Using Competitions to Advance the Development of Standard Test Methods for Response Robots

Adam Jacoff, Raymond Sheh & Ann-Marie Virts
National Institute of Standards and Technology (NIST)
100 Bureau Drive, Gaithersburg, MD 20899, USA
+1 301 975 [4235|5068|3533]
adam.jacoff@nist.gov, raymond.sheh@robolit.com, ann.virts@nist.gov

Tetsuya Kimura
Nagaoka University of Technology, Fujihashi, Nagaoka, Niigata, Japan
+81 258 47 9708
kimura@mech.nagaokaut.ac.jp

Johannes Pellenz
Bundeswehr Technical Center for Engineer and General Field Equipment, 56070 Koblenz, Germany, +49 261 876 7470
pellenz@uni-koblenz.de

Sören Schwertfeger
Jacobs University Bremen
28759 Bremen, Germany
+49 421 200 3155
s.schwertfeger@jacobs-university.de

Jackrit Suthakorn
Mahidol University, 25/25 Puttamonthon 4 Road, Salaya, Nakorn Pathom 73170, Thailand, +66 2 889 2138
egjst@mahidol.ac.th

ABSTRACT

Competitions are an effective aid to the development and dissemination of standard test methods, especially in rapidly developing, fields with a wide variety of requirements and capabilities such as Urban Search and Rescue robotics. By exposing the development process to highly developmental systems that push the boundaries of current capabilities, it is possible to gain an insight into how the test methods will respond to the robots of the future. The competition setting also allows for the rapid iterative refinement of the test methods and apparatuses in response to new developments.

For the research community, introducing the concepts behind the test methods at the research and development stage can also help to guide their work towards the operationally relevant requirements embodied by the test methods and apparatuses. This also aids in the dissemination of the test methods themselves as teams fabricate them in their own laboratories and re-use them in work outside the competition.

In this paper, we discuss how international competitions, and in particular the RoboCupRescue Robot League competition, have played a crucial role in the development of standard test metrics for response robots as part of the ASTM International Committee of Homeland Security Applications; Operational Equipment; Robots (E54.08.01). We will also discuss how the competition has helped to drive a vibrant robot developer community towards solutions that are relevant to first responders.

Keywords

autonomous robots, competitions, field robotics, manipulation, mobile robots, performance metrics, response robots

1. INTRODUCTION

The Intelligent Systems Division of the National Institute of Standards and Technology (NIST) has been making use of international competitions as part of its standards development process for response robots. These standards, part of ASTM International Committee of Homeland Security Applications; Operational Equipment; Robots (E54.08.01), are developed with the active input of robotics researchers, developers, and test administrators and are based on requirements formulated in consultation with first responders. They measure the performance of operationally relevant aspects of whole robot systems, as a deployable configuration.

Recent world events, ranging from natural disasters such as the 2011 Tohoku Earthquake through to various counterterrorism operations and military deployments in places like Iraq and Afghanistan, have thrown the spotlight onto the use of robots order to reduce the risk to humans. The ability to evaluate the different aspects of response robot performance in a relevant, objective manner is crucial to ensuring that first responders and other end users have an accurate understanding of the capabilities of current robots and are able to make an informed choice when procuring such robots.

In administering competitions, NIST gains a valuable understanding of the performance of upcoming best-in-class technologies and an opportunity to perform rapid iterative refinement on the test methods and apparatuses. The competitions also serve as a proving ground where new test methods and apparatuses may be conceived and refined in the presence of researchers and developers that have a deep understanding of new capabilities still at the research stage. In return, the development community gains an insight into the needs of first responders and an opportunity to finetune their approach and promote their work. As they replicate the test methods, they also help to disseminate their use

through the academic community.

Academic competitions are unique in their ability to gather a wide variety of research platforms in the one location, at the same time, to tackle the same problem. As a target for standard test metrics, response robots present a unique challenge since they must be tested end-to-end as complete robotic systems and evaluated in terms of their performance in a particular aspect of their task. Unlike other fields, it is not practical to derive this performance from the performance of the individual components. For example, the test methods for vision are affected not only by the quality of the camera but also the power supply, communications system, human-robot interface, and any directed perception mechanisms. Competitions inherently test such systems end-to-end and are ideal for gaining exposure to complete, highly innovative, experimental implementations.

