



(12) 发明专利

(10) 授权公告号 CN 108367436 B

(45) 授权公告日 2021.05.18

(21) 申请号 201680070415.3

(22) 申请日 2016.11.02

(65) 同一申请的已公布的文献号
申请公布号 CN 108367436 A

(43) 申请公布日 2018.08.03

(30) 优先权数据
62/262,249 2015.12.02 US
15/069,834 2016.03.14 US

(85) PCT国际申请进入国家阶段日
2018.05.31

(86) PCT国际申请的申请数据
PCT/US2016/060079 2016.11.02

(87) PCT国际申请的公布数据
W02017/095580 EN 2017.06.08

(73) 专利权人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 R·B·托瓦 S·奥米德萨菲伊
A·阿加默汉马蒂

(74) 专利代理机构 上海专利商标事务所有限公
司 31100

代理人 陈炜 袁逸

(51) Int.Cl.
B25J 9/16 (2006.01)
G06T 7/50 (2006.01)

审查员 任大林

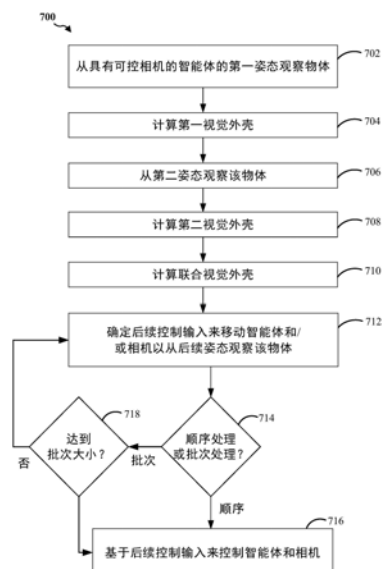
权利要求书3页 说明书9页 附图6页

(54) 发明名称

针对三维空间中的对象位置和范围的主动
相机移动确定

(57) 摘要

一种运动规划方法包括：从具有可控相机的智能体的第一姿态观察对象。该方法还包括：确定一个或多个后续控制输入来移动智能体和相机以从至少一个后续姿态观察该对象。(诸)后续控制输入被确定以最小化基于从相机收集到的视觉数据对对象的预期包围测量。该方法进一步包括：基于(诸)后续控制输入来控制智能体和相机。



1. 一种针对机器人设备进行运动规划的方法,所述机器人设备包括相机,所述方法包括:

使用所述机器人设备的所述相机从多个不同姿态观察对象;

基于多个三维3D外壳的交集来确定所述对象的预期体积,所述多个3D外壳基于在所述多个不同姿态处观察的所述对象的可见像素来确定;

针对基于多个潜在控制输入中的潜在控制输入而确定的每一潜在后续姿态,基于经估计会在所述潜在后续姿态处观察到的所述对象的可见像素来预测后续3D外壳,每一潜在后续姿态是在维持所述机器人设备的当前姿态的情形下确定的;

确定移动所述机器人设备以从后续姿态观察所述对象的控制输入,所述控制输入对应于与所述多个3D外壳具有最小预估交集的后续3D外壳,以使基于在所述后续姿态从所述相机收集的视觉数据所述对象的预期体积得以最小化;以及

基于所述控制输入来控制所述机器人设备移动到与所述后续姿态相对应的位置。

2. 如权利要求1所述的方法,其中,基于使得用于最小化所述预期体积的成本最小化来确定所述控制输入。

3. 如权利要求1所述的方法,其中,所述控制输入是使用对多个潜在控制输入的批处理来确定的。

4. 如权利要求1所述的方法,其中,所述控制输入是顺序确定的。

5. 如权利要求1所述的方法,其中,所述机器人设备被进一步地使用最小数目的控制输入来控制以围绕所述对象移动。

6. 一种针对机器人设备进行运动规划的装置,所述机器人设备包括相机,所述装置包括:

存储器;以及

耦合至所述存储器的至少一个处理器,所述至少一个处理器被配置成:

使用所述机器人设备的所述相机从多个不同姿态观察对象;

基于多个三维3D外壳的交集来确定所述对象的预期体积,所述多个3D外壳基于在所述多个不同姿态处观察的所述对象的可见像素来确定;

针对基于多个潜在控制输入中的潜在控制输入确定的每一潜在后续姿态,基于经估计会在所述潜在后续姿态处观察到的所述对象的可见像素来预测后续3D外壳,每一潜在后续姿态是在维持所述机器人设备的当前姿态的情形下确定的;

确定移动所述机器人设备以从后续姿态观察所述对象的控制输入,所述控制输入对应于与所述多个3D外壳具有最小预估交集的后续3D外壳,以使基于在所述后续姿态从所述相机收集的视觉数据所述对象的预期体积得以最小化;以及

基于所述控制输入来控制所述机器人设备移动到与所述后续姿态相对应的位置。

7. 如权利要求6所述的装置,其中,所述至少一个处理器被进一步配置成:

基于使得用于最小化所述预期体积的成本最小化来确定所述控制输入。

8. 如权利要求6所述的装置,其中,所述至少一个处理器被进一步配置成:

使用对多个潜在控制输入的批处理来确定所述控制输入。

9. 如权利要求6所述的装置,其中,所述至少一个处理器被进一步配置成:

使用顺序处理来确定所述控制输入。

10. 如权利要求6所述的装置,其中,所述至少一个处理器被进一步配置成:

使用最小数目的控制输入来控制所述机器人设备围绕所述对象移动。

11. 一种针对机器人装置进行运动规划的设备,所述机器人装置包括相机,所述设备包括:

用于使用所述机器人装置的所述相机从多个不同姿态观察对象的装置;

用于基于多个三维3D外壳的交集来确定所述对象的预期体积的装置,所述多个3D外壳基于在所述多个不同姿态处观察的所述对象的可见像素来确定;

用于针对基于多个潜在控制输入中的潜在控制输入确定的每一潜在后续姿态,基于经估计会在所述潜在后续姿态处观察到的所述对象的可见像素来预测后续3D外壳的装置,每一潜在后续姿态是在维持所述机器人装置的当前姿态的情形下确定的;

用于确定移动所述机器人装置以从后续姿态观察所述对象的控制输入的装置,所述控制输入对应于与所述多个3D外壳具有最小预估交集的后续3D外壳,以使基于在所述后续姿态从所述相机收集的视觉数据所述对象的预期体积得以最小化;以及

用于基于所述控制输入来控制所述机器人装置移动到与所述后续姿态相对应的位置的装置。

12. 如权利要求11所述的设备,其中,用于确定的装置包括用于基于使得用于最小化所述预期体积的成本最小化来确定所述控制输入的装置。

13. 如权利要求11所述的设备,其中,用于确定的装置包括用于使用对所述多个潜在控制输入的批处理来确定所述控制输入的装置。

14. 如权利要求11所述的设备,其中,用于确定的装置包括用于使用顺序处理来确定所述控制输入的装置。

15. 如权利要求11所述的设备,其中,用于控制的装置包括用于使用最小数目的控制输入来控制所述机器人装置围绕所述对象移动的装置。

16. 一种其上编码有针对机器人设备进行运动规划的程序代码的非瞬态计算机可读介质,所述机器人设备包括相机,所述程序代码由处理器执行时使得所述处理器:

使用所述机器人设备的所述相机从多个不同姿态观察对象;

基于多个三维3D外壳的交集来确定所述对象的预期体积,所述多个3D外壳基于在所述多个不同姿态处观察的所述对象的可见像素来确定;

针对基于多个潜在控制输入中的潜在控制输入确定的每一潜在后续姿态,基于经估计会在所述潜在后续姿态处观察到的所述对象的可见像素来预测后续3D外壳,每一潜在后续姿态是在维持所述机器人设备的当前姿态的情形下确定的;

