



(12)发明专利申请

(10)申请公布号 CN 109564592 A

(43)申请公布日 2019.04.02

(21)申请号 201780050335.6

(22)申请日 2017.08.09

(30)优先权数据

15/238,660 2016.08.16 US

(85)PCT国际申请进入国家阶段日

2019.02.15

(86)PCT国际申请的申请数据

PCT/US2017/046051 2017.08.09

(87)PCT国际申请的公布数据

W02018/034904 EN 2018.02.22

(71)申请人 电子湾有限公司

地址 美国加利福尼亚州

(72)发明人 布拉多克·加斯基尔

阿迪·圭拉·哈维维

(74)专利代理机构 中科专利商标代理有限责任公司 11021

代理人 倪斌

(51)Int.Cl.

G06F 17/28(2006.01)

G06N 3/04(2006.01)

G06N 5/02(2006.01)

G10L 15/22(2006.01)

G06Q 30/02(2006.01)

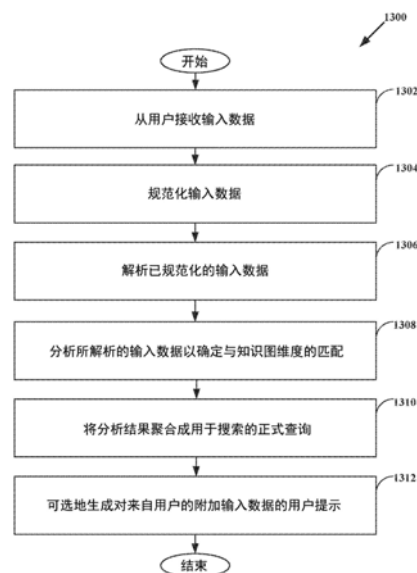
权利要求书2页 说明书24页 附图15页

(54)发明名称

在多层对话中生成接下来的用户提示

(57)摘要

用于在多层交互式对话中生成对来自用户的其他数据的提示的系统和方法。实施例通过用户输入及机器生成提示的已处理序列改善了对电子市场中可供购买的最相关项目的搜索。可以基于用户查询是否已经充分指明、用户意图是否模糊或搜索任务是否已经变化来选择性地生成问题类型提示、验证陈述类型提示和推荐类型提示。对新的主要对象的检测表示搜索任务发生变化。维护提示与用户回复之间的上下文关联,但是搜索任务变化导致先前的上下文数据被忽略。提示可基于知识图中的数据元素关联强度值、相对数据元素位置和深度而指向未指定的知识图维度,并且遵循数据元素知识图维度类型的预定顺序来生成。



1. 一种用于在多轮对话中生成对附加自然语言输入的提示的方法,所述方法包括:
接收知识图中的维度与对用户查询数据的分析的结果之间的已排名匹配,所述知识图维度至少包括类别、属性和属性值各一个,并且所述结果包括用户感兴趣的主要对象、用户意图和相关参数;
搜索库存并且将搜索结果并入到所述知识图中;
确定是否实现所述分析的结果与直接或间接链接到所述主要对象的知识图维度之间的预定的足够级别的匹配;以及
如果已经实现了所述足够级别的匹配,则基于所述搜索结果来生成并输出推荐类型提示。
2. 根据权利要求1所述的方法,其中,对话中的提示的数量受限于预定的最大值,所述预定的最大值是通过将常识知识并入到所述分析中以更好地推断用户意图来最小化的。
3. 根据权利要求1所述的方法,其中,所述推荐类型提示推荐电子市场中待售的至少一个项目。
4. 根据权利要求1所述的方法,还包括:如果尚未实现所述足够级别的匹配但是已经实现近乎足够级别的匹配,则替代地生成并输出提供多个推荐的不同的推荐类型提示,其中所述多个推荐中的每个推荐跨越至少一个未指定的已链接知识图维度。
5. 根据权利要求1所述的方法,其中,所述推荐类型提示包括至少一个表征电子市场中待售的项目的图像。
6. 根据权利要求1所述的方法,其中,所述推荐类型提示包括文本输出和口语输出中的一项。
7. 根据权利要求1所述的方法,还包括:如果在所述对话期间改变所述主要对象,则替代地生成并输出关于所述主要对象的验证陈述类型提示,并且忽略来自之前的搜索任务的上下文数据,其中,改变所述主要对象表示搜索任务的改变。
8. 一种计算机可读存储介质,所述计算机可读存储介质中嵌入有指令集,所述指令集当由计算机的一个或多个处理器执行时,使所述计算机执行用于在多轮对话中生成对附加自然语言输入的提示的下列操作:
接收知识图中的维度与对用户查询数据的分析的结果之间的已排名匹配,所述知识图维度至少包括类别、属性和属性值各一个,并且所述结果包括用户感兴趣的主要对象、用户意图和相关参数;
搜索库存并且将搜索结果并入到所述知识图中;
确定是否实现所述分析的结果与直接或间接链接到所述主要对象的知识图维度之间的预定的足够级别的匹配;以及
如果已经实现了所述足够级别的匹配,则基于所述搜索结果来生成并输出推荐类型提示。
9. 根据权利要求8所述的介质,其中,对话中的提示的数量受限于预定的最大值,所述预定的最大值是通过将常识知识并入到所述分析中以更好地推断用户意图来最小化的。
10. 根据权利要求8所述的介质,其中,所述推荐类型提示推荐电子市场中待售的至少一个项目。
11. 根据权利要求8所述的介质,还包括:如果尚未实现所述足够级别的匹配但是已经

实现近乎足够级别的匹配,则替代地生成并输出提供多个推荐的不同的推荐类型提示,其中所述多个推荐中的每个推荐跨越至少一个未指定的已链接知识图维度。

12.根据权利要求8所述的介质,其中,所述推荐类型提示包括至少一个表征电子市场中待售的项目的图像。

13.根据权利要求8所述的介质,其中,所述推荐类型提示包括文本输出和口语输出中的一项。

14.根据权利要求8所述的介质,还包括:如果在所述对话期间改变所述主要对象,则替代地生成并输出关于所述主要对象的验证陈述类型提示,并且忽略来自之前的搜索任务的上下文数据,其中,改变所述主要对象表示搜索任务的改变。

15.一种用于在多轮对话中生成对附加自然语言输入的提示的系统,所述系统包括:

自然语言理解组件,被配置为提供知识图中的维度与对用户查询数据的分析的结果之间的已排名匹配,所述知识图维度至少包括类别、属性和属性值各一个,并且所述结果包括用户感兴趣的主要对象、用户意图和相关参数;

搜索组件,被配置为搜索库存并且将搜索结果并入到所述知识图中;

对话管理器组件,被配置为确定是否实现所述分析的结果与直接或间接链接到所述主要对象的知识图维度之间的预定的足够级别的匹配;以及

如果已经实现了所述足够级别的匹配,则由所述对话管理器组件基于所述搜索结果来生成并输出推荐类型提示。

16.根据权利要求15所述的系统,其中,对话中的提示的数量受限于预定的最大值,所述预定的最大值是通过将常识知识并入到所述分析中以更好地推断用户意图来最小化的。

17.根据权利要求15所述的系统,其中,所述推荐类型提示推荐电子市场中待售的至少一个项目。

18.根据权利要求15所述的系统,还包括:如果尚未实现所述足够级别的匹配但是已经实现近乎足够级别的匹配,则替代地由所述对话管理器组件生成并输出提供多个推荐的不同的推荐类型提示,其中所述多个推荐中的每个推荐跨越至少一个未指定的已链接知识图维度。

19.根据权利要求15所述的系统,其中,所述推荐类型提示包括至少一个表征电子市场中待售的项目的图像。

20.根据权利要求15所述的系统,还包括:如果在所述对话期间改变所述主要对象,则替代地由所述对话管理器组件生成并输出关于所述主要对象的验证陈述类型提示,并且忽略来自之前的搜索任务的上下文数据,其中,改变所述主要对象表示搜索任务的改变。

21.一种承载指令集的机器可读介质,所述指令集当由计算机的一个或多个处理器执行时,使所述计算机执行根据权利要求1至7中任一项所述的方法。

在多轮对话中生成接下来的用户提示

[0001] 相关申请的交叉引用

[0002] 该国际申请要求于2016年8月16日提交的题为“Generating Next User Prompts In An Intelligent Online Personal Assistant Multi-Turn Dialog”的美国专利申请序列号15/238,660的优先权,通过引用将其全文并入本文。

背景技术

[0003] 传统搜索是没有人情味的。人们不能以标准的语言与传统的浏览引擎对话。传统搜索是耗时的,存在太多的选择并且可能浪费很多时间来浏览结果的页面。受制于传统工具的技术限制,用户难以传达意图,例如用户不能共享产品的照片以帮助搜索。随着选择激增至数十亿的在线项目,比较搜索变得比以往任何时候都更加重要,而目前的解决方案并不是针对这种规模而设计的。通常显示不相关的结果,而不会出现最佳结果。传统形式的比较搜索(搜索+细化+浏览)不再有用。

发明内容

[0004] 在一个示例中,智能个人助理系统包括可扩缩的人工智能(AI),其渗透到现有的消息传递平台的基本结构(fabric)中以提供智能在线个人助理(或“机器人”)。该系统可以利用现有的库存(inventory)和经组织的数据库,以在人类用户和智能在线个人助理之间的预测性多轮通信中提供智能的个性化答案。智能个人助理系统的一个示例包括知识图。机器学习组件可以连续地识别用户意图并从用户意图中学习,从而随着时间的推移增强用户身份和理解能力。由此,所提供的用户体验是鼓舞人心的、直观的、独特的,并且可以关注于某些年龄组(例如,千禧一代)的使用和行为模式。

附图说明

[0005] 附图通过示例而非限制的方式一般性地示出本文档中讨论的各种实施例。为了更容易地识别对任何特定元素或动作的讨论,附图标记中的最高有效的一个或多个数字指的是首先引入该元素的图号。

[0006] 图1示出了根据一些示例实施例的联网系统。

[0007] 图2示出了根据一些示例实施例的智能个人助理系统的总体架构。

[0008] 图3示出了根据一些示例实施例的语音识别组件的组件。

[0009] 图4示出了代表性软件架构软件架构,其可以结合本文中所描述的各种硬件架构来使用。

[0010] 图5示出了根据一些示例实施例的机器的组件,所述机器能够从机器可读介质(例如,计算机可读存储介质)中读取指令并执行本文中所讨论的方法中的任何一种或多种。

[0011] 图6示出了根据一些示例实施例的可以在其中部署智能在线个人助理的示例环境。

[0012] 图7示出了根据一些示例实施例的智能个人助理系统处理自然语言用户输入以生

成电子市场中的项目推荐的概述。

[0013] 图8示出了根据一些示例实施例的自然语言理解 (NLU) 组件、该组件的子组件以及与该组件交互的其他组件。

[0014] 图9示出了根据一些示例实施例的各种分析的结果。

[0015] 图10示出了根据一些示例实施例的知识图。

[0016] 图11A和图11B示出了根据一些示例实施例的具有项目类别、一些项目属性和一些项目属性值的简明知识图。

[0017] 图12示出了根据一些示例实施例的智能个人助理系统处理自然语言用户输入以生成建议性提示的概述。

[0018] 图13示出了根据一些示例实施例的用于处理自然语言用户输入以生成项目推荐的方法的流程图。

具体实施方式

[0019] 该上下文中的“载波信号”指能够存储、编码或携带用于由机器执行的指令的任何无形介质,并且包括数字或模拟通信信号或用于促进这种指令的传送的其它无形介质。可以经由网络接口设备使用传输介质并使用多种已知传输协议中的任何一种来通过网络发送或接收指令。

[0020] 该上下文中的“客户端设备”指的是接口连接到通信网络以从一个或多个服务器系统或其他客户端设备获得资源的任何机器。客户端设备可以是:移动电话、台式计算机、膝上型计算机、便携式数字助理 (PDA)、智能电话、平板计算机、超级本、上网本、笔记本电脑、多处理器系统、基于微处理器或可编程的消费电子产品、游戏机、机顶盒或用户可以用来接入联网系统102的任何其他通信设备,但不限于此。

[0021] 该上下文中的“通信网络”指网络的一个或多个部分,该网络可以是自组织 (ad hoc) 网络、内联网、外联网、虚拟专用网 (VPN)、局域网 (LAN)、无线LAN (WLAN)、广域网 (WAN)、无线WAN (WWAN)、城域网 (MAN)、互联网、互联网的一部分、公共交换电话网络 (PSTN) 的一部分、普通老式电话服务 (POTS) 网络、蜂窝电话网络、无线网络、**Wi-Fi®**网络、另一类型的网络、或两个或更多个这样的网络的组合。例如,网络或网络的一部分可以包括无线或蜂窝网络,并且耦接可以是码分多址 (CDMA) 连接、全球移动通信系统 (GSM) 连接或其他类型的蜂窝或无线耦接。在该示例中,耦接可以实现各种类型的数据传输技术中的任何一种,例如单载波无线电传输技术 (1xRTT)、演进数据优化 (EVDO) 技术、通用分组无线电服务 (GPRS) 技术、GSM演进增强数据速率 (EDGE) 技术、包括3G的第三代合作伙伴计划 (3GPP)、第四代无线 (4G) 网络、通用移动通信系统 (UMTS)、高速分组接入 (HSPA)、全球微波接入互操作性 (WiMAX)、长期演进 (LTE) 标准、由各种标准设置组织定义的其他标准、其他远程协议或其他数据传输技术。

[0022] 该上下文中的“组件”指具有由功能或子例程调用、分支点、应用程序接口 (API)、或提供特定处理或控制功能的划分或模块化的其他技术限定的边界的设备、物理实体或逻辑。组件可以经由其接口与其他组件进行组合,以实施机器处理。组件可以是封装功能性硬件单元,被设计与其他组件或程序中通常执行相关功能中的特定功能的一部分一起使用。组件可以构成软件组件 (例如,机器可读介质上体现的代码) 或硬件组件。“硬件组件”是能

够执行特定操作的有形单元,并且可以按照特定的物理方式来配置或布置。在各种示例实施例中,一个或多个计算机系统(例如独立的计算机系统、客户端计算机系统或服务器计算机系统)或者计算机系统的一个或多个硬件组件(例如处理器或处理器组)可以由软件(例如应用或应用部分)配置为进行操作以执行本文中描述的特定操作的硬件组件。也可以机械地、电性地或用其任何合适的组合来实现硬件组件。例如,硬件组件可以包括永久地被配置为执行特定操作的专用电路或逻辑。硬件组件可以是专用处理器,例如现场可编程门阵列(FPGA)或专用集成电路(ASIC)。硬件组件还可以包括由软件临时配置为执行特定操作的可编程逻辑或电路。例如,硬件组件可以包括由通用处理器或其他可编程处理器执行的软件。一旦由这样的软件配置,硬件组件就变成特定的机器(或机器的特定组件),其被专门定制用于执行所配置的功能,而不再是通用处理器。应理解,以机械方式、以专用和永久配置的电路或以临时配置的电路(例如由软件配置)实现硬件组件的决定可以出于成本和时间的考虑。因此,短语“硬件组件”(或“硬件实现的组件”)应理解为涵盖有形实体,其是在物理上构造、永久配置(例如硬线连接)或临时配置(例如编程)为以特定方式操作或执行本文中描述的特定操作的实体。考虑临时配置(例如编程)硬件组件的实施例,无需在任一时刻配置或实例化硬件组件中的每一个。例如,在硬件组件包括被软件配置成为专用处理器的通用处理器的情况下,通用处理器可以在不同时间被配置为分别不同的专用处理器(例如包括不同的硬件组件)。因此,软件将特定的一个或多个处理器例如配置为在一个时刻构成特定硬件组件并且在不同的时刻构成不同的硬件组件。硬件组件可以向其他硬件组件提供信息并从其他硬件组件接收信息。因此,所描述的硬件组件可以被看作通信地耦接。如果同时存在多个硬件组件,则可以通过两个或更多个硬件组件之间的信号传输(例如通过适当的电路和总线)实现通信。在多个硬件组件在不同时间被配置或实例化的实施例中,可以例如通过存储并获取多个硬件组件可访问的存储器结构中的信息来实现这样的硬件组件之间的通信。例如,一个硬件组件可以执行操作并在与其通信耦接的存储器设备中存储该操作的输出。另一硬件组件接着可以稍后访问存储器设备,以获取并处理所存储的输出。硬件组件还可以发起与输入或输出设备的通信,并且能够对资源(例如信息的集合)进行操作。此处描述的示例方法的各种操作可以至少部分地由临时配置(例如通过软件)或永久配置为执行相关操作的一个或多个处理器执行。无论是临时还是永久配置,这样的处理器可以构成进行操作以执行本文描述的一个或多个操作或功能的处理器实现的组件。如本文所使用的,“处理器实现的组件”指使用一个或更多个处理器实现的硬件组件。类似地,本文描述的方法可以至少部分地由处理器实现,其中特定处理器或多个处理器是硬件的示例。例如,方法的至少一些操作可以由一个或多个处理器或处理器实现的组件来执行。此外,一个或多个处理器还可操作以支持在“云计算”环境中或作为“软件即服务”(SaaS)执行有关操作。例如,操作中的至少一些可由计算机(作为包括处理器的机器的示例)组执行,这些操作可经由网络(例如互联网)并经由一个或多个适当接口(例如应用程序接口(API))来访问。某些操作的执行可以分布在处理器中,并不只驻留在单个机器内,而是部署在多个机器中。在一些示例实施例中,处理器或处理器实现的组件可以位于单个地理位置中(例如,在家庭环境、办公环境或服务器群中)。在其他示例实施例中,处理器或处理器实现的组件可以分布在多个地理位置中。

