



(12) 发明专利申请

(10) 申请公布号 CN 105320726 A

(43) 申请公布日 2016. 02. 10

(21) 申请号 201510289544. 9

(22) 申请日 2015. 05. 29

(30) 优先权数据

62/005, 760 2014. 05. 30 US

14/502, 737 2014. 09. 30 US

(71) 申请人 苹果公司

地址 美国加利福尼亚

(72) 发明人 P·P·皮埃尔诺特 J·G·宾德尔

(74) 专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 罗亚男

(51) Int. Cl.

G06F 17/30(2006. 01)

G10L 15/22(2006. 01)

权利要求书5页 说明书26页 附图4页

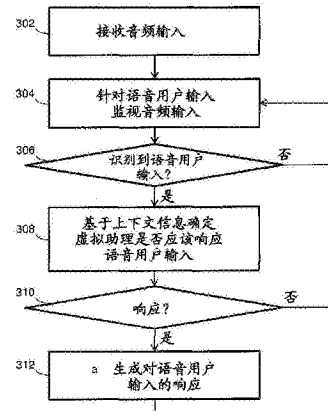
(54) 发明名称

降低对手动开始 / 结束点和触发短语的需求

(57) 摘要

本公开涉及降低对手动开始 / 结束点和触发短语的需求。提供了用于选择性处理和响应语音用户输入的系统 and 处理。在一个示例中, 能在用户设备处接收包含语音用户输入的音频输入。能通过识别语音用户输入的开始点和结束点来从音频输入识别语音用户输入。能基于上下文信息确定语音用户输入是否针对虚拟助理。能使用基于规则的系统或概率系统来进行确定。如果确定语音用户输入意图针对虚拟助理, 则能处理语音用户输入并且能生成合适的响应。相反如果确定语音用户输入不意图针对虚拟助理, 则能忽视语音用户输入和 / 或可不生成响应。

处理 300



1. 一种用于在电子设备上操作虚拟助理的方法,该方法包括:
 - 在所述电子设备处接收音频输入;
 - 监视所述音频输入以识别第一语音用户输入;
 - 识别所述音频输入中的所述第一语音用户输入;
 - 基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入;
 - 响应于确定响应所述第一语音用户输入:
 - 生成对所述第一语音用户输入的响应;和
 - 监视所述音频输入以识别第二语音用户输入;以及
 - 响应于确定不响应所述第一语音用户输入,监视所述音频输入以识别所述第二语音用户输入而不生成对所述第一语音用户输入的响应。
2. 根据权利要求 1 所述的方法,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入不包括识别所述第一语音用户输入的开始处的一个或多个预定词。
3. 根据权利要求 1 所述的方法,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入不包括识别在接收所述第一语音用户输入之前接收的物理或虚拟按钮输入。
4. 根据权利要求 1 所述的方法,其中生成对所述第一语音用户输入的响应包括以下中的一个或多个:
 - 对所述第一语音用户输入执行语音到文本转换;
 - 基于所述第一语音用户输入确定用户意图;
 - 基于所述第一语音用户输入确定要执行的任务;
 - 基于所述第一语音用户输入确定要执行的任务的参数;
 - 执行要执行的任务;
 - 显示对所述第一语音用户输入的文本响应;和
 - 输出对所述第一语音用户输入的音频响应。
5. 根据权利要求 1-3 中的任一项所述的方法,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入包括:
 - 评估取决于与所述第一语音用户输入相关联的所述上下文信息的一个或多个条件性规则。
6. 根据权利要求 1-3 中的任一项所述的方法,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入包括:
 - 基于与所述第一语音用户输入相关联的所述上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分;和
 - 将所述可能性得分与阈值进行比较。
7. 根据权利要求 6 所述的方法,其中所述上下文信息包括接收所述第一语音用户输入与之前用户输入之间经过的时间,其中计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:
 - 响应于所述经过的时间的值大于阈值持续时间而减小所述可能性得分;和

响应于所述经过的时间的值小于阈值持续时间而增大所述可能性得分。

8. 根据权利要求 6 所述的方法,其中所述上下文信息包括接收所述第一语音用户输入时用户与所述电子设备之间的距离,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述距离大于阈值距离而减小所述可能性得分;和

响应于所述距离小于阈值距离而增大所述可能性得分。

9. 根据权利要求 6 所述的方法,其中所述上下文信息包括接收所述第一语音用户输入时用户眼睛的方向,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于用户眼睛的方向指向所述电子设备而增大所述可能性得分;和

响应于用户眼睛的方向指离所述电子设备而减小所述可能性得分。

10. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入是否被自动语音识别器识别的指示,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述指示表示所述第一语音用户输入被自动语音识别器识别而增大所述可能性得分;和

响应于所述指示表示所述第一语音用户输入未被自动语音识别器识别而减小所述可能性得分。

11. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入与之前语音用户输入之间的语义关系,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述语义关系的值大于语音用户输入语义阈值而增大所述可能性得分;和

响应于所述语义关系的值小于语音用户输入语义阈值而减小所述可能性得分。

12. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入的长度,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述第一语音用户输入的长度小于阈值长度而增大所述可能性得分;和

响应于所述第一语音用户输入的长度大于阈值长度而减小所述可能性得分。

13. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入的发言者的标识。

14. 根据权利要求 6 所述的方法,其中所述上下文信息包括接收所述第一语音用户输入的时间,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述时间在预定时间集合内而增大所述可能性得分;和

响应于所述时间不在预定时间集合内而减小所述可能性得分。

15. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入与接收所述第一语音用户输入时所述电子设备显示的内容之间的语义关系,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述语义关系的值大于内容语义阈值而增大所述可能性得分;和

响应于所述语义关系的值小于内容语义阈值而减小所述可能性得分。

16. 根据权利要求 6 所述的方法,其中所述上下文信息包括接收所述第一语音用户输入时用户执行的手势,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述手势是预定手势集合中的一个而增大所述可能性得分;和

响应于所述手势不是预定手势集合中的一个而减小所述可能性得分。

17. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入与所述电子设备的之前输出之间的语义关系,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述语义关系的值大于之前输出语义阈值而增大所述可能性得分;和

响应于所述语义关系的值小于之前输出语义阈值而减小所述可能性得分。

18. 根据权利要求 6 所述的方法,其中所述上下文信息包括所述第一语音用户输入与之前电子邮件之间的语义关系,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述语义关系的值大于之前电子邮件语义阈值而增大所述可能性得分;

响应于所述语义关系的值小于之前电子邮件语义阈值而减小所述可能性得分。

19. 一种用于在电子设备上操作虚拟助理的装置,包括用于实现权利要求 1-18 中任一项所述的方法的步骤的部件。

20. 一种电子设备,所述电子设备包括:

音频接收单元 (502),配置为接收音频输入;

处理单元 (508),其耦接到音频接收单元 (502) 并且被配置为:

监视所述音频输入以识别第一语音用户输入;

识别所述音频输入中的所述第一语音用户输入;

基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入;

响应于确定响应所述第一语音用户输入:

生成对所述第一语音用户输入的响应,和

监视所述音频输入以识别第二语音用户输入;以及

响应于确定不响应所述第一语音用户输入,监视所述音频输入以识别所述第二语音用户输入而不生成对所述第一语音用户输入的响应。

21. 根据权利要求 20 所述的电子设备,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入不包括识别所述第一语音用户输入的开始处的一个或多个预定词。

22. 根据权利要求 20 所述的电子设备,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入不包括识别在接收所述第一语音用户输入之前接收的物理或虚拟按钮输入。

23. 根据权利要求 20 所述的电子设备,其中生成对所述第一语音用户输入的响应包括以下中的一个或多个:

对所述第一语音用户输入执行语音到文本转换;

基于所述第一语音用户输入确定用户意图；
基于所述第一语音用户输入确定要执行的任务；
基于所述第一语音用户输入确定要执行的任务的参数；
执行要执行的任务；
显示对所述第一语音用户输入的文本响应；和
输出对所述第一语音用户输入的音频响应。

24. 根据权利要求 20-22 中的任一项所述的电子设备,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入包括:

评估取决于与所述第一语音用户输入相关联的所述上下文信息的一个或多个条件性规则。

25. 根据权利要求 20-22 中的任一项所述的电子设备,其中基于与所述第一语音用户输入相关联的上下文信息确定是否响应所述第一语音用户输入包括:

基于与所述第一语音用户输入相关联的所述上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分;和

将所述可能性得分与阈值进行比较。

26. 根据权利要求 25 所述的电子设备,其中所述上下文信息包括接收所述第一语音用户输入与之前用户输入之间经过的时间,其中计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述经过的时间的值大于阈值持续时间而减小所述可能性得分;和

响应于所述经过的时间的值小于阈值持续时间而增大所述可能性得分。

27. 根据权利要求 25 所述的电子设备,其中所述上下文信息包括接收所述第一语音用户输入时用户与所述电子设备之间的距离,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述距离大于阈值距离而减小所述可能性得分;和

响应于所述距离小于阈值距离而增大所述可能性得分。

28. 根据权利要求 25 所述的电子设备,其中所述上下文信息包括接收所述第一语音用户输入时用户眼睛的方向,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于用户眼睛的方向指向所述电子设备而增大所述可能性得分;和

响应于用户眼睛的方向指离所述电子设备而减小所述可能性得分。

29. 根据权利要求 25 所述的电子设备,其中所述上下文信息包括所述第一语音用户输入是否被自动语音识别器识别的指示,并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括:

响应于所述指示表示所述第一语音用户输入被自动语音识别器识别而增大所述可能性得分;和

响应于所述指示表示所述第一语音用户输入未被自动语音识别器识别而减小所述可能性得分。

30. 根据权利要求 25 所述的电子设备,其中所述上下文信息包括所述第一语音用户输入与之前语音用户输入之间的语义关系,并且其中基于上下文信息计算所述虚拟助理应该

响应所述第一语音用户输入的可能性得分包括：

响应于所述语义关系的值大于语音用户输入语义阈值而增大所述可能性得分；和
响应于所述语义关系的值小于语音用户输入语义阈值而减小所述可能性得分。

31. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括所述第一语音用户输入的长度，并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括：

响应于所述第一语音用户输入的长度小于阈值长度而增大所述可能性得分；和
响应于所述第一语音用户输入的长度大于阈值长度而减小所述可能性得分。

32. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括所述第一语音用户输入的发言者的标识。

33. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括接收所述第一语音用户输入的时间，并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括：

响应于所述时间在预定时间集合内而增大所述可能性得分；和
响应于所述时间不在预定时间集合内而减小所述可能性得分。

34. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括所述第一语音用户输入与接收所述第一语音用户输入时所述电子设备显示的内容之间的语义关系，并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括：

响应于所述语义关系的值大于内容语义阈值而增大所述可能性得分；和
响应于所述语义关系的值小于内容语义阈值而减小所述可能性得分。

35. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括接收所述第一语音用户输入时用户执行的手势，并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括：

响应于所述手势是预定手势集合中的一个而增大所述可能性得分；和
响应于所述手势不是预定手势集合中的一个而减小所述可能性得分。

36. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括所述第一语音用户输入与所述电子设备的之前输出之间的语义关系，并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括：

响应于所述语义关系的值大于之前输出语义阈值而增大所述可能性得分；和
响应于所述语义关系的值小于之前输出语义阈值而减小所述可能性得分。

37. 根据权利要求 25 所述的电子设备，其中所述上下文信息包括所述第一语音用户输入与之前电子邮件之间的语义关系，并且其中基于上下文信息计算所述虚拟助理应该响应所述第一语音用户输入的可能性得分包括：

响应于所述语义关系的值大于之前电子邮件语义阈值而增大所述可能性得分；
响应于所述语义关系的值小于之前电子邮件语义阈值而减小所述可能性得分。

降低对手动开始 / 结束点和触发短语的需求

技术领域

[0001] 这总体涉及语音处理,更具体地,涉及使用上下文数据选择性处理语音用户输入。

背景技术

[0002] 智能自动助理(或虚拟助理)提供用户与电子之间的直观交互。这些助理能允许用户以语音和/或文本形式使用自然语言与设备或系统交互。例如,用户能通过向与电子设备相关联的虚拟助理提供语音用户输入来访问电子设备的服务。虚拟助理能从语音用户输入解读用户的意图,并且将用户的意图实施为任务。然后可通过执行电子设备的一个或多个功能执行任务并且可将相关输出以自然语音形式返回用户。

[0003] 为了使虚拟助理适当处理并响应语音用户输入,虚拟助理能首先使用一般称为开始点和结束点的处理分别识别音频输入流内的语音用户输入的开始和结束。传统的虚拟助理能基于接收的音频流的能量等级和/或声学特性或者用户的手动识别来识别这些点。例如,一些虚拟助理能需要用户通过在向虚拟助理说话之前按下物理或虚拟按钮或者通过在向虚拟助理说话之前以自然语言形式说出特定触发短语来输入开始点标识符。响应于接收到这些开始点标识符中的一个,虚拟助理能将随后接收的音频解读为语音用户输入。尽管能使用这些技术来清楚地识别针对虚拟助理的语音用户输入,但以此方式与虚拟助理交互对于用户可能是不自然或困难的。例如,在虚拟助理与用户之间的来回对话中,可能需要用户在各个语音用户输入之前输入开始点标识符(例如按下按钮或重复相同触发短语)。

