



(12) 发明专利申请

(10) 申请公布号 CN 117062695 A

(43) 申请公布日 2023. 11. 14

(21) 申请号 202280024436.7

(22) 申请日 2022.03.02

(30) 优先权数据

63/156,832 2021.03.04 US

(85) PCT国际申请进入国家阶段日

2023.09.25

(86) PCT国际申请的申请数据

PCT/US2022/018567 2022.03.02

(87) PCT国际申请的公布数据

W02022/187395 EN 2022.09.09

(71) 申请人 特优英特智能公司

地址 美国马萨诸塞州

(72) 发明人 J·A·格林斯坦

A·Z·科索夫斯基-萨克斯

(74) 专利代理机构 北京市铸成律师事务所

11313

专利代理师 王珺 李文颖

(51) Int.Cl.

B25J 9/16 (2006.01)

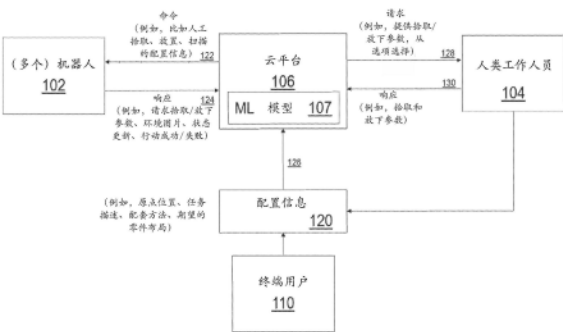
权利要求书2页 说明书12页 附图6页

(54) 发明名称

机器人系统

(57) 摘要

本公开总体上涉及机器人系统,并且更具体地涉及用于包括按需智能组件的机器人平台的系统和方法。一种用于操作机器人的示例性计算机实现的方法包括:获得给机器人的指令,其中,指令与第一用户相关联;基于指令识别任务;将任务传输给机器人;从机器人接收与任务相关联的请求;确定请求是否可以通过一个或多个经过训练的机器学习算法解决;如果请求不可以通过一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;接收来自第二用户的对查询的响应;以及使得任务由机器人基于响应来执行。



1. 一种用于操作机器人的计算机实现的方法,所述方法包括:
获得给所述机器人的指令,其中,所述指令与第一用户相关联;
基于所述指令识别任务;
将所述任务传输给所述机器人;
从所述机器人接收与所述任务相关联的请求;
确定所述请求是否能够通过一个或多个经过训练的机器学习算法解决;
如果所述请求不能够通过所述一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;
接收来自所述第二用户的对所述查询的响应;以及
使得所述任务由所述机器人基于所述响应来执行。
2. 如权利要求1所述的方法,其中,所述指令是自然语言指令,并且所述自然语言指令指导所述机器人拾取和/或放置一个或多个物体。
3. 如权利要求1至2中任一项所述的方法,其中,所述任务包括多个子任务,并且其中,所述多个子任务包括拾取子任务和放下子任务。
4. 如权利要求3所述的方法,其中,所述请求包括图像和对一个或多个拾取参数的查询。
5. 如权利要求4所述的方法,其中,所述拾取参数包括拾取点、抓取角度、抓取深度或以上各项的任意组合。
6. 如权利要求3所述的方法,其中,所述请求包括图像和对一个或多个放下参数的查询。
7. 如权利要求6所述的方法,其中,所述放下参数包括放下点、旋转角度、放下高度或以上各项的任意组合。
8. 如权利要求1至7中任一项所述的方法,其中,确定所述请求是否能够通过一个或多个经过训练的机器学习算法解决包括:
将所述机器人捕获的数据输入到一个或多个经过训练的机器学习模型,以获得解决方案和与所述解决方案相关联的置信度得分;
确定所述置信度得分是否低于预定义阈值。
9. 如权利要求8所述的方法,其中,所述查询包括由所述一个或多个机器学习模型给出的所述解决方案。
10. 如权利要求8所述的方法,其中,所述机器人捕获的所述数据包括图像。
11. 如权利要求8所述的方法,进一步包括:基于来自所述第二用户的所述响应来训练所述一个或多个机器学习模型。
12. 如权利要求9所述的方法,进一步包括:使得显示第一图形用户界面,所述第一图形用户界面包括由所述机器人捕获的一个或多个图像。
13. 如权利要求11所述的方法,其中,所述第一图形用户界面进一步包括:
对由所述一个或多个机器学习模型给出的所述解决方案的指示;以及
用于接受所述解决方案的一个或多个用户界面控件。
14. 如权利要求1至13中任一项所述的方法,进一步包括:使得在所述第一用户的电子设备上显示第二图形用户界面,用于接收所述指令。

15. 如权利要求14所述的方法,其中,所述第二图形用户界面基于所述机器人的目标应用来选择。

16. 一种电子设备,包括:

一个或多个处理器;

存储器;以及

一个或多个程序,其中,所述一个或多个程序存储在所述存储器中,并且被配置成由所述一个或多个处理器执行,所述一个或多个程序包括用于以下操作的指令:

获得给机器人的指令,其中,所述指令与第一用户相关联;

基于所述指令识别任务;

向所述机器人传输所述任务;

从所述机器人接收与所述任务相关联的请求;

确定所述请求是否能够通过一个或多个经过训练的机器学习算法解决;

如果所述请求不能够通过所述一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;

接收来自所述第二用户的对所述查询的响应;以及

使得所述任务由所述机器人基于所述响应来执行。

17. 一种非暂时性计算机可读存储介质,存储有一个或多个程序,所述一个或多个程序包括指令,当由电子设备的一个或多个处理器执行时,所述指令使得所述电子设备执行:

获得给机器人的指令,其中,所述指令与第一用户相关联;

基于所述指令识别任务;

向所述机器人传输所述任务;

从所述机器人接收与所述任务相关联的请求;

确定所述请求是否能够通过一个或多个经过训练的机器学习算法解决;

如果所述请求不能够通过所述一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;

接收来自所述第二用户的对所述查询的响应;以及

使得所述任务由所述机器人基于所述响应来执行。

机器人系统

[0001] 相关申请的交叉引用

[0002] 本申请要求于2021年3月4日提交的美国临时申请63/156,832的优先权,其全部内容通过引用并入本文,用于所有目的。

技术领域

[0003] 本公开总体上涉及机器人系统,并且更具体地涉及用于包括按需智能组件的机器人平台的系统和方法。

背景技术

[0004] 传统自动化系统可以被编程为在严格结构化的环境中与有限数量的特定物体一起工作。然而,生产环境中的机器人会遇到各种各样的物体、环境和任务,这使得基于经典计算机视觉或监督机器学习技术的解决方案变得外强中干,并且范围有限。例如,它们不适合在非结构化环境中操纵(例如,拾取、放置)任意物体。此外,传统的自动化系统开发和集成的成本很高,并且容易受到工作流程变化的影响。因此,只有不到5%的仓库使用先进的自动化技术来操纵物体。

[0005] 因此,公司仍然依靠人工来拾取和包装产品。人类拣货员难以培训和雇佣,而且他们经常超负荷工作、昂贵且不可预测。因此,仓库和物流网络因人类工作人员的次优表现而遭遇瓶颈。

发明内容

[0006] 本公开包括一种机器人平台,其包括按需智能组件。智能组件可以包括执行机器人工作流程(pipeline)的未解决部分的远程人类工作人员池,允许机器人例如在非结构化环境中处理任意物体。不是让人类工作人员手动操作机器人,而是人类工作人员对与机器人流水线的子组件相对应的结构良好的查询提供答案。智能组件以(例如,与雇用人类工作人员或开发定制的自动化解决方案相比)极低的成本为比如拾取和放置之类的非结构化任务提供了即时解决方案。此外,它具有高速度、多功能、灵活性和低成本集成的优势。机器人平台可以用于自动化任何任务,比如仓库管理、制造、交付、检查、物流等任务。

[0007] 此外,由智能组件产生的数据可以反过来用于训练受监督的机器学习模型,以逐步将人类工作人员排除在循环之外。例如,该系统可以聚集人类决策的数据集,将棘手的机器人问题转化为可解决的监督学习问题。

[0008] 一种用于操作机器人的示例性计算机启动方法,包括:接收机器人的自然语言指令,其中,自然语言指令与第一用户相关联;基于自然语言指令识别包括多个子任务的任务;确定多个子任务中的每一者是否需要第二用户的输入;根据特定子任务需要第二用户输入的确定,生成预定义格式的查询;将查询传输到第二用户的电子设备;接收来自第二用户的对查询的响应;以及使得给定任务由机器人基于响应来执行。