In the rest of this paper, we will discuss the role of the RoboCupRescue Robot League (RoboCup RRL) competition, community, and associated events, in the development of standard test methods for response robots. In particular, we will focus on several examples of test methods that were conceived in, or refined at, the competition and have since become or are soon to become standards. For an overview of the competition itself, including performances from the most recent competition, the reader is invited to refer to [8]. More detailed information about the competition and arena are available from the rules outline and arena construction manual [6, 10].

2. THE ROBOCUPRESCUE ROBOT LEAGUE

The International RoboCup Competition is best known for its soccer playing robots, where the challenge is to build robots that are able to play soccer, according to World Cup rules, better than the winners of the 2050 human World Cup. However, since 2001, RoboCup has also played host to the RoboCup RRL, a NIST-administered event where the task is to develop robots to solve challenges from the field of Urban Search and Rescue robotics. The RoboCup RRL sees over 100 teams of undergraduate and graduate students and researchers from around the world compete in regional competitions. These culminate in between 15 and 25 international teams competing over a week of intense competition, development, evaluation, and collaboration. These competitions ''test the test methods'' in the presence of cutting edge, experimental technologies.

The competition takes place in an arena that represents a building in various stages of collapse. It consists of a variety of standard test method apparatuses for response robots and an example is shown in Figure 1. Many of the test methods and apparatuses that make up current ASTM standards started in these competitions and were tested and refined at these events. The goal of each team is to reach simulated victims, report their state, and build a map that would allow a human rescuer to reach them. The victims are strategically placed such that in order to reach them, robots must overcome the test methods. To add structure to the competition, the arenas are separated into three sub-arenas, denoted by the colors Yellow, Orange, and Red and representing challenges posed by Autonomy, Structured Obstacles, and Advanced Mobility Terrain.

The RoboCup RRL is unique in its emphasis not on finding a champion, but rather on building a community that works together to advance the state of the art in Urban Search and Rescue robotics: ''A League of Teams with one goal: to Develop and Demonstrate Advanced Robotic Capabilities for Emergency Responders.'' To this end, the competition is carefully administered to encourage the participation of a wide cross-section of the developer community that has something to contribute to this application domain, even if they do not have the resources for a championship team.

Specialized teams, which often produce highly sought-after Best-in-Class solutions to particular challenges in this domain, rarely have the broad based resources to compete well across the whole competition. As we will discuss, mechanisms that encourage teams to collaborate and special awards for Best-in-Class solutions, have resulted in many such specialized teams competing. From the perspective of the standards process, the participation of these teams is highly desirable as exposure to a wide variety of the best approaches to each test method have proven to be invaluable in their refinement. These specialized developers are often able to suggest, and in many cases demonstrate, many alternative approaches to the test method, and contribute to improving its ability to properly represent real world performance. In some cases these specialist developers have become part of the standards process. A particular advantage to incorporating these groups into the standard process is that, as academic institutions, they often have more freedom to experiment and develop innovative implementations to solve problems, without the commercial pressures and restrictions that affect commercially developed solutions. In the background, their exposure to the test methods encourages development towards the operationally relevant requirements embedded in the test methods and encourages the dissemination of the test methods through their adoption by the academic community.

The broader goals of promoting and advancing the state of the art are also well served by supporting such a vibrant community of robot developers as significant cross-pollination of ideas and capabilities occurs at such events. Many of the more general teams have begun to demonstrate capabilities that were formerly only demonstrated by highly specialized teams; conversely several highly specialized teams have begun to branch out into more general capabilities. As they gain a greater understanding of the challenges faced by the first responders, through their exposure to the test methods, they have also become more involved in the test method development process. In Section 4 we will discuss how this has included bringing advanced robotic hardware to robot evaluation exercises to demonstrate to first responders, government, and robot developers what the state of the art may look like in the future.

In the rest of this section we will briefly outline the way in which the competition is run; for a more detailed discussion please refer to the League Overview [8], Rules Outline [6], and Arena Construction Manual [10]. We will then discuss, in Section 3, some salient examples of test methods that were conceived or refined during the competition, while in Section 4 we will discuss examples of further integration between the competition and the wider standards process.