[0023] 该上下文中的“机器可读介质”是指能够暂时或永久地存储指令和数据的组件、设

备或其他有形介质,并且可以包括但不限于随机存取存储器(RAM)、只读存储器(ROM)、缓冲存储器、闪速存储器、光学介质、磁性介质、高速缓冲存储器、其它类型的存储设备(例如,可擦除可编程只读存储器(EEPROM))和/或其任何合适的组合。术语“机器可读介质”应被视为包括能够存储指令的单个介质或多个介质(例如,集中式或分布式数据库、或相关联的缓存和服务)。术语“机器可读介质”还将被视为包括能够存储被机器执行的指令(例如,代码)的任何介质或多个介质的组合,使得指令在被机器的一个或多个处理器执行时,使机器执行本文中所描述的方法中的任何一个或多个。因此,“机器可读存储介质”指单个存储装置或设备、以及包括多个存储装置或设备的“基于云”的存储系统或存储网络。术语“机器可读存储介质”不包括信号本身。机器可读介质包括机器可读存储介质和传输介质或载波信号。

[0024] 该上下文中的“处理器”指根据控制信号(例如,“命令”、“op代码”、“机器代码”等)操纵数据值并且产生被用来操作机器的对应输出信号的任何电路或虚拟电路(由在实际处理器上执行的逻辑仿真的物理电路)。处理器可以是例如中央处理单元(CPU)、精简指令集计算(RISC)处理器、复杂指令集计算(CISC)处理器、图形处理单元(GPU)、数字信号处理器(DSP)、专用集成电路(ASIC)、射频集成电路(RFIC)、或其任何组合。处理器还可以是具有两个或更多个独立处理器(有时也称为“核”)的多核处理器,其中该两个或更多个独立处理器可以同时执行指令。

[0025] 本专利文档的公开的一部分包含受到版权保护的内容。版权所有人不反对任何人对专利文档或专利公开(如其在专利和商标局中的专利文件或记录中出现的原样)进行复制再现,但是在其他情况下版权所有人保留所有版权。以下注意适用于下文描述的且形成本文档一部分的软件和数据:版权所有2016,eBay Inc.保留所有权利。

[0026] 以下描述包括体现本公开的示意性实施例的系统、方法、技术、指令序列和计算机程序产品。在下文的描述中,为了解释的目的,阐述了很多细节以提供对本发明主题的各种实施例的理解。然而,对本领域技术人员将显而易见的是,本发明主题的实施例可以在没有这些具体细节的情况下实施。一般地,不必详细示出众所周知的指令实例、协议、结构和技术。

[0027] 参考图1,示出了高级SaaS网络架构100的示例实施例。联网系统116经由网络110(例如,互联网或广域网(WAN))向客户端设备108提供服务器侧功能。在客户端设备108上托管并执行应用104的示例形式的web客户端102和编程客户端。联网系统116包括应用服务器122,应用服务器122进而托管智能个人助理系统106,智能个人助理系统106向访问联网系统116的应用104提供多种功能和服务。应用104还提供本文中所描述的多个接口,所述多个接口向客户端设备108的用户呈现跟踪和分析操作的输出。

[0028] 客户端设备108使用户能够访问联网系统116并与联网系统116交互。例如,用户向客户端设备108提供输入(例如,触摸屏输入或字母数字输入),并且经由网络110将输入传送给联网系统116。在这种情况下,响应于接收到来自用户的输入,联网系统116经由网络110将信息传送回给客户端设备108以呈现给用户。

[0029] 应用程序接口(API)服务器118和web服务器120分别耦接到应用服务器122,并分别向应用服务器122提供编程接口和web接口。应用服务器122托管包括组件或应用的智能个人助理系统106。应用服务器122进而被示为耦接到数据库服务器124,数据库服务器124促进对信息存储库(例如,数据库/云126)的访问。在示例实施例中,数据库/云126包括存储

由智能个人助理系统106访问和生成的信息的存储设备。

[0030] 另外,在第三方服务器112上执行的第三方应用114被示为具有经由通过应用程序接口(API)服务器118所提供的编程接口对联网系统116的编程访问。例如,第三方应用114使用从联网系统116获取的信息,可以支持由第三方托管的网站上的一个或多个特征或功能。

[0031] 现在将具体地转向由客户端设备108托管的应用,web客户端102可以经由web服务器120支持的web接口来访问各种系统(例如,智能个人助理系统106)。类似地,应用104(例如,“app”)经由应用程序接口(API)服务器118提供的编程接口来访问由智能个人助理系统106提供的各种服务和功能。例如,应用104可以是在客户端设备108上执行的“app”,例如iOS或Android OS应用,以使用户能够以离线方式访问联网系统116和在联网系统116上输入数据,并执行编程客户端应用104与联网系统116之间的批处理模式通信。

[0032] 此外,虽然图1所示出的SaaS网络架构100采用客户端-服务器架构,但是本发明主题当然不限于此种架构,而是可以同样良好地应用于例如分布式或对等架构系统。智能个人助理系统106还可以被实现为独立的软件程序,其不一定具有联网能力。

[0033] 图2是示出了根据一些示例实施例的智能个人助理系统106的总体架构的框图。具体地,智能个人助理系统106被示为包括前端组件202(FE),智能个人助理系统106通过该前端组件202(例如,通过网络110)与SaaS网络架构100内的其他系统进行通信。前端组件202可以与现有的消息传递系统的基本结构进行通信。如本文中所使用的,术语“消息传递基本结构”指的是可以为诸如Facebook消息器、微软Cortana和其他“机器人”之类的第三方平台提供支撑(power)的API和服务的集合。在一个示例中,消息传递基本结构可以支持允许用户与商业意图交互的在线商务生态系统。作为与智能个人助理或“机器人”的接口的一部分,前端组件202的输出可以在客户端设备(例如图1中的客户端设备108)的显示器中渲染。

[0034] 智能个人助理系统106的前端组件202耦接到针对前端的后端组件(BFF) 204,该BFF 204用于将前端组件202与人工智能框架128链接。人工智能框架128可以包括如下所讨论的若干组件。在各种组件之间交换的数据和每个组件的功能可以在一定程度上变化,这取决于具体实现。

[0035] 在智能个人助理系统106的一个示例中,AI协调器206协调人工智能框架128内部和外部的组件之间的通信。例如,AI协调器206的输入模态可以从计算机视觉组件208、语音识别组件210和文本规范化组件导出,该文本规范化组件可以形成语音识别组件210的一部分。计算机视觉组件208可以从视觉输入(例如,照片)中识别对象和属性。语音识别组件210可以将音频信号(例如,口语话语)转换成文本。例如,文本规范化组件可以操作以进行输入规范化,例如通过将表情符号渲染为文本来进行语言规范化。其他规范化是可能的,例如拼字正确(orthographic)规范化、外语规范化、交谈文本规范化等。

[0036] 人工智能框架128还包括自然语言理解或NLU组件214,其进行操作以提取用户意图和各种意图参数。从图8开始更详细地描述NLU组件214。

[0037] 人工智能框架128还包括对话管理器216,其进行操作以理解(例如,诸如搜索查询或话语之类的输入的)“特异性的完整性(completeness of specificity)”,并且决定下一个动作类型和相关参数(例如,“搜索”或“向用户请求其他信息”)。为方便起见,无论是文本、语音还是图像相关格式,本说明书中的所有用户输入都可以被称为“话语”。

[0038] 在一个示例中,对话管理器216与上下文管理器218以及自然语言生成(NLG)组件212相关联地操作。关于智能在线个人助理(或“机器人”)和助理的相关联的人工智能,上下文管理器218管理用户的上下文和通信。上下文管理器218保留用户交互的短期历史。如下所述,用户偏好的较长期历史可以保留在身份服务222中。例如,这些历史中的一个或二者中的数据条目可以包括给定输入、机器人交互或通信轮转的相关意图和所有参数以及所有相关结果。NLG组件212进行操作以用AI消息组成自然语言话语,以呈现给与智能机器人交互的用户。

[0039] 搜索组件220也被包括在人工智能框架128内。搜索组件220可以具有前端单元和后端单元。后端单元可以进行操作以管理项目或产品库存,并提供针对库存进行搜索、针对用户意图与意图参数的特定元组进行优化的功能。搜索组件220被设计为针对非常大的高质量库存每天全球服务数十亿次查询。搜索组件220可以适应文本或适应人工智能(AI)编码的语音和图像输入,并基于显式的和导出的查询意图为用户识别相关的库存项目。

[0040] 身份服务222组件进行操作以管理用户简档,例如用户属性形式的显式信息,例如“姓名”、“年龄”、“性别”、“地理位置”,以及诸如“信息馏出物”之类的形式的隐式信息,例如,“用户兴趣”或“类似的人物”等。人工智能框架128可以包括身份服务222的一部分或与身份服务222相关联地操作。身份服务222包括策略、API和服务的集合,其巧妙地集中所有用户信息,帮助人工智能框架128来具有对用户意图的“智能”见解。身份服务222可以保护在线零售商和用户免于欺诈或私人信息的恶意使用。

[0041] 本公开的身份服务222提供许多优点。身份服务222是包含用户身份和简档数据的单个中央存储库。它可以通过新的见解和更新来不断丰富用户简档。它使用帐户链接和身份联合来映射用户与公司、家庭、其他帐户(例如,核心帐户)、以及用户的人物和关系的社交图的关系。身份服务222形成丰富的通知系统,该通知系统在用户选择的时间和媒体上传送他们想要的所有信息且仅传送这样的信息。

[0042] 在一个示例中,身份服务222专注于统一清算所中用于搜索、AI、商品推销和机器学习模型的尽可能多的用户信息,以最大化每个组件向每个用户提供见解的能力。单个中央存储库以细致详细的方案包含用户身份和简档数据。在初始(onboarding)阶段,身份服务222通过机器人应用中的强制认证来填充用户简档和理解力。可以加载可从认证源(例如,社交媒体)获得的任何公共信息。在中间(sideboarding)阶段,身份服务222可以利用从公共来源收集的关于用户的信息、用户行为、交互以及用户告知AI的显式目的集(例如,购物任务、灵感、偏好)来扩充该简档。当用户与人工智能框架128交互时,身份服务222收集并推断更多关于用户的信息且存储显示数据和导出的信息,并更新其他统计推断的概率和估计。随着时间的推移,在简档丰富阶段,身份服务222还挖掘行为数据,例如点击、印象和浏览活动,以得到诸如品味、偏好和购物行业之类的导出的信息。在身份联合和帐户链接阶段中,当被传送或推断时,身份服务222更新用户的家庭、雇主、组、从属关系、社交图和其他帐户(包括共享帐户)。

[0043] 人工智能框架128的功能可以被分组为多个部分,例如决策部分和上下文部分。在一个示例中,决策部分包括AI协调器206、NLU组件214、对话管理器216、NLG组件212、计算机视觉组件208和语音识别组件210的操作。AI功能的上下文部分涉及与用户和所传达的意图(例如,针对给定的库存或其他)有关的参数(隐式和显式)。为了随时间的推移测量和改善

AI质量,可以使用样本查询(例如,开发集)来训练人工智能框架128,并且在不同的查询集(例如,评估集)上测试人工智能框架128,其中该两个集合可以通过人类的综合管理来开发。此外,可以在由经验丰富的综管理专家或人类品味营造者覆盖规则224限定的交易和交互流程上进行训练人工智能框架128。在人工智能框架128的各种组件内编码的流程和逻辑定义了智能助理基于所识别的用户意图进行的后续话语或呈现(例如,问题,结果集)。

[0044] 以上进一步参考了智能个人助理系统106中的智能在线个人助理或机器人的示例输入模态。智能个人助理系统106试图理解用户的意图(例如,针对性的搜索、比较、购物/浏览等)以及任何强制性参数(例如,产品、产品类别、项目等)和/或可选的参数(例如,诸如项目/产品的属性、场合之类的显式信息)以及隐式信息(例如,地理位置、个人偏好、年龄和性别等),并且用深思熟虑或“智能的”响应来回应用户。显式输入模态可以包括文本、语音和视觉输入,并且可以利用用户的隐式知识(例如,地理位置、先前的浏览历史等)来丰富。输出模态可以包括文本(诸如语音或自然语言语句,或产品相关信息)以及智能设备(例如,客户端设备108)的屏幕上的图像。因此,输入模态指的是用户可以与机器人通信的不同方式。输入模态还可以包括键盘或鼠标导航、触敏手势等。

[0045] 关于计算机视觉组件208的模态,照片通常可以比文本更好地表示用户正在寻找的东西。用户可能不知道什么是项目的名称,或者可能很难甚至不可能使用文本来获得只有专家才可能知道的细致的详细信息,例如服饰的复杂图案或家具的某种风格。此外,在移动电话上键入复杂的文本查询是不方便的,并且长文本查询通常具有差的回应(recall)。因此,计算机视觉组件208的关键功能可以包括对象定位、对象识别、光学字符识别(OCR)以及基于来自图像或视频的视觉线索与库存的匹配。当在具有内置相机的移动设备上运行时,启用了计算机视觉的机器人是有利的。强大的深度神经网络可以用于实现计算机视觉应用。

[0046] 在一个示例中,对话管理器216具有上下文管理器218和NLG组件212作为子组件。如上所述,对话管理器216操作以理解“特异性的完整性”并决定下一个动作类型和参数(例如,“搜索”或“请求来自用户的进一步信息”)。上下文管理器218进行操作以管理给定用户对机器人及其AI的上下文和通信。上下文管理器218包括两个部分:长期历史和短期记忆。每个上下文管理器条目可以描述相关意图和所有参数以及所有相关结果。上下文是关于库存以及其他未来的知识来源的。NLG组件212进行操作以用AI消息组成自然语言话语,以呈现给与智能机器人交互的用户。

[0047] 人与机器之间流畅的、自然的、富含信息的、甚至是娱乐性的对话是一个困难的技术问题,其在过去的一个世纪的大部分时间里已经被研究过,但仍然被认为是未解决的。然而,AI的最新发展产生了有用的对话系统,例如Siri™和Alexa™。

[0048] 在智能机器人的电子商务示例中,寻求解决此问题的最初非常有帮助的元素是利用大量的电子商务数据集合。这些数据中的一些可以被保留在专有数据库中或云(例如数据库/云126)中。关于该数据的统计可以作为上下文从搜索组件220向对话管理器216进行传送。人工智能框架128可以直接对来自用户的话语起作用,所述话语可以经过语音识别组件210,然后经过NLU组件214,并在然后作为半解析数据被传递给上下文管理器218。因此,NLG组件212可以帮助对话管理器216生成类人问题(human-like question),并且以文本或语音来向用户响应。上下文管理器218维持用户和人工智能框架128之间的多轮且长期的交

谈的一致性。

[0049] 可以推荐进行区别对待来对庞大的电子商务数据集进行轮询,以仅寻找相关的有用信息。在一个示例中,人工智能框架128使用来自搜索组件220的结果以及搜索组件220内的智能来提供该信息。该信息可以与来自上下文管理器218的交互历史进行组合。然后,人工智能框架128可以决定下一轮对话,例如,它是否应该是问题,或是用于验证例如现有理解或用户意图的“基础陈述”,或是项目推荐(或者,例如,所有三者的任意组合)。这些决定可以通过数据集、用户的聊天历史和用户理解能力的模型的组合来做出。NLG组件212可以基于这些决定生成用于对用户的文本或口语答复的语言。

[0050] 由本发明主题提供的技术解决方案允许用户以自然的交谈来与智能在线个人助理进行通信。该助手是高效的,因为随着时间的推移,它越来越理解特定的用户偏好,并且对各种产品了如指掌。例如,通过各种方便的输入模态,用户可以共享照片,或使用语音或文本,并且所辅助的用户体验可以类似于与高端商店中的可信任的、知识渊博的人类购物助理进行交谈。

[0051] 传统上,在线购物系统使用的方法和数据通过生硬的、简化的假设来瞄准不知名的购买者群体,以最大化短期收入。传统网站和app不了解用户希望收到通知的方式、原因和时间。通知可能是令人烦恼的、不适当的和没有人情味的,无视每个用户的偏好。一个人与单个帐户不同。人们共享帐户和设备。密码使平台既不安全也不易于使用。弱的在线身份和忽略环境信号(例如设备、位置、异常行为后的通知)的问题使得易于在市场中进行欺诈。

[0052] 参考图3,现在描述语音识别组件210的所示组件。特征提取组件进行操作以将原始音频波形转换为表示声音的某个维度的数字向量。该组件使用深度学习来将原始信号投影到高维度语义空间中。声学模型组件进行操作以托管语音单元(例如音素和异音素)的统计模型。虽然可以使用深度神经网络,它们可以包括高斯混合模型(GMM)。语言模型组件使用语法的统计模型来限定单词如何在句子中排放在一起。这些模型可以包括基于n元语法(n-gram)的模型或除单词嵌入外构建的深度神经网络。使用隐藏马尔可夫模型(HMM)框架中的特征提取组件、声学模型组件和语言模型组件从特征序列中导出单词序列,语音到文本(STT)解码器组件通常可以利用从原始信号中导出的特征来将语音话语转换成单词序列。在一个示例中,云(例如,数据库/云126)中的语音到文本服务使这些组件部署在具有API的云框架中,该API允许针对语音话语发布音频样本并取回对应的单词序列。控制参数可用于定制或影响语音到文本过程。

[0053] 在人工智能框架128的一个示例中,提供了用于语音识别组件210的两个附加部分,即说话者适配组件和语言模型(LM)适配组件。说话者适配组件允许STT系统(例如,语音识别组件210)的客户端针对每个说话者/用户定制特征提取组件和/或声学模型组件。这可能是重要的,因为大多数语音到文本系统都是在来自目标区域中的代表性说话者集合的数据上进行训练的,并且通常系统的准确度在很大程度上取决于目标说话者与训练池中的说话者匹配的程度。说话者适配组件允许语音识别组件210(并且因此允许人工智能框架128)通过持续地学习用户的语调、发音、重音和其他语音因素的特有风格来对说话者变化稳健,并将这些应用于语音依赖性组件(例如,特征提取组件)和声学模型组件。虽然这种方法可能需要针对每个说话者创建并存储小的语音简档,但准确度的潜在好处通常远远超过存储缺陷。

