Grasping the Performance: Facilitating Replicable Performance Measures via Benchmarking and Standardized Methodologies

Joe Falco Karl Van Wyk Shuo Liu Stefano Carpin

I. Introduction

"It is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult to give them the skills of a one-year old when it comes to perception and dexterity" [1]. More than fifteen years after it was first stated, Moravec's paradox still holds true today. Fueled by vigorous research in machine learning, the gap has consistently narrowed on the perception side. But most of the fine manual motor skills displayed by a toddler are to-date far beyond what robots can do. It is of course true that many valuable tasks involving physical interaction with objects can be solved by contemporary robots as indicated by a thriving industrial robotics sector. However, in the future, robots are expected to work side-by-side with humans in unstructured environments, and the ability to reliably grasp and manipulate objects used in everyday activities will be an unavoidable requirement. Today's robots are far from being ready for this challenge.

Multiple reasons account for this persistent gap — some are technological, whereas others are methodological. We postulate that one of the major obstacles inhibiting progress in this area surrounds the continued inability to replicate and compare results generated by the grasping community. As in other robotics sub-domains, too many researchers lack the ability to contrast their findings with those of their peers using a principled, accepted methodology. With a surge of robotic hand designs [2]–[5] and algorithmic developments, it is difficult to make informed decisions or draw quantitative conclusions regarding commercial products or lab prototype performances in various task-related settings.

Standardized performance testing is an emerging tool within the robotics community that is proving itself worthy in other robotics community disciplines and offers unbiased evaluation methods to assess how well a system performs in a particular ability [6]. The results of such evaluations and benchmarks help match capabilities to end-user needs, and provide developers insight for improving their product designs. Accordingly, we created a subset of grasp performance metrics, test methods, and example experiments to be vetted by the robotic hand research community. Together, accelerated progress can be made in this critical area that will lead to a set of standard measures and performance test methods. Such standards will guide the development of future grasping technologies and allow for seamless system benchmarking.

This document presents the beginnings of a framework for robotic hand performance benchmarking. Many of the concepts presented are the results of an informal working group[1] organized by the National Institute of Standards and Technology (NIST) that continues as part of the recently formed IEEE Robotics and Automation Society (RAS) Robotic Hand Grasping and Manipulation (RHGM) Technical Committee[2]. A successful grasp is the combination of an appropriate coupling of hardware (i.e., robot hands) and algorithmic components (e.g., grasp planners and grasp control). Both sides have to be benchmarked and the assessment process needs to be repeatable. This manuscript introduces a subset of hardware and control benchmarks that are demonstrated using a set of robotic hand platforms[3]. A parallel effort aimed at defining best practices for benchmarking and comparing grasp planning algorithms is also being developed, but will not be discussed here due to space limitations.

This paper is organized as follows: a) proposed test methods to characterize robotic hand performance are introduced in Section II, and b) detailed descriptions and experimental results for a subset of these methods are presented in Section III. The material presented in this paper is meant to propose a path to develop replicable[4] performance measures for robotic grasping, but is not intended as the definitive methodology. It is foreseeable and

Joe Falco and Karl Van Wyk are with the National Institute of Standards and Technology, Gaithersburg, MD, USA.
Shuo Liu and Stefano Carpin are with the School of Engineering, University of California, Merced, CA, USA.
S. Liu and S. Carpin are partially supported by the National Institute of Standards and Technology under cooperative agreement 70NANB12H143. Any opinions, findings, and conclusions or recommendations expressed in these materials are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the funding agencies of the U.S. Government.

[1]Participants included Peter Allen, Colombia University; Stefano Carpin, UC Merced; Mark Cutkosky, Stanford University; Aaron Dollar, Yale University; Erik Engeberg, University of Akron; Joe Falco, NIST; Karl Van Wyk, NIST; Jeremy Fishel, SynTouch; JP Jobin, Robotiq; Gerald Loeb, SynTouch; Elena Messina, NIST; Jane Shi, General Motors; Cathy Parker, re2.

[2]IEEE RAS Technical Committee on Robotic Hand Grasping and Manipulation; http://www.rhgm.org

[3]This particular set of hands was chosen based on their availability, and it is expected that more hands will be tested by NIST as they become available.

[4]More details regarding the actual tests and data analysis will be available at http://www.nist.gov/el/isd/grasp.cfm (under construction).

desirable that the core ideas presented in this manuscript will be extended through a community-driven approach. To the best of our knowledge, no comparable effort has been formulated in the past and this is the first paper detailing a comprehensive methodology for repeatable research in robotic hand technology that spans both hardware and software components. Repeatability in grasping or any other robotic subdomain requires conscious commitment by the researchers to share information (designs, data, models, code, etc.) in an open, understandable format – a mentality we hope to inspire with the ideas presented here.

II. Grasp Performance Tests

Presented herein are the physical measurements for assessing performance of robotic hands using measurement techniques external to the system under test. In the literature, physical results of grasping are reported using both qualitative and quantitative data. Qualitative measures are easily found in robotic grasping research literature, however, examples of applying quantitative measures to evaluate grasp performance are sparse and have only been developed formally for prosthetics [7]–[11].

When evaluating the capabilities of a robotic hand, performance tests should be agnostic to the other system components such as the robot arm and perception system. While it is possible to access data directly from a robotic hand and derive the defined metrics, these measurements would be based on the inherent properties of the system under test. Therefore, independent measurement systems must be developed to support testing to allow for comparative metrics between systems to establish extrinsic ground truths.

Breaking down a problem into its parts can provide novel insights towards its solution. In particular, consider the underlying tasks associated with a robotic pick and place operation for a fully integrated multi-fingered robotic hand as shown in Fig. 1[5]. Each task in this particular operation possesses a number of associated problems that can serve as a basis for extracting performance measures. Furthermore, identifying the significance of particular performance measures for different grasping tasks would provide valuable knowledge on necessary functionalities towards task completion.

Robotic hands are an integrated system of mechatronics, sensors, and control algorithms with considerable variability in their number of degrees of actuation, degrees of freedom, and joint types. Furthermore, there exists a variety of touch sensing strategies across different platforms. Advanced sensing capabilities include fingertip embedded 6-axis load cells [12], pressure-sensitive tactile sensors, vision-based

[5]This concept was introduced within the NIST-organized grasp metrics informal working group by SynTouch LLC. The terminology chosen here was based on that used throughout the community and is not the result of any single author.