确定移动所述机器人设备以从后续姿态观察所述对象的控制输入,所述控制输入对应于与所述多个3D外壳具有最小预估交集的后续3D外壳,以使基于在所述后续姿态从所述相机收集的视觉数据所述对象的预期体积得以最小化;以及

基于所述控制输入来控制所述机器人设备移动到与所述后续姿态相对应的位置。

17. 如权利要求16所述的非瞬态计算机可读介质,其中,用于确定的程序代码进一步包括用于如下的程序代码:

基于使得用于最小化预期体积的成本最小化来确定所述控制输入。

18. 如权利要求16所述的非瞬态计算机可读介质,其中,用于确定的程序代码进一步包

括用于使用对所述多个潜在控制输入的批处理来确定所述控制输入的程序代码。

19. 如权利要求16所述的非瞬态计算机可读介质,其中,用于确定的程序代码进一步包括用于使用顺序处理来确定所述控制输入的程序代码。

20. 如权利要求16所述的非瞬态计算机可读介质,其中,用于控制的程序代码进一步包括用于使用最小数目的控制输入来控制所述机器人设备围绕所述对象运动的程序代码。

针对三维空间中的对象位置和范围的主动相机移动确定

[0001] 相关申请的交叉引用

[0002] 本申请要求于2015年12月2日提交且题为“ACTIVE CAMERA MOVEMENT DETERMINATION FOR OBJECT POSITION AND EXTENT IN THREE-DIMENSIONAL SPACE (针对三维空间中的对象位置和范围的主动相机移动确定)”的美国临时专利申请No.62/262,249的权益,其公开内容通过援引全部明确纳入于此。

[0003] 背景

[0004] 领域

[0005] 本公开的某些方面一般涉及机器学习,尤其涉及改进运动规划的系统和方法。

背景技术

[0006] 移动智能体(例如,机器人)在环境周围移动时经常面临未标识的对象和障碍物。当选择要在环境中采取的动作时,智能体可以选择减少它们与目标状态的距离、增加它们关于环境的信息、以及避免障碍物碰撞、或其某种组合的动作。

[0007] 在尝试以有限资源或在未知环境中采取动作时,收集关于环境的信息是重要目标。在经常遍历域中可能有用的一种类型的环境信息是该域中的对象的形状和范围。出于导航的目的,对象在智能体本身的大小数量级上的粗略范围是有用的。这一比例上关于对象的形状和范围的信息经常通过基于深度的传感器(诸如结构化光、激光和雷达类型的感测技术)或通过由许多相机视图中(从不止一个相机或一个相机在不同时间点)共同记录的特征所作出的深度计算来收集。然而,用于估计对象的形状和范围的大多数技术简单地利用了智能体的移动,但是没有将有利于估计对象的形状和范围的移动进行优先化。

[0008] 概述

[0009] 在本公开的一方面,给出了一种运动规划方法。该方法包括:从具有可控相机的智能体的第一姿态观察对象。该方法还包括:确定一个或多个后续控制输入来移动智能体和相机以从至少一个后续姿态观察该对象。(诸)后续控制输入被确定以使得基于从相机收集到的视觉数据对对象的预期包围测量被最小化。该方法进一步包括:基于(诸)后续控制输入来控制智能体和相机。

[0010] 在本公开的另一方面,给出了一种用于运动规划的装置。该装置包括存储器以及耦合至该存储器的至少一个处理器。该一个或多个处理器被配置成:从具有可控相机的智能体的第一姿态观察对象。(诸)处理器还被配置成:确定一个或多个后续控制输入来移动智能体和相机以从一个或多个后续姿态观察该对象。(诸)后续控制输入被确定以使得基于从相机收集到的视觉数据对对象的预期包围测量被最小化。(诸)处理器被进一步配置成:基于(诸)后续控制输入来控制智能体和相机。

[0011] 在本公开的又一方面,给出了一种用于运动规划的设备。该设备包括:用于从具有可控相机的智能体的第一姿态观察对象的装置。该设备还包括:用于确定一个或多个后续控制输入来移动智能体和相机以从一个或多个后续姿态观察该对象的装置。(诸)后续控制输入被确定以使得基于从相机收集到的视觉数据对对象的预期包围测量被最小化。该设备

进一步包括：用于基于(诸)后续控制输入来控制智能体和相机的装置。

[0012] 根据本公开的再一方面，给出了一种非瞬态计算机可读介质。该非瞬态计算机可读介质上编码有用于运动规划的程序代码。该程序代码由处理器执行，并且包括用于从具有可控相机的智能体的第一姿态来观察对象的程序代码。该程序代码还包括：用于确定一个或多个后续控制输入来移动智能体和相机以从一个或多个后续姿态观察该对象的程序代码。(诸)后续控制输入被确定以使得基于从相机收集到的视觉数据对对象的预期包围测量被最小化。该程序代码进一步包括：用于基于(诸)后续控制输入来控制智能体和相机的程序代码。

[0013] 本公开的附加特征和优点将在下文描述。本领域技术人员应当领会，本公开可容易被用作修改或设计用于实施与本公开相同的目的的其他结构的基础。本领域技术人员还应认识到，这样的等效构造并不脱离所附权利要求中所阐述的本公开的教导。被认为是本公开的特性的新颖特征在其组织和操作方法两方面连同进一步的目的是和优点在结合附图来考虑以下描述时将被更好地理解。然而，要清楚理解的是，提供每一幅附图均仅用于解说和描述目的，且无意作为对本公开的限定的定义。

[0014] 附图简述

[0015] 在结合附图理解下面阐述的详细描述时，本公开的特征、本质和优点将变得更加明显，在附图中，相同附图标记始终作相应标识。

[0016] 图1解说了根据本公开的某些方面的使用片上系统(SOC) (包括通用处理器) 来设计神经网络的示例实现。

[0017] 图2解说了根据本公开的各方面的系统的示例实现。

[0018] 图3解说了根据本公开的各方面的用于估计对象的形状的示例性技术。

[0019] 图4A-B是解说根据本公开的各方面的示例性视觉外壳的示意图。

[0020] 图5A-B是解说根据本公开的各方面的示例性联合视觉外壳的示意图。

[0021] 图6-7解说根据本公开的各方面的用于运动规划的方法。

[0022] 详细描述

[0023] 以下结合附图阐述的详细描述旨在作为各种配置的描述，而无意表示可实践本文中所描述的概念的仅有配置。本详细描述包括具体细节以便提供对各种概念的透彻理解。然而，对于本领域技术人员将显而易见的是，没有这些具体细节也可实践这些概念。在一些实例中，以框图形式示出众所周知的结构和组件以避免湮没此类概念。

[0024] 基于本教导，本领域技术人员应领会，本公开的范围旨在覆盖本公开的任何方面，不论其是与本公开的任何其他方面相独立地还是组合地实现的。例如，可以使用所阐述的任何数目的方面来实现装置或实践方法。另外，本公开的范围旨在覆盖使用作为所阐述的本公开的各个方面的补充或者与之不同的其他结构、功能性、或者结构及功能性来实践的此类装置或方法。应当理解，所披露的本公开的任何方面可由权利要求的一个或多个元素来实施。

[0025] 措辞“示例性”在本文中用于表示“用作示例、实例、或解说”。本文中描述为“示例性”的任何方面不必被解释为优于或胜过其他方面。

[0026] 尽管本文描述了特定方面，但这些方面的众多变体和置换落在本公开的范围之内。虽然提到了优选方面的一些益处和优点，但本公开的范围并非旨在被限定于特定益处、

