

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 10 月 15 日 (15.10.2020)



(10) 国际公布号
WO 2020/207017 A1

(51) 国际专利分类号:
B25J 9/16 (2006.01) *B25J 15/02* (2006.01)
B25J 13/08 (2006.01) *B25J 15/08* (2006.01)

(21) 国际申请号: PCT/CN2019/119079

(22) 国际申请日: 2019 年 11 月 18 日 (18.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910289751.2 2019 年 4 月 11 日 (11.04.2019) CN

(71) 申请人: 上海交通大学 (SHANGHAI JIAO TONG UNIVERSITY) [CN/CN]; 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。

(72) 发明人: 刘成良 (LIU, Chengliang); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 贡亮

(GONG, Liang); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 李旭东 (LI, Xudong); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 汪韬 (WANG, Tao); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 沈晓晔 (SHEN, Xiaoye); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 林晨晖 (LIN, Chenhui); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 陶建峰 (TAO, Jianfeng); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。

(74) 代理人: 上海汉声知识产权代理有限公司 (SHANGHAI HANGSOME INTELLECTUAL PROPERTY LTD.); 中国上海市闵行区银都路 3828 弄 56 号 307 室, Shanghai 201108 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,

(54) **Title:** METHOD AND DEVICE FOR COLLABORATIVE SERVO CONTROL OF UNCALIBRATED MOVEMENT VISION OF ROBOT IN AGRICULTURAL SCENE

(54) 发明名称: 农业场景无标定机器人运动视觉协同伺服控制方法与设备

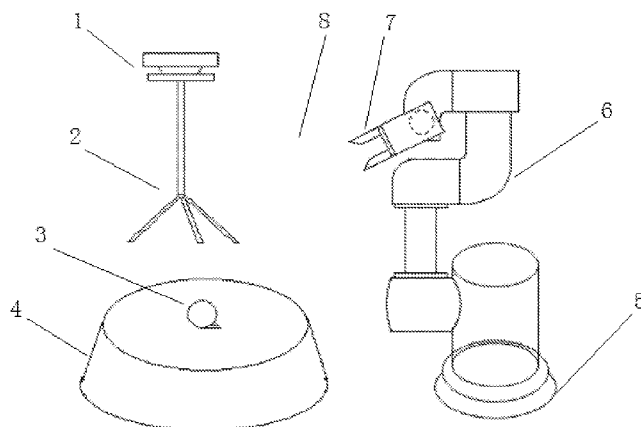


图 1

(57) **Abstract:** Provided are a method and device for the collaborative servo control of the uncalibrated movement vision of a robot in an agricultural scene. A mechanical gripper (7) is mounted at an arm tail end of a mechanical arm (6) of the device, and a target grabbing object (3) is located within the grabbing range of the mechanical arm (6); a control module (9) drives the mechanical gripper (7) to grab the target grabbing object and controls an image sensor (1) to carry out image sampling for the process of the mechanical arm (6) grabbing the target grabbing object (3); and the image sensor (1) sends sampled image data to the control module (9). The target grabbing object (3) and related environment within a scene do not need to be precisely spatially calibrated for the device, and the mechanical arm carries out strategy guidance according to a trained network so as to complete a grabbing task; moreover, the requirements for a space sensing device are lower, environment adaptation is high, and the device may be transferred to multiple tasks.

BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 提供了一种农业场景无标定机器人运动视觉协同伺服控制设备与方法, 其中, 设备的机械臂(6)在臂末端安装有机械抓手(7), 目标抓取物(3)处在机械臂(6)的可抓取范围内; 控制模块(9)驱动机械抓手(7)抓取目标抓取物, 并控制图像传感器(1)对机械臂(6)抓取目标抓取物(3)的过程进行图像采样; 图像传感器(1)将采样的图像数据发送给控制模块(9)。这种设备无需对于场景内的目标抓取物(3)及相关环境进行精确的空间标注, 机械臂将按照训练好的网络进行策略引导完成抓取任务, 对于空间感知设备的要求更低, 环境适应性强, 并可迁移至多种任务。

农业场景无标定机器人运动视觉协同伺服控制方法与设备

技术领域

本申请涉及机器人领域，特别涉及农业场景无标定机器人运动视觉协同伺服控制方法与设备。

背景技术

机器人作为精密高效的执行机械在军事、医学、制造业等领域有着广泛的应用，集成了电子、传感、控制等多种技术，机器人得以在工业场景中实现不同的功能；为了应对更为复杂多变的任务需求、提升机器人的性能与智能，智能型机器人的概念被提了出来，即应用了如强化学习控制、规划模式识别等学习型控制算法与图像视觉计算或深度神经网络的智能感知技术，具备自适应、自学习功能的一类机器人，在保障了机器人的工作精度与鲁棒性的同时，环境适应性与任务柔性得以提升。

经过对现有技术的检索发现，中国专利文献号 CN106041941A，公开日 2016.10.26，公开了一种工业机械臂的轨迹规划方法与装置，为同样领域内实现工业机械臂驱动的方法，该技术针对 SCARA 机械臂的每个关节构建坐标系与工作区域，通过预先计算输入控制方向射线与工作区域边界的交点，优化机械臂的速度规划过程。但该技术需要精确标定场景，获得目标终点的空间坐标后进行模式化的轨迹规划驱动，对于标定设备及技术要求高，对于不同场景适应能力差，尤其无法适应复杂多变的农业非结构化场景。

对于另一项现有技术文献，中国专利文献号 CN105353772A，公开日 2016.02.24，公开了一种无人机定位跟踪的视觉伺服控制方法，该技术通过在无人机上安装定位装置，惯性测量单元与摄像机获得大地、机体、相机与图像坐标系数据，通过运算各坐标系之间的相对过渡关系，以控制无人机拍

摄目标物位于图像中心。该技术结合了视觉传感实现基于视觉伺服的无人机控制，但该技术仅能计算简单场景下无人机的目标姿态规划问题，难以迁移应用至农业机器人的应用领域，给出无标定农业场景下的执行策略。

