Benchmarking Protocols for Evaluating Grasp Strength, Grasp Cycle Time, Finger Strength, and Finger Repeatability of Robot End-effectors

Joe Falco[1], Daniel Hemphill[2], Kenny Kimble[1], Elena Messina[1], Adam Norton[2], Rafael Ropelato[1], and Holly Yanco[2]

Abstract - This paper describes a set of metrics and supporting benchmarking protocols for determining the performance characteristics of robot end-effectors. In the short-term, these tools are proving useful as a common ground for assessing and comparing end-effectors. The long-term goal is a standard framework for providing technical specifications for robotic end-effectors to help pair technologies to application spaces. This paper presents a subset of the metrics – grasp strength, grasp cycle time, finger strength, and finger repeatability – with accompanying measurement techniques and supporting test artifacts. The application of these metrics and protocols is demonstrated using example implementations to characterize a variety of robot end-effectors, with example data sets and test designs provided for downloading.

I. Introduction

While researchers sometimes accompany their robotic end-effector research with performance-based testing, experimental procedures are one-offs and there is little to no duplication of experiments (standardized benchmarking) in order to compare research results within the robotics community. To fill this gap, the National Institute of Standards and Technology (NIST) has initiated a community-driven approach through the Institute of Electrical and Electronics Engineers (IEEE) Robotics and Automation Society (RAS) Robotic Hand Grasping and Manipulation (RHGM) to define metrics and test methods so that researchers can benchmark and make comparisons between end-effector designs as well as sensor and algorithm implementations. This paper describes benchmarking protocols with implementation examples for measuring four of the eleven performance criteria: grasp strength, grasp cycle time, finger strength, and finger repeatability. For additional details, see [1], a working document which houses these formalized performance test methods (for eventual standardization).

In order to design relevant performance metrics and associated test methods for characterizing robotic end-effectors used for grasping and manipulation applications (i.e., hands and grippers), it helps to understand the issues surrounding robotic grasping, position accuracy, applied force, and manipulation. Regardless of the actual task being performed, any of these problems can be broken down into their first principles – kinematics and kinetics – or, simply, motion and effort. Building test methods from this fundamental point of view will ultimately lead to relevant performance capture, and will span from lower-level capabilities including primitive sensing and control to higher-level capabilities including manipulation, perception, and decision making. These metrics and test methods form part of a measurement science framework inspired by examination of the technical directions and inputs from researchers, developers, and users [2]. The described benchmarking protocols characterize elemental performance of robotic end-effectors to help build a more complete picture of their overall performance in application-relevant tasks. While it may be possible to measure kinematic and kinetic performance characteristics directly from the internal sensors of the end-effector, these measurements would be based on the inherent properties of the end-effector under test. Rather, independent measurement systems are used to benchmark forces and motion for comparative metrics between end-effectors to establish extrinsic ground truths. Examples of these independent measurement systems are described with each protocol.

Throughout this article, several commercially available end-effector grasping technologies[1] were used to verify the benchmarking protocols described. The models and manufacturers of these systems are identified, but the corresponding performance data is anonymized since the intent of example implementations is not to characterize commercially available end-effectors, but rather to demonstrate the utility of these performance benchmarks. It is the task of the community to make these comparisons.

Manuscript received: August, 15, 2019; Revised: November, 18, 2019; Accepted: December, 16, 2019.

This paper was recommended for publication by Editor Han Ding upon evaluation of the Associate Editor and Reviewers comments. This work was supported in part by the National Institute of Standards and Technology (NIST) under 70NANB17H256.

[1]Joe Falco, Kenny Kimble, Rafael Ropelato, and Elena Messina are with the National Institute of Standards and Technology (NIST), 100 Bureau Drive, Gaithersburg, MD 20899 USA, (e-mail: joseph.falco@nist.gov, kenneth.kimble@nist.gov, rafael.ropelato@nist.gov, elena.messina@nist.gov).

[2]Daniel Hemphill, Adam Norton, and Holly Yanco are with the New England Robotics Validation and Experimentation (NERVE) Center, University of Massachusetts Lowell, 110 Canal St., Lowell, MA 01852 USA, (e-mail: daniel_hemphill@student.uml.edu, adam_norton@uml.edu, holly@cs.uml.edu).

Digital Object Identifier (DOI): see top of this page.

[1]Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

All protocol tools and performance data can be downloaded here: https://tinyurl.com/UML-NIST-Manip-Bench.

II. Prior Work

Many of these test methods are derived, based on informal experiments found within the robotic end-effector research literature. In [3], the authors propose a benchmark to measure the kinematic ability of an end-effector to grasp cylindrical objects of different diameters as well as its ability to resist external forces acting on the object under grasp. In [4] the authors conduct similar experiments with grasps as in [3] but also independently measure the contact force. The authors in [5] implemented a test apparatus to test the power grasp capabilities of an end-effector that used tendons to actuate finger motion.