发明内容

[0004] 公开了用于操作虚拟助理的系统和处理。一个示例处理能包括在电子设备处接收音频输入,监视音频输入以识别第一语音用户输入,识别音频输入中的第一语音用户输入,以及基于与第一语音用户输入相关联的上下文信息确定是否响应第一语音用户输入。处理还能包括,响应于确定响应第一语音用户输入:生成对第一语音用户输入的响应;以及监视音频输入以识别第二语音用户输入。处理还能包括,响应于确定不响应第一语音用户输入,监视音频输入以识别第二语音用户输入而不生成对第一语音用户输入的响应。

附图说明

[0005] 图 1 说明能根据各种示例操作的虚拟助理的示例性环境。

[0006] 图 2 说明根据各种示例的示例性用户设备。

[0007] 图 3 说明根据各种示例的用于使用上下文信息选择性处理和响应语音用户输入的示例性处理。

[0008] 图 4 说明根据各种示例的用于使用上下文信息选择性处理和响应语音用户输入的另一示例性处理。

[0009] 图 5 说明根据各种示例的配置为使用上下文信息选择性处理和响应语音用户输入的电子设备的功能框图。

具体实施方式

[0010] 在下面的示例描述中,参考通过说明方式示出可以实现的特定示例的附图。要理解,在不脱离各种示例的范围的情况下,可以使用其它示例,并且可以进行结构改变。

[0011] 这涉及用于选择性地处理并响应语音用户输入的系统和处理。在一个示例处理中,可以在用户设备处接收包括语音用户输入的音频输入。能通过识别语音用户输入的的开始点和结束点来识别语音用户输入。可以基于上下文信息确定语音用户输入是否针对用户设备上运行的虚拟助理并且虚拟助理是否应该响应语音用户输入。能使用基于规则的系统或概率(例如机器学习)系统进行确定。如果确定语音用户输入针对虚拟助理并且虚拟助理应该响应语音用户输入,则可处理语音用户输入并且能生成合适的响应。如果确定语音用户输入不针对虚拟助理,则能忽略语音用户输入和/或不会生成响应。使用上下文信息确定语音用户输入是否针对虚拟助理能有利地允许用户与虚拟助理交互而无需在各个语音用户输入之前手动识别开始点(例如,通过按下按钮或说出触发短语)。

[0012] 系统概述

[0013] 图1说明根据各种示例的用于实现虚拟助理的示例性系统100。术语“虚拟助理”、“数字助理”、“智能自动助理”或“自动数字助理”可以指解释语音和/或文字形式的自然语言输入以推断用户意图并基于推断的用户意图执行动作的任何信息处理系统。例如,为了按照推断的用户意图行动,系统能执行下述中的一个或多个:识别具有设计为实现推断的用户意图的步骤和参数的任务流;输入从推断的用户意图到任务流的特定要求;通过调用程序、方法、服务、API等执行任务流;生成可听(例如语音)和/或可视形式的对用户的输出响应。

[0014] 虚拟助理可以能够接受至少部分处于自然语言命令、请求、陈述、叙述和/或查询形式的用户请求。典型地,用户请求寻求或者信息性答案或者虚拟助理执行任务。对用户请求的满意响应能包括或者提供请求的信息性答案、请求的任务的执行,或者二者的组合。例如,用户能问虚拟助理问题,例如“我现在在哪里?”基于用户的当前位置,虚拟助理能回答“你在中央公园。”用户也能请求执行任务,例如,“请提醒我今天下午4点给妈妈打电话。”在响应中,虚拟助理能确认请求然后在用户的电子日程中创建合适的提醒项。在执行请求的任务期间,虚拟助理有时能通过扩展时段在涉及信息的多次交换的连续对话中与用户交互。存在多种其它方式用于与虚拟助理交互以请求信息或各种任务的执行。除了提供言语响应和采取编程动作,虚拟助理还能以其它视觉或听觉形式(例如作为文本、警告、音乐、视频、动画等)可能使用多个设备提供响应(例如,经由头戴耳机输出文本到语音并且在电视上显示文本)。

[0015] 申请人2011年1月10日提交的美国实用新型申请序列号12/987,982“Intelligent Automated Assistant”中描述了虚拟助理的示例,上述申请的全部公开通过引用在此并入。

[0016] 如图1所示,在一些示例中,能根据客户端-服务器模型实现虚拟助理。虚拟助理能包括在用户设备102上运行的客户端方部分和在服务器系统100上运行的服务器方部分。用户设备102能包括任何电子设备,例如移动电话、平板电脑、便携式媒体播放器、台式计算机、膝上型计算机、PDA、电视、电视机顶盒、可穿戴电子设备等,并且能通过一个或多个

网络 108 与服务器系统 110 通信,网络 108 能包括互联网、内联网或任何其它有线或无线公共或私有网络。在用户设备 102 上运行的客户端方部分能提供客户端方功能,例如面向用户输入和输出处理以及与服务器系统 110 通信。服务器系统 110 能为驻留在相应用户设备 102 上的任意数量的客户端提供服务器方功能。

[0017] 服务器系统 110 能包括一个或多个虚拟助理服务器 114,虚拟助理服务器 114 能包括面向客户端的 I/O 接口 122,一个或多个处理模块 118,数据和模型存储器 120,以及到外部服务的 I/O 接口 116。面向客户端的 I/O 接口 122 能便利用于虚拟助理服务器 114 的面向客户端的输入和输出处理。一个或多个处理模块 118 能利用数据和模型存储器 120 来基于自然语言输入确定用户的意图并且基于推断的用户意图执行任务运行。在一些示例中,虚拟助理服务器 114 能通过网络 108 与诸如电话服务、日程服务、信息服务、消息服务、导航服务等的外部服务 124 通信,用于任务完成或信息获取。到外部服务的 I/O 接口 116 能便利这种通信。

[0018] 服务器系统 110 能在一个或多个独立数据处理设备或分布式计算机网络上实现。在一些示例中,服务器系统 110 能利用各种虚拟设备和 / 或第三方服务提供商 (例如第三方云服务提供商) 的服务来提供服务器系统 110 的底层计算资源和 / 或架构资源。

[0019] 尽管图 1 示出虚拟助理的功能包括客户端方部分和服务器方部分,但在一些示例中,助理的功能可实现为安装在用户设备上的独立应用。此外,虚拟助理的客户端和服务器部分之间的功能划分在不同示例中可变化。例如,在一些示例中,在用户设备 102 上运行的客户端可以是仅提供面向用户输入和输出处理功能并将虚拟助理的所有其它功能委托给后端服务器的精简客户端。

[0020] 用户设备

[0021] 图 2 是根据各种示例的用户设备 102 的框图。如图所示,用户设备 102 能包括存储器接口 202,一个或多个处理器 204,以及外围接口 206。用户设备 102 中的各种组件可通过一个或多个通信总线或信号线耦接在一起。用户设备 102 还能包括各种传感器、子系统、以及连接到外围接口 206 的外设。传感器、子系统、以及外设收集信息和 / 或便利用户设备 102 的各种功能。

[0022] 例如,用户设备 102 能包括耦接到外围接口 206 以便利方向、光和接近感测功能的运动传感器 210、光传感器 212 和接近传感器 214。一个或多个其它传感器 216,例如定位系统 (例如 GPS 接收器)、温度传感器、生物传感器、陀螺仪、罗盘、加速计等,也可连接到外围接口 206,以便利相关功能。

[0023] 在一些示例中,能利用相机子系统 220 和光学传感器 222 便利相机功能,例如拍照和记录视频剪辑。能通过一个或多个有线和 / 或无线通信子系统 224 便利通信功能,有线和 / 或无线通信子系统 224 能包括各种通信端口、射频接收器和发送器和 / 或光学 (例如红外) 接收器和发送器。音频子系统 226 能耦接到扬声器 228 和麦克风 230 以便利语音启用功能,例如语音识别、语音复制、数字记录和电话功能。

[0024] 在一些示例中,用户设备 102 还能包括耦接到外围接口 206 的 I/O 子系统 240。I/O 子系统 240 能包括触摸屏控制器和 / 或其它输入控制器 244。触摸屏控制器 244 能耦接到触摸屏 246。触摸屏 246 和触摸屏控制器 244 例如能使用多种触摸感测技术 (例如电容、电阻、红外、表面声波技术、接近传感器阵列等) 中的一种检测接触和运动或其间断。其它

输入控制器 244 能耦接到其它输入 / 控制设备 248, 例如一个或多个按钮、摇臂开关、指轮、红外端口、USB 端口和 / 或指点设备 (例如触摸笔)。

[0025] 在一些示例中, 用户设备 102 还能包括耦接到存储器 250 的存储器接口 202。存储器 250 能包括任意电子的、磁的、光学的、电磁的、红外的或半导体的系统、装置或设备、便携式计算机盘 (磁的)、随机访问存储器 (RAM) (磁的)、只读存储器 (ROM 的) (磁)、可擦除可编程只读存储器 (EPROM) (磁的)、便携式光盘 (例如 CD, CD-R, CD-RW, DVD, DVD-R 或 DVD-RW)、或闪存 (例如致密闪存卡)、安全数字卡、USB 存储器设备、存储棒等。在一些示例中, 能使用存储器 250 的非瞬时性计算机可读存储介质存储由指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合使用的指令 (例如用于执行处理 300 和 / 或 400, 下面描述), 指令执行系统、装置或设备例如是基于计算机的系统, 包含处理器的系统, 或能从指令执行系统、装置或设备取出指令并执行指令的其它系统。在其它示例中, 指令 (例如用于执行处理 300 和 / 或 400, 下面描述) 能存储在服务器系统 110 的非瞬时性计算机可读存储介质上, 或能在存储器 250 的非瞬时性计算机可读存储介质和服务器系统 110 的非瞬时性计算机可读存储介质之间划分。在本文档的背景下, “非瞬时性计算机可读存储介质” 可以是能包含或存储由指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合使用的程序的任何介质。

[0026] 在一些示例中, 存储器 250 能存储操作系统 252, 通信模块 254, 图形用户接口模块 256, 传感器处理模块 258, 电话模块 260, 和应用模块 262。操作系统 252 能包括用于处理基本系统服务和用于执行硬件相关任务的指令。通信模块 254 能便利与一个或多个附加设备、一个或多个计算机和 / 或一个或多个服务器的通信。图形用户接口模块 256 能便利图形用户接口处理。传感器处理模块 258 能便利传感器相关处理和功能。电话模块 260 能便利电话相关处理和功能。应用模块 262 能便利用户应用的各种功能, 例如电子消息收发、网页浏览、媒体处理、导航、成像和 / 或其它处理和功能。

[0027] 如这里描述的, 存储器 250 也能存储客户端方虚拟助理指令 (例如, 在虚拟助理客户端模块 264 中) 和各种用户数据 266 (例如, 用户专用词汇数据, 喜好数据和 / 或其它数据, 例如客户的电子地址书、待办事项列表、购物列表等), 以提供虚拟助理的客户端方功能。

[0028] 在各种示例中, 虚拟助理客户端模块 264 能够通过用户设备 104 的各种用户接口 (例如, I/O 子系统 240、音频子系统 226 等) 接受语音输入 (例如, 言语输入)、文本输入、触摸输入和 / 或姿势输入。虚拟助理客户端模块 264 还能够提供听觉 (例如, 言语输出)、视觉和 / 或触觉形式的输出。例如, 输出可提供为语音、声音、警告、文本消息、菜单、图形、视频、动画、振动和 / 或以上两种或更多种的组合。在操作期间, 虚拟助理客户端模块 264 能使用通信子系统 224 与虚拟助理服务器通信。另外, 虚拟助理客户端模块 264 能与诸如家庭自动设备的其它设备通信, 并且能因此对物理世界具有物理效果 (例如, 打开门) 或者能嵌入在这种设备中。

[0029] 在一些示例中, 虚拟助理客户端模块 264 能利用各种传感器、子系统和外设来从用户设备 102 的周围环境收集附加信息, 以建立与用户、当前用户交互和 / 或当前用户输入相关联的上下文。在一些示例中, 虚拟助理客户端模块 264 能利用用户输入向虚拟助理服务器提供上下文信息或其子集以帮助推断用户的意图。虚拟助理还能使用上下文信息来确

定如何准备输出并且向用户传递输出。如以下更详细讨论的,上下文信息能进一步被用户设备 102 或服务器系统 110 使用来确定语音用户输入是否针对虚拟助理并且确定合适的响应。

