The Elliott and Connolly Benchmark: A Test for Evaluating the In-Hand Dexterity of Robot Hands

Ryan Coulson[1], Chao Li[2], Carmel Majidi[1,3], and Nancy S. Pollard[1]

***Abstract*— Achieving dexterous in-hand manipulation with robot hands is an extremely challenging problem, in part due to current limitations in hardware design. One notable bottleneck hampering the development of improved hardware for dexterous manipulation is the lack of a standardized benchmark for evaluating in-hand dexterity. In order to address this issue, we establish a new benchmark for evaluating in-hand dexterity, specifically for humanoid type robot hands: the Elliott and Connolly Benchmark. This benchmark is based on a classification of human manipulations established by Elliott and Connolly, and consists of 13 distinct in-hand manipulation patterns. We define qualitative and quantitative metrics for evaluation of the benchmark, and provide a detailed testing protocol. Additionally, we introduce a dexterous robot hand - the CMU Foam Hand III - which is evaluated using the benchmark, successfully completing 10 of the 13 manipulation patterns and outperforming human hand baseline results for several of the patterns.**

I. Introduction

For several decades, roboticists have endeavoured to develop autonomous systems capable of dexterous manipulation [1], [2]. Although substantial progress has been made towards the ultimate goal of achieving human-level dexterity, this still remains an extremely challenging and unsolved problem. One critical aspect of this challenge is the design of suitable hardware platforms - i.e. robot hands - for dexterous manipulation.

In order to properly evaluate the design of a dexterous robot hand, it is essential to have tools for objectively assessing its dexterity. However, while there are a number of manipulation benchmarks present in the literature (see Table I), no single method has been adopted as a "gold standard" for benchmarking robot dexterity. Rather, authors generally follow the approach of providing videos or photo sequences of their robot hands performing various, ad-hoc dexterous manipulation tasks in order to demonstrate the robot's dexterity. Because each hand is evaluated using a unique set of manipulation tasks, there is no way of comparing the dexterity of one robot hand to another.

Clearly, there is a need for a standardized benchmark that can be used to evaluate the dexterity of robot hands.

*This research was partially supported by National Science Foundation awards IIS-1637853 and CMMI-1925130. Chao Li was supported by a NSS-BS scholarship from the Agency for Science, Technology and Research (A*STAR) Singapore.

[1]Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA ryanlcoulson@gmail.com, cmajidi@andrew.cmu.edu, nsp@cs.cmu.edu

[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA chaoli2@andrew.cmu.edu

[3]Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

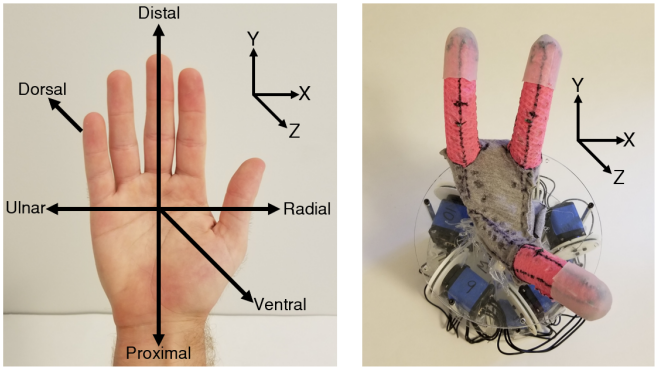


Fig. 1. Coordinate axes for a human hand (left) and the CMU Foam Hand III (right).

Accordingly, we propose a new benchmark - the Elliott and Connolly Benchmark - which consists of 13 in-hand manipulation patterns from a manipulation classification established by Elliott and Connolly [14]. This benchmark is specifically intended for evaluating the in-hand dexterity of humanoid type robot hands. It is tied to motions observed in daily life, and covers the full range of potential in-hand manipulation primitives.

A. Establishing a Dexterity Benchmark

Throughout the rest of this paper, we will define dexterity as "the capability of changing the position and orientation of the manipulated object from a given reference configuration to a different one, arbitrarily chosen within the hand workspace" [1]. Specifically, we focus on *in-hand* dexterity, which involves controlling the pose of a grasped object using *only* the hand, without relying on the arm or body. It is also important to make a distinction between in-hand dexterity (as discussed in this paper) and extrinsic dexterity [15], which entails relying upon external forces - such as gravity and contact forces between the manipulated object and the environment - in order to change the pose of the manipulated object.

We can also define the *range of dexterity* of a hand using a classification established by Bullock et. al [16], which categorizes in-hand manipulation tasks based on rotations and translations along hand coordinate axes. They further categorize in-hand manipulations according to whether or not there is motion at contact - in other words, whether the initial contacts made between the hand and object remain fixed throughout a given manipulation. In total, this classification establishes 12 distinct categories of in-hand manipulation,

TABLE I

BENCHMARKS, TESTS, AND METRICS USED TO EVALUATE DEXTERITY.

