



## (12)发明专利

(10)授权公告号 CN 103226949 B

(45)授权公告日 2017.02.01

(21)申请号 201210599203.8

G10L 15/22(2006.01)

(22)申请日 2012.10.08

G10L 15/26(2006.01)

(65)同一申请的已公布的文献号

H04M 1/725(2006.01)

申请公布号 CN 103226949 A

(43)申请公布日 2013.07.31

(30)优先权数据

13/250,854 2011.09.30 US

(73)专利权人 苹果公司

地址 美国加利福尼亚

(72)发明人 T·R·格鲁博尔 C·D·布瑞哈姆

D·S·科恩 G·诺威克

B·S·菲利普斯

(74)专利代理机构 北京市金杜律师事务所

11256

代理人 王茂华

(51)Int.Cl.

G06F 17/28(2006.01)

(56)对比文件

US 2007/0282595 A1,2007.12.06,全文.

CN 101535983 A,2009.09.16,全文.

US 6324512 B1,2001.11.27,全文.

US 6353794 B1,2002.03.05,全文.

CN 101939740 A,2011.01.05,说明书第

[0001],[0035]-[0043],[0050],[0055]-

[0058],[0068],[0088],[0090]-[0091]段,权利要求1、14,图1.

US 2003/0233230 A1,2003.12.18,全文.

US 2007/0208579 A1,2007.09.06,全文.

US 7302392 B1,2007.11.27,全文.

US 2008/0015864 A1,2008.01.17,全文.

审查员 康丹丹

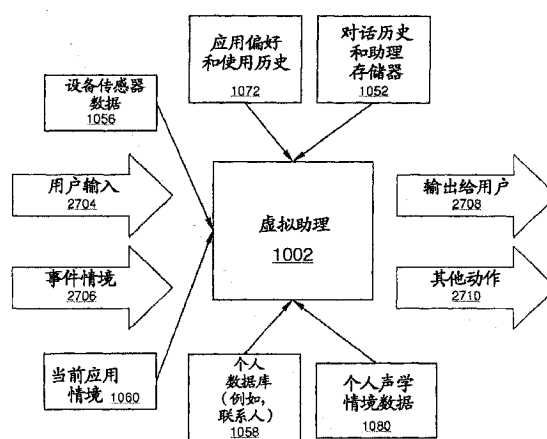
权利要求书5页 说明书35页 附图32页

(54)发明名称

在虚拟助理中使用情境信息来促进命令的处理

(57)摘要

本发明涉及在虚拟助理中使用情境信息来促进命令的处理。虚拟助理使用情境信息来补充来自用户的自然语言或手势输入。情境帮助澄清用户的意图并减少该用户输入的候选解释的数量,并且减少该用户要提供过多澄清输入的需要。情境可以包括任意可以得到的信息,该信息可被所述助理用来补充明确的用户输入以约束信息处理问题和/或个人化结果。可以在处理的不同阶段使用情境来约束解决方案,所述处理包括,例如,语音识别、自然语言处理、任务流处理和对话生成。



1. 一种用于解释用户输入以在计算设备上执行任务的计算机实施的方法,包括:

提示用户进行输入;

接收语音用户输入;

从情境源接收情境信息,接收的所述情境信息包括描述声学环境的数据,所述语音用户输入在所述声学环境中被接收;

解释接收的语音用户输入以得到用户意图的表示;

至少部分地基于所得到的用户意图的表示,识别至少一个任务和用于该任务的至少一个参数;

使用所述至少一个参数执行所述至少一个任务,以得到结果;

基于所得到的结果生成对话响应;以及

输出所生成的对话响应;

其中使用所接收的情境信息执行提示用户进行输入、解释接收的语音用户输入、识别至少一个任务和用于该任务的至少一个参数、以及生成对话响应中的至少一个步骤。

2. 根据权利要求1所述的方法,其中所接收的情境信息还包括从包括以下各项的组中选择的至少一个:

应用情境;

关联于该用户的个人数据;

来自关联于该用户的数据库的数据;

从对话历史获得的数据;

从至少一个传感器接收的数据;

应用偏好;

应用使用历史;

描述事件的数据;

当前对话状态;

由该用户先前提供的输入;

位置;

本地时间;以及

环境条件。

3. 根据权利要求1所述的方法,其中:

提示用户包括经由会话接口提示用户;以及

接收所述语音用户输入包括经由该会话接口通过以下操作来接收所述语音用户输入:

接收所述语音用户输入;以及

将所述语音用户输入转换为文本表示,并且其中将所述语音用户输入转换为文本表示包括:

生成所述语音用户输入的多个候选文本解释;以及

至少排序所生成的候选文本解释的一个子集;

其中使用所接收的情境信息执行所述生成和排序步骤中的至少一个,并且其中在所述生成和排序步骤的至少一个中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:

描述在其中接收所述语音用户输入的声学环境的所述数据；  
从至少一个传感器接收的数据；  
从关联于该用户的数据库获得的词汇；  
关联于应用偏好的词汇；  
从使用历史获得的词汇；以及  
当前会话状态。

4. 一种用于解释用户输入以执行任务的系统，包括：

输出设备，被配置为提示用户进行输入；  
输入设备，被配置为接收语音用户输入；  
至少一个处理器，通信地耦合于该输出设备和该输入设备，被配置为执行以下操作：  
从情境源接收情境信息，接收的所述情境信息包括描述声学环境的数据，所述语音用户输入在所述声学环境中被接收；  
解释所接收的语音用户输入以得到用户意图的表示；  
至少部分地基于所得到的用户意图的表示，识别至少一个任务和用于该任务的至少一个参数；

使用所述至少一个参数执行所述至少一个任务，以得到结果；以及  
基于所得到的结果生成对话响应；

其中所述输出设备被进一步配置为输出所生成的对话响应；

并且其中使用所接收的情境信息执行提示用户进行输入、解释所接收的语音用户输入、识别所述至少一个任务和用于该任务的至少一个参数、和生成对话响应中的至少一个。

5. 根据权利要求4所述的系统，其中所接收的情境信息还包括从包括以下各项的组中选择的至少一个：

应用情境；  
关联于该用户的个人数据；  
来自关联于该用户的数据库的数据；  
从对话历史获得的数据；  
从至少一个传感器接收的数据；  
应用偏好；  
应用使用历史；  
描述事件的数据；  
当前对话状态；  
由该用户先前提供的输入；  
位置；  
本地时间；以及  
环境条件。

6. 根据权利要求4所述的系统，其中：

该输出设备被配置为经由会话接口提示用户；并且

该输入设备被配置为通过接收所述语音用户输入而经由该会话接口来接收所述语音用户输入；

以及其中所述至少一个处理器被配置为将所述语音用户输入转换为文本表示。

7. 根据权利要求6所述的系统,其中所述至少一个处理器被配置为通过以下操作将所述语音用户输入转换为文本表示:

生成所述语音用户输入的多个候选文本解释;以及

至少排序所生成的候选文本解释的一个子集;

其中使用所接收的情境信息执行所述生成和排序操作中的至少一个。

8. 根据权利要求7所述的系统,其中在所述生成和排序操作中的至少一个中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:

描述在其中接收所述语音用户输入的声学环境的所述数据;

从至少一个传感器接收的数据;

从关联于该用户的数据库获得的词汇;

关联于应用偏好的词汇;

从使用历史获得的词汇;以及

当前对话状态。

9. 根据权利要求4所述的系统,其中所述输出设备被配置为通过至少部分地基于所接收的情境信息生成至少一个提示,来提示该用户。

10. 根据权利要求4所述的系统,其中所述至少一个处理器被配置为通过至少部分地基于所接收的情境信息对所接收的语音用户输入执行自然语言处理,来解释所接收的语音用户输入以得到用户意图的表示。

11. 根据权利要求10所述的系统,其中在解释所接收的语音用户输入中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:

描述事件的数据;

应用情境;

由该用户先前提供的输入;

关于该用户的已知信息;

位置;

日期;

环境条件;以及

历史。

12. 根据权利要求4所述的系统,其中所述至少一个处理器被配置为通过至少部分地基于所接收的情境信息识别至少一个任务和用于该任务的至少一个参数,来识别至少一个任务和用于该任务的至少一个参数。

13. 根据权利要求12所述的系统,其中在识别至少一个任务和用于该任务的至少一个参数中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:

描述事件的数据;

来自关联于该用户的数据库的数据;

从至少一个传感器接收的数据;

应用情境;

由该用户先前提供的输入;

关于该用户的已知信息；  
位置；  
日期；  
环境条件；以及  
历史。

14. 根据权利要求4所述的系统，其中所述至少一个处理器被配置为通过至少部分地基于所接收的情境信息生成对话响应，来生成对话响应。

15. 根据权利要求14所述的系统，其中在生成对话响应中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个：

来自关联于该用户的数据库的数据；  
应用情境；  
由该用户先前提提供的输入；  
关于该用户的已知信息；  
位置；  
日期；  
环境条件；以及  
历史。

16. 根据权利要求4所述的系统，其中所接收的情境信息包括从包括以下各项的组中选择的至少一个：

存储在服务器处的情境信息；以及  
存储在客户端处的情境信息。

17. 根据权利要求4所述的系统，其中所述至少一个处理器被配置为通过以下操作从情境源接收情境信息：

从情境源请求所述情境信息；以及  
响应于所述请求接收所述情境信息。

18. 根据权利要求4所述的系统，其中所述至少一个处理器被配置为通过在接收所述语音用户输入之前接收情境信息的至少一部分，来从情境源接收情境信息。

19. 根据权利要求4所述的系统，其中所述至少一个处理器被配置为通过在接收所述语音用户输入之后接收情境信息的至少一部分，来从情境源接收情境信息。

20. 根据权利要求4所述的系统，其中所述至少一个处理器被配置为通过以下操作从情境源接收情境信息：

作为初始化操作的一部分接收静态情境信息；以及  
在接收所述语音用户输入之后接收附加的情境信息。

21. 根据权利要求4所述的系统，其中所述至少一个处理器被配置为通过以下操作从情境源接收情境信息：

接收情境信息中的变化的推送通知；以及  
响应于所述推送通知，更新本地存储的情境信息。

22. 根据权利要求4所述的系统，其中所述输出设备、输入设备和至少一个处理器被实施为从包括以下各项的组中选择的至少一个的组件：

电话；  
智能电话；  
平板电脑；  
膝上计算机；  
个人数字助理；  
台式计算机；  
信息亭；  
消费者电子设备；  
消费者娱乐设备；  
音乐播放器；  
照相机；  
电视机；  
电子游戏单元；以及  
机顶盒。

23. 一种用于解释用户输入以在计算设备上执行任务的装置，包括：

第一接收单元，用于从情境源接收情境信息；

提示单元，用于提示用户进行输入；

第二接收单元，用于接收语音用户输入，其中由所述第一接收单元接收到的所述情境信息包括描述声学环境的数据，所述语音用户输入在所述声学环境中被接收；

解释单元，用于解释所接收的语音用户输入以得到用户意图的表示；

识别单元，用于至少部分地基于所得到的用户意图的表示，识别至少一个任务和用于该任务的至少一个参数；

执行单元，用于使用所述至少一个参数执行所述至少一个任务以得到结果；

生成单元，用于基于所得到的结果生成对话响应；以及

输出单元，用于输出所生成的对话响应；

其中使用所接收的情境信息执行提示用户进行输入、解释所接收的语音用户输入、识别至少一个任务和用于该任务的至少一个参数和生成对话响应中的至少一个。

24. 根据权利要求23所述的装置，其中：

所述提示单元被配置为经由会话接口提示用户；以及

所述第二接收单元被配置为通过接收所述语音用户输入而经由所述会话接口来接收所述语音用户输入；

以及其中所述装置进一步包括被配置为将所述语音用户输入转换为文本表示的转换单元。

25. 根据权利要求24所述的装置，其中所述转换单元被配置为通过以下操作将所述语音用户输入转换为文本表示：

生成所述语音用户输入的多个候选文本解释；以及

至少排序所生成的候选文本解释的一个子集；

其中使用所接收的情境信息执行所述生成和排序操作中的至少一个。

## 在虚拟助理中使用情境信息来促进命令的处理

[0001] 对相关申请的交叉引用

[0002] 本申请要求优先权,作为题为“Contextual Voice Commands”、律师案卷号为P7393US1、申请日为2009年6月5日的、美国专利申请序列号12/479,477的部分继续,其全部公开内容通过引用包含于此。

[0003] 本申请进一步要求优先权,作为题为“IntelligentAutomated Assistant”、律师案卷号为P10575US1、申请日为2011年1月10日的、美国专利申请序列号12/987,982的部分继续,其全部公开内容通过引用包含于此。

[0004] 美国专利申请序列号12/987,982要求题为“Intelligent Automated Assistant”、律师案卷号为SIRIP003P、申请日为2010年1月18日的、美国临时专利申请序列号12/987,982的优先权,其全部公开内容通过引用包含于此。

[0005] 本申请进一步要求题为“Generating and Processing Data Items That Represent Tasks to Perform”、律师案卷号为P11337P1、申请日为2011年6月3日的、美国临时专利申请序列号61/493,201的优先权,其全部公开内容通过引用包含于此。

[0006] 本申请与题为“Generating and Processing Task Items that Represent Tasks to Perform”、律师案卷号为P11337US1、与本申请同日提交的、美国专利申请序列号\_\_\_\_/\_\_\_\_相关联,其全部公开内容通过引用包含于此。

[0007] 本申请与题为“Automatically Adapting User Interfaces for Hands-Free Interaction”、律师案卷号为P11357US1、与本申请同日提交的、美国专利申请序列号\_\_\_\_/\_\_\_\_相关联,其全部公开内容通过引用包含于此。

### 技术领域

[0008] 本发明涉及虚拟助理,并且更具体地涉及用于改进对提供给这样的助理的命令的解释和处理的机制。

### 背景技术

[0009] 现今的电子设备能够经由因特网并从其它的源访问大量、不断增加并且多种多样的功能、服务和信息。用于这样的设备的功能正迅速增加,众多消费者设备、智能电话、平板电脑等能够运行软件应用以执行多种任务并提供不同类型的信息。通常,每个应用程序、功能、网点、或特征具有其自己的用户接口和其自己的操作范例,对用户而言,其中的许多难以学习或无法控制。另外,许多用户可能甚至难以发现什么功能和/或信息可在其电子设备上或各种网站上得到;因此,这些用户可能遭受挫折或受到打击,或可能完全无法以有效的方式使用它们可得到的资源。

[0010] 特别地,初学者用户、或以一些方式受到伤害的或伤残的、和/或老年的、繁忙的、心烦意乱的、和/或操作车辆的个人可能难于有效地与他们的电子设备进行接口,和/或有效地使用在线服务。这样的用户特别可能难于应对对于它们的使用可能有用的大量多变且不一致的功能、应用和网站。

[0011] 因此,现有系统经常难以使用和操纵,并且经常呈现给用户不一致且遭受打击的接口,这样的接口经常阻碍用户有效地使用技术。

[0012] 智能自动化助理,在此也称为虚拟助理,可提供人与计算机之间的改进的接口。这样的助理允许用户以语音和/或文本形式使用自然语言与系统或设备接口,所述助理可以如申请日为2011年1月10日、律师案卷号为P10575US1、题为“IntelligentAutomatedAssistant”的相关的美国专利申请序列号12/987,982所描述的那样实施,该申请的全部公开内容通过引用包含于此。这样的助理解释用户输入,将用户的意图实施为任务和针对那些任务的参数,执行服务以支持那些任务,并产生用户可理解的输出。

[0013] 虚拟助理可利用多个信息源中的任意一个以处理用户输入,包括例如知识基础、模型、和/或数据。在许多情况下,单独的用户的输入不足以清楚地定义用户的意图和要执行的任务。这可能由于输入流中的噪声、用户之间的个体差异、和/或自然语言的固有的含糊。例如,电话上的文本消息应用程序的用户可调用虚拟助理并说出命令“呼叫她(call her)”。当这样的命令为完全合理的英语时,其并不是精确的、可执行的陈述,因为对于该请求存在多种解释以及多种可能的解决方案。因此,在没有进一步的信息的情况下,虚拟助理可能不能正确地解释和处理这样的输入。该类型的含糊会导致执行错误、执行不正确的动作、和/或通过澄清输入的请求过多地麻烦用户。

## 发明内容

[0014] 根据本发明的多个实施例,虚拟助理使用情境信息(此处也称为“情境”)以补充来自用户的自然语言或手势输入。这有助于澄清用户的意图并减少该用户输入的候选解释的数量,并减少对该用户提供过多澄清输入的需要。情境可以包括任何可以得到的信息,该助理使用该信息来补充明确的用户输入以约束信息处理问题和/或个人化结果。例如,如果来自该用户的输入包括代词(诸如在命令“call her(呼叫她)”中的“her(她)”,该虚拟助理可以使用情境推断代词的所指,例如确定被呼叫人的身份和/或要使用的电话号码。在此描述情境的其它使用。

[0015] 根据本发明的多个实施例,在电子设备上实施的虚拟助理中可以实施多种机制中的任意一种,用于获取并应用情境信息来执行计算。在多个实施例中,所述虚拟助理是如申请日为2011年1月10日的、律师案卷号为P10575US1、题为“Intelligent Automated Assistant”的美国专利申请序列号12/987,982中所描述的智能自动化助理,该申请的全部公开内容通过引用包含于此。这样的助理使用自然语言对话以集成的、会话的方式与该用户交流,并且在适当时候调用外部服务来获取信息或执行各种动作。根据此处描述的技术,在这样的助理中使用情境信息,例如,在执行信息处理功能(诸如语音识别、自然语言处理、任务流处理和对话生成)时来减少歧义。

[0016] 根据本发明的多个实施例,可以配置、设计和/或操作虚拟助理来使用情境以执行各种不同类型的操作、功能和/或特征,和/或来组合其上安装了虚拟助理的电子设备的多个特征、操作和应用。在某些实施例中,当执行以下各项中的任意项或所有项时,本发明的虚拟助理可以使用情境:主动引导来自用户的输入、解释用户意图、在竞争的解释中去歧义、根据需要请求和接收澄清信息和/或基于已辨别的意图执行(或开始)动作。

[0017] 例如,通过激活和/或与在电子设备上可获得的任意应用或服务以及在电子网络(诸如因特网)上可获得的服务进行接口,可以执行动作。在多个实施例中,可以经由应用编程接口(API)或通过任意其它适当的机制执行外部服务的这种激活。如此,根据本发明的多个实施例实现的虚拟助理可以统一、简化并改善关于电子设备的多个不同应用和功能以及关于在因特网上可获得的服务的用户体验。因此减轻了用户学习在所述设备上和在网络连接的服务上什么功能可获得、如何与这些服务接口以得到他或她想要的以及如何解释从这些服务接收到的输出的负担;更确切地,本发明的助理可以作为该用户和这些多样服务之间的中介。

[0018] 另外,在多个实施例中,本发明的虚拟助理提供了会话的接口,所述用户可以发现与传统图形用户界面相比,该会话接口更加直观并且更少繁琐。所述用户可以使用多个可用的输入和输出机制,诸如语音、图形用户界面(按钮和链接)、文本输入等中的任意一个以会话对话的形式与所述助理交流。可以使用多个不同平台(例如设备API、网络、电子邮件等)中的任意一个,或其任意组合实施本系统。可以在这样的会话的情境中将针对附加输入的请求呈现给用户。可以使用短期和长期存储器(short and long term memory),以便在考虑先前的事件和在特定会话中的通信、以及关于该用户的历史信息和简档信息的适当的情境中解释用户输入。

[0019] 此外,在多个实施例中,从用户与在设备上的特征、操作或应用接口得到的情境信息可以用于简化在该设备上或在其它设备上的其它特征、操作或应用的操作。例如,所述虚拟助理可以使用电话呼叫的情境(诸如被呼叫人)来简化文本消息的开始(例如确定该文本消息应当发送给同一个人,而无须用户必须明确地指定该文本消息的接收方)。本发明的虚拟助理因此能够解释诸如“send him a text message(向他发送文本消息)”的指令,其中根据从当前电话呼叫和/或从在该设备上的任意特征、操作或应用得到的情境信息解释“him(他)”。在多个实施例中,所述虚拟助理考虑多种类型的可用的情境数据来确定要使用的地址簿联系人、要使用的联系人数据、对于该联系人要使用的电话号码等,因此用户不需要手动重新指定这样的信息。

[0020] 情境信息源包括,例如但不限于:用作为对上述助理的接口的设备的当前状态,诸如当前时间、位置、应用或数据对象;个人数据,诸如用户的地址簿、日历和应用使用历史;以及在该用户和所述虚拟助理之间的对话的状态,诸如最近提及的人和/或地点。

[0021] 情境可以应用于在所述虚拟助理的操作中的各种各样的计算和推断。例如,当处理用户输入时可以使用情境来减少歧义或者另外地约束解决方案的数量。因此情境可以用于在处理的不同阶段期间约束所述解决方案,所述处理例如包括但不限于:

[0022] 语音识别-接收语音输入并且生成文本形式的候选解释,例如,“call her(呼叫她)”、“collar”和“call Herb(呼叫Herb)”。可以使用情境约束语音识别模块考虑哪些词或短语、如何排列它们以及哪些由于大于考虑的阈值而被接受。例如,该用户的地址簿可以增加个人姓名到其它的语音的语言通用模型,使得这些姓名可以被识别并且被给予优先权。

[0023] 自然语言处理(NLP)-解析文本并将词与句法和语义角色关联,例如,确定用户输入是关于打电话给代词“her(她)”所指的人,并且寻找到该人的特定数据表示。例如,文本消息应用的情境可以帮助约束“her(她)”的解释为代表“我正在与之进行文本交谈的人”。

[0024] 任务流处理-识别用户任务、任务步骤和用于帮助该任务的任务参数,例如,用于

被称为“her(她)”的人的电话号码。此外,文本消息应用的情境可以约束该电话号码的解释来指示系统应当使用当前或最近用于文本消息会话的号码。

[0025] 对话生成-生成助理响应,作为与该用户会话的与其任务有关的一部分,例如,用响应“OK,I'll call Rebecca on her mobile...(好的,我将呼叫Rebecca的移动电话...)”来解释该用户的意图。冗长的和非正式的语气的级别是可以通过情境信息引导的选择。

[0026] 在多个实施例中,本发明的虚拟助理可以控制电子设备的各种特征和操作。例如,该虚拟助理可以经由API或通过其它手段调用与在设备上的功能和服务,来执行该设备上使用传统用户接口可以另外地启动的功能和操作。这些功能的操作可以包括,例如,设置闹钟、打电话、发送文本消息或电子邮件消息、增加日历事件等。在用户和助理之间的会话的对话的情境中,这些功能的操作可以实现为附加功能。在这样的对话的情境中,这些功能的操作可以由该用户指定,或者基于对话的情境它们可以被自动执行。本领域技术人员将认识到该助理可以因此被用作为用于启动和控制电子设备上的不同操作的控制机构,其可以被用作为传统机构(诸如按钮或图形用户界面)的替代。如此处所述,情境信息可以用于关于作为控制机构使用的虚拟助理的通知和改善。

## 附图说明

[0027] 附图示出本发明的多个实施例,并且连同说明书,用以根据所述实施例解释本发明的原理。本领域技术人员将认识到附图所示出的特定的实施例仅仅是示例性的,并不意图限定本发明的范围。

[0028] 图1是描绘根据一实施例的虚拟助理和能够影响其操作的情境源的一些例子的框图。

[0029] 图2是描绘根据一实施例的用于在虚拟助理中的处理的多个阶段使用情境的方法的流程图。

[0030] 图3是描绘根据一实施例的用于在语音引导和解释中使用情境的方法的流程图。

[0031] 图4是描绘根据一实施例的用于在自然语言处理中使用情境的方法的流程图。

[0032] 图5是描绘根据一实施例的用于在任务流处理中使用情境的方法的流程图。

[0033] 图6是描绘根据一实施例的分布于客户端和服务端之间的情境的源的例子的框图。

[0034] 图7a至7d是描绘根据多个实施例的用于获取和协调情境信息的机制的例子的事件图。

[0035] 图8a至8d描绘了可以与本发明多个实施例结合使用的情境信息的各种表示的例子。

[0036] 图9描绘了根据一实施例的指定用于多个情境信息源的通信和高速缓存策略的配置表的例子。

[0037] 图10是描绘根据一实施例的在接口程序处理期间访问在图9中配置的情境信息源的例子的事件图。

[0038] 图11至13是描绘根据一实施例的在文本消息域(domain)中使用应用情境来为代词获得所指对象的例子的一系列截图。

- [0039] 图14是示出根据一实施例的用于名称歧义消除的虚拟助理提示的截图。
- [0040] 图15是示出根据一实施例的使用会话情境以推断用于命令的位置的虚拟助理的截图。
- [0041] 图16是描绘根据一实施例的使用电话收藏列表作为情境源的的例子截图。
- [0042] 图17至20是描绘根据一实施例的使用当前应用情境来解释并实施命令的例子的一系列截图。
- [0043] 图21是描绘使用当前应用情境来解释调用了不同应用的命令的例子截图。
- [0044] 图22至24是描绘根据一实施例的使用到来文本消息形式的事件情境的例子的一系列截图。
- [0045] 图25A至25B是描绘根据一实施例的使用先前对话情境的例子的一系列截图。
- [0046] 图26A至26B是描绘根据一实施例的用于在候选解释中进行选择的用户接口的例子截图。
- [0047] 图27是描绘虚拟助理系统的一实施例的例子框图。
- [0048] 图28是描绘根据至少一实施例的适用于实现虚拟助理至少一部分的计算设备的框图。
- [0049] 图29是描绘根据至少一实施例的用于在独立计算系统上实现虚拟助理的至少一部分的架构的框图。
- [0050] 图30是描绘根据至少一实施例的用于在分布式计算设备上实现虚拟助理的至少一部分的架构的框图。
- [0051] 图31是描绘示出一些不同类型的客户端和操作模式的系统架构的框图。
- [0052] 图32是描绘根据一实施例的通过互相通信来实现本发明的客户端和服务器的框图。
- [0053] 图33是根据某些实施例的用于解释用户输入以执行任务的系统的功能框图。
- [0054] 图34是根据某些实施例的用于解释用户输入以在计算设备上执行任务的装置的功能框图。

## 具体实施方式

[0055] 根据本发明多个实施例,为支持虚拟助理的操作,获得并且应用各种各样的情境信息来执行信息处理功能。出于描述的目的,“虚拟助理”等同于术语“智能自动化助理”,二者都指执行一个或多个以下功能的任意信息处理系统:

- [0056] ●解释语音的和/或文本的形式的人类语言输入;
- [0057] ●将用户意图的表示操作为能够执行的形式,诸如具有步骤和/或参数的任务的表示;
- [0058] ●通过调用程序、方法、服务、API等执行任务表示;以及
- [0059] ●以语言和/或图形形式向用户生成输出响应。

