Development of Standard Test Methods for Evaluation of ROV/AUV Performance for Emergency Response Applications

Adam Jacoff and Kamel Saidi
National Institute of Standards and Technology
Intelligent Systems Division
Gaithersburg, Maryland, USA
kamel.saidi@nist.gov

Robert Von Loewenfeldt
South Carolina Law Enforcement Division
Columbia, South Carolina, USA
rvonloewenfeldt@sled.sc.gov

Yukio Koibuchi
University of Tokyo,
Tokyo, Japan
koi@k.u-tokyo.ac.jp

Abstract— This paper discusses the National Institute of Standards and Technology (NIST) and the Department of Homeland Security (DHS) efforts to develop standard test methods for aquatic response robot performance. Different remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) were used to evaluate the test methods and the tests were refined accordingly. Experiments were conducted in order to evaluate the validity of the test methods. Results of those experiments as well as future work are discussed herein.

Keywords—ROV; AUV; underwater robot; test methods; performance; emergency response robot

I. Introduction

Remotely operated vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs) have multiple emergency response applications such as bomb disposal, search and rescue, and disaster response. For example, ROVs and AUVs were used to conduct underwater cleanup and victim recovery following the 2011 earthquake and tsunami in Japan [1]. Bomb squads, search and rescue teams, and disaster response teams need to know the performance of the ROVs and/or AUVs that they own or plan to purchase in order to understand their capabilities and to match those capabilities with the scenarios in which they will be deployed.

A few standards exist for ROVs for industrial applications [2, 3], but these only relate to ROV interfaces and a few basic measures of performance, primarily within the petroleum and natural gas industries. Although a few guides exist [4] there are currently no standard tests to evaluate ROV or AUV performance for emergency response applications.

The National Institute of Standards and Technology (NIST), in cooperation with other organizations, has been leading efforts by the Robot Task Group under the ASTM Sub-Committee E54.08 on Operational Equipment for Homeland Security Applications. The task group has published 15 international standards for response robots [5] to date and these standards have been replicated by dozens of organizations worldwide to measure and evaluate response robot capabilities. The standards address critical needs by helping to inform response robot procurement and deployment decisions with statistically significant robot-capabilities data for a variety of mission-essential tasks. These standards also help guide robot manufacturers toward innovations that answer responder needs while encouraging hardening of developmental systems. To date, ASTM E54.08.01 standards have been used to specify more than $50M worth of response robot procurements for firefighters, bomb squads, and soldiers. These standards are now also beginning to enhance operator training by supporting newly developed measures of operator proficiency.

Although some of the above standards may be applied to land, aquatic, or aerial response robots, the primary focus of the subcommittee so far has been on terrestrial (or ground) robots. The work presented in this paper may be adopted by ASTM E54.08 or by another standards committee or organization.

II. Background

NIST is presently developing a suite of tests that can be used to evaluate basic capabilities of ROVs and AUVs. This paper discusses the subcommittee's past and present efforts to develop test methods for aquatic response robots. The tests include, among others, tasks that are intended for measuring the ROV/AUV's cutting, inspection, station keeping, object retrieval, object placement, sonar resolution, visual acuity, and mapping capabilities. In addition, the maximum thrust and payload carrying capacities are also being considered.

ROV/AUV requirements for search and rescue applications were developed based on input from the

This work was partially sponsored by the US Department of Homeland Security (DHS) Science and Technology Directorate, Test, Evaluation & Standards Division.

user community [6]. The requirements include: structural inspection, leak localization/mitigation, object (body) recovery, water traverse, rapid current station keeping, payload delivery, and object recovery. Bomb squad requirements are somewhat more specialized since ROVs/AUVs must deal with underwater explosives. Bomb squad ROV/AUV requirements were also developed based on user feedback.

In addition to the above requirements, the objective of these underwater test methods, as well as for all the test methods developed for ground and aerial robots, is to provide quick and easy ways to measure capabilities that anybody can replicate and practice to evaluate their own ROV/AUV.

III. Descriptions of the Test Methods and Apparatuses

The proposed test methods and apparatuses described below were developed at NIST and during field exercises. These test methods and apparatuses are still evolving and are considered early prototypes. Changes to initial designs are discussed in Section VI.

