



(12) 发明专利申请

(10) 申请公布号 CN 113767379 A

(43) 申请公布日 2021. 12. 07

(21) 申请号 201980096007.9

(22) 申请日 2019.05.29

(30) 优先权数据

62/843,874 2019.05.06 US

(85) PCT国际申请进入国家阶段日

2021.11.01

(86) PCT国际申请的申请数据

PCT/US2019/034359 2019.05.29

(87) PCT国际申请的公布数据

WO2020/226663 EN 2020.11.12

(71) 申请人 谷歌有限责任公司

地址 美国加利福尼亚州

(72) 发明人 萨尔夫吉特·辛格

古斯塔沃·梅内塞斯·蓬特·莫雷拉

格雷迪·西蒙 彼得·布兰特

图沙尔·钱德拉

(74) 专利代理机构 中原信达知识产权代理有限责任公司 11219

代理人 邓聪惠 周亚荣

(51) Int.Cl.

G06F 16/242 (2006.01)

G06F 16/332 (2006.01)

G06F 16/732 (2006.01)

G06F 16/74 (2006.01)

G06F 16/93 (2006.01)

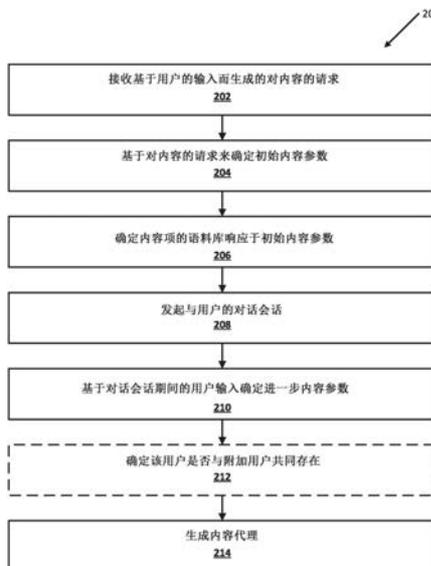
权利要求书4页 说明书17页 附图6页

(54) 发明名称

使用内容代理和/或存储的内容参数来渲染内容

(57) 摘要

公开了一种技术,该技术能够基于从对内容的初始用户请求以及对话会话而确定的内容参数来生成内容代理,以进一步细化该对内容的请求。各种实现方式包括响应于附加用户请求而使用内容代理来渲染附加内容。附加或可替代的实现方式包括使用内容代理来向用户主动渲染内容。



1. 一种由一个或多个处理器实现的方法,所述方法包括:
  - 接收对内容的请求,所述对内容的请求基于用户经由客户端设备的一个或多个用户接口输入设备的输入而生成;
  - 基于用户请求来确定一个或多个初始内容参数;
  - 确定内容项的语料库响应于所述初始内容参数;
  - 响应于确定所述内容项的语料库响应于所述初始内容参数;
  - 经由所述客户端设备发起与所述用户的对话会话,以及
  - 基于所述对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数,所述进一步内容参数细化所述请求并限制所述语料库中响应于所述请求的内容项的数量;
  - 确定在所述请求时,所述用户与至少一个附加用户共同存在;
  - 基于与所述请求相关联的初始内容参数以及基于所述对话会话期间的用户输入确定的进一步内容参数来生成内容代理;
  - 使所述内容代理与与所述附加用户共同存在的用户相关;
  - 在生成所述内容代理之后:
  - 确定所述用户与所述附加用户共同存在的附加实例;
  - 响应于所述内容代理与与所述附加用户共同存在的用户相关;
  - 使用所述内容代理来主动渲染符合所述初始内容参数和所述进一步内容参数的给定内容。
2. 根据权利要求1所述的方法,其中经由所述客户端设备发起与所述用户的对话会话以及基于所述对话会话期间的一个或多个用户输入来确定所述一个或多个进一步内容参数还响应于:
  - 确定所述内容项的语料库中的内容项的数量超过阈值数量。
3. 根据任一前述权利要求所述的方法,其中经由所述客户端设备发起与所述用户的对话会话包括:
  - 基于所述内容项的语料库中的内容项中的一个或多个的一个或多个特征,生成针对所述内容项中的一个或多个定制的提示;以及
  - 响应于所述用户请求,并且在所述对话会话期间的一个或多个用户输入之前渲染所述提示。
4. 根据任一前述权利要求所述的方法,还包括:
  - 在生成所述内容代理之后:
  - 确定响应于所述初始内容参数的一个或多个新的内容项以包括在所述内容项的语料库中,其中所述新的内容项是在确定所述内容项的语料库之后创建的;
  - 基于所述新的内容项中的一个或多个的一个或多个新特征来生成针对所述新的内容项中的一个或多个定制的新提示;
  - 通过提供所述新提示经由所述客户端设备发起与所述用户的后续对话会话;
  - 基于所述后续对话会话期间的一个或多个后续用户输入来确定一个或多个后续内容参数,所述后续内容参数进一步细化所述请求并限制所述语料库中响应于所述请求的新的内容项的数量;以及
  - 基于所述后续内容参数来修改所述内容代理。

5. 根据任一前述权利要求所述的方法,还包括:

在生成所述内容代理之后并在主动渲染符合所述初始内容参数和所述进一步内容的给定内容之前:

使用所述内容代理来确定响应于所述用户请求的内容;以及

使得计算系统响应于所述用户请求来渲染所述内容。

6. 根据任一前述权利要求所述的方法,还包括:

在使用所述内容代理来主动渲染符合所述初始内容参数和所述进一步内容参数的给定内容之后:

通过经由所述客户端设备发起与所述用户的附加对话会话来确定一个或多个附加内容参数,所述附加内容参数进一步细化所述请求并进一步限制所述语料库中响应于所述请求的内容项。

7. 根据任一前述权利要求所述的方法,其中所述内容项的语料库包括至少两个媒体类型的内容项。

8. 根据任一前述权利要求所述的方法,其中所述内容项的语料库包括所述用户未查看的内容项,并且其中使用所述内容代理来主动渲染所述给定内容包括在从所述内容项的语料库中选择所述给定内容时对所述用户未查看的内容项进行优先排序。

9. 根据任一前述权利要求所述的方法,还包括:

在使用所述内容代理来主动渲染符合所述初始内容参数和所述进一步内容参数的给定内容之后:

接收指示所述用户是否与所述给定内容交互的一个或多个用户反馈信号;以及

基于所述一个或多个用户反馈信号来更新所述内容代理。

10. 根据权利要求9所述的方法,其中所述一个或多个用户反馈信号包括指示所述用户与所述内容交互的正反馈信号。

11. 根据权利要求9所述的方法,其中所述一个或多个用户反馈信号包括指示所述用户未与所述内容交互的负反馈信号。

12. 根据任一前述权利要求所述的方法,其中所述客户端设备包括自动助理客户端,其中所述请求在所述自动助理客户端的助理接口处被接收,并且其中所述对话会话是与所述用户和所述客户端设备的自动助理客户端的对话会话。

13. 根据权利要求12所述的方法,其中所述用户与所述自动助理客户端的用户简档相关联,并且还包括:

确定所述用户与使用附加客户端设备的所述附加用户共同存在的另一实例,

其中所述附加客户端设备与所述客户端设备不同,并且

其中所述用户与所述附加用户共同存在的所述另一实例是在所述用户与所述附加用户共同存在的所述附加实例之后;以及

使用所述内容代理来使用所述附加客户端设备主动渲染附加给定内容。

14. 根据权利要求12所述的方法,其中所述附加用户与所述客户端设备的已知用户简档相关联,并且其中确定所述用户与所述附加用户共同存在的所述附加实例包括:

确定所述用户与与所述客户端设备的已知用户简档相关联的附加用户共同存在的附加实例。

15. 根据权利要求12所述的方法,其中所述附加用户不与所述客户端设备的已知用户简档相关联。

16. 根据任一前述权利要求所述的方法,还包括:

确定在所述请求时的请求的附加场境;以及

其中使所述内容代理与与所述附加用户共同存在的用户相关还包括,使所述内容代理与在所述请求的附加场境中与所述附加用户共同存在的用户相关。

17. 根据权利要求16所述的方法,其中所述附加场境自动地被确定,并且至少部分基于来自所述客户端设备的一个或多个传感器的传感器数据加以确定。

18. 一种由一个或多个处理器实现的方法,所述方法包括:

接收对内容的请求,所述对内容的请求基于用户经由客户端设备的一个或多个用户接口输入设备的输入而生成;

基于所述用户请求来确定一个或多个初始内容参数;

确定内容项的语料库响应于所述初始内容参数;

响应于确定所述内容项的语料库响应于所述初始内容参数:

经由所述客户端设备发起与所述用户的对话会话,以及

基于所述对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数,所述进一步内容参数细化所述请求并限制所述语料库中响应于所述请求的内容项的数量;

确定在所述请求时,所述用户与至少一个附加用户共同存在;

基于与所述请求相关联的初始内容参数以及基于所述对话会话期间的用户输入确定的进一步内容参数来生成内容代理;

使所述内容代理与与所述附加用户共同存在的用户相关;

在生成所述内容代理之后:

接收对内容的附加请求,所述对内容的附加请求基于所述用户经由所述客户端设备的一个或多个用户接口输入设备的附加输入而生成;

确定所述用户与所述附加用户共同存在的附加实例;

响应于所述内容代理与与所述附加用户共同存在的用户相关:

响应于所述附加用户请求,使用所述内容代理来确定附加内容;以及

使得所述客户端设备渲染所述附加内容。

19. 一种由一个或多个处理器实现的方法,所述方法包括:

接收对内容的请求,所述对内容的请求基于用户经由客户端设备的一个或多个用户接口输入设备的输入而生成;

基于所述请求来确定一个或多个初始内容参数;

确定内容项的语料库响应于所述初始内容参数;

响应于确定所述内容项的语料库响应于所述初始内容参数:

经由所述客户端设备发起与所述用户的对话会话,以及

基于所述对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数,所述进一步内容参数细化所述请求并限制所述语料库中响应于所述请求的内容项的数量;

基于所述对内容的请求和/或基于来自所述客户端设备的传感器数据自动地确定在所述请求时的一个或多个场境条件;

基于与所述请求相关联的初始内容参数以及基于所述对话会话期间的用户输入确定的进一步内容参数来生成内容代理；

使所述内容代理与所述一个或多个场境条件相关；

在生成所述内容代理之后：

确定所述一个或多个场境条件的出现的附加实例；

响应于所述内容代理与所述一个或多个场境条件相关；

使用所述内容代理来主动渲染符合所述初始内容参数和所述进一步内容参数的给定内容。

20. 根据权利要求19所述的方法，其中所述一个或多个场境条件包括所述用户在所述请求时正在参加特定活动。

21. 根据权利要求19或权利要求20所述的方法，其中所述一个或多个场境条件包括所述用户与至少一个附加用户在所述请求时共同存在。

22. 一种包括指令的计算机程序，所述指令在由计算系统的一个或多个处理器执行时使所述计算系统执行根据任一前述权利要求所述的方法。

23. 一种计算系统，其被配置成执行权利要求1至21中的任一项所述的方法。

24. 一种存储指令的计算机可读存储介质，所述指令可由计算系统的一个或多个处理器执行以执行根据权利要求1至21中的任一项所述的方法。

## 使用内容代理和/或存储的内容参数来渲染内容

### 背景技术

[0001] 人类可以使用本文称为“自动助理”（也称为“数字代理”、“聊天机器人”、“交互式个人助理”、“智能个人助理”、“助理应用”、“对话代理”等）的交互式软件应用进行人机对话。例如，人类（当它们与自动助理交互时，它们可以被称为“用户”）可以使用口语自然语言输入（即，话语）和/或通过提供文本（例如，键入的）自然语言输入向自动助理提供命令和/或请求，在一些情况下，该口语自然语言输入可以被转换成文本并且然后被处理。自动助理通过提供响应用户接口输出来响应于请求，该输出可以包括听觉和/或视觉接口输出。