[0060] 在申请日为2011年1月10日的、律师案卷号为P10575US1、题为“Intelligent Automated Assistant”的相关的美国专利申请序列号12/987,982中描述了这样的虚拟助理的例子,该申请的全部公开内容通过引用包含于此。

[0061] 现在将参考附图所示的示例性实施例详细描述各种技术。在以下描述中,许多特

定的细节被阐明,以提供对此处描述或引用的一个或多个方面和/或特征的透彻理解。然而,对本领域技术人员来说,显然,此处描述或引用的一个或多个方面和/或特征可以在没有这些特定细节中的一些或全部的情况下实践。在其它例子中,为了不模糊此处描述或引用的一些方面和/或特征,没有详细描述公知的处理步骤和/或结构。

[0062] 本申请中会描述一个或多个不同的发明。此外,对于此处描述的一个或多个发明,本专利申请中可能描述许多实施例,并且这些实施例仅为示例性目的给出。所描述的实施例并不意图任何意义上的限定。一个或多个发明可以广泛地应用于多个实施例,就像从本公开中所显而易见的那样。这些实施例被充分描述以使本领域技术人员能够实践一个或多个发明,并且应当理解在不脱离一个或多个发明的范围的情况下,可以使用其它实施例以及作出结构的、逻辑的、软件的、电子的和其它的改变。因此,本领域技术人员将认识到一个或多个发明可以使用各种修改和替代来实践。参照形成本公开一部分的一个或多个特定实施例或附图来描述一个或多个发明的特定特征,并且在其中一个或多个发明的具体实施例以示例性方式被示出。然而应当理解,这些特征并不限定于在所述一个或多个特定实施例或附图中使用,其中参照所述一个或多个特定实施例或附图来描述这些特征。本公开既不是一个或多个发明的所有实施例的文字描述,也不是必须存在于所有实施例中的一个或多个发明的特征清单。

[0063] 本专利申请中提供的节标题以及本专利申请的名称仅为了方便而并不是要以任何方式限制本公开。

[0064] 除非另外明确说明,相互通信的设备不必相互持续通信。此外,相互通信的设备可以直接地通信或通过一个或多个中间媒介间接地通信。

[0065] 对具有相互通信的多个组件的实施例的描述并不意味需要所有这样的组件。相反,描述多个可选组件来示出一个或多个发明的广泛的可能实施例。

[0066] 此外,虽然处理步骤、方法步骤、算法等可能以前后顺序被描述,但是这样的处理、方法和算法可以被配置为以任意适合的顺序工作。换句话说,本专利申请中所描述的步骤的任意次序或顺序本身并不表明以那样的顺序执行步骤是必须的。而且,一些步骤可以同时执行而不管被描述为非同时发生或者暗示是非同时出现(例如,因为一个步骤在另一步骤之后描述)。此外,通过在附图中对其进行描绘的处理的例示并不意味着所示出的处理不包括对它们的其它变体和修改,并不意味着所示出的处理或其步骤中的任意一个对于一个或多个发明是必要的,并且并不意味着所示出的处理是优选的。

[0067] 当描述单个设备或物品时,显而易见的是多于一个的设备/物品(不论它们是否协作)可以使用来替代单个的设备/物品。类似地,在描述多于一个的设备或物品(不论它们是否协作)的情况下,显而易见的是单个设备/物品可以用来替代所述多于一个的设备或物品。

[0068] 设备的功能和/或特征可选地通过一个或多个没有明确描述为具有这样的功能/特征的其它设备体现。因此,一个或多个发明的其它实施例不需要包括该设备自身。

[0069] 为清楚起见,此处描述或引用的技术或机制有时将以单体的形式描述。然而应当注意的是,除非特别注明,特定实施例包括技术的多次迭代或机制的多个实例。

[0070] 虽然在用于实现智能化助理(也称为虚拟助理)的技术的背景下进行了描述,但是应当理解此处所描述的多个方面和技术也可以布置和/或应用于其它技术领域,所

述其它技术领域涉及相对于软件的人和/或计算机化交互。

[0071] 有关虚拟助理技术(例如,可以被在此处描述的一个或多个虚拟助理系统实施例利用、提供和/或实施的技术)的其它方面在以下的一个或多个中公开,它们的全部公开内容通过引用包含于此。

[0072] ●题为“Intelligent Automated Assistant”、律师案卷号为P10575US1、申请日为2011年1月10日的美国专利申请序列号12/987,982;

[0073] ●题为“IntelligentAutomated Assistant”、律师案卷号为SIRIP003P、申请日为2010年1月18日的、美国临时专利申请序列号61/295,774;

[0074] ●题为“Method And Apparatus for Building an Intelligent Automated Assistant”、申请日为2006年9月8日的美国专利申请序列号11/518,292;以及

[0075] ●题为“System and Method for Semantic Anto-Completion”、申请日为2009年6月12日的美国临时专利申请序列号61/186,414。

[0076] 硬件架构

[0077] 一般地,此处公开的虚拟助理技术可以在硬件上或者软件和硬件的组合上实施。例如,它们可以在操作系统内核中、独立的用户处理中、绑定到网络应用的库包中、特殊构造的机器上和/或网络接口卡上实施。在特定实施例中,此处所公开的技术可以在诸如操作系统的软件中或在运行于操作系统上的应用中实施。

[0078] 此处公开的至少一些虚拟助理实施例的软件/硬件的混合实现方式可以在由存储在存储器中的计算机程序选择性激活或重新配置的可编程机器上实施。这些网络设备可以具有多个可被配置或被设计为利用不同类型的网络通信协议的网络接口。用于一些这样的机器的一般性架构出现在此处公开的描述中。根据特定实施例,此处公开的不同虚拟助理实施例的至少一些特征和/或功能可以实施在一个或多个通用网络主机上,诸如终端用户计算机系统、计算机、网络服务器或者服务器系统、移动计算设备(例如,个人数字助理、移动电话、智能电话、膝上计算机、平板计算机等)、消费者电子设备、音乐播放器,或者任何其它适合的电子设备,路由器、开关等,或者其任意组合。在至少一些实施例中,在此公开的不同虚拟助理实施例的至少一些特征和/或功能可以在一个或多个虚拟计算环境(例如,网络计算云等)中实施。

[0079] 现在参考图28,其示出了描绘适于实现此处公开的虚拟助理的特征和/或功能的至少一部分的计算设备60的框图。计算设备60可以是,例如,终端用户计算机系统、网络服务器或服务器系统、移动计算设备(例如,个人数字助理、移动电话、智能电话、膝上计算机、平板计算机等)、消费者电子设备、音乐播放器,或者任何其它适合的电子设备,或者其任意组合或部分。计算设备60可以适于经由通信网络(诸如因特网),使用用于不管无线还是有线的这样的通信的已知协议,与其它计算设备(诸如客户端和/或服务器)通信。

[0080] 在一实施例中,计算设备60包括中央处理单元(CPU)62、接口68和总线67(诸如外围设备互连(PCI)总线)。当在适当的软件或固件的控制下操作时,CPU62可以负责实施与经特殊配置的计算设备或机器的功能相关的特定功能。例如,在至少一实施例中,用户的个人数字助理(PDA)或智能电话可以被配置或被设计成作为使用CPU62、存储器61和65以及接口68的虚拟助理系统运行。在至少一实施例中,在软件模块/组件控制下,可以使所述CPU62执行一个或多个不同类型的虚拟助理功能和/或操作,例如可以包括操作系统和任何适合的

应用软件、驱动等。

[0081] CPU62可以包括一个或多个处理器63例如,来自Motorola或Intel系列的微处理器或MIPS系列的微处理器的处理器。在一些实施例中,处理器63可以包括用于控制计算设备60操作的专门设计的硬件(例如,专用集成电路(ASIC)、电可擦除可编程只读存储器(EEPROM)、现场可编程门阵列(FPGA)等)。在一特定实施例中,存储器61(诸如非易失性随机存取存储器(RAM)和/或只读存储器(ROM))也是CPU62的组成部分。然而,存在多种不同的方式,其中存储器可以耦合到该系统。存储器块61可以用于各种各样目的,例如,高速缓存和/或存储数据、编程指令等。

[0082] 如此处所使用的,术语“处理器”并不仅仅限定于在现有技术中称为处理器的那些集成电路,而是泛指微控制器、微计算机、可编程逻辑控制器、专用集成电路和任何其它可编程电路。

[0083] 在一实施例中,接口68作为接口卡(有时称为“线路卡”)提供。一般地,它们经计算网络控制数据包的发送和接收,并且有时支持与计算设备60一起使用的其它外围设备。可以提供以太网接口、帧中继接口、电缆接口、DSL接口、令牌环接口等接口。此外,可以提供不同类型的接口,例如,通用串行总线(USB)、串口、以太网、火线、PCI、并口(parallel)、射频(RF)、蓝牙、近场通信(例如,使用近场磁)、802.11(WiFi)、帧中继、TCP/IP、ISDN、快速以太网接口、千兆以太网接口、异步传输模式(ATM)接口、高速串行接口(HSSI)接口、销售点(POS)接口、光纤数据分布式接口(FDDI)等。一般地,这些接口68可以包括适合与适当媒介通信的端口。在一些情况下,它们也可以包括独立的处理器,并且在某些情况下,可以包括易失和/或非易失存储器(例如,RAM)。

[0084] 虽然在图28中所示出的系统示出了用于计算设备60的一特定架构来实施此处描述的发明的技术,但是这决不是其上可以实施此处描述的至少一部分特征和技术的唯一的设备架构。例如,可以使用具有一个或任意数量处理器63的架构,并且这些处理器63可以存在于单个设备中或分布于任意数量设备中。在一实施例中,单个处理器63处理通信及路由计算。在多个实施例中,不同类型的虚拟助理特征和/或功能可以在包括客户端装置(诸如运行有客户端软件的个人数字助理或智能电话)和服务器系统(诸如以下更详细描述的服务系统)的虚拟助理系统中实施。

[0085] 不管网络设备配置如何,本发明的系统可以使用一个或多个存储器或存储器模块(例如,存储器块65),其被配置为存储数据、用于通用网络操作的程序指令和/或与此处描述的虚拟助理技术的功能相关的其它信息。例如,程序指令可以控制操作系统和/或一个或多个应用的操作。也可以将存储器配置为存储此处描述的数据结构、关键词分类信息、广告信息、用户点击和按压信息和/或其它特定非程序信息。

[0086] 因为可以使用这些信息和程序指令来实施此处所描述的系统/方法,因此至少一些网络设备实施例可以包括非瞬时(nontransitory)机器可读存储介质,例如,其可以被配置或被设计来存储用于执行此处描述的不同操作的程序指令、状态信息等。这些非瞬时机器可读存储介质的例子包括但不限于:磁介质,诸如硬盘、软盘及磁带;光介质,诸如CD-ROM盘;磁光盘,诸如光磁软盘;以及被特殊配置来存储和执行程序指令的硬件设备,诸如只读存储器设备(ROM)、闪存、忆阻器、随机存取存储器(RAM)等。程序指令的例子包括诸如由编译器产生的机器代码和包含由该计算机使用解释器执行的高级代码的文件二者。

[0087] 在一实施例中,在独立的计算设备上实施本发明的系统。现在参考图29,示出了描绘根据至少一实施例的用于在独立计算系统上实施虚拟助理的至少一部分的架构的框图。计算设备60包括运行用于实现虚拟助理1002的软件的处理器63。输入设备1206可以是适于接收用户输入的任意类型,包括例如键盘、触摸屏、麦克风(例如,用于语音输入)、鼠标、触摸板、跟踪球、五路开关、操纵杆和/或其任意组合。输出设备1207可以是屏幕、扬声器、打印机和/或其任意组合。存储器1210可以是具有现有技术中已知的结构和架构的随机存取存储器,在运行软件期间由处理器63使用。存储设备1208可以是用于以数字形式存储数据的任何磁的、光的和/或电的存储设备;例子包括闪存、磁性硬盘、CD-ROM等。

[0088] 在另一实施例中,在分布式计算网络,诸如具有任意数量的客户端和/或服务器的分布式计算网络上实施本发明的系统。现在参考图30,其示出了描述根据至少一实施例的用于在分布式计算网络上实施虚拟助理的至少一部分的架构的框图。

[0089] 在图30示出的配置中,提供了任意数量的客户端1304;每一客户端1304可以运行用于实施本发明的客户端侧部分的软件。此外,可以提供任意数量的服务器1340来处理从客户端1304接收的请求。客户端1304和服务器1340可以经由电子网络1361(诸如因特网)相互通信。可以使用任意已知的网络协议,包括例如有线和/或无线协议,来实施网络1361。

[0090] 此外,在一实施例中,当需要获得附加信息或涉及存储与特定用户的先前交互有关的数据时,服务器1340可以调用外部服务1360。与外部服务1360的通信可以例如经由网络1361进行。在多个实施例中,外部服务1360包括与硬件设备自身有关的或安装于其上的网络使能的服务和/或功能。例如,在一实施例中,在智能电话或其它电子设备上实施助理1002,助理1002可以获得存储在日历应用(“app”)、联系人和/或其它源中的信息。

[0091] 在多个实施例中,助理1002可以控制其上安装有该助理1002的电子设备的多个特征和操作。例如,助理1002可以经由API或通过其它方式调用与设备上的功能和应用接口的外部服务1360,以执行功能和操作,所述功能和操作否则可能使用在设备上的传统用户接口启动。这些功能和操作可以包括,例如,设置闹钟、打电话、发送文本消息或电子邮件消息、增加日历事件等。这些功能和操作可以作为用户和助理1002之间的会话对话的情境中的附加功能执行。这些功能和操作可以在这样的对话的情境中由该用户指定,或者它们可以基于所述对话的情境自动执行。本领域技术人员将认识到助理1002可以因此被用于发起并控制电子设备上的多个操作的控制机制,其可以用作传统机制(诸如按钮或图形用户界面)的替代。

[0092] 例如,用户可以向助理1002提供输入,诸如“I need to wake tomorrow at 8am(我需要在明天上午8点钟叫醒)”。一旦助理1002已经确定该用户的意图,使用此处所描述的技术,助理1002可以调用外部服务1340来与该设备上的闹钟功能或应用进行接口。助理1002代表用户设置闹钟。如此,该用户可以使用助理1002取代用于在该设备上设置闹钟或执行其它功能的传统机制。如果该用户的请求是模糊(ambiguous)的或者需要进一步明确,助理1002可以使用此处描述的各种技术,包括主动引导、解释、建议等,并且包括获得情境信息,从而调用正确的服务1340并且采取所意图的动作。在一实施例中,在调用服务1340来执行功能之前,助理1002可以提示用户进行确认和/或请求来自任何适当的源的情境信息。在一实施例中,用户可以选择性地禁止助理1002调用特定服务1340的能力,或者根据需要,可以禁止所有这样的服务调用。

[0093] 本发明的系统可以被实施为具有任意数量的不同类型的客户端1304和操作模式。现在参考图31,其示出了一框图,所述框图描绘了示出一些个不同类型的客户端1304和操作模式的系统架构。本领技术人员将认识到图31中所示出的各种类型的客户端1304和操作模式仅仅是示例性的,并且可以使用除所描绘的那些以外的客户端1304和/或操作模式实施本发明的系统。另外,该系统可以独立地或以任意组合的方式包括任意的或所有的这种客户端1304/或操作模式。所描绘的例子包括:

[0094] ●具有输入/输出设备和/或传感器1402的计算机设备。客户端组件可以被部署在任意的这样的设备1402上。可以使用网络浏览器1304A或用于使能经由网络1361与服务器1340通信的其它软件应用来实施至少一个实施例。输入和输出通道可以是任意类型的,包括例如视觉和/或听觉通道。例如,在一实施例中,可以使用基于语音的通信方法实施本发明的系统,使得用于盲人的、等同于网络浏览器的助理的实施例由语音驱动并且使用语音来输出。

[0095] ●具有I/O和传感器1406的移动设备,所述客户端可以被实施为该移动设备1304B上的应用。这包括,但不限于,移动电话、智能电话、个人数字助理、平板设备、网络游戏控制台等。

[0096] ●具有I/O和传感器1410的消费者设备,所述客户端可以被实施为该设备1304C上的嵌入式应用。

[0097] ●具有仪表板接口和传感器1414的车辆和其它交通工具,所述客户端可以被实施为嵌入式系统应用1304D。其包括,但不限于,车辆导航系统、声音控制系统、车内娱乐系统等。

[0098] ●网络计算设备,诸如路由器1418,或者驻留在网络上或与网络接口的任何其它设备,所述客户端可以被实施为设备驻留应用1304E。

[0099] ●电子邮件客户端1424,针对其所述助理的实施例经由电子邮件形式(modality)服务器1426被连接。电子邮件形式服务器1426充当通信桥梁,例如把该用户输入作为电子邮件消息发送给该助理和发送该助理的输出给该用户作为答复。

[0100] ●即时消息客户端1428,针对其所述助理的实施例经由消息形式服务器1430被连接。消息形式服务器1430充当通信桥梁,把该用户输入作为消息发送给该助理和发送该助理的输出给该用户作为答复的合成语音。

[0101] ●语音电话1432,针对其所述助理的实施例经由互联网协议电话(VoIP)形式服务器1430被连接。VoIP形式服务器1430充当通信桥梁,把该用户输入作为说话的声音发送给该助理和发送该助理的输出给该用户,例如作为答复的合成语音。

[0102] 对于发送消息平台,包括但不限于,电子邮件、即时消息、讨论论坛、群组聊天会话、实时帮助或者客户支持会话等,助理1002可以充当该会话中的参与者。助理1002可以监控该会话,并且使用此处描述的一个或多个技术和方法答复个人或所述群组,以进行一对一交互。

[0103] 在多个实施例中,用于实施本发明技术的功能可以分布在任意数量的客户端和/或服务器组件之中。例如,不同软件模块可以被实施用于执行与本发明有关的不同功能,并且这些模块可以被不同地实施,从而运行在服务器/客户端组件上。用于这种配置的进一步细节在申请日为2011年1月10日的、律师案卷号为P10575US1、题为“Intelligent

Automated Assistant”的相关的美国专利申请序列号12/987,982中被描述,该申请的全部公开内容通过引用包含于此。

[0104] 在图32的例子中,输入引导功能和输出处理功能分布于客户端1304和服务器1340之中,输入引导的客户端部分2794a和输出处理的客户端部分2792a位于客户端1304上,而输入引导的服务器部分2794b和输出处理的服务器部分2792b位于服务器1340上。以下组件位于服务器1340上:

[0105] ●完整词汇2758b;

[0106] ●语言模式识别器的完全库 (libaray) 2760b;

[0107] ●短期个人存储器的主版本 (master version) 2752b;

[0108] ●长期个人存储器的主版本2754b。

[0109] 在一实施例中,客户端1304在本地保持这些组件的子集和/或部分,以提高响应度并降低对网络通信的依赖。可以根据已知的高速缓存管理技术来保持和更新这些子集和/或部分。这些子集和/或部分包括,例如:

[0110] ●词汇子集2758a;

[0111] ●语言模式识别器库的子集2760a;

[0112] ●短期个人存储器的高速缓存2752a;

[0113] ●长期个人存储器的高速缓存2754a;

[0114] 另外的组件可以被实施为服务器1340的一部分,包括例如:

[0115] ●语言解释器2770;

[0116] ●对话流处理器2780;

[0117] ●输出处理器2790;

[0118] ●域 (domain) 实体数据库2772;

[0119] ●任务流模型2786;

[0120] ●服务编排 (orchestration) 2782;

[0121] ●服务性能模型2788。

[0122] 这些组件的每一个将在以下更详细地描述。当需要时服务器1340通过与外部服务1360接口获得附加信息。

[0123] 概念架构

[0124] 现在参考图27,示出了虚拟助理1002的特定示例实施例的简化框图。如以上引入的相关的美国专利申请中更详细地描述的,虚拟助理1002的不同实施例可以被配置、设计和/或操作成提供与虚拟助理技术总体相关的各种不同类型的操作、功能和/或特征。此外,如此处更详细地描述的,在此公开的虚拟助理1002的多种不同类型的操作、功能和/或特征可以为与虚拟助理1002接口的不同实体使能或提供不同类型的优点和/或益处。可以使用以上描述的任意的硬件架构或使用不同类型的硬件架构实施图27中示出的实施例。

[0125] 例如,根据不同的实施例,虚拟助理1002可以被配置、设计和/或操作来提供多种不同类型的操作、功能和/或特征,例如,以下的一个或多个(或其组合):

[0126] ●自动化经由因特网可获得的数据和服务的应用来发现、寻找、从中挑选、购买、保留或订购产品和服务。除自动化使用这些数据和服务的处理之外,虚拟助理1002也可以立即使能数据和服务的一些源的组合使用。例如,其可以组合来自一些评论站点的关于产

品的信息、检查来自多个批发商的价格和可获得性,并且检查它们的位置和时间约束,以及帮助用户发现对他们的问题的个性化解决方案。

[0127] ●自动化经由因特网可获得的数据和服务的使用来发现、调查、从中挑选、保留和另外地了解要做的东西(包括但不限于电影、事件、表演、展览、展示和景点);要去的地方(包括但不限于旅行目的地、旅馆和其它要停留的地方、地标和其它感兴趣的地方等);要去吃或去喝的地方(诸如饭店和酒吧),要约会的时间和地点,以及在因特网上可以找到的任意其它娱乐或社交源。

[0128] ●经由自然语言对话使能应用和服务的操作,所述操作另外地通过具有图形用户界面的专用应用提供,包括搜索(包括基于位置的搜索);导航(地图和方向);数据库查询(诸如通过名字或其它属性寻找企业或人);获得天气情况和预报,检查市场商品的价格或金融交易的状态;监视航班的流量或状态;访问和更新日历和日程表;管理提醒、闹钟、任务和项目;经由电子邮件或其它消息平台通信;以及本地或远程操作设备(例如,拨打电话、控制灯光和温度、控制家用安全设备、播放音乐或视频等)。在一实施例中,虚拟助理1002可以被用来启动、操作和控制在该设备上的多个可用的功能和应用。

[0129] ●为活动、产品、服务、娱乐源、时间管理或任意其它类型的推荐服务提供个人推荐,所述推荐服务受益于采用自然语言的交互对话和对数据和服务的自动化访问。

[0130] 根据不同实施例,由虚拟助理1002提供的各种类型的功能、操作、动作和/或其它特征中的至少一部分可以实施在一个或多个客户端系统、一个或多个服务器系统和/或其组合处。

[0131] 根据不同实施例,由虚拟助理1002提供的各种类型的功能、操作、动作和/或其它特征中的至少一部分可以在解释及操作用户输入中使用情境信息,如此处更详细地描述的。

[0132] 例如,在至少一个实施例中,当执行特定任务和/或操作时,虚拟助理1002可操作成使用和/或生成各种不同类型的数据和/或其它类型的信息。这可以包括,例如,输入数据/信息和/或输出数据/信息。例如,在至少一实施例中,虚拟助理1002可操作成访问、处理和/或另外地使用来自一个或多个不同类型源的信息,所述不同类型源例如一个或多个本地和/或远程存储器、设备和/或系统。另外,在至少一个实施例中,虚拟助理1002可操作成生成一个或多个不同类型的输出数据/信息,所述输出数据/信息例如可以被存储在一个或多个本地和/或远程设备和/或系统的存储器中。

[0133] 可以由虚拟助理1002访问和/或使用的不同类型的输入数据/信息的例子可以包括,但不限于,以下的一个或多个(或其组合):

[0134] ●语音输入:来自移动设备,诸如移动电话和平板计算机、具有麦克风的计算机、蓝牙耳机、车辆语音控制系统,经电话系统、应答服务上的记录、集成消息服务上的音频语音邮件、具有语音输入(诸如时钟收音机)的消费者应用、电话站、家用娱乐控制系统及游戏控制台。

[0135] ●来自计算机或移动设备上的键盘、遥控器或其它消费者电子设备上的键板的文本输入,发送给该助理的电子邮件消息,发送给该助理的即时消息或类似短消息,从多用户游戏环境的玩家接收的文本以及消息馈送中的文本流。

[0136] ●来自传感器或基于位置的系统的位置信息。例子包括移动电话上的全球定位系

统 (GPS) 和辅助GPS (A-GPS)。在一实施例中,位置信息与明确的用户输入相结合。在一实施例中,基于已知的地址信息和当前位置确定,本发明系统能够检测何时用户在家。如此,相对于不在家,可以作出关于在家时该用户可能感兴趣的信息类型的某些推断,以及根据他或她是否在家来推断应该代表该用户调用的服务和动作的类型。

[0137] ●来自客户端设备上的时钟的时间信息。这可以包括,例如,来自电话或其它客户端设备的指示本地时间和时区的时间。此外,时间可以被使用在用户请求的情境中,例如,解释诸如“一小时内”和“今晚”的短语。

[0138] ●指南针、加速计、陀螺仪和/或行进速度数据,以及来自移动或手持设备或嵌入式系统 (诸如车辆控制系统) 的其它传感器数据。这也包括从遥控器到设备和游戏控制台的设备定位数据。

[0139] ●来自具有图形用户界面GUI的任意设备上的GUI的点击和菜单选择以及其它事件。进一步例子包括触摸触摸屏。

[0140] ●来自传感器和其它数据驱动触发器的事件,诸如闹钟、日历提醒、价格变化触发器、位置触发器、来自服务器的推送到设备上的通知等。

[0141] 对于此处描述的实施例,输入也包括该用户交互历史的情境,包括对话和请求历史。

[0142] 如以上交叉引用的相关美国专利申请中所描述的,多个不同类型输出数据/信息可以由虚拟助理1002生成。这些输出数据/信息可以包括,但不限于,以下的一个或多个 (或其组合):

[0143] ●直接发送给输出设备和/或设备的用户接口的文本输出;

[0144] ●经电子邮件发送给用户的文本和图形;

[0145] ●经消息服务发送给用户的文本和图形;

[0146] ●语音输出,其可以包括以下的一个或多个 (或其组合):

[0147] ○合成语音;

[0148] ○采样语音;

[0149] ○记录消息;

[0150] ●具有图片、富文本、视频、声音和超链接的信息的图形布局 (例如,网络浏览器中呈现的内容);

[0151] ●执行器 (actuator) 输出,用以控制设备上的物理动作,例如促使其开启或关闭、发声、变更颜色、振动、控制灯光等;

[0152] ●调用设备上的其它应用,诸如调用映射应用、语音拨叫电话、发送电子邮件或者即时消息、播放媒体、在日历中制作条目、任务管理器和便签应用及其它应用;

[0153] ●执行器输出,用以控制由设备控制或附着到设备的设备的物理动作,诸如操作远程照相机、控制轮椅、在远程扬声器上播放音乐、在远程显示器上播放视频等。

[0154] 应当认识到,图27的虚拟助理1002仅仅是来自可被实施的广泛的虚拟助理系统实施例的一个例子。与例如在图27的虚拟助理系统实施例的例子中示出的组件/特征相比,该虚拟助理系统的其它实施例 (未示出) 可以包括附加的组件/特征、较少和/或不同的组件/特征。