发明内容

本发明的目的在于提供对空间感知设备的要求更低，环境适应性强的农业场景无标定机器人运动视觉协同伺服控制方法和设备。

为了解决上述问题，本发明提供了一种农业场景无标定机器人运动视觉协同伺服控制设备，包括机械臂、目标抓取物、图像传感器和控制模块，其中，所述机械臂在臂末端安装有机械抓手；所述目标抓取物处在所述机械臂的可抓取范围内；所述控制模块分别与所述机械臂和图像传感器电连接，所述控制模块驱动所述机械抓手抓取所述目标抓取物，并控制所述图像传感器对所述机械臂抓取所述目标抓取物的过程进行图像采样；所述图像传感器将采样的图像数据发送给所述控制模块。

优选地，所述机械臂为六自由度机械臂。

本发明还提供了一种农业场景无标定机器人运动视觉协同伺服控制方法，包括如下步骤：

构建场景空间特征向量获取网络，获取场景空间特征特征向量；

获取示教动作样本；

构建逆强化回报值策略网络；

逆强化回报值策略网络迁移训练；

基于视觉特征提取网络与逆强化回报值策略网络，采用策略引导驱动算法获得正向引导规划结果。

优选地，所述场景空间特征向量获取网络为视觉卷积神经网络。

优选地，所述获取场景空间特征特征向量具体为：

图像传感器对机械臂抓取目标抓取物的过程进行图像采样，并提取 RGB

图像信息；

以所述图像信息作为所述场景空间特征向量获取网络的输入量，输出向量即为场景空间特征特征向量。

优选地，所述获取示教动作样本具体为：

牵引机械臂完成对目标抓取物的抓取，获取一次示教抓取的示教抓取动作数据；

驱动机械臂模拟示教抓取动作数据，自主完成对目标抓取物的抓取动作，用以拍摄获取示教抓取场景图像特征数据；

基于所述示教抓取动作数据和示教抓取场景图像特征数据整合得到示教动作样本。

优选地，所述构建逆强化回报值策略网络具体为：

构建用于拟合表示回报值的逆强化回报值策略网络；

通过仿真域随机化算法生成仿真参数；

使用 ROS 规划库规划模拟虚拟抓取动作，并采样得到模拟抓取路径；

逆强化回报值策略网络仿真预训练。

优选地，所述逆强化回报值策略网络迁移训练具体为：使用所述示教动作样本进行所述逆强化回报值策略网络的优化训练。

与现有技术相比，本发明存在以下技术效果：

1、本发明实施例农业场景无标定机器人运动视觉协同伺服控制方法无需对于场景内的目标抓取物及相关环境进行精确的空间标注，机械臂将按照训练好的网络进行策略引导完成抓取任务，对于空间感知设备的要求更低，环境适应性强，并可迁移至多种任务。

2、本发明实施例农业场景无标定机器人运动视觉协同伺服控制方法构建了场景空间特征向量获取网络用于获取场景特征，并在仿真环境内通过域随机化算法模拟抓取，使用仿真数据进行逆强化回报值策略网络的预训练；场

景空间特征向量获取网络和逆强化回报值策略网络分别的进行预训练，将传统复杂视觉运动伺服问题解耦处理，降低了训练网络的复杂度。

3、本发明实施例农业场景无标定机器人运动视觉协同伺服控制方法域随机化算法可以快速的生成大量的训练数据，减少示教员的示教操作数量，在有限的时间与资源内提升网络的训练效果。

当然，实施本发明的任一产品并不一定需要同时达到以上所述的所有优点。

附图说明

为了更清楚地说明本发明实施例的技术方案，下面将对实施例描述中所需要使用的附图作简单的介绍，显而易见，下面描述中的附图仅仅是本发明的一些实施例，对于本领域技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。附图中：

图1为本发明实施例农业场景无标定机器人运动视觉协同伺服控制设备结构示意图；

图2为本发明实施例农业场景无标定机器人运动视觉协同伺服控制设备硬件连接示意图；

图3为本发明实施例农业场景无标定机器人运动视觉协同伺服控制设备软件配置层次图；

图4为本发明实施例农业场景无标定机器人运动视觉协同伺服控制方法流程图；

图5为本发明实施例场景空间特征向量获取网络的网络结构。

具体实施方式

以下将结合附图对本发明提供的农业场景无标定机器人运动视觉协同伺服控制方法与设备进行详细的描述，本实施例在以本发明技术方案为前提下

进行实施，给出了详细的实施方式和具体的操作过程，但本发明的保护范围不限于下述的实施例，本领域技术人员在不改变本发明精神和内容的范围内，能够对其进行修改和润色。

本发明实施例构建场景空间特征向量获取网络，即视觉卷积神经网络，用于提取场景与目标抓取物的空间特征；构建逆强化回报值策略网络间接描述可能的驱动抓取策略；同时，在仿真环境内通过域随机化算法模拟抓取，使用仿真数据进行逆强化回报值策略网络的预训练，场景空间特征向量获取网络与逆强化回报值策略网络可以分别的进行预训练，将传统复杂视觉运动伺服问题解耦处理，降低了网络预训练的复杂度。其中，域随机化算法可以快速的生成大量的训练数据，减少了人工示教的操作数量，在有限的时间与资源内提升网络的训练效果。最后，通过真实场景与示教数据的融合，修正系统网络以使其适应真实的场景与任务。在网络训练完成后，通过引导性策略搜索算法给出规划结果。在最终的应用过程中，无需对场景内的目标抓取物及相关环境进行精确的空间标注，机械臂只需按照训练好的网络进行策略引导，完成抓取任务，训练好的网络对空间感知设备的要求更低，环境适应性强，并可迁移至多种任务。

实施例一

本发明实施例提供了农业场景无标定机器人运动视觉协同伺服控制设备，包括机械臂、目标抓取物、图像传感器和控制模块，请参考图1，

机械臂为 UR5 机械臂 6，UR5 机械臂 6 为六自由度机械臂，并在臂末端安装了机械抓手 7，机械抓手 7 可以通过夹紧、张开运动完成目标抓取物 3 的抓取，UR5 机械臂 6 通过一底座机座 5 固定在场景环境 8 中；

目标抓取物 3 优选为蔬果，如西红柿等，放置于一工位平台 4 上，工位平台 4 为稳定的具有一定高度的工作平面，如桌子等，工位平台 4 放置在场景环境 8 中，目标抓取物 3 处在 UR5 机械臂 6 的可抓取范围内；

