Proceedings of the SPIE Defense and Security Symposium, Orlando, FL, April 17-21, 2006.

Performance Standards for Urban Search and Rescue Robots

Elena Messina and Adam Jacoff
Intelligent Systems Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8230

ABSTRACT

In this paper, we describe work in performance standards for urban search and rescue (USAR) robots begun in 2004 by the Department of Homeland Security. This program is being coordinated by the National Institute of Standards and Technology and will result in consensus standards developed through ASTM International, under the Operational Equipment Subcommittee of their Homeland Security Committee. The first phase of the program involved definition of requirements by subject matter experts. Responders participated in a series of workshops to identify deployment categories for robots, performance categories, and ranges of acceptable or target performance in the various categories. Over one hundred individual requirements were identified, within main categories such as Human-System Interaction, Logistics, Operating Environment, and System (which includes Chassis, Communications, Mobility, Payload, Power, and Sensing). To ensure that the robot developers and eventual end users work closely together, "responders meet robots" events at situationally relevant sites are being held to refine and extend the performance requirements and develop standard test methods. The results of these standard performance tests will be captured in a compendium of existing and developmental robots with classifications and descriptors to differentiate particular robotic capabilities. This, along with ongoing efforts to categorize situational USAR constraints such as building collapse types or the presence of hazardous materials, will help responders match particular robotic capabilities to response needs. In general, these efforts will enable responders to effectively use robotic tools to enhance their effectiveness while reducing risk to personnel during disasters.

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

1. INTRODUCTION

Urban Search and Rescue (USAR) is defined as "the strategy, tactics, and operations for locating, providing medical treatment, and extrication of entrapped victims."[1] USAR teams exist at local and state levels. On the federal level, the Federal Emergency Management Agency (FEMA) has 28 Task Forces that structure local emergency responders into integrated disaster response forces. Well-known examples of events in which FEMA USAR Task Forces were deployed include the collapse of the World Trade Center towers and Hurricanes Katrina and Rita.

Understandably, the work of a responder is extremely arduous and dangerous. USAR teams are adopting various technologies to help them better do their jobs. For instance, they use cameras at the end of a flexible pole and seismic sensors to help them locate victims. There is much potential for effective, robust, and cost-effective technologies to be integrated into the search processes. In studies that focused on technology needs, robots have been specifically identified as holding promise [2,3]. Blitch noted that USAR is a domain "that is a very dangerous job for human rescuers, poses an almost infinitely difficult spectrum of challenges, and yet provides an opportunity for robots to play a pivotal support role in helping to save lives. [4] However, at this time, the state of robot technology overall is not very mature and typically purchase and maintenance costs are prohibitive for most responder organizations.

Any new candidate technological solution must be proven useful to the responder community prior to deployment in the field. Standardized test methods generated directly from responder requirements can ensure that applicable technologies are relatively easy to use, integrate efficiently into existing infrastructure, and provide demonstrable utility to response operations. Being able to characterize the performance of a new technology under specified – yet representative – conditions, will also enable funding agencies, such as FEMA, to obtain best value in their procurements. Another benefit of having standard performance evaluations is the acceleration of the needed technology developments.

To address these salient needs, the Department of Homeland Security Science and Technology Directorate initiated an effort in 2004 with NIST to develop comprehensive standards to support development, testing, and certification of effective robotic technologies for USAR applications. These standards will address robot mobility, sensing, navigation, planning, integration into operational caches, and human system interaction. Such standards will allow DHS to provide guidance to local, state, and federal homeland security organizations regarding the purchase, deployment, and use of robotic systems for USAR applications.

Standard test methods generated from explicit requirements for USAR robots, with objective performance metrics and repeatable performance testing, will accelerate the development and deployment of mobile robotic tools for USAR responders. Currently, no such standards or performance metrics exist, although some guidelines for performance, capabilities, and human-system interactions have been identified. [5, 6]

This paper is organized as follows. Section 2 describes briefly the process model employed in this program. In Section 3, we discuss the capture of performance requirements. Test method development, including a discussion of the general guiding principles, is described in Section 4. The necessary step of validating the test methods and gauging the relevance of robots to particular deployment scenarios is then discussed in Section 5. Section 6 introduces the concept of validation exercises. Related work is covered in Section 7. We conclude with a brief discussion on future directions.

2. OVERVIEW AND TECHNICAL APPROACH

The entire program is structured to ensure that the end users' needs are captured and addressed. The plan includes annual workshops to monitor progress as well as several exercises that allow responders to work with emerging robotic equipment in realistic environments while helping to refine proposed test methods. The requirements defined by the responders are the foundation for constructing robot performance measures along with testing and evaluation (T&E) protocols that provide reproducible methods for assessing and comparing the effectiveness of overall robotic systems and key components. Test sites will be built that realistically evaluate these robot's capabilities. Supporting measurement infrastructure to facilitate characterization of the test sites and to capture robot performance during test administration will also be developed. Ultimately, the goal is to have one or more sites certified to perform this program's standard test methods and provide ongoing robot performance testing. Initially, each site may focus on specific aspects of the overall robotic systems (mobility for example), to avoid issues of conformity between test sites.