Benchmark/Test/Metric	Protocol	Purpose
Purdue Pegboard Test (1948) [3]	Insert pins into pegboard; stack collars and washers on pins.	Assess fine manual dexterity for vocational evaluation and rehabilitation.
Box and Blocks Test (1957) [4]	Transfer blocks from one box to another.	Assess gross manual dexterity for vocational evaluation and rehabilitation.
Jebsen Taylor Hand Function Test (1969) [5]	Writing, simulated page-turning, pick-and-place of small objects, simulated feeding, stacking checkers, and lifting large objects (lightweight and heavy).	Assess ability to complete activities of daily living (ADLs).
Kapandji Test (1986) [6]	The tip of the thumb is used to contact different locations around the hand.	Test opposition and counter-opposition of the thumb.
Sollerman Hand Function Test (1995) [7]	Consists of 20 tasks based on activities of daily living (ADLs).	Assess ability to complete ADLs.
SHAP Test (2002) [8]	Six abstract objects are grasped using various grasp types; 14 activities of daily living (ADLs) are also performed.	Assess the effectiveness of upper limb prostheses; rehabilitation.
Block Pick and Place Benchmark (2015) [9]	Blocks are picked and placed into specific locations and poses.	Assess the dexterity of a robotic manipulator.
Reachable Configuration Manifold Metric (2015) [10]	Map all possible object positions that can be reached from some initial grasp configuration.	Evaluate the dexterous workspace of objects held within the fingertips in a precision grasp.
NIST In-Hand Manipulation Test (2018) [11]	An object is manipulated along as many independent Cartesian axes as possible (up to six), along a desired Cartesian trajectory. Error is calculated between desired trajectory and measured trajectory.	Assess how well a robotic hand can control the pose of an object.
In-Hand Manipulation Benchmark (2020) [12]	Objects are manipulated from an initial grasp to a desired grasp. Error between reached grasp and desired grasp is calculated.	Evaluate the planning and control aspects of robotic in-hand manipulation systems.
Rubik's Cube Manipulation Benchmark (2020) [13]	Rotate a Rubik's cube in prescribed patterns.	Evaluate a robotic system's overall manipulation accuracy, speed, and robustness.

corresponding to individual rotations and translations along each axis in 3D space (i.e. θ_x, Δ_x, θ_y, Δ_y, θ_z, Δ_z - refer to Figure 1 for coordinate axis configuration), for cases of either motion at contact (A) or no motion at contact (NA). It should be noted that the authors could not find a human manipulation pattern corresponding to θ_y (A), reducing the size of the set of possible in-hand manipulation categories to 11. Thus, the range of dexterity of a given humanoid type robot hand can be evaluated based on how many of the 11 possible in-hand manipulation categories established in [16] are feasible.

Table I summarizes a variety of benchmarks, tests, and metrics that have been used to evaluate dexterity. It is important to note that many of these tests - the Purdue Pegboard Test, Box and Blocks Test, Jebsen Taylor Hand Function Test, Sollerman Hand Function Test, SHAP Test, and Block Pick and Place Benchmark - were not developed with the explicit goal of measuring in-hand dexterity. Rather, these tests are meant to evaluate an ability to perform pick-and-place tasks or simple tool usage, often focusing on activities of daily living (ADLs). These tasks primarily depend upon grasping ability and whole-arm dexterity, as opposed to in-hand manipulation. It should also be noted that robot manipulation benchmarks often take the form of competitions, such as the IROS Robotic Grasping and Manipulation Competition [17] and the DARPA Autonomous Robotic Manipulation (ARM) Program [18]. However, these competitions are generally intended to test the competency of robotic systems as a whole (including perception capabilities, planning algorithms, learning/training strategies, etc.), rather than comparing different hardware designs.

Of the benchmarks, tests, and metrics from Table I which *are* explicitly intended for evaluation of in-hand dexterity - the Reachable Configuration Manifold, NIST In-hand Manipulation Test, and In-hand Manipulation Benchmark - only the NIST In-hand Manipulation Test explicitly evaluates the full range of dexterity of the hand based on rotations and translations about hand coordinate axes, as categorized by Bullock et al. [16]. However, this test does not evaluate the hand's ability to complete manipulations which require the initial contacts between the hand and object to be altered, nor is it tied to motions observed in daily life.

Our proposed benchmark - the Elliott and Connolly Benchmark - is based on a classification which consists of 13 distinct in-hand manipulation patterns, derived from observations of multiple human subjects [14]. Here, manipulation patterns are defined as "independently coordinated digit movements intended to move objects within the hand" [14]. We find that this classification is suitable for adaptation as a dexterity benchmark test for several reasons. First, and most importantly, across the 13 manipulation patterns established by Elliot and Connolly, each of the in-hand manipulation categories from Bullock et al. [16] is represented at least once. Thus, we can claim that any hand capable of performing all of these patterns is accordingly capable of performing the full range of possible in-hand manipulation primitives. Second, across the 13 patterns, all of the primary sub-classes of in-hand manipulation - namely finger gaiting, rolling, and sliding [1] - are represented. Finally, because these manipulation patterns are derived from observations of human subjects, they are relevant to the completion of real-world tasks.