[0155] 虚拟助理1002可以包括多个不同类型的组件、设备、模块、处理和系统等,其例如,

可以通过使用硬件和/或硬件和软件的组合来实施和/或实例化。例如,如图27的实施例的例子中所示出的,助理1002可以包括以下类型的系统、组件、设备、处理等的一个或多个(或其组合):

- [0156] ●一个或多个活动本体(active ontologies) 1050;
- [0157] ●主动输入引导组件2794(可以包括客户端部分2794a和服务器部分2794b);
- [0158] ●短期个人存储器组件2752(可以包括主版本2752b和高速缓存2752a);
- [0159] ●长期个人存储器组件2754(可以包括主版本2754b和高速缓存2754a;可以包括,例如,个人数据库1058、应用偏好和使用历史1072等);
- [0160] ●域模型组件2756;
- [0161] ●词汇组件2758(可以包括完整词汇2758b和子集2758a);
- [0162] ●语言模式识别器组件2760(可以包括完全库2760b和子集2760a);
- [0163] ●语言解释器组件2770;
- [0164] ●域实体数据库2772;
- [0165] ●对话流处理器组件2780;
- [0166] ●服务编排组件2782;
- [0167] ●服务组件2784;
- [0168] ●任务流模型组件2786;
- [0169] ●对话流模型组件2787;
- [0170] ●服务模型组件2788;
- [0171] ●输出处理器组件2790。

[0172] 在某些基于客户端/服务器的实施例中,这些组件的一些或者全部可以分布于客户端1304和服务器1340之间。

[0173] 在一实施例中,虚拟助理1002经由任意适合的输入形式接收用户输入2704,包括例如触摸屏输入、键盘输入、语音输入和/或其任意组合。在一实施例中,助理1002还接收情境信息1000,其可以包括如在此更详细描述的事件情境2706和/或任意的其它类型的情境。

[0174] 在处理根据此处描述的技术的用户输入2704和情境信息1000时,虚拟助理1002生成用于呈现给用户的输出2708。可以根据任意适当的输出形式生成输出2708,在合适的情况下,其可以由情境1000及其它因素通知。输出形式的例子包括呈现在显示器上的视觉输出、听觉输出(其可以包括语音输出和/或蜂鸣声和其它声音)、触觉输出(诸如振动)和/或其任意组合。

[0175] 涉及图27中描述的不同组件的操作的附加细节在申请日为2011年1月10日的、律师案卷号为P10575US1、题为“Intelligent Automated Assistant”的相关的美国专利申请序列号12/987,982中被描述,该申请的全部公开内容通过引用包含于此。

[0176] 情境

[0177] 如以上所描述的,在一实施例中虚拟助理1002获取并且应用各种各样情境信息来执行信息处理功能。以下描述阐明:

- [0178] ●用于由虚拟助理1002使用的情境信息的源的范围;
- [0179] ●用于表示、组织和搜索情境信息的技术;
- [0180] ●情境信息可以支持虚拟助理的多个功能的操作的方法;和

[0181] ●用于在分布式系统中有效获取、访问和应用情境信息的方法。

[0182] 本领域技术人员将认识到以下对使用情境信息的源、技术和方法的描述仅仅是示例性的,并且其它源、技术和方法可以被使用而不脱离本发明的本质特征。

[0183] 情境源

[0184] 在由虚拟助理1002执行的信息处理的各个阶段中,可以使用一些不同类型的情境以减少用户输入的可能解释。例子包括应用情境、个人数据情境和在先对话历史。本领域技术人员将认识到其它的情境的其它源也是可用的。

[0185] 现在参考图1,示出了描绘根据一实施例的虚拟助理1002和能够影响其操作的情境的源的一些例子的框图。虚拟助理1002获得用户输入2704,诸如语音的或打字语言,处理所述输入,并且向用户生成输出2708和/或代表用户执行2710动作。应认识到,如图1中描绘的虚拟助理1002仅仅是来自可以被实施的广泛的虚拟助理系统实施例中的一个例子。相比于例如在图1中描绘的示例虚拟助理1002的组件/特征,虚拟助理系统的其它实施例(未示出)可以包括附加的组件/特征,较少和/或不同的组件/特征。

[0186] 如此处更详细描述,虚拟助理1002可以利用知识和数据的任意不同源,诸如字典、域模型和/或任务模型。从本发明的观点来看,这些源(被称为背景源)对于助理1002是内在的。除了用户输入2704和背景源之外,虚拟助理1002也可以利用来自一些情境源的信息,包括例如设备传感器数据1056、应用偏好和使用历史1072、对话历史和助理存储器1052、个人数据库1058、个人声学情境数据1080、当前应用情境1060和事件情境2706。这些将在此详细描述。

[0187] 应用情境1060

[0188] 应用情境1060指的是用户正在做某件事的应用或者类似软件状态。例如,用户可能正在使用文本消息应用与特定个人交谈。虚拟助理1002不需要专用于该文本消息应用的用户接口或其部分。相反,虚拟助理1002可以从任意数量的应用接收情境,其中每一个应用贡献其情境来通知虚拟助理1002。

[0189] 当虚拟助理1002被调用时,如果用户当前正使用应用,该应用的状态可以提供有用的情境信息。例如,如果虚拟助理1002从电子邮件应用中被调用,情境信息可以包括发送者信息、接收者信息、发送日期和/或时间、主题、从电子邮件内容提取的数据、邮箱或者文件夹名称等。

[0190] 现在参考图11至13,示出了描绘根据一实施例在文本消息域中使用应用情境以得到代词所指对象的例子的一组截图。图11示出了当用户使用文本消息应用时可以被显示的屏幕1150。图12示出了在文本消息应用中虚拟助理1002已经被激活后的屏幕1250。在该例子中,虚拟助理1002呈现提示1251给用户。在一实施例中,该用户可以通过在麦克风图标1252上进行敲击提供语音输入。在另一实施例中,助理1002能够在任意时间接受语音输入,并且在提供输入之前不需要用户在麦克风图标1252上进行敲击;因此,图标1252可以是助理1002正在等待语音输入的提示物。

[0191] 在图13中,该用户已经参与和虚拟助理1002的对话,如屏幕1253上示出的。该用户的语音输入“call him(呼叫他)”已经被回应,并且虚拟助理1002正在响应,该响应将以特定电话号码呼叫特定的人。为了解释该用户的模糊输入,虚拟助理1002使用多个情境源的组合来得到代词的所指对象,如此处更详细地描述的。

[0192] 现在参考图17至20,示出了根据一实施例的使用当前应用情境以解释和实施命令的另一例子。

[0193] 在图17中,呈现给用户他或她的电子邮件收件箱1750,并且用户选择特定电子邮件消息1751以查看。图18示出了在已经被选择用于查看之后的电子邮件消息1751;在该例子中,电子邮件消息1751包括图像。

[0194] 在图19中,当查看电子邮件应用内的电子邮件消息1751时,用户已经激活了虚拟助理1002。在一实施例中,电子邮件消息1751的显示在该屏幕上向上移动以为来自虚拟助理1002的提示150腾出位置。该显示加强了虚拟助理1002在当前已查看的电子邮件消息1751的情境中提供帮助的意向(notion)。相应地,到虚拟助理1002的该用户的输入将在电子邮件消息1751正被查看的当前情境中被解释。

[0195] 在图20中,该用户已经提供了命令2050:“Reply let's get this to marketing right away (答复:我们立刻将这带入市场)”。情境信息,包括关于电子邮件消息1751和其中显示该电子邮件消息的所述电子邮件应用的信息,用于解释命令2050。可以使用该情境来确定命令2050中的词“reply (答复)”和“this (这)”的含义,并且解决如何在特定消息线程上针对特定接收者建立电子邮件撰写事务。在这种情况下,虚拟助理1002能够访问情境信息,从而确定“marketing (市场)”指示名字为Jone Applecore的接收者,并能够确定用于接收者的电子邮件地址。相应地,虚拟助理1002为用户撰写电子邮件2052,以得到认可并且发送。如此,虚拟助理1002能够基于用户输入连同描述当前应用的状态的情境信息实施任务(撰写电子邮件消息)。

[0196] 应用情境也可以帮助跨应用识别用户意图的含义。现在参考图21,示出了其中在查看电子邮件消息(诸如电子邮件消息1751)的情境中用户调用虚拟助理1002的例子,但是用户的命令2150是“Send him a text (发送文本给他)…”。命令2150被虚拟助理1002解释为指示文本消息,而不是电子邮件,应当被发送。然而,使用该词“him (他)”指示希望同一接收者(Jones Applecore)。虚拟助理1002因此认识到应当与该接收者进行通信,但是通过不同的通道上(文本消息发送给该人的电话号码,该电话号码从存储在设备上的联系人信息获得)。相应地,虚拟助理1002为用户撰写文本消息2152,以获得认可并且发送。

[0197] 可以从应用获得的情境信息的例子包括,但不限于:

[0198] ●应用的身份;

[0199] ●当前对象或者在应用中正在操作的对象,诸如当前电子邮件消息、当前歌曲或播放列表或正在播放的通道、当前书籍或电影或照片、当前日历日/星期/月份、当前提示列表、当前电话呼叫、当前文本消息会话、当前地图位置、当前网页或搜索查询、当前城市或用于位置敏感的应用的其它位置、当前社交网络分布或当前对象的任意的其它特定应用概念;

[0200] ●名称、地点、日期和能从当前对象提取的其它可识别的实体或值。

[0201] 个人数据库1058

[0202] 情境数据的另一源是在设备(诸如电话)上的该用户的个人数据库1058,例如包含名称和电话号码的地址簿。现在参考图14,示出了根据一实施例的其中虚拟助理1002正在进行提示以消除姓名歧义的截图1451的例子。这里,该用户已经说“Call Herb (呼叫 Herb)”;虚拟助理1002提示用户在该用户的地址簿中的匹配联系人中进行选择。因此,该地

址簿被用作个人数据情境的源。

[0203] 在一实施例中,从个人数据库1058获得该用户的个人信息,以用作情境,来解释和/或实施用户的意图或者虚拟助理1002的其它功能。例如,当该用户仅通过名字指向某人时,可以使用用户联系人数据库中的数据来减少解释用户命令时的歧义。可以从个人数据库1058获得的情境信息的例子包括,但不限于:

[0204] ●用户联系人数据库(地址簿)-包括关于名称、电话号码、物理地址、网络地址、账户标识、重要日期-关于人、公司、组织、地点、网址和该用户可能指向的其它实体的信息;

[0205] ●用户自己的姓名、优选发音、地址、电话号码等;

[0206] ●用户的已命名的关系,诸如母亲、父亲、姐妹、老板等。

[0207] ●用户的日历数据,包括日历事件、特定日期的名称或者该用户可能指向的任意其它已命名的条目;

[0208] ●用户的提示或任务列表,包括要做的事情的列表、回忆或得到该用户可能指向的东西;

[0209] ●歌曲名称、流派、播放列表、以及与用户音乐库有关的用户可能指向的其它数据;

[0210] ●人、地点、类别、标签、标志或在照片或视频或该用户媒体库中的其它媒体上的其它符号名称;

[0211] ●标题、作者、流派或在该用户个人库中的书籍或者其它作品中的其它符号名称。

[0212] 对话历史1052

[0213] 情境数据的另一源是该用户与虚拟助理1002的对话历史1052。这些历史包括,例如,对域、人物、地点等等的引用。现在参考图15,示出了根据一实施例虚拟助理1002使用对话情境来推断针对命令的地点的例子。在屏幕1551中,该用户首先询问“What's the time in New York(在纽约是什么时间)”;虚拟助理1002通过提供纽约市的当前时间做出响应1552。该用户接着询问“What's the weather(天气怎样)”。虚拟助理1002使用在先会话历史来推断针对天气查询的期望的位置是会话历史中提及的最后位置。因此其响应1553提供关于纽约市的天气信息。

[0214] 作为另一例子,如果该用户说“find camera shops near here(寻找这附近的照相馆)”并接着,在检查结果之后,说“How about in San Francisco(在旧金山怎么样?)”,助理可以使用对话情境来确认“How about(怎么样)”意指“做相同任务(寻找照相馆)”,以及“in San Francisco(在旧金山)”意指“将搜索地点从这里改变到旧金山”。虚拟助理1002也可以使用对话的在先细节,诸如提供给该用户的在先输出,作为情境。例如,如果虚拟助理100使用了旨在幽默的聪明的响应,例如“Sure thing,you're the boss(当然,你是老板)”,其可以记得它已经说了这话,并且可以避免在对话会话中重复该短语。

[0215] 来自对话历史和虚拟助理存储器的情境信息的例子包括,但不限于:

[0216] ●对话中提及的人;

[0217] ●对话中提及的地点和位置;

[0218] ●所关注的当前时限;

[0219] ●所关注的当前应用域,诸如电子邮件或者日历;

[0220] ●所关注的当前任务,诸如阅读电子邮件或创建日历条目;

- [0221] ●所关注的当前域对象,例如刚刚阅读的电子邮件消息或刚刚创建的日历条目;
- [0222] ●对话或事务流的当前状态,诸如问题是否被提出以及期望什么样的可能答案;
- [0223] ●用户请求的历史,诸如“good Italian restaurants (好的意大利餐馆)”;
- [0224] ●用户请求的结果的历史,诸如返回的餐馆集合;
- [0225] ●在对话中的该助理使用的短语的历史;
- [0226] ●由该用户告知该助理的事实,诸如“my mother is Rebecca Richards (我妈妈是Rebecca Richards)”和“I like that restaurant (我喜欢那家餐馆)”。

[0227] 现在参考图25A至25B,示出了描绘根据一实施例的使用在先对话情境例子的一系列截图。在图25A中,该用户已经输入对来自John的任意新电子邮件的请求2550。虚拟助理1002通过显示来自John的电子邮件消息2551进行响应。在图25B中,用户输入命令2552“Reply let's get this to marketing right away (答复:我们立刻将这带入市场)”。虚拟助理1002使用在先对话情境解释命令2552;特别地,该命令被解释为指示图25中显示的电子邮件消息2551。

[0228] 设备传感器数据1056

[0229] 在一实施例中,运行虚拟助理1002的物理设备可以具有一个或多个传感器。这样的传感器可以提供情境信息源。这样的信息的例子包括,但不限于:

- [0230] ●用户当前位置;
- [0231] ●该用户当前位置的当地时间;
- [0232] ●设备的位置、方位和移动;
- [0233] ●当前灯光级别、温度和其它环境测量结果;
- [0234] ●使用中的麦克风和照相机的属性;
- [0235] ●正被使用的当前网络和所连接的网络的签名(signatures),包括以太网、Wi-Fi和蓝牙。签名包括网络接入点的MAC地址、分配的IP地址、设备标识符(诸如蓝牙名称)、频率通道和无线网络的其它属性。
- [0236] 传感器可以是任意类型,包括例如:加速计、指南针、GPS单元、高度检测器、光传感器、温度计、气压计、时钟、网络接口、电池测试电路等。

[0237] 应用偏好和使用历史1072

[0238] 在一实施例中,描述用于多个应用的用户偏好和设置的信息,以及他或她的使用历史1072,被用作情境,以解释和/或实施该用户意图或虚拟助理1002的其它功能。这些偏好和历史1072的例子包括,但不限于:

- [0239] ●快捷方式、收藏、书签、朋友列表或关于人的用户数据的任意其它收集,公司、地址、电话号码、地点、网站、电子邮件消息或任意其它引用;
- [0240] ●设备上的最近呼叫;
- [0241] ●最近的文本消息会话,包括该会话的参与方;
- [0242] ●对地图或方向的最近请求;
- [0243] ●最近的网络搜索和URL;
- [0244] ●股票应用中所列出的股票;
- [0245] ●最近的歌曲或视频或其它播放的媒体;
- [0246] ●提醒应用上设置的闹钟的名称;

[0247] ●设备上的应用的名称或其它数字对象；

[0248] ●用户优选的语言或在用户位置使用的语言。

[0249] 现在参考图16,示出了根据一实施例的使用电话收藏列表作为情境源的例子。在屏幕1650中,收藏联系人1651的列表被示出。如果用户提供输入“呼叫John”,该收藏联系人1651的列表可以被用来确定“John”指代John Appleseed的移动电话号码,因为那个号码出现在该列表中。

[0250] 事件情境2706

[0251] 在一实施例中,虚拟助理1002能够使用与异步事件相关联的情境,该异步事件的发生与该用户和虚拟助理1002的交互独立。现在参考图22至24,示出了例示根据一实施例、在事件发生之后激活虚拟助理1002的例子,所述事件能够提供事件情境或提醒情境。在该情况下,该事件是到来的文本消息2250,如图22中示出的。在图23中,虚拟助理1002已经被调用,并且文本消息2250连同提示1251被示出。在图24中,该用户已经输入命令“call him (呼叫他)”2450。虚拟助理1002使用事件情境将“him(他)”解释为表示发送到来的文本消息2250的那个人,从而消除该命令的歧义。虚拟助理1002进一步使用该事件情境来确定用于该向外呼出使用的电话号码。确认消息2451被显示以指示该呼叫正在被拨打。

[0252] 提醒情境信息的例子包括,但不限于:

[0253] ●到来的文本消息或页面;

[0254] ●到来的电子邮件消息;

[0255] ●到来的电话呼叫;

[0256] ●提示通知或任务提醒;

[0257] ●日历提醒;

[0258] ●闹钟、定时器或其它基于时间的提醒;

[0259] ●比分(scores)通知或来自比赛的其它事件;

[0260] ●金融事件的通知,诸如股票价格提醒;

[0261] ●新闻快讯或其它广播通知;

[0262] ●来自任意应用的推送通知。

[0263] 个人声学情境数据1080

[0264] 当解释语音输入时,虚拟助理1002也可以考虑语音所进入的声学环境。例如,安静的办公室的噪声分布不同于车辆或公共场所的噪声分布。如果语音识别系统可以识别并存储声学分布数据,这些数据也可以作为情境信息提供。当结合诸如使用中的麦克风的属性的其它情境信息、当前位置和当前对话状态时,声学情境可以帮助输入的识别和解释。

[0265] 表示并访问情境

[0266] 如以上描述的,虚拟助理1002可以使用来自任意不同源中的情境信息。可以使用任意的不同机制中表示情境,使得其是虚拟助理1002可用的。现在参考图8a至8d,它们示出了可以与本发明多个实施例结合使用的情境信息的表示的例子。

[0267] 表示人、地点、时间、域、任务和对象

[0268] 图8a示出了情境变量的例子801-809,所述情境变量表示简单属性,如用户当前位置的地理坐标。在一实施例中,可以保持当前值作为一组核心情境变量。例如,可以是当前用户、所关注的当前位置、所关注的当前时限、所关注的当前应用域、所关注的当前任务和

所关注的当前域对象。数据结构 (诸如图8a中示出的) 可以被用于这样的表示。

[0269] 图8b示出了更复杂的表示的例子850,其可以用于存储用于联系人的情境信息。还示出了包括用于联系人的数据的表示的例子851。在一实施例中,联系人(或人)可以被表示为具有姓名、性别、地址、电话号码的属性和可能保存于联系人数据库中的其它属性的对象。类似的表示可以被用于地点、时间、应用域、任务、域对象等。

[0270] 在一实施例中,给定类型的当前值的集合被表示。这些集合指示当前人、当前地点、当前时间等。

[0271] 在一实施例中,情境值按历史布置,因此在第N次重复处,存在当前情境值的帧,并且还存在于第N-1次重复的当时情境值的帧,回到对所期望的历史长度的某些限制。图8c示出了包括情境值的历史的阵列811的例子。特别地,图8c的每一列表示情境变量,而行对应于不同的时间。

[0272] 在一实施例中,如图8d所示,键入的情境变量的集合按历史布置。在该例中,示出了指示人的情境变量的集合861,以及指示地点的情境变量的另一集合871。因此,针对历史中特定时间的相关的情境数据可以被检索并应用。

[0273] 本领域技术人员将认识到图8a至8d中示出的特定表示仅仅是示例性的,并且用于表示情境的多个其它机制和/或数据格式可以被使用。例子包括:

[0274] ●在一实施例中,该系统的当前用户可以一些特定方式被表示,从而该虚拟助理1002知道如何获得该用户地址并指示该用户的家、工作、移动电话等。

[0275] ●在一实施例中,人之间的关系可以被表示,允许虚拟助理1002理解诸如“我妈妈”或“我老板的房子”的引用。

[0276] ●地点可以被表示为具有诸如名称、街道地址、地理坐标等属性的对象。

[0277] ●时间可以被表示为具有包括世界时间、时区偏移(time zone offset)、分辨率(resolution) (诸如年、月、日、时、分或秒) 的属性的对象。时间对象也可以表示象征性时间,诸如“今天”、“这星期”、“这[即将来临的] 周末”、“下周”、“Annie的生日”等。时间对象也可以表示持续时间或时间点。

[0278] ●情境也可以根据应用域提供,该应用域表示服务或应用或话语(discourse)的域,诸如电子邮件、文本消息、电话、日历、联系人、照片、视频、地图、天气、提示、时钟、网络浏览器、Facebook、Pandora等等。当前域指示关注这些域中的哪一个。

[0279] ●情境也可以定义一个或多个任务或者操作以在域中执行。例如,在电子邮件域中存在诸如读取电子取邮件消息、搜索电子邮件、撰写新电子邮件等任务。

[0280] ●域对象是与各种域关联的数据对象。例如,电子邮件域在电子邮件消息上操作,日历域在日历时间上操作,等等。

[0281] 处于在此处提供描述的目的,情境信息的这些表示被称为给定类型的情境变量。例如,当前用户的表示是Person(个人)类型的情境变量。

[0282] 表示情境起源

[0283] 在一实施例中,情境变量的起源被明确地表示,因此其可以在信息处理中使用。情境信息的起源是该源和/或推理集合的表征,所述推理用于推断或者获取信息。例如,如图8b中所示的Person(个人)情境值851可以从文本消息域对象得到,该文本消息域对象从事件情境2760获得。情境值851的该源可以被表示。

[0284] 表示用户请求和/或意图的历史

[0285] 在一实施例中,用户请求的历史可以被存储。在一实施例中,用户意图(如从自然语言处理中得到的)的深层结构表示的历史也可以被存储。这允许虚拟助理1002在之前解释的输入的情境中理解新的输入。例如,如果该用户询问“what is the weather in New York(纽约的天气如何)?”,语言解释器2770可以参考纽约的位置解释该问题。如果该用户接着说“what is it for this weekeed(这周末其如何)?”,虚拟助理1002可以参考该在先解释,从而确定“what is it(其如何)”应被解释为意指“what is the weather(天气如何)”。

[0286] 表示结果的历史

[0287] 在一实施例中,用户请求的结果的历史可以以域对象的形式被存储。例如,用户请求“find me some good Italian restaurants(为我找一些好的意大利餐馆)”可能返回一组表示餐馆的域对象。如果该用户接着输入命令,诸如“call Amilio's(呼叫Amilio的)”,虚拟助理1002可以在所述结果进行搜索,以在所述搜索结果中寻找名称为Amilio的餐馆,与可以被呼叫的所有可能的地方相比,该搜索结果是较小的集合。

[0288] 情境变量的延迟绑定(Delayed Binding)

[0289] 在一实施例中,情境变量可以表示根据需要被获取或被得到的信息。例如,当被访问时,表示当前位置的情境变量可以调用API,所述API从设备获取当前位置数据并且接着做其它处理以计算例如街道地址。依赖于高速缓存策略,该情境变量的值可以被保持一段时间。

[0290] 搜索情境

[0291] 虚拟助理1002可以使用任意不同的方法来搜索相关的情境信息以解决信息处理问题。不同类型的搜索的例子包括,但不限于:

[0292] ●通过情境变量名称进行搜索。如果所请求的情境变量的名称是已知的,诸如“当前用户名字”,虚拟助理1002可以对其实例进行搜索。如果历史被保持,虚拟助理1002可以首先搜索当前值,并接着考虑先前的数据直到发现匹配。

[0293] ●通过情境变量类型进行搜索。如果所请求的情境变量类型是已知的,诸如Person(个人),虚拟助理1002可以对该类型的情境变量的实例进行搜索。如果历史被保持,虚拟助理1002可以首先搜索当前值,并接着考虑先前的数据直到发现匹配。

[0294] 在一实施例中,如果当前信息处理问题要求单一匹配,一旦匹配被发现则终止搜索。如果允许多个匹配,则可以依次获取多个匹配结果直到达到某种限制。

[0295] 在一实施例中,如果合适,虚拟助理1002可以将其搜索约束为具有特定起源的数据。例如,如果在用于电子邮件的任务流中查找People(人物)对象,虚拟助理1002可以仅考虑其起源是关联于那个域的应用的情境变量。

[0296] 在一实施例中,使用情境变量的任意可用的属性,虚拟助理1002根据启发法使用规则来对匹配进行排序。例如当处理包括命令“tell her I'll be late(告诉她我会迟到)”的用户输入时,虚拟助理1002通过参考情境来解释“her(她)”。通过这么做,虚拟助理1002可以应用排序,从而指示用于People(人物)对象的优选,所述对象的起源是针对诸如文本消息和电子邮件的通信应用的应用使用历史。如另一例子,解释命令“call her(呼叫她)”时,虚拟助理1002可以应用排序,从而相对于其电话号码不被知道的人员对象倾向于

具有电话号码的People(人物)对象。在一实施例中,排序规则可以与域关联。例如,可以使用不同的排序规则来对用于电子邮件和电话域的个人变量进行排序。依赖于特定的表示和对所需的情境信息的访问,本领域技术人员将认识到可以创建和/或应用任意这样的排序规则。

[0297] 使用情境改善虚拟助理处理

[0298] 如以上所描述的,情境可以被应用于与虚拟助理1002的操作关联的各种各样的计算和推断。现在参考图2,其示出了描绘根据一实施例的用于在虚拟助理1002中的在处理的的不同阶段使用情境的方法10。

[0299] 方法10可以结合虚拟助理1002的一个或多个实施例来实施。

[0300] 在至少一实施例中,方法10可以操作成执行/或实施不同类型的功能、操作、动作和/或其它特征,例如以下的一个或多个(或其组合):

[0301] ●执行用户和虚拟助理1002之间的会话接口的接口控制流循环。方法10的至少一次迭代可以作为该会话中的一层(ply)。会话接口是用户和助理1002以会话方式通过进行来回表达交流的接口。

[0302] ●提供用于虚拟助理1002的执行控制流。即,该过程控制输入的收集、输入的处理、输出的生成和对用户的输出表示。

[0303] ●协调虚拟助理1002的组件之间的通信。即,其可以指示一组组件的输出反馈到另一组组件中的位置,以及来自该环境的全部输入和在该环境上可能发生的动作的位置。