A test is used in [6] to assess an end-effector's ability to control delicate manipulation tasks using slippage measures as well as a crushing measure to indicate the ability to control delicate manipulation tasks. The authors in [7] present an experimental setup to test grasp improvements achieved when integrating piezo-film contact sensors with a reactive control algorithm. The developers of the iHY also developed a test for measuring the compliance of planar and spherical pinch grasps using a 6-axis force-torque sensor to a mill head-stock with the end-effector fixtured in the mill's vice [5]. The authors in [8] report on experiments used to evaluate the in-hand manipulation capabilities of an under actuated hand by tracking the position of the object relative to an initial fingertip grasped position. Preliminary efforts to define these metrics and test methods into a framework for benchmarking end-effectors, with an emphasis on robotic hands is reported in [9].

III. Grasp Strength

Grasp strength is a kinetic metric defined as the maximum force a grasping end-effector can impose on an object [1], [10]. This measure will yield information regarding the payload capability of an end-effector for various object sizes as well as its limits in resisting pulling or pushing forces during a grasp operation. Some manufacturers of end-effectors may provide measures of grasp strength on a product specification sheet, but the methods used to determine this metric may not be specified or may not be usable by end-effectors of different morphology. For example, the Robotiq 2F-85 instruction manual notes force measurements derived using a S-type load cell (presumably positioned between the fingers) [11], which would be not usable by an end-effector with fingers that are not in-line with one another.

A. Test Artifact and Measurement Method

The test artifact used is a split cylinder or split block design (see Figure 1), consisting of two three-dimensional (3D) printed structures that hold two or more single-axis force load cells. The load cells are oriented parallel to one another to measure the force exerted on the artifact by the end-effector along the measurement axis. The block design is used for pinch grasps and the cylindrical design is used for wrap grasps. We define these grasp types as follows:

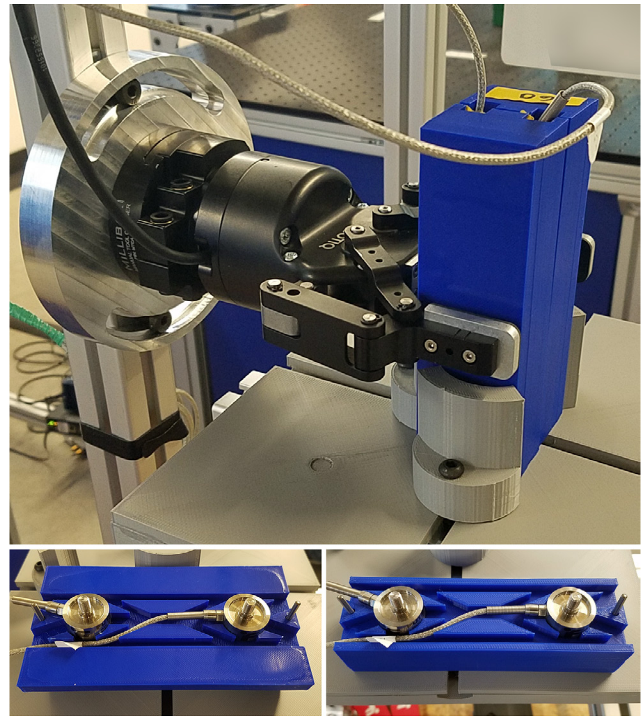

Fig. 1. Top: Test setup for conducting grasp strength and grasp cycle time with a Robotiq 2F-85 Gripper. Bottom: Examples of the inside, with varying size and shape test artifacts.

- Pinch: also called a parallel grasp, wherein the contacting links exert a single axis of opposing force on the artifact.
- Wrap: also called a power grasp, wherein multiple contacting surfaces, such as finger links and palm, exert a vector sum of opposing force on the artifact.

Different artifact designs should be used based on the characteristics of the end-effector and the type of grasp being evaluated. See Figure 2 for examples of each grasp and corresponding artifact design. Note that the test artifact can only be used to evaluate grasps where all fingers that make contact with the object travel along planes parallel to one another (i.e., concentric grasps cannot be evaluated). In addition to the load cells, dowel pins are used to limit shearing effects when testing. Rubber bands may also be placed with minimum tension around the outside of the artifacts to prevent the two halves from separating.

A range of different-sized artifacts is used to generate a spread of performance results for each end-effector and grasp type. This range is determined empirically or in simulation based on the minimum- and maximum-sized artifacts that satisfy the stated grasp type criteria. It is recommended that at least three sizes be used: minimum, maximum, and median (midway between minimum and maximum). Constraints on artifact size include a minimum size based on load cell dimensions and the inability to satisfy grasp type conditions. Design files for fabricating test artifacts are provided ranging from 45 - 135mm in width in 5mm increments, but any size artifact can be used.

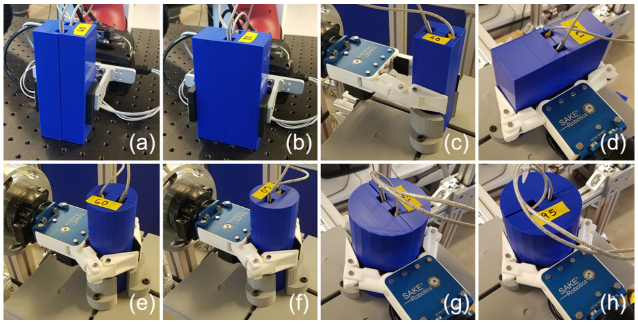

Fig. 2. Grasps and varying artifact sizes and shapes. Parallel jaw gripper performing pinch grasps on min (a) and max (b) size rectilinear artifacts. Tendon-driven multi-link gripper performing pinch grasps on a min (c) and max (d) size rectilinear artifact, and wrap grasps on min 0° (e), min 90° (f), max 0° (g), and max 90° (h) cylindrical artifacts.

