Robotic Grasping and Manipulation Competition: Future Tasks to Support the Development of Assembly Robotics

[1]Karl Van Wyk, [1]Joe Falco, and [1]Elena Messina

[1]National Institute of Standards and Technology (NIST)

Abstract. The Robot Grasping and Manipulation Competition, held during the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) in Daejeon, South Korea was sponsored by the IEEE Robotic and Automation Society (RAS) Technical Committee (TC) on Robotic Hands Grasping and Manipulation (RHGM) [1]. This competition was the first of a planned series of grasping and manipulation-themed events of increasing difficulty that are intended to spur technological developments and advance test methods and benchmarks so that they can be formalized for use by the community. The coupling of standardized performance testing with robot competitions will promote the use of unbiased evaluation methods to assess how well a robot system performs in a particular application space. A strategy is presented for a series of grasping and manipulation competitions that facilitate objective performance benchmarking of robotic assembly solutions. This strategy is based on test methods that can be used for more rigorous assessments and comparison of systems and components outside of the competition regime. While competitions have proven to be useful mechanisms for assessing the relative performance of robotic systems with measures of success, they often lack a methodical measurement science foundation. Consequently, scientifically sound and statistically significant metrics, measurement, and evaluation methods to quantify performance are missing. Using performance measurement methods in a condensed format will accommodate competition time limits while introducing the methods to the community as tools for benchmarking performance in the developmental and deployment phases of a robot system. The particular evaluation methods presented here are focused on the mechanical assembly process, an application space that is expected to accelerate with the new robot technologies coming to market.

Keywords: robot, grasping, manipulation, competition, benchmarks, performance measures, manufacturing

1 Introduction

Robot competitions [2–4] are plentiful and provide an excellent opportunity for researchers and developers to benchmark task-based solutions in a vying environment where the final score is based on degree of task completion followed by a subjective analysis of winners and losers to determine relative advantages and disadvantages of the competing systems. These competitions provide a common problem space to demonstrate advancement of the state-of-the-art in new software and hardware solutions of an integrated system while promoting the field of robotics for both educational and general audiences. We address concerns that competitions often lack scientifically sound and statistically significant metrics, measurement, and evaluation methods [5, 6].

The 2016 competition featured two main categories of challenges with a mixture of service-oriented (e.g., home assistant) and manufacturing-relevant tasks and objects. A pick-and-place challenge involved removing items from a shopping bin and placing them on target surfaces. A series of manipulation tasks that were manufacturing oriented included twisting a bolt into a threaded hole with a nut driver, hammering a nail, and sawing open a cardboard box. Retrospectively, these tasks served as a good starting point, but needed maturity in various key ways. For instance, the diversity of the tasks (particularly manufacturing related) were not sufficient and scoring was unforgiving if a particular step in the process was unachievable. Various features of the tasks (e.g., fastener sizes) were arbitrarily chosen, and initialization of tests, although randomized, were not controlled across competitors. Consequently, some teams experienced much more difficult starting scenarios. Further details about the experiences with the inaugural competition can be found in this book's chapter on Competition Feedback and Lessons Learned. Moving forward, a more rigorous approach that is better-aligned with manufacturing tasks is being undertaken. This is intended to help advance robotic grasping and manipulation specifically towards addressing assembly tasks. Therefore, an approach to competitions that employs a more rigorous assembly-centric performance evaluation methodology and artifacts is described in the remainder of this chapter.

Standardized performance testing is an emerging and necessary tool within the robotics community providing unbiased evaluation methods that assess how well a system performs a particular ability. These performance evaluations can be used to assess a system's individual components, as well as its system level operation. The National Institute of Standards and Technology (NIST) works to develop technical foundations for performance standards in several key areas of robotics including emergency response robots, perception, grasping and manipulation, and agility [7–12]. In addition, this NIST work is often introduced at competition venues as a mechanism to disseminate, as well as evaluate, the performance test methods prior to the standardization phase.

