失性存储区域1768中的信息,诸如由电子邮件应用使用的电子邮件或其他消息,等等。同步应用(未示出)还驻存于系统1702上并且被编程为与驻存在主机计算机上的对应同步应用交互以保持存储于非易失性存储区域1768中的信息与存储于主机计算机处的对应信息同步。如应当认识到的,其他应用可以被加载到存储器1762中并且运行于移动计算设备1700上。

[0116] 根据一方面,系统1702具有电源1770,其被实施为一个或多个电池。根据一方面,电源1770还包括外部电源,诸如补充电池或者对电池进行再充电的AC适配器或电源对接支架。

[0117] 根据一方面,系统1702包括执行发射和接收射频通信的功能的无线电装置1772。无线电装置1772支持经由通信载体或服务提供者在系统1702与“外界”之间的无线连接性。到和自无线电装置1772的传输在操作系统1764的控制下进行。换言之,由无线电装置1772接收的通信可以经由操作系统1764被散布到应用程序1750,并且反之亦然。

[0118] 根据一方面,视觉指示器1720用于提供视觉通知和/或音频接口1774用于经由音频换能器1725提供可听通知。在示出的示例中,视觉指示器1720是发光二极管(LED),并且音频换能器1725是扬声器。这些设备可以被直接耦合到电源1770,使得当被激活时,这些设备在由通知机制指示的持续时间内保持接通,即使处理器1760和其他部件可能关机以节省电池电力。LED可以被编程为无限期地保持接通直到用户采取动作以指示设备的接通状态。音频接口1774用于向用户提供可听信号并且从用户接收可听信号。例如,除了被耦合到音频换能器1725,音频接口1774还可以被耦合到麦克风以接收可听输入,诸如支持电话对话。根据一方面,系统1702还包括使得板上相机1730能够记录静态图像、视频流、等等的操作的视频接口1776。

[0119] 根据一方面,实施系统1702的移动计算设备1700具有额外的特征或功能。例如,移动计算设备1700包括额外数据存储设备(可移除的和/或不可移除的),诸如,磁盘、光盘、或磁带。在图17B中通过非易失性存储区域1768示出了这样的额外存储设备。

[0120] 根据一方面,由移动计算设备1700生成或捕获并且经由系统1702存储的数据/信息被本地存储在移动计算设备1700上,如以上所描述的。根据另一方面,数据被存储在由设备经由无线电装置1772或经由移动计算设备1700和与移动计算设备1700相关联的单独的计算设备(例如,诸如因特网的分布式计算网络中的服务器计算机)之间的有线连接可访问的任何数量的存储介质上。如应当认识到的,这样的数据/信息经由移动计算设备1700经由无线电装置1772或者经由分布式计算网络可访问。类似地,根据一方面,根据公知的数据/信息传输和存储方式,包括电子邮件和合作性数据/信息共享系统,这样的数据/信息在计算设备之间容易地传输以用于存储和使用。

[0121] 图18示出如以上所描述的用于解决对话理解系统中的歧义的系统的体系结构的一个示例。使得利用解决歧义的对话理解1852和语音界面1854开发的、或与解决歧义的对话理解1852和语音界面1854交互的、或结合解决歧义的对话理解1852和语音界面1854编辑的内容能够被存储在通信通道中或者以其他存储类型被存储。例如,各种文档可以使用目录服务1822、web端口1824、邮箱服务1826、即时通讯存储1828、或社交网站1830来存储。解决歧义的对话理解1852可操作用于使用这些类型的系统中的任何等用于解决对话理解中的歧义,如本文中所描述的。根据一方面,服务器1820向客户端1805a、b、c提供解决歧义的对话理解1852和语音界面1854。作为一个示例,服务器1820是通过web提供解决歧义的对话理解1852和语音界面1854的web服务器。服务器1820通过网络1840在web上向客户端1805提供解决歧义的对话理解1852和语音界面1854。通过举例的方式,客户端计算设备被实施并体现在个人计算机1805a、平板计算设备1805b或移动计算设备1805c(例如,智能电话)、或其他计算设备中。客户端计算设备的这些示例中的任何示例可以操作用于从存储1816获得内容。

[0122] 例如,以上参考根据各方面的方法、系统和计算机程序产品的框图和/或操作说明来描述实施方式。在框中指出的功能/动作可以以与任何流程图中示出的不同的顺序来进行。例如,取决于涉及的功能/动作,连续示出的两个框实际上可以基本上同时地被运行,或各框可以有时以相反的顺序被运行。

[0123] 对本申请中提供的一个或多个示例的描述和说明不旨在以任何方式限制或约束要求保护的发明范围。本申请中提供的各方面、示例或细节被认为足以传达所有权并且使得其他人能够利用最好模式。实现方式不应当被解释为限于本申请中提供的任何方面、示例或细节。不管是组合地还是单独地示出和描述,各种特征(结构和方法两者)旨在被选择性地包括或省略以产生具有特定特征集的示例。已经提供了对本申请的描述和说明,本领域技术人员可以设想落在本申请中体现的总体发明构思的更宽方面的精神内的未脱离更宽范围的变型、修改和备选示例。

