Emergency Response Robot Evaluation Exercise

Adam Jacoff[i]
301 975 4235
adam.jacoff@nist.gov

Hui-Min Huang[i]
301 975 3427
hui-min.huang@nist.gov

Ann Virts[i]
301 975 5068
ann.virts@nist.gov

Anthony Downs[i]
301 975 3436
anthony.downs@nist.gov

Raymond Sheh[ii]
raymond.sheh@robolit.com

[i]National Institute of Standards & Technology
100 Bureau Drive MS 8230
Gaithersburg, MD 20899

[ii]Robolit LLC
1829 Pine Street, Suite 404
Philadelphia, PA 19103

ABSTRACT

More than 60 robot test methods are being developed by a team led by the National Institute of Standards and Technology (NIST) with the sponsorship of U.S. Department of Homeland Security (DHS). These test methods are being specified and standardized under the standards development organization ASTM International. These standards are developed for the purposes of identifying the capabilities of mobile robots to help emergency response organizations assess the applicability of the robots.

The test methods are developed using an iterative process during which they are prototyped and validated by the participating researchers, developers, emergency response users, and robot manufacturers. We have conducted a series of evaluation exercises based on the test method implementations. These events were participated by representatives from all the different segments of the community. As such, these events present a unique opportunity for advancing the test methods, collecting capability data, and identifying robotic technology focusing issues. This paper describes an exercise event that this effort recently conducted.

Categories and Subject Descriptors

J.2 [physical sciences and engineering] unmanned systems performance

General Terms

Measurement, Performance, Design, Human Factors, Standardization, Verification

Keywords

capability, emergency response, evaluation, human-robot interaction, HRI, measure, metrics, mobility, power, radio communications, robot, performance, repetition, sensor, standard, task, test, test method, test suite, trial

1. INTRODUCTION

National Institute of Standards and Technology (NIST), with sponsorship from the Department of Homeland Security (DHS) Science and Technology Directorate, is developing a suite of DHS-NIST-ASTM International Standard Test Methods for Response Robots to quantitatively identify the capabilities of robots for emergency response applications, independent of robot size. These standard test methods identify robot capabilities in mobility/maneuvering, energy/power, sensing, radio communication, manipulation, human-robot interaction, logistics, and safety to provide point of comparison for a variety of robot sizes and configurations prior to testing in more realistic scenarios. Statistically significant test results captured within standard test methods measure incremental system improvements, highlight break-through capabilities, and support procurement and deployment decisions. More than sixty such test methods are under development with associated apparatuses, procedures, and performance metrics. They are being standardized through the ASTM International Standards Committee on Homeland Security Applications, Subcommittee on Operational Equipment, Robots Task Group (E54.08.01). Earlier publications [1, 2, 3] described these test methods development efforts.

1.1 Key Definitions

The term emergency response robot, or response robot, must be defined first. We define it as: a remotely deployed device intended to perform operational tasks at operational tempos that can serve as an extension of the operator to:

- improve remote situational awareness,
- provide means to project operator intent through the equipped capabilities,
- improve effectiveness and efficiency of the mission, and
- reduce risk to the operator.

Key features of a response robot include:

- Rapidly deployed
- Remotely operated from an appropriate standoff
- Mobile in complex environments

- Sufficiently hardened against harsh environments
- Reliable and field serviceable
- Durable and/or cost effectively disposable
- Equipped with operational safeguards

Repetition is a fundamental term used in the effort. It is defined as a robot's completion of the task as specified in the test method and readiness for repeating the same task when required.

Test event or event is defined as a set of testing activities—test methods at various stages of maturity or scenario tasks—that are planned and organized by the test sponsor and to be held at the designated test site(s).

Trial is defined as the identified number of repetitions to be performed by a testing robot for the test results to reach required statistical significance.

1.2 Test Method Focus

These test methods address high-priority tasks identified by emergency responders, including:

- Fast, light, and mobile reconnaissance tasks for throwable robots;
- Wide area survey tasks for hazardous material (HAZMAT) or other events for packable or luggable robots;
- Counter Improvised Explosive Devices (C-IED), Vehicle Borne IED (C-VBIED), and Personal Borne (C-PBIED) tasks for mobile manipulators;
- Aerial reconnaissance for small unmanned aerial systems (sUAS) conforming to the emerging Federal Aviation Administration (FAA) Group I class weighing less than 2 kg (4.4 lbs), less than 30 knots (56 km/hour) maximum speed in horizontal flight, and harmless upon impact;
- Underwater reconnaissance for small remotely operated vehicles (ROV).

For each of these application domains, the standard test methods enable quantitative robot evaluations, provide practice tasks, and help measure operator proficiency.

1.3 Development Process

The standards development process involves hosting periodic robot requirements workshops, standards committee meetings, and response robot evaluation exercises at responder training facilities. Emergency responders, robot developers, and test administrators are gathered around draft standard test methods to practice deployment scenarios. The evaluation exercise events allow emergency responders to articulate essential robot capabilities, validate proposed test methods, and refine performance thresholds and objectives based on objective performance data captured across a class of robots. Emergency responders involved in the process learn about the state-of-the-science in robotic capabilities and help ensure that the test method apparatuses and procedures address their application needs. These events also inform robot developers regarding the reliability and applicability of their robots for actual deployment scenarios, and the ease of use of their systems as they train responders within the test apparatuses. Robot developers involved in the process learn about emerging operational requirements and can demonstrate robotic capabilities by capturing statistically significant performance data within the resulting standard test methods.