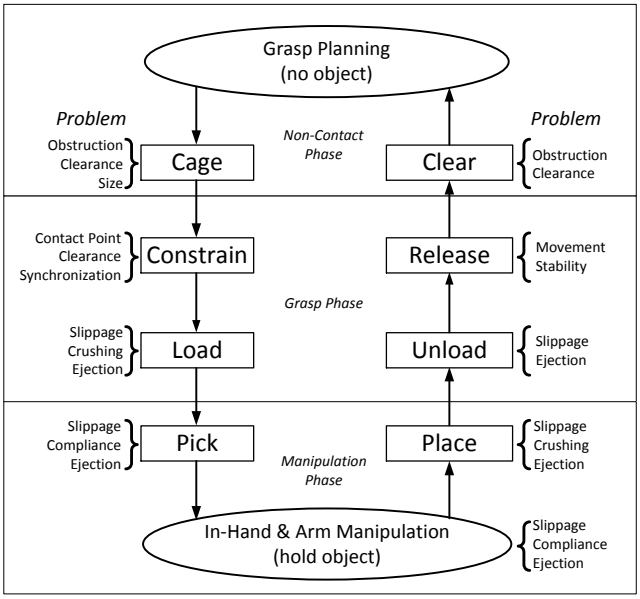

Fig. 1. Pick and place task segmentation and transitions between grasped and un-grasped states, and examples of potential problems.

contact sensing [13], and bio-inspired impedance tactile sensing capable of resolving point of contact, shear, and normal forces, as well as other sensing modalities such as vibration and temperature [14]. Consequently, the design space for a robotic hand is enormous and requires a modular set of performance metrics and associated test methods that can be used to draw direct comparisons between different hand designs at various levels of application.

To address this need, a framework for benchmarking the performance of robotic hands is proposed and shown in Fig. 2. As depicted, a modular set of performance test methods can be chosen based on various considerations: a taxonomy of grasp types the hand can perform well, a scheme for classifying a hand that includes sensing and control capabilities, relevant test metrics, and a common set of test objects (artifacts). The desirable and relevant performance metrics can then be extracted at various levels with component, system, and functional tests.

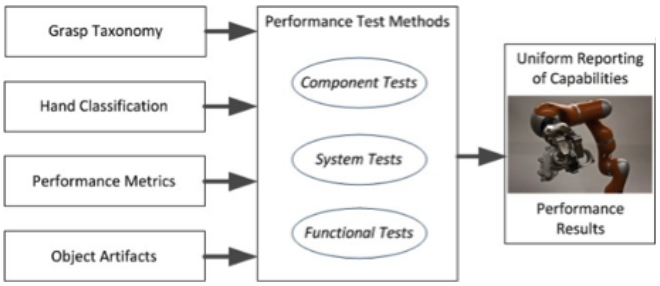

Fig. 2. Framework for standardized benchmarking of robotic hands

Component-based performance characteristics include

kinematic properties such as volumetric capabilities and grasp configurations, and kinetic properties such as hand strength (measured by force). Sensors can be tested at their stock sensing modalities for properties such as resolution, sensitivity, and latency. System tests seek to characterize realizable functionalities for the robotic hand pending proper implementation of additional controllers, machine learning algorithms, and calibration. Finally, functional tests are designed to evaluate a robotic hand at the task level. This includes purposeful grasping and manipulation, and can be paired with additional systems such as a vision system and a robotic arm. A summary of proposed component and system level tests are listed in Tables I and II, respectively. Functional tests are not being addressed in this paper. Additional tests will be added as this effort progresses.

TABLE I

PROPOSED COMPONENT LEVEL TESTS TO SUPPORT KINETIC AND KINEMATIC HAND PROPERTIES.

Component Test Method	**Description**
Volumetric Capability	Minimum and maximum volumetric capabilities based on primitive shape and grasp types supported by the robotic hand.
Part Acquisition	Minimum sized primitive artifact that can be grasped from a flat surface.
Touch Sensitivity	Smallest contact force that indicates the presence of touch on a robotic finger (see Section III-A).
Finger Strength	Maximum force that can be exerted by a fully extended finger at the fingertip (see Section III-B).
Grasp Strength	Maximum pinch and power grasp forces on a standard test artifact (see Section III-C).
Closing Time	Time it takes to completely close a fully-extended robotic finger.
Cartesian Range of Motion	Localized operating volume for a finger.
Repeatability	Position repeatability of a finger at a set of points within the Cartesian Range of Motion.
Slip Resistance	Force associated with the onset of slip on a defined surface over a range of fingertip forces, up to the maximum force (see Section III-D).
Surface Covering Compliance	Stiffness properties for key contact surfaces on the robotic hand that include compliant coverings/pads and surfaces of sensing components.
Bandwidth	Measures of cycle times for reading sensors and writing controls

III. TEST METHODS AND PERFORMANCE MEASURES

In order to design relevant performance metrics and methods for characterizing robotic hands, it helps to understand the issues surrounding robotic grasping and manipulation. Regardless of the actual task at hand, any grasping and manipulation problem can be broken down into its first principles, kinetics and kinematics, or more simply, effort and motion. Kinetics are the forces acting on bodies or particles that are responsible for causing their motion. In particular, any kinetic metric or test method will be evaluating force, torques, and any other measure of effort such as electrical current. Kinematics is the geometry of motion of bodies or particles with complete disregard for the forces that cause such motion. Therefore,

TABLE II

PROPOSED SYSTEM LEVEL TESTS TO SUPPORT TESTING THE FUNCTIONALITY.

System Test Method	**Description**
Sensor Based Grasp Efficiency	A measure of the ability to maintain an efficient grasp on an object while adapting to external forces.
In-Hand Manipulation	The positioning accuracy and range of motion when manipulating an object within a grasp.
Hand Stiffness	The stiffness properties of a grasp inducing external forces and measuring associated displacements on a grasped object.
Finger Force Tracking	The performance of fingertip force tracking (see Section III-E).
Force Calibration	The accuracy of calibrated sensors in determining contact force magnitude and directions (see Section III-F).

any kinematic metric or test method will be concerned with evaluating positions, velocities, or accelerations of bodies, parts, or particles, and will typically be in units of length and time. Candidate entities of interest include geometric descriptions of palms, fingers, and parts under grasp as well as locations of points of contact. Building test methods from this fundamental point of view will lead to relevant performance capture, and will span from lower-level capabilities including primitive sensing and control to higher-level capabilities including touch-based manipulation and perception.

Presented below are descriptions for a subset of the metric and test method pairs listed in Table I and Table II with accompanying experimental implementations and results. Those not covered are still under development. Experiments were conducted for touch sensitivity, finger strength, grasp strength, slip resistance, force tracking, and sensor calibration test methods using three robotic hand configurations with different mechanical design, sensing, and control paradigms. More specifically, two hands[6] were used called "Hand 1" and "Hand 2", where Hand 1 has the capability to support two exchangeable touch sensory suites (impedance or resistance-based touch sensing), and Hand 2 incorporated current-based contact sensing at the motor drives. Furthermore, Hand 1 has three fingers of length 13.54 cm and 7 degrees-of-freedom (DOF) while Hand 2 has three fingers of length 13.5 cm and 4 DOF. Some of these experiments show a stationary robot supporting the hand under test, where in all cases the hand could have been supported using a static fixture instead[7].

[6]Certain commercial equipment, instruments, or materials are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

[7]The testing of additional robotic hand technologies will be included in these efforts to ensure that the benchmarks under development support the full spectrum of evolving robotic hand designs. In addition, some of the benchmarks will also be applicable to fingerless gripping technologies.