At the outset of the program, supporting work efforts were established to address other prerequisites for the successful introduction of robots into the urban search and rescue application domain. In recognition that these novel tools need to be integrated into existing responder operations, new standard operating procedures will be developed, along with corresponding training and deployment plans. An understanding of where and how to deploy robots within various response situations is needed in order to develop standard operating procedures. To this end, characterizations of both the deployment environments and the robots themselves must be developed.

Researchers at NIST's Building and Fire Research Laboratory are developing a structural collapse taxonomy. A taxonomy is " a scheme that partitions a body of knowledge and defines the relationships among the pieces. It is used for classifying and understanding the body of knowledge."[7] Existing sources of building classifications and how they relate to collapse are being studied.[8] Currently, for responders' purposes (victim identification and recovery), the use of the structure and time of day (e.g. school / night) are deemed more important data than specific construction types. Physical characterization of rubble is also being attempted. Experiments using laser scanning and range image analysis of rubble piles at training facilities are underway. Data are collected on both the exterior and interior of piles using very high resolution three-dimensional scanners (Figure 1).

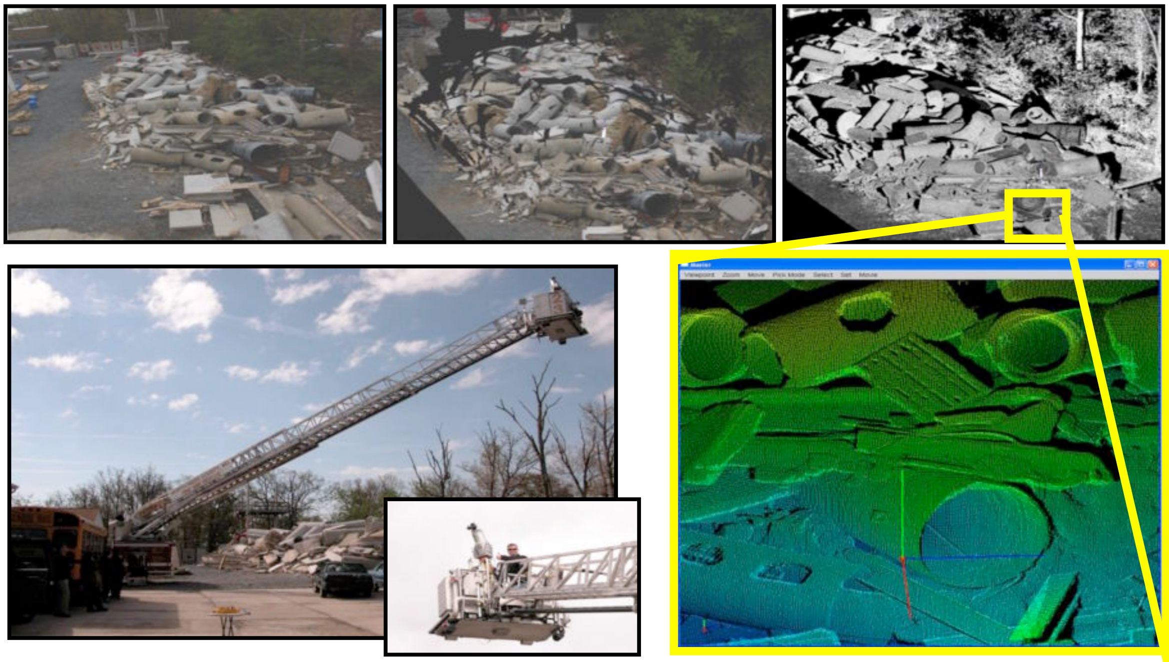


Figure 1: Data Collection Process for Collapse Taxonomy Rubble Characterization. This set of images shows the overhead views of the rubble pile, the data capture rig, and a resulting point cloud captured by the three-dimensional scanner.

The characterization of the robots is being captured through the use of an ontology.[9] The goal of this Robot Ontology effort is to develop and to begin to populate a neutral knowledge representation (the data structures) capturing relevant information about robots and their capabilities. This knowledge representation must be flexible enough to adapt as the robot requirements evolve. An ontology provides the necessary flexibility and expandability. In this context, an ontology can be thought of as a knowledge representation approach that represents key concepts, their properties, their relationships, and their rules and constraints. Whereas taxonomies usually provide only a set of vocabulary and a single type of relationship between terms (usually a parent/child type of relationship), an ontology provides a much richer set of relationship and allows for constraints and rules to govern those relationships. In general, ontologies make all pertinent knowledge about a domain explicit and are represented in a computer-interpretable fashion that allows software to reason over that knowledge to infer addition information.