2.1 Preliminary Missions

The RoboCupRescue Robot League is a point-scoring exercise. Teams run several time-limited ''missions'' within the

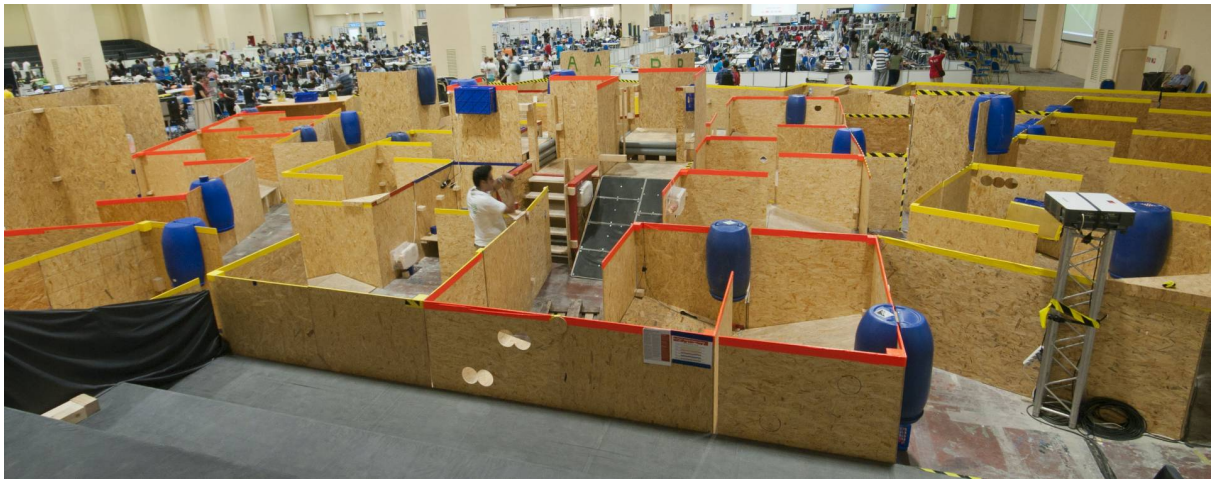

Figure 1: The RoboCupRescue Robot League arena from the 2011 International competition, held in Istanbul, Turkey. Several test method apparatuses may be seen in this photo. The blue barrels are the fiducials used in the Map Quality metric, together with the walls that form the maze. These will be discussed in Section 3.5. The ramp, stairs and raised platform in the center, along with the mismatched ramps in areas highlighted with orange tape, form the structured obstacles of the Orange arena. The continuous pitch-and-roll ramps highlighted by walls with yellow tape form the Yellow arena. The boxes hung against the walls house victims that robots must reach and are discussed in Section 3.4.

arena over the course of several preliminary, semi-final, and final rounds. In each mission, they deploy robots in order to find and characterize the victims, which are distributed throughout the arena. Teams are awarded points based on the quality of the information that they obtain about each victim. This includes the ability to bring back high resolution imagery of the victim, take their temperature or return a thermal image, sense the presence of carbon dioxide, and detect if the victim is speaking. Points are also awarded for the quality of the map in which the victim is reported.

Depending on the number of teams, all teams participate in four to six Preliminary missions, each taking around 15 to 20 minutes. These missions are held in half of the arena, allowing two teams to run missions concurrently. This gives all teams the best opportunity to demonstrate their capabilities to their full potential and allows them to gain experience with the test method apparatuses. In the process, valuable data is generated on the performance of a wide variety of robots in the test apparatuses.

2.2 Championship

The Preliminary missions act as a qualifying round for the Championship, which selects the Champion, 2nd, and 3rd place teams. Usually run as a Semi-finals and Finals round, the Championship takes place in an arena twice the size of that in the preliminary missions, giving teams a greater incentive to rapidly cover as much of the arena as possible. The determination of the Championship is also based on points, set to zero for all teams at the start of the Championship and earned in the same way as in the Preliminaries.

A unique aspect of the League is in the qualification process, which aims to be as inclusive and forgiving as possible to ensure that teams have the freedom to experiment and push their implementations. In the process, teams are able to push the test method apparatuses to their limits. The qualification process ignores the worst of each team's preliminary missions and the qualification cutoff is decided once the distribution of preliminary scores is known. While this means that the number of teams in the Championship is variable, it ensures that there is a clear performance gap between the best performing eliminated team and the worst performing qualified team.