图像传感器为 kinect 图像传感器 1，具体为 Kinect2.0 图像传感器，Kinect 图像传感器 1 固定在一 kinect 安装支架 2 上，Kinect 安装支架 2 为可以将 kinect 图像传感器 1 固定在一定高度的装置，优选使用铝型材搭建，kinect 安装支架 2 放置于 UR5 机械臂 6 与目标抓取物 3 的侧部，kinect 图像传感器 1 可以拍摄到 UR5 机械臂 6、目标抓取物 3 及场景环境 8；

控制模块为 Jetson TX1 控制板 9，Jetson TX1 控制板 9 分别与 UR5 机械臂 6 和 kinect 图像传感器 1 电连接，Jetson TX1 控制板 9 驱动 UR5 机械臂 6 通过机械抓手 7 抓取目标抓取物 3，并控制 kinect 图像传感器 1 对 UR5 机械臂 6 抓取目标抓取物 3 的过程进行图像采样，kinect 图像传感器 1 将采样的图像数据发送给 Jetson TX1 控制板 9。

具体地，请参考图 2，Kinect 图像传感器 1 通过 Kinect 适配转换器 10 将接口转换为 USB3.0 接口，Kinect 适配转换器 10 通过 USB3.0 与 Jetson TX1 控制板 9 进行连接；UR5 机械臂 6 通过连接机械臂控制箱 12 获取电源，机械臂控制箱 12 与 Jetson TX1 控制板 9 通过网线连接，Jetson TX1 控制板 9 通过网线接口向机械臂控制箱 12 输入机械臂控制信号。

优选地，Jetson TX1 控制板 9 通过 HDMI 接口连接一显示屏 11。

进一步地，请参考图 3，在 Jetson TX1 控制板 9 内安装 ubuntu 操作系统、驱动组件；通过安装 Jetpack 开发工具来为 Jetson TX1 控制板 9 安装其他软件；通过安装 Kinect 支持库来使得控制模块 Jetson TX1 控制板 9 可以驱动 Kinect 图像传感器，并使用相关的图像处理工具与算法；安装数据库，通过安装 python 依赖库与 MongoDB 数据库软件，完成 Jetson TX1 控制板 9 内嵌入式数据库的搭建，用于保存之后的训练用相关数据；安装 Docker 容器引擎以创建独立的软件运行环境，并将 ROS 操作系统与 Tensorflow 框架安装在 Docker 容器内，使得 Jetson TX1 控制板 9 内包含一个具备本实施例完整开发环境的容器引擎，并可以快速迁移至其他的硬件系统。

其中，ROS 操作系统内包含有 RGB-D（彩色深度图像）采样处理的算法节点与 UR5 机械臂的采样控制节点；Tensorflow 框架内包含有 GPS 引导策略算法控制程序，以及训练好的视觉空间特征提取与强化回报值策略网络。

实施例二

基于实施例一的农业场景无标定机器人运动视觉协同伺服控制设备，本发明提供了农业场景无标定机器人运动视觉协同伺服控制方法，请参考图 4，包括如下步骤：

S101：构建场景空间特征向量获取网络，获取场景空间特征特征向量。

本实施例中，场景空间特征向量获取网络为视觉卷积神经网络，请参考图 5，场景空间特征向量获取网络迁移使用了 CIFAR-1000 VGG16 的前五层作为图像特征提取网络，为卷积神经网络结构，具体地，

卷积神经网络的卷积层计算遵循下式：

$$x_j^l = f(\sum_{i \in M_j} x_i^{l-1} k_{ij}^l + b_j^l)$$

其中， x_j^l 表示第 l 层的第 j 个特征图， $\sum_{i \in M_j} x_i^{l-1} k_{ij}^l$ 表示的是针对 $l-1$ 层有所关联的特征图 x_i^{l-1} 和第 l 层的第 j 个卷积核 k_{ij}^l 做卷积运算并求和， b_j^l 是对第 l 层的第 j 个特征图补充的偏置参数， $f(\dots)$ 为激励函数，用于将式子右侧的卷积结果生成输出；

卷积神经网络的池化层计算遵循下式：

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}))$$

其中， x_j^l 表示第 l 层的第 j 个特征图， β_j^l 表示权重，在本实施例中的 2×2 池化过程中， β_j^l 的值取为 $1/4$ ， $\text{down}(\dots)$ 为下采样函数， $f(\dots)$ 为激励函数，用于将式子右侧的池化结果生成输出。

本实施例中，使用的场景空间特征向量获取网络构造如下表：

No	Name	参数
1	输入	240x240x3

2	conv1+pool1	卷积:7x7x64, 滑动步长 1, 池化:2x2, 滑动步长 1
3	conv2+pool2	卷积:5x5x32, 滑动步长 1, 池化:2x2, 滑动步长 1
4	conv3+pool3	卷积:5x5x64, 滑动步长 1, 池化:2x2, 滑动步长 1
5	Softmax	32
6	全连接 5	64
7	全连接 6	64
8	全连接 7	40

表 1

本实施例中, Jetson TX1 控制板 9 控制 kinect 图像传感器 1 拍摄抓取场景, 并提取 RGB 图像信息, 得到的图像数据为 240x240x3 的 3 通道 RGB 彩色图像, 以图像数据作为场景空间特征向量获取网络的输入量, 场景空间特征向量获取网络的最终输出 40 维的稀疏向量 F , 用于表示场景图像特征。

S102: 获取示教动作样本。

本实施例中, 获取获取示教动作样本包括如下步骤:

S1021: 牵引机械臂完成对目标抓取物的抓取, 获取一次示教抓取的示教抓取动作数据;

采用人为牵引 UR5 机械臂 6 完成 UR5 机械臂 6 的示教抓取路径, 使 UR5 机械臂 6 末端的机械抓手 7 到达可以直接抓取目标抓取物 3 的位置, 抓取过程中, Jetson TX1 控制板 9 以频率 f 对运动过程中的关节状态信息连续采样, 获得一次示教抓取的示教抓取动作数据;

本实施例中, UR5 机械臂 6 包括六个自由度关节, 每个自由度关节的状态信息记为 $S_{robot}(\theta_i, \omega_i, a_i, v_i, a_i', x_i)$, 包括: 关节转角 θ_i 、转速 ω_i 、关节角加速度 a_i 、关节节点中心的空间运动速度 v_i 、关节节点中心的空间运动加速度 a_i' , 相对初始位置的位移 x_i ; 其中, $S_{robot}^{direct}(\theta_i, \omega_i)$ 为可直接获取的关节状态信息, 包括: 关节转角 θ_i 和转速 ω_i , 关节初始零点 $\theta_i = 0, \omega_i = 0$; $S_{robot}^{indirect}(a_i, v_i, a_i', x_i)$ 为间接获取的关节状态数据, 根据 $S_{robot}^{direct}(\theta_i, \omega_i)$ 信息与采样步长 $T = 1/f$ 计算得到的关节状态, f 为采样频率。