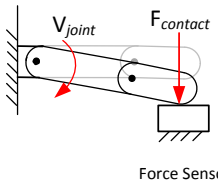

Fig. 3. Robotic finger is commanded to close on an object that is attached to a reference force sensor.

A. Touch Sensitivity

1) Metric and Test Method: Touch sensitivity is a kinetic measure of the smallest self-registered contact force exerted by a robotic finger on an object. The significance of this trait revolves around the hands ability to delicately interact with minimal disturbance to the immediate environment as well as detect small force perturbations. Direct applications would include part acquisition with object location or shape uncertainties as well as touch-based grasp planning. This characteristic is a function of the hand's sensor capabilities, motion controllers, bandwidth, joint speed, finger size, and finger-object configuration.

In order to most accurately capture the performance of a hand in this category, a dynamic test is needed. Of the previously listed dependencies, only the joint speed and finger-object configuration are assumed controllable. In particular, the robotic finger is commanded to close on an object that is attached to a reference force sensor[8] such that forces are measured before, during, and after finger-object contact (see Fig. 3. Once contact is detected by the hand, the active finger is commanded to stop its motion. In order to reduce the performance search space, only the worst-case finger-object configuration was investigated. Specifically, the finger is commanded to close at a specified "base-joint"[9] speed, V_{joint}, while any remaining joints are controlled such that finger extension is preserved. Thus, fingertip-object collision occurs with full-finger extension, which maximizes fingertip Cartesian velocity, and yields more aggressive finger-object impacts. Meanwhile, by commanding different base-joint speeds, a spread of behavior can be generated that will provide valuable insight on the trade-off between speed and touch sensitivity for any particular robotic hand.

A specific performance measure that can be computed for this test method involves contact force. To begin, compute the resultant magnitude of contact forces from the sensor data for each trial run, $F_{contact}$, by computing the L_2 norm of the three-dimensional contact force in time at a particular V_{joint}. Following, extract the peak $F_{contact}$, $F_{contact,max}$, for each touch test cycle (see Fig. 4. Collect the values of $F_{contact,max}$ at each V_{joint} for several repetitions, and compute the mean and 95% confidence intervals.

[8]A reference force sensor capable of resolving forces in three dimensions is suggested to make a more accurate capture of the full contact force.

[9]The base-joint is the first joint in the finger kinematic linkage.

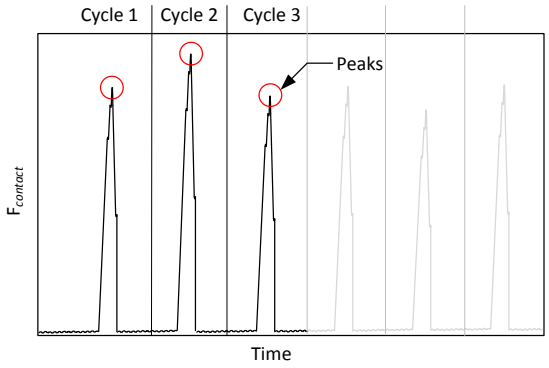

Fig. 4. Spike train of contact forces as a robotic finger periodically makes contact with a reference force sensor.

2) Experiments: The test method measures the peak contact forces of a single finger at various base-joint speeds with a 6-axis load cell when collision occurs at approximately full-finger extension as shown in Fig. 5. Once the robotic hand sensed contact, the actuated finger was commanded to hold position. The finger is then retracted to prepare for the next run. Three-dimensional force data was saved for each trial run at a 3 kHz sampling rate. Each finger for each chosen value of V_{joint} for every robotic hand was tested 10 times to create uncertainty bounds regarding the measure.

Results of these experiments are shown in Fig. 6, and the most sensitive fingers are now discussed as examples. Hand 1 with impedance sensing shows the highest sensitivity with a mean $F_{contact,max}$ of 0.26 N at a V_{joint} of 1 deg/s for Finger 1. These contact forces increased with higher values of V_{joint} with a maximum of 12.50 N for Finger 1 at a V_{joint} of 50 deg/s. Hand 1 with resistance-based contact sensing yielded a mean $F_{contact,max}$ of approximately 2.30 N for Finger 2 at 1 deg/s closing speed. This measure grew to a mean $F_{contact,max}$ of 18.70 N at 50 deg/s for Finger 2. Hand 2 with current sensing had a mean $F_{contact,max}$ of of 8.43 N at 10 deg/s for Finger 3 (1 deg/s was not possible on this hand). A $F_{contact,max}$ of 30.05 N at 50 deg/s was generated with Finger 1. All fingers were capable of detecting and reacting to contact, with some showing better touch sensitivity than others.

B. Finger Strength

1) Metric and Test Method: 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 hand during grasping or manipulation capabilities. The reasons for measuring strength on a single finger basis are two-fold: 1.) grasping and manipulation can occur with any number of fingers which means that the most independent measure of strength would be finger strength, and 2.) there can be inherent variability in finger strength across different fingers even in cases where they are

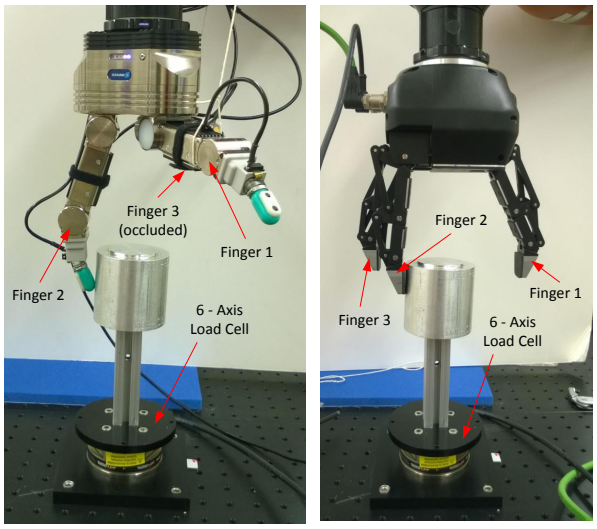

Fig. 5. Touch sensitivity test using a load cell to record the maximum contact force produced using Hand 1 tactile sensing (left) and Hand 2 current sensing (right).

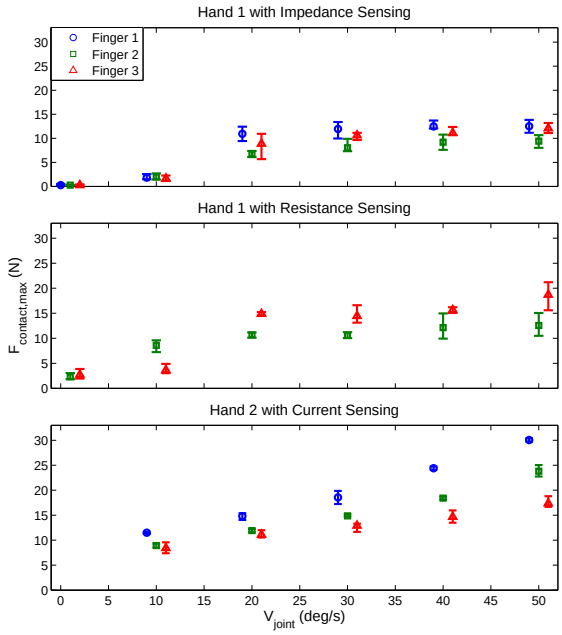

Fig. 6. Mean and 95% confidence intervals of the maximum contact force of every robotic finger across three robotic hands with different touch sensing strategies. Note that in the 2nd plot, Finger 1 of Hand 1 was not tested due to a faulty resistance sensor.