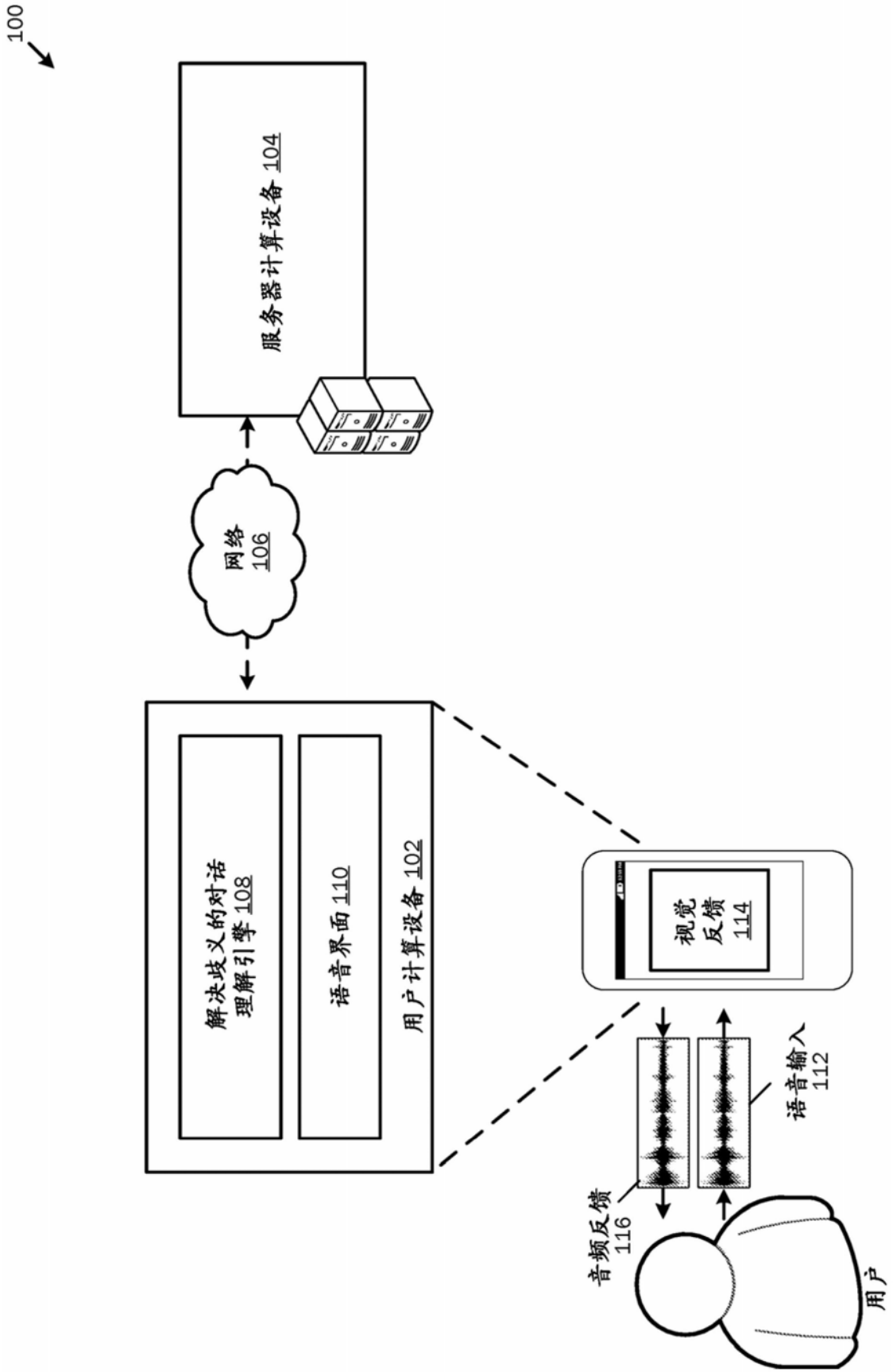


图1

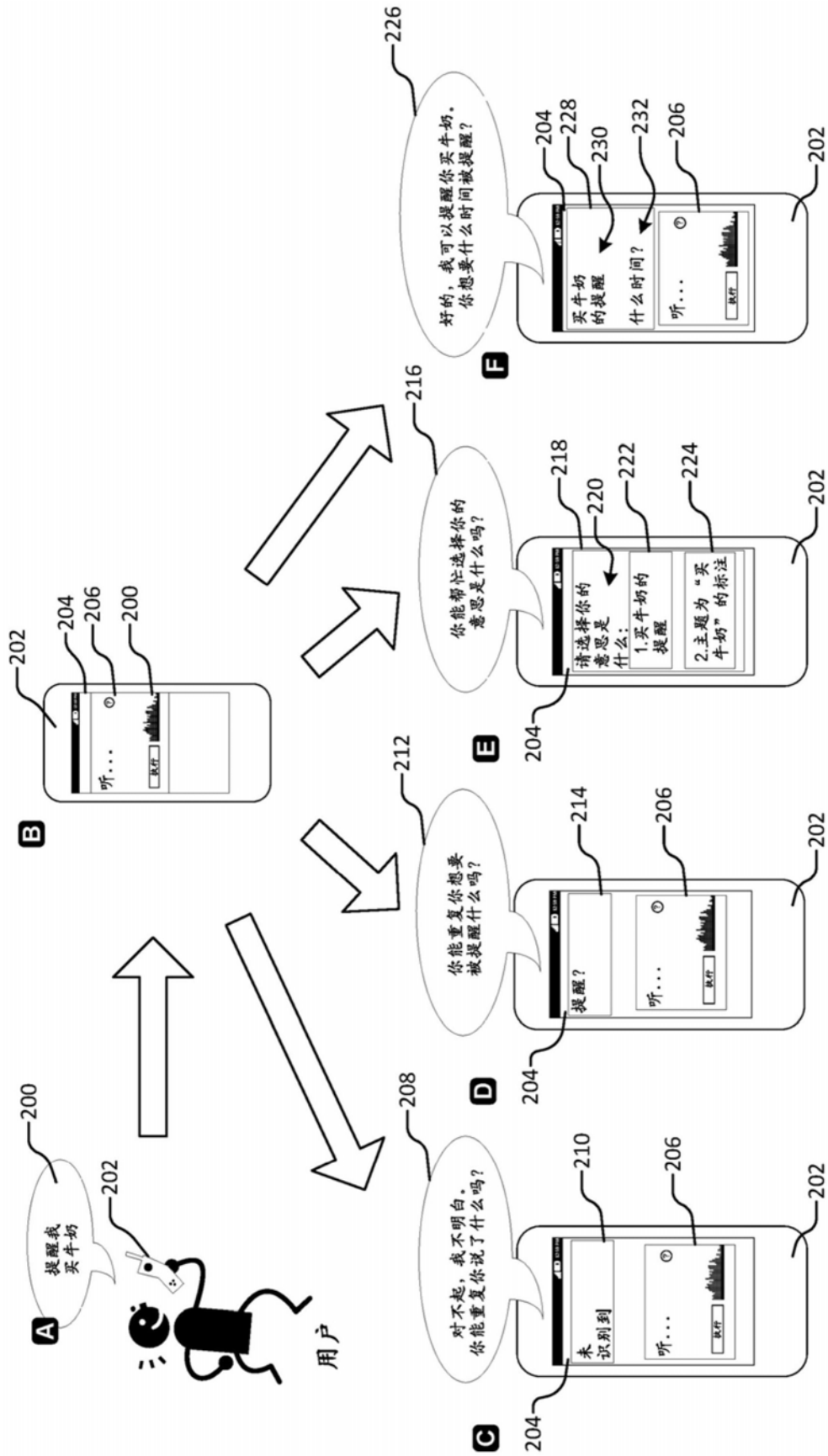


图2



图3



图4