[0002] 用户可以与自动助理和/或其他应用交互，以请求各种数字内容，包括：视频、音频、图像、网页、新闻文章、视频、食谱、博客帖子、电子书、模因、社交媒体帖子和/或附加类型的内容。

### 发明内容

[0003] 本文描述的实现方式涉及使用从用户请求确定的内容参数以及在用户与客户端设备之间的对话会话期间确定的内容参数来生成内容代理。如本文所描述，遵循用户请求的用户与客户端设备之间的对话会话可以由客户端设备的应用（任选地与远程应用组件对接）发起和/或引导。该应用可以是例如本文更详细描述自动助理。此外，在对话会话期间提供的提示可以由应用基于响应于用户请求的初始内容语料库的特征来生成，并且可以以限制该初始内容语料库的目标来生成，从而限制在确定内容代理响应于用户请求和/或响应于内容代理（主动的和/或响应于未来的用户请求）对内容的未来渲染而渲染哪些内容时所考虑的内容数量。以这些和其他方式，可以减少由内容代理考虑的内容项的数量，导致对要渲染的内容的计算效率更高的考虑和选择。此外，在各种实现方式中，试图限制初始内容语料库的对话会话的发起可以是选择性的，并且仅响应于确定存在一个或多个条件而发生。此类条件可以包括，例如，内容项的初始语料库包括至少阈值数量的内容项，内容项的初始语料库（和/或相关联的索引）至少是阈值存储大小，和/或其他条件。以这些和其他方式，对话会话只能被选择性地发起，从而平衡通过约束内容语料库节省的计算效率和参加对话会话所需的资源。因此，可以提高用户与客户端设备交互和控制客户端设备的效率，减少用户与客户端设备响应之间的摩擦，并减少计算开销两者。

[0004] 在本文描述的许多实现方式中，所生成的内容代理及其相关联的内容参数是持久的，并且继续在多个不同的时间跨多个不同的会话选择性地渲染来自语料库的内容，如由所确定的内容参数所约束的。通过内容代理和/或内容参数的持久性，可以避免再次建立内容参数的后续请求和对话会话（以及它们相关联的使用网络和/或计算资源）。此外，由于符合内容参数的新内容项被添加到语料库，内容代理可以选择性地发起新的对话会话来建立附加的内容参数。在新对话会话期间提供的提示可以任选地基于这些新的内容项的特征（例如，先前不存在于语料库的内容项中的特征）来生成，以确定是限制还是包括来自该语料库的此类新的内容项。这使得语料库能够继续受到约束，并且对话会话可以任选地在提供任何新的内容项之前被发起，以防止渲染用户不寻求的内容项。此外，对话会话可以任选

地响应于条件(例如,上文提到的那些)的存在而被选择性地发起,从而平衡对话会话资源需求和通过受约束语料库获得的资源节省。

[0005] 作为一些实现方式的一个示例,用户可以请求“向我展示如何设置我的路由器”。可以从“向我展示如何设置我的路由器”的初始请求中确定一个或多个初始参数,诸如“设置”和“路由器”的初始内容参数。可以经由客户端设备发起与用户的对话会话以细化该内容的请求。在请求“向我展示如何设置我的路由器”后发起的示例对话会话可以包括:客户端设备-“您的路由器是什么品牌的”;用户-“假想路由器品牌”;客户端设备-“您的路由器是什么型号”;用户-“假想的型号”;客户端设备“您在设置路由器的特定步骤中遇到问题了吗”;用户-“否”。可以从该对话会话中确定进一步内容参数,诸如“假想路由器品牌”、“假想型号”、“所有设置步骤”以及附加的和/或替代的内容参数。初始内容参数(例如,设置、路由器)可以用于识别响应于初始请求的内容语料库。另外地或替代地,进一步内容参数(例如,假想路由器品牌、假想型号、所有设置步骤)可以用于细化内容语料库。在许多实现方式中,进一步内容参数可以用于通过例如限制内容语料库中的内容项数量来细化该内容语料库。在多种实现方式中,该内容代理可以基于与请求相关联的初始参数以及基于对话会话期间的用户输入而确定的进一步内容参数来生成。内容参数可以另外地或替代地基于用户是否和/或如何与内容项交互来确定。例如,如果用户选择(例如,通过触摸输入)与特定路由器型号相关的内容项,则系统可以推断用户可能具有该型号的路由器,并确定与路由器类型相关的内容参数。作为另一示例,如果用户选择与特定路由器型号相关的多个内容项和/或在至少阈值持续时间内查看内容项,则系统可以确定与路由器类型相关的内容参数。

[0006] 内容代理可以用于渲染内容语料库中与内容代理相对应的一个或多个内容项。在各种实现方式中,用户提供对内容的附加请求(其中由对内容的附加请求所请求的内容被包括在内容语料库中),并且内容代理可以用于确定该内容语料库中响应于附加请求的内容项。例如,可以基于对内容的初始请求和/或对“向我展示晚餐食谱”的请求的对话会话来生成食谱内容代理。当用户进行“向我展示早餐食谱”的后续请求时,食谱内容代理可以用于确定内容项。后续请求可以在初始请求之后的多个时间进行,包括:后续请求在与初始请求相同的对话会话中进行;后续请求与初始请求在同一天的晚些时候进行;后续请求是在初始请求的第二天进行;后续请求在初始请求一周后进行;和/或后续请求在初始请求之后的附加时间长度内进行。

[0007] 另外地或替代地,内容代理可以用于向用户主动渲染内容项。例如,当用户在附加的时间实例与客户端设备交互时,上述食谱内容代理可以用于主动渲染附加食谱。可以使用内容代理在多个时间渲染主动内容,包括:在与初始请求相同的对话会话中;在下一个对话会话中;初始请求后的第二天;初始请求后的一周;和/或附加的时间长度。

[0008] 在许多实现方式中,用于生成内容代理的内容参数(即,从初始请求确定的内容参数和/或从后续对话会话确定的内容参数)被用于细化附加和/或替代内容语料库中的内容。例如,当生成餐馆选择内容代理时,可以确定“对花生过敏”的内容参数。在许多实现方式中,该内容参数(即,对花生过敏)可以用于生成和/或细化附加和/或替代内容代理。例如,对花生过敏内容参数可以用于细化与食谱内容代理相关的内容语料库。换句话说,在生成第一内容代理(餐馆选择内容代理)时使用的用户对花生的过敏(对花生过敏内容参数)可以被主动用于生成第二内容代理(食谱内容代理)。另外地或替代地,内容参数可以任选

地随时间变化,诸如内容参数随时间衰减。换句话说,内容代理可以通过随时间减少一个或多个内容参数的影响和/或随时间移除一个或多个内容参数来更新。作为一个示例,指示对辛辣食物的偏好的内容参数可以从“我想吃辛辣食物”的初始请求中确定。如果没有偏好的再确认(例如,通过选择“辛辣食物”内容项和/或对辛辣食物的另一请求来再确认),则指示对辛辣食物的偏好的内容参数会随着时间而衰减。可以更新内容代理,使得在一段时间(例如,一天、一周、一个月和/或附加的时间长度)内,可以从与内容代理相关联的内容语料库中移除辛辣食物推荐。

[0009] 随着时间过去,新的内容可以被添加到内容语料库。如本文所使用,新的内容指示:最初不包括在内容语料库中的内容;在确定内容语料库之后创建的内容;用户没有与之交互的内容语料库内的内容(例如,未观看的视频、未阅读的文章和/或用户尚未与之交互的附加内容);和/或对于用户来说是新的附加内容。在许多实现方式中,用户与内容的交互可以包括选择、选择和观看至少一个阈值持续时间,和/或与内容的其他交互。在若干实现方式中,当向用户渲染内容时,内容代理可以对新的内容进行优先排序。例如,新闻内容代理可以对渲染新的新闻内容进行优先排序。

[0010] 在许多实现方式中,附加对话会话可以用于进一步细化内容代理的内容语料库。用户可以发起该附加对话会话。例如,用户可以响应于内容代理被用于渲染不想要的内容而发起附加的对话会话。另外地或替代地,可以将新的内容添加到内容语料库(例如,在生成内容代理时不可用的内容),并且可以基于该新的内容发起附加的对话会话。例如,家用机器人套件内容代理可以使用“向我展示我的假想家用机器人套件的项目”的初始请求(以及对应的对话会话)来生成。可以为假想家用机器人套件发布新的部分,并且可以识别与新发布的部分相关的对应内容。在多种实现方式中,系统可以主动发起与用户的对话会话,请求该新的内容(即,与新发布的机器人套件部分相关的内容)是否应被添加到与家用机器人套件内容代理相对应的内容语料库中。类似地,在附加和/或替代实现方式中,用户可以发起对话会话,用于在家用机器人套件内容代理的内容语料库中包括和/或排除新的内容。例如,用户可以通过说“我没有新发布的机器人套件”来提供明确排除内容的反馈。

[0011] 内容代理可以与用户的用户简档相关联,该用户进行对内容的初始请求并参与对话会话以细化该对内容的请求。已知用户简档可以基于一个或多个确定的特性而与用户相关联,诸如语音签名(例如,文本无关语音识别、文本相关语音识别和/或其他语音识别)、面部辨识、图像馈送、运动特性和/或其他数据。在许多实现方式中,内容代理可以与用户简档相关联,从而使得内容代理能够用于跨用户的许多设备向用户推荐内容。例如,用户可以向第一客户端设备(例如,独立交互式扬声器)提供初始请求和对话会话。内容代理可以与用户的用户简档相关联。当用户与第二客户端设备(例如,移动电话)交互时,与用户简档相关联的内容代理可以用于为第二客户端设备的用户渲染内容。

[0012] 附加或替代实现方式中的内容代理与用户相关,该用户在与一个或多个附加用户共同存在的同时创建内容代理。例如,凯瑟琳(Katherine)可以进行对内容的初始请求以及参加对话会话,以进一步细化该对内容的请求。客户端设备可以使用传感器数据在请求时确定/检测凯瑟琳与安(Ann)(即,附加用户)共同存在。传感器数据可以例如包括一个或多个语音签名和/或捕获的图像/视频。例如,检测另一用户共同存在可以包括以下各项中的一者或多者:检测多个语音签名(例如,文本无关语音识别、文本相关语音识别和/或其他语