B. Developing a Dexterous Robot Hand

In addition to establishing a new dexterity benchmark, this paper introduces a dexterous robot hand - the CMU

Foam Hand III - which is evaluated using the benchmark. The design of this hand builds upon the work of King et al. [19], [20], [21], who established techniques for fabricating and controlling tendon driven hands made almost entirely of foam. This class of robot hand is promising for dexterous manipulation, since the foam-based architecture largely eliminates constraints on morphology, and permits the hand to be actuated with a theoretically unlimited number of tendons in any desired configuration. Of course, there are practical limitations on the number of motors that can be used to drive these tendons, and the dexterous hands developed in [19], [20], [21] used 10 motors to drive an equal number of tendons. The contribution of this work is to modify the design of these hands in order to enable a high level of dexterity while requiring a relatively low number of motors for operation.

Within the literature, there are several examples of dexterous robot hands which use a limited number of motors. These include the Pisa/IIT SoftHand 2 [22], which uses two motors; various hands developed as a part of the Yale OpenHand Project [23], [24], [10], which use between two and four motors; the highly biomimetic hand developed by Xu and Todorov [25], which uses 10 motors; and the RBO Hand 2 [26], which uses 7 pneumatic actuators. However, due to the aforementioned lack of a standardized dexterity benchmark, it is difficult to objectively assess the level of dexterity of any of these hands. This challenge motivates the introduction of the Elliott and Connolly Benchmark, which would allow the relative dexterity of various robot hands to be compared.

II. Benchmark Test

A. Metrics

In order to establish a benchmark test based on the 13 manipulation patterns from Elliott and Connolly [14], it is necessary to define metrics for evaluation. First, we define a qualitative, binary metric - *Success* or *Failure* - for each pattern. In order to evaluate this metric, we have developed a list of criteria for each manipulation pattern, provided in the "Criteria" column of Appendix I. These criteria were developed based on our understanding of the descriptions and illustrations provided by Elliott and Connolly for each manipulation pattern. We have also recorded videos of a human hand performing each manipulation pattern (see link in Multimedia Document), which can be referenced when determining whether a pattern has been performed successfully.

In addition to this binary designation, failed manipulation patterns can also be labeled as *Incomplete*, and successful patterns can be labeled as *Non-anthropomorphic*. The only manipulation patterns which are eligible to be labeled as incomplete are Rotary Step and Interdigital Step. This is due to the fact that each of these patterns requires a certain amount of rotation ($360°$, per the criteria in Appendix I) in order to be evaluated as successful, meaning it is possible to achieve partial completion by rotating the object less than $360°$. A successful manipulation pattern is labeled as *Non-anthropomorphic* if the object is rotated or translated along the correct axis, but in a manner that is dissimilar to human baseline examples.

We have also defined quantitative metrics for the evaluation of each manipulation pattern. These metrics - *Normalized Average Translation* and *Average Rotation* - are used to assess translations and rotations of an object along hand coordinate axes. The goal for the hand is to maximize the value of each of these metrics.

Normalized Average Translation is defined as the distance an object is translated along a particular hand coordinate axis, normalized by the average finger length of the hand and averaged over N trials. The rationale for normalizing by average finger length is to eliminate any bias which arises from the fact that larger hands will be capable of achieving proportionally larger translations than smaller hands when transitioning between identical postures. Average finger length is defined as the average distance between the fingertips and the center of the hand. The center of the hand is defined as the point of intersection on the palm by lines extended from the fingertips. *Normalized Average Translation* can be calculated using the following equation:

$$T = \frac{\sum_{i=1}^{N} d_i}{N * F}, \tag{1}$$

where N is the number of trials, d_i is the translation distance for an individual trial, and F is the average finger length of the hand.

Average Rotation is defined as the degree to which an object is rotated about a particular hand coordinate axis, averaged over N trials. *Average Rotation* can be calculated using the following equation:

$$R = \frac{\sum_{i=1}^{N} \theta_i}{N}, \tag{2}$$

where N is the number of trials and θ_i is the degree of rotation for an individual trial.

Additionally, we define performance scores, which are used to compare the performance of the robot hand to the performance of a human hand. The *Translation Performance Score* for a particular manipulation pattern m can be calculated as follows:

$$P_m^T = \frac{T_m^{robot}}{T_m^{human}} \times 100. \tag{3}$$

Likewise, the *Rotation Performance Score* can be calculated as:

$$P_m^R = \frac{R_m^{robot}}{R_m^{human}} \times 100. \tag{4}$$

Thus, the performance score for a particular manipulation pattern indicates the percentage of translation or rotation achieved by the robot, relative to the human baseline.

TABLE II
YCB OBJECTS REQUIRED FOR THE ELLIOTT AND CONNOLLY BENCHMARK.

Manipulation Pattern	Object	YCB ID
Pinch (P)	Bolt & Nut	46, 47
Dynamic Tripod (DT)	Small Marker	41
Squeeze (S)	Syringe	N/A
Twiddle (T)	Bolt & Nut	46, 47
Rock (R)	Cup (yellow)	64
Rock II (RII)	Small Marker	41
Radial Roll (RR)	Marble (green)	62
Index Roll (IR)	Marble (green)	62
Full Roll (FR)	Wood Block	69
Rotary Step (RS)	Cup (yellow)	64
Interdigital Step (IS)	Small Marker	41
Linear Step (LS)	Large Marker	40
Palmar Slide (PS)	Large Marker	40

B. Protocol

In order to perform this benchmark test, the required materials are a robot hand, a camera, a printer which can be used to print various AprilTags [27], and a set of objects for manipulation. The objects required for this benchmark have been selected from the YCB Object Set [28] - a correspondence between manipulation patterns and objects is shown in Table II. In the case of the Squeeze manipulation pattern, a suitable object could not be found within the YCB Object Set, so we use a commercially available $5\,\mathrm{mL}$ syringe. In order to ensure consistency and repeatability across tests, it is important that experimenters do not deviate from the set of objects listed in Table II.

