Standardized Engineering Assessment of Field Robots in a Mars-like Environment

Graham Mann*
Nicolas Small
Kevin Lee
Applied Artificial Intelligence Laboratory
Murdoch University, South St.
Murdoch, Western Australia 6150
{g.mann,n.small,k.lee}@murdoch.edu.au

Jonathan Clarke
Mars Society Australia
P.O. Box 372 Clifton Hill
Victoria, Australia 3068
president@marssociety.org.au

Raymond Sheh
Dept. of Computing
Curtin University, Kent St.
Bentley, Western Australia 6102
raymond.sheh@curtin.edu.au

Abstract

Controlled testing on standard tasks and within standard environments can afford meaningful performance comparisons even between robots of heterogeneous design. But because they must perform practical tasks in unstructured, and therefore non-standard, environments, the benefits of this approach have barely begun to accrue for field robots. This paper describes a desert trial of six student prototypes of astronaut-support robots organized by the Mars Society Australia (MSA) that presented an opportunity to apply a set of standardized engineering tests developed by the US National Institute of Standards and Technology (NIST), along with three operational tests in natural Mars-like terrain. The results suggest that the same standards developed for emergency response robots are also applicable to the fieldwork support domain. They can yield useful insights into the differences in capabilities between robots, resulting in real design improvements. In particular, the number of WiFi errors was observed to be greater than with analog RC, leading to a better communication design for a participating robot. FoV was observed to affect end-to-end visual acuity and target-finding performance, leading one student team to add a second, narrow FoV camera to their robot. The high performance of a small quadrotor convinced members of the MSA board to consider an experimental program to develop them for future use in the atmosphere of Mars. The exercise shows the value of combining repeatable engineering tests with task-specific application testing in the field.

1 Introduction

By their nature, field robots are difficult to evaluate objectively. Apart from the complexity of the machines themselves, they must operate in natural, unstructured environments, which cannot be easily characterized

*Corresponding author

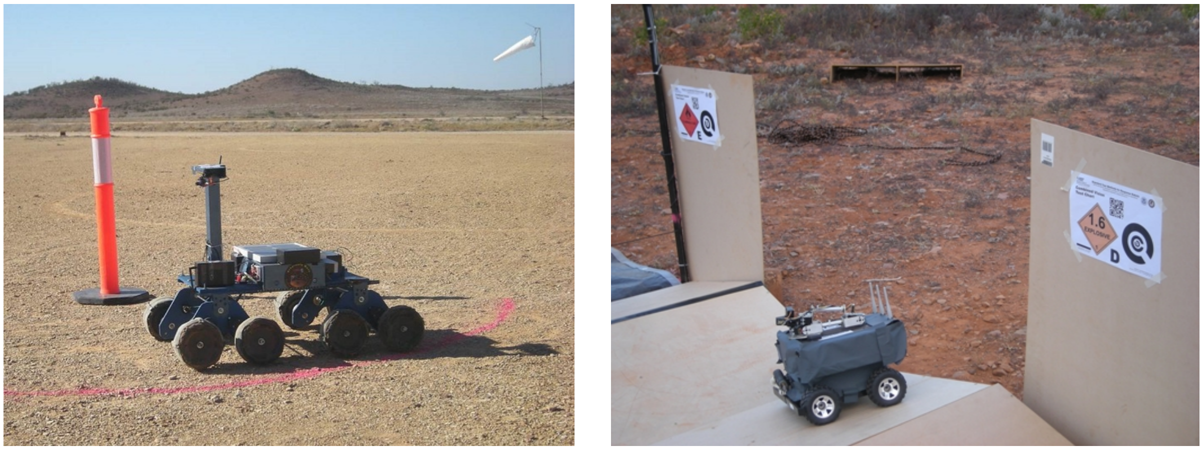

Figure 1: Miner during sustained speed test (left). Corobot on pitch/roll ramps (right).

or measured. The kinds of tasks they must undertake can be uncommon and poorly described. We may expect robots to be flexible in the kinds of behaviours they exhibit, which means that describing a typical task could be inadequate. Worse still, machine design, task and environment cannot be considered as independent factors, since they might interact in complex ways. Another complication is that the robot may be teleoperated, which adds a human controller and user interface to the system, with the attendant problems of evaluating these elements.

The need for standardized benchmarks has been understood for years [Cleary et al., 2005]; [Madhavan et al., 2009] and considerable progress has been made towards widely accepted standards [Moon et al., 2013]. A well-known contribution in this regard is the practice of staging robot competitions that try to hold constant the task, environment and expected behaviours, such as the Robocup events [Birk, 2010], and the DARPA Grand Challenges [Iagnemma, 2006]. In some cases (e.g. LAGR Project [Jackel et al., 2006]; DARPA ARM-S [Hudson et al., 2012]; Robocup Standard Platform League [Barrett et al., 2013]), much of the robot hardware is also fixed, leaving only software solutions and some details of sensors or manipulators as the key design differences to be compared.

This approach has proven quite productive, but it has limitations. The most serious for field robots is that published work in this area tends to be focused on demonstrating the robot's fitness for purpose based on the requirements for the machine, often according to the contingencies of practical funding for such applied engineering. That commits the studies of performance to tasks which are not necessarily standard, or even particularly well-described, and to measurements within environments that cannot easily be described or duplicated. A second, less serious concern is that competition between rival teams might tend to suppress sharing of information about possible problems and solutions, and thus slow progress in the field. In practice, however, it would seem that this can be overcome by encouraging teams to think of the opposition to be overcome as the task itself, rather than the other teams as in a sporting event [Sheh et al., 2012]. That these events customarily end in research conferences attests to the generally cooperative and open-handed spirit of the participants. Thirdly, competition events tend to be held indoors, under quite controlled lighting, weather and ground surface conditions, as well as clearly marked and well-laid out task setups. These can yield comparable results, yet are far from ideal for field robot testing.