A. Bollard Thrust

The bollard thrust test method measures the zero-speed pulling capability of an ROV (i.e., it is not affected by drag or other factors). This is equivalent to the maximum thrust of the ROV. This simple test is important because robot manufacturers often provide values for maximum thrust on their specifications sheets based on theoretical calculations. Instead, the simple apparatus described below measures the ROV's maximum thrust directly.

The test apparatus consists of a stationary mounting point, a cable redirected through a pulley, and a digital force gauge connected to the stationary end of the cable (Fig. 1). The apparatus can be secured to the side of a water tank or to a dock for testing.

B. Inspection/Station Keeping

The inspection (or station keeping) test method measures an ROV's ability to maintain its position in the water under specific conditions (e.g., turbidity or current). This is accomplished by providing a set of targets that the ROV has to inspect (Fig. 2). Each target consists of a series of nested acuity optotypes of decreasing size. The ROV must attempt to identify the orientation of the smallest optotype possible in each target. ROVs with better station-keeping capabilities should be able to perform this task faster.

C. Visual Acuity

The visual acuity test method is equally important in open air as it is under water (where visibility is measurably more impaired). This test consists of placing

Fig. 1. The bollard thrust apparatus.

several charts (Fig. 3), with specific acuity optotypes on them, at a certain distance from an ROV and evaluating the ability of the ROV/operator to read the charts through the ROV's interface.

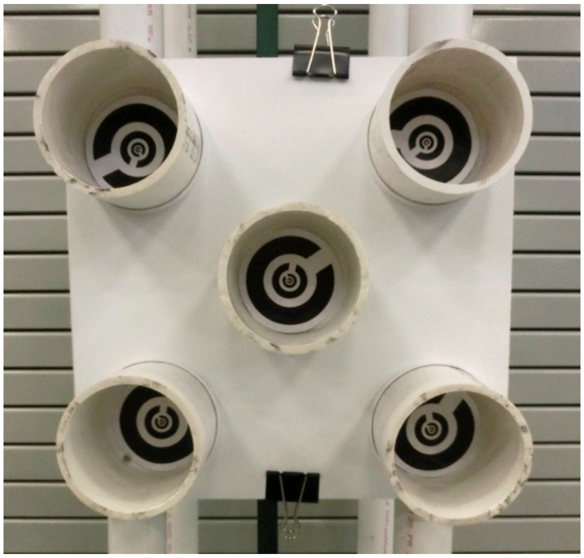
D. Mapping

Underwater mapping of a harbor (e.g., after a hurricane) is a very important task, especially because there are few other options. The same mapping fiducial concept that is used for ground robots [7] was adapted and submerged underwater.

The initial apparatus that was developed consisted of plastic cylinders with aluminum wings that bisected the cylinder (Fig. 4). The apparatus was held up in the water using buoys (orange spheres in Fig. 4) at the top and weights at the bottom. The apparatus depicted in Fig. 4 also incorporated visual acuity targets on the wings in order to simultaneously test mapping and visual acuity.

The cylinder with wings makes for a very distinctive shape when scanned and mapped and its location and orientation should be readily apparent in the sonar image. They can be deployed in single height or double height configurations depending on the water depth available. When they were originally deployed in the double height configuration, the wings were offset 90 degrees from one another. This was changed in subsequent revisions of the apparatus.

Fig. 2. The inspection/station keeping apparatus.



The apparatus is buoyed from the top and anchored to the bottom with two separate anchors so as to minimize translation or rotation while floating submerged below the water's surface.

Map evaluations using these mapping fiducials for ground robots include metrics for coverage, consistency, local accuracy, and global accuracy [7] that can be directly applied to underwater map evaluations.

E. Sonar Resolution

The sonar resolution test method measures the resolving capability of a sonar sensor, which is the prevailing navigation sensor on underwater robots (given that visual sensors are so hindered by water conditions such as turbidity).

The initial sonar resolution apparatus that was developed consisted of a 1 m square sheet of aluminum with a series of 9 holes, of 3 different sizes, cut out of it (Fig. 5). The apparatus is buoyed from the top and anchored to the bottom with two separate anchors so that it does not translate or rotate while floating submerged below the water's surface. Placing the apparatus at a certain standoff distance from a wall would then show reflections from the aluminum and the wall (which would also allow us to determine the sonar's angle of incidence as well as its range resolution). The sizes of the holes could be varied based on the expected resolution of the sonar.

Another design of the sonar resolution apparatus consists of simple 3D geometric objects (such as cubes, cylinders, and pyramids) of varying sizes. The objects are placed anywhere within an ROVs environment or in predetermined orientations (Fig. 6).