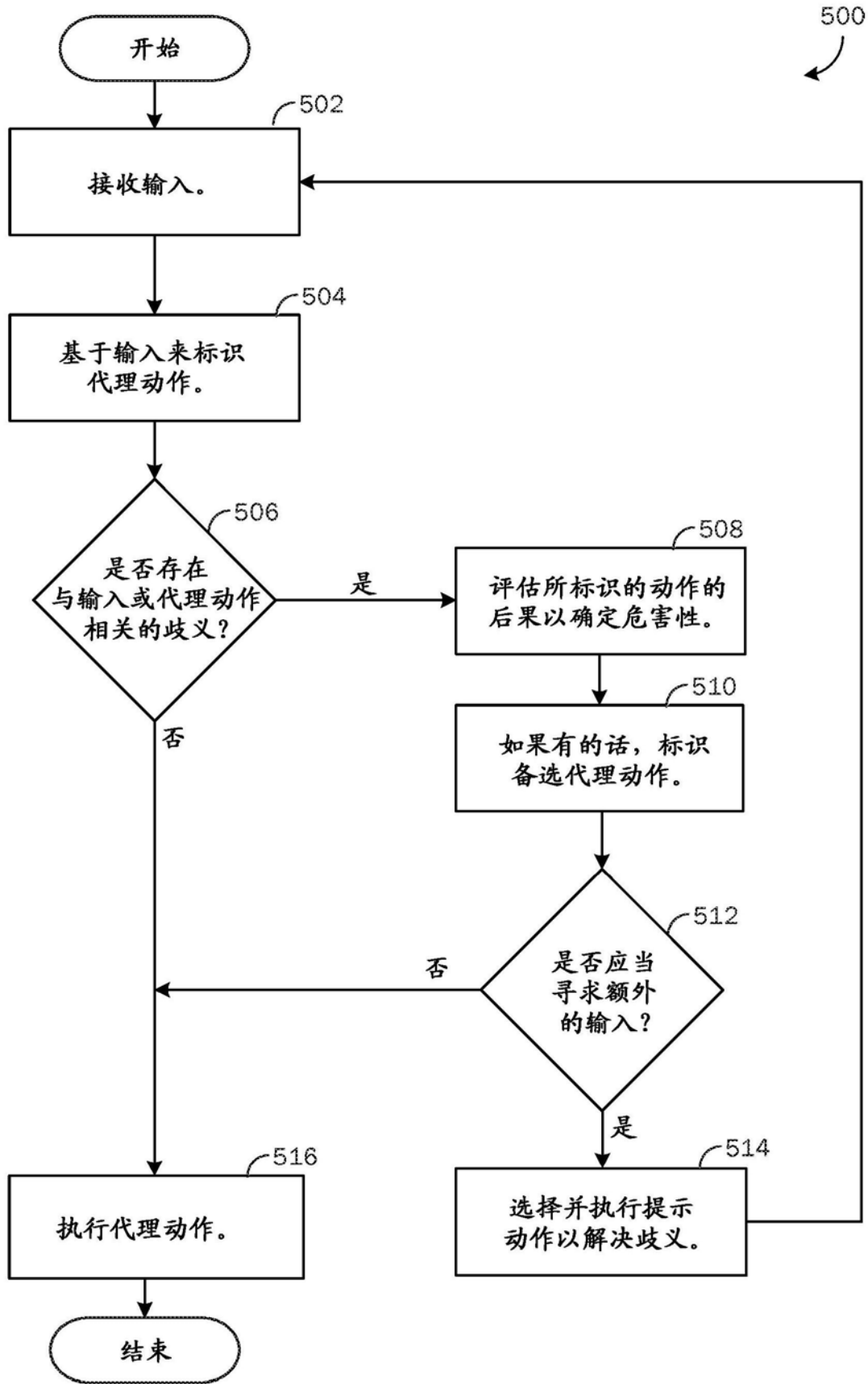


图5

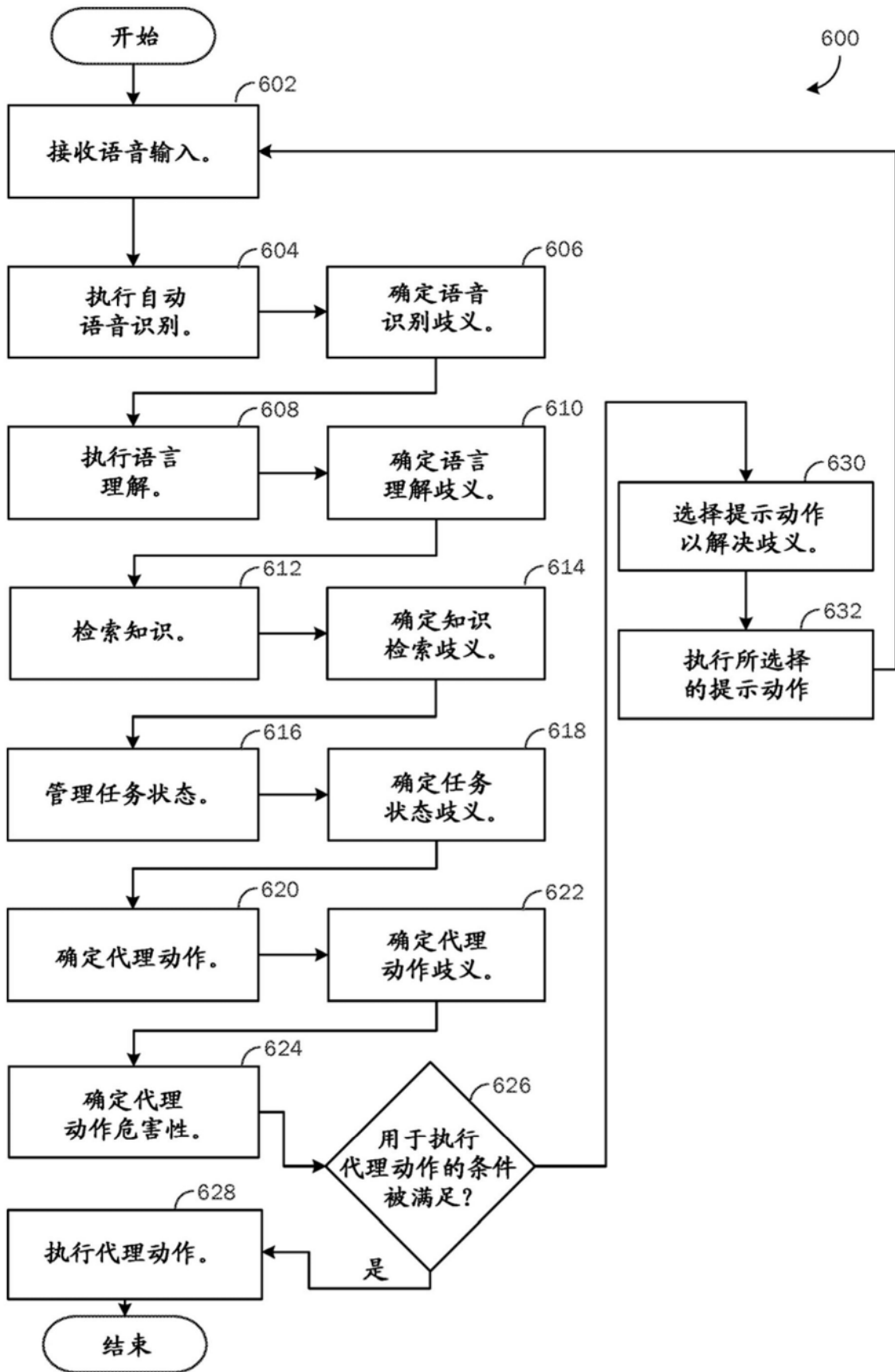


图6

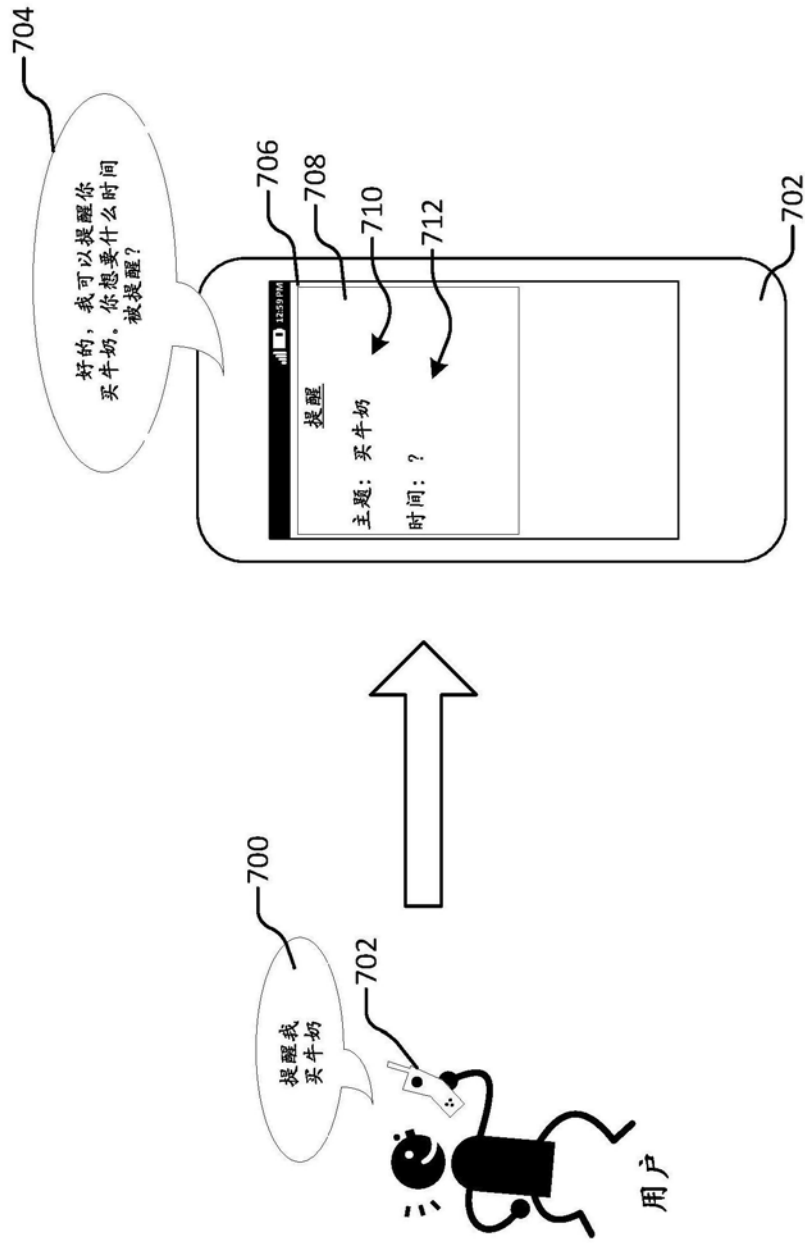