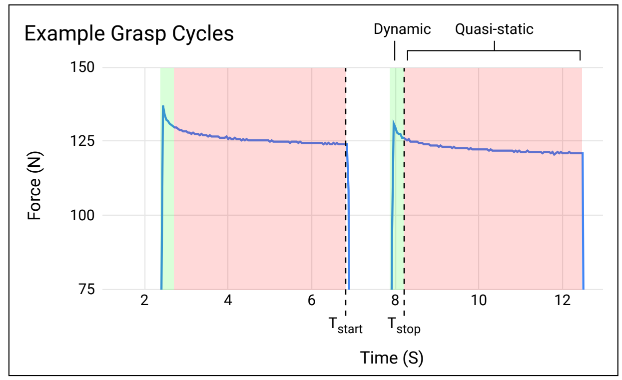

Fig. 3. Example of two grasp cycles showing the dynamic and quasi-static force regions for calculating grasp strength and the start and stop timing for calculating grasp cycle time.

B. Test Method Protocol

The same protocol is used for pinch and wrap grasps. During all repetitions, the orientation of the artifact relative to the end-effector should be maintained. Adjustments to the artifact orientation may be necessary if the end-effector moves it out of alignment. It is also necessary that the artifact is unconstrained throughout testing, in order to prevent any external forces that will result in erroneous readings. For pinch grasps, the axis of the load cells should be perpendicular to the finger surfaces when in contact with the artifact (0° orientation). For wrap grasps, two orientations of the load cells are possible: when their axis is parallel to the palm surface (0° orientation) and when it is perpendicular (90° orientation). The dominant orientation is determined to be the maximum force that the end-effector exerts on the artifact for a particular grasp type. This may be determined empirically prior to running the test method. The artifact should then be rotated 90°to test the non-dominant artifact orientation. Force readings for 0° orientation and 90° orientation should be reported separately (not combined).

The test artifact is placed in a standing position on a platform. The end-effector should be mounted such that its position is secure, its palm surface is parallel to the surface of the test artifact, and the center of its contact points with the artifact aligns approximately with the center of the artifact. Figure 1 shows a test setup utilizing an adjustable aluminum frame to maintain position of the test artifact and the end-effector.

The following steps should be performed for each artifact size and orientation:

1) Position and orient the artifact within grasping reach of the end-effector such that it can be grasped without causing the artifact to move significantly.
2) Under position control, command the end-effector to open completely.
3) Under position control, command the end-effector to close completely to induce control saturation producing the maximum force exertion grasp.
4) Once maximum force exertion is established for five seconds, the end-effector is retracted to its start position in its fully opened state. Reposition the artifact if necessary.
5) Repeat steps 2-4 for a minimum of 32 cycles.
6) Record force sensor data throughout the test.
7) Calculate the performance measures.

C. Performance Measures

For each set of instantaneous force readings, add forces across all load cells since they are in-line to yield a total grasp force, F_{total}:

$$F_{total} = \sum_{i=1}^{n} F_i \tag{1}$$

Next, the quasi-static force for each grasp cycle should be extracted. Quasi-static grasp forces (see Figure 3) are chosen for evaluation as they remove impact effects and give a more accurate estimate of the true strength of the end-effector. For the tools we provide to automate calculation of quasi-static force (see Section III-F), the first and last 10% of all force readings greater than 0 for each grasp cycle are removed and the trimmed middle 80% is used to calculate the mean F_{total} of each grasp cycle. Using the calculated means of all grasp cycles, compute the mean, standard deviation, and 95% confidence interval of the test results.

The type of grasp performed, artifact shape, size, and orientation should be reported alongside the mean force, standard deviation, and 95% confidence interval results of a test. For comparing the performance of two different end-effectors, the same grasp type and size type (min, med, max) should be used.

D. Example Implementation and Results

To demonstrate this protocol, we performed grasp strength testing of two end-effectors (Robotiq 2F-85 Gripper and SAKE Robotics EZ Gripper) and used the three artifact sizes (min, median, max) scaled according to each end-effector's characteristics. See Table I for a spread of anonymized comparative data between the two robotic end-effectors and multiple artifact types. This is an example of how data generated from this protocol should be reported.

TABLE I
EXAMPLE GRASP STRENGTH DATA FROM TWO ROBOTIC END-EFFECTORS. ARTIFACT SIZE DIMENSIONS ARE NOT SHOWN IN ORDER TO FURTHER ANONYMIZE THE DATA.