Fig. 3. The visual acuity apparatus.

Fig. 4. A double-height underwater mapping fiducial with plastic cylinders and aluminum wings.

F. Rope Cutting

The rope-cutting test method measures the ROV's ability to cut through rope or cable of different materials and thicknesses and in different orientations.

The initial rope-cutting apparatus consisted of 2 different sizes of rope with 10 repetitions of each size, 5 horizontal and 5 vertical cuts. Each rope was equipped with a small shape-/color-coded buoy to indicate success and timing at the surface. Two separate anchors and a floatation buoy centered above the apparatus ensured that it didn't change position and orientation once placed, and that it floated upright (Fig. 7).

Fig. 5. The flat sonar resolution apparatus.

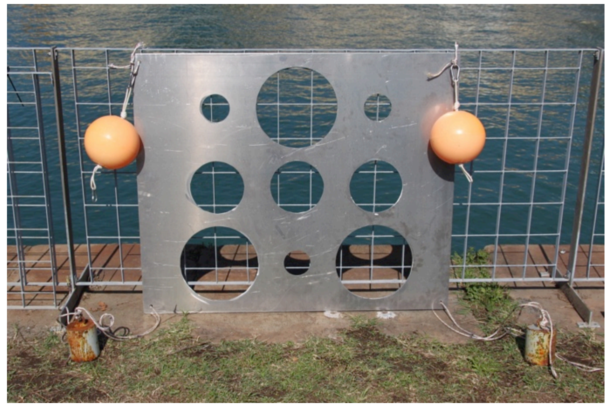
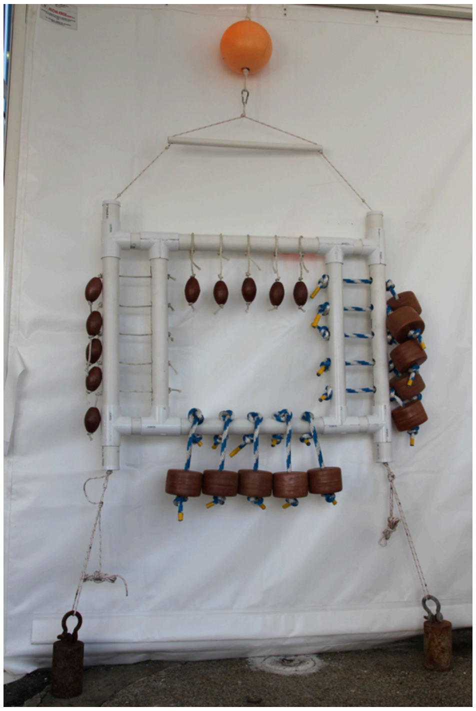


Fig. 6. The three-dimensional sonar resolution apparatus.

Fig. 7. The rope-cutting apparatus.



G. Rod Cutting

The rod-cutting test method measures the ROV's ability to cut through rods of different materials and thicknesses and in different orientations.

The initial rod-cutting apparatus consisted of wooden dowels placed in an apparatus that accommodated two trials of 10 repetitions, 5 horizontal and 5 vertical cuts (Fig. 8). Much like the rope-cutting apparatus, the rod-cutting apparatus also included a buoy attached to each rod.

H. Hooking

The hooking test method measures the ROV's ability to deploy a carabiner (or similar device) to hook onto an object underwater. This type of task is often conducted in search and rescue scenarios.

The hooking apparatus consists of 5 U-bolts arranged in different orientations as shown in Fig. 9. The ROV must attempt to hook the carabiner onto all 5 U-bolts.

I. Soft Grab

The soft grab test method is similar to the hooking test method in that it measures the ability of the ROV to deploy a tool that can attach to something underwater. In the case of the soft grab test method, the tool is an alligator clip (Fig. 10) and the object to hook onto is a soft target that can be made out of foam or cloth (Fig. 11). This type of task is meant to simulate the retrieval of a bag or a victim.

Fig. 8. The rod-cutting apparatus.

