

复杂场景下机械臂多动作协同抓取策略

李德平¹, 洪楷宣¹, 柳宁¹, 王高²

1. 暨南大学智能科学与工程学院, 广东 珠海 519070;

2. 暨南大学信息科学技术学院, 广东 广州 510632

基金项目: 国家自然科学基金项目(62172188, 61775172, 62276114); 韶关市科技计划项目(230615178031046); 珠海市科技计划项目(2220004002542, 2220004002325, ZH22017001210107PWC)

通信作者: 王高, twangg@jnu.edu.cn 收稿/录用/修回: 2024-07-16/2024-09-03/2024-09-09

摘要

在复杂和不确定的现实场景中, 被抓取物体的无序、遮挡和自遮挡关系阻碍了机器人完成场景感知和精确抓取。研究人员提出了主动视觉框架来增强场景感知能力, 通过协同视点调整和抓取来完成抓取任务, 旨在紧密堆叠的场景中更好地获得场景信息并分离对象。然而, 目前大多数主动视觉与抓取动作都是自上而下的, 这限制了机器人对场景的感知能力。本文提出了一种6自由度(DoF)姿态空间中主动视觉感知和抓取协同策略。该策略基于深度强化学习构建视点调整网络(VNet)、4DoF抓取网络(4GNet)和6DoF抓取网络(6GNet)来学习最优协作策略, 并根据Q函数和约束选择合适原始动作。为了将多个视点所捕获的信息融合, 还提出了一种视点调整后场景融合方法, 通过将多个视点捕获的场景信息集成到固定大小的高度图中来提高机器人的感知面积和质量。实验结果表明, 与自上而下的视点调整动作相比, 单次视点调整的情况下, 该方法获取到的场景面积增大了8.93%, 为抓取提供了更全面的场景信息。在包含10个目标对象的杂乱场景中, 抓取成功率达到了89.53%, 与先进算法VPG相比, 本文提出的方法在抓取成功率上提升了12.02%。

关键词

深度强化学习

机械臂抓取

主动视觉感知

视点融合

中图法分类号: TP18

文献标志码: A

Multi-action Cooperative Grasping Strategy for Robotic Arms in Complex Scenes

LI Deping¹, HONG Kaixuan¹, LIU Ning¹, WANG Gao²

1. College of Intelligent Systems Science and Engineering, Jinan University, Zhuhai 519070, China;

2. College of Information Science and Technology, Jinan University, Guangzhou 510632, China

Abstract

In complex, uncertain real-world environments, challenges such as disorder, occlusion, and self-occlusion of grasped objects hinder robots from effectively perceiving scenes and executing precise grasps. To tackle these issues, researchers have proposed an active visual framework to enhance scene perception strategies. By coordinating viewpoint adjustments with grasping tasks, this approach aims to improve scene information acquisition and object separation. However, most current methods employ top-down actions, limiting the robot's scene perception capabilities. This

Keywords

deep reinforcement learning;

robotic manipulator grasping;

active visual perception;

viewpoint fusion

study introduces an active visual perception and grasping coordination strategy in a 6-degree-of-freedom (6DoF) pose space. Utilizing deep reinforcement learning, we develop a viewpoint adjustment network, a 4-degree-of-freedom (4DoF) grasping network, and a 6DoF grasping network to learn optimal collaborative strategies. Actions are determined using Q-functions and constraints to execute suitable primitive actions. To enhance scene perception, we propose a scene fusion method following viewpoint adjustment, which integrates information from multiple viewpoints into fixed-size height maps. Experiment results demonstrate an 8.93% increase in captured scene area compared to top-down methods in single viewpoint scenarios, providing comprehensive information for grasping tasks. In cluttered scenes containing ten target objects, the grasping success rate reaches 89.53%. Compared to the state-of-the-art VPG algorithm, our proposed method achieves a 12.02% increase in grasping success rate.

0 引言

在复杂和不确定的现实场景中,被抓取物体的分布的无序性、物体之间的遮挡和自遮挡等因素阻碍了机器人完成场景感知和精确的抓取动作^[1]。为了使机器人能针对这些复杂场景进行鲁棒抓取,研究人员采用了强化学习和深度学习技术来提高抓取精度和鲁棒性^[2-4],并取得了显著的效果,如 DexNet 以及其后续改进网络^[5-7],使用超过 10 000 个 3 维模型和 250 万个抓取位姿数据集中训练端到端抓取模型,模型可根据处理后的深度图端到端的生成抓取位姿,在独立物体场景中取得了极佳的性能。KALASHNIKOV 等^[8]使用深度强化学习方法学习动态闭环控制策略,使用在 7 台真实机器人上采集的 580 000 个抓取位姿数据进行训练,通过最大化累积期望奖励来实现机器人抓取位姿选取。最终抓取精度达到了 96.00%, MAHLER 等^[7]提出的 DexNet4.0 在经过 500 万个合成深度图像、抓取位姿和从大量 3 维物体后可以达到每小时 300 次以上的平均抓取速率,并能保证大于 95.00% 的抓取成功率。

然而,这些方法都需要大量的数据和时间来帮助模型训练^[8-9]。随着场景复杂程度的不断增加,仅依靠单一抓取动作渐渐展现出了局限性^[10]。为了进一步提升抓取的性能,更好地面对被抓取物体紧密分布的场景,一些在抓取过程中引入其他原始动作协助抓取的方法被提出。如 ZENG 等^[11]所提出的 VPG (Visual Pushing for Grasping) 通过引入推动作来解决紧密堆叠现象,将堆叠在一起的物体用推动作分离,并将单一固定视点获取的场景信息输入 DQN (Deep Q-Learning) 学习框架,计算 q 值选择最优动作执行。YANG 等^[12]在 ZENG 等的基础上

对策略进行改进,在深度强化学习框架中引入注意力机制,并通过推动作前后局部 q 值的改变来评估推动作的好坏,以抑制不合理的推动作。李鑫等^[13]基于 DDQN (Double Deep Q-Network) 学习得到更为优秀的抓推策略,并通过掩码函数促进模型进行高效率学习。但是推动作可能会将物体推出固定相机的视野,导致很难完成场景清空任务。为解决这些问题,MOHAMMED 等^[14]通过增加固定相机的数量来解决固定相机的视野不足的问题。但会造成很多设备资源的浪费。近些年一些基于主动视觉框架的多动作协同抓取方法被提出。如 MORRISON 等^[15]提出的 MVP (Multi-View Picking),该策略通过主动感知方法,可以直接根据实时获取的姿势估计分布来选择信息视点,帮助机械臂更好地获取视觉信息。LIU 等^[16]基于深度强化学习,提出 VAGS (Viewpoint Adjusting and Grasping Synergy),将视点调整与抓取直接协调起来,通过最大 q 值决定动作执行位置,使机器人具有独立调整摄像机视点的功能,并提出了 DAES (Dynamic Action Exploration Space) 探索方法,使机器人能在 6 000 次训练内达到收敛,大幅降低了训练时间。

然而,大多数主动视觉和抓取都是自上而下的,且大多数基于主动视觉框架的抓取判断局限于当前视点,使模型对抓取位姿的判断容易陷入局部最优,这极大地限制了机器人获取场景信息和抓取的能力,并且浪费了之前视点所获取的场景信息。为此,针对复杂场景中的抓取问题,本文提出了一种 6DoF 姿态空间中主动视觉感知和抓取的协同策略。本文的主要贡献有:

- 1) 为了提高机器人的抓取能力,采用 DQN 深度强化学习方法来联合训练视点调整网络 (VNet)、4DoF 抓取网络 (4GNet) 和 6DoF 抓取网络 (6GNet)

以达到最大化奖励函数。

2) 提出了一种视点调整后场景融合方法整合不同视点捕获的信息融合并保存成固定大小的高度图, 用于保证视点调整前后网络输入的一致性。

3) 提出了一种视点调整动作映射方法, 用于将视点调整动作从自上而下映射到 6DoF 空间。并通过实验证明能提高机器人的感知面积和质量。

1 算法原理

1.1 策略整体架构

如图 1 所示, 本文提出的抓取视点调整协同策略由 3 个主要的模块组成: 数据预处理模块、动作预测模块、动作执行模块。系统所需的场景信息由安装在机械臂上的 RGB-D 相机以预设的姿势捕获, 随后将 RGB-D 图像传递进数据预处理模块。数据