2. EMERGENCY RESPONSE ROBOT EVALUATION EXERCISE

The seventh in a series of DHS/NIST Response Robot Evaluation Exercises was hosted at the emergency responder training facility known as Disaster City in College Station, Texas (TX). Thirty emergency responders from across the country participated—half representing DHS Federal Emergency Management Agency (FEMA) Urban Search and Rescue (US&R) teams and half representing bomb squads. They helped validate emerging standard robot test methods, became familiar with available robot capabilities, and advised robot developers regarding operational requirements. All applicable robots were invited to take part in these exercises including ground, aquatic, and the aforementioned sUAS. Robots' capabilities were identified within the implemented emerging standard test methods before being used to familiarize and train the responders with the capabilities and being deployed with the responders to perform operational tasks in the implemented practice scenarios. These designed correspondences between the test methods and the scenarios include:

- Test methods for energy endurance, mobility—covering obstacles and terrains, radio communications—covering line of sight, non-line of sight, and structure penetration, and sensors—covering video acuity, pan-tilt-zoom tasks, 2-way speech intelligibility, range imager resolution, and thermal imager resolution will prepare robots to perform operational tasks for down-range reconnaissance of hazardous material and passenger train wrecks from stand-offs greater than 150 m (500 ft). Figure 1 illustrates a mobility test course featuring crossing ramp terrains.
- Test methods for navigating, searching, and mapping (2D and 3D) complex environments will prepare robots for operational tasks in building interiors and exteriors, partially collapsed structures, and confined spaces in rubble piles.
- Test methods for mobile manipulation—covering non-contact inspection, access tool for window breaking and boring, and grasping/removal tasks will prepare robots for operational tasks to C-IED, C-VBIED, and C-PBIED.
- Test methods for towing trailers and gripper-dragging objects will prepare robots for operational tasks in C-IED, C-VBIED, C-PBIED, and US&R scenarios.
- Test methods for underwater navigation, station-keeping, and sensor acuity will prepare for operational tasks in vehicle reconnaissance in the onsite pond.
- Test methods for air-worthiness, station-keeping, and sensor acuity will prepare small unmanned aerial systems with vertical takeoff, hovering, and landing capabilities for operational tasks supporting several scenarios noted above.

These response robot evaluation exercises introduced emerging robotic capabilities to emergency responders within their own training facilities, while educating robot developers about the performance requirements necessary to be effective in these rigorous application domains. They also helped correlate the draft standard test methods with envisioned deployment tasks and laid the foundation for usage guides identifying a robot's applicability to particular scenarios. The results were the following:

- Refined and validated draft standard test apparatuses, procedures, and metrics

- Quantitative robot capability data to support test method balloting within the ASTM International Committee on Homeland Security
- Feedback for robot developers who allow user training/practice within test apparatuses
- Updated/Expanded Response Robot Capabilities Compendium capturing the trade-offs involved in tested robot configurations showing what robots can and cannot be expected to do reliably in the field

Disaster City is a 210,000 squared meter (52-acre) training facility designed to deliver the full array of skills and techniques needed by urban search and rescue professionals. As part of the Texas Engineering Extension Service (TEEX) at Texas A&M University and a training site for FEMA Texas Task Force (TF) 1 (TX-TF1), the facility features full-size collapsible structures that replicate community infrastructure, including a strip mall, office building, industrial complex, assembly hall/theater, single family dwelling, train derailment, three rubble piles, a C-VBIED scenario, and an underwater vehicle reconnaissance scenario.

2.1 Agenda

This event was held on Monday through Friday, including two days of robot practice and testing within the DHS-NIST-ASTM International Standard Test Methods for Response Robots, two days of robots deploying in operational scenarios with responders, and a final half-day ASTM standards committee meeting to capture feedback.

Day 1 and Day 2: Robot Practice and Testing
November 14-15, 8:00 am Safety Briefing - 5:00 pm Hot-wash

On site were robot developers and test administrators only. All participating robots ran through all applicable test methods, providing practice sessions prior to arrival of the emergency responders. "Expert" operators, chosen by the robot developers to capture baseline performance data and provide developer feedback regarding the test apparatuses and test methods, operated the robots. The robot capability data identified was not to be published. Rather the robot developers were exposed to the entire suite of responder-validated test methods and provided an opportunity to help refine the test methods prior to standardization. In other words, this event was the final opportunity for such refinement for this set of tests.

Day 3: Robot Testing and Operator Training
November 16, 8:00 am Safety Briefing - 5:00 pm Hot-wash

On site were emergency responders representing FEMA Task Force Teams and bomb squads from across the country, robot developers, and NIST administrators. The assembled responders rotated in small groups through all test methods to train on robots prior to deploying them into the US&R training props on site. They became familiar with robotic capabilities using the best performing robots in any given test method. While being exposed to the latest emerging technologies, the responders provided feedback to developers regarding necessary capabilities, operator interfaces, and realistic usage scenarios.

A lunchtime presentation focused on the use of robots in response to Japan's multiple disasters this year. It was presented by the leadership of Japan's International Rescue Systems Institute, who was also a professor at Tohoku University in Sendai where the devastating earthquake and tsunami did the most damage. In addition, a professor from the University of Tokyo discussed the response at the Fukushima Daiichi nuclear facility.

Day 4: Operational Scenarios
November 17, 8:00 am Safety Briefing - 5:00 pm Hot-wash

The emergency responders focused on the most applicable robots to perform targeted tasks in the operational practice scenarios around the site, which included embedded test methods practiced in the previous days. Robot developers accompanied the responders on scenario deployments as observers, advisors, and as operators in particularly difficult deployments to show the potential of robot capabilities. Robot developers onsite, including those whose robots were not selected by responders for deployment, watched the incident response scenarios and observed the robot deployments and absorb the lessons.