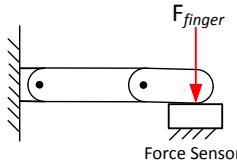

Fig. 7. Robotic finger exerting maximum fingertip force as measured by a reference force sensor.

mechanically equivalent. Finger strength is a function of the hands actuator capabilities, motion controllers, mechanical design, and finger-object configuration.

Of the previously listed dependencies, only the finger-object configuration is used as a test variable. While using an extended (near-singular) finger configuration, position the finger just above the force sensor and verify a zero force reading (see Fig. 7). Under position control, command the finger to close completely which should induce control saturation. This configuration was chosen as it is easily replicated across a diverse set of robot hands and results in a maximum moment arm at the point of contact which generates smaller contact forces when compared to other, more curled finger states. Record force sensor data throughout the test. Use of a three-axis reference force sensor is encouraged to more accurately capture the total contact force.

The particular performance measure regarding fingertip contact force magnitude, F_{finger}, should be computed by the L_2 norm for each set of force readings given by the external force sensor. Next, the contact force magnitude from the quasi-static (settled) force region (see Fig. 8) should be extracted for each load cycle (several cycles or repetitions should be conducted), and then averaged to yield the maximum finger strength, $F_{finger,max}$. Collect maximum forces for several cycles, and compute the mean and 95% confidence intervals to estimate finger strength.

2) Experiments: Experimental results were obtained in this category by commanding each finger (one finger per test) of a robotic hand to close completely while on a collision course with an object attached to an external 6-axis load cell. Again, the robotic hand and load cell were positioned such that contact took place at the fingertip when the finger was fully extended and perpendicular to the palm as shown in Fig. 9. Finger-object contact was established and removed 32 times per finger for each hand while recording interaction forces via the load cell to generate performance distributions. The results from these experiments are shown in Fig. III. Hand 1 exhibits $F_{finger,max}$ values of 8.24 N to 14.29 N, while Hand 2 exhibits $F_{finger,max}$ values ranging from 30.44 N to 33.89 N.

C. Grasp Strength

1) Metric and Test Method: Grasp strength is a kinetic measure of the maximum internal force a robotic hand can impose on an object. This measure will yield information regarding a particular hands payload capabilities for various object sizes as well as its limits in resisting pulling or pushing forces during a grasp operation. Grasp strength is a function of the hands actuator capabilities, motion controllers, mechanical design, grasp configuration, and object size.

Of the previously listed dependencies, only the grasp

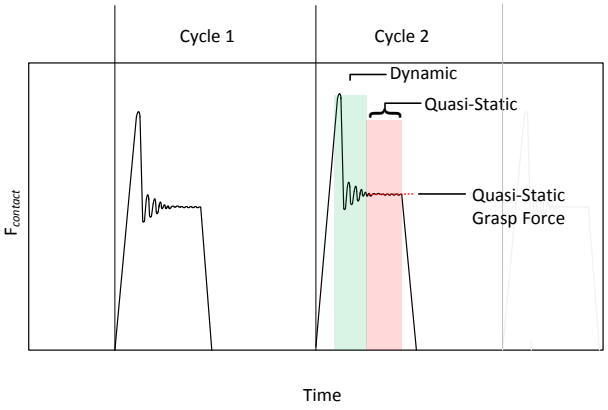

Fig. 8. Depiction of dynamic and quasi-static force regions during load cycles.

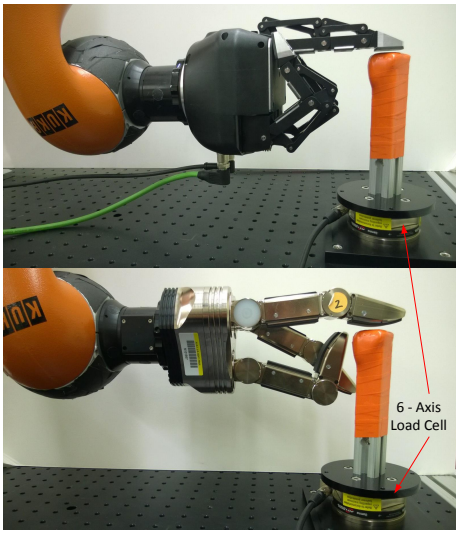

Fig. 9. Finger strength test setup, where Hand 1 (bottom) and Hand 2 (top) are positioned with the load cell such that fingertip contact takes place when the finger is fully extended and perpendicular to the palm. Note, the orange covering protects the fingertip from sharp edges on the vertical load cell extension.

configuration and object size are assumed controllable. For this particular test, a wrap grasp is issued on cylinders with different diameters with embedded force sensing (see Fig. 10). The wrap grasp is a typical power grasp that should yield maximum grasp strength performance. The cylinder diameters are varied in order to capture grasp strength performance across differently-sized artifacts that induce changes in the settled grasp configuration. The artifact-embedded force sensors are used as the reference sensor for resolving cross-sectional internal force transmission. In this particular design, since the embedded force sensing can only resolve internal forces in one direction, measuring force transmissions in the 0 deg and 90 deg configurations allows for the calculation of an approximate cross-sectional internal force. Finally, the cylinder is positioned axially within the grasp such that the force sensing artifact is centered among the contact points.

TABLE III

MEAN AND 95% CONFIDENCE INTERVALS OF THE THE MAXIMUM FINGERTIP CONTACT FORCE MAGNITUDE EXERTED BY EACH FINGER FOR HAND 1 AND HAND 2 MEASURED USING A LOAD CELL.

	Hand 1 $F_{finger,max}$ (N)		Hand 2 $F_{finger,max}$ (N)	
Finger	Mean	95% Confidence Interval	Mean	95% Confidence Interval
1	14.29	[13.66, 14.73]	30.44	[29.39, 30.95]
2	8.24	[8.00, 8.80]	30.84	[30.63, 31.04]
3	10.64	[10.41, 10.83]	33.89	[33.21, 34.73]

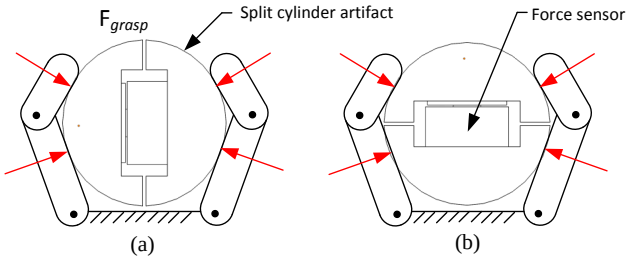

Fig. 10. Wrap grasp on a split-cylinder artifact with embedded force sensors placed in (a) 0 degree and (b) 90 degree orientations.