音识别)；使用面部辨识来检测附加用户；在图像馈送中识别多个人；识别多个人的运动特性；和/或指示附加用户的共同存在的其他数据。响应于凯瑟琳的初始请求而生成的内容代理可以与与安共同存在的凯瑟琳相关。在一些实现方式中，内容代理可以对应于与任何附加用户共同存在的凯瑟琳(例如，如果凯瑟琳与山姆(Sam)、汤姆(Tom)、玛丽(Mary)和/或任何附加用户共同存在，则利用该内容代理)。在一些实现方式中，内容代理可以对应于与特定附加用户共同存在的凯瑟琳(例如，凯瑟琳必须与安共同存在，内容代理才能用于为凯瑟琳渲染附加内容)。以这些和其他方式，可以为给定用户生成不同的内容代理，每一内容代理具有不同的对应内容参数和不同的共同存在要求。例如，给定用户可以具有：第一音乐代理，其具有第一内容参数和给定用户单独存在的要求(即，不与任何其他用户共同存在)；第二音乐代理，其具有第二内容参数和给定用户与一个或多个特定用户(例如，用户的孩子)共同存在的要求；以及第三音乐代理，其具有第三内容参数和给定用户与一个或多个附加特定用户(例如，配偶)但不与一个或多个特定用户(例如，孩子)共同存在的要求。因此，每一内容代理可以仅响应于给定的共同存在要求而被激活(例如，用于主动或响应性地渲染内容)。针对不同共同存在情况的不同内容代理的此类利用使得能够在适当的情况下渲染适当的内容，并且防止对于给定共同存在情况用户不寻求的内容的浪费性渲染。在一些实现方式中，不同的内容代理可以用于各种场境(context)条件，包括除共同存在之外的场境条件。例如，用户在进行初始请求时参加的活动可以被考虑并与内容代理相关联。可以利用用户的活动来代替共同存在或者与共同存在一起使用。例如，活动可以包括驾驶车辆、乘坐公共交通工具、在健身房锻炼、在家看电影和/或附加活动。来自客户端设备的一个或多个信号可以用于确定场境，诸如客户端设备的当前位置、客户端设备的运动和/或附加信号。作为一个示例，捕获的运动信号(例如，来自客户端设备的加速度计的数据、快速变化的GPS坐标等)可以指示用户正在驾驶他们的汽车。可以确定汽车中任何附加乘客的共同存在。驾驶汽车时的这种共同存在可以用于激活内容代理。例如，确定用户在驾驶时与他们的家人共同存在可以激活第一内容代理，而用户在驾驶时与他们的朋友共同存在可以激活第二内容代理。另外地或替代地，用户可以在进行初始请求和/或参与对话会话时提供场境，以细化内容语料库。例如，初始请求“我正在和我的配偶一起看电视。请向我展示电影推荐”可以被处理以确定“看电视”(特定活动)和用户与其配偶共同存在的场境条件。场境条件可以与内容代理相关，并且内容代理随后响应于确定一个或多个(例如，所有)场境条件的再次出现而被激活。在多种实现方式中，基于对内容的请求和/或基于来自客户端设备的传感器数据，可以自动确定初始请求时的场境。

[0013] 因此，各种实现方式阐述了用于实现到用户期望内容的导航的技术，并且通过使用使用用户定义的内容参数生成的内容代理，以更高效的方式实现内容导航。例如，基于从初始用户对内容的请求中导出的初始内容参数来识别候选期望内容的语料库。在与用户的对话会话期间确定附加内容参数，并且初始内容参数以及附加内容参数可以用于生成内容代理。这些附加内容参数进一步细化内容语料库(例如，通过基于附加内容参数移除内容项来细化内容语料库)。附加对话会话(用户发起的对话和/或内容代理发起的对话)可以用于识别进一步内容参数，这些参数又可以用于更新内容代理并进一步细化所需用户内容的语料库。随着内容代理(以及对应内容的语料库)被进一步细化，内容代理可以用于更有效地识别用户期望的内容。

[0014] 通过使用内容代理确定用户期望的内容,可以节省系统资源(例如,处理器周期、存储器、电池电量和/或计算系统的附加资源)。例如,通过在渲染内容时利用生成的内容代理,用户对期望内容的未来请求可以是更短的请求,并且可以通过内容参数的持久性来避免未来请求中的对话会话。此类较短的请求和/或缺少对话会话可以缩短整个用户-系统交互的持续时间,通过使交互更早停止来节省各种资源。此外,内容代理可以用于为用户主动渲染期望的内容。期望内容的该主动渲染减少了用户与客户端设备的交互(例如,通过消除用户请求期望内容的需要),这又进一步减少了系统资源的使用。

[0015] 其他实现方式可以包括存储指令的非暂时性计算机可读存储介质,该指令可由一个或多个处理器(例如,中央处理单元(CPU))、图形处理单元(GPU)和/或张量处理单元(TPU)执行以执行诸如上文和/或在本文别处所描述的一个或多个方法的方法。其他实现方式可以包括一个或多个计算机和/或一个或多个机器人的系统,该系统包括一个或多个处理器,这些处理器可操作来执行存储的指令以执行诸如上文和/或在本文别处所描述的一个或多个方法的方法。

[0016] 应了解,本文更详细描述的前述概念和附加概念的所有组合都被认为是本文所公开的主题的一部分。例如,出现在本公开末尾的要求保护的的主题的所有组合被认为是本文所公开的主题的一部分。

## 附图说明

[0017] 图1A示出其中可以实现本文所公开的多个实现方式的示例环境。

[0018] 图1B示出根据本文描述的多个实现方式的示例状态图。

[0019] 图2是示出根据本文描述的多个实现方式生成内容代理的示例过程的流程图。

[0020] 图3是示出根据本文描述的多个实现方式使用内容代理来确定内容的示例过程的流程图。

[0021] 图4示出其中可以实现本文所公开的实现方式的另一示例环境。

[0022] 图5示出计算设备的示例架构。

## 具体实施方式

[0023] 图1A示出其中可以实现本文所公开的实现方式的示例环境。图1A的示例环境包括一个或多个用户接口输入设备102、一个或多个用户接口输出设备104和客户端设备106。另外地或替代地,客户端设备106可以与内容116和/或内容代理118相关联。客户端设备106可以包括内容代理引擎108,用于生成内容代理118以及用于为用户渲染来自内容116的语料库的内容。在许多实现方式中,内容代理引擎108可以包括多种模块,包括代理模块110、对话模块112、内容模块114和/或附加模块(未描绘)。用户接口输入设备102可以包括例如,物理键盘、触摸屏(例如,实现虚拟键盘或其他文本输入机制)、麦克风和/或相机。用户接口输出设备104可以包括例如显示屏和/或扬声器。用户接口输入设备102和用户接口输出设备104可以与用户的一个或多个客户端设备106合并。例如,用户的移动电话可以包括用户接口输入设备102和用户接口输出设备104;独立个人助理硬件设备可以包括用户接口输入设备102和用户接口输出设备104;或者第一计算设备可以包括用户接口输入设备102,并且单独计算设备可以包括用户接口输出设备104;等。

[0024] 尽管客户端设备106在图1A中被示为与用户接口输入设备102和用户接口输出设备104分离,但在一些实现方式中,计算系统106的所有或方面可以实现在客户端设备上,该客户端设备还包含用户接口输入设备102和/或用户接口输出设备104。在一些实现方式中,客户端设备106可以包括自动助理(未描绘),并且自动助理的所有或方面可以实现在与计算设备分离和远离的计算设备上,该计算设备包含用户接口输入设备102和/或用户接口输出设备104(例如,所有或方面可以“在云中”实现)。在那些实现方式中的一些中,自动助理的那些方面可以经由一个或多个网络,诸如局域网(LAN)和/或广域网(WAN)(例如,因特网),与计算设备通信。

[0025] 可以包括用户接口输入设备102和/或用户接口输出设备104的客户端设备106的一些非限制性示例包括以下各项中的一者或多者:台式计算设备、膝上型计算设备、至少部分专用于自动助理的独立硬件设备、平板型计算设备、移动电话计算设备、用户车辆的计算设备(例如,车载通信系统、车载娱乐系统、车载导航系统),或者包括计算设备的用户可穿戴装置(例如,具有计算设备的用户手表、具有计算设备的用户眼镜、虚拟或增强现实计算设备)。可以提供附加和/或替代的计算系统。

[0026] 客户端设备106可以包括用于存储数据和软件应用的一个或多个存储器、用于访问数据和执行应用的一个或多个处理器、以及促进网络通信的其他组件。由客户端设备106执行的操作可以分布在多个计算设备上。例如,在一个或多个位置的一个或多个计算机上运行的计算程序可以通过网络相互耦合。

[0027] 如图1A中所示,用户经由用户接口输入设备102向客户端设备106提供输入。客户端设备经由用户接口输出设备104(任选地在由一个或多个组件进一步处理之后)提供用于呈现给用户的响应输出。响应内容可以包括数字内容,包括:视频、音频、图像、网页、新闻文章、视频、食谱、博客帖子、电子书、模因、社交媒体帖子和/或用户可以经由客户端设备请求渲染的附加类型的内容。

[0028] 客户端设备106可以利用内容代理引擎108来生成内容代理,在为用户渲染内容时使用内容代理,和/或使用内容代理采取附加动作。在许多实现方式中,内容代理108可以包括多种模块,包括:代理模块110、对话模块112、内容模块114和/或附加模块(未示出)。

[0029] 内容代理引擎108可以利用代理模块110:识别用户经由用户接口输入设备102对内容的初始请求中的内容参数;识别用户与客户端设备之间的对话中的内容参数以细化内容语料库(例如,识别由对话模块112发起的对话中的内容参数);生成一个或多个内容代理118;和/或更新现有内容代理118。例如,用户可以经由“在我的家用机器人套件上给我播放视频”的用户接口输入设备提供对内容的初始口头请求。代理模块110可以确定初始口头请求(例如,视频、机器人套件)中的一个或多个内容参数。另外地或替代地,代理模块110可以确定对话中的一个或多个内容参数来细化请求(诸如由对话模块112发起的对话)。在多种实现方式中,代理模块110可以使用初始内容参数以及从与用户的对话中确定的附加内容参数来生成内容代理。此外,代理模块110可以更新现有内容代理(例如,基于从附加对话会话确定的其他内容参数来更新现有内容代理)。

[0030] 在多种实现方式中,代理模块110可以用于当用户做出对内容的初始请求时,确定用户与附加用户共同存在。例如,使用用户接口输入设备102捕获的数据可以用于确定附加用户的共同存在,诸如:使用麦克风捕获的附加用户的口语自然语言、使用相机捕获的附加

用户的图像、和/或指示附加用户的共同存在的附加捕获的传感器数据。在多种实现方式中,代理模块110可以使附加用户的共同存在与所生成的内容代理相关联。换句话说,当确定用户与用户共同存在的附加实例时,对应于附加用户的共同存在的内容代理被用于渲染未来内容。例如,可以使用来自用户詹姆斯(James)的初始请求来生成内容代理,其中确定詹姆斯与用户迈克尔(Michael)共同存在。在一些实现方式中,内容代理可以与与任何附加用户共同存在的詹姆斯相关联(例如,当詹姆斯与玛丽(Mary)和/或附加的其他用户共同存在时,内容代理可以用于渲染内容)。在其他实现方式中,所生成的内容代理可以与与迈克尔共同存在的詹姆斯相关联(例如,当詹姆斯与迈克尔共同存在时,内容代理可以用于渲染内容,而当詹姆斯与玛丽和/或并非迈克尔的附加其他用户共同存在时,内容代理不用于渲染内容)。

[0031] 对话管理器112可以用于发起与用户的对话,诸如经由用户接口输入设备102和用户接口输出设备104与用户的对话。例如,对话可以包括经由扬声器渲染的听觉输出以及使用麦克风捕获的来自用户的口头响应。在许多实现方式中,在对内容的初始请求之后发起对话,以在生成新的内容代理时细化对内容的请求。例如,用户可以提供对“向我展示早午餐食谱”内容的初始请求。可以响应于初始请求来确定初始内容语料库116(诸如由内容模块114确定)(例如,初始内容语料库116是响应于请求“向我展示早午餐食谱”)。对话管理器112可以发起与用户的对话来细化该内容的请求,这又可以细化内容语料库116。例如,在请求“向我展示早午餐食谱”之后的对话可以包括:客户端设备“您想要热的还是冷的早午餐物”;用户“都要”;客户端设备“有没有食物过敏”;用户“有”;客户端设备“你对什么食物过敏”;用户“鸡蛋”。在许多实现方式中,可以向代理模块110提供该对话以确定附加的用户参数(在许多实现方式中,内容模块114可以利用这些参数来细化内容语料库116)。

