

# DHS/NIST Response Robot Evaluation Exercises

Adam Jacoff

National Institute of Standards and  
Technology  
100 Bureau Drive  
Gaithersburg, MD 20899  
Adam.jacoff@nist.gov

Elena Messina

National Institute of Standards and  
Technology  
100 Bureau Drive  
Gaithersburg, MD 20899  
Elena.messina@nist.gov

**Abstract**— The Department of Homeland Security, through the Science and Technology Directorate Standards Program, is developing performance standards for robots applied to urban search and rescue (US&R). The National Institute of Standards and Technology (NIST) is leading this effort with collaboration from subject matter experts within the Federal Emergency Management Agency (FEMA) US&R Task Forces and other response organizations, along with robot manufacturers and robot researchers intent on this application domain. NIST organizes events that bring emergency responders together with a broad variety of robots and the engineers that developed them to work within actual responder training facilities. These informal response robot evaluation exercises provide collaborative opportunities to experiment and practice, while refining stated requirements and performance objectives for robots intended for search and rescue tasks. This paper summarizes the experiences from a recent exercise held in Texas.

**Keywords:** *performance metrics, standards, robots, urban search and rescue, response robots*

## I. INTRODUCTION

Response robot evaluation exercises introduce emerging robotic capabilities to emergency responders while educating robot developers regarding the performance requirements necessary to be effective, along with the environmental conditions and operational constraints necessary to be useful. They also provide an opportunity to refine emerging test methods and associated test artifacts being developed to measure robot performance in ways that are relevant to emergency responders. Conducting these events in actual US&R training scenarios helps correlate the proposed standard test methods with envisioned deployment tasks and lays the foundation for the usage guides which will identify which robot categories appear best suited for particular response tasks. The resulting standard test methods and usage guides for US&R robots will be generated within the ASTM International Homeland Security Committee through the E54.08 Subcommittee on Operational Equipment.

The second in an ongoing series of response robot evaluation exercises for FEMA US&R teams was hosted at the Texas Task Force 1 (TX-TF1) training facility known as Disaster City, which is located at Texas A&M University, College Station, TX. Applicable robots and supporting technologies (e.g., sensors), purchasable or developmental, were invited to take part in this exercise which highlighted operationally relevant US&R scenarios specifically devised for ground, aerial, and underwater response robots. The robots themselves were not formally evaluated during this exercise.

The event included three days of robot evaluations in available US&R training props. The first two days allowed the assembled responders to deploy the robots within the training props, become familiar with emerging technologies likely to provide benefits in the near term, and provide feedback to developers regarding realistic usage. On the third day, the emergency responders chose the most successful robots from the previous 2 days to perform targeted (and practiced) tasks in a 4 hour mock incident response exercise, which also included several canine teams as well. The robot developers acted as advisors/observers for the US&R teams during this exercise. An informal after action briefing was held on the morning of the fourth day to distill applicable knowledge gained during the event and to refine the design parameters for the test methods proposed for standardization. All stakeholders were encouraged to provide feedback on the proposed test methods. This paper summarizes the event; for a more complete rendition, please see [1].

## II. EXERCISE PARTICIPANTS

NIST's team of test engineers and support personnel worked closely with the TEEX/TX-TF1 personnel throughout the planning, setup, and administration of this event, which accommodated roughly seventy people and more than thirty robots across ten different scenario props at Disaster City. The

TEEX/TX-TF1 personnel very ably managed the overall logistics on site, which contributed greatly to the success of this event and ensured safe operations throughout.

The primary participants from the emergency responder community were representatives from FEMA US&R Task Forces, as has been the case throughout the DHS/NIST performance standards program for US&R robots. Some non-FEMA responders who are members of the ASTM standards task group also participated. One canine team participated throughout the event and was joined by several more canine teams for the final day mock incident response.

As for robot participation, there were 16 different models of ground vehicles, 2 models of wall climbers, 7 models of aerial vehicles including a helicopter, and 2 models of underwater vehicles. Two dynamic simulation environments were also available for visualization of high-fidelity robot models within realistic practice environments (including props at Disaster City). The robots represented 9 of the 13 envisioned US&R deployment categories identified in earlier workshops. [2][3] There were multiple instances of some of the more mature models available. Representatives from the robot developers/manufacturers typically deployed their own robots, but some were deployed by the Alliance for Robotic Assisted Crisis Assessment & Response (ARACAR), a non-profit group that has a large cache of robots and is collaborating on the overall robot performance standards effort.

### III. SCENARIOS

This section briefly describes the training scenarios, sometimes referred to as props, that were used during this exercise. Responders identified access points within each scenario during the initial orientation, but had some flexibility regarding how to approach the search mission once they had a robot in hand. Some scenarios had multiple entry points. Figures 1 through 7 illustrate the scenarios.