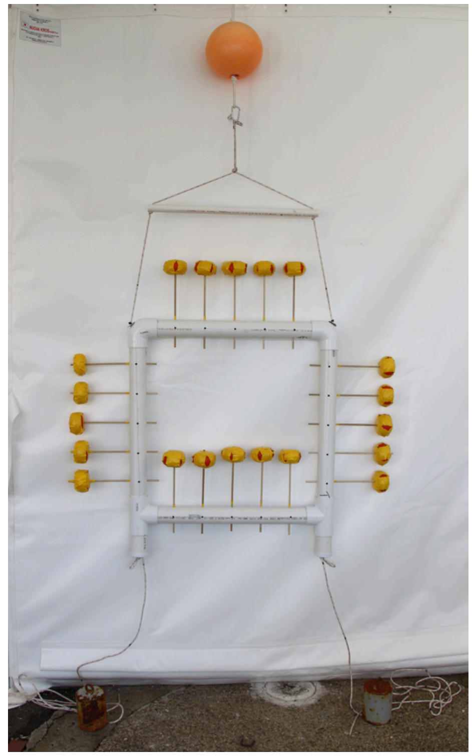
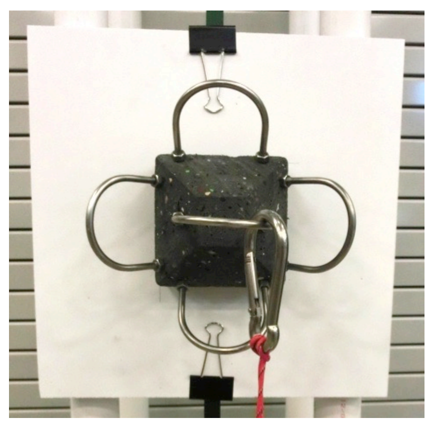


Fig. 9. The hooking apparatus.



J. Grasp, Surface, Swim, and Place

The grasp, surface, swim, and place test measures the ability of an ROV to retrieve an object, surface with the object in its gripper, swim across the surface a certain distance, and then place the object back on the bottom.

The apparatus consists of a weight, 2 target areas (from which to pick and in which to place the weight), and 2 pylons that mark the distance over which the ROV must swim at the surface (Fig. 12).

Fig. 10. The foam version of the soft grab apparatus.

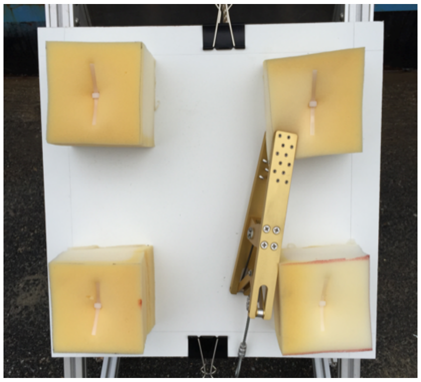


Fig. 11. The cloth version of the soft grab apparatus.

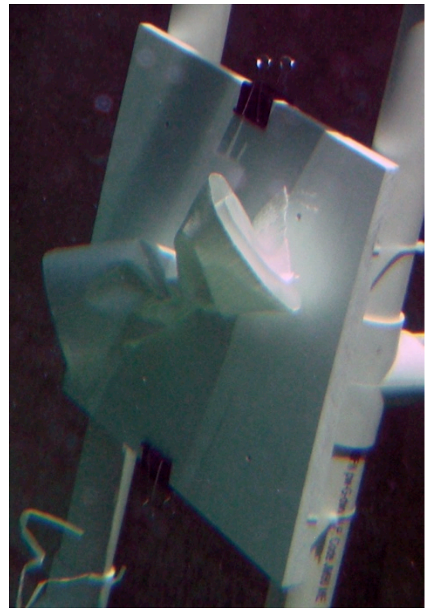
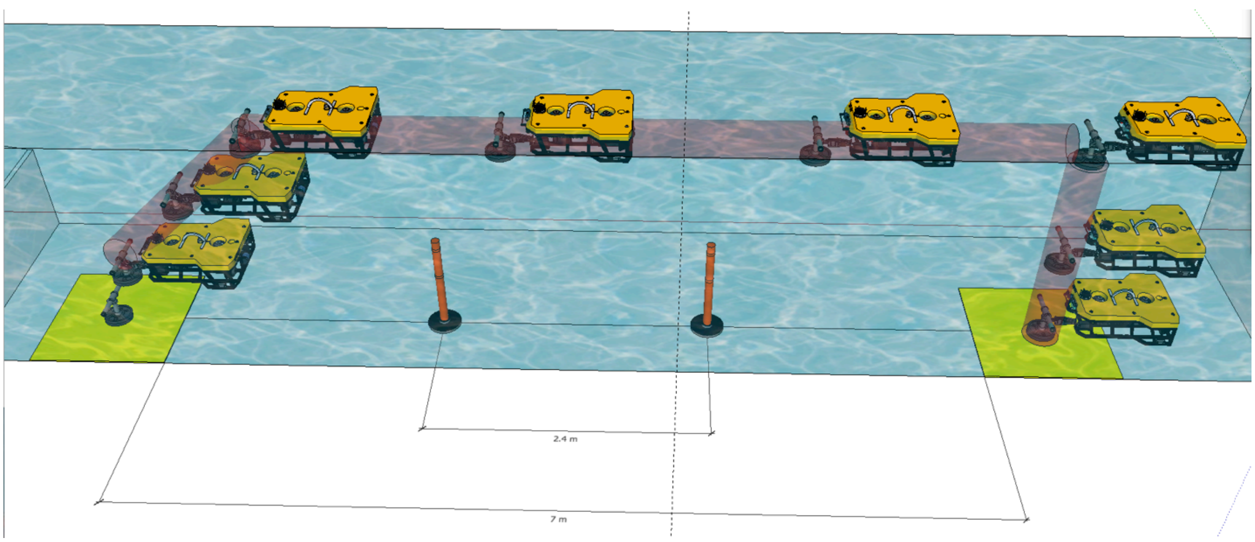