预处理模块将接受到的图像转换为点云和高度图并送入动作预测模块。动作预测模块将不同的信息送入不同的网络以获得不同动作原语的位姿和评分。其中, 点云信息被传递进 6GNet 模块以预测 6DoF 的抓取位姿和该位姿的评分 (6DoF Grasp Value); 高度图信息被送入 VNet 和 4GNet 用于生成像素级 q 值映射并预测相应的动作执行位置。随后, 动作执行模块根据 6GNet 预测得到的 6DoF Grasp Value 来判断执行动作, 如果判断执行抓取动作, 则机械臂会按照动作预测模块输出的抓取位姿进行抓取, 并在抓取成功后结束该轮任务。如果执行视点调整动作, 则机械臂会在调整到新视点后将在新视点捕获的 RGB-D 送入数据预处理模块, 进行两视点场景信息融合。以此循环, 直至完成抓取动作。

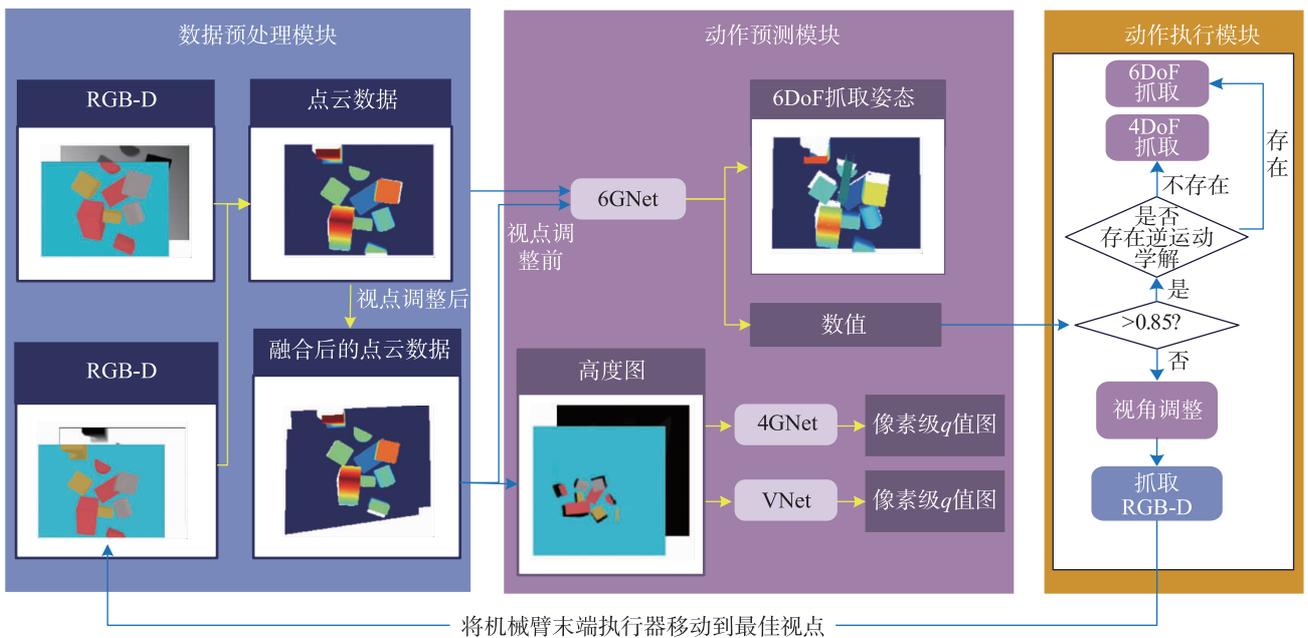


图 1 6DoF 姿态空间中主动感知和抓取的协同策略

Fig.1 Collaborative strategy of active perception and grasping in 6DoF pose space

1.2 MDP

继多动作协同抓取策略 (Visual Pushing for Grasping, VPG) 之后, 将这种顺序决策过程表述为马尔可夫决策过程 (Markov Decision Process, MDP)。定义 MDP 为一个元组 $M = (S, A, T, R, Y)$, 其中 S 表示状态空间, 存在 $s_t \in S$, 定义为在 t 时刻通过 RGB-D 相机采集到的信息, 动作 $a \in A$ 表示动作原语的选择、末端执行器的位置和方向。

$$a = (\varphi, x, \theta), x, \theta \in \mathbb{R}^3 \quad (1)$$

其中, φ 被定义为机械臂选择的行为原语, x 、 θ 分别被定义为末端执行器的位置和方向, T 、 R 和 Y

分别表示转移函数、奖励函数和折扣因子。

系统使用 DQN 算法来学习视点调整和抓取的不确定性策略。将作为输入, 依次经过数据预处理模块、动作预测模块、动作执行模块采取行为 a_t 并返回奖励 r_t 。在执行抓取动作时, 抓取 (包括 4DoF 抓取和 6DoF 抓取) 奖励 R_g 被设计为二进制稀疏奖励函数。定义 t 时刻抓取奖励 $r_{gt} \in R_g$, 其值如式 (2) 决定:

$$r_{gt} = \begin{cases} 1, & \text{GS} \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中, GS 表示成功抓取。在执行视点调整动作时,

视点调整奖励同样被设计为二进制稀疏奖励函数。定义在 t 时刻视点调整奖励, 其值如式(3)决定:

$$r_{vt} = \begin{cases} \alpha + Q_{4g}^{\text{improve}} + Q_{6g}^{\text{improve}}, & \text{GS} \\ Q_{4g}^{\text{improve}} + Q_{6g}^{\text{improve}}, & \text{其他} \end{cases} \quad (3)$$

其中, Q_{4g}^{improve} 是视点调整前后 4GNet 输出的 q 值映射图中的最大值 q 的差值, Q_{6g}^{improve} 是视点调整前后的 6DoF 抓取值的差值, 如果 Q_{4g}^{improve} 和 Q_{6g}^{improve} 大于 0, 则认为视点调整为一次抓取提供了更好的抓取条件; 若 Q_{4g}^{improve} 和 Q_{6g}^{improve} 小于 0, 则可认为这次视点调整对下一次抓取没有起到帮助。如果在视点调整后的抓取动作成功抓取目标物体, 认为该视点调整动作对下一次抓取有所帮助, 故定义 α 表示视点调整对抓取的影响系数, 参照 LIU 等^[16] 的研究确定其值为 0.5, 用于加入抓取是否成功对 r_{vt} 的影响。综上所述, 奖励的定义如式(4)所示:

$$r = \begin{cases} r_{gt}, & \varphi = 6\text{DoF 抓取} \parallel 4\text{DoF 抓取} \\ r_{vt}, & \varphi = \text{视点调整} \end{cases} \quad (4)$$

2 数据预处理模块

数据预处理模块主要完成 RGB-D 数据转换成点云数据、视点调整前后点云数据融合、点云数据转换成高度图三个任务, 旨在在不同情况、不同场景下, 都能为动作预测模块提供格式相同的输入数据。

视点调整后场景融合方法: 为了有效统筹数据预处理模块的任务, 并充分利用多个视点获取的信息以确保视点调整前后网络输入的一致性, 本文提出了一种视点调整后场景融合方法, 旨在使大小相同的高度图能够存储从多个视点捕获的场景信息。在本文中, VNet 和 4GNet 所生成的像素级 q 值映射图和高度图的分辨率皆为 224×224 像素, 每个像素所代表的物理尺寸为 2 mm。如图 2 所示, 该方法由 4 个部分组成:

首先, 将视点调整前后的 RGB-D 图像转换为点云数据(见图 2(a)), 包括点云位置数据和点云颜色数据, 并将其转换到机器人的坐标系中。其次, 由于分辨率为 224×224 的图像只能保存一个视点的场景信息, 无法充分利用多个视点的信息。因此, 本文通过创建一个分辨率为 448×448 的画布来存储多个视点捕获的场景信息。将转换后的彩色点云数据转换成高度图并保存在该画布中, 然后进行灰度化处理, 如图 2(b)所示。接着, 对高度图进行二值化处理、轮廓检测、膨胀, 并填充最大轮

廓, 以获得蒙版, 如图 2(c)所示。最后, 对非蒙版部分及蒙版部分的非工作空间区域进行背景颜色填充。图 2(a)中的灰色点云部分表示非工作空间区域。完成上述步骤后, 对高度图进行一次降采样操作, 使其分辨率达到 224×224 , 从而获得新的彩色高度图与深度高度图。