The responders were organized into four different teams that rotated across each scenario. Similarly, four teams of robots were created, primarily based on compatibility of their wireless communications. One of the challenges of deploying robots is the fact that many use the same radio communication frequencies, which can cause debilitating interference on site. Some robots used tethered communications at times to avoid these issues. The robot teams rotated through two different scenarios each day. The responder teams rotated twice as fast, through four scenarios each day, to work with as many different combinations of robots and scenarios as possible over the three days. Responders rotated to each scenario for 90 min, spending 45 min at two different start points within the scenario working with two different robots. During aerial operations, all interested responders were at that scenario to work with the aerial robots sequentially. Ground robots that could run tethered to avoid any radio conflicts with the aerial

vehicles were allowed to operate simultaneously on any other scenario. Every possible combination of responders/robots/scenarios was not quite achieved given the limited time available.



Figure 1: House of Pancakes Operational Scenario



Figure 2: Single Family Dwelling



Figure 3: Strip Mall



Figure 4: Partial View of Rubble Pile



Figure 5: Wood Pile



Figure 6: HazMat Train



Figure 7: Passenger Train

A mock incident response on the afternoon of the last day allowed the responders to focus on specific scenarios employing the robots of their choice. Tethered operations were encouraged to limit radio interference and to ensure that responders had experience with the benefits and challenges of using tethered robots.

In each of these scenarios, NIST embedded simulated victims, or “victim props,” that the responders were to locate using the robots. These simulated victims emitted assorted combinations of signs of life: human form (mannequin parts), thermal signature (heating blankets and pads stuffed into clothing), movement (waving or shifting), sound (yelling or moaning), CO<sub>2</sub> (in confined spaces). Examples of victims embedded in the scenarios are shown in Fig. 8 below.



Figure 8: Example simulated victims

#### IV. EMERGING TEST METHODS

A set of test methods designed to address specific responder-defined robot requirements were set up in and around the theater building and embedded into several scenarios. This provided an opportunity to refine these test methods based on feedback from responders and developers as they used them for practice and operator training. The initial test methods and artifacts are described briefly below. Another iteration will take place late Summer 2006 to incorporate feedback from the Disaster City event and the resulting test methods will be introduced into the standardization process through the ASTM International E54.08.01 task group.

#### VI. DATA COLLECTION AND CONCLUSIONS

This event provided a focused opportunity to capture feedback from responders and manufacturers. Questionnaires regarding the scenarios and the test methods captured the impressions of all the stakeholders. Further feedback was collected from the responders only during a “hot wash” review meeting immediately following the event. Copious images and video of the robots in action were also collected. This section describes briefly the data collected.

The organizers collected images and videos of robots and personnel participating in the event. Each robot developer will receive all media related to their robots. Highlight images and generally successful robot videos can be found on the NIST project home page:

[http://www.isd.mel.nist.gov/US&R\\_Robot\\_Standards/](http://www.isd.mel.nist.gov/US&R_Robot_Standards/).

Responder feedback was captured regarding the relevance of the scenarios as training props and the operation of robots within the scenarios. The focus was on how effective different robots were within the scenarios to get a general sense of how well the responders felt they operated as a team with the robots. The numeric responses to the questionnaire shown were averaged. Analysis of the responses suggested certain trends:

- All scenarios were considered representative, with the Hazmat Train, House of Pancakes, and Rubble Pile scoring the highest.
- Concerning how representative the tasks they were expected to perform within the scenarios, the Rubble



Pile and House of Pancakes scored highest.

- Teams scored their performance most highly at the HazMat Train, House of Pancakes, and Strip Mall.
- The robot capabilities were given the highest scores for the HazMat Train, Strip Mall, and Wood Pile. The utility of the robots to the scenario was rated highest in the HazMat Train, Wood Pile, and House of Pancakes.
- The time to complete the scenario was considered most appropriate for the Dwelling and House of Pancakes.
- The robot-responder team performance was highest at the Strip Mall, Wood Pile, and House of Pancakes.
- The quality of the operator interface and interactions was highest at the HazMat Train, Dwelling, Wood Pile, and House of Pancakes.

Data were also captured using draft test method forms. Operators self-declared their level of expertise operating the robots and ran through some of the test methods. Times and other measures of performance were captured, primarily to begin characterizing the operational ranges that the tests ought to support and to pilot the procedures for the test methods themselves.