用途或目标。相反,本公开的各方面旨在能宽泛地应用于不同的技术、系统配置、网络和协议,其中一些作为示例在附图以及以下对优选方面的描述中解说。详细描述和附图仅仅解说本公开而非限定本公开,本公开的范围由所附权利要求及其等效技术方案来定义。

[0027] 本公开的各方面涉及运动规划,尤其涉及改进在选择装备有单个相机的智能体(agent)(例如,机器人、无人机、或机动车)在移动序列中的移动时的效率。在本公开中解决的一个问题是,如何选择智能体的移动来最佳地估计未知环境中的对象或障碍物的形状和范围。例如,如果部署了无人机并且其观察到场景中的对象,则控制输入可被确定以移动无人机,从而使用视觉传感器(例如,单个相机)(而非深度传感器)来确定对象的形状和范围。在这样做时,对对象形状和范围的估计可以比(例如,采用随机或不相关移动的)常规方法更快和/或更准确地被执行。

[0028] 根据本公开的各方面,可以提供智能体的当前位置和环境的视觉相机视图作为输入。进而,本公开的系统和方法可以输出动作命令。例如,动作命令可以是对一组致动器的速度命令、或指定有限时间窗口上在空间中的轨迹的预编程运动原语的形式。

[0029] 在一些方面,可以处理相机帧以确定围绕环境中的可能对象的一个或多个边界框。因为至对象的距离可能是未知的,所以每个边界框(例如,二维边界框)可以限定矩形金字塔,其尖端以相机的焦点为中心并在图像平面处延伸通过该矩形。金字塔的底部可被限制成离相机图像平面存在一定距离。在一些方面,该距离可基于相机的已知分辨率或环境的比例来设置。例如,在室内操作的移动设备相机(例如,蜂窝电话相机)可具有比在室外操作的专业单镜头反光(SLR)相机小的最大金字塔范围。如此,用于移动设备相机的矩形金字塔可包括基于单个帧对对象的形状和范围的估计。

[0030] 随后可基于对对象形状和范围的该当前估计来选择移动。智能体(例如,机器人)可以采取所选移动。第二相机帧可被处理以确定围绕可能对象的附加边界框,并且可以确定对每个对象的形状和位置的新估计。该第二帧也产生矩形金字塔,然而,因为存在两个金字塔估计,所以对象位于这两个矩形金字塔的交集(其是比初始估计小的区域)内的置信度可增加。

[0031] 随着动作被采取,这个过程可在时间上重复。以此方式,对对象位置的新估计可被生成,并且对象的形状和范围可被确定。在一个示例性方面,可以选择下一动作,以使得下一相机测量之后的预期相交区域被最小化,受制于整个边界框在相机视图中保持可见的约束。相应地,可以选择比任意一组移动更有可能随时间减小(或甚至最小化)相交区域的移动。

[0032] 图1解说了根据本公开的某些方面使用片上系统(SOC) 100进行前述的运动规划的示例实现,SOC 100可包括通用处理器(CPU)或多核通用处理器(CPU) 102。变量(例如,神经信号和突触权重)、与计算设备相关联的系统参数(例如,带有权重的神经网络)、延迟、频率槽信息、以及任务信息可被存储在与神经处理单元(NPU) 108相关联的存储器块、与CPU 102相关联的存储器块、与图形处理单元(GPU) 104相关联的存储器块、与数字信号处理器(DSP) 106相关联的存储器块、专用存储器块118中,或可跨多个块分布。在通用处理器102处执行的指令可从与CPU 102相关联的程序存储器加载或可从专用存储器块118加载。

[0033] SOC 100还可包括为具体功能定制的附加处理块(诸如GPU 104、DSP 106、连通性块110(其可包括第四代长期演进(4G LTE)连通性、无执照Wi-Fi连通性、USB连通性、蓝牙连

通性等))以及例如可检测和识别姿态的多媒体处理器112。在一种实现中,NPU实现在CPU、DSP、和/或GPU中。SOC 100还可包括传感器处理器114、图像信号处理器 (ISP)、和/或导航120 (其可包括全球定位系统)。

[0034] SOC 100可基于ARM指令集。在本公开的一方面,加载到通用处理器102中的指令可包括用于从具有可控相机的智能体的第一姿态观察对象的代码。加载到通用处理器102中的指令还可包括用于确定至少一个后续控制输入来移动智能体和相机以从后续姿态观察该对象以减小基于从相机收集的视觉数据对对象的预期包围测量的代码。加载到通用处理器102中的指令还可包括用于基于后续控制输入来控制智能体和相机的代码。

[0035] 图2解说了根据本公开的某些方面的系统200的示例实现。如图2中所解说的,系统200可具有可执行本文中所描述的方法的各种操作的多个局部处理单元202。每个局部处理单元202可包括局部状态存储器204和可存储神经网络的参数的局部参数存储器206。另外,局部处理单元202可具有用于存储局部模型程序的局部(神经元)模型程序(LMP)存储器208、用于存储局部学习程序的局部学习程序(LLP)存储器210、以及局部连接存储器212。此外,如图2中所解说的,每个局部处理单元202可与用于为该局部处理单元的各局部存储器提供配置的配置处理器单元214对接,并且与提供各局部处理单元202之间的路由的路由连接处理单元216对接。

[0036] 在一种配置中,机器学习模型被配置成用于从具有可控相机的智能体的第一姿态观察对象。该模型还被配置成用于确定后续控制输入来移动智能体和相机以从后续姿态观察该对象,以最小化基于从相机收集的视觉数据对对象的预期包围测量。该模型被进一步配置成用于基于(诸)后续控制输入来控制智能体和相机。该模型包括观察装置、确定装置、和/或控制装置。在一个方面,观察装置、确定装置、和/或控制装置可以是配置成执行所叙述功能的通用处理器102、与通用处理器102相关联的程序存储器、存储器块118、局部处理单元202、和/或路由连接处理单元216。在另一配置中,前述装置可以是配置成执行由前述装置所叙述的功能的任何模块或任何装置。

[0037] 根据本公开的某些方面,每个局部处理单元202可被配置成基于模型的一个或多个期望功能特征来确定模型的参数,以及随着所确定的参数被进一步适配、调谐和更新来使这一个或多个功能特征朝着期望的功能特征发展。

[0038] 图3是解说用于估计对象的形状的示例性技术的示图。参照图3,在图像(例如,红绿蓝(RGB)图像)中示出了单个对象(例如,示出了阴影308的桌子302)。当然,这仅仅是为了便于解说和理解的示例,并且附加对象可被包括在图像中。使用对象定位过程,可以生成对象或边界框304的二维(2D)轮廓。彩色或灰度解决方案可被用于边界框检测。2D轮廓306可以由 $s_k(x_i, y_i) \in B$ 来表示,其中 $B = \{0, 1\}$ 是二元空间。例如,如果 $s_k(x_i, y_i) = 1$,则对象在图像 I_k 的像素 (x_i, y_i) 中是可见的。否则,对象在像素 (x_i, y_i) 中是不可见的。一个目标是,使用2D图像来确定对象或轮廓图像306的3D轮廓(例如,三维边界框)。

[0039] 给定RGB图像序列 $I_{0:t} = \{I_0, \dots, I_t\}$ 及相关联的相机姿态 $p_{0:t} = \{p_0, \dots, p_t\}$,可以计算对象轮廓序列 $s_{0:t} = \{s_0, \dots, s_t\}$ 。