Best practice in robot testing can now be found in the DHS-NIST-ASTM International Standard Test Methods for Response Robots [Jacoff et al., 2014]. The April 2014 version of this document describes a comprehensive program of tests for emergency response robots that consist of elemental tasks, carried out in elemental apparatuses. The results are then combined in different ways to represent the expected performance in a wide variety of possible applications. The testing procedures also allow for purpose-built operational tests, which put the robots into realistic scenarios suited to their special functions. For instance, bomb disposal robots can be tested by hiding one or more suspicious packages on a bus, which the robots must find and deal with. An opportunity to see if these tests can be adapted to a different purpose arose in the Arkaroola Mars Robot Challenge, a joint venture between a commercial aerospace contractor, Saber Astronautics, and the Mars Society Australia (MSA), a non-profit, incorporated body that conducts research in support of future manned Mars missions. Four teams, mostly consisting of engineering students, brought six field robots to the MSA's test site in Arkaroola, a remote desert station in central Australia. The machines represented the students' design concepts for robots capable of assisting astronauts performing tasks on the Martian surface. In this context, the operational tasks were designed to model support operations in the harsh Mars-like terrain of Arkaroola. The intention was to encourage such innovation among the engineering students in friendly, low-competition field trials, make useful measurements on their prototypes, use those to learn lessons by comparing the performances of the various designs, and gain experience in constructing, deploying and carrying out these relatively new tests.

Before describing the test regimen, a few caveats are in order. First, unlike the original NIST evaluation exercise at Disaster City in College Station, Texas [Jacoff et al., 2012], most of the participating robots were not manufactured to commercial standards of reliability, but were built by students with limited budgets. Five of the machines should thus be considered to belong an experimental class. The sixth machine is a low-cost commercial quadrotor weighing less than 2kg, with a top speed of less than 30 knots (15.4 m/sec) and frangible, so would be classified by NIST (following the FAA Aviation Rule Committee system) as sUAS (Group 1). Second, these tests were not the only activity planned during the event, and available time with the teams was a constraint on what could be achieved. Third, the tests themselves are a work in progress. Not all DHS-NIST-ASTM specifications and procedures are fully developed at this point, and not many researchers are yet experienced with them. Fourth, in some cases, it was necessary to adapt the tests to local conditions and resource limits prevailing at the test site, though every effort was made to preserve the essential standards. Despite these limitations, we were able to gather useful data from the robots in a difficult outdoor environment and use it. This event represents one of the first examples of the use of the DHS-NIST-ASTM process, task and apparatuses outside the emergency response domain.

2 The Test Program

The participating robots were not envisaged as autonomous explorers that are sent to Mars to accomplish scientific missions, such as the Mars Science Laboratory, Curiosity. Instead they were designed to support human science activity on the surface. As such they are to be controlled locally, rather than over planetary distances. Therefore issues of delay and shared autonomy were not considered for the purposes of these tests. Although they might be used for remote science sampling or photographic surveys, they may equally be required to perform routine maintenance tasks or to fetch-and-carry tools and/or instruments.

The selection of suitable tests for this event began with a list of desirable attributes for small robots to aid human explorers.

- **All terrain capability.**
- **Reliable automatic navigation.**
- **Ability to carry sizeable payloads.**
- **Sufficient endurance for the working day.**
- **Ability to locate and image target objects visually and by GPS.**
- **Radio control, vision and sound over kilometers.**
- **Ability to manipulate small objects (via teleoperation, possibly automated).**

These are not specific enough to be called requirements, and could not serve as a starting point for designs in this context. Indeed, one motivation for field studies of this kind is to better understand the requirements of real future surface operations. However, details of the selected tests were sent to participating teams as soon as they were settled, so some provisions could be made in the time available. At present there is no satellite location system working at Mars, but it is likely that a precision location service would be operating on the surface by the time human explorers arrive, operating either by microsatellite constellation [Pirondini and Fernández, 2006] or by a network of ground stations initially deployed by the robots themselves [Matsuoka et al., 2004]. Current concepts for astronaut-assistance robots are organized around collaborative human-robot networks involving one or more robots that either follow behind the astronauts or can be quickly summoned. Tested hardware ranges from small NASA K-10 four-wheelers capable of carrying 13.6kg of science equipment [Fong et al., 2006] to the Centaur humanoid torso on a wheeled base [Mehling et al., 2007] and more recently, the golf-cart sized RAVEN, able to carry an injured person in an emergency [Akin et al., 2010]. Under this latter requirement, a useful working payload for a personal carrier should extend to the weight of an average person wearing a Z-series spacesuit (approximately 130kg on Earth) - beyond the reach of these prototypes. With standardised testing we may at least measure the maximum load each machine is currently capable of dragging. Not many commercial products would have sufficient endurance for an 8 hour working day at present, though fast battery charging [Armutlulu et al., 2012]; [Medeiros et al., 2012] and fast change out [Suzuki et al., 2012]; [Choi and Lee, 2013] options are being developed. Although automation will play an important role in future designs, teleoperation is very likely to be at least one mode of operation for astronaut-support robots [Podnar et al., 2010], and is at present the simplest option for any testing purpose.