[0009] 在一些实施例中,自然语言指令指导机器人拾取和放置一个或多个物体。

[0010] 在一些实施例中,多个子任务包括拾取子任务和放下子任务。子任务还可以包括扫描(环境或预先指定的原点位置)、推动、条形码读取或在笛卡尔或关节空间中执行任意轨迹。

[0011] 在一些实施例中,查询包括图像和对一个或多个选择参数的查询。

[0012] 在一些实施例中,拾取参数包括拾取点、抓取角度、抓取深度或它们的任意组合。

[0013] 在一些实施例中,查询包括图像和对一个或多个放下参数的查询。

[0014] 在一些实施例中,放下参数包括放下点、旋转角度、放下高度或它们的任意组合。

[0015] 在一些实施例中,确定多个子任务中的每一者是否需要第二用户的输入包括:将机器人捕获的数据输入到机器学习模型,以获得解决方案和与参数相关联的置信度得分;根据置信度得分超过预定义阈值的确定,使机器人基于该解决方案执行特定子任务;以及根据置信度得分没有超过预定义阈值的确定,生成查询。

[0016] 在一些实施例中,由机器人捕获的数据包括图像。

[0017] 在一些实施例中,方法进一步包括:基于第二用户的响应来训练机器学习模型。

[0018] 一种示例性电子设备包括:一个或多个处理器;存储器;以及一个或多个程序,其中,一个或多个程序存储在存储器中,并且被配置成由该一个或多个处理器执行,该一个或多个程序包括用于以下操作的指令:接收机器人的自然语言指令,其中,自然语言指令与第一用户相关联;基于自然语言指令识别包括多个子任务的任务;确定多个子任务中的每一者是否需要第二用户的输入;根据特定子任务需要第二用户输入的确定,生成预定义格式的查询;将查询传输到第二用户的电子设备;接收来自第二用户的对查询的响应;以及使得给定任务由机器人基于响应来执行。

[0019] 一种示例性非暂时性计算机可读存储介质,存储有一个或多个程序,一个或多个程序包括指令,当由电子设备的一个或多个处理器执行时,指令使电子设备:接收机器人的自然语言指令,其中,自然语言指令与第一用户相关联;基于自然语言指令识别包括多个子任务的任务;确定多个子任务中的每一者是否需要第二用户的输入;根据特定子任务需要第二用户输入的确定,生成预定义格式的查询;将查询传输到所述第二用户的电子设备;接收来自第二用户的对查询的响应;以及使得给定任务由机器人基于响应来执行。

附图说明

[0020] 图1A示出了根据一些实施例的示例性机器人平台。

[0021] 图1B示出了根据一些实施例的示例性工作流程。

[0022] 图2示出了根据一些实施例的示例性电子设备。

[0023] 图3示出了根据一些实施例的示例性用户界面。

[0024] 图4示出了根据一些实施例的示例性用户界面。

[0025] 图5示出了根据不同示例的用于操作机器人的过程。

具体实施方式

[0026] 本公开包括一种机器人平台,该机器人平台包括按需智能组件。智能组件可以包括执行机器人工作流程的未解决部分的远程人类工作人员池,允许机器人在非结构化环境中处理任意物体。智能组件以(例如,与雇用人类工人或开发定制的自动化解决方案相比)

极低的成本为拾取和放置等非结构化任务提供了即时解决方案。此外,它具有高速度、多功能、灵活性和低成本集成的优势。机器人平台可以用于自动化任何任务,比如仓库管理、制造、交付、检查、物流等任务。

[0027] 此外,由智能部分产生的数据可以反过来用于训练受监督的机器学习模型,以逐步将人类工作人员排除在循环之外。例如,该系统可以聚集人类决策的数据集,将棘手的机器人问题转化为可解决的监督学习问题。

[0028] 给出以下描述是为了使本领域普通技术人员能够制造并使用各种实施例。特定设备、技术和应用的描述仅作为示例提供。对于本领域普通技术人员来说,对本文描述的示例的各种修改将是显而易见的,并且在不脱离各种实施例的精神和范围的情况下,本文定义的一般原理可以应用于其他示例和应用。因此,各种实施例不旨在限于本文描述并示出的示例,而是旨在符合与权利要求一致的范围。

[0029] 尽管下面的描述使用术语“第一”、“第二”等来描述各种元素,但是这些元素不应该受到术语的限制。这些术语仅用于区分一个元素与另一个元素。例如,在不脱离各种描述的实施例的范围的情况下,第一图形表示可以称为第二图形表示,并且类似地,第二图形表示可以称为第一图形表示。第一图形表示和第二图形表示都是图形表示,但是它们不是相同的图形表示。

[0030] 在本文中描述的各种实施例的描述中使用的术语仅用于描述特定实施例的目的,而不旨在进行限制。如在各种描述的实施例和所附权利要求书的描述中所使用的,单数形式“一”、“一个”和“该/所述”也旨在包括复数形式,除非上下文清楚地另外指示。还应理解,本文中使用的术语“和/或”指的是并包含相关联列出的项中的一项或多项的任何和所有可能的组合。应当进一步理解,当在本说明书中使用术语“包括(includes)”、“包括(including)”、“包括(comprises)”和/或“包括(comprising)”指定了所述特征、整数、步骤、操作、元素和/或部件的存在,但不排除一个或多个其他特征、整数、步骤、操作、元素、部件和/或其组的存在或添加。

[0031] 术语“如果”任选地解释为意指“当……时”或“在……时”或“响应于确定”或“响应于检测到”,这取决于上下文。类似地,短语“如果确定”或“如果检测到”任选地解释为“一旦确定”或“响应于确定”或“一旦检测到”或“响应于检测到”,这取决于上下文。

[0032] 图1A示出了根据一些实施例的示例性系统。系统包括一个或多个机器人102、响应于查询的一个或多个人类工作人员104、以及与机器人和人类工作人员通信耦联的云平台106。任选地,系统进一步包括配置应用程序108和一个或多个终端用户110。

[0033] 机器人102包括传感模块(例如,摄像机、激光雷达(LiDAR)传感器)和驱动模块(例如,机械臂)。在一些实施例中,机械臂包括位于末端执行器的摄像机。在一些实施例中,机器人的一个或多个部件(例如,摄像机)连接到互联网。

[0034] 在一些实施例中,机器人102为拾取放置机器人。每个机器人可以包括具有从表面法线抓取物体的吸盘的一个或多个真空抓取器(例如,Robotiq AirPick)、具有从侧面抓取的两个手指的平行爪抓取器(例如,Robotiq 2f-85)、或其任意组合。两种模式的抓取器需要不同类型的拾取点规格,并且物体常更适合一种类型的抓取器而非另一种类型的抓取器。在一些实施例中,机器人可以向云平台106查询使用哪个抓取器(如下面描述的请求形式),并且可以相应地切换抓取器。应当理解,机器人102中的任一者可以是用于执行一

个或多个任务的任何类型的机器人,比如具有任何类型的抓取机构的拾取放置机器人。

[0035] 在一些实施例中,机器人102可以在执行任务前使用配置信息进行配置。如图1A所示,配置信息可以由终端用户110(例如,经由配置应用程序108)来指定。附加地或替代地,配置信息也可以由另一用户(例如,人类工作人员104)指定,或者由不同的计算机系统自动指定(例如,经由API)。

[0036] 配置信息在配置期间提供足够的信息,使得机器人可以独立操作。例如,终端用户可以为机器人指定宽泛的指令/命令,比如自然语言的高级任务、工作空间可见的原点位置以及附加的高级任务设置(例如,机器人是否需要能够旋转物体)。例如,宽泛的指令可以是“将苹果分类放入左边的箱子,并且将香蕉分类放入右边的箱子”或者“将UPS包裹分类放入左边的箱子,并且将联邦包裹包裹分类放入右边的箱子”。

[0037] 在一些实施例中,机器人102通过配置应用程序108被注册并对终端用户110可见。可以使用用户设备(例如,移动设备、台式计算机)来访问配置应用程序108。终端用户可以查看他们所有机器人的状态(例如,运行、停止、离线或紧急停止)。在一些实施例中,终端用户110经由配置应用程序108的用户界面来提供指令(例如,自然语言指令)。例如,用户可以通过在配置应用程序108的用户界面中键入自然语言文本串,通过文本输入来提供指令。作为另一个示例,用户可以通过语音输入提供指令。作为另一个示例,用户可以通过从预设选项选择来提供指令。应当理解,配置应用程序108可以提供任何类型的用户界面,以允许输入比如自然语言指令的配置信息,例如图形用户界面(例如,web应用程序的图形用户界面)或编程界面。