Fig. 12. The grasp, surface, swim, and place apparatus showing a simulated test trajectory.



K. *Disruptor Placement*

The disruptor placement test method measures the ROV's ability to place a magnetized disruptor (an explosive ordinance defeat device) close to an IED (improvised explosive device) attached to a ship's hull.

The apparatus consists of a 40 cm square thin sheet of metal affixed to a plastic board of similar size (Fig. 13). The disruptor is simulated by a short (25 cm) metal pipe with 2 magnets attached to it. The IED is simulated by a plastic box with a single magnet attached to it.

IV. Common Apparatus Carrier Design

Five of the test apparatuses (inspection/station keeping, rope and rod cutting, hooking, soft grab, and disruptor placement) described in the previous section are implemented onto a common apparatus carrier (Fig. 14). The common carrier is made out of polyvinylchloride (PVC) pipe[1] and allows 5 test apparatuses to be mounted onto it at the same time and at 5 different angles (Fig. 15). This provides the ROV with the opportunity to conduct the same test in 5 different orientations, which not all ROVs are capable of achieving.

[1] The original common carrier was constructed out of wood (Fig. 6), but it was found to be too buoyant. The second design was constructed out of aluminum, but it was found to be relatively expensive to replicate.

Fig. 13. The disruptor placement apparatus.

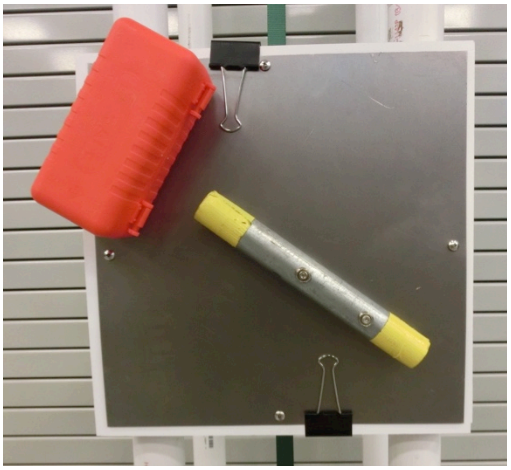


Fig. 14. The common apparatus carrier.

Fig. 15. The common apparatus carrier with 5 hooking tasks.