Logistics was also a factor to be considered for the test program. The DHS-NIST-ASTM equipment is well-specified and designed to be relatively simple to construct [Virts and Downs, 2011], but to reduce transport costs for the Arkaroola Mars Robot Challenge, it was important to minimize the bulk and weight of freight items. This meant redesigning some of the equipment without compromising the standards. The mobility test environment, in particular, needed to be reduced to its minimum elements for economical packing and ease of assembly: 24 wooden ramps, and a 7.2m x 2.4 skeletal enclosure consisting of metal fence pickets, plastic sheeting and ropes. Some desirable tests (e.g. Symmetric Step Fields) had to be left out to minimize costs and construction time, while others (e.g. Radio Comms: Non-Line of Sight Environments) were too large and heavy to be practically transported and set up at the test site. Although as a rule DHS-NIST-ASTM tests could easily be conducted in convenient urban areas, any event involving a gathering of many field robots on location might find the need to economize on time, money and effort.

2.1 Selected DHS-NIST-ASTM tests

Based on these considerations, the following tests were selected from the DHS-NIST-ASTM document [Jacoff et al., 2014].

1. Logistics: Robot Test Config and Cache Packing. The process required the completion of forms for every participating machine to capture details of the physical properties, equipment specifications, configurations, toolkit, packing and transport logistics. These reflect not only the US government sponsor's need to understand the deployability of the robot, but to help characterize each for classification and comparison. The information includes specific photographs of a robot, in different poses and from various angles, against a calibrated background. The information is particularly important for managing the configuration of robots from one test to another. The elemental tests can be taken in combination to represent performance in an operational scenario. Thus it is vital that a given robot undergoes a the various tests without modification.

2. Energy/Power: Endurance : Terrains: Pitch/Roll Ramps. A test rig consisting of 24 15° wooden ramps measuring 1200 x 600mm was laid out in a specified alternating sawtooth pattern to repeatably measure the robots' performance on discontinuous terrain (see Figure 3). Participants guided the robots around a 15m figure-eight path on the ramps around two suspended pylons. Distance d_i and time t_i from full battery charge to inoperability are measured. Because this test had to be conducted in the field, it was

Figure 2: Summary design features of participating robots

necessary to almost completely eliminate the side walls of the environment to save weight and reduce wind stresses (see Figure 3, left). It turned out to be impractical to bench test sets of batteries through multiple charge-recharge cycles in the field.

3. Mobility: Terrains: Flat/Paved Surfaces (100m). Two pylons were placed 50m apart on a flat surface. The ground around each was marked with a circle 2m in diameter. The robots were to make 10 timed figure-of-eight laps around this course, without deviating from the circumscribed path. Thus both speed and control are important. The NIST form suggests reporting the average time per repetition in minutes, but meters-kilogram-second (MKS) units are preferable here, so Table 1 reports their average speeds in meters per second.

4. Mobility: Towing: Grasped Sleds (100m). The robots dragged an aluminum sled, carrying an operator-designated payload, around 10 figure-of-eight laps on the 100m course specified in test 3. Average velocities v_{av} and maximum achieved weights m were recorded. Ideally, the test should be conducted on a concrete paved surface, but this was not available at the test site, so a flat roadway of limestone gravel had to serve. To compare these performances to those of any test on concrete, the different coefficients of sliding friction μ_k between the metal sled and the two surfaces must be taken into account. Assuming that the overall motive force F achieved by the robot (and thus its speed) is dominated by the force it exerts during the constant velocity phases that were observed in the two straight legs of the course, then the key relationship is simply

$$F = \mu_k m g$$

The two coefficients were measured experimentally by dragging the sled loaded with a known mass ten times on both concrete and limestone gravel surfaces at a constant speed, yielding averages of μ_k (Al-concrete) of 0.70 and μ_k (Al-gravel) of 0.42. Thus for a given mass, 40% more force would be required to achieve the same performance on concrete.

5. Radio Comms:Line-Of-Sight Environments. The robots were tested for navigation control and video feed on a straight course at 50m, then stations every 100m thereafter. The robot circumnavigated each station at a radius of 2m, reading a 35 x 35mm bold letter and identifying a standard 100 x 100 mm hazardous material label on the four vertical faces of a box atop a pylon. The last station at which both navigation control and video were perfectly reliable (complete circle and all four visual tests correct) was reported.

6. Sensors:Video:Acuity Charts and Field of View Measures. The robots were placed on a 15° ramp 6m from a far-field Landolt-C vision chart (Figure 3). The operator viewed the chart at their control station via the robot's camera and read down the chart to the smallest line at which the orientations of the C shapes were discernible. No more than two errors were permitted on a line. This is reported as l_{far} a percentage of the 6-6 (20:20) vision standard. The same procedure was used for the near-field Landolt-C chart, except that the distance was then 40cm. The horizontal field of view fov_h was calculated by measuring the distance between the far-field chart and the camera at the point where the long sides of the chart are at the edges of the video screen.

Somewhat lacking here is a specific test for the quality of human-robot interaction. The DHS-NIST-ASTM process tests the performance of the robot system end-to-end, from the robot's sensors and actuators through to the user interface. This approach ensures that all factors influencing the performance of the complete robot system, from the quality of the cameras through to the resolution of the screen, are reflected in the measurement. This is in contrast to most other testing regimes, which seek to test components in isolation and often make the assumption, implied or otherwise, that the performance of the individual components are a reflection of the performance of the whole system. In a sense, this means that all of our tests have a human-robot interface component. The DHS-NIST-ASTM now has specific tests that focus on human-robot interaction, such as the Pan-Tilt-Zoom tasks. These are as affected by factors such as communications delay and sensor mount geometry as they are by the skill of the operator and the design of the control unit. However, they were not sufficiently specified in time to be incorporated in this event.

