MUNICH INSTITUTE OF ROBOTICS AND MACHINE INTELLIGENCE

TECHNICAL UNIVERSITY OF MUNICH

Report Submitted to Seminar: Optimal Control and Reinforcement Learning for Robotics

Examining prevalent Challenges for Contact-Rich Manipulation

D. Zipperstein, J. A. Sol Navarro, L. Oehler, M. Abdelhamid, Y. Lin

MUNICH INSTITUTE OF ROBOTICS AND MACHINE INTELLIGENCE

TECHNICAL UNIVERSITY OF MUNICH

Report Submitted to Seminar: Optimal Control and Reinforcement Learning for Robotics

Examining prevalent Challenges for Contact-Rich Manipulation

Untersuchung vorherrschender Herausforderungen für kontaktreiche Manipulationen

Author:	D. Zipperstein, J. A. Sol Navarro, L. Oehler, M. Abdelhamid, Y. Lin
Lectuers:	Dr. Fan Wu, Xiao Chen, Prof. Dr.-Ing. Sami Haddadin
Submission Date:	January 31, 2024

We confirm that this report is our own work and we have documented all sources and material used.

Munich, January 31, 2024
D. Zipperstein, J. A. Sol Navarro, L. Oehler, M. Abdelhamid, Y. Lin

Abstract

In the rapidly evolving field of robotics, contact-rich manipulation tasks have emerged as a critical part of many real-world applications. Traditional robot control methods often struggle to handle these tasks due to their complexity and uncertainty. However, optimal control and reinforcement learning have shown their potential for improving the performance and enhancing the efficiency of these tasks. This survey first gives a concise overview of the fundamental concepts and principles of contact-rich manipulation, reinforcement learning, and optimal control. In addition, this paper primarily focuses on several challenges that are prevalent in this area, including object perception and estimation, generalization, reality gap, efficient data representation and safety. Each of these challenges poses distinct obstacles that need to be carefully addressed for the effective and successful implementation of contact-rich manipulation tasks. The primary objectives of the present survey are threefold: (1) to identify the main challenges of contact-rich manipulation tasks in the context of robot control, (2) to provide some state-of-the-art approaches to face these challenges, (3) to present a brief conclusion of contact-rich manipulation task challenges and potential solutions.

Contents

1 Introduction

1.1 Background

As of today, robots already perform various manipulation tasks autonomously, demonstrating the rapid advancement in the field of contact-rich manipulation. Common tasks include pick and place operations, assembly line tasks, material handling, packaging and sorting, welding and cutting, medical surgery, and food industry tasks, among others.

Theoretical methods, particularly in preprogrammed factory robotics, involve predefined motion paths, feedback control systems, kinematic models, simulation-based programming, and vision systems. These methods have been effective for repetitive tasks in controlled, structured environments, offering precision and efficiency.

However, recent advancements in reinforcement learning and optimal control present new possibilities for robotic manipulation. These approaches enable robots to learn and adapt, enhancing their capabilities to handle varying conditions and contributing to the evolution of contact-rich manipulation.

1.2 Motivation

Successfully achieving high-level manipulation tasks aided by reinforcement learning and optimal control offers a large number of benefits. It enables robots to adapt to dynamic environments, handle diverse tasks and objects with versatility, reduce programming effort through learning-based approaches, and efficiently solve problems through trial and error.

The integration of Reinforcement Learning and optimal control enhances robotic dexterity, precision, and adaptability, enabling seamless collaboration with humans. This approach contributes to cost-efficiency and scalability, as robots can be deployed across various applications. Additionally, it empowers robots with autonomy, continuous improvement, and real-world applicability, making them valuable assets in industries such as manufacturing, healthcare, and logistics.

2 Preliminaries

2.1 Contact-Rich Manipulation

Elguea-Aguinaco et al. [1] define a contact-rich manipulation task as “any task that involves close interaction between the robot and its environment and comprises complex, high-dimensional and even nonlinear contact dynamics“. Contact-rich robotic manipulations are used in various industries such as manufacturing and healthcare, for tasks involving precise handling of objects, delicate materials, and complex environments [2].

A typical contact-rich manipulation task of robots follows a well-defined pipeline that encompasses various stages, which are depicted in Figure 2.1. The initial phase involves sensors gathering essential data, providing the robot with a foundation for interaction with its surroundings. Subsequently, the robot’s software stack comes into play, where the robot has to do perception first. This includes data processing and object recognition. Further, the robot engages in modeling the object that it wants to interact with. Hereby it has to estimate different properties such as pose and geometry as well as dynamic properties like the center of mass and inertia. This is followed by the planning phase, where the robot has to plan the path to reach the target object and its interaction with it, like grasping or pushing. Afterwards the actuators, like motors, pneumatic or hydraulic actuators need to be controlled. This is usually done by closed-loop control algorithms, like PID control. Lastly, safety is of crucial importance throughout the entire process to avert potentially harmful situations. Furthermore, the robot needs to be able to adapt to different environments and scenarios. This can be either achieved by considering system dynamics and constraints or by learning from real-world scenarios or simulations. These methods are summarized under the terms ’Optimal control’ and ’Reinforcement learning’, which will be introduced in the following chapters.

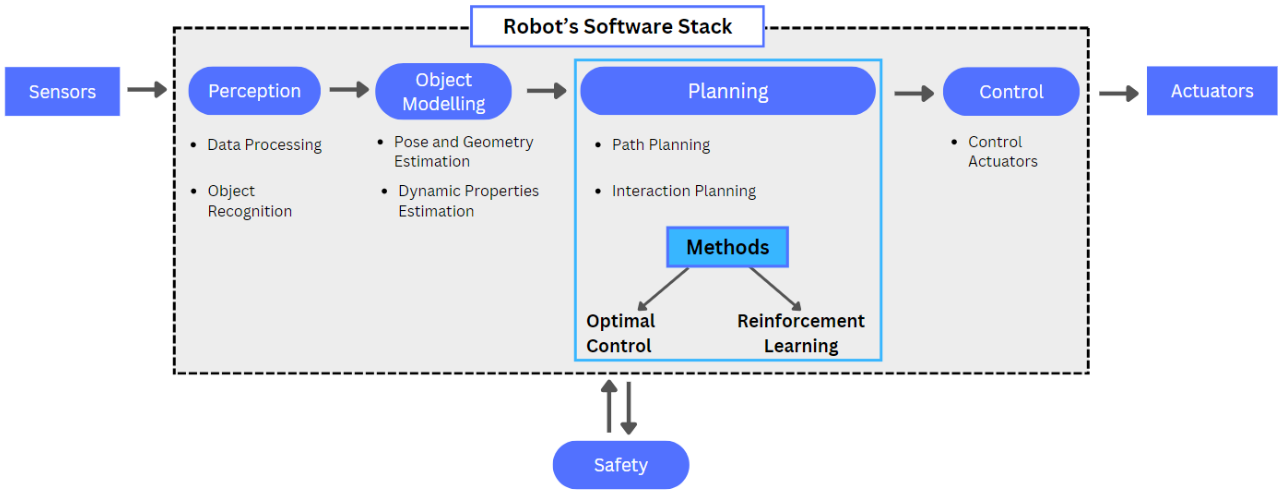

Figure 2.1: Pipeline for Contact-Rich Manipulation Tasks

2.2 Reinforcement Learning

Reinforcement Learning (RL) is a subfield of machine learning that focuses on developing intelligent agents capable of learning optimal behaviors by interacting with an environment. In RL, an agent perceives the environment's states, takes actions, receives rewards as feedback, and adjusts its strategy to improve long-term performance. The agent's goal is to learn a policy, which is a mapping from states to actions that maximizes the expected cumulative reward over time. Q-learning is one of several algorithms in RL used to learn a policy. Q-learning estimates the value of state-action pairs, known as Q-values, to guide the agent's decision-making. The Q-value of a state-action pair represents the expected cumulative reward when taking that action in a particular state and following the optimal policy. The policy update rule Equation 2.1, follows the Bellman equation, is a key aspect of Q-learning.

$$q_\pi(s,a) = \sum_{s',r} p(s',r|s,a)[r + \gamma \sum_{a'} \pi(a'|s')q_\pi(s',a')] \tag{2.1}$$

Where $\sum_{s',r} p(s',r|s,a)$ is the expected policy with $p(s',r|s,a)$ as the probability of getting to state s' given action a, r is the reward for state s' and $\sum_{a'} \pi(a'|s')q_\pi(s',a')$ are the future rewards following the policy discounted with γ. This update rule iteratively refines the Q-values, converging towards an optimal policy that provides optimal actions for each state, leading to optimal behavior of the agent [3].