A lunchtime presentation focused on the use of standard test methods to provide rapid evaluations of ultra lightweight reconnaissance robots to identify the overall capabilities of the class in support of a rapid fielding initiative by the DoD's Joint Improvised Explosive Device Defeat Organization (JIEDDO). It was presented by a representative from JIEDDO.

Day 5: ASTM International Standards Committee Symposium
November 18, 8:00 am 1:00 pm

The ASTM International Standards Committee on Homeland Security Applications; Operational Equipment; Robots (E54.08.01) hosted a Symposium for all participants to provide feedback on the proposed standard test methods, assess potential operational impact of robots, and define necessary improvements for robots to become useful tools for responders. Presentations included robot developers and other parties have used the standard test methods to measure, refine, and ultimately advertise their capabilities. Robot researchers presented cases where standard test methods helped refine assumptions about the domain tasks and focused their innovation, especially through international robot competitions which used the test methods as challenge arenas. Recent robot procurement efforts were also discussed which have used the test methods to quantify a class of robots or to specify certain combinations of capabilities demonstrated to statistical significance.

2.2 Test Stations and Test Methods

The following subsections describe the test stations and the associated test methods that were set up at the test site. Each of the test methods is noted with its standardization status, as follows:

- (ASTM ####): The document specifying the test method has completed its standardization process and is a published standard.
- (B): The draft document specifying the test method is in the balloting process.
- (V): The draft document specifying the test method is being validated within the ASTM Committee. Robots have begun testing within the test method and results are being collected for analysis.
- (P): The test method is being prototyped. Apparatuses might have been designed or developed.

A test method might also be noted with a Work Item number (WK####), which indicates that the test method has been registered officially with ASTM as a candidate standard and has received the designation. We typically do the registration when the test method is estimated to be about 12 months away from a ballot.

2.2.1 Dispatch Station

- Standard Terminology for Urban Search and Rescue Robotic Operations (ASTM E 2521–07A) [4]
- Standard Terminology for Federal/State/Local Bomb Squads (P)
- Standard Practice for Evaluating Cache Packaged Weight and Volume of Robots for FEMA Urban Search and Rescue Teams (ASTM E2592-07) [5]
- Standard Practice for Evaluating Cache Packaged Weight and Volume of Robots for Federal/State/Local Bomb Squads (P)

2.2.2 Mobility Terrains Station

- Maneuvering Tasks: Sustained Speed (ASTM E2829) [13]
- Maneuvering Tasks: Towing Grasped/Hitched Sleds (ASTM E2830) [14]
- Confined Area Terrains: Continuous Pitch/Roll Ramps (ASTM E2826) [10]
- Confined Area Terrains: Crossing Pitch/Roll Ramps (ASTM E2827) [11]
- Confined Area Terrains: Symmetric Stepfields (ASTM E2828) [12]
- Confined Area Terrains: Gravel (V)
- Confined Area Terrains: Sand (V)
- Confined Area Terrains: Mud (P)

2.2.3 Mobility Obstacles Station

- Confined Area Obstacles: Gaps (ASTM E2801) [6]
- Confined Area Obstacles: Hurdles (ASTM E2802) [7]
- Confined Area Obstacles: Inclined Planes (ASTM E2803) [8]
- Confined Area Obstacles: Stair/Landings (ASTM E2804) [9]
- Vertical Insertion/Retrieval Stack with Drops (V)

2.2.4 Energy/Power Station

- Endurance: Confined Area Terrains: Continuous Pitch/Roll Ramps (V) (W34433)
- Peak Power: Confined Area Obstacles: Stairs/Landings (P)

2.2.5 Radio Communications Station

The test site is at the Riverside Campus Airstrip, 20 minutes away.

- Control and Inspection Tasks: Line-of-Sight Environment (ASTM E2854) [15]
- Control and Inspection Tasks: Non-Line-of-Sight Environment (ASTM E2855) [16]
- Control and Perception Tasks: Structure Penetration (P)
- Control and Perception Tasks: Urban Canyon (P)
- Control and Perception Tasks: Interference Signal (P)

2.2.6 Manipulation Station

- Confined Area Inspection Tasks: Recessed Targets on Elevated Surfaces (V) (WK27851)
- Confined Area Grasping and Removal Tasks: Weighted Cylinders on Elevated Surfaces (V) (WK27852)
- Door Opening and Traversal Tasks (V) (WK27852)

2.2.7 Human-System Interaction Station

- Search Tasks: Random Mazes with Complex Terrain (B) (WK33259)
- Navigation Tasks: Random Mazes with Complex Terrain (ASTM E2853) [17]
- Mapping Tasks: Hallway Labyrinths with Complex Terrain (P)
- Mapping Tasks: Sparse Feature Environments (P)
- Operator Interface Constraints: PPE; Posture; Lighting (P)
- Operator Interface Indicators: Low Battery; Robot Tilt (P)

2.2.8 Sensors Station

- Video: Acuity Charts and Field of View Measures (ASTM E2566-08) [18]
- Video: Pan-Tilt-Zoom Tasks (V) (WK33261)
- Audio: Speech Intelligibility (Two-Way) (V) (WK34435)
- Audio: Spectrum Response Tones (Two-Way) (P)
- Range Imager Resolution (P)
- Thermal Imager Resolution (P)

2.2.9 Safety and Environmental Station

- Water Fording (V)
- Throw Distance Over a 2.4m (8ft) Wall (V)
- Washdown/Decontamination (V) (WK33262)
- Lost Communications Behaviors (P)

2.2.10 Aerial: Small Unmanned Aerial Systems (sUAS) Station

The initial stage were for Vertical Takeoff and Landing, FAA Group I, <2kg, 30knots, frangible.