An initial structure for the Robot Ontology has been developed. This initial structure can be broken down into the follow primary categories of knowledge:

- Structural Characteristics – describes the physical and structural aspects of a robot
- Functional Capabilities – describes the behavioral features of the robot
- Operational Considerations – describes the interactions of the robot with the human and the interoperability with other robots

Examples of knowledge captured in the structural characteristics category include (but are not limited to) size, weight, tethering, power source, locomotion mechanism (e.g. wheeled, walking, crawling, flying, jumping), and sensors (e.g. color camera, range imaging, Sonar, Global Positioning Satellites, Audio, Thermal).

Examples of knowledge captured in the functional capabilities category include (but are not limited to):

- Locomotion Capabilities (e.g., maximum speed, maximum step climbing, maximum slope climbing, etc.)
- Sensory Capabilities (e.g., minimum visibility level, map building capability, self-localization, system health, etc.)
- Operational Capabilities (e.g., working time, setup time, max. force available to push, Mean Time Between Failure.)
- Environmental Resistance (e.g., maximum operating temp, maximum submergibility level, etc.)

- Degree of Autonomy (e.g., joint level dependency, drive level dependency, navigation level dependency, etc.)
- Rubble Compatibility (e.g., ability to historically operate well in certain terrains)
- Communications (e.g., communication media, communication channel frequency, content standards, information content, communication locking, communication encryption, etc.)

Robot capabilities can be described using the ontology as a framework. The ontology can be used by computer programs that can, for instance, incorporate the building collapse taxonomy, and be used as decision aids for determining which robot(s) are best suited to a particular deployment and how they ought to be configured. The robot capabilities descriptions are also meant to support the definition of the performance test methods. A human-readable expression of the ontology that captures the results of the standard tests is also being developed. This compendium will detail all the information about a particular robot, including its performance results, and in conjunction with the usage guides, can assist in making robot procurement decisions.

3. THE STANDARDS DEVELOPMENT PROCESS

The Department of Homeland Security specified that the standards body within which performance standards for USAR robots must be accredited by the American National Standards Institute (ANSI). Accreditation by ANSI signifies that the procedures used by the standards body in connection with the development of American National Standards meet the Institute's essential requirements for openness, balance, consensus and due process. These principles mean that

- the standards development group must be open to representatives from all materially affected and interested parties
- there is public review and comment on draft standards
- all comments submitted by voting members and public must be considered and addressed in some fashion
- submitted changes that meet the consensus requirements must be incorporated into a draft standard
- there is an appeals process for those who feel that these principles were not respected during the standards-development process.

Based on a review of interested standards development organizations (SDOs), NIST and DHS selected ASTM International to host the performance standards for urban search and rescue robots. ASTM International is ANSI accredited. The USAR robot standards are being developed under the jurisdiction of Subcommittee E54.08 on Operational Equipment, within the E54 Committee on Homeland Security Applications. Since robots involve so many different subsystems and technologies, leverage of existing standards will be emphasized. Referencing appropriate standards, even if they are not of ASTM International origin, is encouraged.

The standards development process is not necessarily a linear one. In many cases, it is essential to have an iterative approach, wherein the inputs from all the stakeholders are considered throughout the entire timeline. Draft standards should be subject to review and modification by the stakeholders prior to a committee vote that formalizes the draft into a published standard. The iterative review process of the requirements, metrics, and testing methods is designed to produce measurements of capabilities that are pre-requisites to fieldable robots.

In the case of an emerging, multi-disciplinary system of systems such as a robot, it is not possible to define all of the necessary performance requirements at once nor is it desirable to constrain technology development by being too restrictive in the definition of performance standards. A robot applied to urban search and rescue applications is comprised of many components, each of which requires in-depth scrutiny of the performance needs. The components have to integrate amongst themselves; these interactions may create further performance requirements. Furthermore, robots are not tools currently used by responders. The best modes of employing these tools is not currently understood. Methods of usage and expected deployment conditions, of course, affect the definition of requirements for performance.

USAR is a multi-faceted application domain. There are multiple stages during a USAR operation and teams perform a variety of functions. Examples of functions that a FEMA USAR team can perform include: conducting physical search and rescue in collapsed buildings, providing emergency medical assessments and care to trapped victims, assessment and control of hazards, such as gas or electric service, evaluation and stabilization of damaged structures. Robots could potentially support rescue personnel in carrying out all of these functions. To further complicate the picture, the types of disasters that responders encounter are varied as well. Consider, for example three well-known disasters: the collapsed World Trade Center towers, submerged post-Katrina New Orleans, and the post-bombing Murrah building in Oklahoma City. Certainly, it would be nearly impossible, given the state of current technology, to expect a single robot design to be able to respond to all three situations. In the case of the World Trade Center collapse, the rubble pile was massive and there were extremely tight (confined) spaces that required a very small robot. Due to the fires, a robot responding to this event would have to be highly heat-resistant as well. In a New Orleans response, the robot would have to be completely submersible and deal with thick mud. A whole façade of the Murrah building was removed by the explosion. In this response, a robot that could hover next to the different floors and look for victims may have been helpful. This very simplified look at different disaster scenarios demonstrates that a "standard USAR robot" is not a practical goal. In fact, as is described in Section 4, at least thirteen different deployment categories were defined for robots by responders.