2.3 Optimal Control

Optimal Control is a field of control theory that aims to determine the most effective way to manipulate a system to achieve a desired outcome. It addresses problems where a control strategy must be designed to minimize a specific objective, subject to constraints and the system's dynamics. To achieve this, a cost function is formulated to quantify the system's performance over a given time horizon. The cost function for discrete-time is usually defined as in Equation 2.2.

$$J = \sum_{k=0}^{N-1} f(x_k, u_k) + \beta h(x_N) \tag{2.2}$$

With $f(x_k, u_k)$ being the instantaneous cost at time k and $h(x_N)$ being the terminal cost at the final time step N with weighting factor β. The goal is to minimize the value of J for each control input u_k, while adhering to the system dynamics, additional constraints and initial conditions (constraints omitted in Equation 2.2). This approach optimizes the system towards the given goal for the next N time steps [4].

3 Challenges

3.1 Structure

Today, numerous challenges persist in advancing contact-rich manipulation to its full potential. In this report, five major challenges have been selected and analyzed in the following chapters. How these challenges fit into the robot's pipeline can be seen in Figure 3.1.
The researched challenges are:

- 3.2 Object perception and estimation
- 3.3 Generalization
- 3.4 Reality-Gap
- 3.5 Data efficiency
- 3.6 Safety

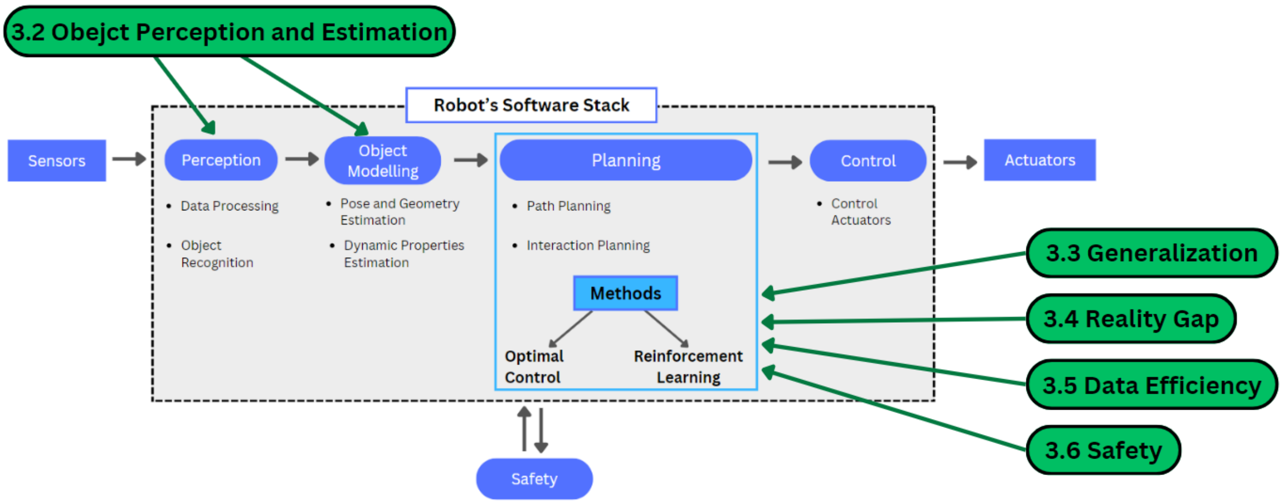

Figure 3.1: Overview of the researched Challenges

3.2 Object perception and estimation

The Challenge

Object perception and estimation play a crucial role in enabling contact-rich manipulations. As depicted in Figure 2.1 it is the first and second stage of the robot's software stack. These fields involve the ability of machines to interpret and understand the surrounding environment through sensory input. Object perception focuses on recognizing and categorizing objects within an image or a scene, while object estimation involves determining the spatial and quantitative properties of these recognized objects. Integrating these capabilities, enables robots to interact with objects with varying shapes, textures, and sizes and dynamic properties.

Due to the predominantly passive sensor deployment in robots nowadays, many challenges are encountered in this domain. One major challenge is object perception in cluttered or occluded environments, where robots must navigate past obstacles or where objects are not fully visible. Another substantial challenge centers around interacting with unknown object properties or dynamics. By not knowing an object's properties the planning of the interaction with it, is very complicated. In real-world scenarios, robots often encounter objects with diverse shapes, materials, and physical properties. Dealing with perception uncertainty and noise is another big challenge. The inherent unpredictability in the environment, coupled with sensor inaccuracies and external disturbances, requires the implementation of robust control strategies and state estimation techniques to ensure precise and reliable manipulation outcomes. Furthermore, coping with high data requirements emerges as a pressing challenge in this domain. Today's perception algorithms need a lot of data to be able to successfully percept. Hence, also high computational power and storage is required, which needs to be provided by expensive hardware.

Many of these challenges can be overcome by active utilization of sensing modalities. These methods can be divided into Active Perception (AP) and Interactive Perception (IP).

Active Perception

Active Perception involves engaging in activities aimed at controlling the sensory apparatus's geometric parameters to enhance perceptual quality. It does that without physical interaction with the environment. An active perceiver, in this context, comprehends the purpose behind sensing, selects what to perceive, and determines how, when, and where to achieve that perception. Methods encompass mechanical alignment, sensor alignment, priming, viewpoint selection, temporal selection, and scene selection [5].

One prominent field of AP is active vision. In Active Vision, the system doesn't just passively receive visual information but actively selects where to look, zoom in, or adjust its viewpoint to gather relevant data by actively controlling sensors, like cameras or microphones. This approach was affirmed by Le Q. and A. [6], who analyzed probabilities of correctly identifying objects based on different viewing angles. They could show that for some angles the probability was much higher than for others. This was due to the over-representation of some object orientations in the training datasets. These orientations are also more common to humans and are called canonical. By implementing an AP approach they could confirm their hypothesis and substantially improve the object detection performance [6].
One prominent more recent active object detection approach that uses machine learning, was

introduced by Fang et al. [7]. Here a framework for more efficient and robust object detection was presented. It was deployed on a system, where a camera attached to a robot arm was used to find the best position and angle that maximizes the object recognition probability, as seen in Figure 3.2. Here self-supervised reinforcement learning was used to learn the action type (direction of the movement) and the action range (step size) to find optimal trajectories. This lead to enhanced detection accuracy, faster convergence, and more robust performance, especially in complex environments with densely sampled viewpoints. Furthermore, the required data for successful object detection was reduced drastically [7].

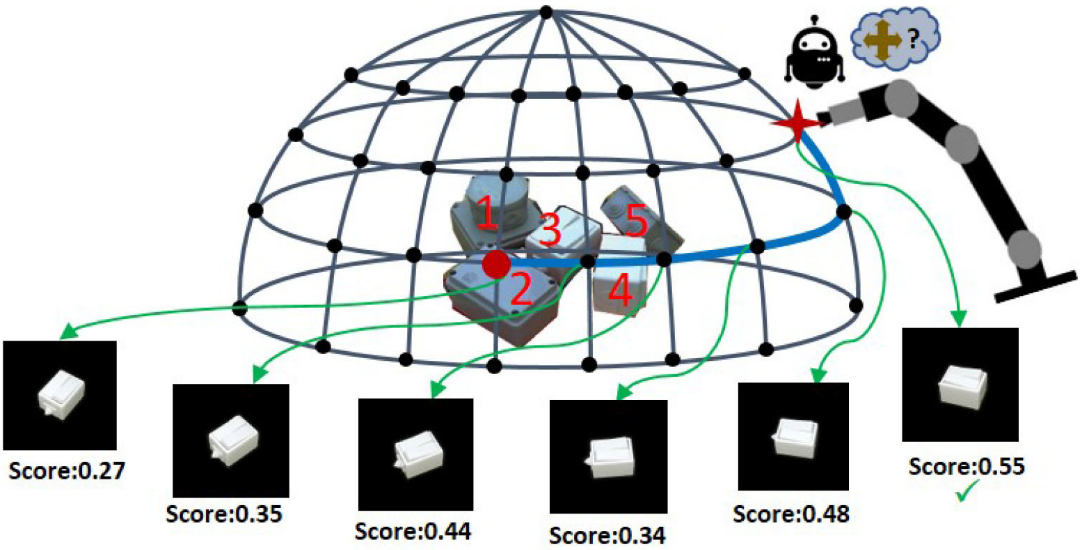

Figure 3.2: Example of Active vision [7]

Interactive Perception

Interactive Perception, on the other hand, utilizes physical interaction with the environment to simplify and enhance perception, which leads to robust and perceptually guided manipulation behavior. Two main benefits arise from this. Firstly, interactions with the environment create novel and rich sensory signals that would otherwise not be present. Secondly, knowledge of the regularity in the combined space of sensory data and action parameters facilitates the prediction and interpretation of the sensory signal [8].