[0054] LM适配组件进行操作以利用来自目标域(例如,库存类别或用户角色)的新单词和代表性句子来定制语言模型组件和语音到文本词汇。这种能力允许人工智能框架128在支持新类别和角色时可扩缩。

[0055] 图3还示出了用于人工智能框架128中的文本规范化的流程序列302。在一个示例中,执行流程序列302的文本规范化组件被包括在语音识别组件210中。流程序列302中的关键功能包括拼写规范化(用于处理标点符号、数字、大小写等)、交谈文本规范化(用于处理具有首字母缩略词、缩写、不完整片段,俚语等的非正式聊天类型文本)、以及机器翻译(用于将已规范化的外语单词序列转换成操作语言(例如包括但不限于英语)中的单词序列)。

[0056] 人工智能框架128促进现代通信。例如,千禧一代经常希望经由照片、语音和文本进行交流。人工智能框架128使用多种模态的技术能力允许意图而不仅仅是文本的表达。人工智能框架128提供技术解决方案并且是有效的。在许多情况下,与使用文本相比,使用语音命令或照片更快地与智能个人助理进行交互。

[0057] 图4是示出了示例软件架构406的框图,示例软件架构406可以结合本文中所描述的各种硬件架构来使用。图4是软件架构的非限制性示例,并且将理解,可以实现许多其他架构以促进本文中描述的功能。软件架构406可以在诸如图5的机器500的硬件上执行,机器500包括处理器504、存储器514和输入/输出(I/O)组件518。代表性的硬件层452被示出,并且可以表示例如图5的机器500。代表性的硬件层452包括具有相关联的可执行指令404的处理单元454。可执行指令404表示软件架构406的可执行指令,包括对本文中所描述的方法、组件等的实现。硬件层452还包括存储器和/或存储模块存储器/存储设备456,所述存储器/存储设备456也具有可执行指令404。硬件层452还可以包括其他硬件458。

[0058] 在图4的示例性架构中,软件架构406可以被概念化为层的堆栈,其中每层提供特定的功能。例如,软件架构406可以包括诸如操作系统402、库420、应用416和表示层414之类的层。在操作上,应用416和/或层内的其它组件可以通过软件堆栈来调用应用编程接口(API)调用408,并且响应于API调用408来接收响应。所示出的层在本质上具有代表性,并不是所有的软件架构都具有所有层。例如,一些移动或专用操作系统可能不提供框架/中间件层418,而其它系统可以提供这样的层。其它软件架构可以包括附加层或不同层。

[0059] 操作系统402可以管理硬件资源并提供公共服务。操作系统402可以包括例如内核422、服务424和驱动426。内核422可以用作硬件和其他软件层之间的抽象层。例如,内核422可以负责存储器管理、处理器管理(例如调度)、组件管理、联网、安全设置等。服务424可以为其它软件层提供其他公共服务。驱动426负责控制底层硬件或与底层硬件接口连接。例如,取决于硬件配置,驱动426可以包括显示器驱动、相机驱动、蓝牙®驱动、闪存驱动、串行通信驱动(例如通用串行总线(USB)驱动),Wi-Fi®驱动、音频驱动、电力管理驱动等等。

[0060] 库420提供由应用416和/或其它组件和/或层使用的公共基础设施。库420可以提供允许其他软件组件以与底层操作系统402功能(例如,内核422、服务424和/或驱动426)直接接口连接相比更容易的方式来执行任务的功能。库420可以包括可以系统库444(例如,C标准库),系统库444提供诸如存储器分配功能、字符串操纵功能、数学功能之类的功能。另外,库420可以包括API库446,例如媒体库(例如,用于支持各种已知的媒体格式(诸如MPREG4、H.264、MP3、AAC、AMR、JPG、PNG)的呈现和操纵的库)、图形库(例如,可以用于在显示器上渲染2D和3D图形内容的OpenGL框架)、数据库(例如,可以提供各种关系数据库功能的

SQLite)、web库(例如,可以提供网络浏览功能的WebKit)等。库420还可以包括各种各样的其它库448,以提供到应用416和其它软件组件/模块的许多其他API。

[0061] 框架/中间件418(有时也称为中间件)可以提供可以由应用416和/或其它软件组件/模块使用的更高级的公共基础设施。例如,框架/中间件418可以提供各种图形用户界面(GUI)功能、高级资源管理、高级位置服务等。框架/中间件418可以提供可以由应用416和/或其它软件组件/模块利用的广泛的其它API,其中一些可以特定于特定操作系统或平台。

[0062] 应用416包括内置应用438和/或第三方应用440。代表性的内置应用438的示例可以包括但不限于联系人应用、浏览器应用、书籍阅读器应用、位置应用、媒体应用、消息传递应用和/或游戏应用。第三方应用440可以包括由与特定平台的供应商不同的实体使用ANDROID™或IOS™软件开发工具包(SDK)而开发的任何应用,并且可以是在移动操作系统(例如IOS™、ANDROID™、WINDOWS® Phone或其他移动操作系统)上运行的移动软件。第三方应用440可以调用由移动操作系统(例如操作系统402)提供的API调用408,以促进本文中所描述的功能。

[0063] 应用416可以使用内置操作系统功能(例如,内核422、服务424和/或驱动426)、库420和框架/中间件418来创建用户接口以与系统的用户交互。备选地或附加地,在一些系统中,与用户的交互可以通过表示层(例如表示层414)发生。在这些系统中,应用/组件“逻辑”可以与和用户交互的应用/组件的各个方面分开。

[0064] 一些软件架构使用虚拟机。在图4的示例中,这由虚拟机410示出。虚拟机410创建软件环境,在该软件环境中,应用/组件可以像它们在硬件机器(例如图5的机器500)上执行一样执行。虚拟机410由主机操作系统(图4中的操作系统(OS)436)托管,并且通常(尽管并不总是)具有虚拟机监视器460,该虚拟机监视器460管理虚拟机的操作和与主机操作系统(例如,操作系统402)连接的接口。软件架构在虚拟机410(例如,操作系统436、库434、框架432、应用430和/或表示层428)内执行。在虚拟机410内执行的这些软件架构的层可以与先前描述的对应层相同,或者可以不同。

[0065] 图5是示出了根据一些示例实施例的机器500的组件的框图,机器500能够从机器可读介质(例如,机器可读存储介质)中读取指令并执行本文中所讨论的方法中的一种或多种。具体地,图5示出了计算机系统的示例形式的机器500的示意图,其中可以执行指令510(例如,软件、程序、应用、小应用、app或其他可执行代码)以使机器500执行本文中所讨论的方法中的任何一种或多种。因此,指令可以被用来实现本文中所描述的模块或组件。指令将通用的未编程的机器转换成被编程为以所描述的方式执行所描述和示出的功能的特定机器。在备选实施例中,机器500作为独立设备操作或可以耦接(例如,联网)到其他机器。在联网部署中,机器500可以在服务器-客户端网络环境中以服务器机器或客户端机器的容量操作,或者作为对等(或分布式)网络环境中的对等机器操作。机器500可以包括但不限于服务器计算机、客户端计算机、个人计算机(PC)、平板计算机、膝上型计算机、上网本、机顶盒(STB)、个人数字助理(PDA)、娱乐媒体系统、蜂窝电话、智能电话、移动设备、可穿戴设备(例如,智能手表)、智能家居设备(例如,智能家电)、其他智能设备、web设备、网络路由器、网络交换机、网络桥接器、或能够顺序地或以其它方式执行指定机器500要采取的动作的指令510的任何机器。此外,尽管仅示出了单个机器500,但是术语“机器”也将被视为包括机器的

集合,其单独地或联合地执行指令510以执行本文中所讨论的方法中的任何一种或多种。

[0066] 机器500可以包括可以被配置为例如经由总线502彼此通信的处理器504、存储器/存储设备506和I/O组件518。存储器/存储设备506可以包括存储器514(比如,主存储器或其它存储器/存储设备)、以及存储单元516,存储器514和存储单元516两者都可例如经由总线502由处理器504访问。存储单元516和存储器514存储体现本文中所描述的方法或功能中的任何一种或多种的指令510。在机器500执行指令510期间,指令510还可以完全地或部分地驻留在存储器514内、存储单元516内、处理器504中的至少一个内(例如,处理器的高速缓存存储器内)、或其任何合适的组合内。因此,存储器514、存储单元516和处理器504的存储器是机器可读介质的示例。

[0067] I/O组件518可以包括用于接收输入、提供输出、产生输出、发送信息、交换信息、捕捉测量等的各种各样的组件。包括在特定机器中的特定I/O组件518将取决于机器的类型。例如,诸如移动电话的便携式机器将可能包括触摸输入设备或其他这样的输入机构,而无头服务器机器将可能不包括这样的触摸输入设备。应理解,I/O组件518可以包括图5中未示出的许多其他组件。I/O组件518根据功能被分组,以便简化以下讨论,并且分组不以任何方式进行限制。在各种示例实施例中,I/O组件518可以包括输出组件526和输入组件528。输出组件526可以包括视觉组件(例如,显示器,诸如等离子体显示面板(PDP)、发光二极管(LED)显示器、液晶显示器(LCD)、投影仪或阴极射线管(CRT))、声学组件(例如扬声器)、触觉组件(例如振动马达、电阻机构)、其他信号发生器等。输入组件528可以包括字母数字输入组件(例如,键盘、配置为接收字母数字输入的触摸屏、光-光学键盘或其他字母数字输入组件)、基于点的输入组件(例如,鼠标、触摸板、轨迹球、操纵杆、运动传感器或其他定点仪器)、触觉输入组件(例如,物理按钮、提供触摸或触摸手势的位置和/或力的触摸屏或其他触觉输入组件)、音频输入组件(例如,麦克风)等。

[0068] 在另一些示例实施例中,I/O组件518可以包括生物测定组件530、运动组件534、有关环境的环境组件536或位置组件538、以及许多其他组件。例如,生物测定组件530可以包括用于检测表达(例如,手表达、面部表情、语音表达、身体姿势或眼睛跟踪)、测量生物信号(例如,血压、心率、体温、汗水或脑波)、识别人(例如,语音识别、视网膜识别、面部识别、指纹识别或基于脑电图的识别)等的组件。运动组件534可包括加速度传感器组件(例如,加速度计)、重力传感器组件、旋转传感器组件(例如,陀螺仪)等。环境组件536可以包括例如照明传感器组件(例如光度计)、温度传感器组件(例如,检测环境温度的一个或多个温度计)、湿度传感器组件、压力传感器组件(例如气压计)、声学传感器组件(例如,检测背景噪声的一个或多个麦克风)、接近传感器组件(例如,检测附近物体的红外传感器)、气体传感器(例如,为了安全而检测危险气体的浓度或测量大气中污染物的气体检测传感器)或可以提供与周围物理环境相对应的指示、测量或信号的其他组件。定位组件538可以包括位置传感器组件(例如,全球定位系统(GPS)接收机组件)、高度传感器组件(例如,高度计或检测气压的气压计(根据气压可以导出高度))、取向传感器组件(例如,磁力计)等。

[0069] 可以使用各种各样的技术来实现通信。I/O组件518可以包括通信组件540,通信组件540可操作以分别经由耦接522和耦接524来将机器500耦接到网络532或设备520。例如,通信组件540可以包括网络接口组件或与网络532接口连接的其他合适设备。在另一些示例中,通信组件540可以包括有线通信组件、无线通信组件、蜂窝通信组件、近场通信(NFC)组

件、蓝牙®组件(例如蓝牙®低功耗)、Wi-Fi®组件、以及经由其他模态提供通信的其他通信组件。设备520可以是另一机器或各种外围设备中的任一种(例如,经由通用串行总线(USB)耦接的外围设备)。

[0070] 此外,通信组件540可以检测标识符或包括可操作以检测标识符的组件。例如,通信组件处理器通信组件540可以包括射频识别(RFID)标签读取器组件、NFC智能标签检测组件、光学读取器组件(例如,用于检测以下各项的光学传感器:一维条形码(例如通用产品代码(UPC)条形码)、多维条形码(例如快速响应(QR)码)、阿兹台克码、数据矩阵、Dataglyph、MaxiCode、PDF417、超码、UCC RSS-2D条形码和其他光学码)、或声学检测组件(例如,用于识别标记的音频信号的麦克风)。此外,可以经由通信组件540来导出各种信息,例如,经由互联网协议(IP)地理位置的位置、经由Wi-Fi®信号三角测量的位置、经由检测可以指示特定位置的NFC信标信号的位置等等。

[0071] 现在参考图6,示出了可以在其中部署由智能个人助理系统106提供的智能在线个人助理的示例环境600。在环境600的中心处,出现具有AI的智能机器人602。机器人利用计算机视觉组件208、语音识别组件210、NLU组件214、对话管理器216、NLG组件212、搜索组件220和身份服务222来使用户参与有效的、令人感兴趣的和起作用的对话,以解码他们的意图并传递个性化结果。

[0072] 相关联的应用604可以通过引人注目的移动设计能力和元素来展示机器人602的全部能力和智能。基本结构606与Facebook Messenger™、Skype™以及Cortana™(例如)集成,以使用户能够在他们已经花费时间的地方进行交易。智能通知610平台经由任何数量的渠道(例如,SMS、推送通知、电子邮件、消息传递)在适当的时间向用户传递适当的信息,以鼓励他们与机器人602和相关联的市场互动。社区608功能使用户能够使用他们已经在其中花费大部分时间的相同消息传递系统来与他们的朋友、品味营造者和品牌进行连接、互动和交互。其他功能包括团购和礼品购买。奖励612平台激励用户更深入地与机器人602互动。奖励可以包括产品的大幅折扣、对特有库存的访问、以及app中通过分数、级别等进行的褒奖。在营销614处,执行传统的社交营销和其他营销的组合以便以更私人的方式赢得一些人群的关注(例如,千禧一代)。常规技术可以包括商品推销、电子邮件、搜索引擎优化(SEO)和搜索引擎营销(SEM)以及用于针对目标新用户和现有用户的诸如社交广告、病毒式优惠券之类的实验性技术。

[0073] 图7示出了智能个人助理系统106处理自然语言用户输入以生成电子市场中的项目推荐的概述。虽然智能个人助理系统106不限于这种使用场景,但在这种情况下它可以是特别有用的。如前所述,文本、图像和语音数据的任何组合可以由人工智能框架128接收。图像数据可以由计算机视觉组件208处理以提供图像属性数据。语音数据可以由语音识别组件210处理成文本。

[0074] 可以将所有这些输入和其他输入提供给NLU组件214以进行分析。NLU组件214可以操作以解析用户输入并帮助确定用户意图和意图相关参数。例如,NLU组件214可以辨别用户感兴趣的主要对象,以及与该主要对象相关的各种属性和属性值。NLU组件214还可以确定其他参数,例如用户输入类型(例如,问题或陈述)和目标项目接收者。NLU组件214可以将所提取的数据提供给对话管理器216以及先前示出的AI协调器206。

[0075] NLU组件214通常可以将正式和非正式的自然语言用户输入转换为用户查询的更

正式的机器可读的结构化表示。对话管理器216可以进一步增强该正式化的查询。在一种场景下,NLU组件214处理用户输入序列,包括原始查询和由用户在多轮交互式对话中响应于来自对话管理器216的机器生成的提示而提供的其他数据。该用户-机器交互可以改善可在电子市场中购买的最相关项目的一个或多个自动搜索的效率和准确性。搜索可以由搜索组件220执行。

[0076] 提取用户意图对于AI机器人确定所需的进一步操作是非常有帮助的。在一个与电子商务相关的示例中,在最高级别下,用户意图可以是购物、闲聊、笑话、天气等。如果用户意图是购物,则可能涉及追求特定的购物任务、为不同于该用户的目标接收者赠送项目、或仅浏览可供购买的项目的库存。一旦识别出高级意图,人工智能框架128的任务就是确定用户正在寻找什么;也就是说,宽泛的需求(例如,鞋子、服装)或更具体的需求(例如,两双新的黑色的Nike™10号胶底鞋)或介于两者之间(例如,黑色胶底鞋)?