To help progress the use of robots for assembly operations, we present a strategy for a grasping and manipulation competition track that promotes objective performance benchmarking of robotic assembly solutions based on test methods that can be used for more rigorous assessments and comparison of systems and components outside of the competition regime. Using these methods in a condensed format will accommodate competition time limits while introducing the methods to the community as tools for benchmarking performance. Competitions have also proven useful to help advance the development of performance benchmarking methods and we expect that the use of these methods during competitions will help to further develop them for use by the robotics community. NIST anticipates that such research will lead to a principled way of specifying robot system characteristics and will help smaller organizations to determine which robot system components are best suited for their application space [13].

2 Why Robotic Assembly?

The International Federation of Robotics indicated in its 2016 World Robotics Report that by 2019, more than 1.4 million new industrial robots will be installed in factories around the world [14]. They also emphasize that these new robots will not only support traditional large manufacturers, but small and medium-sized enterprises (SMEs) as well. In order for robotic solutions to benefit SME-based manufacturing operations, where it is cost prohibitive to employ robotics experts, the robots must be programmable by line operators and easy to redeploy to support low volume, high mixture production runs. Analysis of robot implementations in the auto industry estimates that assembly accounts for 50 % of all manufacturing costs yet it only accounts for 7.3 % of robot sales [14]. A Price Waterhouse (PwC) survey of 107 respondents, conducted in conjunction with the Manufacturing Institute, found that the most common task amongst US manufacturers was assembly (25 %) followed by machining (21 %), and the least common tasks were warehousing and performing dangerous tasks (both 6.5 %). The survey also indicates that assembly was the most common task that manufacturers planned to invest in robotic technology to support (27 %) [15].

As early as the 1970s there were expectations that robots would be able to perform assembly operations to alleviate humans from what were thought to be onerous, dangerous, repetitive, and tedious tasks. While this seemed achievable in concept, robot technologies of the time could not cost-effectively support the tight tolerances and component variability associated with the assembly process. Despite many advancements in hardware and control software, the limitations encountered in the early days of attempting robotic assembly operations still persist after many decades. Due to their highly rigid designs and position-based control, most industrial robots require customized fixtures that are tailored to a particular assembly operation and component geometry in order to perform assembly tasks. These specialized fixtures introduce costs and add time to the setup of every new assembly job. Even more expensive and sophisticated approaches were conceived that compensated for motion errors using force sensing at the end-effector. These methods required 6-axis force-torque sensing at the tool point, low-level force feedback to the robot position or force controller, and the highly application specific algorithmic support for accomplishing assembly operations. Mechanisms and methods to help enable robotic assembly are surveyed in [16].

Recent progress in technologies for robotic arms and end-effectors hold potential to overcome the problems with robotic assembly. For instance, collaborative robots or Co-Bots are designed to safely work alongside human workers in both manufacturing and service sectors [17]. These robots are equipped with force sensing and/or compliance in order to limit contact forces and prevent injury to humans working in their proximity. These capabilities also prove advantageous for facilitating assembly operations.

Concurrently, robotic hand technology is emerging as a next generation end-effector technology with advanced force control and manipulation capabilities. Some existing robotic hand cutaneous sensors coupled with the latest advances in artificial intelligence are approaching and even exceeding the sensing capabilities of the human hand. Moreover, the enhanced reconfigurability of robotic hands promise new ways of tackling the small parts assembly field for manufacturing operations.