[0040] 在每个时间步阶 k 处,相机固有特征 M 、相机姿态 p_k 和轮廓图像 s_k 可被用来计算逆向投影锥体 $C_k(x_w, y_w, z_w) \in B$ 或对象的视觉外壳。在一个示例中, (x_i, y_i) 可以是点 (x_w, y_w, z_w) 在当前时间步阶的相机图像帧上的投影。在该情形中, $C_k(x_w, y_w, z_w) = s_k(x_i, y_i)$ 。换言之,如果

$C_k(x_w, y_w, z_w) = 1$, 则轮廓 s_k 已经指示3D对象可能潜在地包含点 (x_w, y_w, z_w) 。否则, 对象必然不包含点 (x_w, y_w, z_w) 。相应地, 视觉外壳测量模型可被表达为 $C_k = (p_k, s_k, M)$ 。

[0041] 相机可以从多个不同姿态 $p_{0:t}$ 拍摄对象的多张快照或照片, 从而获得对应的轮廓 $s_{0:t}$ 。相机固有特征 M 、轮廓和/或相机姿态可进而被用来在每个时间步阶计算对应的视觉外壳 $C_{0:t}$ 。联合视觉外壳 V_t 可随后被计算为由下式给出的各视觉外壳的交集:

$$[0042] \quad V_t = \bigcap_{k=0}^t C_k \quad (1)$$

[0043] 联合视觉外壳提供3D空间中的形状和位置的近似。

[0044] 在一些方面, 视觉外壳可被用来确定用于移动相机的后续控制输入, 以使得关于对象的联合视觉外壳的测量 m 可被减小或最小化。用于产生这种所谓的active shape-from-silhouette (主动从轮廓到形状) (ASfS) 重构的后续控制输入可由下式给出:

$$[0045] \quad u_{t+1}^* = \underset{u \in U}{\operatorname{argmin}} m\left(\bigcap_{k=0}^{t+1} C_k\right) \text{ 受约束于 } C_k = h(p_k, s_k, M) \quad (2)$$

$$[0046] \quad p_{k+1} = f(p_k, u_k)$$

$$[0047] \quad s_k \neq 0 \quad \forall k,$$

[0048] 其中 u 是来自可能移动集合 U 的用于移动相机和/或智能体的控制输入, 并且 h 和 f 是相机动态和观察模型的函数。控制输入 u 可以是包括方向和速度的向量。变量 C_k 是锥体或视觉外壳, p_k 是姿态 (例如, 相机在3D空间中的位置和/或取向), s_k 是轮廓 (或2D边界框), k 是时间步阶或锥体数目)。变量 t 是时间, 并且 M 是相机属性 (例如, 镜头类型)。变量 m 是减小的测量 (例如, 体积)。测量 m 还可以是表面积、高度、或宽度——例如, 如果在单个方向上移动并且目标是避开对象。

[0049] 使用对联合视觉外壳的最小化测量 (例如, 体积), 可确定后续控制输入以将智能体和/或相机移动到从其观察对象的后续位置。新的联合视觉外壳可以使用后续位置处的视觉数据来确定。新的联合视觉外壳可被类似地最小化以确定后续控制输入。以此方式, 对象的形状和范围可被高效地确定。在一种配置中, 该信息被存储, 从而创建地图以供稍后重新使用。

[0050] 在式2中定义的后续控制输入可以使用批次处理或以流线或顺序方式来确定。在一个示例中, 当采用批次处理时, 所有的控制输入 $u_{0:t}$ 可被离线地确定。在另一示例中, 当后续控制输入以流线或顺序方式确定时, 在给定输入 $u_{0:t}$ 和观察到的视觉外壳 $C_{0:t}$ 的历史的情况下, 后续控制输入 u_{t+1} 可在线递增地确定。

[0051] 图4A-4B是解说根据具有姿态的相机406从图像平面进行视觉外壳投影的示图。参照图4A, 在现实世界环境中, 在具有姿态 p_k 的相机406的视场410内观察到月牙形对象402。相机406的姿态 p_k 产生月牙形对象的图像平面轮廓404的投影。图4B示出了对象402的视觉外壳412。在一些方面, 视觉外壳412指示对象的潜在位置。视觉外壳412可基于相机固有特征 (例如, 相机镜头类型 (例如, 鱼眼镜头))、相机406的姿态 (例如, 位置和/或取向)、和/或对象的轮廓404来确定。视觉外壳在3D中, 而轮廓在2D中。

[0052] 图5A-5B解说了根据本公开的各方面的示例性联合视觉外壳。如图5A中所示, 在第一姿态的相机 (例如, 相机406) 的第一视图中观察到月牙形对象508, 从而产生对象轮廓502。在第二姿态的相机的第二视图也观察到该月牙形对象, 从而产生第二对象轮廓504。对象轮廓502和504、相机固有特征、和/或对应的相机姿态可被用来分别确定视觉外壳512

和514,如图5B中所示。

[0053] 在图5B中,联合视觉外壳510被确定为视觉外壳512和514的交集。联合视觉外壳510使用由相机产生的2D图像来提供对3D空间中的对象形状和位置的近似。此外,根据本公开的各方面,视觉外壳512和514可被用来确定后续控制输入以将相机(和/或智能体)移动到用于捕捉对象的图像的位置以使得关于联合视觉外壳510的测量 m 可被最小化。

[0054] 图6解说了根据本公开的各方面的用于运动规划的方法600。在框602,该过程从具有可控相机的智能体的第一姿态观察对象。

[0055] 在框604,该过程确定后续控制输入来移动智能体和/或相机以从后续姿态观察该对象,以减少或甚至最小化基于从相机收集的视觉数据对对象(例如,视觉外壳)的预期包围测量(例如,体积)。在一些方面,后续控制输入可以顺序地或使用顺序处理(例如,一次处理一个控制输入)或通过使用对潜在后续控制输入的批次处理来确定。批次处理可以使用诸如滚动时域控制(例如,预测将来的10步,并且随后执行下一分析)之类的技术、或其他分析技术。另外,可确定后续控制输入以最小化预期包围体积(例如,联合视觉外壳)或降低该最小化的成本。在一些方面,成本可以包括移动智能体或相机来确定对象的形状和范围时花费的努力、时间、工作、和/或能量。

[0056] 此外,在框606,该过程基于后续控制输入来控制智能体和相机。在一些方面,可使用最小数目个控制输入来控制智能体和相机围绕对象移动。

[0057] 图7是解说根据本公开的各方面的运动规划方法700的框图。在框702,该过程从具有可控相机的智能体的第一姿态观察对象以产生第一2D对象轮廓或轮廓图像。在框704,该过程计算第一视觉外壳。第一视觉外壳可基于相机固有特征(诸如相机镜头类型)、相机的姿态、第一对象轮廓、或其组合来计算。第一视觉外壳是三维的并且可以包括对象可位于其中的体积。

[0058] 在框706,可以从智能体或相机的第二姿态观察该对象。在一些方面,第二姿态可被随机地选择。第二2D对象轮廓被产生。在框708,该过程计算第二视觉外壳。类似于第一视觉外壳,第二视觉外壳可基于相机固有特征、第二对象轮廓或其组合来计算。

[0059] 在框710,该过程基于第一视觉外壳和第二视觉外壳来计算联合视觉外壳。联合视觉外壳可被计算为第一视觉外壳和第二视觉外壳的交集。如此,联合视觉外壳可以提供对象位于由该交集限定的空间内的更大置信度。此外,联合视觉外壳还可以提供3D空间中的对象形状的近似。

[0060] 在框712,该过程确定后续控制输入来将智能体和/或相机移动到下一姿态以最小化联合视觉外壳。也就是说,并非根据基于随机选择过程或基于递增步骤过程所确定的控制输入来移动,后续控制输入可被选择成使得最小化联合视觉外壳,并且由此更高效地确定感兴趣的对象的形状和范围。