2.2 Operational tests

The following three operational tests were designed to evaluate the robots in tasks approximating their intended purpose. Some of these were inspired by the University Rover Challenge, a competition event held annually in Utah [Post and Lee, 2011]. Others came from our own prototyping of such a system in the Mascot hexapod [Mann and Baumik, 2011].

1. Irregular Terrain Traversal. A 106m course consisting of four gates (1.2m pylons spaced 2m apart) was arranged over rough, natural, Mars-like terrain. It included a slopes of between approximately 20° to 40°, loose sand, and large irregular stones (see Figure 3, right). Operators were allowed to walk the course before robot testing. The robots were video recorded and timed during their traversal of the course.

2. Context Imaging. A small, brightly painted 100g target object was placed at a random locations on roughly level ground at distances of between 43 and 76m from the starting point. The operator was given the object's GPS coordinates. The operator was to locate the object as quickly as possible, then photograph it in context. Time to locate the target t_{loc} and distance to target d_t were recorded. Each operator chose his best four images to be rated for quality. Each image was made anonymous, then examined by three expert field geologists who rated each according to five criteria: object in context (shows surrounding structure), image composition, brightness and contrast, sharpness of focus and image resolution. The mean rating over all images, experts and criteria was then computed from the ratings and expressed as a percentage of the perfect score q_{av}.

3. Sample Return. Operators of robots equipped with a manipulator had the option to use it in a variation of the Context Imaging task. The robots had to carry a small geologist's scale, place it alongside the located target object, photograph the object in context, collect the object then return it to the starting location. Time to return t_{ret} was reported.

Table 1: Summary of NIST tests

	Endurance		Mobility	Sled Towing		Comm Range	Visual Acuity		Horiz.FoV
	$d_i(m)$	$t_i(s)$	$v_{av}(m/s)$	$m(kg)$	$v_{av}(m/s)$	$d_s(m)$	$l_{far}(\%)$	$l_{near}(\%)$	$fov_h(deg)$
Little Blue	105	2363	0.47	abstained		200	15	25	39.3
Miner	abstained		0.34	6.47	0.33	200	15	20	68.6
UNSW Rover	abstained		1.59	31.47	1.27	500	20	25	66.9
Corobot	150	2032	0.36	3.47	0.27	50	<10	5	120
Mascot	-	-	-	-	-	-	12	15	63.3
Phantom 2	>1720*	-	0.58	abstained		>860*	<10	<5	125.3

*result of non-standard test

3 Results

The tests were conducted in July, 2014, in three locations: a flat camping area near the Arkaroola station facilities, a gravel airstrip used by the neighbouring Wooltana station (Figure 1, left) and a disused quarry with a variety of Mars-like ground conditions including a curved, gullied slope for the operational tests. Conditions were generally favourable but wind, dust and, on one occasion, rain created problems for the test program. In particular, very fine dust combined with dry air (relative humidity range 21-44%) caused a number of failures of robot and test electronics. The most serious of these were a malfunctioning Arduino board on the Miner which took it out of service and the catastrophic failure of a compact laptop in the Mascot hexapod which eliminated it from all but the visual acuity tests.

To gain the full benefit of comparisons between the robot designs, it is important to describe the design features of the competing machines. One limitation of the published results of the Emergency Response Robot Evaluation Exercise [Jacoff et al., 2012] was that for reasons of commercial confidentiality, and because the NIST must not appear to endorse or disendorse products or companies, the names and design details of machines were not disclosed. Instead, individual machines were represented by nominal codes. The Arkaroola Mars Robot Challenge was under no such constraints, and so it was possible to provide these details, using data collected in the Logistics: Robot Test Config and Cache Packing task. It then became possible to study how specific design features contributed to the performances of the individual entries. Table 1 briefly summarizes the results. Other details are provided in the discussion section as needed.

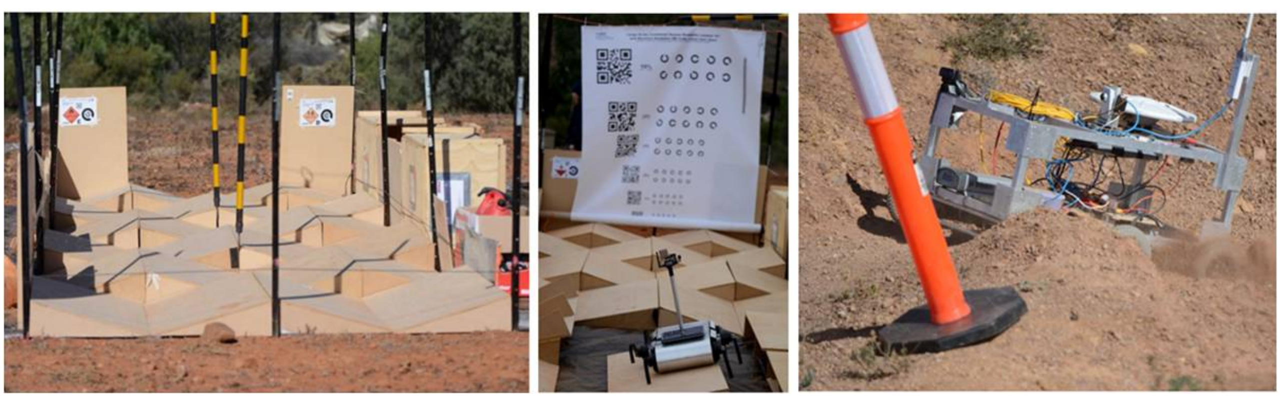

Figure 3: (Left to right) Continuous 15° pitch/roll ramps. Mascot at the far-field acuity test. UNSW Rover traversing irregular terrain.