[0077] 在优于该领域中的现有技术的新颖且明显的改进中,人工智能框架128可以将用户请求映射到最佳地表征所期望的可用项目的某个主维度,例如类别、属性和属性值。这使机器人能够与用户互动,以在必要时进一步细化搜索约束。例如,如果用户向机器人询问与服装相关的信息,则需要说明的顶层属性可以是颜色、材料和样式。此外,随着时间的推移,机器学习可以向系统添加更深层次的语义和更广泛的“常识知识”,以更好地理解用户意图。例如,输入“我在寻找用于意大利6月份的婚礼的服装”意味着服装应该在给定的时间和地点适合特定的天气条件,并且应该适合正式场合。另一个示例可以包括用户询问机器人“我侄子的礼物”。人工智能框架128在训练时将理解送礼物是一种特定类型的意图,基于“侄子”的含义目标接收者是男性,以及应该澄清诸如目标接收者的年龄、场合和爱好/喜好之类的属性。

[0078] 图8示出了根据一些示例实施例的NLU组件214、其子组件以及与其交互的其他组件。在一些实施例中,通过将这个通常复杂的技术问题分解成多个部分来由NLU组件214执行提取用户意图。提取用户意图的整体问题的各个部分中的每一个可以由NLU组件214的特定子组件处理,有时单独地处理并且有时组合地处理。

[0079] 子组件可以例如包括拼写校正器(拼写器)802、机器翻译器(MT)804、解析器806、知识图808、命名实体识别(NER)子组件810、词义检测器(WSD)812、意图检测器813和解释器814。在一个实施例中,NLU组件214可以例如经由AI协调器206接收文本、视觉选择器和图像属性,并且单独地或组合地处理每一个。视觉选择器通常是由用户提供的图形选择,诸如从多个所呈现的颜色样本中选择颜色,或者选择具有相关联并且因此被选择的精神状态的表情符号。在一个实施例中,NLU组件214可以将其要描述的各种输出提供给AI协调器206,以分发给人工智能框架128的其他组件,例如对话管理器216。

[0080] NLU组件214考虑的其他输入可以包括对话上下文816(例如,来自上下文管理器218)、用户身份信息818(例如,来自身份服务222)、项目库存相关信息820(例如,来自电子市场的核心搜索引擎220功能)、和外部常识知识822,以改善根据用户输入对用户意图的语义推断。对这些输入的不同类型的分析可以各自产生可以以聚合的方式解释并经由知识图808协调的结果。知识图808可以例如基于过去用户的交互、库存相关数据或两者。

[0081] 拼写器802可以识别并校正用户输入的文本中的拼写错误。用户文本可以包括但不限于用户查询和项目标题。机器翻译器804可以可选地将用户输入从用户的自然语言翻

译成操作语言,包括但不限于例如英语。拼写器802和机器翻译器804还可以与其他规范化子组件和/或解析器806相配合,以将缩写、首字母缩略词和俚语处理成更正式的数据以用于改进的分析。

[0082] 解析器(或依赖性解析器)806可以通过寻找用户的输入查询的主要对象来帮助检测用户的意图。该过程可以涉及解析器识别和分析例如来自多轮对话的名词短语,包括介词和直接和间接宾语、动词、以及用户输入中的肯定和否定。在一些实施例中,可以在意图检测器子组件813中检测肯定和否定,或者通过诸如词义检测器812之类的不同子组件来检测确认和否定。

[0083] 在一个实施例中,解析器806从用户输入中的可以完全剖析出(resolve)的最长片段中寻找用户感兴趣的主要对象。解析器806还可以丢弃具有低内容的用户输入项,例如“嘿,你好”和“你能帮助我”等,和/或用不那么机器混淆的短语来替换它们。解析器806还可以识别各种场合(例如,婚礼、母亲节等)。

[0084] 意图检测器813可以通过识别感兴趣的主要对象(其通常但不总是项目类别)和解析器806所建议的结果的相应最佳属性来进一步细化对用户意图的识别。例如,如果用户意图是购买特定的项目,则知识图808可以使用其映射到的给定项目库存(例如,eBay库存或数据库/云126)中的主要项目类别。知识图808还可以使用与该项目类别有关的主要(例如,用户最频繁地查询或在项目库存中最频繁地出现的)属性,以及这些属性的主要值。因此,NLU组件214可以提供主要对象、用户意图和知识图808作为其输出,该知识图808是沿着可能与用户查询相关的维度而制定的。关于是否缺失完全剖析针对项目推荐的用户查询所需的信息,以及因此接下来是否(以及如何)提示用户经由附加输入来进一步细化用户的要求,该信息可以帮助到对话管理器216。

[0085] 知识图808的背景信息可被从项目库存中提取,作为从手工组织的目录导出的信息以及从历史用户行为(例如,一段时间内用户与电子市场的所有先前交互的历史)提取的信息的混合项。知识图还可以包括从外部来源(例如,互联网百科全书(例如,维基百科)、在线词典、辞典和词汇数据库(例如,WordNet))提取的常识知识。例如,关于术语相似性和关系的数据可以用于确定术语“女孩”、“女儿”、“姐妹”、“女人”、“阿姨”、“侄女”、“祖母”和“母亲”都指的是女性并且指的是不同的特定亲属血缘关系。这些附加关联可以澄清用户查询术语的一种或多种含义,并且有助于防止生成可以教育机器人但是使用户恼火的提示。焦点小组(focus group)研究表明,一些用户不希望提供比预定数量(例如三个)更多的提示回复,因此这些提示中的每一个都应该尽可能精辟。

[0086] 在一些实施例中,知识图808可以例如通过AI协调器206动态地更新。也就是说,如果项目库存改变或者如果新的用户行为或新的常识知识数据已经导致成功的用户搜索,则智能在线个人助理106能够利用这些改变来进行将来的用户搜索。进行学习的助理可以促进进一步的用户交互,特别是对于那些不太倾向于进行大量交谈的用户而言。因此,实施例可以修改知识图808以调整其包含并且与NLU组件214内的其他子组件和在外部(例如与对话管理器216)共享的信息。

[0087] NER子组件810可以从所解析的用户输入(例如,品牌名称、尺寸信息、颜色和其他描述符)中提取更深的信息,并且帮助将用户自然语言查询转换成包括这种解析出的数据元素的结构化查询。NER子组件还可以利用常识知识来帮助剖析所提取的术语的含义。例

如,根据在线词典和百科全书,对“波尔多”的查询可以更成功地确定查询术语可以指代项目类别(葡萄酒)、属性(类型、颜色、原产地位置)和相应的对应属性值(波尔多、红色、法国)。类似地,地名(例如,太浩湖(Lake Tahoe))可以对应于可以帮助用户寻找相关项目的给定地理位置、天气数据、文化信息、相对成本和流行活动。结构化查询深度(例如,针对给定用户话语长度剖析出的标签的数量)可以帮助对话管理器216选择其应该采取什么进一步的动作来改善搜索组件220执行的搜索的排名。

[0088] 词义检测器812可以处理多义的(即,具有基于上下文的多个不同含义)单词。例如,输入术语“(bank)”可以指地理意义上的“河流边缘”或购买交易支付意义上的“金融机构”。词义检测器812检测这样的单词,并且如果词义依然是模糊的,则可以触发对话管理器216从用户寻求进一步的剖析。词义检测器812或意图检测器子组件813还可以分别从示例性短语中辨别出肯定和否定等,该示例性短语包括但不限于“显示更多”或“不,我不喜欢”。因此,解析器804、意图检测器813和词义检测器812的功能可以在一定程度上重叠或交互,这取决于特定实现。

[0089] 解释器814协调来自各种NLU子组件的已分析的信息并且准备输出。输出可以例如包括用户查询的主要对象,以及关于以下项所剖析出的信息:相关知识图维度(例如,项目类别、项目属性、项目属性值)、用户的意图(例如,在购物的情况下是否购买特定项目、寻找礼物或一般性浏览)、所识别的用户陈述的类型、预期的目标项目接收者等。通过组合对共享的、增强的和已处理的用户输入执行的单独分析,人工智能框架128的组件提供可信任的个人购物者(机器人),其既理解用户意图又了解广泛的产品。因此,NLU组件214将自然语言用户查询变换成结构化查询,以帮助向用户提供最相关的结果。

[0090] 因此,NLU组件214通过减少错误、增加正确预测用户查询下的用户意图的可能性以及产生更快和更好的针对性搜索和项目推荐来整体上改善智能个人助理系统106的操作。NLU组件214,特别是与多轮对话场景中的对话管理器216一起,通过提供用于执行的更多的聚焦用户交互历史和/或聚焦项目库存的搜索查询来有效地管理搜索组件220的操作。这种独特的功能经由所描述的元件的特定有序组合而超越了当前的技术现状。

[0091] 现在更一般性地描述NLU组件214和智能个人助理系统106的使用示例,该使用示例用于处理来自用户的输入数据。用户可以提供诸如“我在为我的妻子寻找一副太阳镜”之类的口语陈述。NLU组件214可以处理该自然语言用户输入以生成要提供给搜索引擎220和/或对话管理器216的更正式的查询。该更正式的查询可以包括将一个或多个剖析出的句柄中的每一个句柄与对应的剖析值相关联的一组标签。例如,更正式的查询可以包括“<意图:送礼物,陈述类型:陈述,主要对象:太阳镜,目标:妻子,目标性别:女性>”。搜索引擎可以基于对这些标签的搜索提供比搜索最初提交的用户输入会产生的结果更相关的结果。

[0092] 在该示例中,智能个人助理系统106确定用户意图是送礼物(与仅仅是自我购物或浏览相对)、用户提供了陈述(与问题相对)、以及用户感兴趣的主要对象是太阳镜。虽然用户正在购物,但意图是将项目送给特定的目标项目接收者,即他的妻子。已知送礼物任务是一种特定类型的购物任务,其可以在一定程度上与一般的库存浏览或用户为自己购买项目不同地处理。

[0093] 智能个人助理系统106通常还可以通过命名实体识别器子组件810辨别出“妻子”指的是女性。例如,可以从身份服务212提供的数据中找到作为目标项目接收者的特定个

体。此外,通过使用常识知识,智能个人助理系统106可以确定术语“妻子”是指已婚女性,并且儿童通常未婚。该信息可以有助于相对于其他类型的太阳镜(例如,男式太阳镜、儿童太阳镜)来约束对女性太阳镜的搜索,以生成更相关的项目推荐,而无需用户提示来获得相同的信息。

[0094] 图9示出了根据一些示例实施例的各种分析的结果。在一个示例中,用户可以键入文本输入“嗨,你能为我寻找一双红色的nike鞋吗?”所产生的正式查询可以包括“<意图:购买,陈述类型:问题,主要对象:鞋,目标:自己,颜色:红色,品牌:nike>”。这里,用户输入是一个问题,并且用户正在购买特定项目(与仅浏览项目库存或为其他人寻找礼物相对地)。解析器806可以确定术语“嗨,你能为我寻找”不提供大量的有用内容,并且因此可以被忽视。

[0095] 拼写器子组件802可以确定“nikey”是术语“nike”的已知错误拼写,并进行适当的校正。解析器子组件806可以通过识别动词、介词和名词短语来在语法上分析已规范化的输入数据。单个单词之间的语法关系可以说明一个单词如何依赖于或修饰另一个单词,并且该信息可以提供用于变换用户查询的线索。

[0096] 解析器子组件806还可以执行名词短语组块,并从最长的已解析查询片段“红色nike鞋”中辨别出用户感兴趣的主要对象是鞋。也就是说,鞋被确定为最大数量的修饰语的对象,并且处于所产生的组块结构的最深级别。请注意,虽然主要对象通常是项目类别,但情况并非总是如此。这里的主要对象还由修饰语(“红色”和“nike”)描述,命名实体识别器810可以确定该修饰语分别与颜色和品牌相关。

[0097] 还要注意,在这种情况下,为主要对象提供了两个属性(颜色、品牌)和对应的属性值(红色、nike),而在前面的示例中,最多提供了一个属性(例如,经由推理间接地指定了女性的太阳镜)。作为结果,对话管理器216可以判定用户的原始查询被充分约束,适当的提示可以是一个或多个项目推荐,而不是向用户询问将会进一步缩窄后续搜索的附加约束的问题。相反,对于先前的查询,可能需要关于女性的太阳镜的更多细节,因此对话管理器216可以在多轮对话中为此生成多个提示。然而,一些用户对大量提示感到恼火,并且宁愿应对能够从每一个回合中自己提取更多信息的机器人。因此,通过从每个用户话语中收集尽可能多的信息来最小化多轮对话中的轮数是有利的。

[0098] 例如,NLU组件214可以确定在所搜索的项目库存中存在红色nike鞋的许多不同的列表,和/或确定在用户进行项目选择之前先前用户的交互确定了附加属性值。因此,NLU组件214可以查阅知识图808以确定用户感兴趣的该主要对象的最有用的属性。知识图808可以具有指示对于项目类别“鞋”而言,最有帮助和/或被频繁指定的属性是颜色、品牌和大小,并具有示出在寻找相关项目时每个属性的重要性的相对相关性或关联强度或条件概率的对应条件概率值。情况可能是所有这些属性可能需要参数化以使查询被认为足够特定以导致搜索成功。然而,情况也可能是仅需要参数化充分覆盖预定百分比的可用关联的有限数量的属性。

[0099] 在该示例中,用户已经提供了颜色和品牌的属性值,但是没有提供尺寸的属性值,因此对话管理器216因而可以询问用户“你想要什么尺寸?”并等待进一步的用户输入。假设用户回复“我想要10”。这是什么意思?智能个人助理系统106可以将“我想要10”解释为意味着用户想要10个先前指定的红色nike鞋。常识知识可能提供鞋通常成对出现的信息,因此

针对用户对提示的响应的重新解释可能在一定程度上被细化为如下的观点：用户代之以想要十双红色nike鞋。然而，这两种解释都不正确，因为它们都没有考虑会话的上下文。也就是说，“我想要10”用户输入是对生成来收集关于先前话语的更多信息（在这种情况下是尺寸属性的值）的提示的回复。如果智能个人在线助理106不能将回复与任何先前的用户输入相关联，则它可以输出对其不能剖析交谈上下文进行指示的出错陈述。

[0100] 通过针对购物任务不仅跟踪用户交互的长期历史而且还跟踪当前用户交互的短期记忆，上下文管理器218可以防止这种混淆。对多轮对话中的提示的回复不一定是孤立的用户话语，而是通常在上下文意义上与对话中的先前的用户话语和先前的提示（如果有的话）相关。因此，智能个人助理系统106适合于这样的用户交谈：该用户交谈导致累积的搜索约束，该累积的搜索约束足以使得细化的搜索查询在寻找要推荐的相关项目时更成功。

[0101] 然而，在一些情况下，NLU组件214可以确定用户已经放弃先前的查询任务并且现在有兴趣寻找其他东西。因此，在一些实施例，对话管理器216可以从NLU组件214接收对该确定的指示，并相应地改变其行为。该对话管理器216的行为可以包括保存用于当前搜索任务的交互以便例如可能在以后使用，以及基于当前的用户话语开始新的对话而不使用与先前的搜索任务有关的任何上下文信息。在一个实施例中，当检测到用户感兴趣的新的主要对象时，NLU组件214可以确定发生了这种任务变化。

[0102] 图10示出了根据一些示例实施例的知识图808。知识图808通常是表示多个节点（这里以椭圆形示出）的数据库或文件。对于处理自然语言用户输入以生成项目推荐的示例性场景，每个节点可以表示项目类别、项目属性或项目属性值。在该示例中，项目类别包括“男式运动鞋”、“汽车和卡车”和“女士运动鞋”。如图所示，可能已经通过库存跟踪系统或通过智能个人助理系统106为每个项目类别分配了标识号。

[0103] 在该示例中的知识图808中示出的项目属性包括“产品线”、“品牌”、“颜色”和“样式”。项目属性通常直接链接到项目类别，但情况并非总是如此。在该示例中的知识图808中示出的项目属性值包括“乔丹气垫”、“科比布莱恩特”、“空军1号”、“亚瑟士”、“耐克”，“新百伦”、“阿迪达斯”、“蓝色”、“白色”、“红色”、“黑色”、“金属黑色”、“跑步”、“篮球”和“胶底鞋”。项目属性值通常直接链接到项目属性，但情况并非总是如此。

[0104] 知识图808节点之间示出的链接是有向边，其可以具有指示两个特定节点之间的关系强度的相关性的相关性或关联值。在图10中指示了知识图808的一些相关值。可以通过各种方式创建相关值，并且可以将相关值用于各种目的。

[0105] 例如，在一个实施例中，相关值可以从可供购买的项目的库存中导出。库存可以是当前的或历史的。当卖方列出待售的项目时，卖方可以指定一个或多个项目类别、属性和/或属性值作为描述该项目并且因此是可以由想要购买该项目的用户提供的有用搜索术语的元数据。在一些情况下，电子市场可以通过各种方式对卖方的项目进行分类，例如通过向卖方提供描述可用的预定项目类别和常用的描述性术语的指南。

[0106] 例如，卖方可能有一双鞋子要出售，并且可以指定它们是由阿迪达斯制造的男式蓝色运动跑鞋。卖方可以向市场指定项目类别是“男式运动鞋”，并且可以提示卖方例如从项目属性列表中指定项目属性。备选地，电子市场可以识别卖方已经提供了多个项目属性值，并且可以自动地将这些项目属性值与各种项目属性相关，例如，与那些具有这些值作为指定可能性（可能在元数据中）的属性相关联。电子市场可以例如确定“男式运动鞋”实际上

是更宽泛的类别“鞋”的子类别或属性,因为例如卖方或电子市场已经为该类别定义了子类别或属性。

[0107] 电子市场可以周期性地分析其可供销售的项目的库存,并以知识图808的形式提供描述该库存的汇总数据。在该方法中,示例性知识图808可以注意到在“男式运动鞋”类别中的所有库存项目中,百分之三十(或0.3)的项目由项目属性“产品线”表征,百分之四十(或0.4)的项目由项目属性“品牌”表征,并且百分之二十(0.2)的项目由项目属性“颜色”表征,如图所示。在由项目属性“产品线”表征的项目中,百分之二十(或0.2)具有项目属性值“科比布莱恩特”,百分之十(或0.1)具有项目属性值“空军1号”,如图所示。因此,在该实施例中,知识图808可以包括描述可用项目的实际库存的条目。

[0108] 对于可能具有数百万的项目可供购买的非常大的电子市场,对整个项目库存的详细分析,特别是其在任何给定时刻的状态,可能在计算上是昂贵的。因此,这种分析可以仅偶尔地或定期地进行。统计抽样方法还可以产生知识图808,其描述项目库存的特征的近似估计。

[0109] 在处理用户查询期间,可以将来自用户查询的解析的输入数据元素与知识图808的维度进行匹配,以帮助将用户的需求与可用的项目供应相匹配。知识图808的维度可以包括描述可供购买的项目的项目类别、项目属性和项目属性值。如果用户已经表达了对男式运动鞋的兴趣,则用户期望智能个人助理系统106帮助该用户从可供购买的项目的库存中寻找相关项目。找到无法购买的项目可能导致用户完全失去购物兴趣,这是一个令人非常担忧的结果。因此,相关性值可以表示给定项目类别中的由给定的项目属性描述或者具有给定的项目属性值的项目的相对数量。相关性值可以基于条件概率,例如,假如指定了特定的项目属性,那么指定特定的项目属性值的概率是多少。

[0110] 在不同的实施例中,知识图808可以基于一段时间内所有用户与电子市场的历史交互。也就是说,节点可以包括由许多先前的用户在他们的话语或与市场的导航历史中提供的搜索术语。该方法不是如前所描述的分析库存,而是分析用户行为,例如,当买方与市场互动以寻找库存中的相关项目时买方说了什么和做了什么。