The first step in performing this benchmark is to measure the average finger length of the robot hand being used. This can be achieved by attaching threads to the fingertips of the robot and extending the threads along the fingers, moving towards the palm (see Multimedia document for examples). The threads are then cut where they intersect (defined as the center of the palm), and their resultant lengths are measured and averaged. Note that in many cases, the threads will not all intersect at a single point - in these cases, each thread can be cut at the first intersection that occurs.

Prior to performing a manipulation pattern, an AprilTag must be attached to the object being used. In most cases, the AprilTags can be attached to objects in a way that allows them to be manipulated without causing interference (see Multimedia Document for examples). The Marble is an exception. In order to attach an AprilTag to this object without causing interference, we attach a Bolt to the marble and then attach the AprilTag to the end of the bolt. Another bolt is attached to the opposite side of the marble in order to balance the weight. In the case of the Rock II and Interdigital Step manipulation patterns, no AprilTag is needed since the tilt angle of the marker can be measured directly. Additionally, no AprilTag is required for Squeeze since linear displacement of the syringe plunger can be directly measured. Note that while the usage of AprilTags is recommended for this protocol, it is not explicitly required – any similarly precise method of motion tracking is acceptable.

When performing a manipulation pattern, the first step is to configure the hand into a suitable initial pose, with the object in the grasp of the hand. There is no precisely "correct" initial pose for any of the manipulation patterns - rather the experimenter should refer to the criteria in Appendix I, as well as the human hand baseline video linked in the Multimedia Document. The manipulation begins as soon as the hand moves out of its initial pose, and lasts until the hand reaches its target pose. As with the initial pose, there is no precise target pose to be achieved for any of the manipulation patterns. However, given that the dexterity of the hand is being evaluated in part based on the *Average Normalized Translation* or *Average Rotation* achieved for each manipulation pattern, experimenters are encouraged to select target poses which maximize these metrics, while still adhering to the criteria provided in Appendix I. It is also important that the experimenter notes the overall orientation of their robot hand relative to gravity (e.g. palm facing upwards) for each manipulation pattern, since the ability of the hand to complete a given pattern may depend on its orientation.

During each manipulation, the camera records the rotation and translation of the AprilTag (when applicable). The resultant data must then be analyzed in order to determine a value for rotation or translation about a particular hand coordinate axis. These values are then used to calculate the *Average Normalized Translation* or *Average Rotation* for each manipulation pattern. In some cases, a single pattern may involve both translation and rotation (see "Category" column in Appendix I). Each manipulation must be repeated at least three times in order to calculate a mean and standard deviation from the data.

No autonomous perception or planning is required in order to complete this benchmark. This is because we are only concerned with the capabilities of the robot hardware, regardless of whether these capabilities can be implemented in an autonomous system using current algorithms. In other words, this benchmark is only assessing the *potential* dexterity of robot hands. For the same reason, each manipulation pattern can be repeated as many times as necessary in order to achieve at least three successful trials - there is no need to record individual failed trials. Additionally, the time required to complete each manipulation pattern does not need to be recorded, since this benchmark is not focused on manipulation efficiency.

III. ROBOT HAND

As explained in Section I-B, a tendon driven foam hand - the CMU Foam Hand III - was also developed for this study and evaluated using the Elliott and Connolly Benchmark. The design goal for this hand was to achieve a high level of dexterity while requiring a limited number of motors for operation. The design process primarily consisted of testing and iteration of physical prototypes - ultimately, a bio-inspired, 3-fingered hand with 10 tendons actuated by an equal number of motors was found to strike the best balance between high dexterity and few motors.

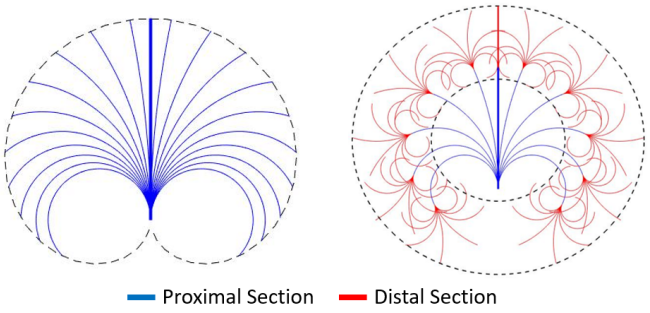


Fig. 2. Two-dimensional workspace of continuum fingers. Workspace is represented by the area between dotted lines. (left) Finger actuated using only proximal tendons. (right) Finger actuated using proximal and distal tendons. Note that the finger on the left (proximal tendons only) has zero workspace, since there is only one dotted line, hence zero area between dotted lines.