图7

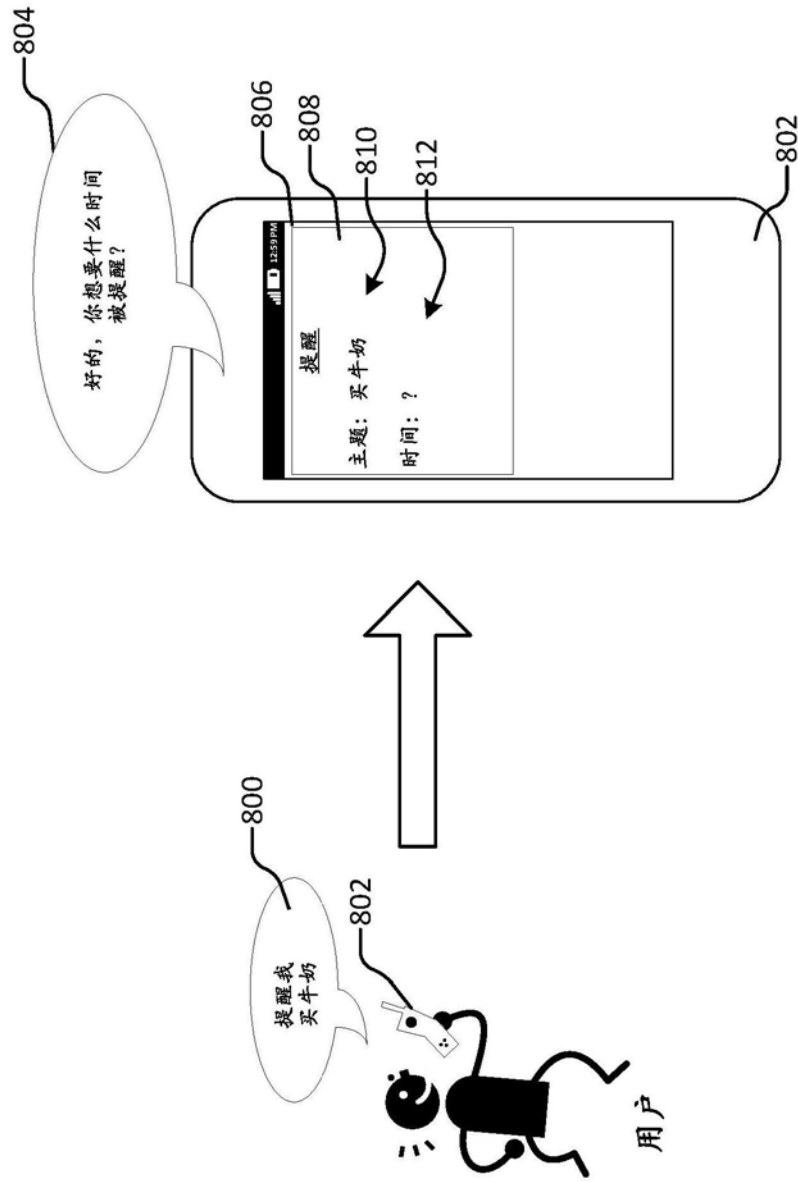


图8

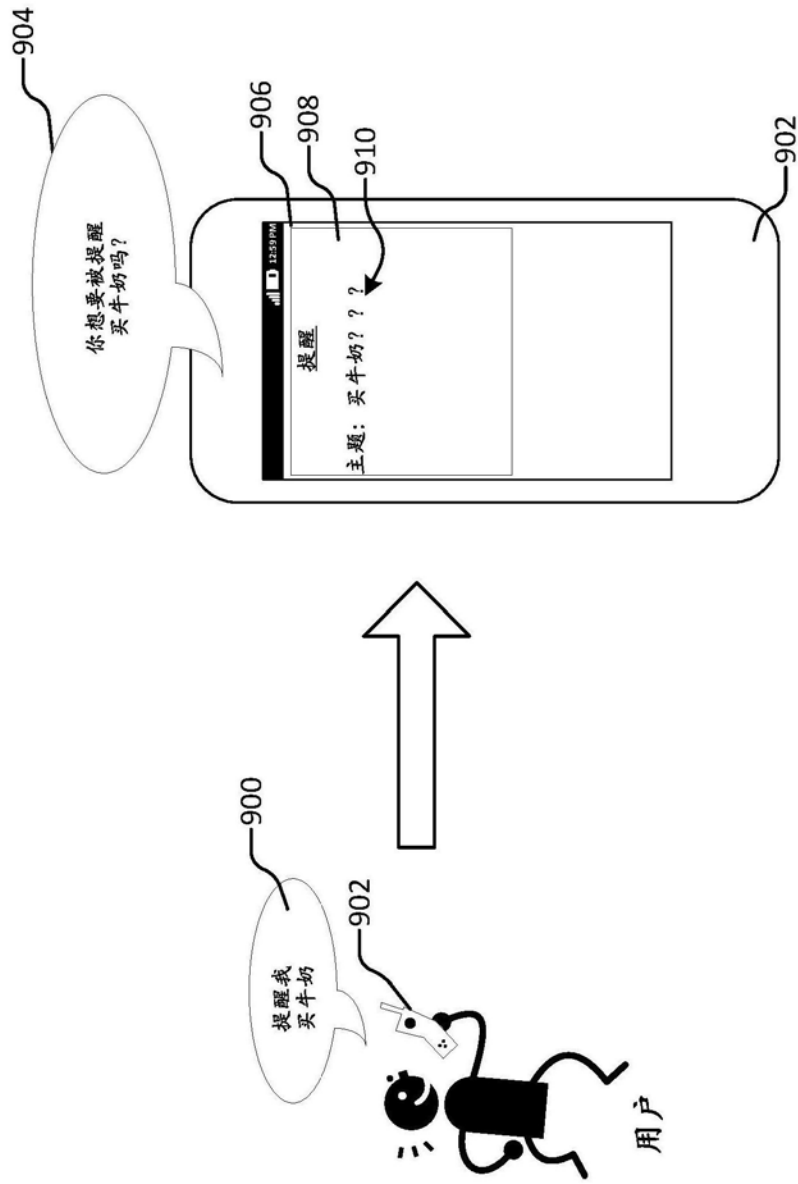


图9