3 Measuring the Assembly Capabilities of Robot Systems

An assembly consists of a set of operations that join together individual parts or subassemblies. For the purposes of the robotic grasping and manipulation competition, we focus on assemblies that incorporate small part insertions and fastening methods such as threading, snap fitting, and gear meshing using standard components including screws, nuts, washers, gears and electrical connectors. Since robot system designs can vary greatly, a goal in developing standardized performance tests for assembly robotics is to provide a modular set of task-based tests to support a full spectrum of robotic solutions. On one end of the spectrum, a robot system and its components can be designed to suit a specific application task and perform this one task in a structured environment very efficiently. The structure comes in the form of specialized fixtures, part feeders, end-effectors, and tools that provide the necessary compliance to accommodate the assembly tolerances in the presence of robot position errors. On the other end of the spectrum, a robot system can be designed to be flexible and adaptive for handling a variety of parts, variations in similar parts, and multiple assembly process types in an unstructured environment. With respect to the task-based performance tests, a robot system designed to solve a particular task may excel at an individual test module, whereas a flexible system will be proficient across multiple test modules.

Manual assembly efficiencies take into account the time associated with individual actions such as grasp, orient, insert, and fasten, as performed by a human with decades of experience and practice using their hands, eyes, and brains. One avenue for methodically designing task-level tests within manufacturing leverages factors identified by Boothroyd-Dewhurst (B-D) design for assembly (DFA) studies [18]. These studies have already identified and tabulated various important factors based on manual human performance in an assembly task. For instance, size and symmetry of parts, tool usage, fixturing, mechanical resistance, mechanical fastening processes, visual occlusion, and physical obstruction all influence time-based human performance. Designing benchmarking tasks that efficiently sample this design space greatly aids the assessment of a robotic system as a whole, and quickly identify its strengths and weaknesses.

Aside from designing the physical tests, relevant performance metrics must also be carefully considered. For most applications, two very simple metrics that are most important capture speed and reliability. Speed is typically measured as the completion time for a particular task or sub-task. Reliability is captured as the probability of successfully completing a task or sub-task. The theoretical upper bound probability for successfully inserting a component (PS) is calculated given a confidence level (CL), the number of successes (m), and the number of independent trials (n). Given the binomial cumulative distribution function,

$$F(m-1; n, PS) = \sum_{i=0}^{\mathbf{m-1}} \binom{n}{i} PS^i (1-PS)^{n-i} \geq CL, \tag{1}$$

the PS is its minimum value to some precision while still satisfying the above inequality. Both of these metrics are intuitive and relatively inexpensive to measure. Other subsidiary metrics can include the measurement of transmitted forces by the robot during the assembly process, cost-effectiveness of the robotic solution, and energy efficiency. We focus on speed and reliability metrics for the competition.

Another important aspect of performance measurement is providing confidence in the measured results. Consequently, multiple test repetitions of a particular task are required to generate a sufficient amount of data for benchmarking comparisons. Moreover, the use of various statistical tests including tests for correlation, distribution, variance, and mean help identify significant comparative differences in performance data. Conducting these tests can also help reduce the number of false claims that may be issued regarding a robot's level of performance.

4 Proposed Assembly Performance Tests

We present the concept of manufacturing task boards, where each task board design has a manufacturing theme such as insertion, threaded fastening, gear meshing, and electrical connectors (to be designed). The task boards are designed to incorporate standard off-the-shelf components of varying sizes that are representative of components typically used in assemblies. The parts can be presented to the robot system for grasping with various degrees of difficulty ranging from placement in known locations to randomized placement. Note, the use of tools is permitted, although manual changing of end-effector components is not. Although the tests are designed to be accomplished using a single robot arm, any number of arms may be used.

4.1 Insertion Task Board

The insertion task board (Figure 1) is designed to quantify a robot system's capability in performing “simple” peg-in-hole insertions. Relevant experiment design factors include 1) size of peg, 2) cross-sectional shape, and 3) position of peg. Peg-hole clearances are designed to be standard sliding fits with fixed peg lengths. The pegs are of standard metric sizes, and are commercially available in the form of bar stock. Specifically, the edge lengths of the square cross-sectional pegs are 5 mm, 10 mm, 15 mm, and 20 mm. The diameters of the circular pegs are 5 mm, 10 mm, 15 mm, and 20 mm, as well.