抓取过程中, UR5 机械臂驱动节点程序以频率 f 对运动过程中的关节状态

信息进行多次连续采样。每次采样，UR5 机械臂驱动节点程序将采集可直接获取的关节状态信息 $S_{robot}^{direct}(\theta_i, \omega_i)$ ，并同步计算间接获取关节状态信息 $S_{robot}^{indirect}(a_i, v_i, a_i', x_i)$ ，并一次采样中的 $S_{robot}^{direct}(\theta_i, \omega_i)$ 与 $S_{robot}^{indirect}(a_i, v_i, a_i', x_i)$ 将合并为一次关节状态信息采样结果 $S_{robot}(\theta_i, \omega_i, a_i, v_i, a_i', x_i)$ 。

然后，将抓取过程中多次采样获得的每个关节状态信息采样结果 $S_{robot}(\theta_i, \omega_i, a_i, v_i, a_i', x_i)$ 按照采样时间先后顺序排布，形成一个连续关节状态信息数据序列。该序列即为一次示教抓取的示教抓取动作数据。

S1022：驱动机械臂模拟示教抓取动作数据，自主完成对目标抓取物的抓取动作，用以拍摄获取示教抓取场景图像特征数据；

在完成一次示教动作后，人员离开场景环境，基于示教抓取动作数据中包含的 UR5 机械臂 6 的六个自由度关节状态信息，Jetson TX1 控制板 9 驱动 UR5 机械臂 6 模拟示教过程完成一次模拟抓取目标抓取物 3 的动作，同时，Jetson TX1 控制板 9 以频率 f 驱动 kinect 图像传感器 1 对抓取过程进行图像采样，获得一次示教抓取场景图像特征数据。

S1023：基于示教抓取动作数据和示教抓取场景图像特征数据整合得到示教动作样本；

将示教抓取动作数据、示教抓取场景图像特征数据、机械臂与任务固有条件参数同步记录在 MongoDB 数据库，整合得到示教动作样本 $D_i(\{\gamma_i\}, g, d)$ ，其中， $\{\gamma_i\} = \{S_i, P_i\}$ ， $\{S_i\}$ 为六个自由度关节状态信息数据， $\{P_i\}$ 为图像特征数据序列， g 为目标抓取物状态信息（包括目标抓取物的大小、距离）， d 为机械臂动力学信息（包括机械臂模型构件的质量、机械臂模型初始的关节姿态）及控制参数。

S103：构建逆强化回报值策略网络。

本实施例中，构建逆强化回报值策略网络包括如下步骤：

S1031：构建用于拟合表示回报值的逆强化回报值策略网络；

本实施例中，逆强化回报值策略网络为 DNN 结构深度网络，该深度网络用于拟合表示引导策略中的回报值函数，从而避免建模人工选取特征参数。

本实施例中使用的逆强化回报值策略网络构造如下表：

NO	Name	参数
1	输入	40 维特征向量
2	全连接 1	50
3	全连接 2	30
4	全连接 3	12

表 2

然后，通过均匀随机产生逆强化回报值策略网络的初始权重参数初值 θ_0 。此时，可用深度网络表示一个未经学习训练优化的回报值函数。

S1032：通过仿真域随机化算法生成仿真参数；

首先，设置可行参数域 C ，标示域随机化算法参数的可能范围。参数域 C 内包括目标抓取物 3 相关参数的可行参数域 C_g ，UR5 机械臂 6 的相关动力学参数的可行参数域 C_d 。

具体地，在具有 GTX1080 显卡的训练机上安装 ubuntu 系统，并移植上述构建在 Jetson TX1 控制板 9 内的 Docker 容器；同时，在训练机中的 ROS 操作环境内导入 UR5 机械臂 6 的真实模型与目标抓取物 3 的抽象模型，并通过域随机化算法，随机生成 UR5 机械臂 6 的初始状态与目标抓取物 3 的大小空间位置，并决定仿真环境中的拍摄观测视角。

本实施例中，域随机化算法中所使用的参数如下表：

NO	Name	参数	参数域
1	机械臂模型构件的质量	$M_{Link}^i (i=1,2,...,6)$	$[M_{min}, M_{max}] kg$
2	机械臂模型初始的关节姿态	$\beta_i (i=1,2,...,6)$	$[\beta_{min}, \beta_{max}] rad$

3	初始的阻尼系数	$D_{Joint}^i (i=1,2,...,6)$	$[D_{min}, D_{max}]$
4	目标抓取物大小	$ObjectSize(l \times l \times l)$	$[L_{min}, L_{max}] m$
5	主视窗视角	Vpangle	$[\alpha_{min}, \alpha_{max}]$
6	目标抓取物的距离	Location	$[X_{min}, X_{max}] m$
7	控制器的增益系数	Gain	$[G_{min}, G_{max}]$
8	时间步长	T	$[T_{min}, T_{max}] s^{-1}$

表 3

然后，在参数域 C 内随机生成一组域随机化算法参数 $(M_{link}^i, \beta_i, D_{Joint}^i, ObjectSize, Vpangle, Location, Gain, T)$ ，令 $g(ObjectSize, Vpangle, Location)$ ， $d(M_{link}^i, D_{Joint}^i, Gain, T)$ ，以及状态 $S(\beta_i)$ 。

S1033：使用 ROS 规划库规划模拟虚拟抓取动作，并采样得到模拟抓取路径。

基于上述域随机化算法参数设定仿真环境内任务目标、初始状态与执行条件，通过 ROS 规划库规划模拟完成该仿真环境下的虚拟抓取动作，并对虚拟抓取动作路径进行虚拟采样，得到模拟抓取路径状态数据；同时，根据主视窗视角参数 Vpangle，调整仿真中的观测视角，并进行连续图像采样获取模拟抓取场景图像数据。