The majority of the feedback captured was in the form of informal qualitative input from the responders themselves. One of the goals of this exercise was to determine which categories of robots were thought to be the most close to being deployable (both in terms of technological maturity and relevance to disaster response). The responders felt that the following categories of robots were nearly ready to field:

- Small, throwable, so-called “peek bots.” Robots that are able to be deployed into very confined spaces and send video or potentially sensor data back to the operators.
- Aerial survey robots that could “look over the hill” to assess the situation and determine at least which roads are passable. US&R teams can save valuable time if they can determine whether a roadway is blocked. They don’t necessarily expect aerial robots to assess structural integrity or even detect victims. They would like to be able to monitor atmospheric conditions from these platforms as well.
- Wide-area survey robots, which could support a Type II downrange reconnaissance mission. These robots don’t necessarily have to enter confined spaces or traverse rubble piles, but they do need to be able to climb stairs or at least curbs and modest irregular terrain. They need GPS tracking with info overlaid onto a map. They would typically move quickly down range (at least 1 km) to assess the situation and deploy multiple sensors (chemical, biological, radiological, nuclear, and explosive) with telemetry.

As a result of this input, the focus of the first wave of test methods will be on capturing performance for robots in these three categories.

Several other constructive comments covered other aspects of robot capabilities and performance. This summary includes observations during the course of the event, as well as those that were noted during the hot wash. The responders identified the following improvements for current implementations:

#### Sensors

- Thermal/infrared capabilities, to help locate victims as well as to identify fires and hot spots. This is particularly critical when there is smoke.
- Onboard mapping of environments when navigating through smoke.
- Better navigation aids, such as global positioning system (GPS) with the ability to show the robot coordinates and direction of view.
- Better placement of cameras, so they provide better depth perception. Responders sometimes view the same location from two different camera perspectives in an attempt to gain depth perception. Cameras should view the robot’s own tracks or wheels to help with situational awareness.
- Better far field visual acuity, up to 305 m (1000 feet), to help with planning.

#### Mobility

- Better mobility over loose debris. Random stepfields provide reasonable abstracted rubble, but should be looser to allow displacing individual steps. Wires and strings should be added to snag tracks.
- Continuous driving after throwing a track, especially if throwing tracks is a periodic.
- Minimum speed of 6.4 km (4 miles) per hour.

#### Communications

- Better radio communications, should allow choices of frequencies if one becomes problematic.
- Indication of radio communications signal strength and/or bandwidth – maybe even automatically detected to change frequencies and improve signal quality.
- Longer radio communications ranges both in-sight and beyond line-of-sight.
- Tethered communications presented a clear signal for long range and beyond line-of-sight problem, but the tether implementations introduced mobility complications and the additional workload for the operator to remember the tether, not run over it, and reel it in or spool it out.

#### Human/System Interaction

- Easier operator interfaces. Some are too complex (too many modes), while others were easier to learn. There was a lot of variation in the “usability” of the controllers for the robots.
- Better feedback on the robot state, such as arm position and runtime remaining, etc. The amount of information available to the operator at the control station also varied.
- Better operator control unit (OCU) displays for daylight conditions. Responders resorted to draping

their jackets over their heads and the OCU at times.

- Better audio feedback to the operator, to listen to the robot's actions more than to search the environment. Directional audio (stereo) and headphones were very helpful.
- The usability test method should be modified to separate out camera manipulation (which was part of the procedure in the version piloted at this exercise).

#### Manipulation

- Independent base rotation joints for manipulators, to remove reliance on mobility (tracks or wheels) to rotate the manipulator. Especially helpful when the robot is on uneven terrain.
- Test methods for opening doors, which are important tasks for conducting searches.

#### Logistics

- Easier track replacement in the field, especially if the tracks get thrown periodically.
- Easier wash-down and decontamination when necessary. Many of the robot designs would pool fluids in body features. Smoother designs would allow fluids to run off.

#### Concepts of Operation

Several responder teams paired up different robots in their scenarios. They used them collaboratively in the following manners:

- A larger robot carried a smaller robot to a particular location and released it. The smaller robot then conducted the search in more confined spaces.
- Multiple robots were positioned to provide multiple views of a location, or one robot's cameras observed a second robot as it moved for better remote situational awareness.
- One robot assisted another robot, either by opening a door or removing debris from the robot's tracks.

Additional types of quantitative data collected were aimed at future environment characterization and performance measurement infrastructure. Very high-resolution point clouds of some of the scenarios were produced. [4]. Three-dimensional tracking of robots, responders, and canines was conducted to evaluate the relevance of the tracking and ascertain under which circumstances it is viable.

NIST researchers produced laser scan data of three of the training scenarios at Disaster City. These scenarios included the rubble pile, the wood pile, and the passenger train derailment. 3D image data sets were collected using two commercially available laser scanners over the course of three days. Each scenario was scanned from multiple locations and each scan location was registered using targets placed in the environment. The data collected will be used to help visualize the environments within which the robots are expected to operate and characterize them.