[0030] 在一些示例中,伴随用户输入的上下文信息能包括传感器信息,例如照明、环境噪声、环境温度、周围环境的图像或视频、与另一对象的距离等。上下文信息还能包括与用户设备 102 的物理状态相关联的信息(例如设备方向、设备位置、设备温度、功率等级、速度、加速度、运动模式、蜂窝信号强度等)或用户设备 102 的软件状态(例如运行的进程、安装的程序、过去和当前的网络活动、背景服务、错误日志、资源使用、最前面的应用等)。这些类型的上下文信息中的任意可被提供到虚拟助理服务器作为与用户输入相关联的上下文信息。另外,上下文信息还能包括生物用户数据,例如心率、手温、语音质量、面部表情等。

[0031] 在一些示例中,虚拟助理客户端模块 264 能响应于来自虚拟助理服务器的请求选择性地提供存储在用户设备 102 上的信息(例如,用户数据 266)。虚拟助理客户端模块 264 还能在虚拟助理服务器 114 请求时经由自然语言对话或其他用户接口来引出来自用户的附加输入。虚拟助理客户端模块 264 能将附加输入传送到虚拟助理服务器 114 以帮助虚拟助理服务器 114 进行意图推断和/或满足用户请求中表达的用户意图。

[0032] 在各种示例中,存储器 250 能包括附加指令或更少指令。此外,用户设备 102 的各种功能能用硬件和/或固件实现,包括一个或多个信号处理和/或专用集成电路。

[0033] 用于操作虚拟助理的处理

[0034] 图 3 说明根据各种示例的能被执行以使用上下文信息选择性地处理并响应用户输入(例如语音用户输入)的示例性处理 300。在一些示例中,能使用与图 1 所示的系统 100 类似或相同的系统执行处理 300。

[0035] 在框 302,能在用户设备处接收音频输入。音频输入能包括任何可检测的声音,例如音乐、用户语音、背景噪声、其组合等。在一些示例中,用户设备(例如用户设备 102)能经由麦克风接收包括用户的自然语言言语的音频输入。麦克风能将音频输入转换为模拟或数字表示并将音频数据提供到一个或多个处理器(例如处理器 204)。尽管示出为与处理 300 的其它框区分,但应该理解,在一些示例中,能在处理 300 的一些或全部其它框被执行的同时继续在框 302 接收音频输出。

[0036] 在框 304,能监视在框 302 接收的音频输入以识别音频输入中包括或潜在包括语音用户输入的片段。在一些示例中,这能包括监视音频输入的一个或多个特性以识别音频输入内的语音用户输入的开始点和结束点。能使用任何已知开始/结束点算法识别开始点和结束点,例如依靠音频输入的能量特征(例如短时能量和过零率)来在音频输入中将用户语音与背景噪声区分的算法。在一些示例中,用户设备的处理器能分析从设备的麦克风接收的音频数据的能量以识别音频输入中能量足够高并且具有用户语音的过零率特性的片段。在其它示例中,用户设备能将音频数据发送到能够确定语音用户输入的开始点和结束点的远程服务器(例如虚拟助理服务器 114)。

[0037] 在一些示例中,框 304 还能包括对检测的语音用户输入执行语音到文本转换操作,这种操作或者在设备上本地进行,或者通过将音频数据输入到能够进行这种操作的远程服务器来进行。在其它示例中,框 304 可不包括执行语音到文本转换操作。而是,可以在框 308 确定虚拟助理应该响应语音用户输入之后在框 312 执行语音到文本转换操作。

[0038] 在框 306,可以在框 304 监视音频输入的同时确定是否识别到语音用户输入。如果未识别到语音用户输入,则处理可返回框 304。然而,如果识别到语音用户输入,处理可继续至框 308。

[0039] 在框 308,可以基于上下文信息,通过确定在框 304 识别的语音用户输入是否针对虚拟助理来确定虚拟助理是否应该响应语音用户输入(例如,用户使语音用户输入指向虚拟助理并且期望虚拟助理执基于语音用户输入行任务或者提供响应)。下面描述可以用于在框 308 确定语音用户输入是否针对虚拟助理的上下文信息各种示例源。框 308 能由用户设备、远程服务器(例如虚拟助理服务器 114)或其组合执行。

[0040] 在一些示例中,概率系统能用于通过确定用户意图使语音用户输入指向虚拟助理的可能性或置信度得分来确定虚拟助理是否应该响应语音用户输入。概率系统能包括机器学习系统或分类器,例如神经网络。另外,概率系统能学习并使用反馈回路适应用户。在这些概率系统示例中,可能性或置信度得分能包括用户意图使语音用户输入指向虚拟助理的计算概率的数字或其它表示。然后能将计算的可能性或置信度得分与阈值进行比较以确定虚拟助理是否应该响应语音用户输入。例如,如果计算的可能性或置信度得分大于阈值,则可以确定语音用户输入意图针对虚拟助理。然而,如果计算的可能性或置信度得分不大于阈值,则可以确定语音用户输入不意图针对虚拟助理。

[0041] 能用任意数量的方式确定可能性或置信度得分。例如,确定能总体包括加和来自任意数量不同类型的上下文信息的正、负和/或中性贡献。例如,能使用通式 $P = C_1 + C_2 + C_3 + \dots + C_n$ 计算可能性或置信度得分,其中代表 P 代表语音用户输入是意图针对用户设备的可能性或置信度得分, $C_1 \dots C_n$ 可以是代表对来自 N 种不同类型的上下文信息的可能性或置信度得分的正、负或中性贡献的正、负或零值。正贡献能代表建议语音用户输入意图针对虚拟助理的上下文信息类型,负贡献能代表建议语音用户输入不意图针对虚拟助理的上下文信息类型,中性贡献能代表关于语音用户输入意图针对虚拟助理的可能性为中性的上下文信息类型。因此,大 P 值能指示语音用户输入有可能针对虚拟助理,而小或负 P 值能指示语音用户输入有可能不针对虚拟助理。各上下文信息贡献添加到可能性或置信度得分确定的权重或值可以是均匀的或不均匀的。另外,各贡献添加到可能性或置信度得分确定的权重或值能取决于特定类型上下文信息的值。例如,如果贡献 C_1 取决于用户声音量,则 C_1 的符号(例如 +/-)和/或大小能取决于用户声音量的数字表示。

[0042] 尽管以上提供了示例概率系统,但应该理解,可以对描述的系统进行修改和/或可以使用其它计分惯例。例如,正贡献相反能代表建议语音用户输入不意图针对虚拟助理的上下文信息类型,负贡献相反能代表建议语音用户输入意图针对虚拟助理的上下文信息类型。在其它示例中,来自不同类型上下文信息的贡献可以全部为正,大的正值指示上下文信息建议语音用户输入意图(另选地,不意图)针对虚拟助理。在另外其它示例中,来自不同类型上下文信息的贡献可以全部为负,大的负值指示上下文信息建议语音用户输入意图(另选地,不意图)针对虚拟助理。

[0043] 在其它示例中,能使用基于规则的系统,通过评估基于上下文信息的任意数量的条件性规则以确定语音用户输入是否意图针对虚拟助理,来确定虚拟助理是否应该响应语音用户输入。在一些示例中,基于规则的系统能包括决策树的使用。在其它示例中,可基于用户行为学习由基于规则的系统使用的规则。为了说明基于规则的系统示例,第一规则能

包括以下条件：如果用户面对设备并且用户声音量高于阈值量，则可以确定用户意图使语音用户输入指向虚拟助理。第二规则能包括以下条件：如果根据用户的日程用户在开会，则可以确定用户不意图使语音用户输入指向虚拟助理。可以使用包含取决于任意类型的上下文信息的任意数量条件的其它类似规则来使设备确定语音用户输入是否意图针对虚拟助理。在一些示例中，能对规则排序，使得如果多个规则评估为真，则可以使用较高排序规则的输出，作为在框 308 执行的确定操作的结果。另外，在一些示例中，如果没有规则评估为真，则可以进行语音用户输入意图针对虚拟助理（或语音用户输入不意图针对虚拟助理）的默认确定。

[0044] 在框 310，如果因为语音用户输入不意图针对虚拟助理而在框 308 确定虚拟助理不应响应语音用户输入，则处理可返回框 304 以针对语音用户输入监视音频输入。在一些示例中，处理 300 可从框 310 继续至框 304，而不生成语音用户输入的响应。例如，处理 300 可从框 310 继续至框 304，而不执行以下中的一个或多个：执行语音到文本转换，推断用户意图，识别具有设计为实现推断的用户意图的步骤和参数的任务流，输入从推断的用户意图到任务流的特定要求，通过调用程序、方法、服务、API 等运行任务流，以可听（例如语音）和 / 或可视形式生成对用户的输出响应。如果相反因为语音用户输入意图针对虚拟助理而在框 308 确定虚拟助理应该响应语音用户输入，处理可继续至框 312。

[0045] 在框 312，能由用户设备和 / 或远程服务器生成对语音用户输入的响应。在一些示例中，生成对语音用户输入的响应能包括以下中的一个或多个：执行语音到文本转换，推断用户意图，识别具有设计为实现推断的用户意图的步骤和参数的任务流，输入从推断的用户意图到任务流的特定要求，通过调用程序、方法、服务、API 等运行任务流，以可听（例如语音）和 / 或可视形式生成对用户的输出响应。例如，框 312 能包括执行用户请求的操作（例如，打开应用，发送消息，呼叫联系人，执行搜索查询，创建日程约会等），提供用户请求的信息（例如，返回搜索查询结果），执行导致物理环境改变的动作（例如，与家用电器通信以锁门）等。可在用户设备上本地执行操作，或通过向用于处理的远程服务器发送数据执行操作，或通过其组合执行操作。在处理语音用户输入以在框 312 提供合适响应之后，处理能返回框 304。

[0046] 使用处理 300，用户设备实现的虚拟助理能以允许用户用自然语言向虚拟助理说话而不必例如通过在向虚拟助理说话之前按下物理或虚拟按钮或者通过在向虚拟助理说话之前用自然语言说出特定触发短语（例如，预定词语或词语序列，例如“Hey Siri”）而手动输入开始点标识符的方式来选择性忽视或响应语音用户输入。在一些示例中，能使用处理 300 来处理用户设备 102 接收的所有语音用户输入。

[0047] 为了说明图 3 的操作，用户与虚拟助理之间的一个示例交互能开始于用户在言语上问问题，例如“纽约现在几点了？”，用户无需例如通过在向虚拟助理说话之前按下物理或虚拟按钮或者通过说出特定触发短语（例如，预定词语或词语序列，例如“Hey Siri”）来手动输入开始点标识符。在框 302，用户设备能接收包括用户问题的音频输入。在框 304，能监视音频输入以检测语音用户输入。由于音频输入包括用户问题，所以可以在框 306 确定语音用户输入被识别。在框 308，可以基于与识别的语音用户输入相关联的上下文信息，确定虚拟助理是否应该响应用户问题。在此示例中，可以确定（使用基于规则的系统或概率系统）虚拟助理应该响应用户问题，因为上下文信息指示用户在说问题时在看用户设

备并且用户声音的量高于阈值量。因此,处理能继续至框 310 和 312,在此能生成对用户问题的响应。例如,在框 312,能处理用户问题以确定用户的意图,识别要执行的任务,执行功能以响应于用户询问确定并显示时间“5:01p. m.”。处理 300 然后能返回框 304。

[0048] 用户然后能在言语上问另一个问题,例如“那里天气如何?”,用户无需例如通过在向虚拟助理说话之前按下物理或虚拟按钮或者通过说出特定触发短语(例如,预定词语或词语序列,例如“Hey Siri”)来手动输入开始点标识符。能在框 304 监视用户设备接收的并且包括用户的第二个问题的音频输入。由于音频输入包括用户第二个问题,所以在框 306 确定识别到语音用户输入。在框 308,可基于与识别的语音用户输入相关联的上下文信息,确定虚拟助理是否应该响应用户问题。在此示例中,可以确定(使用基于规则的系统或概率系统)虚拟助理应该响应用户问题,因为上下文信息指示用户在从接收到第一个问题的答案后的阈值长度时间内问了第二个问题,这意味着第二个问题是同一会话的一部分。因此,处理能继续至框 310 和 312,在此能生成对用户问题的响应。例如,在框 312,能处理用户问题以确定用户的意图,识别要执行的任务,执行功能以响应于用户询问确定并显示消息“天气晴朗”。处理 300 然后能返回框 304 以针对另外的语音用户输入监视音频输入。

[0049] 在其它示例中,用户设备 102 能在处理 300 被调用之前要求用户手动输入开始点标识符。例如,能要求用户在开始向虚拟助理说话之前说出触发短语或按下物理或虚拟按钮。响应于手动开始点标识符,能如上所述执行处理 300,并且能处理后续语音用户输入,而不要求用户输入附加的开始点标识符。图 4 示出用于执行此修改交互的示例处理。