[0304] 在至少一些实施例中,方法10的一部分也可以被实施在计算机网络的其它设备和/或系统处。

[0305] 根据特定实施例,方法10的多个实例或线程可以通过使用一个或多个处理器63和/或硬件的其它组合和/或硬件和软件的其它组合,并发实施和/或启动。在至少一实施例中,方法10的一个或多个或所选择的部分可以被实施在一个或多个客户端1304、在一个或多个服务器1340和/或其组合处。

[0306] 例如,在至少一些实施例中,可以通过软件组件、网络服务、数据库等,或其任意组合来执行、实施和/或启动方法10的不同方面、特征和/或功能。

[0307] 根据不同实施例,响应于对满足用于触发方法10的至少一个实例的启动的一个或多个不同类型的标准(例如,最小阈值标准)的一个或多个条件或事件的检测,可以启动方法10的一个或多个不同线程或实例。可以触发方法的一个或多个不同线程或实例的启动和/或实施的不同类型的条件或事件的例子可以包括但不限于以下的一个或多个(或其组合):

[0308] ●用户与虚拟助理1002的实例的会话,例如但不限于以下的一个或多个:

[0309] ○正在启动的移动设备应用,例如,实施虚拟助理1002的实施例的移动设备应用;

[0310] ○正在启动的计算机应用,例如,实施虚拟助理1002的实施例的应用;

[0311] ○被按压的在移动设备上专用按钮,诸如“语音输入按钮”;

[0312] ○附接到计算机或移动设备的外围设备(诸如耳机、电话听筒或基站、GPS导航系统、消费者设备、遥控器或具有可以与调用帮助关联的按钮的任意其它设备)上的按钮;

[0313] ○从网络浏览器开始到实施虚拟助理1002的网站的网络会话;

[0314] ○从现有的网络浏览器会话中开始到实施虚拟助理1002的网站的接口,在该网站

中,例如,虚拟助理1002服务被请求;

[0315] ○发送到形式服务器1426的电子邮件消息,该形式服务器1426协调与虚拟助理1002的实施例的通信;

[0316] ○发送到形式服务器1426的文本消息,该形式服务器1426协调与虚拟助理1002的实施例的通信;

[0317] ○拨打到形式服务器1434的电话,该形式服务器1434协调与虚拟助理1002的实施例的通信;

[0318] ○发送到应用的事件(诸如提醒或通知),该应用提供虚拟助理1002的实施例。

[0319] ●当提供虚拟助理1002的设备被打开和/或启动时。

[0320] 根据不同实施例,可以手动地、自动地、静态地、动态地、同时地和/或以上述各项组合的方式启动和/或实施方法10的一个或多个不同线程或实例。此外,可以以一个或多个不同的时间间隔(例如,在特定时间间隔期间、以固定的周期间隔、以不规则的周期间隔、根据需求等)启动方法10的不同实例和/或实施例。

[0321] 在至少一实施例中,当执行特定任务和/或操作,方法10的给定实例可以使用和/或生成多种不同类型的数据和/或其它类型的信息,包括如此处描述的情境数据。数据还可以包括任意其它类型的输入数据/信息和/或输出数据/信息。例如,在至少一实施例中,方法10的至少一实例可以访问、处理和/或以其它方式使用来自一个或多个不同类型的源(例如,一个或多个数据库)的信息。在至少一实施例中,可以经由与一个或多个本地和/或远程存储器设备的通信访问数据库信息的至少一部分。另外,方法10的至少一实例可以生成一个或多个不同类型的输出数据/信息,其例如可以被存储在本地存储器和/或远程存储器设备中。

[0322] 在至少一实施例中,可以使用一个或多个不同类型的初始化参数执行方法10的给定实例的初始配置。在至少一实施例中,可以经由与一个或多个本地和/或远程存储器设备的通信访问初始化参数的至少一部分。在至少一实施例中,提供给方法10的实例的初始化参数的至少一部分可以对应于输入数据/信息和/或从输入数据/信息得到。

[0323] 在图2的特定示例中,假设单个用户正通过网络从具有语音输入能力的客户端应用访问虚拟助理1002的实例。

[0324] 语音输入被引导并且解释100。引导可以包括以任意合适的模式呈现提示。在多个实施例中,该客户端的用户接口提供多个输入模式。这些可以包括,例如:

[0325] ●用于打字输入的接口,其可以调用主动的打字输入引导过程;

[0326] ●用于语音输入的接口,其可以调用主动的语音输入引导过程;

[0327] ●用于从菜单选择输入的接口,其可以调用主动的基于GUI的输入引导。

[0328] 在以上引用的相关专利申请中描述了用于执行这些模式中的每一个的技术。本领域技术人员将认识到可以提供其它输入模式。步骤100的输出是输入语音的一组候选解释190。

[0329] 该组候选解释190由语言解释器2770(也称为自然语言处理器,或NLP)处理200,语言解释器2770解析所述文本输入并生成一组用户意图290的可能解释。

[0330] 在步骤300中,用户意图290的表示被传送给对话流处理器2780,对话流处理器2780实施如结合图5描述的对话和流分析过程的实施例。会话流程处理器2780确定意图的

哪一种解释是最有可能的,将该解释映射成域模型的实例和任务模型的参数,并且确定任务流中的下一流程步骤。

[0331] 在步骤400中,执行所识别的流程步骤。在一实施例中,通过服务编排组件2782执行该流程步骤的调用,该服务编排组件2782代表该用户的请求调用一组服务。在一实施例中,这些服务为共同结果贡献一些数据。

[0332] 在步骤500中生成对话响应。在步骤700中,该响应被发送给客户端设备以在其上输出。设备上的客户端软件将其呈现在该客户端设备的屏幕(或其它输出设备)上。

[0333] 如果在查看该响应之后,该用户完成790,则该方法结束。如果用户没有完成,通过返回到步骤100而开始循环的另一次迭代。

[0334] 可以在方法10中的不同点通过该系统的不同组件使用情境信息1000。例如,如图2中所描述的,可以在步骤100、200、300和500使用情境1000。以下提供对在这些步骤中使用情境1000的进一步描述。然而,本领域技术人员将认识到,使用情境信息并不限于这些特定步骤,并且系统也可以在其它点使用情境信息,而不脱离本发明的本质特征。

[0335] 另外,本领域技术人员将认识到方法10的不同实施例可以包括不同于图2中所示出的特定实施例中示出的附加的特征和/或操作,和/或可以省略如在图2的特定实施例中示出的方法10的至少一部分特征和/或操作。

[0336] 在语音引导和解释中使用情境

[0337] 现在参考图3,示出了描绘根据一实施例的用于在语音推导和解释100中使用情境以改善语音识别的方法的流程图。可以使用情境1000,例如,用于在语音识别中去除歧义来引导候选假设的生成、排序和过滤,该候选假设把音素和词匹配。不同的语音识别系统使用生成、排序和过滤的不同混合,但是情境1000通常可以在任意阶段应用以减少假设空间。

[0338] 方法开始100。助理1002接收121听觉信号形式的话音或语音输入。语音到文本服务122或者处理器生成一组听觉信号的候选文本解释124。在一实施例中,使用例如可从马萨诸塞州的伯灵顿的Nuance通讯公司(Nuance Communications, Inc. of Burlington, Massachusetts)获得的Nuance识别器实施语音到文本服务122。

[0339] 在一实施例中,助理1002利用统计语言模型1029来生成语音输入121的候选文本解释124。在一实施例中,应用情境1000来偏置由语音到文本服务122生成的候选解释124的生成、过滤和/或排序。例如:

[0340] ● 语音到文本服务122可以使用来自用户个人数据库1058的词汇来偏置统计语言模型1029。

[0341] ● 语音到文本服务122可以使用对话状态情境来选择定制统计语言模型1029。例如,当询问是/否的问题时,统计语言模型1029可以选择偏向听到这些词。

[0342] ● 语音到文本服务122可以使用当前应用情境来偏向相关词。例如在文本消息应用情境中,可以相对于“collar”优先选择“call her (呼叫她)”,因为这样的情境提供可能被呼叫的Person(个人)对象。

[0343] 例如,给定的语音输入可能引导语音到文本服务122来生成解释“call her (呼叫她)”和“collar”。通过统计语言模型(SLM) 1029的引导,可以在其听到“call (呼叫)”后通过语法约束将语音到文本服务122调整成听到姓名。也可以基于情境1000调整语音到文本服务122。例如,如果“Herb”是该用户地址簿中的名字,则可以使用该情境以降低用于考虑

“Herb”作为该第二音节的解释的阈值。即,该用户个人数据情境中的姓名的存在会影响统计学语言模型1029用以生成假设的选择和调整。该姓名“Herb”可以是常规SLM1029的部分或者其可以由情境1000直接添加。在一实施例中,其可以被添加为附加的SLM1029,附加的SLM1029基于情境1000被调整。在一实施例中,其可以是现有的SLM1029的调整,其中现有的SLM1029基于情境1000被调整。

[0344] 在一实施例中,统计语言模型1029也被调整以从可以被存储在长期个人存储器2754中的应用偏好和使用历史1072和/或个人数据库1058中查找词、名称和短语。例如,可以将联系人/地址簿内的to-do项、列表项、个人便签、日历条目、人名,联系/地址簿内提及的电子邮件地址、街道或城市名称等中的文本提供给统计语言模型1029。

[0345] 根据候选解释适合虚拟助理1002的句法的、和/或语义的模型的程度,排序组件对候选解释进行分析124和排序126。可以使用关于用户输入的约束的任意源。例如,在一实施例中,根据在句法和/或语义含义、域模型、任务流模型、和/或会话模型等中解析该解释的程度,助理1002可以对该语音到文本解释器的输出进行排序:其评估在候选解释124中的词的各种组合与活动本体(active ontology)及其关联的模型的概念、关系、实体和属性匹配的程度,如以上引用的相关美国专利申请中所描述的。

[0346] 候选解释的排序126也受情境1000影响。例如,当虚拟助理1002被调用时,如果用户当前正在文本消息应用中进行会话,与所述词“collar”相比较,所述短语“call her(呼叫她)”更可能是正确的解释,因为存在可能在情境中要呼叫的“her(她)”。在当前的应用情境指示应用可以提供“可呼叫实体”时,可以通过将假设的排序126调整为偏爱短语,诸如“call her(呼叫她)”或“call<contact name>(呼叫<联系人名称>)”,来实现这样的偏置。

[0347] 在多个实施例中,助理1002用于解释文本输入的算法或过程,包括图3中示出的自然语言处理过程的任意实施例,可以被用于对由语音到文本服务122所生成的候选文本解释124进行排序和评分。

[0348] 作为约束候选解释的生成或影响候选解释的排序的代替或附加地,也可以使用情境1000来过滤候选解释124。例如,过滤规则可以规定用于“Herb”的地址簿条目的情境是以指示包含其的短语应被考虑为最高候选130,即使其不这样将在过滤阈值之下也是如此。依赖于被使用的特定语音识别技术,在生成、排序和/或过滤阶段可以应用基于情境偏置的约束。

[0349] 在一实施例中,如果排序组件126确定128来自解释124的最高排序语音解释排在特定阈值之上,则可以自动地选择130最高排序解释。如果没有解释排在特定阈值之上,向用户呈现132语音134的可能的候选解释。用户接着可以在所显示的选择中进行选择136。

[0350] 现在也参考图26A和26B,示出了描绘根据一实施例的用于在候选解释中选择的用户接口的例子的截图。图26A示出了在模糊解释2651下面具有圆点的该用户语音的表示。如果该用户在文本上轻敲,其示出如图26B中所示的可替代解释2652A、2652B。在一实施例中,情境1000会影响候选解释2652A、2652B中的哪一个是优选的解释(其如图26A中被示为初始默认值)以及对如在图26B中用以呈现的一组有限的备选的选择。

[0351] 在多个实施例中,所显示的选择中的用户选择136可以通过任意模式的输入(包括例如多种方式输入)实现。这些输入模式包括,但不限于,主动引导打字输入、主动引导语音输入、主动呈现用于输入的GUI等。在一实施例中,用户可以在候选解释134中进行选择,例

如通过敲击或说话。在说话的情况下,新的语音输入的可能解释受到所提供的一小组选择134的强约束。

[0352] 不管自动选择130或由该用户选择136输入,返回所产生的一个或多个文本解释190。在至少一实施例中,返回的输入是有注释的,所以关于在步骤136中选定的哪个选择的信息连同所述情境输入一起被保存。这使得当字符串返回时,例如字符串之下的语义的概念或实体关联于所述字符串,其改善了随后的语言解释的精确度。

[0353] 结合图1所描述的源中的任意一个可以将情境1000提供给图3中所示的语音引导和解释方法。例如:

[0354] ●个人声学情境数据1080,用于从可能的SLM1029中选择或以其它方式调整它们以优化所识别的声学情境。

[0355] ●设备传感器数据1056,描述使用中的麦克风和/或照相机的属性,可以被用以从可能的SLM1029中选择或以其它方式调整它们以优化所识别的声学情境。

[0356] ●来自个人数据库1058及应用偏好和使用历史1072的词汇可以被用作情境1000。例如,媒体的标题和艺术家的名称可以被用于调整语言模型1029。

[0357] ●通过文本语音服务122可以使用当前对话状态、对话历史和助理存储器1052的一部分来偏置候选解释124的生成/过滤/排序。例如,一类会话状态为询问是/否的问题。当处于这样的状态,过程100可以选择偏向听到这些词的SLM1029,或其可以在122在特定情境调整中偏置这些词的排序和过滤。

[0358] 在自然语言处理中的情境的使用

[0359] 可以使用情境1000来促进自然语言处理(NLP)-将文本输入解析为语义结构,该语义结构表示可能的解析。现在参考图4,示出了描绘根据一实施例在可以由语言解释器2770执行的自然语言处理中使用情境的方法的流程图。

[0360] 方法开始200。接收输入文本202。在一实施例中,使用模式识别器2760、词汇数据库2758、本体和其它模型1050将输入文本202与词和短语进行匹配,以便识别用户输入和概念之间的关联。步骤210生成一组候选句法解析212,其与语义相关性220匹配,生成候选语义解析222。在230,候选解析接着被处理以移除模糊的备选,通过相关性232进行过滤并排序,并且返回。

[0361] 在自然语言处理的整个过程中,可以应用情境信息1000来减小假设空间并且约束可能的解析。例如,如果语言解释器2770接收到两个候选“call her(呼叫她)”和“call Herb(呼叫Herb)”,那么语言解释器2770会找到针对词“call(呼叫)”、“her(她)”和“Herb”的绑定212。可以使用应用情境1060,以将对于“call(呼叫)”的可能的词含义约束为意指“phone call(电话呼叫)”。也可以使用情境来寻找针对“her(她)”和“Herb”的指示。对于“her(她)”,为了得到可呼叫实体源可以搜索情境源1000。在该例子中,文本消息会话中的当事人是可呼叫的实体,并且该信息是来自文本消息应用的情境的一部分。在“Herb”的情况下,用户的地址簿是去歧义情境源,还有其它个人数据,诸如应用偏好(诸如来自域实体数据库2772的收藏号码)和应用使用历史(诸如来自域实体数据库2772的最近电话呼叫)的其它个人数据。在一例子中,当前文本消息当事人是RebeccaRichards并且在用户地址簿中存在HerbGowen,由语言解释器2770创建的两个解析将是表示“PhoneCall(RebeccaRichards)”和“PhoneCall(HerbGowen)”的语义结构。

[0362] 在生成候选句法解析212时,语言解释器2770也可以使用来自应用偏好和使用历史1072、对话历史和助理存储器1052和/或个人数据库1058的数据。可以例如从短期和/或长期存储器2752、2754获得这些数据。如此,在同一会话中之前提供的输入和/或关于用户的已知信息可以被使用,从而改善性能、减少歧义并增强交互的会话自然性。也可以使用来自主动本体1050、域模型2756和任务流模型2786的数据,以在确定有效候选句法解析212时实施证据推理。

[0363] 在语义匹配220中,语言解释器2770根据其匹配语义模型(诸如域模型和数据库)的程度考虑可能的解析结果的组合。语义匹配220可以使用来自例如主动本体1050、短期个人存储器2752和长期个人存储器2754的数据。例如,语义匹配220可以使用以下数据,该数据来自对在对话中的场所或本地事件(来自对话历史和助理存储器1052)或个人喜欢的场所(来自应用偏好和使用历史1072)的先前引用。语义匹配220步骤还使用情境1000以将短语解释为域意图的结构。生成222一组候选,或可能的语义的解析结果。

[0364] 在去歧义步骤230,语言解析器2770对候选语义解析结果222的证据强度进行加权。去歧义230包括通过去除不太可能的或多余的备选选项来减少候选语义解析222的数量。去歧义230可以使用来自例如主动本体1050的结构的数据。在至少一实施例中,主动本体中的节点之间的连接提供用于在候选语义解析结果222中去歧义的证据支持。在一实施例中,使用情境1000来帮助这样的去歧义。这样的去歧义的例子包括:确定具有相同名字的一些人员中的一个;确定诸如“reply (答复)”(电子邮件或文本消息)的命令的指示对象;代词去指示等。

[0365] 例如,诸如“call Herb (呼叫Herb)”的输入潜在地指示匹配“Herb”的任意实体。可能存在任意数量的这样的实体,不仅在用户地址簿(个人数据库1058)中,而且在来自个人数据库1058和/或域实体数据库2772的业务名称的数据库中。在步骤232中,一些情境的源可以约束匹配“Herb”的组,和/或排列和过滤它们。例如:

[0366] ●其它应用偏好和使用历史1072,诸如在收藏电话号码列表上的Herb,或者最近呼叫,或文本消息会话或电子邮件的线程的最近当事人;

[0367] ●个人数据库1058中提及的Herb,诸如被命名为例如父亲或兄弟关系,或在最近日历事件中所列出的参与者的Herb。如果任务是播放媒体而不是电话呼叫,则来自媒体标题、创建者等的名字将作为约束的源。

[0368] ●在请求或者在结果中的对话1052的最近层。例如,如以上结合图25A至25B描述的,在搜索来自John的电子邮件之后,使用仍在该对话情境中的搜索结果,该用户可以撰写答复。助理1002可以使用该对话情境来识别特定应用域对象情境。

[0369] 情境1000还可以帮助减少词中除了正确名称之外的歧义。例如,如果电子邮件应用的用户告知助理1002去“reply (答复)”(如图20中所描绘的),该应用的情境帮助确定应与电子邮件答复而不是文本消息答复有关的词。

[0370] 在步骤232中,语言解释器2770过滤并排序232上述最高语义解析作为用户意图290的表示。可以使用情境1000来通知这样的过滤和排序232。结果为用户意图290的表示。

[0371] 在任务流处理中的情境的使用

[0372] 现在参考图5,示出了描绘根据一实施例、用于在可以由对话流处理器2780执行的任务流处理中使用情境的方法的流程图。在任务流处理中,图4的方法生成的候选解析被排

序,并且被例示为生成可以被执行的操作任务描述。

[0373] 方法开始300。接收用户意图290的多个候选表示。如结合图4所描述的,在一实施例中,用户意图290的表示包括一组语义解析。

[0374] 在步骤312中,基于用户意图的确定,对话流处理器2780利用其它信息确定语义解析的优选解释以确定要执行的任务及其参数。例如,可以从域模型2756、任务流模型2786和/或对话流模型2787或其任意组合获得信息。例如,任务可以是电话呼叫,而任务参数是要呼叫的电话号码。

[0375] 在一实施例中,在执行步骤312中使用情境1000,以通过推断默认值和解决歧义来引导参数312的绑定。例如,情境1000可以引导该任务描述的例示并且确定是否存在用户意图的最佳解释。

[0376] 例如,假设所述意图输入290是“PhoneCall (RebeccaRichards) (电话呼叫 (RebeccaRichards))”和“PhoneCall (HerbGowen) (电话呼叫 (HerbGowen))”。该电话呼叫任务请求电话号码的参数。可以应用情境100的一些源以确定用于Rebecca和Herb的哪一电话号码有效。在该例子中,在联系人数据库中的用于Rebecca的地址簿条目具有两个电话号码,而用于Herb的条目没有电话号码只有一电子邮件地址。使用来自个人数据库1058 (诸如联系人数据库)的情境信息1000允许虚拟助理1002相对于Herb更偏向于Rebecca,因为存在用于Rebecca的电话号码并且没有用于Herb的电话号码。为了确定哪一个电话号码用于Rebecca,可以考虑应用情境1060以选择正在被使用以与Rebecca进行文本消息会话的号码。虚拟助理1002由此可以确定与Rebecca Richards的文本消息会话的情境中的“call her (呼叫她)”意指拨打电话给Rebecca正在用于文本消息的移动电话。在步骤390返回该特定信息。

[0377] 情境1000可以不只是用于减少电话号码的歧义。每当存在用于任务参数的多个可能值,只要具有用于该参数的值的情境1000的任意源可用,那么就可以使用情境1000。情境1000可以减少歧义 (并且避免必须提示用户在候选中进行选择) 的其它例子包括,但不限于:电子邮件地址;物理地址;时间和日期;地点;列表名称;媒体标题;艺术家名称;业务名称;或任意其它值空间。

[0378] 任务流处理300所需要的其它类型推断也可以受益于情境1000。例如,默认值推断可以使用当前位置、时间和其它当前值。默认值推断有益于确定在用户请求中不明确的任务参数的值。例如,如果有人说“天气如何?”,它们暗示周围的当前天气如何。

[0379] 在步骤310中,对话流处理器2780确定用户意图的这种解释是否被足够强地支持以继续下去,和/或是否其比备选的模糊的解析更好地被支持。如果存在竞争歧义或者足够的不确定性,接着执行步骤322,以设置对话流步骤,以便该执行阶段使对话输出提示,以获得来自用户的更多信息。用于提示用户以解决歧义的截图的例子在图14中示出。在步骤322中可以使用情境1000来排序以及注释所显示的候选项目菜单,以供用户从中进行选择。

[0380] 在步骤320中,考虑该任务流模型以确定适当的下一步骤。例如,可以从域模型2756、任务流模型2786和/或对话流模型2787或其任意组合中获得信息。

[0381] 步骤320或步骤322的结果是用户请求390的表示,其可以包括足以用于对话流处理器2780的任务参数和用来分派所述适当的的服务的服务编排2782。

[0382] 使用情境来改善对话生成

[0383] 在对话响应生成500期间,助理1002可以将其对该用户意图的理解及其如何在任务中执行解释回去。这样的输出的例子是“OK,I'll call Rebecca on her mobile...(好的,我将拨打Rebecca的手机)”。这允许用户授权助理1002执行所关联的任务自动化,诸如拨打呼叫。在对话生成步骤500中,助理1002确定在其对用户意图的理解的解释中有多少细节更传递回用户。

[0384] 在一实施例中,情境1000也可以用于引导在对话中对细节的适当级别的选择,以及基于先前的输出的过滤(为了避免重复信息)。例如,助理1002可以使用从情境1000中推断的个人和电话号码的知识,以确定是否提及姓名和电话号码以及详细到什么级别。可以应用的规则的例子包括,但不限于:

[0385] ●当通过情境解析代词时,通过姓名提及要呼叫的个人。

[0386] ●当从类似情境(诸如文本消息)推断个人时,仅使用名字。

[0387] ●当从应用或者个人数据情境推断电话号码时,使用电话号码的符号名(诸如“移动电话”),而不是要拨号的实际的电话号码。

[0388] 除了引导细节的适当级别,在对话生成步骤500中也可以使用情境1000,例如,为了过滤先前的表达,以避免重复,以及为了指示该会话中的先前提及的实体。

[0389] 本领域技术人员将认识到还可以以其它方式使用情境1000。例如,结合此处描述的技术,可以根据题为“Contextual Voice Commands”、律师案卷号为P7393US1、申请日为2009年6月5日的、美国专利申请序列号12/479,477中描述的机制使用情境1000,该申请的全部公开内容通过引用包含于此。

[0390] 情境收集和通信机制

[0391] 在多个实施例中,在虚拟助理1002中使用不同的机制,以便收集和通信情境信息。例如,在一实施例中,其中在客户端/服务器环境中实施虚拟助理1002,使得在所述客户端和所述服务器之间分布其服务,也可以分布情境1000的源。

[0392] 现在参考图6,示出了根据一实施例在客户端1304和服务器1340之间分布情境1000的源的例子。客户端设备1304可以是移动计算设备或其它设备,其可以是诸如设备传感器数据1056、当前应用情境1060、事件情境2706等的情境信息1000的源。情境1000的其它源可以分布在客户端1304或者服务器1340或二者的一些组合上。例子包括应用偏好和使用历史1072c、1072s;对话历史和助理存储器1052c、1052s;个人数据库1058c、1058s;和个人声学情境数据1080c、1080s。在这些例子的每一个中,情境1000的源可以存在于服务器1340上、客户端1304上或二者上。此外,如以上所描述的,图2中描述的各种步骤可以通过客户端1304或服务器1340,或二者的一些组合执行。

[0393] 在一实施例中,情境1000可以在分布式组件(诸如客户端1304和服务器1340)中进行通信。这些通信可以经由本地API或经由分布式网络,或通过某些其它方式。

[0394] 现在参考图7a至7d,示出了描绘根据多个实施例的用于获取和协调情境信息1000的机制的例子的事件图。存在多种技术用于加载或通信情境,使得对于虚拟助理1002当需要或有用时情境是可用的。关于虚拟助理1002的操作,根据四个可以设置的事件描述这些机制中的每一个:设备或应用初始化601;初始用户输入602;初始输入处理603和依赖情境的处理604。

[0395] 图7a示出了一旦用户输入已经开始602则使用“拉(pull)”机制加载情境信息1000

的方法。一旦用户调用虚拟助理1002并且提供至少某些输入602,虚拟助理1002就加载610情境1000。可以通过从适当源请求或获取情境信息1000来执行加载610。一旦已经加载610情境1000则输入处理603开始。

[0396] 图7b描绘了当初始化601设备或应用时,加载620某些情境信息1000的方法;一旦用户输入已经开始602,则使用拉机制加载附加的情境信息1000。在一实施例中,在初始化时被加载620的情境信息1000可以包括静态情境(即,不频繁变化的情境);一旦用户输入开始602则被加载621的情境信息1000包括动态情境(即,自从加载620静态情境后可能已经变化的情境)。通过从系统运行时性能中去除加载静态情境信息1000的成本,这样的方法可以改善性能。

[0397] 图7c描述了图7b方法的变形。在该例子中,动态情境信息1000允许在输入处理开始603后继续加载621。因此,加载621可以与输入处理并行发生。虚拟助理1002进程仅在处理依赖于所接收的情境信息1000时在步骤604被阻挡。

[0398] 图7d示出了完全可配置版本,其以多达五种不同方式中的任意一种处理情境:

[0399] ●静态情境信息1000是从情境源到运行虚拟助理1002的环境或设备在一个方向同步640。由于情境源中的数据变化了,该变化被推送给虚拟助理1002。例如,当地址簿被初始创建或启动时,可以将该地址簿同步到虚拟助理1002。每当该地址簿被修改时,就立即地或以批处理方式,将变化被推送给虚拟助理1002。如图7d中所描述的,这样的同步640可以在任意时刻发生,包括在用户输入开始602之前。

[0400] ●在一实施例中,当用户输入开始602时,静态情境源的同步状态可以被检查。如果必要,开始641同步保持静态情境信息1000的处理。

[0401] ●当用户输入开始602时,加载642某些动态情境1000,如在610和621那样,使用情境1000的处理仅仅被阻挡,以等待尚未加载的它们需要的情境信息1000。

[0402] ●当处理需要其它情境信息1000时,处理根据需求643加载其它情境信息1000。

[0403] ●当事件发生时,将事件情境2706从源发送644到运行虚拟助理1002的设备。使用事件情境2706的处理仅等待事件高速缓存就绪,并且可以继续前进而不在其后任意时刻阻挡。以该方式加载的事件情境2706包括以下的任意一个:

[0404] ●在用户输入开始602之前加载事件情境2706,例如未读消息通知。例如,使用同步高速缓存,可以保持这种信息。

[0405] ●与用户输入开始602同时地或在用户输入开始602之后加载事件情境2706。作为例子,当用户正与虚拟助理1002交互时,文本消息可能到达;可以与助理1002的处理并行地推送事件情境,所述事件情境通知助理1002该事件。

[0406] 在一实施例中,对于情境信息1000的每个源,通过规定通信策略和访问API,实现在获取和协调情境信息1000时的灵活性,这样相对于具有针对每个请求可用的信息的值平衡了通信成本。例如,与每个语音到文本请求相关的变量,诸如个人声学情境数据1080或描述麦克风参数的设备传感器数据1056,可以针对每个请求加载。这样的通信策略可以例如在配置表中被指定。

[0407] 现在参考图9,示出了根据一实施例、可以用于为情境信息1000的不同源指定通信和高速缓存策略的配置表900的例子。对于多个不同情境源(包括用户名称、地址簿名称、地址簿号码、SMS事件情境和日历数据库)中的每一个,针对图2的步骤中的每一个指定特定类

型的情境加载：引导并解释语音输入100、解释自然语言200、识别任务300及生成对话响应500。表900中的每个条目指示以下之一：

[0408] ●同步：在设备上同步情境信息1000；

[0409] ●按需：响应于虚拟助理1002对情境信息1000的请求，提供情境信息1000；

[0410] ●推送：将情境信息1000推送到设备。

[0411] 可完全配置的方法使得大范围的潜在相关情境信息1000成为可用，以简化所述在人和机器之间的自然语言接口。将某些信息保持在情境源和虚拟助理1002两者中，而按需查询其它信息，而不是加载所有时间的全部信息（这样可能导致低效）。例如，如上所述，在本地保持用于实时操作（诸如语音识别）的信息（诸如名称），而按需查询仅被某些可能的请求使用的信息，诸如用户个人日历。在用户调用助理时，不能预期的数据（诸如到来的SMS事件）在当它们发生时被推送。

[0412] 现在参考图10，其示出了描绘根据一实施例、在助理1002与用户进行对话的交互序列的处理期间，访问图9中所配置的情境信息源的例子的事件图950。

[0413] 图10中示出的序列表示以下交互序列：

[0414] ●T1:助理1002:“Hello Steve, what I can do for you? (你好Steve,我能为你做什么?)”