[0032] 对话管理器112可以附加地和/或可替代地发起对话,以根据用户的请求细化内容语料库。用户随后可以使用对话模块112来发起附加对话以进一步细化该内容语料库116。例如,用户可以响应于内容代理被用于渲染不期望的内容而发起附加的对话会话。例如,可以使用“向我展示早午餐食谱”的初始请求和对应的初始对话会话来生成早午餐内容代理。早午餐内容代理随后可以用于为用户渲染香蕉煎饼食谱。不喜欢香蕉的用户可以使用对话模块112发起附加的对话会话,以从内容语料库116中排除包含香蕉的食谱。在多种实现方式中,内容参数可能具有固有的不确定性。对话模块112可以确定是否发起对话会话来消除不确定性。例如,用户可以发起与设置路由器相关的内容代理,其中用户没有指定型号。与特定路由器相关的一个或多个内容参数可能包括由于系统不知道路由器型号而导致的固有不确定性。通常,内容代理可以最初提供与设置路由器相关的响应内容。如果用户提供了对需要了解路由器型号的内容的请求(诸如设置型号特定特征的请求),则对话模块112可以发起对话会话来提示用户路由器型号。

[0033] 另外地或替代地,对话模块112可以响应于向内容语料库116添加新的内容、响应于指示用户对内容语料库116的渲染内容不满意的信号、和/或响应于附加或替代事件,抢先发起对话会话。例如,新的内容可以被识别为响应于初始内容参数以及附加内容参数(例如,使用如下所述的内容模块114来识别),以包括在与内容代理相关联的内容语料库116中。对话管理器112可以基于添加到内容语料库116的新的内容而抢先发起附加对话会话。

[0034] 可以利用内容模块114以响应于初始内容参数(诸如使用代理模块110识别的内容

参数)来识别初始内容语料库,基于一个或多个内容参数(诸如由代理模块110从对话会话中确定的附加内容参数)来细化内容语料库,向现有内容语料库添加新的内容,选择内容语料库中的内容项以渲染给用户,和/或执行与内容语料库相关的附加动作。

[0035] 在许多实现方式中,内容模块114可以响应于来自用户的对内容的请求来识别初始内容语料库。例如,内容模块114可以响应于对应于经由用户对内容的初始请求的初始内容参数(例如,由上述代理模块110确定的初始内容参数)来识别内容1、内容2、内容3、内容4、内容5、内容6、内容7和内容8。另外地或替代地,内容模块114可以基于内容参数(诸如代理模块110从对话模块112模拟的对话中确定的内容参数)来细化初始内容语料库。例如,内容代理114可以使用从与用户的对话中确定的内容参数,通过移除例如内容2、内容4和内容5来细化内容语料库(即,基于附加的内容参数,内容语料库从8项内容减少到5项内容(内容1、内容3、内容5、内容7和内容8))。内容语料库的减少可以包括附加的或替代的减少,例如基于进一步的内容参数将内容语料库减少10%、25%、50%和/或内容语料库中项数量的附加减少。

[0036] 可以利用内容模块114来向现有内容语料库添加新的内容。例如,如果内容语料库中的项数量下降到阈值以下,如果识别出先前未包括在内容语料库中的新的内容,如果在确定初始内容语料库被识别之后在创建新的内容的地方识别出新的内容(例如,响应于内容参数的今天发布的新闻文章可以被添加到上周被确定为响应于内容参数的内容语料库),和/或如果识别出附加的新的内容,则可以添加新的内容。

[0037] 另外地或替代地,内容模块114可以选择内容项来为用户渲染。例如,用户可以请求附加的内容项,并且内容模块114可以选择该内容项来渲染。另外地或替代地,内容模块114可以抢先向用户渲染内容项。例如,当用户被确定为与附加用户共同存在时,内容模块114可以选择内容项来为用户渲染。在许多实现方式中,内容可以在用户可能想要重访的内容(即,可重访的内容)和/或用户不太可能想要重访的内容(即,不可重访的内容)之间进行区分。例如,用户可能想要重访他们最喜欢的乐队的歌曲(即,音乐是可重访的内容)。当选择用户之前与之交互的内容116的项时,内容模块114可以改变先前与内容交互的顺序。例如,内容模块114可以在第一次为用户渲染内容时按顺序A、B、C、D、E选择音乐内容,并且内容模块114可以在第二次为用户渲染内容时按顺序B、D、E、C、A选择音乐内容。另外地或替代地,用户不太可能想要重访他们已经阅读过的新闻文章(即,新闻文章是不可重访的内容)。内容模块114可以从内容语料库排除该先前渲染的内容。在许多实现方式中,当选择内容以向用户渲染时,内容模块114可以对新的内容进行优先排序。例如,当选择内容项时,内容模块114可以通过用户喜爱的乐队对新发布的歌曲进行优先排序。

[0038] 图1B示出根据多个实现方式的示例状态图150。在图1B中,使用一个或多个用户接口输入设备152(例如,图1A的用户接口输入设备102)捕获对内容158的初始请求,并将其传输到客户端设备(例如,图1A的客户端设备106)的内容代理引擎156(例如,图1A的内容代理引擎108)。例如,用户可以提供由麦克风捕获的“向我展示如何使用我的假想家用机器人套件”的口头请求。另外地或替代地,用户可以使用客户端设备的触摸屏键入“向我展示如何使用我的假想家用机器人套件”的初始请求。

[0039] 在160,内容代理引擎156基于所接收的初始请求来确定初始内容参数。例如,“假想的家用机器人套件”、“如何使用”的初始内容参数和/或附加内容参数可以从“向我展示

如何使用我的假想家用机器人套件”的初始请求中确定。

[0040] 在162,内容代理引擎156可以确定初始内容语料库。在许多实现方式中,初始内容语料库包括响应于初始内容参数的内容。例如,初始内容语料库可以包括图像、视频、文章和/或关于假想的家用机器人套件的其他内容项。在164,内容代理引擎156可以通过经由用户接口输出设备154传输要渲染的对话来发起对话会话。在166,响应于经由用户接口输出设备154渲染的对话会话,用户可以向用户接口输入设备154提供输入。在许多实现方式中,可以向内容代理引擎156传输所接收的对话输入。在多种实现方式中,对话会话可以包括对话中的多个回合。例如,关于“向我展示如何使用我的假想家用机器人套件”内容的初始请求的对话会话可以包括:客户端设备“您有什么版本?”;用户“版本号1.0”;客户端设备“您想要包括视频吗?”;用户“是”。

[0041] 在许多实现方式中,内容代理引擎156可以任选地确定提供对内容158的初始请求的用户在进行初始请求时和/或在参与对话会话时与附加用户共同存在,以进一步细化对内容的请求。

[0042] 在168,内容代理156可以从对话会话来确定附加内容参数。例如,可以从在164发起的对话会话中确定“版本号1.0”、“包括视频”的附加内容参数和/或附加内容参数。

[0043] 在170,内容代理引擎156可以使用在160确定的初始内容参数和在168确定的附加内容参数来生成内容代理。在许多实现方式中,所生成的内容代理可以任选地与与附加用户共同存在的用户相关。

[0044] 在172,内容代理引擎156可以使用在168确定的附加内容参数来细化内容语料库。例如,“版本号1.0”的附加内容参数可以用于细化内容语料库,以排除与版本号2.0相关的内容。

[0045] 在174,内容代理引擎156可以通过从细化的内容语料库中选择内容项并将该内容传输到一个或多个用户接口输出设备154用于渲染,来利用所生成的内容代理来渲染内容。

[0046] 在176,响应于初始内容参数和附加内容参数,内容代理156可以监视要包括在内容语料库中的新的内容。在许多实现方式中,新的内容可以与不包括在初始内容参数和/或附加内容参数中的内容参数相关联。新的内容和/或对应的新的内容参数的这种识别可以使得内容代理156在164发起与用户的附加对话会话,并且继续进行到166、168、170、172和/或174。在许多实现方式中,从附加对话中确定新的内容参数,并且这些新的内容参数可以用于进一步细化内容语料库。例如,内容代理引擎156可以确定与假想家用机器人套件版本号3.0相关的新的内容。在164,可以发起关于新版本(版本号3.0)的对话会话,并且所生成的对话会话可以用于识别附加的内容参数,以进一步细化内容语料库。

[0047] 图2是示出根据本文公开的实现方式生成内容代理的示例过程200的流程图。为了方便起见,参考执行操作的系统来描述流程图的操作。该系统可以包括各种计算机系统的各种组件,诸如图1A的客户端设备106的一个或多个组件。此外,虽然过程200的操作以特定顺序展示,但这并不意图是限制性的。一个或多个操作可以被重新排序、省略和/或添加。

[0048] 在块202,系统接收对内容的请求,其中该请求是基于用户的输入而生成的。例如,用户可以提供多种类型的请求,包括对“向我展示晚餐餐厅”内容的口头请求。对内容的请求可以包括口头请求、键入的请求、通过按下按钮(物理按钮和/或选择用户接口上的虚拟按钮)发起的请求、手势请求和/或附加类型的用户请求。用户可以请求多种类型的内容,包

括：视频、音频、图像、网页、新闻文章、视频、食谱、博客帖子、电子书、模因、社交媒体帖子和/或用户可以经由客户端设备请求渲染的附加类型的内容。在多种实现方式中，对内容的请求可以识别特定类型的内容（例如，“给我播放新闻视频”指定渲染视频）。另外地或替代地，对内容的请求不必指定内容的类型，并且可以响应于该请求来确定各种类型的内容（例如，各种类型的内容可以响应于“向我展示新闻”的请求，包括视频、文章、网页和/或附加类型的内容）。

[0049] 此外，在与用户的对话会话期间，内容的类型可以被确定为内容参数。例如，对话会话可以包括：客户端设备“您想看视频吗”；用户“否”；客户端设备“您想看新闻文章吗”；用户“是”，这可以排除一种类型的内容（视频），而包括另一类型的内容（新闻文章）。在一些实现方式中，允许的内容类型仅是由内容参数指定的内容类型。在其他实现方式中，除了内容参数中明确排除的类型之外，所有类型的内容都是允许的。

[0050] 在块204，系统基于对内容的请求来确定初始内容参数。例如，“假想运动队”和“新闻”的初始内容参数可以从对“向我展示关于假想运动队的新闻”的内容的初始请求中确定。

[0051] 在块206，系统确定内容项的语料库是响应于初始内容参数。例如，可以确定包括与假想运动队相关的新闻文章、与假想运动队相关的视频和/或与假想运动队相关的附加内容的内容语料库。

[0052] 在块208，系统发起与用户的对话会话。例如，对话会话可以包括经由客户端设备的扬声器渲染的音频内容以及由客户端设备的麦克风捕获的口头输入。附加和/或可替代的对话会话可以包括多种用户接口输入和输出设备的使用。例如，如上文关于块202所述，对话会话可以包括：客户端设备“您想看视频吗”；用户“否”；客户端设备“您想看新闻文章吗”；用户“是”。

