



(12) 发明专利申请

(10) 申请公布号 CN 114556333 A

(43) 申请公布日 2022. 05. 27

(21) 申请号 202080073037.0

(22) 申请日 2020.09.26

(30) 优先权数据

62/923,342 2019.10.18 US

16/659,363 2019.10.21 US

16/703,700 2019.12.04 US

16/733,044 2020.01.02 US

16/741,642 2020.01.13 US

(85) PCT国际申请进入国家阶段日

2022.04.18

(86) PCT国际申请的申请数据

PCT/US2020/052966 2020.09.26

(87) PCT国际申请的公布数据

W02021/076305 EN 2021.04.22

(71) 申请人 脸谱科技有限责任公司

地址 美国加利福尼亚州

(72) 发明人 尼古拉斯·迈克尔·安德鲁·本森

丽萨·晓艺·黄 埃里克·肖

盛雅婷 何子健 艾哈迈德·阿利

阿伦·巴布 阿尔门·阿加贾亚恩

弗朗西斯拉夫·P·佩诺夫

刘洪磊 刘小虎 彭福春 李季航

于劲松 迈克尔·罗伯特·汉森

阿努杰·库马尔 刘白杨

保罗·安东尼·克鲁克

帕拉尔斯·帕雷什·沙

文森特·查尔斯·张 邵瑞

阿尼米什·辛哈

凯施蒂兹·马里克 文胜焕

詹弘远

(74) 专利代理机构 北京安信方达知识产权代理

有限公司 11262

专利代理师 陆建萍 杨明钊

(51) Int. Cl.

G06F 16/9535 (2019.01)

G06F 16/435 (2019.01)

G06F 3/01 (2006.01)

G06F 3/16 (2006.01)

H04N 5/232 (2006.01)

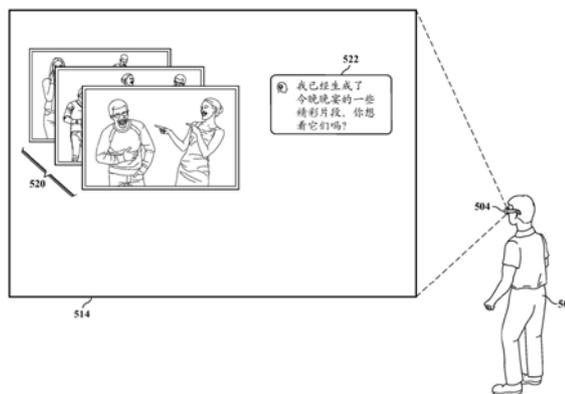
权利要求书3页 说明书43页 附图13页

(54) 发明名称

由助理系统启用的智能摄像机

(57) 摘要

一种系统、介质和方法,包括从与用户相关联的客户端系统接收来自用户的用户输入,其中客户端系统包括一个或多个摄像机,基于一个或多个机器学习模型和由一个或多个摄像机捕获的感测数据确定一个或多个摄像机的视场中的一个或多个兴趣点,基于一个或多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或多个兴趣点中的至少一个的记录,基于多个媒体文件生成一个或多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件,以及向客户端系统发送用于呈现一个或多个精彩片段文件的指令。



1. 一种方法,包括由一个或更多个计算系统:

与用户相关联的客户端系统接收来自所述用户的用户输入,其中所述客户端系统包括一个或更多个摄像机;

基于一个或更多个机器学习模型和由所述一个或更多个摄像机捕获的感测数据,确定所述一个或更多个摄像机的视场中的一个或更多个兴趣点;

基于所述一个或更多个兴趣点生成多个媒体文件,其中每个媒体文件是所述一个或更多个兴趣点中的至少一个的记录;

基于所述多个媒体文件生成一个或更多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件;和

向所述客户端系统发送用于呈现所述一个或更多个精彩片段文件的指令。

2. 根据权利要求1所述的方法,其中所述感测数据基于文本信号、视觉信号或音频信号中的一个或更多个;和/或优选地,其中所述用户输入基于文本输入、音频输入、图像输入、视频输入、眼睛凝视、手势或运动中的一个或更多个;和/或优选地,其中所述多个媒体文件中的每一个包括图像或视频剪辑中的一个或更多个。

3. 根据权利要求1或权利要求2所述的方法,其中确定所述兴趣点包括:

检测所述视场中的一个或更多个人;和

基于一个或更多个面部识别算法,确定检测到的人中的一个或更多个人的一个或更多个标识符;并且优选地,其中确定所述兴趣点是基于所述检测到的人中的一个或更多个人的兴趣度的度量,其中所述兴趣度的度量是由所述一个或更多个机器学习模型基于所述标识符中的一个或更多个来确定的。

4. 根据权利要求1或权利要求2所述的方法,其中确定所述兴趣点包括:

检测所述视场中的一个或更多个人;和

确定检测到的人中的一个或更多个人的一个或更多个面部表情;并且优选地,其中确定所述兴趣点是基于所述检测到的人中的一个或更多个人的兴趣度的度量,其中所述兴趣度的度量是由所述一个或更多个机器学习模型基于所述面部表情中的一个或更多个来确定的。

5. 根据权利要求1或权利要求2所述的方法,其中确定所述兴趣点包括:

检测所述视场中的一个或更多个对象;并且优选地,其中确定所述兴趣点是基于检测到的对象中的一个或更多个的兴趣度的度量,其中所述兴趣度的度量是由所述一个或更多个机器学习模型基于所述检测到的对象中的一个或更多个来确定的。

6. 根据权利要求1至5中任一项所述的方法,其中确定所述兴趣点基于由所述客户端系统捕获的所述用户的眼睛凝视数据。

7. 根据权利要求1至6中任一项所述的方法,其中所述预定义质量标准基于模糊度、亮度或颜色的鲜艳度中的一个或更多个。

8. 根据权利要求1至7中任一项所述的方法,还包括:

响应于所述精彩片段文件,从所述客户端系统接收来自所述用户的用户查询;

访问与所述用户相关联的多个情节记忆;

将所访问的情节记忆中的一个或更多个情节记忆识别为与所述用户查询相关;

检索对应于所识别的情节记忆的一个或更多个媒体文件,其中每个媒体文件包括帖

子、评论、图像或视频剪辑中的一个或多个；和

向所述客户端系统发送用于呈现对应于所识别的情节记忆的一个或多个媒体文件的指令。

9. 根据权利要求1至8中任一项所述的方法,还包括:

向所述客户端系统发送用于拉近所述摄像机中的一个或多个以将所述兴趣点中的一个或多个定位在所述视场的中心的指令。

10. 根据权利要求1至8中任一项所述的方法,还包括:

向所述客户端系统发送用于拉远所述摄像机中的一个或多个以将所述兴趣点中的一个或多个定位在所述视场的中心的指令。

11. 根据权利要求1至10中任一项所述的方法,其中基于以下一个或多个为所述用户个性化所述精彩片段文件:

与所述用户相关联的用户简档数据;

与所述用户相关联的用户偏好;

所述用户的先前用户输入;或者

社交图中用户与其他用户的关系。

12. 根据权利要求1至11中任一项所述的方法,还包括:

从所述客户端系统接收来自所述用户的与一个或多个其他用户共享所述精彩片段文件中的一个或多个的用户请求;和

分别向与所述一个或多个其他用户相关联的一个或多个其他客户端系统发送用于呈现共享的精彩片段文件的指令。

13. 根据权利要求1至12中任一项所述的方法,还包括:

检测所述客户端系统的移动;和

将一个或多个视觉稳定算法应用于由所述一个或多个摄像机捕获的感测数据。

14. 一个或多个计算机可读非暂时性存储介质,其包含软件,所述软件在被执行时能够操作来:

从与用户相关联的客户端系统接收来自所述用户的用户输入,其中所述客户端系统包括一个或多个摄像机;

基于一个或多个机器学习模型和由所述一个或多个摄像机捕获的感测数据,确定所述一个或多个摄像机的视场中的一个或多个兴趣点;

基于所述一个或多个兴趣点生成多个媒体文件,其中每个媒体文件是所述一个或多个兴趣点中的至少一个的记录;

基于所述多个媒体文件生成一个或多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件;和

向所述客户端系统发送用于呈现所述一个或多个精彩片段文件的指令。

15. 一种系统,包括:一个或多个处理器;以及耦合到所述处理器的非暂时性存储器,所述存储器包括能够由所述处理器执行的指令,所述处理器当执行所述指令时能够操作以:

从与用户相关联的客户端系统接收来自所述用户的用户输入,其中所述客户端系统包括一个或多个摄像机;

基于一个或多个机器学习模型和由所述一个或多个摄像机捕获的感测数据,确定所述一个或多个摄像机的视场中的一个或多个兴趣点;

基于所述一个或多个兴趣点生成多个媒体文件,其中每个媒体文件是所述一个或多个兴趣点中的至少一个的记录;

基于所述多个媒体文件生成一个或多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件;和

向所述客户端系统发送用于呈现所述一个或多个精彩片段文件的指令。

由助理系统启用的智能摄像机

技术领域

[0001] 本公开总体上涉及在网络环境内的数据库和文件管理,且具体涉及用于智能助理 (smart assistant) 系统的硬件和软件。

[0002] 背景

[0003] 助理系统可以基于用户输入、位置感知和从各种在线源访问信息 (如天气状况、交通拥堵、新闻、股票价格、用户日程、零售价格等) 的能力的组合来为用户提供信息或服务。用户输入可以包括文本 (例如, 在线聊天) (尤其是在即时消息传递应用或其他应用中的文本), 声音 (voice)、图像、运动 (motion) 或它们的组合。助理系统可以基于用户输入来执行礼宾类型的服务 (例如, 预订晚餐、购买活动门票、安排旅行) 或者提供信息。助理系统还可以基于在线信息和事件来执行管理或数据处理任务, 而无需用户启动或交互。可以由助理系统执行的那些任务的示例可以包括日程管理 (例如, 对晚餐约会发送用户由于交通状况而要迟到的告警信息, 更新双方的日程, 以及改变餐馆预订时间)。助理系统可以通过计算设备、应用编程接口 (API)、以及用户设备上应用激增 (proliferation) 的组合来实现。

[0004] 可以包括社交网络网站的社交网络系统可以使它的用户 (例如个人或组织) 能够与它互动以及通过它彼此互动。社交网络系统可以利用来自用户的输入来在社交网络系统中创建并存储与用户相关联的用户简档 (user profile)。用户简档可以包括用户的人口统计信息、通信渠道信息以及关于个人兴趣的信息。社交网络系统还可以用来自用户的输入来创建并存储用户与社交网络系统的其他用户的关系的记录, 以及提供服务 (例如, 简档/动态消息 (news feed) 帖子、照片分享、事件组织、消息传递、游戏或广告) 以便于在用户之间或当中的社会交互。

[0005] 社交网络系统可以通过一个或更多个网络向用户的移动设备或其他计算设备发送与它的服务相关的内容或消息。用户还可以在用户的移动设备或其他计算设备上安装软件应用, 用于访问用户的用户简档和在社交网络系统内的其他数据。社交网络系统可以生成一组个性化的内容对象以显示给用户, 例如关连 (connect) 到该用户的其他用户的汇集的动态 (story) 的动态消息 (newsfeed)。

[0006] 特定实施例的概述

[0007] 在特定实施例中, 助理系统可以帮助用户获得信息或服务。助理系统可以使用户能够在有状态和多轮会话 (stateful and multi-turn conversation) 中利用多模态用户输入 (例如声音、文本、图像、视频、运动) 与它进行交互来获取帮助。作为示例而非限制, 助理系统可以支持音频 (口头) 输入和非口头输入, 例如视觉、位置、手势、运动或混合/多模态输入。助理系统可以创建和存储用户简档, 该用户简档包括与用户相关联的个人信息和上下文信息 (contextual information)。在特定实施例中, 助理系统可以使用自然语言理解来分析用户输入。分析可以基于用户的用户简档, 以获得更个性化和上下文感知的理解。助理系统可以基于分析来解析 (resolve) 与用户输入相关联的实体。在特定实施例中, 助理系统可以与不同的代理交互, 以获得与解析的实体相关联的信息或服务。助理系统可以通过使用自然语言生成为用户生成关于信息或服务的响应。通过与用户的交互, 助理系统可以

使用对话管理技术来管理和推进与用户的会话流。在特定实施例中,助理系统还可以通过汇总(summarize)信息来辅助用户有效且高效地消化所获得的信息。助理系统还可以通过提供帮助用户与在线社交网络交互(例如,创建帖子、评论、消息)的工具,来帮助用户更好地参与在线社交网络。助理系统另外可以帮助用户管理不同的任务,如持续跟踪事件。在特定实施例中,助理系统可以在没有用户输入的情况下,基于用户简档,在与用户相关的时间主动执行与用户兴趣和偏好相关的任务。在特定实施例中,助理系统可以检查隐私设置,以确保根据用户的隐私设置来允许访问用户的简档或其他用户信息以及执行不同的任务。

[0008] 在特定实施例中,助理系统可以通过建立在客户端进程和服务器端进程上的混合架构来帮助用户。客户端进程和服务器端进程可以是用于处理用户输入和向用户提供帮助的两个并行工作流。在特定实施例中,客户端进程可以在与用户相关联的客户端系统上本地执行。相比之下,服务器端进程可以在一个或更多个计算系统上远程执行。在特定实施例中,客户端系统上的仲裁器可以协调接收用户输入(例如,音频信号),确定是使用客户端进程还是服务器端进程或两者来响应用户输入,以及分析来自每个进程的处理结果。仲裁器可以基于上述分析指示客户端或服务器端的代理执行与用户输入相关联的任务。执行结果可以进一步作为输出渲染给客户端系统。通过利用客户端和服务器端进程,助理系统可以有效地帮助用户优化计算资源的使用,同时保护用户隐私和增强安全性。

[0009] 在特定实施例中,当使用助理驱动的启用摄像机的客户端系统(例如,具有摄像机的智能电话、AR/VR眼镜、智能显示器)时,助理系统可以通过向用户提供不用手的图片/视频捕获、共享和记忆回忆来改善助理系统的用户体验。助理系统可以具有对象/人识别、兴趣点确定和精彩片段(highlight)生成的附加功能。助理系统可以被动地(reactively)和主动地帮助用户。在被动场景中,可存在从用户经由客户端系统到助理系统的显式命令或提示(例如,“拍照”)。然后,助理系统可以将其转换为机器标准意图,并通过启用了摄像机的客户端系统执行相关联的任务来为用户服务。在主动场景中,助理系统本身可以基于其通过客户端系统接收的信号,例如来自AR眼镜的用户的眼睛凝视,来确定其可以代表用户通过启用摄像机的客户端系统采取的潜在有趣的动作。尽管本公开描述了通过特定系统以特定方式在特定基于摄像机的任务中帮助用户,但是本公开设想通过任何合适的系统以任何合适的方式在任何合适的基于摄像机的任务中帮助用户。

[0010] 在特定实施例中,助理系统可以从与用户相关联的客户端系统接收来自用户的用户输入。客户端系统可以包括一个或更多个摄像机。在特定实施例中,助理系统然后可以基于一个或更多个机器学习模型和由一个或更多个摄像机捕获的感测数据,确定一个或更多个摄像机的视场中的一个或更多个兴趣点。助理系统然后可以基于一个或更多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或更多个兴趣点中的至少一个的记录。在特定实施例中,助理系统然后可以基于多个媒体文件生成一个或更多个精彩片段文件。每个精彩片段文件可以包括满足预定质量标准的媒体文件。在特定实施例中,助理系统还可以向客户端系统发送用于呈现一个或更多个精彩片段文件的指令。

[0011] 通过助理系统启用智能摄像机存在某些技术挑战。一个技术挑战可能包括准确确定兴趣点。本文公开的实施例提出的解决上述挑战的解决方案可以是使用不同的机器学习算法分析由客户端系统捕获的感测数据,并基于该分析确定兴趣度,这可以导致兴趣点的更精确的确定,因为机器学习算法有助于定位用户在视场中的兴趣,并且兴趣度进一步细

化了这种兴趣的确定。另一个技术挑战可能包括生成对于用户有意义和感兴趣的精彩片段文件。由本文公开的实施例提出的解决这一挑战的解决方案可以是基于用户简档数据、用户偏好、先前的用户输入和用户关系数据来个性化精彩片段文件,这些都提供了关于当用户回顾由助理系统捕获的时刻时什么对用户来说是有意义和有趣的特定见解。

[0012] 本文公开的某些实施例可以提供一个或多个技术优势。这些实施例的技术优势可以包括帮助用户方便地以高质量捕获感兴趣和有意义的时刻,因为助理系统可以使用户能够使用各种输入并应用不同的校准和处理算法来确保对应于这些时刻的媒体文件具有高质量。实施例的另一个技术优势可以包括帮助用户将当前捕获的时刻与用户的记忆相关连,从而用户可以更好地理解它们,因为助理系统可以执行记忆回忆来检索与当前捕获的时刻相关的媒体文件。本文公开的某些实施例可以不提供、提供一些或全部上述技术优势。鉴于本公开的附图、说明书和权利要求,一个或多个其他技术优势对于本领域技术人员来说可以是显而易见的。

[0013] 根据一些实施例,提供了一种方法,方法包括由一个或多个计算系统:从与用户相关联的客户端系统接收来自用户的用户输入,其中客户端系统包括一个或多个摄像机;基于一个或多个机器学习模型和由一个或多个摄像机捕获的感测数据,确定一个或多个摄像机的视场中的一个或多个兴趣点;基于一个或多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或多个兴趣点中的至少一个的记录;基于多个媒体文件生成一个或多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件;以及向客户端系统发送用于呈现一个或多个精彩片段文件的指令。

[0014] 在一些实施例中,感测数据可以基于文本信号、视觉信号或音频信号中的一个或多个。

[0015] 在一些实施例中,用户输入可以基于文本输入、音频输入、图像输入、视频输入、眼睛凝视、手势或运动中的一个或多个。

[0016] 在一些实施例中,多个媒体文件中的每一个可以包括图像或视频剪辑中的一个或多个。

[0017] 在一些实施例中,确定兴趣点可以包括:检测视场中的一个或多个个人;以及基于一个或多个面部识别算法,确定检测到的人中的一个或多个个人的一个或多个标识符。

[0018] 在一些实施例中,确定兴趣点可以基于检测到的人中的一个或多个人的兴趣度的度量,其中兴趣度的度量由一个或多个机器学习模型基于标识符中的一个或多个来确定。

[0019] 在一些实施例中,确定兴趣点可以包括:检测视场中的一个或多个个人;以及确定检测到的人中的一个或多个人的一个或多个面部表情。

[0020] 在一些实施例中,确定兴趣点可以基于检测到的人中的一个或多个人的兴趣度的度量,其中兴趣度的度量由一个或多个机器学习模型基于面部表情中的一个或多个来确定。

[0021] 在一些实施例中,确定兴趣点可以包括:检测视场中的一个或多个对象。

[0022] 在一些实施例中,确定兴趣点可以基于检测到的对象中的一个或多个的兴趣度的度量,其中兴趣度的度量由一个或多个机器学习模型基于检测到的对象中的一个或更

多个来确定。

[0023] 在一些实施例中,确定兴趣点可以基于客户端系统捕获的用户的眼睛凝视数据。

[0024] 在一些实施例中,预定义的质量标准可以基于模糊度、亮度或颜色的鲜艳度中的一个或更多个。

[0025] 在一些实施例中,该方法可以进一步包括:响应于精彩片段文件,从客户端系统接收来自用户的用户查询;访问与用户相关联的多个情节记忆(episodic memory);将所访问的情节记忆中的一个或更多个情节记忆识别为与用户查询相关;检索对应于所识别的情节记忆的一个或更多个媒体文件,其中每个媒体文件包括帖子、评论、图像或视频剪辑中的一个或更多个;以及向客户端系统发送用于呈现对应于所识别的情节记忆的一个或更多个媒体文件的指令。

[0026] 在一些实施例中,该方法可以进一步包括:向客户端系统发送用于拉近摄像机中的一个或更多个以将兴趣点中的一个或更多个定位在视场的中心的指令。

[0027] 在一些实施例中,该方法可以进一步包括:向客户端系统发送用于拉远摄像机中的一个或更多个以将兴趣点中的一个或更多个定位在视场的中心的指令。

[0028] 在一些实施例中,可以基于以下一个或更多个为用户个性化精彩片段文件:与用户相关联的用户简档数据;与用户相关联的用户偏好;用户的先前用户输入;或者用户与社交图中其他用户的关系。

[0029] 在一些实施例中,该方法可以进一步包括:从客户端系统接收来自用户的用户请求,以与一个或更多个其他用户共享精彩片段文件中的一个或更多个;以及分别向与一个或更多个其他用户相关联的一个或更多个其他客户端系统发送用于呈现共享的精彩片段文件的指令。

[0030] 在一些实施例中,该方法可以进一步包括:检测客户端系统的移动;以及将一个或更多个视觉稳定算法应用于由一个或更多个摄像机捕获的感测数据。

[0031] 根据本发明的实施例,提供了一种或更多种计算机可读非暂时性存储介质,其包含软件,该软件在被执行时可操作来:从与用户相关联的客户端系统接收来自用户的用户输入,其中该客户端系统包括一个或更多个摄像机;基于一个或更多个机器学习模型和由一个或更多个摄像机捕获的感测数据,确定一个或更多个摄像机的视场中的一个或更多个兴趣点;基于一个或更多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或更多个兴趣点中的至少一个的记录;基于多个媒体文件生成一个或更多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件;以及向客户端系统发送用于呈现一个或更多个精彩片段文件的指令。

[0032] 根据本发明的实施例,提供了一种系统,系统包括:一个或更多个处理器;以及耦合到处理器的非暂时性存储器,其包括可由处理器执行的指令,处理器在执行指令时可操作以:从与用户相关联的客户端系统接收来自用户的用户输入,其中客户端系统包括一个或更多个摄像机;基于一个或更多个机器学习模型和由一个或更多个摄像机捕获的感测数据,确定一个或更多个摄像机的视场中的一个或更多个兴趣点;基于一个或更多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或更多个兴趣点中的至少一个的记录;基于多个媒体文件生成一个或更多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件;以及向客户端系统发送用于呈现一个或更多个精彩片段文件的指令。

[0033] 本文公开的实施例仅仅是示例,并且本公开的范围不限于它们。特定实施例可以包括本文公开的实施例的部件、元件、特征、功能、操作或步骤中的全部、一些,或不包括这些部件、元件、特征、功能、操作或步骤。根据本发明的实施例在所附权利要求中具体公开,涉及方法、存储介质、系统和计算机程序产品,其中在一个权利要求类别(例如方法)中提到的任何特征,也可以在另一个权利要求类别(例如系统)中要求保护。在所附权利要求中的从属性或往回引用仅为了形式原因而被选择。然而,也可以要求保护由对任何先前权利要求的有意往回引用(特别是多个从属关系)而产生的任何主题,从而权利要求及其特征的任何组合被公开了并且可以被要求保护,而不管所附权利要求中选择的从属关系如何。可以被要求保护的主体不仅包括如在所附权利要求中阐述的特征的组合,而且还包括在权利要求中的特征的任何其他组合,其中,在权利要求中提到的每个特征可以与在权利要求中的任何其他特征或其他特征的组合相结合。此外,本文描述或描绘的实施例和特征中的任何一个可以在单独的权利要求中和/或以与本文描述或描绘的任何实施例或特征的任何组合或以与所附权利要求的任何特征的任何组合被要求保护。

[0034] 附图简述

[0035] 图1示出了与助理系统相关联的示例网络环境。

[0036] 图2示出了助理系统的示例架构。

[0037] 图3示出了助理系统的服务器端进程的示例流程图。

[0038] 图4示出了由助理系统处理用户输入的示例流程图。

[0039] 图5A示出了智能摄像机可以帮助用户的示例场景。

[0040] 图5B示出了兴趣点的示例确定。

[0041] 图5C示出了兴趣点的另一示例确定。

[0042] 图5D示出了示例精彩片段文件。

[0043] 图6示出了通过助理系统启用智能摄像机的示例方法。

[0044] 图7示出了示例社交图。

[0045] 图8示出了嵌入空间的示例视图。

[0046] 图9示出了示例人工神经网络。

[0047] 图10示出了示例计算机系统。

[0048] 示例实施例的描述

[0049] 系统概况

[0050] 图1示出了与助理系统相关联的示例网络环境100。网络环境100包括通过网络110彼此连接的客户端系统130、助理系统140、社交网络系统160和第三方系统170。尽管图1示出了客户端系统130、助理系统140、社交网络系统160、第三方系统170和网络110的特定布置,但是本公开设想了客户端系统130、助理系统140、社交网络系统160、第三方系统170和网络110的任何合适的布置。作为示例而不是作为限制,客户端系统130、社交网络系统160、助理系统140和第三方系统170中的两个或更多个可以绕过网络110彼此直接连接。作为另一个示例,客户端系统130、助理系统140、社交网络系统160和第三方系统170中的两个或更多个可以全部或部分地在物理上或逻辑上彼此位于同一位置。此外,尽管图1示出了特定数量的客户端系统130、助理系统140、社交网络系统160、第三方系统170和网络110,但是本公开设想了任何合适数量的客户端系统130、助理系统140、社交网络系统160、第三方系统170

和网络110。作为示例而不是作为限制,网络环境100可以包括多个客户端系统130、助理系统140、社交网络系统160、第三方系统170和网络110。

[0051] 本公开设想任何合适的网络110。作为示例而不是作为限制,网络110的一个或多个部分可以包括自组织网络、内联网、外联网、虚拟专用网络(VPN)、局域网(LAN)、无线LAN(WLAN)、广域网(WAN)、无线WAN(WWAN)、城域网(MAN)、互联网的一部分、公共交换电话网(PSTN)的一部分、蜂窝电话网、或这些中的两个或多个的组合。网络110可以包括一个或多个网络110。