[0415] ●T2:用户:“When is my next meeting? (我的下个会在什么时候?)”

[0416] ●T3:助理1002:“Your next meeting is at 1:00pm in the boardroom. (你的下个会议在会议室,下午1点。)”

[0417] ●T4: [到来的SMS消息的声音]

[0418] ●T5:用户:“Read me that message. (为我读取那个消息。)”

[0419] ●T6:助理1002:“Your message from Johnny says ‘How about lunch’ (你的消息来自Johnny,内容是“午餐怎样”)”

[0420] ●T7:用户:“Tell Johnny I can’t make it today. (告诉Johnny我今天不能吃。)”

[0421] ●T8:助理1002:“OK, I’ll tell him. (好,我会告诉他。)”

[0422] 在时刻T0,在交互开始之前,同步770用户姓名并且同步771地址簿名称。如图7d的组成部分640所示,这些是在初始化时加载的静态情境的例子。这允许助理1002通过他的名字(“Steve”)指示该用户。

[0423] 在时刻T1,完成同步步骤770和771。在时刻T2,该用户说出请求,该请求根据图2的步骤100、200和300被处理。在任务识别步骤300,虚拟助理1002查询774作为情境1000的源的用户个人数据库1058:特别地,虚拟助理1002请求来自该用户日历数据库的信息,所述用户日历数据库被配置为根据表900按需访问。在时刻T3,执行步骤500并生成对话响应。

[0424] 在时刻T4,接收SMS消息;这是事件情境2706的例子。基于表900中的配置,将该事件的通知推送773给虚拟助理1002。

[0425] 在时刻T5,该用户要求虚拟助理1002读取SMS消息。在执行步骤200时,事件情境2706的存在引导NLP组件以将“that message (那个消息)”解释为新的SMS消息。在时刻T6,可以通过任务组件执行步骤300以调用API来将SMS消息读给用户。在时刻T7,该用户以模糊的动词(“tell (告诉)”)和姓名(“Johnny”)做出请求。通过使用包括在步骤773中所接收的

事件情境2706的情境1000的不同源解决这些歧义,该NLP组件解释自然语言200;其告知该NLP组件命令指的是来自名为Johnny的个人的SMS消息。在步骤T7执行流程步骤400,包括通过从所接收到的事件情境对象中查找要使用的号码以匹配该姓名771。助理1002由此能够撰写新的SMS消息并且将其发送给Johnny,如在步骤T8中所确认的。

[0426] 已经相对于可能的实施例以具体的细节描述了本发明。本领域技术人员将认识到发明可以在其它实施例中实践。首先,组件的特定命名、术语的大写、属性、数据结构,或任意其它编程或结构化的方面不是强制的或重要的,并且实施本发明或其特征的机制可以具有不同的名称、格式或协议。此外,如所描述的,系统可以通过硬件和软件的组合实施,或完全以硬件元件,或完全以软件元件实施。而且,此处描述的所述不同系统组件之间的功能的具体划分仅仅是示例性的,并不是强制性的;由单个系统组件执行的功能可以代替地由多个组件执行,并且由多个组件执行的功能可以代替地由单个组件执行。

[0427] 在不同实施例中,本发明可以被实施为用于单独地或以任意组合执行以上描述的技术的系统或方法。在另一实施例中,本发明可以被实施为包括非瞬时性计算机可读存储媒体和计算机程序代码的计算机程序产品,所述计算机程序代码编码在非瞬时性计算机可读存储媒体上,使得计算设备或其它电子设备中的处理器执行以上描述的技术。

[0428] 本说明书中关于“一实施例”或“实施例”的引用意指在本发明的至少一实施例中包括结合所述实施例所描述的特定特征、结构或特性。在本说明书中不同地方出现的短语“在一实施例中”不一定全部都指示同一实施例。

[0429] 关于对在计算设备的存储器内的数据比特的操作的算法和符号的表示,给出了以上的某些部分。这些算法的描述和表示是数据处理领域技术人员用以最有效地将他们的工作的实质传递给本领域其它技术人员的手段。这里的算法通常被认为是导致期望结果的前后一致的步骤(指令)序列。所述步骤是物理量的那些需要的物理操作。通常,虽然不是必须的,这些量采用能被存储、传送、组合、比较和以其它方式操作的电的、磁的或光学的信号的形式。主要出于共同使用的原因,有时将这些信号指示为比特、值、元素、符号、字符、术语、号码等是方便的。此外,有时将需要对物理量进行物理操作的步骤的某些安排指示为模块或代码装置也是方便的,而不失一般性。

[0430] 然而,应当记住所有这些和类似术语与适当的物理量相关联,并且仅仅是应用于这些量的便利标记。除非从以下论述中显然地另外地特殊声明,应当认识到在说明书中,利用诸如“处理”或“计算”或“算出”或“显示”或“确定”等术语的论述,指的是计算机系统或类似电子计算模型和/或设备的动作和处理,其操作并且转换在计算机系统存储器或寄存器或其它这样的信息存储、传输或显示设备内的被表示为物理(电子)量的数据。

[0431] 本发明的特定方面包括此处描述的采用算法形式的处理步骤和指令。应当指出,本发明的处理步骤和指令可以嵌入在软件、固件和/或硬件中,并且当嵌入在软件中时,可以下载所述处理步骤和指令以驻留在各种各样操作系统所使用的不同平台上,并且从所述平台操作所述处理步骤和指令。

[0432] 本发明还涉及用于执行本文的操作的装置。为所要求的目的,该装置可以被专门构造的,或其可以包括由在该计算设备中存储的计算机程序选择性激活或重新配置的通用计算设备。这样的计算机程序可以存储在计算机可以读存储介质中,例如但不限于,任意类型的盘,包括软盘、光盘、CD-ROM、磁光盘、只读存储器(ROM)、随机存取存储器(RAM)、EPROM、

EEPROM、磁或光卡、专用集成电路 (ASIC) 或者适于存储电指令的任意类型的介质,并且每一个都耦合到计算机系统总线。此外,此处涉及的计算设备可以包括单个处理器或可以是使用多处理器设计的架构,以获得增加的计算能力。

[0433] 此处所呈现的算法和显示并不是固有地与任意的特定计算设备、虚拟化系统或其它装置有关。依照此处的教导,各种通用系统也可以与程序一起使用,或可以证明构造更专用的装置来执行所要求的方法步骤是方便的。根据此处提供的描述,用于各种各样这样的系统的所要求的结构将是明显的。此外,本发明不是参照任意特定的编程语言描述的。应当认识到,可以使用各种各样的编程语言来实施此处描述的本发明的教导,并且为揭示本发明的实现和最佳模式,才提供了以上对特定语言的任意参照。

[0434] 相应地,在各个实施例中,本发明可以被实施为软件、硬件、和/或用于控制计算机系统、计算设备或其它电子设备的其它元件,或其任意组合或其多个。根据本领域公知的技术,这样的电子设备包括,例如,处理器、输入设备(诸如键盘、鼠标、触摸板、跟踪板、操纵杆、跟踪球、麦克风和/或其任意组合)、输出设备(诸如屏幕、扬声器等)、存储器、长期存储器(例如磁存储装置、光存储装置等)和/或网络连接。这样的电子设备可以是便携式或非便携式的。可以使用以实施本发明的电子设备的例子包括:移动电话、个人数字助理、智能电话、信息亭(kiosk)、台式计算机、膝上计算机、平板计算机、消费者电子设备、消费者娱乐设备;音乐播放器;照相机;电视机;机顶盒;电子游戏单元等。用于实施本发明的电子设备可以使用任意操作系统,例如,可以从加利福尼亚的库珀蒂诺的苹果公司(Apple Inc. of Cupertino, California)获得的iOS或MacOS,或适合在设备上使用的任意其它操作系统。

[0435] 根据某些实施例,图33示出了根据以上描述的本发明的原理配置的、用于解释用户输入来执行任务的系统3000的功能框图。该系统3000的功能块可以由用于实现本发明的原理的硬件、软件或者硬件和软件的组合实施。本领域技术人员理解图33中所描述的功能块可以被组合或被拆分为子块来实施如上所述的本发明的原理。因此,本说明书可以支持此处描述的功能块的任意可能的组合或者拆分或者进一步的限定。

[0436] 如图33中所示,用于解释用户输入以执行任务的系统3000包括:输出设备3100、输入设备3200和至少一个处理器3300。输出设备3100被配置为提示用户进行输入。输入设备3200被配置为接收用户输入。所述至少一个处理器3300通信地耦合到该输出设备3100和该输入设备3200。所述至少一个处理器3300包括接收单元3310、解释单元3320、识别单元3330、执行单元3340和生成单元3350。接收单元3310被配置为从情境源接收情境信息。解释单元3320被配置为解释所接收到的用户输入以得到用户意图的表示。确定单元3330被配置为至少部分地基于所得到的用户意图的表示,识别至少一个任务和用于该任务的至少一个参数。执行单元3340被配置为使用所述至少一个参数来执行至少一个任务,以得到结果。生成单元被配置为基于所述得到的结果生成对话响应。输出设备3100被进一步地配置为输出所生成的对话响应。使用所接收到的情境信息执行提示该用户进行输入、解释所述接收的用户输入、识别至少一个任务和用于该任务的至少一个参数和生成所述对话响应中的至少一个。

[0437] 在某些实施例中,所接收的情境信息包括从包括以下各项的组中选择的至少一个:应用情境;关联于该用户的个人数据;来自关联于用户的数据库的数据;从对话历史获得的数据;从至少一传感器接收的数据;应用偏好;应用使用历史;描述事件的数据;描述其

中接收语音输入的声学环境的数据;当前对话状态;由该用户先前提提供的输入;位置;本地时间;和环境条件。

[0438] 在某些实施例中,输入设备3100被配置为经由会话接口提示用户;并且输入设备3200被配置为通过接收语音输入经由该会话接口接收用户输入;并且其中所述至少一个处理器3300进一步包括转换单元3330,该转换单元3330被配置为将语音输入转换为文本表示。

[0439] 在某些实施例中,转换单元3330被配置为通过以下操作将语音输入转换为文本表示:生成语音输入的多个候选文本解释,以及对所生成的候选文本解释的至少一子集进行排序;其中使用所接收的情境信息执行所述生成和排序操作中的至少一个。

[0440] 在某些实施例中,在所述生成和排序操作的至少一个中使用的接收的情境信息包括从包括以下各项的组中选择的至少一个:描述其中接收语音输入的声学环境的数据;从至少一个传感器接收的数据;从关联于用户的数据库获得的词汇;关联于应用偏好的词汇;从使用历史获得的词汇;和当前对话状态。

[0441] 在某些实施例中,输出设备3100被配置为至少部分地基于所接收的情境信息,通过生成至少一个提示来提示用户。

[0442] 在某些实施例中,解释单元3330被配置为至少部分地基于所接收的情境信息,通过对所接收的用户输入执行自然语言处理来得到用户意图的表示。

[0443] 在某些实施例中,在解释所接收的用户输入中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:描述事件的数据;应用情境;由用户先前提提供的输入;关于用户的已知信息;位置;日期;环境条件;和历史。

[0444] 在某些实施例中,确定单元3330被配置为通过至少部分地基于所接收的情境信息识别所述至少一个任务和用于该任务的至少一个参数,来识别所述至少一个任务和用于该任务的至少一个参数。

[0445] 在某些实施例中,在确定至少一个任务和用于该任务的至少一个参数中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:描述事件的数据;来自关联于用户的数据库的数据;从至少一个传感器接收的数据;应用情境;用户先前提提供的输入;关于用户的已知信息;位置;日期;环境条件;和历史。

[0446] 在某些实施例中,生成单元3350被配置为通过至少部分地基于所接收的情境信息生成对话响应来生成对话响应。

[0447] 在某些实施例中,在生成对话响应中使用的所接收的情境信息包括从包括以下各项的组中选择的至少一个:来自关联于用户的数据库的数据;应用情境;用户先前提提供的输入;关于用户的已知信息;位置;日期;环境条件;和历史。

[0448] 在某些实施例中,所接收的情境信息包括从包括以下各项的组中选择的至少一个:存储在服务器处的情境信息;和存储在客户端处的情境信息。

[0449] 在某些实施例中,接收单元3310被配置为通过请求来自情境源的情境信息以及响应于所述请求接收所述情境信息,来从情境源接收情境信息。

[0450] 在某些实施例中,接收单元3310被配置为通过在接收用户输入之前接收所述情境信息的至少一部分,来从情境源接收情境信息。

[0451] 在某些实施例中,接收单元3310被配置通过在接收用户输入之后接收所述情境信

息的至少一部分,来从情境源接收情境信息。

[0452] 在某些实施例中,接收单元3310被配置为通过作为初始化操作的一部分接收静态情境信息和在接收用户输入之后接收附加情境信息,来从情境源接收情境信息。

[0453] 在某些实施例中,接收单元3310被配置为通过接收情境信息中的变化的推送通知,以及响应于所述推送通知,更新本地存储的情境信息,来从情境源接收情境信息。

[0454] 在某些实施例中,输出设备3100,输入设备3200和至少一个处理器3300被实施为从包括以下各项的组中选择的至少一个的组件:电话;智能电话;平板电脑;膝上计算机;个人数字助理;台式计算机;信息亭;消费者电子设备;消费者娱乐设备;音乐播放器;照相机;电视机;电子游戏单元;和机顶盒。

[0455] 根据某些实施例,图34示出了根据如上所述的本发明的原理配置的用于解释用户输入以在计算设备上执行任务的装置4000的功能框图。该装置4000的功能块可以由用以实现本发明的原理的硬件、软件、或硬件和软件的组合实施。本领域技术人员理解图34中所描述的功能块可以被组合或被拆分为子块来实施如以上所述的发明的原理。因此,此处的描述可以支持此处描述的功能块的任意可能组合或拆分或进一步的限定。

[0456] 如图34中所示,用于解释用户输入以在计算设备上执行任务的装置4000包括:第一接收单元4100,提示单元4200,第二接收单元4300,解释单元4400,识别单元4500,执行单元4600,生成单元4700,输出单元4800和转换单元4900。第一接收单元4100被配置为从情境源接收情境信息。提示单元4200被配置为提示用户进行输入。第二接收单元4300被配置为接收用户输入。解释单元4400被配置为解释所接收的用户输入来得到用户意图的表示。识别单元4500被配置为至少部分地基于所得到的用户意图的表示,识别至少一个任务和用于该任务的至少一个参数。执行单元4600被配置为使用所述至少一个参数执行所述至少一个任务,以得到结果。生成单元4700被配置为基于所得到的结果生成对话响应。输出单元4800被配置为输出所生成的对话响应。使用所接收到的情境信息执行提示用户进行输入、解释所接收的用户输入、识别所述至少一个任务和用于所述任务中的至少一个参数和生成对话响应中的至少一个。

[0457] 在某些实施例中,提示单元被配置为经由会话接口提示该用户;并且第二接收单元被配置为通过接收语音输入经由该会话接口接收用户输入;并且其中该装置进一步包括被配置为将语音输入转换为文本表示的转换单元4900。

[0458] 在某些实施例中,转换单元4900被配置通过生成该语音输入的多个候选文本解释以及对所生成的候选文本解释的至少一子集进行排序,来讲该语音输入转换为文本表示;其中使用所接收的情境信息执行所述生成和排序操作中的至少一个。

[0459] 虽然本发明已经就有限数量的实施例进行了描述,但是受益于以上描述的本领域技术人员将认识到可以设计不脱离如此处描述的本发明的范围的其它实施例。此外,应当指出,为了可读性和指导目的,原则性地选择了本说明书中使用的语言,并且该语言可以不被选择以描绘或限制本发明的主题。因此,本发明的公开内容意在示例而不是限制本发明的范围,本发明的范围在权利要求中阐明。