[0053] 在块210，系统基于对话会话期间的用户输入确定进一步的内容参数。例如，包括新闻文章在内的无视频的进一步内容参数可以从以下对话会话中确定：客户端设备“您想看视频吗”；用户“否”；客户端设备“您想看新闻文章吗”；用户“是”。

[0054] 在块212，系统任选地确定该用户是否与附加用户共同存在。例如，系统可以基于使用用户接口输入设备捕获到的数据（诸如：使用麦克风捕获的附加用户的口语自然语言、使用相机捕获的附加用户的图像、和/或指示附加用户的共同存在的附加捕获的传感器数据）来确定附加用户的共同存在。

[0055] 在块214，系统基于初始参数（即，从对内容的初始请求中确定的内容参数）以及进一步内容参数（即，从与用户的对话会话中确定的内容参数）而生成内容代理。在许多实现方式中，如果在块212系统确定用户与附加用户共同存在，则系统可以使内容代理与与附加用户共同存在的该用户相关。在一些此类实现方式中，内容代理可以与与任何附加用户共同存在的用户相关。在其他实现方式中，内容代理可以与在块212中被确定为与用户共同存在的特定附加用户共同存在的用户相关。

[0056] 图3是示出根据本文公开的实现方式使用内容代理来确定附加内容的示例过程300的流程图。为了方便起见，参考执行操作的系统来描述流程图的操作。该系统可以包括各种计算机系统的各种组件，诸如图1A的客户端设备106的一个或多个组件。此外，虽然过程300的操作以特定顺序展示，但这并不意图是限制性的。一个或多个操作可以被重新排

序、省略和/或添加。

[0057] 在块302,该系统任选地基于来自用户的输入生成内容代理。例如,可以根据图2的过程200来生成内容代理。在许多实现方式中,所生成的内容代理可以与与附加用户共同存在的用户相关。例如,可以根据过程200的块212来确定用户与附加用户的共同存在。

[0058] 在块304,系统确定用户与附加用户共同存在的附加实例。例如,系统可以基于使用用户接口输入设备捕获到的数据(诸如:使用麦克风捕获的附加用户的口语自然语言、使用相机捕获的附加用户的图像、和/或指示附加用户的共同存在的附加捕获的传感器数据)来确定附加用户的共同存在的附加实例。

[0059] 在块306,系统使用内容代理以经由一个或多个用户接口输出设备为用户渲染内容。如块306A和306B所示,可以利用根据本文描述的实现方式的内容代理以各种方式渲染内容。

[0060] 在块306A,系统使用内容代理来渲染内容,其中该内容是响应于对内容的附加用户请求。例如,可以使用“向我展示晚餐食谱”的初始请求来生成食谱内容代理。食谱内容代理可以用于为“向我展示午餐食谱”用户的后续请求渲染内容。

[0061] 在块306B,系统使用内容代理来渲染内容,其中该内容是被主动渲染的。例如,系统可以确定用户与附加用户共同存在的附加实例,并且可以使用食谱内容代理来主动渲染早午餐食谱。

[0062] 转到图4,可以实现本文公开的实现方式的示例环境。图4包括客户端计算设备402,其执行自动助理客户端404的实例。一个或多个基于云的自动助理组件410可以在一个或多个计算系统(统称为“云”计算系统)上实现,该计算系统经由大体以408指示的一个或多个局域网和/或广域网(例如,因特网)通信地耦合到客户端设备402。

[0063] 自动助理客户端404的实例借助于其与一个或多个基于云的自动助理组件410的交互,可以形成从用户的角度来看似乎是用户可以用来参加人机对话的自动助理400的逻辑实例。因此应理解,在一些实现方式中,与在客户端设备402上执行的自动助理客户端404进行交互的用户实际上可以与他或她自己的自动助理400的逻辑实例进行交互。为了简洁和简单性起见,本文用作“服务”特定用户的术语“自动助理”通常指在由用户操作的客户端设备402上执行的自动助理客户端404和一个或多个基于云的自动助理组件410(其可以在多个客户端计算设备的多个自动助理客户端之中共享)的组合。还应理解,在一些实现方式中,自动助理400可以响应于来自任何用户的请求,而不管该用户实际上是否由自动助理400的特定实例“加以服务”。

[0064] 客户端计算设备402可以是例如:台式计算设备、膝上型计算设备、平板型计算设备、移动智能电话计算设备、独立交互式扬声器、智能家电、和/或包括计算设备的用户可穿戴装置(例如,具有计算设备的用户的手表、具有计算设备的用户的眼镜、虚拟或增强现实计算设备)。可以提供附加和/或替代性的客户端计算设备。另外地或替代地,客户端计算设备402的操作可以分布在多个计算设备之间。例如,客户端计算设备402的一个或多个操作可以分布在移动智能电话和车辆计算设备之间。此外,客户端计算设备402的操作可以在多个计算设备之间重复(在一些情况下,这些计算设备可以通信地耦合)。作为另一示例,移动智能电话以及车辆接口设备可以各自实现自动助理400的操作,诸如移动智能电话和车辆接口设备两者都包括调用引擎(如下文描述)。在各种实现方式中,客户端计算设备402可以

任选地操作除自动助理客户端404之外的一个或多个其他应用,诸如消息交换客户端(例如,SMS、MMS、在线聊天)、浏览器等。在那些各种实现方式中的一些中,一个或多个其他应用可以任选地与自动助理404对接(例如,经由应用编程接口),或者包括它们自己的自动助理应用实例(其也可以与基于云的自动助理组件410对接)。

[0065] 自动助理400经由客户端设备(未示出)的用户接口输入和输出设备来参与与用户的人机对话会话。为了保护用户隐私和/或节省资源,在许多情况下,用户必须在自动助理完全处理口头话语之前,经常明确地调用自动助理400。自动助理400的明确调用可以响应于在客户端设备402处接收的某些用户接口输入而发生。例如,可以经由客户端设备402调用自动助理400的用户接口输入可以任选地包括客户端设备402的硬件和/或虚拟按钮的致动。此外,自动助理客户端可以包括一个或多个本地引擎406,诸如可操作来检测一个或多个口头调用短语的存在的调用引擎。调用引擎可以响应于检测到一个或多个口头调用短语来调用自动助理400。例如,调用引擎可以响应于检测到诸如“嘿助理”、“OK助理”和/或“助理”之类的口头调用短语来调用自动助理400。调用引擎可以连续处理(例如,如果不处于“非活动”模式)基于来自客户端设备402的一个或多个麦克风的输出的音频数据帧的流,以监视口头调用短语的出现。当监视口头调用短语的出现时,调用引擎丢弃(例如,在缓冲器中临时存储之后)不包括口头调用短语的任何音频数据帧。然而,当调用引擎检测到在已处理的音频数据帧中出现口头调用短语时,调用引擎可以调用自动助理400。如本文所使用,“调用”自动助理400可以包括使自动助理400的一个或多个先前非活动的功能被激活。例如,调用自动助理400可以包括使一个或多个本地引擎406和/或基于云的自动助理组件410进一步处理基于其检测到调用短语的音频数据帧,和/或一个或多个后续音频数据帧(而在调用之前没有发生音频数据帧的进一步处理)。

[0066] 自动助理404的一个或多个本地引擎406是可选的,并且可以包括例如上述调用引擎、本地语音到文本(“STT”)引擎(将捕获的音频转换成文本)、本地文本到语音(“TTS”)引擎(将文本转换成语音)、本地自然语言处理器(确定音频和/或从音频转换的文本的语义)和/或其他本地组件。因为客户端设备402在计算资源(例如,处理器周期、存储器、电池等)方面相对受限,本地引擎406相对于包括在基于云的自动助理组件410中的任何对应物可能具有有限的功能性。

[0067] 自动助理客户端404可以另外包括内容代理引擎(未示出),其可以由自动助理客户端404根据各种实现方式来使用,包括:生成内容代理、使用内容代理确定与用户请求相关的内容、使用内容代理确定内容而不接收用户请求等。

[0068] 相对于本地引擎406的任何对应物,基于云的自动助理组件410利用云的几乎无限的资源来执行音频数据和/或其他用户接口输入的更稳健和/或更准确的处理。此外,在各种实现方式中,响应于调用引擎检测到口头调用短语,或者检测到自动助理400的一些其他明确调用,客户端设备402可以向基于云的自动助理组件410提供音频数据和/或其他数据。

[0069] 所示出的基于云的自动助理组件410包括基于云的TTS模块412、基于云的STT模块414和自然语言处理器416。在一些实现方式中,自动助理400的一个或多个引擎和/或模块可以被省略、组合和/或实现在与自动助理400分离的组件中。此外,在一些实现方式中,自动助理400可以包括附加和/或替代的引擎和/或模块。

[0070] 基于云的STT模块414可以将音频数据转换成文本,然后可以将其提供给自然语言

处理器416。在各种实现方式中,基于云的STT模块414可以至少部分地基于由指派引擎(未示出)提供的扬声器标签和指派的指示,将音频数据转换成文本。

[0071] 基于云的TTS模块412可以将文本数据(例如,由自动助理400制定的自然语言响应)转换为计算机生成的话音输出。在一些实现方式中,TTS模块412可以将计算机生成的话音输出提供给客户端设备402,以便例如使用一个或多个扬声器直接输出。在其他实现方式中,由自动助理400生成的文本数据(例如,自然语言响应)可以被提供给本地引擎406中的一者,该本地引擎然后将文本数据转换成在本地输出的计算机生成的话音。

[0072] 自动助理400的自然语言处理器416处理自由形式的自然语言输入,并基于自然语言输入生成注释输出,供自动助理400的一个或多个其他组件使用。例如,自然语言处理器416可以处理作为文本输入的自然语言自由形式输入,该文本输入是由STT模块414对用户经由客户端设备402提供的音频数据的转换。所生成的注释输出可以包括自然语言输入的一个或多个注释以及任选地自然语言输入的一个或多个(例如,所有)术语。在一些实现方式中,自然语言处理器416被配置成识别和注释自然语言输入中的各种类型的语法信息。例如,自然语言处理器416可以包括被配置成利用术语的语法角色来注释术语的话音标记器(未描绘)的一部分。而且,例如,在一些实现方式中,自然语言处理器416可以另外地和/或替代地包括被配置成确定自然语言输入中的术语之间的句法关系的依存语法分析器(未描绘)。

[0073] 在一些实现方式中,自然语言处理器416可以另外地和/或替代地包括实体标记器(未描绘),该实体标记器被配置成注释一个或多个片段中的实体引用,诸如对人(包括例如,文学人物、名人、公众人物等)、组织、地点(真实的和想象的)等等的引用。自然语言处理器416的实体标记器可以以高粒度级别(例如,能够识别对诸如人的实体类的所有引用)和/或以低粒度级别(例如,能够识别对诸如特定人的特定实体的所有引用)注释对实体的引用。实体标记器可以依赖自然语言输入的内容来解析特定实体,和/或可以任选地与知识图或其他实体数据库通信来解析特定实体。

[0074] 在一些实现方式中,自然语言处理器416可以另外地和/或替代地包括共指解析器(未描绘),该共指解析器被配置成基于一个或多个上下文线索来对同一实体的引用进行分组或“聚类”。例如,共指解析器可以用于将自然语言输入“我喜欢上次我们在那里吃饭的假想咖啡馆”中的术语“那里”解析为“假想咖啡馆”。