The performance measure for this method involves calculating the maximum internal grasp force magnitude, $F_{grasp,max}$. This value is determined by computing the L_2 norm of forces acting in the cross-sectional plane of the artifact after the grasp has settled yielding quasi-static grasp forces. Using the settled force is the same strategy taken in the previous section (see Fig. 8 for illustration). Several grasp cycles should be carried out for a particular artifact diameter such that a mean performance and 95% confidence interval can be calculated for a particular hand grasp configuration and artifact size.

2) Experiments: The grasp strength metric was extracted by measuring the internal force imparted by the robotic hand on cylinders of two different diameters – 50 mm and 80 mm — as shown in Fig. 11. Each cylinder possessed two internal single-axis load cells to measure internal forces. Two different sizes were used in order to capture any variation in grasp force based on object size. Forces were measured at a 0 and 90 degree orientation as shown in Fig. 11 in order to compute an approximate resultant internal force. Each cylinder was oriented within the grasp such that its axis was parallel to the plane of the palm. For Hand 1, the natural settling location for the cylinder was between the proximal and distal phalanges, while Hand 2 had a settling location against the palm. Tests were repeated 32 times per cylinder and orientation.

As shown in Table IV, the mean and 95% confidence intervals of the internal grasp force for both hands and cylinders were captured. Noticeably, Hand 1 consistently imparted lower grasp forces than Hand 2, which was expected based on the results surrounding finger strength as presented

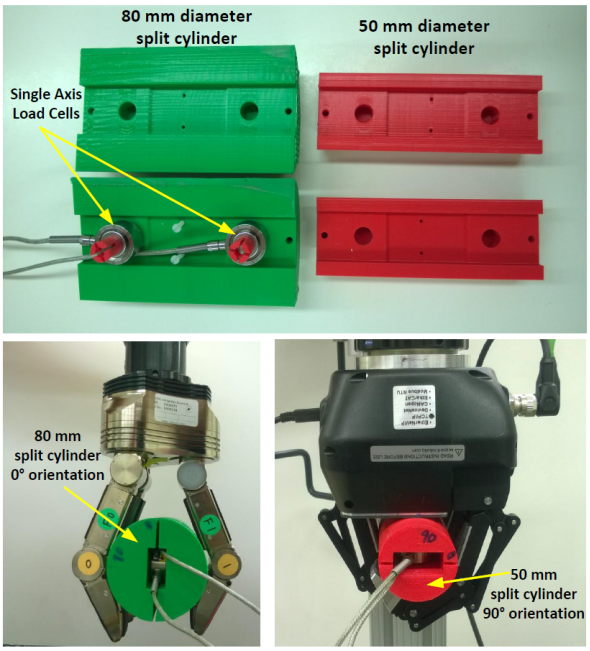

Fig. 11. 80 mm and 50 mm diameter split cylinder configurations for determining grasp forces (top). Robotic Hand 1 performing cylindrical grasp on 80 mm split cylinder oriented at 0 deg (bottom left). Robotic Hand 2 performing cylindrical grasp on 50 mm split cylinder oriented at 90 deg (bottom right).

in the previous section. For the 50 mm diameter cylinder, Hand 1 imparted a mean $F_{grasp,max}$ of 47.02 N (confidence interval of 44.37 N to 49.47 N) and 76.11 N (confidence interval of 70 N to 84.32 N) on the 50 mm and 80 mm cylinders, respectively. Meanwhile, Hand 2 imparted a mean $F_{grasp,max}$ of 118.98 N (confidence interval of 101.26 N to 137.84 N) and 92.97 N (confidence interval of 84.81 N to 100.67 N) on the 50 mm and 80 mm cylinders, respectively. Interestingly, Hand 1 imparted higher internal forces on the larger diameter cylinder while the opposite is true for Hand 2. Consequently, imparted internal forces on the 80 mm cylinder by both hands were much closer to each other.

TABLE IV

THE MEAN AND 95% CONFIDENCE INTERVALS OF THE MAXIMUM INTERNAL GRASP FORCE FOR HAND 1 AND HAND 2.

	Hand 1 $F_{grasp,max}$ (N)		Hand 2 $F_{grasp,max}$ (N)	
Cylinder Diameter (mm)	Mean	95% Confidence Interval	Mean	95% Confidence Interval
50	47.02	[44.37, 49.47]	118.98	[101.26, 137.84]
80	76.11	[70.00, 84.32]	92.97	[84.81, 100.67]

D. Slip Resistance

1) Metric and Test Method: Slip resistance is a kinetic measure of a robotic hands ability to resist slip. The focus of this test method is to investigate the inherent surface friction properties of the hand. With higher friction coefficients, robotic fingers will possess wider friction cones at the areas of contact with an object. This behavior would

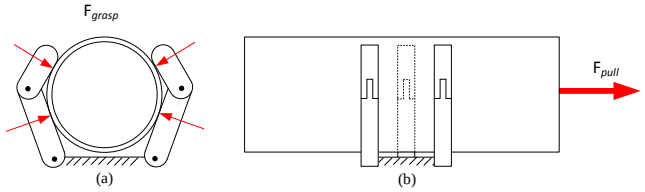

Fig. 12. Test setup for slip resistance where a cylindrical artifact is placed in a wrap grasp at maximum hand power. The pipe is then pulled at an increasing force until slip is observed.

ultimately allow friction forces to contribute more greatly to the overall grasping effort yielding greater resistances to slipping, and generally enhanced energy efficiency during the grasping operation. This characteristic is a function of the hands actuator capabilities, motion controllers, mechanical design, grasp configuration, object size, and object surface properties.

Of the previously listed dependencies, only the grasp configuration, object size, and object surface properties are assumed controllable. Given this large performance search space, some variables are fixed to make testing more tractable while still providing useful results. Specifically, the wrap grasp on a cylindrical artifact was chosen to investigate slip resistance capabilities under maximum power and highest number of hand-object points of contact (see Fig. 12). Furthermore, use of a cylindrical artifact under a wrap grasp eliminates the undesirable behavior of object-finger locking. Finally, it is encouraged to construct artifacts with existing materials to ensure relatively consistent artifact surface properties. The general test procedure consists of the following three steps: 1.) place the cylindrical artifact in the robotic hand using a wrap grasp at maximum power, 2.) pull on the pipe at a controlled rate of increasing force while recording force until gross slipping is visually confirmed between the hand and artifact, and 3.) repeat this process several times over a range of artifact sizes that the robotic hand is capable of grasping.

The measure of interest in this test is the maximum obtainable pull force before gross slip of a given hand and pipe size under a specified grasp. For each test cycle, record the pull force, F_{pull}, over time. Extract the maximum pull force, $F_{pull,max}$ from the force-time plot (see Fig. 13). Calculate the mean and 95% confidence intervals for each pipe diameter size from several trial runs. It should be noted that "micro" slipping may occur at times during these tests that are characterized by periods of brief "necking" or plateauing in the force-time plot (see Fig. 13). These are generated as the artifact settles within the grasp, and should be ignored.