[0111] 在该示例实施例中,图10中所示的相关值可以在条件概率方面指示最普遍或频繁发生的用户交互。例如,如果用户指示对女性运动鞋感兴趣,则知识图808可以指示在百分之三十(或0.3)的这样的买方交互中,买方指定项目属性“样式”,在百分之二十(或0.2)的这样的买方交互中,买方指定项目属性“品牌”,并且在百分之三十(0.3)的这样的买方交互中,买方指定项目属性“颜色”。因此,不考虑可用库存,知识指南808表征用户的搜索行为,例如,用户如何尝试寻找相关项目。

[0112] 如在先前的实施例中,在处理用户查询期间,可以将来自用户查询的解析的输入数据元素与知识图808的维度进行匹配,以帮助将用户的需求与可用的项目供应相匹配。然而,知识图808的维度现在可以包括当先前的用户查找要购买的相关项目时通过他们的查询输入提供的类别、属性和属性值。例如,如果用户已经表达了对女性运动鞋感兴趣,则智能个人助理系统106可以通过确定其他用户如何从可供购买的项目的库存中找到与女性运动鞋项目相关的项目来最佳地继续进行。因此,知识图808中的相关值或分数可以表示在导致令人满意的搜索结论时给定类别、给定属性或给定属性值的相对次数。换句话说,相关值可以有效地表示对在知识图808中从一个搜索术语横穿到另一个搜索术语的用户交互路径

的给定部分如何“被击打”的测量。

[0113] 不考虑如何制定知识图808,用户输入术语与知识图维度(例如,项目类别、项目属性和项目属性值)之间的匹配可以用于将原始用户查询变换为改善的查询。该匹配可以例如帮助确定在多轮对话中应该为用户生成什么(如果有的话)提示以最佳地寻找相关的搜索结果。因此,为此目的,NLU组件214可以将来自知识图808的信息提供给对话管理器216。也就是说,NLU组件214可以向对话管理器216递送简明知识图808,以及用户兴趣的主要对象、用户意图和相关参数,其中简明知识图808的维度具有一些相关性。

[0114] 图11A和11B示出了根据一些示例实施例的具有项目类别、一些项目属性和一些项目属性值的简明知识图808。为清楚起见,单独示出和讨论每个图,但是实际上可以一起参考两个图共有的知识图808。在图11A中,如先前所描述的,已规范化和解析的用户查询已经为用户感兴趣的主要对象“鞋”提供了项目属性/值标签<颜色:红色,品牌:nike>。知识图表808指示“鞋”与“男式运动鞋”之间存在百分之四十(0.4)的相关性,并且“男式运动鞋”与“品牌”之间存在百分之四十(0.4)的相关性,并且“男式运动鞋”与“颜色”之间存在百分之二十(0.2)的相关性。“男式运动鞋”与“产品线”之间也存在百分之三十(0.3)的相关性,并且各种项目属性值(例如,“乔丹气垫”、“科比布莱恩特”和“空军1号”)的各种相关性是已知的。因此,无论是基于库存还是用户行为,“男式运动鞋”和“产品线”的尚未指定的查询术语与成功搜索具有显著的关联。因此,对话管理器216可以根据这些尚未指定的可能性的关联或相关值、或它们在知识图808层次结构中的相对位置、或两者的组合,通过用户提示来对这些尚未指定的可能性的参数化进行排序和优先化。

[0115] 类似地,对于图11B,知识图808指示“鞋”与“女式运动鞋”之间存在百分之三十(0.3)的相关性,并且“女式运动鞋”与“样式”之间存在百分之三十(0.4)的相关性。用户没有指定“女式运动鞋”和“样式”,也没有指定“样式”(例如“篮球”、“跑步”和“胶底鞋”)的相关项目属性值。因此,对话管理器216还可以根据这些尚未指定的可能性的关联或相关值、或它们在知识图808层次结构中的相对位置、或两者的组合,通过用户提示来对这些尚未指定的可能性的参数化进行优先化。

[0116] 在一种提示生成策略中,对话管理器216可以从最宽的类别进行到子类别或属性,并且然后到属性值,以按该顺序确定提示话题序列。也就是说,假如已经指定了类别“鞋”,对话管理器216可以直接进行到剖析用户是对“男式运动鞋”还是“女式运动鞋”感兴趣,因为这两种可能性在知识图808中具有最高(或仅有)的可用关联强度。这种分层引导的搜索方法可以吸引那些不想回答多于有限数量的提示的用户,以校正到(zero in)相关项目上。

[0117] 在另一种提示生成策略中,对话管理器216可以从出现在知识图808中的所有未指定的属性和属性值中更随机地选择提示话题。虽然这种方法在某种程度上是无向的,但是当用户在浏览库存而不是追求特定的购物任务时它可能是合适的。不因为与智能个人助理系统106聊天而恼火的用户可能更喜欢这种更具探索性或谈话性的方法,其在某种意义上遍游知识图808的可能性。

[0118] 在图11A和图11B中,可以根据用户是对男式还是女式运动鞋更感兴趣来选择用于进一步用户输入的候选提示,并且因此还可以根据用户是对特定产品线还是样式感兴趣来选择用于进一步用户输入的候选提示。注意,知识图808中的较窄属性(在这种情况下为产品线或样式)在某些情况下实际上可能是用户提示的更好候选,这取决于每个候选者的决

定性。也就是说,样式和产品线各自在知识图808中与它们中的每一个上方的相应的项目属性或子类别等同地相关联,但是存在可用于产品线属性值可能性的更多数据。因此,询问用户是否对乔丹气垫鞋感兴趣的提示也隐式地询问用户是否对特定产品线和男式运动鞋感兴趣。因此,用户的单个肯定或否定可以帮助辨别出用户一次性地明确接受或拒绝多种可能性(例如,属性和属性值)的意图。

[0119] 图12示出了根据一些示例实施例的智能个人助理系统106处理自然语言用户输入以生成建议性提示的概述。由于已知给用户的不精辟的提示(例如,提供可在不询问用户的情况下确定的信息)会使一些用户烦恼,因此一些实施例可以使用附加数据来缩小用户明确给出的可能的搜索约束的领域。例如,NLU组件214已经辨别出用户对购买红色nike鞋感兴趣,并且知识图808指示男性运动鞋和女性运动鞋是可能的提示主题(等等)。

[0120] 然而,可以存在指示用户是对男式运动鞋还是女式运动鞋感兴趣而无需进行询问的额外数据。例如,当前用户与电子市场的交互历史可以指示该用户的大部分或全部购买都是针对与女性相关联的项目。例如,这可能是因为当前用户是执行另一个自我购物任务的女性,或者可能是因为当前用户经常执行送礼物任务,其中预期的目标接收者是女性。此外,常识知识或其他可能相关的外部上下文信息可以通过对话管理器216来调整提示可能性的权重。例如,关于位置、天气、成本、文化和场合的外部数据可以在调整对下一个提示的确定以得到最大敏锐度时起到类似的作用。

[0121] 因此,智能个人助理系统106可以得出用户可能对女性运动鞋比对男性运动鞋更感兴趣的结论,而无需生成确认这一点的提示。因此,对话管理器216可以基于已处理的用户输入以及知识图808继续进行到下一个最可能敏锐的提示话题。在图11B的示例中,假如用户对女性运动鞋感兴趣并且已经指定了品牌和颜色的属性的值,则最佳候选提示可以涉及尚未指定的属性,即样式。

[0122] 因此,对话管理器216可以简单地询问用户“你喜欢什么类型的样式?”然而,该方法没有利用知识图808中关于项目属性值的可用附加知识,无论是来自项目库存数据还是过去的用户交互数据。因此,在一个实施例中,对话管理器可以生成对附加用户输入的提示,该提示还陈述知识图808中可用的备选项,和/或还可以具有知识图808中可用的关联值。

[0123] 例如,提示1202可以替代地询问用户“你喜欢什么类型的样式,比如胶底鞋或跑鞋?”这种类型的问题提示制定既向用户通知可以与成功搜索相关的(例如,由于库存或过去的用户交互行为)建议,也收集附加的用户输入。注意,并不需要建议知识图中的所有已知项目属性值,并且并非条目之间的所有有向边都可以具有指定的分数值。如前所述,智能个人助理系统106可以使用其他数据来为那些更具慧眼的人筛选可能性。

[0124] 此外,对话管理器216甚至可以提供所建议的精确用户输入措辞,当在回复中使用该措辞可能导致相关的搜索结果。例如,提示1202可以替代地询问用户“你想要‘胶底鞋样式’还是‘跑鞋样式’?”。这样的措辞建议可以产生如下的回复(特别是口语回复):该回复具有易于处理的形式的所有剩余的尚未指定的约束(例如,“胶底鞋样式”指定属性值“胶底鞋”和属性“样式”)。

[0125] 在另一示例中,对话管理器216可以具有足以生成进行建议性项目推荐的提示的数据,该数据来自对用户输入的分析 and 来自其他数据。在这种情况下,对话管理器可以具有

指示用户可能对胶底鞋感兴趣的数据。对话管理器216不使用问题类型提示来直接确认,而是可以继续搜索并向用户输出一些可能相关的库存项目的文本和/或图像。因此,提示1204可以宣布“我找到了这些胶底鞋:”并且显示可供购买的特定项目或项目组的图像(或更一般地,表征该特定项目或项目组的图像)。这种方法使得提供不完全约束查询的用户容易确定或否定单个建议类型提示。例如,肯定可以是口头答复或对特定显示项目的选择。

[0126] 在另一示例中,对话管理器216可以选择包括验证陈述的提示,例如“我理解你想寻找红色nike鞋”或“好的,现在我可以帮你寻找红色nike鞋”,以便以会话的方式引导用户提供对用户感兴趣的主要对象的进一步确认和启示性讨论。该提示类型允许用户剖析智能个人助理系统106可能无法自动剖析的模糊性,而无需询问可能引起混淆的问题类型提示。例如,如果在用户文本输入中存在许多不寻常的拼写错误,或者如果在嘈杂的环境中接收到用户的语音,则可能发生这种模糊性,使得规范化不能很好地工作。

[0127] 当用户提供了指示用户兴趣变化的话语时,验证陈述类型提示也可以是特别有用的。也就是说,机器人可以做出验证陈述以允许用户确认新的搜索任务已经开始,并且先前的搜索任务的上下文不再适用。例如,先前寻找红色鞋的机器人可以用“好的,现在让我们寻找伞而不是红色nike鞋”来响应关于伞的用户输入。如果用户不打算改变兴趣,则用户很可能将提供更详细的回复,该回复总结相关的查询术语以使机器人“回到目标”。

[0128] 在另一示例中,对话管理器216可以生成提示,该提示不仅指示在库存中没有找到满足所有指定搜索约束的项目,而且还指示已经经由搜索找到可用的满足一些或大部分指定搜索约束的项目。例如,如果在库存中没有用户查询的红色nike鞋,则对话管理器216可以说“目前没有红色nike鞋,但现在有蓝色或绿色nike鞋”。因此,该提示方法避免了可能导致用户完全失去搜索兴趣的死端(dead-end)结果,并且鼓励用户继续进行已经确定可能成功的略微扩展或修改的搜索。因此,对话管理器216可以鼓励用户“回溯”,并经由相关项目属性值、项目属性或甚至项目类别继续进行搜索。这种提示生成方法对于正在为目标接收者(不熟悉其偏好)浏览或搜索礼物的人来说可能是特别有用的。

[0129] 类似地,如果用户正在寻找黑色nike鞋,但通过搜索确定库存中只有红色、蓝色和绿色nike鞋可用,则询问用户该用户是否对黑色nike鞋感兴趣可能会适得其反并且实际上令人恼火。因此,在一个实施例中,如果这样的提示在被用户的回复确认时将引向库存中不可用的项目,则对话管理器216不生成任何类型的提示。也就是说,该版本的智能在线个人助理106不会主动引导用户进入死端。

[0130] 图13示出了根据一些示例实施例的用于处理自然语言用户输入以生成项目推荐的方法的流程图。该方法可以经由先前所描述的结构元件以及经由计算机中的处理器执行的指令来实现。在1302处,该方法可以从用户接收输入数据。在1304处,该方法可以规范化所接收的输入数据。在1306处,该方法可以解析已规范化的输入数据,例如,以从所解析的输入数据中识别用户感兴趣的主要对象和相关参数。

[0131] 在1308处,该方法可以分析所解析的输入数据以寻找知识图808的维度与主要对象和相关参数之间的匹配。在1310处,该方法可以将分析结果聚合成用于搜索的正式查询。在1312处,该方法可以可选地生成针对来自用户的附加输入数据的一个或多个用户提示。

[0132] 尽管已经参照特定示例实施例描述了本主题,将显而易见的是:可以在不脱离所公开主题的更宽泛的范围的情况下对这些实施例做出各种修改和改变。因此,说明书和附

图应被认为是说明性的而不是限制性的。形成可以实现主题的具体实施例的一部分的附图是作为说明而不是限制的方式示出的。充分详细地描述了所示出的实施例,以使得本领域技术人员能够实践本文中所公开的教导。可以利用并得出其他实施例,从而可以在不脱离本公开的范围的情况下做出结构和逻辑上的替换和改变。因此,该说明书不应当看做是限制意义,并且各种实施例的范围仅通过任意所附权利要求以及权利要求的等同物的全部范围来限定。

[0133] 本发明主题的这些实施例被单独地和/或统一地由术语“发明”来指代,其仅是为了方便,而不是旨在主动将本申请的范围限制为任意单个发明或发明构思(如果实际上不止一个被公开的话)。因此,尽管本文示出并描述了特定实施例,应当理解,适于实现相同目的的任意设置都可以用于替换所示出的特定实施例。本公开旨在覆盖各种实施例的任意和所有的适应性修改或变化。通过研究上述内容,上述实施例的组合以及本文中没有具体描述的其它实施例对于本领域技术人员来说将是明显的。

[0134] 下面的编号示例是实施例。

[0135] 1、一种用于在多轮对话中生成对附加自然语言输入的提示的方法,所述方法包括:

[0136] 接收知识图中的维度与对用户查询数据的分析的结果之间的已排名匹配,所述知识图维度至少包括类别、属性和属性值各一个,并且所述结果包括用户感兴趣的主要对象、用户意图和相关参数;

[0137] 搜索库存并且将搜索结果并入到所述知识图中;

[0138] 确定是否实现所述分析的结果与直接或间接链接到所述主要对象的知识图维度之间的预定的足够级别的匹配;以及

[0139] 如果已经实现了所述足够级别的匹配,则基于所述搜索结果来生成并输出推荐类型提示。

[0140] 2、根据示例1所述的方法,其中,对话中的提示的数量受限于预定的最大值,所述预定的最大值是通过将常识知识并入到所述分析中以更好地推断用户意图来最小化的。

[0141] 3、根据示例1或示例2所述的方法,其中,所述推荐类型提示推荐电子市场中待售的至少一个项目。

[0142] 4、根据示例1至3中任一项所述的方法,还包括:如果尚未实现所述足够级别的匹配但是已经实现近乎足够级别的匹配,则替代地生成并输出提供多个推荐的不同的推荐类型提示,其中所述多个推荐中的每个推荐跨越至少一个未指定的已链接知识图维度。

[0143] 5、根据示例1至4中任一项所述的方法,其中,所述推荐类型提示包括至少一个表征电子市场中待售的项目的图像。

[0144] 6、根据示例1至5中任一项所述的方法,其中,所述推荐类型提示包括文本输出和口语输出中的一项。

[0145] 7、根据示例1至6中任一项所述的方法,还包括:如果在所述对话期间改变所述主要对象,则替代地生成并输出关于所述主要对象的验证陈述类型提示,并且忽略来自之前的搜索任务的上下文数据,其中,改变所述主要对象表示搜索任务的改变。

[0146] 8、一种计算机可读存储介质,所述计算机可读存储介质中嵌入有指令集,所述指令集当由计算机的一个或多个处理器执行时,使所述计算机执行用于在多轮对话中生成对

附加自然语言输入的提示的下列操作：

[0147] 接收知识图中的维度与对用户查询数据的分析的结果之间的已排名匹配,所述知识图维度至少包括类别、属性和属性值各一个,并且所述结果包括用户感兴趣的主要对象、用户意图和相关参数;

[0148] 搜索库存并且将搜索结果并入到所述知识图中;

[0149] 确定是否实现所述分析的结果与直接或间接链接到所述主要对象的知识图维度之间的预定的足够级别的匹配;以及

[0150] 如果已经实现了所述足够级别的匹配,则基于所述搜索结果来生成并输出推荐类型提示。

[0151] 9、根据示例8所述的介质,其中,对话中的提示的数量受限于预定的最大值,所述预定的最大值是通过将常识知识并入到所述分析中以更好地推断用户意图来最小化的。

[0152] 10、根据示例8或示例9所述的介质,其中,所述推荐类型提示推荐电子市场中待售的至少一个项目。

[0153] 11、根据示例8至10中任一项所述的介质,还包括:如果尚未实现所述足够级别的匹配但是已经实现近乎足够级别的匹配,则替代地生成并输出提供多个推荐的不同的推荐类型提示,其中所述多个推荐中的每个推荐跨越至少一个未指定的已链接知识图维度。

[0154] 12、根据示例8至11中任一项所述的介质,其中,所述推荐类型提示包括至少一个表征电子市场中待售的项目的图像。

[0155] 13、根据示例8至12中任一项所述的介质,其中,所述推荐类型提示包括文本输出和口语输出中的一项。

[0156] 14、根据示例8至13中任一项所述的介质,还包括:如果在所述对话期间改变所述主要对象,则替代地生成并输出关于所述主要对象的验证陈述类型提示,并且忽略来自之前的搜索任务的上下文数据,其中,改变所述主要对象表示搜索任务的改变。

