

基于视觉感知和触觉先验知识学习的机器人稳定抓取

薛 腾, 刘文海, 潘震宇, 王伟明

(上海交通大学机械与动力工程学院, 上海 200240)

摘 要: 针对机器人抓取过程中需要实时评估抓取质量以动态调整抓取构型的问题, 提出了基于触觉先验知识的机器人稳定抓取方法. 首先, 根据抓取过程中物体抵抗外界扰动的能力, 提出了一种基于触觉信息的抓取质量评估方法. 在此基础上, 建立了视触觉联合数据集并学习触觉先验知识. 其次, 提出了融合视觉图像和触觉先验知识的稳定抓取构型生成方法. 最后, 在搭建的机器人抓取系统对 10 种目标物体进行了实验验证. 结果表明, 相比传统的视觉方法, 该方法的抓取稳定性提高了 55%; 针对已知物体和未知物体, 稳定抓取成功率分别为 86% 和 79%, 体现了较好的泛化效果.

关键词: 稳定抓取; 触觉先验知识; 视触觉融合; 抓取质量评估

中图分类号: TP242

文献标识码: A

文章编号: 1002-0446(2021)-01-0001-08

Stable Robotic Grasp Based on Visual Perception and Prior Tactile Knowledge Learning

XUE Teng, LIU Wenhai, PAN Zhenyu, WANG Weiming

(School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Considering the problem that the robot needs to estimate the grasp quality in real time to dynamically adjust the grasp configuration in grasping process, a stable robotic grasp method based on prior tactile knowledge learning is proposed. Firstly, a tactile information based method is put forward for evaluating grasp quality according to the capacity of resisting the external perturbations in grasping process. Based on it, a visual-tactile joint dampest is built, and the prior tactile knowledge is learned. Secondly, an architecture is proposed to generate stable grasp configurations by fusing the visual image and the prior tactile knowledge. Finally, an experimental verification with 10 kinds of objects is carried out on the proposed robotic grasp system. The results show that the grasp stability is improved by 55% with the proposed method over traditional vision-based methods, and the success rates of stable grasp are 86% and 79% respectively on known and unknown objects, demonstrating good generalization performance of the proposed method.

Keywords: stable grasp; prior tactile knowledge; visual tactile fusion; grasp quality evaluation

1 引言 (Introduction)

随着传感器硬件与多模态融合技术的发展, 机器人的感知与决策越来越高效, 可执行的自动化操作越来越复杂. 然而, 其操作的灵巧性仍旧远远不如人类, 如: 稳定抓取.

显然, 视觉在机器人抓取规划阶段具有重要作用^[1], 有些研究基于全局感知信息直接生成抓取构型^[2]. 然而, 机器人抓取是一个与物体固有属性、机械手特征以及相对摩擦密切相关的物理接触操作过程. 为了更好地表征机械手与目标物体之间的物理接触, 触觉被用于感知操作过程中接触力的细微变化以实时评估抓取质量. 研究人员提出了大

量基于触觉感知的抓取质量指标. Calandra^[3] 利用 GelSight 触觉传感器和 RGB 相机获取视触觉联合信息, 通过布尔量表征抓取质量. 类似地, Allen^[4] 也提出了一个基于 Barrett 机械手触觉信息的布尔量评价指标. Hogan 等^[5] 则通过晃动末端执行器并观测物体状态来评估抓取质量稳定性, 以 0~1 的连续值分别表示抓取失败、跌落与稳定抓取. 除此之外, 抓取过程中的滑动也是一个值得考虑的因素, Howe 和 Cutkosky 等^[6] 指出加速度感知可以用来表征物体滑动, Melchiorri 等^[7] 通过测量接触点的法向力和切向力, 并将二者的比值与摩擦系数进行比较来评估滑动情况. Su 和 Chebotar 等^[8-9] 使用 BioTac (一种由 19 个电极组成的刚性核心外壳组

成的仿生触觉传感器)感知切向力和振动的变化,并通过时空特征描述子表达不同抓取情况下 BioTac 的感知情况,利用支持向量机(SVM)进行滑动预测分类学习.类似地,James 等^[10]通过测量嵌入仿生触觉传感器内部的指针位置,利用 SVM 方法判别物体是静态的还是滑动的.Yuan 等^[11]则通过检测 GelSight 传感器的弹性形变来感知正向力、切向力与扭转载荷从而判断物体的抓取状态.

在当前基于触觉的抓取质量评估研究中,触觉传感器的数据获取方式和表现形式随着传感器硬件种类的不同而表现出较大的差异性,不同的抓取质量指标对特定传感器有较大的硬件依赖性.考虑到本文使用的触觉传感器只能测量 5 个线性排列的接触点的正向压力值,无法获取上述文献中抓取质量指标所需要的原始数据,因此有必要面向该硬件提出更具针对性的评价指标.