2) Experiments: Experimental results for this section were obtained by closing a robotic hand around various diameter polyvinyl chloride (PVC) pipes (a standard cylindrical product) at maximum grip force in a cylindrical

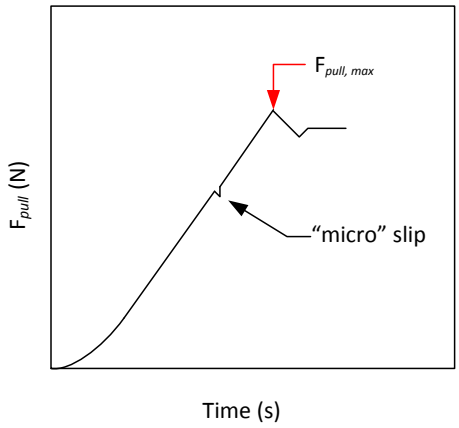

Fig. 13. Depiction of maximum pull force and "micro" slip events.

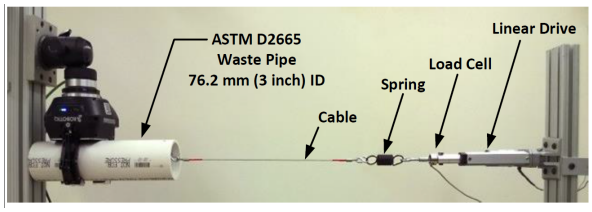

Fig. 14. Power and resistance to slip test using a length of ASTM D2665 waste pipe. A linear drive pulls attached to a cable provides and incremental load on the pipe. The load rate is decreased using an in-line spring and force is recorded using a single axis load cell.

grip configuration as shown in Fig. 14. Once grasped, the pipe was pulled by a linear actuator with a connected cable and attached load cell to measure pull forces. The tension force in the wire climbed as a function of time during actuation until a peak force was reached resulting in an immediate drop indicating a shift from static Coulomb friction to dynamic Coulomb friction. This procedure was repeated 10 times for each hand and each cylinder diameter to generate performance distributions. The following results reveal the peak tension force obtained across both hands and various PVC pipe diameters. As seen in Fig. 15, Hand 2 consistently yields higher maximum pull forces in comparison to Hand 1. As expected, there is also a variation in the maximum pull force depending on the size of the cylinder diameter. With different object sizes, robotic hands will enclose more or less compactly around an object. With more points of contact and more heavily curled fingers, higher pull forces should be expected. In this particular case, both hands experienced peak mean $F_{pull,max}$ values of 82.64 N for Hand 1 and 164.73 N for Hand 2 for the 7.62 cm (3 in) inner diameter pipe.

E. Finger Force Tracking

1) Metric and Test Method: Finger force tracking is a kinetic measure regarding the fingers ability to impose desired contact forces on its environment. This capability is particularly important for many state-of-the-art robotic

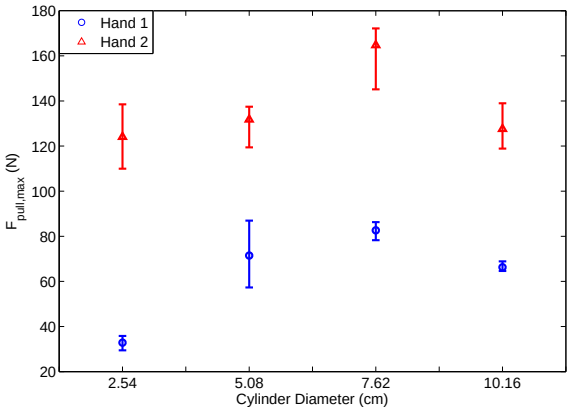

Fig. 15. Mean and 95% confidence intervals of the the maximum pull force achieved by each hand across several PVC pipes of diameters ranging from 2.54 to 10.16 cm (1 to 4 inches.)

grasping and manipulation control algorithms that use force-based control approaches [15]–[20]. Moreover, this capability can be used for touch-based grasp planning, controlled interaction for texture discrimination and object localization. This characteristic is a function of a hand's actuator capabilities, sensor calibration, motion and force controllers, control and sensing bandwidth, mechanical design, finger-artifact configuration, and the parameters of the selected contact force trajectory.

This test method seeks to capture the force tracking performance of an individual finger of a robotic hand. Of the previously listed dependencies, only the finger-artifact configuration and the parameters of the desired contact force profile are assumed controllable. The test begins by commanding the finger under test to track a desired force profile by making contact with an artifact attached to a reference force sensor. It is encouraged to use a reference sensor that is capable of resolving three-dimensional contact forces for greater measurement accuracy. The parameters of this desired force profile can vary in contact force direction as well as magnitude, and therefore the contact force can exist anywhere within the contact friction cone (see Fig. 16). Even though different finger-artifact configurations may be explored, it is important to ensure that force control is engaged when the finger configuration is in a curled state. This finger configuration will avoid force control stability issues that can occur near singular configurations. During the test, the desired force profile ($F_d \in \mathbb{R}^{3x1}$) and the contact forces measured by the load cell ($F_L \in \mathbb{R}^{3x1}$) are all recorded for extracting performance measures.

There are three relevant performance metrics for test method involving force magnitude, force direction, and force peak overshoot. When considering force magnitude, calculate the Root Mean Squared Error (RMSE) between the desired force magnitude ($||F_d|| \in \mathbb{R}$) and for the forces measured by the reference force sensor ($||F_L|| \in \mathbb{R}$). When

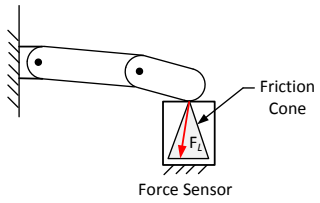

Fig. 16. Finger imposing forces on a reference force sensor within the contact friction cone.

considering force direction, calculate the RMSE between the desired force magnitude ($\hat{F}_d \in \mathbb{R}^{3x1}$) and for the forces measured by the reference force sensor ($\hat{F}_L \in \mathbb{R}^{3x1}$). When considering peak overshoot, calculate the peak overshoot ($F_{peak} \in \mathbb{R}$) between $||F_d||$ and $||F_L||$.

2) Experiments: Each robotic finger was commanded to impart a certain contact force magnitude and direction on a 6-axis load cell as shown in Fig. 17. A nonlinear admittance control algorithm was implemented on Hand 1 to yield force tracking, while only the stock force control capabilities were used on Hand 2. Specifically, a contact force magnitude of 1, $\frac{F_{finger,max}}{2}$, and $F_{finger,max}$ were commanded where $F_{finger,max}$ is based on the maximum end-effector force capability as measured in the previous section ($F_{finger,max}$ is 10 N for Hand 1 and 30 N for Hand 2). These values were chosen to investigate force tracking capabilities at extrema (and approximately their average) to reduce the performance search space. A time-varying force profile was also issued for those fingers with force-tracking capabilities and is defined by

$$||F_{d,z}|| = 5\log\left(\sin\left(\frac{\pi(t+3)}{2}\right) + \cos\left(\frac{t}{4}+\pi\right) + 3\right) + 1 \quad (1)$$

where $t \in \mathbb{R}$ is time and $F_{d,z} \in \mathbb{R}$ is the desired force trajectory in the world coordinate system's z-axis. This spread of force magnitudes was chosen to test the hands' abilities to follow set-point force tracking for small, medium, and large forces that scale with the capabilities of the hand. The time-varying force profile contains multiple frequencies and a range of magnitude shifts to more thoroughly test force-tracking performance. All measures were obtained by recording force tracking data for 60 seconds of continuous control operation for each finger, and averaged over all fingers per hand layout.