The plate insertion geometry cutouts with the necessary sliding fit tolerances can be inexpensively manufactured using an on-line, laser cutting service based on a NIST-supplied design. In the standard test configuration, the plate is fastened to a rigid surface with the gravity vector parallel to the plane of the plate. The test begins with all pegs inserted as shown in Figure 1. The goal is to remove all pegs from one side of the board, and re-insert them from the other side.

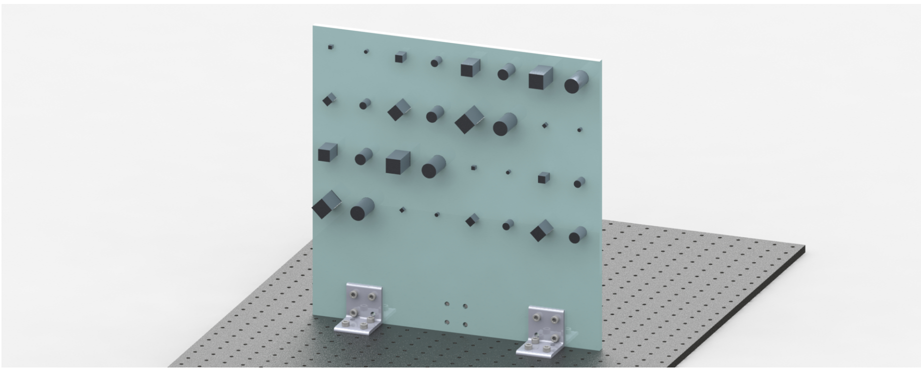

Fig. 1. Insertion Taskboard.

Completion time (CT) is the time required to grasp, move, and insert an individual peg. From the B-D handling table [18], the insertion task board has a '00' handling code for the grasping and manipulation of the pegs with an associated time of 1.13 s by humans. Furthermore, the B-D insertion table indicates a '00' insertion code with an associated time of 1.5 s. Therefore, the theoretical completion time for each peg by a human is 2.63 s. With 32 pegs, the total board should be completed by a human within 84.16 s. Note, both CT and PS can be analyzed with data collected across all pegs simultaneously, or compartmentalized, by dividing data into square pegs and circular pegs, pegs of different sizes, or some combination thereof. Compartmentalization of data can help shed light on the robot system's performance sensitivity with regards to different features of the pegs.

4.2 Fastener Task Board

The fastener task board (Figure 2) is designed to quantify a robot system's capability for fine sensorimotor control. For manufacturing applications, this test seeks to measure a robot's performance at inserting and removing threaded fasteners. Relevant experiment design factors include 1) size of fastener, 2) shape of fastener, and 3) position of fastener.

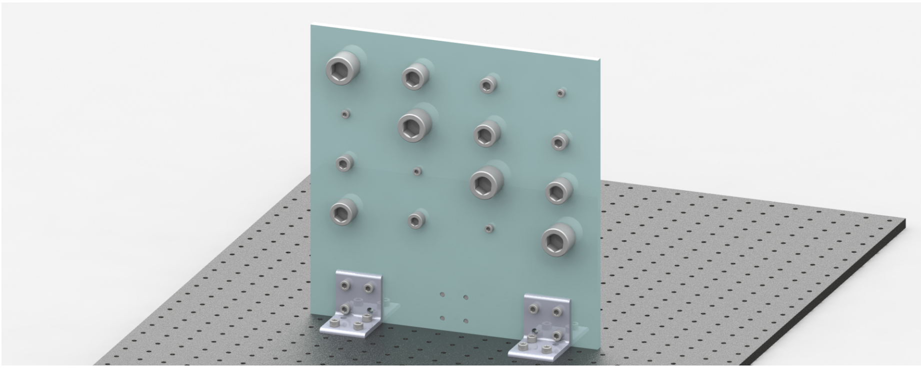

Fig. 2. Fastener Taskboard.