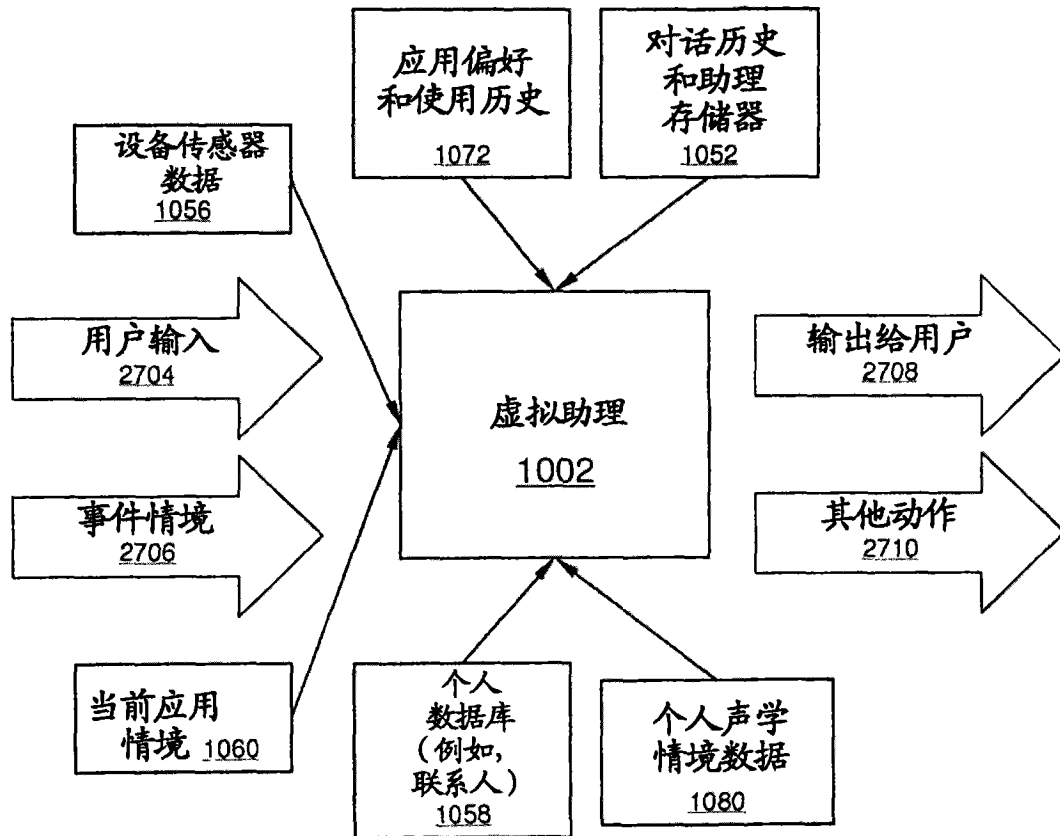


图1

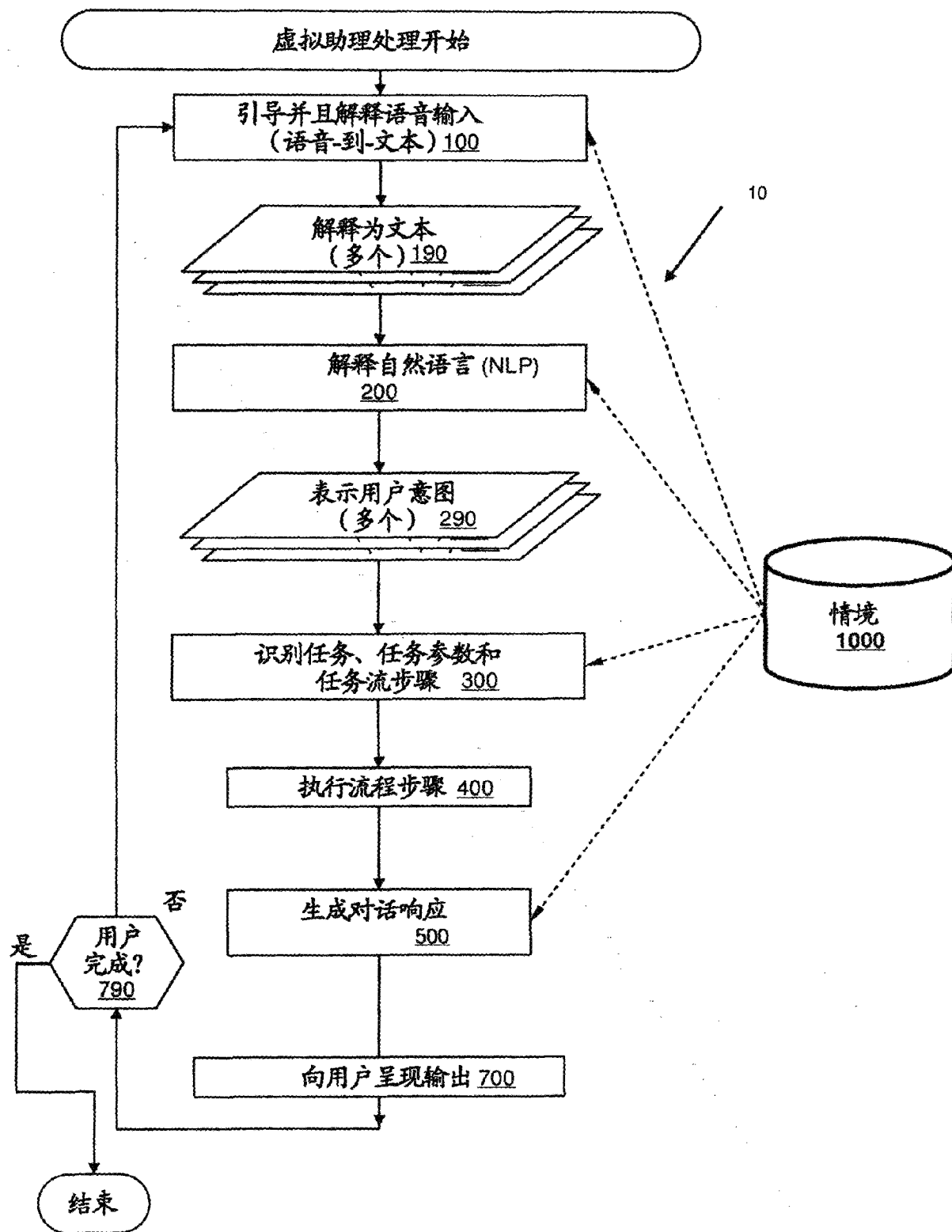


图2

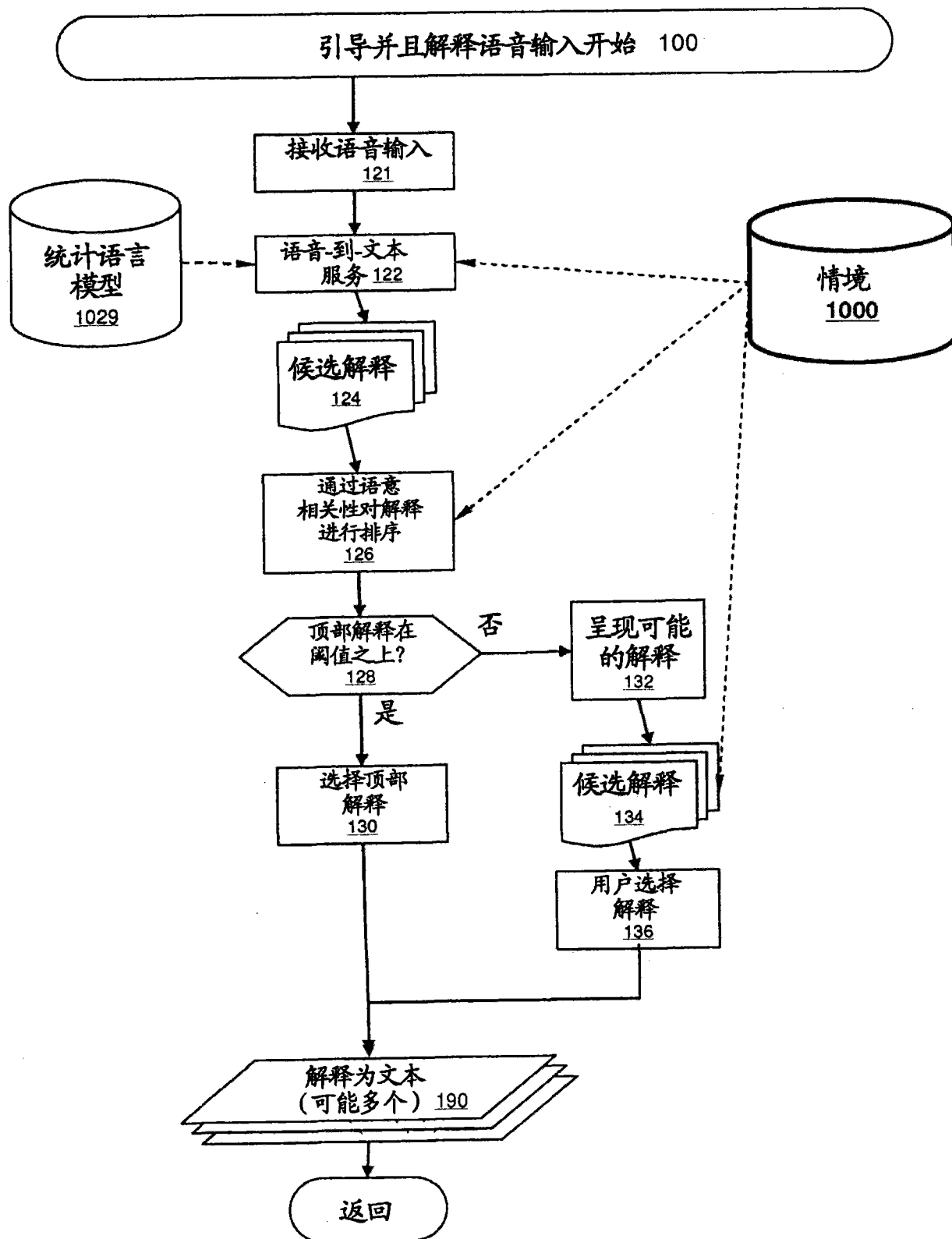


图3

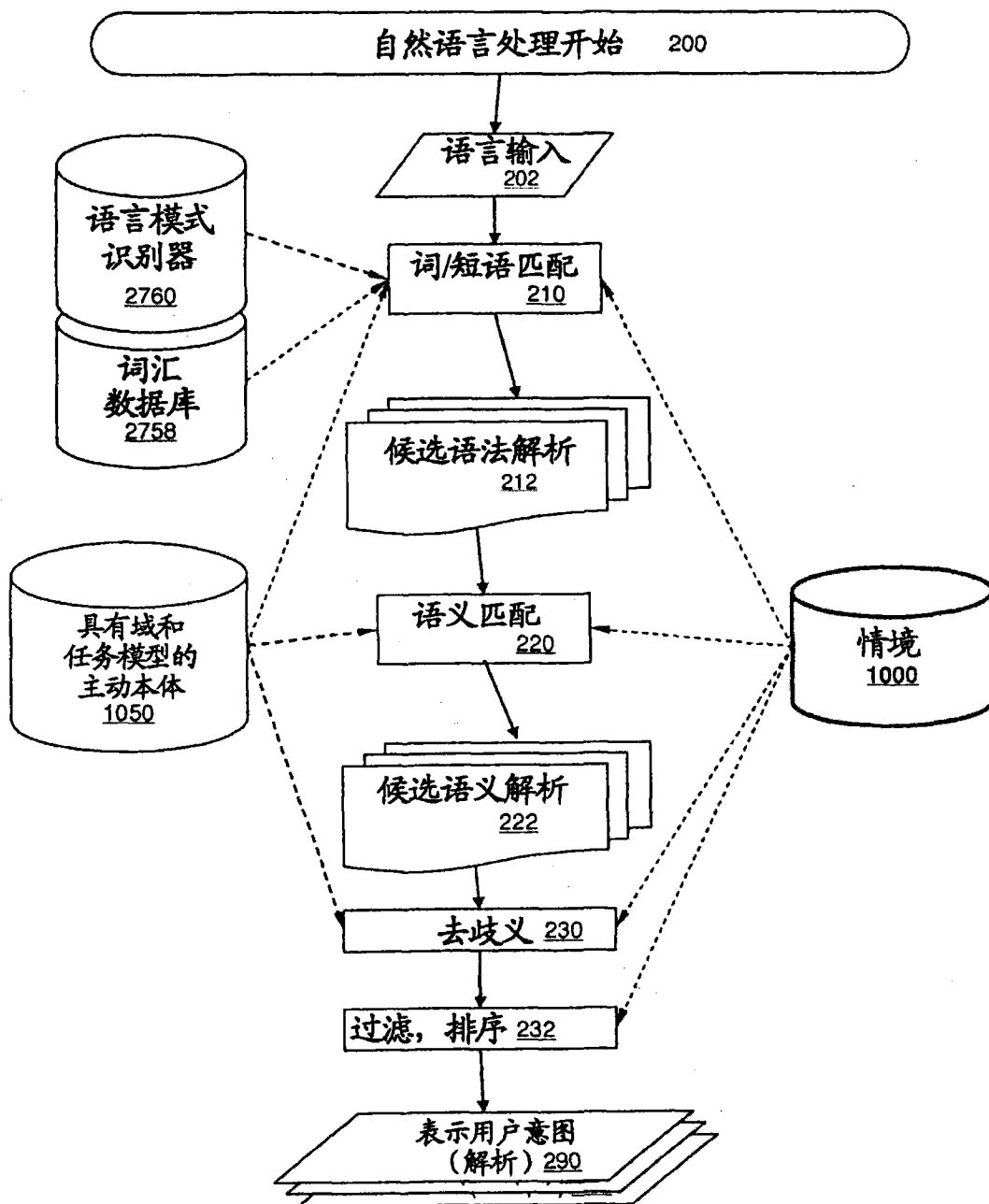


图4

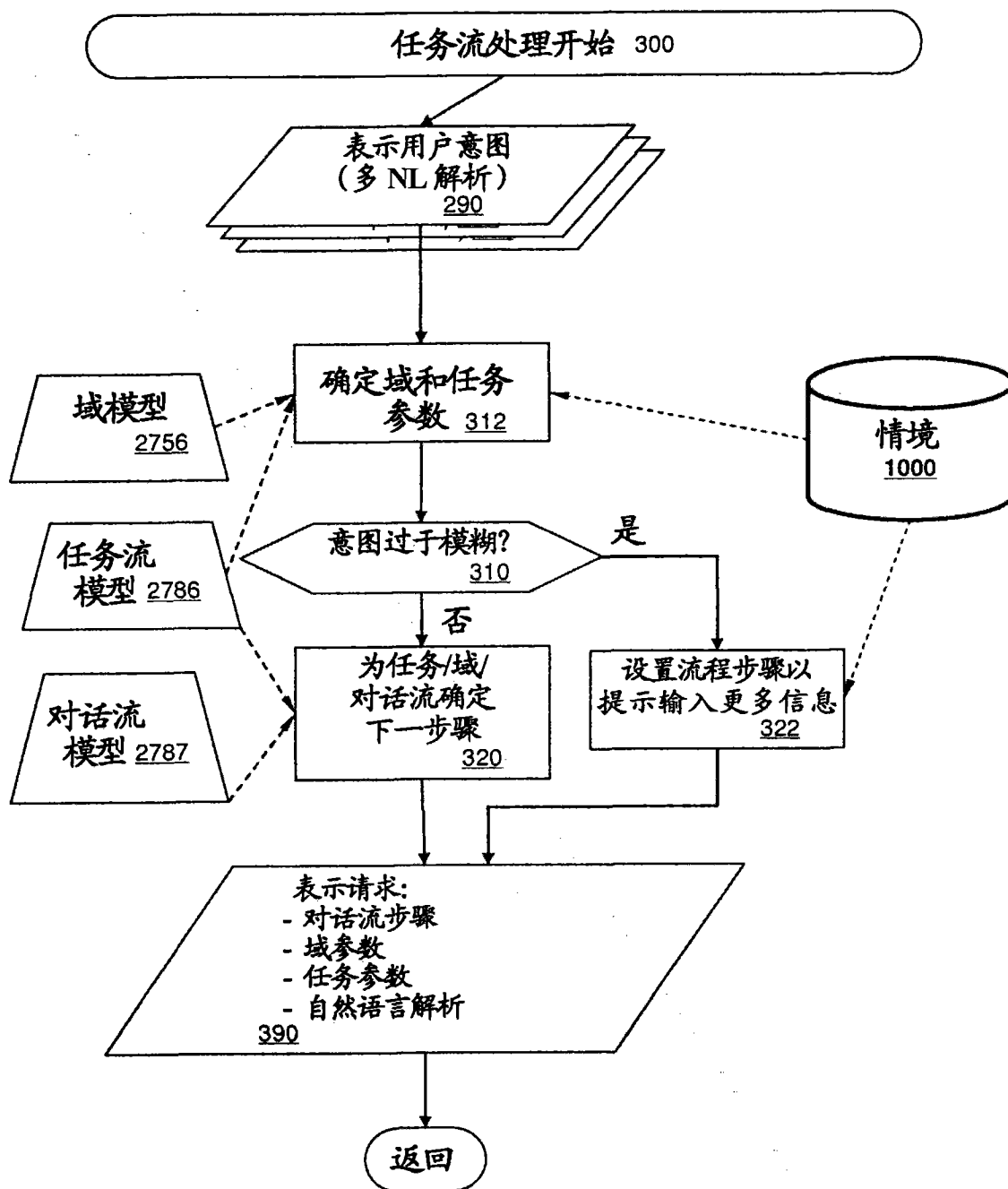


图5

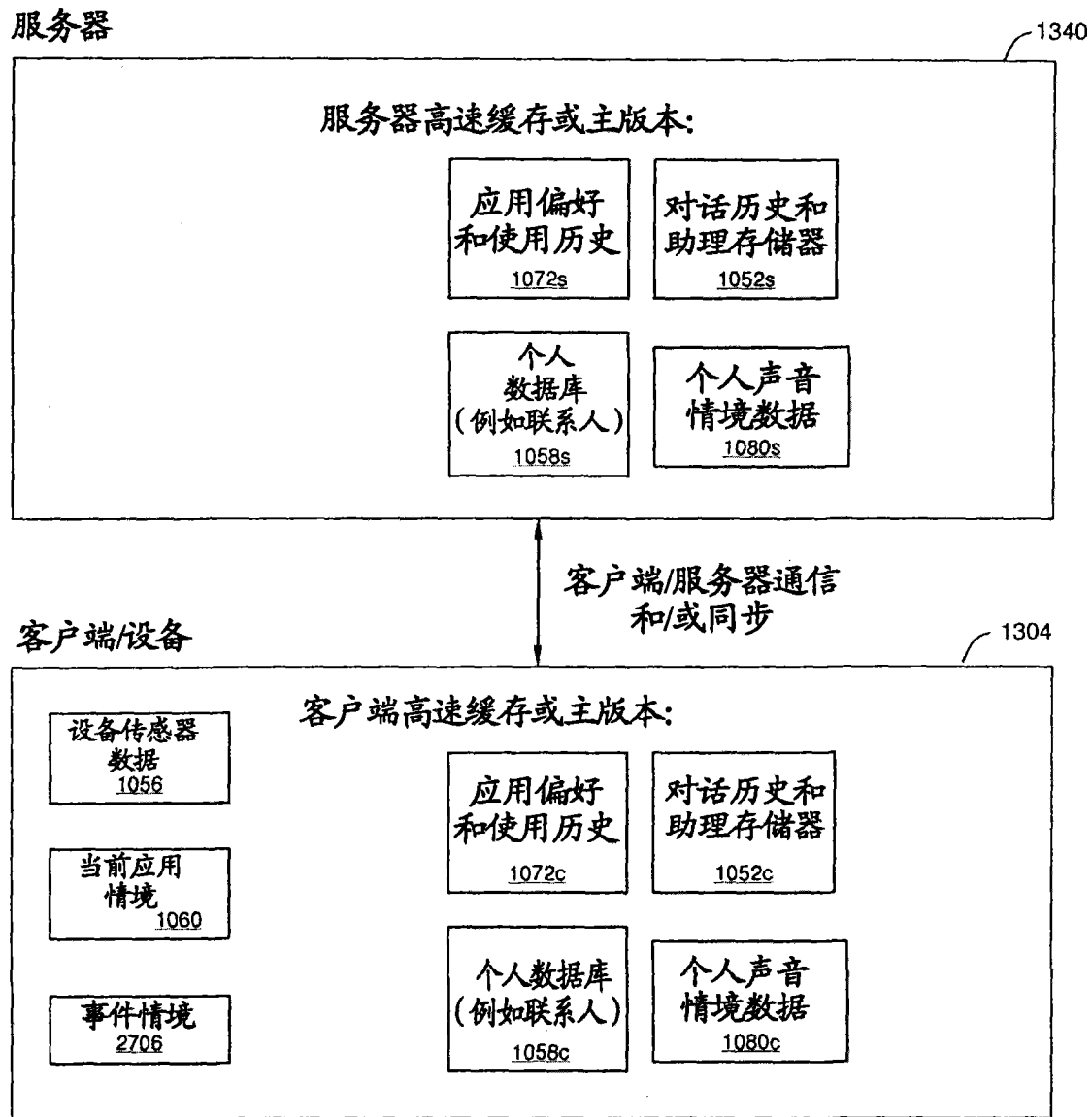


图6

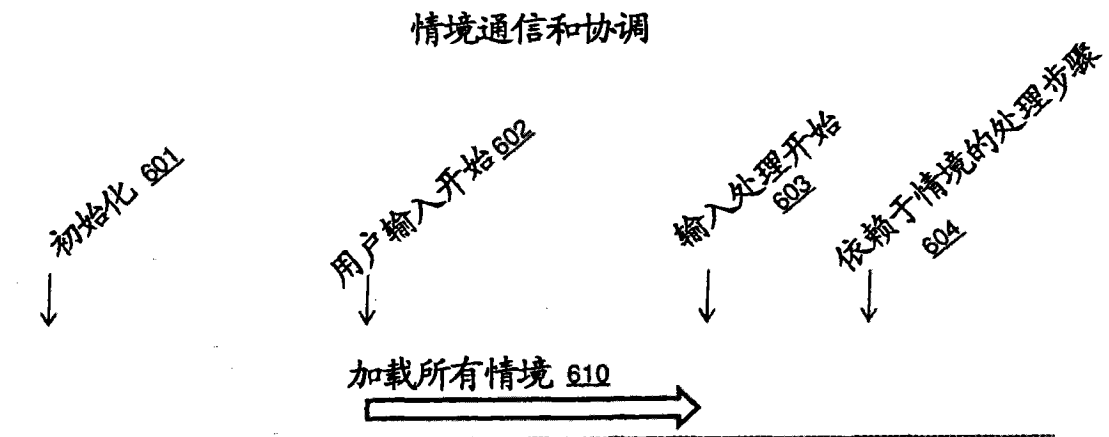


图7a

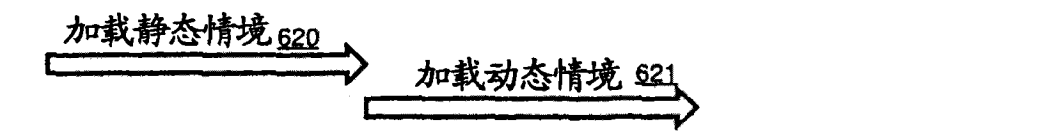


图7b

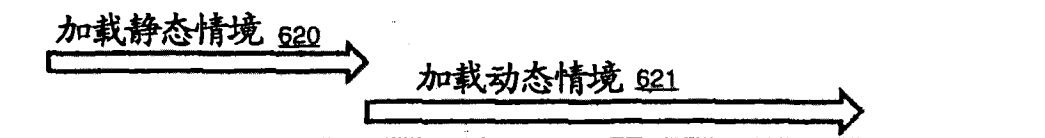


图7c

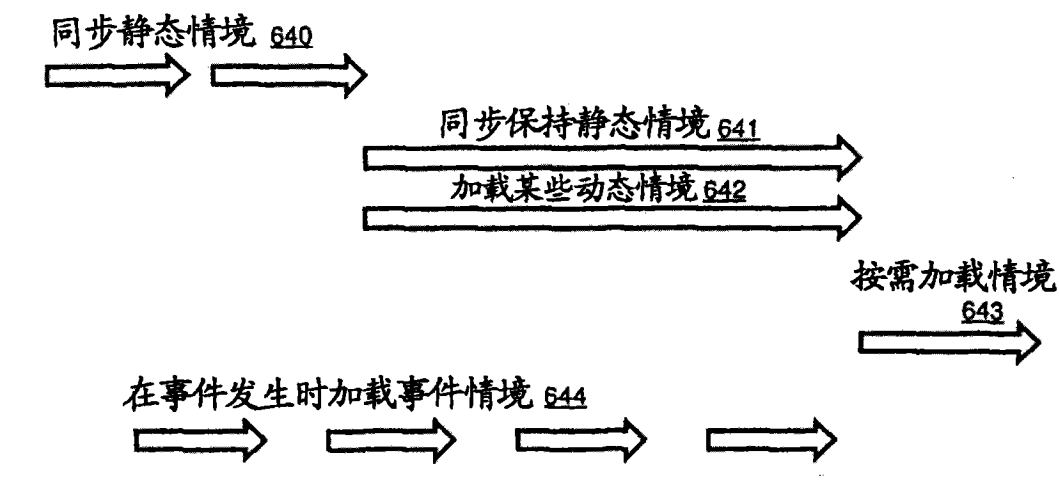


图7d

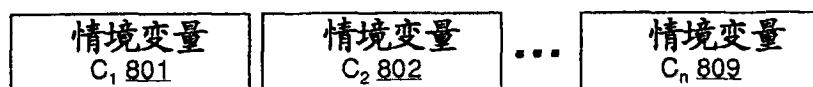


图8a

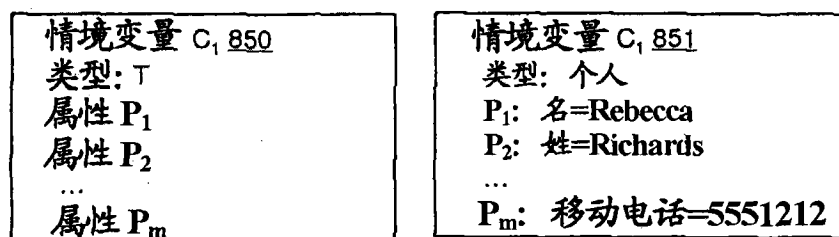


图8b

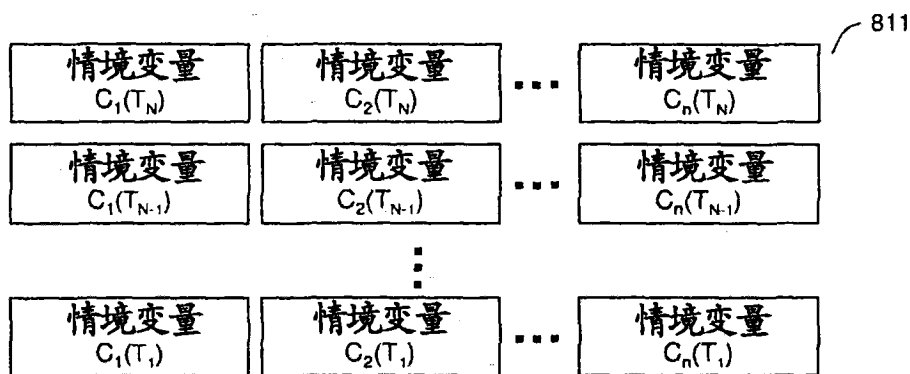


图8c

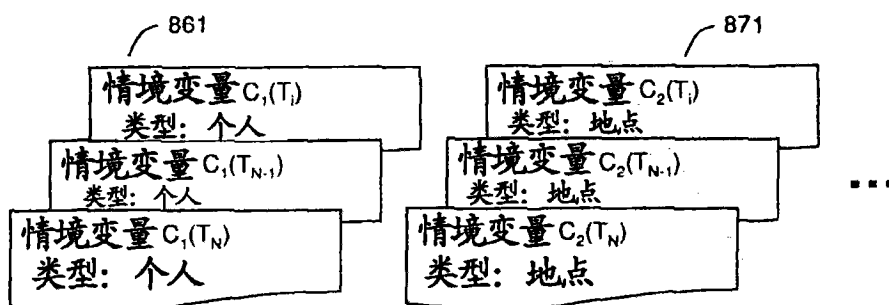


图8d

900

| 情境源      | 语音<br>100 | NLP<br>200 | 任务<br>300 | 对话<br>500 |
|----------|-----------|------------|-----------|-----------|
| 用户名      | 同步        | 同步         | 同步        | 同步        |
| 地址簿名称    | 同步        | 同步         | 同步        | -         |
| 地址簿号码    | -         | -          | 按需        | -         |
| SMS 事件情境 | -         | 推送         | 推送        | -         |
| 日历数据库    | -         | 按需         | 按需        | -         |