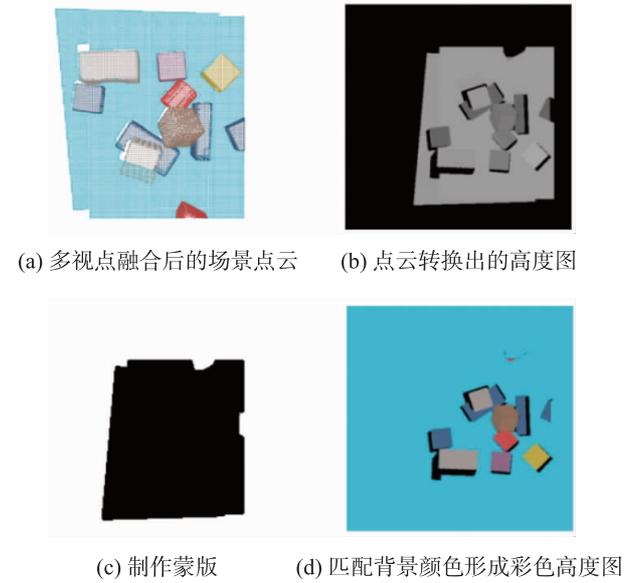


图 2 用于视点调整的场景融合方法示意图

Fig.2 Schematic diagram of scene fusion method for viewpoint adjustment

通过上述方式, 数据预处理模块可以有效融合不同视点的信息。确保在视点调整前后为动作预测模块提供点云输入和相同形状的高度图输入。

3 动作预测模块

动作预测模块的目标是利用深度强化学习端到端生成动作原语的位姿和评分, 它被划分成 3 个部分, 用于生成抓取动作和评分(6DoF Grasp Value)的 6GNet。用于生成像素级 q 值映射并预测位置和方向的 4GNet 和 VNet。本节分别介绍动作预测模块的 3 个组成部分与其输出。

1) 6GNet: 6GNet 的实现主要分成 3 个部分, 预处理、抓取位姿生成和评分。6GNet 以点云作为输入, 定义点云 $\mathbf{p} \in \mathbb{R}^{3 \times n}$, 代表该点在机器人坐标系中位置 (p_x, p_y, p_z) , 以 p_{z_value} 作为阈值将点云集中 $p_z < p_{z_value}$ 点云去除以去除代表桌面的点云, 仅仅保留下表示被抓取物体的点云, p_z 的阈值 p_{z_value} 设置公式如下:

$$p_{z_value} = \frac{\text{mean}(p_z)}{n} \quad (5)$$

其中, n 为超参数, 由实验得出当 $n=5$ 时可以将代表桌面的点云滤除并对 \mathbf{p} 进行均匀采样 n_1 个点, 其中 $n_1=0.2n$, 本文定义采样后的点云集为 C 。随后使用李会军等^[17]所使用的抓取候选生成方法生成抓取位姿。该 6 自由度抓取动作可以表示为 $\{x, y, z, r_x, r_y, r_z\}$, x, y, z 为相对与原点沿对应轴的移动距离, r_x, r_y, r_z 为绕对应轴旋转的欧拉角。对于每一个采样点 $p_i \in C$, 在其八叉树空间所划分的半径中使用使用近邻搜索算法得到 p_i 的 k 个近邻点构成局部点云 $C_r(p_i)$, 其中 r 为八叉树空间中的划分半径。根据 $C_r(p_i)$ 构建协方差矩阵 $\mathbf{M}_c(p_i)$, 公式如式(6)所示:

$$\mathbf{M}_c(p_i) = \frac{1}{k} \sum_{b \in C_r(p_i)} (\mathbf{b} - \mathbf{p}_c)(\mathbf{b} - \mathbf{p}_c)^T \quad (6)$$

其中 \mathbf{p}_c 是局部点云的质心点, 又由式(7)计算得到:

$$\left\{ \begin{array}{l} \mathbf{p}_c \{ \mathbf{p}_x, \mathbf{p}_y, \mathbf{p}_z \} \in C_r(p_i) \\ p_x = \frac{\sum_{i=1}^n x}{n} \\ p_y = \frac{\sum_{i=1}^n y}{n} \\ p_z = \frac{\sum_{i=1}^n z}{n} \end{array} \right. \quad (7)$$

求解协方差矩阵 $\mathbf{M}_c(p_i)$ 的特征值与其对应的特征向量分别命名为 $\lambda_1, \lambda_2, \lambda_3$ 和 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, 这里规定 $\lambda_1 < \lambda_2 < \lambda_3$ 。根据主成分分析理论, $\mathbf{v}_2, \mathbf{v}_3$ 分别为 $C_r(p_i)$ 在 p_i 处的法线和主向量。其中法线方向为末端执行器执行器的 X 轴的正方向。随后, 根据法线和主向量可反推末端执行器执行器的 Y 轴和 Z 轴。由于 $C_r(p_i)$ 在 p_i 处的主向量和抓取位姿正交时抓取较为稳定^[18], 故 $C_r(p_i)$ 在 p_i 处的主向量和 X 轴向量之间的外积为末端执行器执行器的 Y 轴, 最后通够右手法则确定 Z 轴。末端执行器各轴正方向的单位向量 $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$, 以及末端执行器执行器原点 O 可通过式(8)表示, 其中偏移量 d 为末端执行器长度 l 的关系为 $d = \frac{1}{2}$ 。

$$\left\{ \begin{array}{l} \mathbf{e}_x = \frac{\mathbf{v}_1}{|\mathbf{v}_1|} \\ \mathbf{e}_y = \frac{\mathbf{v}_1 \times \mathbf{v}_3}{|\mathbf{v}_1 \times \mathbf{v}_3|} \\ \mathbf{e}_z = \frac{\mathbf{e}_x \times \mathbf{e}_y}{|\mathbf{e}_x \times \mathbf{e}_y|} \\ \mathbf{O} = \mathbf{p}_i - d \cdot \mathbf{e}_x \end{array} \right. \quad (8)$$

随后, 根据抓取位姿模拟该位置夹持器的闭合动作, 判断夹持器与环境是否发生碰撞, 进一步筛选出抓取位姿。最后, 将剩余的抓取位姿输入到预训练好的 PointNet^[19] 网络中进行评分, 输出抓取评分(6DoF 抓取评分), 并选择抓取评分最高的抓取位姿作为 GNet 的输出。GNet 生成的抓取位姿和 6DoF 抓取评分示例如图 3 所示。其中, 红色的点云区域代表生成抓取位姿的可视化。PointNet 在整个训练过程不做参数更新。通过这些步骤, 为每个抓取姿势获得分数, 从而能够选择最佳抓取解决方案, 本轮最优 6DoF 抓取动作。

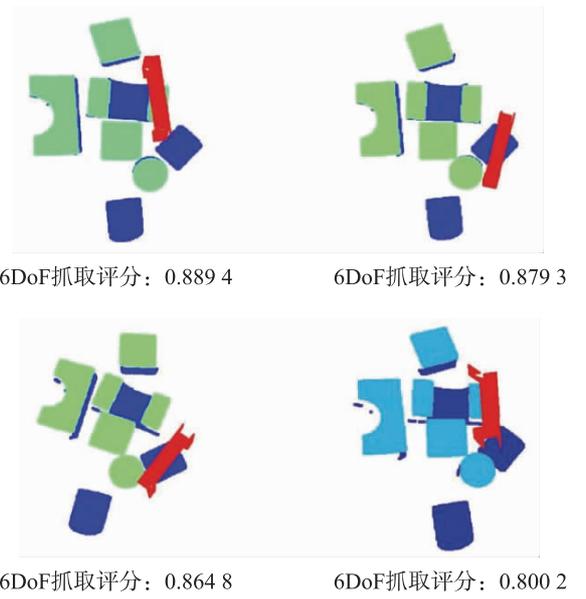


图3 GNet 生成的抓取位姿和 6DoF 抓取评分

Fig.3 The grasp pose and 6DoF grasp value generated by GNet

2) 4GNet、VNet: 4GNet、VNet 具有相同的网络结构, 它们都以彩色高度图和深度高度图作为输入, 使用并行的 MobileNetV2^[20] 提取特征。选用 MobileNetV2 作为特征提取的主干网络有两个主要的原因。第一, MobileNetV2 通过 Expansion layer 和 Projection layer 实现倒残差结构, 提高了模型获取特征的能力。第二, 在 Expansion layer 对数据进行升维之后使用深度可分离卷积来提取特征, 以减小倒残差结构的计算量, 以做到特征量和计算量的平衡。这些特征经过层归一化 (Batch Normalization, BN)^[21] 处理, 以稳定网络训练过程中的数值分布。并利用整流线性单元 (Rectified Linear Unit, ReLU)^[22] 激活函数增强非线性处理能力。此外, 采用 1×1 卷积来调整通道数, 进一步细化了特征表示, 使用双线性上采样将数值映射高度图相同分辨率的