关于机器人稳定抓取策略,Hogan 等^[5]提出了一种数据驱动的稳定抓取模型,基于重复抓取策略和抓取质量评价方法动态规划最佳的二次抓取.这种方法可以避免直接从大量的原始数据中学习重复抓取策略.然而,该模型需要根据当前抓取构型的预测值不断调整接触点,是一个动态的抓取过程.与之类似,Hyttinen 等^[12]基于一个布尔型的分类器判断抓握是否稳定,并根据预测值调整抓取位姿.Guo 等^[13]通过视觉感知生成初始抓取矩形框,并利用触觉信息和应变计评估抓取质量,重复完成多次抓取,直至超过预先设定的阈值.Calandra 等^[3]提出了一个端到端的抓取规划模型,用于预测当前抓取方案条件下的输出值,并通过多次迭代选择最佳的抓取方案.

可以发现,上述方法通过感知当前抓取过程中的触觉信息来评估抓取质量,并结合稳定性评价指标与抓取适应性方法进行多次重复抓取,严重依赖于多次试错,需要连续不断地评估与调整抓取姿态,直至满足稳定性要求.



图 1 婴儿与儿童不同的稳定抓取方式

Fig.1 Illustration of stable grasp of an infant and a child

图 1 展示了普通婴儿以及儿童在稳定性抓取方

式上的不同.图 1(a)描述了一个婴儿尝试抓取一个玩具积木的过程,通过观察发现,婴儿无法直接确定应该抓取哪个位置,会经历很多次掉落或者滑动.图 1(b)描述了一个儿童通过稳定抓取完成搭建完整积木组合体的过程.正如 Goodale^[14]指出,2 种情况的区别是婴儿在抓取的瞬间不能立刻判断出合适的抓取点与抓取角度,也可被认定为缺乏相应的触觉先验知识,而儿童通过不断的抓取尝试,学习到了视觉和触觉在抓取过程中的相互联系,仅通过视觉就能判断出满足稳定抓取的抓取点和角度,特别是在抓握后能从触觉感知判断出抓取是否稳定,这有效避免了重复抓取,并对后续操作,例如搭积木,具有重要的作用.

基于此,本文认为婴儿阶段的稳定性抓取更多地依赖于不断重复,后婴儿阶段则是基于先前积累的触觉经验,直接通过视觉实现精准稳定抓取,避免了不断试错的过程.本文的目的是赋予机器人类似的触觉先验能力,通过将触觉先验知识应用于机器人稳定抓取构型生成过程,使机器人在稳定抓取的感知与决策中具有类人的灵巧操作能力.本文技术框架如图 2 所示.

本文的主要创新点包括以下两部分:

- 1) 抓取质量评价与触觉先验知识学习.提出了面向多点阵列式触觉传感器的数据处理与质量评价方法,基于此,提出了触觉先验知识学习模型.
- 2) 触觉先验抓取规划.提出了一个多阶段的稳定性抓取策略,能够基于触觉先验知识与视觉图像生成稳定的抓取构型.

2 触觉先验知识学习 (Prior tactile knowledge learning)

2.1 抓取质量评价指标

本文使用由 Righthand Lab 生产的 TakkStrip 触觉传感器.受文 [5] 启发,设计了一组机器人扰动行为以评估抓取质量,包括 2 次直线移动(图 3 中的步骤 5)与 2 次末端抖动(图 3 中的步骤 6).机械臂抓取物体到达指定高度后,先记录一次当前压力值,而后执行扰动行为,每当发生扰动行为时记录一次压力值.由于 TakkStrip 传感器可一次性记录 5 个接触点的压力值,因此记录 5 次数据后,可生成 5×5 的数据矩阵,将其以 RGB 图像的形式表示,定义为触觉图像.

本文定义 3 种抓取状态以及相应的质量评估方式:

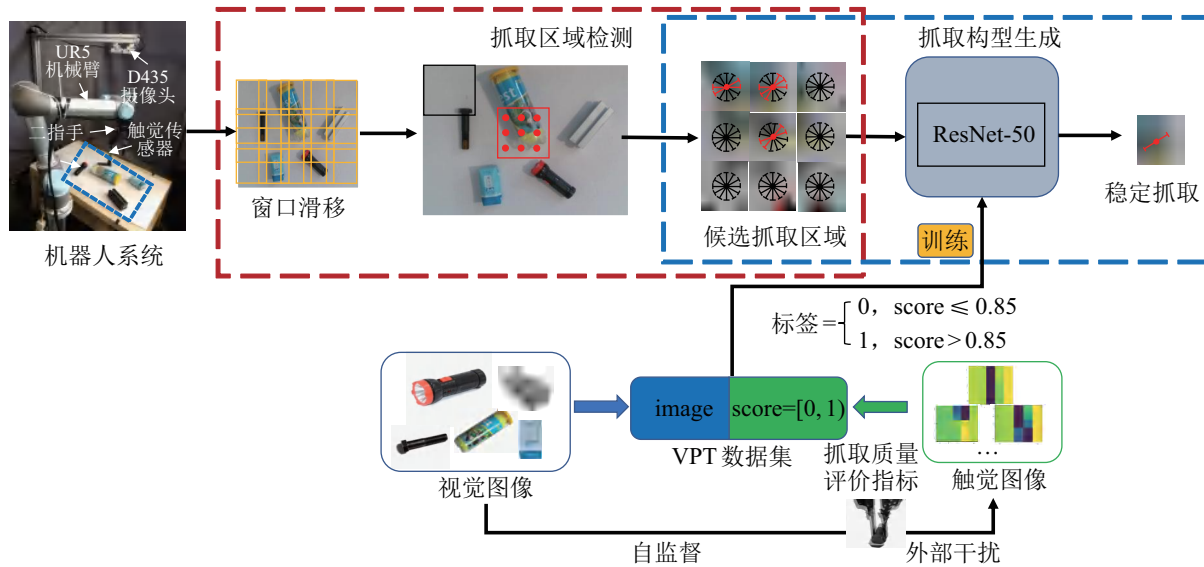


图 2 基于触觉先验知识的稳定抓取技术框架

Fig.2 Overview of the stable grasp based on prior tactile knowledge

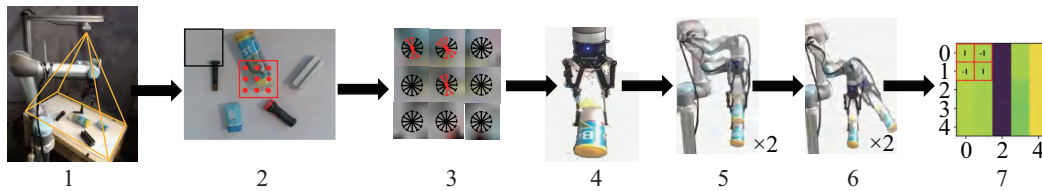


图 3 数据采集过程

Fig.3 Data collection process

1) 跌落 (图 4(a)): 通过实验观察, 跌落的压力值表现为一个接触点的巨大数值变化。基于此, 设计了一个 1×2 的卷积核, 并与触觉图像做卷积运算以提取特征。由于运算结果表现的是同一压力点在不同时刻的压力值变化, 该 1×2 卷积核被定义为时间核 (图 5(a))。卷积结果是一个 4×5 矩阵, 每一行都代表了一次扰动行为后的压力值变化。本文认为最大值出现的行值代表了跌落时间, 即式 (1) 中的 i 。 i 越小, 表示跌落得越早, 说明其抓取稳定性越差, 抓取质量分数相对越低。

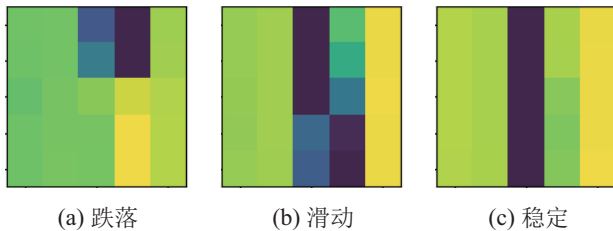
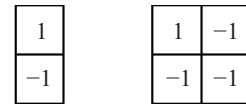


图 4 触觉图像

Fig.4 Tactile images

2) 滑动 (图 4(b)): 当发生滑动时, 相邻接触点的压力值将会反向变化。举例而言, 假设物体从接触点 1 滑动到接触点 2, 则接触点 1 的压力值将

会降低而接触点 2 的压力值则会增加。本文设计了一个 2×2 的卷积核用于触觉图像卷积运算。该卷积核由“1”“-1”组成, 通过卷积运算即可计算相邻接触点在经历一次扰动行为后的压力值变化。因此, 定义该卷积核为空间核 (图 5(b))。运算结果是一个 4×4 的矩阵, 所有的元素求和代入式 (1), 即 x 。 x 越小, 说明压力值变化越小, 即滑动的强度相对较低, 物体更稳定。



(a) 时间核 (b) 空间核

图 5 时间核与空间核

Fig.5 Time kernel and space kernel

3) 稳定 (图 4(c)): 本文认为, 不存在理想的完全稳定, 定义滑动程度较小的状态为稳定状态, 其抓取质量计算过程与“滑动”状态一致。

图 4 显示了在物体跌落、滑动与稳定抓取 3 种状态下的触觉图像。

基于触觉传感器硬件特性与上述抓取状态分类, 本文提出了一套抓取质量评价指标, 通过 0~1

的连续值表征抓取稳定性. 评价指标如下:

$$\begin{cases} 0, & \text{抓取失败} \\ 0.5 \times \frac{i}{5}, & \text{跌落} \\ 0.5 \times e^{-x/1000} + 0.5, & \text{滑动或稳定} \end{cases} \quad (1)$$

式中, i 表示跌落物体的触觉图像与时间核卷积运算后最大值出现的行值, x 表示未跌落物体的触觉图像与空间核卷积运算后的元素和.