Two main IP methods are utilized the most, namely **haptic exploration** and **auditory exploration**.
Haptic exploration involves the use of haptic or tactile sensing to explore and understand the physical characteristics of objects and environments. This sensory modality enables robots to gather detailed information about surfaces, textures, and shapes through physical interaction. Auditory exploration refers to the utilization of sound-based sensing mechanisms for robots to comprehend and navigate their environment. By leveraging microphones, robots can interpret auditory cues such as echoes, spatial location of sound sources and environmental noises. This sensory modality aids in object detection, localization and navigation, particularly in scenarios where visual information may be limited or obscured.

By deploying IP various perception and estimation subareas can be enhanced significantly. The main subareas in which IP is beneficial are presented in the following paragraphs.

Object Segmentation By haptic interaction with the object, real-time feedback is provided to refine and improve the segmentation results. The interaction induces scene motion and this provides an additional clue for associating observed parts of the same object. For example, Patten, Zillich, and Vincze [9] used a probabilistic segmentation approach, which was recursively updated after interaction with the objects to reduce segmentation uncertainty.

Object Recognition To detect object instances or objects of a specific category, robots have to learn the appearance or shape of these objects under various conditions, like occlusions, different lighting conditions, or the scale of the images. The state-of-the-art approaches in Computer Vision require enormous amounts of training data to handle these variations. IP approaches allow a robot to move objects and hence reveal previously hidden features. Thereby, it can resolve some of the challenges autonomously and may alleviate the need for enormous amounts of training data [8].
Improved object recognition can also be achieved by Auditory Exploration, which was done by Dou et al. [10]. They used an auditory exploration method based on reinforcement learning, which enables the robot to actively explore the operational behavior of interest and establish the coupling relationship between perception and action to reduce the ambiguity of target recognition. The robot interacts with visually indistinguishable bottles by adopting multiple action behaviors to generate sound data from which the perceptual model learns to classify object contents. This is similar to human beings actively exploring and accumulating experience through sound in an environment where vision cannot be judged [10]. This auditory exploration approach is depicted in Figure 3.3.

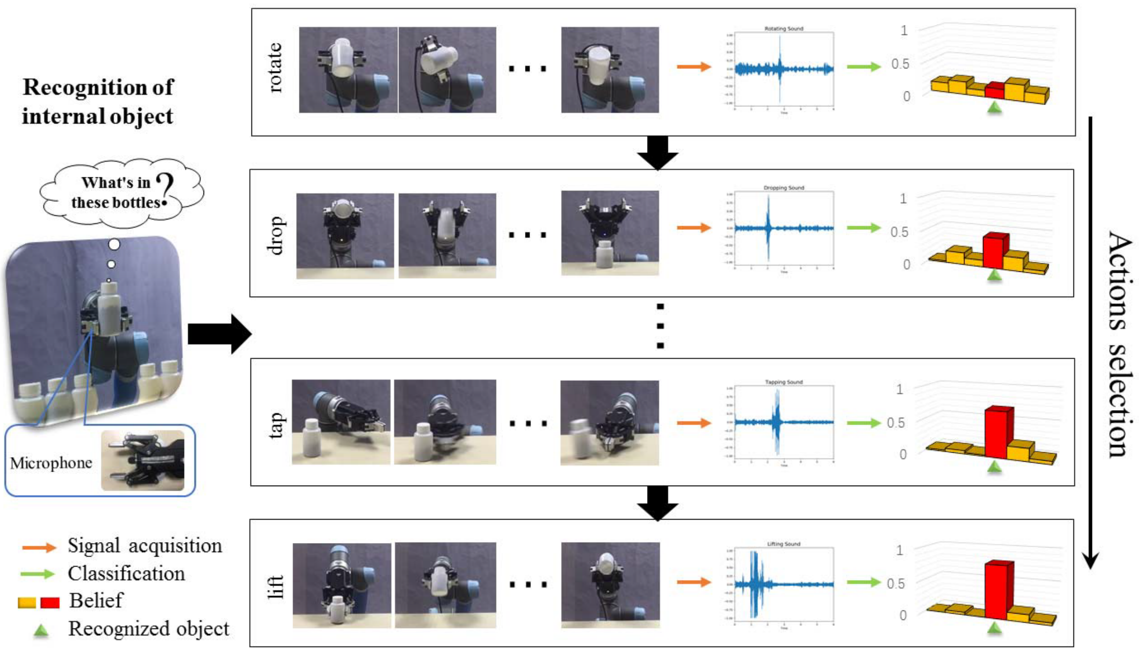

Figure 3.3: Example of Auditory Exploration for internal object recognition[10]

Object Dynamics and Haptic Properties Learning Estimating or learning objects' dynamics and haptic properties without interacting with the object can be quite difficult. Therefore haptic exploration needs to be used to gain more information by physical interaction. Ziyan, Elibol, and Chong [11] showed how this can be implemented. Here a robot arm executed pushes to an object while capturing the resulting movements of the object through an RGBD camera. The object itself could have various properties such as shapes, center-of-mass distributions, inertia values, or friction properties. The robotic system employed an encoder-decoder structured model featuring a cascaded residual attention mechanism. This model was trained on a simulation dataset, allowing it to integrate prior knowledge and make predictions about novel objects. It was shown that by applying pushes to novel objects the model could successfully predict their properties.

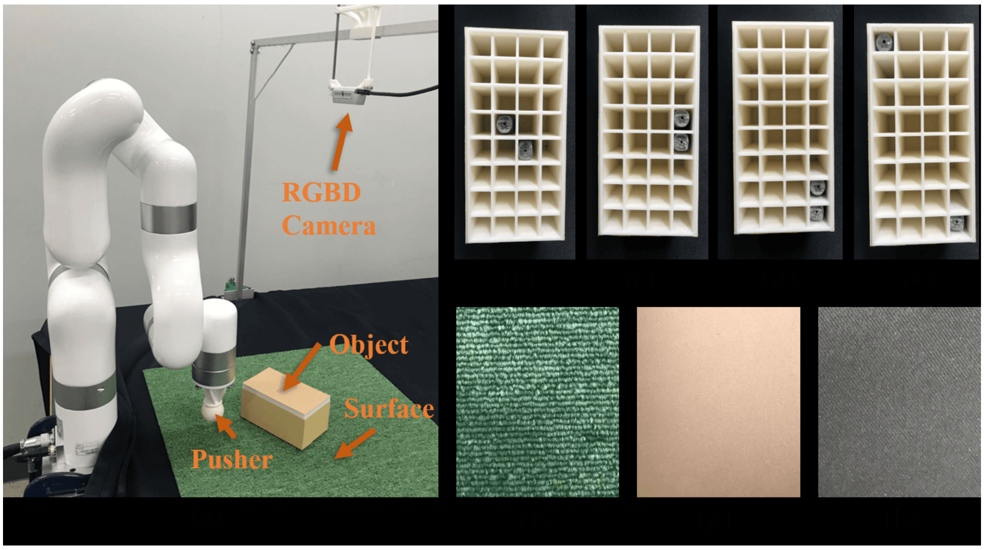

Figure 3.4: Example of Haptic Exploration for estimating object properties [11]

Articulation Estimation The estimation of object articulation mechanisms can also be simplified through the use of IP. The robot has to determine whether the relative movement of two objects is constrained or not and needs to understand the resulting physical limitations [8]. One IP approach to how this can be simplified was shown by Hausman et al. [12], where a particle filter-based approach was used to represent and actively reduce uncertainty over articulated motion models. The presented method integrated visual observations with feedback from manipulation actions to best characterize a distribution of possible articulation models.