[0052] 链路150可以将客户端系统130、助理系统140、社交网络系统160和第三方系统170连接到通信网络110或连接到彼此。本公开设想任何合适的链路150。在特定实施例中,一个或多个链路150包括一个或多个有线(诸如例如数字用户线路(DSL)或基于电缆的数据服务接口规范(DOCSIS))链路、无线(诸如例如Wi-Fi或全球互通微波接入(WiMAX))链路、或光(诸如例如同步光网络(SONET)或同步数字体系(SDH))链路。在特定实施例中,一个或多个链路150各自包括自组织网络、内联网、外联网、VPN、LAN、WLAN、WAN、WWAN、MAN、互联网的一部分、PSTN的一部分、基于蜂窝技术的网络、基于卫星通信技术的网络、另一链路150、或两个或多个这种链路150的组合。链路150不需要在整个网络环境100中一定是相同的。一个或多个第一链路150可以在一个或多个方面上不同于一个或多个第二链路150。

[0053] 在特定实施例中,客户端系统130可以是一种电子设备,该电子设备包括硬件、软件或嵌入式逻辑部件、或两个或多个这样的部件的组合,并且能够执行由客户端系统130实现或支持的适当功能。作为示例而不是作为限制,客户端系统130可以包括计算机系统,例如台式计算机、笔记本或膝上型计算机、上网本、平板计算机、电子书阅读器、GPS设备、摄像机、个人数字助理(PDA)、手持电子设备、蜂窝电话、智能手机、智能扬声器、虚拟现实(VR)头戴式装置、增强现实(AR)智能眼镜、其他合适的电子设备、或其任何合适的组合。在特定实施例中,客户端系统130可以是智能助理设备。关于智能助理设备的更多信息可以在2018年4月9日提交的美国专利申请第15/949011号、2018年10月5日提交的美国专利申请第16/153574号、2018年1月3日提交的美国外观设计专利申请第29/631910号、2018年1月2日提交的美国外观设计专利申请第29/631747号、2018年1月3日号提交的美国外观设计专利申请第29/631913号和2018年1月3日提交的美国外观设计专利申请第29/631914号中找到,本文将讨论这些专利申请中的每一个。本公开设想任何合适的客户端系统130。客户端系统130可以使在客户端系统130处的网络用户能够访问网络110。客户端系统130可以使它的用户能够与在其他客户端系统130处的其他用户进行通信。

[0054] 在特定实施例中,客户端系统130可以包括web浏览器132,并且可以具有一个或多个附加件、插件或其他扩展件。在客户端系统130处的用户可以输入统一资源定位符(URL)或将web浏览器132引导到特定的服务器(例如服务器162或与第三方系统170相关联的服务器)的其他地址,并且web浏览器132可以生成超文本传输协议(HTTP)请求并将HTTP请求传递到服务器。服务器可以接受HTTP请求,并响应于HTTP请求而向客户端系统130传递一个或多个超文本标记语言(HTML)文件。客户端系统130可以基于来自服务器的HTML文件来显现web界面(例如网页)用于呈现给用户。本公开设想任何合适的源文件。作为示例而不是作为限制,可以根据特定的需要根据HTML文件、可扩展超文本标记语言(XHTML)文件

或可扩展标记语言 (XML) 文件来显现web界面。这样的界面还可以执行脚本、标记语言和脚本的组合等。在本文,在适当的情况下,对web界面的引用包括一个或更多个相应的源文件(浏览器可以使用这些源文件来显现web界面),反之亦然。

[0055] 在特定实施例中,客户端系统130可以包括安装在客户端系统130上的社交网络应用134。客户端系统130处的用户可以使用社交网络应用134来访问在线社交网络。客户端系统130处的用户可以使用社交网络应用134来与用户的社交关连(connection)(例如,朋友、关注者(follower)、关注账户、联系人等)进行通信。客户端系统130处的用户还可以使用社交网络应用134来与在线社交网络上的多个内容对象(例如,帖子、新闻文章、临时内容等)交互。作为示例而不是作为限制,用户可以使用社交网络应用134浏览热门话题和突发新闻。

[0056] 在特定实施例中,客户端系统130可以包括助理应用136。客户端系统130的用户可以使用助理应用136来与助理系统140交互。在特定实施例中,助理应用136可以包括独立的应用。在特定实施例中,助理应用136可以被集成到社交网络应用134或另一个合适的应用(例如,消息传递应用)中。在特定实施例中,助理应用136也可以集成到客户端系统130、助理硬件设备或任何其他合适的硬件设备中。在特定实施例中,可以经由web浏览器132访问助理应用136。在特定实施例中,用户可以经由不同的模态提供输入。作为示例而非限制,模态可以包括音频、文本、图像、视频、运动、定向等。助理应用136可以将用户输入传送给助理系统140。基于用户输入,助理系统140可以生成响应。助理系统140可以将生成的响应发送给助理应用136。然后,助理应用136可以向客户端系统130的用户呈现响应。所呈现的响应可以基于不同的模态,如音频、文本、图像和视频。作为示例而非限制,用户可以通过对着客户端系统130的麦克风说话来口头就交通信息询问助理应用136(即,经由音频模态)。然后,助理应用136可以将请求传送给助理系统140。助理系统140可以相应地生成响应并将其发送回助理应用136。助理应用136还可以在客户端系统130的显示器上以文本和/或图像向用户呈现响应。

[0057] 在特定实施例中,助理系统140可以帮助用户从不同的源检索信息。助理系统140还可以帮助用户向不同的服务提供者请求服务。在特定实施例中,助理系统140可以经由客户端系统130中的助理应用136接收用户对信息或服务的请求。助理系统140可以使用自然语言理解来基于用户简档和其他相关信息分析用户请求。分析的结果可以包括与在线社交网络相关联的不同实体。然后,助理系统140可以检索信息或请求与这些实体相关联的服务。在特定实施例中,当为用户检索信息或请求服务时,助理系统140可以与社交网络系统160和/或第三方系统170交互。在特定实施例中,助理系统140可以使用自然语言生成技术为用户生成个性化的通信内容。个性化通信内容可以包括例如检索到的信息或所请求服务的状态。在特定实施例中,助理系统140可以使用户能够通过使用对话管理技术来在有状态和多轮会话中与它进行关于信息或服务的交互。在下面图2的讨论中更详细地描述了助理系统140的功能。

[0058] 在特定实施例中,社交网络系统160是可以托管在线社交网络的网络可寻址计算系统。社交网络系统160可以生成、存储、接收并发送社交网络数据(诸如例如,用户简档数据、概念简档数据、社交图信息或与在线社交网络相关的其他合适的的数据)。社交网络系统160可以由网络环境100的其他部件直接地或经由网络110来访问。作为示例而不是作

为限制,客户端系统130可以使用web浏览器132或与社交网络系统160相关联的原生(native)应用(例如,移动社交网络应用、消息传递应用、另一合适的应用或其任何组合)直接地或经由网络110来访问社交网络系统160。在特定实施例中,社交网络系统160可以包括一个或更多个服务器162。每个服务器162可以是整体式服务器(unitary server)或跨越多台计算机或多个数据中心的分布式服务器。服务器162可以具有各种类型,例如且没有限制,web服务器、新闻服务器、邮件服务器、消息服务器、广告服务器、文件服务器、应用服务器、交换服务器、数据库服务器、代理服务器、适合于执行本文描述的功能或过程的另一服务器、或者其任何组合。在特定实施例中,每个服务器162可以包括硬件、软件或嵌入式逻辑部件、或用于执行由服务器162实现或支持的适当功能的两个或更多个这样的部件的组合。在特定实施例中,社交网络系统160可以包括一个或更多个数据储存器164。数据储存器164可以用于存储各种类型的信息。在特定实施例中,可以根据特定的数据结构来组织存储在数据储存器164中的信息。在特定实施例中,每个数据储存器164可以是关系数据库、纵列(columnar)数据库、相关性数据库或其他合适的数据库。尽管本公开描述或示出了特定类型的数据库,但是本公开设想了任何合适类型的数据库。特定实施例可以提供使客户端系统130、社交网络系统160、助理系统140或第三方系统170能够管理、检索、修改、添加或删除存储在数据储存器164中的信息的接口。

[0059] 在特定实施例中,社交网络系统160可以在一个或更多个数据储存器164中存储一个或更多个社交图。在特定实施例中,社交图可以包括多个节点——其可以包括多个用户节点(各自对应于特定用户)或多个概念节点(各自对应于特定概念)——以及连接节点的多条边。社交网络系统160可以向在线社交网络的用户提供与其他用户通信和交互的能力。在特定实施例中,用户可以经由社交网络系统160加入在线社交网络,且然后添加与社交网络系统160中的他们想要关连到的多个其他用户的关连(例如,关系)。在本文,术语“朋友”可以指用户经由社交网络系统160与其形成关连、关联(association)或关系的社交网络系统160的任何其他用户。

[0060] 在特定实施例中,社交网络系统160可以向用户提供对由社交网络系统160支持的各种类型的项目或对象采取动作的能力。作为示例而不是作为限制,项目和对象可以包括社交网络系统160的用户可以属于的群组或社交网络、用户可能感兴趣的事件或日历条目、用户可以使用的基于计算机的应用、允许用户经由服务来购买或销售商品的交易、用户可以执行的与广告的交互、或其他合适的项目或对象。用户可以与能够在社交网络系统160中或者由第三方系统170的外部系统表示的任何事物进行交互,第三方系统170与社交网络系统160分离并且经由网络110耦合到社交网络系统160。

[0061] 在特定实施例中,社交网络系统160能够链接各种实体。作为示例而不是作为限制,社交网络系统160可以使用户能够彼此交互以及从第三方系统170或其他实体接收内容,或者允许用户通过应用编程接口(API)或其他通信渠道与这些实体交互。

[0062] 在特定实施例中,第三方系统170可以包括一种或更多种类型的服务器、一个或更多个数据储存器、一个或更多个接口(包括但不限于API)、一个或更多个web服务、一个或更多个内容源、一个或更多个网络或任何其他合适的部件(例如,服务器可以与这些部件通信)。第三方系统170可以由与操作社交网络系统160的实体不同的实体进行操作。然而,在特定实施例中,社交网络系统160和第三方系统170可以结合彼此来操作以向社交网络系统

160或第三方系统170的用户提供社交网络服务。在这个意义上,社交网络系统160可以提供平台或骨干网,其他系统(例如第三方系统170)可以使用该平台或骨干网来在整个互联网上向用户提供社交网络服务和功能。

[0063] 在特定实施例中,第三方系统170可以包括第三方内容对象提供者。第三方内容对象提供者可以包括可以被传递到客户端系统130的内容对象的一个或更多个源。作为示例而不是作为限制,诸如,内容对象可以包括关于用户感兴趣的事情或活动的信息,例如电影放映时间、电影评论、餐馆评论、餐馆菜单、产品信息和评论或其他合适的信息。作为另一示例而不是作为限制,内容对象可以包括激励内容对象(例如优惠券、折扣券、礼品券或其他合适的激励对象)。在特定实施例中,第三方内容提供商可以使用一个或更多个第三方代理来提供内容对象和/或服务。第三方代理可以是在第三方系统170上托管和执行的实现。

[0064] 在特定实施例中,社交网络系统160还包括用户生成的内容对象,其可以增强用户与社交网络系统160的交互。用户生成的内容可以包括用户可以添加、上传、发送或“发布”到社交网络系统160的任何内容。作为示例而不是作为限制,用户将帖子从客户端系统130传送到社交网络系统160。帖子可以包括数据,例如状态更新或其他文本数据、位置信息、照片、视频、链接、音乐或其他类似数据或媒体。内容还可以由第三方通过“通信渠道”(例如动态消息或流)来添加到社交网络系统160。

[0065] 在特定实施例中,社交网络系统160可以包括各种服务器、子系统、程序、模块、日志和数据储存器。在特定实施例中,社交网络系统160可以包括下列项中的一个或更多个:web服务器、动作记录器、API请求服务器、相关性和排序引擎、内容对象分类器、通知控制器、动作日志、第三方内容对象暴露日志、推理模块、授权/隐私服务器、搜索模块、广告定位模块(advertisement-targeting module)、用户界面模块、用户简档储存器、关连储存器、第三方内容储存器或定位储存器。社交网络系统160还可以包括合适的部件,如网络接口、安全机构、负载平衡器、故障转移服务器、管理和网络操作控制台、其他合适的部件、或其任何合适的组合。在特定实施例中,社交网络系统160可以包括用于存储用户简档的一个或更多个用户简档储存器。用户简档可以包括例如,传记信息、人口统计信息、行为信息、社交信息或其他类型的描述性信息(例如工作经历、教育历史、爱好或偏好、兴趣、亲和力或定位)。兴趣信息可以包括与一个或更多个类别相关的兴趣。类别可以是一般的或特定的。作为示例而不是作为限制,如果用户“赞(like)”关于一种品牌的鞋的文章,该类别可以是品牌,或者“鞋”或“衣服”的一般类别。关连储存器可以用于存储关于用户的关连信息。关连信息可以指示具有相似或共同的工作经历、组成员资格、爱好、教育历史或者以任何方式相关或共享共同属性的用户。关连信息还可以包括在不同用户和内容(内部和外部)之间的用户定义的关连。web服务器可以用于经由网络110将社交网络系统160链接到一个或更多个客户端系统130或一个或更多个第三方系统170。web服务器可以包括邮件服务器或用于在社交网络系统160和一个或更多个客户端系统130之间接收并路由(routing)消息的其他消息传递功能。API请求服务器可以允许例如助理系统140或第三方系统170通过调用一个或更多个API来访问来自社交网络系统160的信息。动作记录器可以用于从web服务器接收关于用户在社交网络系统160上或之外的动作的通信。结合动作日志,可以维护用户暴露于第三方内容对象的第三方内容对象日志。通知控制器可以向客户端系统130提供关于内容对象的信息。信息可以作为通知被推送到客户端系统130,或者信息可以响应于从客户端系统130接

收的请求而从客户端系统130中被拉取。授权服务器可以用于实施社交网络系统160的用户的一个或多个隐私设置。用户的隐私设置确定与用户相关联的特定信息可以如何被分享。授权服务器可以诸如例如通过设置适当的隐私设置来允许用户选择加入或选择退出使他们的动作由社交网络系统160记录或者与其他系统(例如,第三方系统170)共享。第三方内容对象存储器可以用于存储从第三方(例如第三方系统170)接收的内容对象。定位存储器可以用于存储从与用户相关联的客户端系统130接收的定位信息。广告定价模块可以组合社交信息、当前时间、定位信息或其他合适的信息以用通知的形式向用户提供相关广告。

[0066] 助理系统

[0067] 图2示出了助理系统140的示例架构。在特定实施例中,助理系统140可以帮助用户获得信息或服务。助理系统140可以使用户能够在有状态和多轮会话中用多模态用户输入(如声音、文本、图像、视频、运动)与其进行交互,以获得帮助。作为示例而非限制,助理系统140可以支持音频输入(口头)和非口头输入,例如视觉、位置、手势、运动或混合/多模态输入。助理系统140可以创建和存储包括与用户相关联的个人信息和上下文信息的用户简档。在特定实施例中,助理系统140可以使用自然语言理解来分析用户输入。分析可以基于用户的用户简档,以获得更个性化和上下文感知的理解。助理系统140可以基于分析来解析与用户输入相关联的实体。在特定实施例中,助理系统140可以与不同的代理交互,以获得与解析的实体相关联的信息或服务。助理系统140可以通过使用自然语言生成来为用户生成关于信息或服务的响应。通过与用户的交互,助理系统140可以使用对话管理技术来管理和转发与用户的会话流。在特定实施例中,助理系统140还可以通过汇总信息来帮助用户有效和高效地消化所获得的信息。助理系统140还可以通过提供帮助用户与在线社交网络交互(例如,创建帖子、评论、消息)的工具,来帮助用户更多地参与在线社交网络。助理系统140另外可以帮助用户管理不同的任务,如持续跟踪事件。在特定实施例中,助理系统140可以在与用户相关的时间,在没有用户输入的情况下,基于用户简档主动执行与用户兴趣和偏好相关的预授权任务。在特定实施例中,助理系统140可以检查隐私设置,以确保根据用户的隐私设置,访问用户的简档或其他用户信息以及执行不同的任务是允许的。在2018年11月6日提交的美国专利申请第16/182542号中可以找到更多关于帮助受隐私设置约束的用户的信息,在本文对其进行讨论。

[0068] 在特定实施例中,助理系统140可以通过建立在客户端进程和服务器端进程上的混合架构来帮助用户。客户端进程和服务器端进程可以是用于处理用户输入和向用户提供帮助的两个并行工作流。在特定实施例中,客户端进程可以在与用户相关联的客户端系统130上本地执行。相比之下,服务器端进程可以在一个或多个计算系统上远程执行。在特定实施例中,客户端系统130上的助理编排器(assistant orchestrator)可以协调接收用户输入(例如,音频信号)并确定是使用客户端进程还是服务器端进程或两者来响应用户输入。对话仲裁器可以分析来自每个进程的处理结果。对话仲裁器可以基于上述分析指示客户端或服务器端的代理执行与用户输入相关联的任务。执行结果可以进一步作为输出向客户端系统130渲染。通过利用客户端和服务器端进程,助理系统140可以有效地帮助用户优化计算资源的使用,同时保护用户隐私并增强安全性。

[0069] 在特定实施例中,助理系统140可以从与用户相关联的客户端系统130接收用户输入。在特定实施例中,用户输入可以是用户生成的输入,其在单轮中被发送到助理系统140。

用户输入可以是口头的、非口头的或其组合。作为示例而非限制,非口头用户输入可以基于用户的声音、视觉、位置、活动、手势、运动或其组合。如果用户输入基于用户的声音(例如,用户可以对客户端系统130说话),则这种用户输入可以首先由系统音频API 202(应用编程接口)处理。系统音频API 202可以进行回声消除、噪声去除、波束形成和自用户(self-user)声音激活、说话者识别、声音活动检测(VAD)以及任何其他声学技术,以生成可由助理系统140容易地处理的音频数据。在特定实施例中,系统音频API 202可以根据用户输入执行唤醒词(wake-word)检测204。作为示例而非限制,唤醒词可以是“嘿,助理”。如果检测到这样的唤醒词,则可以相应地激活助理系统140。在替代实施例中,用户可以通过视觉信号激活助理系统140,而无需唤醒词。视觉信号可以在能够检测各种视觉信号的低功率传感器(例如,摄像机)处接收。作为示例而非限制,视觉信号可以是由客户端系统130检测到的条形码、QR代码或通用产品代码(UPC)。作为另一个示例而非限制,视觉信号可以是用户对对象的凝视。作为又一个示例而非限制,视觉信号可以是用户手势,例如,用户指向对象。

[0070] 在特定实施例中,来自系统音频API 202的音频数据可以被发送到助理编排器206。助理编排器206可以在客户端系统130上执行。在特定实施例中,助理编排器206可以确定是通过使用客户端进程还是使用服务器端进程来响应用户输入。如图2所示,客户端进程在虚线207下方示出,而服务器端进程在虚线207上方示出。助理编排器206还可以确定通过同时使用客户端进程和服务器端进程来响应用户输入。尽管图2将助理编排器206示出为客户端进程,但是助理编排器206可以是服务器端进程,或者可以是在客户端和服务器端进程之间分离的混合进程。

[0071] 在特定实施例中,在从系统音频API 202生成音频数据之后,服务器端进程可以如下。助理编排器206可以将音频数据发送到远程计算系统,该远程计算系统托管助理系统140的不同模块以响应用户输入。在特定实施例中,音频数据可以在远程自动语音识别(ASR)模块208处接收。ASR模块208可以允许用户口述并使语音转录为书面文本,使文档合成为音频流,或者发出被系统识别为这样的命令。ASR模块208可以使用统计模型来确定对应于由助理系统140作为音频输入接收的给定语音部分的最可能的单词序列。模型可以包括隐马尔可夫模型、神经网络、深度学习模型或其任意组合中的一个或更多个。接收的音频输入可以以特定的采样率(例如,16、44.1或96kHz)被编码成数字数据,并且具有代表每个样本的特定数量的比特(例如,24比特中的8、16比特)。

[0072] 在特定实施例中,ASR模块208可以包括不同的部件。ASR模块208可以包括字形到音素(G2P, grapheme-to-phoneme)模型、发音学习模型、个性化声学模型、个性化语言模型(PLM)或端点模型中的一个或更多个。在特定实施例中,G2P模型可用于确定用户的字形到音素的风格,例如,当特定用户说特定单词时它听起来像什么。个性化声学模型可以是音频信号和语言中语音单元(phonetic unit)的声音之间关系的模型。因此,这种个性化声学模型可以识别用户的声音听起来如何。个性化声学模型可以使用训练数据生成,例如作为音频输入接收的训练语音和对应于该语音的相应语音单元。可以使用特定用户的声音来训练或细化个性化声学模型,以识别该用户的语音。在特定实施例中,个性化语言模型然后可以确定对应于特定音频输入的所识别的语音单元的最可能的短语。个性化语言模型可以是语言中可能出现各种单词序列的概率模型。可以使用个性化语言模型将音频输入中的语音单元的声音与单词序列进行匹配,并且可以将更大的权重分配给更可能是该语言中的短语的

单词序列。然后可以选择具有最高权重的单词序列作为对应于音频输入的文本。在特定实施例中，个性化语言模型还可以用于预测给定上下文，用户最有可能说什么单词。在特定实施例中，端点模型可以检测何时到达话语的终点。

[0073] 在特定实施例中，ASR模块208的输出可以被发送到远程自然语言理解(NLU)模块210。NLU模块210可以执行命名实体解析(NER)。当分析用户输入时，NLU模块210可以另外考虑上下文信息。在特定实施例中，意图和/或槽(slot)可以是NLU模块210的输出。意图可以是语义意图的预定义分类中的元素，其可以指示用户与助理系统140交互的目的。NLU模块210可以将用户输入分类为预定义分类的成员，例如，对于输入“播放贝多芬的第五交响乐”，NLU模块210可以将输入分类为具有意图[IN:播放_音乐]。在特定实施例中，域可以表示交互的社交上下文，例如教育，或者一组意图的命名空间，例如音乐。槽可以是对应于用户输入中的字符串的命名子串，表示基本语义实体。例如，“比萨”的槽可以是[SL:菜肴]。在特定实施例中，有效或预期命名槽的集合可以基于分类的意图。作为示例而非限制，对于意图[IN:播放_音乐]，有效的槽可以是[SL:歌曲_名称]。在特定实施例中，NLU模块210可以另外从社交图、知识图或概念图中的一个或多个提取信息，并且从一个或多个远程数据存储器212检索用户简档。NLU模块210还可以通过确定聚合什么信息、注释用户输入的n-gram(n元语法)、基于聚合的信息用置信度分数对n-gram进行排序、将排序后的n-gram公式化为可以被NLU模块210用于理解用户输入的特征，来处理来自这些不同源的信息。

[0074] 在特定实施例中，NLU模块210可以以个性化和上下文感知的方式从用户输入中识别域、意图或槽中的一个或多个。作为示例而不是作为限制，用户输入可以包括“告诉我如何去咖啡店(show me how to get to the coffee shop)”。NLU模块210可以基于用户的个人信息和相关联的上下文信息来识别用户想要去的特定咖啡店。在特定实施例中，NLU模块210可以包括特定语言词典(lexicon of language)、语法分析器(parser)和语法规则(grammar rule)，以将句子划分成内部表示。NLU模块210还可以包括一个或多个程序，该程序使用语用学(pragmatics)来执行朴素(naive)语义或随机语义分析，以理解用户输入。在特定实施例中，语法分析器可以基于包括多个长短期记忆(LSTM)网络的深度学习架构。作为示例而不是作为限制，语法分析器可以基于循环神经网络语法(RNNG)模型，这是一种类型的循环和递归LSTM算法。关于自然语言理解的更多信息可以在2018年6月18日提交的美国专利申请第16/011062号、2018年7月2日提交的美国专利申请第16/025317号、和2018年7月17日提交的美国专利申请第16/038120号中找到，本文将讨论这些专利申请中的每一个。

[0075] 在特定实施例中，NLU模块210的输出可以被发送到远程推理模块214。推理模块214可以包括对话管理器和实体解析部件。在特定实施例中，对话管理器可以具有复杂的对话逻辑和产品相关的业务逻辑。对话管理器可以管理用户和助理系统140之间的对话状态和会话流。对话管理器可以另外存储用户和助理系统140之间的先前会话。在特定实施例中，对话管理器可以与实体解析部件通信以解析与一个或多个槽相关联的实体，这支持对话管理器推进用户和助理系统140之间的会话流。在特定实施例中，当解析实体时，实体解析部件可以访问社交图、知识图或概念图中的一个或多个。实体可以包括例如独特的用户或概念，每个用户或概念可以具有唯一的标识符(ID)。作为示例而不是作为限制，知识图可以包括多个实体。每个实体可以包括与一个或多个属性值相关联的单个记录。特定

记录可以与唯一的实体标识符相关联。对于实体的一个属性,每个记录可以有不同的值。每个属性值可以与置信度概率相关联。属性值的置信度概率表示该值对于给定属性是准确的概率。每个属性值也可以与语义权重相关联。属性值的语义权重可以表示考虑到所有可用信息,该值在语义上适合给定属性的程度。例如,知识图可以包括电影“火星救援(The Martian)”(2015)的实体,该实体包括已经从多个内容源(例如,Facebook、在线百科全书、电影评论源、媒体数据库和娱乐内容源)提取,然后被去重复(deduped)、解析和融合以生成知识图的单个唯一记录的信息。实体可以与指示电影“火星救援”(2015年)的类型(genre)的空间属性值相关联。关于知识图的更多信息可以在2018年7月27日提交的美国专利申请第16/048049号和2018年7月27日提交的美国专利申请第16/048101号中找到,本文将讨论这些专利申请中的每一个。