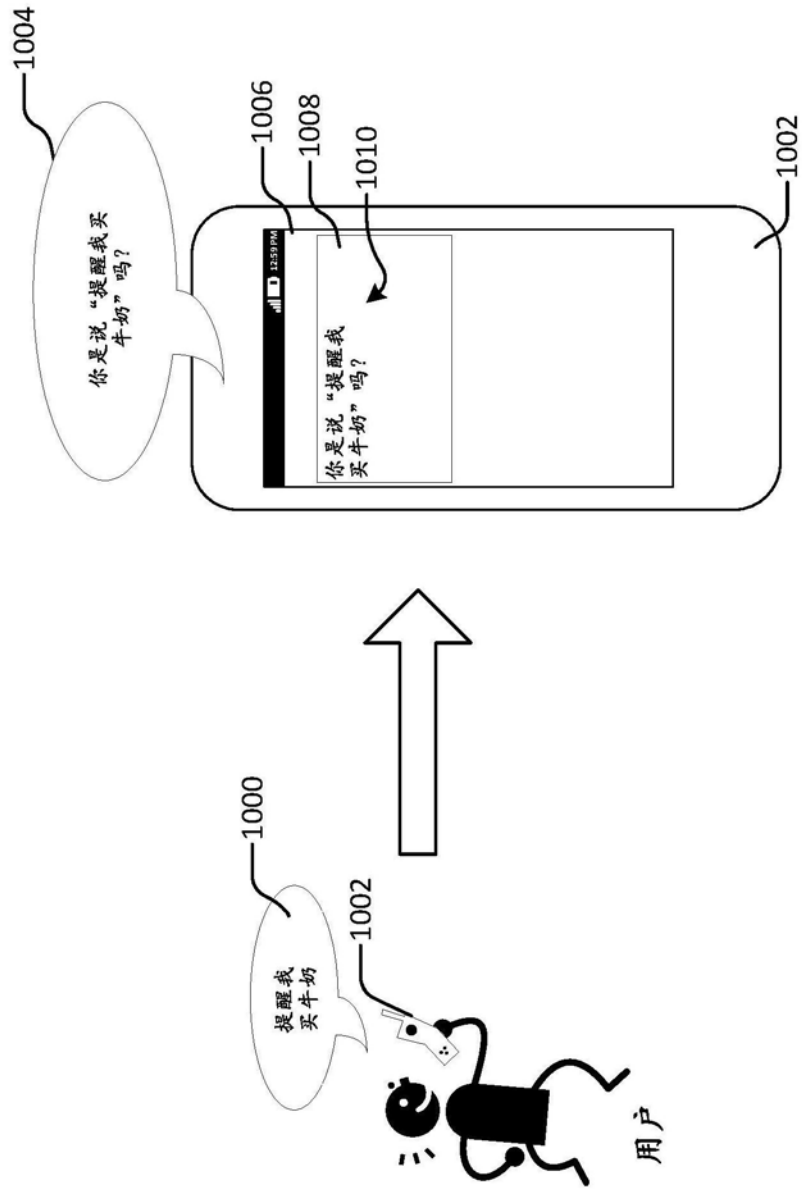


图10

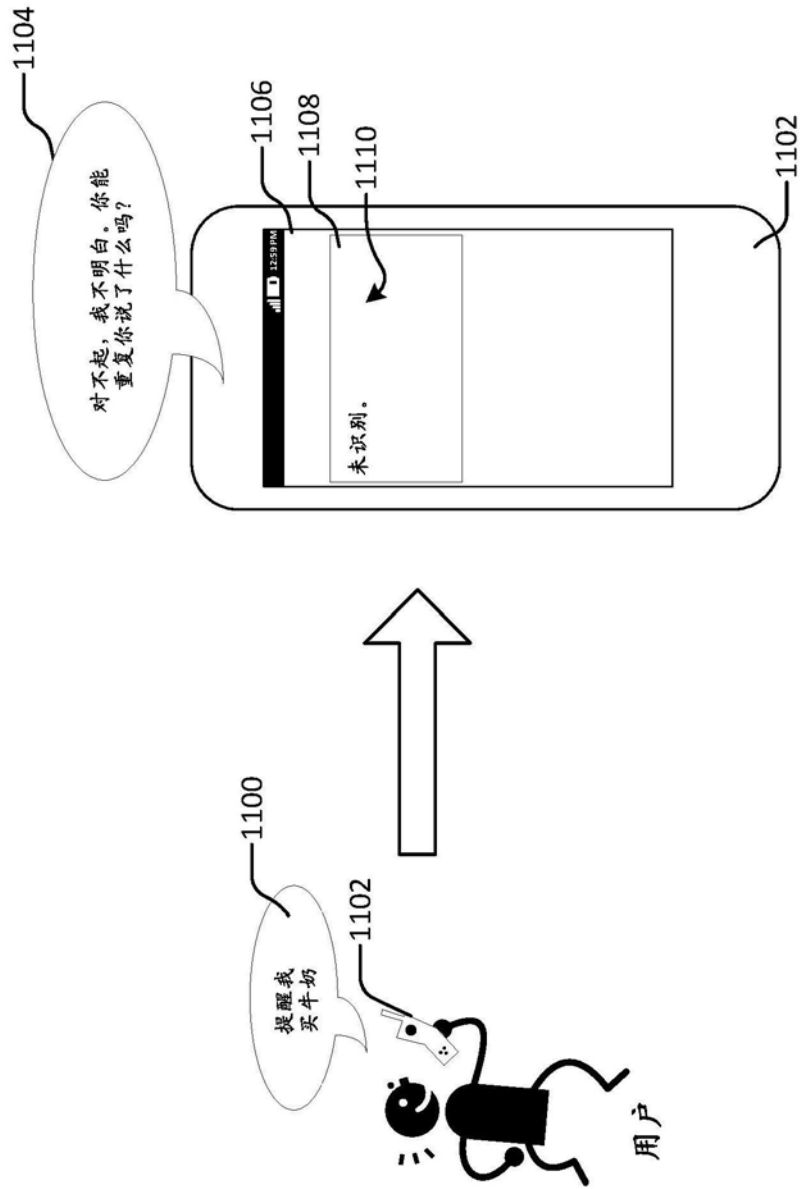


图11

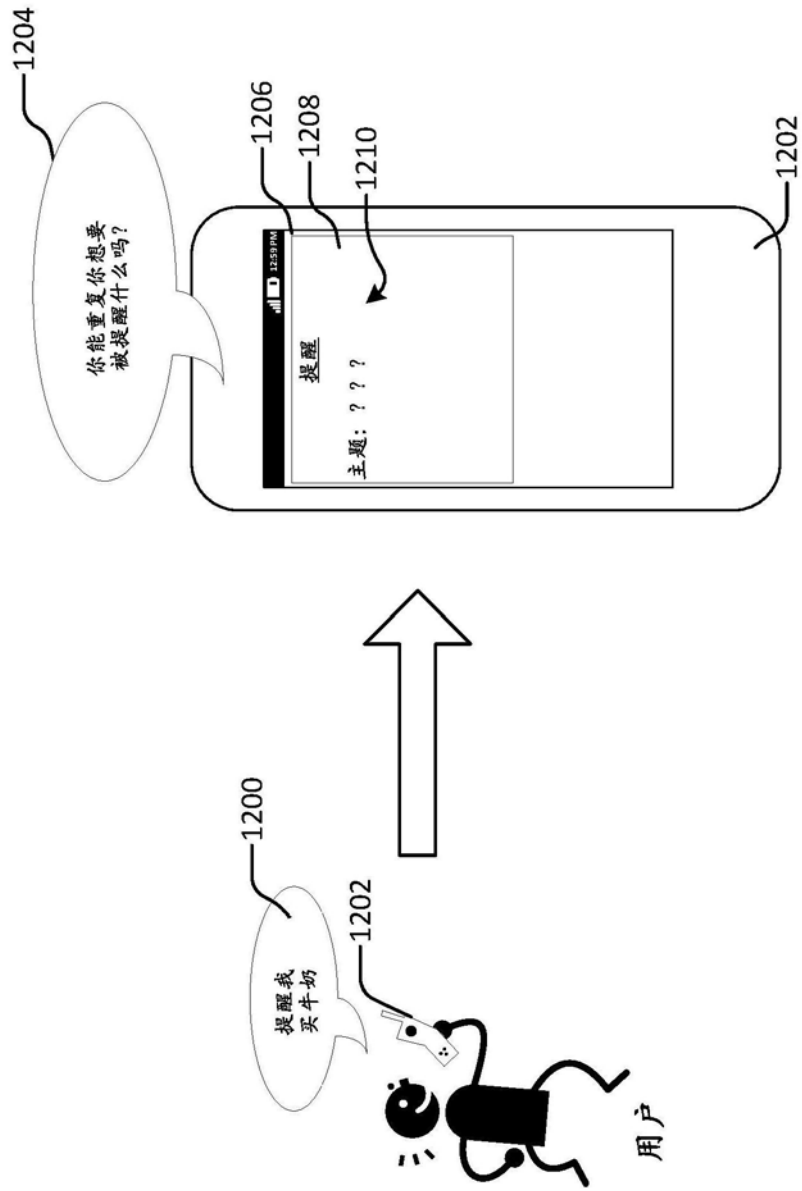