Object Pose Estimation Object pose estimation can also be improved by IP. Here object pose uncertainty can be reduced by either touching or moving the object. Michael C. Koval and Srinivasa [13] showed an approach where manifold particle filters and tactile sensors were used to track the pose of a pushed object.

The presented methods promise to elevate the capabilities of robots in recognizing, understanding and interacting with objects. Future improvements will lead to improved efficiency and productivity and enhanced safety and reliability and enable new applications.

3.3 Generalization

The Challenge

Generalization in the context of Contact-Rich Manipulation is a critical aspect as it enables the use and adaption of learned skills from diverse past experiences and applies them effectively to new, unseen environments or tasks [1]. Contact-rich manipulation in robotics faces several challenges when it comes to achieving Skill Generalization. Among many others, some of the most prominent difficulties arise from :

- Diverse Environments: Contact-rich manipulation involves interacting with various objects in constantly changing environments. Generalizing across diverse conditions, such as different object shapes, materials, and environmental setups, can be challenging.
- Sensory Variability: Sensors provide information about the environment and the objects being manipulated. However, sensory data can vary widely due to differences in lighting, object appearance, or sensor noise.
- Uncertainty in Dynamics: Objects in the real world often have complex and uncertain dynamics. The requirement to adapt to different physical properties and dynamics of objects adds complexity to the manipulation task.
- Contact Forces: Manipulation involves physical interactions with objects, and the contact forces during these interactions can vary. Generalizing control policies to handle different force profiles and contact scenarios is a non-trivial task.
- Object Properties: Objects come in various shapes, sizes, and materials. Models and algorithms that can adapt to the diverse properties of objects encountered in the real world are required.
- Task Variability: Manipulation tasks can vary widely, from simple pick-and-place actions to more complex tasks requiring precise control and coordination.

By pursuing generalization in contact-rich manipulation, researchers and engineers aim to create robots that are not only capable of handling a wide range of tasks but can also adapt and learn in a manner that mimics human-like flexibility in interacting with the physical world. These motivations drive advancements in robotics to make robots more capable, efficient, and applicable in diverse real-world settings.

Motivation

Although robots outperform most humans in repeatable well-defined tasks, they fail to complete easy tasks autonomously in dynamic scenarios. For instance, in the agricultural field, the complexity increases when dealing with natural objects, such as fruits or leaves. This is due to the high variability of many of the parameters that affect robot behavior, many of which cannot be pre-determined[14]. Being able to generalize learned skills will help empower the robotics systems with the versatility required for a future real-world deployment.

On the other hand, a cost reduction during the learning phase can be expected. Collecting training data for every possible scenario is a time-consuming and intricate task. Generalization tackles this issue by allowing robots to learn with limited data, reducing the cost and time required for training. A general-purpose robot with robust generalization capabilities can be deployed across a range of applications, making it a more economical and scalable solution.

Furthermore, Generalization reduces the need to explicitly program every specific manipulation task. Instead of manually specifying how the robot should handle each object or situation, the robot can learn from experience and generalize its knowledge, reducing the programming effort required for new scenarios.

Approaches

Achieving Generalization in Contact-Rich Manipulation within robotics involves various approaches. Five methodologies and their core concepts for accomplishing Skill Generalization are introduced.

Reinforcement Learning RL approaches show great potential as they offer a way to solve control tasks in unstructured scenarios [15]. These methods allow the robotic system to learn through interaction with its environment and enable generalization by transferring the learned behavior to unseen scenarios.

The SOft Data Augmentation (SODA) approach, proposed in [16], enhances training stability by decoupling data augmentation from policy learning in reinforcement learning. SODA maximizes mutual information between latent representations of augmented and non-augmented data, promoting effective generalization and improved policy optimization.

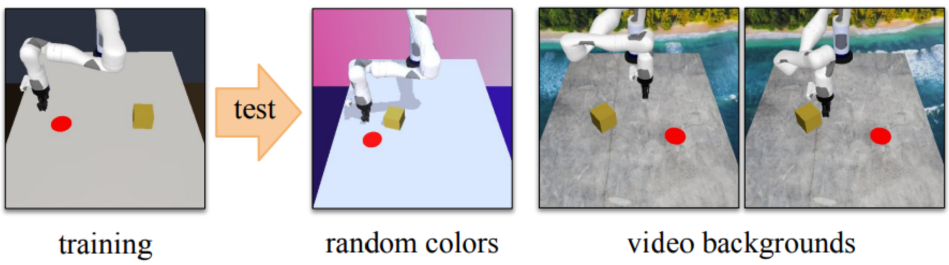

Figure 3.5: Testing Generalization in unseen environments with random colors and backgrounds [16]

Transfer Learning The robot first learns a skill in a source domain and transfers the acquired knowledge to a target domain, adapting to the new context. The learning process is improved by avoiding expensive data-labeling efforts. This approach leverages knowledge gained from one task to improve performance in a related, but different, task, enhancing Generalization.

Neural Networks and Function Approximation The main advantage of neural networks is that they enable robots to approximate complex functions, allowing them to generalize manipulation policies based on learned representations. Neural networks are employed to represent the mapping between sensory inputs and manipulation actions, providing a flexible framework for generalization.

In [17], the use of a neural network with an encoder-decoder architecture for learning variable impedance manipulation skills in robotic systems is indeed a noteworthy application. Variable impedance control is crucial for enabling robots to adapt to different tasks and environments by adjusting their stiffness or compliance.

Learning from Demonstrations Robots learn by imitating a teacher, commonly a human, acquiring a set of manipulation behaviors that can be applied to similar tasks. State-action pairs recovered from the teacher's demonstration are used by LfD algorithms to derive a policy that reproduces the demonstrated behavior [18]. In this method, the policy is acquired from experience, diverging from RL approaches where the policy relies on exploration.

In [19], Learning from Demonstration for contact-rich manipulation tasks is introduced. It recognizes that numerous demonstrations often lack sufficient data for a general task policy. To address this, information is augmented from a single demonstration autonomously, emphasizing the role of environmental constraints in generalization. Policies extracted from one augmented human demonstration generalize across similar mechanisms and diverse environmental setups. The workflow of the proposed approach is shown in Figure 3.6.

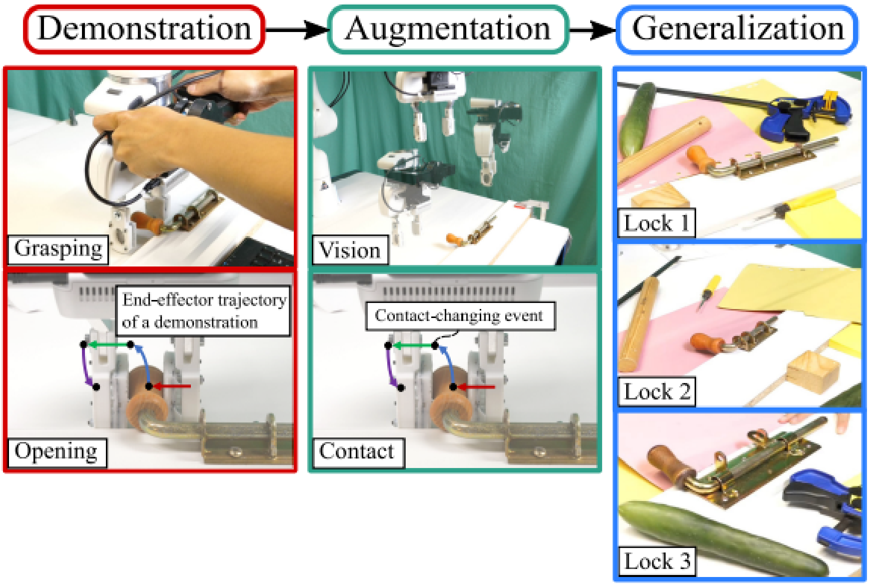

Figure 3.6: One-Shot Generalization achieved through LfD for Contact-Rich Manipulation [19]

Optimal Control In contact-rich manipulation tasks in robotics, optimal control is employed to achieve generalization by formulating and solving control problems that adapt to varying contact scenarios. Instead of relying on fixed trajectories, optimal control algorithms optimize control inputs based on the system's dynamics, contact forces, and task objectives. This allows robots to generalize their actions across different contact-rich scenarios, handling uncertainties and environmental variations. By optimizing control policies, the robot can adapt its behavior to achieve efficient and robust performance.