- Maneuvering Tasks: Station-Keeping: Horizontal and Vertical (V)
- Energy/Power: Endurance (V)
- Safety: Crash Impact Forces (V)
- Safety: Lost Communications Behaviors (P)

2.2.11 Aquatic: Small Remotely Operated Vehicles Station

- Maneuvering Tasks: Sustained Speed (P)
- Maneuvering Tasks: Station-Keeping in a Current (P)
- Maneuvering Tasks: Bollard Thrust (P)
- Manipulation: Cutting Tasks: Rigid and Flexible (P)
- Manipulation: Lifting and Placing Tasks (P)
- Sensors: Video Acuity and Field of View (P)
- Sensors: Sonar Resolution (P)
- Safety: Gripper Drag

2.2.12 Counter Vehicle-Borne Improvised Explosive Devices Station

- Non-Contact Inspection Tasks:
  - Elevated Surfaces with Recessed Targets (0 and 90 degree approach) (P)

- Convex Surfaces with Recessed Targets (Vertical and Horizontal) (P)
- Vehicle Cabs (through window) (P)
- Vehicle Underbody (P)

- Grasping and Removal Tasks:
  - Elevated Surfaces with Weighted Cylinders (0 and 90 degree approach) (P)
  - Elevated Surfaces with Fuel Cans and Propane Tanks (P)
  - PBIED Gripper Drag and Roll-over (P)

- Payload Placement Tasks:
  - Vehicle Underbody Expulsion Disruptors (P)
  - Vehicle Interior Bottle Disruptors (P)
  - Vehicle Interior Overpressure Disruptors (P)

- Tool Deployment Tasks: (part of the robot configuration)
  - Window Breaking and Boring Drills (P)
  - PAN Disruptor Aiming (P)
  - Cutting Straps/Cloth (P)

- Trailer Towing and Placement:
  - Large Vehicle Bomb Disruptors (P)

2.2.13 Operational Scenarios

- Passenger Train Search and Package Removal Tasks
- Hazmat Train Reconnaissance and Retrieval Tasks
- Pancake Collapsed Structure
- Municipal Building and Parking Garage Collapse
- Rubble Piles #1, #2, and the Wood Rubble Pile
- Strip Mall Reconnaissance
- Aerial: Exterior Building Reconnaissance
- Aquatic: Submerged Vehicle Reconnaissance in the Lake

2.3 Test Administration Policy

The suite of standard test methods characterizes the capabilities of robots intended to operate in human-scale, complex environments with variable terrains, lighting, temperature, etc. These current tests are all teleoperation based, although new tests aiming for autonomy are being developed [19, 20]. Each test was assigned an operator station, which was positioned in such a manner as to insulate the operator from the sights and sounds generated at the test apparatus but was within the robot's communications range, except for the radio communications test methods. The operator was required to stay there and use her/his OCU to test the robot—see Section 2.6 for field reset situations. The robot configuration as tested shall be specified in detail to include its size, mass, manipulators, payloads, batteries, communications, etc. This configuration is subjected to the entire suite of test methods. Any variation in robot configuration must be retested across the entire suite of test methods to provide a comprehensive overview of performance characteristics and trade-offs for that particular robot variant. Systems with assistive capabilities or autonomous behaviors should demonstrate improved remote operator/robot performance, efficiency, or survivability of the robot under test. Although these test methods were developed for response robots, they may be applicable to other application domains with modest variations in terrains, targets, or tasks.

2.4 Apparatuses and Targets

The apparatuses associated with these test methods challenge specific robot capabilities in repeatable ways to facilitate direct comparisons of different robot models and particular configurations of similar robot models. Many of the test apparatuses use terrains, targets, and tasks that are intentionally abstract to facilitate the standardization process, which requires capture of repeatable results within a specific test facility and reproducible results across different test facilities. They are generally fabricated using readily available materials to facilitate fabrication by robot developers to support system innovation, refinement, and hardening, and for robot users to support robot evaluation and proficiency training. For example, many test apparatuses are constructed with oriented strand board (OSB) to provide a common friction surface similar to dust covered concrete. The specific terrains, targets, and tasks used can be modified or replaced with more operationally representative examples while using the same apparatuses and procedures to further support training, practice, and comparison of specific system capabilities. These test methods should be considered baseline evaluations and performed prior to more relevant operational tasks defined by robot users. Such operational tasks should leverage a specific set of test methods to establish that robots can perform the necessary capabilities to statistical significance.

Visual targets are used within the test apparatuses to evaluate the visual and color acuity of robots under test in lighted and dark conditions. Visual targets consist of Snellen visual acuity charts, also known as Tumbling E's, and standard hazardous material labels and placards. Snellen Tumbling E's are essentially line resolution tests that can be read through the remote operator station and announced by a robot operator to the test administrator. The test administrator then verifies the reading before scoring the result on the form. A correct reading of a particular line of four Tumbling E's produces a numeric measurement of the visual acuity that can be referenced to average human vision. The visual acuity test method uses comprehensive sets of Tumbling E charts to identify the robot's far field and near field visual acuity. Three line labels shown in Figure 4 are used within other test apparatuses as visual targets to provide an indication of the robot's visual acuity relative to human vision.