将模拟抓取路径状态数据、模拟抓取场景图像数据与域随机化算法参数结合，生成一次仿真规划动作样本数据 $Z_t(\{\gamma_t'\}, g', d')$ ，并将其保存在 MongoDB 数据库中。其中， $\{\gamma_t'\} = \{S_t', P_t'\}$ ， $\{S_t'\}$ 为六个自由度关节状态信息数据， $\{P_t'\}$ 为图像特征数据序列， g' 为目标抓取物状态信息（包括目标抓取物的大小、距离），

d' 为机械臂动力学信息（包括机械臂模型构件的质量、机械臂模型初始的关节姿态）及控制参数。

S1034: 逆强化回报值策略网络仿真预训练;

使用仿真规划动作样本数据 $Z_t(\{\gamma'_t\}, g', d')$ 对逆强化回报值策略网络进行预训练。

首先, 以随机生成的逆强化回报值策略网络的参数权重初始初值 θ 作为迭代初值, 即 $\theta^1 = \text{initial_weights}() = \theta_0$ 。

开始迭代循环, 循环特征量 n 从 1 开始执行至迭代上限 $n_{\max}$:

以第 n 次循环当前的网络权重参数 θ^n , 以空间图像特征 F 为输入量, 计算当前的回报值分布情况, 计算如下式:

$$\gamma^n = \text{nn_forward}(F, \theta^n)$$

然后, 根据当下回报值分布, 计算 MDP 最优策略 π^n :

$$\pi^n = \text{solve_mdp}(\gamma^n)$$

然后, 计算期望状态频率 $IE[\mu^n]$ 与专家示教损失项 L_D , D 表示以示教数据为专家动作;

$$IE[\mu^n] = \text{propagate_policy}(\pi^n)$$

$$L_D^n = \log(\pi^n) \times \mu_D^a$$

计算专家示教损失函数对回报值 r 的导数 $\frac{\partial L_D^n}{\partial \gamma^n}$ 与专家示教损失函数对网络模型参数的导数 $\frac{\partial L_D^n}{\partial \theta_D^n}$, 式中 μ_D 为专家状态动作频率:

$$\frac{\partial L_D^n}{\partial \gamma^n} = \mu_D - IE[\mu^n]$$

$$\frac{\partial L_D^n}{\partial \theta_D^n} = \text{nn_backprop}(\frac{\partial L_D^n}{\partial \gamma^n})$$

根据 $\frac{\partial L_D^n}{\partial \theta_D^n}$ 梯度修正网络模型参数, 完成一次优化迭代:

$$\theta^{n+1} = \text{update_weights}(\theta^n, \frac{\partial L_D^n}{\partial \theta_D^n})$$

算法迭代至最大迭代次数或专家示教损失项 L_D 小于可容忍限度，网络收敛得到 θ_{end} 。以此作为网络权重参数的回报值策略网络将在模拟环境中引导机械臂模型执行与 ROS 规划库规划期望策略相近的执行策略。

S104: 逆强化回报值策略网络迁移训练。

首先，以 S103 中预训练的逆强化回报值策略网络的权重参数 θ_{end} 为初始条件，使用 S102 中采样获得的示教动作样本 $D_i(\{\gamma_i\}, g, d)$ 替代 S103 的仿真规划动作样本数据 $Z_i(\{\gamma_i'\}, g', d')$ ，进行逆强化回报值策略网络训练，以进行网络训练修正，实现仿真环境模型到真实策略模型的模型迁移；

具体地，令 $\theta = \theta_{end}$ ，基于示教样本数据作为专家动作，以真实场景经视觉特征提取网络计算后得到的图像特征向量作为特征输入量，进行逆强化回报值策略网络的迁移训练优化，具体的算法执行步骤与 S1034 的优化过程相同，获得优化的网络权重 θ_{end}^* ；

然后，以 θ_{end}^* 作为网络权重参数得到的逆强化回报值策略网络，即为具备有人类策略认知的环境内策略引导回报网络，以此网络评估机械臂策略的回报，并可引导机械臂在农业复杂环境任务中做出与人类认知近似的决策。

S105: 基于视觉特征提取网络与逆强化回报值策略网络，采用策略引导驱动算法获得正向引导规划结果。

基于上述学习算法训练后的视觉特征提取网络与逆强化回报值策略网络，采用策略引导驱动算法（GPS）进行正向的引导规划。

具体的引导规划流程如下：

首先，生成若干差异性动态规划策略 $(DDP)\pi_{g_1}, \dots, \pi_{g_n}$ ；

然后，对于所述若干差异性动态规划策略进行采样得到策略路径数据

$\varsigma_1, \dots, \varsigma_m$, 多个策略路径数据可以通过 $q(\varsigma) = \frac{1}{n} \sum_i \pi_{g_i}(\varsigma)$ 计算组合平均策略以进行同时引导计算提高效率;

对以上策略的参数进行极大似然估计 $\theta^* \leftarrow \arg \max_{\theta} \sum_i \log \pi_{\theta^*}(\varsigma_i)$;

组合 $\pi_{g_1}, \dots, \pi_{g_n}$ 与 π_{θ^*} 得到初始样本集 S ;

基于场景特征提取后的向量状态参数, 使用逆强化回报值策略网络对样本集 S 进行评估, 评估 $\pi_{g_1}, \dots, \pi_{g_n}$ 与 π_{θ^*} 的回报值。

若 π_{g_i} 回报大于 π_{θ^*} 的回报, 则将 π_{g_i} 赋予 π_{θ^*} , 并对应的减少策略函数中的正则项参数;

若 π_{g_i} 回报小于 π_{θ^*} 的回报, 则增大策略函数中的正则项参数;

对 $\pi_{g_1}, \dots, \pi_{g_n}$ 重复执行上述评估, 最终得到一个引导后的最优策略, 即为正向引导规划结果。

本发明方法基于运动视觉双伺服驱动, 通过自适应的学习算法训练机器人获得智能的空间感知与任务规划能力。在最终的驱动过程中, 无需对于场景内的目标抓取物及相关环境进行精确的空间标注, 机械臂将按照训练好的网络进行策略引导完成抓取任务, 对于空间感知设备的要求更低, 环境适应性强, 并可迁移至多种任务。