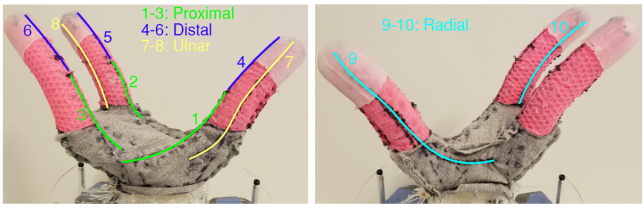


Fig. 3. Tendon routing diagram for the CMU Foam Hand III.

The morphology of the CMU Foam Hand III is based on the thumb, index, and middle fingers of the human hand. The robot hand's fingers are curved away from its palm in its rest pose in order to eliminate the need for extensor tendons - the hand relies upon the elasticity of the foam for finger extension. The hand's tendon routing, coupled with its continuum structure, enables its fingers to bend in almost any direction. Importantly, the inclusion of both proximal and distal tendons enables a multi-jointed behavior that greatly increases the workspace of the hand's fingertips, relative to a case with only proximal tendons (see Figure 2). A tendon routing diagram for the hand is shown in Figure 3. For more details on the design of the CMU Foam Hand III, refer to Coulson [29].

IV. Experiments

The Elliott and Connolly Benchmark was completed using the CMU Foam Hand III. We used a 5MP Raspberry Pi Zero W camera module placed $35\,\mathrm{cm}$ away from the hand to record each manipulation. Prior to performing any manipulations, the camera was calibrated using the camera calibration library in OpenCV (Open Computer Vision Library). Additionally, after determining the average finger length of the robot hand, that length was sketched onto a piece of graph paper, which was placed $35\,\mathrm{cm}$ away from the camera. The average finger length was then measured from the undistorted image in terms of pixels. In postprocessing of manipulation footage, each frame was corrected for lens distortion. The resulting undistorted video was then used to obtain quantitative measurements.

The CMU Foam Hand III was controlled open-loop by moving the hand between predefined keyframes. Here, a keyframe is defined as a hand posture in tendon space. Since the hand's tendons are actuated using servo motors, a keyframe consists of position values for each of the hand's 10 motors. Keyframes were empirically derived via trial-and-error testing. Many manipulation patterns required only two keyframes - one for the initial pose and one for the target pose - but some patterns required up to 5 keyframes for completion (including the initial and target keyframes). We chose to use keyframes because it was the simplest option for control of our hand, but any control strategy may be used in order to complete the benchmark.

Prior to completing the benchmark with the robot hand, we completed the benchmark using a human hand in order to establish a baseline for each manipulation pattern. Because the CMU Foam Hand III has only 3 fingers, we used only 3 fingers for the human hand baseline. Videos from the human hand performance were referenced while deriving keyframes for the robot hand. Other groups performing this benchmark are encouraged to use the human hand videos and data provided in this paper for reference (see Multimedia Document). They are also encouraged to collect their own human hand videos and data, although this is not required.

V. Results and Discussion

The CMU Foam Hand III was able to successfully complete 10 of the 13 manipulation patterns from the Elliott and Connolly Benchmark, based on the criteria provided in Appendix I. The successfully completed manipulation patterns included Pinch, Dynamic Tripod, Twiddle, Rock, Rock II, Radial Roll, Index Roll, Full Roll, Linear Step, and Palmar Slide. Photo sequences of each manipulation pattern are included in Appendix I.

Quantitative results from the experiments are shown in Figure 4. It should be noted that in most cases, the robot was outperformed by the human. This is to be expected, since humans exhibit extraordinarily high levels of dexterity. However, there are some cases - namely the *Normalized Average Translation* for Dynamic Tripod, Twiddle, and Index Roll - where the robot outperforms the human. This result suggests that the design of the robot hand may have some kinematic advantages over the human hand when considering certain aspects of dexterity. Finally, note that the performance scores can be used to assess the robot's relative proficiency for each manipulation pattern. For example, Rock and Pinch both received relatively low performance scores compared to other manipulation patterns, implying that it may be prudent to focus on these patterns (among others) when considering how improvements might be made to the robot hand design.

There are several aspects of the design of the CMU Foam Hand III that we believe contributed to its success in performing the Elliott and Connolly Benchmark. First, the bio-inspired morphology and pragmatic tendon routing of the robot enabled it to closely imitate postures used in the human hand baseline for most manipulation patterns. Second, the robot's mechanical compliance - resulting from its foam