[0075] 在一些实现方式中,自然语言处理器416的一个或多个组件可以依赖于来自自然语言处理器416的一个或多个其他组件的注释。例如,在一些实现方式中,命名实体标记器可以依赖于来自共指解析器和/或依存语法分析器的注释来注释特定实体的所有提及。而且,例如,在一些实现方式中,共指解析器可以依赖于来自依存语法分析器的注释来对同一实体的引用进行聚类。在一些实现方式中,在处理特定自然语言输入时,自然语言处理器416的一个或多个组件可以使用相关的先前输入和/或特定自然语言输入之外的其他相关数据来确定一个或多个注释。

[0076] 图5是可以任选地用以执行本文所述技术的一个或多个方面的示例计算设备510的框图。在一些实现方式中,客户端计算设备和/或其他组件中的一者或多者可以包括示例计算设备510的一个或多个组件。

[0077] 计算设备510通常包括经由总线子系统512与多个外围设备通信的至少一个处理

器514。这些外围设备可以包括存储子系统524(包括例如,存储器子系统525和文件存储子系统526)、用户接口输出设备520、用户接口输入设备522和网络接口子系统516。输入和输出设备允许用户与计算设备510交互。网络接口子系统516提供到外部网络的接口,并且被耦合到其他计算设备中的对应接口设备。

[0078] 用户接口输入设备522可以包括键盘、指点设备(诸如鼠标、轨迹球、触摸板或图形输入板)、扫描仪、并入到显示器中的触摸屏、音频输入设备(诸如语音辨识系统、麦克风)和/或其他类型的输入设备。通常,使用术语“输入设备”旨在包括将信息输入到计算设备510或通信网络上的所有可能类型的设备和方式。

[0079] 用户接口输出设备520可以包括显示器子系统、打印机、传真机或非视觉显示器(诸如音频输出设备)。显示器子系统可以包括阴极射线管(“CRT”)、诸如液晶显示器(“LCD”)的平板设备、投影设备或用于创建可视图像的一些其他机构。显示器子系统还可以诸如经由音频输出设备提供非视觉显示器。通常,使用术语“输出设备”旨在包括从计算设备510输出信息到用户或到另一机器或计算设备的所有可能类型的设备和方式。

[0080] 存储子系统524存储提供本文描述的一些或所有模块的功能性的编程和数据结构。例如,存储子系统524可以包括用以执行图2和/或图3的一个或多个过程的所选择方面以及用以实现图1A中所描绘的各种组件的逻辑。

[0081] 这些软件模块通常由处理器514单独或者结合其他处理器来执行。存储子系统524中使用的存储器525可以包括多个存储器,包括用于在程序执行期间存储指令和数据的主随机存取存储器(“RAM”)530和其中存储固定指令的只读存储器(“ROM”)532。文件存储子系统526可以为程序和数据文件提供持久性存储,并且可以包括硬盘驱动器、软盘驱动器以及相关可移动介质、CD-ROM驱动器、光学驱动器或可移除介质盒。实现某些实现方式的功能性的模块可以由文件存储子系统526存储在存储子系统524中,或者处理器514可访问的其他机器中。

[0082] 总线子系统512提供用于让计算设备510的各种组件和子系统按照预期相互通信的机制。尽管总线子系统512示意性展示为单个总线,但总线子系统的替代实现方式可以使用多个总线。

[0083] 计算设备510可以是不同类型的,包括工作站、服务器、计算集群、刀片服务器、服务器群或任何其他数据处理系统或计算设备。由于计算机和网络的不断变化的性质,图5中描绘的计算设备510的描述仅旨在作为用于说明一些实现方式的特定的特定示例。计算设备510的许多其他配置可能具有比图5中所描绘的计算设备更多或更少的组件。

[0084] 另外,一些实现方式包括一个或多个计算设备的一个或多个处理器(例如,中央处理单元(CPU)、图形处理单元(GPU)和/或张量处理单元(TPU)),其中一个或多个处理器可操作来执行存储在相关联的存储器中的指令,并且其中指令被配置成导致执行本文描述的任一方法。一些实现方式还包括存储计算机指令的一个或多个非暂时性计算机可读存储介质,该计算机指令可由一个或多个处理器执行以执行本文所描述的任一方法。

[0085] 在本文所描述的系统收集关于用户(或者本文经常称作的“参与者”)的个人信息或可以利用个人信息的情况下,可以向用户提供机会来控制程序或特征是否收集用户信息(例如,关于用户的社交网络、社交动作或活动、职业、用户的偏好或用户的当前地理位置的信息),或者控制是否和/或如何从内容服务器接收可能与用户更相关的内容。而且,某些数

据可以在被存储或使用之前以一种或多种方式加以处理,以便消除个人身份信息。例如,可以处理用户的身份,使得不能为用户确定个人身份信息,或者可以在获得地理位置信息的地方一般化用户的地理位置(诸如城市、邮政编码或州级),使得不能确定用户的特定地理位置。因此,用户可以控制如何收集和/或使用关于用户的信息。

[0086] 在一些实现方式中,提供了一种由一个或多个处理器实现的方法,包括接收对内容的请求。对内容的请求基于用户经由客户端设备的一个或多个用户接口输入设备的输入而生成。该方法还包括基于用户请求来确定一个或多个初始内容参数。该方法还包括确定内容项的语料库是响应于初始内容参数。响应于确定内容项的语料库是响应于初始内容参数,该方法还包括经由客户端设备发起与用户的对话会话,基于对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数,该进一步内容参数细化该请求并限制语料库中响应于该请求的内容项的数量。该方法还包括确定在请求时,用户与至少一个附加用户共同存在。该方法还包括基于与请求相关联的初始内容参数以及基于对话会话期间的用户输入确定的进一步内容参数来生成内容代理。该方法还包括使内容代理与附加用户共同存在的用户相关。在生成内容代理之后,该方法还包括确定用户与附加用户共同存在的附加实例。响应于内容代理与附加用户共同存在的用户相关,该方法还包括使用内容代理来主动渲染符合初始内容参数和进一步内容参数的给定内容。

[0087] 本文所公开的技术的这些和其他实现方式可以包括以下特征中的一者或多者。

[0088] 在一些实现方式中,经由客户端设备发起与用户的对话会话以及基于对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数还包括,确定内容项的语料库超过内容项的阈值数量。

[0089] 在一些实现方式中,经由客户端设备发起与用户的对话会话以及基于对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数还响应于:确定内容项的语料库中的内容项的数量超过阈值数量。

[0090] 在一些实现方式中,经由客户端设备发起与用户的对话会话包括:基于内容项的语料库中的内容项中的一者或多者的一个或多个特征,生成针对内容项中的一者或多者定制的提示。在那些实现方式的一些版本中,该方法还包括响应于用户请求,并且在对话会话期间的一个或多个用户输入之前渲染该提示。

[0091] 在一些实现方式中,在生成内容代理之后,该方法还包括确定响应于初始内容参数的一个或多个新的内容项以包括在内容项的语料库中,其中新的内容项是在确定内容项的语料库之后创建的。在那些实现方式的一些版本中,该方法还包括基于新的内容项中的一者或多者的一个或多个新特征来生成针对新的内容项中的一者或多者定制的新提示。在那些实现方式的一些版本中,该方法还包括通过提供该新提示经由客户端设备发起与用户的后续对话会话。在那些实现方式的一些版本中,该方法还包括基于后续对话会话期间的一个或多个后续用户输入来确定一个或多个后续内容参数,该后续内容参数进一步细化该请求并限制语料库中响应于该请求的新内容项的数量。在那些实现方式的一些版本中,该方法还包括基于后续内容参数来修改内容代理。

[0092] 在一些实现方式中,在生成内容代理之后并在主动渲染符合初始内容参数和进一步内容的给定内容之前,该方法还包括:使用内容代理来确定响应于用户请求的内容。在那些实现方式的一些版本中,该方法还包括使得计算系统响应于用户请求来渲染该内容。

[0093] 在一些实现方式中,在使用内容代理来主动渲染符合初始内容参数和进一步内容参数的给定内容之后,该方法还包括:通过经由客户端设备发起与用户的附加对话会话来确定一个或多个附加内容参数,该附加内容参数进一步细化该请求且进一步限制该语料库中响应于该请求的内容项。

[0094] 在一些实现方式中,内容项的语料库包括至少两个媒体类型的内容项。

[0095] 在一些实现方式中,内容项的语料库包括用户未查看的内容项,并且其中使用内容代理来主动渲染给定内容包括在从内容项的语料库中选择给定内容时对用户未查看的内容项进行优先排序。

[0096] 在一些实现方式中,在使用内容代理来主动渲染符合初始内容参数和进一步内容参数的给定内容之后,该方法还包括:接收指示用户是否与给定内容交互的一个或多个用户反馈信号。在那些实现方式的一些版本中,该方法还包括基于一个或多个用户反馈信号来更新内容代理。

[0097] 在一些实现方式中,一个或多个用户反馈信号包括指示用户与内容交互的正反馈信号。在那些实现方式的一些版本中,一个或多个用户反馈信号包括指示用户不与内容交互的负反馈信号。

[0098] 在一些实现方式中,客户端设备包括自动助理客户端,其中在自动助理客户端的助理接口处接收请求,并且其中对话会话是与用户和客户端设备的自动助理客户端的对话会话。在那些实现方式的一些版本中,用户与自动助理客户端的用户简档相关联,并且该方法还包括:使用附加客户端设备来确定用户与附加用户共同存在的另一实例,其中附加客户端设备不同于客户端设备,并且其中用户与附加用户共同存在的另一实例是在用户与附加用户共同存在的附加实例之后。在那些实现方式的一些版本中,该方法还包括使用内容代理来使用附加客户端设备主动渲染附加给定内容。在那些实现方式的一些版本中,附加用户与客户端设备的已知用户简档相关联,并且其中确定用户与附加用户共同存在的附加实例包括:确定用户与客户端设备的已知用户简档相关联的附加用户共同存在的附加实例。在那些实现方式的一些版本中,附加用户不与客户端设备的已知用户简档相关联。

[0099] 在一些实现方式中,该方法还包括确定在请求时,请求的附加场境。在一些实现方式中,使内容代理与与附加用户共同存在的用户相关还包括,使内容代理与在该请求的附加场境中与附加用户共同存在的用户相关。附加场境可以是例如用户参加的活动。该活动可以基于来自用户的明确输入(例如,“我正在做周末家务”)来确定,和/或可以基于各种传感器信号(例如,GPS、加速度计等)来确定。在一些实现方式中,附加场境自动地被确定,并且至少部分基于来自客户端设备的一个或多个传感器的传感器数据来确定。

[0100] 在一些实现方式中,提供一种由一个或多个处理器实现的方法,该方法包括接收对内容的请求,该对内容的请求基于用户经由客户端设备的一个或多个用户接口输入设备的输入而生成。该方法还包括基于用户请求来确定一个或多个初始内容参数。该方法还包括确定内容项的语料库是响应于初始内容参数。响应于确定内容项的语料库是响应于初始内容参数,该方法还包括经由客户端设备发起与用户的对话会话,并且基于对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数,该进一步内容参数细化该请求并限制语料库中响应于该请求的内容项的数量。该方法还包括确定在请求时,用户与至少一个附加用户共同存在。该方法还包括基于与请求相关联的初始内容参数以及基于对话会