[0061] 在框714,该过程评估用于确定控制输入的处理类型。如果在框714中处理类型是顺序处理,则在框716,该过程基于所确定的后续控制输入来控制智能体和/或相机移动。

[0062] 另一方面,如果指示批次处理,则在框718,该过程评估是否已经达到期望批次大小。批次大小可根据设计偏好来任意地确定。如果尚未达到期望批次大小(例如,10个后续控制输入),则该过程返回到框712以确定下一后续控制输入。在这一场景中,相机在下一时间步阶中并未实际上移动到下一位置。相反,在一些方面,对象的投影视觉外壳是针对相机

的下一姿态来确定的,并且可被用来确定下一后续控制输入。

[0063] 如果在框718已经达到了期望批次大小,则在框716,该过程基于所确定的后续控制输入来控制智能体和/或相机移动。在这一场景中,智能体是基于批次中最后确定的控制输入来移动的。

[0064] 在一些方面,方法600和700可由SOC 100 (图1) 或系统200 (图2) 执行。即,举例而言但不作为限定,方法600和700的每个元素可由SOC 100或系统200、或者一个或多个处理器(例如,CPU 102和局部处理单元202) 和/或其中包括的其他组件来执行。

[0065] 以上所描述的方法的各种操作可由能够执行相应功能的任何合适的装置来执行。这些装置可包括各种硬件和/或(诸) 软件组件和/或(诸) 模块,包括但不限于电路、专用集成电路(ASIC)、或处理器。一般而言,在附图中有解说的操作的场合,那些操作可具有带相似编号的相应配对装置加功能组件。

[0066] 如本文所使用的,术语“确定”涵盖各种各样的动作。例如,“确定”可包括演算、计算、处理、推导、研究、查找(例如,在表、数据库或其他数据结构中查找)、探知及诸如此类。另外,“确定”可包括接收(例如接收信息)、访问(例如访问存储器中的数据)、及类似动作。此外,“确定”可包括解析、选择、选取、确立及类似动作。

[0067] 如本文中所使用的,引述一系列项目中的“至少一个”的短语是指这些项目的任何组合,包括单个成员。作为示例,“a、b或c中的至少一个”旨在涵盖:a、b、c、a-b、a-c、b-c、以及a-b-c。

[0068] 结合本公开所描述的各种解说性逻辑框、模块、以及电路可用设计成执行本文中所描述的功能的通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列信号(FPGA) 或其他可编程逻辑器件(PLD)、分立的门或晶体管逻辑、分立的硬件组件或其任何组合来实现或执行。通用处理器可以是微处理器,但在替换方案中,处理器可以是任何市售的处理器、控制器、微控制器、或状态机。处理器还可以被实现为计算设备的组合,例如,DSP与微处理器的组合、多个微处理器、与DSP核心协同的一个或多个微处理器、或任何其他此类配置。

[0069] 结合本公开描述的方法或算法的步骤可直接在硬件中、在由处理器执行的软件模块中、或在这两者的组合中实施。软件模块可驻留在本领域所知的任何形式的存储介质中。可使用的存储介质的一些示例包括随机存取存储器(RAM)、只读存储器(ROM)、闪存、可擦除可编程只读存储器(EPROM)、电可擦除可编程只读存储器(EEPROM)、寄存器、硬盘、可移动盘、CD-ROM,等等。软件模块可包括单条指令、或许多条指令,且可分布在若干不同的代码段上,分布在不同的程序间以及跨多个存储介质分布。存储介质可被耦合到处理器以使得该处理器能从/向该存储介质读写信息。在替换方案中,存储介质可以被整合到处理器。

[0070] 本文所公开的方法包括用于达成所描述的方法的一个或多个步骤或动作。这些方法步骤和/或动作可以彼此互换而不会脱离权利要求的范围。换言之,除非指定了步骤或动作的特定次序,否则具体步骤和/或动作的次序和/或使用可以改动而不会脱离权利要求的范围。

[0071] 所描述的功能可在硬件、软件、固件或其任何组合中实现。如果以硬件实现,则示例硬件配置可包括设备中的处理系统。处理系统可以用总线架构来实现。取决于处理系统的具体应用和整体设计约束,总线可包括任何数目的互连总线和桥接器。总线可将包括处

理器、机器可读介质、以及总线接口的各种电路链接在一起。总线接口可用于尤其将网络适配器经由总线连接至处理系统。网络适配器可用于实现信号处理功能。对于某些方面，用户接口（例如，按键板、显示器、鼠标、操纵杆，等等）也可以被连接到总线。总线还可以链接各种其他电路，诸如定时源、外围设备、稳压器、功率管理电路以及类似电路，它们在本领域中是众所周知的，因此将不再进一步描述。

[0072] 处理器可负责管理总线和一般处理，包括执行存储在机器可读介质上的软件。处理器可用一个或多个通用和/或专用处理器来实现。示例包括微处理器、微控制器、DSP处理器、以及其他能执行软件的电路系统。软件应当被宽泛地解释成意指指令、数据、或其任何组合，无论是被称作软件、固件、中间件、微代码、硬件描述语言、或其他。作为示例，机器可读介质可包括随机存取存储器 (RAM)、闪存存储器、只读存储器 (ROM)、可编程只读存储器 (PROM)、可擦式可编程只读存储器 (EPROM)、电可擦式可编程只读存储器 (EEPROM)、寄存器、磁盘、光盘、硬驱动器、或者任何其他合适的存储介质、或其任何组合。机器可读介质可被实施在计算机程序产品中。该计算机程序产品可以包括包装材料。

[0073] 在硬件实现中，机器可读介质可以是处理系统中与处理器分开的一部分。然而，如本领域技术人员将容易领会的，机器可读介质或其任何部分可在处理系统外部。作为示例，机器可读介质可包括传输线、由数据调制的载波、和/或与设备分开的计算机产品，所有这些都可由处理器通过总线接口来访问。替换地或补充地，机器可读介质或其任何部分可被集成到处理器中，诸如高速缓存和/或通用寄存器文件可能就是这种情形。虽然所讨论的各种组件可被描述为具有特定位置，诸如局部组件，但它们也可按各种方式来配置，诸如某些组件被配置成分布式计算系统的一部分。

[0074] 处理系统可以被配置为通用处理系统，该通用处理系统具有一个或多个提供处理器功能性的微处理器、以及提供机器可读介质中的至少一部分的外部存储器，它们都通过外部总线架构与其他支持电路系统链接在一起。替换地，该处理系统可以包括一个或多个神经元形态处理器以用于实现本文中所描述的神经元模型和神经系统模型。作为另一替换方案，处理系统可以用带有集成在单块芯片中的处理器、总线接口、用户接口、支持电路系统、和至少一部分机器可读介质的专用集成电路 (ASIC) 来实现，或者用一个或多个现场可编程门阵列 (FPGA)、可编程逻辑器件 (PLD)、控制器、状态机、门控逻辑、分立硬件组件、或者任何其他合适的电路系统、或者能执行本公开通篇所描述的各种功能性的电路的任何组合来实现。取决于具体应用和加诸于整体系统上的总设计约束，本领域技术人员将认识到如何最佳地实现关于处理系统所描述的功能性。