[0076] 在特定实施例中,实体解析部件可以检查隐私约束,以保证实体的解析不违反隐私策略。作为示例而非限制,要解析的实体可以是在他/她的隐私设置中指定他/她的身份不应该在在线社交网络上可搜索的另一个用户,因此实体解析部件可能不会响应于请求而返回该用户的标识符。基于从社交图、知识图、概念图 and 用户简档获得的信息,并且服从适用的隐私策略,实体解析部件因此可以以个性化的、上下文感知的和隐私感知的方式解析与用户输入相关联的实体。在特定实施例中,每个解析的实体可以与由社交网络系统160托管的一个或更多个标识符相关联。作为示例而非限制,标识符可以包括对应于特定用户的唯一用户标识符(ID)(例如,唯一用户名或用户ID号码)。在特定实施例中,每个解析的实体也可以与置信度分数相关联。关于解析实体的更多信息可以在2018年7月27日提交的美国专利申请第16/048049号和2018年7月27日提交的美国专利申请第16/048072号中找到,本文将讨论这些专利申请中的每一个。

[0077] 在特定实施例中,对话管理器可以进行对话优化和助理状态跟踪。对话优化是使用数据来理解对话中最可能的分支应该是什么的问题。作为示例而非限制,利用对话优化,助理系统140可能不需要确认用户想要呼叫谁,因为助理系统140具有基于对话优化推断的人很可能是用户想要呼叫的人的高置信度。在特定实施例中,对话管理器可以使用强化学习来进行对话优化。助理状态跟踪旨在跟踪随着用户与世界交互以及助理系统140与用户交互而随时间变化的状态。作为示例而非限制,根据适用的隐私策略,助理状态跟踪可以跟踪用户正在谈论什么、用户与谁在一起、用户在哪里、当前正在进行什么任务以及用户的凝视在哪里等。在特定实施例中,对话管理器可以使用一组操作符来跟踪对话状态。操作符可以包括必要的数据和逻辑来更新对话状态。在处理传入的请求后,每个操作符都可以充当对话状态的增量(delta)。在特定实施例中,对话管理器可以进一步包括对话状态跟踪器和动作选择器。在替代实施例中,对话状态跟踪器可以替换实体解析部件,并解析指代/提及,并跟踪状态。

[0078] 在特定实施例中,推理模块214可以进一步进行错误触发缓解。错误触发缓解的目标是检测助理请求的错误触发(例如,唤醒词),并避免当用户实际上不打算调用助理系统140时产生错误记录。作为示例而非限制,推理模块214可以基于无意义检测器来实现错误触发缓解。如果无意义检测器确定唤醒词在与用户交互的这一点上没有意义,则推理模块214可以确定推断用户打算调用助理系统140可能是不正确的。在特定实施例中,推理模块214的输出可以被发送到远程对话仲裁器216。

[0079] 在特定实施例中,ASR模块208、NLU模块210和推理模块214中的每一个都可以访问远程数据储存器212,该远程数据储存器212包括用户情节记忆,以确定如何更有效地帮助用户。关于情节记忆的更多信息可以在2019年8月27日提交的美国专利申请第16/552559号中找到,该申请在本文讨论。数据储存器212可以另外存储用户的用户简档。用户的用户简档可以包括用户简档数据,该用户简档数据包括与用户相关联的人口统计信息、社交信息和上下文信息。用户简档数据还可以包括用户对通过动态消息、搜索日志、消息平台等上的会话来聚集的多个主题的兴趣和偏好。用户简档的使用可能会受到隐私限制,以确保用户的信息只能用于他/她的利益,而不能与任何其他人共享。关于用户简档的更多信息可以在2018年4月30日提交的美国专利申请第15/967239号中找到,该申请在此讨论。

[0080] 在特定实施例中,与前述涉及ASR模块208、NLU模块210和推理模块214的服务器端进程并行,客户端进程可以如下。在特定实施例中,助理编排器206的输出可以被发送到客户端系统130上的本地ASR模块216。ASR模块216可以包括个性化语言模型(PLM)、G2P模型和端点模型。由于客户端系统130的有限计算能力,助理系统140可以在客户端进程期间在运行时优化个性化语言模型。作为示例而非限制,助理系统140可以为用户可能谈论的多个可能主题预先计算多个个性化语言模型。当用户请求帮助时,助理系统140然后可以快速交换这些预先计算的个性化语言模型,使得个性化语言模型可以由助理系统140在运行时基于用户活动进行本地优化。结果,助理系统140可以具有在有效地确定用户可能在谈论什么的同时节省计算资源的技术优势。在特定实施例中,助理系统140还可以在运行时快速重新学习用户发音。

[0081] 在特定实施例中,ASR模块216的输出可以被发送到本地NLU模块218。在特定实施例中,与服务器端支持的远程NLU模块210相比,本文的NLU模块218可以更紧凑。当ASR模块216和NLU模块218处理用户输入时,它们可以访问本地助理存储器220。出于保护用户隐私的目的,本地助理存储器220可以不同于存储在数据储存器212上的用户记忆。在特定实施例中,本地助理存储器220可以经由网络110与存储在数据储存器212上的用户记忆同步。作为示例而非限制,本地助理存储器220可以将用户客户端系统130上的日历与和用户相关联的服务器端日历同步。在特定实施例中,本地助理存储器220中的任何安全数据仅可由在客户端系统130上本地执行的助理系统140的模块访问。

[0082] 在特定实施例中,NLU模块218的输出可以被发送到本地推理模块222。推理模块222可以包括对话管理器和实体解析部件。由于有限的计算能力,推理模块222可以进行基于特别为客户端系统130定制的学习算法的设备上学习。作为示例而非限制,推理模块222可以使用联邦学习。联邦学习是分布式机器学习方法的一个特定类别,它使用驻留在移动电话等终端设备上的分散数据来训练机器学习模型。在特定实施例中,推理模块222可以使用特定的联邦学习模型,即联邦用户表示学习,来将现有的神经网络个性化技术扩展到联邦学习。联邦用户表示学习可以通过学习特定于任务的用户表示(即嵌入)或通过个性化模型权重来个性化联邦学习中的模型。联邦用户表示学习简单、可扩展、隐私保护和资源高效。联邦用户表示学习可以将模型参数分为联邦参数和私有参数。私有参数,例如私有用户嵌入,可以在客户端系统130上本地训练,而不是传送到远程服务器或在远程服务器上平均。相比之下,联邦参数可以在服务器上远程训练。在特定实施例中,推理模块222可以使用另一个特定的联邦学习模型,即主动联邦学习,以将在远程服务器上训练的全局模型传输

到客户端系统130,并在这些客户端系统130上本地计算梯度。主动联邦学习可以使推理模块最小化与下载模型和上传梯度相关的传输成本。对于主动联邦学习,在每一轮中,客户端系统不是均匀地随机选择的,而是以当前模型和客户端系统上的数据为条件的概率选择的,以最大化效率。在特定实施例中,推理模块222可以使用另一个特定的联邦学习模型,即联邦Adam。传统的联邦学习模型可以使用随机梯度下降(SGD)优化器。相比之下,联邦Adam模型可能使用基于矩的优化器。联邦Adam模型可以使用平均模型来计算近似梯度,而不是像传统工作那样直接使用平均模型。这些梯度然后可以被馈送到联邦Adam模型中,该模型可以去除随机梯度的噪声并使用每个参数的自适应学习速率。联邦学习产生的梯度可能比随机梯度下降噪声更多(因为数据可能不是独立且同分布的),因此联邦Adam模型可能有助于更好地处理噪声。联邦Adam模型可以使用梯度来采取更聪明的步骤来最小化目标函数。实验表明,在基准上的传统联邦学习在ROC(接收机工作特性)曲线上有1.6%的下降,而联邦Adam模型只有0.4%的下降。此外,联邦Adam模型没有增加通信或设备上的计算。在特定实施例中,推理模块222还可以执行错误触发缓解。当用户的语音输入包括受隐私约束的数据时,这种错误触发缓解可以帮助检测客户端系统130上的错误激活请求,例如唤醒词。作为示例而非限制,当用户处于语音呼叫中时,用户的会话是私密的,并且基于这种会话的错误触发检测只能在用户的客户端系统130上本地发生。

[0083] 在特定实施例中,助理系统140可以包括本地上下文引擎224。上下文引擎224可以处理所有其他可用信号,以向推理模块222提供更多信息提示。作为示例而非限制,上下文引擎224可以具有与人相关的信息、来自客户端系统130传感器(例如,麦克风、摄像机)的感测数据(其通过计算机视觉技术进一步分析)、几何构造、活动数据、惯性数据(例如,由VR头戴装置收集的数据)、位置等。在特定实施例中,计算机视觉技术可以包括人体骨骼重建、面部检测、面部识别、手跟踪、眼睛跟踪等。在特定实施例中,几何构造可以包括使用由客户端系统130收集的数据来构造用户周围的对象。作为示例而非限制,用户可戴着AR眼镜,并且几何构造可以旨在确定地板在哪里、墙壁在哪里、用户的手在哪里等。在特定实施例中,惯性数据可以是与线性和角运动相关联的数据。作为示例而非限制,可以通过测量用户身体部位如何移动的AR眼镜来捕获惯性数据。

[0084] 在特定实施例中,本地推理模块222的输出可以被发送到对话仲裁器216。对话仲裁器216可以在三种场景下不同地运行。在第一种场景下,助理编排器206确定使用服务器端进程,为此,对话仲裁器216可以将推理模块214的输出传输到远程动作执行模块226。在第二种场景下,助理编排器206确定使用服务器端进程和客户端进程,为此,对话仲裁器216可以聚集来自两个进程的两个推理模块(即,远程推理模块214和本地推理模块222)的输出并分析它们。作为示例而非限制,对话仲裁器216可以执行排序并选择最佳推理结果来响应用户输入。在特定实施例中,对话仲裁器216还可以基于分析来确定是使用服务器端还是客户端上的代理来执行相关任务。在第三种场景下,助理编排器206确定使用客户端进程,并且对话仲裁器216需要评估本地推理模块222的输出,以确定客户端进程是否能够完成处理用户输入的任务。

[0085] 在特定实施例中,对于上述第一和第二场景,对话仲裁器216可以确定服务器端的代理对于响应用户输入执行任务是必要的。因此,对话仲裁器216可以向动作执行模块226发送关于用户输入的必要信息。动作执行模块226可以调用一个或更多个代理来执行任务。

在替代实施例中,对话管理器的动作选择器可以确定要执行的动作,并相应地指示动作执行模块226。在特定实施例中,代理可以是充当一个域的多个内容提供者之间的中介者(broker)的一种实现。内容提供者可以是负责执行与意图相关联的动作或完成与意图相关联的任务的实体。在特定实施例中,代理可以包括第一方代理和第三方代理。在特定实施例中,第一方代理可以包括可由助理系统140访问和控制的内部代理(例如,与在线社交网络提供的服务相关联的代理(Messenger、Instagram、Whatsapp))。在特定实施例中,第三方代理可以包括助理系统140无法控制的外部代理(例如,第三方在线音乐应用代理、门票销售代理)。第一方代理可以与第一方提供者关联,该第一方提供者提供由社交网络系统160托管的内容对象和/或服务。第三方代理可以与第三方提供者相关联,该第三方提供者提供由第三方系统170托管的内容对象和/或服务。在特定实施例中,第一方代理或第三方代理中的每一个可以被指定用于特定域。作为示例而非限制,该域可以包括天气、交通、音乐等。在特定实施例中,助理系统140可以协同使用多个代理来响应用户输入。作为示例而不是作为限制,用户输入可以包括“指引我到我的下一个会议(direct me to my next meeting)”。助理系统140可以使用日历代理来检索下一个会议的位置。助理系统140然后可以使用导航代理来指引用户到下一个会议。

[0086] 在特定实施例中,对于上述第二和第三场景,对话仲裁器216可以确定客户端上的代理能够响应于用户输入执行任务,但是需要附加信息(例如,响应模板),或者任务只能由服务器端上的代理来处理。如果对话仲裁器216确定任务只能由服务器端的代理处理,则对话仲裁器216可以向动作执行模块226发送关于用户输入的必要信息。如果对话仲裁器216确定客户端上的代理能够执行任务,但是需要响应模板,则对话仲裁器216可以向远程响应模板生成模块228发送关于用户输入的必要信息。响应模板生成模块228的输出还可以被发送到在客户端系统130上执行的本地动作执行模块230。

[0087] 在特定实施例中,动作执行模块230可以调用本地代理来执行任务。与服务器端的代理相比,客户端系统130上的本地代理能够执行更简单的任务。作为示例而不是作为限制,多个特定于设备的实现(例如,对客户端系统130的或客户端系统130上的消息传递应用的实时调用)可以由单个代理在内部处理。可替代地,这些特定于设备的实现可以由与多个域相关联的多个代理来处理。在特定实施例中,动作执行模块230可以另外执行一组通用可执行对话动作。这组可执行对话动作可以与代理、用户和助理系统140本身交互。这些对话动作可以包括用于槽请求、确认、歧义消除、代理执行等的对话动作。对话动作可以独立于动作选择器或对话策略的底层实现。基于树的策略和基于模型的策略都可以生成相同的基本对话动作,回调函数隐藏了任何特定于动作选择器的实现细节。

[0088] 在特定实施例中,来自服务器端的远程动作执行模块226的输出可以被发送到远程响应执行模块232。在特定实施例中,动作执行模块226可以向对话仲裁器216发回更多信息。响应执行模块232可以基于远程会话理解(CU)编写器。在特定实施例中,来自动作执行模块226的输出可以被公式化为 $\langle k, c, u, d \rangle$ 元组,其中k指示知识源、c指示通信目标、u指示用户模型以及d指示话语模型。在特定实施例中,CU编写器可以包括自然语言生成器(NLG)和用户界面(UI)有效载荷生成器。自然语言生成器可以使用不同的语言模型和/或语言模板基于动作执行模块226的输出来生成通信内容。在特定实施例中,通信内容的生成可以是特定于应用的,并且对于每个用户也是个性化的。CU编写器还可以使用UI有效载荷生成器

来确定所生成的通信内容的模态。在特定实施例中,NLG可以包括内容确定部件、句子规划器和表层实现(surface realization)部件。内容确定部件可以基于知识源、通信目标和用户的期望来确定通信内容。作为示例而不是作为限制,确定可以基于描述逻辑。描述逻辑可以包括例如三个基本观念(notion),它们是个体(表示域中的对象)、概念(描述个体集合)和角色(表示个体或概念之间的二元关系)。描述逻辑可以由一组构造器来表征,该构造器允许自然语言生成器从原子(atomic)概念/角色构建复杂概念/角色。在特定实施例中,内容确定部件可以执行以下任务来确定通信内容。第一任务可以包括翻译任务,其中对自然语言生成器的输入可以被翻译成概念。第二任务可以包括选择任务,其中可以基于用户模型从翻译任务产生的概念中选择相关概念。第三任务可以包括验证任务,其中可以验证所选概念的一致性。第四任务可以包括实例化任务,其中被验证的概念可以被实例化为可由自然语言生成器处理的可执行文件。句子规划器可以确定通信内容的组织,使其为人所理解。表层实现部件可以确定要使用的特定单词、句子的顺序以及通信内容的风格。UI有效载荷生成器可以确定要呈现给用户的通信内容的优选模态。在特定实施例中,CU编写器可以检查与用户相关联的隐私约束,以确保通信内容的生成遵循隐私策略。关于自然语言生成的更多信息可以在2018年4月30日提交的美国专利申请第15/967279号和2018年4月30日提交的美国专利申请第15/966455号中找到,本文将讨论这些专利申请中的每一个。

[0089] 在特定实施例中,来自客户端系统130上的本地动作执行模块230的输出可以被发送到本地响应执行模块234。响应执行模块234可以基于本地会话理解(CU)编写器。CU编写器可以包括自然语言生成(NLG)模块。由于客户端系统130的计算能力可能有限,出于计算效率的考虑,NLG模块可以是简单的。因为NLG模块可以是简单的,所以响应执行模块234的输出可以被发送到本地响应扩展模块236。响应扩展模块236可以进一步扩展响应执行模块234的结果,以使响应更加自然并包含更丰富的语义信息。

[0090] 在特定实施例中,如果用户输入基于音频信号,则服务器端的响应执行模块232的输出可以被发送到远程文本到语音(TTS)模块238。类似地,客户端上的响应扩展模块236的输出可以被发送到本地TTS模块240。两个TTS模块都可以将响应转换为音频信号。在特定实施例中,来自响应执行模块232、响应扩展模块236或两端的TTS模块的输出可以最终被发送到本地渲染输出模块242。渲染输出模块242可以生成适合于客户端系统130的响应。作为示例而非限制,响应执行模块232或响应扩展模块236的输出可以包括自然语言串、语音、带有参数的动作或可以在VR头戴装置或AR智能眼镜中显示的渲染图像或视频中的一个或多个。结果,渲染输出模块242可以基于CU编写器的输出来确定要执行什么任务,以适当地渲染响应,用于在VR头戴装置或AR智能眼镜上显示。例如,响应可以是基于视觉的模态(例如,图像或视频剪辑),其可以通过VR头戴装置或AR智能眼镜显示。作为另一个例子,响应可以是用户可以通过VR头戴装置或AR智能眼镜播放的音频信号。作为又一示例,响应可以是增强现实数据,其可以由VR头戴装置或AR智能眼镜渲染以增强用户体验。

[0091] 在特定实施例中,助理系统140可以具有各种能力,包括音频认知、视觉认知、信号智能、推理和记忆。在特定实施例中,音频识别的能力可以使助理系统140能够理解与不同语言的各种域相关联的用户输入,理解会话并能够对其进行汇总,对复杂的命令执行设备上的音频认知,通过语音识别用户,从会话中提取主题并自动标记会话的部分,实现没有唤醒词的音频交互,从环境噪声和会话中过滤并放大用户语音,理解用户正在与哪个客户端

系统130(如果多个客户端系统130在附近)通话。

[0092] 在特定实施例中,视觉认知的能力可以使助理系统140能够执行面部检测和跟踪,识别用户,以不同的角度识别主要大都市地区的大多数感兴趣的人,通过现有的机器学习模型和一次性学习的组合识别世界上大多数感兴趣的对象,识别感兴趣的时刻并自动捕获它,实现对跨不同的时间段的多个视觉帧的语义理解,为人、地点、对象识别的附加能力提供平台支持,识别全套设置和包括个性化位置的微位置(micro-location)、识别复杂活动、识别复杂手势以控制客户端系统130、处理来自以自我为中心的摄像机的图像/视频(例如,具有运动、捕获角度、分辨率等)、实现关于具有较低分辨率的图像的相似级别的准确度和速度,进行人、地点和对象的一次性配准和识别,并在客户端系统130上执行视觉识别。

[0093] 在特定实施例中,助理系统140可以利用计算机视觉技术来实现视觉认知。除了计算机视觉技术之外,助理系统140可以探索能够补充这些技术的选项,以扩大对象的识别。在特定实施例中,助理系统140可以使用补充信号,例如对象标签的光学字符识别(OCR)、用于地点识别的GPS信号、来自用户客户端系统130的信号来识别用户。在特定实施例中,助理系统140可以执行一般场景识别(家庭、工作、公共空间等)来为用户设置上下文并缩小计算机视觉搜索空间以识别最可能的对象或人。在特定实施例中,助理系统140可以指导用户训练助理系统140。例如,众包可用于让用户标记并帮助助理系统140随时间识别更多的对象。作为另一个例子,当使用助理系统140时,用户可以注册他们的个人对象作为初始设置的一部分。助理系统140还可以允许用户为他们交互的对象提供正/负信号,以便训练和改进用于他们的个性化模型。

[0094] 在特定实施例中,信号智能的能力可以使助理系统140能够确定用户位置、理解日期/时间、确定家庭位置、理解用户的日历和未来期望的位置、集成更丰富的声音理解以仅通过声音来识别设置/上下文、在运行时建立信号智能模型,该模型可以针对用户的个人例程进行个性化。

[0095] 在特定实施例中,推理能力可以使助理系统140能够具有以下能力:在未来的任何时间点拾取任何先前的会话线程,合成所有信号以理解微观和个性化的上下文,从用户的历史行为中学习交互模式和偏好并准确地建议他们可能重视的交互,基于微观上下文理解生成高度预测性的主动建议,理解用户可能希望在一天的什么时间看到什么内容,理解场景中的变化以及这可能如何影响用户期望的内容。

[0096] 在特定实施例中,记忆的能力可以使助理系统140能够记住用户先前调用或交互了哪些社交关连,随意写入存储器和查询存储器(即,公开听写和自动标签),基于先前的交互和长期学习提取更丰富的偏好,记住用户的生活史,从以自我为中心的数据流和自动目录中提取丰富的信息,并以结构化形式写入存储器以形成丰富的短期、情节和长期记忆。

[0097] 图3示出了助理系统140的服务器端进程的示例图。在特定实施例中,服务器助理服务模块301可以在接收到用户请求时访问请求管理器302。在替代实施例中,如果用户请求基于音频信号,则用户请求可以首先由远程ASR模块208处理。在特定实施例中,请求管理器302可以包括上下文提取器303和会话理解对象生成器(CU对象生成器)304。上下文提取器303可以提取与用户请求相关联的上下文信息。上下文提取器303还可以基于在客户端系统130上执行的助理应用136来更新上下文信息。作为示例而不是作为限制,上下文信息的更新可以包括在客户端系统130上显示内容项目。作为另一个示例而不是作为限制,上下文

信息的更新可以包括客户端系统130上是否设置了闹钟。作为另一个示例而不是作为限制，上下文信息的更新可以包括客户端系统130上是否在播放歌曲。CU对象生成器304可以生成与用户请求相关的特定内容对象。内容对象可以包括与用户请求相关联的对话会话数据和特征，该对话会话数据和特征可以与助理系统140的所有模块共享。在特定实施例中，请求管理器302可以将上下文信息和生成的内容对象存储在数据存储器212中，数据存储器是在助理系统140中实现的特定数据存储器。

[0098] 在特定实施例中，请求管理器302可以将生成的内容对象发送到远程的NLU模块210。NLU模块210可以执行多个步骤来处理内容对象。在步骤305，NLU模块210可以生成内容对象的白名单(whitelist)。在特定实施例中，白名单可以包括匹配用户请求的解释数据。在步骤306，NLU模块210可以基于白名单执行特征化。在步骤307，NLU模块210可以基于从特征化产生的特征对用户请求执行域分类/选择，以将用户请求分类到预定义的域中。还可以基于两个相关的过程进一步处理域分类/选择结果。在步骤308a，NLU模块210可以使用意图分类器来处理域分类/选择结果。意图分类器可以确定与用户请求相关联的用户意图。在特定实施例中，每个域可以有一个意图分类器来确定给定域中最可能的意图。作为示例而不是作为限制，意图分类器可以基于机器学习模型，该模型可以将域分类/选择结果作为输入，并计算该输入与特定预定义意图相关联的概率。在步骤308b，NLU模块210可以使用元意图(meta-intent)分类器来处理域分类/选择结果。元意图分类器可以确定描述用户意图的类别。在特定实施例中，多个域共有的意图可以由元意图分类器处理。作为示例而不是作为限制，元意图分类器可以基于机器学习模型，该模型可以将域分类/选择结果作为输入，并计算输入与特定预定义元意图相关联的概率。在步骤309a，NLU模块210可以使用槽标记器(slot tagger)来注释与用户请求相关联的一个或多个槽。在特定实施例中，槽标记器可以为用户请求的n-grams注释一个或多个槽。在步骤309b，NLU模块210可以使用元槽标记器为来自元意图分类器的分类结果注释一个或多个槽。在特定实施例中，元槽标记器可以标记通用槽，例如对项目(例如，第一个项目)的指代、槽的类型、槽的值等。作为示例而非限制，用户请求可以包括“将我账户中的500美元兑换成日元”。意图分类器可以将用户请求作为输入，并将其公式化为向量。意图分类器然后可以基于表示用户请求的向量和表示不同预定义意图的向量之间的向量比较来计算用户请求与不同预定义意图相关联的概率。以类似的方式，槽标记器可以将用户请求作为输入，并将每个单词公式化为向量。意图分类器然后可以基于表示单词的向量和表示不同预定义槽的向量之间的向量比较来计算每个单词与不同预定义槽相关联的概率。用户的意图可以被分类为“兑换钱(changing money)”。用户请求的槽可以包括“500”、“美元(dollars)”、“账户(account)”和“日元(Japanese yen)”。用户的元意图可以被分类为“金融服务(financial service)”。元槽(meta slot)可以包括“金融(finance)”。

[0099] 在特定实施例中，NLU模块210可以包括语义信息聚合器310。语义信息聚合器310可以通过提供语义信息来帮助NLU模块210改进内容对象的域分类/选择。在特定实施例中，语义信息聚合器310可以以下述方式聚合语义信息。语义信息聚合器310可以首先从用户上下文引擎315检索信息。在特定实施例中，用户上下文引擎315可以包括离线聚合器和在线推断服务。离线聚合器可以处理从先前时间窗口收集的与用户相关联的多个数据。作为示例而非限制，数据可以包括在预定的时间范围内(例如，从先前的90天窗口)收集的动态消