A square aluminum plate is drilled and tapped with a pattern of holes to support four each of M5 x 0.5, M10 x 1.25, M16 x 2.0, and M20 x 4.0 ISO Standard metric bolts, nuts, and washers (Figure 2). In the standard test configuration, the plate is fastened to a rigid surface with the gravity vector parallel to the plane of the plate. The test begins with all fasteners attached to the plate as shown in Figure 2. The robot system under test must then autonomously remove the fasteners and refasten them to the other side of the plate.

Completion time (CT) is the time required to remove fasteners at a particular location, and re-fasten them from the other side of the board. This process requires 1) unfastening nut, 2) grasping and moving nut off to the side, 3) grasping and moving washer off to the side, 4) unfastening bolt, 5) grasping and moving bolt to other side of plate, 6) fastening bolt to plate, 7) grasping and moving washer, 8) inserting washer, 9) grasping and moving nut, and 10) fastening nut. An underlying assumption for the subsequent calculations is that the time to complete an unfastening and fastening step is approximately the same.

An example calculation of CT for a set of M5 fasteners using the above process and the B-D handling codes (Table 1) includes 6 s for step 1, 1.43 s for step 2, 1.69 s for step 3, 6 s for step 4, 1.5 s for step 5, 6 s for step 6, 1.69 s for step 7, 1.5 s for step 8, 1.43 s for step 9, and 6 s for step 10. The total CT for a set of M5 fasteners for a human is then estimated to be 33.24 s. Similar calculations can be made for the other fasteners, and a total CT for the entire board can be estimated.

The theoretical upper bound probability for successfully rerouting a set of fasteners can be calculated using the same inequality as listed before. Again, both the CT and PS measures can be calculated including all fasteners simultaneously, or compartmentalized by subdividing by size of fastener or type of fastener.

Table 1. Handling and insertion codes and times for various fasteners as indicated by B-D tables.

Part(s)	Handling Code	Handling Time (s)	Insertion Code	Insertion Time (s)
M5 Nut	'01'	1.43	'38'	6
M10, M16, M20 Nuts	'00'	1.13	'38'	6
All Bolts	'10'	1.5	'38'	6
M5 Washer	'03'	1.69	'00'	1.5
M10, M16, M20 Washers	'00'	1.13	'38'	6

4.3 Gear Task Board

The gear task board (Figure 3) is designed to quantify a robot systems capability for performing gear meshing. Relevant experiment design factors include 1) gear pitch diameter, 2) gear pitch, and 3) position of gear. The gears are of standard metric sizes, and are commercially available. The design results in four clusters of gears, where each cluster involves gears of the same pitch.

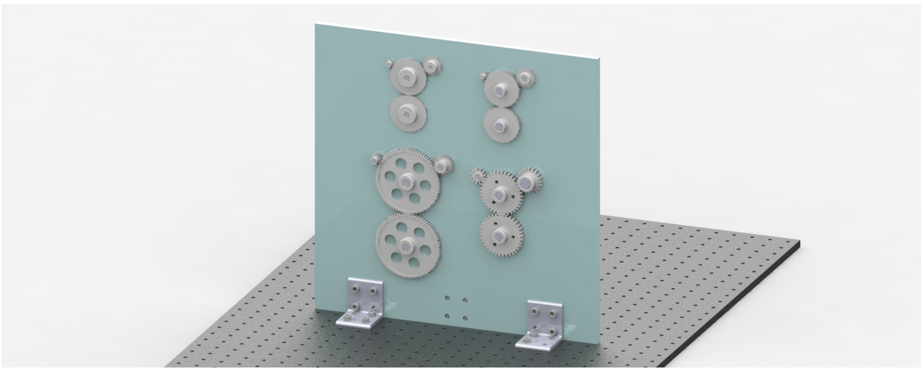

Fig. 3. Gear Taskboard.

The test begins with all gears inserted and meshed as shown in Figure 3. The goal is to remove all gears from one side of the board, and re-insert and re-mesh them from the other side. Note, the use of tools is permitted, although manual changing of end-effector components is not. Moreover, the test is designed to be accomplished using a single robot arm, although any number of arms may be used to accomplish the gear task board.