To further encourage the participation of specialist teams, which often fail to qualify due to their narrow focus, and to promote the dissemination of Best-in-Class implementations, the League encourages qualified teams to combine with a team that was eliminated and progress through the Championship as a joint team. On winning or placing, awards are given to both teams.

2.3 Best-in-Class Awards

The Best-in-Class awards, which rank equal in status to the championship, are designed to reward the demonstration of Best-in-Class performance in specific challenges posed by Urban Search and Rescue robotics. Currently, there are three Best-in-Class awards for Mobility, Autonomy, and Manipulation. Each of these awards is also decided on the basis of points, half of them coming from the demonstration of the relevant capabilities in the preliminary missions and half coming from a special Best-in-Class round of the competition, for which an entire day is often dedicated.

2.3.1 Best-in-Class Mobility

The Best-in-Class Mobility award is given to the team that demonstrates proficiency in the test method apparatuses relating to advanced robot mobility. Half of the score for this award is based on the number of victims located by the team in the Red part of the arena during the preliminary missions. This part of the arena tests the mobility of the robots and consists of stepfields, which will be discussed in detail in Section 3.1. The second half of the score is based

on points scored during a special Best-in-Class Mobility run, the nature of which varies from year to year to expose new test method apparatuses to a variety of robots and highly motivated operators. In recent years, this run has been based on the number of laps of the Mobility: Stepfields standard test method within a 10 minute time limit, with the operator out of sight of the robot. In this way the Mobility run becomes an iteration of the standard test. Prior to that, points have instead been awarded for the number of times they can traverse particular test method apparatuses in a 10 minute time limit. For example, traversing the Mobility: Obstacles: Stairs, the Mobility: Obstacles: Hurdles, the Mobility: Inclined Plane, and each pallet of the Mobility: Terrains: Stepfields in each direction would earn a point each.

2.3.2 Best-in-Class Autonomy

The Best-in-Class Autonomy award is given to the team that demonstrates proficiency in challenges relating to autonomous navigation, autonomous detection of objects of interest, and autonomous map building. Half of the score for this award is based on the number of victims found by completely autonomous systems in the preliminary missions. The second half is based on the combined quality and coverage score resulting from a dedicated Best-in-Class Autonomy run through an enlarged maze. This is a direct application of the Map Quality metric [7], currently under development as a standard test method. Autonomous robot mapping is still a highly specialized area so the competition provides a valuable opportunity to gather data from a wide variety of very different approaches from all over the world on the same apparatus and perform rapid iterative refinement as part of the standard test method development process. It also involves researchers into the process; most of the teams that participate in this challenge have also been active in providing input to the standard test method development process.

2.3.3 Best-in-Class Manipulation

The Best-in-Class Manipulation award is given to the team that demonstrates proficiency in challenges relating to manipulating objects in the arena. Half of the score is determined based on the number of objects that teams are able to place with victims in the arena -- teams may retrieve objects, representing such things as radios, water, or supplies, from a shelf and place them with victims that they find. This task by itself is analogous to the placement task in the Manipulation: Grasping Dexterity test. This leads in to the second half of the score, where teams must retrieve objects from one shelf and place them in holes in another shelf as many times as possible in a fixed time period.

3. EMBEDDED TEST METHODS AND APPARATUSES

The field of Urban Search and Rescue provides many challenges. The test method apparatuses represent these challenges, as gathered through extensive consultation with first responders and distilled into separate, reproducible physical challenges. Many of these appear in the RoboCupRescue Robot League arena. Due to the variety of challenges, it is rare for a single team to be able to perform well across all of them. Indeed, it is often the case that good performance in a particular set of challenges is an open research problem and the team that demonstrates best-in-class performance in those challenges needs to dedicate all of their effort and expertise towards solving that particular challenge.