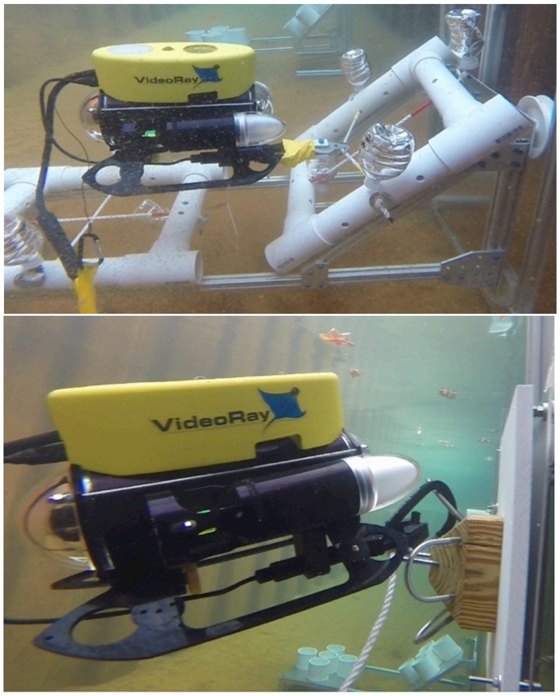
V. Validation Exercises

The majority of the test methods and corresponding apparatuses described above were deployed at multiple locations over the past year. The locations include a 70 m^3 tank at NIST in Gaithersburg, Maryland (Fig. 16); a harbor at the 2014 Eurathlon competition, in La Spezia, Italy (Fig. 17 and Fig. 18); a 70 m^3 tank at the San Diego Fire-Rescue Training Facility in San Diego, California (Fig. 19); and a 70 m^3 tank at the Defense Advanced Research Projects Agency (DARPA) Robotics Challenge (DRC) Finals Expo in Pomona, California (Fig. 20).

Earlier versions of some of the apparatuses were also deployed at Disaster City in College Station, Texas in 2011 (Fig. 21).

During the above exercises, the primary objectives were to validate the practicality and effectiveness of the apparatuses, to start developing procedures for conducting the tests, and to understand the limitations and challenges of the environments into which the apparatuses were deployed.

Fig. 16. An ROV performing a rope-cutting task (top) and a hooking task (bottom) in the NIST tank[2].

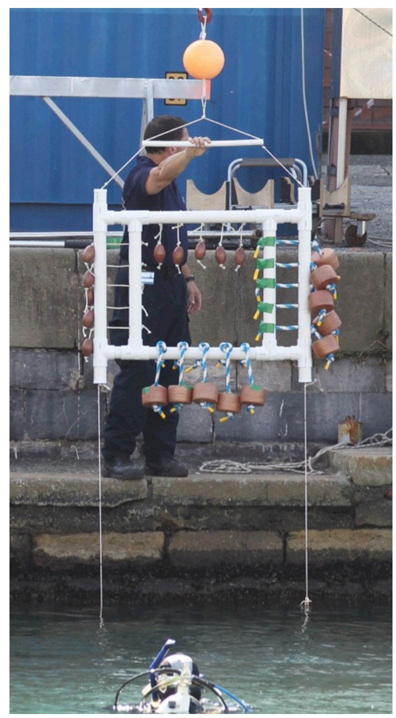


VI. Lessons Learned

Several improvements to and variations of the test methods and apparatuses came about as the result of the validation exercises described above. Some of these changes are summarized below:

- Changed from single-task apparatus that provides only one task orientation at a time to a common carrier with 5 simultaneous task orientations for easier statistically significant testing.
- Changed from a wood common carrier design to an aluminum design with rollers for each task (seen in Fig. 16) for ease of task administration (e.g., replacing cut ropes) and reducing apparatus buoyancy.
- Changed from rollers to slides for ease of fabrication.
- Changed from aluminum common carrier design to PVC for reducing cost and ease of replication.
- Changed from foam soft grab objects to cloth for reducing apparatus buoyancy.
- Changed from multiple sizes of U-bolts, for the hooking task, to one medium-size U-bolt for reducing complexity and as a starting baseline.
- Changed from single height mapping fiducials to double height for increased visibility in deeper water.
- Changed from 90° offset wings in the double height mapping fiducials to non-offset wings for reducing complexity in the sonar data (Fig. 22).
- Changed from plastic to metal cylinders in the mapping fiducials for better sonar reflection (Fig. 22).
- Made various refinements to the common carrier design including quick task attachment using binder clips and different methods of securing the carriers inside a tank.
- Decided that all tests should be conducted under ideal water visibility conditions as a baseline and that turbidity should be measured during each test.
- Decided to implement current to increase station-keeping difficulty.
- Changed from an outboard boat engine (Fig. 23) to submersible pumps (Fig. 24) for generating current to reduce safety concerns and for finer current control.

Fig. 17. A rope-cutting task consisting of 20 tasks in 2 different diameter ropes being deployed in a harbor at the 2014 Eurathlon competition.



---

[2] Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure 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 materials or equipment identified are necessarily the best available for the purpose.

Fig. 18. A combined mapping fiducial and visual acuity target. Top left inset shows a view from an AUV camera while the top right inset shows the output of a sonar system with the target circled in white. The bottom picture shows the target being deployed in a harbor at the 2014 Eurathlon competition.