[0157] 15、一种用于在多轮对话中生成对附加自然语言输入的提示的系统,所述系统包括:

[0158] 自然语言理解组件,被配置为提供知识图中的维度与对用户查询数据的分析的结果之间的已排名匹配,所述知识图维度至少包括类别、属性和属性值各一个,并且所述结果包括用户感兴趣的主要对象、用户意图和相关参数;

[0159] 搜索组件,被配置为搜索库存并且将搜索结果并入到所述知识图中;

[0160] 对话管理器组件,被配置为确定是否实现所述分析的结果与直接或间接链接到所述主要对象的知识图维度之间的预定的足够级别的匹配;以及

[0161] 如果已经实现了所述足够级别的匹配,则由所述对话管理器组件基于所述搜索结果来生成并输出推荐类型提示。

[0162] 16、根据示例15所述的系统,其中,对话中的提示的数量受限于预定的最大值,所述预定的最大值是通过将常识知识并入到所述分析中以更好地推断用户意图来最小化的。

[0163] 17、根据示例15或示例16所述的系统,其中,所述推荐类型提示推荐电子市场中待售的至少一个项目。

[0164] 18、根据示例15至17中任一项所述的系统,还包括:如果尚未实现所述足够级别的匹配但是已经实现近乎足够级别的匹配,则替代地由所述对话管理器组件生成并输出提供

多个推荐的不同的推荐类型提示,其中所述多个推荐中的每个推荐跨越至少一个未指定的已链接知识图维度。

[0165] 19、根据示例15至18中任一项所述的系统,其中,所述推荐类型提示包括至少一个表征电子市场中待售的项目的图像。

[0166] 20、根据示例15至19中任一项所述的系统,还包括:如果在所述对话期间改变所述主要对象,则替代地由所述对话管理器组件生成并输出关于所述主要对象的验证陈述类型提示,并且忽略来自之前的搜索任务的上下文数据,其中,改变所述主要对象表示搜索任务的改变。