图12

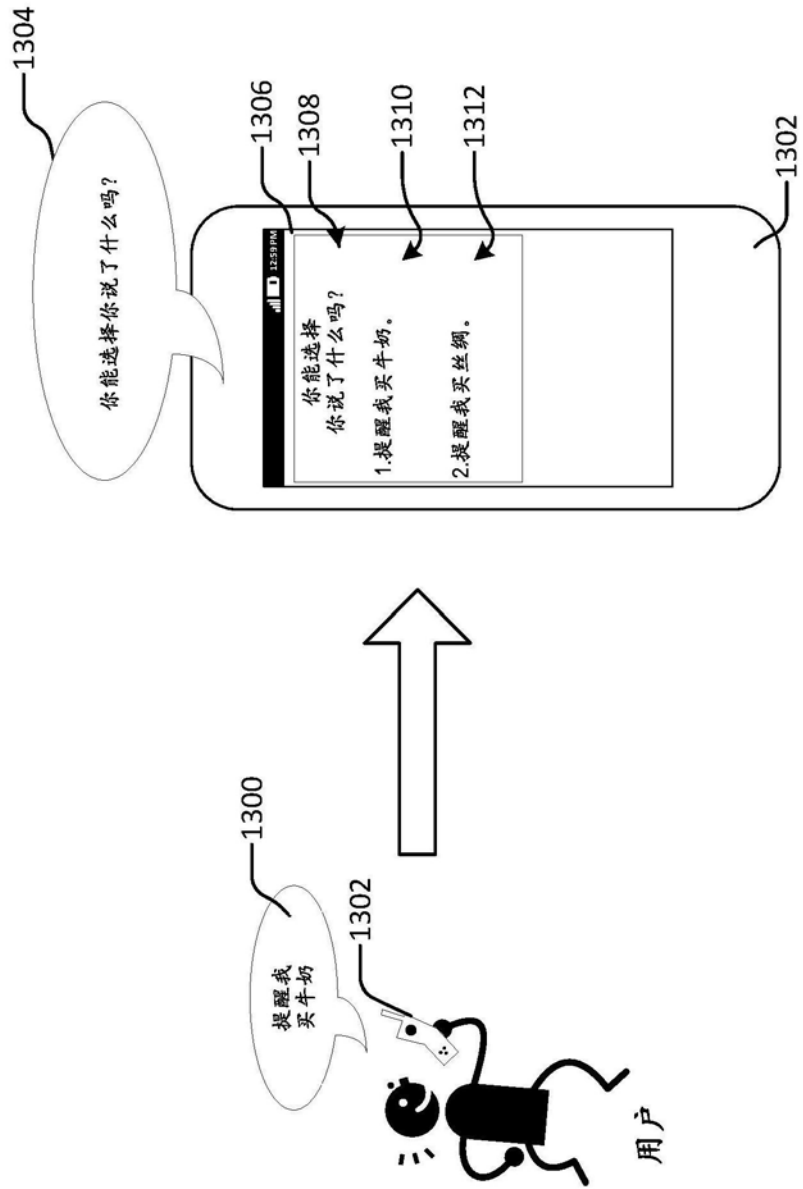


图13

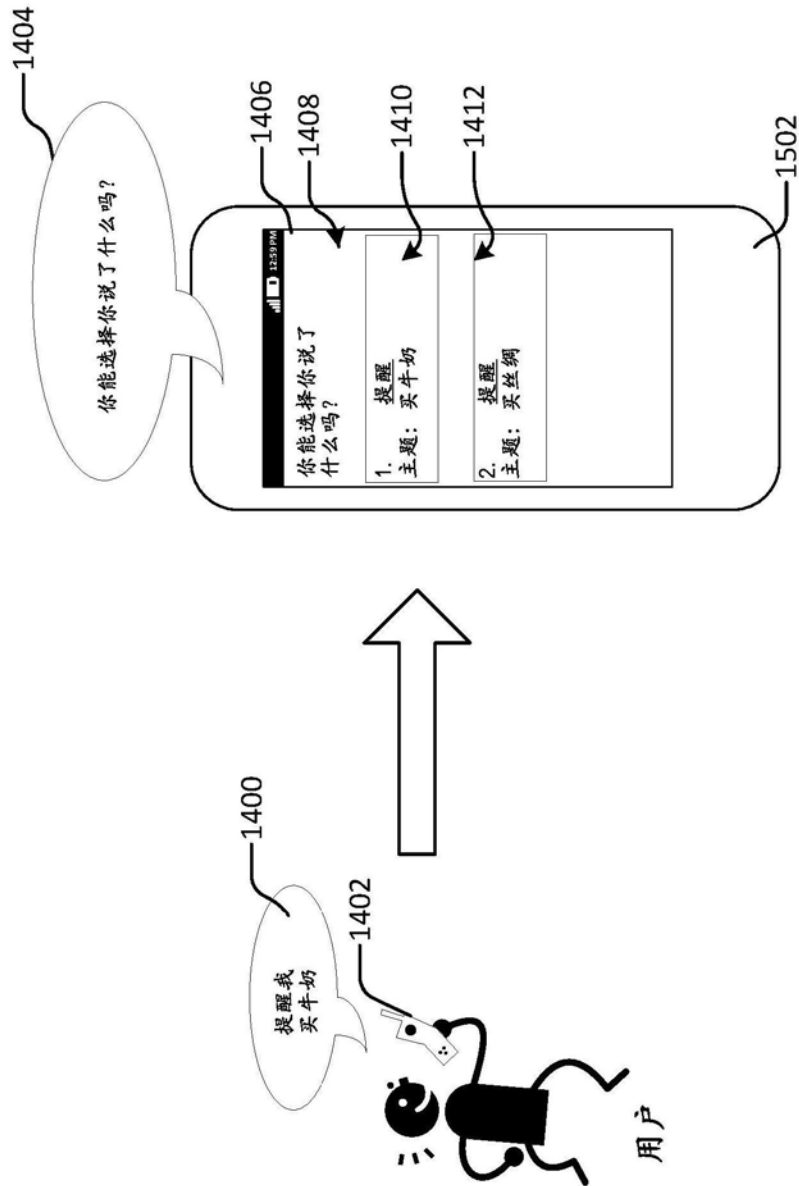


图14