NIST researchers have been working with an asset tracking

system to capture continuous location data for robots, personnel, and/or dogs operating within training scenarios. This tracking system requires equipping the perimeter of the scenario with antennas in carefully measured locations to receive signals and triangulate the position of active radio tags affixed to moving assets within the scenario. It also requires one or more reference tags placed at known locations within the scenario for calibration. NIST uses this tracking system to capture quantitative performance data (2D or 3D positions over time) during training operations to compare particular technologies, approaches, and/or methods of deployment. This kind of quantitative data capture enables performance metrics such as: deviations from intended paths, dwell locations and durations, percent of area searched, completeness of collaborative searches, etc. Robots could be tracked as they performed initial reconnaissance on the street surrounding the wood pile, although the robots could not see the interior from the street because the highest points of the pile are around the perimeter. As the robots entered the wood pile through buried concrete culverts, the tracking data disappeared, probably due to the robot's inability to exit the culvert into to the complex interior of the pile. Similarly, all attempts to track robots within the interior of the pile failed, probably due to signal attenuation from the very low robot positions within the pile, which required a straight-line path through the densely packed wood pile perimeter to reach the ground level receivers outside. Elevated antenna positions would certainly have helped, but we were unable to fit it into the schedule. Responders and dogs were tracked more successfully as they maneuvered in and around the wood pile. Radio tags placed on the canine's collar and the responder's helmet were able to better communicate with the receivers due to being at a more elevated position than on the robots.

## VI. SUMMARY

The exercise held at Disaster City helped advance understanding of the performance issues relevant to application of robots to urban search and rescue missions. Working closely with subject matter experts within relevant training scenarios, NIST and other organizations were able to further develop and refine performance requirements and test methods for US&R robots. The responders gained insights pertaining to the current status of robotic technology as well as the future potential. They determined the three initial robot categories for Wave 1 standard test methods and deployment. The categories address small, throwable "Ground Peek Robots"; "Wide-Area Survey Robots"; and "High Altitude Loiter Robots" (which may require a change in description since effective altitudes were demonstrated at Disaster City to be 91 m – 300 feet – above ground level). The responders were able to begin developing new concepts of operation, which will be essential once Task Forces and other response organizations begin to integrate robots into their deployments. The manufacturers were able to gain firsthand knowledge of the expectations that responders have for robots used in search

and rescue missions. They received direct feedback from the responders on their systems, and better correlated the stated performance requirements with the expected environments. And the various working groups responsible for developing test methods under the ASTM US&R robot standards task group collected data. The test methods and artifacts will be refined as a result of the lessons learned.

As planned, the proposed test methods for Wave 1 will be submitted into the balloting process this year. The stakeholders will have an opportunity to review the test methods in their near-final incarnations in August 2006, when the next response robot evaluation exercise will be held in Maryland. At that event, responders will have an opportunity to begin experimenting with sensor payloads on the robots. Per their feedback at Disaster City, NIST is working with manufacturers, sensor standards groups, and others to devise scenarios wherein sensors for detecting chemicals, radiological materials, and possibly biological and other hazards are mounted on robots.

#### ACKNOWLEDGEMENT

The authors would like to thank the Department of Homeland Security Science and Technology Directorate for funding this work. We are indebted to the members of the FEMA Task Forces that provide such valuable insights and to the robot manufacturers and developers who voluntarily come to these events. We also would like to acknowledge the Texas A&M University and Texas Task Force 1 for outstanding support during the Disaster City event. Thanks to the rest of the NIST team, who work tirelessly to make these events possible.

#### REFERENCES

- [1][http://www.isd.mel.nist.gov/US&R\\_Robot\\_Standards/disaster\\_city/Workshop5\\_April2006.pdf](http://www.isd.mel.nist.gov/US&R_Robot_Standards/disaster_city/Workshop5_April2006.pdf)
- [2] E. Messina, A. Jacoff, et al., Statement of Requirements for Urban Search and Rescue Robot Performance Standards (Preliminary Version), May 2005. Available at:  
[http://www.isd.mel.nist.gov/US&R\\_Robot\\_Standards/Requirements\\_Report\(prelim\).pdf](http://www.isd.mel.nist.gov/US&R_Robot_Standards/Requirements_Report(prelim).pdf).
- [3] Messina, E. and Jacoff, A., "Performance Standards for Urban Search and Rescue Robots," Proceedings of the 2006 SPIE Defense and Security Symposium, Orlando, FL, April 2006.
- [4] A. Lytle et al., "US&R Data Characterization Sets," Proceedings of the 2006 IEEE Safety, Security, and Rescue Robotics.