[0167] 21、一种承载指令集的机器可读介质,所述指令集当由计算机的一个或多个处理器执行时,使所述计算机执行根据示例1至7中任一项所述的方法。

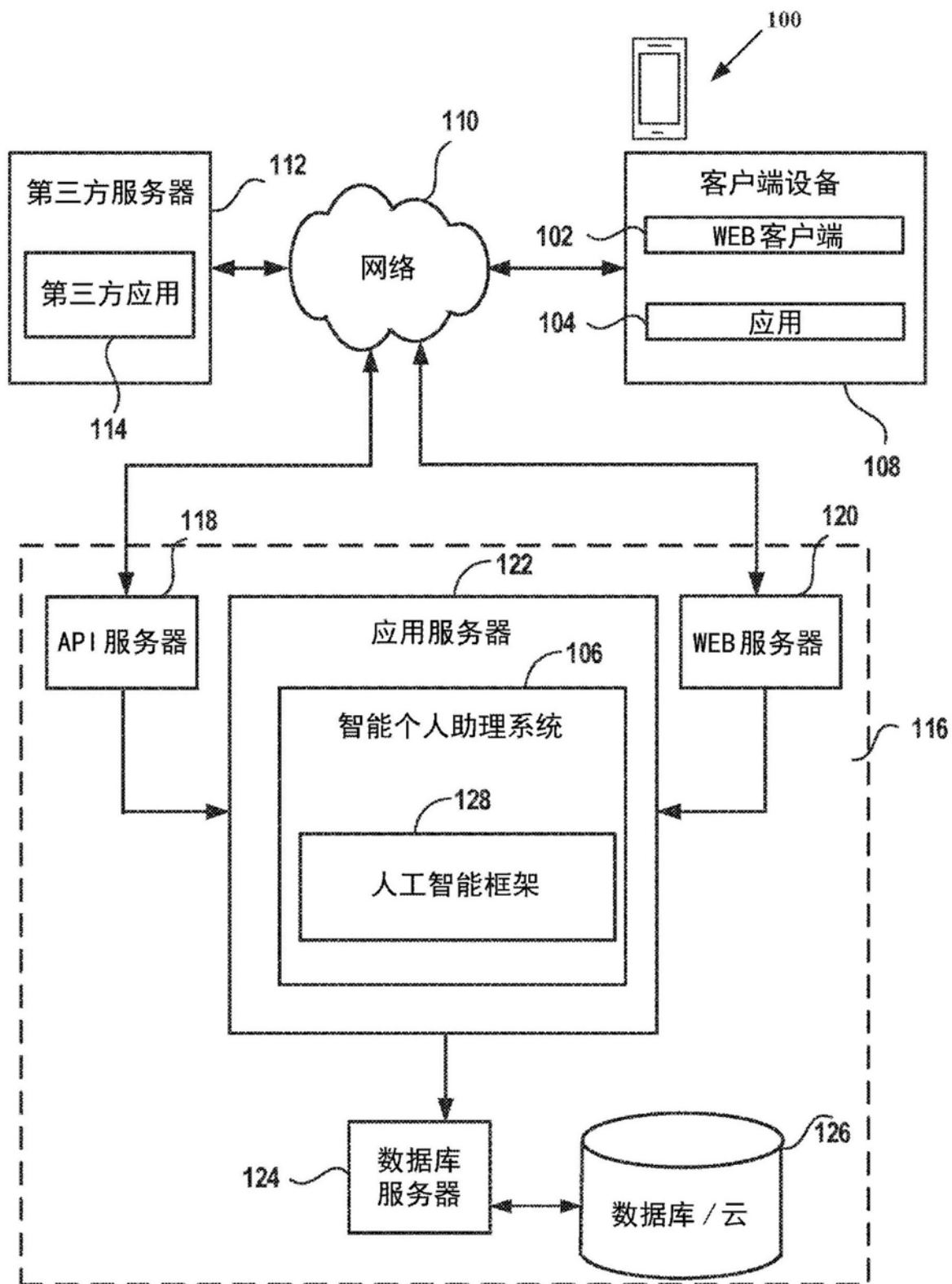


图1

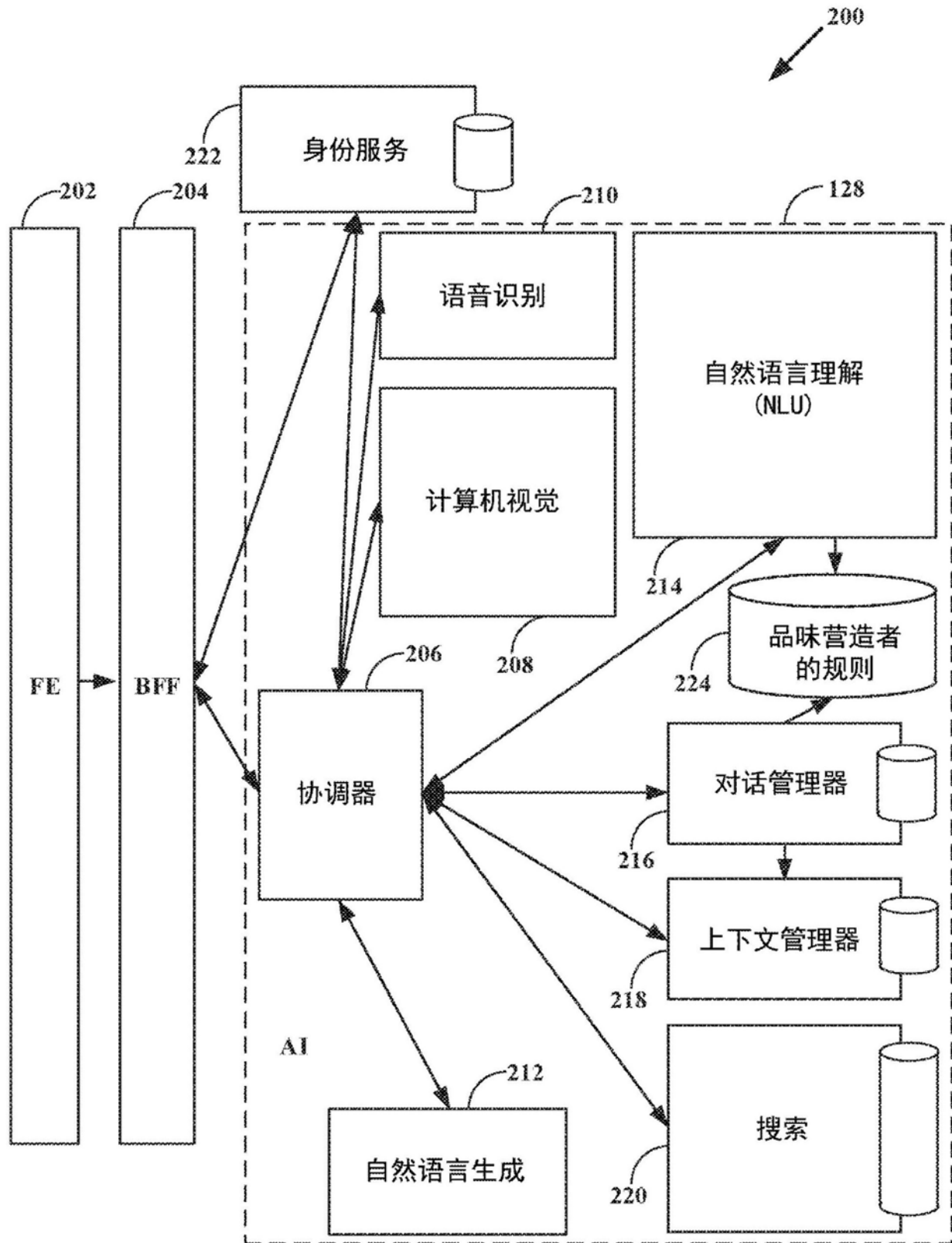


图2

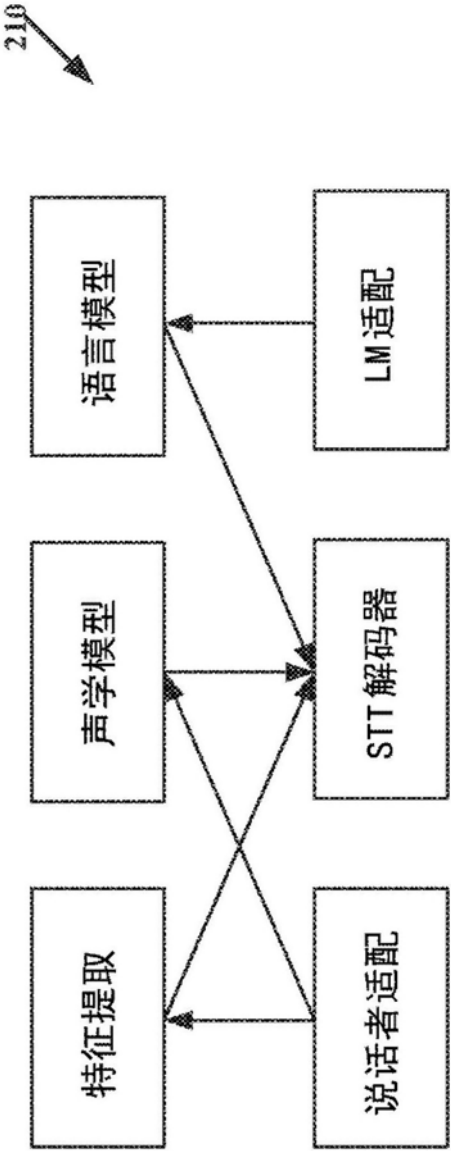


图3A

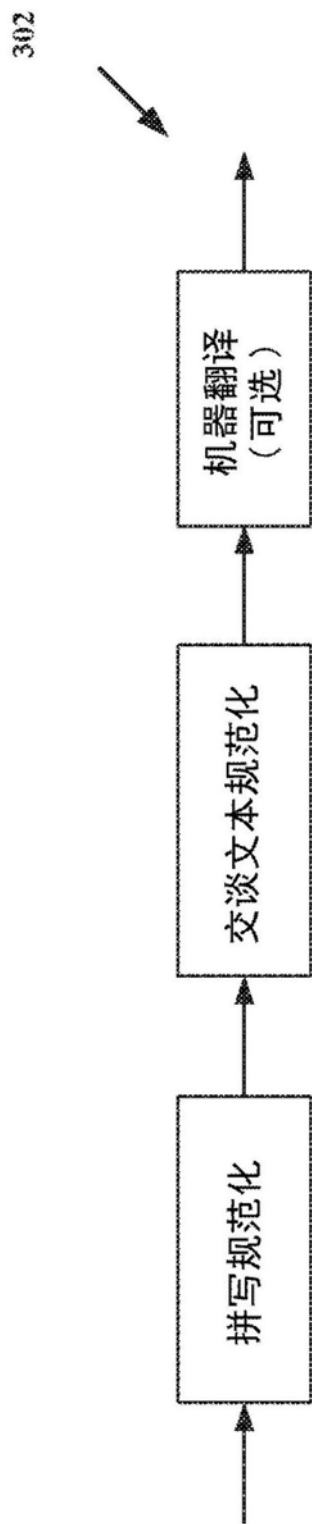


图3B

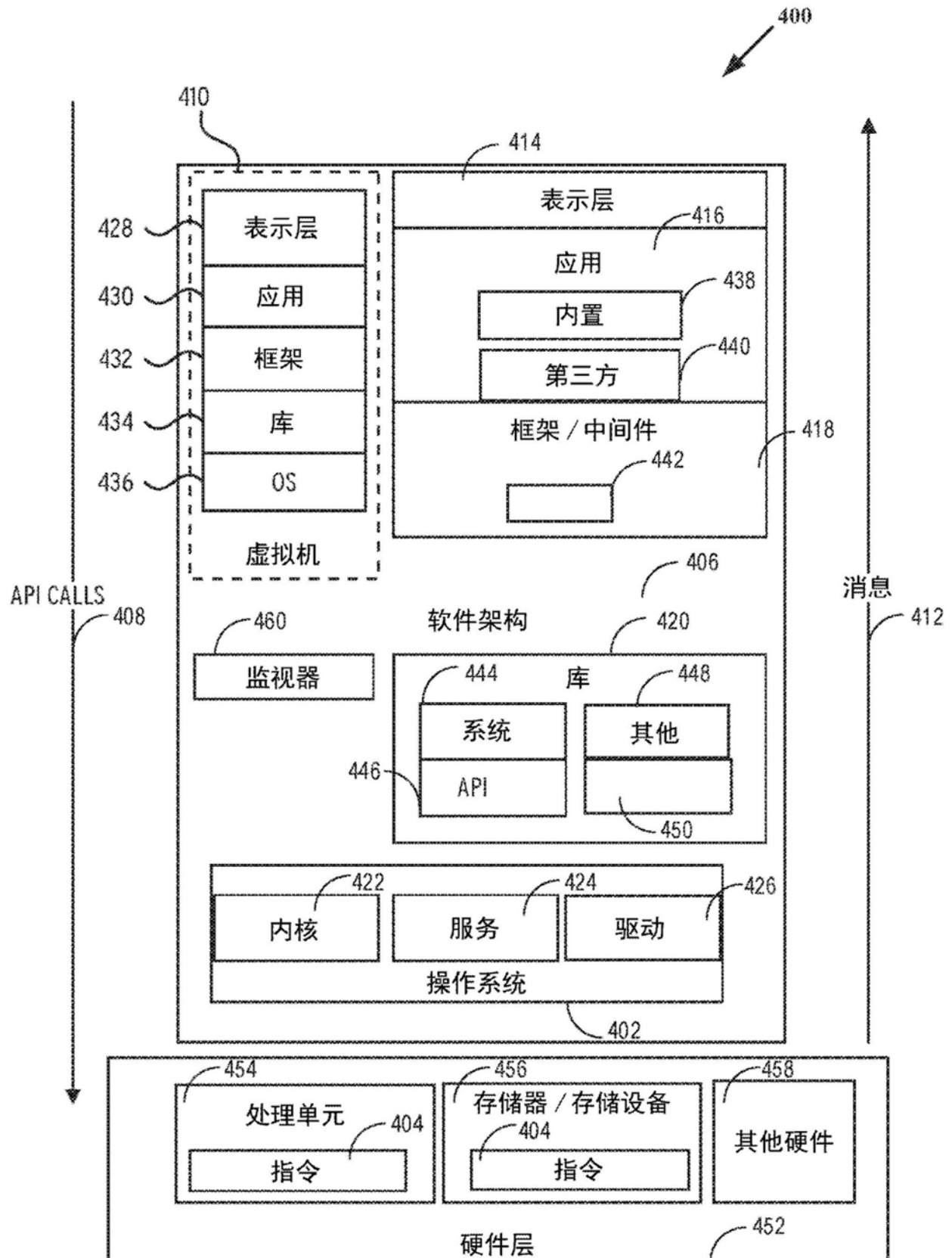


图4

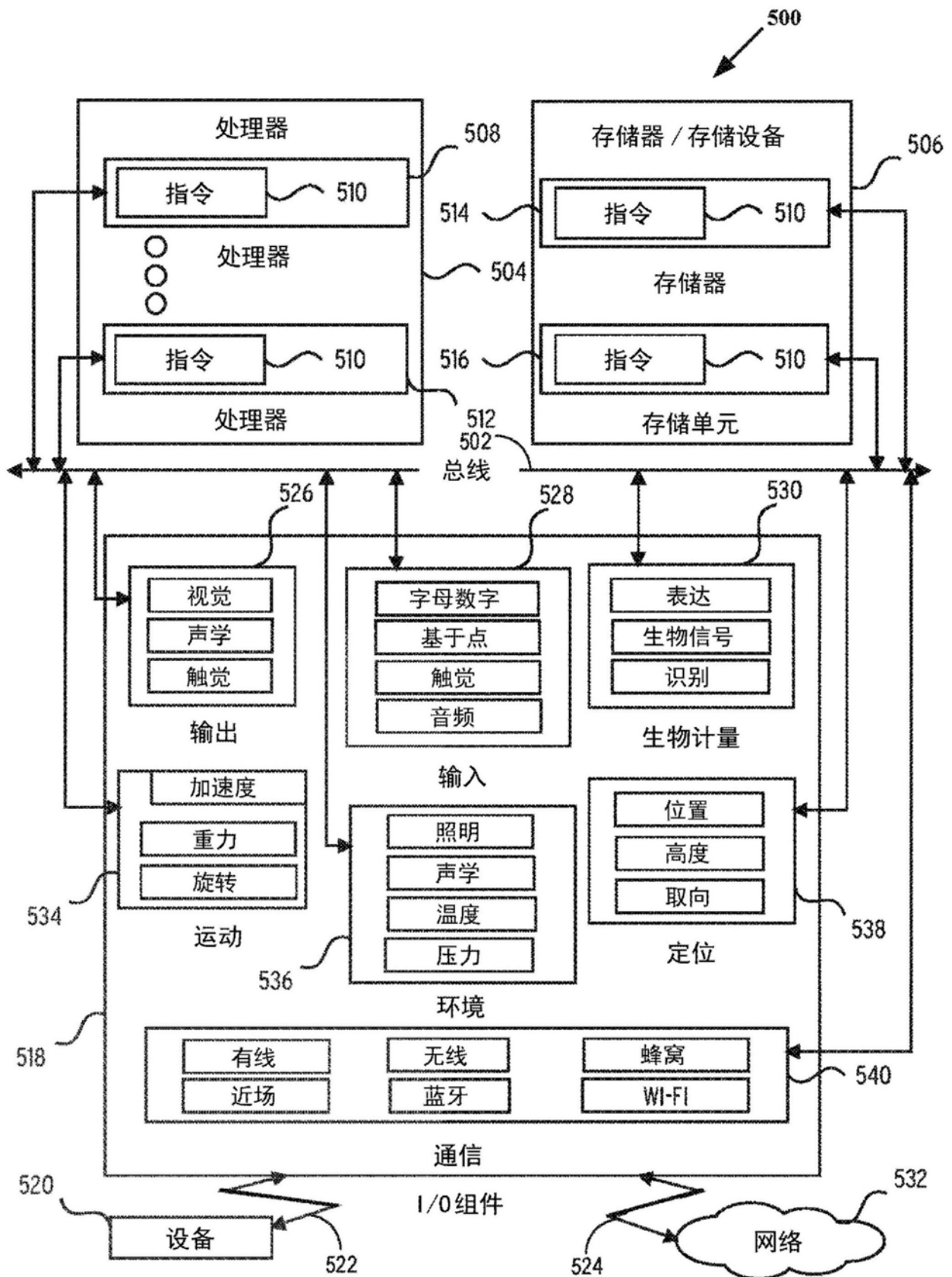


图5

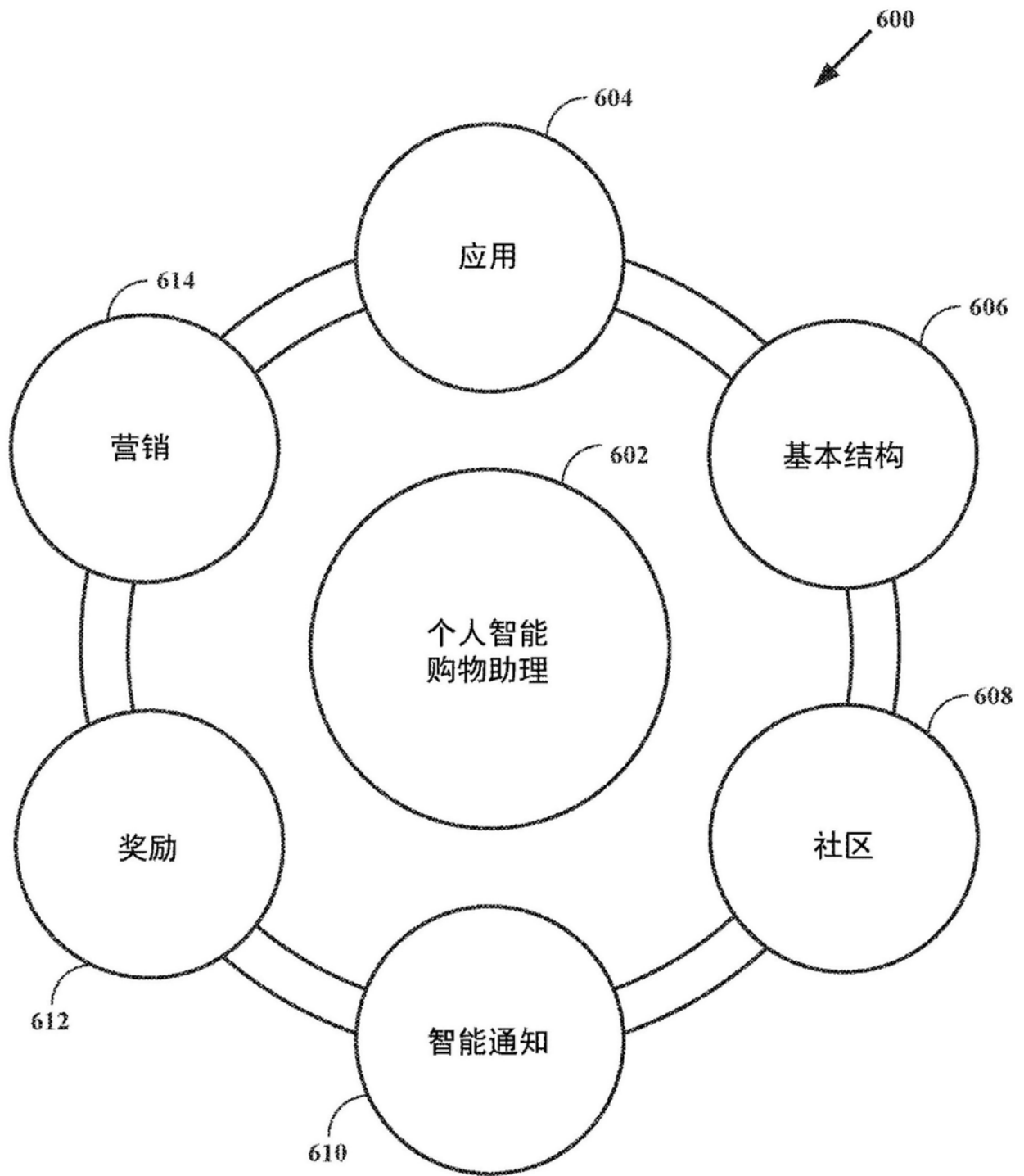