[0038] 在一些实施例中,配置过程包括两个步骤。在第一步中,机器人被定位到初始位置(或原点位置)。例如,机器人可以被配置成定点在其工作空间(例如,其上有箱子的桌子、传送带),使得所有要被操纵的物体对于感测模块都是可见的。在第二步中,可以向机器人提供指令(例如,自然语言指令),指示机器人应该做什么(例如,“将苹果分类放入左边的箱子,并且将香蕉分类放入右边的箱子”,“将UPS包裹分类放入左边的箱子,将联邦包裹分类放入右边的箱子”)。在一些实施例中,可以仅在机器人停止时进行配置。

[0039] 在一些实施例中,基于机器人的目标应用(例如,组装、包装、箱子拾取、检查)来定制配置过程,因此,配置应用程序108可以根据机器人的目标应用提供不同的用户界面,以便于输入针对机器人的配置信息。例如,如果机器人的目标应用是要制造零件套件,则配置应用程序可以提供用户界面,允许用户选择零件箱子以及应该拾取多少个每种零件来形成套件。这种配置将通知高级机器人程序,以及高级操作(比如拾取、放置和推动)的顺序和参数化。作为另一个示例,如果机器人的目标应用是制造零件套件,则配置应用程序可以被配置成接收并分析自然语言输入,以识别零件箱子以及应该拾取多少个每种零件来形成套件。在一些实施例中,为了确定机器人的目标应用,配置应用程序可以接收指示待配置的机器人的目标应用的输入,并且基于目标应用提供对应的用户界面。在一些实施例中,为了确定机器人的目标应用,配置应用程序可以自动分析待配置的机器人,识别机器人的目标应用,并提供相应的用户界面来相应地配置机器人。

[0040] 一旦配置了机器人,机器人就可以启动并开始执行其主循环。在任何时候,可以从配置应用程序中停止机器人。例如,终端用户可以通过配置应用程序手动起动和停止机器人。在一些实施例中,机器人不断查询云平台106以确定其状态(例如,起动或停止),并相应

地行动。在一些实施例中,机器人从云平台接收命令指令和状态更新,而不是向配置应用程序查询信息和指令。如果机器人状态从停止变为运行,则它查询云服务以找到(或被自动发送)其配置数据(例如,工作空间姿势和自然语言指令)。如果机器人意外停止(例如,由于安全问题或环境配置错误),则终端用户将通过应用程序得到通知。

[0041] 在一些实施例中,配置过程包括由人类工作人员104执行的附加配置步骤,以修改终端用户110的配置抑或执行附加配置步骤。组合起来,由终端用户110和人类工作人员104执行的配置步骤可以使用较低技能的按需劳动力来替代或增强传统的高技能编程系统集成工作。

[0042] 机器人102可以运行软件程序来执行任务,以完成命令(例如,由终端用户提供的配置信息指定)。在一些实施例中,机器人102包括运行软件程序的嵌入式平台。程序可以被构造成循环来重复执行任务。示例性任务包括拾取和放置物体、验证图像是否符合定义的条件集合(例如,电子商务包是否包含所有必需的项目)等。每个任务可以包括在循环中执行的多个子任务。这个循环的一些子任务可以在本地执行(即,使用机器人推断的参数),而其他子任务通过调用链接到机器人软件的专有API被外包给云软件。在一些实施例中,不是机器人运行独立的循环并将子任务外包给云执行,而是主活动循环在云上运行,并且将子任务外包给机器人进行本地执行。

[0043] 云平台106可以接收机器人102的请求。附加地或替代地,云平台被配置成基于活动循环的状态(例如,外包子任务)自动向机器人提供信息。示例性请求或信息可以包括根据指令选择在哪里拾取物体以及在图像中的哪里放置物体、确定图像中物体的脆弱性等。

[0044] 在一些实施例中,请求采用的是预定义的形式。例如,机器人提供的请求包括:工作空间的图像、一个或多个自然任务语言指令(通过配置从终端用户接收)、以及对拾取参数和放下参数的查询。更复杂的请求表格可以包括来自机器人的附加数据(比如可达到的姿势、候选选择、更多的终端用户配置设置)以及向服务/人类工作人员查询更多的信息(使用哪个抓取器来选择、抓取的角度、放下的角度、放下的高度等)。

[0045] 在一些实施例中,每个请求表格具有由该表格发出的所有请求及其由人类工作人员做出的响应的相关数据集,以及由该数据监督的相关机器学习模型,有时按任务或应用进行分类。例如,请求表格可以用于识别图像中的拾取点,并且它可以与包括所有请求(包括图像)和所有响应(包括在那些图像中识别的点)的数据集相关联。可以使用数据集来训练机器学习模型,以接收输入图像并识别输入图像中的拾取点。

[0046] 收到请求后,云平台可以查询对应的机器学习模型,以决定出模型是否可产生高质量的结果,或是否需要查询一个或多个人类工作人员。例如,向模型提供图像,并且模型可以输出预测的物品易碎性并输出置信度得分。如果表格模型对于请求具有高确定性或置信度(例如,高于预定义阈值),则云服务使用模型来生成响应并将其返回给用户。如果模型不确定,则可以将请求添加到队列中,由远程工作人员回答,并在完成后将其返回给机器人(并将请求添加到相关联的数据集,然后用于训练模型)。

[0047] 在一些实施例中,可以使用另外的算法来双重检查由人类抑或模型产生的结果,例如,通过向另外的人类询问共识。算法还可以用于向提供更高质量结果的工人提供更高的报酬。

[0048] 在一些实施例中,如果有一个以上的人工可以用于处理请求队列中的请求,则可

以使用另外的算法来最佳匹配人类工作人员和机器人请求。

[0049] 图1B示出了根据一些实施例的拾取放置操作的示例性过程。如图所示,终端用户110可以经由配置应用程序为机器人提供配置信息120(例如,自然语言指令、设置)。在一些实施例中,配置信息120可以附加地或替代地由人类工作人员104提供。在一些实施例中,配置信息120可以附加地或替代地由另一计算机系统或软件程序自动提供(例如,经由API)。在一些实施例中,配置数据由终端用户、人类工作人员和/或API输入,并被传输到云平台用于存储和未来访问。

[0050] 参考图1B,可以向云平台106提供配置信息120,如箭头126所示。配置信息然后可以由云平台106传输到机器人102,使得机器人102可以基于配置信息执行任务,如箭头122所示。机器人可以生成响应并将响应传输到云平台106,如箭头124所示。例如,响应可以包括对机器人成功执行任务所需的信息的请求。示例性请求或信息可以包括根据指令选择在哪里拾取物体以及在图像中的哪里放置物体、确定图像中物体的脆弱性等。

[0051] 收到机器人102的请求后,云平台106可处理请求并获取所请求的信息。在一些实施例中,云平台106使用一个或多个机器学习模型107来处理请求。例如,来自机器人102的请求可以包括物品的图像,并且请求是要确定在哪里拾取物体,并且云平台106可以将关于物品的图像和/或信息输入到机器学习模型中。机器学习模型可以输出表示拾取箭头的原始点和第二点的点。机器学习模型还可以输出指示模型对拾取箭头有多确信的置信度得分。基于机器学习模型的输出(例如,拾取箭头),云平台可以参数化动作(例如,拾取参数)以传输给机器人102。

[0052] 作为另一个示例,来自机器人102的请求可以包括场景的图像,并且请求用于确定在场景中放置物品的位置,并且云平台106可以将关于物品的图像和/或信息输入到机器学习模型中。机器学习模型可以输出表示位置箭头的原始点和第二点的点。机器学习模型还可以输出指示模型对放置箭头有多确信的置信度得分。基于机器学习模型的输出(例如,放置箭头),云平台可以参数化动作(例如,放置参数)以传输给机器人102。

[0053] 机器学习模型的输入不限于场景图像和物品信息。输入可以是点、数字/分类数据(例如,像拾取深度、接近角的选项),或者一些其他形式的输入。机器学习模型的输出不限于拾取/放置箭头和相关联的置信度得分。输出可以是数字/分类预测,或者一些其他形式的输出。