2.2 视触觉联合数据集

目前, 有多个公开的视触觉联合数据集, 比如南加州大学的 Chebotar 等^[15] 基于 BioTac 建立的 BiGS 抓取稳定性数据集, 麻省理工学院的 Luo 等^[16] 基于 GelSight 建立的 ViTac 数据集, 清华大学 Wang 等^[17] 基于英特尔中国研究院的 Eagle Shoal 机械手建立的视触觉联合数据集. 但是, 触觉数据的获取与表现形式严重依赖于触觉传感器本身, 在不同的传感器之间难以实现泛化. 本文基于 TakkStrip 触觉传感器和 D435 深度相机建立了 VPT 视触觉联合数据集.

以 5 类常见生活物品作为抓取对象, 采用 Intel 公司的 D435 相机采集视觉图像, 采用 TakkStrip 触觉传感器采集触觉信息. 抓取对象被随机摆放在桌面上, 相机采集图片, 而后机械臂执行扰动操作, 与此同时, 触觉传感器记录压力值. 通过式 (1) 计算抓取质量, 标记原始图片. 数据采集过程如图 3 所示.

图 6 为 VPT 数据集的一部分, 包括视觉图像、正负标签抓取点与正负标签抓取角度, 其中正样本的抓取构型标记为红色, 负样本的抓取构型标记为黑色. 正负样本的标记是一个自动计算过程, 主要通过机械臂执行扰动行为并采集触觉传感器信息, 根据触觉信息计算抓取质量, 实现自动标记. 本文规定抓取质量 $q \geq 0.85$ 为稳定抓取的正样本, 其余为负样本, 包括滑动、跌落与抓取失败等情况.



图 6 VPT 数据集

Fig.6 VPT dataset

考虑二指手的抓取特性, 执行器末端旋转 180° 的抓取效果与原始角度一致, 因此, 本文采用图片旋转的方式进行数据增强. 每一个训练样本旋转 180° , 其正负标签保持不变, 从而样本容量扩大 2 倍. 机械臂执行抓取与扰动操作 1500 次, 经过数据增强, 最终建立包含 3000 个样本的 VPT 数据集.

2.3 触觉先验知识学习

基于 VPT 数据集, 本节提出触觉先验知识学习模型, 如图 7 所示.

模型采用 ResNet-50^[18] 作为特征提取器, 包含 16 个残差模块, 建立了视觉图像与抓取质量之间的映射关系. 网络以 VPT 数据集中的 100×100 像素彩色图像正负样本作为输入, 包含 R、G、B 三个通道, 经特征提取器生成 512 通道的特征图, 再经由自适应池化和线性回归得到单一回归量, 即触觉图像对应的抓取质量分数. 网络训练采用 Adam 优化方法, 以均方差 MSEloss 作为损失函数, 采用 PyTorch 框架, 使用 1 块 GTX 1080 Ti GPU 进行训练, 整个训练时间 15 min.

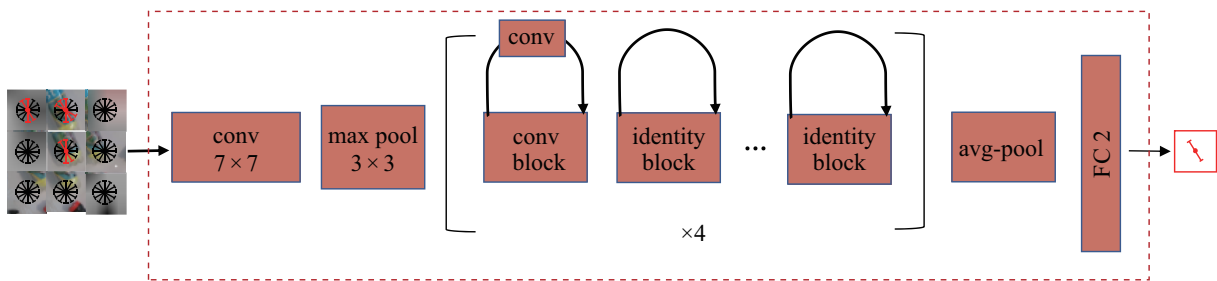


图 7 触觉先验知识学习模型

Fig.7 The learning model for prior tactile knowledge

3 触觉先验抓取规划 (Grasp planning based on prior tactile knowledge)