[0075] 机器可读介质可包括数个软件模块。这些软件模块包括当由处理器执行时使处理系统执行各种功能的指令。这些软件模块可包括传送模块和接收模块。每个软件模块可以驻留在单个存储设备中或者跨多个存储设备分布。作为示例，当触发事件发生时，可以从硬驱动器中将软件模块加载到RAM中。在软件模块执行期间，处理器可以将一些指令加载到高速缓存中以提高访问速度。可随后将一个或多个高速缓存行加载到通用寄存器文件中以供处理器执行。在以下述及软件模块的功能性时，将理解此类功能性是在处理器执行来自该软件模块的指令时由该处理器来实现的。此外，应领会，本公开的各方面产生对处理器、计算机、机器或实现此类方面的其它系统的机能的改进。

[0076] 如果以软件实现，则各功能可作为一条或多条指令或代码存储在计算机可读介质

上或藉其进行传送。计算机可读介质包括计算机存储介质和通信介质两者,这些介质包括促成计算机程序从一地向另一地转移的任何介质。存储介质可以是能被计算机访问的任何可用介质。作为示例而非限定,此类计算机可读介质可包括RAM、ROM、EEPROM、CD-ROM或其他光盘存储、磁盘存储或其他磁存储设备、或能用于携带或存储指令或数据结构形式的期望程序代码且能被计算机访问的任何其他介质。另外,任何连接也被正当地称为计算机可读介质。例如,如果软件是使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)、或无线技术(诸如红外(IR)、无线电、以及微波)从web网站、服务器、或其他远程源传送而来,则该同轴电缆、光纤电缆、双绞线、DSL或无线技术(诸如红外、无线电、以及微波)就被包括在介质的定义之中。如本文中所使用的盘(disk)和碟(disc)包括压缩碟(CD)、激光碟、光碟、数字多用碟(DVD)、软盘、和蓝光[®]碟,其中盘(disk)常常磁性地再现数据,而碟(disc)用激光来光学地再现数据。因此,在一些方面,计算机可读介质可包括非瞬态计算机可读介质(例如,有形介质)。另外,对于其他方面,计算机可读介质可包括瞬态计算机可读介质(例如,信号)。上述的组合应当也被包括在计算机可读介质的范围内。

[0077] 因此,某些方面可包括用于执行本文中给出的操作的计算机程序产品。例如,此类计算机程序产品可包括其上存储(和/或编码)有指令的计算机可读介质,这些指令能由一个或多个处理器执行以执行本文中所描述的操作。对于某些方面,计算机程序产品可包括包装材料。

[0078] 此外,应当领会,用于执行本文中所描述的方法和技术的模块和/或其它恰适装置能由用户终端和/或基站在适用的场合下载和/或以其他方式获得。例如,此类设备能被耦合至服务器以促成用于执行本文中所描述的方法的装置的转移。替换地,本文中所描述的各种方法能经由存储装置(例如,RAM、ROM、诸如压缩碟(CD)或软盘等物理存储介质等)来提供,以使得一旦将该存储装置耦合至或提供给用户终端和/或基站,该设备就能获得各种方法。此外,可利用适于向设备提供本文中所描述的方法和技术的任何其他合适的技术。

[0079] 将理解,权利要求并不被限定于以上所解说的精确配置和组件。可在以上所描述的方法和装置的布局、操作和细节上作出各种改动、更换和变形而不会脱离权利要求的范围。