Emerging, draft, and standard test method apparatuses that have appeared, in full or adapted form, in the RoboCupRescue Robot League arenas over the past years include those for the following ASTM standard, validating (V), balloting (B), and prototyping (P) test methods:

- Confined Area Terrains: Continuous Pitch/Roll Ramps (ASTM E2826)
- Confined Area Terrains: Crossing Pitch/Roll Ramps (ASTM E2827)
- Confined Area Terrains: Symmetric Stepfields (ASTM E2828)
- Confined Area Obstacles: Hurdles (ASTM E2802)
- Confined Area Obstacles: Inclined Planes (ASTM E2803)
- Confined Area Obstacles: Stair/Landings (ASTM E2804)
- Confined Area Inspection Tasks: Recessed Targets on Elevated Surfaces (V) (WK27851)
- Confined Area Grasping and Removal Tasks: Weighted Cylinders on Elevated Surfaces (V) (WK27852)
- Search Tasks: Random Mazes with Complex Terrain (B) (WK33259)
- Navigation Tasks: Random Mazes with Complex Terrain (V) (WK33260)
- Mapping Tasks: Hallway Labyrinths with Complex Terrain (P)
- Video: Acuity Charts and Field of View Measures (ASTM E2566-08)
- Audio: Speech Intelligibility (Two-Way) (V) (WK34435)
- Thermal Imager Resolution (P)

In the rest of this section, we will discuss several salient examples of standard test methods that were conceived at, or saw rapid development within, the RoboCupRescue Robot League competition, and which have subsequently become standard test methods or are in the process of becoming standard test methods.

3.1 Symmetric Stepfields

Stepfields are blocks of wood of specified lengths, arranged in a grid such that the tops form an uneven surface. They represent rubble, steps, and other arbitrary terrain in a way that is easy to reproduce [5]. First used to analyze the movement of cockroaches [4], stepfields have since been used in robot evaluations to evaluate the performance of robots of all sizes. A major challenge in the use of stepfields as a standard test method apparatus is specifying the dimensions and standard configuration. Since 2005, the RoboCupRescue Robot League has played a vital part in the evolution of the patterns up to the present day standard. Several iterations of the stepfield appear in Figure 2.

The initial Random Stepfield apparatuses, sized for Urban Search and Rescue robots, consisted of blocks with a footprint of 10x10 cm and varying in length from 5 cm to 40 cm. Early iterations of the stepfields consisted of these blocks arranged in a 10x10 pallet with a prescribed pattern of tall blocks and surrounded by randomly placed blocks, following the rule that adjacent blocks should differ by no more than 20 cm. The tall blocks therefore form ridges or pillars that the robots

(a) (b) (c) (d)

Figure 2: The evolution of the Stepfields standard test method apparatus through the course of the RoboCup competitions. (a) The initial Random Stepfield, as introduced in the 2005 competition. (b) The Symmetric Stepfields, the patterns for which were developed and refined during the 2008 competition. (c) The latest version of the Symmetric Stepfields, adapted for competition. (d) The final form of the Symmetric Stepfields in the test method.

need to navigate around or over and represent larger objects such as pipes or large rocks among smaller, random rubble. The fluid nature of the RoboCupRescue Robot League arena, the large number of missions and the wide variety of robot geometries enabled an immediate comparison between the different schemes and the ability to identify and address shortcomings in the design of the apparatus that were not apparent until they were exposed to particular, unusual robot geometries. Through the course of several competitions, different variants evolved.

These random stepfields proved to be very effective in producing a challenging terrain for advanced mobility robots and helped shape the evolution of the robot geometries from those that were mostly optimized for climbing stairs and curbs to those that could also handle more general rough terrain. However, their randomness impaired the repeatability of the trials and made it difficult to use as a standard test method apparatus. To overcome this, in the 2008 competition symmetric stepfields were introduced where the entire stepfield pallet consisted of a prescribed, symmetric pattern. The competition enabled the evaluation and refinement of a variety of patterns over the course of over 100 missions with a wide range of robot geometries. The final incarnation of the Stepfields test method apparatus appeared in the 2010 competition. The terrain and figure-of-8 pattern, which now forms ASTM Standard E2828 [3], was tested both in the main competition as well as in the Best-in-Class Mobility competition.

3.2 Continuous Pitch/Roll Ramps

In its early days of the RoboCup RRL, the competition was dominated by wheeled robots, usually variants of floor robots used in the lab for research into navigation and planning. In order to make the environment more challenging and closer to what might be encountered in the real world, continuous pitch-and-roll ramps were introduced. Several incarnations of this apparatus appear in Figure 3. These ramps force robots to demonstrate sufficient power and control to position themselves on a non-flat surface, enough degrees of freedom to direct perception when the base is not horizontal, and 3D-aware sensing and mapping in order to generate maps that are correctly registered despite the attitude of the robot changing. However, they are not so hard as to act as a barrier to entry for teams that are not specialized in mechanical engineering.