Fig. 19. ROV tests deployed in a tank at the San Diego Fire-Rescue Training Facility.

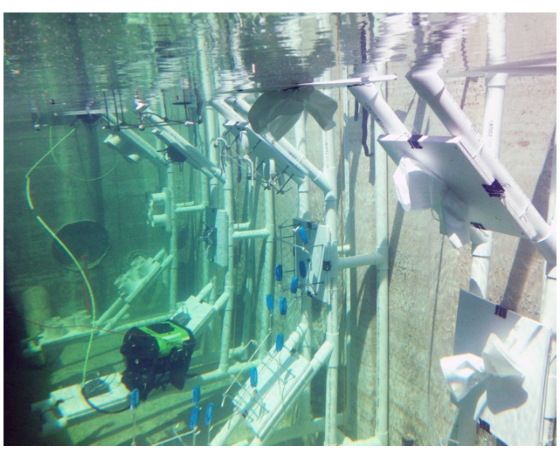


Fig. 20. An ROV performing an inspection task in a tank at the DRC Finals Expo.

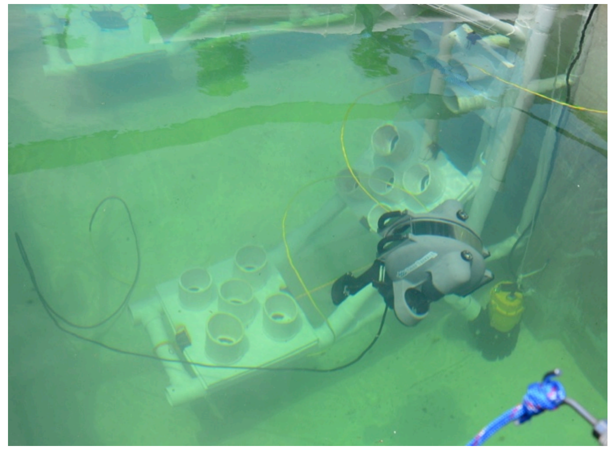
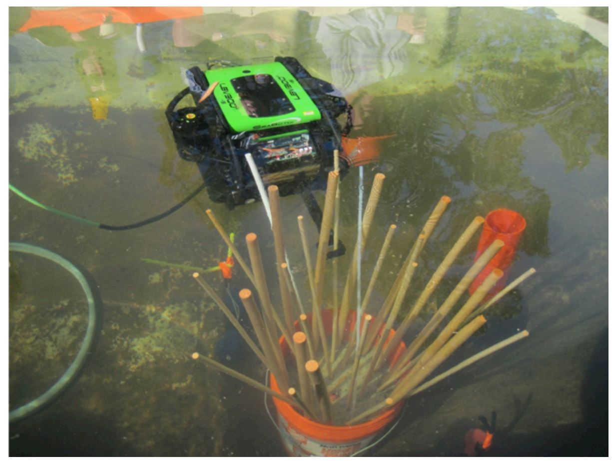


Fig. 21. An ROV performing a rod-cutting task in an early implementation of the apparatus at Disaster City.



VII. Conclusions and Future Work

Several apparatuses for testing ROV/AUV basic performance have been developed and demonstrated. These tests are a first step in developing standard test methods for evaluating ROV/AUV performance. Based on several validation exercises conducted over the past year, the tests were deemed to be challenging enough for many currently available ROVs. The tests do not cover all aspects of ROV/AUV performance, but they are a good first step at quantifying performance.

Future work includes additional exercises to validate the test methods already developed. Specifically, more validations are needed for tests that some small ROVs are not able to complete (such as the grasp, surface, swim, and place and the disruptor placement tasks).

Current apparatuses will also continue to be refined to make them easier and cheaper to implement. In addition, the steps involved in conducting these tests will be further developed into more formal procedures.

Furthermore, although testing ROV/AUV basic capabilities is good for understanding performance, sometimes conducting tasks found in actual scenarios presents unforeseen challenges and helps operators and manufacturers better understand their system's capabilities. To that end, more realistic scenarios will be developed and deployed to assess an ROV's performance after it has successfully navigated the set of basic tests.

Fig. 22. Design for a new double-height mapping fiducial using steel drums.