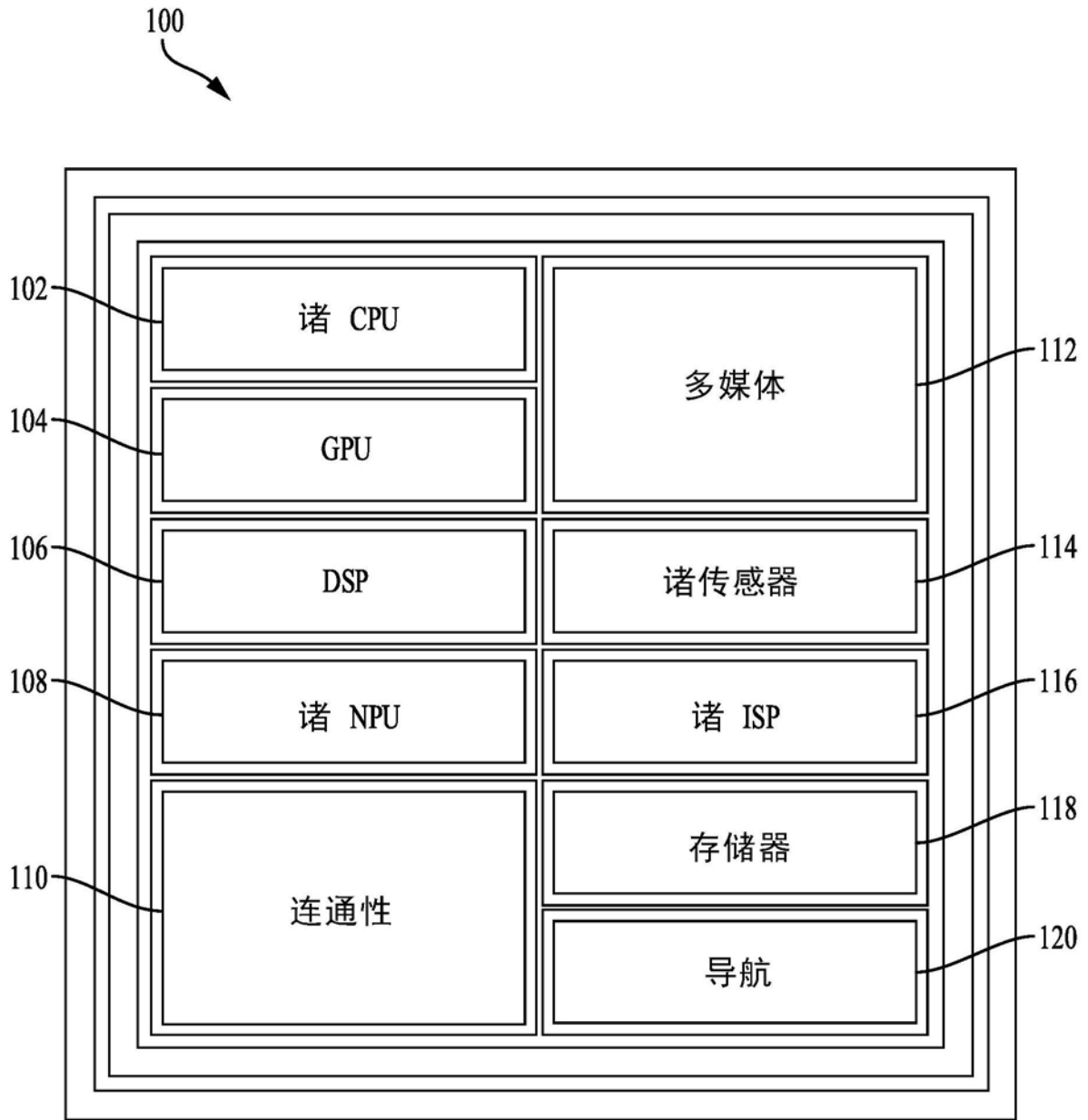


图1

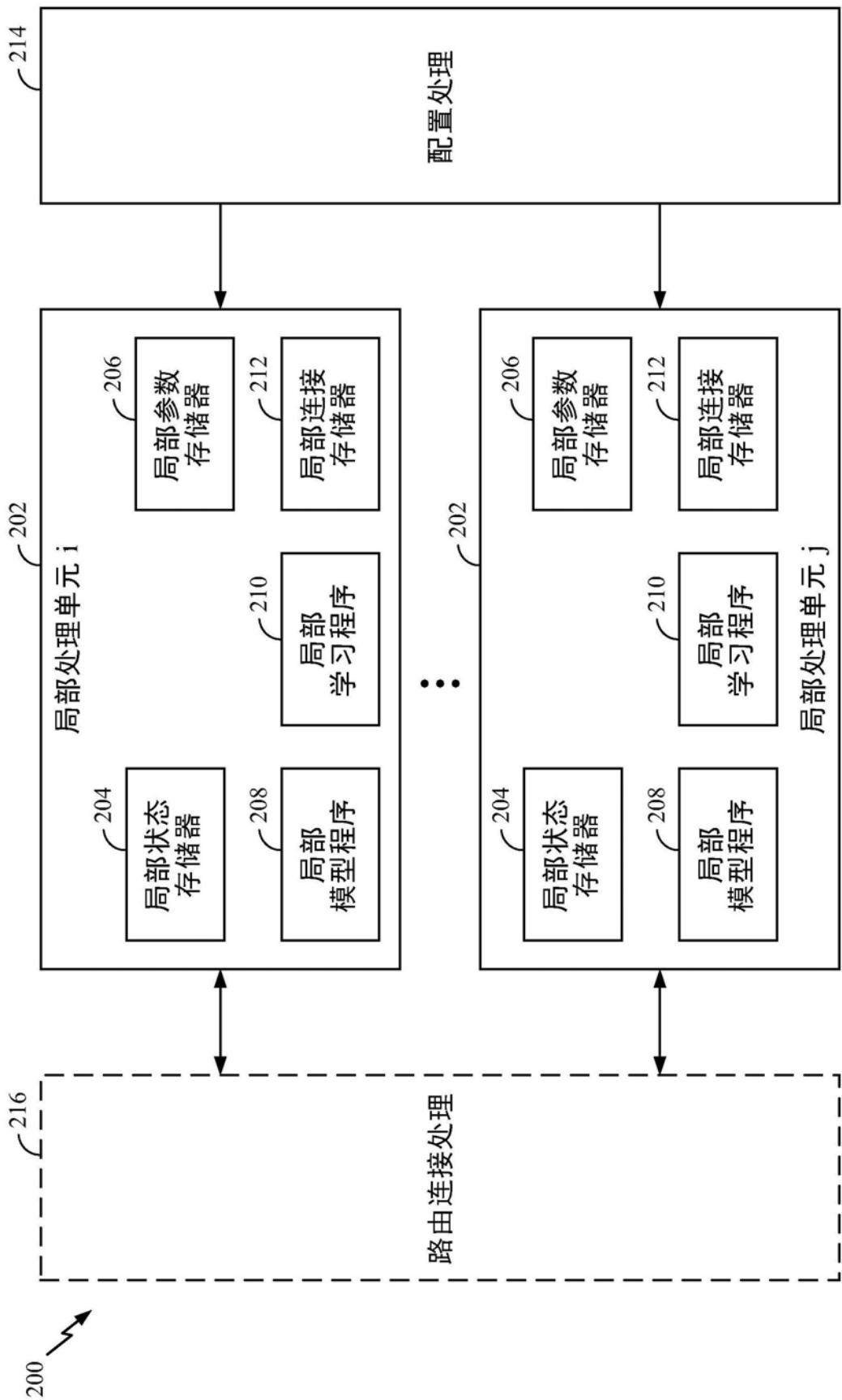


图2

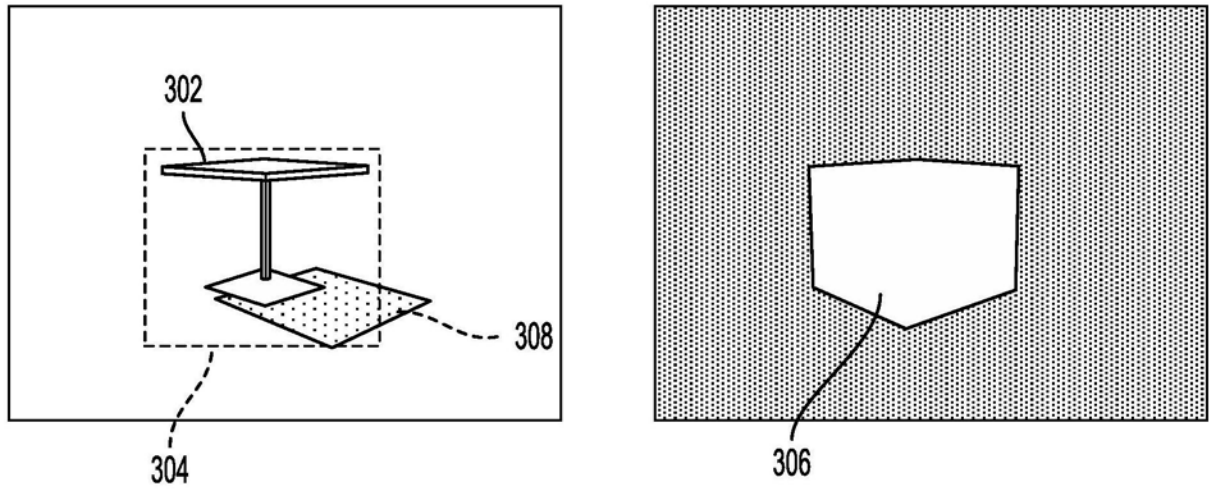


图3

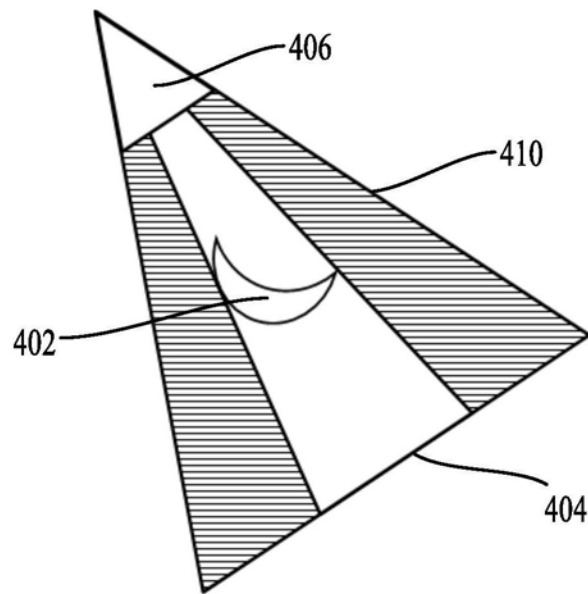


图4A

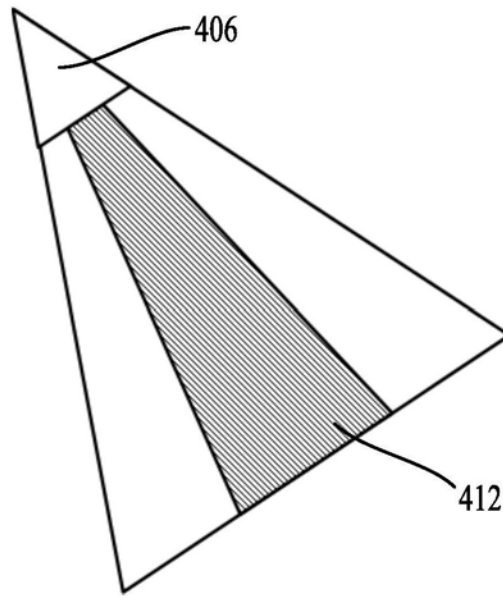