3.4 Reality-Gap

The Challenge

The development of autonomous robots capable of operating effectively in complex and dynamic real-world environments hinges on the ability to train them on massive data to learn complex behaviors. Traditional methods for training robots often involve lengthy and costly real-world experiments, which expose the robot to the risks of physical harm and damage. To address these limitations, simulation-based reinforcement learning (RL) has emerged as a promising alternative.

Simulation-based RL offers several compelling advantages over training them directly in the real world. It provides a safe, controlled, and efficient environment for robots to learn complicated tasks without causing damage or danger. Additionally, simulations can generate vast amounts of training data much more efficiently than real-world tests, enabling faster and more effective training for such robots.

Despite these significant advantages, transferring the learned policy from simulation to real-world robots presents a huge challenge known as the "reality gap." This gap arises from the inherent differences between the simulated and real-world environments. These discrepancies can include data fidelity inconsistencies, physics differences, sensor noise, lighting conditions and object properties, etc.

Motivation

Sim-to-real transfer, the process of transferring a reinforcement learning (RL) policy learned in a simulated environment to a real-world environment, holds immense promise for advancing the development of autonomous robots capable of operating effectively in real-world settings. By harnessing the advantages of simulation, which offers a safe, controlled, and efficient environment for learning, while bridging the gap to the real world, sim-to-real transfer has the potential to revolutionize the field of robotics.

While simulation provides a valuable training environment, it inherently differs from the real world. Discrepancies in physics simulations, sensor readings, and varied environmental factors can lead to sub-optimal performance or even failure when transferring the learned policy to the real world directly. Sim-to-real transfer aims to bridge this reality gap, enabling robots to adapt effectively to the real-world environment and achieve their desired behaviors.

Methodologies For Addressing Reality-Gap

Within the motivations mentioned above, we propose some methodologies for addressing the problem of sim-to-real transfer. In this section, this survey categorizes approaches into three broad categories and briefly introduces the principal idea of each one respectively.

Domain Randomization Methods Domain randomization is a technique that aims to reduce the sim-to-real gap by increasing the diversity and robustness of the synthetic data. The idea is to randomize the parameters of the simulation, such as textures, lighting, physics, and noise, to create a large set of varied scenarios that cover the possible variations in the real world.

This way, the model can learn to generalize to unseen situations and environments, without relying on the exact match between the simulation and the reality. Domain randomization can be seen as a form of data augmentation, where the goal is to increase the effective size and diversity of the training data. However, unlike conventional data augmentation methods, domain randomization does not require any prior knowledge or human intervention to generate realistic variations. Instead, it relies on the assumption that the simulation can capture the essential features of the real world and that the randomization ranges are sufficiently large to encompass the real-world variations. The principal idea of domain randomization is shown in Figure 3.7.

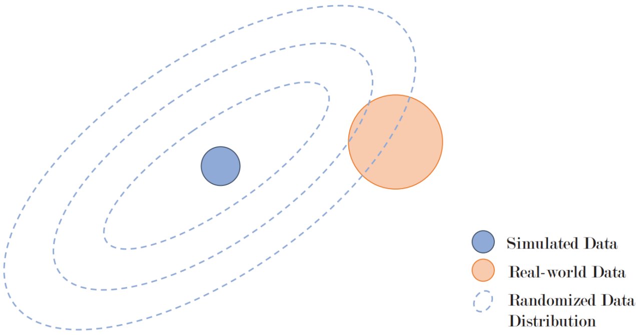

Figure 3.7: Principal idea of domain randomization methods [20]

There are two general types of domain randomization [20]: visual randomization and dynamics randomization. Visual randomization refers to randomizing the appearance of the simulation, such as the colors, textures, lighting, and camera parameters [21]. This can help the model to cope with different visual conditions and improve its perception skills. Dynamics randomization refers to randomizing the behavior of the simulation, such as the masses, inertias, frictions, and forces of the objects and the robot. This can help the model to cope with different physical conditions and improve its control skills.

Domain Adaptation Methods In the context of sim-to-real transfer in robotics, domain adaptation methods aim to improve the performance of a model on the target domain, namely the real world, by using a reinforcement learning policy learned in the source domain, namely the simulation. Given the ease of generating data in simulated environments compared to real-world environments, researchers employ the technique of training RL policies in simulators to reap the benefits of this approach. As shown in Figure 3.8, we represent source domain data and target domain data in high-dimensional feature space respectively, which encodes the data into a more convenient form for future feature extraction and learning. The goal of domain adaptation is to define a unified feature space that contains both the source domain features and the target domain feature. This unified feature space should be able to preserve the discriminative information of the data while minimizing the domain shift.

Domain adaptation methods can be classified into three main categories: supervised DA, semi-supervised DA and unsupervised DA [22].

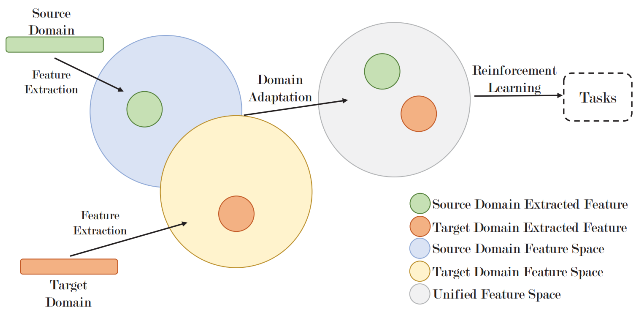

Figure 3.8: Principal idea of domain adaptation methods
[20]

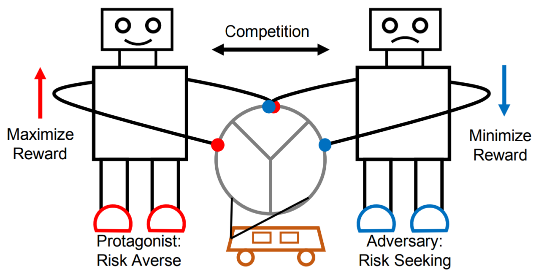

Figure 3.9: Principal idea of adversarial reinforcement learning
[24]

Adversarial Reinforcement Learning The adversarial reinforcement learning approach aims to train a more robust agent by exposing it to an adversary that applies perturbation or disturbance to the system. The idea is to create a competitive scenario between the protagonist, an agent trying to fulfill the task in order to maximize its reward, and the adversary, an agent trying to add extra forces or disturbance to the system in order to minimize the reward of the other one. This way, the agent can learn a more robust and stable policy, thereby enhancing the transferability of the reinforcement learning policy to the real world.

A recent paper proposes a novel framework in which a robust agent can learn an optimal destabilization policy by training along with an adversary applying random and systematic forces to the system[23]. Another paper introduces a framework called Risk Averse Robust Adversarial Reinforcement Learning that constructs an adversary that actively seeks catastrophic outcomes [24]. As shown in Figure 3.9, in an autonomous driving scenario, a risk-averse protagonist is trying to maximize the reward by driving the car safely, while a risk-seeking adversary is trying to minimize it by slowing or crashing the car. By modeling a stronger adversary which brings more risks and difficulties to the protagonist, a more robust policy will be learned by the protagonist.