End-effector	Grasp type	Orientation	Artifact size	Avg total force (N)	Stdev (N)	95% confidence interval (N)
A	Pinch	0°	Min	16.21	1.11	[15.83, 16.59]
			Med	17.17	0.40	[17.03, 17.31]
			Max	23.88	0.38	[23.75, 24.01]
B	Pinch	0°	Min	171.53	3.85	[170.2, 172.86]
			Med	235.31	4.93	[233.6, 237.02]
			Max	274.33	2.59	[273.43, 275.23]
A	Wrap	0°	Min	19.71	0.46	[19.55, 19.87]
		90°		80.78	2.70	[79.84, 81.72]
		0°	Med	26.79	1.74	[26.19, 27.39]
		90°		39.55	1.65	[38.98, 40.12]
		0°	Max	14.77	0.50	[14.6, 14.94]
		90°		22.09	0.73	[21.84, 22.34]
B	Wrap	0°	Min	194.45	1.67	[193.87, 195.03]
		90°		252.75	2.64	[251.84, 253.66]
		0°	Med	162.55	0.89	[162.24, 162.86]
		90°		199.16	1.94	[198.49, 199.83]
		0°	Max	138.84	8.04	[136.05, 141.63]
		90°		160.29	3.19	[159.18, 161.4]

For both end-effectors, dominant wrap grasps occurred on the 90° orientation artifact with lower force readings on the 0° orientation. Also, pinch grasp strength increased as artifact size increased and wrap grasp strength decreased as artifact size increased. These relationships may not be true for all end-effectors; therefore, it is recommended that both orientations of wrap grasps be tested and that multiple artifact sizes be used.

E. Discussion

When performing pinch tests, some end-effectors (in particular those that have multi-link fingers and are tendon-driven) are observed to pull objects inwards towards the palm after making contact with its distal links. For testing purposes, in such cases it is important not to constrain the artifact as this will cause an unwanted shearing force, even if this means that a pinch grasp is not possible because the artifact is naturally pulled into a wrap grasp. Additionally, tendon-driven end-effectors can show decreased efficiency in repeated cycles. To establish a baseline grasp strength for a tendon-driven end-effector, it may be necessary to perform an endurance test with a cycle lasting 2-3 minutes and/or measuring grasp strength after a few hundred cycles in addition to the proposed method. The result of this endurance test should be reported alongside test data.

F. Protocol Tools Available for Download

1) Split block/cylinder artifact design files. A variety of artifact sizes are provided (45 - 135mm in 5mm increments), but the model dimensions can be further modified to create new artifact sizes. Note: this design is made to work with FUTEK LCM300 load cells, but can be modified to support other types of sensors.
2) Software to support data collection and automated analysis of load cell data to extract quasi-static force and calculate grasp strength.
3) Raw and processed data from the example implementation tests.

IV. GRASP CYCLE TIME

Grasp cycle time is a kinematic metric defined as the minimum time required for an end-effector to cycle from a pre-grasp configuration, to a grasp position, and back to the pre-grasp configuration [1], [10]. This measure will yield information regarding the opening and closing speeds of an end-effector for a given object. This test protocol can be used to benchmark the grasp cycle time for comparative metrics between end-effectors. This metric is particularly important for manufacturing operations as it will influence the resulting throughput rate that is achievable by a robot system. Metrics for grasp cycle time can be extracted from the performance of grasp strength tests, provided that the time in between grasps is minimized as described in the protocol.

A. Test Artifact and Measurement Method

The test artifact used is the split cylinder or split block design consisting of two 3D printed structures that hold two or more single-axis force load cells (see Figure 1). Considerations for artifact dimensions and shape can be found in Section III-A. The size of the artifact will directly influence grasp cycle time, as larger artifacts will require less travel time for the fingers to make contact and therefore grasp.

B. Test Method Protocol

The setup for the grasp cycle time test method is the same as that of grasp strength: maintaining test artifact orientation (0° or 90°), positioning the test artifact (upright), positioning the end-effector (parallel to the surface of the test artifact, contact points centered), etc. (see Section III-B). The protocol is also very similar; however, timing between open and closed states should be minimized. Whenever possible, use the end-effector's communication protocol to command it to close as soon as its open state is reached. In the case that an end-effector does not register its open state, the state must be determined empirically.

The following steps should be performed for each artifact size:

1) Position and orient the artifact within grasping reach of the end-effector such that it can be grasped without causing the artifact to move significantly.
2) Under position control, command the end-effector to open completely. The next step should occur as soon as the end-effector has finished opening.
3) Under position control, command the end-effector to close completely to induce control saturation producing the maximum force exertion grasp.
4) Once maximum force exertion is established for five seconds, the end-effector is retracted to its start position in its fully opened state. Reposition the artifact if necessary.
5) Repeat steps 2-4 for a minimum of 32 cycles.
6) Record force sensor data throughout the test.
7) Calculate the performance measures.

TABLE II
EXAMPLE GRASP CYCLE TIME DATA FROM FOUR ROBOTIC END-EFFECTORS.

End-effector	Grasp type	Orientation	Artifact size (mm)	Avg cycle time (s)	Stdev (s)	95% confidence interval (s)
C	Pinch	0	60	1.01	0.08	[0.98, 1.04]
D				0.60	0.06	[0.58, 0.62]
E	Wrap	90	50	1.41	0.10	[1.37, 1.45]
F				4.04	0.10	[4.00, 4.08]

C. Performance Measures

A grasp cycle begins as soon as the robotic end-effector begins closing its fingers from its initial pose and ends once it has finished opening its fingers after performing a grasp. $T_{grasp\ cycle}$ is determined using two adjacent cycles from T_{start} and T_{stop} as shown in Figure 3. T_{start} occurs at a cycle end where the quasi-static force is removed from the object and T_{stop} occurs at the point where the dynamic force has settled to a quasi-static state.

$$T_{grasp\ cycle} = T_{stop} - T_{start} \tag{2}$$

Given the timing of the quasi-static grasp forces, compute the mean cycle time and 95% confidence intervals across all recorded grasp cycle times. When comparing the performance of two different end-effectors, the same artifact shape and size should be used.