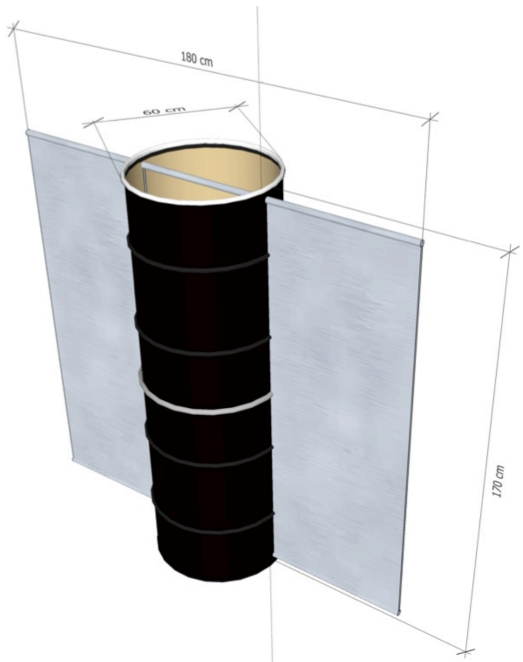
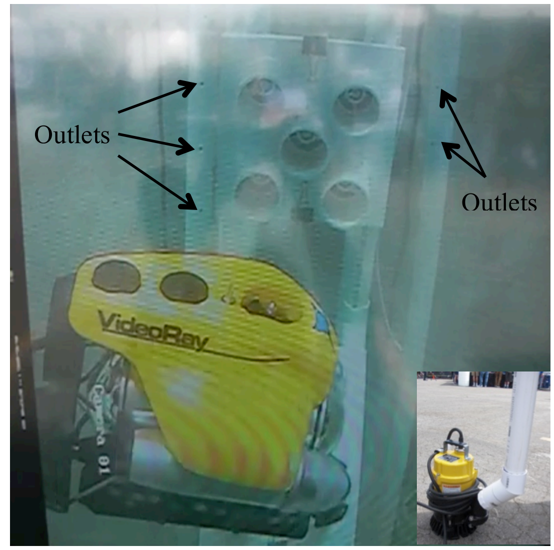


Fig. 23. An outboard boat engine mounted inside a shallow water tank to generate current.

Fig. 24. Current generation using a submersible pump (inset) in front of an inspection/station keeping apparatus inside a tank.



Finally, the water current generation methods will be refined to generate stronger flows as needed; turbidity will be varied in a controlled fashion to assess how an ROV's performance is affected by reduced visibility; and new test methods and apparatuses will be developed as the need for them arises.

REFERENCES

[1] Robin R. Murphy, Karen L. Dreger, Sean Newsome, Jesse Rodocker, Brian Slaughter, Richard Smith, Eric Steimle, Tetsuya Kimura, Kenichi Makabe, Kazuyuki Kon, Hisashi Mizumoto, Michinori Hatayama, Fumitoshi Matsuno, Satoshi Tadokoro, and Osamu Kawase, "Marine heterogeneous multirobot systems at the great Eastern Japan Tsunami recovery," Journal of Field Robotics, Vol. 29, No. 5, pp. 819-831, September 2012.

[2] ISO 13628-8:2002, Petroleum and natural gas industries -- Design and operation of subsea production systems -- Part 8: Remotely Operated Vehicle (ROV) interfaces on subsea production systems, The International Organization for Standardization (ISO), Geneva, Switzerland.

[3] NORSOK U-102:2012, Remotely operated vehicle (ROV) services, Edition 2, September 2012, Standards Norway, Lysaker, Norway.

[4] Committee F41 on Unmanned Maritime Vehicle Systems (UMVS), http://www.astm.org/COMMIT/SUBCOMMIT/F41.htm, Accessed: August 10, 2015.

[5] ASTM Subcommittee E54.08 on Operational Equipment for Homeland Security Applications: http://www.astm.org/COMMIT/SUBCOMMIT/E5408.htm, Accessed: August 10, 2015.

[6] Messina, Elena R. and Jacoff, Adam S., "Performance Standards for Urban Search and Rescue Robots," Proceedings of the SPIE Defense and Security Symposium, Orlando, FL, USA, 04/17/2006 to 04/21/2006.

[7] Sören Schwertfeger, Adam Jacoff, Chris Scrapper, Johannes Pellenz and Alexander Kleiner, Evaluation of Maps using Fixed Shapes: The Fiducial Map Metric, 2010, In Proc. of the Int. Workshop on Performance Metrics for Intelligent Systems (PerMIS), 344-351.