3.1 抓取规划问题描述

抓取规划阶段以彩色图像作为输入, 目的是基于学到的触觉先验知识从图像空间找到满足稳定抓取指标的抓取构型. 当前, 基于概率预测的抓取区域^[19]与包含物体位姿特征^[20]的抓取矩形是主要的构型方式. 受其启发, 本文主要考虑刚体二指手抓取, 提出了相应的抓取构型表示. 刚体二指手由电机控制, 可张开不同的宽度, 并可随着机械臂运动至不同的位姿, 所以一个完备的二指手抓取构型包括抓取位置、抓取角度、抓取宽度和抓取深度. 如图 8 所示, 平面抓取表示为垂直于 x - y 桌面的 $\mathbf{g} = (\mathbf{p}, \theta, q)$, 其中 \mathbf{p} 表示二指手中心点的位置 (x, y, z) , θ 表示二指手绕 z 轴旋转的角度, w 表示抓取宽度, q 表示抓取质量, 即本文提到的抓取稳定性. 对于视觉图像空间 $\mathbf{I} = (\tilde{\mathbf{p}}, \tilde{\theta}, \tilde{w})$, 其中, $\tilde{\mathbf{p}} = (u, v, d)$ 代表图像空间内目标物体的像素点以及深度值, $\tilde{\theta}$ 是图像空间中目标物体相对于摄像头坐标系的旋转角度, \tilde{w} 是图像空间内目标物体的像素宽度. 对于触觉空间 $\mathbf{T} = (r, c)$, 根据上文提到的触觉信息处理方式, 将压力值转换为触觉图像, 其中 (r, c) 为触觉图像空间的像素点坐标.

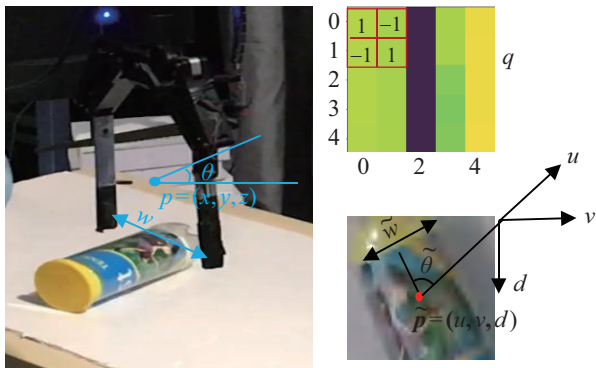


图 8 二指手抓取构型

Fig.8 Grasp configuration of 2-finger manipulator

综上, 在基于触觉先验信息的二指手平面抓取中, 可通过下式将视觉空间信息转化为世界坐标系下机械臂运动规划信息:

$$\mathbf{g} = ({}^R\mathbf{T}_c \mathbf{f}(\mathbf{I}), q(\mathbf{T}))$$

其中, ${}^R\mathbf{T}_c$ 表示摄像头坐标系到机器人坐标系的坐标转换, 可由摄像头外参标定结果得到; $\mathbf{f}(\mathbf{I})$ 代表从 2D 图像坐标系到 3D 摄像头坐标系的变换矩阵, 可由摄像头内参标定结果得到; $q(\mathbf{T})$ 表示利用式 (1) 计算出的抓取质量.

3.2 触觉先验抓取规划模型

由于本文主要考虑家用常规物品的平面抓取, 所抓取的物体具有不易碎的性质, 因此将抓取深度统一设置为桌面高度, 抓取宽度统一设置为较小固定值以简化抓取构型变量, 即抓取构型为 $\mathbf{g} = (\mathbf{p}, \theta, q)$.

基于上文的触觉先验知识学习方法, 为实现从图像空间到稳定抓取指令的直接映射, 本文提出触觉先验抓取规划模型 (图 9), 包括抓取区域检测和抓取构型生成 2 个阶段. 抓取区域检测以 400×300 的彩色图像作为输入, 采用 80×78 的步长, 100×100 的提取框滑动整个图像, 生成 20 个 100×100 的候选区域, 以 (u_i, v_i, η_i) 表示, 其中 (u_i, v_i) 表示中心点的像素坐标, 表示该像素空间中物体像素与背景像素的比值, 最大比值区域记为抓取候选区.

针对该候选区, 为了更好地实现稳定抓取构型回归, 在 100×100 的抓取区域中重新进行均匀采样 9 个点, 并且以它们为中心重新获取 9 个 100×100 大小的候选区. 每个候选区以 30° 的角度差值围绕各自的中心点旋转 6 次, 最终得到 54 个包含抓取点与抓取角度的候选抓取构型集合, 以 $\mathbf{g}_i = (\mathbf{p}_i, \theta_i, q_i)$ 表示. 然后, 将它们输入抓取构型生成器, 即上文的触觉先验知识学习模型, 输出得到稳定的抓取构型.