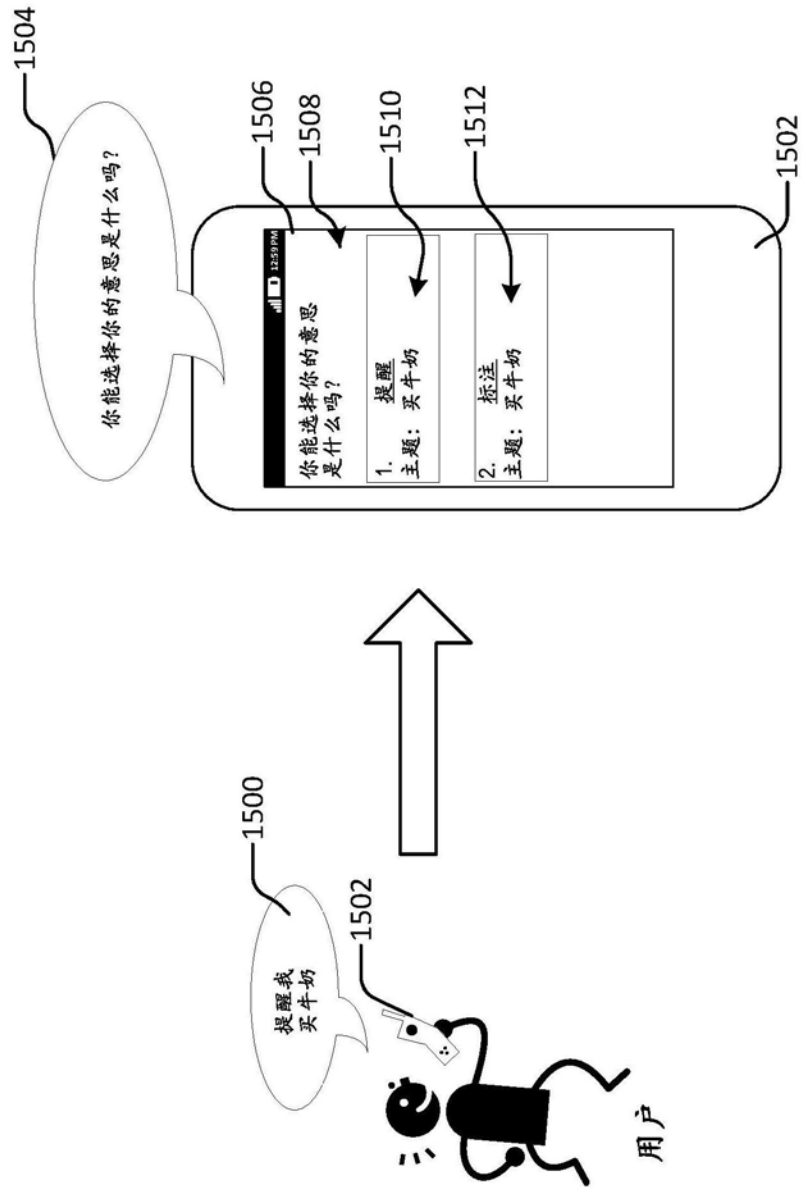


图15

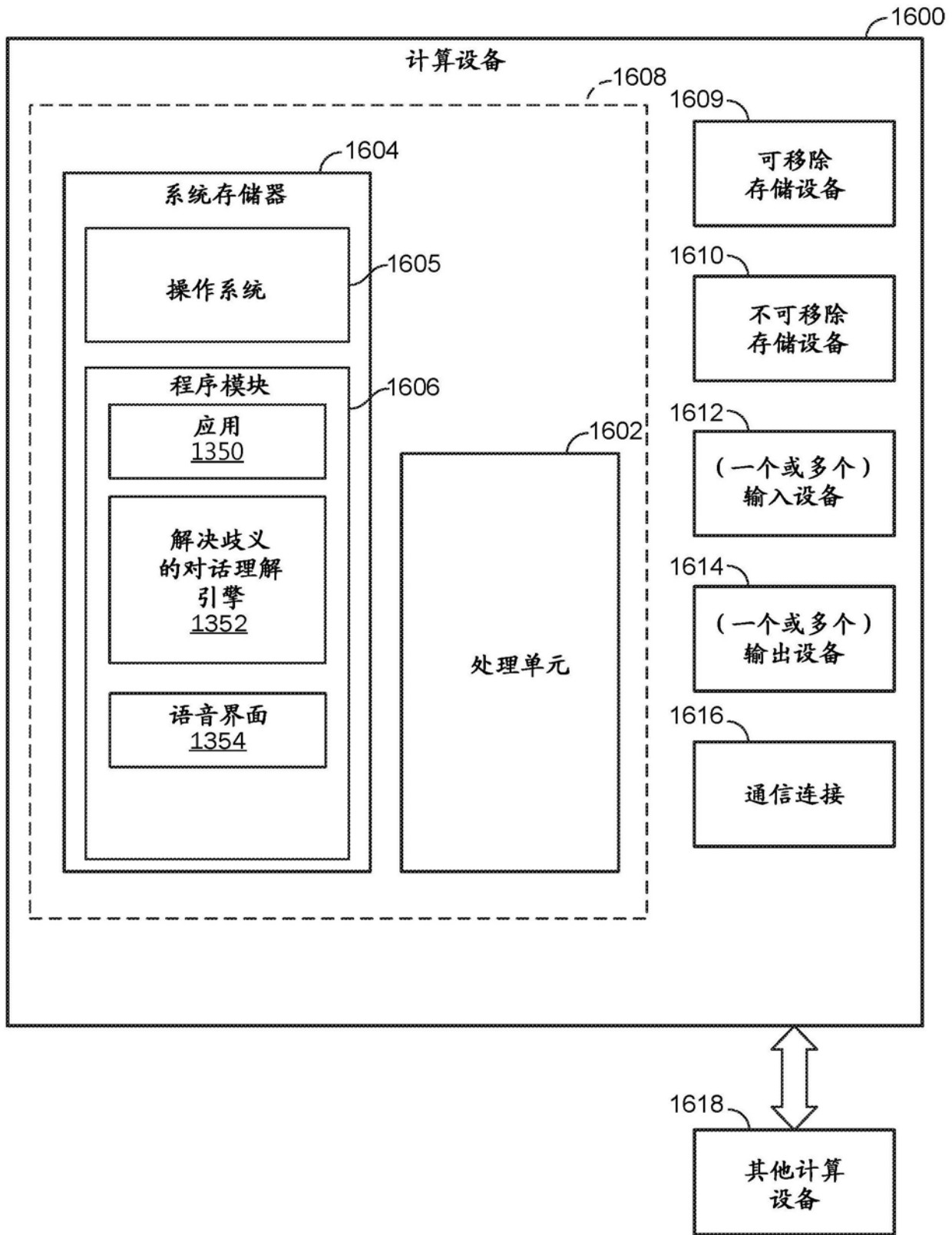


图16

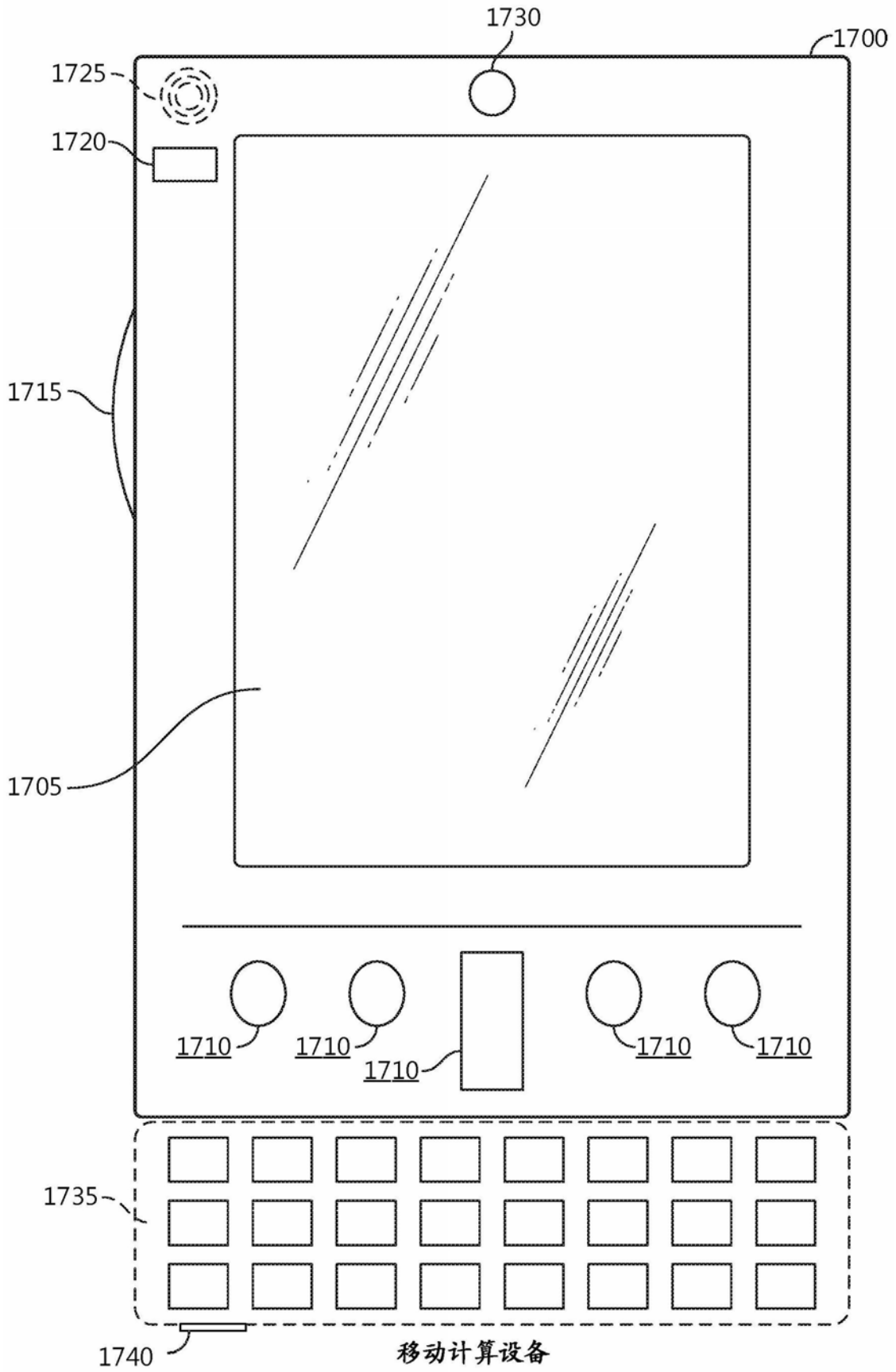