3.5 Data efficiency

The Challenge

Contact-rich manipulation tasks, where the robot interacts with objects and the environment through contact, present unique challenges for data-efficient model learning and prediction in real-world scenarios. These tasks often involve complex dynamics, non-linear interactions, and vast state space, requiring large amounts of data to train accurate models [25]. However, acquiring such extensive data in real-world scenarios can be expensive, time-consuming, and impractical. This data scarcity poses a significant hurdle for developing robust and data-efficient manipulation algorithms [26].

- Data Scarcity: Acquiring enough data for training in contact-rich manipulation can be challenging due to the dynamic nature of these tasks and the variability of environmental conditions. Furthermore, the complexity of contact interactions and the large number of possible states increase the data requirements. This lack of data limits the applicability of data-driven approaches and hinders the development of effective manipulation algorithms.
- Explosion of Modes: Another challenge is due to the multitude of possible contact configurations and interaction patterns between the robot, objects, and the environment, resulting in an explosion of modes. This leads to an extensive state space, making it difficult to effectively learn and predict the dynamics. This often necessitates an impractically large amount of data for training.
- Irrelevant Information: In addition to data scarcity and the explosion of modes, contact-rich manipulation tasks often involve irrelevant or noisy information from the environment. This irrelevant data can hinder the learning process and lead to sub-optimal performance. For instance, sensor noise, background clutter, and irrelevant object features can contribute to over-fitting and reduce the generalization of the learned models.

Motivation

In recent years, robotics has made remarkable advancements, allowing robots to perform a variety of tasks, from assembling complex machinery to assisting humans in surgical procedures. However, data-driven algorithms face significant challenges in contact-rich manipulation, where robots interact with objects and the environment through physical contact, due to data scarcity, mode explosion, and irrelevant information. Data efficiency is crucial for contact-rich manipulation tasks for several reasons:

- Reduced Training Cost and Time: Data-efficient algorithms require less data to train, resulting in significant cost savings for data acquisition, labeling, and processing.
- Faster Task Learning: By reducing the need for extensive data collection, data-efficient algorithms enable faster task learning and the development of new manipulation skills.
- Improved Generalization: Data-efficient algorithms can generalize better to unseen situations due to their ability to learn from limited data.

Methodology

This section introduces methodologies for achieving data-efficient models in the field of contact-rich manipulation. The presented methodologies aim to overcome the challenges of Data Scarcity, Explosion of Modes, and Irrelevant Environmental Information, which are mentioned above.

Data Scarcity The data scarcity problem is addressed in [26] for contact-rich manipulation tasks by employing a hybrid automaton-based model structure and utilizing Gaussian processes for uncertainty representation.

- Hybrid automaton-based model structure: The hybrid automaton-based model structure captures the inherent discrete and continuous dynamics of contact-rich manipulation tasks. It decomposes the system's behavior into distinct modes, each representing a specific state of the system, such as idle, grasping, or pushing. This decomposition allows the model to efficiently represent the complex transitions between these modes, which are often governed by contact interactions.
- Gaussian processes: Gaussian processes are employed to represent the uncertainty associated with the model's predictions. This uncertainty quantification is crucial in contact-rich manipulation scenarios, where the robot's interactions with the environment can introduce significant variability. By explicitly modeling uncertainty, the robot can make informed decisions about its actions, taking into account the potential outcomes and their associated probabilities.

The automaton structure provides a compact representation of the system's dynamics, reducing the required training data. By combining the hybrid automaton-based model structure and Gaussian processes, efficient learning and prediction in low-data regimes can be achieved.

Explosion of Modes This challenge is addressed in [27] where concepts of local smoothing, sampling-based motion planning, and quasi-dynamic contact models are presented, to tackle problems like stiff, non-smooth contact dynamics, non-convexity of the planning problem, as well as the explosion of contact modes.

- Quasi-dynamic contact models: These models provide a simplified representation of contact dynamics that is less computationally demanding than full dynamic models. These models can tolerate some level of model uncertainty, making them more robust to changes in the environment and the robot's dynamics.
- Local smoothing: Smooth out the trajectories generated by the quasi-dynamic models using local smoothing techniques. This helps to reduce oscillations and improve the smoothness of the robot's movements. The smooth surrogate is used to provide a computationally efficient approximation of complex functions.
- Sampling-based motion planning: Sampling-based motion planning algorithms are used to search for feasible plans in the vicinity of locally smoothed trajectories. This approach allows the planner to take advantage of the smoothness of the trajectories while exploring the full range of possible contact configurations.

By utilizing these concepts, it is possible to reduce the number of discrete contact modes through the smoothing process. This makes it more computationally feasible to represent and plan for contacts. Additionally, the data required to learn the quasi-dynamic models can be significantly reduced, as the model can capture the essential dynamics of contact interactions without requiring detailed knowledge of all possible contact configurations. Combining these concepts, iterative MPC with smoothing can be used to plan and control complex manipulation tasks. iMPC can handle model uncertainty, improve trajectory smoothness, and adapt to unforeseen events. To achieve tractable global motion planning for highly contact-rich manipulation, combining smoothing with Rapidly-Exploring Random Tree (RRT) is very useful. In addition, utilizing a local Mahalanobis metric for RRT allows the algorithm to concentrate on state space regions that are more pertinent to the current state, thereby reducing search time.

Irrelevant Environmental Information Filtering only the relevant information from the environment is quite a challenge. By focusing on the information that is relevant to the robot's current task, these techniques can enable robots to operate more effectively and efficiently in real-world environments. In [28], this challenge is covered by introducing two useful concepts, which are task-specific environment representations and task-oriented object pose estimation.

Task-specific environment representations are tailored to the specific requirements of the robot's task. They are constructed by identifying and extracting the relevant features of the environment that are necessary for the robot to perform its task effectively. This eliminates the need to store and process unnecessary information, reducing the overall data requirements. In the context of object pose estimation, task-specific representations can be particularly beneficial.

For task-specific environment representations, both 2D occupancy grid maps (OGMs) and Real-time Octree-based 3D representations were used. This hybrid representation allows the robot to efficiently plan and navigate in real time while maintaining low memory consumption. On the other hand, Property-driven Pose Estimation using Fiducial Markers and Ontology Technology was used for efficient pose estimation.

- 2D OGM: The probabilistic 2-D occupancy grid is efficiently modeled using 2-D OGM, to accurately represent the geometric properties of the environment. It provides a quick and compact representation of the environment's topological structure, which enables the robot to quickly identify potential obstacles and plan paths.
- Real-time Octree-based 3D Representations: Used for safe manipulation tasks, where this representation is required for collision checking between the robot and the environment. An efficient octree-based representation method is applied to selectively construct a probabilistic 3-D occupancy grid with low memory consumption, using a pruning measure. The 3D representation is used for more detailed planning and navigation tasks, ensuring accuracy and robustness in complex environments.
- Property-driven Pose Estimation: This estimation is done based on fiducial markers and ontology technology. Fiducial markers are physical markers placed in the environment that can be detected and tracked by the robot. Ontology technology provides a structured representation of objects and their properties. By combining the detection of fiducial markers with ontology reasoning, the robot can infer and estimate the pose of task-related target objects. This enables the robot to perform task-oriented manipulation tasks more efficiently and accurately.

3.6 Safety

The Challenge

Safety in contact-rich manipulation tasks is concerned with avoiding physical harm to the environment, the robot or even humans, by keeping the robot in non-harmful states during the manipulation task [29]. Given a state representation, this may involve avoiding unwanted positions, joint angles, forces or contacts with the environment.

While robustness is sometimes regarded as a sub-field of safety, since it is concerned with reliability and accuracy across a range of conditions (see chapter Generalization), the definition of safety in this chapter is focused on ensuring the safe operation of a robot. The stability of the system in a control theory sense is not considered under the term safety here as well.

Safe robot actions guarantee a safe transition to a safe state or even ensure a safe time horizon of safe states further on in time [30]. Depending on the inclusiveness of the state representation, the safety of the manipulated object needs to be considered as well. That means that for a given time horizon, the manipulated object should not cause harm, e.g. preventing a pushed object from rolling towards an unsuspecting human.

Sources of safety concerns

The time it takes for the system to respond to changes, uncertainty about the environment and system states, complex contact dynamics, formal verification and behavior prediction are all sources of potential safety concerns.