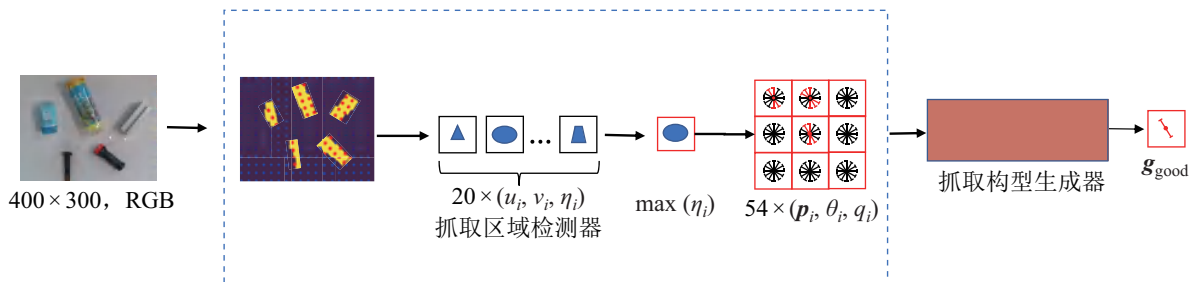


图 9 基于触觉先验知识的抓取规划模型

Fig.9 The grasp planning model based on prior tactile knowledge

4 实验结果与分析 (Experiment results and analysis)

为验证本文提出的抓取质量评估方法与基于触觉先验知识的机器人稳定抓取规划的有效性, 将其与现有的触觉抓取评估方法进行对比, 并在自建数据集上对比分析了触觉先验抓取与视觉抓取的抓取效果, 同时利用实际机器人进行了二指手抓取, 验证本方法的有效性.

4.1 触觉抓取质量评估对比

Hogan 等^[5]提出了不依赖特定触觉传感器的抓取质量评估方法. 该方法主要依赖于机器人抓取过程中的外部抖动, 并定义在抖动时间内物体不跌落为稳定抓取, 物体跌落则根据时间先后顺序评估抓取质量.

基于本文的评估方法与 Hogan 提出的方法, 进行了 50 次抓取实验, 并评估抓取质量. 将抓取质量按照从低到高顺序排列, 比较同一次抓取的不同质量分数, 如图 10 所示. 其中, 红色折线为 Hogan 方法, 蓝色折线为本文提出的评估方法. 可以发现, 2 种方法具有较好的一致性, 本文提出的方法在物体滑动检测方面有更好的评估表现.

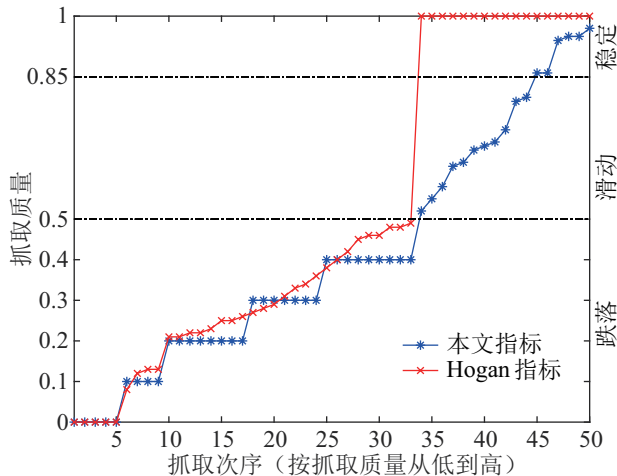


图 10 触觉抓取质量评估对比

Fig.10 Comparison of tactile grasp quality metrics

4.2 触觉先验知识学习方法的有效性

针对由 3000 个样本组成的 VPT 数据集, 按 3:1 的比例将其分为训练集与测试集, 在测试集上得到的预测准确性为 86%. 为了验证本方法的泛化效果, 选取 5 种新物体通过训练好的触觉先验知识学习模型预测其抓取构型, 其结果如表 1 所示. 结果表明, 本方法具有一定的泛化性, 在跨模态学习的基础上, 实现了共享特征的迁移, 并可以在更广的模态范围推广.

然而, 不同物体之间的准确率差异表明, 形状和重量的多样性对模型泛化提出了较大的挑战.

表 1 触觉先验知识学习的泛化效果

Tab.1 Generalization of prior tactile knowledge learning

| 抓取物体 | 准确率 |
|---------------|-----|
| 乐事薯片盒 | 66% |
| RealSense 包装盒 | 82% |
| 男士剃须泡沫瓶 | 85% |
| 光明瓶装牛奶 | 77% |
| 工具刀 | 83% |

4.3 机器人抓取实验

本文以 6 自由度 UR5 机械臂、RealSense 系列的 D435 摄像头、TakkTile 生产的 TakkStrip 触觉传感器以及 Robotiq 二指手搭建机器人抓取系统. 其中, D435 摄像头固定在距桌面 800 mm 高的支架上, 视角向下. TakkStrip 触觉传感器是由 5 个相距 8 mm 的压力单元组成的横向阵列, 可同时测量 5 个接触点的正向压力值, 安装在二指手的内侧. 机器人硬件系统如图 11 所示.