D. Example Implementation and Results

To demonstrate this protocol, we performed grasp cycle time testing of four robotic end-effectors; two performed pinch grasps on the same size rectilinear artifact (Rethink Robotics Pneumatic Gripper and Robotiq 2F-85 Gripper) and two performed wrap grasps on the same size cylindrical artifact (Robotiq 3-Finger Gripper and Robotiq 2F-85 Gripper). T_{start} and T_{stop} as shown in Figure 3 were extracted from the recorded artifact force profiles. To anonymize the performance data, each end-effector has been code named with a letter. See Table II for comparison data between the robotic end-effectors.

E. Discussion

The determination of the transition from the dynamic grasp state to quasi-static grasp is still under development. The settling time is a function of the end-effector controller and sensory feedback. A protocol is currently in development to establish an end-effector-agnostic understanding of when a grasp cycle enters a quasi-static range. In cases where a grippers post-contact reconfiguration moves the artifact out of place, manual readjustment of the artifact after the grasp may be required.

F. Protocol Tools Available for Download

1) Split block/cylinder artifact design files (same as those specified in Section III-F).
2) Software to support data collection and automated analysis of grasp cycle time.
3) Raw and processed data from the example implementation tests.

V. FINGER STRENGTH

Finger strength is a kinetic measure of the maximum force a robotic finger can impose on its environment. This measure relates to the overall strength of the end-effector during grasping [1], [10]. Individual finger strength can be used to resolve grasp capability where all fingers are not in contact with an object in comparison to the grasp strength metric. These tests also identify variability across different fingers, including mechanically equivalent fingers. This metric is related to grasp strength as the individual strengths of each finger on an end-effector will influence its overall grasp strength.

A. Test Artifact and Measurement Method

The strength of each finger is measured by making contact with a force sensor. A single axis force sensor may be used; however, a six-axis force torque sensor will compensate for off-axis alignment error. The sensor contact surface height may need to be adjustable (e.g., by use of extruded aluminum) to accommodate pose and contact for each finger and prevent collisions with the sensor or the environment. The contact surface shape should also be designed such that the force exerted from the robotic finger is normal to the sensor.

A finger strength test setup can be seen in Figure 5 that uses a six-axis force torque sensor with custom mounting plates fixing the sensor to a table surface and an extruded aluminum column to the top of the sensor. A 3D printed cap is attached to the top of the aluminum column.

B. Test Method Protocol

The artifact placement relative to the tested finger is vital to testing the force consistently and correctly. Consider this relationship of location the “finger-object orientation.” The significant finger-object configuration for benchmarking occurs when the induced moment arm from making contact is at its maximum which means the maximum attainable contact force will be at a minimum for the finger under test. For most robotic hand designs, this occurs when a finger is fully extended and all finger links are extended in the same direction. This is the preferred orientation for comparing different hand designs, or specifying hand characteristics.

Due to the varying designs of end-effectors, the finger-object orientation may be different for each finger. Regardless of initial orientation, each of the fingers should contact the artifact at a position perpendicular to the axis of the load cell so that the applied force in the testing direction is at its maximum possible.

The following steps should be performed for each finger under test:

1) Configure the end-effector such that only the finger being tested is within reach of the test artifact, position the finger under test just above the force sensor, and verify a zero force reading from the test artifact sensor.
2) Under position control, command the finger to open completely.
3) Under position control, command the finger to close completely to induce control saturation and apply maximum force to the load cell.

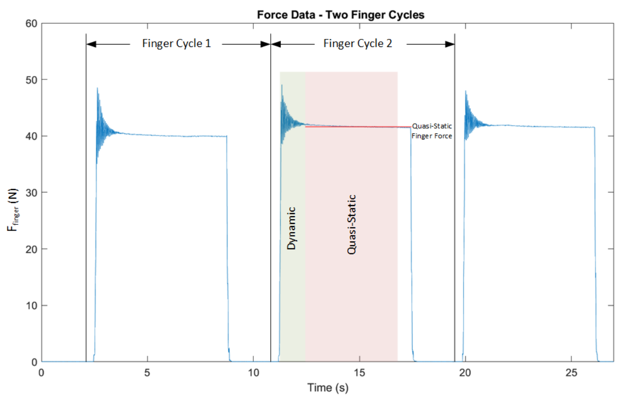

Fig. 4. Example of two finger cycles showing the dynamic and quasi-static force regions. The quasi-static force region is used to calculate finger strength.

4) Once the maximum force is established for a few seconds, the finger is then commanded to open completely to fully disengage the artifact.
5) Repeat steps 2-4 for a minimum of 32 cycles.
6) Record force sensor data throughout the test.
7) Calculate the performance measures.