Just like there are many disciplines required within a search and rescue team, the components within a robot are also quite diverse. The disciplines involved in the various components that comprise robots are specialized and different enough that a different set of expertise is required to adequately study the requirements and develop the corresponding performance tests.

Challenges such as those briefly listed above cannot, however, impede progress towards the goal of having well-understood performance goals and means of measuring whether systems meet them The approach taken in developing performance standards for USAR robots is to break the problem down into logical, cohesive, manageable categories, and for each of these categories, produce standard test methods that are accompanied by usage guides. The test methods are a set of means by which to objectively measure a robot's performance in a particular area. The usage guides provide suggested performance ranges (test results) desired for different application scenarios.

4. REQUIREMENTS CAPTURE

Although the potential for utilizing robots to assist rescuers in USAR operations was recognized prior to this program's inception, a methodical capture of responders' view of how they would use robots and what the detailed performance requirements were for robots had not occurred previously. NIST worked closely with DHS Science and Technology and FEMA to initiate a series of workshops that defined the initial set of performance requirements for robots applied to USAR. The first three workshops deliberately did not include robot technologists and vendors, so as to not initially bias the input from the end users with knowledge of existing technologies or approaches. Once a substantial body of requirements was gathered from responders, in subsequent workshops, robot technology providers (researchers, vendors, other government programs) were encouraged to participate.

The requirements definition process during the initial set of workshops was comprised of identifying and describing individual requirements, defining how a robot's performance with respect to a given requirement is to be measured, and, where possible, specifying the objective (desired) and threshold (minimum or maximum) performance values. The resulting list of requirements totaled over 100. These were grouped into several broad major categories. One major category, System, was further decomposed into sub-categories. These are shown in Table 1. Examples of individual requirements are shown in Figure 2. A draft report detailing the process, the initial set of requirements, and the robot deployment categories is found at the NIST web site.[10]

Table 1: Major Performance Categories for USAR Robots

Human-System Interaction	Pertaining to the human interaction and operator(s) control of the robot
Logistics	Related to the overall deployment procedures and constraints in place for disaster response
Operating Environment	Surroundings and conditions in which the operator and robot will have to

	operate
Safety	Pertaining to the safety of humans and potentially property in the vicinity of the robots
System:	Overall physical unit comprising the robot. This consists of the sub-components below:
- Chassis	The main body of the robot, upon which additional components and capabilities may be added. This is the minimum set of capabilities (base platform).
- Communications	Pertaining to the support for transmission of information to and from the robot, including commands for motion or control of payload, sensors, or other components, as well as underlying support for transmission of sensor and other data streams back to operator
- Mobility	The ability of the robot to negotiate and move around the environment
- Payload	Any additional hardware that the robot carries and may either deploy or utilize in the course of the mission
- Power	Energy source(s) for the chassis and all other components on board the robot
- Sensing	Hardware and supporting software which sense the environment

Type:	**CHASSIS**
Sub-Type:	**ILLUMINATION**
Requirement:	**ADJUSTABLE**
Metric:	**YES/NO**
Description:	This requirement captures the responders' expectation to use video in confined spaces and for short-range object identification, which can wash out from excessive illumination of the scene.

Type:	**COMMUNICATIONS**
Sub-Type:	**N/A**
Requirement:	**RANGE – BEYOND LINE OF SIGHT**
Metric:	**METERS**
Description:	This requirement captures the responders' expectation to project remote situational awareness into compromised or collapsed structures or to convey other types of information. They specifically noted that the robot should be able to ingress a specified number of meters into the worst case collapse, which was further defined as a reinforced steel structure. This requirement also covers operations around corners of buildings and other locations beyond line of sight. The responders made no distinction regarding tethered or wireless implementations to address this requirement.

Type:	**HUMAN-SYSTEM INTERACTION**
Sub-Type:	**CONTEXT**
Requirement:	**PROTECTIVE CLOTHING**
Metric:	**SCALE 1-5** 1 = No protection 3 = Minimum protection (threshold) 5 = Complete protection (objective)
Description:	This requirement captures the responders' expectation to be operating the system while wearing personal protective equipment such as gloves, helmet, eye protection, ear protection, etc. The operator should be able to maintain acceptable usability (discussed in greater detail in the Test Methods: Human-System Interaction section of this report) of the system while wearing the stated level of personal protective equipment

Figure 2: Example of Performance Requirements. A category (type/sub-type) is identified, along with a metric to be used in measuring the robot's performance, which can be binary, length, time, or a scale defined by the responders.