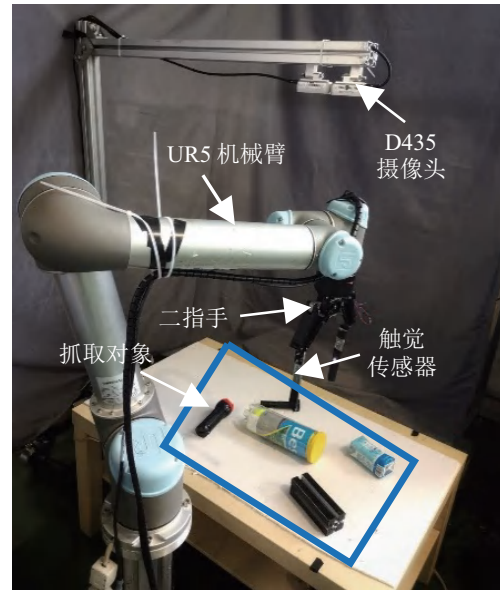


图 11 机器人硬件系统

Fig.11 Hardware of the robot system

为了验证触觉先验抓取规划方法, 本文将其与基于视觉的机器人抓取方法进行对比试验. 后者主要是在抓取构型生成阶段不利用触觉信息, 而是基于单纯的视觉信息判断抓取是否稳定, 即将物体不跌落视为稳定.

对比 2 种抓取方式, 触觉先验抓取的抓取位置、分布与视觉抓取明显不同. 基于学到的触觉先验知识, 机器人在抓取构型生成过程中可以有效剔

除会引起物体跌落以及滑动的候选构型, 从而实现稳定抓取. 而基于视觉的抓取方法只能判断物体

是否跌落, 难以有效避免物体滑动, 从而难以实现较高的抓取稳定性.

表 2 抓取实验结果
Tab.2 Results of grasp experiments

| 物体名称 | 已知物体 | | | | | 未知物体 | | | | |
|-----------------|---|---|---|---|---|---|---|---|---|---|
| | 螺钉 | 酸奶盒 | 网球筒 | 手电筒 | 铝型材 | 乐事薯片盒 | RealSense 包装盒 | 男士剃须泡沫瓶 | 光明牛奶 | 工具刀 |
| 质量 /g | 292.52 | 211.72 | 221.10 | 76.52 | 222.62 | 128.48 | 130.90 | 134.36 | 304.29 | 71.25 |
| 轴向长度 /mm | 112 | 100 | 210 | 131 | 130 | 210 | 145 | 130 | 142 | 130 |
| 外形 |  |  |  |  |  |  |  |  |  |  |
| 视觉稳定 抓取成功率 | 34% | 64% | 52% | 70% | 42% | 62% | 54% | 66% | 32% | 70% |
| 触觉先验稳定 抓取成功率 | 90% | 92% | 76% | 96% | 88% | 70% | 86% | 88% | 76% | 94% |
| 相对提高 | 165% | 44% | 46% | 37% | 110% | 13% | 59% | 33% | 138% | 34% |

本文定义稳定抓取为在添加外界扰动行为后, 目标物体不发生跌落与滑动, 保持相对稳定, 即根据抓取质量计算得到的 $q \geq 0.85$. VPT 数据集的 5 种已知物体与另外 5 种未知常见家用物品被用于实验验证, 2 种方法的稳定抓取成功率如表 2 所示. 可以发现, 与传统视觉方法相比, 引入触觉先验知识后成功率提高了 31%.

另外, 从表中不同物体的稳定抓取成功率分析, 质量较轻、体积较大的物体相对于质量较重、体积较小的物体的抓取成功率更高, 这是因为前者对抓取点与抓取角度的分类结果不敏感, 对于误差有较好的容错性. 如图 12 所示, 图 12(a) 的抓取点选择并不是最佳, 但是由于物体质量较小, 手爪与物体间的摩擦力足够抵消重力影响, 因此较易实现稳定抓取. 而图 12(c) 中的螺钉, 则只有在精准的抓取构型下, 才能不受抓取过程中外界扰动行为的影响.

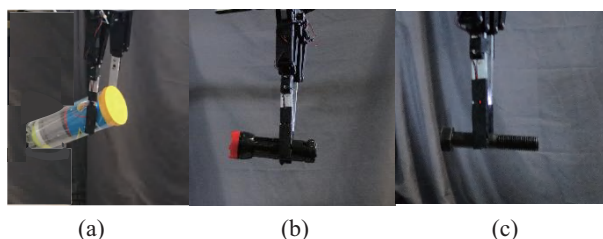


图 12 不同物体的抓取表现

Fig.12 Grasp performances for different objects

除此之外, 可以发现, 在 2 种不同的抓取方法下, 对于质量较大的物体, 其抓取准确率的提升幅度更大, 这是由于触觉先验规划相对于传统的视觉

方法, 借助学习到的触觉先验知识, 可根据当前视觉图像推理生成准确的稳定抓取构型, 从而避免了跌落与相对滑动, 从侧面反映了引入触觉先验知识对机器人稳定抓取的重要性.