矩阵中,生成像素级 q 值图。唯一不同的是,在 4GNet 中,对输入的彩色高度图和深度高度图先进行 16 次旋转,每次旋转 22.5° 。因此 4GNet 的输出为 16 个像素级 q 值图,用来映射抓取的角度和位置。而 VNet 的输出为 1 个像素级 q 值图用来映射视点调整的位置。随后,根据 q 值图中最大 q 值点映射得到本轮最佳 4DoF 抓取动作位置和最佳动作位置。VNet L_{vi} 和 4GNet L_{4gi} 的损失函数如下:

$$L_{4gi} = \begin{cases} \frac{1}{2}(Q_{4g}^{\theta_i} - y_i^{\bar{\theta}}), & |(Q_{4g}^{\theta_i} - y_i^{\bar{\theta}})| < 1 \\ |Q_{4g}^{\theta_i} - y_i^{\bar{\theta}}| - \frac{1}{2}, & \text{其他} \end{cases} \quad (9)$$

$$L_{vi} = \begin{cases} \frac{1}{2}(Q_v^{\theta_i} - y_i^{\bar{\theta}}), & |(Q_v^{\theta_i} - y_i^{\bar{\theta}})| < 1 \\ |Q_v^{\theta_i} - y_i^{\bar{\theta}}| - \frac{1}{2}, & \text{其他} \end{cases} \quad (10)$$

其中, θ_i 表示当前网络的参数, $Q_{4g}^{\theta_i}, Q_v^{\theta_i}$ 分别表示网络参数为 θ_i 时 4GNet 和 VNet 的输出, $\bar{\theta}$ 是目标网络参数, $y_i^{\bar{\theta}}$ 是目标网络的输出。

图 6 展示了 4GNet 与 VNet 训练与参数更新流程。机器人通过 RGB-D 相机获取到环境 s_t , 并将其作为 DQN 的输入, 得到 6DoF 抓取评分, 4GNet, VNet 各自生成的最大 q 值 q_{max} , 并判断出动作 a_t 。

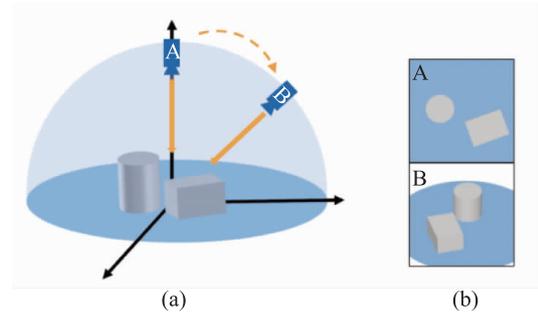


图 4 视点调整示例

Fig.4 Example of viewpoint adjustment

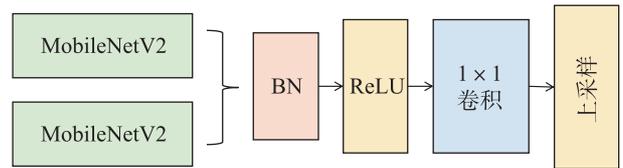


图 5 4GNet 与 VNet 的网络结构

Fig.5 Network structure of 4GNet and VNet

随后根据式(4)求出动作 a_t 的奖励 r_t 。得到一组动作 a_t 所对应的经验 $[s_t, a_t, r_t, a_{t+1}]$ 并存入对应经验池。然后,在经验池中随机选择一组其他经验作为 DQN 的输入, 计算生成目标 q 值。根据式(9)、式(10)计算损失值并更新 DQN 参数。

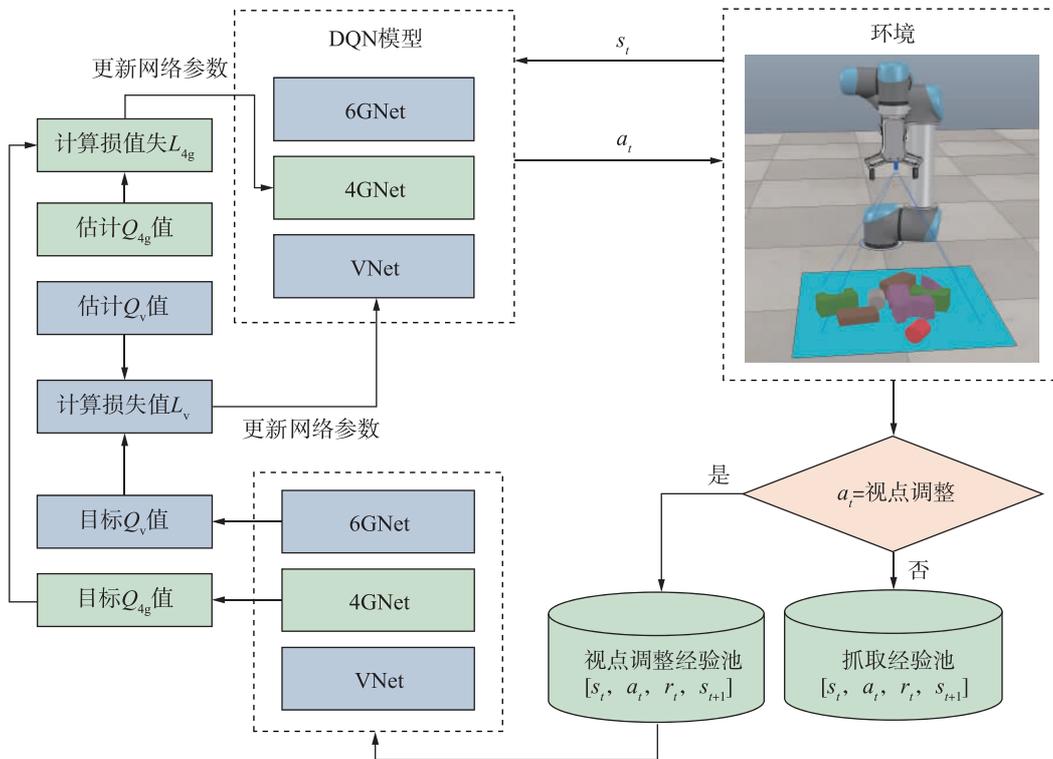


图 6 4GNet 与 VNet 训练与参数更新流程

Fig.6 Training and parameter update process for 4GNet and VNet

4 动作执行模块

通过动作预测模块得到各个原始动作的姿态和评分后, 将各个姿态数据输入动作执行模块进行决策, 决定最后执行的动作。如图 2 所示, 本文根据 6DoF 抓取评分来判断执行动作。如果 6DoF 抓取评分大于 0.85 且该抓取位姿存在无碰撞可达路径, 执行 6DoF 抓取动作; 如果 6DoF 抓取评分大于 0.85 但该抓取位姿不存在无碰撞可达路径, 执行 4GNet 所生成的自上而下的抓取动作; 若 6DoF 抓取评分小于 0.85, 则认为当前视点所捕获的场景信息不足以生成一个合理的抓取位姿, 则按照 VNet 所输出位置执行视点调整动作。定义视点调整动作包括 6 个自由度 $\{x, y, z, r_x, r_y, r_z\}$, 其中 x, y, z 表示末端执行器的位置, 其被限定在球体半径 R_s 为 18 cm 的上半球点的上半球面上。如图 4 所示, r_x, r_y, r_z 被定义为从位置 x, y, z 指向球心方向的欧拉角。如图 4(a) 所示, 视点调整可以到达位姿, 视点调整动作空间被限制在上半球, 方向指向中心。图 4(b) 表示机械臂分别从视点 A 和 B 捕获的图像。如果执行视点调整动作, 机械臂选择 VNet 输出的像素级 q 值图中 q 值最高的点的坐标来确定视点调整动作的 x 和 y 坐标。然后将 R_s, x 和 y 代入球面方程 $x^2 + y^2 + z^2 = R^2$ 求解出 z 。随后求出末端执行器的朝向向量 $(x - x_0, y - y_0, z - z_0)$, 其中 x_0, y_0, z_0 为上半球的球心坐标。转换得到 r_x, r_y, r_z 的值。控制机械臂达到指定位置后, 机械臂捕获 RGB-D 并将数据送入数据预处理模块, 进行视点场景信息融合。以此循环, 直至抓取成功。

5 实验

在实验部分, 本文在仿真环境中对提出的策略进行验证, 并在真实环境完成迁移。首先, 将本文提出的多动作协同抓取策略与已有的抓取策略进行比较, 证明该方法的性能和有效性。然后, 对该策略设计了消融实验, 证明 6DoF 视点调整的能有效增加机器人的感知面积和质量, 并测试不同自由度抓取方式的性能。最后, 在真实环境部署将该方法, 测试该方法抓取性能。