Table 2: Robot Deployment Categories

Robot Category	Employment Role(s)	Deployment Method(s)	Tradeoffs
Ground: Peek Robots	Provide rapid audio visual situational awareness; provide rapid HAZMAT detection; data logging for subsequent team work	Tossed, chucked, thrown pneumatically, w/surgical tubing; marsupially deployed	Trade mobility, duration, sensing for increased expendability
Ground: Collapsed Structure-- Stair/Floor climbing, map, spray, breach Robots	Stairway & upper floor situational awareness; mitigation activities; stay behind monitoring	Backpacked; self driven; marsupially deployed	Experience form factor for increased mobility, sensing, manipulation; mapping variant; spraying variant; breaching variant
Ground: Non-collapsed Structure-- Wide area Survey Robot	Long range, human access stairway & upper floor situational awareness; contaminated area survey; site assessment; victim identification; mitigation activities; stay behind monitoring	Backpacked; self driven; marsupially deployed	Experience form factor for increased mobility, sensing, manipulation; mapping variant; spraying variant; breaching variant
Ground: Wall Climbing Deliver Robots	Deliver Payloads to upper floors; provide expanded situational awareness when aerial platforms are unavailable or untenable	Placed; thrown pneumatically, w/surgical tubing; marsupially deployed	Trade payload capacity for vertical mobility and stable perching
Ground: Confined Space, Temporary Shore Robots	Adaptive, temporary shoring; provide stay behind monitoring; victim triage & support	Placed: lowered via tether	Trade mobility and payload capacity for shoring capacity
Ground: Confined Space Shape Shifters	Search; provide stay behind monitoring	Placed; lowered via tether	Trade payload capacity for confined space access
Ground: Confined Space Retrieval Robots	Retrieve objects from confined spaces; provide stay behind monitoring	Placed; lowered via tether	Trade sensing capacity for manipulators, confined space access
Aerial:High Altitude Loiter Robots	Provide overhead perspective & sit. awareness; provide HAZMAT plume detection; provide communications repeater coverage	Released: balloon or fixed wing; tethered LTAF (kite)	Trade penetration capacity for vertical perspective
Aerial: Rooftop Payload Drop Robots	Payload delivery to rooftops; provide overhead perspective; provide communications repeater coverage	Launched F/W; tethered LTAF (kite)	Trade penetration capacity & loiter time for vertical drop
Aerial: Ledge Access Robot	Object retrieval from upper floors; crowd control with a loudspeaker object attached, provide situational awareness	Launched Vertical Take-off and Landing (VTOL); VTOL	Trade simplicity, penetration capacity, loiter time for precise vertical drop
Aquatic: Variable Depth Sub Robot	Structural inspection; leak localization/mitigation; object (body) recovery	Dropped into water; lowered via tether	Trade ground mobility for sub surface access & free swim capacity
Aquatic: Bottom Crawler Robot	Water traverse; rapid current station keeping; object recovery	Driven across water; lowered via tether	Pursue amphibious mobility at cost of other performance
Aquatic: Swift Water Surface Swimmer	Upstream access and station keeping; payload delivery; object recovery	Dropped into water; marsupially deployed	Pursue swift water capacity at cost of other performance

As noted above, there is no typical USAR scenario. During the initial three requirements definition workshops, potential robot deployment categories (which could correspond to different disaster types or aspects of a response) were enumerated. Thirteen categories were defined, which detailed the capabilities that the robot should have, along with the deployment method, and tradeoffs. Ground, aerial, and aquatic robot deployments are represented. The deployment categories are listed in Table 2.

5. TEST METHOD DEVELOPMENT

The principal output of this program will be standard tests methods and metrics for the various performance requirements and characteristics defined by the responders. The test methods should be objective and clearly defined. The standard tests will be hosted by one or more certified sites. Ideally, the test methods will also be reproducible by robot developers and manufacturers to provide tangible goals for system capabilities. This will enable robot and component developers to exercise their systems in their own locations in order to attain the required performance.

A test method will be developed for each of the performance requirements generated. The draft test methods are being designed by NIST, in close consultation with responders, developers, and technical experts. In this section, we provide some details on possible performance tests. As can be seen from the first example, test methods can be designed and executed so as to evaluate multiple performance requirements. The test methods shown in this paper are examples; the standard test methods produced through ASTM International may be different.

The visual acuity example test method (Figure 3) and associated results reporting sheet (Figure 4) show one way to test the performance of the robot's vision system. The method outlined is a timed test to read standard eye charts in sequence from a variety of distances. Just as human vision is tested per our ability to read different sized letters from a pre-determined distance, this test measures the minimum-sized feature that a robot operator can perceive through the robot's system. This includes the on-board camera(s) that are capturing the images of the eye chart, the transmittal of the images from the robot to the operator station, and finally the display of the images on a screen viewed by the operator. Other performance characteristics can be measured through variations of this test. The distance to the eye charts can measure both near and far field perception. Since charts are also placed to the side of the robot, the repositioning of the sensors and the time it takes the operator are also measured. Several initial robot orientations can be prescribed so as to differentiate articulated pan/tilt systems from fixed cameras and the operator's ability to remotely direct and re-direct the sensors. Ambient lighting can be varied to test the robot's onboard illumination capabilities or low light cameras. The test could also be conducted using both wired and wireless communication modes (radio noise can devastate image resolution). The goal of this test method is to isolate easy-to-measure metrics: time to perform all readings and average acuity across a variety of relevant situations. This allows direct comparison of performance capabilities without necessarily stating what level of performance is acceptable for a given implementation, user, or role.