Hazardous materials labels provide a variety of standard visual targets that introduce modest complexity for visual identification tasks and operational relevance for some users. The labels contain four attributes including color, icon, text, and number. The text and numbers are sized for average human acuity. Identification of any three of four attributes is considered successful identification of the target.

More operationally relevant objects are used to provide targets for reconnaissance tasks, including simulated pipe bombs, simulated artillery shells, timer devices, cell phones, detonation cords, power sources, etc. Non-visual targets can also be used to test the capabilities of onboard sensors. For example, we have placed trace chemical, radiological, and explosive sources along with these visual acuity targets within the test apparatuses to identify proximity at initial detection and then localization accuracy of sources.

2.5 Test Trials and Statistical Significance

Performance data collections are conducted using the test apparatuses and associated test procedures to capture robot and remote operator performance across a statistically significant number of repetitions. Robots are tested to completion of certain tasks with "expert" operators designated by the developer to capture a task-based capability for a given robot in a given apparatus. The number of repetitions for each test method is determined by ASTM (or the test sponsor) using statistical principles while considering test administration practicalities for longer tests, such as the Endurance test method. The elapsed time of each test is typically not included as a standard metric to de-emphasize speed in favor of task completeness, although the test duration is captured for secondary comparison purposes. Timing measures are typically reported as an average time to perform each repetition, or as an average time to perform a particular sub-task within a test method that can produce varying levels of completeness so that novice operators can quantitatively establish their proficiency as a percentage of "expert" performance within the same test method.

Test trials typically consist of 30 repetitions to demonstrate statistical significance to at least 80 % reliability with 80 % confidence. Successful and failed trials are specifically noted. During the first trial at a particular apparatus setting, the Test Administrator may stipulate that the robot was dominating the apparatus at that setting after demonstrating the first 10 successful repetitions with no failures. However, if there are any failed repetitions, a second or even a third set of 10 repetitions would be required. For a trail to be noted as statistically significant, no more than 1 failure in 20 repetitions, or 3 failures in 30 repetitions, are allowed. This enables setting the apparatus to some known capability and quickly moving toward more aggressive apparatus settings to determine the limit of the robot's capabilities. All subsequent trials must be tested to 30 repetitions for a given apparatus setting.

2.6 Field Maintenance Resets During Test Trials

During a test trial robots may become stuck, inverted, or inoperable. The operator has the option to call a Field Maintenance Reset, which allows the operator to leave the operator station, reset the robot to the start position, and perform routine maintenance for up to 10 minutes (or other limits set by the sponsor). The goal is to allow some interaction with the robot in order to continue the trial to completion. The toolset captured in the cache packaging tools picture and list is allowed with the robot at the start point. No spare parts are allowed (excluding commonly available supply items such as tape and cable ties). A Maintenance/Repair form is to be filled out to include the information on the test method, indication of failure, the remedy, tools used, and overall time to perform the maintenance or repair. The maintenance interaction may be captured on video as well to be used later for training or other purpose. This is intended to be a field maintenance procedure, so the robot is considered to be downrange with some limited number of tools and personnel. However, any person or team of people may interact with the robot at the start point but the robot may not be removed from the start point. The actual list of field maintenance tools necessary to keep the robot operational is evident after the testing is complete along with likely points of failure.

2.7 Abstentions from Test Methods

Each robot configuration should be tested in all applicable test methods and may attempt each test as many times as necessary to attain a satisfactory result. Robots may abstain (through the developer's designee) from a particular test method when considered not applicable or choose not to release the resulting data from a specific test trial when considered not successful. This encourages robot developers to attempt test methods and learn about their systems. In either instance, the page is to be marked as "ABSTAINED" to indicate that the test method was available at test time and the manufacturer acknowledges the omission of performance data. Although some robot implementations may not be designed or equipped for particular test methods, (e.g., robots without manipulators in the manipulator test methods) this testing methodology makes no assumptions regarding capabilities. Specifics of particular robot configurations should be considered when the robot has abstained from a given test method. If the test method is considered critical to the operational needs of the sponsor or user, the test should be considered failed until the robot can demonstrate satisfactory performance at a later date.

If a robot returns to the test facility at a later date to quantify improvements in capability for a particular robot configuration, the robot is to be subjected to a subset of tests representing each of the test method suites. For example: Energy/Power: Endurance; Radio Comms: LOS & NLOS; Mobility Terrains: Crossing Ramps; Mobility Obstacles: Inclined Plane; Sensors: Pan-Tilt-Zoom Tasks, Human-Robot Interaction: Random Maze Navigation.

3. RESULTS

The event was conducted according to the schedule and was actively participated by all who registered. Over 150 test trials were conducted with the results captured to support the respective purposes. The following elaborates the results in detail.

3.1 Participation

An evaluation event like this presented a unique environment where participants with different roles integrated to evolve the technology and test methods for emergency response robots. The following are the composition of the participation (See Figure 2 and Figure 3 for group pictures):

(A) DHS Sponsor
The project sponsor is onsite to provide guidance.

(B) Robots
There were over 25 robot configurations (i.e., some particular robot models brought multiple units to the event) or robotic special tools participated, which can be categorized as:

- Over 20 ground robots or robotic special tools
- Two aerial robots Small Unmanned Aerial Vehicles (sUAS) (FAA ARC Group I under 2 kg, 30 knots, frangible)
- Four aquatic robots, or customarily called ROVs
- Over 22 robots or robotic special tools from the U.S. and 4 from overseas
- Over 21 commercially available and 5 from research organizations

Note that a robot might belong to multiple categories.