5.1 实施细节

本文通过部署 CoppeliaSim^[23] 仿真系统, 物理引擎为 Bullet Physics 2.83, 使用 CoppeliaSim 自带的逆运动学路径规划器。该仿真环境包括一个配备了 RG2 抓手的 UR5 机械臂 (UR5 机械臂与 Yaskawa

MOTOMAN-GP8 拥有相同的运动自由度, 因此选用 UR5 机械臂进行仿真实验)。在 RG2 抓手上安装有分辨率为 640×480 的 Kinect RGB-D 摄像机。设置初始摄像机在工作空间坐标系中的位置为 $[-0.45, 0, 0.33]$ 。在配备了 NVIDIA GTX 3060 GPU 的 Ubuntu18.04 服务器上进行训练和仿真。VNet, 4-GNet 都训练了 6 000 步, 折扣因子 γ 设置为 0.5。本策略使用两阶段的方式进行训练。第一阶段用于训练 4GNet, 在 4GNet 的训练中, 设置机械臂仅执行 4DoF 抓取操作, 而无需结合 6DoF 抓取和视点调整。每执行一次抓取动作都被视为一次迭代。第二阶段用于训练 VNet, 在 VNet 的训练过程中, 设置 3 种行为语: 4DoF 抓取、6DoF 抓取和视点调整共同执行。每个抓取都被视为一次迭代, 一次迭代可能包括一次抓取和多次视点调整。

5.2 基线模型

MVP (Multi-View Picking): MVP 是 MORRISON 等提出的一种主动学习方法。通过多视图拾取提高了抓取成功率, 根据当前捕获场景图像的熵值计算次优视角, 然后将手持摄像机机器人移动到目标视角。通过多次预测视角, 直到满足终止条件, 机器人执行抓取动作。

VPG (Visual Pushing for Grasping): VPG 是用于复杂抓取环境的抓退协同策略。该方法根据端到端生成的像素级 q 值图来决定执行的动作。使用推动作用来重定向难以抓握的对象。以此增加抓取成功率。

VAGS (Viewpoint Adjusting and Grasping Synergy): VAGS 用深度强化学习来协调视点调整和抓取动作。引入动态动作探索空间 (DAES) 方法规避一些不合理的探索位置, 加快模型收敛速度, 提高训练效率。

5.3 评估指标

为了更好地评价策略的性能, 使用以下指标来评估和验证模型的性能^[24-25]:

抓取成功率: 机械臂成功抓取的次数与尝试抓取的总次数之比。用来评估机械臂在完成抓握任务时的准确性和可靠性。

完成率^[26]: 本文规定完成一次实验为机械臂在连续抓取失败的次数不超过 5 次的情况下成功地拾取了所有物体。完成率的计算方法为完成实验的轮数比总的实验轮数。

可见面积增加率 (Visible Area Increase Rate, VAIR): 机械臂进行调整视点之后捕获的有效面积与调整视点之前捕获的有效区域的比率。有效面积

是指工作空间内的面积。本文的模拟环境俯视图如图 7(a)所示,蓝色区域属于工作空间。

有效点云增加率(Effective Point Cloud Increase Rate, EPCI):将表示物体的点云数据视为有效点云数据,表示桌面的点云视为劣质数据。引入了有效点云增加率评价视点调整的质量,该指标被定义为调整视点前后有效点云数量的差除以调整视点前有效点云的数量。图 7(b)显示机械臂捕获的点云。图 7(c)显示去除表示桌面的点云后留下的有效点云(被抓取物体表面的点)。

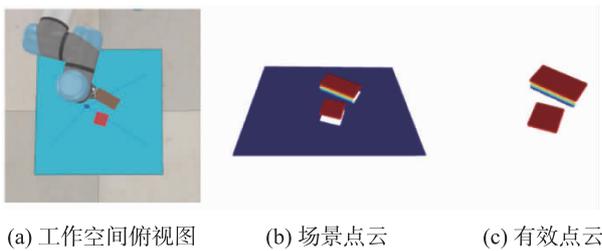


图 7 评价指标相关定义图解

Fig.7 Illustration of related definitions of evaluation metrics

5.4 仿真实验

在仿真实验中,建立了一个包含 8 个基本形状的被抓取目标库,包括圆柱、半圆柱、三角棱柱、长方体、正方体。每次在工作空间中添加 1~10 个物体,目标物体的形状,颜色,位置都是随机的。机械臂每种物体数量进行 50 轮抓取实验。

本文首先测试运动执行模块中抓取评分(6DoF 抓取评分)对抓取成功率的影响。将抓取评分设置为 0.80~0.90,步长设置为 0.01,记录机械臂在不同阈值下,抓取工作空间中包含 10 个被抓取物体时的抓取成功率。如图 8 所示,当抓取阈值在 0.85 的时候,机械臂的抓取表现最好,故本文将抓取评分设置为 0.85。

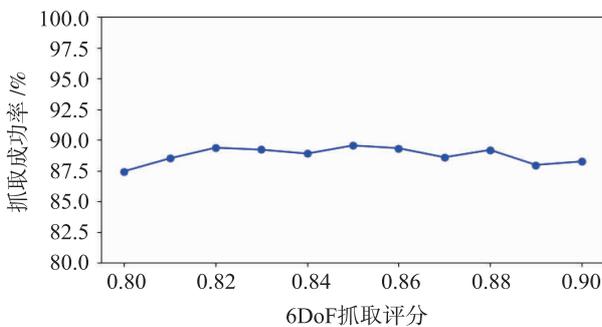


图 8 抓取成功率随抓取阈值的变化曲线

Fig.8 Grasping success rate curve with variation of grasping threshold

随后,将本文提出的方法与多个基线模型进行比较,实验结果如 1 表所示。图 9 演示了在仿真环境下的一轮抓取实验。本文提出的方法在每轮实验中都展示出更好的性能,并且随着场景中被抓取物体数量的增加,本文的方法的优势变得更大,在工作空间中被抓取物体数量达到 10 个时,本文的策略与 VPG 相比抓取成功率提高了 12.02%,与先进主动感知抓取策略 VAGS 相比抓取成功率提高了 9.38%,这表明本文的方法可以适应抓取物体更多的复杂场景。

表 1 不同抓取策略的抓取成功率

Tab.1 Grasping success rat of different grasping strategies

单位: %

被抓取物体的数量	MVP	VPG	VAGS	本文方法
1	92.33	93.67	96.00	100
2	88.73	91.00	96.06	97.06
3	87.03	88.30	90.50	95.83
4	84.21	86.80	91.63	96.11
5	79.61	86.01	86.15	91.53
6	74.91	84.50	84.55	91.67
7	74.73	81.40	82.48	88.07
8	73.21	78.09	84.95	91.49
9	71.22	80.50	80.41	92.29
10	70.96	77.51	80.15	89.53

仿真实验中,在工作空间中被抓取物体数量达到 10 个时,本文通过记录抓取一个物体需要的平均运动次数以及机器人从全局拍照位处为起点成功抓取物体所需的平均时间两个指标来验证抓取效率。实验结果如表 2 所示,本文提出的方法抓取一个物体需要的平均只需要执行 1.36 次运动,原因是随着抓取的进行,场景中物体不断减少,当场景中的物体不存在遮挡关系时,机器人可以在不进行视点调整的情况下进行抓取。但本文所用的模型结构相交于 VAGS 更加复杂,所以平均抓取时间慢于 VAGS。

表 2 抓取效率实验结果

Tab.2 Experiment results of grasping efficiency

方法	抓取成功率 / %	平均运动次数	平均时间 /s
MVP	70.96	3.38	17.8
VAGS	80.15	1.43	9.0
本文方法	89.53	1.36	11.7