A second example of potential test methods is from the Logistics requirements. FEMA Task Forces need to know what the impact of adding robots to their cache will be. The volume and weight are key metrics. The Task Forces use specific containers to carry their equipment and supplies to a response. Therefore, they would like to know the transport requirements for the robots in terms of the containers with which they are already familiar. Figure 5a shows the requirement as defined by the responders. A straightforward test method is to determine the minimum size container(s) from the listed ones necessary to hold the robot and all necessary accessories. Figure 5b shows a robot next to a Hardigg case. Not shown are peripheral devices necessary for the operation of the robot. These would also have to be included in the evaluation of what is the minimum packaging requirement for shipping the robot system.

6. VALIDATION EXERCISES

This program employs an iterative development approach to ensure that the performance requirements are appropriate and that the vendor and technology communities are able to interact with the end users on a frequent basis. Regular "responders meet robot" events at representative USAR training sites also present opportunities to dry-run testing protocols to an audience of responders and technologists. Comments from these communities help refine and strengthen the tests. The events also serve to provide feedback on a frequent basis to the technology developers, who are able to see how their systems perform informally against the emerging performance standards.

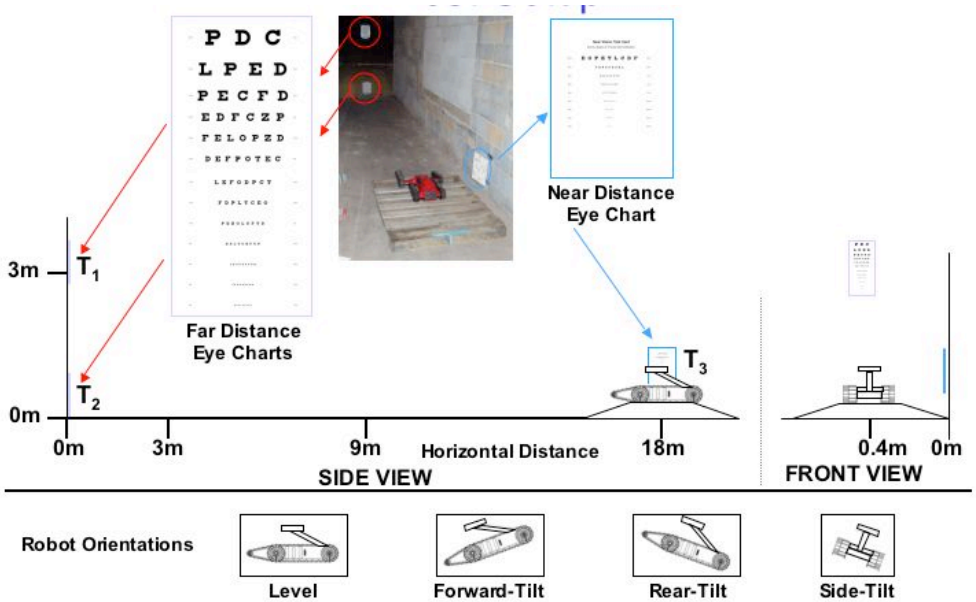
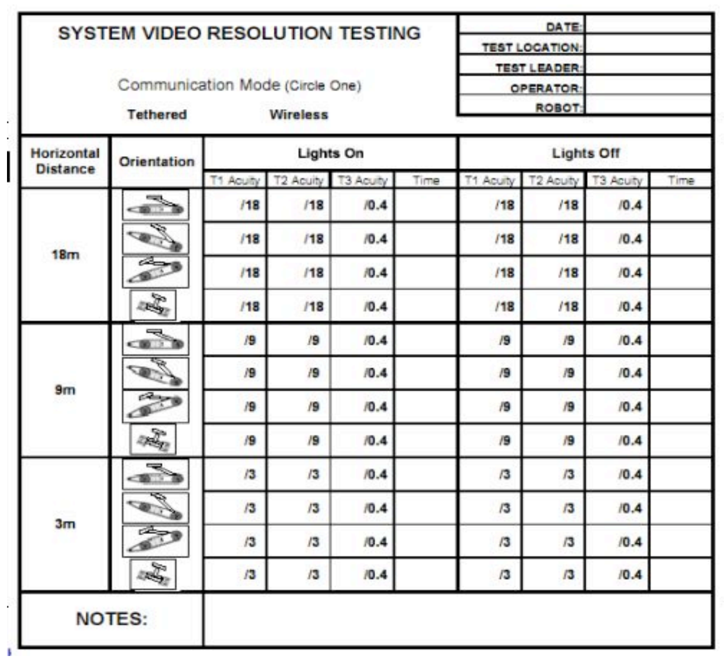


Figure 3: Example Set Up for Visual Acuity Test Method. Eye charts are placed at defined distances and robot is positioned per defined orientations (level and non-level). Test measures smallest resolution text or symbols that operator can read at the operator station. Time to position cameras and read line is also measured. Additional test methods can be based on this setup, including changing ambient lighting and degrading communications between robot and operator station.