5 结论 (Conclusion)

本文提出了一种新的抓取质量评估方法, 设计了针对二指手的完备抓取构型, 并采用触觉先验抓取规划提高了机器人稳定抓取的成功率. 通过自建数据库验证了抓取质量评估方法的有效性与触觉先验知识学习的泛化效果. 最后, 通过机器人抓取实验验证了本方法相对于传统视觉抓取法可以有效提高稳定抓取的成功率. 试验结果显示, 本文提出的触觉先验抓取法相对于传统的视觉抓取法, 成功率提高了 55%, 避免了抓取过程中的滑动和跌落, 并对新物体有较好的泛化效果.

由于触觉信息是动态且连续的, 前后时序具有较大的关联性, 未来计划考虑更多的触觉时序特征, 以实现更复杂的长时序灵巧操作.

参考文献 (References)

- [1] 李宇飞, 高朝辉, 申麟. 基于视觉的机械臂空间目标抓取策略研究 [J]. 中国科学: 技术科学, 2015, 45(1): 31-35. Li Y F, Gao Z H, Shen L. Study of vision-based space target capturing strategy for manipulators[J]. Scientia Sinica: Technologica, 2015, 45(1): 31-35.
- [2] Saxena A, Driemeyer J, Ng A Y. Robotic grasping of novel objects using vision[J]. International Journal of Robotics Research, 2008, 27(2): 157-173.
- [3] Calandra R, Owens A, Upadhyaya M, et al. The feeling of success: Does touch sensing help predict grasp outcomes[DB/OL]. (2017-10-16) [2020-02-01]. <https://arxiv.org/abs/1710.05512>.

- [4] Dang H, Allen P K. Stable grasping under pose uncertainty using tactile feedback[J]. *Autonomous Robots*, 2014, 36(4): 309-330.
- [5] Hogan F R, Bauza M, Canal O, et al. Tactile regrasp: Grasp adjustments via simulated tactile transformations[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2018: 2963-2970.
- [6] Howe R D, Cutkosky M R. Sensing skin acceleration for slip and texture perception[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 1989: 145-150.
- [7] Melchiorri C. Slip detection and control using tactile and force sensors[J]. *IEEE/ASME Transactions on Mechatronics*, 2000, 5(3): 235-243.
- [8] Su Z, Hausman K, Chebotar Y, et al. Force estimation and slip detection/classification for grip control using a biomimetic tactile sensor[C]//IEEE-RAS International Conference on Humanoid Robots. Piscataway, USA: IEEE, 2015: 297-303.
- [9] Chebotar Y, Hausman K, Su Z, et al. Self-supervised regrasp-ing using spatio-temporal tactile features and reinforcement learning[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2016: 1960-1966.
- [10] James J W, Pestell N, Lepora N F. Slip detection with a biomimetic tactile sensor[J]. *IEEE Robotics and Automation Letters*, 2018, 3(4): 3340-3346.
- [11] Yuan W, Li R, Srinivasan M A, et al. Measurement of shear and slip with a GelSight tactile sensor[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2015: 304-311.
- [12] Hyttinen E, Kragic D, Detry R. Estimating tactile data for adaptive grasping of novel objects[C]//IEEE-RAS International Conference on Humanoid Robotics. Piscataway, USA: IEEE, 2017: 643-648.
- [13] Guo D, Sun F C, Liu H P, et al. A hybrid deep architecture for robotic grasp detection[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2017: 1609-1614.
- [14] Goodale M A. Vision and action: The control of grasping[M]. Bristol, UK: Intellect Books, 1990.
- [15] Chebotar Y, Hausman K, Su Z, et al. BiGS: BioTac grasp stability dataset[C]//ICRA 2016 Workshop on Grasping and Manipulation Datasets. 2016.
- [16] Luo S, Yuan W, Adelson E, et al. ViTac: Feature sharing between vision and tactile sensing for cloth texture recognition[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2018: 2722-2727.
- [17] Wang T, Yang C, Kirchner F, et al. Multimodal grasp data set: A novel visual-tactile data set for robotic manipulation[J]. *International Journal of Advanced Robotic Systems*, 2019, 16(1). DOI: 10.1177/1729881418821571.
- [18] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2016: 770-778.
- [19] Zeng A, Song S, Yu K T, et al. Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2018: 3750-3757.
- [20] Jiang Y, Moseson S, Saxena A. Efficient grasping from RGB-D images: Learning using a new rectangle representation[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2011: 3304-3311.

作者简介:

薛 腾 (1994—), 男, 硕士生. 研究领域: 机器人灵巧操作, 多模态融合.

刘文海 (1991—), 男, 博士生. 研究领域: 机器人抓取, 深度学习.

潘震宇 (1994—), 男, 博士生. 研究领域: 人机协作, 力位控制.