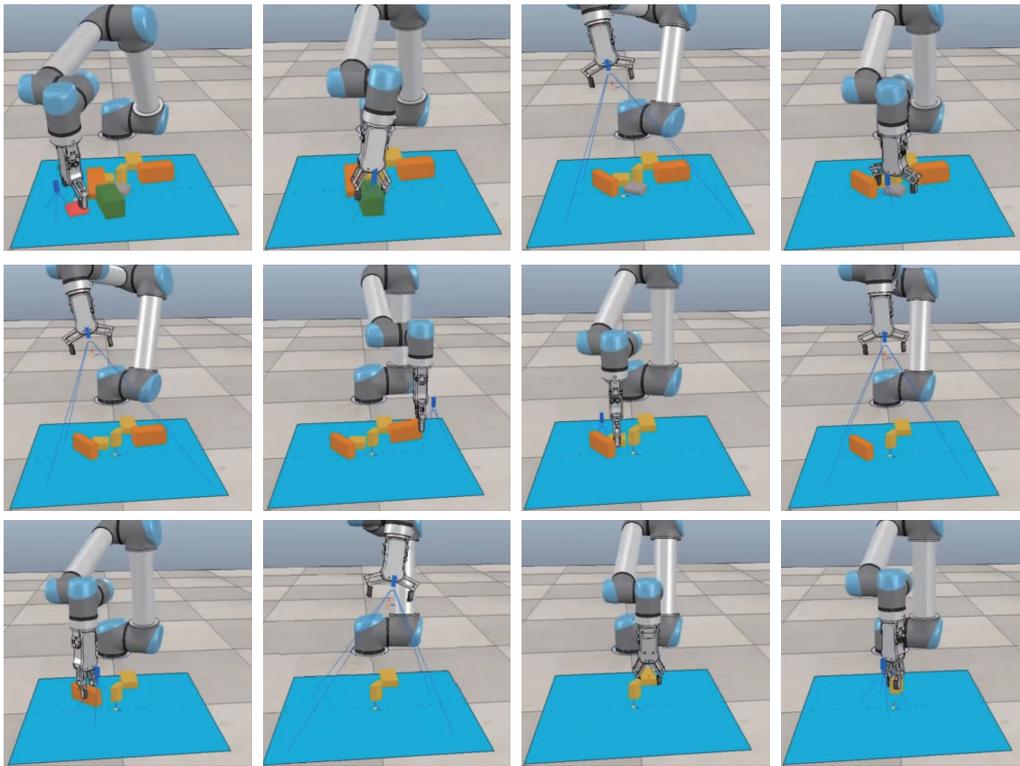


图9 仿真环境下实验过程

Fig.9 The experimental process in a simulated environment

5.5 消融实验

在本节中, 将本文提出的方法与消融方法进行比较。包含两个测试目标: 1) 测试视点后调整的场景融合方法是否能帮助机械臂更好的获得场景信息。2) 测试将抓取姿势从自上而下上升到6DoF是否能帮助机械臂完成更好的抓取。为了验证以上猜想, 本文设置了两个实验:

视点调整实验: 该实验比较各种视点调整方式的场景信息获取能力。本文使用可见面积增加率(VAIR), 有效点云增加率(EPCI)两个指标进行场景信息获取能力评估。可见面积增加率越高代表该视点调整可以带来更大的有效区域增加; 有效点云增加率越高代表视点调整动作获取到的点云质量越高。本文对50个具有随机被抓取物体个数、颜色、位置、方向和形状的场景进行测试, 从50次实验运行中平均可见面积增加率和有效点云增加率, 实验结果如表3所示。

如图10所示, 6DoF实验使用的是本文所提出的视点调整动作, 获取的场景点云如图10(a)所示; 4DoF实验使用的视点调整动作是自上而下。获取的场景点云如图10(b)所示。两种方法都使用

表3 视点调整实验结果

Tab.3 The experimental results of viewpoint adjustment

视点调整方法	VAIR	EPCI
6DoF	1.22	0.256
4DoF	1.12	0.052

视点调整后场景融合方法处理捕获的场景信息。为了更好地对比两种方法, 本文将两种方法在同一位置获取的点云信息转到机器人坐标系, 如图10(c)所示。其中, 被蓝色形状框住的区域为6DoF视点调整动作获取的点云信息。红色区域为自顶向下(4DoF)视点调整动作获取的点云信息。可以看出, 6DoF视点调整方法不仅可以覆盖更大的面积, 还可以更好地捕获目标侧面的点云信息。表3的实验结果验证了这一推测。6DoF视点调整方法捕获了更大的有效区域和更有效的点云, 这有助于模型生成更好的抓取姿势。

抓取策略对比实验: 该实验设置了3个多模型抓取策略进行比较: 1) 本文所用多动作协同抓取策略。2) 不使用6GNet, 抓取采用4GNet输出的抓取位姿进行抓取。3) 在数据预处理模块不使用视

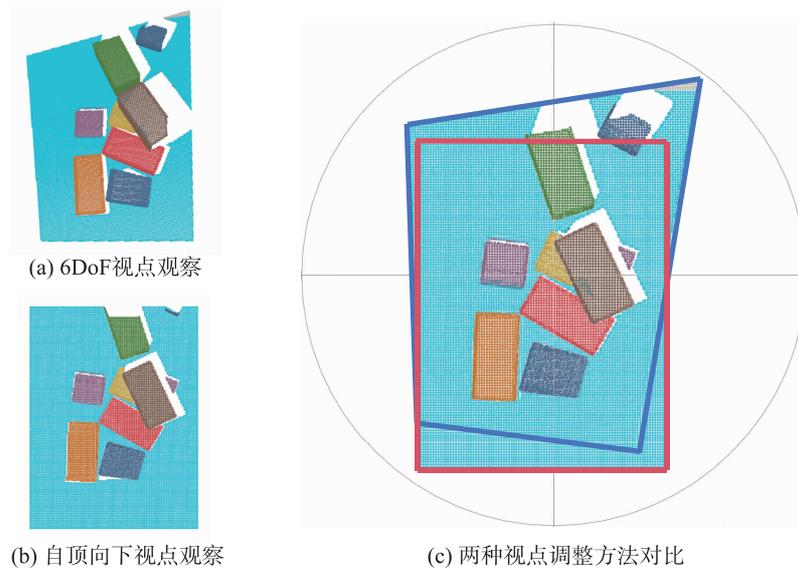


图 10 不同视点调整方法获取的点云示意图

Fig.10 Schematic diagram of the point cloud obtained by different viewpoint adjustment methods

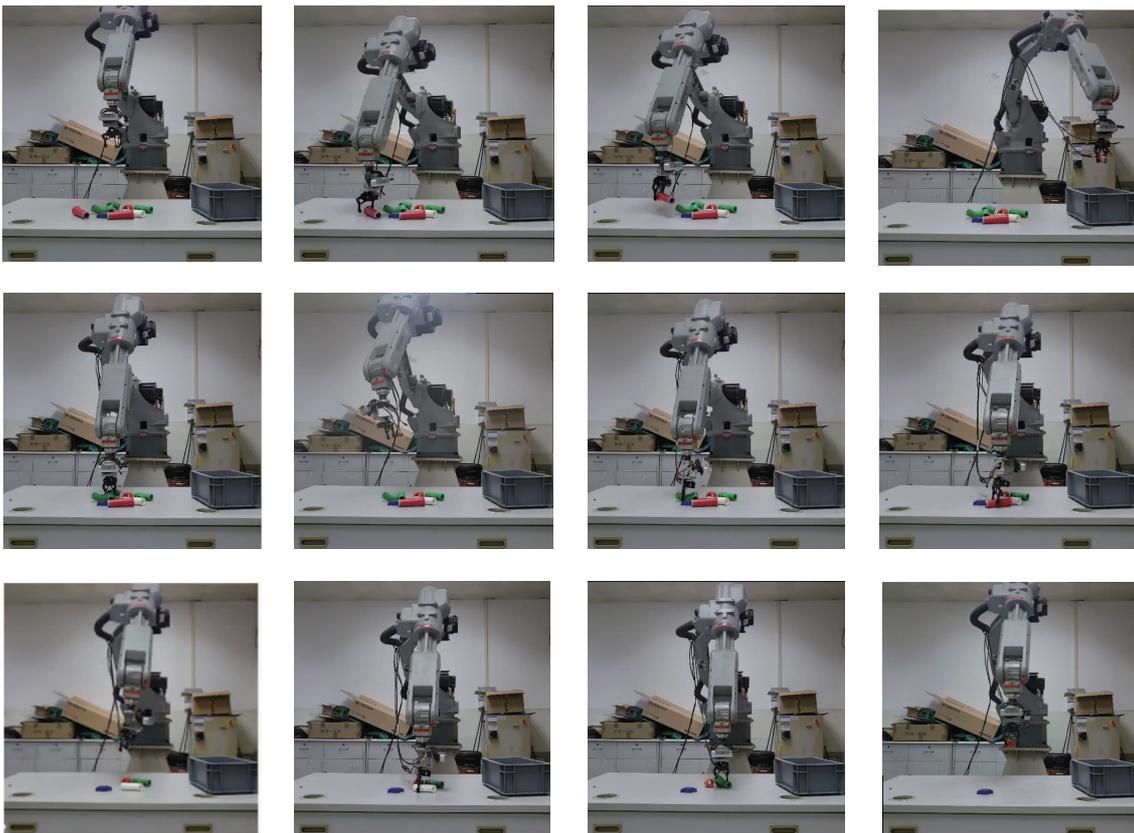


图 11 真实环境下实验过程

Fig.11 Experimental process in a real-world environment

点调整后场景融合方法，策略只使用当前视点获取的场景信息来进行抓取。实验流程与仿真实验的流程相同，即在工作空间中添加 1~10 个物体，目标物体的形状，颜色，位置都是随机的。机械臂每种