[0054] 在一些示例中,原始点和第二点中的每一者可以是概率分布的形式。可以基于对应于原始点和第二点的概率分布来计算置信度得分。

[0055] 如果云平台无法基于其学习的模型对动作进行参数化(例如,如果置信度得分低于预定义阈值),则云平台可以转发请求或询问人类工作人员,如箭头128所示。在图3和图4中提供了示出给人类工作人员的示例性用户界面。对请求的解决方案(例如,拾取和放下参数)可以由人类工作人员提供给云平台,如箭头130所示,并且云平台将解决方案中继给机器人以完成任务。

[0056] 例如,拾取放置机器人可以在其运行状态下执行以下步骤,其运行状态由机器人自身或云平台驱动。在下面的过程中,一些步骤可以任选地组合,一些步骤的顺序可以任选地改变,并且一些步骤可以任选地省略。例如,在一些情况下(如成像),比如缓存之类的优化可以允许跳过其中的一些步骤。在一些示例中,可以结合过程执行附加步骤。因此,所示

出的操作(以及下面更详细描述的操作)本质上是示例性的,因此,不应被视为限制。

[0057] 1.移动到初始位置(即原点位置),使得工作区可见。在一些实施例中,这可以使用机器人规划算法(例如,IK或RRT)来完成。

[0058] 2.用安装到机器人末端执行器的摄像机对工作空间进行成像。

[0059] 3.给定来自配置和工作空间图像的自然语言指令,查询云平台106中的ML模型以寻找拾取点和放下点。如果ML模型未能产生高置信度的结果,则机器人或云平台可以向人类工作人员询问拾取点和放下点。

[0060] 4.导航到拾取点,拾取物体,然后导航到放下点,并放下物体。

[0061] 5.转到步骤1。

[0062] 如上所示,这个循环的一些步骤可以由机器人在本地执行(例如,步骤1和步骤2),而其他步骤(例如,步骤3)由云软件执行。这可以通过机器人运行循环并向云模型查询拾取/放下参数来实现,抑或通过云模型运行循环并向机器人查询以执行动作来实现。

[0063] 关于步骤3,拾取和放下是具有不同参数的两个不同的子任务。对于放下,可以查询多个放下参数。可以向人类工作人员呈现图像,并要求其在图像上指定要放下的物体的点。如果机器人还需要能够旋转物体,则也可以要求人类工作人员指定角度,指定角度很可能是作为从指定放下点拖动的箭头输入的。如果物体易碎,则机器人可能还需要知道物体放下的高度,或者接近的角度。

[0064] 对于拾取,不同类型的抓取器可以查询不同的拾取参数。对于真空抓取器,机器人需要物体上的单点来参数化拾取。在远程人类工作人员的用户界面中,可以向他们呈现机器人所看到的图像,并且他们可以点击物体上的某个点。对于平行爪抓取器,机器人可能需要抓取中心、抓取角度和抓取深度的点(例如,对于立方体,抓取应该靠近顶部还是靠近底部)。

[0065] 在一些实施例中,如果步骤4不能成功执行,则过程返回到步骤3,以实施成功的拾取。例如,机器人可能发现返回的拾取点在机器人的范围之外,因此在运动学上是无效的,并且它将重复步骤3以获得替代解决方案。

[0066] 关于步骤4,一旦拾取和放置被参数化,机器人就可以导航到拾取点,根据拾取参数拾取物体,然后导航到放下点。这里的导航步骤可以通过多种方式解决。

[0067] 在一些实施例中,方法基于的是如下基于计划的途径:

[0068] 1.机器人使用安装在末端执行器上的深度摄像机拍摄工作空间(在 origin 位置)的深度图像。

[0069] 2.机器人从工作空间的深度图像中构建点云。

[0070] 3.机器人向云平台106查询拾取和放下参数,或者由云平台给指示以使用给定参数执行拾取和放下。如上所述,参数可以由云平台使用机器学习模型或人类工作人员来获得。

[0071] 4.机器人使用点云将返回的拾取点投影到3D空间中的真实点。然后计算这个点的曲面法线。

[0072] 5.然后,机器人查询逆运动学(I.K.)解算器(其受到约束)以识别机器人关节角度,机器人关节角度将末端效应器放置在拾取点处,面向表面法线,或以由云平台指定的一些其他角度放置。

[0073] 6. 机器人然后查询运动规划算法 (例如, RRT、插值IK或关节空间插值) 来计算机器人关节角度随时间的轨迹, 这将 (在物理上似乎) 允许机器人从其所在位置到达拾取点。

[0074] 7. 一旦机器人得到放下点, 它就将该点投射到3D点和表面法线中, 向上移位放下高度。然后, 它运行IK和运动规划来计算从拾取点到放下点的路径 (类似于上面的三个步骤), 以及从放下点到原点位置的路径。

[0075] 8. 然后, 机器人执行预先计算好的拾取点计划, 并根据拾取参数执行拾取。

[0076] 9. 然后机器人执行预先计算好的计划到达放下点, 并放下物品。

[0077] 10. 然后, 机器人执行预先计算好的计划到达原点位置。

[0078] 11. 转到步骤1。

[0079] 上述方法可以依靠高质量深度感测来计算3D点和表面法线。在现实世界的设置中, 深度摄像机常常具有次优的性能, 这可能会影响使用上述流水线的性能。使用深度摄像机的3D点投影的替代方法是如下的“视觉伺服”:

[0080] 1. 机器人使用安装在末端执行器上的摄像机拍摄工作空间 (在 origin 位置) 的图像。

[0081] 2. 然后, 机器人使用控制器移动机器人, 其方式使得指定的拾取点保持在摄像机的中心, 并且机器人更靠近拾取点 (视觉伺服)。

[0082] 3. 一旦机器人与物体发生碰撞 (或使用传感器预先阻止碰撞), 它就根据指定的参数执行拾取。如上所述, 参数可以由云平台使用机器学习模型或人类工作人员来获得。

[0083] 4. 机器人要么 (a) 使用上述流水线执行到达放下点的计划 (放下点的误差比拾取点的误差问题小得多), 要么 (b) 执行返回原点位置的计划并视觉伺服到放下位置。

[0084] 5. 机器人执行放下, 并使用基于计划的方法导航回原点位置。

[0085] 6. 转到步骤1。

[0086] 在一些实施例中, 在上述过程期间, 如果机器人/云服务发现不可恢复的错误或安全错误 (例如, 机器人被卡住, 待分拣的箱子丢失), 则机器人可以将其状态设置为停止 (或在发生安全错误的情况下紧急停止), 并通过配置应用程序108适当地警告终端用户。此外, 在过程的任何步骤中, 如果机器人需要信息来执行任何子任务, 则它可以查询云平台, 并且云平台可以使用一个或多个机器学习模型和人类工作人员来响应请求。

[0087] 在一些实施例中, 机器人可以访问一个或多个机器学习模型 (例如, 本地存储的模型) 的第一集合, 并使用一个或多个机器学习模型的第一集合来确定参数。例如, 机器人可以将一个或多个图像输入到第一模型集合中, 以确定拾取参数和与输出相关联的第一置信度得分。如果第一置信度得分没有超过第一阈值, 则机器人生成将被发送到中央平台的请求, 该请求可以包括相同的图像。在中央平台处, 可以使用第二组机器学习模型来确定拾取参数和与输出相关联的第二置信度得分。如果第二置信度得分没有超过第二阈值, 则可以如本文所述询问人类工作人员。

[0088] 随着时间的推移, 这个过程产生标记的数据集。例如, 数据集可以包括多个图像, 每个图像包括标记的拾取点 (例如, 由人类工作人员标记)。数据集可以用于训练机器学习模型 (例如, 由上述云平台106使用的表单模型中的一者), 该模型被配置成在给定输入图像时识别拾取点。因此, 随着时间的推移, 系统可以将人类逐步排除在循环之外。

[0089] 该云平台可以与不同类型的机器人系统 (包括预先存在的拾取放置系统) 集成。在一些实施例中, 硬件部件可以插入到第三方机器人中, 以允许该机器人如本文所述操作。在