C. Performance Measures

The fingertip contact force magnitude, F_{finger}, should be computed as

$$F_{finger} = \sqrt{(F_x^2) + (F_y^2) + (F_z^2)} \tag{3}$$

for each set of force readings given by the sensor. Next, the contact force magnitude from the quasi-static force region (see Figure 4) should be extracted for each load cycle, and then averaged to yield the maximum finger strength, $F_{finger,max}$. Collect the maximum forces and compute the mean, standard deviation, and 95% confidence interval.

D. Example Implementation and Results

This protocol was demonstrated using two different commercially available end-effectors, a Robotiq 3-Finger Adaptive Robot Gripper and a Schunk Dexterous Hand. Finger force performance data for the two end-effectors under test was calculated and analyzed. To anonymize the data, each end-effector is given a numerical value: end-effector 1 (EE1) and end-effector 2 (EE2)). Each end-effector's fingers are labeled as F1, F2, or F3. See Figure 6 and Table III for comparison data between the six fingers. The results indicate an overall greater finger strength of EE2 over EE1 with a variation between finger strength in the order of 2.5 N for EE1 and 4 N for EE2.

VI. Finger Repeatability

Finger repeatability is a kinematic measure of the difference in pose results when a finger is commanded to a position multiple times from the same direction. This measure will yield information regarding the ability of an end-effector to re-establish a pose. The mechanical design of an end-effector will

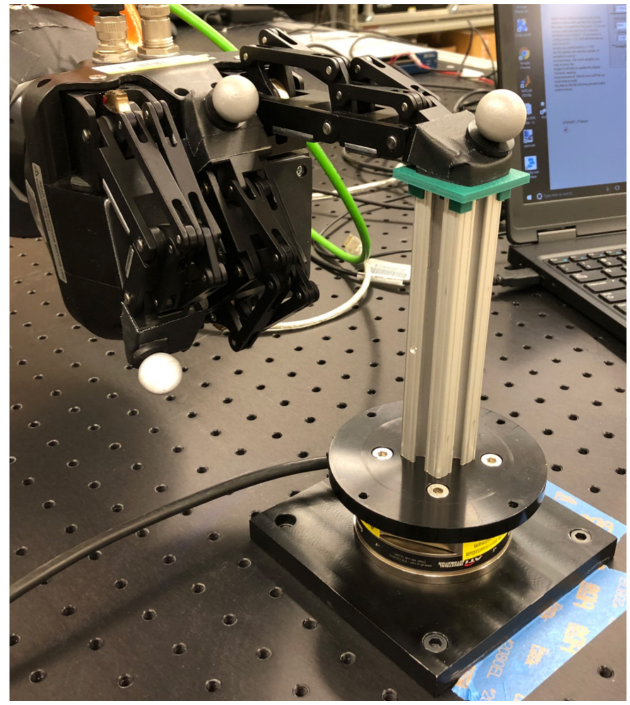

Fig. 5. Test setup for finger strength using a Robotiq 3-Finger Adaptive Robot Gripper. A fully extended finger applies its maximum force to a six-axis load cell through an extruded aluminum stand-off with a 3D printed cap.

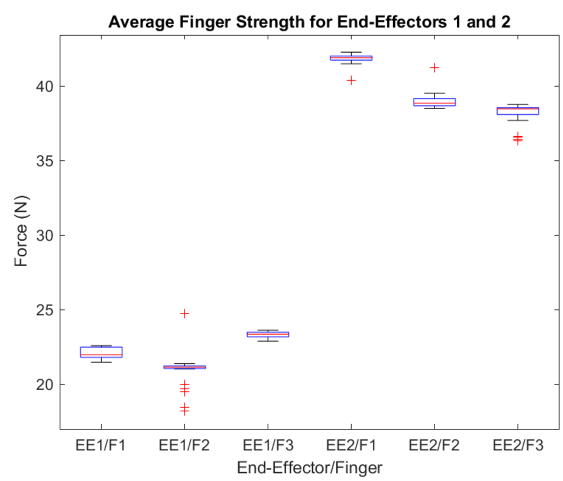

Fig. 6. Example finger strength data for six fingers across two end-effectors.

TABLE III
EXAMPLE FINGER STRENGTH DATA FOR SIX FINGERS ACROSS TWO END-EFFECTORS.

Finger	Avg finger strength (N)	Stdev (N)	95% confidence interval (N)
EE1 / F1	22.09	0.34	[21.97, 22.22]
EE1 / F2	20.93	1.08	[20.54, 21.32]
EE1 / F3	23.32	0.20	[23.24, 23.39]
EE2 / F1	41.84	0.32	[41.72, 41.96]
EE2 / F2	38.97	0.50	[38.79, 39.14]
EE2 / F3	38.13	0.75	[37.86, 38.40]

affect its ability to consistently obtain the same pose. Design considerations include motor/encoder selection as well as the drive technology used to position components of the hand or gripper. In addition, robotic hands performing fine motor-controlled tasks may require a high level of repeatability. While the degree of repeatability required by a robotic hand varies depending on the task to be performed, it can be advantageous to understand this characteristic prior to selecting the end-effector to support a particular application. Moreover, performance of a task that requires achieving the same pose continuously over a length of time can slowly degrade. In order to fully understand the repeatability of an end-effector, each of its individual fingers should be tested using a series of continuous poses.