Table 1 gives the results of the selected DHS-NIST-ASTM tests. As specified in NIST documents and required by our human research ethics permit, teams could abstain from tests that they did not wish to enter. The data reported here represent the target metrics specified by NIST for the tests, but the units have been

Table 2: Summary of operational tests

	Irreg. Terrain $v_{av}(m/sec)$	Context Imaging $t_{loc}(s)$	$d_t(m)$	$q_{av}(\%)$	Sample Return $t_{ret}(s)$
Little Blue	0.2	417	43.1	65.0	N/A
Miner	0.11	-	-	-	N/A
UNSW Rover	0.27	789	47.1	82.5	1569
Corobot	0.08	1980	52.1	79.7	-
Mascot	-	-	-	-	N/A
Phantom 2	0.74	163	76.0	76.7	311

altered to the conventional MKS system in some cases. In the event, the teams found the pitch/roll ramps a challenging environment, either because it was hard on robots with small wheels and no suspension, or because some robots seemed too large to be compatible with the apparatus. The latter may be a shortcoming of this test. Only two robots participated, both requiring multiple operator interventions (from jamming, toppling, or loss of wheels) and only one (Corobot) completed the entire 150 meter course. All of the wheeled ground robots were able to participate in the other tests, but during days with many tests a shortage of spare fully-charged battery packs sometimes limited what could be done. Little Blue was the least massive (7.5 kg) and had no convenient point for hitching a sled, so the operator abstained from that task.

The UNSW rover lacked any axle suspension and had a clearance of less than 10cm, yet was able to outperform the other vehicles in speed, load-carrying capacity, radio range and visual acuity. Miner initially suffered poor traction due to inadequate tread on its eight plastic tires until the operators improvised a repair from strips of rubber fixed axially to the wheels which overcame the slippage. Results on the visual acuity tests were generally poor (but see Section 4). As an unmanned aerial vehicle (UAV), the Phantom 2 defied many of the test protocols, but some roughly equivalent, though non-standard test results were included for informal comparison. It had sufficient range to exhaust the expanse of flat ground available for formal range-testing, but the operator flew the robot to a peak approximately 860m from the launch point and circumnavigated a cairn there several times with full video feedback before returning. He claimed to have flown it out to its imposed safety range limit of 2km on occasion.

Even after the wheel slippage was corrected, the Miner experienced difficulties in the operational tests because not all the wheels were powered, so that traction was still erratic. Close study of the suspension system from video shot during the Irregular Terrain Traversal task suggested that the rocker-bogie suspension might perform better on uneven surfaces if i) the center bearings of the rocker arms were improved to allow them to rotate more freely and ii) all eight wheels were driven. If the driven wheel of a rocker left the ground, the other tended to stay in contact, but because it was not powered, that corner of vehicle lost control and the vehicle tended to yaw. Driving all eight wheels might increase power consumption, but lower-powered motors could be used since there would be more torque available at the ground. The Phantom 2 was opportunistically tested on the operational tasks, and displayed a high level of performance on tests where speed and manoeuvrability were important. It was able to successfully complete the Sample Return task, by suspending a rare-earth magnet on a 190cm line to collect the steel target object.

4 Analysis and Design Implications

These results were compiled into machine-specific reports that were then communicated to the participating teams for their use. In the case of machines from the Applied Artificial Intelligence Laboratory (Mascot and Corobot), these have already begun to change the design thinking of two developers. Here, we focus on three outcomes that have emerged so far.

Reliability of radio link. Two kinds of radio communication links between the operator control unit (OCU) and the robot were on display: analog radio control (RC) transmitter/receiver pairs operating in the FM band (Little Blue, Miner, Mascot and Phantom 2), and orthodox 802.11b/g/n WiFi, sometimes boosted by special antennae, between a laptop and a wireless modem (UNSW Rover and Corobot). During the Arkaroola Robot Challenge trials, we noticed that the reliability of the WiFi links, in particular was poor, with long setup times and several observed data packet losses with subsequent interruptions to control and telemetry services, while the analog RC suffered no such problems. We further note that many emergency response robots tend to prefer analog radio to any kind of digital wireless system, for reasons of fast setup, simplicity and reliability. Model RC systems suit experimental class robots because they tend to be compact, lightweight, rugged, low-powered and low-cost. The best of today's hardware uses frequency hopping to avoid cross-system interference and can operate at ranges of over 15 kilometers (e.g. DragonLink). A disadvantage for teleoperation systems based on such hardware has traditionally been its lack of data transfer capability, but developments in model control products now make available bidirectional telemetry links that can be adapted to this purpose. For example, we noted poor OCU control the Mascot before its computer failed. We plan to modify the Mascot's control system from a simple six-channel RC link (two of which differentially signal the gait software via A/D converters) to an eight-channel RC link using a Taranis X9D transmitter and a FrSky D8R-XP receiver. As well as the RC servo channels, this pair offers duplex RS-232C data transfer at up to 9600bps, which is to be used to form a command link between the Mascot's on-board computer and a 7-inch tablet at the control station running a bespoke touch-screen interface (Figure 6) with large buttons. In both control unit and robot, the data connection requires an RS232C-to-USB converter, since USB is the only connector available on our laptops. Although convenient, USB connectors are not generally rugged enough for fieldwork in rough terrain, so special care needs to be taken to make them reliable [Foster, 2012].