图4B

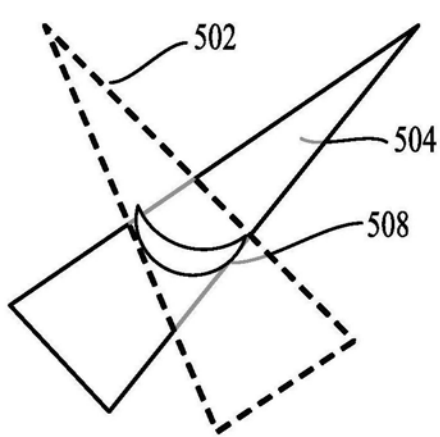


图 5A

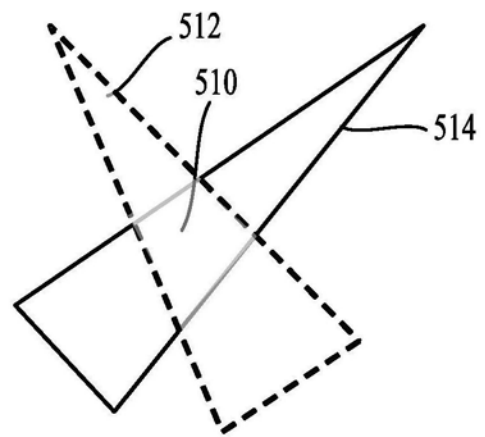


图 5B

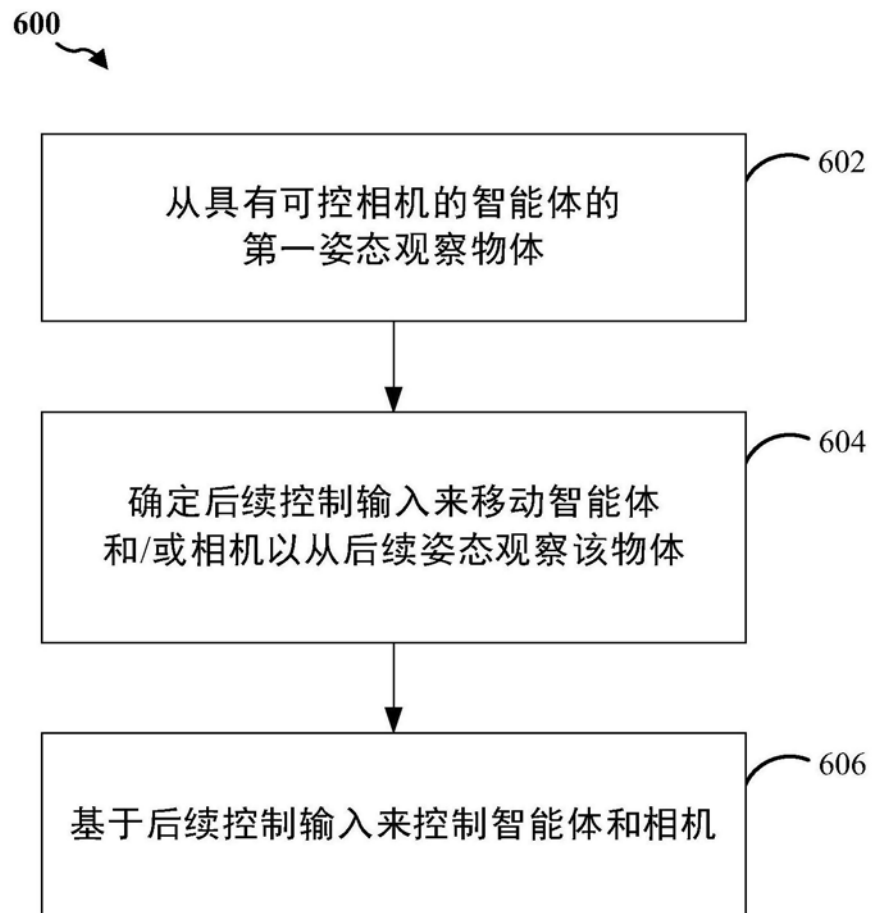


图6

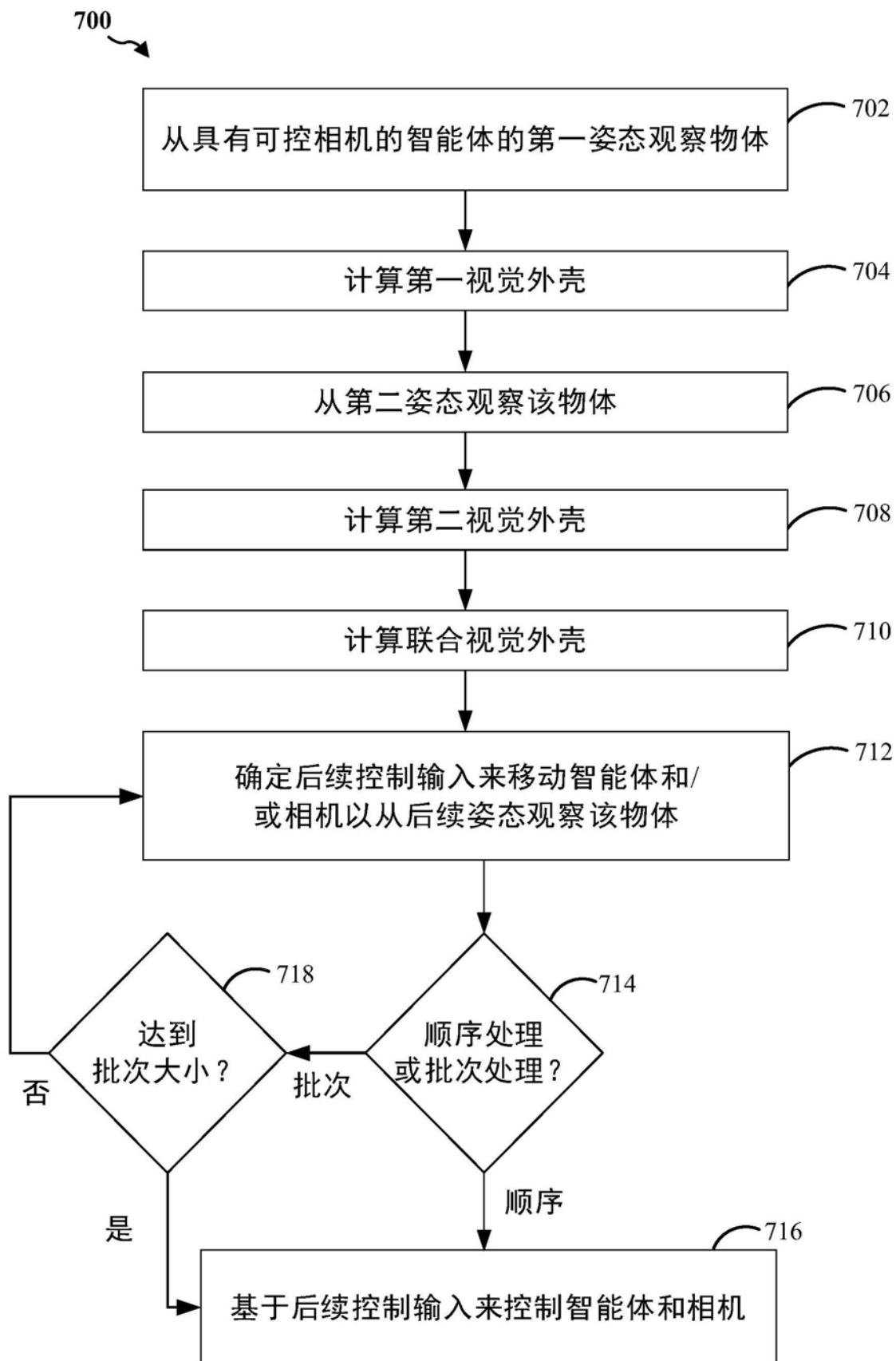


图7