(C) Participating Emergency Responders
15 FEMA US&R members, representing the following teams: California (CA)-TF1, CA-TF2, CA-TF3, CA-TF6, Colorado (CO)-TF1, Florida (FL)-TF2, Indiana (IN)-TF1, New York (NY)-TF1, TX-TF1, Virginia (VA)-TF1, and Washington (WA)-TF1.

15 Bomb technicians, representing the following teams: Boca Raton, FL Police Department (Dept.), Chico, CA Police Dept., Florida State Fire Marshal's Office, Garland, TX Police Dept., Jacksonville, FL Sheriff's Office, Michigan State Police, Montgomery County, Maryland Fire/Explosive Investigations, New Jersey State Police Arson/Bomb Unit, Odessa, TX Police Dept., Sacramento, CA Sheriff Office, Santa Clara, CA Sheriff's Office, and Seattle, WA Police Arson/Bomb Squad.

(D) Research and Development, Test Method Design, Set Up, and Administration Personnel

- Representatives from Southwest Research Institute, USA and from Mitre Corp., USA – Mobility Test Methods
- Representatives from a USA robot company and from US Army Aberdeen Test Center, USA – Sensors Test Methods
- Representatives from Pennsylvania State University – Energy/power Endurance Test Method
- Representative from NIST, Boulder, Colorado site – Radio Communications Test Methods
- Representative from Jacobs University, Germany – Mapping Test Methods
- Representative from Nagaoka University of Technology, Japan – Safety and Environment Test Methods
- Representative from a robot company – Aquatic Test Methods
- Representative from Ryerson Univ., Canada – Aerial Test Methods
- Representative from Bureau of Procurement, Germany – General support and advice
- NIST team from the Gaithersburg, Maryland site – host of the event

(E) Site Support
Disaster City administration, TEEX, assigned a team to support the operation.

(F) General Audience
Many participated for general interests, representing various DOD organizations and other Government agencies, USA and International industries, and various research organizations.

3.2 Resulting Capabilities Compendium

Test results, over 150 sets, were organized for different purposes. They are used to support the repeatability analysis in the test method standards. They are also extremely valuable information for the emergency response communities. The following subsections describe these in detail.

3.2.1 Bar Charts

The graphical test forms associated with each test method provide an intuitive understanding of the robot's capabilities in order to facilitate side-by-side comparisons. However, there are dozens of test methods in the suite and users of the data benefit from comparisons across the entire class of robots. Bar charts such as those shown in Figures 5 through 8 help identify Best-In-Class robots in specific test methods, and allow initial identification of trade-offs for particular robot configurations. But once a search is narrowed to several robots, a detailed study of the associated performance data forms is recommended.

In Figure 5, each bar along the X axis clearly represents the robot's tested average speed in the continuous ramp test terrain. Figures 6 and 7 represent the robots' capabilities in the increasingly more difficult crossing ramp and stepfields terrains, respectively. For example, the leftmost robot's speeds were (10, 5 and 0) meters/minute from Figure 5 through 7. 0 means that the robot was not able to complete the test. Figure 8 shows the robots' combined test results in these three and all the other terrains as listed Section 2.2.2. Our goals of providing intuitive representations to facilitate capability identifications have been achieved through the illustrations of these charts.

3.2.2 Comparison and Trade-Off Software Tool

We are also developing a software tool called Response Robot Capabilities Compendium, which contains capability data from all robots that achieve statistically significance within the DHS-NIST-ASTM International Standard Test Methods for Response Robots (See Figure 9). Currently, NIST has conducted all the testing as part of the standards development process. Additional test facilities recently opened in San Antonio, Texas, Kobe, Japan, and Koblenz, Germany. They will start contributing tested robot capability data, soon. Yet additional organizations from various parts of the world are anticipated to interact with this effort and request our help to establish similar testing facilities, in the near future. Given the myriad combinations of robot sizes, weights, and capabilities, a software interface into the database is the best way to understand the implications of specifying certain attributes or performance thresholds. This interface allows the user to see which robots have demonstrated statistically significant performance for the highest priority capabilities necessary to perform their intended mission. They can quickly see the effects of specifying too stringent a requirement in any particular capability or attribute as the number of robots that have successfully demonstrated the specified combination are filtered. Backing off on the threshold for even one requirement can bring several more robots into consideration. So users quickly learn the trade-offs involved and what the state of the science can deliver with regard to the combination of attributes and capabilities they have in mind.

Figure 9 illustrates that the candidate robots were filtered through after a user identified the requirements in LOS, endurance, and stair traverse. The tool allows for setting up multiple levels of detail on the information display. Any confidential information will be activated or de-activated properly before the tool is delivered to a user.

4. SUMMARY AND FUTURE EFFORTS

The event was implemented as an integrated exercise environment. The emergency responders enjoyed great learning experience of the robotic capabilities. The robotic developers were provided great opportunities to exercise the robotic systems and to explore technology advancement opportunities. The test method developers and administrators were immersed in a great environment to evolve the test methods.

Overall, the event exercised our emphasis on repeatable and scalable testing and evaluation processes. The user communities have already enjoyed successes in applying the results, including robot acquisition and responder proficiency training. We plan to continue expanding the scope of this process and methodology to include testing robots with various levels of autonomy, to further explore advanced robotic requirements, and to cover robots applying to additional domains.