图6

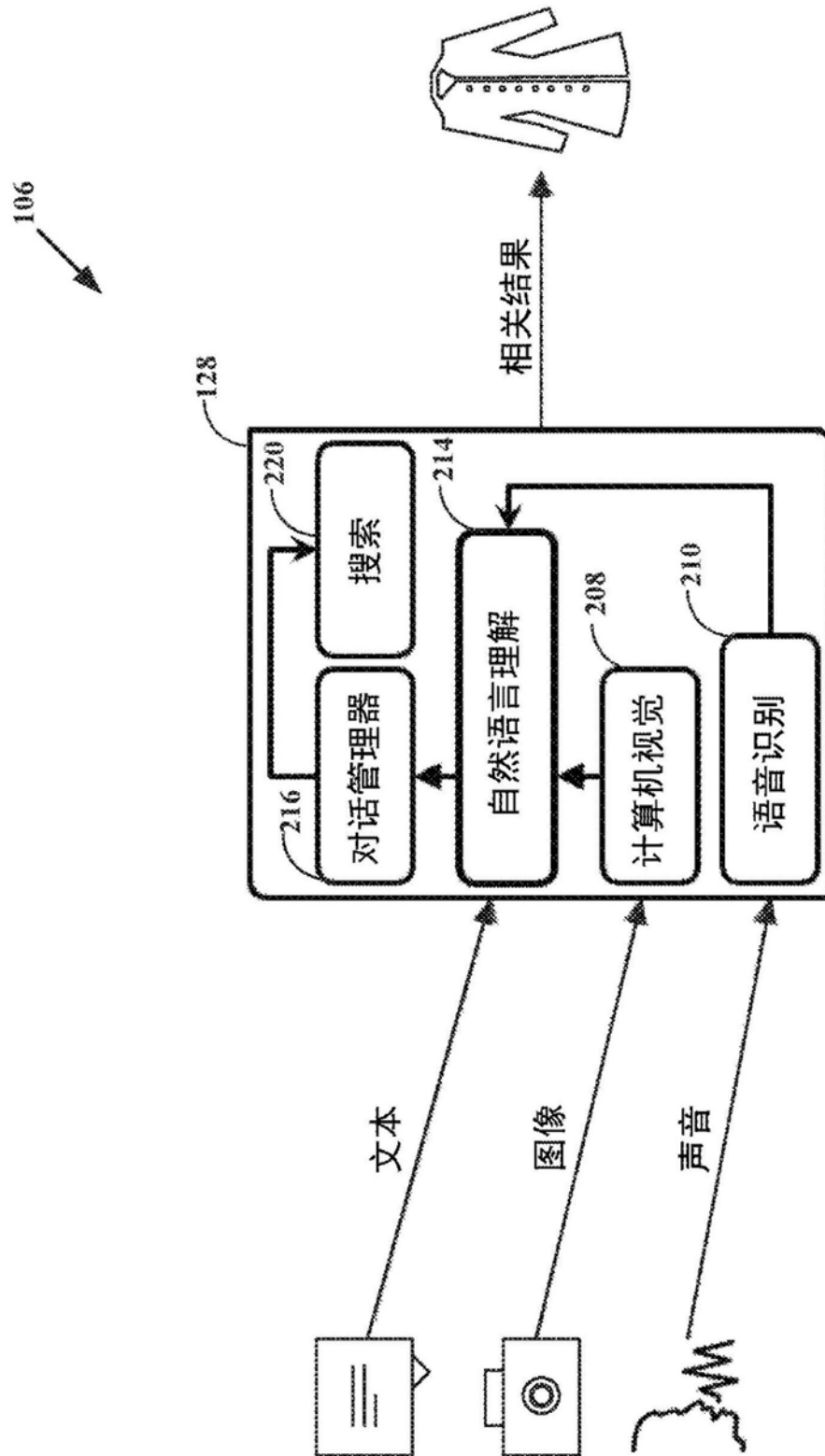


图7

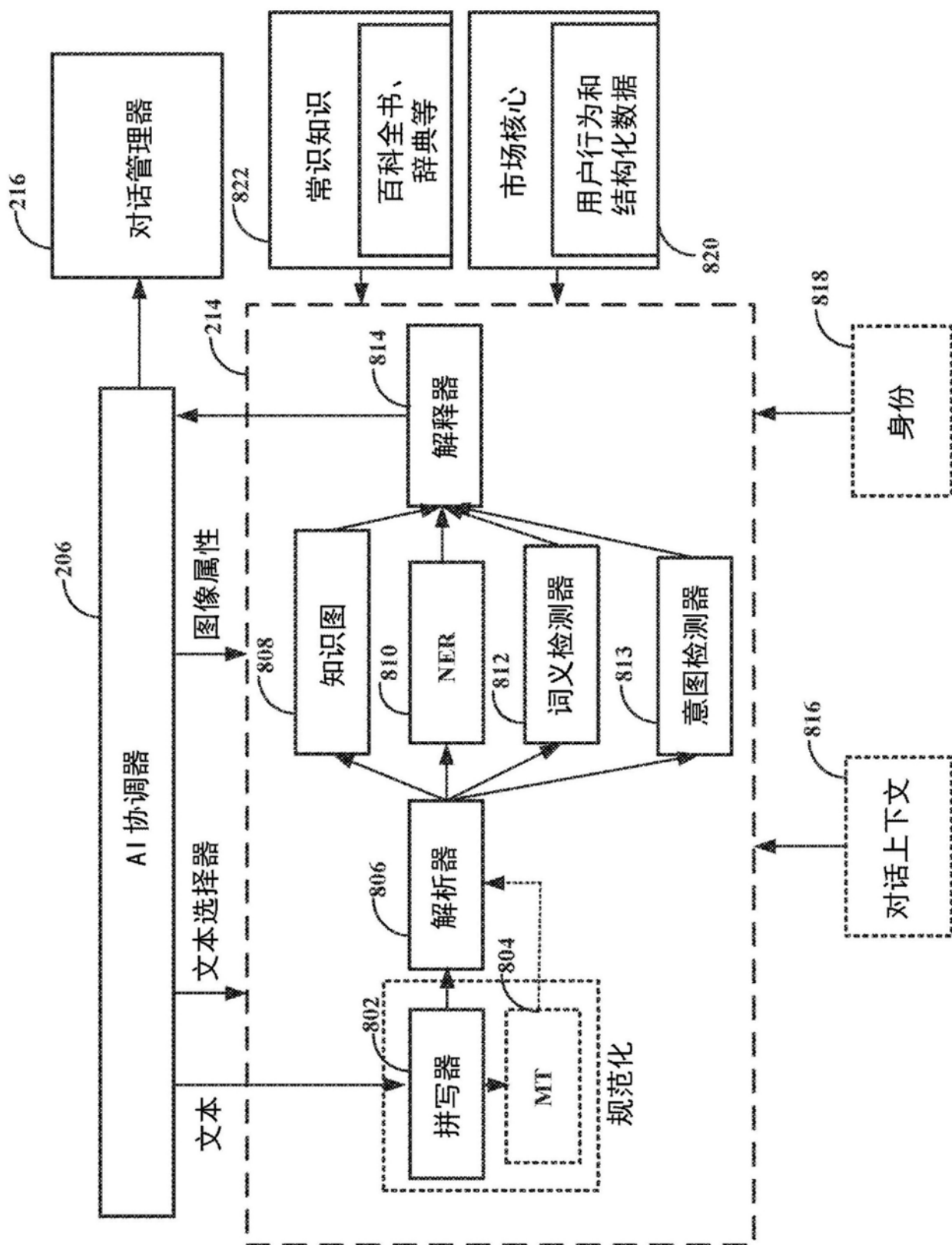


图8

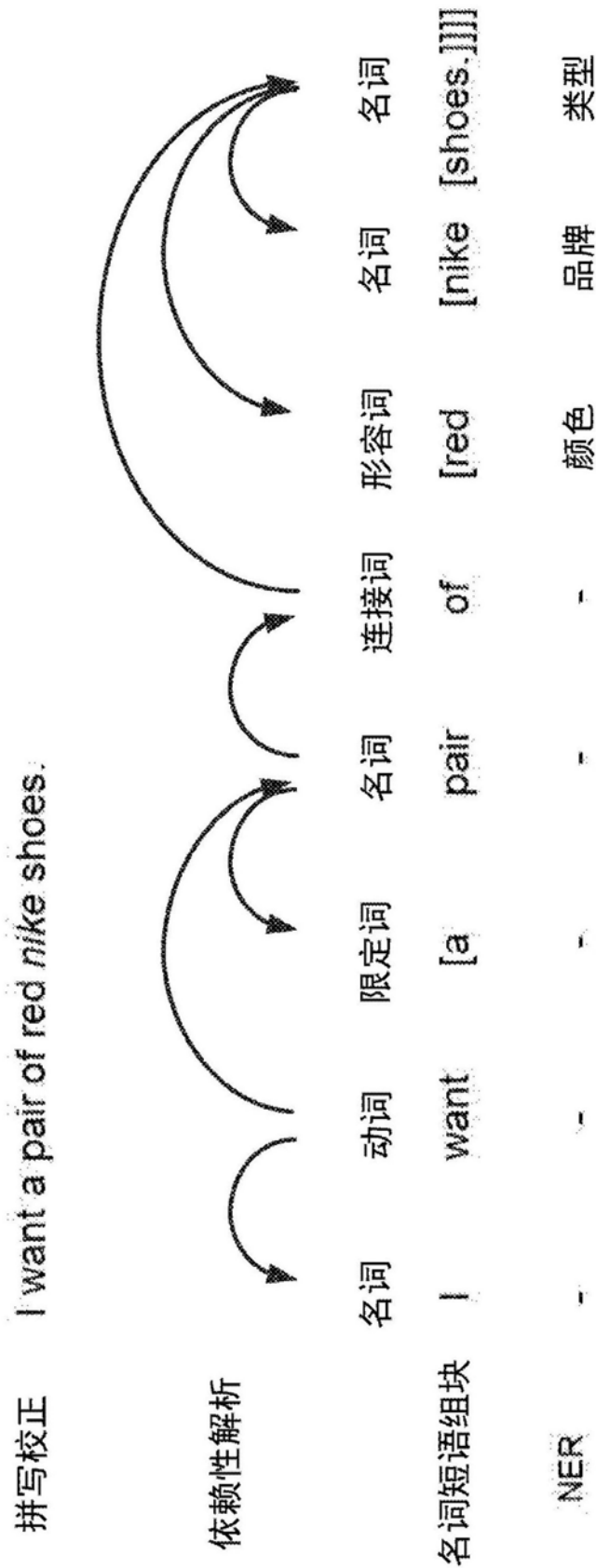


图9

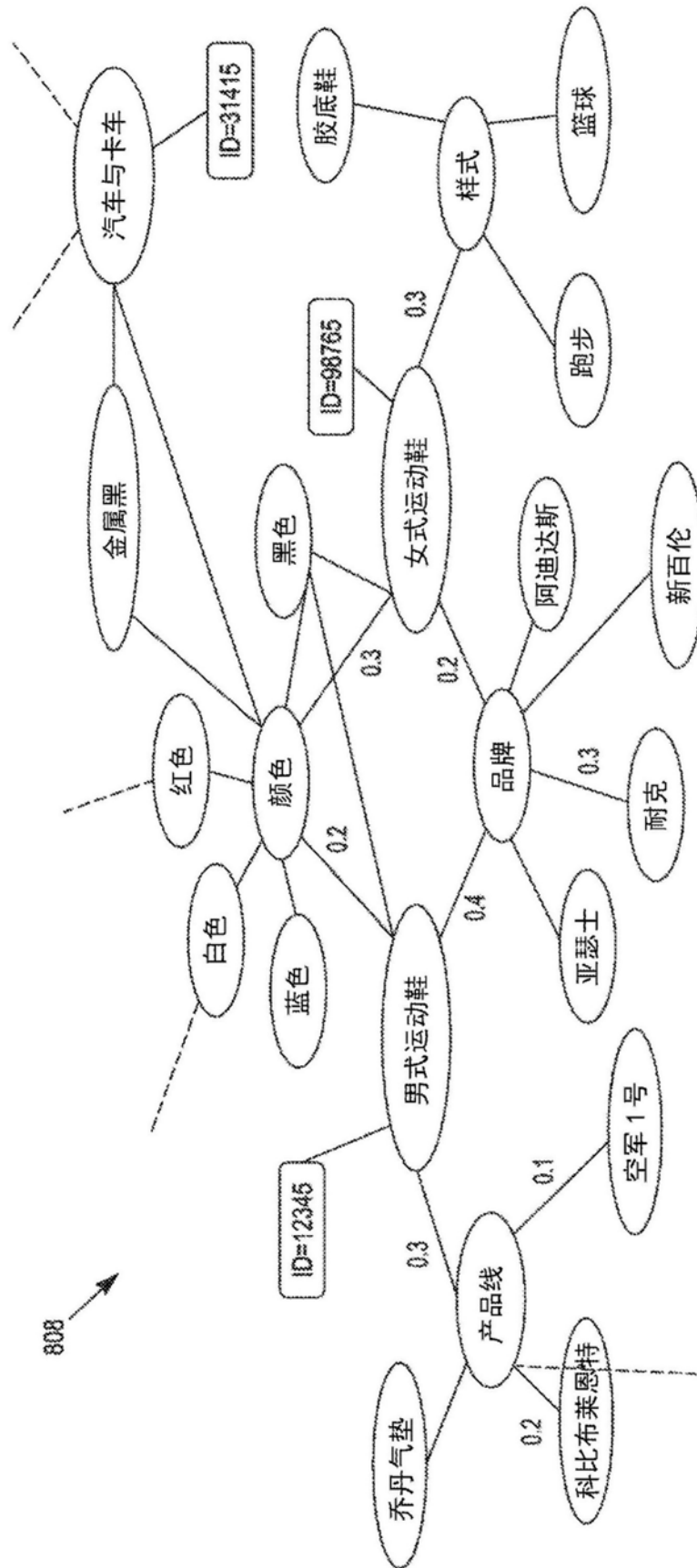


图10

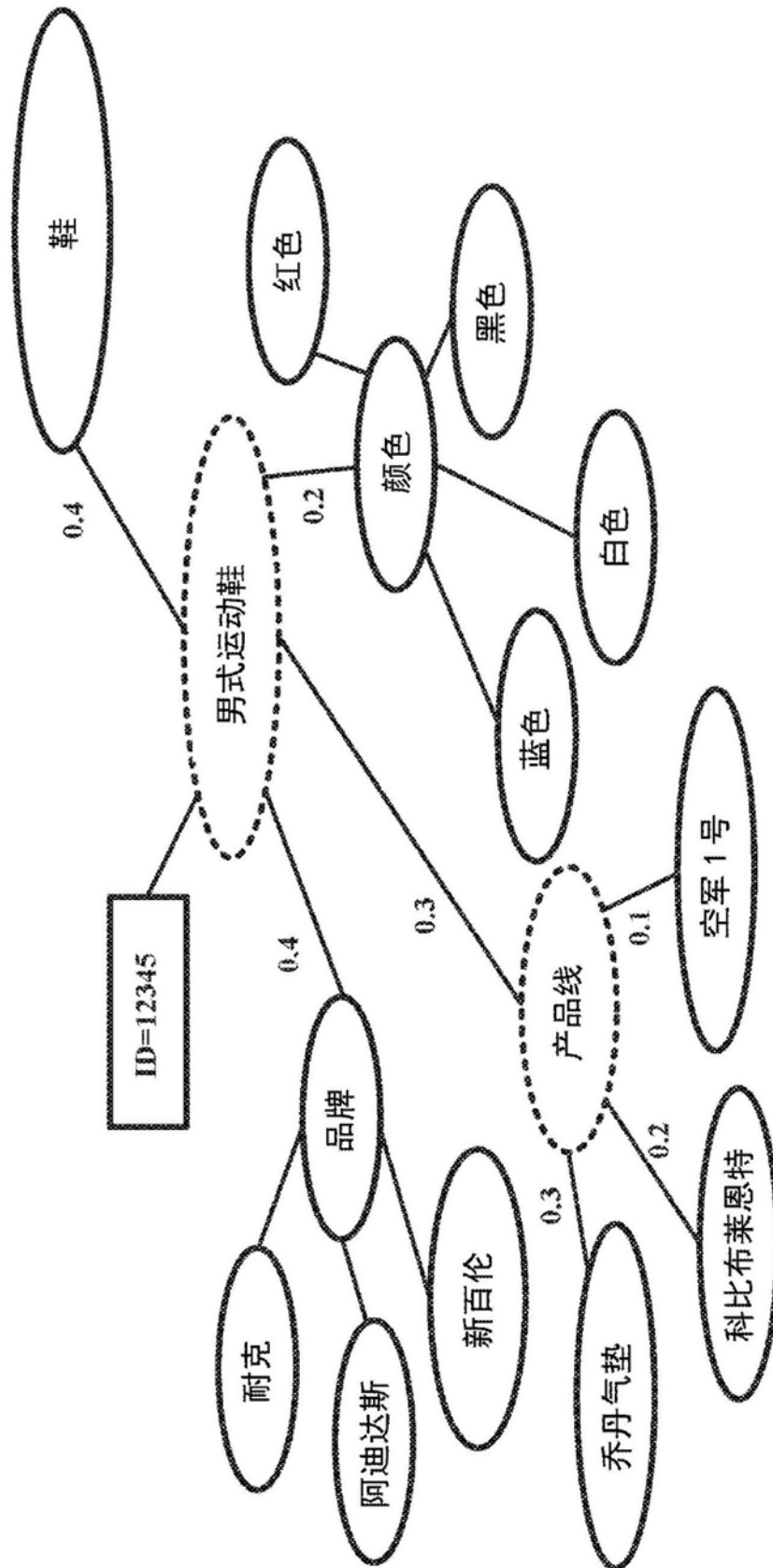


图11A

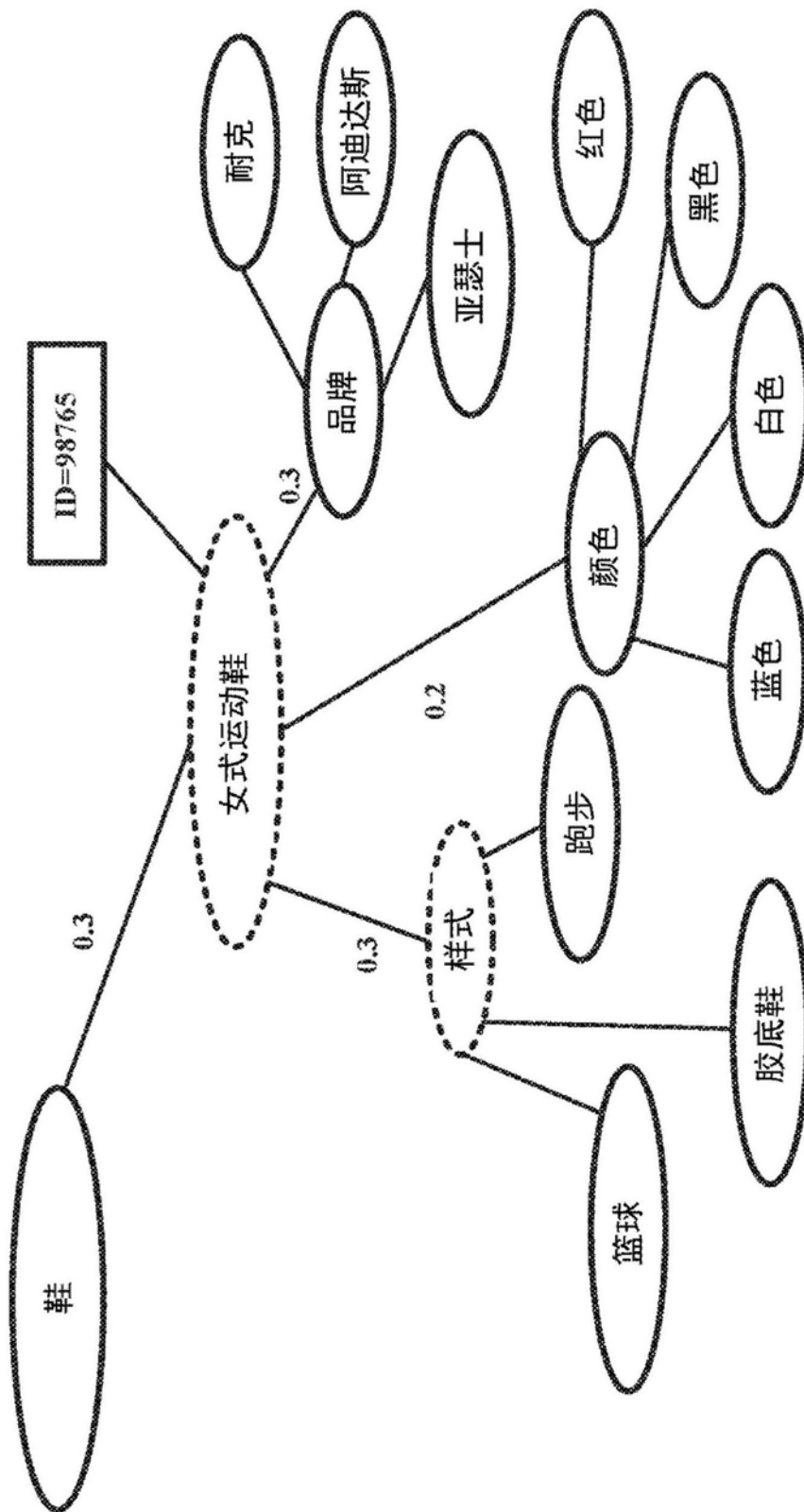


图11B

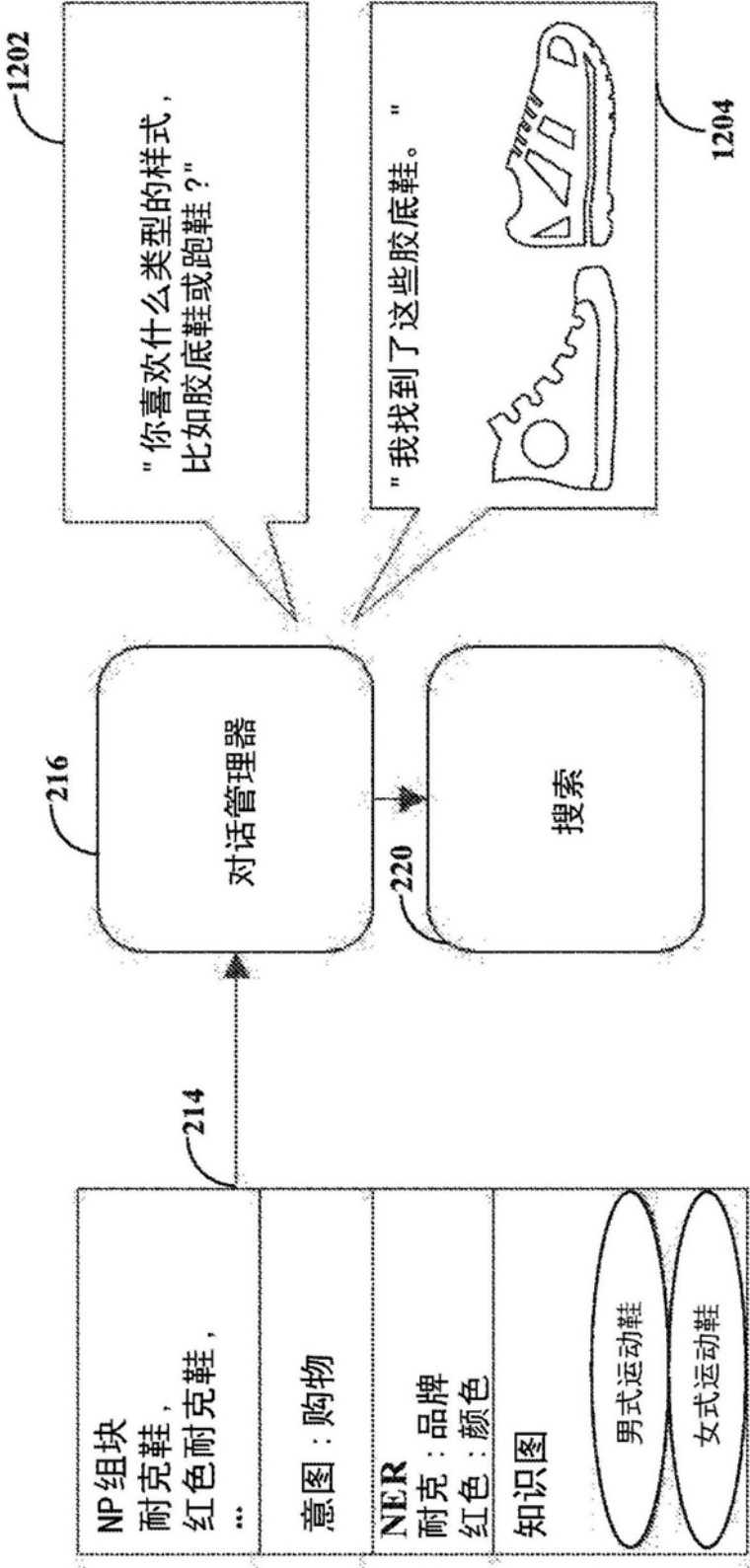


图12

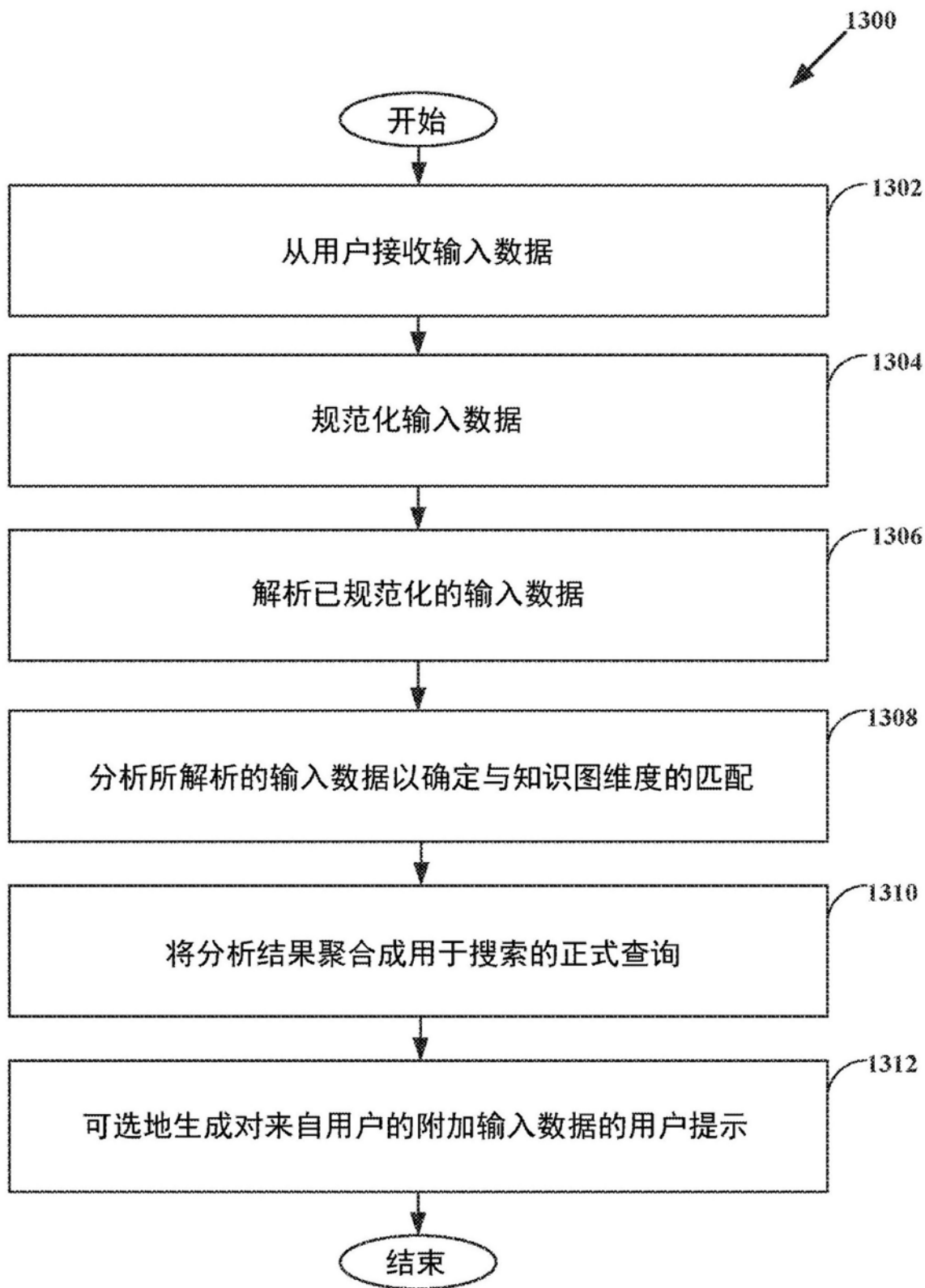


图13