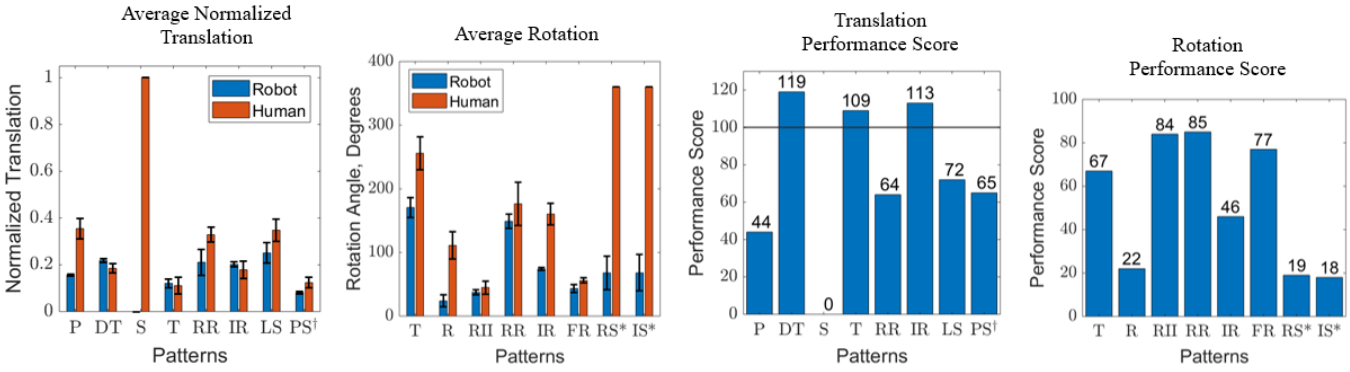


Fig. 4. Quantitative results from benchmark testing. Error bars represent one standard deviation from the mean. Patterns marked with an asterisk (*) are evaluated as incomplete. Patterns marked with a dagger (†) are evaluated as non-anthropomorphic. For performance scores, a score of 100 indicates equivalence between the robot and human.

structure - allows for a considerable margin of error in its control. This is because the robot's fingers inherently deform in order to accommodate objects which they come into contact with, eliminating the need for minute adjustments in posture which might otherwise be necessary in order to establish sufficient contacts with manipulated objects. Finally, the silicone end-caps placed over the robot's fingertips (see Figure 3) greatly increase the friction between the robot hand and the manipulated objects, leading to a substantial decrease in the frequency of objects being dropped during manipulation.

With regard to the manipulation patterns which the robot hand was unable to perform successfully - Squeeze, Rotary Step, and Interdigital Step - there were several factors which contributed towards these failures. The first of these limiting factors was the control strategy used, i.e. empirically derived keyframes. While this control strategy is straightforward to implement, it is limited by the physical intuition of the programmer, and has no theoretical groundings which drive it towards success. Especially for more complicated, multi-step manipulation patterns such as Rotary Step and Interdigital Step, it is possible that a more sophisticed control strategy, e.g. one derived via reinforcement learning, could produce successful results where empirically derived keyframes failed.

Another limiting factor was the fact that the morphological design and tendon routing of the robot hand were constrained by the goal of limiting the number of motors required for actuation. Specifically, the robot hand could have benefitted from the addition of a fourth finger, as well as additional tendons. Determining the optimal balance between dexterity and motor quantity is a non-trivial problem that merits exploration in future work.

The final limiting factor was the low stiffness and correspondingly poor force transmission capability of the CMU Foam Hand III. This is one of the primary drawbacks of a hand made entirely from soft materials, and was the main factor in the hand's inability to complete the Squeeze pattern. The task associated with this pattern involves compressing the plunger of a syringe. When attempting this task, the hand's fingers would consistently buckle before they could apply enough force to displace the plunger.

In addition to the limitations of the robot hand itself, there are several conceptual limitations of this study. First, it should be noted that while the objective of this benchmark is to assess robot hardware designs independent of software implementation, the performance of robot hands ultimately cannot be decoupled from the control strategy that is employed. This is evident from the fact that the CMU Foam Hand III failed several manipulation patterns in part due to limitations associated with keyframe-based control. Second, it is important to note that this benchmark allows for the initial contacts between the robot hand and the manipulated object to be established manually prior to manipulation. This may artificially over-emphasize abilities that the hand would not be able to perform without careful oversight by the researcher. Ultimately, dexterous manipulation becomes a much more challenging problem when the robot must autonomously grasp - and potentially re-grasp - an object in order to establish suitable contacts before performing in-hand manipulations.

VI. Conclusion

This study focused on establishing a new benchmark for evaluating the in-hand dexterity of humanoid type robot hands: the Elliott and Connolly Benchmark. This benchmark consists of 13 distinct in-hand manipulation patterns that are tied to motions observed in daily life. These patterns span the full range of possible in-hand manipulation primitives, and include examples of finger gaiting, rolling, and sliding. Qualitative and quantitative metrics are defined for evaluation of the Elliott and Connolly Benchmark, and a detailed testing protocol is established.

Additionally, a dexterous robot hand - the CMU Foam Hand III - is introduced. The CMU Foam Hand III is tendon driven and made almost entirely of foam, building upon the work of King et al. [19], [20], [21]. The CMU Foam Hand III is evaluated using the Elliott and Connolly Benchmark, successfully completing 10 of the 13 manipulation patterns,

and outperforming human hand baseline results for several of the patterns.

In future work, we aim to reduce the number of motors required for operation of the CMU Foam Hand III by implementing mechanical synergies [30], which allow for one-to-many motor-driven actuation of tendons. Additionally, we hope to increase the dexterity of the CMU Foam Hand III by implementing reinforcement learning-based control, taking inspiration from the outstanding in-hand manipulation results achieved via reinforcement learning by OpenAI [31]. Finally, we plan to explore alternative morphologies and tendon routings for the hand through continued iteration and testing.

ACKNOWLEDGEMENT

Thanks to Dominik Bauer, Jonathan King, and Abhinav Gupta - all members of the CMU community - for their input and insight on the design of the CMU Foam Hand III. Additionally, thanks to Ella Moore (CMU) and Karmesh Yadav (CMU) for their help in the hand's development.