Completion time (CT) is the time required to grasp, move, and insert a gear. From the B-D handling table [18], the grasping and transportation of the gears in the lower left quadrant have a '00' handling code with an associated time of

1.13 s by humans. Furthermore, the B-D insertion table indicates a '03' insertion code with an associated time of 3.5 s. Therefore, the theoretical completion time for each gear by a human is 4.63 s. This gear cluster should then be completed by a human within 18.52 s.

Once again, the theoretical upper bound probability for migrating gears can be calculated using the previously listed inequality. CT and PS measures can be calculated across all gears, per gear cluster, or per gear.

4.4 Challenge Task Board

The concept of task boards can be extended to support competitions. Competitors are supplied with a set of task boards, one for each of the mechanical assembly topics mentioned above. These are used to develop and test their robotic applications where we provide them with test methods and evaluation techniques to measure their progress. In such a scenario, a subset of each manufacturing topic defined on the boards described above is included in a competition board (including electrical connectors) as shown in Figure 4. Furthermore, the challenge task board is designed to be low cost with readily available components, and NIST will potentially supply them as kits to the competitors at no cost in order to help promote the use of these benchmarking tools.

At the competition, the task board presented to the competitors contains a mixture of assembly components using a subset of the same components defined in practice boards and in a layout previously unknown to the competitors. To accommodate time limitations, teams will only perform one test cycle on the challenge task board. One point will be awarded for every part removed from the taskboard, and one point for every part re-inserted or re-fastened from the other side of the board. The maximum total points is achieved when all parts are migrated from one side of the board to the other.

The rules for completing this task board include 1) no manual end-effector changes, 2) no manual relocation of robot base after initialization, 3) board must remain in upright configuration (but can be re-located by the robot along the working surface), and 4) any number of robotic arms may be used.

We believe there are many benefits to this approach including easy expandability of assembly topics, good initial coverage of particular assembly tasks to gauge competitor capabilities, and benchmarking tools for the assessment of assembly robotics both inside and outside competitions with feedback from users to help improve them. In addition, the modularity of the competition task board facilitates the selection and administration of suitable difficulty levels based on the progress of competitors prior to the competition.

5 Conclusions

The progress of technological advancement and adoption in robotics can be accelerated through rigorous benchmarks and performance evaluations. Competitions have been shown to support the development and dissemination of concepts and draft versions of benchmarks and test methods. To help stimulate a

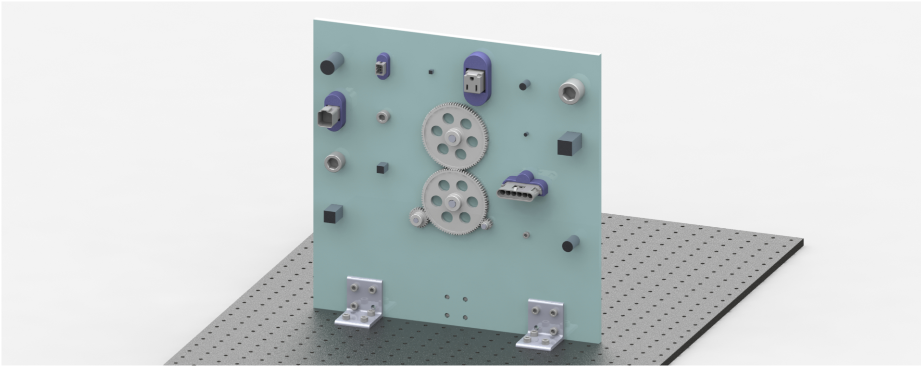

Fig. 4. Challenge Taskboard.

broader understanding of performance requirements for robotic assembly, NIST is participating in a robotic hand grasping and manipulation competition. This chapter described four task boards which present common assembly operations: insertion, fastening, and gear meshing. Given how widespread these operations are, human-based time benchmarks exist and can be used for comparison with robotic solutions. Robotic grasping and manipulation solutions are expected to improve and these task boards provide a means for quantifying this technological progress. Future competitions will incorporate additional assembly-relevant tasks and may expand the metrics captured beyond time and reliability. Aside from competitions, these metrics and task board-based test methods can be useful for understanding the strengths and weaknesses of different hardware and software solutions, yielding a trustworthy foundation for comparing and selecting robotic systems for assembly operations.

