Hand Manipulation Suite: A Benchmark for Dexterous Manipulation

Vikash Kumar[1], Aravind Rajeswaran[1,2], Abhishek Gupta[3], Emanuel Todorov[2], Sergey Levine[3]
[1]OpenAI, [2]University of Washington, [3]UC Berkeley

***Abstract*—Dexterous hand manipulation is among the most challenging problems in robotics, and remains largely unsolved. This is due to a combination of factors: high dimensionality, intermittent contact dynamics, inadequate tactile sensing information, and under-actuation in the case of dynamic object manipulation. Recent advancements in data-driven methods have led to significant advancements in multiple fields, including robotics. However, these methods have had limited impact on dexterous manipulation due to challenges in modeling the details of hands and the lack of adequate simulation platforms to track and share progress. We propose an experimental platform called the Hand Manipulation Suite (HMS). The goal of HMS is to serve as a shared common platform for research, and to provide the means to scale modern data-driven approaches to dexterous hand manipulation. HMS provides detailed models of dexterous hands and a suite of tasks with associated evaluation metrics, which can be used to directly compare different methods and human performance. A few high fidelity demonstrations, collected using teleoperation, are also provided to illustrate the tasks.**

I. Introduction

Large scale datasets and simulators have been instrumental for progress in fields ranging from computer vision [4] to reinforcement learning [2]. Similarly, in stochastic optimal control (SOC) and reinforcement learning (RL), simulators and benchmarks such as MuJoCo [9] and OpenAI gym [3] have been quickly adopted as evaluation standards by the community. These platforms, through easy simulation, have facilitated the development of algorithms for continuous control tasks like locomotion. However, more challenging control tasks like dexterous manipulation have been under-explored in recent work, in large part due to the limited availability of suitable manipulation tasks, detailed hand models, and common evaluation metrics. In this work, we aim to bridge this gap by introducing the Hand Manipulation Suite (HMS), a set of standardized simulated benchmarks for dexterous manipulation.

Task specific custom manipulators are prevalent in industry for solving isolated manipulation tasks. However, if we need robots to be effective in less structured and human centric environments like home care, elderly care, rescue missions etc, a single manipulator capable of handling multiple complex situations is necessary. Owing to the human centric design of our environments, this manipulator will likely need to be anthropomorphic with high dexterity. However, controlling such complex manipulators is exceedingly difficult. Recent advances in RL and SOC are starting to deliver initial breakthroughs for simple skills. The time is ripe to consolidate and leverage these modern approaches to solve complex tasks such as dexterous manipulation.

The challenges due to high dimensionality, constrained workspace, and limited sensing will likely necessitate customization of existing techniques, leveraging multiple sensing modalities, and development of newer algorithms. HMS introduces a platform to consolidate various efforts on hand manipulation. HMS uses MuJoCo physics engine [9] to provides detailed models of existing anthropomorphic hands, a diverse set of task with robust evaluation metrics, and illustrative demonstrations (via teleoperation). HMS is conceptualized and designed to be interesting to a wide variety of communities including: (1) researchers interested in algorithmic developments for hard continuous control problems[1] (2) roboticists interested in solving hand manipulation in general (3) researchers interested in estimation under heavy occlusion (4) researchers interested in discovering new effective sensing modalities and desirable features (e.g. spatial-temporal properties of tactile skin) (5) researchers interested in investigating hand designs, to name a few. HMS will be released as open-source software for non-commercial use.

II. Hand Manipulation Suite

Despite attention across multiple fields like neuroscience, graphics, and robotics hand manipulation largely remains unsolved. HMS aims to take steps towards addressing this challenge and facilitating research on dexterous hand manipulation by providing an effective platform with necessary details and significantly lower barrier of entry to research in this field.

A. Hands

HMS comes with detailed model of two hands developed independently – *ADROIT* and *MPL* . Both are five fingered anthropomorphic hands. However, they differ significantly in their kinematics and actuation mechanism, thereby resulting in different behaviors. It is recommended to work simultaneously on both models to avoid overfitting to a single platform. Simultaneous study of diverse systems ensures that we understand the underlying principles of manipulation instead of custom design-specific solutions.

B. Tasks

HMS includes several tasks that evaluate hand capabilities across various manipulation skills. To evaluate dexterous manipulation in direct comparison with human performance, we modeled Southampton Hand Assessment Procedure (SHAP) [6]

[1]a wrapper compatible with OpenAI gym is provided to facilitate adoption and use of currently existing algorithms

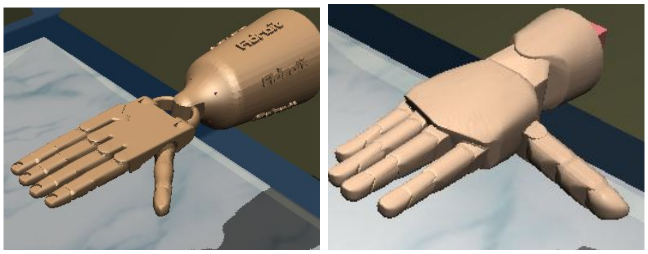
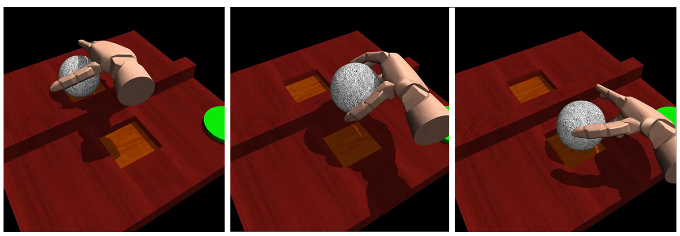