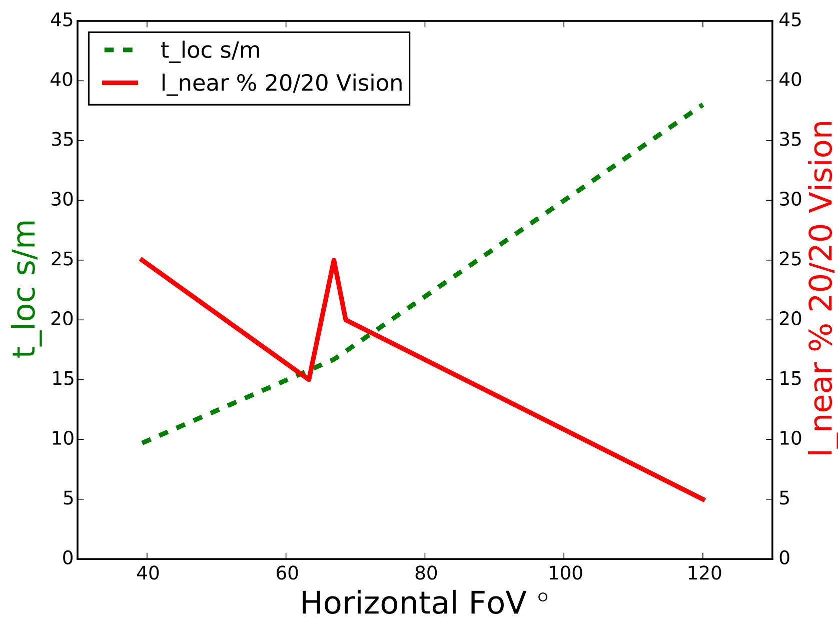

Figure 4: Near-field visual acuity and target locating ability as functions of camera field of view.

Implications of the FoV/acuity tradeoff. Figure 4 plots measured acuity (lowest readable lines on the Landolt-C near-field chart) against the camera's horizontal field of view (FoV). The relationship is non-linear, possibly due to the high quality of the UNSW Rover's Axis M10 camera, but considering that this test actually records the judgments of different human operators viewing the outputs of different cameras on screens of various sizes and quality, it nevertheless shows some tendency to fall as the field of view increases. In a way this is not surprising: for a given number of pixels, lenses projecting the the image of an object over a smaller area of an image array will focus less diffuse information about shape and edges. The FoV/resolution tradeoff is well known [Greene and Vollmerhausen, 1988]; [Lai et al., 2012], but the variable measured here is not simple image resolution. The acuity achievable by a human operator viewing the scene transmitted from the robot also depends on the sharpness and contrast available at the screen as well as their age-related quality of vision [Barten, 1999].

Furthermore, we observed that FoV mattered for locating targets. This variable is even less simply related to image resolution: the operators skill at pointing the camera and the speed of the robot will affect target acquisition times. All the ground vehicle operators reported depending heavily on visual imaging, rather than GPS location, during the context imaging task. Apparently the error radii of the GPS equipment used was too great to be of much assistance at this scale. To normalise the location times t_{loc} with respect to the different target distances in each trial, this was divided by distance from the starting point d_t. As fov_h fell, so approximately did t_{loc}/d_t. Although there are only three observations (the Phantom 2 quadrotor is excluded here because of its relatively high speed, its advantageous aerial viewpoint and because its operator reported depending on GPS localisation), this is broadly consistent with the above observation. Yet a narrow field of view is likely to be detrimental to situational awareness [Hughes et al., 2003]. We observed that the Corobot with its 120° Genius camera allowed its operator to avoid numerous obstacles and snares that bedevilled Little Blue. While it is possible to build foveated cameras that with higher pixel densities concentrated near the optical axis like biological retinae [Ude et al., 2006], these are complex and not readily available. Now, with improvements in cost, weight and quality of camera technology, providing multiple cameras with suitable FoVs and mounting points is clearly a simpler, less expensive and more practical solution for field robots. Based on these considerations, the Corobot operator has now added a small 640 x 480 pixel EXOO camera with a 30° FoV to the arm (Figure 6). The operator can switch between this and the wide FoV Genius

Figure 5: UNSW Rover collecting target object during Sample Return task (left). Phantom 2 quadrotor above horizontal visual targets (right).

camera as appropriate for situational awareness or better target-finding, or to obtain close-up views of manipulated objects.

Value of quadrotor for astronaut support. In many ways, the Phantom 2 quadrotor was in a class of its own. It showed a comparatively high level of performance on most tests. Its GoPro Hero3+ camera could record over 2 hours of 1080p/24fps video for later playback. A special-purpose DHS-NIST-ASTM test, Aerial:sUAS (Group 1) VTOL Station-Keeping, was attempted for this robot. The task requires the operator to keep the machine hovering 2m from five pairs of visual targets arranged at the corners and center of a 5 x 5m square as they are identified. Each target combined a 35 x 35mm bold letter, a standard 100 x 100 mm hazardous material diamond and concentric Landolt-C figures representing feature sizes of 20, 8, 3.2, 1.3 and 0.52mm. The operator is to visit each pair twice, but never the same target successively, while identifying the letter, Hazmat sign and the orientation of the smallest possible Landolt-C on one target each visit. Mean time per repetition and average of smallest feature sizes identified are to be reported. The DHS-NIST-ASTM document specifies that the targets are to be arranged both vertically and horizontally to separately test forward and downward cameras. However, since local buildings or trees did not offer stable support for the vertical structures of that height, and because the same gimbal-mounted GoPro camera would be used in both cases, measurements were carried out with horizontal targets only (Figure 5, right). The operator used the camera's still mode, which captured 12 megapixel images that could be quickly displayed on a large video screen. The results show good performance: a mean repetition rate of 75 seconds, and average feature size of 3.2mm identifiable at the 2m specified altitude.