The competition provided a vital proving ground for the pitch-and-roll ramps, where different layouts could be tested relative to targets that the robots had to approach and inspect or paths that the robots needed to traverse. Different heights of ramps were also tried before the current standard was settled on. Much experience was gained from observing a wide variety of robot geometries perform, from tracked to wheeled to legged robots, performing in the arena. This experience has guided subsequent test method development incorporating pitch-and-roll ramps [1].

3.3 Crossing Pitch/Roll Ramps

Crossing pitch-and-roll ramps, shown in Figure 4, are an evolutionary branch from the continuous variety, first introduced in the 2008 competition in response to the need to develop a terrain that of a difficulty between that of the continuous pitch-and-roll ramps and the stepfields. The resulting terrain should be traversable by wheeled robots only if driven carefully, providing an incentive for teams that were focusing on autonomy to add terrain analysis and more advanced autonomous terrain negotiation.

Once again the fluid nature of the competition arena allowed the rapid evaluation of a wide variety of configurations of the crossing ramps and the ability to observe the way in which a variety of different robot geometries and control methodologies responded to them. These observations have shaped the final crossing ramps test method apparatus that now appears as a middle difficulty test terrain in the standard test method suite [2].

3.4 Inspection Tasks, Acuity Charts, and Thermal Imager Resolution

Points in the main RoboCupRescue Robot League competition are scored based on the robots ability to get close to the "victims" in the arena and obtain information about their state. Examples of victims are shown in Figure 5. All teams make extensive use of visible light cameras and many teams make use of novel thermal sensing techniques.

All of the teams currently in the League are academic teams or in some way associated with universities and high schools. The sensing abilities of the robots that they develop differ significantly from those encountered in the field because as previously mentioned they are significantly less constrained by commercial and practical limitations and are instead focused on research into particular areas of specialization. This has resulted in a wide variety of exotic and experimental sensors, coupled with innovative ways of transferring this

(a) (b) (c)

Figure 3: The evolution of the continuous pitch-and-roll ramps. (a) Their first appearance in 2006. (b) Their subsequent refinement and use throughout the arena. (c) Continuous ramps now appear in many of the standard tests.

(a) (b) (c)

Figure 4: The evolution of the pitch-and-roll crossing ramps. (a) Their conception during the 2008 competition, where they were called "Wacky World". (b) A different form during the 2009 competition matching full and half height ramps. (c) Their final form in the standard test method.

(a) (b) (c)

Figure 5: Examples of victims placed throughout the arena as targets for teams to reach, identify, and localize in their map. (a) An open victim in the stepfield terrain. (b) An open victim in a car being inspected by a robot. (c) A victim hidden in a wall, accessible through holes. The visual acuity eyechart and heating pad (white, back of box) can be seen; the doll in the foreground acts as a secondary visual target and provides an occlusion that shadows the heating pad for ad-hoc thermal imager evaluation.

data, fusing it, and presenting it to the remote operator.

Exposing the test method apparatuses to these robots yields a glimpse at how deployed robots of the future may perform and ways in which the tests may need to scale. For example, several of the top performing robots in the League have vision capabilities that far exceed that of humans, a capability that only exists in the largest of commercially available response robots. Several teams are also experimenting with unusual thermal sensors, some of which detect objects of interest without producing conventional images. Making thermal imager tests relevant to these classes of devices ensures that new sensors that may become widespread in the near future can be meaningfully compared with those that responders are already familiar with.

Another innovation that is well developed in the league but almost unheard of in deployed robots is assistive and full autonomous behavior. These range from controllers to help steady a camera and assist the operator in directing it to where they desire, right through to vision algorithms that can detect, recognize, and interpret objects of interest in the scene. By encouraging teams to incorporate these developments into their robotic entries and fielding them in the emerging test methods, the test method development process gains valuable early insight into the capabilities that are possible and may soon become available. This helps to ensure that when these capabilities are being fielded and marketed that the standards are ready for them. For example, there are now draft test method apparatuses available that test the visual acuity of autonomous and semi-autonomous systems, in a way that is directly comparable to that of teleoperated systems.