一些实施例中,可以提供包括定制机械臂、摄像机和与云平台通信的软件能力的定制包。

[0090] 图3示出了根据一些实施例提供给人类工作人员的示例性用户界面300。用户界面300提示人类工作人员解决任务310(“将形状分类放入箱子”)。用户界面可以包括工作空间的一个或多个图像。在所描绘的示例中,用户界面300包括描绘要被分类的物体的图像312,以及描绘物体需要被分类到的箱子的图像314。如上所述,图像312和314都可以由负责将图像312中表示的物理物体分类放入图像314中表示的物理箱子中的机器人来捕获。在一些实施例中,图像被预处理以消除敏感信息。

[0091] 在一些实施例中,拾取和放下参数的部件可以表示为工作区图像(或其他图形表示)上的图形注释。例如,拾取可以表示为箭头,其中,选择的第一个点是抓取器的中心点,并且第二个点表示抓取器的旋转,以及在具有可变宽度的两指抓取器情况下的延伸。如图3所示,人类工作人员可以在图像312上添加箭头302来指示物体的拾取参数。人类工作人员可以将箭头304添加到图像314上,以指示物体将被放下的位置。人类工作人员的解决方案(即,拾取和放下参数)可以被传输到云平台,然后传输给机器人,使得机器人可以物理地执行拾取和分类任务。附加设置316可以提供给人类工作人员,并且它们可以被修改。

[0092] 应理解,机器人任务可以任何合适的方式参数化。例如,在接近拾取或放下位置之前,可以使用附加点来确定起始点,使得人类工作人员能够控制接近的角度和路径航路点。

[0093] 如图3所示,用于选择拾取和放置的标记界面可以包含在拾取注释期间生成的拾取物品的快照,作为裁剪抑或不同形式的分段。这个快照可以在放置选择期间用作新位置中的物品的预览,利用不同的放置点选择进行平移,并且利用不同的放置箭头选择进行旋转。

[0094] 在一些实施例中,发送给人类工作人员的查询包括由机器学习模型(例如,图1B中的机器学习模型107)产生的推论。人类工作人员的任务可以是要从一组推论中进行选择,修正推论,或者拒绝它们。图4示出了根据一些实施例的提供给人类工作人员的示例性用户界面。用户界面400提供由机器学习模型生成的试验性解决方案:分别用于拾取/放置三个物体的解决方案1、解决方案2、解决方案3。在一些实施例中,这些解决方案通过机器学习模型与低置信度得分相关联,并且因此被呈现给人类工作人员以供进一步确认。

[0095] 对于每种解决方案,人类工作人员可以接受、拒绝或修改。例如,人类工作人员可以确定解决方案1和解决方案2提供了正确的解决方案并接受它们。另一方面,解决方案3为正确拾取星形物体提供了不正确的拾取参数,并为星形物体提供了不正确的放置位置。因此,人类工作人员可以拒绝并覆盖它们(例如,通过在图像上绘制新的箭头)。

[0096] 在一些实施例中,机器人对指令的执行被自动并行化和流水线化,因此多条指令可同时运行,且指令可以由可以同时或同步运行的许多子指令组成。

[0097] 图5示出了根据各种示例操作机器人的过程500。例如,使用实现软件平台的一个或多个电子设备来执行过程500。在一些示例中,使用客户机-服务器系统来执行过程500,并且过程500的框以任何方式在服务器与客户机设备之间划分。在其他示例中,过程500的框在服务器与多个客户端设备之间划分。因此,虽然过程500的部分在本文被描述为由客户端-服务器系统的特定设备执行,但是将会理解,过程500不如此限制。在其他示例中,仅使用客户端设备(例如,用户设备100)或仅多个客户端设备来执行过程500。在过程500中,一些框可以任选地组合,一些框的顺序可以任选地改变,并且一些框可以任选地省略。在一些

示例中,可以结合过程500执行附加步骤。因此,所示出的操作(以及下面更详细描述的操作)本质上是示例性的,因此,不应被视为限制。

[0098] 在框502处,示例性系统(例如,一个或多个电子设备)获取机器人的自然语言指令,其中,自然语言指令与第一用户相关联。在框504处,系统基于自然语言指令识别任务。在框506处,系统将任务传输给机器人。在框508处,系统从机器人接收与任务相关联的请求。在框510处,系统确定该请求是否可以通过一个或多个经过训练的机器学习算法来解决。在框512处,如果所述请求不可以通过一个或多个经过训练的机器学习算法解决,则系统向第二用户的电子设备传输查询。在框514处,系统接收来自第二用户的对查询的响应。在框516处,系统使得任务由机器人基于响应来执行。

[0099] 上述操作任选地由图2所示的部件实施。图2示出了根据一个实施例的计算设备的示例。设备200可以是连接到网络的主计算机。设备200可以是客户端计算机或服务器。如图2所示,设备200可以是任何合适类型的基于微处理器的设备,诸如个人计算机、工作站、服务器、或诸如电话或平板电脑等的手持计算设备(便携式电子设备)。设备可以包括例如处理器210、输入设备220、输出设备230、存储装置240和通信设备260中的一者或多者。输入设备220和输出设备230通常可以对应于上述的那些设备,并且可以与计算机连接或集成。

[0100] 输入设备220可以是提供输入的任何合适的设备,诸如触摸屏、键盘或小键盘、鼠标、或者语音识别设备。输出设备230可以是提供输出的任何合适的设备,诸如触摸屏、触觉设备、或者扬声器。

[0101] 存储装置240可以是提供存储的任何合适的设备,诸如包括RAM、高速缓存、硬盘驱动器或可移除存储盘的电、磁或光存储器。通信设备260可以包括能够通过网络发送并接收信号的任何合适的设备,诸如网络接口芯片或设备。计算机的组件可以以任何合适的方式连接,诸如通过物理总线或无线连接连接。

[0102] 可以存储在存储装置240中并由处理器210执行的软件250可以包括例如体现本公开的功能的编程(例如,体现在如上所述的设备中)。

[0103] 软件250还可以被存储在任何非暂时性计算机可读存储介质内和/或在任何非暂时性计算机可读存储介质内传输,以供指令执行系统、装置或设备(诸如上述那些)使用或与指令执行系统、装置或设备结合使用,软件250可以从指令执行系统、装置或设备获取与软件相关联的指令并执行指令。在本公开的上下文中,计算机可读存储介质可以是包含或存储供指令执行系统、装置或设备使用或者与指令执行系统、装置或设备结合使用的编程的任何介质,比如存储装置240。

[0104] 软件250也可以在任何传输介质内传播,以供指令执行系统、装置或设备(诸如上述那些)使用或与指令执行系统、装置或设备结合使用,软件250可以从指令执行系统、装置或设备获取与软件相关联的指令并执行指令。在本公开的上下文中,传输介质是可以通信、传播或传输编程以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备相关联使用的任何介质。传输可读介质可以包括但不限于电子、磁、光、电磁或红外有线或无线传播介质。

[0105] 设备200可以连接到网络,网络可以是任何合适类型的互连通信系统。网络可以实现任何合适的通信协议,并且可以由任何合适的安全协议来保护。网络可以包括可以实现网络信号的发送和接收的任何合适布置的网络链路,诸如无线网络连接、T1或T3线路、有线

网络、DSL或电话线路。

[0106] 设备200可以实现适合于在网络上操作的任何操作系统。软件250可以用任何合适的编程语言,诸如C、C++、Java或Python,来编写。在各种实施例,体现本公开的功能的应用程序软件可以以不同的配置部署,比如在客户机/服务器布置中或通过网络浏览器部署为基于网络的应用程序或网络服务。

[0107] 实施例

[0108] 示例性方法、非暂时性计算机可读存储介质、系统和电子设备在以下项目中列出:

[0109] 1.一种用于操作机器人的计算机实现的方法,该方法包括:

[0110] 获得给机器人的指令,其中,指令与第一用户相关联;

[0111] 基于指令识别任务;

[0112] 向机器人传输任务;

[0113] 从机器人接收与任务相关联的请求;

[0114] 确定请求是否可以通过一个或多个经过训练的机器学习算法解决;

[0115] 如果请求不可以通过一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;

[0116] 接收来自第二用户的对所述查询的响应;以及

[0117] 使得任务由机器人基于响应来执行。

[0118] 2.如项目1所述的方法,其中,指令是自然语言指令,并且自然语言指令指导机器人拾取和/或放置一个或多个物体。

[0119] 3.如项目1至2中任一项所述的方法,其中,任务包括多个子任务,并且其中,多个子任务包括拾取子任务和放下子任务。

[0120] 4.如项目3所述的方法,其中,请求包括图像和对一个或多个拾取参数的查询。

[0121] 5.如项目4所述的方法,其中,拾取参数包括拾取点、抓取角度、抓取深度或以上各项的任意组合。

[0122] 6.如项目3所述的方法,其中,请求包括图像和对一个或多个放下参数的查询。

[0123] 7.如项目6所述的方法,其中,放下参数包括放下点、旋转角度、放下高度或以上各项的任意组合。

[0124] 8.如项目1至项目7中任一项所述的方法,其中,确定请求是否可以通过一个或多个经过训练的机器学习算法解决包括:

[0125] 将机器人捕获的数据输入到一个或多个经过训练的机器学习模型,以获得解决方案和与解决方案相关联的置信度得分;

[0126] 确定置信度得分是否低于预定义阈值。

[0127] 9.如项目8所述的方法,其中,查询包括由一个或多个机器学习模型给出的解决方案。

[0128] 10.如项目8所述的方法,其中,由机器人捕获的数据包括图像。

[0129] 11.如项目8所述的方法,进一步包括:基于来自第二用户的响应来训练一个或多个机器学习模型。

[0130] 12.如项目9所述的方法,进一步包括:使得显示第一图形用户界面,第一图形用户界面包括由机器人捕获的一个或多个图像。

- [0131] 13.如项目11所述的方法,其中,第一图形用户界面进一步包括:
- [0132] 对由一个或多个机器学习模型给出的解决方案的指示;以及
- [0133] 用于接受该解决方案的一个或多个用户界面控件。
- [0134] 14.如项目1至13中任一项所述的方法,进一步包括:使得在第一用户的电子设备上显示第二图形用户界面,用于接收指令。
- [0135] 15.如项目14所述的方法,其中,基于机器人的目标应用选择第二图形用户界面。
- [0136] 16.一种电子设备,包括:
- [0137] 一个或多个处理器;
- [0138] 存储器;以及
- [0139] 一个或多个程序,其中,一个或多个程序存储在存储器中,并且被配置为由一个或多个处理器执行,一个或多个程序包括用于以下操作的指令:
- [0140] 获得给机器人的指令,其中,指令与第一用户相关联;
- [0141] 基于指令识别任务;
- [0142] 向机器人传输任务;
- [0143] 从机器人接收与任务相关联的请求;
- [0144] 确定请求是否可以通过一个或多个经过训练的机器学习算法解决;
- [0145] 如果请求不可以通过一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;
- [0146] 接收来自第二用户的对查询的响应;以及
- [0147] 使得任务由机器人基于响应来执行。
- [0148] 17.一种非暂时性计算机可读存储介质,存储有一个或多个程序,一个或多个程序包括指令,当由电子设备的一个或多个处理器执行时,指令使得电子设备执行:
- [0149] 获得给机器人的指令,其中,指令与第一用户相关联;
- [0150] 基于指令识别任务;
- [0151] 向机器人传输任务;
- [0152] 从机器人接收与任务相关联的请求;
- [0153] 确定请求是否可以通过一个或多个经过训练的机器学习算法解决;
- [0154] 如果请求不可以通过一个或多个经过训练的机器学习算法解决,则向第二用户的电子设备传输查询;
- [0155] 接收来自第二用户的对查询的响应;以及
- [0156] 使得任务由机器人基于响应来执行。
- [0157] 尽管已经参考附图充分描述了本公开和示例,但是要注意的是,对于本领域技术人员来说,各种变化和修改将变得显而易见。这样的改变和修改应理解为包括在由权利要求书限定的本公开和示例的范围内。
- [0158] 为了解释的目的,已经参考特定实施例描述了前述描述。然而,上面的说明性讨论并不旨在穷举或将本发明限制于所公开的精确形式。鉴于上述教导,许多修改和变化是可能的。选择并描述实施例是为了最好地解释技术的原理及其实际应用。因此,本领域的其他技术人员能够最好地利用具有适合于预期的特定用途的各种修改的技术和各种实施例。