物体数量进行 50 轮抓取实验。

实验结果如 4 表所示。

3 种方法的完成率均为 100%。通过比较可以看出，在使用用于视点调整的场景融合方法后，场

景融合方法的利用有效地提高了机械臂的抓取性能。此外, 6DoF 抓取的引入为机械臂提供了杂乱环境中抓取选项更多, 与仅使用 4DoF 抓取的情况相比, 抓取成功率提高了 4.5%。

表 4 抓取策略对比实验

Tab.4 Comparison experiment of grasping strategies

单位: %

抓取策略	完成率	抓取成功率
策略 1	100	93.46
策略 2	100	88.98
策略 3	100	86.64

5.6 真实环境实验

为了更好地验证所提出策略的可行性, 本文在真实的机器人环境中进行了实验。如图 12 所示, 真实机器人环境配置由一个带伺服驱动的机械夹持器的 Yaskawa MOTOMAN-GP8 6DoF 工业机器人和一个固定在机械臂上的 Intel RealSense D435 摄像头组成, 由 MOTOCOM32 控制。被抓取物体选用标准工业工艺件。真实实验设置初始摄像机位置为 $[0.99, -1.21, -0.10]$, 单位为 m。

图 13 为真实环境中个行为原语演示, 其中, 橙色箭头代表机械臂移动方向, 白色机械臂掩模代表

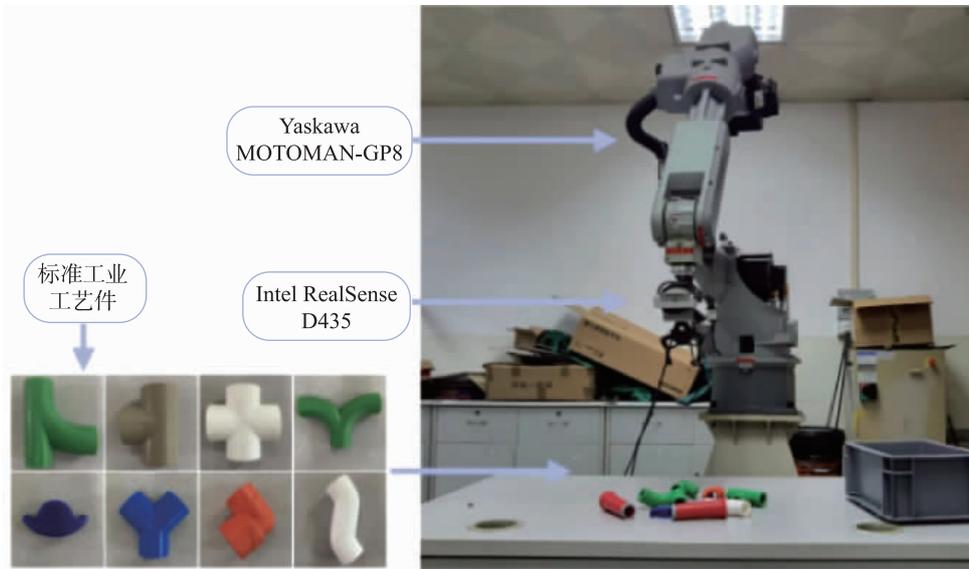


图 12 真实环境实验设备与抓取物体

Fig.12 Real-world experimental equipment and the object to be grasped

机械臂本次移动的起始位置。图 13(a) 演示了视点调整并在视点调整后进行一次 6DoF 抓取; 图 13(b) 演示了直接进行一次 6DoF 抓取; 图 13(c) 演示了视点调整后的 4DoF 抓取。随机放置 5 ~ 10 个标准工业工艺件在机械臂工作空间内。测试 VPG、VAGS 及策略 2、策略 3 四种方法的性能, 每论方法设置了 20 个场景, 并且所有模型都没有经过重新训练, 都是从模拟环境之间迁移到真实环境。图 11 演示一轮真实环境下的抓取实验。每次抓取成功后都将抓取物体放入图右下放的收纳箱中。实验结果如表 5 所示, 表中 gs 表示抓取成功率, cp 表示完成率。实验结果显示, 策略 1、策略 2 都在各指标上都优于其他算法, 但对比仿真环境有一定的差距, 这是由于真实环境在进行多视点场景信息融合时存

在误差所导致的。实验结果说明, 本文的策略有良好的泛化能力, 确保模型可以用于抓取不同形状, 不同纹理的物体。

表 5 真实环境实验结果

Tab.5 Real-world experimental results

物体个数	VPG		VAGS		策略 1		策略 2	
	gs	cp	gs	cp	gs	cp	gs	cp
5	78.5	90	86.8	95	91.6	100	89.9	95
6	72.5	85	83.8	95	92.1	100	86.4	95
7	68.3	75	80.6	90	86.7	95	84.8	95
8	66.4	70	80.8	90	87.4	95	84.9	95
9	62.0	65	82.5	90	85.3	90	80.8	90
10	60.5	60	83.5	90	84.6	90	82.1	90
平均	68.0	74	83	91	87.9	95	84.8	93

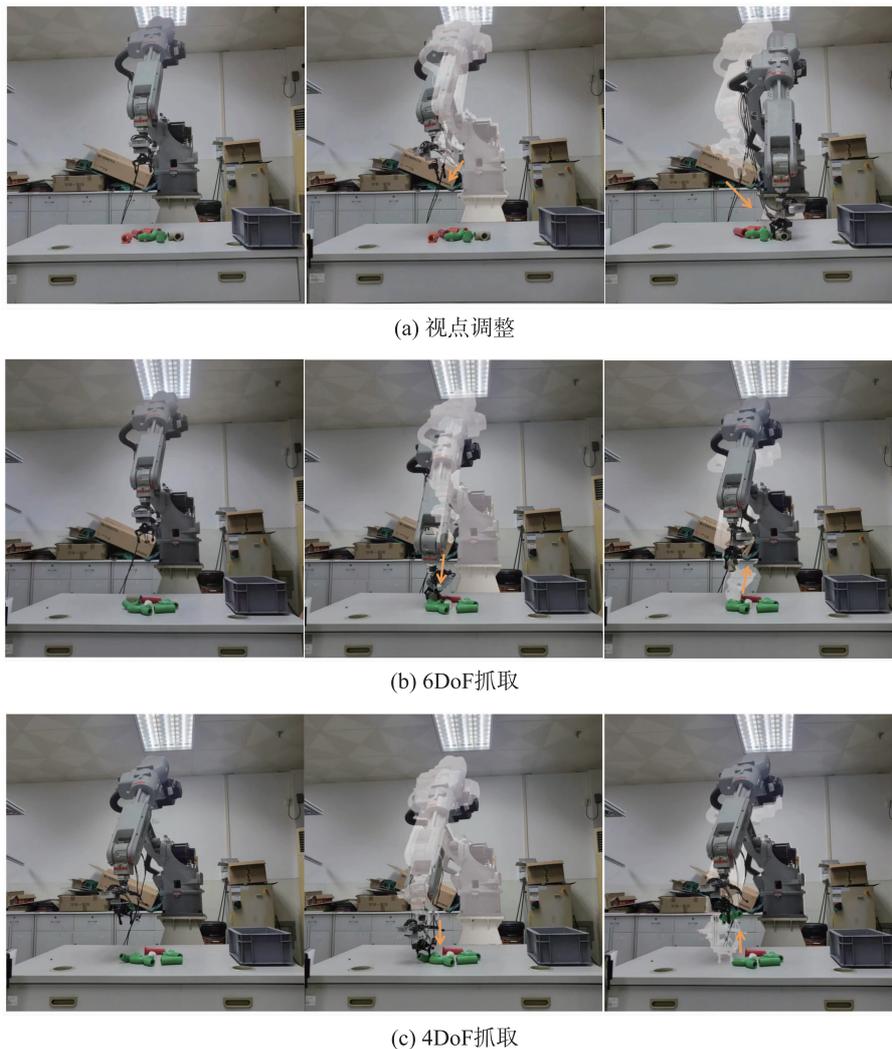


图 13 真实环境下各行为原语演示

Fig.13 Demonstration of various primitive behaviors in a real environment

6 总结

本文提出了一种 6DoF 空间中物体抓取和视点调整的协同抓取策略。首先,该策略能使机械臂根据场景信息自主选择合适的原始动作来执行。其次,本文通过扩展机械臂工作空间到 6DoF,为抓取动作提供了更多的可能性。同时,6DoF 姿态下的视点调整动作相较于自上向下的视点调整动作能捕获更大、更有效的场景信息。此外,为了更好地利