息帖子/评论、与动态消息帖子/评论的交互、搜索历史等。处理结果可以作为用户简档的一部分存储在用户上下文引擎315中。在线推断服务可以分析由助理系统140在当前时间接收的与用户相关联的会话数据。分析结果也可以作为用户简档的一部分存储在用户上下文引擎315中。在特定实施例中，离线聚合器和在线推断服务都可以从多个数据中提取个性化特征。提取的个性化特征可以被助理系统140的其他模块使用，以更好地理解用户输入。在特定实施例中，语义信息聚合器310然后可以在以下步骤中处理从用户上下文引擎315检索的信息，即用户简档。在步骤311，语义信息聚合器310可以基于自然语言处理(NLP)来处理从用户上下文引擎315检索的信息。在特定实施例中，语义信息聚合器310可以：通过文本标准化来切分(tokenize)文本，从文本中提取句法(syntax)特征，并基于NLP从文本中提取语义特征。语义信息聚合器310另外可以从上下文信息中提取特征，该上下文信息是从用户和助理系统140之间的对话历史访问的。语义信息聚合器310还可以基于上下文信息进行全局单词嵌入、特定于域的嵌入和/或动态嵌入。在步骤312，处理结果可以由实体标记器用实体来注释。在步骤313，基于注释，语义信息聚合器310可以为检索到的信息生成字典。在特定实施例中，字典可以包括可以离线动态更新的全局字典特征。在步骤314，语义信息聚合器310可以对由实体标记器标记的实体进行排序。在特定实施例中，语义信息聚合器310可以与包括社交图、知识图或概念图中的一个或多个的不同图320通信，以从用户上下文引擎315提取与检索到的信息相关的本体数据。在特定实施例中，语义信息聚合器310可以聚合用户简档、排序后的实体和来自图320的信息。语义信息聚合器310然后将聚合的信息提供给NLU模块210，以促进域分类/选择。

[0100] 在特定实施例中，NLU模块210的输出可以被发送到远程推理模块214。推理模块214可以包括共指(co-reference)部件325、实体解析部件330和对话管理器335。NLU模块210的输出可以首先在共指部件325处被接收，以解释与用户请求相关联的内容对象的指代。在特定实施例中，共指部件325可用于识别用户请求所指的项目。共指部件325可以包括指代创建326和指代消解(reference resolution) 327。在特定实施例中，指代创建326可以由NLU模块210确定的实体创建指代。指代消解327可以准确地消解这些指代。作为示例而非限制，用户请求可以包括“给我找到最近的杂货店并指引我到那里”。共指部件325可以将“那里”解释为“最近的杂货店”。在特定实施例中，共指部件325可以在必要时访问用户上下文引擎315和对话引擎335，以用提高的准确性来解释指代。

[0101] 在特定实施例中，所识别的域、意图、元意图、槽和元槽以及所消解的指代可以被发送到实体解析部件330以解析相关实体。实体解析部件330可以执行通用和特定于域的实体解析。在特定实施例中，实体解析部件330可以包括域实体解析331和通用实体解析332。域实体解析331可以通过将槽和元槽归类到不同的域来解析实体。在特定实施例中，可以基于从图320中提取的本体数据来解析实体。本体数据可以包括不同槽/元槽和域之间的结构关系。本体也可以包括槽/元槽可以如何在较高级别包括域的层次结构内被分组、相关，并根据相似性和差异细分的信息。通用实体解析332可以通过将槽和元槽归类到不同的通用主题来解析实体。在特定实施例中，解析也可以基于从图320中提取的本体数据。本体数据可以包括不同槽/元槽和通用主题之间的结构关系。本体也可以包括槽/元槽可以如何在较高级别包括主题的层次结构内被分组、相关，并根据相似性和差异细分的信息。作为示例而非限制，响应于对特定品牌电动汽车的优点的查询的输入，通用实体解析332可以将电动汽

车的指代品牌解析为车辆,并且域实体解析331可以将电动汽车的指代品牌解析为电动汽车。

[0102] 在特定实施例中,实体解析部件330的输出可以被发送到对话管理器335,以推进与用户的会话流。对话管理器335可以是异步状态机,其重复更新状态并基于新状态选择动作。对话管理器335可以包括对话意图解析336和对话状态跟踪器337。在特定实施例中,对话管理器335可以执行所选择的动作,然后再次调用对话状态跟踪器337,直到所选择的动作需要用户响应,或者不再有动作要执行。所选择的每个动作可取决于先前动作的执行结果。在特定实施例中,对话意图解析336可以基于用户和助理系统140之间的对话历史来解析与当前对话会话相关联的用户意图。对话意图解析336可以将NLU模块210确定的意图映射到不同的对话意图。对话意图解析336还可以基于来自NLU模块210、实体解析部件330的信号以及用户和助理系统140之间的对话历史来对对话意图进行排序。在特定实施例中,对话状态跟踪器337可以是无附带后果部件,并且生成建议对话状态更新的对话状态更新操作符的n个最佳候选,而不是直接改变对话状态。对话状态跟踪器337可以包括意图解析器,其包含基于对话状态处理不同类型的NLU意图并生成操作符的逻辑。在特定实施例中,逻辑可以由意图处理器来组织,该意图处理器例如为当助理系统140请求歧义消除时处理意图的歧义消除意图处理器、包括处理确认的逻辑的确认意图处理器等。意图解析器可以将轮意图(turn intent)与对话状态组合在一起,以生成与用户会话的上下文更新。然后,槽解析部件可以用包含知识图和域代理的解析提供者递归地解析更新操作符中的槽。在特定实施例中,对话状态跟踪器337可以更新/排序当前对话会话的对话状态。作为示例而非限制,如果对话会话结束,对话状态跟踪器337可以将对话状态更新为“完成”。作为另一示例而非限制,对话状态跟踪器337可以基于与对话状态相关联的优先级来对对话状态进行排序。

[0103] 在特定实施例中,推理模块214可以分别与远程动作执行模块226和对话仲裁器216通信。在特定实施例中,推理模块214的对话管理器335可以与动作执行模块226的任务完成部件340就对话意图和相关联的内容对象进行通信。在特定实施例中,任务完成模块340可以针对不同的对话意图对不同的对话假设进行排序。任务完成模块340可以包括动作选择器341。在替代实施例中,动作选择器341可以包括在对话管理器335中。在特定实施例中,对话管理器335还可以对照包含在对话仲裁器216中的关于对话状态的对话策略345进行检查。在特定实施例中,对话策略345可以包括描述代理350的动作执行计划的数据结构。对话策略345可以包括通用策略346和任务策略347。在特定实施例中,通用策略346可以用于不特定于单个任务的动作。通用策略346可以包括处理低置信度意图、内部错误、不可接受的用户重试响应、基于ASR或NLU置信度分数跳过或插入确认等。通用策略346还可以包括对来自对话状态跟踪器337输出的对话状态更新候选进行排序并挑选要更新的一个(例如挑选排序最高的任务意图)的逻辑。在特定实施例中,助理系统140可以具有用于通用策略346的特定接口,其允许将分散的跨域策略/业务规则,尤其是在对话状态跟踪器337中找到那些,合并到动作选择器341的功能中。通用策略346的接口还可以允许创作可以绑定到特定情况或客户端的独立子策略单元,例如可以基于客户端、情况等容易地打开或关闭的策略功能。通用策略346的接口还可以允许提供具有退避的策略分层,即多个策略单元,处理特定情况的高度专业化的策略单元由在更广泛的环境中适用的更通用的策略346来支持。在这种情况下,通用策略346可替代地包括意图或任务特定的策略。在特定实施例中,任

务策略347可以包括基于任务和当前状态的动作选择器341的逻辑。在特定实施例中,可以有以下四种类型的任务策略347:1)手工创建的基于树的对话计划;2)直接实现生成动作的接口的编码策略;3)配置器指定的槽填充任务;以及4)从数据中学习的基于机器学习模型的策略。在特定实施例中,助理系统140可以用基于规则的逻辑引导新域,并且稍后用机器学习模型来细化任务策略347。在特定实施例中,对话策略345可以是基于树的策略,其是预先构建的对话计划。基于当前对话状态,对话策略345可以选择节点来执行和生成相应的动作。作为示例而非限制,基于树的策略可以包括主题分组节点和对话动作(叶)节点。

[0104] 在特定实施例中,动作选择器341可以采用对话状态的候选操作符,并咨询对话策略345来决定应该执行什么动作。助理系统140可以使用分层对话策略,其中通用策略346处理跨域业务逻辑,任务策略347处理任务/域特定逻辑。在特定实施例中,通用策略346可以从候选操作符中选择一个操作符来更新对话状态,随后通过任务策略347选择面向用户的动作。一旦任务在对话状态中是活动的,可以咨询相应的任务策略347来选择正确的动作。在特定实施例中,对话状态跟踪器337和动作选择器341都可以不改变对话状态,直到所选择的动作被执行。这可以允许助理系统140执行对话状态跟踪器337和动作选择器341,用于处理推测的ASR结果,并利用模拟运行(dry runs)进行n最佳排序。在特定实施例中,动作选择器341可以将对话状态更新操作符作为输入的一部分来选择对话动作。对话动作的执行可以生成一组期望,以指示对话状态跟踪器337处理未来轮。在特定实施例中,当处理来自下一轮的用户输入时,期望可用于向对话状态跟踪器337提供上下文。作为示例而非限制,槽请求对话动作可能期望证明所请求的槽的值。

[0105] 在特定实施例中,对话管理器335可以支持槽提及的多轮组成解析。对于来自NLU 210的组成分析,解析器可以递归地解析嵌套槽。对话管理器335可以另外支持嵌套槽的歧义消除。作为示例而非限制,用户请求可以是“提醒我给Alex打电话”。在创建可动作的待办事项提醒实体之前,解析器可能需要知道要给哪个Alex打电话。当特定槽需要进一步的用户澄清时,解析器可以暂停解析并设置解析状态。通用策略346可以检查解析状态,并为用户澄清创建相应的对话动作。在对话状态跟踪器337中,基于用户请求和最后的对话动作,对话管理器可以更新嵌套槽。这种能力可以允许助理系统140与用户交互,不仅收集缺失的槽值,而且减少更复杂/模糊的话语的模糊性以完成任务。在特定实施例中,对话管理器可以进一步支持请求嵌套意图和多意图用户请求中的缺失槽(例如,“拍摄这张照片并将其发送给爸爸”)。在特定实施例中,对话管理器335可以支持机器学习模型,以获得更健壮的对话体验。作为示例而非限制,对话状态跟踪器337可以使用基于神经网络的模型(或任何其他合适的机器学习模型)来对任务假设上的信念进行建模。作为另一个示例而非限制,对于动作选择器341,最高优先级策略单元可以包括白名单/黑名单覆盖,这可能必须通过设计来发生;中等优先级单元可以包括为动作选择设计的机器学习模型;并且当机器学习模型选择不处理情况时,较低优先级单元可以包括基于规则的后退。在特定实施例中,基于机器学习模型的通用策略单元可以帮助助理系统140减少冗余的歧义消除或确认步骤,从而减少执行用户请求的轮数。

[0106] 在特定实施例中,动作执行模块226可以调用不同的代理350来执行任务。代理350可以在注册的内容提供者中进行选择来完成该动作。数据结构可以由对话管理器335基于意图和与该意图相关联的一个或更多个槽来构造。对话策略345还可以包括通过逻辑操作

符互相关的多个目标。在特定实施例中，目标可以是对话策略的一部分的输出结果，并且它可以由对话管理器335构造。目标可以由具有一个或更多个命名参量的标识符（例如，字符串）来表示，该一个或更多个命名参量将目标参数化。作为示例而不是作为限制，目标及其关联的目标参量可以表示为{确认_艺术家, 参量: {艺术家: “Madonna”}}。在特定实施例中，对话策略可以基于树形结构表示，其中目标被映射到树叶。在特定实施例中，对话管理器335可以执行对话策略345来确定要执行的下一个动作。对话策略345可以包括通用策略346和特定于域的策略347，它们都可以指导如何基于对话状态选择下一个系统动作。在特定实施例中，动作执行模块226的任务完成部件340可以与包含在对话仲裁器216中的对话策略345通信，以获得下一个系统动作的指导。在特定实施例中，动作选择部件341因此可以基于对话意图、相关联的内容对象和来自对话策略345的指导来选择动作。

[0107] 在特定实施例中，动作执行模块226的输出可以被发送到远程响应执行模块232。具体地，动作执行模块226的任务完成部件340的输出可以被发送到响应执行模块226的CU编写器355。在替代实施例中，所选择的动作可能需要一个或更多个代理350参与。因此，任务完成模块340可以向代理350通知所选择的动作。同时，对话管理器335可以接收更新对话状态的指令。作为示例而不是作为限制，更新可以包括等待代理350的响应。在特定实施例中，CU编写器355可以基于任务完成模块340的输出，使用自然语言生成器(NLG) 356为用户生成通信内容。在特定实施例中，NLG 356可以使用不同的语言模型和/或语言模板来生成自然语言输出。自然语言输出的生成可以是特定于应用的。自然语言输出的生成也可以针对每个用户进行个性化。CU编写器355还可以使用UI有效载荷生成器357来确定所生成的通信内容的模态。由于所生成的通信内容可以被认为是用户对请求的响应，所以CU编写器355可以另外使用响应排序器358来对所生成的通信内容进行排序。作为示例而不是作为限制，排序可以指示响应的优先级。

[0108] 在特定实施例中，响应执行模块232可以基于CU编写器355的输出来执行不同的任务。这些任务可以包括写入（即，存储/更新）从数据储存器212检索的对话状态361并生成响应362。在特定实施例中，CU编写器355的输出可以包括自然语言串、语音、带有参数的动作、或者可以在VR头戴装置或AR智能眼镜中显示的渲染图像或视频中的一个或更多个。结果，响应执行模块232可以基于CU编写器355的输出来确定要执行什么任务。在特定实施例中，生成的响应和通信内容可以由响应执行模块232发送到本地渲染输出模块242。在替代实施例中，如果所确定的通信内容的模态是音频，则CU编写器355的输出可以另外发送到远程TTS模块238。由TTS模块238生成的语音和由响应执行模块232生成的响应然后可以被发送到渲染输出模块242。

[0109] 图4示出了由助理系统140处理用户输入的示例流程图。作为示例而非限制，用户输入可以基于音频信号。在特定实施例中，客户端系统130的麦克风阵列402可以接收音频信号（例如，语音）。音频信号可以以音频帧的格式传输到处理环路(loop) 404。在特定实施例中，处理环路404可以发送音频帧用于声音活动检测(VAD) 406和声音唤醒(WoV) 检测408。检测结果可以返回到处理环路404。如果WoV检测408指示用户想要调用助理系统140，则音频帧连同VAD 406结果可以被发送到编码单元410以生成编码的音频数据。在编码之后，出于隐私和安全目的，编码的音频数据可以被发送到加密单元412，随后是链接单元414和解密单元416。解密之后，音频数据可以被发送到麦克风驱动器418，麦克风驱动器418可以进

一步将音频数据传输到音频服务模块420。在替代实施例中,可以在与客户端系统130配对的无线设备(例如,蓝牙设备)处接收用户输入。相应地,音频数据可以从无线设备驱动器422(例如,蓝牙驱动器)发送到音频服务模块420。在特定实施例中,音频服务模块420可以确定用户输入可以由在客户端系统130上执行的应用来实现。因此,音频服务模块420可以将用户输入发送到实时通信(RTC)模块424。RTC模块424可以向视频或音频通信系统(例如VOIP或视频呼叫)传送音频分组。RTC模块424可以调用相关应用(App)426来执行与用户输入相关的任务。

[0110] 在特定实施例中,音频服务模块420可以确定用户正在请求需要助理系统140响应的帮助。因此,音频服务模块420可以通知客户端助理服务模块426。在特定实施例中,客户端助理服务模块426可以与助理编排器206通信。助理编排器206可以确定是使用客户端进程还是服务器端进程来响应用户输入。在特定实施例中,助理编排器206可以确定使用客户端进程,并向客户端助理服务模块426通知这样的决定。因此,客户端助理服务模块426可以调用相关模块来响应用户输入。

[0111] 在特定实施例中,客户端助理服务模块426可以使用本地ASR模块216来分析用户输入。ASR模块216可以包括字形到音素(G2P)模型、发音学习模型、个性化语言模型(PLM)、端点模型和个性化声学模型。在特定实施例中,客户端助理服务模块426可以进一步使用本地NLU模块218来理解用户输入。NLU模块218可以包括命名实体解析(NER)部件和基于上下文会话的NLU部件。在特定实施例中,客户端助理服务模块426可以使用意图中介者428来分析用户的意图。为了准确了解用户的意图,意图中介者428可以访问包括与用户和世界相关联的实体的实体存储器430。在替代实施例中,用户输入可以经由在客户端系统130上执行的应用432来提交。在这种情况下,输入管理器434可以接收用户输入,并由应用环境(App Env)模块436对其进行分析。分析结果可以被发送到应用432,应用432可以进一步将分析结果发送到ASR模块216和NLU模块218。在替代实施例中,用户输入可以经由在客户端系统130上执行的助理应用438直接提交给客户端助理服务模块426。然后,客户端助理服务模块426可以基于如上所述的模块,即ASR模块216、NLU模块218和意图中介者428,执行类似的过程。

[0112] 在特定实施例中,助理编排器206可以向用户确定服务器端进程。因此,助理编排器206可以将用户输入发送到托管助理系统140的不同模块的一个或更多个计算系统。在特定实施例中,服务器助理服务模块301可以从助理编排器206接收用户输入。服务器助理服务模块301可以指示远程ASR模块208分析用户输入的音频数据。ASR模块208可以包括字形到音素(G2P)模型、发音学习模型、个性化语言模型(PLM)、端点模型和个性化声学模型。在特定实施例中,服务器助理服务模块301可以进一步指示远程NLU模块210理解用户输入。在特定实施例中,服务器助理服务模块301可以调用远程推理模型214来处理来自ASR模块208和NLU模块210的输出。在特定实施例中,推理模型214可以执行实体解析和对话优化。在特定实施例中,推理模型214的输出可以被发送到代理350,用于执行一个或更多个相关任务。

[0113] 在特定实施例中,代理350可以访问本体模块440,以准确地理解来自实体解析和对话优化的结果,从而它可以准确地执行相关任务。本体模块440可以提供与多个预定义的域、意图和槽相关联的本体数据。本体数据还可以包括不同槽和域之间的结构关系。本体数据还可以包括槽可以如何在较高级别包括域的层次结构内被分组、相关,并根据相似性和差异细分的信息。本体数据也可以包括槽可以如何在较高级别包括主题的层次结构内被分

组、相关,并根据相似性和差异细分的信息。一旦任务被执行,代理350可以将执行结果连同任务完成指示一起返回给推理模块214。

[0114] 本文公开的实施例可以包括或结合人工现实系统来实现。人工现实是一种在呈现给用户之前已经以某种方式进行了调整的现实形式,其可以包括例如虚拟现实(VR)、增强现实(AR)、混合现实(mixed reality,MR)、混杂现实(hybrid reality)或其某种组合和/或衍生物。人工现实内容可以包括完全生成的内容或与采集内容(例如,真实世界的照片)组合的生成内容。人工现实内容可以包括视频、音频、触觉反馈、或其某种组合,且其中任何一个都可以在单个信道中或在多个信道中被呈现(例如向观看者生成三维效果的立体视频)。另外,在一些实施例中,人工现实可以与应用、产品、附件、服务或其某种组合相关联,这些应用、产品、附件、服务或其某种组合例如用于在人工现实中创建内容和/或在人工现实中被使用(例如,在人工现实中执行活动)。提供人工现实内容的人工现实系统可以在各种平台上实现,这些平台包括连接到主计算机系统的头戴式显示器(HMD)、独立的HMD、移动设备或计算系统、或者能够向一个或更多个观看者提供人工现实内容的任何其他硬件平台。

[0115] 助理系统启用的智能摄像机

[0116] 在特定实施例中,当使用助理驱动的启用摄像机的客户端系统130(例如,具有摄像机的智能电话、AR/VR眼镜、智能显示器)时,助理系统140可以通过向用户提供不用手的图片/视频捕获、共享和记忆回忆来改善助理系统140的用户体验。助理系统140可以具有对象/人识别、兴趣点确定和精彩片段生成的附加功能。助理系统140可以被动地和主动地帮助用户。在被动场景中,可以有从用户经由客户端系统130到助理系统140的明确命令或提示(例如,“拍照”)。然后,助理系统140可以将其转换为机器标准意图,并通过启用摄像机的客户端系统130执行相关联的任务来为用户服务。在主动场景中,助理系统140本身可以基于其经由客户端系统130接收的信号,例如来自AR眼镜的用户的眼睛凝视,来确定其可以代表用户经由启用摄像机的客户端系统130采取的潜在有趣的动作。尽管本公开描述了通过特定系统以特定方式在特定基于摄像机的任务中帮助用户,但是本公开设想通过任何合适的系统以任何合适的方式在任何合适的基于摄像机的任务中帮助用户。

[0117] 在特定实施例中,助理系统140可以从与用户相关联的客户端系统130接收来自用户的用户输入。客户端系统130可以包括一个或更多个摄像机。在特定实施例中,助理系统140然后可以基于一个或更多个机器学习模型和由一个或更多个摄像机捕获的感测数据,确定一个或更多个摄像机的视场中的一个或更多个兴趣点。助理系统140然后可以基于一个或更多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或更多个兴趣点中的至少一个的记录。在特定实施例中,助理系统140然后可以基于多个媒体文件生成一个或更多个精彩片段文件。每个精彩片段文件可以包括满足预定义质量标准的媒体文件。在特定实施例中,助理系统140还可以向客户端系统130发送用于呈现一个或更多个精彩片段文件的指令。

[0118] 在特定实施例中,用户输入可以基于文本输入、音频输入、图像输入、视频输入、眼睛凝视、手势或运动中的一个或更多个。在特定实施例中,手势是由用户的一个或更多个身体部位有意执行的。相比之下,运动可与用户的移动或对象的移动相关联,这可能是有意的,也可能是无意的。作为示例而非限制,用户输入可以是“开始,智能摄像机”的文本输入。作为另一个示例而非限制,用户可戴着AR眼镜,并且用户输入可以是手指在眼镜上的敲击。

作为另一个示例而非限制,用户输入可以是“助理,捕获这个聚会的最佳时刻”的语音输入。在替代实施例中,用户可以通过与助理系统140相关联的设置选择开启自动捕获模式,这允许助理系统140自动检测摄像机视场中的人和对象并记录它们。在特定实施例中,如果用户输入具有不明确的提及,则助理系统140可以消除这种提及的歧义。作为示例而非限制,用户可以说“拍一张那个的照片”,为此,助理系统140可以识别视场中最重要的对象,自动裁剪视场以仅将对象显示为“那个”的歧义消除,并且一旦接收到来自用户的确认就拍摄照片。尽管本公开以特定的方式描述了特定的用户输入,但是本公开设想以任何合适的方式的任何合适的用户输入。

[0119] 在特定实施例中,助理系统可以利用各种技术来确定兴趣点。在特定实施例中,感测数据可以基于文本信号、视觉信号或音频信号中的一个或多个。作为示例而非限制,感测数据可以由客户端系统130捕获的视觉信号。因此,确定兴趣点可以包括检测视场中的一个或多个个人(服从隐私设置)以及基于一个或多个面部识别算法确定检测到的人中的一个或多个个人的一个或多个标识符。具体而言,助理系统140可以依赖于不同角度的面部识别,受制于客户端系统130的摄像机的移动、摄像机视场中的多人、远距离(例如,穿过房间)、以及人们四处移动等。此外,助理系统140可以执行设备上的面部识别,考虑到诸如AR眼镜的客户端系统130的有限计算能力,这可能更具挑战性。在特定实施例中,确定兴趣点可以基于检测到的人中的一个或多个人的兴趣度的度量。兴趣度的度量可以由一个或多个机器学习模型基于一个或多个上述标识符来确定。作为示例而非限制,检测到的人的标识符可以指示他们中的一个用户的儿子,而其他用户是用户儿子的老师。因此,用户的儿子可能具有更高的兴趣度的度量,并且助理系统140可以确定用户的儿子应该是兴趣点。尽管本公开描述了由特定系统以特定方式确定特定兴趣点,但是本公开设想由任何合适的系统以任何合适的方式确定任何合适的兴趣点。

[0120] 在特定实施例中,确定兴趣点可以包括检测视场中的一个或多个个人,并确定检测到的人中的一个或多个个人的一个或多个面部表情(服从隐私设置)。具体而言,助理系统140可以依赖于不同角度的面部表情识别,受制于客户端系统130的摄像机的移动、摄像机视场中的多人、远距离(例如,穿过房间)、以及人们四处移动等。此外,助理系统140可以执行设备上的面部表情识别,考虑到诸如AR眼镜的客户端系统130的有限计算能力,这可能更具挑战性。在特定实施例中,确定兴趣点可以基于检测到的人中的一个或多个人的兴趣度的度量。兴趣度的度量可以由一个或多个机器学习模型基于面部表情中的一个或多个来确定。作为示例而非限制,助理系统140可以确定两个检测到的人的面部表情是愉快的,而另一个检测到的人的面部表情是生气的。因此,具有愉快面部表情的两个人可能具有更高的兴趣度,并且助理系统140可以确定他们应该是兴趣点。尽管本公开描述了由特定系统以特定方式确定特定兴趣点,但是本公开设想由任何合适的系统以任何合适的方式确定任何合适的兴趣点。

[0121] 在特定实施例中,确定兴趣点可以包括检测视场中的一个或多个对象。因此,确定兴趣点可以基于检测到的对象中的一个或多个的兴趣度的度量。在特定实施例中,兴趣度的度量可以由一个或多个机器学习模型基于检测到的对象中的一个或多个来确定。作为示例而非限制,用户可以在他/她的房子外面,并且助理系统140可以检测房子、树和敞篷车。助理系统140可以确定敞篷车具有更高的兴趣度,为此,助理系统140可以进一步

将敞篷车确定为兴趣点。尽管本公开描述了由特定系统以特定方式确定特定兴趣点,但是本公开设想由任何合适的系统以任何合适的方式确定任何合适的兴趣点。

[0122] 在特定实施例中,确定兴趣点可以基于客户端系统130捕获的用户的眼睛凝视数据。作为示例而非限制,用户可戴着AR眼镜注视着狗。从注视产生的眼睛凝视数据可以指示用户对狗感兴趣。因此,助理系统140可以确定狗是兴趣点。使用不同的机器学习算法分析客户端系统130捕获的感测数据并基于该分析确定兴趣度可能是解决准确确定兴趣点的技术挑战的有效解决方案,因为机器学习算法有助于定位视场中的用户的兴趣,并且兴趣度进一步细化了这种兴趣的确定。尽管本公开描述了由特定系统以特定方式确定特定兴趣点,但是本公开设想由任何合适的系统以任何合适的方式确定任何合适的兴趣点。