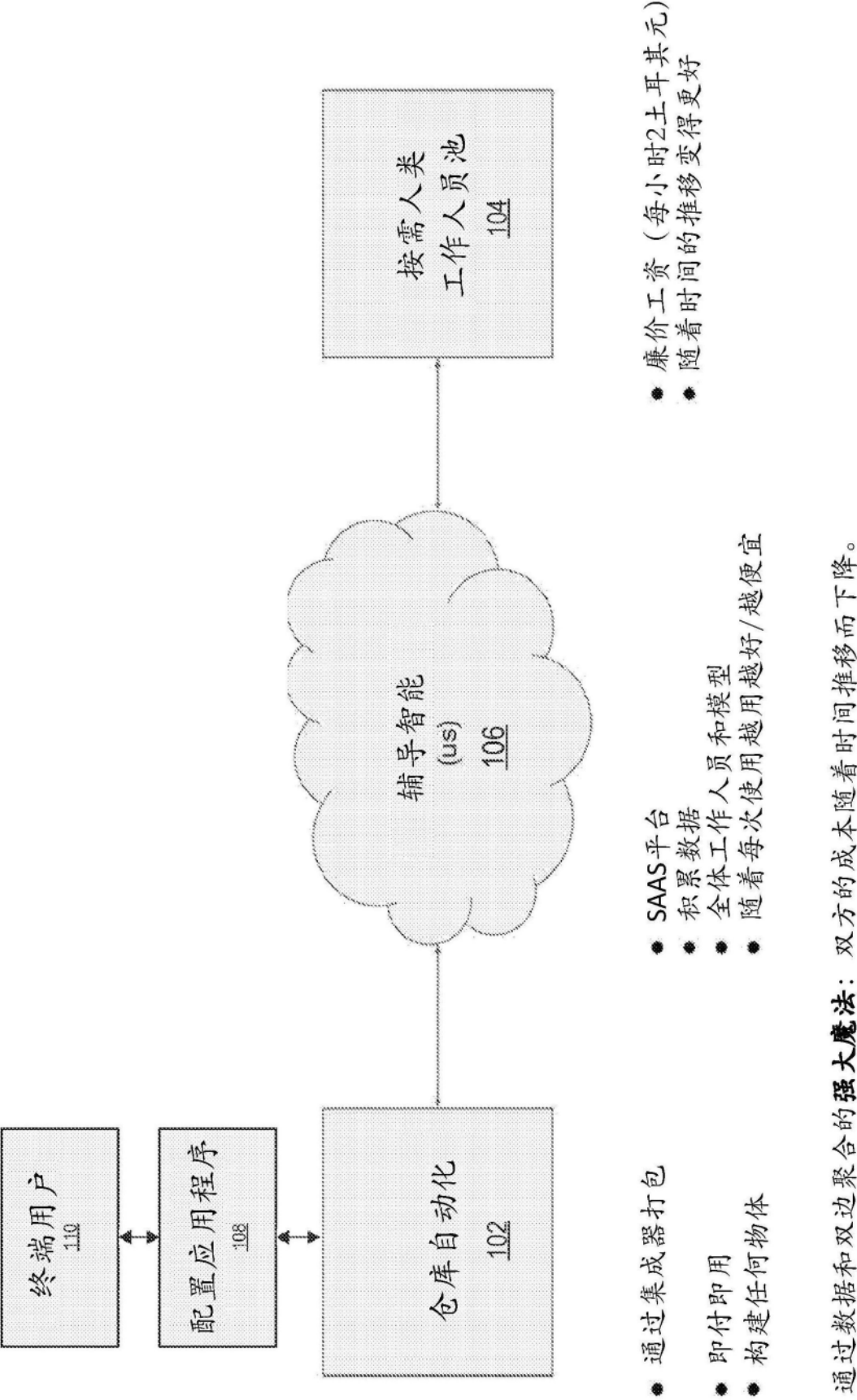


图1A

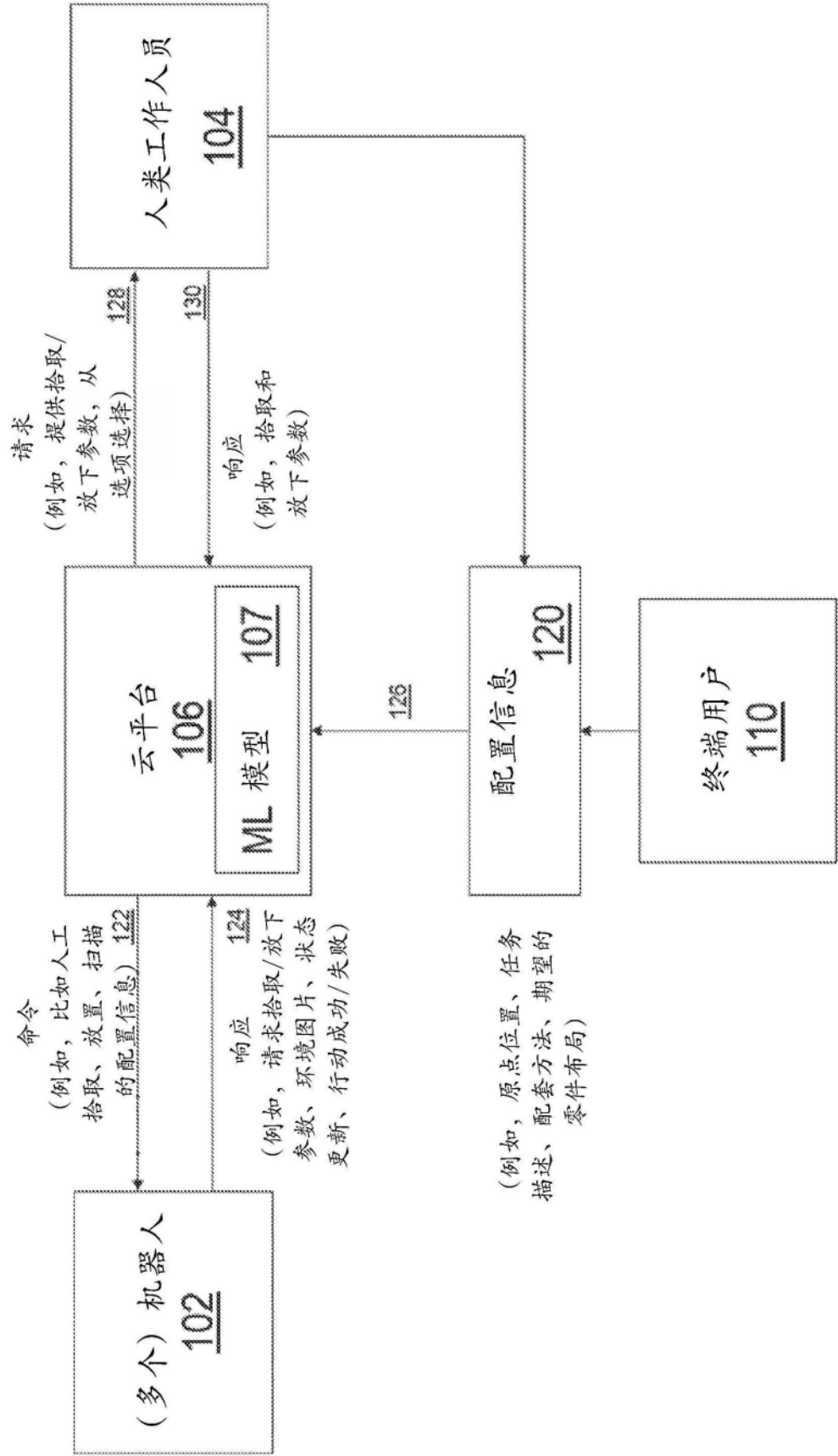


图1B

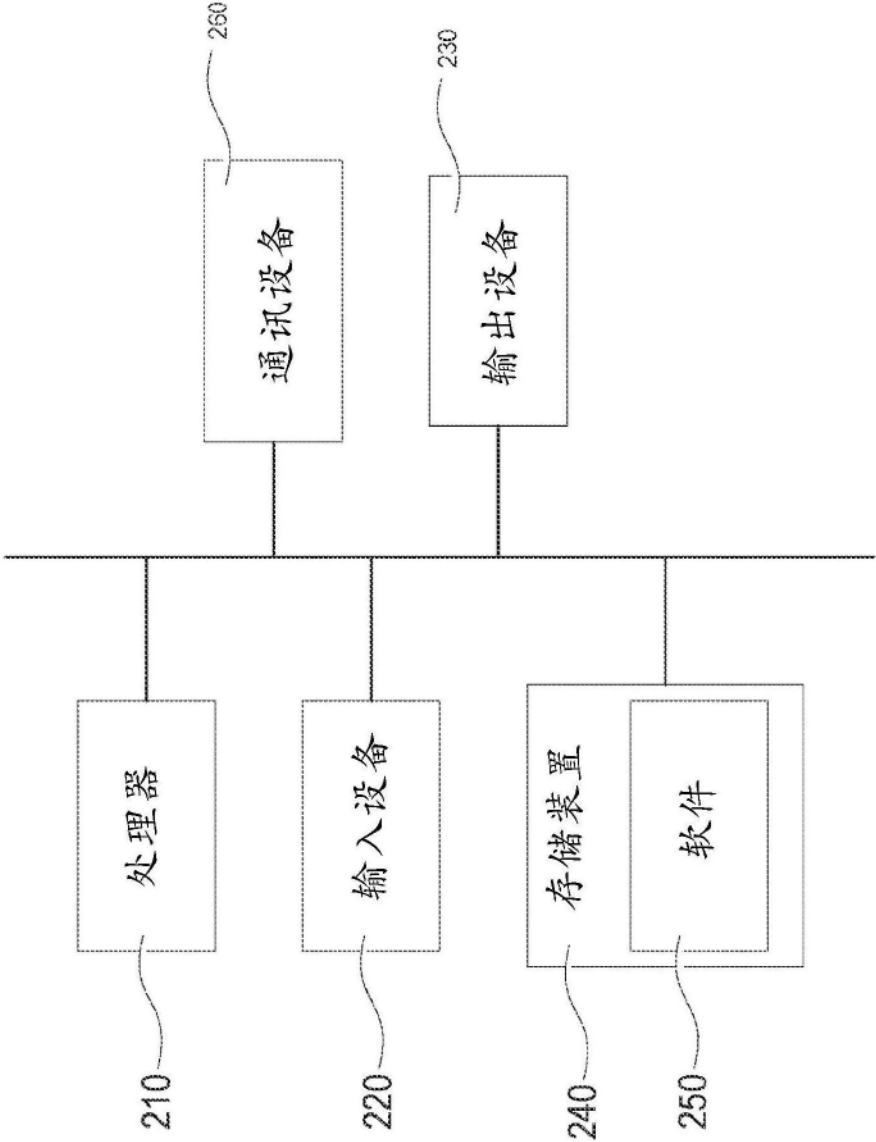
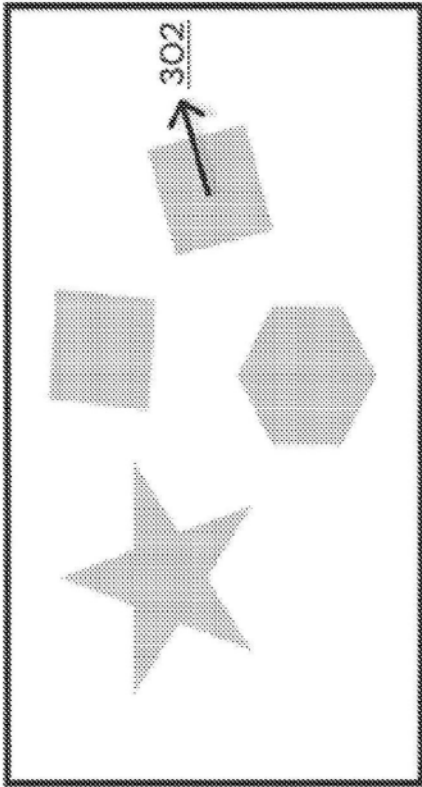