话期间的用户输入确定的进一步内容参数来生成内容代理。该方法还包括使内容代理与附加用户共同存在的用户相关。在生成内容代理之后,该方法还包括接收对内容的附加请求,该对内容的附加请求基于用户经由客户端设备的一个或多个用户接口输入设备的附加输入而生成。该方法还包括确定用户与附加用户共同存在的附加实例。响应于内容代理与附加用户共同存在的用户相关,该方法还包括使用内容代理来确定响应于附加用户请求的附加内容。该方法还包括使得客户端设备渲染附加内容。

[0101] 在一些实现方式中,提供一种由一个或多个处理器实现的方法,该方法包括接收对内容的请求,该对内容的请求基于用户经由客户端设备的一个或多个用户接口输入设备的输入而生成。该方法还包括基于请求来确定一个或多个初始内容参数。该方法还包括确定内容项的语料库是响应于初始内容参数。响应于确定内容项的语料库是响应于初始内容参数,该方法还包括经由客户端设备发起与用户的对话会话。该方法还包括基于对话会话期间的一个或多个用户输入来确定一个或多个进一步内容参数,该进一步内容参数进一步细化该请求并限制语料库中响应于该请求的内容项的数量。该方法还包括基于对内容的请求和/或基于来自客户端设备的传感器数据自动地确定在请求时的一个或多个场境条件。该方法还包括基于与请求相关联的初始内容参数以及基于对话会话期间的用户输入确定的进一步内容参数来生成内容代理。该方法还包括使内容代理与一个或多个场境条件相关。在生成内容代理之后,该方法还包括确定一个或多个场境条件的出现的附加实例。响应于内容代理与一个或多个场境条件相关,该方法还包括使用内容代理来主动渲染符合初始内容参数和进一步内容参数的给定内容。

[0102] 本文所公开的技术的这些和其他实现方式可以包括以下特征中的一者或多者。

[0103] 在一些实现方式中,一个或多个场境条件包括用户在请求时正在参加特定活动。在那些实现方式的一些版本中,一个或多个场境条件包括用户与至少一个附加用户在请求时共同存在。