以上公开的仅为本申请的两个具体实施例, 但本申请并非局限于此, 任何本领域的技术人员能思之的变化, 都应落在本申请的保护范围内。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程, 是可以通过计算机程序来指令相关的硬件来完成, 所述的程序可存储于一计算机可读取存储介质中, 该程序在执行时, 可包括如上述各方法的实施例的流程。其中, 所述的存储介质可为磁碟、光盘、只读存储记忆体 (Read-Only Memory, ROM) 或随机存储记忆体 (Random Access Memory, RAM) 等。

权 利 要 求 书

1、一种农业场景无标定机器人运动视觉协同伺服控制设备，其特征在于，包括机械臂、目标抓取物、图像传感器和控制模块，其中，

所述机械臂在臂末端安装有机械抓手；

所述目标抓取物处在所述机械臂的可抓取范围内；

所述控制模块分别与所述机械臂和图像传感器电连接，所述控制模块驱动所述机械抓手抓取所述目标抓取物，并控制所述图像传感器对所述机械臂抓取所述目标抓取物的过程进行图像采样；

所述图像传感器将采样的图像数据发送给所述控制模块。

2、根据权利要求1所述的农业场景无标定机器人运动视觉协同伺服控制设备，其特征在于，所述机械臂为六自由度机械臂。

3、一种农业场景无标定机器人运动视觉协同伺服控制方法，其特征在于，包括如下步骤：

构建场景空间特征向量获取网络，获取场景空间特征特征向量；

获取示教动作样本；

构建逆强化回报值策略网络；

逆强化回报值策略网络迁移训练；

基于视觉特征提取网络与逆强化回报值策略网络，采用策略引导驱动算法获得正向引导规划结果。

4、根据权利要求3所述的农业场景无标定机器人运动视觉协同伺服控制方法，其特征在于，所述场景空间特征向量获取网络为视觉卷积神经网络。

5、根据权利要求3所述的农业场景无标定机器人运动视觉协同伺服控制方法，其特征在于，所述获取场景空间特征特征向量具体为：

图像传感器对机械臂抓取目标抓取物的过程进行图像采样，并提取 RGB 图像信息；

以所述图像信息作为所述场景空间特征向量获取网络的输入量，输出向