用机械臂所捕获的场景信息,本文提出了用于视点调整的场景融合方法,将从多个视点捕获的信息集成到等分辨率的高度图中,增强了机械臂以整体方式感知目标对象的能力。实验结果表明,本文方法在复杂场景中的表现具备较好的先进性,在包含 10 个被抓取物体的复杂环境中,依然能保持 89.53% 抓取成功率。策略在通过仿真环境训练后,可直接推广到真实测试环境,证明了策略的鲁棒性和泛化能力。

参考文献

- [1] 徐进,柳宁,李德平,等. 一种基于抓取簇和碰撞体系的工业零件抓取姿态检测算法[J]. 机器人, 2022, 44(2): 153 - 160.
- XU J, LIU N, LI D P. et al. A grasping poses detection algorithm for industrial workpieces based on grasping cluster and collision voxels[J]. Robot, 2022, 44(2): 153 - 166.
- [2] HAFNER R, BARTH-MARON G, VECERIK M, et al. Data-efficient deep reinforcement learning for dexterous manipulation

- [EB/OL]. [2023-10-12]. <https://arxiv.org/abs/1704.03073>. DOI: 10.48550/arXiv.1704.03073.
- [3] QUILLEN D, JANG E, NACHUM O, et al. Deep reinforcement learning for vision-based robotic grasping: A simulated comparative evaluation of off-policy methods [C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2018: 6284-6291.
- [4] BREYER M, FURRER F, NOVKOVIĆ T, et al. Comparing task simplifications to learn closed-loop object picking using deep reinforcement learning[J]. IEEE Robotics and Automation Letters, 2019, 4(2): 1549-1556.
- [5] MAHLER J, POKORNY F T, HOU B, et al. Dex-Net 1.0: A cloud-based network of 3D objects for robust grasp planning using a multi-armed bandit model with correlated rewards [C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2016: 1957-1964.
- [6] MAHLER J, LIANG J, NIYAZ S, et al. Dex-Net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics[EB/OL]. [2023-10-16]. <https://arxiv.org/abs/1703.09312>. DOI: 10.48550/arXiv.1703.09312.
- [7] MAHLER J, MATL M, LIU X, et al. Dex-Net 3.0: Computing robust vacuum suction grasping targets in point clouds using a new analytic model and deep learning [C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2018: 5620-5627.
- [8] KALASHNIKOV D, IRPAN A, PASTOR P, et al. Scalable deep reinforcement learning for vision-based robotic manipulation [C]//Conference on Robot Learning. New York, USA: PMLR, 2018: 651-673.
- [9] 王高, 陈晓鸿, 柳宁, 等. 一种基于视角选择经验增强算法的机器人抓取策略[J]. 华南理工大学学报(自然科学版), 2022, 50(9): 126-137.
WANG G, CHEN X H, LIU N, et al. A robot grasping policy based on viewpoint selection experience enhancement algorithm [J]. Journal of South China University of Technology, 2022, 50(9): 126-137.
- [10] DOGAR M R, SRINIVASA S S. A planning framework for nonprehensile manipulation under clutter and uncertainty[J]. Autonomous Robots, 2012, 33: 217-236.
- [11] ZENG A, SONG S, WELKER S, et al. Learning synergies between pushing and grasping with self-supervised deep reinforcement learning [C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2018: 4238-4245.
- [12] YANG Z, SHANG H. Robotic pushing and grasping knowledge learning via attention deep Q-learning network [C]//International Conference on Knowledge Science, Engineering and Management. Piscataway, USA: IEEE, 2020: 223-234
- [13] 李鑫, 沈捷, 曹恺, 等. 深度强化学习的机械臂密集场景多物体抓取方法[J/OL]. 计算机工程与应用, 2024. [2024-03-29]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20231008.1644.008.html>.
LI X, SHEN J, CAO K, et al. Multi-object grasping method for robotic arms in dense scenes using deep reinforcement learning [J/OL]. Computer Engineering and Applications, 2024. [2024-03-29]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20231008.1644.008.html>.
- [14] MOHAMMED M Q, KWEK L C, CHUA S C, et al. Deep reinforcement learning-based robotic grasping in clutter and occlusion [J]. Sustainability, 2021, 13(24): 13686-13713.
- [15] MORRISON D, CORKE P, LEITNER J. Multi-view picking: Next-best-view reaching for improved grasping in clutter [C]//International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2019: 8762-8768.
- [16] LIU N, GUO C, LIANG R, et al. Collaborative viewpoint adjusting and grasping via deep reinforcement learning in clutter scenes[J]. Machines, 2022, 10(12): 1135-1156.
- [17] 李会军, 瞿孝昌, 叶宾. 基于未知物体三维点云特征的机器人六自由度抓取[J]. 控制理论与应用, 2022, 39(6): 1103-1111.
LI H J, QU X J, YE B. Robot six-degree-of-freedom grasping based on 3D point cloud features of unknown objects[J]. Control Theory and Applications, 2022, 39(6): 1103-1111.
- [18] BALASUBRAMANIAN R, LING X, BROOK P D, et al. Human-guided grasp measures improve grasp robustness on physical robot [C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2010: 2294-2300
- [19] QI C R, SU H, MO K, et al. PointNet: Deep learning on point sets for 3d classification and segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2017: 652-660.

- Research, 2013, 32(11): 1231 – 1237.
- [18] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, USA: IEEE, 2021: 2235 – 2239.
- [19] LI Y, YAO T, PAN Y, et al. Contextual transformer networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(2): 1489 – 1500.
- [20] LIU S, HUANG D. Receptive field block net for accurate and fast object detection[C]//European Conference on Computer Vision. Berlin, Germany: Springer, 2018: 385 – 400.
- [21] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834 – 848.
- [22] 盛博莹, 侯进, 李嘉新, 等. 面向复杂交通场景的道路目标检测方法[J]. 计算机工程与应用, 2023, 59(15): 87 – 96.
SHENG B Y, HOU J, LI J X, et al. Road target detection method for complex traffic scenarios[J]. Journal of Computer Engineering & Applications, 2023, 59(15): 87 – 96.
- [23] 董恒祥, 潘江如, 董芙楠, 等. 基于改进 YOLOv5s 模型的车辆及行人检测方法[J]. 北华大学学报(自然科学版), 2024, 25(2): 244 – 254.
DONG H X, PAN J R, DONG F N, et al. Vehicle and pedestrian detection method based on improved YOLOv5s model[J]. Journal of Beihua University (Natural Science Edition), 2024, 25(2): 244 – 254.

作者简介

濮志远(2001 –), 男, 硕士。研究领域为无人驾驶汽车环境感知。

罗素云(1975 –), 女, 硕士, 副教授。研究领域为无人驾驶汽车环境感知及控制。

(上接第 631 页)

- [20] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2018: 4510 – 4520.
- [21] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. New York, USA: PMLR, 2015: 448 – 456.
- [22] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]//International Conference on Machine Learning. New York, USA: PMLR, 2010: 807 – 814.
- [23] ROHMER E, SINGH S P N, FREESE M. V-REP: A versatile and scalable robot simulation framework[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2013: 1321 – 1326.
- [24] 赵亮, 杨铁, 于鹏, 等. 一种用于机械臂拟人化控制的学习框架[J]. 机器人, 2023, 45(5): 513 – 522.
ZHAO L, YANG T, YU P, et al. A Learning framework for anthropomorphic control of robotic arms[J]. Robot, 2023, 45(5): 513 – 522.
- [25] 孙先涛, 陈文杰, 陈伟海, 等. 一种机械手定位装置及其定位方法: CN112873176B[P]. 2021 – 09 – 07.
SUN X T, CHEN W J, CHEN W H, et al. A Mechanical hand positioning device and its positioning method: CN112873176B [P]. 2021 – 09 – 07.
- [26] TANG B, CORSARO M, KONIDARIS G, et al. Learning collaborative pushing and grasping policies in dense clutter[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2021: 6177 – 6184.

作者简介

李德平(1987 –), 男, 博士, 讲师。研究领域为计算机视觉, 3 维视觉, 机器人视觉感知。

洪楷宣(2000 –), 男, 硕士生。研究领域为计算机视觉。