图9

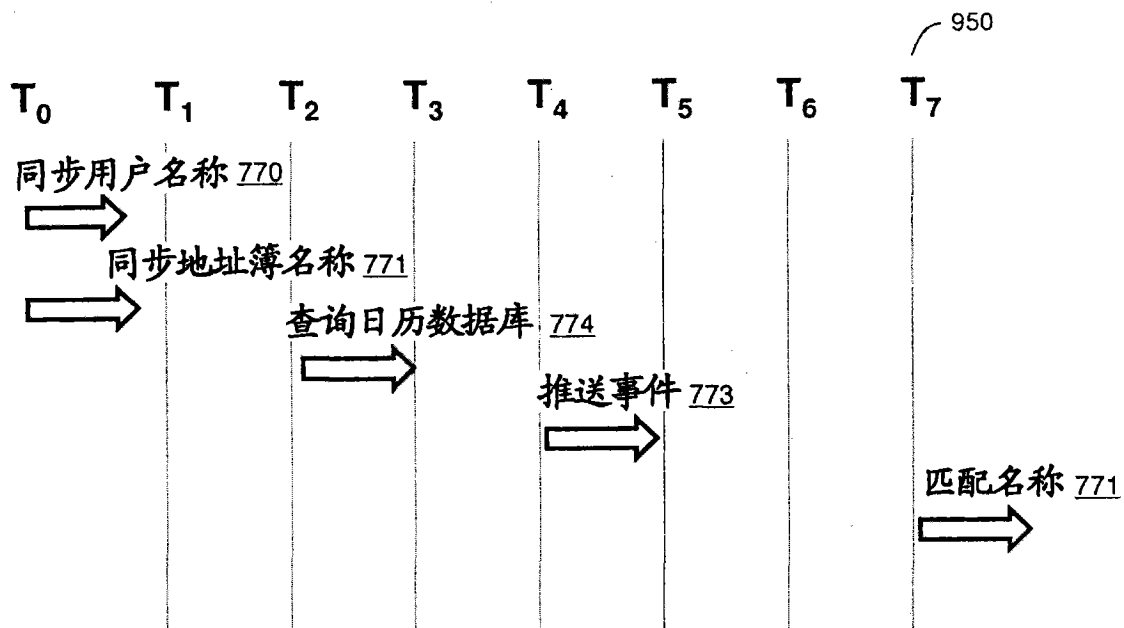


图10

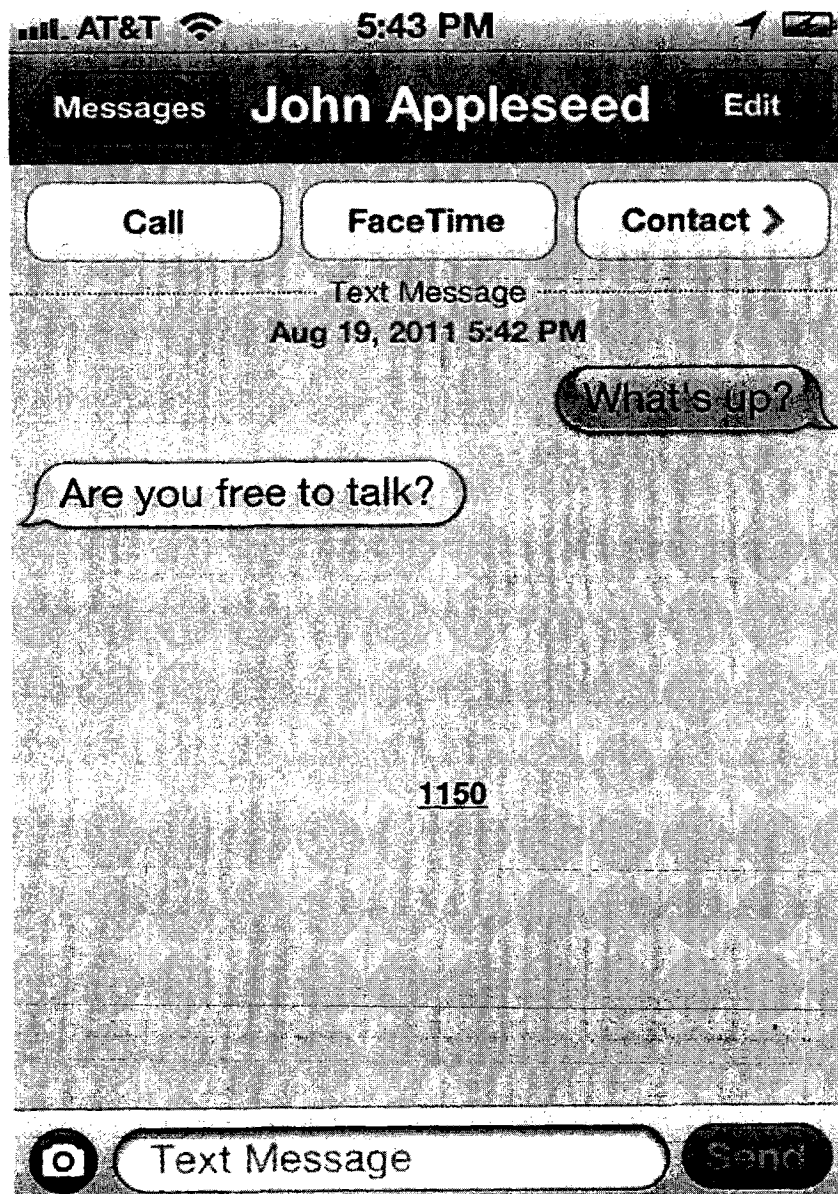


图11

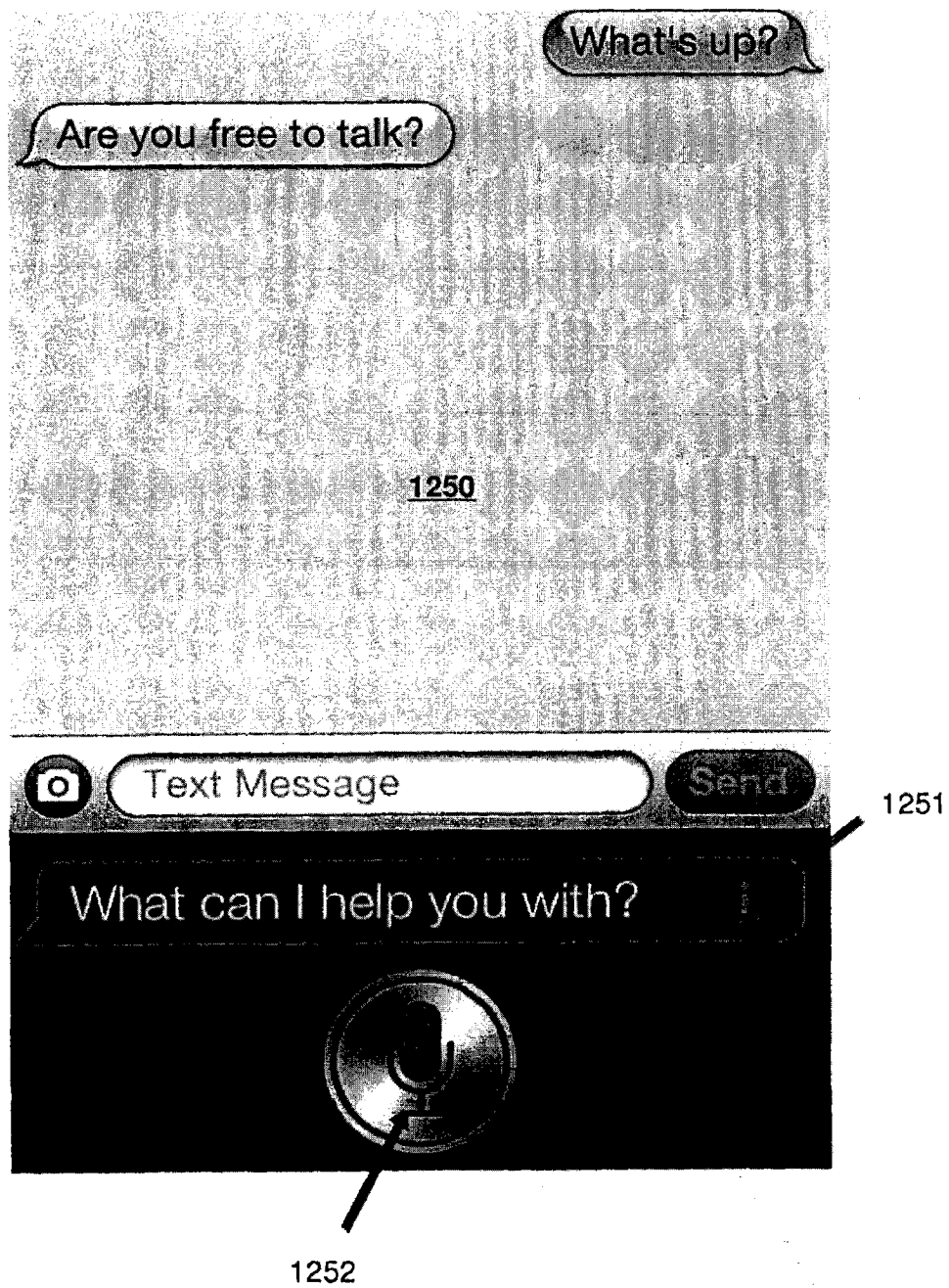


图12

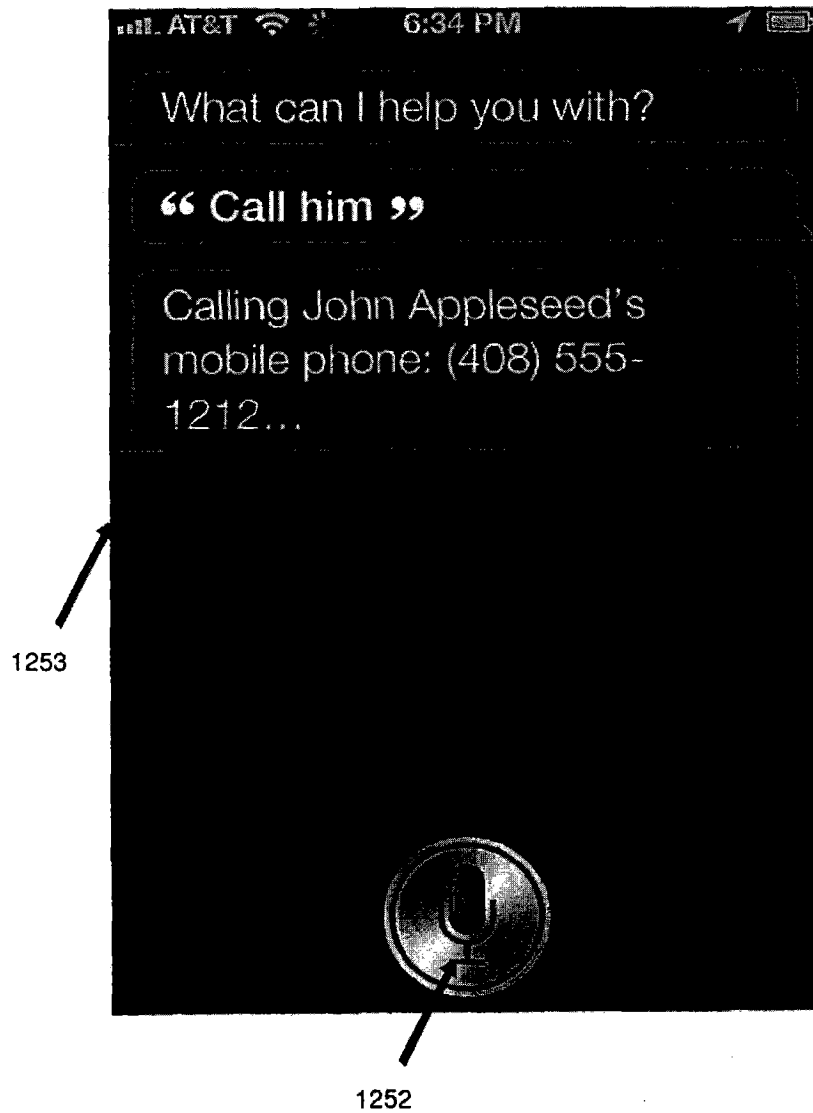


图13

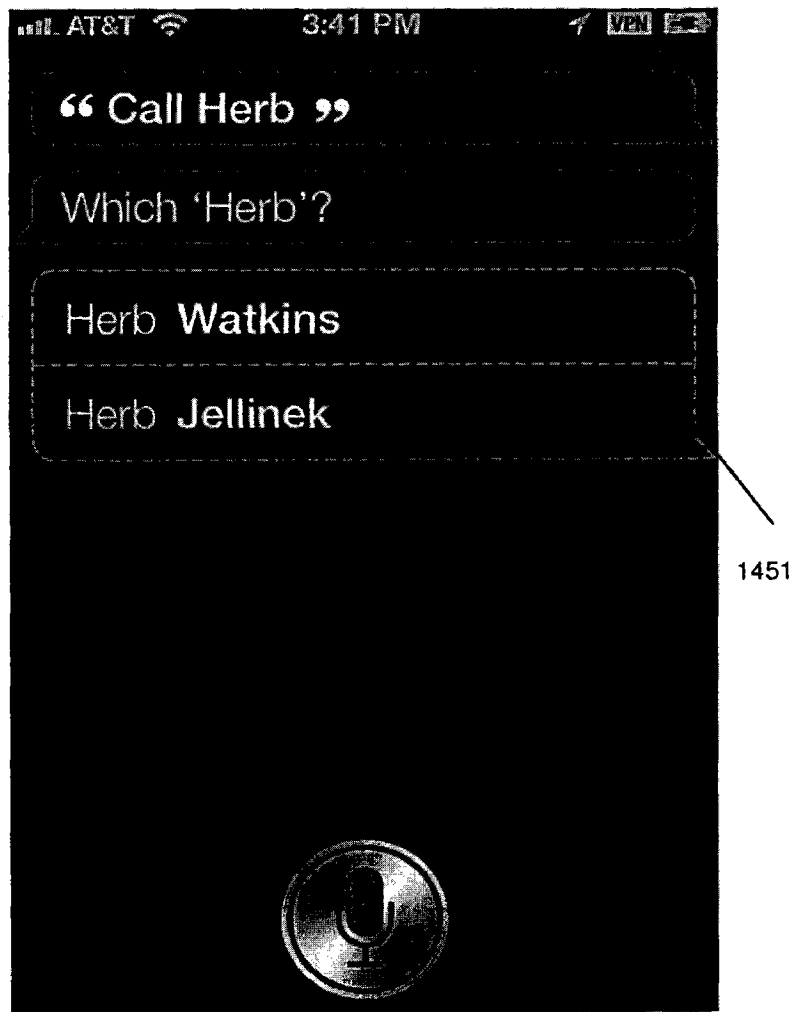


图14



图15

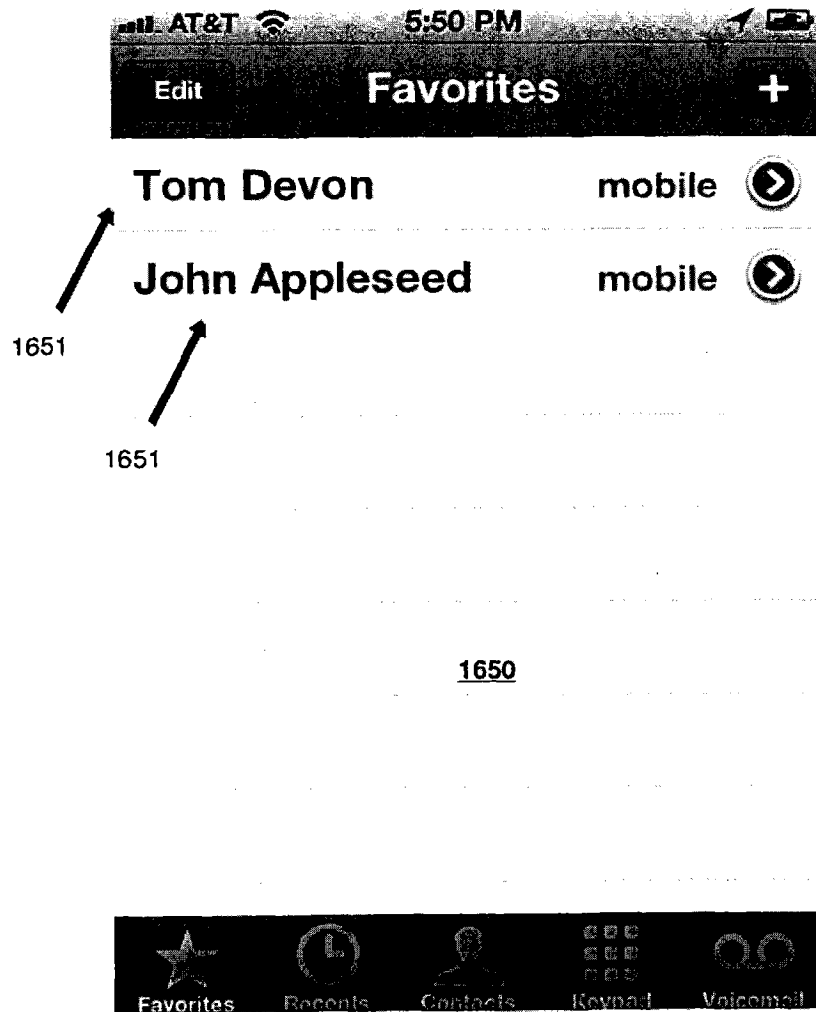


图16



图17

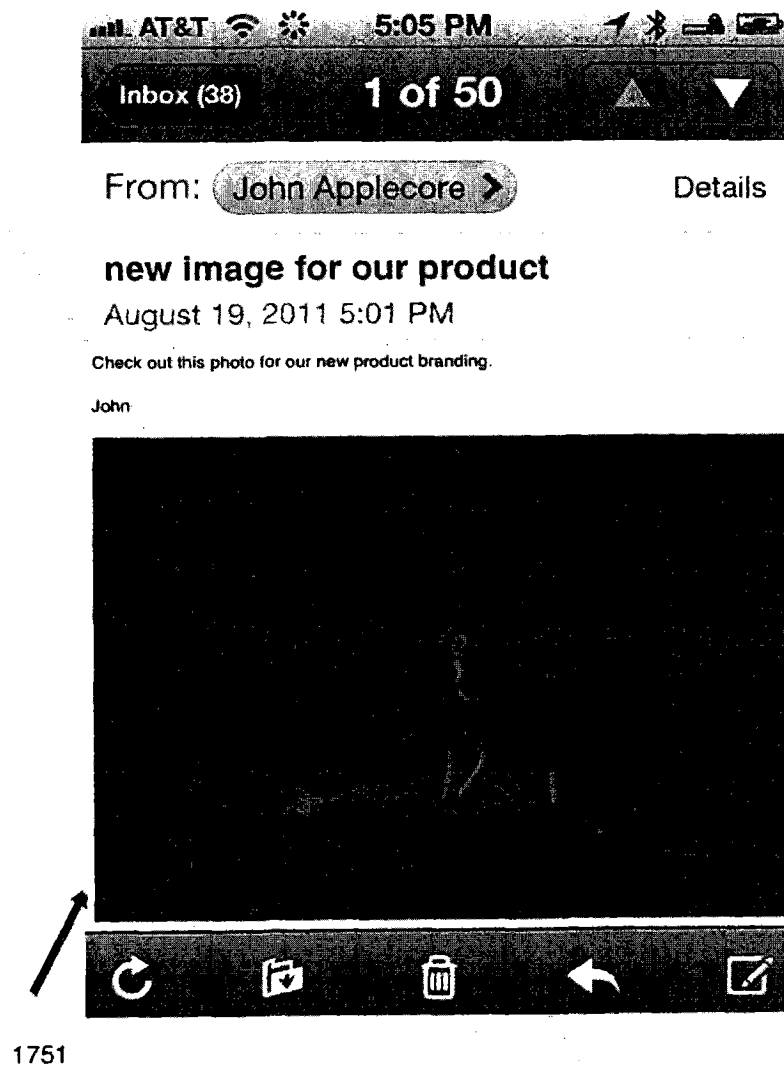


图18

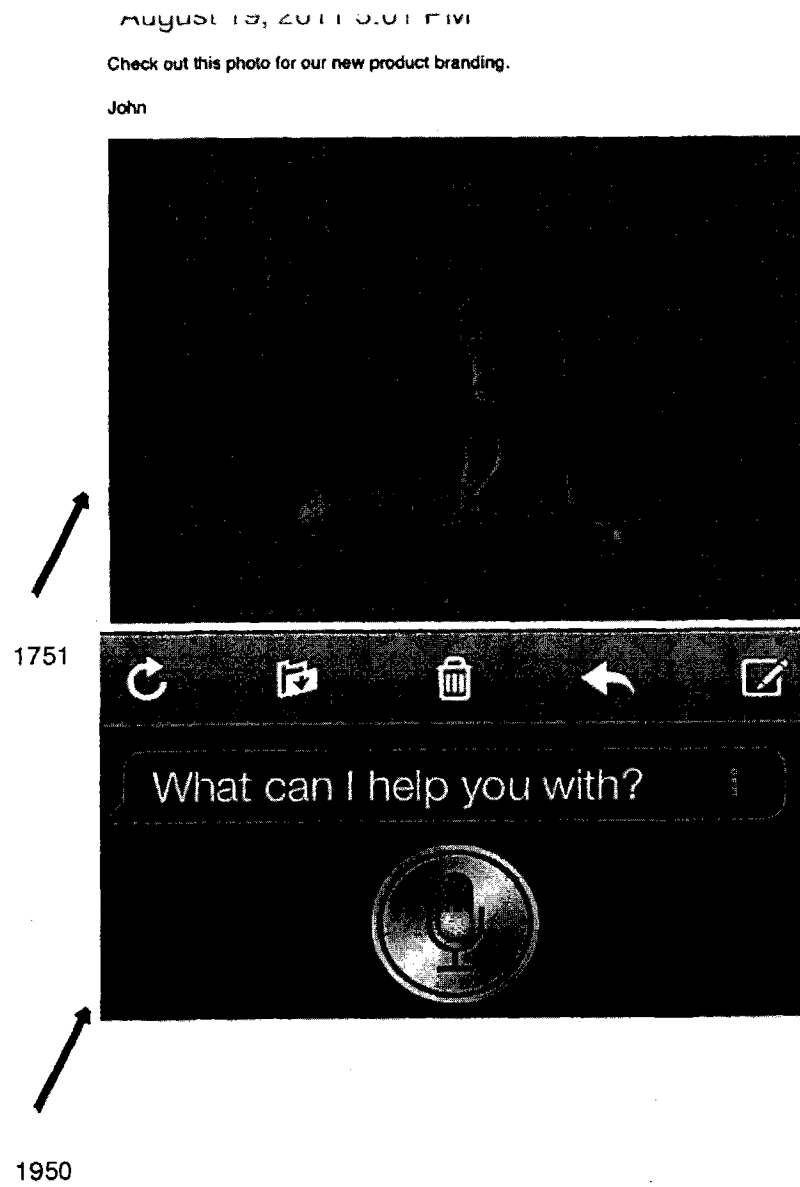


图19

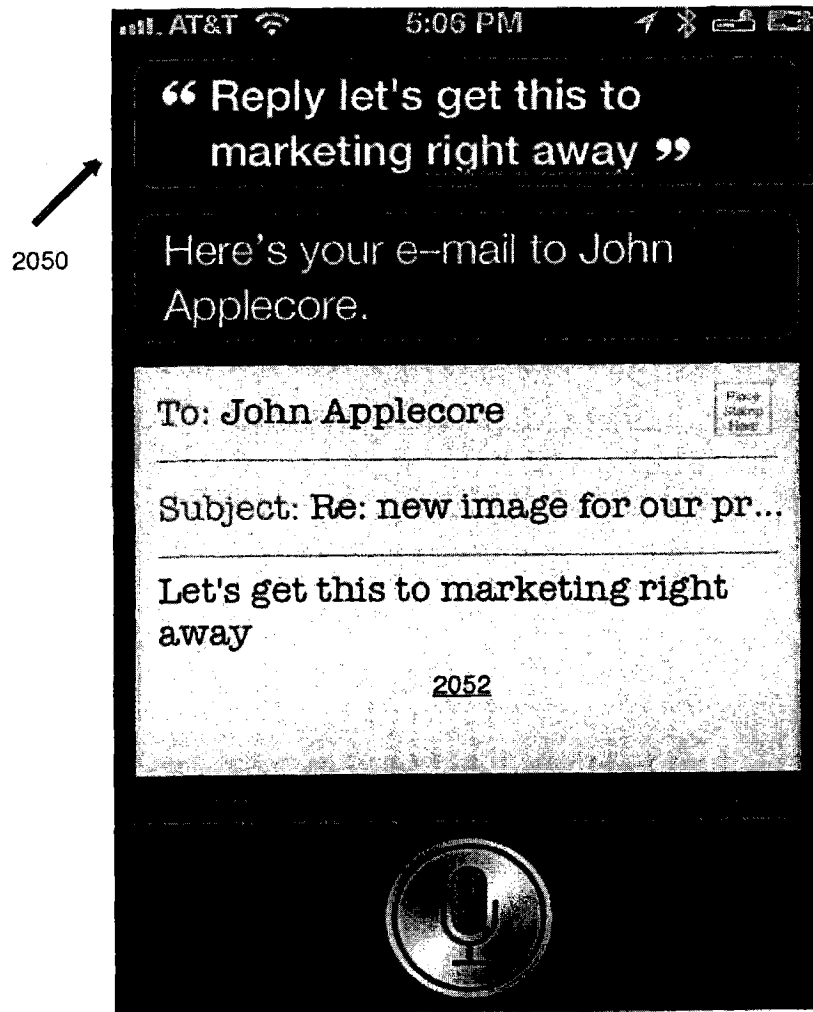


图20



图21



图22

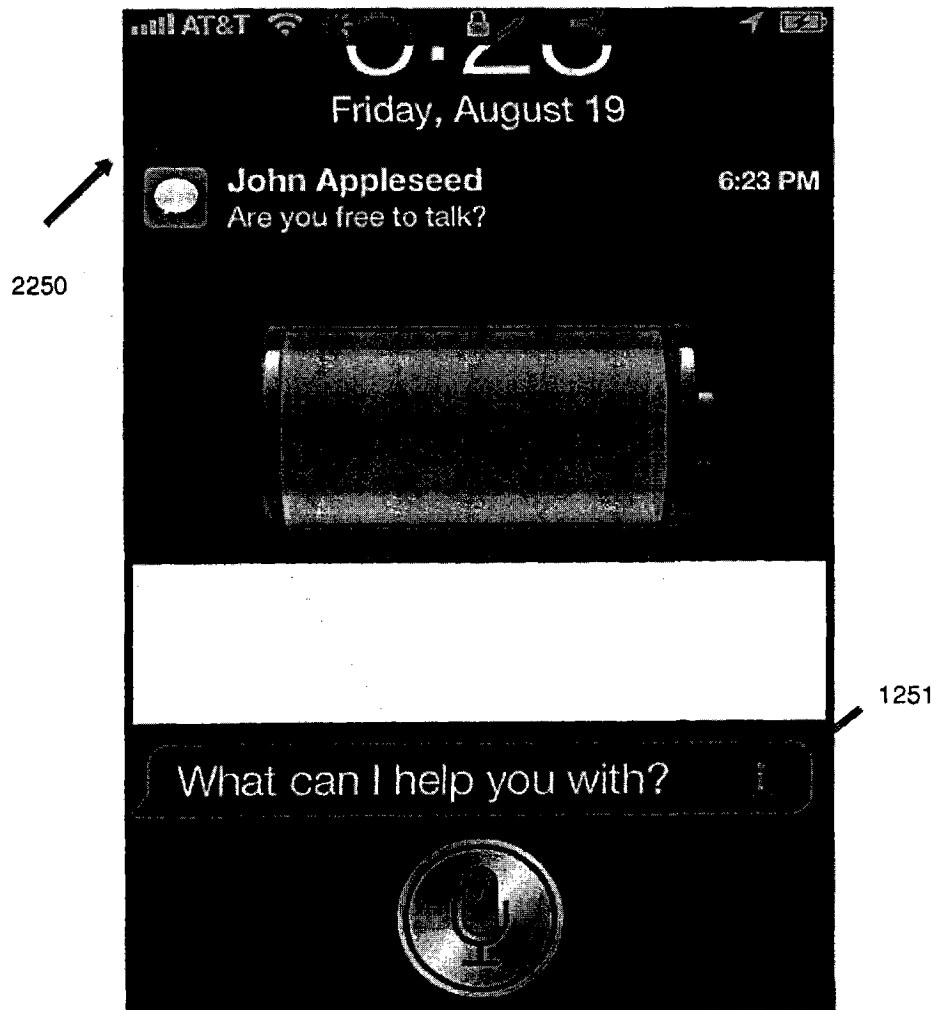


图23