[0123] 在特定实施例中,用于确定兴趣点的机器学习模型可以基于掩模区域卷积神经网络(R-CNN)。掩模R-CNN是一种神经网络架构,其被开发来用于对象检测、对象分割和关键点检测。在特定实施例中,当使用机器学习模型来确定兴趣点时,助理系统140可以另外使用中央凹(foveated)处理,该中央凹处理旨在对客户端系统130的摄像机的特定区域进行优先化和聚焦。中央凹处理有助于提高机器学习模型的灵敏度,同时保证计算效率。在特定实施例中,当使用机器学习模型来确定兴趣点时,助理系统140可以进一步利用自适应背景建模。自适应背景建模可以更好地将环境中的人或对象与环境墙壁上图片中显示的人或对象区分开来。尽管本公开以特定的方式描述了特定的机器学习模型,但是本公开设想以任何合适的方式确定任何合适的机器学习模型。

[0124] 在特定实施例中,助理系统140可以自动修改摄像机的视场,以向用户提供最高质量的视图。在特定实施例中,助理系统140可以向客户端系统130发送用于拉近摄像机中的一个或更多个以将兴趣点中的一个或更多个定位在视场的中心的指令。在特定实施例中,指令可以使摄像机能够定位视场中的多个兴趣点,使得它们的质心在视场的中心。在特定实施例中,助理系统140可以向客户端系统130发送用于拉远摄像机中的一个或更多个以将兴趣点中的一个或更多个定位在视场的中心的指令。类似地,指令可以使摄像机能够定位视场中的多个兴趣点,使得它们的质心在视场的中心。在特定实施例中,助理系统140可以自动对客户端系统130捕获的视觉数据执行必要的校准,以便于媒体文件的后续生成,从而提高媒体文件的质量。在特定实施例中,媒体文件的质量可以是媒体文件的所有视觉显著属性的组合的度量。作为示例而非限制,这些属性可以包括模糊度、亮度或颜色的鲜艳度中的一个或更多个。在特定实施例中,助理系统140可以检测客户端系统130的移动。然后,助理系统140可以将一个或更多个视觉稳定算法应用于由一个或更多个摄像机捕获的感测数据。作为示例而非限制,佩戴AR眼镜的用户可能来回移动他/她的头部,这可能导致视觉数据的低质量。因此,助理系统140可以使用视频稳定技术来校正用户的头部移动,使得视频在处理看起来平滑。因此,助理系统140可以具有帮助用户方便地以高质量捕获有趣和有意义的时刻的技术优势,因为助理系统140可以使用户能够使用各种输入并应用不同的校准和处理算法来确保对应于这些时刻的媒体文件是高质量的。尽管本公开描述了特定系统以特定方式的特定处理技术,但是本公开设想了由任何合适的系统以任何合适的方式进行的任何合适的处理技术。

[0125] 在特定实施例中,每个媒体文件的记录可以是静态图像的记录或视频的记录。换句话说,多个媒体文件中的每一个可以包括图像或视频剪辑中的一个或更多个。在特定实

施例中,助理系统140可以使用计算摄影技术来在捕获期间和捕获后增强媒体文件。作为示例而非限制,助理系统140可以快速顺序拍摄四张图片,并将它们合并以减少噪声。在特定实施例中,助理系统140可以使用户能够通过不同类型的命令编辑媒体文件。作为示例而非限制,用户可以通过不用手的编辑来编辑媒体文件,例如,由用户的语音命令请求的图片增强。尽管本公开以特定的方式描述了特定的文件,但是本公开设想了以任何合适的方式的任何合适的文件。

[0126] 在特定实施例中,助理系统140可以通过处理媒体文件来创建精彩片段文件,并返回最高质量的媒体文件的集合。精彩片段文件中的每个媒体文件可以满足预定义的质量标准。预定义的质量标准可以基于模糊度、亮度或颜色的鲜艳度中的一个或更多个。作为示例而非限制,模糊的图片可能具有低质量。作为另一个示例而非限制,具有良好亮度的图片可以具有高质量。在特定实施例中,助理系统140可以为每个捕获的媒体文件分配分数,并基于它们各自的分数对媒体文件进行排序,以生成精彩片段文件。分数可以由一个或更多个机器学习模型基于鲜艳度、颜色、兴趣度或其他合适的因素中的一个或更多个来计算。在特定实施例中,助理系统140可以另外使用以下技术来生成和个性化精彩片段文件:助理系统140可以执行重复数据消除并确定每个媒体文件的唯一性,使得所选择的媒体文件不是现有精彩片段文件的重复,并且是用户唯一感兴趣的。助理系统140可以将每个媒体文件的背景定向调整为直的,并且自动裁剪媒体文件或使媒体文件居中。助理系统140还可以检测模糊度并丢弃模糊的媒体文件。助理系统140可以根据隐私设置检测媒体文件中的人,并确定用户可能个人认识谁,因此发现相应的媒体文件是有趣的。另外,助理系统140可以检测那些人的情绪或动作(例如,眨眼、微笑、大笑等)并使用检测到的情绪和动作来确定精彩片段文件。助理系统140可以另外执行对于场景、位置和对象的语义理解,以生成精彩片段文件。此外,助理系统140可以自动增强媒体文件的颜色、亮度或清晰度,或者在精彩片段文件的生成期间降低媒体文件的噪声。在特定实施例中,可以基于与用户相关联的用户简档数据、与用户相关联的用户偏好、用户的先前用户输入或用户关系数据(例如,与社交图中的其他用户的关连)中的一个或更多个来为用户个性化精彩片段文件。作为示例而非限制,用户偏好可以指示用户想要节省他们的客户端系统130的储存器,为此助理系统140可以仅选择图像作为精彩片段文件。作为另一个示例而非限制,社交图可以指示用户与Noah和Alex相比与其他人具有更密切的关系。因此,当由助理系统140生成的媒体文件包括Noah和Alex以及一些其他人的图像/视频时,助理系统140可以选择Noah和Alex的那些作为精彩片段文件。基于用户简档数据、用户偏好、先前用户输入和用户关系数据来个性化精彩片段文件可能是解决生成对用户有意义和有兴趣的精彩片段文件的技术挑战的有效解决方案,因为如上所述的这些不同因素都提供了关于当用户回顾由助理系统140捕获的时刻时什么对用户有意义和有兴趣的特定见解。

[0127] 在特定实施例中,助理系统140还可以为精彩片段文件生成个性化字幕(captioning)。通过结合用户的个性,个性化字幕可能对用户更有吸引力。在特定实施例中,助理系统140可以进一步促进精彩片段文件的共享。助理系统140可以从客户端系统130接收来自用户的与一个或更多个其他用户共享精彩片段文件中的一个或更多个的用户请求。作为示例而非限制,用户请求可以是不用手的命令。助理系统140然后可以分别向与一个或更多个其他用户相关联的一个或更多个其他客户端系统130发送用于呈现共享的精彩

片段文件的指令。尽管本公开描述了由特定系统以特定方式生成特定精彩片段文件,但是本公开设想由任何合适的系统以任何合适的方式生成任何合适的精彩片段文件。

[0128] 在特定实施例中,助理系统140可以进一步为用户执行记忆问答(QA)。助理系统140可以从客户端系统130接收来自用户的用户查询,以响应精彩片段文件。助理系统140然后可以访问与用户相关联的多个情节记忆。助理系统140然后将所访问的情节记忆中的一个或更多个情节记忆识别为与用户查询相关。助理系统140然后可以检索对应于所识别的情节记忆的一个或更多个媒体文件。每个媒体文件可以包括帖子、评论、图像或视频剪辑中的一个或更多个。助理系统140还可以向客户端系统130发送用于呈现对应于所识别的情节记忆的一个或更多个媒体文件的指令。作为示例而非限制,在看到精彩片段文件后,用户可能会问“我见过那个穿红色t恤的人吗?”然后,助理系统140可以访问用户的情节记忆,并发现用户几年前在一次会议上见过穿红色t恤的人。助理系统140可以检索在会议上拍摄的团体照片。用户和穿红色t恤的人都在团体照片中。因此,助理系统140可以具有帮助用户将当前捕获的时刻与用户的记忆相关连的技术优势,使得用户可以更好地欣赏它们,因为助理系统140可以执行记忆回忆来检索与当前捕获的时刻相关的媒体文件。尽管本公开描述了由特定系统以特定方式执行特定记忆QA,但是本公开设想由任何合适的系统以任何合适的方式执行任何合适的记忆QA。关于情节记忆的更多信息可以在2019年8月27日提交的美国专利申请第16/552559号中找到,该申请在此讨论。

[0129] 图5A-5D示出了其中助理系统140可以启用客户端系统130上的智能摄像机来帮助用户的示例场景。图5A示出了智能摄像机可以帮助用户的示例场景。在图5A中,用户502可以使用启用摄像机的客户端系统130,例如AR眼镜504。用户502可能正在与一些其他在场的人,即人506、人508、人510和人512一起参加晚宴。图5B示出了兴趣点的示例确定。在图5B中,AR眼镜504的摄像机的视场可以被AR眼镜504的屏幕514反映。助理系统140可以使用由AR眼镜504捕获的视觉信号来确定兴趣点。根据隐私设置,助理系统140可以为晚宴上的每个人执行面部识别。特别地,助理系统140可以识别出人512是用户502的妻子。因此,助理系统140可以将人512确定为兴趣点516。助理系统140可以进一步发送指令来拉近AR眼镜504的摄像机,使得人512在屏幕514的中间。其他人可能会作为视场的背景消失或模糊。例如,人508和人510不再在屏幕514中,并且人506被部分示出并被模糊。一旦兴趣点516被确定,用户502可以使用各种用户输入,例如语音命令,经由助理系统140拍摄图片、视频或跟踪兴趣点516,这受制于适用的隐私策略。图5C示出了兴趣点的另一示例确定。助理系统140可以使用由AR眼镜504捕获的视觉信号来确定兴趣点。助理系统140可以为晚宴上的每个人执行面部表情识别。特别地,助理系统140可以确定兴趣点518应该是人506和人508,因为他们都在笑。助理系统140可以进一步发送指令来拉近AR眼镜504的摄像机,使得人506和人508在屏幕514的中间。其他人可能会被模糊为视场的背景。例如,人510和人512是模糊的。一旦兴趣点518被确定,用户502可以使用各种用户输入,例如语音命令,经由助理系统140拍摄图片、视频或跟踪兴趣点518,这受制于适用的隐私策略。图5D示出了示例精彩片段文件。在图5D中,助理系统140可能已经拍摄了先前确定的兴趣点的照片/视频。为了便于用户502回顾有趣且有意义的时刻,助理系统140可以生成精彩片段文件520,其可以包括图像、视频或其组合。助理系统140可以进一步询问用户是否想要查看它们。例如,助理系统140可以在AR眼镜504的屏幕514上呈现通知522。通知522可以是“我已经生成了今晚晚宴的一些精彩片段。”

你想看它们吗?”尽管本公开以特定方式描述了智能摄像机的特定示例,但是本公开设想以任何合适方式的智能摄像机的任何合适示例。

[0130] 图6示出了用于由助理系统140启用智能摄像机的示例方法600。该方法可以在步骤610开始,其中助理系统140可以从与用户相关联的客户端系统130接收来自用户的用户输入,其中客户端系统130包括一个或更多个摄像机。在步骤620,助理系统140可以基于一个或更多个机器学习模型和由一个或更多个摄像机捕获的感测数据,确定一个或更多个摄像机的视场中的一个或更多个兴趣点。步骤620可以包括以下并行子步骤。在子步骤621,助理系统140可以进一步基于检测视场中的一个或更多个人并且基于一个或更多个面部识别算法和隐私设置来确定检测到的人中的一个或更多个人的一个或更多个标识符来确定兴趣点,其中确定兴趣点基于检测到的人中的一个或更多个人的兴趣度的度量,其中兴趣度的度量由一个或更多个机器学习模型基于标识符中的一个或更多个来确定。在子步骤622,助理系统140可以进一步基于检测视场中的一个或更多个人并确定检测到的人中的一个或更多个人的一个或更多个面部表情来确定兴趣点,其中确定兴趣点基于检测到的人中的一个或更多个人的兴趣度的度量,其中兴趣度的度量由一个或更多个机器学习模型基于面部表情中的一个或更多个来确定。在子步骤623,助理系统140可以进一步基于检测视场中的一个或更多个对象来确定兴趣点,其中确定兴趣点基于检测到的对象中的一个或更多个的兴趣度的度量,其中兴趣度的度量由一个或更多个机器学习模型基于检测到的对象中的一个或更多个来确定。在子步骤624,助理系统140可以进一步基于由客户端系统130捕获的用户的眼睛凝视数据来确定兴趣点。在子步骤625,助理系统140可以进一步基于场景和语义理解来确定兴趣点。在步骤630,助理系统140可以基于一个或更多个兴趣点生成多个媒体文件,其中每个媒体文件是一个或更多个兴趣点中的至少一个的记录。在步骤640,助理系统140可以基于多个媒体文件生成一个或更多个精彩片段文件,其中每个精彩片段文件包括满足预定义质量标准的媒体文件,其中精彩片段文件基于与用户相关联的用户简档数据、与用户相关联的用户偏好、用户的先前用户输入或用户关系数据中的一个或更多个而为用户个性化。在步骤650,助理系统140可以向客户端系统130发送用于呈现一个或更多个精彩片段文件的指令。在适当的情况下,特定实施例可以重复图6的方法的一个或更多个步骤。虽然本公开描述并且示出了图6的方法的特定步骤以特定顺序发生,但是本公开设想到图6的方法的任何合适的步骤以任何合适的顺序出现。此外,尽管本公开描述并示出了包括图6的方法的特定步骤的用于帮助用户使用智能摄像机的示例方法,但是本公开设想了用于帮助用户使用智能摄像机的任何合适的方法,包括任何合适的步骤,在适当的情况下,可以包括图6的方法的所有、一些步骤或没有这些步骤。此外,虽然本公开描述并且示出了执行图6的方法的特定步骤的特定部件、设备或系统,但是本公开设想到执行图6的方法的任何合适步骤的任何合适的部件、设备或系统的任何合适的组合。

[0131] 社交图

[0132] 图7示出了示例社交图700。在特定实施例中,社交网络系统160可以在一个或更多个数据存储器中存储一个或更多个社交图700。在特定实施例中,社交图700可以包括多个节点——该多个节点可以包括多个用户节点702或多个概念节点704——以及关连这些节点的多条边706。每个节点可以与唯一的实体(即,用户或概念)相关联,每个实体可以具有唯一的标识符(ID),例如唯一的号码或用户名。出于教导的目的,以二维视觉映射表示

(two-dimensional visual map representation) 示出了图7所示的示例社交图700。在特定实施例中,社交网络系统160、客户端系统130、助理系统140或第三方系统170可以访问社交图700和相关社交图信息以用于合适的应用。社交图700的节点和边可以作为数据对象被存储在例如数据储存器(例如社交图数据库)中。这种数据储存器可以包括社交图700的节点或边的一个或更多个可搜索或可查询的索引。

[0133] 在特定实施例中,用户节点702可以对应于社交网络系统160或助理系统140的用户。作为示例而不是作为限制,用户可以是与社交网络系统160或助理系统140交互或通信或通过社交网络系统160或助理系统140交互或通信的个人(人类用户)、实体(例如,企业、公司或第三方应用)或(例如,个人或实体的)团体。在特定实施例中,当用户向社交网络系统160注册账户时,社交网络系统160可以创建对应于用户的用户节点702,并将用户节点702存储在一个或更多个数据储存器中。本文描述的用户和用户节点702在适当的情况下可以指注册的用户和与注册的用户相关联的用户节点702。另外或作为替代方案,在适当的情况下,本文描述的用户和用户节点702可以指没有向社交网络系统160注册的用户。在特定实施例中,用户节点702可以与由用户提供的信息或由各种系统(包括社交网络系统160)收集的信息相关联。作为示例而不是作为限制,用户可以提供他或她的姓名、简档图片、联系信息、出生日期、性别、婚姻状况、家庭状况、职业、教育背景、偏好、兴趣或其他人口统计信息。在特定实施例中,用户节点702可以与对应于与用户相关联的信息的一个或更多个数据对象相关联。在特定实施例中,用户节点702可以对应于一个或更多个web界面。

[0134] 在特定实施例中,概念节点704可以对应于概念。作为示例而不是作为限制,概念可以对应于地点(诸如例如,电影院、餐馆、地标或城市);网站(诸如例如,与社交网络系统160相关联的网站或与web应用服务器相关联的第三方网站);实体(诸如例如,个人、企业、团体、运动队或名人);资源(诸如例如,音频文件、视频文件、数字照片、文本文件、结构化文档或应用),其可以位于社交网络系统160内或外部服务器(例如web应用服务器)上;不动产或知识产权(诸如例如,雕塑、绘画、电影、游戏、歌曲、想法、照片或书面作品);游戏;活动;想法或理论;另一个合适的概念;或者两个或更多个这样的概念。概念节点704可以与由用户提供的概念的信息或由各种系统(包括社交网络系统160和助理系统140)收集的信息相关联。作为示例而不是作为限制,概念的信息可以包括名称或标题;一个或更多个图像(例如,书籍的封面的图像);位置(例如,地址或地理位置);网站(其可以与URL相关联);联系信息(例如,电话号码或电子邮件地址);其他合适的概念信息;或者这样的信息的任何合适的组合。在特定实施例中,概念节点704可以与一个或更多个数据对象相关联,该一个或更多个数据对象对应于与概念节点704相关联的信息。在特定实施例中,概念节点704可以对应于一个或更多个web界面。

[0135] 在特定实施例中,社交图700中的节点可以表示web界面(其可以被称为“简档界面”)或者由web界面表示。简档界面可以由社交网络系统160或助理系统140托管或是社交网络系统160或助理系统140可访问的。简档界面也可以在与第三方系统170相关联的第三方网站上被托管。作为示例而不是作为限制,对应于特定外部web界面的简档界面可以是特定外部web界面,并且简档界面可以对应于特定概念节点704。简档界面可以由所有其他用户或其他用户的选定子集可查看。作为示例而不是作为限制,用户节点702可以具有相应的用户简档界面,其中相应的用户可以添加内容、作出声明或以其他方式表达他或她自己。作

为另一示例而不是作为限制,概念节点704可以具有相应的概念简档界面,其中一个或多个用户可以添加内容、作出声明或表达他们自己,特别是关于对应于概念节点704的概念。

[0136] 在特定实施例中,概念节点704可以表示由第三方系统170托管的第三方web界面或资源。第三方web界面或资源可以包括内容、可选择的或其他图标、或代表动作或活动的其他可交互对象以及其他元素。作为示例而不是作为限制,第三方web界面可以包括可选择的图标,例如“赞”、“签到(check-in)”、“吃”、“推荐”,或其他合适的动作或活动。查看第三方web界面的用户可以通过选择图标之一(例如,“签到”)来执行动作,使客户端系统130向社交网络系统160发送指示用户的动作的消息。响应于该消息,社交网络系统160可以在对应于用户的用户节点702和对应于第三方web界面或资源的概念节点704之间创建边(例如,签到类型边),并将边706存储在一个或多个数据储存器中。

[0137] 在特定实施例中,社交图700中的一对节点可以通过一条或更多条边706关连到彼此。连接一对节点的边706可以表示在该对节点之间的关系。在特定实施例中,边706可以包括或表示对应于在一对节点之间的关系的一个或多个数据对象或属性。作为示例而不是作为限制,第一用户可以指示第二用户是第一用户的“朋友”。响应于该指示,社交网络系统160可以向第二用户发送“好友请求”。如果第二用户确认“好友请求”,则社交网络系统160可以在社交图700中创建将第一用户的用户节点702关连到第二用户的用户节点702的边706,并将边706作为社交图信息存储在一个或多个数据储存器164中。在图7的示例中,社交图700包括指示在用户“A”和用户“B”的用户节点702之间的朋友关系的边706、以及指示在用户“C”和用户“B”的用户节点702之间的朋友关系的边。尽管本公开描述或示出了关连特定用户节点702的具有特定属性的特定边706,但是本公开设想关连用户节点702的具有任何合适属性的任何合适的边706。作为示例而不是作为限制,边706可以表示友谊、家庭关系、商业或雇佣关系、粉丝关系(包括例如,赞等)、关注者(follower)关系、访问者关系(包括例如,访问、查看、签到、分享等)、订购者关系、上级/下级关系、互惠关系、非互惠关系、另一种合适类型的关系、或两种或更多种这样的关系。此外,尽管本公开一般将节点描述为被关连,但是本公开也将用户或概念描述为被关连。在本文,对被关连的用户或概念的引用在适当的情况下可以指在社交图700中由一条或更多条边706关连的对应于那些用户或概念的节点。

[0138] 在特定实施例中,在用户节点702和概念节点704之间的边706可以表示由与用户节点702相关联的用户对与概念节点704相关联的概念执行的特定动作或活动。作为示例而不是作为限制,如图7所示,用户可以“赞”、“出席”、“播放”、“收听”、“烹饪”、“工作于”或“观看”概念,其中每个可以对应于边类型或子类型。对应于概念节点704的概念简档界面可以包括例如可选择的“签到”图标(诸如例如,可点击的“签到”图标)或可选择的“添加到收藏夹”图标。类似地,在用户点击这些图标之后,社交网络系统160可以响应于对应于相应动作的用户动作来创建“收藏夹”边或“签到”边。作为另一示例而不是作为限制,用户(用户“C”)可以使用特定的应用(第三方在线音乐应用)来收听特定的歌曲(“想象(Imagine)”)。在这种情况下,社交网络系统160可以在对应于用户的用户节点702和对应于歌曲和应用的节点704之间创建“收听”边706和“使用”边(如图7所示),以指示用户收听了歌曲并使用了应用。此外,社交网络系统160可以在对应于歌曲和应用的节点704之间创建“播放”边706(如图7所示),以指示特定的歌曲被特定的应用播放。在这种情况下,“播放”边706对应

于由外部应用(第三方在线音乐应用)对外部音频文件(歌曲“想象”)执行的动作。尽管本公开描述了连接用户节点702和概念节点704的具有特定属性的特定边706,但是本公开设想连接了连接用户节点702和概念节点704的具有任何适当属性的任何适当边706。此外,尽管本公开描述了表示单个关系的在用户节点702和概念节点704之间的边,但是本公开设想了表示一个或更多个关系的在用户节点702和概念节点704之间的边。作为示例而不是作为限制,边706可以表示用户喜欢并使用了特定概念。可替代地,另一条边706可以表示用户节点702和概念节点704之间(如图7所示,用户“E”的用户节点702和“在线音乐应用”的概念节点704之间)的每种类型的关系(或者多个单一关系)。

[0139] 在特定实施例中,社交网络系统160可以在社交图700中的用户节点702和概念节点704之间创建边706。作为示例而不是作为限制,(诸如例如,通过使用由用户的客户端系统130托管的web浏览器或专用应用)查看概念简档界面的用户可以通过点击或选择“赞”图标来指示他或她喜欢由概念节点704表示的概念,这可以使用户的客户端系统130向社交网络系统160发送指示用户喜欢与概念简档界面相关联的概念的消息。响应于该消息,社交网络系统160可以在与用户相关联的用户节点702和概念节点704之间创建边706,如由在用户节点和概念节点704之间的“赞”边706所示的。在特定实施例中,社交网络系统160可以将边706存储在一个或更多个数据储存器中。在特定实施例中,边706可以由社交网络系统160响应于特定用户动作而自动形成。作为示例而不是作为限制,如果第一用户上传图片、观看电影或收听歌曲,则可以在对应于第一用户的用户节点702和对应于那些概念的概念节点704之间形成边706。尽管本公开描述了以特定方式形成特定边706,但是本公开设想了以任何合适的方式形成任何合适的边706。

[0140] 向量空间和嵌入

[0141] 图8示出了向量空间800的示例视图。在特定实施例中,可以在d维向量空间中表示对象或n-gram,其中d表示任何合适的维数。尽管向量空间800被示为三维空间,但这仅仅是为了说明的目的,因为向量空间800可以具有任何合适的维度。在特定实施例中,n-gram可以在向量空间800中被表示为向量,该向量被称为项目嵌入(term embedding)。每个向量可以包括对应于向量空间800中的特定点(即,向量的终点)的坐标。作为示例而不是作为限制,如图8所示,向量810、820和830可以被表示为向量空间800中的点。n-gram可以被映射到相应的向量表示。作为示例而不是作为限制,通过应用由字典定义的函数 $\vec{\pi}$,n-gram t_1 和n-gram t_2 可以分别映射到向量空间800中的向量 \vec{v}_1 和 \vec{v}_2 ,使得 $\vec{v}_1 = \vec{\pi}(t_1)$ 并且 $\vec{v}_2 = \vec{\pi}(t_2)$ 。作为另一个示例而不是作为限制,可以利用被训练为将文本映射到向量表示的字典,或者这种字典本身可以通过训练来生成。作为另一示例而非限制,单词嵌入模型可用于将n-gram映射到向量空间800中的向量表示。在特定实施例中,可以通过使用机器学习模型(例如,神经网络)将n-gram映射到向量空间800中的向量表示。机器学习模型可能已经使用训练数据的序列(例如,每个包括n-gram的多个对象的语料库(corpus))进行了训练。

[0142] 在特定实施例中,对象可以在向量空间800中被表示为向量,该向量被称为特征向量或对象嵌入。作为示例而不是作为限制,通过应用函数 $\vec{\pi}$,对象 e_1 和 e_2 可以分别映射到向量空间800中的向量 \vec{v}_1 和 \vec{v}_2 ,使得 $\vec{v}_1 = \vec{\pi}(e_1)$ 并且 $\vec{v}_2 = \vec{\pi}(e_2)$ 。在特定实施例中,可以基于对象的一个或更多个特性、属性或特征、对象与其他对象的关系或与对象相关联的任何