REFERENCES

[1] A. Bicchi, "Hands for dexterous manipulation and robust grasping: A difficult road toward simplicity," *IEEE Transactions on robotics and automation*, vol. 16, no. 6, pp. 652–662, 2000.

[2] C. Piazza, G. Grioli, M. Catalano, and A. Bicchi, "A century of robotic hands," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, pp. 1–32, 2019.

[3] J. Desrosiers, R. Hebert, G. Bravo, and E. Dutil, "The purdue pegboard test: normative data for people aged 60 and over," *Disability and rehabilitation*, vol. 17, no. 5, pp. 217–224, 1995.

[4] V. Mathiowetz, G. Volland, N. Kashman, and K. Weber, "Adult norms for the box and block test of manual dexterity," *American Journal of Occupational Therapy*, vol. 39, no. 6, pp. 386–391, 1985.

[5] E. D. Sears and K. C. Chung, "Validity and responsiveness of the jebsen–taylor hand function test," *The Journal of hand surgery*, vol. 35, no. 1, pp. 30–37, 2010.

[6] A. Kapandji, "Clinical test of apposition and counter-apposition of the thumb," *Annales de chirurgie de la main: organe officiel des societes de chirurgie de la main*, vol. 5, no. 1, p. 67, 1986.

[7] C. Sollerman and A. Ejeskär, "Sollerman hand function test: a standardised method and its use in tetraplegic patients," *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*, vol. 29, no. 2, pp. 167–176, 1995.

[8] J. Adams, K. Hodges, J. Kujawa, and C. Metcalf, "Test-retest reliability of the southampton hand assessment procedure," *International Journal of Rehabilitation Research*, vol. 32, p. S18, 2009.

[9] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, "Benchmarking in manipulation research: The ycb object and model set and benchmarking protocols," *arXiv preprint arXiv:1502.03143*, 2015.

[10] L. U. Odhner and A. M. Dollar, "Stable, open-loop precision manipulation with underactuated hands," *The International Journal of Robotics Research*, vol. 34, no. 11, pp. 1347–1360, 2015.

[11] J. Falco, K. Van Wyk, and E. Messina, "Performance metrics and test methods for robotic hands," *DRAFT NIST Special Publication*, vol. 1227, 2018.

[12] S. Cruciani, B. Sundaralingam, K. Hang, V. Kumar, T. Hermans, and D. Kragic, "Benchmarking in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 588–595, 2020.

[13] B. Yang, P. E. Lancaster, S. S. Srinivasa, and J. R. Smith, "Benchmarking robot manipulation with the rubik's cube," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2094–2099, 2020.

[14] J. M. Elliott and K. Connolly, "A classification of manipulative hand movements," *Developmental Medicine & Child Neurology*, vol. 26, no. 3, pp. 283–296, 1984.

[15] N. C. Dafle, A. Rodriguez, R. Paolini, B. Tang, S. S. Srinivasa, M. Erdmann, M. T. Mason, I. Lundberg, H. Staab, and T. Fuhlbrigge, "Extrinsic dexterity: In-hand manipulation with external forces," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 1578–1585.

[16] I. M. Bullock, R. R. Ma, and A. M. Dollar, "A hand-centric classification of human and robot dexterous manipulation," *IEEE transactions on Haptics*, vol. 6, no. 2, pp. 129–144, 2012.

[17] J. Falco, Y. Sun, and M. Roa, "Robotic grasping and manipulation competition: competitor feedback and lessons learned," in *Robotic Grasping and Manipulation Challenge*. Springer, 2016, pp. 180–189.

[18] D. Hackett, J. Pippine, A. Watson, C. Sullivan, and G. Pratt, "An overview of the darpa autonomous robotic manipulation (arm) program," *Journal of the Robotics Society of Japan*, vol. 31, no. 4, pp. 326–329, 2013.

[19] J. P. King, D. Bauer, C. Schlagenhauf, K.-H. Chang, D. Moro, N. Pollard, and S. Coros, "Design. fabrication, and evaluation of tendon-driven multi-fingered foam hands," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–9.

[20] C. Schlagenhauf, D. Bauer, K.-H. Chang, J. P. King, D. Moro, S. Coros, and N. Pollard, "Control of tendon-driven soft foam robot hands," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–7.

[21] D. Bauer, C. Bauer, J. P. King, D. Moro, K.-H. Chang, S. Coros, and N. Pollard, "Design and control of foam hands for dexterous manipulation," *International Journal of Humanoid Robotics*, vol. 17, no. 01, p. 1950033, 2020.

[22] C. Della Santina, C. Piazza, G. Grioli, M. G. Catalano, and A. Bicchi, "Toward dexterous manipulation with augmented adaptive synergies: The pisa/iit softhand 2," *IEEE Transactions on Robotics*, vol. 34, no. 5, pp. 1141–1156, 2018.

[23] L. U. Odhner and A. M. Dollar, "Dexterous manipulation with underactuated elastic hands," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 5254–5260.

[24] R. R. Ma and A. M. Dollar, "An underactuated hand for efficient finger-gaiting-based dexterous manipulation," in *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*. IEEE, 2014, pp. 2214–2219.