SYSTEM VIDEO RESOLUTION TESTING

Communication Mode (Circle One)

Tethered Wireless

DATE:
TEST LOCATION:
TEST LEADER:
OPERATOR:
ROBOT:

Horizontal Distance	Orientation	Lights On				Lights Off			
		T1 Acuity	T2 Acuity	T3 Acuity	Time	T1 Acuity	T2 Acuity	T3 Acuity	Time
18m		/18	/18	/0.4		/18	/18	/0.4	
		/18	/18	/0.4		/18	/18	/0.4	
		/18	/18	/0.4		/18	/18	/0.4	
		/18	/18	/0.4		/18	/18	/0.4	
9m		/9	/9	/0.4		/9	/9	/0.4	
		/9	/9	/0.4		/9	/9	/0.4	
		/9	/9	/0.4		/9	/9	/0.4	
		/9	/9	/0.4		/9	/9	/0.4	
3m		/3	/3	/0.4		/3	/3	/0.4	
		/3	/3	/0.4		/3	/3	/0.4	
		/3	/3	/0.4		/3	/3	/0.4	
		/3	/3	/0.4		/3	/3	/0.4	
NOTES:									

Figure 4: Example Test Results Reporting Sheet for Visual Acuity Test

Type:	**LOGISTICS**
Sub-Type:	**CACHE PACKAGING**
Requirement:	**VOLUME PER CONTAINER**
Metric:	**SCALE 1-5** 1 = Pelican 1650 box 3 = Hardigg box checkable on commercial aircraft 5 = Ropack model 4048, 4039 with drop door
Description:	This requirement captures the responders' expectation to move and store all equipment using existing methods and tools.

(a)

(b)

Figure 5: Example Logistics Test Method. (a) shows the cache packaging requirement as defined by responders. (b) shows a robot next to one of the standard containers used by FEMA Task Forces (the Hardigg).

The first of these exercises was held at the FEMA Nevada Task Force 1 training site in August 2005. A diverse set of almost twenty robots were run by responders through training scenarios or proposed test methods. The most important result of this exercise was the direct communication between the end users and the robot developers *in situ*. Events similar in flavor have been held by other organizations, such as the Center for Robot-Assisted Search and Rescue and the National Aeronautics and Space Administration's Disaster Assistance and Rescue Team (NASA DART), but have not been specifically aimed at development of performance standards.

The somewhat informal tests in the early part of the program will be expanded into technology readiness level evaluation exercises. As the technologies comprising the USAR robotic systems evolve, it will be possible to conduct more rigorous tests under realistic scenario-driven exercises involving responders. These tests will expose any gaps in the existing performance requirements and testing protocols and help refine existing requirements and tests prior to submitting them to the standards process.

Ultimately, the test exercises will perform technology readiness level assessments (TRLs) as the technologies mature. TRLs are a systematic metric and associated measurement system that supports both the assessment of the maturity of a particular technology and the consistent comparison of maturity between different types of technology. First developed by NASA[11], this 9-level scale defines broad categories of deployment readiness of technologies, components, subsystems, or systems. NIST has directed TRL evaluations for the Army Research Laboratory in which the technical maturity of the Autonomous Navigation System developed under the Demo III Program was assessed.[12] Experience gained in this process will inform TRL assessments for USAR robots.

7. RELATED STANDARDS WORK

As robots are becoming more prevalent in the field, especially for military applications, other standardization efforts are underway.[13] As noted above, the standards for USAR robots will reference any relevant results from other standards development organizations. Two main efforts are addressing the need for interoperability of components for robots. The Society of Automotive Engineers (SAE) AS-4 committee and the North Atlantic Treaty Organization (NATO) Standardisation Agency are both developing messaging and communications standards. SAE AS-4 is providing the conduit for standards developed originally within the Joint Architecture for Unmanned Systems Working Group, which is supported by the Office of the Secretary of Defense. The NATO Standardisation Agency has produced a series of standards that include has published a series of standards, called STANAGs. The title of STANAG 4586 is "Standard Interfaces of Unmanned Aerial Vehicle Control System for NATO UAV Interoperability."

In terms of performance standards, the Department of Justice National Institute of Justice (NIJ) has initiated an effort to develop standards for bomb-disposal robots. Building on the extensive foundational work performed by the Technical Support Working Group and others.[14] NIST is working with NIJ to develop standard test methods that are aimed at bomb-disposal robots. Of course, where possible, there will be leverage of the USAR robot performance standards.