Fig. 1: Antagonistically actuated 24 dof *ADROIT* hand (left) and position actuated 22 dof MPL hand (right).

in our virtual environment. Originally developed to assess the effective functionality of upper limb prostheses, SHAP has since also been applied for assessments of unimpaired participants. The supplementary video[2] and Fig 2a demonstrates some of the tasks considered. We omit further details due to space constraints.



(a) Spherical

III. Evaluations

One approach to standardized evaluations is to use a built-in reward function to compare performance and provide supervision to learning-based methods. OpenAI gym [3] takes such an approach and proved to be a great platform for development of reinforcement learning algorithms. An inadvertent consequence was that large differences in final scores could correspond to functionally indistinguishable behavior, while even small changes in scores could correspond to qualitatively very different behaviors. [7] even found time invariant linear control policies capable of solving all the gym tasks.

In order to evaluate success, HMS implements an "oracle" that accepts a *control policy* and outputs a quantitative measure of its effectiveness. This is implemented through constraint satisfactions, such as checking if the object is within a small region around desired location and if there were any collisions with barriers. In addition, time taken, energy usage, and accidental collisions are also analyzed to assess effectiveness of the policy in the form of binary flags. If the policy completes the task without raising any of the flags, the solution is considered successful. In addition to the oracle, teleoperated demonstrations are provided using the MuJoCo Haptix [5] framework. These demonstrations serve multiple purposes:

1) They act as sanity checks to ensure that there is indeed a strategy that can solve the HMS task.
2) These successful demonstrations can be leveraged to warmstart various SOC or RL algorithms. Learning from Demonstrations (LfD) based approaches can also be leveraged.
3) Watching the teleoperated demonstrations visually conveys to the researcher what an effective solution looks like. As a consequence, they are less likely to accept inappropriate behavior as solution.

[2]https://youtu.be/M_q_YH1aB6Q

IV. Discussion

As indicated in Section III, HMS will contain oracles for evaluation and few demonstrations. A naïve RL or SOC approach would involve attempting to directly use the sparse evaluation oracle as the reward function. However, as demonstrated multiple times in robotic control, a well shaped reward function with non-flat gradients has a number of advantages. Thus, we allow for users to device their own reward functions, but the end result must be evaluated using the provided oracle. As mentioned before, the goal in HMS is to understand manipulation, and not find task specific solutions that improve the learning curve. We expect reward shaping to be an integral part of this endeavor. A generic approach to reward shaping that is broadly applicable across manipulation is likely to involve novel general principles, which in itself is a breakthrough. An alternative approach would be to use the provided demonstrations (or collect more) for LfD or Inverse Optimal Control based approaches.

As a preliminary experiment, we tested the current state of the art model-free RL algorithms like TNPG/TRPO [8] on the various tasks in HMS. When the evaluation oracle is directly used as the reward function, the RL methods are not able to make any progress in reasonable time scales. Our first simple attempts at reward shaping were also not very successful, with policies getting stuck in local optima, which though meaningful, do not accomplish the task. It is not yet clear if the source of difficulty is high dimensionality, exploration strategies, lack of well shaped rewards, or some combination of the above. For future work, we plan to study the performance of model-based RL algorithms like Guided Policy Search and LfD [1] approaches. Subsequently, we hope to develop novel algorithms that enables solving these complex tasks. Since there is an existence proof that the problem can be solved through teleoperation, and since dexterous hand manipulation has a plethora of applications, we hope for a wide adoption of HMS to accelerate algorithmic advancements for dexterous manipulation.

References

[1] Pieter Abbeel and Andrew Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, 2004.

[2] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling. The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research*, 47:253–279, jun 2013.

[3] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *CoRR*, abs/1606.01540, 2016.

[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[5] Vikash Kumar and Emanuel Todorov. Mujoco HAPTIX: A virtual reality system for hand manipulation. In *15th IEEE-RAS International Conference on Humanoid Robots, Humanoids 2015, Seoul, South Korea, November 3-5, 2015*, pages 657–663, 2015.

[6] Colin M. Light, Paul H. Chappell, and Peter J. Kyberd. Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: Normative data, reliability, and validity. *Archive of Physical Medicine and Rehabilitation*, 2002.

[7] Aravind Rajeswaran, Kendall Lowrey, Emanuel Todorov, and Sham Kakade. Towards generalization and simplicity in continuous control. *CoRR*, abs/1703.02660, 2017.

[8] John Schulman, Sergey Levine, Philipp Moritz, Michael I. Jordan, and Pieter Abbeel. Trust region policy optimization. *CoRR*, abs/1502.05477, 2015.

[9] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2012, Vilamoura, Algarve, Portugal, October 7-12, 2012*, pages 5026–5033, 2012.