Table V shows the results obtained for the controller's actual performance. In particular, the results show that Hand 1 with impedance sensing yields RMSE values for $||F_d|| - ||F_L||$ of at least half those obtained for Hand 1 with resistance sensing. Hand 2 performs similarly to Hand 1 when $||F_d|| = \frac{F_{finger,max}}{2}$, but generates much larger RMSE values over 5 N at $F_{finger,max}$. In general, Hand 1 with impedance sensing allows for improved control over contact force directions with lower RMSE values in all

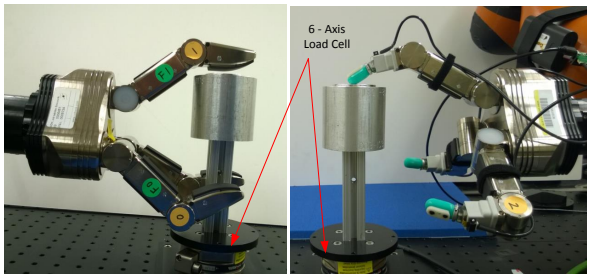

Fig. 17. Test setup for finger force tracking.

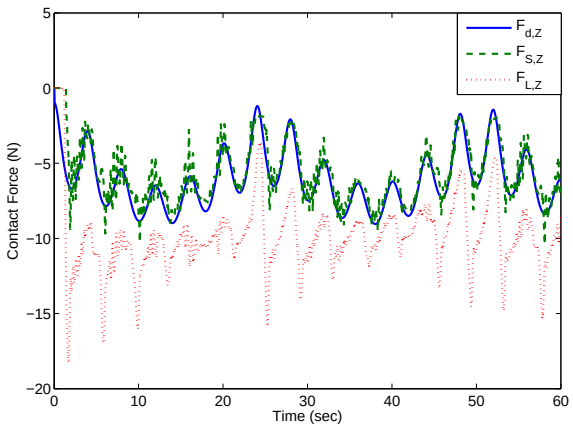

Fig. 18. The desired force profile ($F_{d,Z}$), the contact force as sensed by the onboard sensor ($F_{S,Z}$), and the contact force as sensed by an external load cell ($F_{L,Z}$) for Hand 1, Finger 2 with resistance sensing.

directions for $\hat{F}_d - \hat{F}_L$ when compared to Hand 1 with resistance sensing. Hand 2 is an underactuated hand and does not possess controllability over its fingertip contact force directions and is therefore excluded from this measure. Finally, Hand 1 with impedance sensing has about a third the amount of contact force overshoot when compared to Hand 1 with resistance sensing across all desired contact force profiles. Hand 2 cannot control for forces near 1 N or time-varying forces, and was therefore excluded from those tests. However, for $\frac{F_{finger,max}}{2}$, Hand 2 performs reasonably well with an overshoot of 2.864 N. Undershoot was exhibited with a negative value for $F_{finger,max}$.

In order to more clearly illustrate the behavior of the force controlled system, Fig. 18 shows the desired force profile, the contact force as measured by the intrinsic sensor, and the contact force as measured by an external load cell for Hand 1, Finger 2 with resistance sensing. This figure illustrates the complexity surrounding force control of robotic fingers: the sensed forces closely trace the desired forces, but the actual forces are consistently shifted from both the perceived or desired forces.

F. Force Calibration

1) Metric and Test Method: Force based sensor calibration is important for many state-of-the-art robotic

TABLE V

FORCE TRACKING PERFORMANCE ERRORS FOR THREE FORCE-CONTROLLED HAND LAYOUTS. (*INDICATES SYSTEM UNDERSHOOT).

Robotic Hand	$\|\|F_d\|\|$ N	RMSE (N) $\|\|F_d\|\| - \|\|F_L\|\|$	RMSE $\hat{F}_d - \hat{F}_L$	Force Peak Overshoot (N)
Hand 1 (Impedence Sensing)	1	0.567	[0.124; 0.428; 0.323]	1.046
	$\frac{F_{finger,max}}{2}$	2.182	[0.020; 0.225; 0.102]	4.972
	$F_{finger,max}$	1.773	[0.015; 0.134; 0.021]	5.159
	Eq. 1	2.092	[0.024; 0.204; 0.082]	5.160
Hand 1 (Resistance Sensing)	1	2.121	[0.218; 0.285; 0.483]	6.382
	$\frac{F_{finger,max}}{2}$	4.577	[0.075; 0.398; 0.178]	12.028
	$F_{finger,max}$	4.032	[0.062; 0.283; 0.133]	16.746
	Eq. 1	5.013	[0.093; 0.330; 0.223]	14.223
Hand 2 (Current Sensing)	1	N/A	N/A	N/A
	$\frac{F_{finger,max}}{2}$	1.226	N/A	2.864
	$F_{finger,max}$	5.129	N/A	-3.012*
	Eq. 1	N/A	N/A	N/A

grasping and manipulation control algorithms that use force-based control approaches. That is, in order to control contact forces, force sensor readings must be accurate. Moreover, force capabilities can be used for touch-based grasp planning, controlled interaction for texture discrimination and object localization. This characteristic is a function of the tactile sensor mechanical design, and its calibration.

This test method seeks to capture the performance of force based tactile sensors by comparing the force readings measured by the sensor ($F_S \in \mathbb{R}^{3x1}$) to force data recorded simultaneously using an external force sensor ($F_L \in \mathbb{R}^{3x1}$). Using the desired sensor-object orientation, position the sensor under test just above the force sensor and verify a zero force reading. Press the sensor against the load cell and record both the sensor force reading and the load cell readings. If desired, collect F_S during the finger force tracking test method as well to to extract the necessary information to calculate force calibration performance metrics.

Again, there are three relevant performance metrics for this test method involving force magnitude, force direction, and maximum force error. When considering force magnitude, calculate the RMSE between the tactile sensor force magnitudes ($||F_S|| \in \mathbb{R}$) and those measured by the reference force sensor ($||F_L|| \in \mathbb{R}$) for all data collected. When considering force direction, compute the RMSE between the force direction as measured by the tactile sensor ($\hat{F}_S \in \mathbb{R}^{3x1}$) and the external force sensor ($\hat{F}_L \in \mathbb{R}^{3x1}$). When considering the maximum force error, calculate the absolute maximum error between the contact force magnitude as measured by the hand sensor and the reference force sensor.