A. Test Artifact and Measurement Method

The measurement method and supporting artifacts can vary depending on the expected performance of repeatability. A motion capture system can provide finger position data in 3D space with typical $\pm$ 0.5 mm accuracy for an internally calibrated stereo based camera system. Use of a plunger-style variance indicator (see Figure 7) for finger contact measurements will provide measurements in a single direction within $\pm$ 0.025 mm or $\pm$ 0.0025 mm depending on the accuracy of the indicator. Use of an indicator may require additional tests to resolve deviations in 3D space; trajectories to the final repeatability measurement pose should fall along the plunger axis to avoid shearing forces on the variance indicator that could lead to measurement error or damage to the measurement instrument. In addition, the rigidity of the robot mount and supporting structure should be sufficient as to avoid the introduction of additional repeatability errors by movement of the end-effector base.

A variation in testing could include the addition of weight to each finger based on some overall percentage of finger strength. Such a test could provide insight during the end-effector design process.

B. Test Method Protocol

While there are slight differences in the setup and data recording of the test method depending on the measurement method that is used, the protocols for the test methods are the same. In the case of using a plunger-style variance indicator, the artifact placement relative to the tested finger is vital to testing the repeatability consistently and correctly. Use of a motion capture system requires finger trajectories that avoid occlusions between the camera and the measured targets. To conduct this testing protocol, the finger will be commanded to four unique poses (home, first, second, and third). Each pose should have a unique position and orientation compared to the other poses.

Regardless of measurement method, the following steps should be performed for each finger under test:

1) Command the finger under test to an initial home position and record the home position.
2) Command the finger to go to another separate pose that actuates each of the individual joints to disengage completely from the home position.
3) Command the finger to a second disengaged pose.
4) Command the finger to a third disengaged pose.
5) Command the finger back to its home position and record this position and error relative to the initial home position.
6) Repeat steps 2-5 for a minimum of 32 cycles.
7) Calculate the performance measures.

C. Performance Measures

The displacement offset between initial home position and each evaluated home position, ΔS, should be computed as:

$$\Delta S = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (x-x_0)^2} \qquad (4)$$

For each set of displacement readings given by the pose measurement system, compute ΔS over the desired number of cycles. For each set of cycles, compute the mean, standard deviation, and 95% confidence interval.

D. Example Implementation and Results

Repeatability measurements were conducted using two commercially available three-fingered end-effectors: a Robotiq 3-Finger Adaptive Robot Gripper and a Schunk Dexterous Hand. Two different measurement approaches were investigated and implemented. The first used a motion capture system to locate the end-effector fingers in 3D space by identifying reflective sphere markers attached to the fingertips of the gripper and hand. Initial test results indicated that measured repeatability exceeded the measurement capabilities of the motion capture system.

The second implementation used a digital plunger-style variance indicator with an uncertainty of 0.0025 mm. This particular unit was capable of logging data directly to a computer. The end-effectors are mounted to a structure fabricated using aluminum extrusions to minimize measurement errors associated with movement of the base when contacting the indicator (see Figure 7). The gauge is positioned such that its plunger is resting against a hard, flat surface of the outer most jointed section of the finger, perpendicular to the surface and at least 50 percent engaged. Each finger is commanded to run through 4 poses, each fully disengaging from the previous. Upon returning to the home position a slight pause is given to allow for the finger to settle into a quasi-static position.

Results from testing are shown in Figure 8 and the average, standard deviation and 95% confidence interval for each finger are reported in Table IV. The results indicate that end-effector 1 (EE3) has a better repeatability than end-effector 2 (EE4) for the given test trajectories despite the fact that EE3 has several outliers. In general, both hands show good repeatability characteristics.

VII. Conclusions

This paper describes a set of metrics and supporting benchmarking protocols for determining the performance characteristics of robot end-effectors. The benchmarking protocols

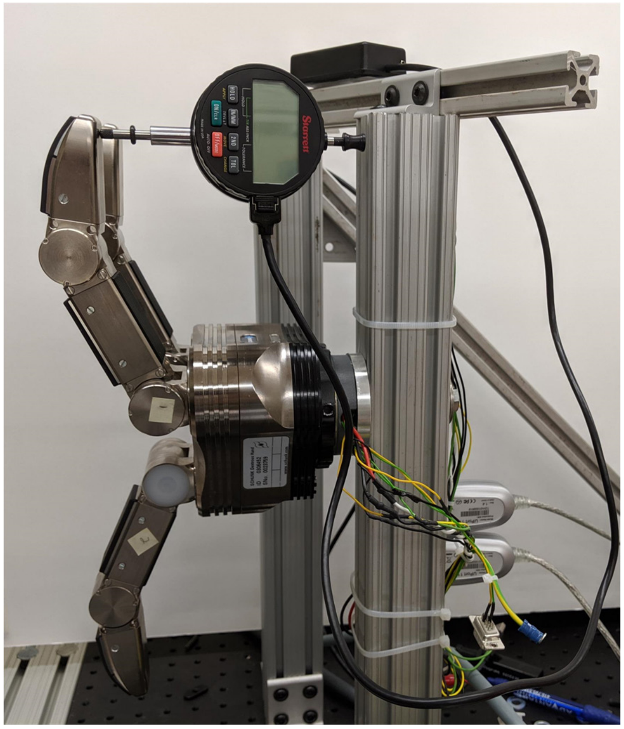

Fig. 7. Test setup for finger repeatability of a Schunk Dexterous Hand.