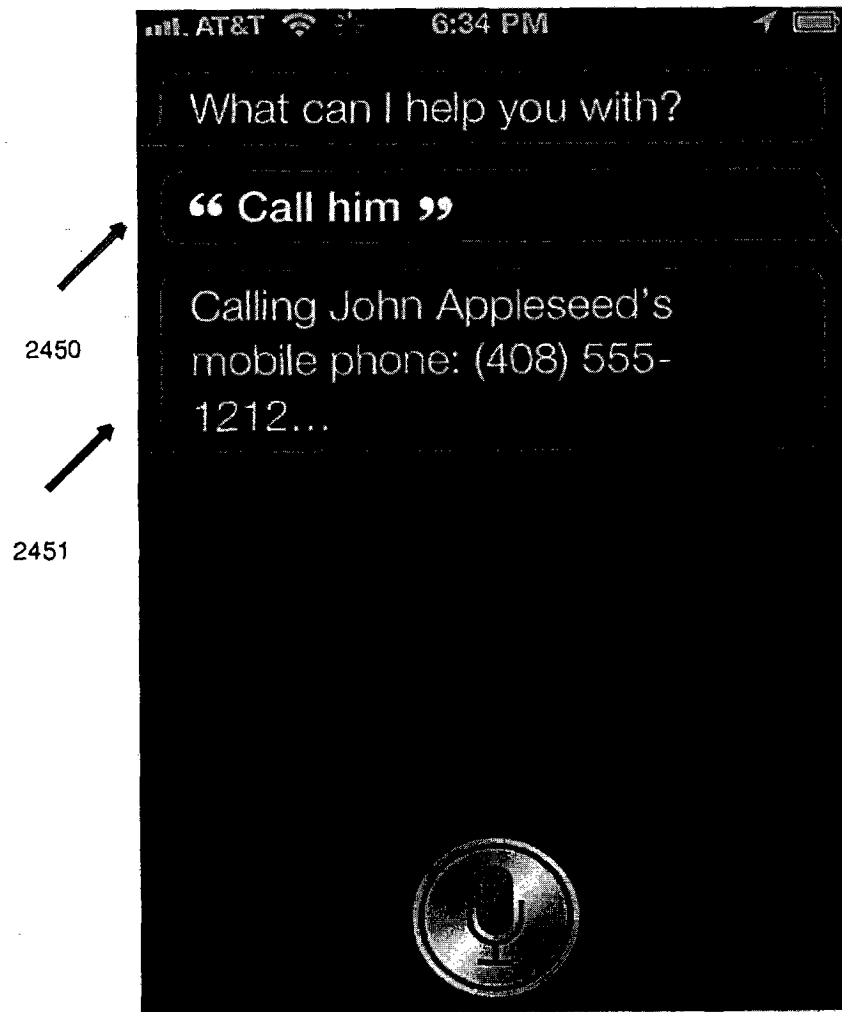


图24

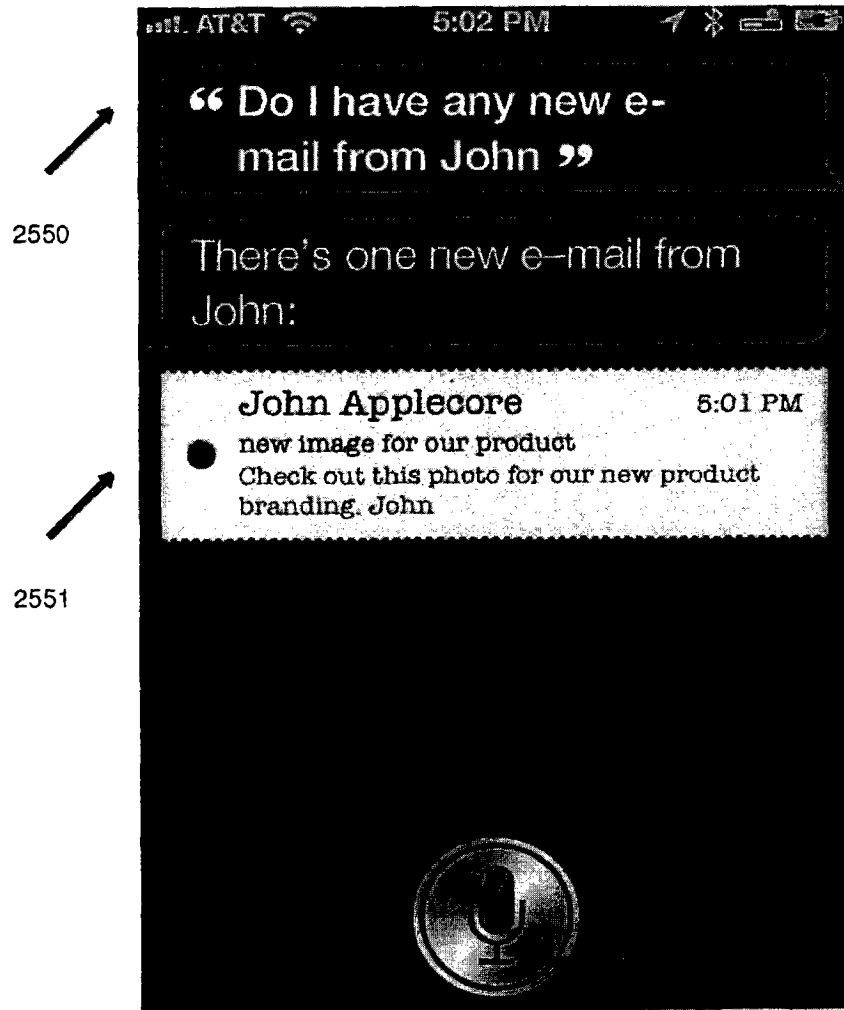


图25A



图25B



2651

图26A

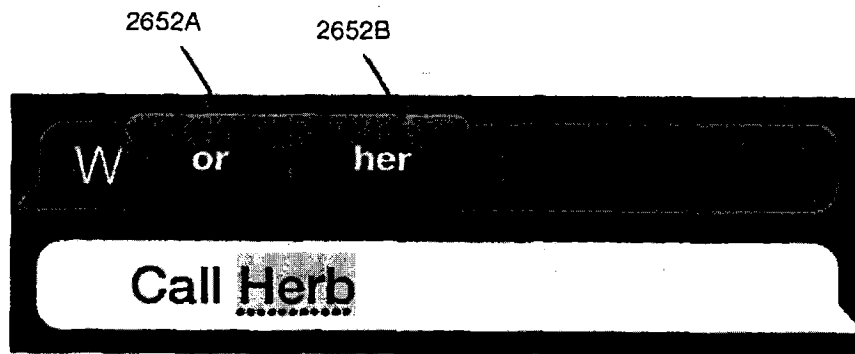


图26B

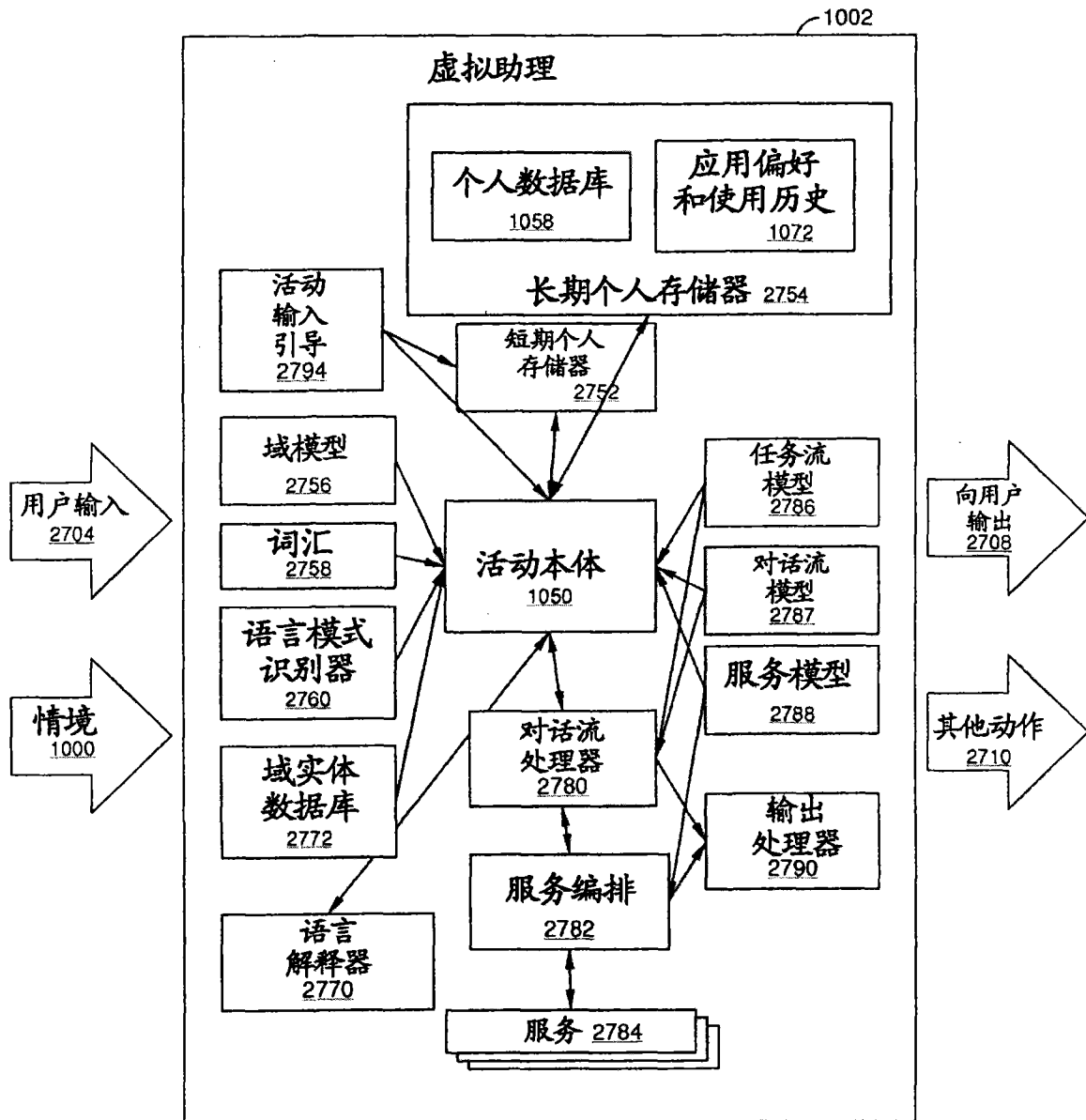


图27

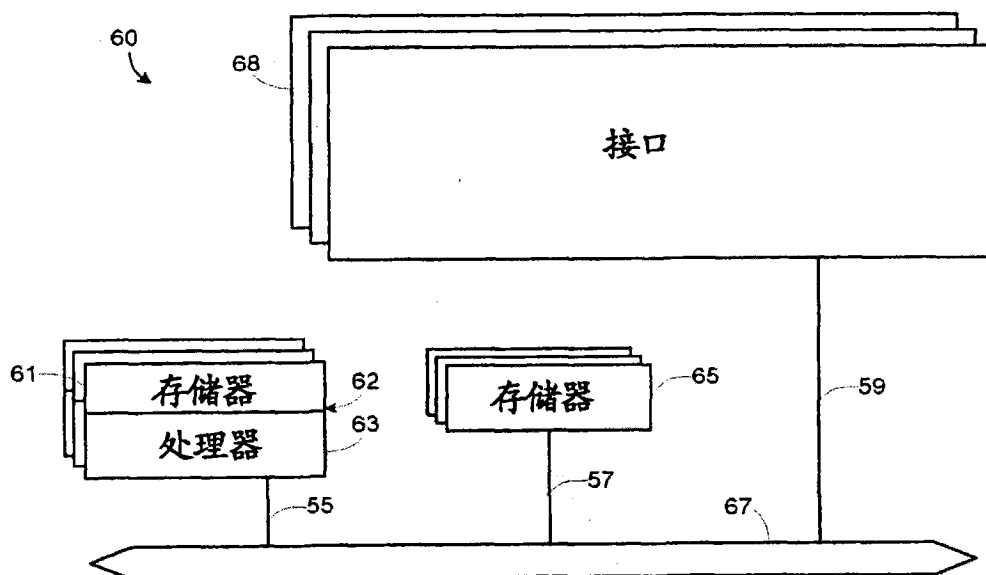


图28

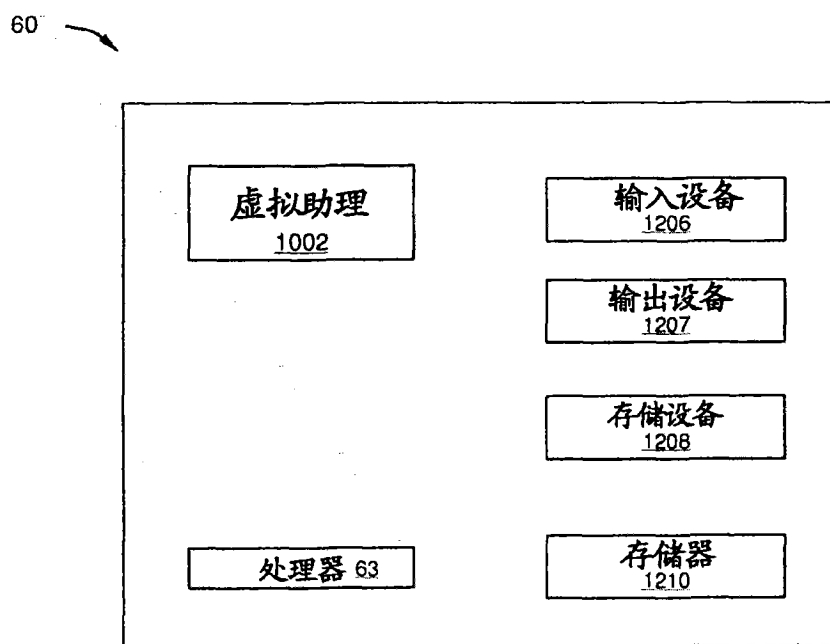


图29

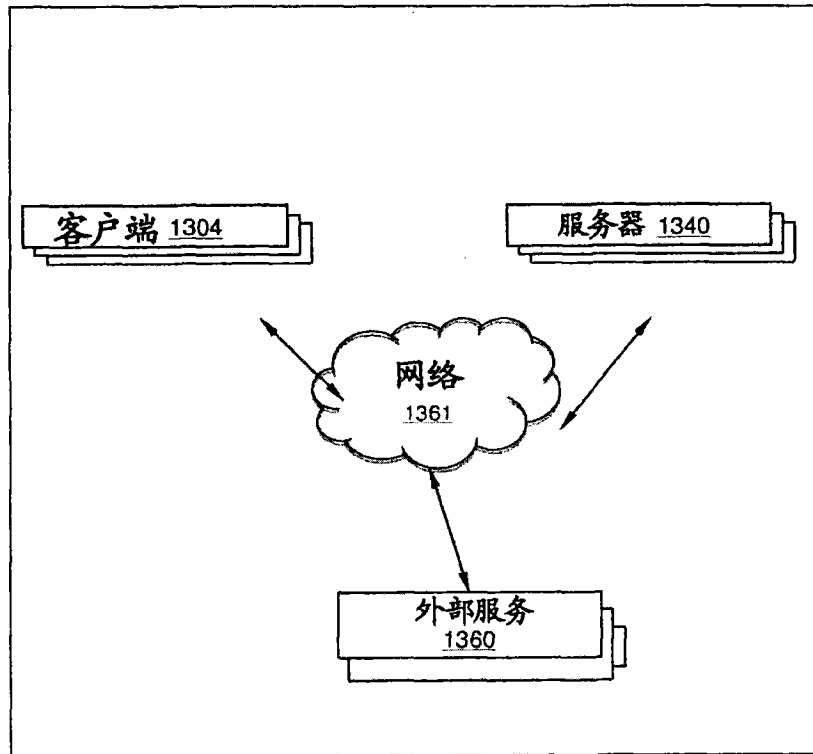


图30

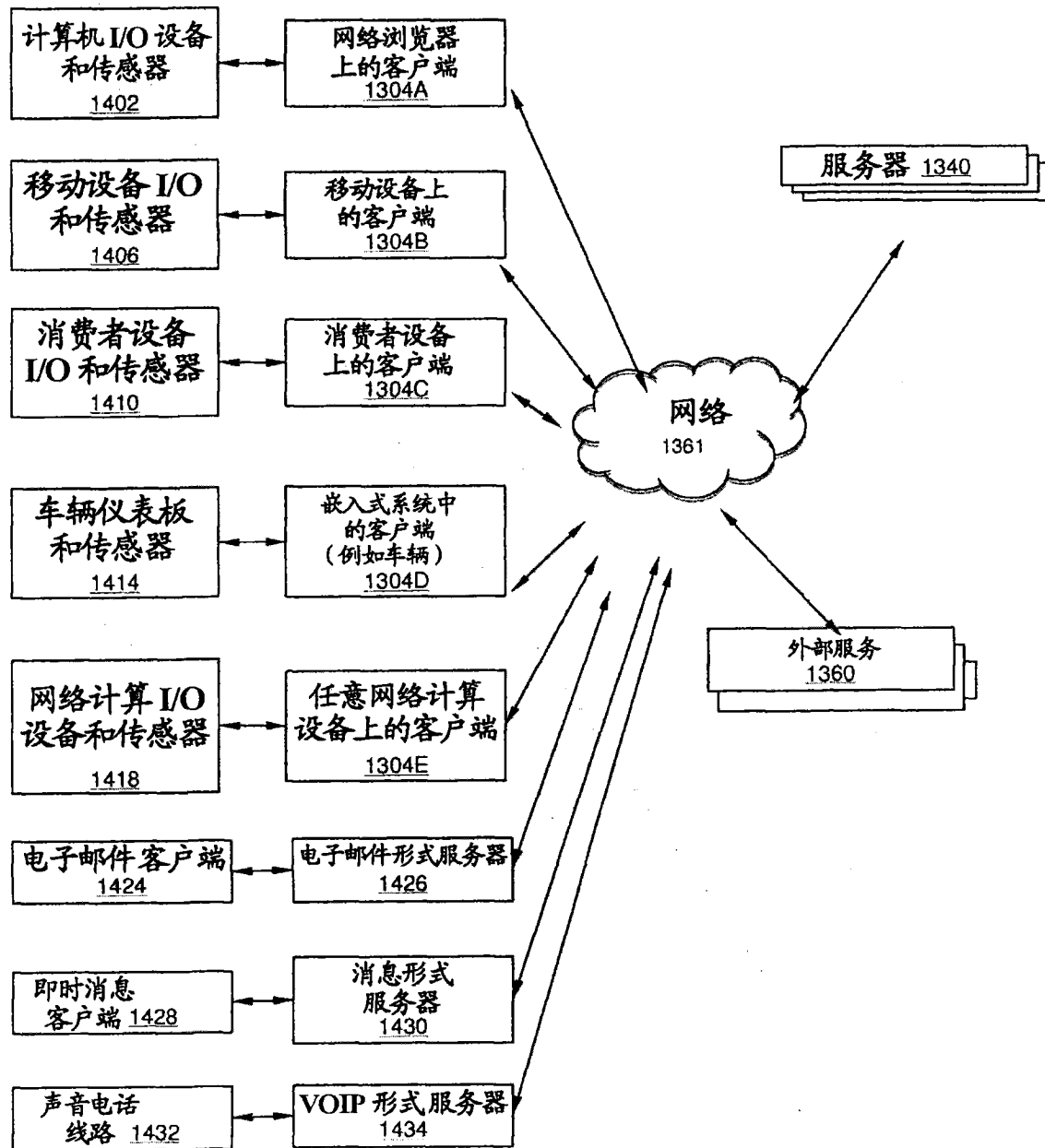


图31

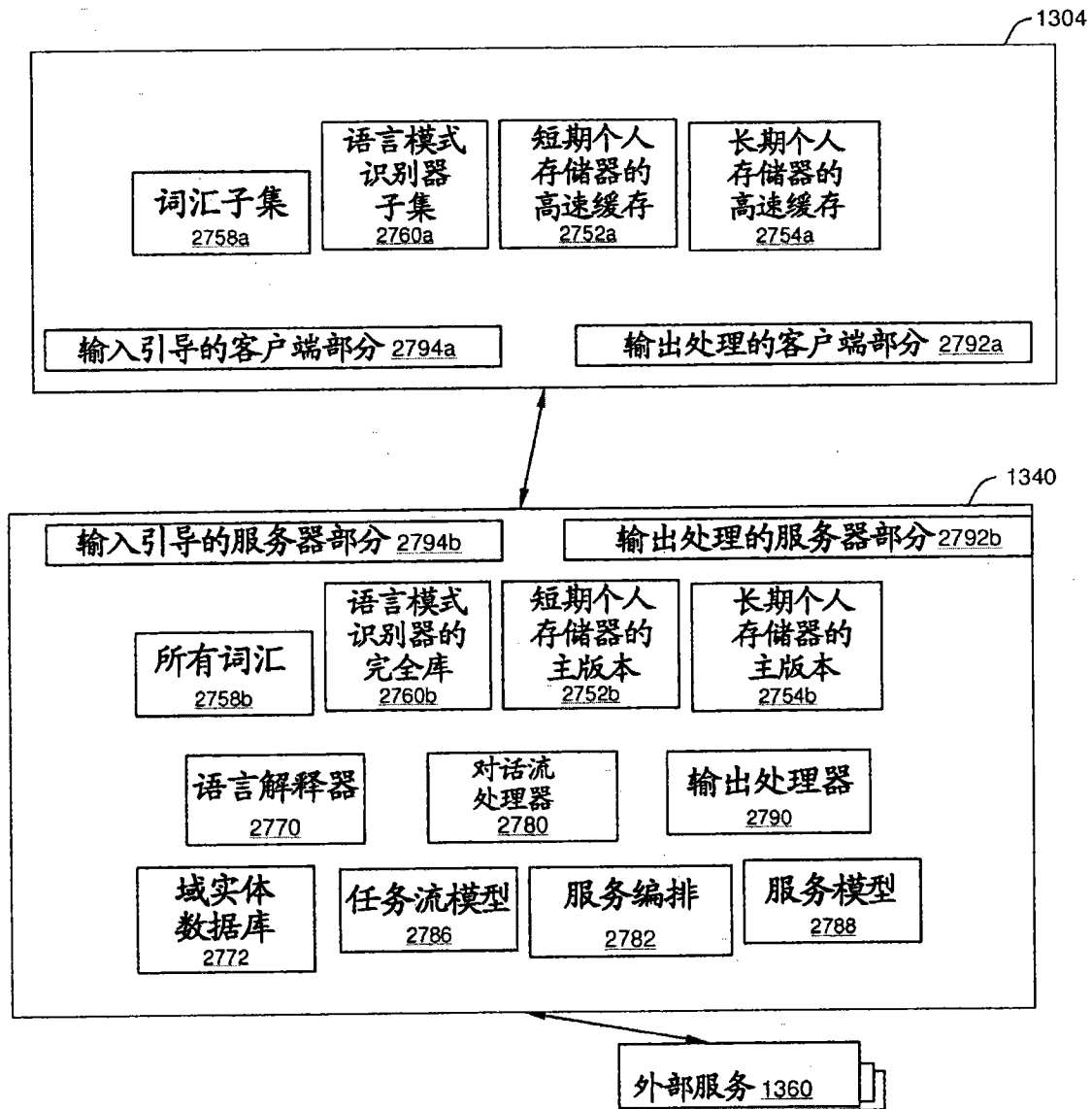


图32

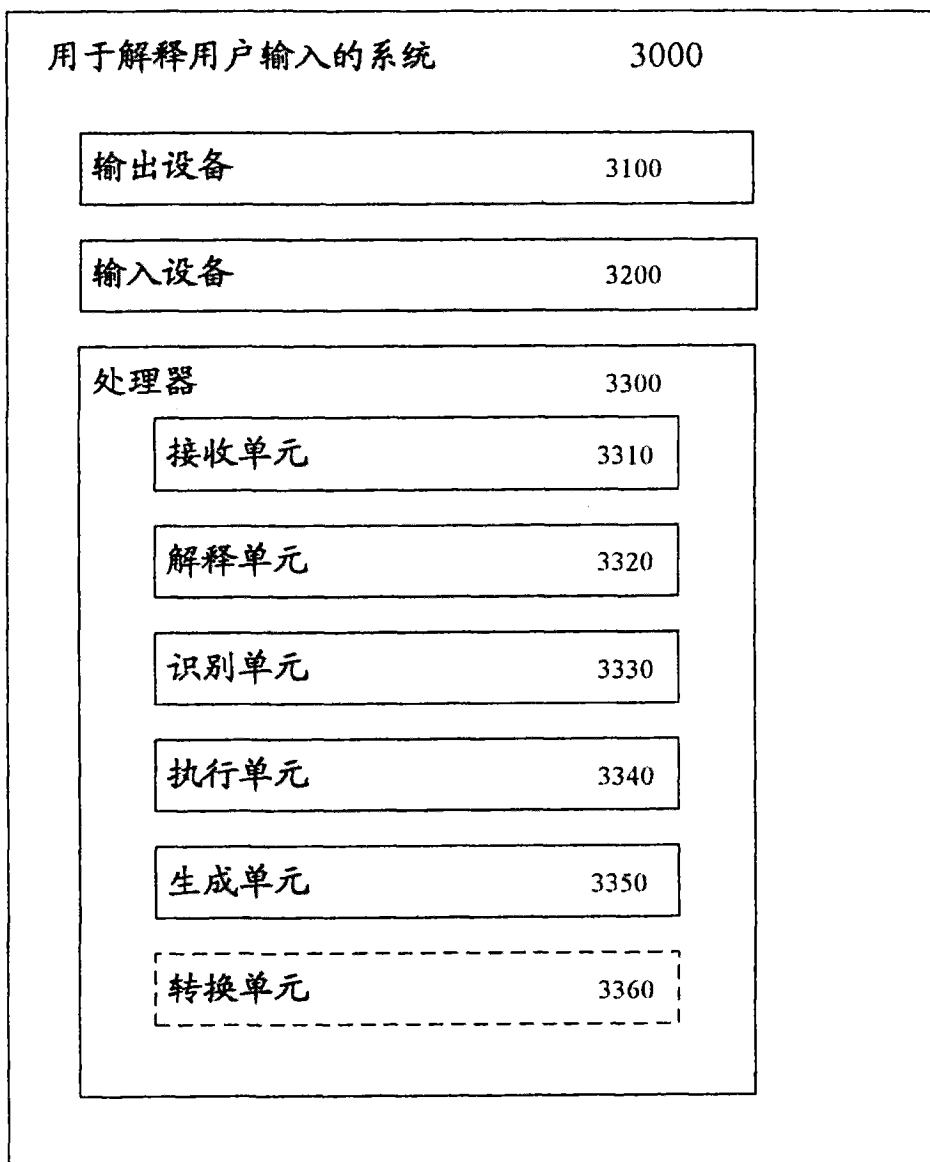


图33

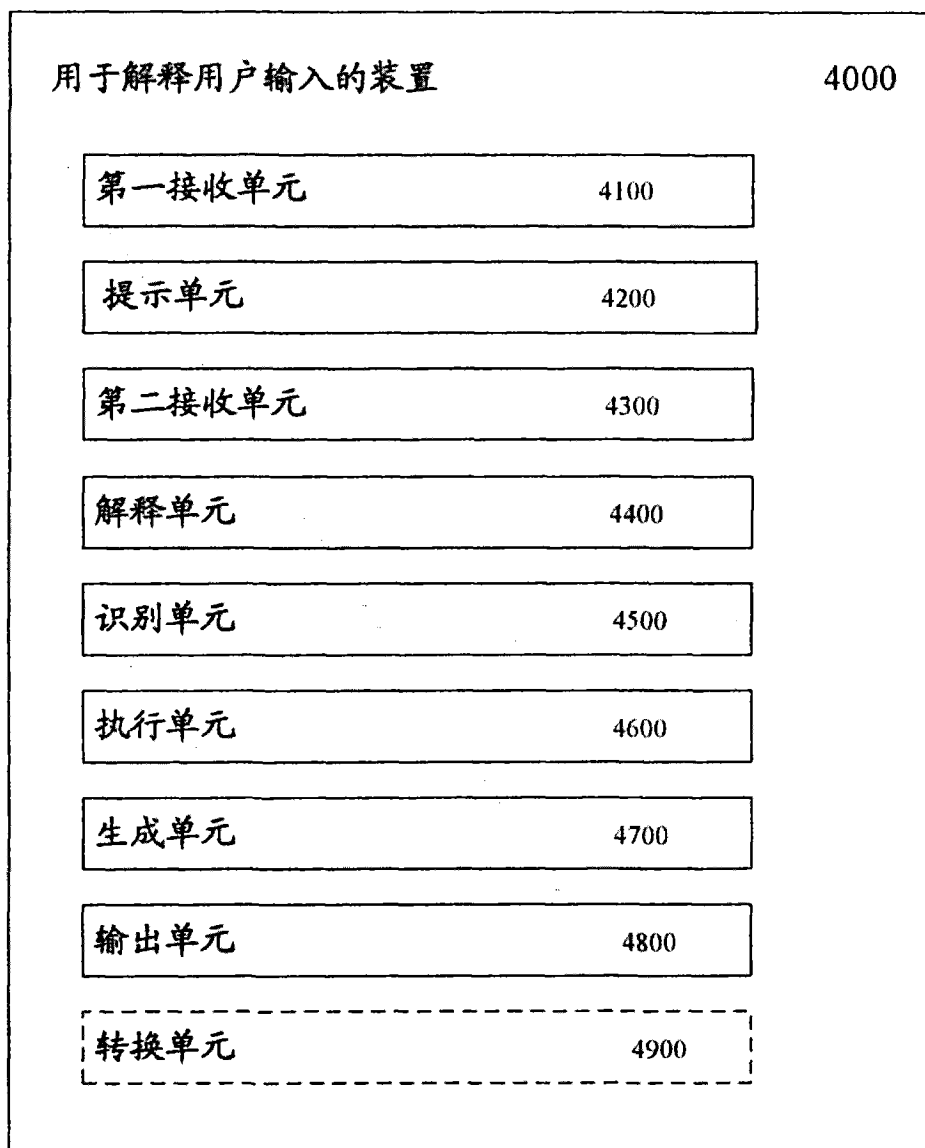


图34