2) Experiments: The particular performance measures are similar to those in the previous section, except that they are calculated between the forces measured by the external load cell and the forces measured by the intrinsic, on-board calibrated sensors, $F_S \in \mathbb{R}^{3x1}$. Also, maximum force error is calculated instead of peak overshoot. Table VI reveals that the impedance sensors of Hand 1 are more accurate in predicting contact force magnitudes with consistently lower RMSE values for $||F_L|| - ||F_S||$ and overshoot values when compared to the resistance sensors. Hand 1 with impedance sensing also appears to predict contact force directionality with higher fidelity only in the primary force controlled axis, Z. Accuracy in the secondary axes appears to be better for Hand 1 with resistance sensing.

TABLE VI

FORCE CALIBRATION PERFORMANCE ERRORS FOR TWO FORCE-CONTROLLED HAND LAYOUTS.

Robotic Hand	$\|\|F_d\|\|$ N	RMSE (N) $\|\|F_L\|\| - \|\|F_S\|\|$	RMSE $\hat{F}_L - \hat{F}_S$	Maximum Force Error (N)
Hand 1 (Impedence Sensing)	1	1.054	[0.412; 0.529; 0.285]	3.004
	$\frac{F_{finger,max}}{2}$	2.855	[0.254; 0.248; 0.118]	7.108
	$F_{finger,max}$	2.170	[0.087; 0.154; 0.038]	9.084
	Eq. 1	2.711	[0.201; 0.257; 0.099]	7.191
Hand 1 (Resistance Sensing)	1	2.586	[0.218; 0.280; 0.815]	6.380
	$\frac{F_{finger,max}}{2}$	4.825	[0.075; 0.427; 0.144]	13.386
	$F_{finger,max}$	4.939	[0.062; 0.336; 0.101]	16.398
	Eq. 1	5.411	[0.093; 0.359; 0.158]	13.003

IV. Conclusions and future work

This document presents the beginnings of a framework for robotic hand performance benchmarking and concepts. We presented a proposed set of component and system level physical tests along with experimental results from a subset of these tests for three robotic hand configurations. The results of test methods like those presented herein can be used to generate a better understanding of robotic hand technology which reduces end-user adoption risks while fostering insight for future product designs. Detailed descriptions of these experimental setups, test methods, and data sets are available at http://www.nist.gov/el/isd/grasp.cfm. This work will be extended within the IEEE RAS RHGM Technical Committee using a community-driven approach (http://www.rhgm.org).

References

[1] H. Moravec. *The future of robot and human intelligence*. Harvard University Press, 1998.

[2] C.S. Lovchik and M.A. Diftler. The robonaut hand: a dexterous robot hand for space. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 2, pages 907–912 vol.2, 1999.

[3] F. Rothling, R. Haschke, J.J. Steil, and H. Ritter. Platform portable anthropomorphic grasping with the bielefeld 20-dof shadow and 9-dof tum hand. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 2951–2956, Oct 2007.

[4] H. Liu, K. Wu, P. Meusel, N. Seitz, G. Hirzinger, M.H. Jin, Y.W. Liu, S.W. Fan, T. Lan, and Z.P. Chen. Multisensory five-finger dexterous hand: The dlr/hit hand ii. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3692–3697, Sept 2008.

[5] Lael U Odhner, Leif P Jentoft, Mark R Claffee, Nicholas Corson, Yaroslav Tenzer, Raymond R Ma, Martin Buehler, Robert Kohout, Robert D Howe, and Aaron M Dollar. A compliant, underactuated hand for robust manipulation. *The International Journal of Robotics Research*, 2014.

[6] F. Bonsignorio, A.P. del Pobil, and E. Messina. Fostering progress in performance evaluation and benchmarking of robotic and automation systems [tc spotlight]. *Robotics Automation Magazine, IEEE*, 21(1):22–25, March 2014.

[7] A.M. Dollar, L.P. Jentoft, J.H. Gao, and R.D. Howe. Contact sensing and grasping performance of compliant hands. *Autonomous Robots*, 28(1):65–75, 2010.

[8] G. A. Kragten, C. Meijneke, and J. L. Herder. A proposal for benchmark tests for underactuated or compliant hands. *Mechanical Sciences*, 1(1):13–18, 2010.

[9] C. Meijneke, G. A. Kragten, and M. Wisse. Design and performance assessment of an underactuated hand for industrial applications. *Mechanical Sciences*, 2(1):9–15, 2011.

[10] V. Wright. Prosthetic outcome measures for use with upper limb amputees: A systematic review of peer-reviewed literature. *Journal of Prosthetics and Orthotics*, 21(4):3–63, 2009.

[11] Lael U. Odhner, Leif P. Jentoft, Mark R. Claffee, Nicholas Corson, Yaroslav Tenzer, Raymond R. Ma, Martin Buehler, Robert Kohout, Robert D. Howe, and Aaron M. Dollar. A compliant, underactuated hand for robust manipulation. *The International Journal of Robotics Research*, 33(5):736–752, 2014.

[12] R. Platt, C. Ihrke, L. Bridgewater, D. Linn, R. Diftler, M. Abdallah, S. Askew, and F. Permenter. A miniature load cell suitable for mounting on the phalanges of human-sized robot fingers. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 5357–5362, May 2011.

[13] G. Obinata, A. Dutta, N. Watanabe, and N. Moriyama. Vision based tactile sensor using transparent elastic fingertip for dexterous handling. In S. Kolski, editor, *Mobile Robot.: Perception and Navigation*, pages 137–148. Pro Literatur Verlag, 2006.

[14] C.H. Lin, T.W. Erickson, J.A. Fishel, N. Wettels, and G.E. Loeb. Signal processing and fabrication of a biomimetic tactile sensor array with thermal, force and microvibration modalities. In *Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on*, pages 129–134, Dec 2009.

[15] K. Van Wyk. *Grasping and Manipulation Force Control for Coordinating Multi-Manipulator Robotic Systems with Proprioceptive Feedback*. PhD thesis, University of Florida, 2014.

[16] T. Wimbock, C. Ott, and G. Hirzinger. Analysis and experimental evaluation of the intrinsically passive controller (ipc) for multifingered hands. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 278–284, May 2008.

[17] J. Markdahl, Y. Karayiannidis, Xiaoming Hu, and D. Kragic. Distributed cooperative object attitude manipulation. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2960–2965, May 2012.

[18] L.U. Odhner, R.R. Ma, and A.M. Dollar. Experiments in underactuated in-hand manipulation. In J.P. Desai, G. Dudek, O. Khatib, and V. Kumar, editors, *Experimental Robotics*, volume 88 of *Springer Tracts in Advanced Robotics*, pages 27–40. Springer International Publishing, 2013.

[19] J.M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K.J. Kuchenbecker. Human-inspired robotic grasp control with tactile sensing. *Robotics, IEEE Transactions on*, 27(6):1067–1079, Dec 2011.

[20] Yu Zhao and Chien-Chern Cheah. Neural network control of multifingered robot hands using visual feedback. *Neural Networks, IEEE Transactions on*, 20(5):758–767, May 2009.