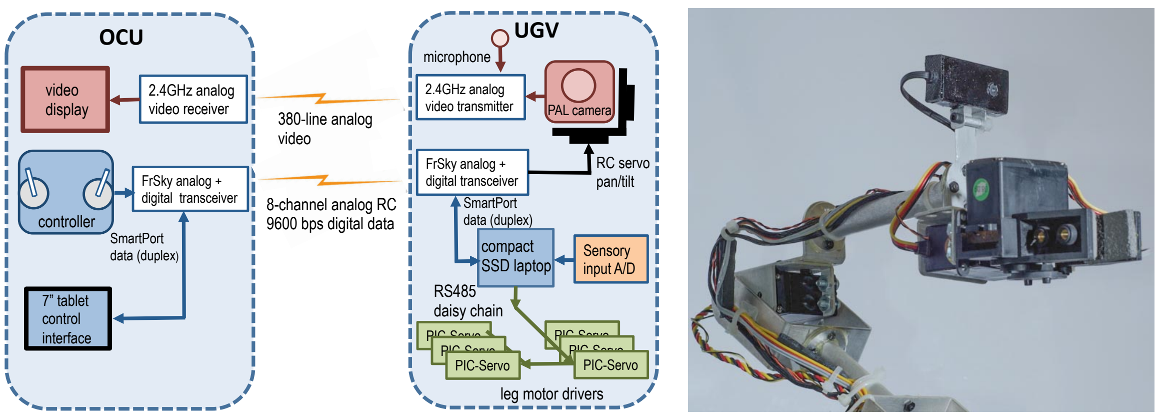

Figure 6: Design modifications to two test robots based on field testing. The Mascot hexapod requires a better RC link which can carry digital data at 9600bps between the robot's laptop and a tablet computer with touch screen command interface (left). Addition of a 30° FoV camera on the Corobot's manipulator arm (right).

5 Conclusions

This event demonstrated that at least some standardized DHS-NIST-ASTM tests can be conducted in the field for little expense and difficulty. Such data can be laid alongside those from equivalent tests made at other times and places, for direct comparison. Furthermore, standard tests can be supplemented by operational tests more specifically suited to the special purposes of field robots and beyond the those of response robots. These tests can be used not only to provide informative field performance data, but also, when a number of different designs can be tested alongside each other, the comparisons can result in a useful sharing of design ideas, and practical suggestions leading to design improvements. This combination of controlled, standardized tasks and actual performance at the worksite provides the advantages of both at the same event.

The most common failure modes for the six robots were physical toppling due on machines with a high center of gravity, sticking on test equipment and wheel slippage and decoupling. Loose connectors, resets or circuit failure due to dust and WiFi dropouts were also common. Orthodox ruggedization methods could deal with many of these problems in these designs. We believe analog RC systems are best for experimental class robots, especially newer designs with frequency-hopping and bidirectional digital data links. The visual capability of the human-robot systems was affected by field of view, sharpness and contrast available at the OCU screen and the operator's visual acuity. Cameras with small FoVs provide for better end-to-end visual acuity and thus better target-finding capability, but less situational awareness, and they may add to operator workload by requiring many skilled steering motions. The pattern is reversed for large FoVs. Clearly, cameras of various FoVs that can be easily selected by at the OCU are the simplest solution here. Future test events of this kind should include specific attention to the usability of OCU layouts and controls.

Finally, the performance of the Phantom 2 quadrotor impressed observers at the Arkaroola Mars Robot Challenge. In a number of simulated short exploratory sorties with two astronauts in simulated spacesuits and one assistance robot on the ground, the Phantom 2 offered valuable support for mission oversight, large target identification and videography. Substantial modification would be necessary to make a quadrotor fly in the thin atmosphere of Mars [Young et al., 2005], but the potential utility is so high that some members of the MSA Board are already considering the possibility of a balloon-borne high-altitude test of such a modified machine and the authors plan to purchase one for experimentation shortly.

Acknowledgments

This research was supported by the Australia-India Council, the Commonwealth Science and Industrial Research Organization and the School of Engineering and Information Technology, Murdoch University.

References

[Akin et al., 2010] Akin, D. L., Bowden, M. L., Saripalli, S., and Hodges, K. (2010). Developing technologies and techniques for robot-augmented human surface science. In *AIAA Space 2010 Conference and Exhibition, Anaheim, California.*

[Armutlulu et al., 2012] Armutlulu, A., Allen, S. B., and Allen, M. (2012). 3d microstructures for fast charge and discharge batteries. *PowerMEMS 2012.*

[Barrett et al., 2013] Barrett, S., Genter, K., He, Y., Hester, T., Khandelwal, P., Menashe, J., and Stone, P. (2013). Ut austin villa 2012: Standard platform league world champions. In *RoboCup 2012: Robot Soccer World Cup XVI*, pages 36–47. Springer.

[Barten, 1999] Barten, P. G. (1999). *Contrast sensitivity of the human eye and its effects on image quality*, volume 72. SPIE press.

[Birk, 2010] Birk, A. (2010). The true spirit of robocup [education]. *Robotics Automation Magazine, IEEE*, 17(4):108–108.

[Choi and Lee, 2013] Choi, Y.-H. and Lee, D.-H. (2013). Battery swap station for mobile robot. In *Electrical Machines and Systems (ICEMS), 2013 International Conference on*, pages 1445–1447. IEEE.

[Cleary et al., 2005] Cleary, K., Kinsella, A., and Mun, S. K. (2005). Or 2020 workshop report: Operating room of the future. In *International Congress Series*, volume 1281, pages 832–838. Elsevier.

[Fong et al., 2006] Fong, T., Kunz, C., Hiatt, L. M., and Bugajska, M. (2006). The human-robot interaction operating system. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 41–48. ACM.