[0050] 在框 402,能接收开始点标识符。开始点标识符能包括用户说的触发短语,选择物理或虚拟按钮,或者从用户接收的其它手动输入。在框 404,能以与如上所述框 302 类似或相同的方式接收音频输入。在框 406,能通过识别第一语音用户输入的结束点,从在框 404 接收的音频输入识别初始语音用户输入。如上所述,能基于音频输入的能量特征识别结束点。在框 408,能以与如上所述框 302 类似或相同的方式生成对第一语音用户输入的响应。然而,与框 312 相比,能执行框 408 而不以与框 308 类似的方式确定虚拟助理是否应该响应第一语音用户输入,因为在框 402 接收到手动开始点标识符。在框 408 生成对第一语音用户输入的响应之后,处理能继续至框 304。在一些示例中,能省略框 302,因为之前是在框 404 接收到音频输入。框 304、306、308、310、312 能重复执行,如以上参照图 3 描述的。

[0051] 在一些示例中,一旦在处理 400 中调用框 304,就能针对所有后续语音用户输入继续执行框 304、306、308、310、312。在其它示例中,如果在大于阈值长度时间内没有接收到语音用户输入(例如在框 304 和 306),则框 304、306、308、310、312 的执行能停止,导致用户需要在输入下一语音用户输入之前在框 402 输入开始点标识符。

[0052] 使用处理 400,用户设备实现的虚拟助理能仅要求用户输入一次开始点标识符,并且能允许虚拟助理选择性忽视或响应后续语音用户输入,而不要求用户在每个后续语音用户输入之前都重复输入手动开始点标识符。

[0053] 为了说明图 4 的操作,用户与虚拟助理之间的示例交互能开始于用户例如通过在向虚拟助理说话之前按下物理或虚拟按钮或者通过说出特定触发短语(例如,预定词语或词语序列,例如“Hey Siri”)来手动输入开始点标识符。能在框 402 接收开始点标识符。然后用户能在言语上问问题,例如“纽约现在几点了?”用户设备能在框 404 接收包括用户问题的音频输入。在框 406,能通过识别用户问题的结束点来从音频输入识别语音用户输

入。在框 408,能生成对用户问题的响应,而不必执行类似于框 308 的确定步骤,因为用户在框 402 手动识别输入是在虚拟助理处进行的。例如,在框 408,能处理用户问题以确定用户的意图,识别要执行的任务,并且响应于用户的询问执行功能以确定并显示时间“5:01p.m”。处理 400 然后能继续至框 304。

[0054] 用户然后能在言语上问另一个问题,例如“那哪里天气如何?”,用户无需例如通过在向虚拟助理说话之前按下物理或虚拟按钮或者通过说出特定触发短语(例如,预定词语或词语序列,例如“Hey Siri”)来手动输入开始点标识符。能在框 304 和 306 重复监视用户设备接收的包括用户的第二个问题的音频输入。由于音频输入包括用户第二个问题,所以在框 306 确定识别到语音用户输入。在框 308,可基于与识别的语音用户输入相关联的上下文信息,确定虚拟助理是否应该响应用户问题。在此示例中,可以确定(使用基于规则的系统或概率系统)虚拟助理应该响应用户问题,因为上下文信息指示用户在从接收到第一个问题的答案后的阈值长度时间内问了第二个问题,这意味着第二个问题是同一会话的一部分。因此,处理能继续至框 310 和 312,在此能生成对用户问题的响应。例如,在框 312,能处理用户问题以确定用户的意图,识别要执行的任务,响应于用户询问执行功能以确定并显示消息“天气晴朗”。处理 300 然后能返回框 304 以针对另外的语音用户输入监视音频输入。

[0055] 在一些示例中,在执行处理 300 或处理 400 的框 304、306、308、310、312 的同时,可以在用户设备 102 的显示器上显示可视指示符以指示用户设备 102 能够接受自然语言形式的语音用户输入,而不必使用触发短语或其它手动输入来指示语音用户输入是意图针对用户设备 102 的。

[0056] 另外,尽管以特定次序示出并描述了处理 300 和 400 的框,应该理解,处理 300 和 400 的框能按其它次序或同时执行。例如,在处理 300 中,在框 304、306、308、310、312 中的一些或全部能够被执行的同时,用户设备 102 能在框 302 继续接收音频输入。类似地,在处理 400 中,在框 304、306、308、310、312、406、408 中的一些或全部能够被执行的同时,用户设备 102 能在框 404 继续接收音频输入。

[0057] 应该理解,处理 300 和 400 的框能在用户设备 102、服务器系统 110 或用户设备 102 和服务器系统 110 的组合上执行。例如,在一些示例中,处理 300 或 400 的所有框能在用户设备 102 上执行。在其它示例中,处理 300 或 400 的所有框能在服务器系统 110 执行。在另外其它示例中,处理 300 或 400 的一些框能在用户设备 102 处执行,而处理 300 或 400 的其它框能在服务器系统 110 执行。

[0058] 上下文信息

[0059] 如上所讨论的,处理器 204 和 / 或服务器系统 110 在处理 300 或 400 的框 308 中可以使用任意数量类型的上下文信息(也能包括分类器或机器学习系统使用的特征)来确定语音用户输入是否意图针对虚拟助理。下面描述一些示例类型的上下文信息以及在处理 300 的框 308 这些类型的上下文信息能用于确定语音用户输入是否意图针对虚拟助理的方式。

[0060] 在一些示例中,上下文信息能包括来自用户设备 102 的时钟或定时器的时间数据。时间数据能表示任意期望的两个或更多个事件之间的时间长度。例如,时间数据能表示接收到语音用户输入与接收到之前用户输入(例如按钮按下,鼠标点击,屏幕触摸,之前

语音用户输入等)的时间之间的时间长度。通常,在一些示例中,两个事件之间的较短时间长度能指示用户更可能意图使当前语音用户输入指向虚拟设备,而事件之间的较长时间长度能指示用户不大可能意图使当前语音用户输入指向虚拟设备。然而,在其它示例中,两个事件之间的较长时间长度能指示用户更可能意图使当前语音用户输入指向虚拟设备,而事件之间的较短时间长度能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0061] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果连续语音用户输入之间的时间长度小于阈值持续时间,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果连续语音用户输入之间的时间长度大于或等于阈值持续时间,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0062] 在一个示例概率系统中,连续语音用户输入之间的时间长度可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与时间长度值具有线性或非线性关系。例如,小于阈值持续时间的长度能对最终可能性或置信度得分贡献正值,其中针对较短时间长度,正值的大小能较大。类似地,大于或等于阈值持续时间的长度能对最终可能性或置信度得分贡献零或负值,其中针对较长时间长度,负值的大小能较大。在一些示例中,连续语音用户输入之间的时间长度可用于训练概率系统的机器学习系统。

[0063] 在一些示例中,上下文信息能包括来自存储器 250 或位于用户设备 102 内或远离用户设备 102 的另一存储设备的会话历史数据。会话历史数据能包括从用户接收的任意数量之前语音用户输入和/或用户设备生成并提供给用户的响应。在一些示例中,之前接收的语音用户输入能与当前语音用户输入比较以确定当前语音用户输入是否与之前接收的语音用户输入相同。在这些示例中,之前语音用户输入和当前语音用户输入之间的匹配(例如,用户重复他或她自己造成)能指示用户更可能意图使当前语音用户输入指向虚拟设备,而之前语音用户输入和当前语音用户输入之间的不匹配能指示用户不大可能意图使当前语音用户输入指向虚拟设备或者关于用户意图使当前语音用户输入指向虚拟设备的可能性是中性的。在一些示例中,用户重复他或她自己能用于反馈回路,以训练概率系统的机器学习系统。

[0064] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果当前语音用户输入与之前接收的语音用户输入相同或匹配,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果当前语音用户输入不与之前接收的语音用户输入相同或匹配,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0065] 在一个示例概率系统中,确定之前接收的语音用户输入与当前语音用户输入相同或匹配可对最终可能性或置信度得分贡献正值,而之前接收的语音用户输入与当前语音用户输入之间的不匹配可对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正贡献或负贡献的大小。

[0066] 在其它示例中,能对当前语音用户输入和一些或全部会话历史数据执行语义相似性分析。在一些示例中,这能包括计算确定的用户意图的相似性(例如,采取元组<命令,

参数>形式的自然语言解释短语的结果)。在其它示例中,执行语义相似性分析以确定语义距离能包括确定与相似性矩阵相组合的编辑距离。在这些示例中,可确定当前语音用户输入与之前接收的语音用户输入或用户设备生成并提供给用户的响应中的一个或多个之间的语义距离,并将其用于在框 308 确定语音用户输入意图针对虚拟助理的可能性或置信度得分。在这些示例中,当前语音用户输入与之前接收的语音用户输入中的一个或多个(例如,前面刚刚的语音用户输入)和/或用户设备生成并提供给用户的响应中的一个或多个之间的小语义距离能指示用户更可能意图使当前语音用户输入指向虚拟设备,而当前语音用户输入与之前接收的语音用户输入中的一个或多个(例如,前面刚刚的语音用户输入)和/或用户设备生成并提供给用户的响应中的一个或多个之间的大语义距离能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0067] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果当前语音用户输入与一个或多个之前的语音用户输入或用户设备生成的响应之间的语义距离小于阈值,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果当前语音用户输入与一个或多个之前的语音用户输入或用户设备生成的响应之间的语义距离大于或等于阈值,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0068] 在一个示例概率系统中,当前语音用户输入与一个或多个之前的语音用户输入或用户设备生成的响应之间的语义距离可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与语义距离具有线性或非线性关系。例如,小于阈值的语义距离能对最终可能性或置信度得分贡献正值,其中针对较小语义距离,正值的大小能较大。类似地,大于或等于阈值的语义距离能对最终可能性或置信度得分贡献零或负值,其中针对较大语义距离,负值的大小能较大。

[0069] 在一些示例中,上下文信息能包括与用户设备 102 的距离传感器(例如接近传感器 214)的距离数据。距离数据能表示用户设备与用户之间的空间距离(例如,用户设备与用户面部之间的距离)。通常,在一些示例中,用户设备与用户之间的较短距离能指示用户更可能意图使当前语音用户输入指向虚拟设备,而用户设备与用户之间的较长距离能指示用户不大可能意图使当前语音用户输入指向虚拟设备。然而,在其它示例中,用户设备与用户之间的较长距离能指示用户更可能意图使当前语音用户输入指向虚拟设备,用户设备与用户之间的较短距离能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0070] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果用户设备与用户之间的距离小于阈值距离,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果用户设备与用户之间的距离大于或等于阈值距离,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0071] 在一个示例概率系统中,用户设备与用户之间的距离可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与用户设备与用户之间的距离值具有线性或非线性关系。例如,小于阈值距离的距离能对最终可能性或置信度得分贡献正值,其中针

对较短距离,正值的大小能较大。类似地,大于或等于阈值距离的距离能对最终可能性或置信度得分贡献零或负值,其中针对较大距离,负值的大小能较大。

[0072] 在一些示例中,上下文信息能包括来自音频子系统 226 的音频数据。音频数据能包括语音用户输入的音量表示。通常,在一些示例中,较高音量的语音用户输入能指示用户更可能意图使当前语音用户输入指向虚拟设备,而较低音量的语音用户输入能指示用户不大可能意图使当前语音用户输入指向虚拟设备。然而,在其它示例中,较低音量的语音用户输入能指示用户更可能意图使当前语音用户输入指向虚拟设备,而较高音量的语音用户输入能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0073] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果语音用户输入的音量大于阈值音量,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果语音用户输入的音量小于或等于阈值音量,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0074] 在一个示例概率系统中,语音用户输入音量可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与语音用户输入音量具有线性或非线性关系。例如,大于阈值音量的音量能对最终可能性或置信度得分贡献正值,其中针对较高音量,正值的大小能较大。类似地,小于或等于阈值音量的音量能对最终可能性或置信度得分贡献零或负值,其中针对较低音量,负值的大小能较大。

[0075] 在一些示例中,上下文信息能包括来自音频子系统 226 的音频数据。音频数据能包括语音用户输入的音量表示。在一些示例中,如果忽视之前语音输入并且当前语音用户输入音量高于之前语音用户输入,则这能指示用户更可能意图使当前语音用户输入指向虚拟设备。

[0076] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果当前语音用户输入音量大于之前语音用户输入音量,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果当前语音用户输入音量小于或等于之前语音用户输入音量,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0077] 在一个示例概率系统中,语音用户输入音量可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与语音用户输入音量具有线性或非线性关系。例如,如果当前语音用户输入音量大于刚刚之前语音用户输入音量,则可以添加对最终可能性或置信度得分的正值。类似地,如果当前语音用户输入音量小于刚刚之前语音用户输入音量,则可以对最终可能性或置信度得分添加零或负值。

[0078] 在其它示例中,能使用已知发言者识别技术分析音频数据,以确定接近或位于用户设备音频距离内的不同发言者的数量。在这些示例中,确定存在多于一个发言者能指示用户不大可能意图使当前语音用户输入指向虚拟设备(而是向附近的另一个人说话),而确定存在仅一个发言者能指示用户更可能意图使当前语音用户输入指向虚拟设备。