If the system response time is too long, it may not be able to respond properly to the rapidly changing dynamic environment [31]. A faster sensor sampling rate and CPU processing time allow a higher frequency controller to overcome this problem. However, the fundamental issue of responding to a continuous-time system in discrete steps prevails.

Predicting an action that will lead to a particular safe state is challenging because it requires a very accurate model of the environment and robot dynamics. These real-world dynamics are usually highly nonlinear and discontinuous. For example, a gripper grasping an object or the on-off surface contact in a wiping task. This makes it difficult to accurately predict what control inputs are needed to bring the robot to a safe state [32].

The unknown and dynamic environment can only be perceived by noisy sensors. Together with unfulfilled model assumptions for the representation of the environment, this leads to a high uncertainty of the robot state and its map. Well-calibrated high-resolution sensors with redundancies, noise reduction techniques, active exploration (see Active Perception), enough training data [32] and good representations of the environment (e.g. a dynamic human behavior model [31]) reduce the uncertainty [33].

A good model of the robot's kinematics and dynamics is also needed to ensure that a given control input will lead the robot to a verifiable safe configuration. Compliance control, which uses forces instead of the position of the end-effector as a reference signal, can help to safely respond to external environmental forces by directly limiting the produced forces [33].

Responding safely to external disturbances or loss of functionality for sub-systems further complicates the safety issue and requires constant self-monitoring [29].

When a human is physically close to the robot, the system should also detect the human's intention and predict his movement to avoid collision [31]. For better communication with humans, the decisions made by the robot should be explainable [34]. An always-accessible robot interface ensures good controllability. An important precursor to safety.

Our research has identified four general approaches to increase safety: Implicit, Explicit, Retroactive, and Preemptive approaches. These are discussed in more detail below.

Implicit approach One way to get a reinforcement algorithm or optimal control method to adapt safe behavior is through reward shaping or risk weighting. Here, a safety penalty is introduced to force the model toward harm avoidance behavior. This safety penalty is added to weigh unsafe states and make them unlikely to be visited by the learned policy or control behavior. Equation 3.1 demonstrates such state weighting in reinforcement learning by adding the safety penalty to the bellman equation in Q-learning.

$$q_\pi(s,a) = \sum_{s',r} p(s',r|s,a)[r + \gamma \sum_{a'} \pi(a'|s') q_\pi(s',a') - \underbrace{F(s,s')}_{\text{safety penalty}}] \tag{3.1}$$

Equation 3.2 shows the cost function to be minimized in model predictive control with additional risk weighting for states and control inputs. With such an optimal control strategy, unsafe system behavior is discouraged.

$$J = \sum_{k=0}^{N-1} (f(x_k, u_k) + \alpha \underbrace{g(x_k, u_k)}_{\text{safety penalty}}) + \beta h(x_N) \tag{3.2}$$

The safety penalty should be proportional to how safety-critical the state is. For example, the paper by Mitsioni et al. (2023) [35] clustered the training set of known safe trajectories in a lower dimensional manifold. To evaluate the safety of a trajectory, it is first projected onto the lower-dimensional manifold. The Euclidean distance between the cluster center and the projected trajectory corresponds to the safety penalty for that trajectory [35]. Formally proving safety with such models, which are guided only by incentives for safe behavior, is challenging. However, the implicit approach is particularly useful when only safe examples are available, which is often the case in real-world environments where unsafe examples can break the system or cause serious harm.

Explicit approach If there are multiple examples of safe and unsafe states, a classifier can be trained to explicitly evaluate the safety of given states. If the explicit criteria for safety are known, this classifier can be a simple algorithm. Once it is possible to perform safety verification for states, a set of possible control inputs can be evaluated to see if they will bring the system to a safe state. Only this set of safe actions will later be considered by the controller to choose the optimal input for the robot [30] (see Figure 3.10).

Accurately predicting the state the system will end up in given a certain action requires a good model of the nonlinear system dynamics. Even with a formally proven save classifier for state safety, guaranteeing that a given action will result in the desired safe state remains a challenge.

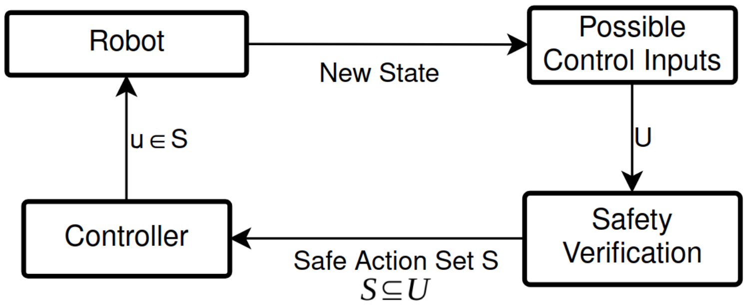

Figure 3.10: Control loop with safety verification given an explicit classifier

Retroactive approach The model of the environment is often insufficient to capture all possible influences, leaving the system to react to unknown outside disturbances. In the event that the system finds itself in a safety-critical state that cannot be prevented, the robot should attempt to minimize the harm caused to safe levels. This typically necessitates a self-monitoring system. The paper by Zhu, Kang, and Chen (2022) [29] presents a method for detecting unexpected collisions in a robot arm by monitoring its joint torques. Upon detection of a collision, an evading controller is activated to minimize the force applied to the collided object while also minimizing the disturbance of the collision on the task space. Although this method has shown promising results, further research is needed to generalize it to a wider range of contact-rich manipulation tasks. Retroactive safety is a viable approach when safety-critical states cannot be avoided, such as during active exploration of the environment.

Preemptive approach Preemptive safety is primarily concerned with planning a safe trajectory or sequence of safe states to a target state. This is particularly difficult in the dynamic real-world application of contact-rich manipulation tasks. Since obstacles may move within the robot's path, a statically planned trajectory may become unsafe to execute. During runtime, a high-frequency planner must compute safe fallback trajectories one time step into the future, so that a trajectory can be followed if the currently followed trajectory turns out to be unsafe. Such a low-level planner is realized by Thumm and Althoff (2022) [31] as a safety shield. Using reachability sets of the robot and its environment, the paper formally proves all-time safety, meaning that the robot will stop moving before a collision can occur [31]. Although it cannot guarantee that the desired safe states are reached, since inverse kinematics is only accurate if the dynamics of the whole system are well modeled, it formally proves that a fail-safe trajectory always exists.

Systems capable of contact-rich manipulation tasks must meet essential safety requirements to be widely commercialized in a legal and ethical manner. Proving safety (for a CE marking compliant with EU standards) will pave the way for robots to assist in or fully automate a variety of demanding manipulation tasks.

4 Conclusion

Contact-rich manipulation tasks represent a significant obstacle in the field of robotics, necessitating resilient, robust, and adaptable control strategies to ensure safe and effective interaction between autonomous robots and their surroundings. This survey paper covered the fundamental background behind contact-rich manipulation task methods, introduced the primary challenges, and feasible approaches for addressing these challenges:

Object Perception and Estimation Using sensing modalities actively is essential for overcoming challenges faced in perception and estimation. This can be achieved through the implementation of active perception, which involves proactively controlling the sensory apparatus to enhance overall perception. Alternatively, interactive perception can be employed, where the robot engages dynamically with its environment, notably through methods such as haptic and auditory exploration.

Generalization In conclusion, our exploration of contact-rich manipulation has revealed that achieving generalization is feasible through several methodologies. Presented approaches have demonstrated the ability to go beyond predefined tasks and adjust to various scenarios, showcasing their potential for skill generalization. The utilization of machine learning algorithms further amplifies this adaptability by enabling robotics systems to learn from experience, refine their strategies, and generalize knowledge to new situations.

Reality Gap Bridging the reality gap requires transferring learned policies to new and unseen real-world robots. Some of the existing and effective methods for such problems are domain randomization, domain adaptation, and adversarial reinforcement learning. Future research should focus on developing more accurate simulation environments that can capture the rich and dynamic nature of real-world environments and effective domain adaptation techniques that can handle the nuanced discrepancies between different environments.