图17A

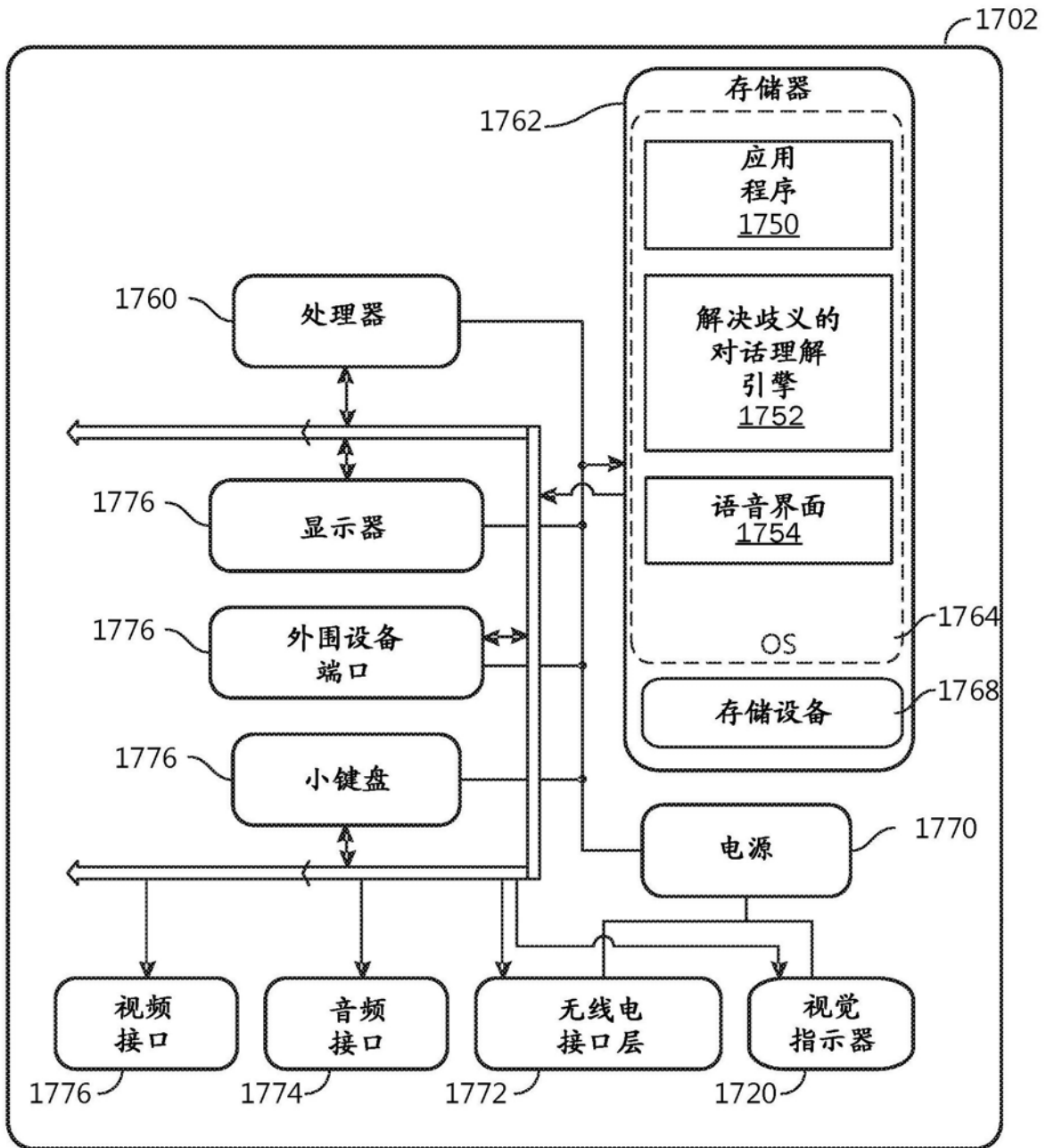


图17B

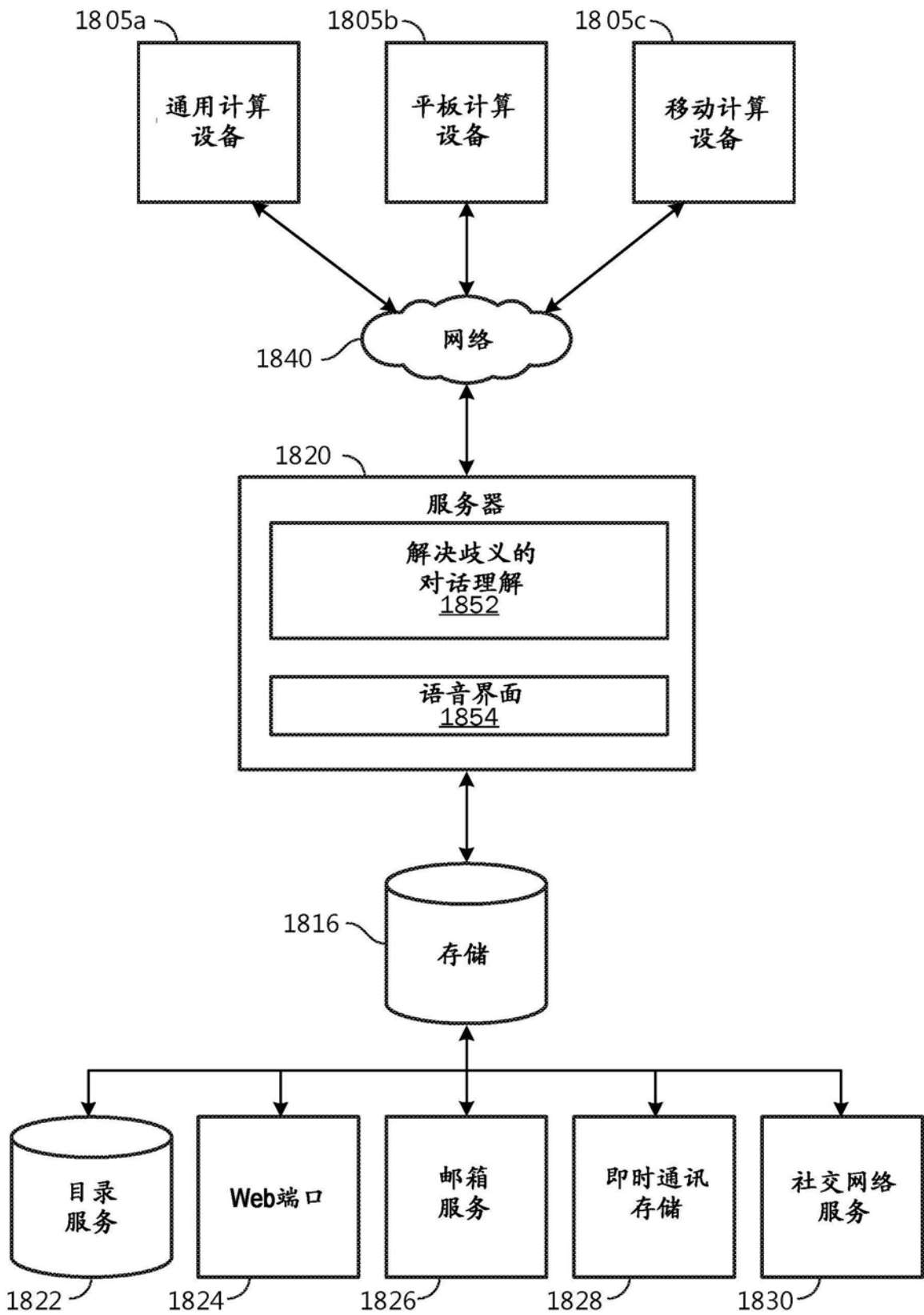


图18