COMPANY/PRODUCT DISCLAIMER

Certain trade names, products, or other types of identifying information might be used to facilitate communications. In no case does such an identification imply recommendation or endorsement by the NIST, nor does it imply that the names or products are necessarily the best available for the purpose. In addition, the data are presented to facilitate prototyping, validation, and standardizing the corresponding ASTM test methods undertaken by its E54.08.01 Robotics Task Group. In no case do the data and the associated representations imply any type of recommendation, endorsement, or judgment by NIST.

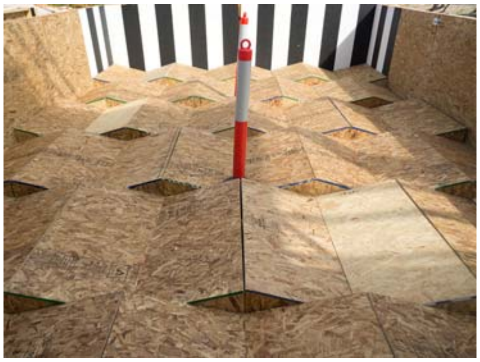

Figure 1: Mobility Test Station Crossing Ramps

Figure 2: Part of the Participants of the 2011 Event

Figure 3: Part of the Participants of the 2010 Event

Figure 4: Visual Target

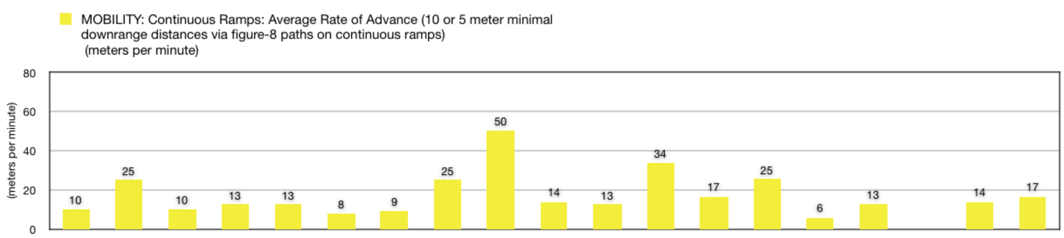

Figure 5: Continuous Ramps Terrain Test Results for Individual Robots

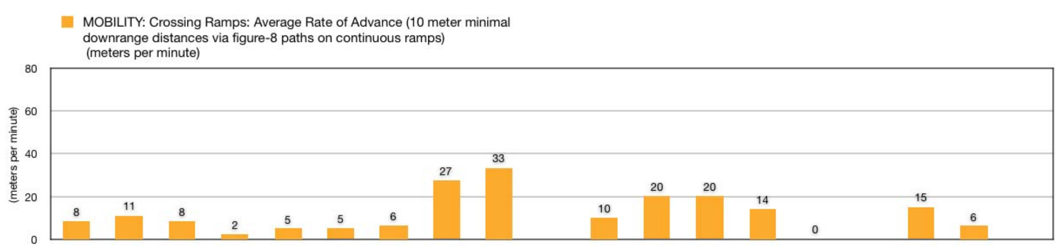

Figure 6: Crossing Ramps Terrain Test Results for Individual Robots

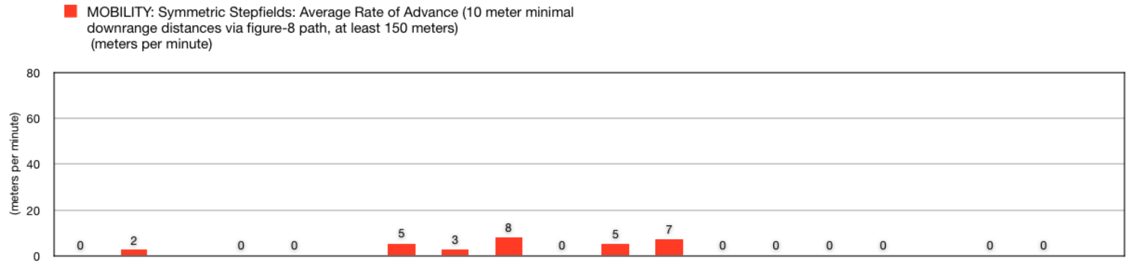

Figure 7: Stepfields Terrain Test Results for Individual Robots

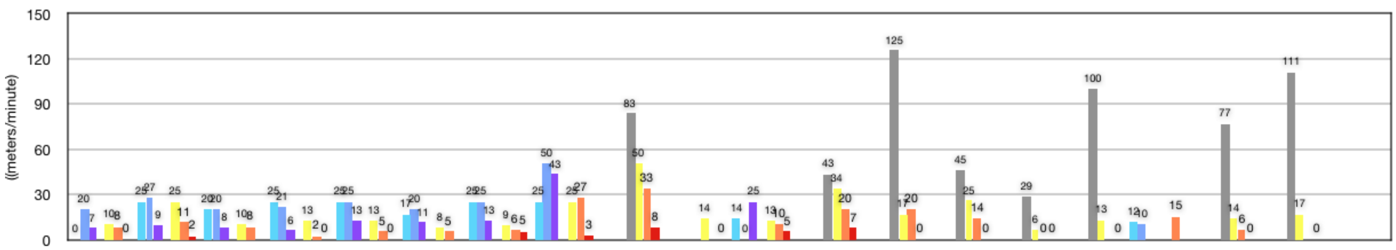

Figure 8: Combined Terrain Test Results for Individual Robots

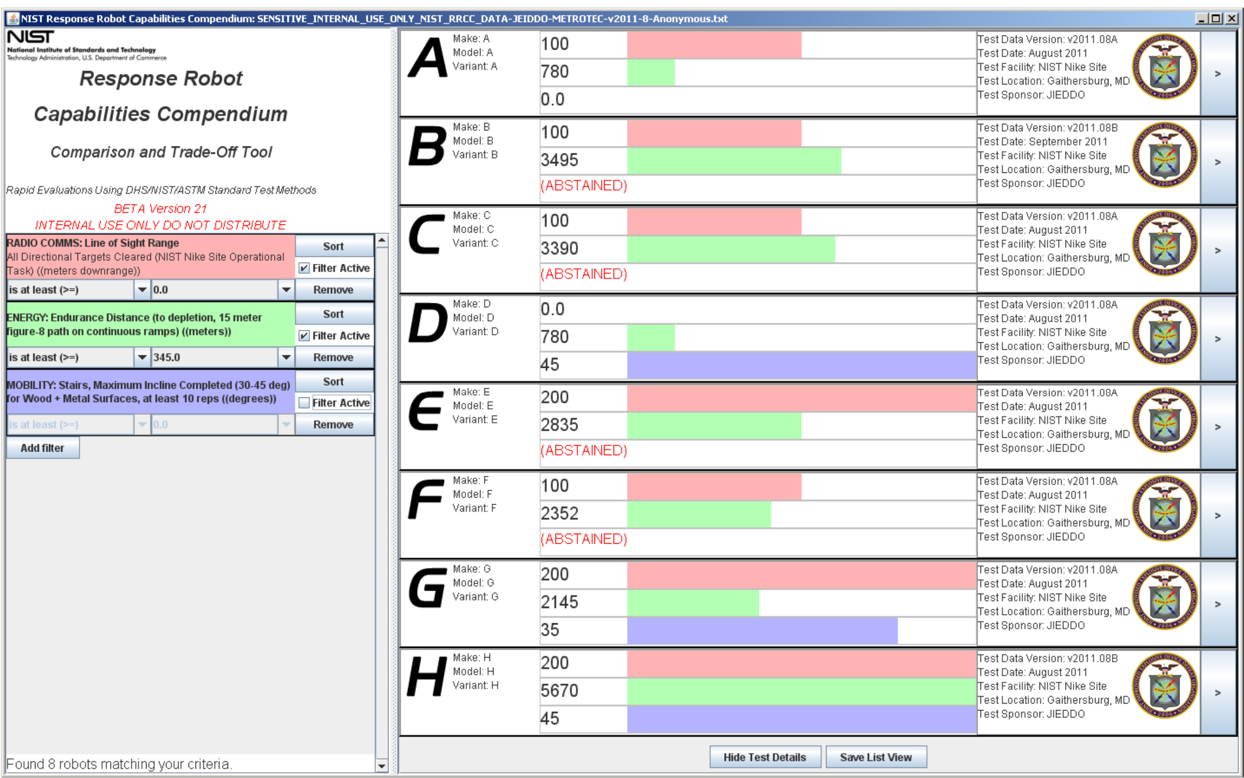

Figure 9: Compendium Illustration

REFERENCES

[1] Jacoff, A., et al., "Comprehensive Standard Test Suites for the Performance Evaluation of Mobile Robots," PerMIS Workshop, 2010

[2] Jacoff, A., and Messina, E, "Urban Search and Rescue Robot Performance Standards: Progress Update," SPIE Defense and Security Conference 2007.

[3] Messina, E. and Jacoff, A. S. "Measuring the Performance of Urban Search and Rescue Robots," IEEE Conference on Homeland Security Technologies, 2007

[4] ASTM International Standard E 2521 – 07a: Standard Terminology for Urban Search and Rescue Robotic Operations

[5] ASTM International Standard E 2592 – 07: Standard Practice for Evaluating Cache Packaged Weight and Volume of Robots for Urban Search and Rescue

[6] ASTM International Standard E2801: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Obstacles: Gap