[Foster, 2012] Foster, B. (2012). Ruggedize usb connections for tough environments. *EE Times Newsletter.*

[Greene and Vollmerhausen, 1988] Greene, D. and Vollmerhausen, R. (1988). Night vision pilotage system field-of-view (FOV)/resolution tradeoff study flight experiment report. *Fort Belvior, VA: US Army Night Vision Laboratory. NV*, pages 1–26.

[Hudson et al., 2012] Hudson, N., Howard, T., Ma, J., Jain, A., Bajracharya, M., Myint, S., Kuo, C., Matthies, L., Backes, P., Hebert, P., et al. (2012). End-to-end dexterous manipulation with deliberate interactive estimation. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2371–2378. IEEE.

[Hughes et al., 2003] Hughes, S., Manojlovich, J., Lewis, M., and Gennari, J. (2003). Camera control and decoupled motion for teleoperation. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, volume 2, pages 1339–1344. IEEE.

[Iagnemma, 2006] Iagnemma, K. (2006). Editorial for journal of field robotics special issue on the darpa grand challenge.

[Jackel et al., 2006] Jackel, L. D., Krotkov, E., Perschbacher, M., Pippine, J., and Sullivan, C. (2006). The darpa lagr program: Goals, challenges, methodology, and phase i results. *Journal of Field Robotics*, 23(11-12):945–973.

[Jacoff et al., 2014] Jacoff, A., Downs, A., Huang, H., Messina, E., Saidi, K., Sheh, R., and Virts, A. (2014). Standard test methods for response robot. ASTM International Committee on Homeland Security Applications,: Operational Equipment; Robots (E54.08.01).

[Jacoff et al., 2012] Jacoff, A., Huang, H.-M., Virts, A., Downs, A., and Sheh, R. (2012). Emergency response robot evaluation exercise. In *Proceedings of the Workshop on Performance Metrics for Intelligent Systems*, pages 145–154. ACM.

[Lai et al., 2012] Lai, J. S., Ford, J. J., Mejias, L., Wainwright, A. L., O'Shea, P. J., and Walker, R. A. (2012). Field-of-view, detection range, and false alarm trade-offs in vision-based aircraft detection.

[Madhavan et al., 2009] Madhavan, R., Lakaemper, R., and Kalmár-Nagy, T. (2009). Benchmarking and standardization of intelligent robotic systems. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–7. IEEE.

[Mann and Baumik, 2011] Mann, G. A. and Baumik, A. (2011). A hexapodal robot for maintenance operations at a future mars base. In *Proc. of the 11th Australian Mars Exploration Conference*.

[Matsuoka et al., 2004] Matsuoka, M., Rock, S. M., and Bualat, M. G. (2004). Autonomous deployment of a self-calibrating pseudolite array for mars rover navigation. In *Position Location and Navigation Symposium, 2004. PLANS 2004*, pages 733–739. IEEE.

[Medeiros et al., 2012] Medeiros, R., Valtchev, S., and Valtchev, S. (2012). The efficient and stable charging of electric vehicle batteries: Simplified instantaneous regulation. In Camarinha-Matos, L., Shahamatnia, E., and Nunes, G., editors, *Technological Innovation for Value Creation*, volume 372 of *IFIP Advances in Information and Communication Technology*, pages 363–374. Springer Berlin Heidelberg.

[Mehling et al., 2007] Mehling, J. S., Strawser, P., Bridgwater, L., Verdeyen, W., and Rovekamp, R. (2007). Centaur: Nasa's mobile humanoid designed for field work. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 2928–2933. IEEE.

[Moon et al., 2013] Moon, S., Rhim, S., Cho, Y.-J., Park, K.-H., and Virk, G. S. (2013). Summary of recent standardization activities in the field of robotics. *Robotica*, 31(02):217–224.

[Pirondini and Fernández, 2006] Pirondini, F. and Fernández, A. J. (2006). A new approach to the design of navigation constellations around mars: the marco polo evolutionary system. In *AIAA 57th Intcrnational Astronautical Congress, IAC*, volume 7, pages 4692–4700.

[Podnar et al., 2010] Podnar, G., Dolan, J., and Elfes, A. (2010). Telesupervised robotic systems and the human exploration of mars. *Journal of Cosmology*, 12:4058–4067.

[Post and Lee, 2011] Post, M. and Lee, R. (2011). Lessons learned from the york university rover team (yurt) at the university rover challenge 2008–2009. *Acta Astronautica*, 68(7):1343–1352.

[Sheh et al., 2012] Sheh, R., Jacoff, A., Virts, A.-M., Kimura, T., Pellenz, J., Schwertfeger, S., and Suthakorn, J. (2012). Advancing the state of urban search and rescue robotics through the robocuprescue robot league competition. In *8th International Conference on on Field and Service Robotics (FSR 2012), Matsushima, Japan, 16-19 July 2012*.

[Suzuki et al., 2012] Suzuki, K. A., Kemper Filho, P., and Morrison, J. R. (2012). Automatic battery replacement system for uavs: Analysis and design. *Journal of Intelligent & Robotic Systems*, 65(1-4):563–586.

[Ude et al., 2006] Ude, A., Gaskett, C., and Cheng, G. (2006). Foveated vision systems with two cameras per eye. In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pages 3457–3462. IEEE.

[Virts and Downs, 2011] Virts, A., J. A. and Downs (2011). *RoboCupRescue Arena Assembly Guide 2011.* National Institute of Standards and Technology.

[Young et al., 2005] Young, L. A., Aiken, E., Lee, P., and Briggs, G. (2005). Mars rotorcraft: possibilities, limitations, and implications for human/robotic exploration. In *Aerospace Conference, 2005 IEEE*, pages 300–318. IEEE.