[0079] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果在接收语音用户输入时确定存在多于一个

发言者,则可以确定用户不意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果在接收语音用户输入时确定存在仅一个发言者,则可以确定用户意图使当前语音用户输入指向虚拟设备。

[0080] 在一个示例概率系统中,在接收语音用户输入时确定存在多于一个发言者能对最终可能性或置信度得分贡献负值,而在接收语音用户输入时确定存在仅一个发言者能对最终可能性或置信度得分贡献零或正值。能基于总体系统设计调整正或负贡献的大小。

[0081] 在另外其它示例中,能使用已知发言者识别技术分析音频数据以确定语音用户输入是否是从用户设备的已知或授权用户(例如,设备拥有者)或者从与之前接收语音用户输入相同的发言者接收的。在这些示例中,确定语音用户输入是从已知或授权用户或者从与之前接收语音用户输入相同的发言者接收的能指示用户更可能意图使当前语音用户输入指向虚拟设备,确定语音用户输入不是从已知或授权用户或者从与之前接收语音用户输入相同的发言者接收的能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0082] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果确定语音用户输入是从已知或授权用户或者从与之前接收语音用户输入相同的发言者接收的,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果确定语音用户输入不是从已知或授权用户或者从与之前接收语音用户输入相同的发言者接收的,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0083] 在一个示例概率系统中,确定语音用户输入是从已知或授权用户或者从与之前接收语音用户输入相同的发言者接收的能对最终可能性或置信度得分贡献正值,而确定语音用户输入不是从已知或授权用户或者从与之前接收语音用户输入相同的发言者接收的能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0084] 在一些示例中,上下文信息能包括来自用户设备 102 的相机子系统 220 的图像数据。图像数据能表示相机子系统 220 捕获的图像或视频。在一些示例中,图像数据能用于估计用户设备与用户之间的距离。例如,图像内用户的大小能用于估计用户设备与用户之间的距离。估计的用户设备与用户之间的距离能按照和如上所述与接近传感器 214 的距离数据类似或相同的方式在基于规则的系统或概率系统中使用。

[0085] 在其它示例中,可以分析(例如使用已知眼部跟踪技术)图像数据以确定在接收到用户语音输入时用户是否在看或面对用户设备。在这些示例中,确定在接收到用户语音输入时用户在看用户设备能指示用户更可能意图使当前语音用户输入指向虚拟设备,而确定在接收到用户语音输入时用户没有在看用户设备能指示用户不大可能意图使当前语音用户输入指向虚拟设备或者关于用户意图使当前语音用户输入指向虚拟设备的可能性可以是中性的。

[0086] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果确定在接收到用户语音输入时用户在看用户设备,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果确定在接收到

用户语音输入时用户没有在看用户设备,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0087] 在一个示例概率系统中,确定在接收到用户语音输入时用户在看用户设备能对最终可能性或置信度得分贡献正值,而确定在接收到用户语音输入时用户没有在看用户设备能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0088] 在另外其它示例中,能分析图像数据以确定设备相对于用户的方向。例如,能使用已知面部识别技术分析图像数据,以基于用户是否出现在光学传感器 222 的视场中确定用户是否位于用户设备前面。类似地,能使用已知图像识别技术分析图像数据以确定用户是否执行特定动作(例如,指向用户设备,在用户设备处做手势等)或者以预定方式定位(例如,坐在电视前面,拿住遥控器等)。在这些示例中,确定当接收用户输入时用户位于用户设备前面、执行特定动作或以预定方式定位能指示用户更可能意图使当前语音用户输入指向虚拟设备,而确定当接收用户输入时用户没有位于用户设备前面、没有执行特定动作或没有以预定方式定位能指示用户不大可能意图使当前语音用户输入指向虚拟设备或者关于用户意图使当前语音用户输入指向虚拟设备的可能性可以是中性的。

[0089] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果确定当接收用户输入时用户位于用户设备前面、执行特定动作或以预定方式定位,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果确定当接收用户输入时用户没有位于用户设备前面、没有执行特定动作或没有以预定方式定位,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0090] 在一个示例概率系统中,确定当接收用户输入时用户位于用户设备前面、执行特定动作或以预定方式定位能对最终可能性或置信度得分贡献正值,而确定当接收用户输入时用户没有位于用户设备前面、没有执行特定动作或没有以预定方式定位能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0091] 在一些示例中,上下文信息能包括来自用户设备 102 的运动传感器 210 的方向数据。运动传感器 210 能包括任意类型的方向传感器,例如倾斜仪、罗盘、陀螺仪等,其能够生成表示用户设备的自由空间方向的方向数据。在一些示例中,用户设备的特定方向(例如,设备正面向上,设备直立,设备处于设备的显示器能被用户观看的方向等)能指示用户更可能意图使当前语音用户输入指向虚拟设备,而用户设备的其它方向(例如,设备正面向下,设备倒转,设备处于设备的显示器不能被用户观看的方向等)能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0092] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果当接收语音用户输入时设备处于一组方向中的一个(例如,设备正面向上,设备直立,设备处于设备的显示器能被用户观看的方向等),则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果当接收语音用户输入时设备没有处于一组方向中的一个,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0093] 在一个示例概率系统中,确定当接收语音用户输入时设备处于一组方向中的一个(例如,设备正面向上,设备直立,设备处于设备的显示器能被用户观看的方向等)能对最终可能性或置信度得分贡献正值,而确定当接收语音用户输入时用户设备没有处于该组方向中的一个能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0094] 在一些示例中,上下文信息能包括来自用户设备 102 的其它传感器 216 的 GPS 接收器的位置数据。位置数据能表示用户设备的地理位置。在一些示例中,用户设备处于特定位置时(例如,在家,在办公室等)接收语音用户输入能指示用户更可能意图使当前语音用户输入指向虚拟设备,而用户设备处于其它特定位置时(例如,在电影院,在会议室等)接收语音用户输入能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0095] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果当接收语音用户输入时设备处于一组位置中的一个(例如,在家,在办公室等),则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果当接收语音用户输入时用户设备没有处于一组位置中的一个,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0096] 在一个示例概率系统中,确定当接收语音用户输入时设备处于一组位置中的一个(例如,在家,在办公室等)能对最终可能性或置信度得分贡献正值,而确定当接收语音用户输入时用户设备没有处于一组位置中的一个能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0097] 在一些示例中,上下文信息能包括来自存储器 250 或位于用户设备 102 内或远离用户设备 102 的另一存储设备的操作状态数据。操作状态数据能包括与用户设备的操作状态有关的任何信息,例如内容是否显示或者呈现给用户,呈现给用户的内容的类型或标识,用户设备运行的应用,最近是否向用户呈现通知,之前或最近联系人,之前或最近电子邮件,之前或最近 SMS 消息,之前或最近电话呼叫,日程条目,网页访问,用户设备显示器开/关状态,用户设备是否接收除语音用户输入之前的用户输入,用户设备上的设置,之前行为等。在一些示例中,当用户设备处于特定操作状态时(例如,内容或其它信息正显示给用户,内容或其它信息正可听地呈现给用户,特定类型内容正呈现给用户,特定内容正呈现给用户,例如用户与虚拟助理之间的会话脚本,用户设备正运行应用,通知最近呈现给用户,用户设备显示器打开,用户设备在接收除语音用户输入之前的用户输入,例如鼠标输入,键盘输入,触敏显示器输入等,电子邮件最近发送到联系人或特定联系人/从联系人或特定联系人接收,SMS 消息最近发送到联系人或特定联系人/从联系人或特定联系人接收,电话呼叫最近发送到联系人或特定联系人/从联系人或特定联系人接收,在用户设备上配置特定设置,执行之前行为等)接收语音用户输入能指示用户更可能意图使当前语音用户输入指向虚拟设备,而当用户设备处于其它特定操作状态时(例如,内容或其它信息没有正显示给用户,内容或其它信息没有正可听地呈现给用户,特定类型内容没有正呈现给用户,特定内容没有正呈现给用户,例如用户与虚拟助理之间的会话脚本,用户设备没有正运行应用,通知没有最近呈现给用户,用户设备显示器关闭,用户设备没有在接收除语音用户输入之前的用户输入,例如鼠标输入,键盘输入,触敏显示器输入等,电子邮件最近没有发送到

联系人或特定联系人 / 从联系人或特定联系人接收, SMS 消息最近没有发送到联系人或特定联系人 / 从联系人或特定联系人接收, 电话呼叫最近没有发送到联系人或特定联系人 / 从联系人或特定联系人接收, 在用户设备上没有配置特定设置, 没有执行之前行为等) 接收语音用户输入能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0098] 在一个基于规则的系统示例中, 可以 (单独, 与其它规则结合, 或作为其它规则中的多个条件中的一个) 使用的一个规则是: 如果当接收语音用户输入时用户设备的显示器打开并且用户设备正在向用户呈现音频信息, 则可以确定用户意图使当前语音用户输入指向虚拟设备。可以 (单独, 与其它规则结合, 或作为其它规则中的多个条件中的一个) 使用的另一个规则是: 如果当接收语音用户输入时用户设备的显示器关闭并且用户设备没有正在向用户呈现音频信息, 则可以确定用户不意图使当前语音用户输入指向虚拟设备。可以类似地使用其它类型的操作状态数据生成导致确定语音用户输入是否针对虚拟助理的规则。

[0099] 在一个示例概率系统中, 当接收语音用户输入时用户设备的显示器打开并且用户设备正在向用户呈现音频信息能对最终可能性或置信度得分贡献正值, 而当接收语音用户输入时用户设备的显示器关闭并且用户设备没有正在向用户呈现音频信息能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。应该理解, 可用类似方式使用其它类型的操作状态数据, 以根据操作状态数据是否指示设备状态是预定状态集合中的一个来进行对最终可能性或置信度得分的正、负或中性贡献。

[0100] 在其它示例中, 能对当前语音用户输入和一些或全部操作状态数据执行语义相似性分析。在这些示例中, 能确定当前语音用户输入与操作状态数据的一个或多个成分之间的语义距离, 并将其用于在框 308 确定语音用户输入是否意图针对用户设备。在这些示例中, 当前语音用户输入与操作状态数据的一个或多个成分之间的小语义距离能指示用户更可能意图使当前语音用户输入指向虚拟设备, 而当前语音用户输入与操作状态数据的一个或多个成分之间的大语义距离能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0101] 在一个基于规则的系统示例中, 可以 (单独, 与其它规则结合, 或作为其它规则中的多个条件中的一个) 使用的一个规则是: 如果当前语音用户输入与操作状态数据 (例如, 用户设备正运行的应用, 呈现给用户的通知, 联系人列表中的名字, 之前联系人, 之前电子邮件, 之前 SMS 消息, 呈现给用户的内容, 期望从用户接收的命令, 例如用户设备运行地图应用时对方向的请求, 用户设备处于眼自由模式时的内容导航指令, 之前接收“停止”或“暂停”指令后的“开始”指令等) 的一个或多个成分之间的语义距离小于阈值, 则可以确定用户意图使当前语音用户输入指向虚拟设备。可以 (单独, 与其它规则结合, 或作为其它规则中的多个条件中的一个) 使用的另一个规则是: 如果当前语音用户输入与操作状态数据的一个或多个成分之间的语义距离大于或等于阈值, 则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0102] 在一个示例概率系统中, 当前语音用户输入与操作状态数据的一个或多个成分之间的语义距离可用于计算对最终可能性或置信度得分的正、负或中性贡献, 其中贡献值能与语义距离具有线性或非线性关系。例如, 小于阈值的语义距离能对最终可能性或置信度得分贡献正值, 其中针对较小语义距离, 正值的大小能较大。类似地, 大于或等于阈值的语

义距离能对最终可能性或置信度得分贡献零或负值,其中针对较大语义距离,正值的大小能较大。

[0103] 在一些示例中,上下文信息能包括来自用户设备 102 的光传感器 212 的照明数据。照明数据能包括光传感器 212 接收的环境光的亮度表示。在一些示例中,当接收语音用户输入时感测的环境光的较高亮度能指示用户更可能意图使当前语音用户输入指向虚拟设备(例如,指示用户处于可接受说话的环境),而当接收语音用户输入时感测的环境光的较低亮度能指示用户不大可能意图使当前语音用户输入指向虚拟设备(例如,指示用户处于不可接受说话的环境,例如电影院)。然而,在其它示例中,当接收语音用户输入时感测的环境光的较低亮度能指示用户更可能意图使当前语音用户输入指向虚拟设备,而当接收语音用户输入时感测的环境光的较高亮度能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0104] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果当接收语音用户输入时感测的环境光的亮度大于阈值亮度,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果当接收语音用户输入时感测的环境光的亮度小于或等于阈值亮度,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0105] 在一个示例概率系统中,当接收语音用户输入时感测的环境光的亮度可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与感测的环境光的亮度值具有线性或非线性关系。例如,小于阈值亮度的亮度能对最终可能性或置信度得分贡献负值,其中针对较低亮度值,负值的大小能较大。类似地,大于或等于阈值亮度的亮度能对最终可能性或置信度得分贡献零或正值,其中针对较高亮度值,负值的大小能较大。