其他合适的信息来将对象映射到向量。作为示例而不是作为限制,函数 $\vec{\pi}$ 可以通过特征提取将对象映射到向量,特征提取可以从初始测量数据集开始,并构建导出值(例如,特征)。作为示例而不是作为限制,通过使用算法以检测或隔离对象的各种期望部分或形状,包括视频或图像的对象可以被映射到向量。用于计算向量的特征可以基于从边检测、拐角(corner)检测、团块(blob)检测、脊线(ridge)检测、尺度不变特征变换、边方向、变化强度、自相关、运动检测、光流、阈值化、团块提取、模板匹配、霍夫(Hough)变换(例如,线、圆、椭圆、任意形状)获得的信息或任何其他合适的信息。作为另一个示例而不是作为限制,包括音频数据的对象可以基于特征(例如频谱斜率、音调系数、音频频谱质心、音频频谱包络、梅尔频率倒频谱(Mel-frequency cepstrum)或任何其他合适的信息)来被映射到向量。在特定实施例中,当对象具有太大而无法有效地被处理的数据或者包括冗余数据时,函数 $\vec{\pi}$ 可以使用所变换的精简特征集(例如,特征选择)来将对象映射到向量。在特定实施例中,函数 $\vec{\pi}$ 可以基于与对象e相关联的一个或多个n-gram来将对象e映射到向量 $\vec{\pi}(e)$ 。尽管本公开描述了以特定方式在向量空间中表示n-gram或对象,但是本公开设想了以任何合适的方式在向量空间中表示n-gram或对象。

[0143] 在特定实施例中,社交网络系统160可以计算向量空间800中的向量的相似性度量。相似性度量可以是余弦相似性、Minkowski距离、Mahalanobis距离、Jaccard相似性系数或任何合适的相似性度量。作为示例而不是作为限制, \vec{v}_1 和 \vec{v}_2 的相似性度量可以是余弦相似性

$\frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|}$ 。作为另一个示例而不是作为限制, \vec{v}_1 和 \vec{v}_2 的相似性度量可以是欧几里德

(Euclidean)距离 $\|\vec{v}_1 - \vec{v}_2\|$ 。两个向量的相似性度量可以表示分别对应于两个向量的两个对象或n-gram彼此之间的相似程度,如通过向量空间800中两个向量之间的距离所测量的。作为示例而不是作为限制,基于相应向量之间的距离,向量810和向量820可以对应于比对应于向量810和向量830的对象与彼此更相似的对象。尽管本公开描述了以特定方式计算向量之间的相似性度量,但本公开设想了以任何合适的方式计算向量之间的相似性度量。

[0144] 关于向量空间、嵌入、特征向量和相似性度量的更多信息可以在2015年11月23日提交的美国专利申请第14/949436号、2016年10月5日提交的美国专利申请第15/286315号和2016年11月30日提交的美国专利申请第15/365789号中找到,本文将讨论这些专利申请中的每一个。

[0145] 神经网络

[0146] 图9示出了示例神经网络(“ANN”)900。在特定实施例中,ANN可以指包括一个或多个节点的计算模型。示例ANN 900可以包括输入层910、隐藏层920、930、940和输出层950。ANN 900的每个层可以包括一个或多个节点,例如节点905或节点915。在特定实施例中,ANN的每个节点可以关连到ANN的另一个节点。作为示例而不是作为限制,输入层910的每个节点可以关连到隐藏层920的一个或多个节点。在特定实施例中,一个或多个节点可以是偏置节点(例如,层中不关连到前一层中的任何节点并且不从其接收输入的节点)。在特定实施例中,每层中的每个节点可以关连到前一层或后一层的一个或多个节点。尽管图9描绘了具有特定层数、特定节点数和节点间特定关连的特定ANN,但是本公开设想了

具有任何合适层数、任何合适节点数和节点间任何合适关连的任何合适的ANN。作为示例而不是作为限制,尽管图9描绘了输入层910的每个节点和隐藏层920的每个节点之间的关连,但是输入层910的一个或更多个节点可以不关连到隐藏层920的一个或更多个节点。

[0147] 在特定实施例中,ANN可以是前馈ANN(例如,没有循环或回路的ANN,其中节点之间的通信从输入层开始在一个方向上流动,并前进到连续层)。作为示例而不是作为限制,隐藏层920的每个节点的输入可以包括输入层910的一个或更多个节点的输出。作为另一个示例而不是作为限制,输出层950的每个节点的输入可以包括隐藏层940的一个或更多个节点的输出。在特定实施例中,ANN可以是深度神经网络(例如,包括至少两个隐藏层的神经网络)。在特定实施例中,ANN可以是深度残差网络。深度残差网络可以是前馈ANN,其包括组织成残差块的隐藏层。第一残差块之后的每个残差块的输入可以是前一个残差块的输出和前一个残差块的输入的函数。作为示例而不是作为限制,到残差块N的输入可以是 $F(x) + x$,其中 $F(x)$ 可以是残差块N-1的输出, x 可以是到残差块N-1的输入。尽管本公开描述了特定的ANN,但是本公开设想任何合适的ANN。

[0148] 在特定实施例中,激活函数可以对应于ANN的每个节点。节点的激活函数可以针对给定输入定义节点的输出。在特定实施例中,节点的输入可以包括输入集合。作为示例而不是作为限制,激活函数可以是恒等函数、二进制阶跃函数、逻辑函数或任何其他合适的函数。作为另一个示例而不是作为限制,节点k的激活函数可以是sigmoid函数

$F_k(s_k) = \frac{1}{1+e^{-s_k}}$,
双曲线正切函数 $F_k(s_k) = \frac{e^{s_k} - e^{-s_k}}{e^{s_k} + e^{-s_k}}$ 、整流器 $F_k(s_k) = \max(0, s_k)$ 或任何其他合适的函数 F_k

(s_k),其中 s_k 可以是节点k的有效输入。在特定实施例中,对应于节点的激活函数的输入可以被加权。每个节点可以使用基于加权输入的对应激活函数来生成输出。在特定实施例中,节点之间的每个关连可以与权重相关联。作为示例而不是作为限制,节点905和节点915之间的关连925可以具有0.4的加权系数,这可以指示节点905的输出乘以0.4被用作节点915的输入。作为另一个示例而不是作为限制,节点k的输出 y_k 可以是 $y_k = F_k(s_k)$,其中 F_k 可以是对应于节点k的激活函数, $s_k = \sum_j (w_{jk} x_j)$ 可以是节点k的有效输入, x_j 可以是关连到节点k的节点j的输出,并且 w_{jk} 可以是节点j和节点k之间的加权系数。在特定实施例中,输入层的节点的输入可以基于表示对象的向量。尽管本公开描述了节点的特定输入和输出,但是本公开考虑了节点的任何合适的输入和输出。此外,尽管本公开可以描述节点之间的特定关连和权重,但是本公开考虑了节点之间的任何合适的关连和权重。

[0149] 在特定实施例中,可以使用训练数据来训练ANN。作为示例而不是作为限制,训练数据可以包括ANN 900的输入和预期输出。作为另一个示例而不是作为限制,训练数据可以包括向量,每个向量表示训练对象和每个训练对象的预期标签。在特定实施例中,训练ANN可以包括通过优化目标函数来修改与ANN的节点之间的关连相关联的权重。作为示例而不是作为限制,可以使用训练方法(例如,共轭梯度法、梯度下降法、随机梯度下降)来反向传播作为表示训练对象的每个向量之间的距离测量的平方和误差(例如,使用最小化平方和误差的成本函数)。在特定实施例中,可以使用丢弃技术来训练ANN。作为示例而不是作为限制,在训练时可以暂时忽略一个或更多个节点(例如,不接收输入并且不生成输出)。对于每个训练对象,ANN的一个或更多个节点都可以有被忽略的一定概率。针对特定训练对象忽略

的节点可以不同于针对其他训练对象忽略的节点(例如,可以逐个对象地临时忽略节点)。尽管本公开描述了以特定方式训练ANN,但是本公开设想了以任何合适的方式训练ANN。

[0150] 隐私

[0151] 在特定实施例中,计算系统的一个或更多对象(例如,内容或其他类型的对象)可以与一个或更多个隐私设置相关联。一个或更多对象可以存储在任何合适的计算系统或应用上,或者以其他方式与任何合适的计算系统或应用相关联,该计算系统或应用诸如例如是社交网络系统160、客户端系统130、助理系统140、第三方系统170、社交网络应用、助理应用、消息传递应用、照片共享应用或者任何其他合适的计算系统或应用。尽管本文讨论的示例是在在线社交网络的上下文中,但是这些隐私设置可以应用于任何其他合适的计算系统。对象的隐私设置(或“访问设置”)可以以任何合适的方式(诸如例如,与对象相关联地、或在授权服务器上的索引、以另一种合适的方式或其任何合适的组合)存储。对象的隐私设置可以指定如何可以在在线社交网络中访问、存储或以其他方式使用(例如,查看、共享、修改、复制、执行、显现(surfaced)或识别)该对象(或与该对象相关联的特定信息)。当对象的隐私设置允许特定用户或其他实体访问该对象时,该对象可以被描述为相对于该用户或其他实体是“可见的”。作为示例而非限制,在线社交网络的用户可以指定关于用户简档页面的隐私设置,该隐私设置识别可以访问在用户简档页面上的工作经历信息的一组用户,因而排除其他用户访问该信息。

[0152] 在特定实施例中,对象的隐私设置可以指定不应当被允许访问与对象相关联的某些信息的用户或其他实体的“黑名单(blocked list)”。在特定实施例中,黑名单可以包括第三方实体。黑名单可以指定一个或更多用户或实体,对象对这些用户或实体是不可见的。作为示例而非限制,用户可以指定不能访问与该用户相关联的相册的一组用户,从而排除这些用户访问相册(同时也可能允许不在指定用户群组内的某些用户访问相册)。在特定实施例中,隐私设置可以与特定社交图元素相关联。社交图元素(例如节点或边)的隐私设置可以指定可以如何使用在线社交网络来访问社交图元素、与社交图元素相关联的信息、或与社交图元素相关联的对象。作为示例而非限制,对应于特定照片的特定概念节点704可以具有指定照片仅可以由在照片中标记的用户以及在照片中标记的用户的访问的隐私设置。在特定实施例中,隐私设置可以允许用户选择加入或选择退出使他们的内容、信息或动作被社交网络系统160或助理系统140存储/记录或者与其他系统(例如,第三方系统170)分享。尽管本公开描述了以特定方式使用特定隐私设置,但是本公开设想了以任何合适的方式使用任何合适的隐私设置。

[0153] 在特定实施例中,隐私设置可以基于社交图700的一个或更多节点或边。可以为社交图700的一条或更多条边706或边类型、或者关于社交图700的一个或更多节点702、704或节点类型,指定隐私设置。应用于关连两个节点的特定边706的隐私设置可以控制对应于这两个节点的两个实体之间的关系对于在线社交网络的其他用户是否可见。类似地,应用于特定节点的隐私设置可以控制对应于该节点的用户或概念对于在线社交网络的其他用户是否可见。作为示例而不是作为限制,第一用户可以向社交网络系统160共享对象。该对象可以与通过边706关连到第一用户的用户节点702的概念节点704相关联。第一用户可以指定应用于关连到对象的概念节点704的特定边706的隐私设置,或者可以指定应用于关连到概念节点704的所有边706的隐私设置。作为另一个示例而不是作为限制,第一用户

可以共享特定对象类型的对象集合(例如,图像集合)。第一用户可以针对与第一用户相关联的该特定对象类型的所有对象将隐私设置指定为具有特定的隐私设置(例如,指定由第一用户发布的所有图像仅对第一用户的朋友和/或在图像中标记的用户可见)。

[0154] 在特定实施例中,社交网络系统160可以向第一用户呈现“隐私向导(privacy wizard)”(例如,在网页、模块、一个或更多个对话框或任何其他合适的界面内),以帮助第一用户指定一个或更多个隐私设置。隐私向导可以显示指令、合适的隐私相关信息、当前隐私设置、用于接受来自第一用户的指定隐私设置的改变或确认的一个或更多个输入的一个或更多个输入字段、或其任何合适的组合。在特定实施例中,社交网络系统160可以向第一用户提供“仪表板(dashboard)”功能,该功能可以向第一用户显示第一用户的当前隐私设置。仪表板功能可以在任何适当的时间向第一用户显示(例如,在来自调用仪表板功能的第一用户的输入之后,在特定事件或触发动作发生之后)。仪表板功能可以允许第一用户在任何时间以任何合适的方式修改第一用户的一个或更多个当前隐私设置(例如,将第一用户重定向到隐私向导)。

[0155] 与对象相关联的隐私设置可以指定允许访问或拒绝访问的任何合适的粒度。作为示例而不是作为限制,可以为特定用户(例如,只有我、我的室友、我的老板)、在特定分离度内的用户(例如,朋友、朋友的朋友)、用户群组(例如,游戏俱乐部、我的家人)、用户网络(例如,特定雇主的雇员、特定大学的学生或校友)、所有用户(“公众”)、无用户(“私人的”)、第三方系统170的用户、特定应用(例如,第三方应用、外部网站)、其他合适的实体、或其任何合适的组合来指定访问或拒绝访问。尽管本公开描述了允许访问或拒绝访问的特定粒度,但是本公开设想允许访问或拒绝访问的任何合适粒度。

[0156] 在特定实施例中,一个或更多个服务器162可以是用于实施隐私设置的授权/隐私服务器。响应于来自用户(或其他实体)的对存储在数据储存器164中的特定对象的请求,社交网络系统160可以向数据储存器164发送对该对象的请求。请求可以识别与该请求相关联的用户,并且对象只有在授权服务器基于与该对象相关联的隐私设置确定该用户被授权访问该对象时才可以被发送给该用户(或者该用户的客户端系统130)。如果请求用户未被授权访问该对象,则授权服务器可以阻止所请求的对象从数据储存器164中被检索,或者可以阻止所请求的对象被发送给用户。在搜索-查询上下文中,只有当查询用户被授权访问对象时,例如,如果对象的隐私设置允许它显现给查询用户、被查询用户发现或以其他方式对查询用户可见,对象才可以被提供作为搜索结果。在特定实施例中,对象可以表示通过用户的动态消息对用户可见的内容。作为示例而非限制,一个或更多个对象对于用户的“热门话题(Trending)”页面可以是可见的。在特定实施例中,对象可以对应于特定用户。对象可以是与特定用户相关联的内容,或者可以是特定用户的账户或存储在社交网络系统160或其他计算系统上的信息。作为示例而不是作为限制,第一用户可以通过在线社交网络的“你可能认识的人(People You May Know)”功能或者通过查看第一用户的朋友列表来查看在线社交网络的一个或更多个第二用户。作为示例而不是作为限制,第一用户可以指定他们不希望他们的动态消息或朋友列表中看到与特定第二用户相关联的对象。如果对象的隐私设置不允许其被显露给用户、被用户发现或对用户可见,则该对象可以从搜索结果中排除。尽管本公开描述了以特定方式实施隐私设置,但是本公开设想以任何合适的方式实施隐私设置。

[0157] 在特定实施例中,与用户相关联的相同类型的不同对象可以具有不同的隐私设置。与用户相关联的不同类型的对象可能具有不同类型的隐私设置。作为示例而非限制,第一用户可以指定第一用户的状态更新是公开的,但是由第一用户共享的任何图像仅对在线社交网络上的第一用户的朋友可见。作为另一个示例而非限制,用户可以为不同类型的实体(例如个人用户、朋友的朋友、关注者、用户群组或公司实体)指定不同的隐私设置。作为另一示例而非限制,第一用户可以指定可以观看由第一用户发布的视频的一组用户,同时防止视频对第一用户的雇主可见。在特定实施例中,可以为不同的用户群组或用户人口统计提供不同的隐私设置。作为示例而非限制,第一用户可以指定与第一用户上同一所大学的其他用户可以观看第一用户的照片,但是作为第一用户的家庭成员的其他用户不可以观看那些相同的照片。

[0158] 在特定实施例中,社交网络系统160可以为特定对象类型的每个对象提供一个或更多默认隐私设置。设置为默认的对象隐私设置可以由与该对象相关联的用户来改变。作为示例而非限制,由第一用户发布的所有图像可以具有仅对第一用户的朋友可见的默认隐私设置,并且对于特定图像,第一用户可以将该图像的隐私设置改变为对朋友和朋友的朋友可见。

[0159] 在特定实施例中,隐私设置可以允许第一用户指定(例如,通过选择退出,通过不选择加入)社交网络系统160或助理系统140是否可以出于任何目的接收、收集、记录或存储与用户相关联的特定对象或信息。在特定实施例中,隐私设置可以允许第一用户指定特定应用或进程是否可以访问、存储或使用与用户相关联的特定对象或信息。隐私设置可以允许第一用户选择加入或选择退出使对象或信息被特定应用或进程访问、存储或使用。社交网络系统160或助理系统140可以访问这样的信息,以便向第一用户提供特定的功能或服务,而社交网络系统160或助理系统140不能出于任何其他目的访问该信息。在访问、存储或使用这样的对象或信息之前,社交网络系统160或助理系统140可以提示用户提供隐私设置,该隐私设置指定哪些应用或进程(如果有的话)可以在允许任何这样的动作之前访问、存储或使用对象或信息。作为示例而不是作为限制,第一用户可以经由与在线社交网络相关的应用(例如,消息传递app)向第二用户传输消息,并且可以指定社交网络系统160或助理系统140不应该存储这种消息的隐私设置。

[0160] 在特定实施例中,用户可以指定社交网络系统160或助理系统140是否可以访问、存储或使用与第一用户相关联的特定类型的对象或信息。作为示例而不是作为限制,第一用户可以指定由第一用户通过社交网络系统160或助理系统140发送的图像不可以被社交网络系统160或助理系统140存储。作为另一个示例而不是作为限制,第一用户可以指定从第一用户发送给特定第二用户的消息不可以被社交网络系统160或助理系统140存储。作为又一个示例而不是作为限制,第一用户可以指定经由特定应用发送的所有对象可以被社交网络系统160或助理系统140保存。

[0161] 在特定实施例中,隐私设置可以允许第一用户指定是否可以从特定客户端系统130或第三方系统170访问与第一用户相关联的特定对象或信息。隐私设置可以允许第一用户选择加入或选择退出从特定设备(例如,用户的智能手机上的电话簿)、从特定应用(例如,消息传递app)或从特定系统(例如,电子邮件服务器)访问对象或信息。社交网络系统160或助理系统140可以提供关于每个设备、系统或应用的默认隐私设置,和/或可以提示第

一用户为每个上下文指定特定的隐私设置。作为示例而不是作为限制,第一用户可以利用社交网络系统160或助理系统140的位置服务特征来提供用户附近的餐馆或其他地方的推荐。第一用户的默认隐私设置可以指定社交网络系统160或助理系统140可以使用从第一用户的客户端设备130提供的位置信息来提供基于位置的服务,但是社交网络系统160或助理系统140不可以存储第一用户的位置信息或将其提供给任何第三方系统170。第一用户然后可以更新隐私设置,以允许第三方图像共享应用使用位置信息,以便对照片进行地理标记。

[0162] 在特定实施例中,隐私设置可以允许用户指定可以从其访问对象的一个或多个地理位置。对对象的访问或拒绝访问可取决于试图访问对象的用户的地理位置。作为示例而非限制,用户可以共享对象并指定只有同一城市中的用户可以访问或查看该对象。作为另一个示例而非限制,第一用户可以共享对象,并且指定只有当第一用户在特定位置时,该对象才对第二用户是可见的。如果第一用户离开该特定位置,则对象可不再对第二用户是可见的。作为另一示例而非限制,第一用户可以指定对象仅对在距第一用户的阈值距离内的第二用户可见。如果第一用户随后改变位置,则原始具有对对象的访问权的第二用户可能会失去访问权,而新的一组第二用户可能会在他们到达第一用户的阈值距离内时获得访问权。

[0163] 在特定实施例中,社交网络系统160或助理系统140可具有可将用户的个人或生物信息用作输入以用于用户认证或体验个性化目的的功能。用户可以选择利用这些功能来增强他们在在线社交网络上的体验。作为示例而不是作为限制,用户可以向社交网络系统160或助理系统140提供个人或生物信息。用户的隐私设置可以指定这样的信息仅可用于特定的进程(如认证),并且还指定这样的信息不能与任何第三方系统170共享,或者不能用于与社交网络系统160或助理系统140相关联的其他进程或应用。作为另一个示例而不是作为限制,社交网络系统160可以为用户提供向在线社交网络提供声纹记录的功能。作为示例而不是作为限制,如果用户希望利用在线社交网络的这一功能,则用户可以提供他或她自己声音的声音记录,以提供在线社交网络上的状态更新。声音输入的记录可以与用户的声纹进行比较,以确定用户说了什么词语。用户的隐私设置可以指定这种声音记录可以仅用于声音输入目的(例如,认证用户、发送声音消息、改进声音识别以便使用在线社交网络的声音操作特征),并且还指定这种声音记录不可以与任何第三方系统170共享,或者不可以被与社交网络系统160相关联的其他进程或应用使用。作为另一个示例而不是作为限制,社交网络系统160可以为用户提供向在线社交网络提供参考图像(例如,面部轮廓、视网膜扫描)的功能。在线社交网络可以将参考图像与稍后接收的图像输入进行比较(例如,用于认证用户,在照片中标记用户)。用户的隐私设置可以指定这种图像仅可用于有限的目的(例如,认证、在照片中标记用户),并且还指定这种图像不能与任何第三方系统170共享,或者不能被与社交网络系统160相关联的其他进程或应用使用。

[0164] 系统和方法

[0165] 图10示出了示例计算机系统1000。在特定实施例中,一个或多个计算机系统1000执行本文描述或示出的一个或多个方法的一个或多个步骤。在特定实施例中,一个或多个计算机系统1000提供本文描述或示出的功能。在特定实施例中,在一个或多个计算机系统1000上运行的软件执行本文描述或示出的一个或多个方法的一个或多个步骤,或者提供本文描述或示出的功能。特定实施例包括一个或多个计算机系统1000

的一个或更多个部分。在本文,在适当的情况下,对计算机系统的引用可以包括计算设备,反之亦然。此外,在适当的情况下,对计算机系统的引用可以包括一个或更多个计算机系统。

[0166] 本公开设想任何合适数量的计算机系统1000。本公开设想了计算机系统1000采取任何合适的物理形式。作为示例而不是作为限制,计算机系统1000可以是嵌入式计算机系统、片上系统(SOC)、单板计算机系统(SBC)(诸如例如,模块上计算机(COM)或模块上系统(SOM))、台式计算机系统、膝上型或笔记本计算机系统、交互式信息亭、大型机、计算机系统网状网、移动电话、个人数字助理(PDA)、服务器、平板计算机系统、或者这些中的两个或更多个的组合。在适当的情况下,计算机系统1000可以包括一个或更多个计算机系统1000;是整体式的或分布式的;跨越多个定位;跨越多台机器;跨越多个数据中心;或者驻留在云中,云可以包括在一个或更多个网络中的一个或更多个云组件。在适当的情况下,一个或更多个计算机系统1000可以在没有实质性空间或时间限制的情况下执行本文描述或示出的一个或更多个方法的一个或更多个步骤。作为示例而不是作为限制,一个或更多个计算机系统1000可以实时地或以批处理模式来执行本文描述或示出的一个或更多个方法的一个或更多个步骤。在适当的情况下,一个或更多个计算机系统1000可以在不同的时间或在不同的定位处执行本文描述或示出的一个或更多个方法的一个或更多个步骤。

[0167] 在特定实施例中,计算机系统1000包括处理器1002、存储器1004、存储装置1006、输入/输出(I/O)接口1008、通信接口1010和总线1012。尽管本公开描述并示出了具有在特定布置中的特定数量的特定组件的特定计算机系统,但是本公开设想了具有在任何合适布置中的任何合适数量的任何合适组件的任何合适的计算机系统。

[0168] 在特定实施例中,处理器1002包括用于执行指令(例如构成计算机程序的那些指令)的硬件。作为示例而不是作为限制,为了执行指令,处理器1002可以从内部寄存器、内部高速缓存、存储器1004或存储装置1006中检索(或取回)指令;将他们解码并执行它们;以及然后将一个或更多个结果写到内部寄存器、内部高速缓存、存储器1004或存储装置1006。在特定实施例中,处理器1002可以包括用于数据、指令或地址的一个或更多个内部高速缓存。在适当的情况下,本公开设想了处理器1002包括任何合适数量的任何合适的内部高速缓存。作为示例而不是作为限制,处理器1002可以包括一个或更多个指令高速缓存、一个或更多个数据高速缓存、以及一个或更多个转译后备缓冲区(TLB)。在指令高速缓存中的指令可以是在存储器1004或存储装置1006中的指令的副本,并且指令高速缓存可以加速处理器1002对那些指令的检索。在数据高速缓存中的数据可以是:在存储器1004或存储装置1006中的数据副本,用于使在处理器1002处执行的指令进行操作;在处理器1002处执行的先前指令的结果,用于由在处理器1002处执行的后续指令访问或者用于写到存储器1004或存储装置1006;或其他合适的数据。数据高速缓存可以加速由处理器1002进行的读或写操作。TLB可以加速关于处理器1002的虚拟地址转译。在特定实施例中,处理器1002可以包括用于数据、指令或地址的一个或更多个内部寄存器。在适当的情况下,本公开设想了处理器1002包括任何合适数量的任何合适的内部寄存器。在适当的情况下,处理器1002可以包括一个或更多个算术逻辑单元(ALU);是多核处理器;或者包括一个或更多个处理器1002。尽管本公开描述并示出了特定的处理器,但是本公开设想了任何合适的处理器。