量即为场景空间特征特征向量。

6、根据权利要求3所述的农业场景无标定机器人运动视觉协同伺服控制方法，其特征在于，所述获取示教动作样本具体为：

牵引机械臂完成对目标抓取物的抓取，获取一次示教抓取的示教抓取动作数据；

驱动机械臂模拟示教抓取动作数据，自主完成对目标抓取物的抓取动作，用以拍摄获取示教抓取场景图像特征数据；

基于所述示教抓取动作数据和示教抓取场景图像特征数据整合得到示教动作样本。

7、根据权利要求3所述的农业场景无标定机器人运动视觉协同伺服控制方法，其特征在于，所述构建逆强化回报值策略网络具体为：

构建用于拟合表示回报值的逆强化回报值策略网络；

通过仿真域随机化算法生成仿真参数；

使用ROS规划库规划模拟虚拟抓取动作，并采样得到模拟抓取路径；

逆强化回报值策略网络仿真预训练。

8、根据权利要求3所述的农业场景无标定机器人运动视觉协同伺服控制方法，其特征在于，所述逆强化回报值策略网络迁移训练具体为：

使用所述示教动作样本进行所述逆强化回报值网络的优化训练。

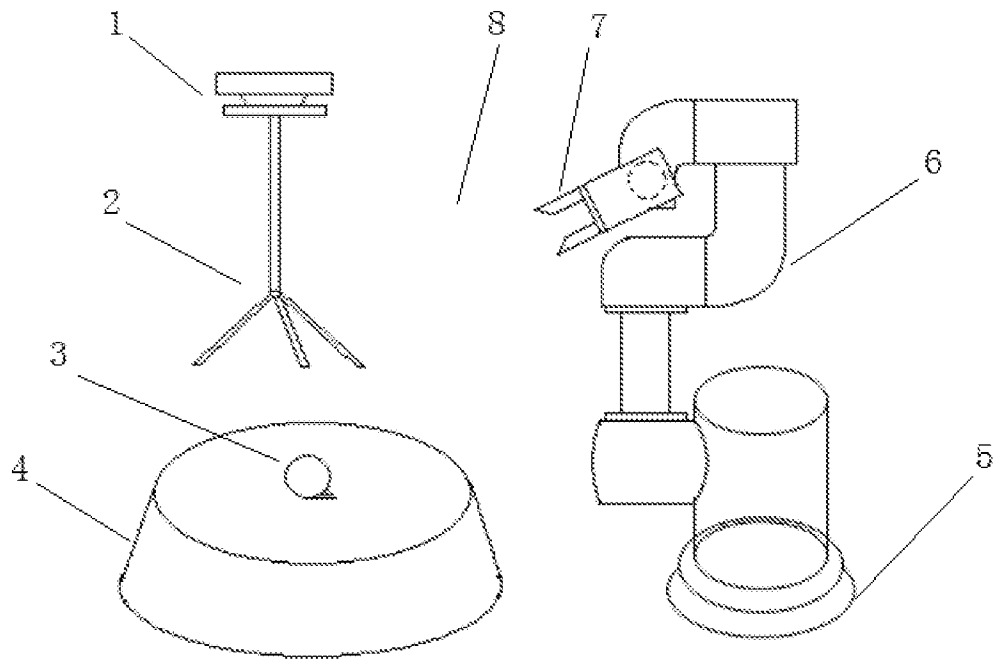


图 1

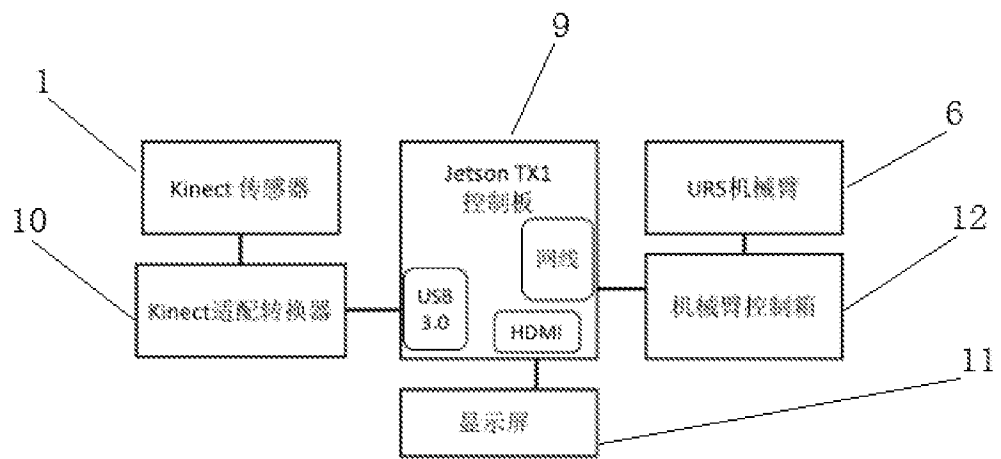


图 2

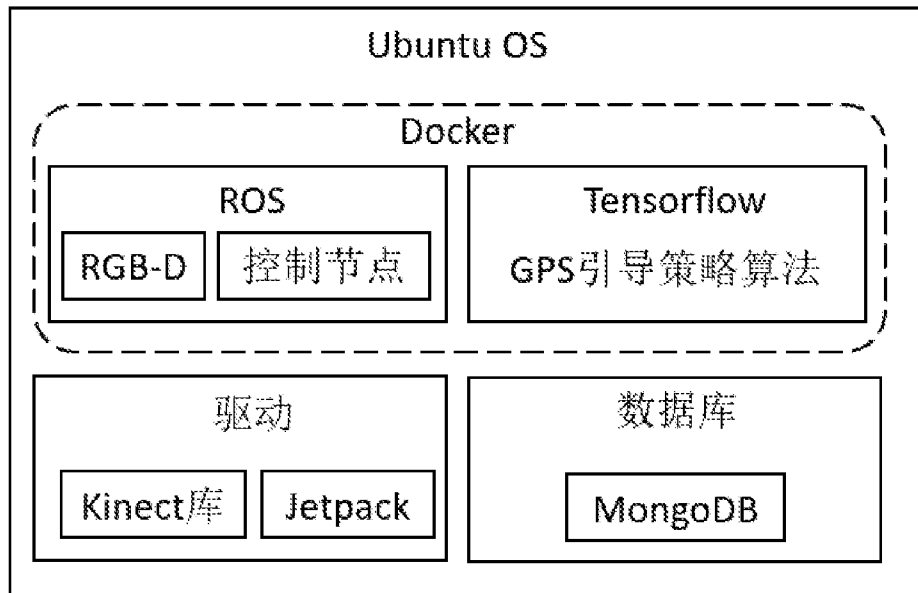


图 3

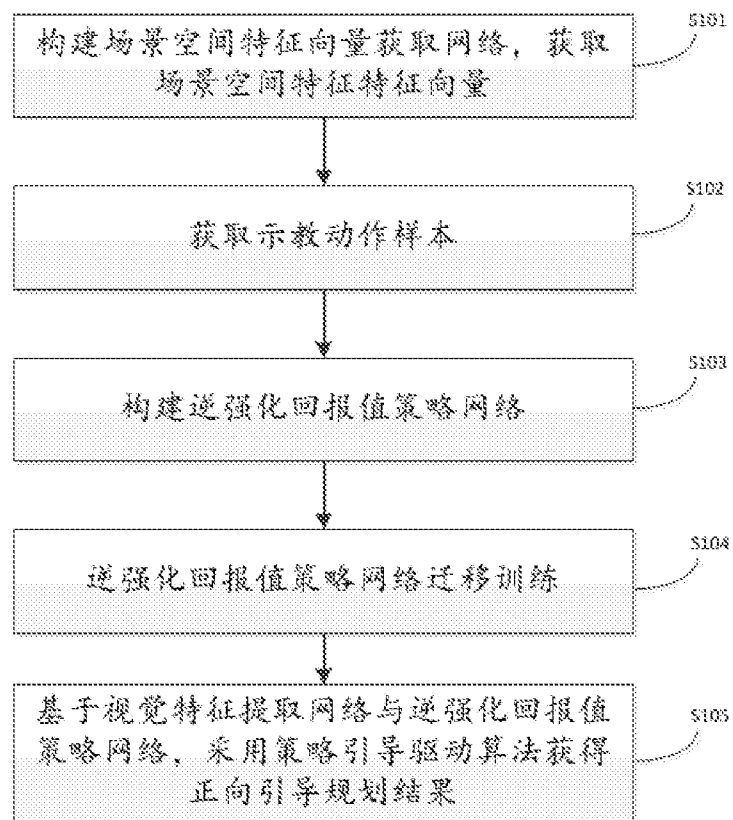


图 4

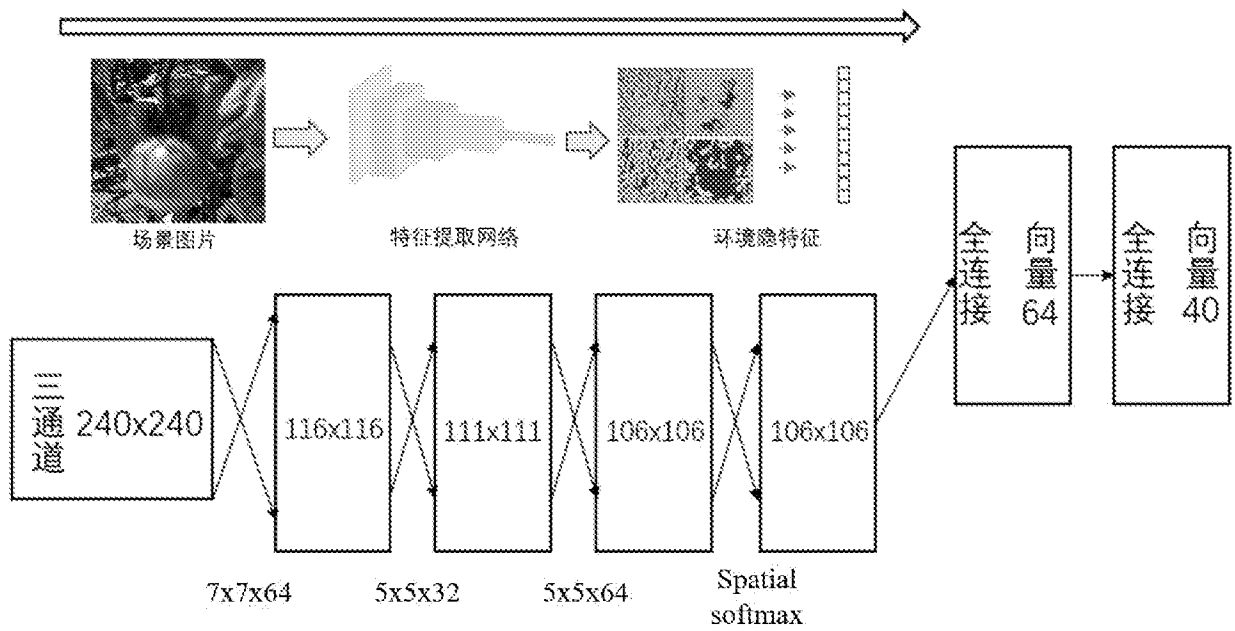


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/119079

A. CLASSIFICATION OF SUBJECT MATTER

B25J 9/16(2006.01)i; B25J 13/08(2006.01)i; B25J 15/02(2006.01)i; B25J 15/08(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