[7] ASTM International Standard E2802: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Obstacles: Hurdles

[8] ASTM International Standard E2803: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Obstacles: Incline Planes

[9] ASTM International Standard E2804: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Obstacles: Stairs/Landings

[10] ASTM International Standard E2826: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Terrains: Continuous Pitch/Roll Ramps

[11] ASTM International Standard E2827: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Terrains: Crossing Pitch/Roll Ramps

[12] ASTM International Standard E2828: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Terrains: Symmetric Stepfields

[13] ASTM International Standard E2829: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Maneuvering Tasks: Sustained Speed

[14] ASTM International Standard E2830: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Maneuvering Tasks: Towing: Grasped/Hitched Sleds

[15] ASTM International Standard E2854: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Radio Communication: Line-of-Sight Range

[16] ASTM International Standard E2855: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Radio Communication: Non-Line-of-Sight Range

[17] ASTM International Standard E2855: Standard Test Method for Evaluating Emergency Response Robot Capabilities: Human-System Interaction (HSI): Search Tasks: Random Mazes with Complex Terrain

[18] ASTM International Standard E2566 – 08: Standard Test Method for Determining Visual Acuity and Field of View of On-Board Video Systems for Teleoperation of Robots for Urban Search and Rescue Applications

[19] NIST Special Publication 1011-I-2.0 Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume I: Terminology, Version 2.0, Huang, H., Ed., 2008 http://www.nist.gov/el/isd/ks/autonomy_levels.cfm

[20] NIST Special Publication 1011-II-1.0 Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume II: Framework Models, Version 1.0, Huang, H., et al., Ed., 2007 http://www.nist.gov/el/isd/ks/autonomy_levels.cfm