[0106] 在一些示例中,上下文信息能包括来自于用户设备 102 内或远离用户设备 102 的自动语音识别(ASR)引擎(例如,来自服务器系统 110)的语音识别数据。语音识别数据能包括 ASR 引擎是否能够识别语音用户输入和/或能够响应语音用户输入的指示。在一些示例中,ASR 引擎能够识别语音用户输入和/或能够响应语音用户输入的指示能指示用户更可能意图使当前语音用户输入指向虚拟设备,而 ASR 引擎不能够识别语音用户输入和/或不能够响应语音用户输入的指示能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0107] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果 ASR 引擎能够识别语音用户输入和/或能够响应语音用户输入,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果 ASR 引擎不能够识别语音用户输入和/或不能够响应语音用户输入的指示,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0108] 在一个示例概率系统中,确定 ASR 引擎能够识别语音用户输入和/或能够响应语音用户输入能对最终可能性或置信度得分贡献正值,而确定 ASR 引擎不能够识别语音用户输入和/或不能够响应语音用户输入能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0109] 在其它示例中,来自 ASR 引擎的语音识别数据还能包括语音用户输入的长度的指示(例如,字数,语音持续时间等)。通常,在一些示例中,较短长度的语音用户输入能指示用户更可能意图使当前语音用户输入指向虚拟设备,而较长长度的语音用户输入能指示用户不大可能意图使当前语音用户输入指向虚拟设备。然而,在一些示例中,较长长度的语音用户输入能指示用户更可能意图使当前语音用户输入指向虚拟设备,而较短长度的语音用户输入能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0110] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果语音用户输入的长度小于阈值长度,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果语音用户输入的长度大于或等于阈值长度,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0111] 在一个示例概率系统中,语音用户输入的长度可用于计算对最终可能性或置信度得分的正、负或中性贡献,其中贡献值能与语音用户输入的长度具有线性或非线性关系。例如,小于阈值长度的长度能对最终可能性或置信度得分贡献正值,其中针对较短长度,正值的大小能较大。类似地,大于或等于阈值长度的长度能对最终可能性或置信度得分贡献零或负值,其中针对较长长度,正值的大小能较大。

[0112] 在其它示例中,来自 ASR 引擎的语音识别数据还能包括从语音用户输入内识别的名词或代词。例如,语音识别数据可以包括名词或代词,诸如“亲爱的”、“他”、“她”或人的名或姓。通常,在一些示例中,存在这些名词或代词中的一个能指示用户不大可能意图使当前语音用户输入指向虚拟设备,而不存在这些名词或代词中的一个(或存在非人类标识符,例如“Siri”)能指示用户更可能意图使当前语音用户输入指向虚拟设备。

[0113] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果语音用户输入包括一组名词或代词中的一个,则可以确定用户不意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果语音用户输入不包括一组名词或代词中的一个(或者包括另一组名词或代词中的一个),则可以确定用户意图使当前语音用户输入指向虚拟设备。

[0114] 在一个示例概率系统中,确定语音用户输入包括一组名词或代词中的一个能对最终可能性或置信度得分贡献负值,而确定语音用户输入不包括一组名词或代词中的一个(或者包括另一组名词或代词中的一个)能对最终可能性或置信度得分贡献正或零值。能基于总体系统设计调整正或负贡献的大小。

[0115] 在一些示例中,上下文信息能包括来自存储器 250 或位于用户设备 102 内或远离用户设备 102 的另一存储设备的用户数据。用户数据能包括与用户相关联的任意类型信息,例如联系人列表,日程,喜好,个人信息,财务信息,家庭信息等。在一些示例中,用户数据能在框 308 与其它类型的上下文信息比较以辅助确定语音用户输入是否针对虚拟助理。例如,接收语音用户输入时的时间能与用户日程比较,以确定用户是否处于用户更可能或不大可能与用户设备的虚拟助理对话的情况,来自 ASR 引擎的语音识别数据能与用户联系人列表中的联系人比较,以确定来自用户联系人列表的名字是否在语音用户输入中提到,来自 ASR 引擎的语音识别数据能与用户喜好比较,以确定语音用户输入是否对应于虚拟助

理应该或不应该忽视的之前定义短语等。

[0116] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果在预定时间集合内的时间(例如,当用户日程指示用户在开会或者从事认为用户不会与虚拟助理对话的活动时)接收到当前语音用户输入,则可以确定用户不意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果在预定时间集合外的时间(例如,当用户日程指示用户没在开会或者没从事认为用户不会与虚拟助理对话的活动时)接收到当前语音用户输入,则可以确定用户意图使当前语音用户输入指向虚拟设备。类似地可以使用其它类型的用户数据来生成导致确定语音用户输入是否针对虚拟助理的规则。

[0117] 在一个示例概率系统中,确定当用户日程指示用户在开会或者从事认为用户不会与虚拟助理对话的活动时接收到当前语音用户输入能对最终可能性或置信度得分贡献负或零值,而确定当用户日程指示用户没在开会或者没从事认为用户不会与虚拟助理对话的活动时接收到当前语音用户输入能对最终可能性或置信度得分贡献负或零值能对最终可能性或置信度得分贡献正值。能基于总体系统设计调整正或负贡献的大小。应该理解,能按类似方式使用其它类型的用户数据,以进行对最终可能性或置信度得分的正、负或中性贡献。

[0118] 在一些示例中,上下文信息能包括来自用户设备 102 的运动传感器或其它传感器 216 中的加速计的运动数据。运动数据能表示用户设备的移动,能用于检测用户摇晃设备造成的设备移动,设备朝向或远离用户的移动(例如,朝向或远离用户嘴部移动),用户穿戴设备造成的移动(例如,作为表或其他可穿戴设备)等。在一些示例中,用户设备体验的特定运动(例如,摇晃,与用户穿戴用户设备相关联的移动,朝向用户移动等)能指示用户更可能意图使当前语音用户输入指向虚拟设备,而用户设备体验的其它移动(例如,远离用户移动)能指示用户不大可能意图使当前语音用户输入指向虚拟设备。

[0119] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果运动数据指示在接收语音用户输入之前用户设备朝向用户嘴部移动,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果运动数据指示在接收语音用户输入之前用户设备没有朝向用户嘴部移动,则可以确定用户不意图使当前语音用户输入指向虚拟设备。类似地能使用其它移动来生成导致确定语音用户输入是否针对虚拟助理的规则。

[0120] 在一个示例概率系统中,确定用户设备在预定移动集合中的一个中移动(例如,接收语音用户输入之前朝向用户嘴部移动)能对最终可能性或置信度得分贡献正值,而确定用户设备没有在预定移动集合中的一个中移动能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。应该理解,可按类似方式使用其它移动以进行对最终可能性或置信度得分的正、负或中性贡献。

[0121] 在一些示例中,上下文信息能包括来自用户设备 102 的其它传感器 216 中的温度传感器的温度数据。温度数据能表示温度传感器感测的温度,并且能用于确定用户设备是否被用户握持。例如,较高温度或特定范围中的温度能表示设备被握持在用户手中,而较低

温度或特定范围外的温度能表示用户设备没有被用户握持。

[0122] 在一个基于规则的系统示例中,可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的一个规则是:如果温度在对应于用户设备被用户握持的特定温度范围内,则可以确定用户意图使当前语音用户输入指向虚拟设备。可以(单独,与其它规则结合,或作为其它规则中的多个条件中的一个)使用的另一个规则是:如果温度不在对应于用户设备被用户握持的特定温度范围内,则可以确定用户不意图使当前语音用户输入指向虚拟设备。

[0123] 在一个示例概率系统中,基于温度数据确定当接收语音用户输入时用户设备正被用户握持能对最终可能性或置信度得分贡献正值,而基于温度数据确定当接收语音用户输入时用户设备没正被用户握持能对最终可能性或置信度得分贡献零或负值。能基于总体系统设计调整正或负贡献的大小。

[0124] 电子设备

[0125] 根据一些示例,图 5 示出根据各种描述示例的原理配置的电子设备 500 的功能框图。设备的功能框图能通过硬件、软件或硬件和软件的组合实现,以实现各种描述示例的原理。本领域技术人员理解,图 5 中描述的功能框能组合或分离为子框以实现各种描述示例的原理。因此,描述此处可选地支持这里描述的功能框的各种可能组合或分离或进一步定义。

[0126] 如图 5 所示,电子设备 500 能包括配置为显示用户界面并且接收触摸输入的触摸屏显示单元 502,以及配置为接收声音输入的声音接收单元 504。在一些示例中,电子设备 500 能包括配置为产生声音的扬声器单元 506。电子设备 500 还能包括耦接到触摸屏显示单元 502 和声音接收单元 504(并且可选地,耦接到扬声器单元 506)的处理单元 508。在一些示例中,处理单元 508 能包括第一监视单元 510、语音识别单元 512、响应确定单元 514、响应生成单元 516、第二监视单元 518、第三监视单元 520。

[0127] 处理单元 508 能配置为接收音频输入(例如,从音频接收单元 504)。处理单元 508 能配置为监视音频输入(例如,使用监视单元 510)以识别音频输入中的第一语音用户输入。在识别了音频输入中的第一语音用户输入时,处理单元 508 能配置为基于与第一语音用户输入相关联的上下文信息确定(例如,使用响应确定单元 514)是否应该生成对第一语音用户输入的响应。响应于确定应该生成响应,处理单元 508 能配置为生成对第一语音用户输入的响应(例如,使用响应生成单元 516)并且针对第二语音用户输入再次监视接收的音频输入(例如,使用第二监视单元 518)。响应于确定不应该生成响应,处理单元 508 能配置为针对第二语音用户输入再次监视接收的音频输入(例如,使用第二监视单元 518)而不生成对第一语音用户输入的响应。

[0128] 在一些示例中,处理单元 508 能配置为确定是否应该响应第一语音用户输入(例如,使用响应确定单元 514)而不识别第一语音用户输入开始处的一个或多个预定词。在其它示例中,处理单元 508 能配置为确定(例如,使用响应确定单元 514)是否应该响应第一语音用户输入而不识别在接收第一语音用户输入之前接收的物理或虚拟按钮输入。

[0129] 在一些示例中,处理单元 508 能配置为通过对第一语音用户输入执行语音到文本转换、基于第一语音用户输入确定用户意图、基于第一语音用户输入确定要执行的任务、基于第一语音用户输入确定要执行的任务的参数、执行要执行的任务、显示对第一语音用户

输入的文本响应、或者输出对第一语音用户输入的听觉响应来生成响应（例如使用响应生成单元 516）。

[0130] 在一些示例中,处理单元 508 能配置为通过评估取决于与第一语音用户输入相关联的上下文信息的一个或多个条件性规则来确定（例如使用响应确定单元 514）是否响应第一语音用户输入。

[0131] 在一些示例中,处理单元 508 能配置为通过基于与第一语音用户输入相关联的上下文信息计算虚拟助理应该响应第一语音用户输入的可能度并且将可能度与阈值进行比较来确定（例如使用响应确定单元 514）是否响应第一语音用户输入。

[0132] 在一些示例中,处理单元 508 能配置为基于与第一语音用户输入相关联的上下文信息确定（例如使用响应确定单元 514）确定是否响应第一语音用户输入,上下文信息包括以下各项中的一个或多个:接收第一语音用户输入与之前用户输入之间经过的时间,之前语音用户输入,接收第一语音用户输入时用户与电子设备之间的距离,接收第一语音用户输入时电子设备的方向,接收第一语音用户输入时用户与电子设备之间的方向,接收第一语音用户输入时用户眼睛方向,第一语音用户输入是否被自动语音识别器识别的指示,第一语音用户输入与之前语音用户输入之间的语义关系,第一语音用户输入的长度,第一语音用户输入的扬声器的标识,接收第一语音用户输入的时间,接收第一语音用户输入时电子设备是否向用户输出信息的指示,期望从用户接收输入,接收第一语音用户输入时电子设备是否被握持的指示,接收第一语音用户输入时电子设备的操作状态,电子设备执行的之前动作,接收第一语音用户输入时电子设备是否在显示内容的指示,第一语音用户输入与接收第一语音用户输入时电子设备显示的内容之间的语义关系,接收第一语音用户输入时用户的位置,接收第一语音用户输入时用户执行的手势,电子设备的之前输出,接收第一语音用户输入时电子设备的位置,接收第一语音用户输入时电子设备运行的应用,之前联系人,之前电子邮件,之前 SMS 消息,接收第一语音用户输入时电子设备的移动,电子设备的用户设置,接收第一语音用户输入时电子设备感测的光量,日程数据。

[0133] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入与之前用户输入之间经过的时间的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分来确定（例如,使用响应确定单元 514）是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于经过的时间的值大于阈值持续时间减小可能性得分并且响应于经过的时间的值小于阈值持续时间增大可能性得分。在一些示例中,之前用户输入能包括之前接收的触敏显示器上的触摸输入,鼠标点击,按钮按下,或语音用户输入。