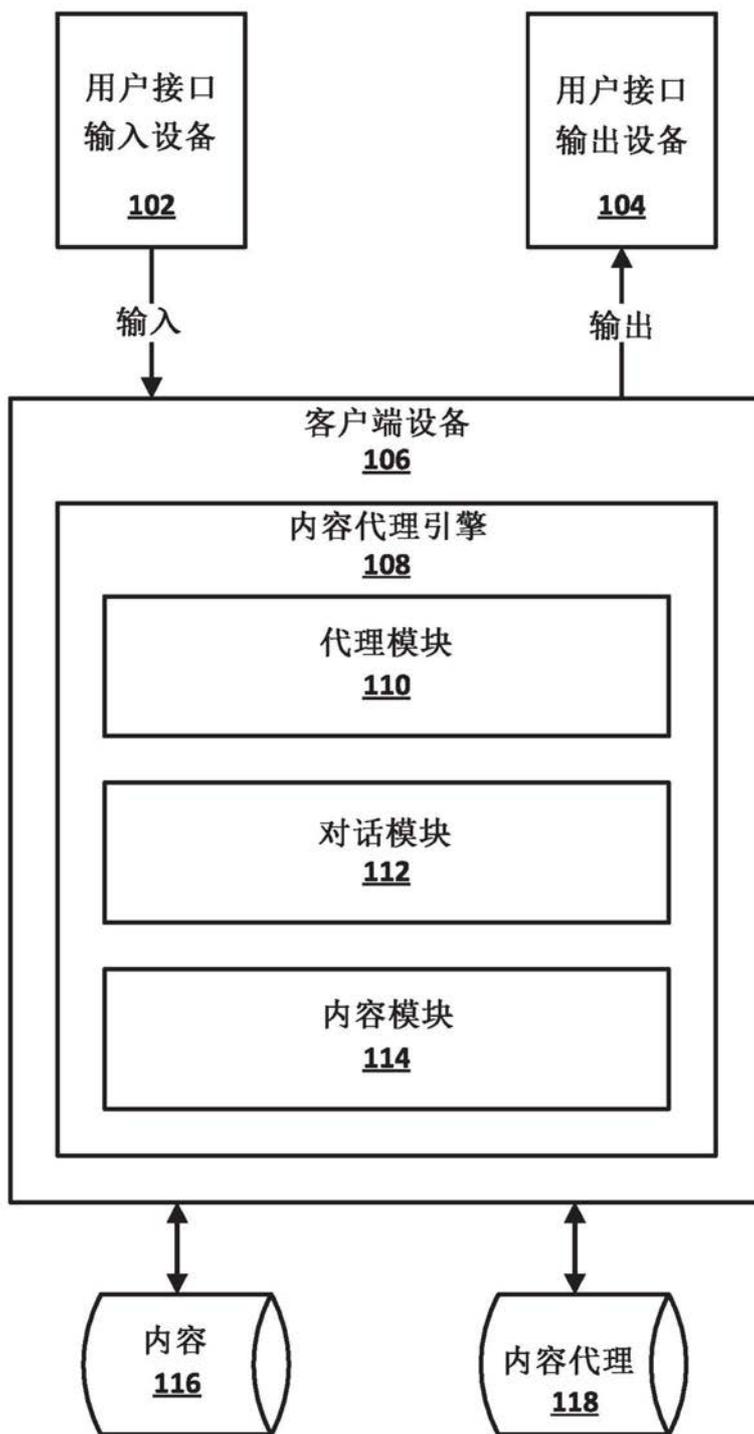


图1A

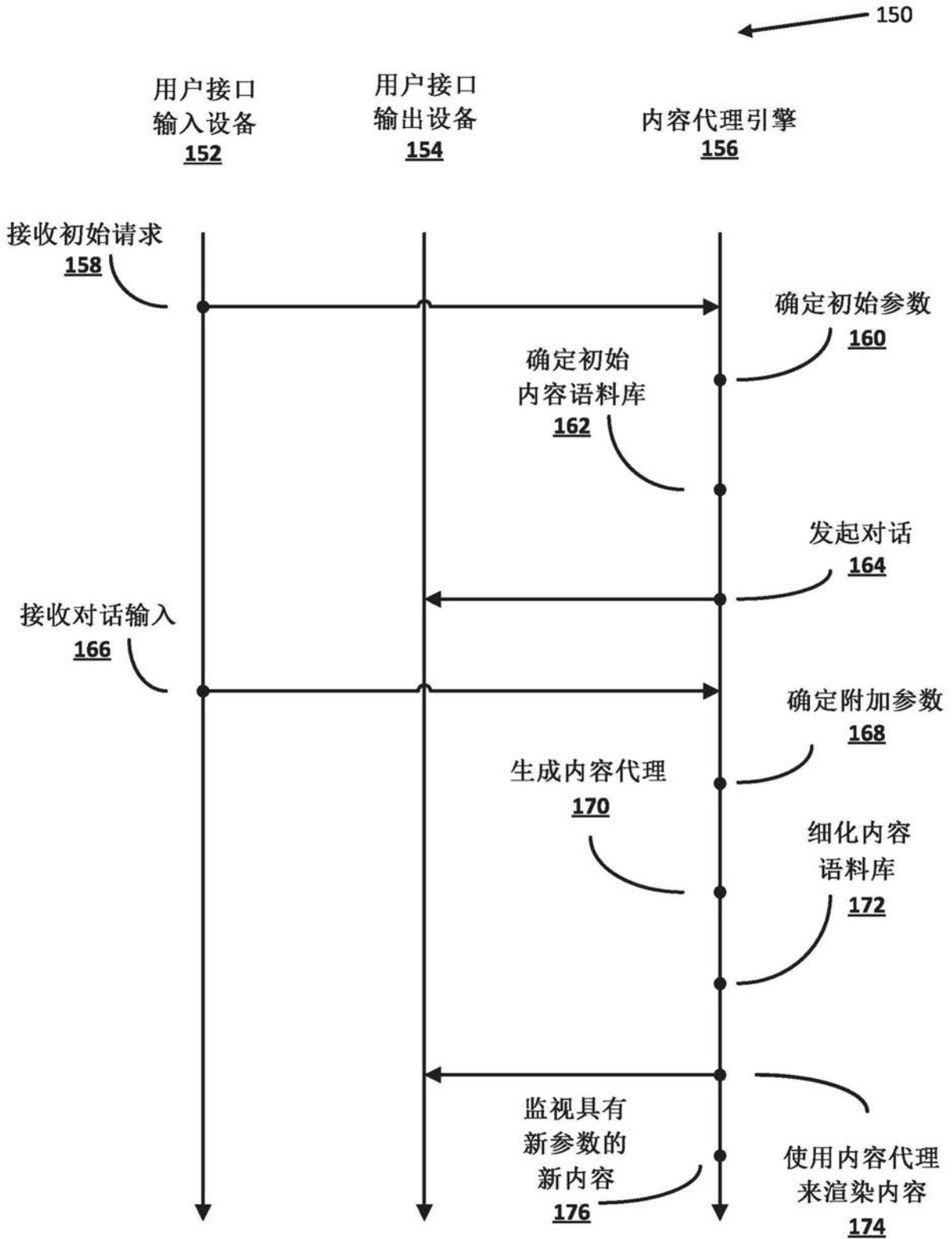


图1B

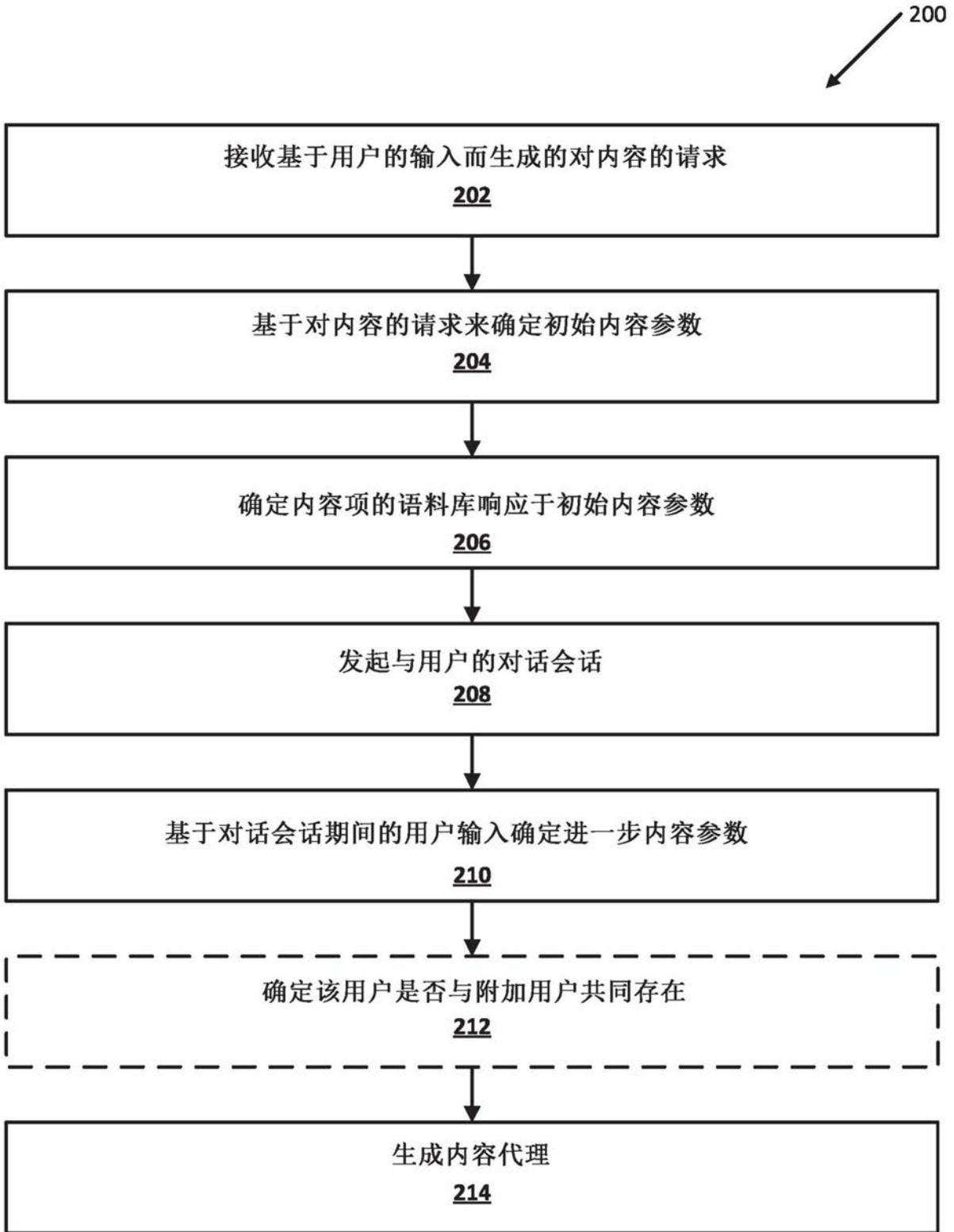


图2

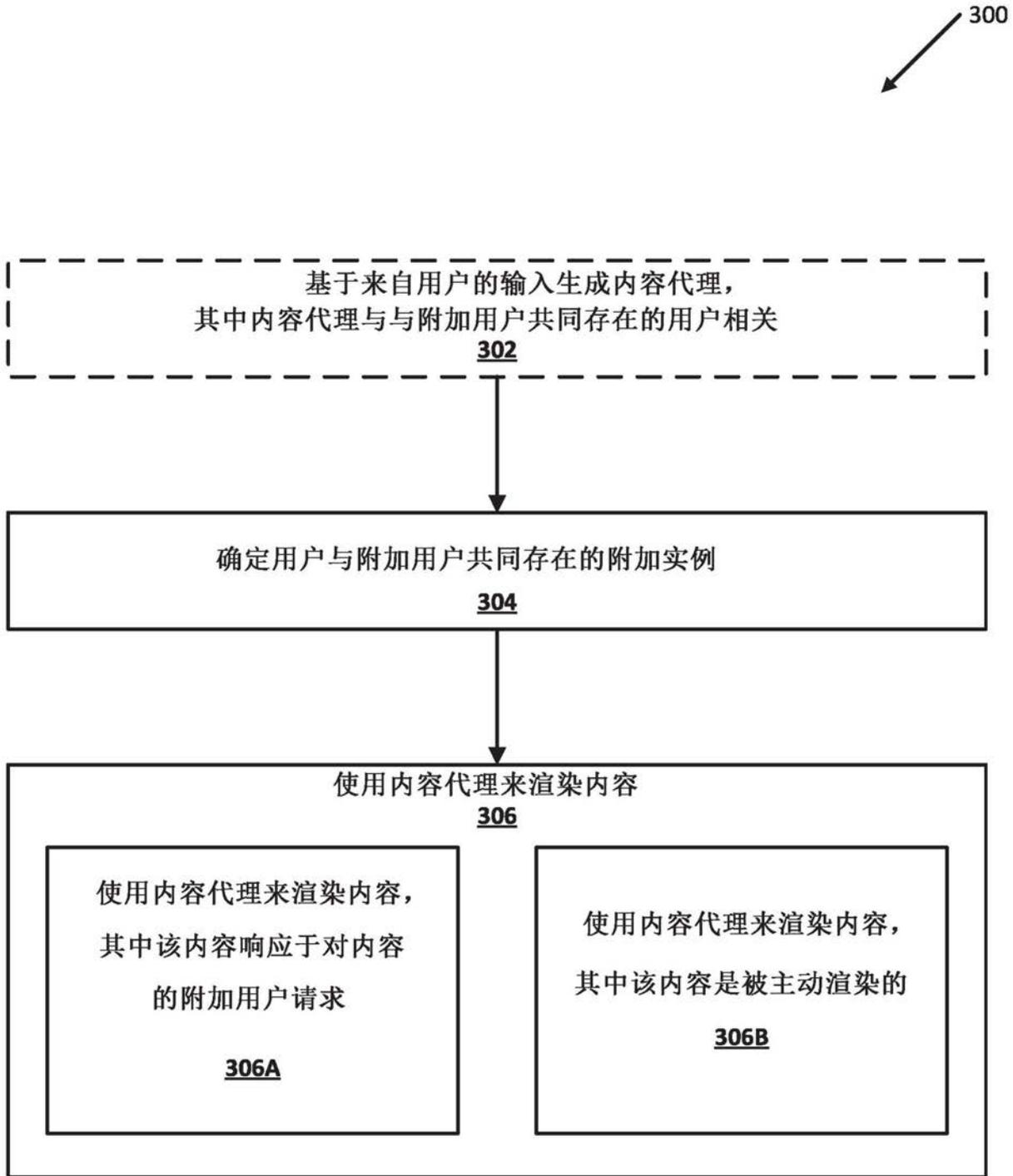


图3

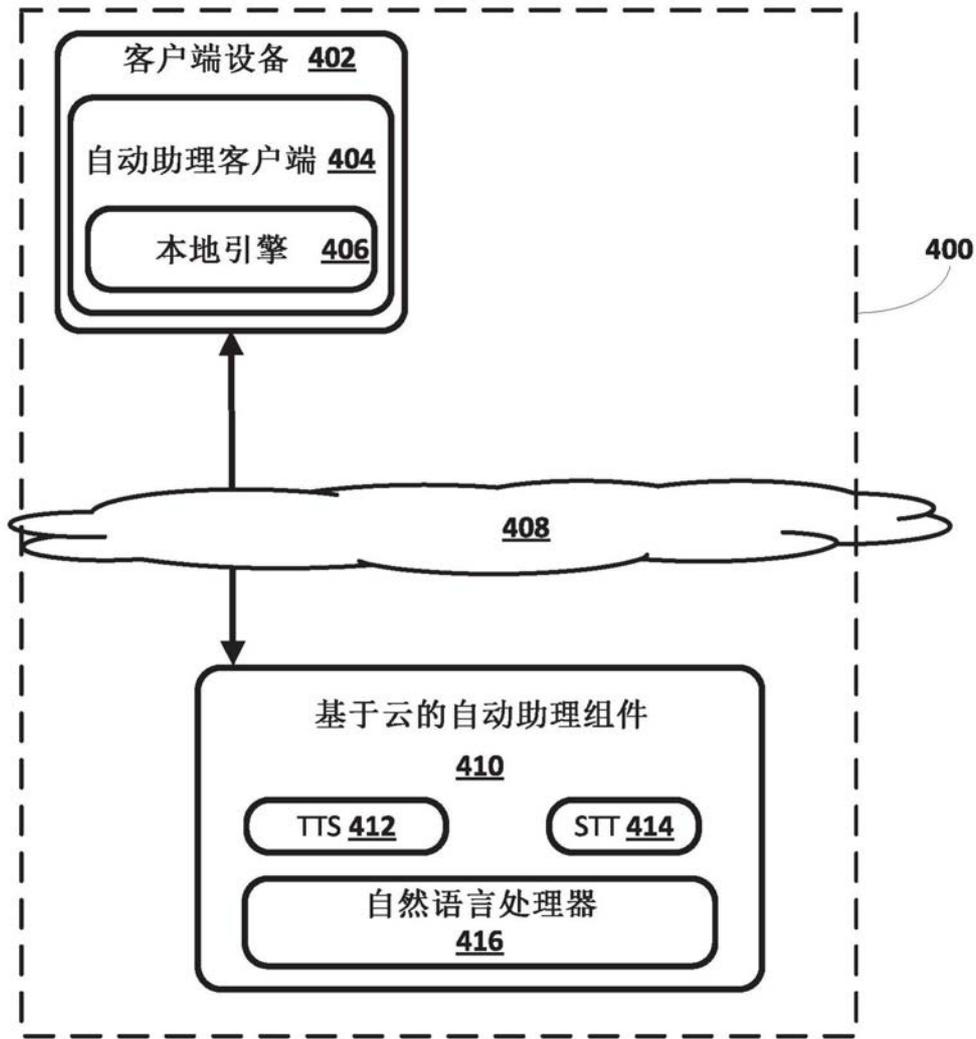


图4

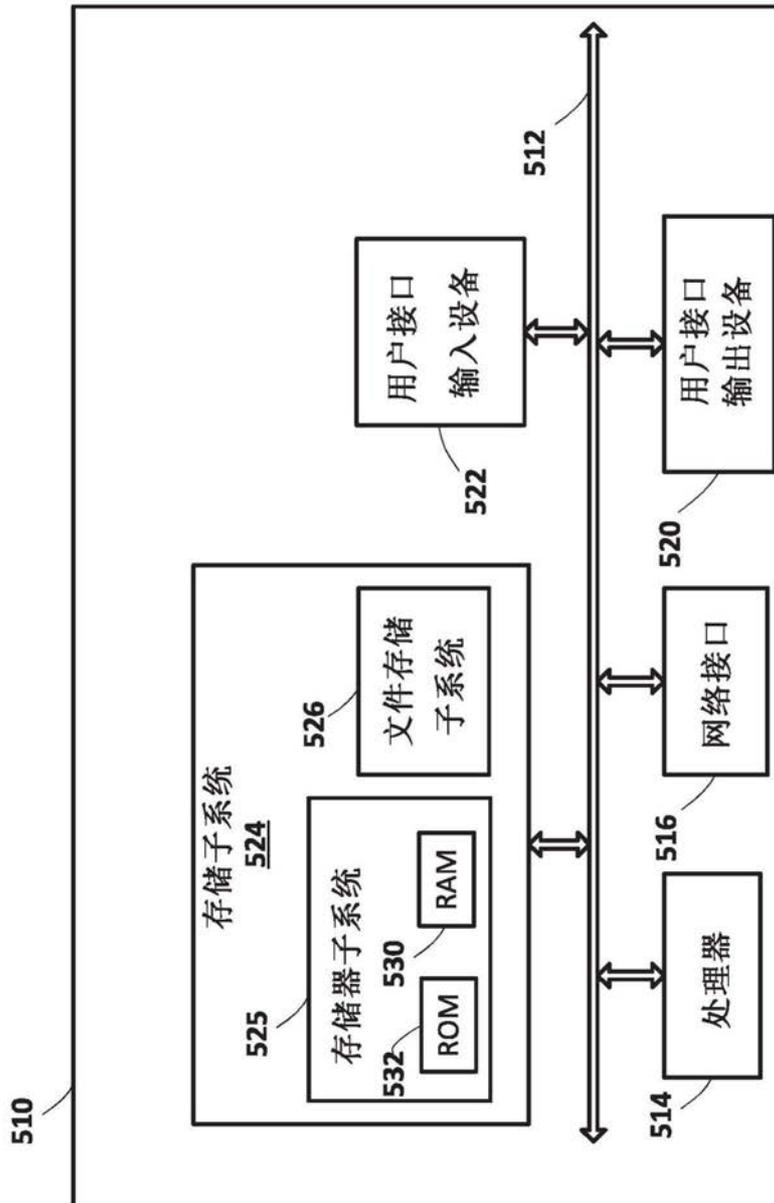


图5