An equally valuable side-effect of this insight is that knowledge of these developments can be passed back through the standards process to the first responders, whose needs direct the whole standards process. It is often the case that their requirements are unmet by commercially available robots. Yet, unbeknownst to both responders and the commercial developers, such problems may have already been solved in the research community and are just waiting for commercialization. This is particularly important in the sensing, inspection from mobile platforms, autonomy, and human-robot interaction fields where the abilities shown by implementations in the lab far exceed those currently in deployment. A push from an end user may be all it takes for a robot vendor to bring such developments to life.

3.5 Human-Robot Interfaces, Mapping, and Autonomy

The arena is arranged as a labyrinth, or maze, of hallways that teams must navigate, with test methods embedded at strategic points; between the test methods the robots must navigate portions of maze consisting of mostly continuous pitch-and-roll ramps. Three groups of test methods make use of the maze: Human-robot interfaces, mapping, and autonomy.

Robots that provide their operators with good levels of situational awareness through their human-robot interface, and which respond to the operator's controls in an appropriate manner, tend to perform well in the maze. This is because they are able to drive through the maze without colliding with the walls, a task that is made more difficult due to the introduction of continuous pitch-and-roll ramps, which can make the robot behave in an unpredictable manner. Particularly good user interfaces and predictable controls, especially those that overlay some autonomous behaviors such as automatically moving downrange without colliding with obstacles, also allow the operator to cut corners closer than they might otherwise, further improving their performance. In contrast, operators using interfaces that do not provide good situational awareness tend to misjudge the positions of corners relative to the robot's edges and thus waste time colliding with the walls or taking wide or slow turns.

The maze is also used to test the ability for robots, teleoperated or autonomous, to build 2D and 3D maps of their environments using a variety of algorithms and sensors such as laser range scanners, range imagers and lidars, and various forms of structure-from-vision techniques. For the purpose of evaluating these abilities, the maze is augmented with fiducials [7] that allow various metrics, such as map coverage and consistency, to be measured in a quantitative manner. Finally, the maze is used to evaluate the performance of robots with the ability to autonomously navigate, search, and map a complex environment.

The RoboCup RRL plays a particularly important role in the development of this test method because many of the capabilities being tested -- advanced human-robot interfaces, robot mapping, and autonomous navigation combined with robot platforms that are able to overcome non-flat flooring -- are almost exclusively available only in research robots.

4. INTEGRATION IN THE STANDARDS PROCESS

Participation by the RoboCup RRL Community in the standards process extends beyond the competition. The standards process is enhanced by the involvement of RoboCup RRL teams at response robot evaluation exercises, teaching camps, and in the standards development process itself. The latter is a natural fit as most teams replicate a subset of the standard test method apparatuses in their own labs in order to practice and aid their research.

4.1 Response Robot Evaluation Exercises

NIST hosts response robot evaluation exercises that bring together robot developers, researchers, first responders, procurement officials, and test method developers and administrators. These events, usually held at fire and rescue training facilities, see developers bringing robots to be tested in current and emerging test method apparatuses as well as more unstructured, operationally significant scenarios. First responders and procurement officers observe the robot performances and experience them hands-on, under test conditions, and within the operationally significant scenarios. In the process, data is collected on robot performance in the test methods which allow them to be further developed, refined, and validated.

The Best-in-Class winners of the RoboCup RRL are invited to bring their robots and equipment to these events, in order to give robot developers, responders, and procurement officials a valuable glimpse at the performances that are possible within the standard test methods. This allows them to put the results of deployable robots into a proper context, relative to what is possible based on emerging technology. It also provides data points for the test methods that are often well beyond those achieved by deployed robots.

4.2 Teaching Camps and Summer Schools

Several key features are incorporated into the competition that encourage teams with a wide variety of specializations to enter and to collaborate with each other, investigate the test methods, and contribute to their testing and development. However, it is still a high pressure environment, with teams usually focused primarily on ensuring their entries do well. Ironically, the competition structure, which gives all teams a chance at competing right to the final day so that they have the best chance of showing off their capabilities, also means that teams don't usually have much truly free time. The RoboCupRescue Robot League hosts teaching camps and summer schools several months after the competition, that allow competitors to reflect on and become more familiar with the test methods and the best-in-class implementations that were demonstrated in them. These events are also a vital part of the standards development process as teams, who are encouraged to bring their robots, are able to experiment with the test methods in greater detail and with more freedom than at the competition.