[0169] 在特定实施例中,存储器1004包括用于存储用于使处理器1002执行的指令或用于

使处理器1002操作的数据的主存储器。作为示例而不是作为限制,计算机系统1000可以将指令从存储装置1006或另一个源(例如,另一个计算机系统1000)加载到存储器1004。处理器1002然后将指令从存储器1004加载到内部寄存器或内部高速缓存。为了执行指令,处理器1002可以从内部寄存器或内部高速缓存中检索指令并将它们解码。在指令的执行期间或之后,处理器1002可以将一个或多个结果(其可以是中间结果或最终结果)写到内部寄存器或内部高速缓存。处理器1002然后将这些结果中的一个或多个写到存储器1004。在特定实施例中,处理器1002仅执行在一个或多个内部寄存器或内部高速缓存中或在存储器1004(而不是存储装置1006或其他地方)中的指令,并且仅对在一个或多个内部寄存器或内部高速缓存中或在存储器1004(而不是存储装置1006或其他地方)中的数据进行操作。一个或多个存储器总线(其可以各自包括地址总线 and 数据总线)可以将处理器1002耦合到存储器1004。如下所述,总线1012可以包括一个或多个存储器总线。在特定实施例中,一个或多个存储器管理单元(MMU)驻留在处理器1002和存储器1004之间,并且便于由处理器1002请求的对存储器1004的访问。在特定实施例中,存储器1004包括随机存取存储器(RAM)。在适当的情况下,该RAM可以是易失性存储器。在适当的情况下,该RAM可以是动态RAM(DRAM)或静态RAM(SRAM)。此外,在适当的情况下,该RAM可以是单端口RAM或多端口RAM。本公开设想了任何合适的RAM。在适当的情况下,存储器1004可以包括一个或多个存储器1004。尽管本公开描述并示出了特定的存储器,但是本公开设想了任何合适的存储器。

[0170] 在特定实施例中,存储装置1006包括用于数据或指令的大容量存储装置。作为示例而不是作为限制,存储装置1006可以包括硬盘驱动器(HDD)、软盘驱动器、闪存、光盘、磁光盘、磁带或通用串行总线(USB)驱动器、或这些中的两个或多个的组合。在适当的情况下,存储装置1006可以包括可移动或不可移动(或固定)介质。在适当的情况下,存储装置1006可以在计算机系统1000的内部或外部。在特定实施例中,存储装置1006是非易失性固态存储器。在特定实施例中,存储装置1006包括只读存储器(ROM)。在适当的情况下,该ROM可以是掩模编程ROM、可编程ROM(PROM)、可擦除PROM(EPROM)、电可擦除PROM(EEPROM)、电可擦除ROM(EAROM)、或闪存、或这些中的两个或多个的组合。本公开设想了采用任何合适的物理形式的大容量存储装置1006。在适当的情况下,存储装置1006可以包括便于在处理器1002和存储装置1006之间的通信的一个或多个存储装置控制单元。在适当的情况下,存储装置1006可以包括一个或多个存储装置1006。尽管本公开描述并示出了特定的存储装置,但是本公开设想了任何合适的存储装置。

[0171] 在特定实施例中,I/O接口1008包括为在计算机系统1000和一个或多个I/O设备之间的通信提供一个或多个接口的硬件、软件或两者。在适当的情况下,计算机系统1000可以包括这些I/O设备中的一个或多个。这些I/O设备中的一个或多个可以实现在人和计算机系统1000之间的通信。作为示例而不是作为限制,I/O设备可以包括键盘、小键盘、麦克风、监视器、鼠标、打印机、扫描仪、扬声器、静态摄像机、触笔、平板计算机、触摸屏、跟踪球、视频摄像机、另一个合适的I/O设备、或这些中的两个或多个的组合。I/O设备可以包括一个或多个传感器。本公开设想了任何合适的I/O设备以及用于它们的任何合适的I/O接口1008。在适当的情况下,I/O接口1008可以包括使处理器1002能够驱动这些I/O设备中的一个或多个的一个或多个设备或软件驱动器。在适当的情况下,I/O接口1008可以包括一个或多个I/O接口1008。尽管本公开描述并示出了特定的I/O接口,但是本公开设想

了任何合适的I/O接口。

[0172] 在特定实施例中,通信接口1010包括提供用于在计算机系统1000和一个或多个其他计算机系统1000或一个或多个网络之间的通信(例如,基于包的通信)的一个或多个接口的硬件、软件或两者。作为示例而不是作为限制,通信接口1010可以包括用于与以太网或其他基于有线的网络进行通信的网络接口控制器(NIC)或网络适配器,或用于与无线网络(例如Wi-Fi网络)进行通信的无线NIC(WNIC)或无线适配器。本公开设想了任何合适的网络和用于它的任何合适的通信接口1010。作为示例而不是作为限制,计算机系统1000可以与自组织网络、个域网(PAN)、局域网(LAN)、广域网(WAN)、城域网(MAN)或互联网的一个或多个部分、或这些中的两个或多个的组合进行通信。这些网络中的一个或多个的一个或多个部分可以是有线的或无线的。作为示例,计算机系统1000可以与无线PAN(WPAN)(例如,蓝牙WPAN)、WI-FI网络、WI-MAX网络、蜂窝电话网络(例如,全球移动通信系统(GSM)网络)、或其他合适的无线网络、或这些中的两个或多个的组合进行通信。在适当的情况下,计算机系统1000可以包括用于这些网络中的任一个的任何合适的通信接口1010。在适当的情况下,通信接口1010可以包括一个或多个通信接口1010。尽管本公开描述并示出了特定的通信接口,但是本公开设想了任何合适的通信接口。

[0173] 在特定实施例中,总线1012包括将计算机系统1000的组件耦合到彼此的硬件、软件或两者。作为示例而不是作为限制,总线1012可以包括加速图形端口(AGP)或其他图形总线、扩展工业标准体系结构(EISA)总线、前端总线(FSB)、HYPERTRANSPORT(HT)互连、工业标准体系结构(ISA)总线、INFINIBAND互连、低引脚数(LPC)总线、存储器总线,微通道体系结构(MCA)总线、外围部件互连(PCI)总线、PCI-Express(PCIe)总线、串行高级技术附件(SATA)总线、视频电子标准协会本地(VLB)总线、或任何其他合适的总线、或这些中的两个或多个的组合。在适当的情况下,总线1012可以包括一个或多个总线1012。尽管本公开描述并示出了特定总线,但是本公开设想了任何合适的总线或互连。

[0174] 在本文,在适当的情况下,一个或多个计算机可读非暂时性存储介质可以包括一个或多个基于半导体的或其他集成电路(IC)(例如,现场可编程门阵列(FPGA)或专用IC(ASIC))、硬盘驱动器(HDD)、混合硬盘驱动器(HHD)、光盘、光盘驱动器(ODD)、磁光盘、磁光盘驱动器、软盘、软盘驱动器(FDD)、磁带、固态驱动器(SSD)、RAM驱动器、安全数字(SECURE DIGITAL)卡或驱动器、任何其他合适的计算机可读非暂时性存储介质、或这些中的两个或多个的任何合适组合。在适当的情况下,计算机可读非暂时性存储介质可以是易失性的、非易失性的或者易失性和非易失性的组合。

[0175] 其他杂项

[0176] 本文中,除非另有明确指示或通过上下文另有指示,否则“或”是包括一切的而非排除性的。因此在本文,除非另有明确指示或通过上下文另有指示,否则“A或B”意指“A、B或两者”。此外,除非另有明确指示或通过上下文另有指示,否则“和”既是联合的又是各自的。因此在本文,除非另有明确指示或通过上下文另有指示,否则“A和B”意指“A和B,联合地或各自地”。

[0177] 本公开的范围包括本领域中的普通技术人员将理解的对本文描述或示出的示例实施例的所有改变、替换、变化、变更和修改。本公开的范围不限于本文描述或示出的示例实施例。此外,尽管本公开将本文的相应实施例描述并示为包括特定的组件、元件、特征、功

能、操作或步骤,但是这些实施例中的任何一个可以包括本领域中的普通技术人员将理解的在本文任何地方描述或示出的任何组件、元件、特征、功能、操作或步骤的任何组合或置换。此外,在所附权利要求中对适合于、被布置成、能够、被配置成、实现来、可操作来、或操作来执行特定功能的装置或系统或装置或系统的组件的引用包括该装置、系统、组件,无论它或那个特定功能是否被激活、开启或解锁,只要该装置、系统或组件是这样被调整、被布置、使能够、被配置、被实现、可操作的、或操作的。此外,尽管本公开将特定实施例描述或示为提供特定优点,但是特定实施例可以提供这些优点中的一些、全部或不提供这些优点。