Data-Efficiency Developing robust, adaptable, and data-efficient manipulation algorithms that can operate effectively in real-world scenarios requires addressing challenges such as data scarcity, an explosion of modes, and irrelevant information from the environment. These challenges and their potential solutions for achieving data-efficient models in contact-rich manipulation tasks were presented.

Safety Four approaches to increase safety have been presented. The explicit and preemptive approaches appear to be the most suitable for formal safety verification. However, it is important to note that the safety of both approaches depends on a precise robot dynamics model.

The objective of current research is to expand existing safety approaches to encompass a wider range of safe manipulation tasks and to develop new non-probabilistic models for contact-rich manipulation tasks that are more suitable for formal verification.

Researchers are continuously advancing the autonomous manipulation abilities of robots to overcome the aforementioned challenges with promising results. Although there are still several challenges to overcome, the rate of development allows for an optimistic outlook on the near future of robots capable of contact-rich manipulation.

References

[1] Íñigo Elguea-Aguinaco et al. "A review on reinforcement learning for contact-rich robotic manipulation tasks". In: *Robotics and Computer-Integrated Manufacturing* 81 (2023), p. 102517. ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2022.102517. URL: https://www.sciencedirect.com/science/article/pii/S0736584522001995.

[2] Markku Suomalainen, Yiannis Karayiannidis, and Ville Kyrki. "A survey of robot manipulation in contact". In: *Robotics and Autonomous Systems* 156 (2022), p. 104224. ISSN: 0921-8890. DOI: https://doi.org/10.1016/j.robot.2022.104224. URL: https://www.sciencedirect.com/science/article/pii/S0921889022001312.

[3] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[4] Jack Macki and Aaron Strauss. *Introduction to optimal control theory*. Springer Science & Business Media, 2012.

[5] Ruzena Bajcsy, Yiannis Aloimonos, and John K. Tsotsos. *Revisiting Active Perception*. 2016. arXiv: 1603.02729 [cs.CV].

[6] Saxena A. Le Q. and Ng A. "Active Perception: Interactive Manipulation for Improving Object Detection". In: *Stanford Journal of Science*. 2010.

[7] Fen Fang et al. "Self-Supervised Reinforcement Learning for Active Object Detection". In: *IEEE Robotics and Automation Letters* 7.4 (2022), pp. 10224–10231. DOI: 10.1109/LRA.2022.3193019.

[8] Jeannette Bohg et al. "Interactive Perception: Leveraging Action in Perception and Perception in Action". In: *IEEE Transactions on Robotics* 33.6 (2017), pp. 1273–1291. DOI: 10.1109/TRO.2017.2721939.

[9] Timothy Patten, Michael Zillich, and Markus Vincze. "Action Selection for Interactive Object Segmentation in Clutter". In: Oct. 2018. DOI: 10.1109/IROS.2018.8593918.

[10] Xiang Dou et al. "Active Auditory Exploration for Identifying Object Contents". In: *2020 5th International Conference on Advanced Robotics and Mechatronics (ICARM)*. 2020, pp. 552–558. DOI: 10.1109/ICARM49381.2020.9195294.

[11] Gao Ziyan, Armagan Elibol, and Nak Young Chong. "Planar Pushing of Unknown Objects Using a Large-Scale Simulation Dataset and Few-Shot Learning". In: *2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)*. 2021, pp. 341–347. DOI: 10.1109/CASE49439.2021.9551513.

[12] Karol Hausman et al. "Active articulation model estimation through interactive perception". In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 2015, pp. 3305–3312. DOI: 10.1109/ICRA.2015.7139655.

[13] Nancy S. Pollard Michael C. Koval and Siddhartha S. Srinivasa. "Pose estimation for planar contact manipulation with manifold particle filters". In: *The International Journal of Robotics Research*. 2015.

[14] ClementVigneault Avital Bechar. "Agricultural robots for field operations. Part 2: Operations and systems". In: *Biosystems Engineering* 153 (2016). URL: https://www.sciencedirect.com/science/article/pii/S1537511015301926?ref=pdf_download&fr=RR-2&rr=844bb7d1bcdf2be2.

[15] Rongrong Liu et al. "Deep reinforcement learning for the control of robotic manipulation: a focussed mini-review". In: *Robotics* 10.1 (2021), p. 22. arXiv: 2102.04148 [cs.CV].

[16] Xiaolong Wang Nicklas Hansen. "Generalization in Reinforcement Learning by Soft Data Augmentation". In: (2021). DOI: https://doi.org/10.48550/arXiv.2011.13389.

[17] Zhang et al. "A neural network based framework for variable impedance skills learning from demonstrations". In: *Robotics and Autonomous Systems* (2023). DOI: https://doi.org/10.1016/j.robot.2022.104312.

[18] Argall et al. "A survey of robot learning from demonstration". In: *Robotics and Autonomous Systems* (2008). DOI: https://doi.org/10.1016/j.robot.2008.10.024.

[19] Manuel Baum Xing Li and Oliver Brock. "Augmentation Enables One-Shot Generalization In Learning From Demonstration for Contact-Rich Manipulation". In: *International Conference on Intelligent Robots and Systems (IROS)*. 2023.

[20] Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. "Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey". In: *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. 2020, pp. 737–744. DOI: 10.1109/SSCI47803.2020.9308468.

[21] OpenAI et al. *Learning Dexterous In-Hand Manipulation*. 2019. arXiv: 1808.00177 [cs.LG].

[22] Lukas Hedegaard, Omar Ali Sheikh-Omar, and Alexandros Iosifidis. "Supervised Domain Adaptation: A Graph Embedding Perspective and a Rectified Experimental Protocol". In: *IEEE Transactions on Image Processing* 30 (2021), 8619–8631. ISSN: 1941-0042. DOI: 10.1109/tip.2021.3118978. URL: http://dx.doi.org/10.1109/TIP.2021.3118978.

[23] Lerrel Pinto et al. *Robust Adversarial Reinforcement Learning*. 2017. arXiv: 1703.02702 [cs.LG].

[24] Xinlei Pan et al. *Risk Averse Robust Adversarial Reinforcement Learning*. 2019. arXiv: 1904.00511 [cs.LG].

[25] Alp Aydinoglu, Adam Wei, and Michael Posa. "Consensus Complementarity Control for Multi-Contact MPC". In: *arXiv preprint arXiv:2304.11259* (2023).

[26] Shahbaz Abdul Khader et al. "Data-efficient model learning and prediction for contact-rich manipulation tasks". In: *IEEE robotics and automation letters* 5.3 (2020), pp. 4321–4328.

[27] Tao Pang et al. "Global planning for contact-rich manipulation via local smoothing of quasi-dynamic contact models". In: *IEEE Transactions on Robotics* (2023).

[28] Ying Zhang, Guohui Tian, and Xuyang Shao. "Safe and efficient robot manipulation: Task-oriented environment modeling and object pose estimation". In: *IEEE Transactions on Instrumentation and Measurement* 70 (2021), pp. 1–12.

[29] Xiang Zhu, Shucheng Kang, and Jianyu Chen. "A contact-safe reinforcement learning framework for contact-rich robot manipulation". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2022, pp. 2476–2482.

[30] Ioanna Mitsioni et al. "Safe data-driven contact-rich manipulation". In: *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2021, pp. 120–127.

[31] Jakob Thumm and Matthias Althoff. "Provably safe deep reinforcement learning for robotic manipulation in human environments". In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE. 2022, pp. 6344–6350.

[32] S Levine, N Wagener, and P Abbeel. "Learning contact-rich manipulation skills with guided policy search. ICRA: Proc. IEEE Intern. Conf. on Robotics and Automation. Seattle, WA, USA". In: *arXiv preprint arXiv:1501.05611* (2015).

[33] Sami Haddadin. *Towards safe robots: approaching Asimov's 1st law*. Vol. 90. Springer, 2013.

[34] Benjamin Beyret, Ali Shafti, and A Aldo Faisal. "Dot-to-dot: Explainable hierarchical reinforcement learning for robotic manipulation". In: *2019 IEEE/RSJ International Conference on intelligent robots and systems (IROS)*. IEEE. 2019, pp. 5014–5019.

[35] Ioanna Mitsioni et al. "Safe data-driven model predictive control of systems with complex dynamics". In: *IEEE Transactions on Robotics* (2023).