4.3 Standards Process Involvement

Virtually all teams that compete in the RoboCup RRL fabricate at least some of the standard test method apparatuses in their own labs. As noted earlier, this is a very effective way of disseminating the use of the standard test methods through academia. This is further amplified by the inevitable sharing of the facilities that happens at academic institutions, resulting in the standard test methods being used in projects that are not directly related to the competition. As these results are published, the test methods become known to research communities outside the standards process itself. As teams become even more intimately familiar with the test methods, they have also become involved in the test method development process and in many cases team members have subsequently worked directly with NIST on developing standard test methods [7]. There have also been examples of new proposed test methods coming up through the RoboCup process [9]. Some labs that participate in the competition have even opened standard testing facilities in their home countries, based on the standard test methods.

5. FUTURE DIRECTIONS AND CONCLUSIONS

The RoboCup RRL continues to see new test method apparatuses rotating in for exposure to the robotic implementations that teams bring and refine at the event. Test methods for autonomy, 2D mapping, and 3D mapping are being further developed with the assistance of expertise from the RoboCup RRL community -- one of few with such a wide variety of expertise in this area. Likewise, improvements to test methods for visual acuity and other vision based sensing are being made with assistance from the League. Existing and new test methods will continue to be refined through the competition, which continues to see new teams joining and contributing their expertise.

Planning is also underway for the development of new test methods for entirely different classes of response robots, that of robots for fighting fires in the home, and the RoboCup RRL will serve as an integral part of this test method development effort. It will integrate currently prototyping sensing and autonomy test methods with new test methods specific to domestic early fire intervention such as the detection of fire-specific signs and the simulated delivery of suppressant.

The RoboCup RRL has been, and continues to be, an effective tool for aiding the development of standard test methods for response robots. In particular, it provides a venue where current and prototypical test method apparatuses and procedures may be evaluated in the presence of a wide variety of implementations, it brings researchers into contact with the test methods and encourages them to assist in their dissemination, and it allows them to contribute their expertise to the test method development process. The competition also assists the wider effort of NIST in promoting research and development in capabilities for robotic Urban Search and Rescue equipment.

6. REFERENCES

[1] ASTM International. E2826 Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Terrains: Continuous Pitch/Roll Ramps. Technical report, ASTM International, 2011.

[2] ASTM International. E2827 Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Terrains: Crossing Pitch/Roll Ramps. Technical report, ASTM International, 2011.

[3] ASTM International. E2828 Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Terrains: Symmetric Stepfields. Technical report, ASTM International, 2011.

[4] R. J. Full, K. Autumn, J. I. Chung, and A. Ahn. Rapid negotiation of rough terrain by the death-head cockroach. *American Zoologist*, 38:81A, 1998.

[5] A. Jacoff, A. Downs, A. Virts, and E. Messina. Stepfield Pallets: Repeatable Terrain for Evaluating Robot Mobility. In *Performance Metrics for Intelligent Systems Workshop*, 2008.

[6] A. Jacoff, S. Tadokoro, E. Mihankhah, T. Kimra, J. Pellenz, A. Birk, J. Suthakorn, M. Hofbauer, and A. L. Gutierrez. RoboCupRescue Robot League: Rules 2011.2. `http://www.nist.gov/el/isd/upload/Robocup_Rules_2011.pdf`, 2011.

[7] S. Schwertfeger, A. Jacoff, C. Scrapper, J. Pellenz, and A. Kleiner. Evaluation of Maps using Fixed Shapes: The Fiducial Map Metric. In *Performance Metrics for Intelligent Systems Workshop*, 2010.

[8] R. Sheh, T. Kimura, E. Mihankhah, J. Pellenz, S. Schwertfeger, and J. Suthakorn. The RoboCupRescue Robot League: Guiding Robots Towards Fieldable Capabilities. In *International Workshop on Advanced Robotics and Social Impacts*, 2011.

[9] K. Shimaoka, K. Ogane, and T. Kimura. An Evaluation Test Field Design for a USAR Robot considering a Collapsed Japanese House. In *International Symposium on Safety, Security and Rescue Robotics*, 2011.

[10] A. Virts, A. Jacoff, and A. Downs. RoboCupRescue Arena Assembly Guide 2011. `http://www.nist.gov/el/isd/upload/2011_Assembly_Guide.pdf`, 2011.