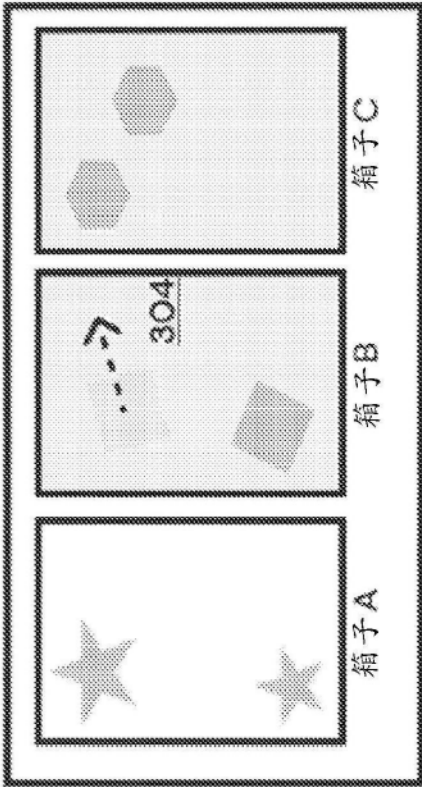
图2

示例拾取放置标记界面 300

任务：将形状分类放入箱子 310



拾取图像 312



放置图像 314

拾取放置设置 316				
拾取深度	2 cm			
放置高度	3 cm			
抓握力	5 N			
接近角度	90°			
抓取器	真空			

箭头 302: 拾取
箭头 304 (虚线): 放置

图3

示例机器学习辅助界面

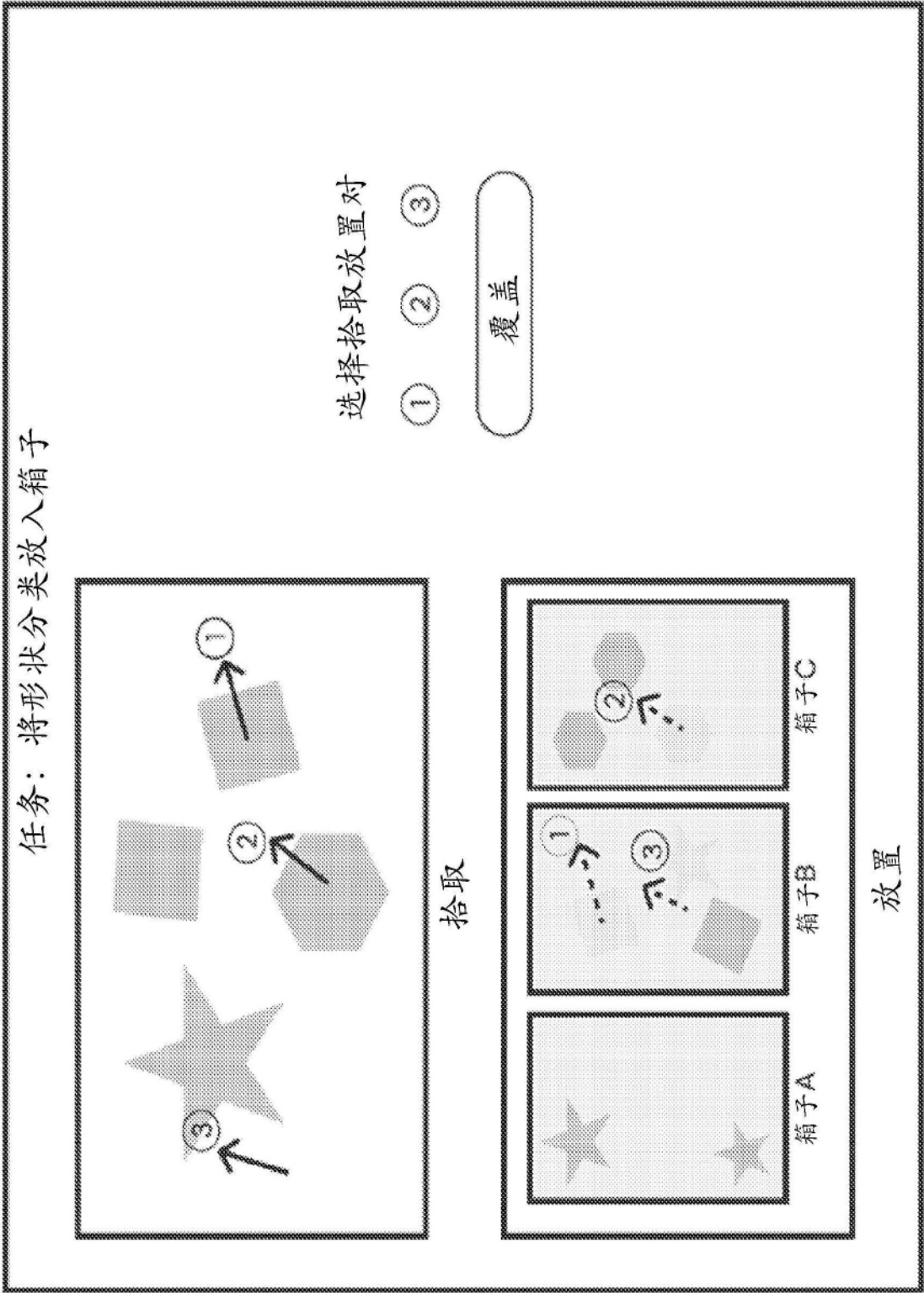


图4

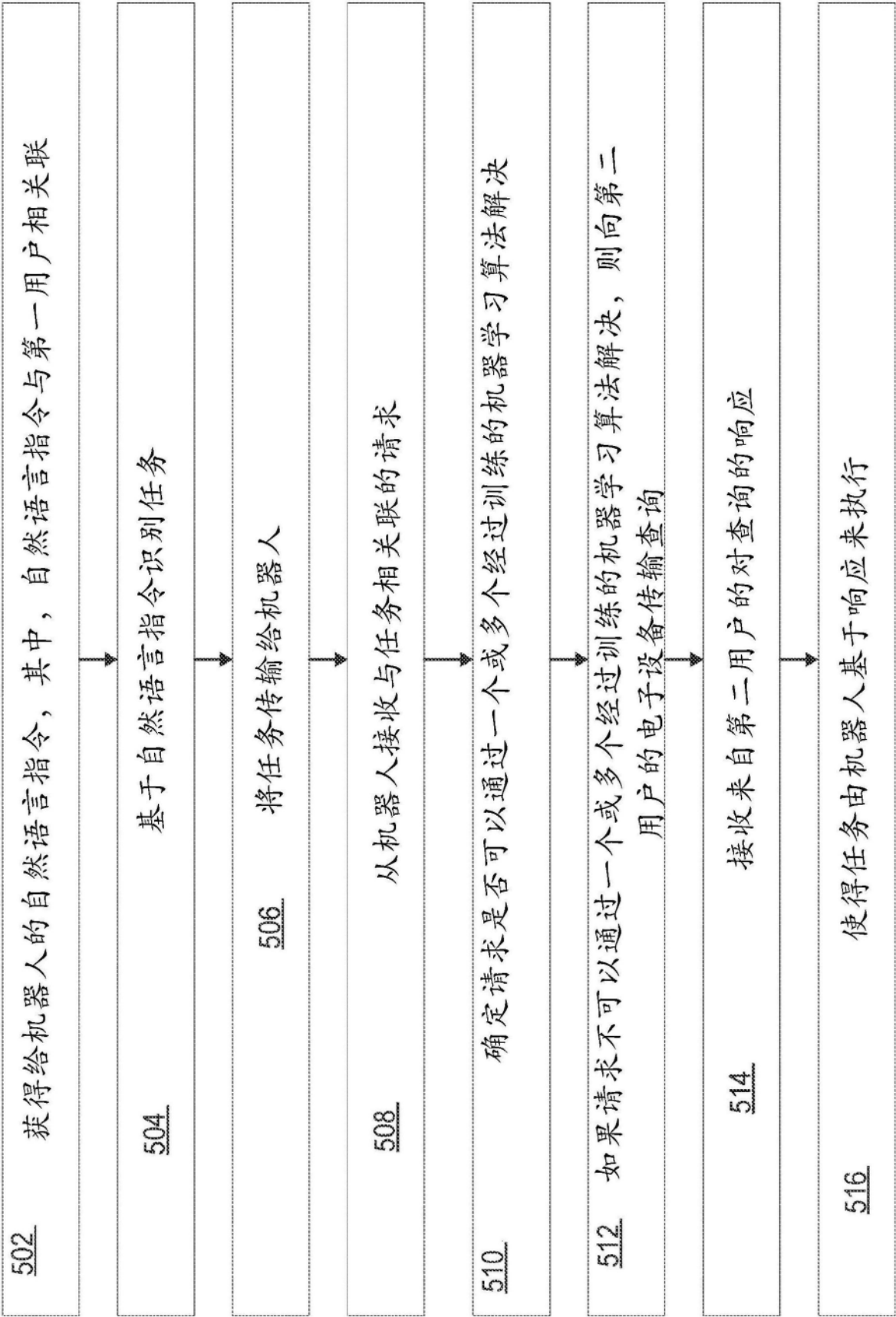


图5