[25] Z. Xu and E. Todorov, "Design of a highly biomimetic anthropomorphic robotic hand towards artificial limb regeneration," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 3485–3492.

[26] R. Deimel and O. Brock, "A novel type of compliant and underactuated robotic hand for dexterous grasping," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 161–185, 2016.

[27] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 3400–3407.

[28] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The ycb object and model set: Towards common benchmarks for manipulation research," in *2015 international conference on advanced robotics (ICAR)*. IEEE, 2015, pp. 510–517.

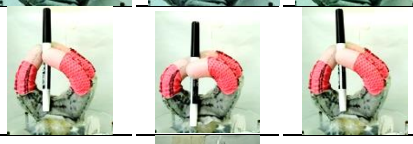
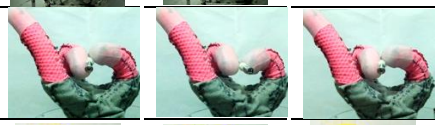
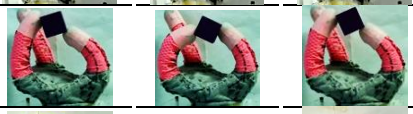
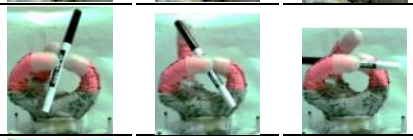
[29] R. Coulson, "Soft materials architectures for robot manipulation," Master's thesis, Carnegie Mellon University, Pittsburgh, PA, August 2020.

[30] C. Y. Brown and H. H. Asada, "Inter-finger coordination and postural synergies in robot hands via mechanical implementation of principal components analysis," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 2877–2882.

[31] O. M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, *et al.*, "Learning dexterous in-hand manipulation," *The International Journal of Robotics Research*, vol. 39, no. 1, pp. 3–20, 2020.

APPENDIX I
MANIPULATION PATTERN INFORMATION

Patterns marked with an asterisk (*) are evaluated as incomplete. Patterns marked with a dagger (†) are evaluated as non-anthropomorphic. See next page for table.

Pattern	Criteria	Sub-Class	Category	Success/ Failure	Sequence
Pinch (P)	Object is held between two fingers. Both fingers are flexed simultaneously in order to translate the object along the ventro-dorsal axis, towards the palm. Fingers are then simultaneously extended to bring the object back to its starting position.	N/A	Δ_Z(NA)	Success	
Dynamic Tripod (DT)	Object is held between three fingers. All three fingers are simultaneously flexed and extended, in repetitive motions, in order to translate the object along the ventro-dorsal axis. Application: writing.	N/A	Δ_Z(NA)	Success	
Squeeze (S)	Deformable object is held between three or more fingers. All fingers are simultaneously flexed towards the object's centroid in order to compress the object, e.g. when squeezing a rubber ball or compressing the plunger of a syringe.	N/A	Δ_Y(NA)	Failure	
Twiddle (T)	Object is held between the distal phalanx of one finger (manipulating finger) and along the side of the proximal phalanx of another finger (stabilizing finger). The manipulating finger is flexed and extended in order to roll the object along the length of the stabilizing finger.	Rolling	Δ_Z(A) θ_X(A)	Success	
Rock (R)	Round object is held between three or more fingers. Fingers are used to rotate the object about ventro-dorsal axis, e.g. when unscrewing the lid of a bottle.	N/A	θ_Z(NA)	Success	
Rock II (RII)	Elongated object is held in opposed grasp between three fingers. Two of the fingers are alternately flexed and extended in order to pivot the object about the third finger.	N/A	θ_Y(NA)	Success	
Radial Roll (RR)	Object is held between the distal phalanx of one finger (stabilizing finger) and along the side of the proximal phalanx of another finger (manipulating finger). The manipulating finger is flexed and extended in order to roll the object along the length of the stabilizing finger.	Rolling	Δ_Z(A) θ_X(A)	Success	
Index Roll (IR)	Object is held between the distal phalanxes of two fingers. One finger is then repetitively flexed and extended in order to roll the object along the length of the other finger.	Rolling	Δ_Z(A) θ_X(A)	Success	
Full Roll (FR)	Object is held between distal phalanxes of two fingers. One finger is then repetitively flexed and extended in order to pivot object about a stationary point on the other finger.	N/A	θ_X(NA)	Success	
Rotary Step (RS)	Finger gaiting sequence during which a round object is incrementally rotated about the ventro-dorsal axis by at least 360 degrees.	Finger Gaiting	θ_Z(A)	Failure*	
Interdigital Step (IS)	Finger gaiting sequence during which an elongated object is incrementally rotated about a pivot point by at least 360 degrees. The rotation occurs about the ventro-dorsal axis.	Finger Gaiting	Δ_Y(A)	Failure*	
Linear Step (LS)	Finger gaiting sequence during which fingers are translated along the length of an elongated object. Sliding between fingers and object is generally required.	Finger Gaiting, Sliding	Δ_X(A)	Success	
Palmar Slide (PS)	Elongated object is held in a palmar grasp. Two fingers, which are initially flexed with their distal phalanxes in contact with the object, are extended in order to translate object along the radio-ulnar axis. Application: removing cap from a pen.	N/A	Δ_X(NA)	Success†