[0134] 在一些示例中,处理单元 508 能配置为通过基于包括之前语音用户输入的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分来确定（例如,使用响应确定单元 514）是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于检测到之前语音用户输入与第一语音用户输入之间的匹配而增大可能性得分。

[0135] 在一些示例中,处理单元 508 能配置为通过基于包括当接收第一语音用户输入时用户与电子设备之间的距离的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定（例如,使用响应确定单元 514）是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于距离大于阈值距离减小可能性得分并且响应于距离小于阈值距离增大可能性得分。在一些示例中,能至少部分基于第一语音用户输入的音量、接

- 近传感器测量的距离、图像传感器生成的图像、或来自加速度计的加速度数据来确定距离。
- [0136] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时电子设备的方向的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于设备方向朝下或颠倒减小可能性得分并且响应于设备方向朝上或直立增大可能性得分。
- [0137] 在一些示例中,处理单元 508 能配置为,通过基于包括接收第一语音用户输入时用户与电子设备之间的方向的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于方向是电子设备的显示器朝向用户增大可能性得到并且响应于方向是电子设备的显示器远离用户减小可能性得分。
- [0138] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时用户眼睛方向的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于用户眼睛方向指向电子设备增大可能性得分并且响应于用户眼睛方向指离电子设备减小可能性得分。
- [0139] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入是否被自动语音识别器识别的指示的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于指示表示第一语音用户输入被自动语音识别器识别增大可能性得分并且响应于指示表示第一语音用户输入不被自动语音识别器识别减小可能性得分。
- [0140] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与之前语音用户输入之间的距离关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于语音用户输入语义阈值增大可能性得分并且响应于语义关系的值小于语音用户输入语义阈值减小可能性得分。
- [0141] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入的长度的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于第一语音用户输入的长度小于阈值长度增大可能性得分并且响应于第一语音用户输入的长度大于阈值长度减小可能性得分。
- [0142] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入的发言者的标识的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于第一语音用户输入的发言者的标识是已知或授权发言者列表中的一个而增大可能性得分并且响应于第一语音用户输入的发言者的标识不是已知或授权发言者列表中的一个而减小可能性得分。在其它示例中,计算可能性得分能包括响应于第一语音用户输入的发言者的标识与之前语音用户输入的发言者的标识相同增大可能性得分并且响应

于第一语音用户输入的发言者的标识与之前语音用户输入的发言者的标识不同减小可能性得分。

[0143] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入的时间的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于时间在预定时间集合内增大可能性得分并且响应于时间不在预定时间集合内减小可能性得分。

[0144] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时电子设备是否正向用户输出信息的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于指示表示接收第一语音用户输入时电子设备正向用户输出信息增大可能性得分并且响应于指示表示接收第一语音用户输入时电子设备没正向用户输出信息减小可能性得分。

[0145] 在一些示例中,处理单元 508 能配置为通过基于包括从用户接收输入的预期的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于从用户接收输入的预期指示预期从用户接收到输入而增大可能性得分并且响应于从用户接收输入的预期指示不预期从用户接收到输入而减小可能性得分。

[0146] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时电子设备是否正被握持的指示的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于指示表示接收第一语音用户输入时电子设备正被握持增大可能性得分并且响应于指示表示接收第一语音用户输入时电子设备未被握持减小可能性得分。

[0147] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时电子设备的操作状态的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于电子设备的操作状态是一组预定操作状态中的一个增大可能性得分并且响应于电子设备的操作状态不是一组预定操作状态中的一个减小可能性得分。

[0148] 在一些示例中,处理单元 508 能配置为通过基于包括电子设备执行的之前动作的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于电子设备执行的之前动作是一组预定动作中的一个增大可能性得分并且响应于电子设备执行的之前动作不是一组预定动作中的一个减小可能性得分。

[0149] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时电子设备是否正显示内容的指示的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于指示表示接收第一语音用户输入时电子设备正显示内容增大可能性得分。

[0150] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与接收第一语音用户输入时电子设备正显示的内容之间的语义关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于内容语义阈值增大可能性得分并且响应于语义关系的值小于内容语义阈值减小可能性得分。

[0151] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时用户的位置的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于用户的位置是预定一组位置中的一个增大可能性得分并且响应于用户的位置不是预定一组位置中的一个减小可能性得分。

[0152] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时用户执行的手势的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于手势是预定手势集合中的一个增大可能性得分并且响应于手势不是预定手势集合中的一个减小可能性得分。

[0153] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与电子设备的之前输出之间的语义关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于之前输出语义阈值增大可能性得分并且响应于语义关系的值小于之前输出语义阈值减小可能性得分。

[0154] 在一些示例中,处理单元 508 能配置为通过基于包括接收第一语音用户输入时电子设备的位置的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于位置是预定位置集合中的一个增大可能性得分并且响应于位置不是预定位置集合中的一个减小可能性得分。

[0155] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与接收第一语音用户输入时电子设备运行的应用之间的语义关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于应用语义阈值增大可能性得分并且响应于语义关系的值小于应用语义阈值减小可能性得分。

[0156] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与之前联系人之间的语义关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于之前联系人语义阈值增大可能性得分并且响应于语义关系的值小于之前联系人语义阈值减小可能性得分。

[0157] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与之前电子邮件之间的语义关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于之前的电子邮件语义阈值增大可能性得

分并且响应于语义关系的值小于之前的电子邮件语义阈值减小可能性得分。

[0158] 在一些示例中,处理单元 508 能配置为通过基于包括第一语音用户输入与之前 SMS 消息之间的语义关系的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于语义关系的值大于之前的 SMS 消息语义阈值增大可能性得分并且响应于语义关系的值小于之前的 SMS 消息语义阈值减小可能性得分。

[0159] 在一些示例中,处理单元 508 能配置为通过基于包括电子设备的移动的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于移动是预定移动集合中的一个增大可能性得分并且响应于移动不是预定移动集合中的一个减小可能性得分。

[0160] 在一些示例中,处理单元 508 能配置为通过基于包括用户设置的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于用户设置是预定用户设置集合中的一个增大可能性得分并且响应于用户设置不是预定用户设置集合中的一个减小可能性得分。

[0161] 在一些示例中,处理单元 508 能配置为通过基于包括电子设备感测的光量的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于光量大于阈值光量增大可能性得分并且响应于光量小于阈值光量减小可能性得分。

[0162] 在一些示例中,处理单元 508 能配置为通过基于包括日程数据的上下文信息计算虚拟助理应该响应第一语音用户输入的可能性得分,来确定(例如,使用响应确定单元 514)是否响应第一语音用户输入。在这些示例中,计算可能性得分能包括响应于日程数据指示在接收第一语音用户输入时用户在忙减小可能性得分。

[0163] 处理单元 508 还能配置为监视音频输入(例如使用第二监视单元 518)以识别音频输入中的第二语音用户输入。在识别到音频输入中的第二语音用户输入时,处理单元 508 能配置为基于与第二语音用户输入相关联的上下文信息确定(例如,使用响应确定单元 514)是否应该生成对第二语音用户输入的响应。响应于确定应该生成响应,处理单元 508 能配置为生成对第二语音用户输入的响应(例如,使用响应生成单元 516),并且针对第三语音用户输入请求再次监视接收的音频输入(例如,使用第三监视单元 520)。响应于确定不应该生成响应,处理单元 508 能配置为针对第三语音用户输入请求再次监视接收的音频输入(例如,使用第三监视单元 520)而不生成对第二语音用户输入的响应。

[0164] 如上所述,本技术的一个方面是收集并使用从各种源可用的数据以改进向用户传送用户可能感兴趣的邀请内容或任何其它内容。本公开设想了一些情况下,此收集的数据能包括唯一标识或者能用于联系或定位特定人的个人信息数据。这种个人信息数据能包括人口统计数据、基于位置的数据、电话号码、电子邮件地址、家庭地址或任何其它识别信息。

[0165] 本公开意识到在本技术这种个人信息数据的使用能用于便利用户。例如,个人信息数据能用于向用户传送具有极大兴趣的目标内容。相应地,这种个人信息数据的使用

使得能够对传送内容进行计算控制。另外,本公开也设想了便利用户的个人信息数据的其它使用。

[0166] 本公开还设想了负责这种个人信息数据的收集、分析、传送、存储或其它使用的实体将遵循已被大家接受的隐私策略和 / 或隐私惯例。特别地,这些实体应该实现并持续使用通常认为满足或超出工业或政府要求的隐私策略和惯例以保持个人信息数据私密和安全。例如,来自用户的个人信息应该被收集用于实体的合法负责的使用,而不应在这些合法使用之外共享或出售。另外,这种收集应该仅在接收到用户的通知意愿之后发生。另外,这些实体会采取任何需要步骤来防护并保护对这种个人信息数据的访问并确保访问个人信息数据的人遵循他们的隐私策略和过程。另外,这些实体能使它们自己被第三方评估以证明它们遵循广泛接受的隐私策略和惯例。

[0167] 尽管上述,本公开还设想了用户选择性阻止使用或访问个人信息数据的示例。即,本公开设想了能提供硬件和 / 或软件元素以防止或阻止访问这种个人信息数据。例如,在广告传送服务的情况下,本技术能配置为允许用户在服务注册期间选择“参加”或“退出”个人信息数据收集的参与。在另一个示例中,用户能选择不提供用于目标内容传送服务的位置信息。在又一个示例中,用户能选择不提供精确位置信息,但允许传送位置区域信息。

[0168] 因此,尽管本公开广义涵盖了使用个人信息数据实现一个或多个公开的各种示例,但本公开还设想了各种示例也能在无需访问这种个人信息数据的情况下实现。即,本技术的各种示例并不会由于缺少全部或部分这种个人信息数据而变得不可操作。例如,能通过基于非个人信息数据或极少量个人信息(例如,与用户相关联的设备请求的内容,内容传送服务可用的其它非个人信息,或公共可用信息)推断喜好选择内容并传送到用户。

[0169] 尽管完全参照附图描述了示例,但应该注意,各种改变和修改将对本领域技术人员变得明显。这些改变和修改要理解为包括在所附权利要求定义的各种示例的范围内。

[0170] 相关申请的交叉引用

[0171] 本申请要求 2014 年 5 月 30 日提交的名称为 REDUCING THE NEED FOR MANUAL START/END-POINTING AND TRIGGER PHRASES 的美国临时序列号 62/005,760 和 2014 年 9 月 30 日提交的名称为 REDUCING THE NEED FOR MANUAL START/END-POINTING AND TRIGGER PHRASES 的美国非临时序列号 14/502,737 的优先权,在此通过引用将它们整体并入用于一切目的。