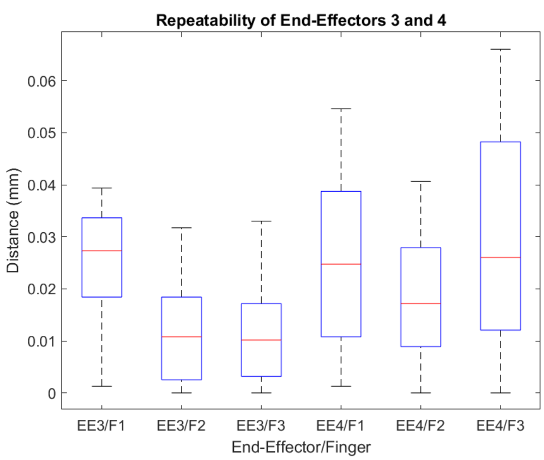

Fig. 8. Example finger repeatability data for six fingers across two end effectors.

TABLE IV
EXAMPLE FINGER REPEATABILITY DATA FOR SIX FINGERS ACROSS TWO END EFFECTORS.

Finger	**Avg offset distance (mm)**	**Stdev (mm)**	**95% confidence interval (mm)**
EE3 / F1	0.0252	0.0101	[0.0216, 0.0288]
EE3 / F2	0.0115	0.0093	[0.0081, 0.0149]
EE3 / F3	0.0116	0.0091	[0.0083, 0.0149]
EE4 / F1	0.0254	0.0155	[0.0198, 0.0310]
EE4 / F2	0.0183	0.0127	[0.0138, 0.0229]
EE4 / F3	0.0295	0.0199	[0.0224, 0.0367]

included here are grasp strength, grasp cycle time, finger strength, and finger repeatability, four of the eleven protocols that have been defined to date using a community-driven approach through the IEEE RAS RHGM technical committee (www.rhgm.org) to define metrics and test methods so that researchers can benchmark and make comparisons between end-effector designs as well as sensor and algorithm implementations. The application of these metrics and protocols is demonstrated using example implementations to characterize a variety of robotic hands and grippers. Example data sets and test designs are provided as supplemental material to help accelerate the use of these tools as well as to better understand the caveats associated with implementing robust benchmarks to evaluate grasp-type end-effectors. The authors welcome continued input from the community using [1] as a working document that will be the eventual starting point for a standardization effort. While a valuable tool for researchers in the short-term, the long-term goal is a standard framework for providing technical specifications for robotic end-effectors to help pair technologies to application spaces.

VIII. ACKNOWLEDGEMENTS

This work is funded in part at the University of Massachusetts Lowell by the National Institute of Standards and Technology (NIST) under 70NANB17H256.

REFERENCES

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

[2] J. Falco, J. Marvel, and E. Messina, "A roadmap to progress measurement science in robot dexterity and manipulation," NIST, Tech. Rep. NIST Report 7993, 2014.

[3] G. A. Kragten, C. Meijneke, and J. L. Herder, "A proposal for benchmark tests for underactuated or compliant hands," *Mechanical Sciences*, vol. 1, no. 1, pp. 13–18, 2010.

[4] C. Meijneke, G. A. Kragten, and M. Wisse, "Design and performance assessment of an underactuated hand for industrial applications," *Mechanical Sciences*, vol. 2, no. 1, pp. 9–15, 2011.

[5] L. U. Odhner, L. P. Jentoft, M. R. Claffee, N. Corson, Y. Tenzer, R. R. Ma, M. Buehler, R. Kohout, R. D. Howe, and A. M. Dollar, "A compliant, underactuated hand for robust manipulation," *The International Journal of Robotics Research*, vol. 33, no. 5, pp. 736–752, 2014.

[6] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, "Human-inspired robotic grasp control with tactile sensing," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1067–1079, Dec 2011.

[7] A. M. Dollar, L. P. Jentoft, J. H. Gao, and R. D. Howe, "Contact sensing and grasping performance of compliant hands," *Autonomous Robots*, vol. 28, no. 1, p. 65, Aug 2009.

[8] L. U. Odhner, R. R. Ma, and A. M. Dollar, *Experiments in Underactuated In-Hand Manipulation*. Heidelberg: Springer International Publishing, 2013, pp. 27–40.

[9] J. Falco, K. Van Wyk, S. Liu, and S. Carpin, "Grasping the performance: Facilitating replicable performance measures via benchmarking and standardized methodologies," *IEEE Robotics & Automation Magazine*, vol. 22, no. 4, pp. 125–136, 2015.

[10] J. Falco, K. Van Wyk, and E. Messina, "Proposed standard terminology for robotic hands and associated performance metrics," NIST, Tech. Rep. DRAFT NIST Special Publication 1229, 2018.

[11] Robotiq, "Robotiq 2f-85 2f-140 for e-series universal robots instruction manual," 2018. [Online]. Available: https://assets.robotiq.com/website-assets/support_documents/document/2F-85_2F-140_Instruction_Manual_e-Series_PDF_20190206.pdf