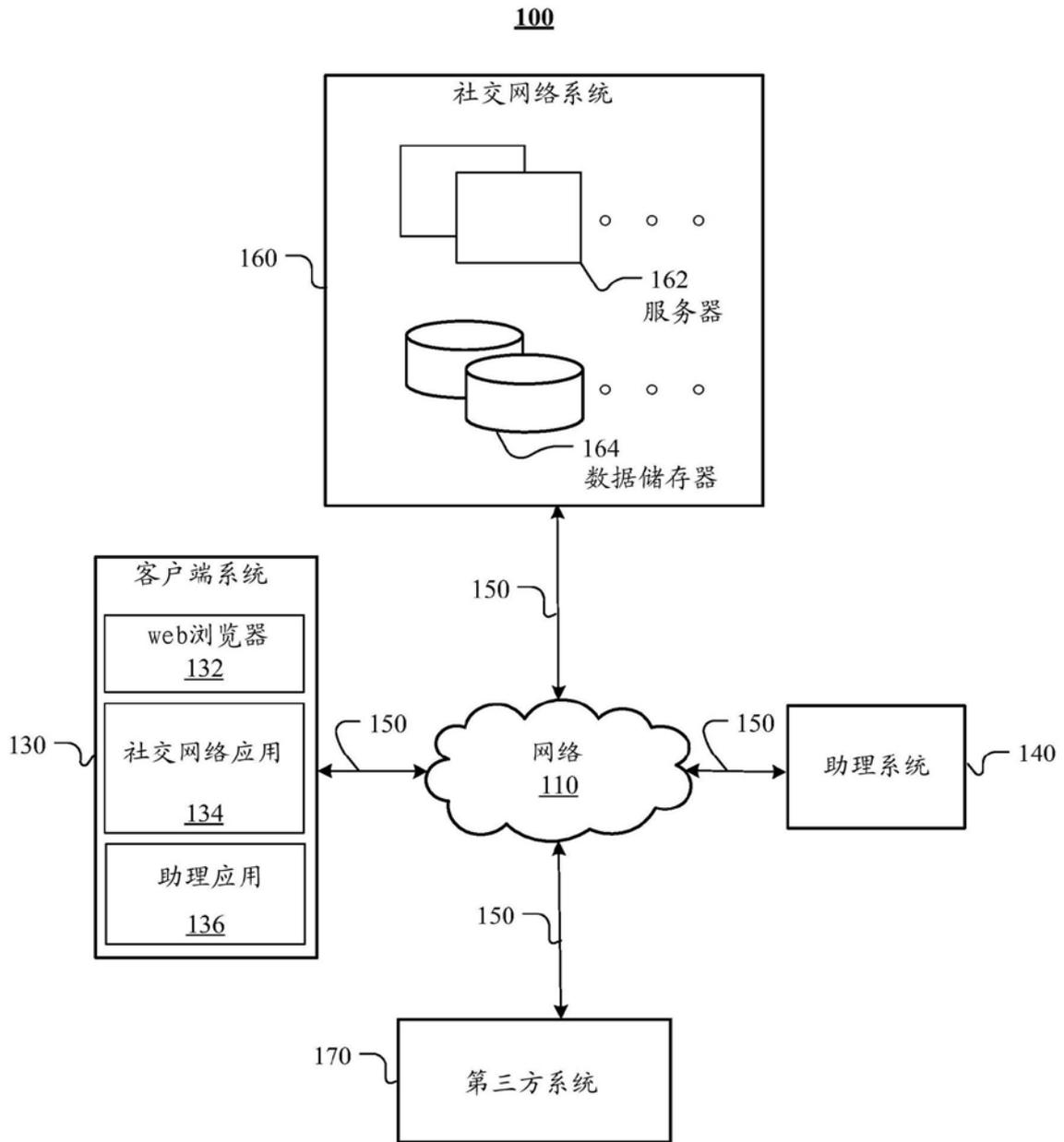


图1

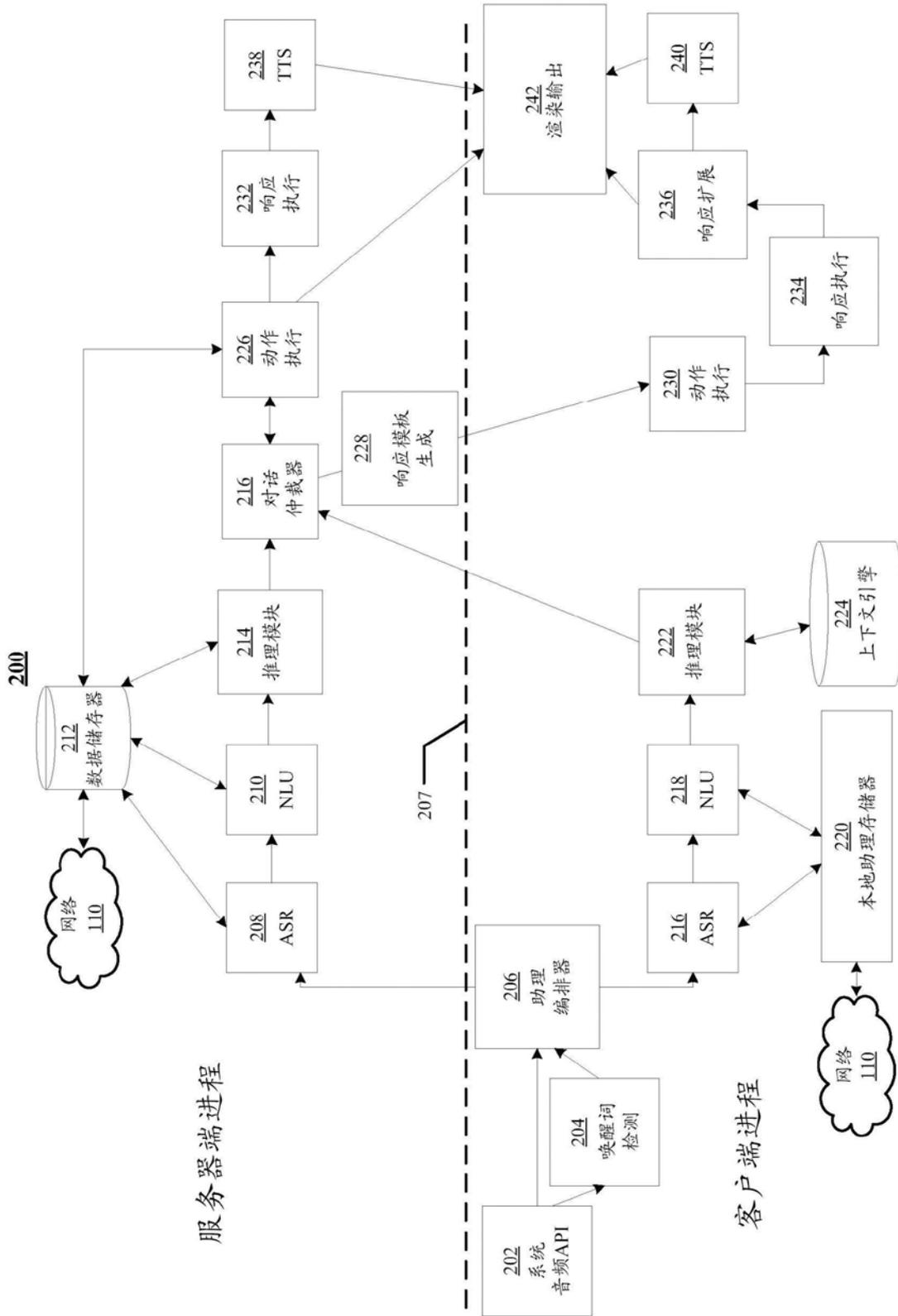


图2

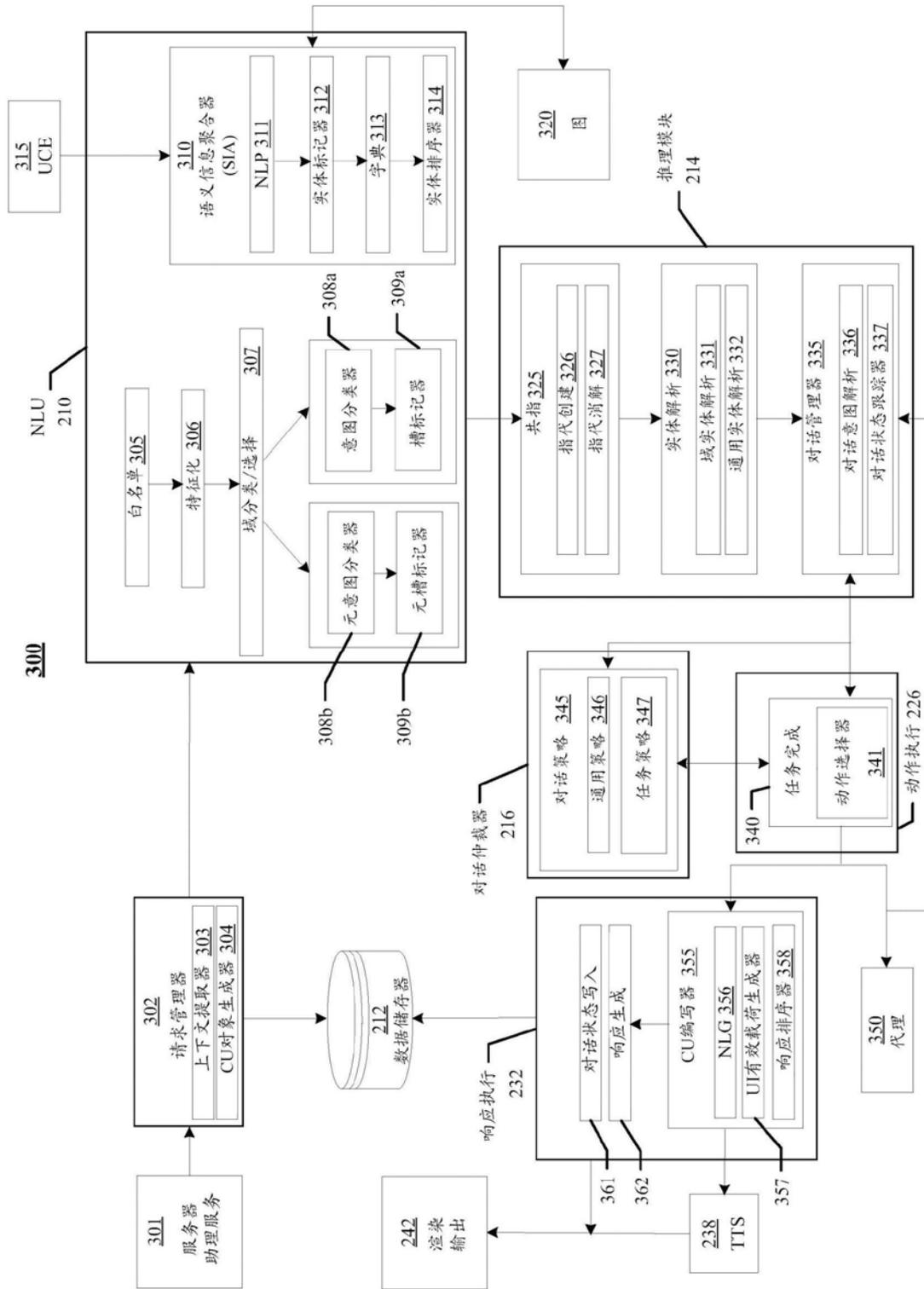


图3

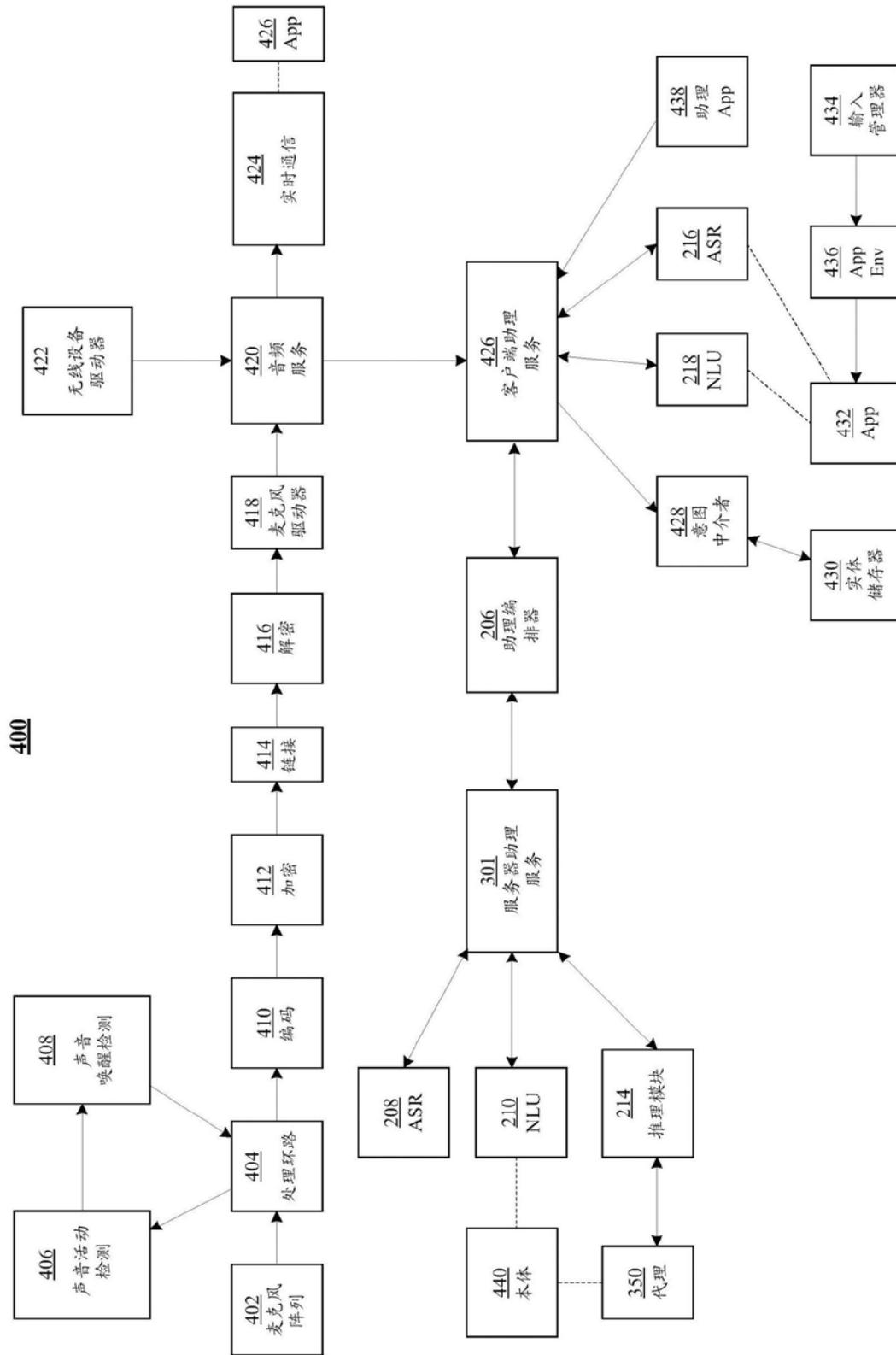


图4

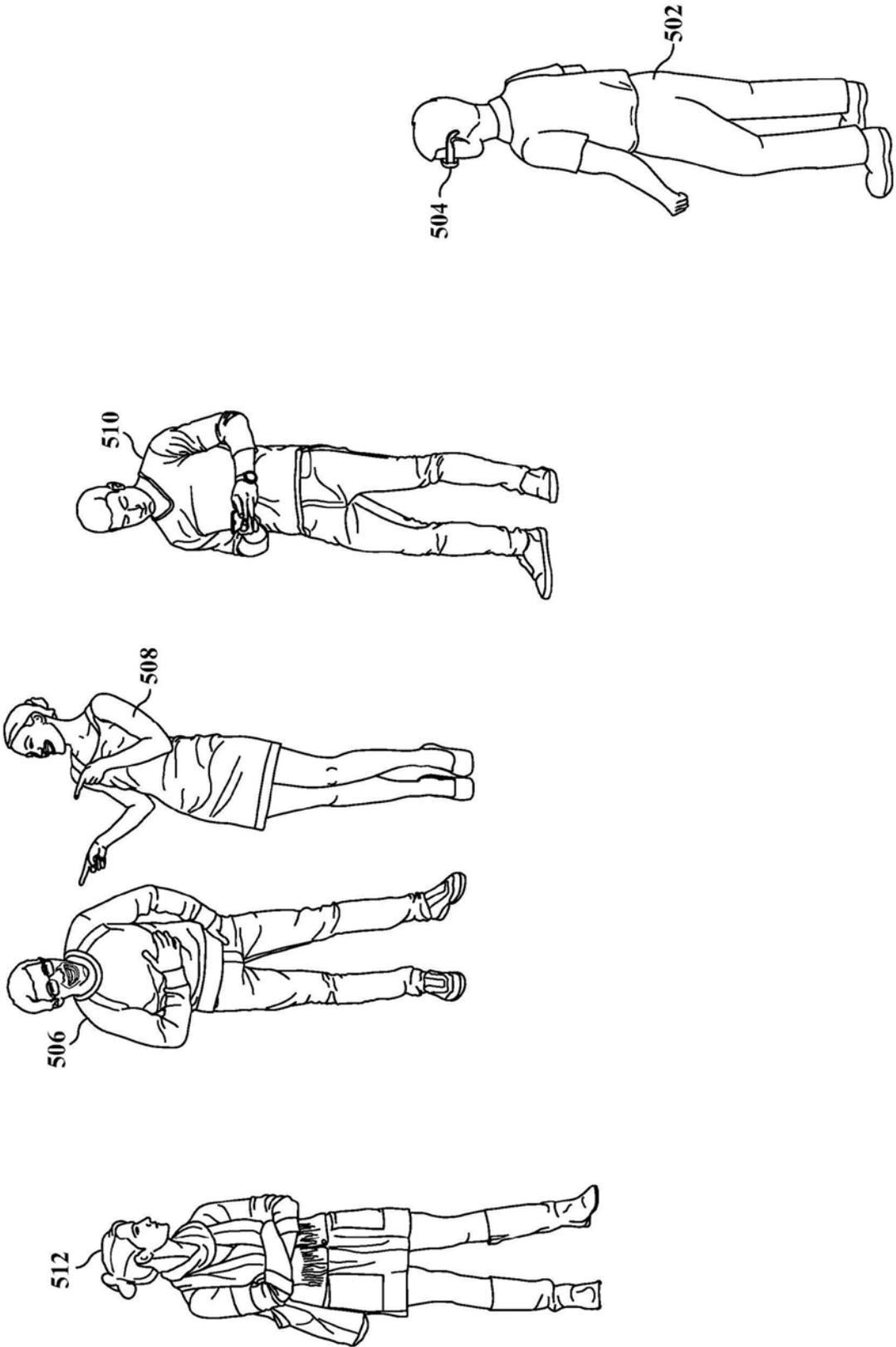


图5A

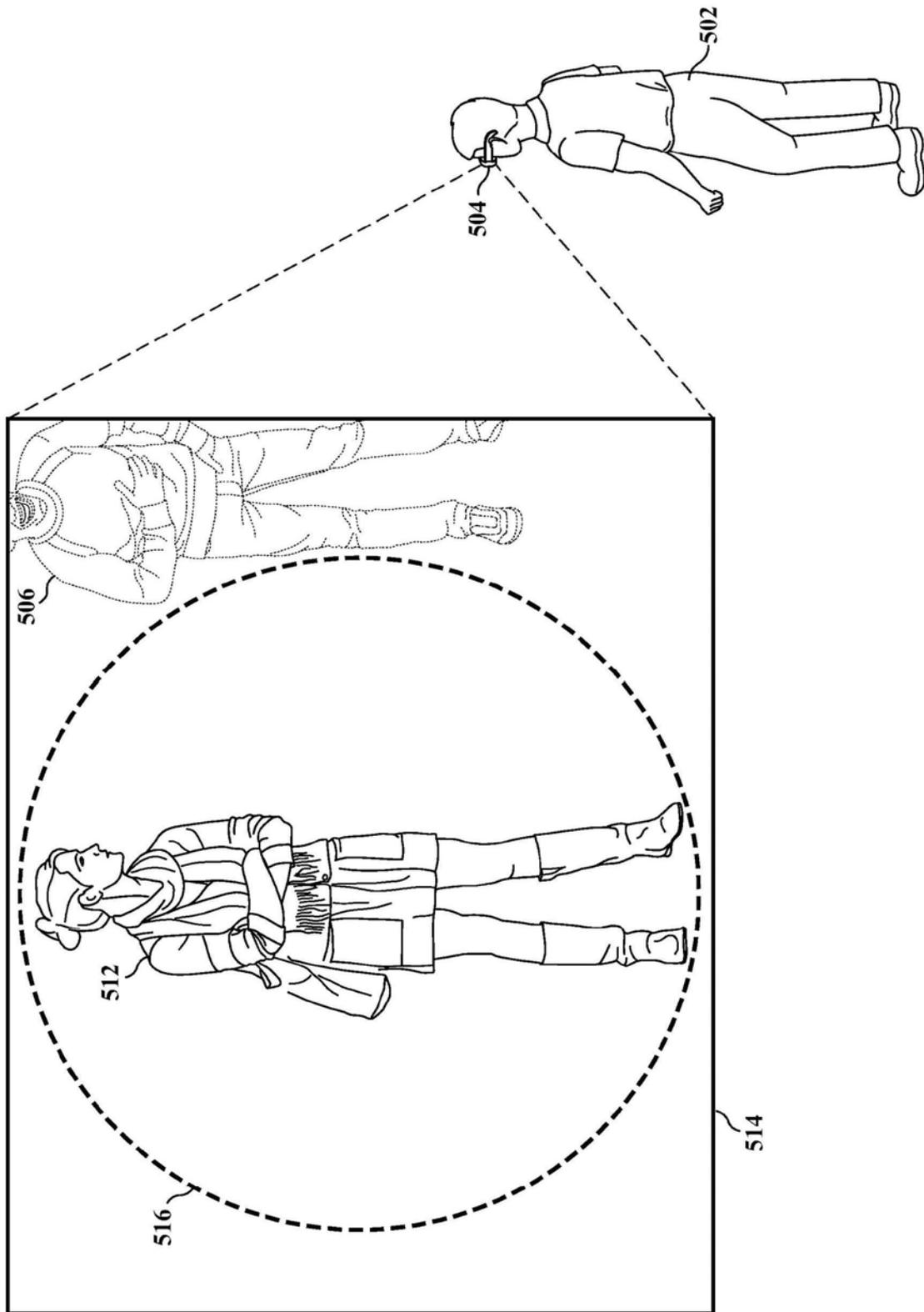


图5B

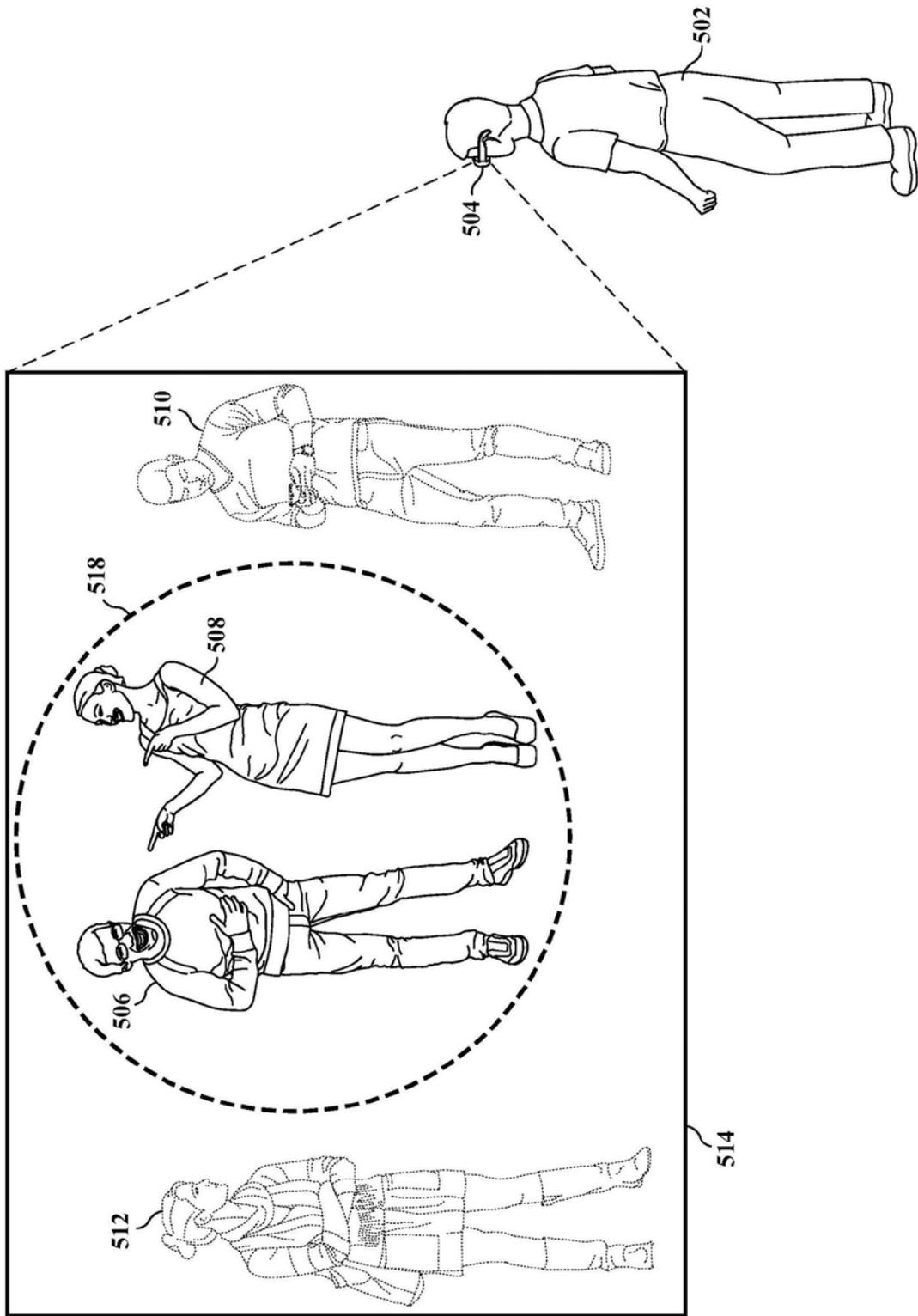


图5C

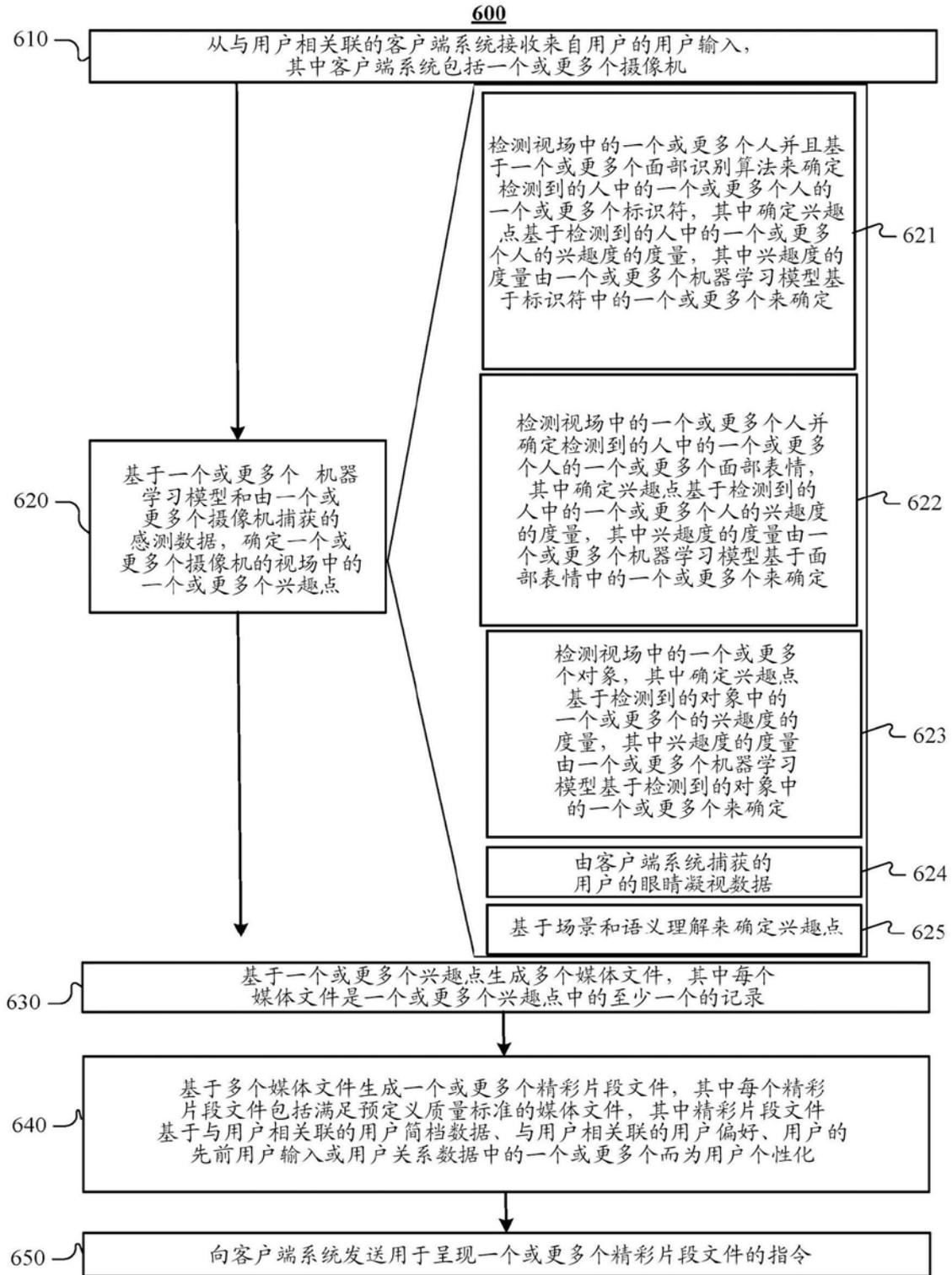


图6

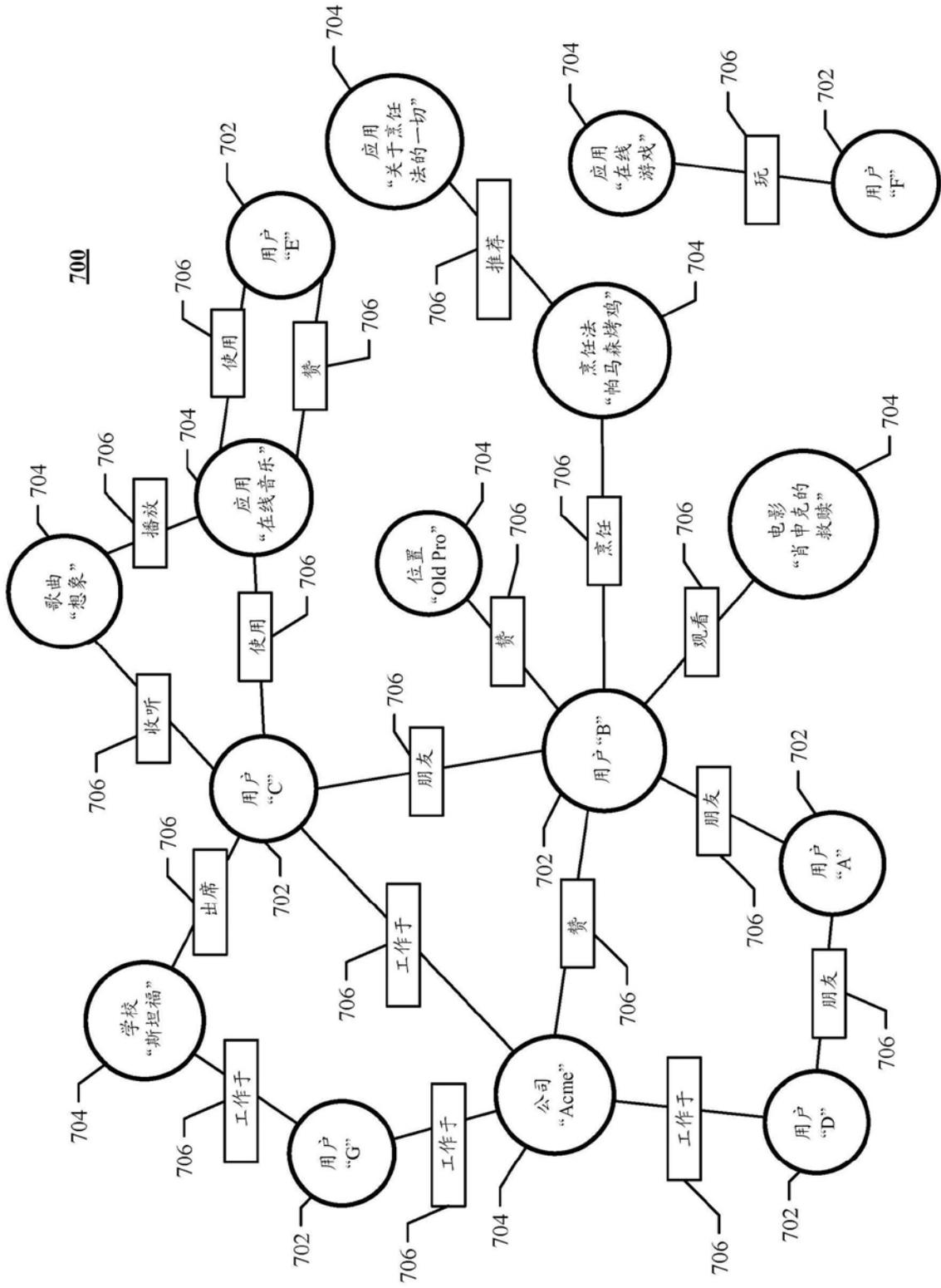


图7

800

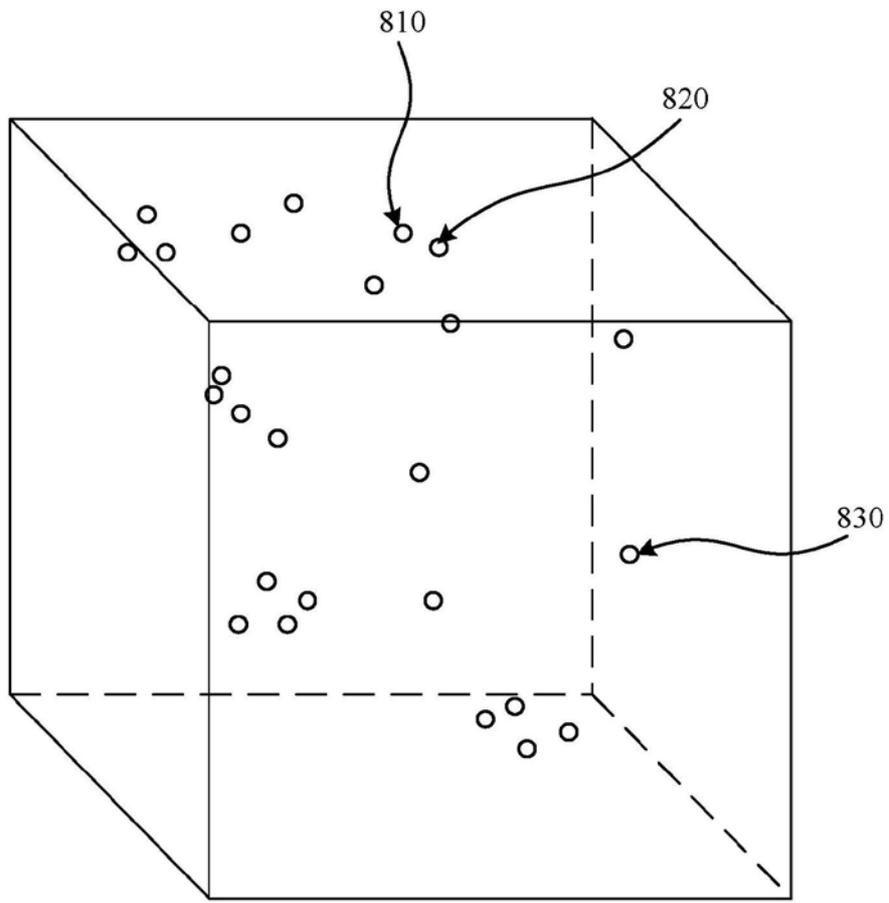


图8

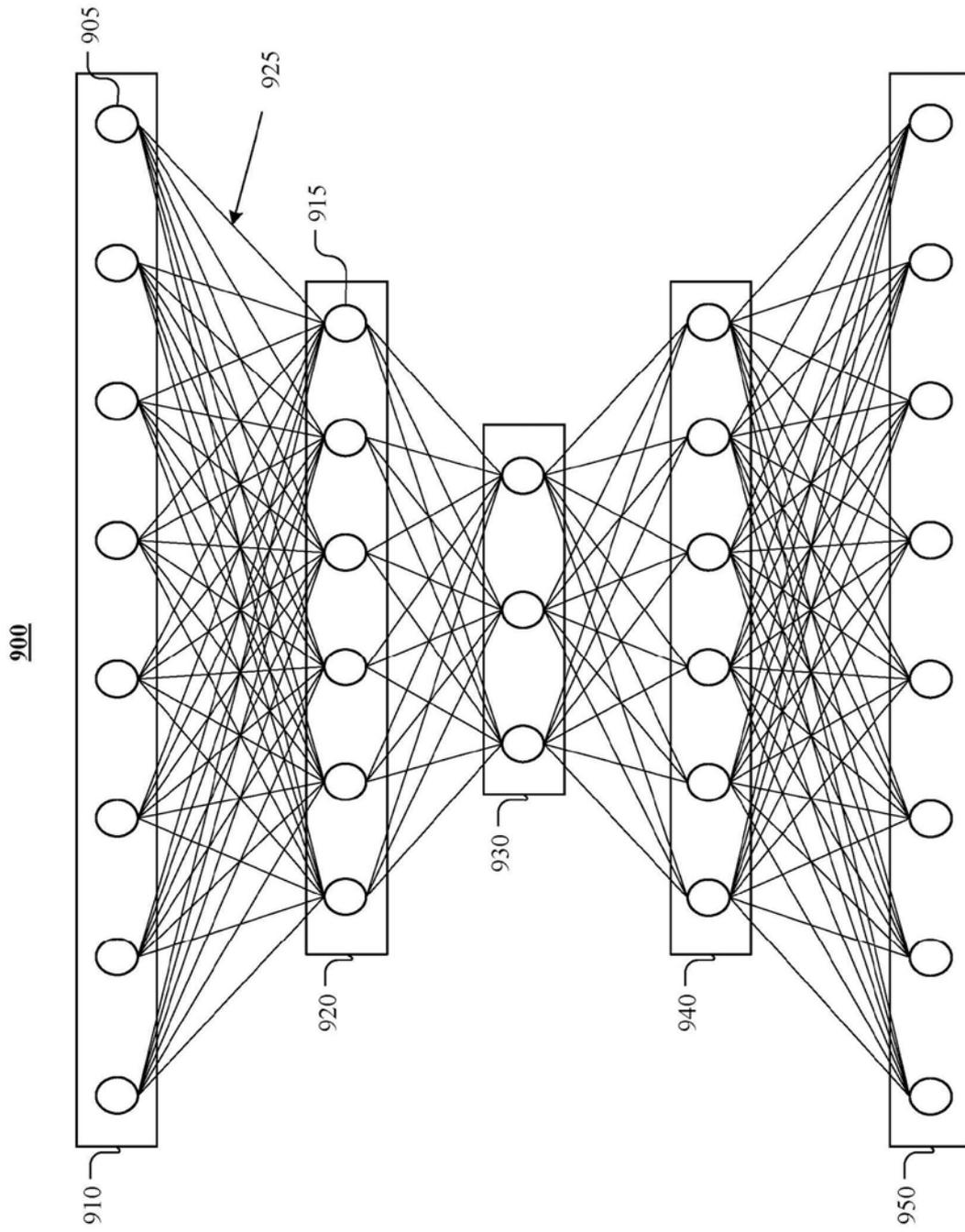


图9

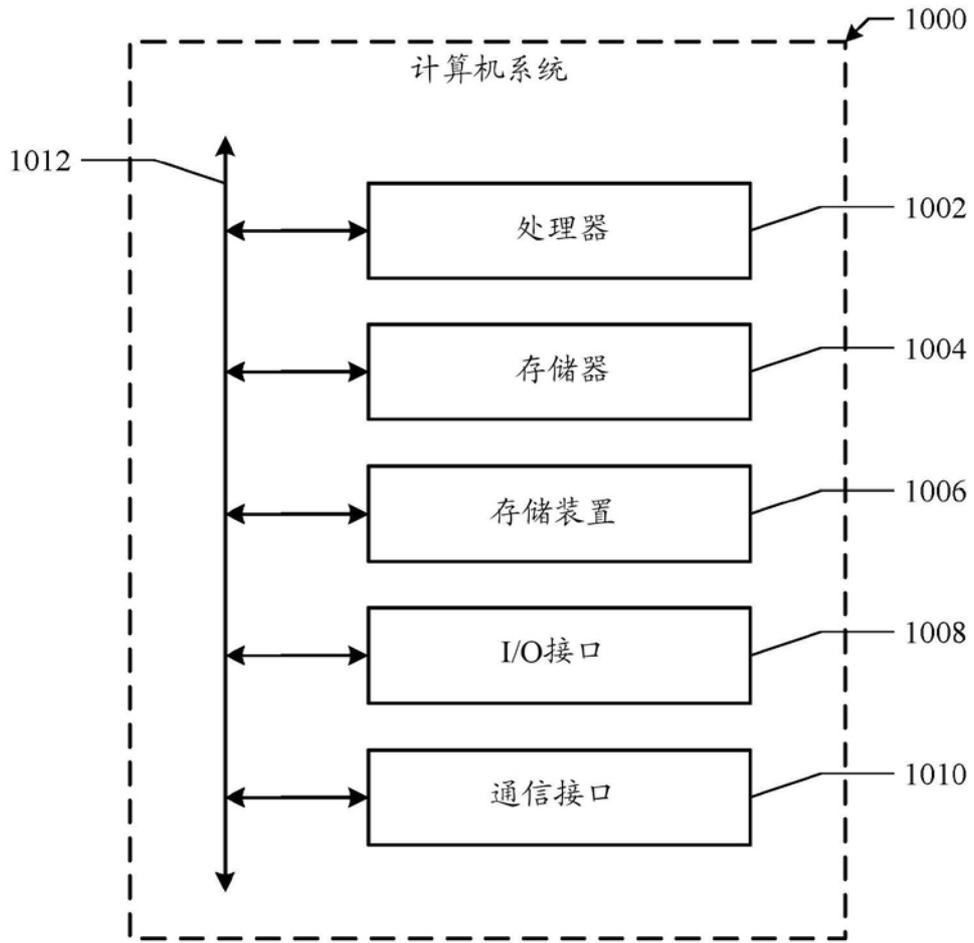


图10