References

1. Robotic Hands Grasping and Manipulation RAS TC, `http://www.rhgm.org/`, Access Date: 6/30/2017
2. Robocup, `http://www.robocup.org/`, Access Date: 6/30/2017
3. DARPA Robotics Challenge `http://archive.darpa.mil/roboticschallenge/`, Access Date: 6/30/2017
4. Amazon Picking Challenge `http://www.amazonrobotics.com/#/roboticschallenge/`, Access Date: 8/25/2017
5. F. Amigoni, A. Bonarini, G. Fontana, Matteo Matteucci, V. Schiaffonati: To What Extent Are Competitions Experiments? A Critical View, `http://rockinrobotchallenge.eu/competitions_experiments.pdf`, Access Date: 6/30/2017
6. M. Anderson, O. C. Jenkins, S. Osentoski: Recasting robotics Challenges as Experiments, IEEE Robotics and Automation Magazine, pp. 10-11, June 2011.
7. Jacoff et al: Using Competitions to Advance the Development of Standard Test Methods for Response Robots, Association of Computing Machinery (ACM),

2012, http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=910662, Access Date: 6/30/2017
8. J. Marvel, T. Hong, E. Messina: 2011 Solutions in Perception Challenge Performance Metrics and Results, Proceedings of the Workshop on Performance Metrics for Intelligent Systems, pp. 59-63, 2012.
9. J. Falco, K. Van Wyk: Grasping the Performance: Facilitating Replicable Performance Measures via Benchmarking and Standardized Methodologies, Robotics and Automation Magazine, Vol. 22, No.4, pp. 125-136, December 2015.
10. Performance Metrics and Benchmarks to Advance the State of Robotic Grasping Web Portal, https://www.nist.gov/programs-projects/performance-metrics-and-benchmarks-advance-state-robotic-grasping, Access Date: 6/30/2017
11. A. Downs, W. Harrison, C. Schlenoff: Test Methods for Robot Agility in Manufacturing, Industrial Robot: An International Journal, Vol. 43 Issue: 5, pp.563 572, June 2016.
12. Agile Robotics for Industrial Automation Competition (ARIAC), https://www.nist.gov/el/intelligent-systems-division-73500/agile-robotics-industrial-automation, Access Date: 6/30/2017
13. S. Shneier, E. Messina, C. Schlenoff, F. Proctor, T. Kramer, J. Falco: Measuring and Representing the Performance of Manufacturing Assembly Robots, NIST Internal Report 8090, November, 2015.
14. International Federation of Robotics, World Robotics Report 2016, http://www.ifr.org/industrial-robots/statistics/, Access Date: 6/30/2017
15. Price Waterhouse, The new hire: How a new generation of robots is transforming manufacturing, https://www.pwc.com/us/en/industrial-products/assets/industrial-robot-trends-in-manufacturing-report.pdf, Access Date: 6/30/2017
16. R. Bostelman and J. Falco: Survey of Industrial Manipulation Technologies for Autonomous Assembly Applications, National Institute of Standards and Technology Report, NISTIR 7844, 2012.
17. F. Tobe, Why Co-Bots Will Be a Huge Innovation and Growth Driver for Robotics Industry, IEEE Spectrum, December 2015, http://spectrum.ieee.org/automaton/robotics/industrial-robots/collaborative-robots-innovation-growth-driver, Access Date: 6/30/2017
18. G. Boothroyd, P. Dewhurst, W. Knight: Product Design for Manufacture and Assembly, Third Edition, 2011, CRC Press.