8. SUMMARY AND FUTURE WORK

We have described the Department of Homeland Security-funded effort to develop performance standards for robots applied to urban search and rescue. The approach taken by the National Institute of Standards and Technology, the organization leading the implementation of the program, is focused on user-specified requirements. The output of the process will be standards test methods published by ASTM International and supplemented by usage guides. Robots are constructed from multi-disciplinary and specialized components – electrical, mechanical, software – and the underlying technologies are still developing. The standardization effort is structured so as to not squelch innovation, but rather to provide direction to the technology developers. Supporting endeavors are aimed at furthering understanding of robot capabilities and deployment situations and at ensuring that robots can be integrated into the existing response infrastructure. They include development of a taxonomy of building collapses, an ontology of robot capabilities, new standard operating procedures, and training programs.

The formal approval of the first set of test methods through ASTM International is envisioned by 2006. This will consist of a subset of performance requirements that either apply broadly to most situations in which robots could be deployed and/or target capabilities that are deemed mature enough for defining operating thresholds for. Workshops and exercises will be held regularly. As the technologies comprising the USAR robotic systems evolve, it will be possible to conduct more rigorous tests under realistic scenario-driven exercises involving responders. These tests will expose any gaps in the existing performance requirements and testing protocols and help refine existing requirements and tests prior to submitting them to the standards process. The exercises will eventually evolve into Technology Readiness Level assessments.

In the coming years, one or more testing and evaluation sites will be selected and certified by NIST as being able to carry out the test methods developed within this program. The test site(s) will conduct any official testing and evaluation of robots for urban search and rescue. The results of the testing will be available through a compendium, which will be supplemented with additional vendor-provided information about each robot, in order to help make buying decisions. Ultimately, the goal is to have a variety of robots with known performance characteristics and an understanding of which robots are best suited to which types of response situations. When this goal is attained, the promise of robotic technology to aid responders in carrying out the USAR missions will hopefully be fulfilled.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Department of Homeland Security Science and Technology Directorate Standards Portfolio (Dr. Bert Coursey, Portfolio Director) for sponsoring this work.

DISCLAIMER

Any mention or display of commercial products is for information only; it does not imply recommendation or endorsement by NIST.

REFERENCES

1. FEMA 9356.1-PR "Urban Search and Rescue Response System In Federal Disaster Operations: Operations Manual" January 2000.
2. Urban Search and Rescue Technology Needs Identification of Needs, DHS/FEMA and National Institute of Justice Report, June 2004.

3. Project Responder: National Technology Plan for Emergency Response to Catastrophic Terrorism, National Memorial Institute for the Prevention of Terrorism and the United States Department of Homeland Security Report, April 2004.
4. Blitch, J.G. (1996), ``Artificial intelligence technologies for robot assisted urban search and rescue'', Expert Systems with Applications, Vol. 11 No. 2, pp. 109-24.
5. Drury, J., Hestand, D., Yanco, H., and Scholtz, J., Design Guidelines for Improved Human-Robot Interaction, Proceedings of CHI2004, Poster Presentation, Vienna, April 2004.
6. Center for Robot Assisted Search and Rescue, http://crasar.csee.usf.edu/rescuerobots/robots.htm, accessed May 6, 2005.
7. IEEE 1986 IEEE Standard Taxonomy for Software Engineering Standards, ANSI / IEEE SM 1002-1987, IEEE Standard Taxonomy for Software Engineering Standards.), 1986.
8. FEMA 310, Handbook for Seismic Evaluation of Buildings—A Prestandard (American Society of Civil Engineers, 1998).
9. Schlenoff, C. and Messina, E., "A Robot Ontology for Urban Search and Rescue," Workshop on Research in Knowledge Representation for Autonomous Systems, part of the Association for Computing Machinery Conference on Information and Knowledge Management," Bremen, Germany, November 2005.
10. Messina, E. et al., "Statement Of Requirements For Urban Search And Rescue Robot Performance Standards," Draft Report, NIST, May 2005 http://www.isd.mel.nist.gov/US&R_Robot_Standards/Requirements Report (prelim).pdf
11. Mankins, J., "Technology Readiness Levels: A White Paper," NASA Office of Space Access and Technology, April 6, 1995. http://www.hq.nasa.gov/office/codeq/trl/trl.pdf
12. Camden, R., Bornstein, J., French, F., Shoemaker, C., Bodt, B., Schipani, S., Runyon, T., Jacoff,, A., Lytle, A., "Autonomous Mobility Technology Assessment Final Report," Army Research Laboratory Technical Report (ARL-TR-3471), April 2005.
13. Huang, H., Messina, E., Albus, J., "Standards-Based Architectural Framework for Intelligent Autonomous Vehicles," in Intelligent Autonomous Vehicle Navigation, Madhavan, R., Messina, E., and Albus, J. (Eds), Nova Science Publishers, to appear in 2006.
14. National Institute of Justice Final Report on Law Enforcement Robot Technology Assessment, http://www.justnet.org/jpsg/robotassessment/robotassessment.html