B25J

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, VEN, CNTXT, CNKI: 农业, 机器人, 机械手, 机械臂, 抓, 视觉, 图像, 传感器, 模拟, 仿真, 网络, 神经, agricultur+, robot+, mechanism+, grab+, vision+, imag+, sensor+, simulat+, emulation+, network+, nerve+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110000785 A (SHANGHAI JIAO TONG UNIVERSITY) 12 July 2019 (2019-07-12) claims 1-8	1-8
X	CN 109531584 A (BEIJING INST RADIO MEASUREMENT) 29 March 2019 (2019-03-29) description, paragraphs [0032]-[0065]	1-8
A	CN 107481247 A (XU, Wenyuan) 15 December 2017 (2017-12-15) entire document	1-8
A	CN 108656107 A (BEIHANG UNIVERSITY) 16 October 2018 (2018-10-16) entire document	1-8
A	CN 108227689 A (HARBIN PAITENG AGRICULTURAL TECHNOLOGY CO., LTD.) 29 June 2018 (2018-06-29) entire document	1-8
A	US 2018147723 A1 (GOOGLE LLC) 31 May 2018 (2018-05-31) entire document	1-8
A	US 2018243914 A1 (INTOUCH TECH INC) 30 August 2018 (2018-08-30) entire document	1-8



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

15 January 2020

Date of mailing of the international search report

03 February 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/119079

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110000785	A	12 July 2019	None			
CN	109531584	A	29 March 2019	None			
CN	107481247	A	15 December 2017	None			
CN	108656107	A	16 October 2018	None			
CN	108227689	A	29 June 2018	None			
US	2018147723	A1	31 May 2018	EP	3414710	A1	19 December 2018
				US	2017252924	A1	07 September 2017
				KR	20180114200	A	17 October 2018
				CN	108885715	A	23 November 2018
				KR	102023149	B1	22 November 2019
				JP	2019509905	A	11 April 2019
				JP	6586243	B2	02 October 2019
				WO	2017151926	A1	08 September 2017
				US	9914213	B2	13 March 2018
US	2018243914	A1	30 August 2018	US	9375843	B2	28 June 2016
				US	2011245973	A1	06 October 2011
				US	7813836	B2	12 October 2010
				US	2013131866	A1	23 May 2013
				US	2016303740	A1	20 October 2016
				US	2011071702	A1	24 March 2011
				US	2005125098	A1	09 June 2005
				US	9296107	B2	29 March 2016
				US	9956690	B2	01 May 2018

国际检索报告

国际申请号

PCT/CN2019/119079

A. 主题的分类

B25J 9/16(2006.01)i; B25J 13/08(2006.01)i; B25J 15/02(2006.01)i; B25J 15/08(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

B25J

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, VEN, CNTXT, CNKI: 农业, 机器人, 机械手, 机械臂, 抓, 视觉, 图像, 传感器, 模拟, 仿真, 网络, 神经, agricultur+, robot+, mechanism+, grab+, vision+, imag+, sensor+, simulat+, emulation+, network+, nerve+

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 110000785 A (上海交通大学) 2019年 7月 12日 (2019 - 07 - 12) 权利要求1-8	1-8
X	CN 109531584 A (北京无线电测量研究所) 2019年 3月 29日 (2019 - 03 - 29) 说明书第[0032]-[0065]段	1-8
A	CN 107481247 A (许文远) 2017年 12月 15日 (2017 - 12 - 15) 全文	1-8
A	CN 108656107 A (北京航空航天大学) 2018年 10月 16日 (2018 - 10 - 16) 全文	1-8
A	CN 108227689 A (哈尔滨派腾农业科技有限公司) 2018年 6月 29日 (2018 - 06 - 29) 全文	1-8
A	US 2018147723 A1 (GOOGLE LLC) 2018年 5月 31日 (2018 - 05 - 31) 全文	1-8
A	US 2018243914 A1 (INTOUCH TECH INC) 2018年 8月 30日 (2018 - 08 - 30) 全文	1-8

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2020年 1月 15日

国际检索报告邮寄日期

2020年 2月 3日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

余娟娟

电话号码 86-(010)-62085091

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/119079

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	110000785	A	2019年 7月 12日	无			
CN	109531584	A	2019年 3月 29日	无			
CN	107481247	A	2017年 12月 15日	无			
CN	108656107	A	2018年 10月 16日	无			
CN	108227689	A	2018年 6月 29日	无			
US	2018147723	A1	2018年 5月 31日	EP	3414710	A1	2018年 12月 19日
				US	2017252924	A1	2017年 9月 7日
				KR	20180114200	A	2018年 10月 17日
				CN	108885715	A	2018年 11月 23日
				KR	102023149	B1	2019年 11月 22日
				JP	2019509905	A	2019年 4月 11日
				JP	6586243	B2	2019年 10月 2日
				WO	2017151926	A1	2017年 9月 8日
				US	9914213	B2	2018年 3月 13日
US	2018243914	A1	2018年 8月 30日	US	9375843	B2	2016年 6月 28日
				US	2011245973	A1	2011年 10月 6日
				US	7813836	B2	2010年 10月 12日
				US	2013131866	A1	2013年 5月 23日
				US	2016303740	A1	2016年 10月 20日
				US	2011071702	A1	2011年 3月 24日
				US	2005125098	A1	2005年 6月 9日
				US	9296107	B2	2016年 3月 29日
				US	9956690	B2	2018年 5月 1日