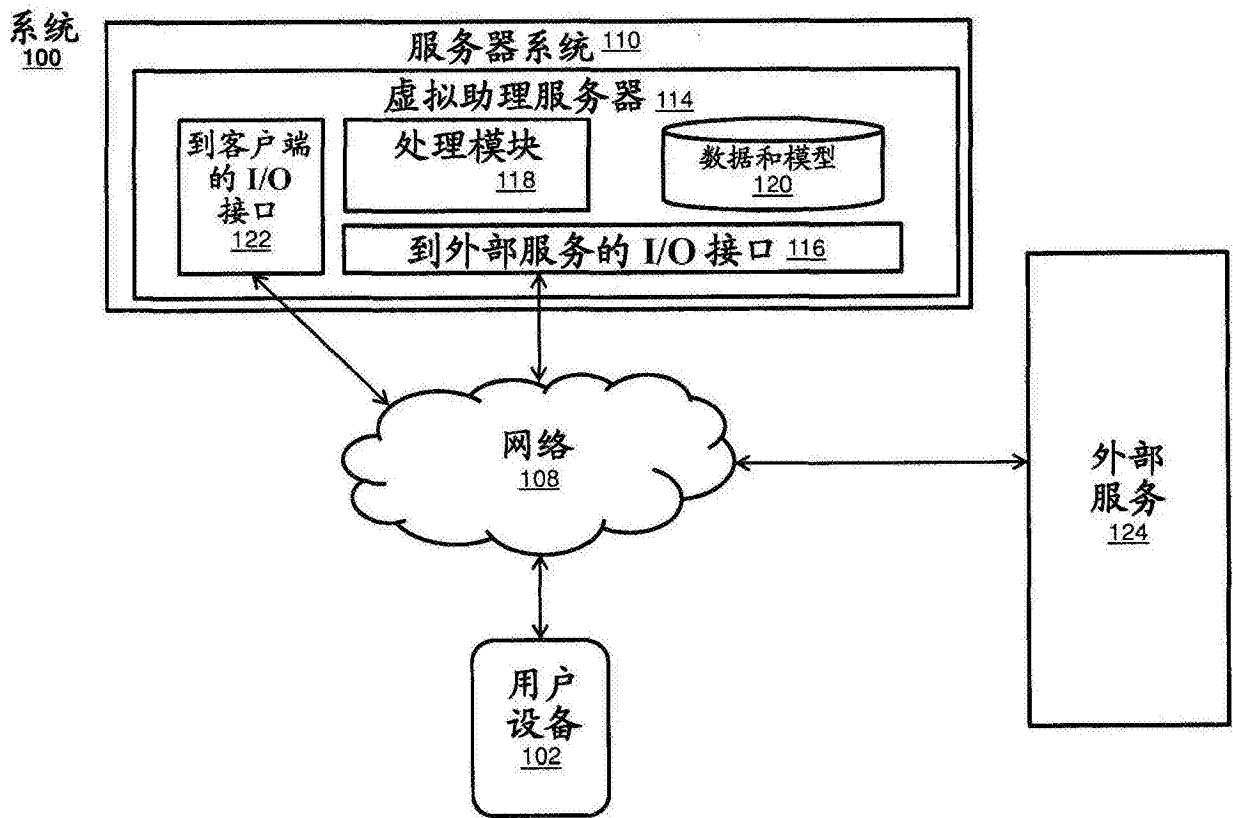


图 1

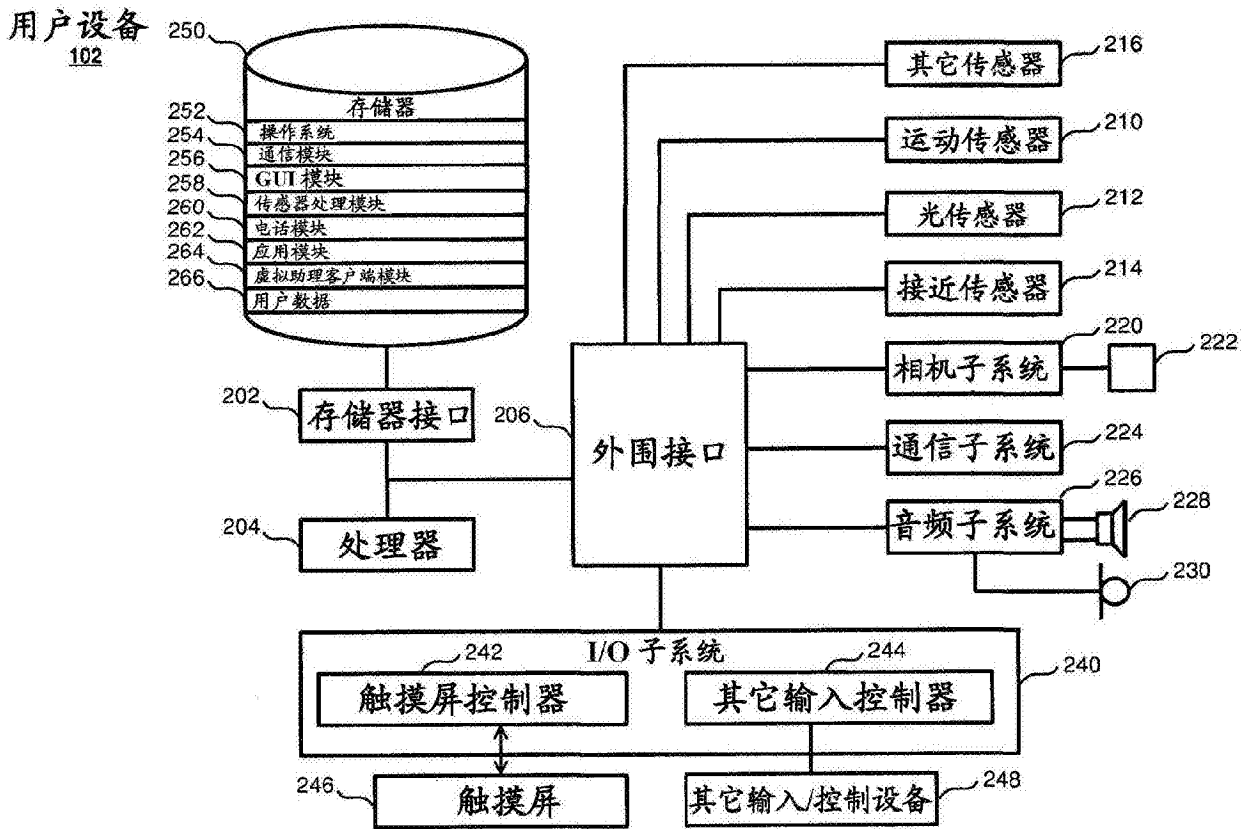


图 2

处理
300

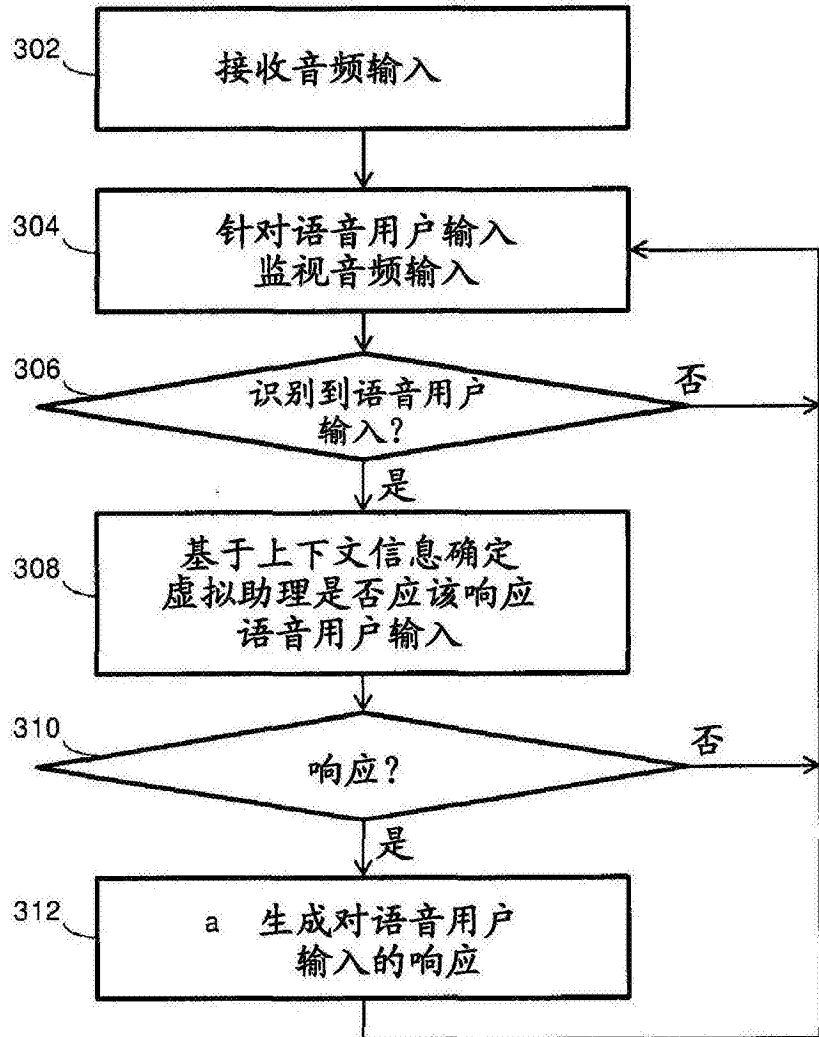


图 3

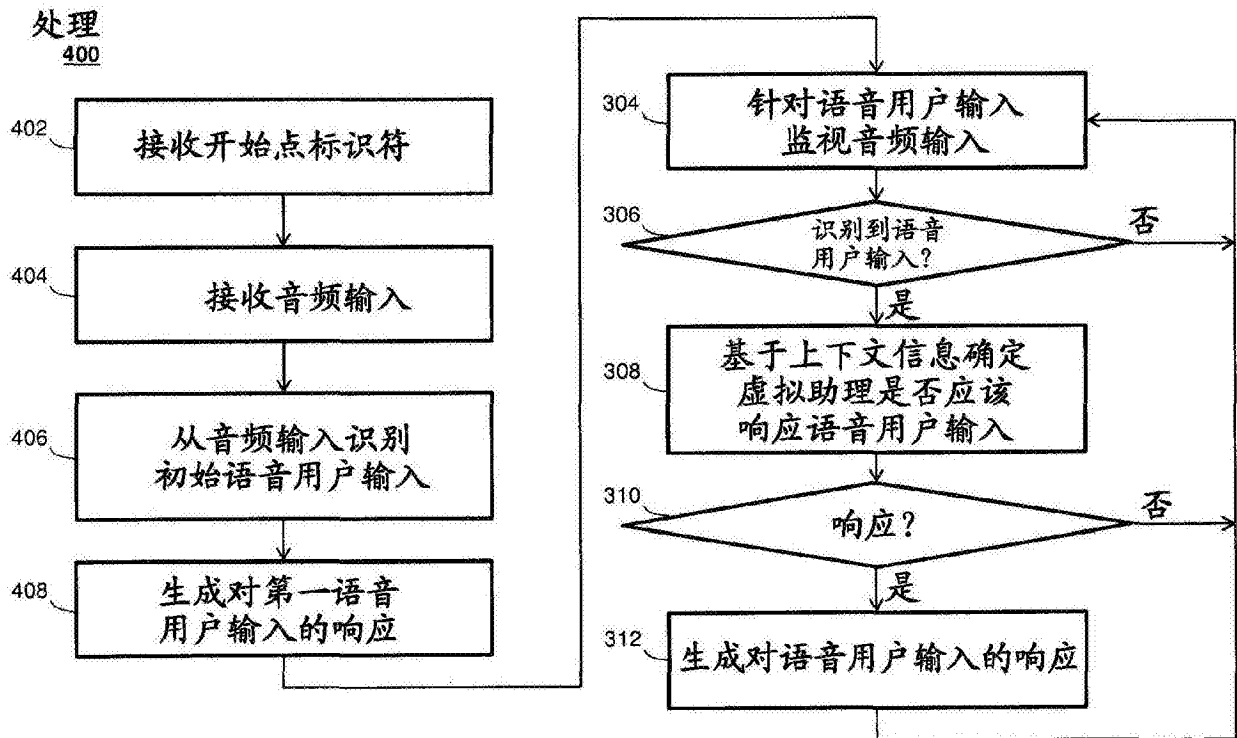


图 4

电子设备 500

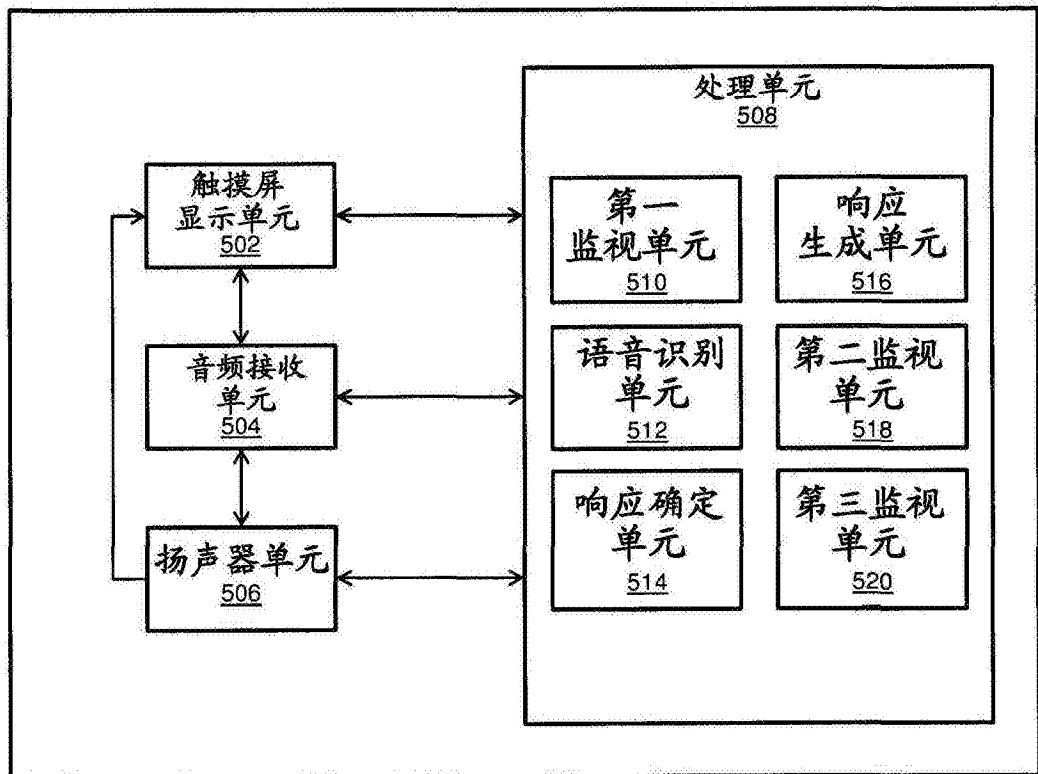


图 5