2008 IEEE/RSJ International Conference on Intelligent Robots and Systems
Acropolis Convention Center
Nice, France, Sept, 22-26, 2008

Scaling Effects in Multi-robot Control

Prasanna Velagapudi, Paul Scerri, Katia Sycara
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Huadong Wang, Michael Lewis, Jijun Wang*
University of Pittsburgh
Pittsburgh, PA 15260, USA
*Quantum Leap Innovations, Newark, DE

Abstract— **The present study investigates the effect of the number of controlled robots on performance of an urban search and rescue (USAR) task using a realistic simulation. Task performance increased in going from four to eight controlled robots but deteriorated in moving from eight to twelve. Workload increased monotonically with number of robots. Performance per robot decreased with increases in team size. Results are consistent with earlier studies suggesting a limit of between 8-12 robots for direct human control. This study demonstrates that these findings generalize to a more realistic setting and complex task.**

I. INTRODUCTION

Applications for multirobot systems (MrS) such as interplanetary construction or cooperating uninhabited aerial vehicles will require close coordination and control between human operator(s) and teams of robots in uncertain environments. Human supervision will be needed because humans must supply the perhaps changing, goals that direct MrS activity. Robot autonomy will be needed because the aggregate demands of decision making and control of a MrS are likely to exceed the cognitive capabilities of a human operator. Controlling robots that must act cooperatively, in particular, will likely be difficult because it is these activities [1] that theoretically impose the greatest decision making load.

Because some functions of a MrS such as identifying victims among rubble depend on human input, evaluating the operator's span of control as the number of controlled entities scale is critical for designing feasible human-automation control systems. Current estimates of human span of control limitations are severe. Miller [2], for example, showed that under expected target densities, a controller who is required to authorize weapon release for a target identified by a UCAV, could control no more than 13 UAVs even in the absence of other tasks. A similar breakpoint of 12 was found by [3] for retargeting Tomahawk missiles. Smaller numbers (3-9) [4] have typically been found for ground robots.

Controlling multiple robots substantially increases the complexity of the operator's task because attention must constantly be shifted among robots in order to maintain situation awareness (SA) and exert control. In the simplest case an operator controls multiple independent robots interacting with each as needed. A search task in which each robot searches its own region would be of this category although minimal coordination might be required to avoid overlaps and prevent gaps in coverage. Control performance at such tasks can be characterized by the average demand of each robot on human attention [4]. Under these conditions increasing robot autonomy should allow robots to be neglected for longer periods of time making it possible for a single operator to control more robots.

Established methods of estimating MrS control difficulty, the neglect tolerance model, NT, [4] and the Fan-out measure [5] are predicated on the independence of robots and tasks. In the NT model the period following the end of human intervention but preceding a decline in performance below a threshold is considered time during which the operator is free to perform other tasks. If the operator services other robots over this period the measure provides an estimate of the number of robots that might be controlled. Fan-out refers to maximum number of robots that can be advantageously controlled under particular conditions. Fan-out can be determined empirically as in [5] by adding robots and measuring performance until a plateau without further improvement is reached or indirectly by predicting the maximum number of robots using parameters from the NT model [4]. Both approaches presume that operating an additional robot imposes an additive demand on cognitive resources. These measures are particularly attractive because they are based on readily observable aspects of behavior: the time an operator is engaged controlling the robot, interaction time (IT), and the time an operator is not engaged in controlling that robot, neglect time (NT).Because of the need to share attention between robots in MRS, teloperation can only be used for one robot out of a team [6] or as a selectable mode [7]. Some variant of waypoint control has been used in most of the MrS studies we have reviewed [4,6,7,8] (see Table 1) with differences arising primarily in behavior upon reaching a waypoint. A more fully autonomous mode has typically been included involving things such as search of a designated area [6], travel to a distant waypoint [8], or executing prescribed behaviors [9]. In studies in which robots did not cooperate and had varying levels of individual autonomy [4, 6, 8] (team size 2-4) performance and workload were both higher at lower autonomy levels and lower at higher ones. So although increasing autonomy in these experiments reduced the cognitive load on the operator, the automation could not perform the replaced tasks as well.

Table 1. Recent Studies on Multi-robots Research

Study	Task	World	Robots	Interaction
Fong et al. (2001): robotic system of collaborative control	Surveillance & reconnaissance	Real world with flat terrain.	2 UGVs (PionnerAT)	Dialog + waypoint control
Trouvain & Wolf (2002): user study of the impact of robot group size	Navigation	2D simulated office world	2, 4, 8 UGVs (homogeneous)	Waypoint
Trouvain et al. (2003): user study of map based and camera based user interface	Exploration	3D simulated outdoor world	1, 2, 4 UGVs (homogeneous)	Supervisory + waypoint control
Nielsen et al. (2003); Crandall et a (2005): user study of interaction schemes	Exploration	2D simulated office like world	3 UGVs (homogeneous)	Point to point and human snapper; Region of interest and sealing.
Olsen et al. (2004): Fan-out Independent study	Exploration	2D simulated office like world	18 UGVs (homogeneous)	Goal setting/plan level automation
	Exploration	Maze like real world	4 real robots (homogeneous)	Direction + collision, Goal setting auto exploration
Cummings & Mitchell (2005): Time management and scheduling	Attack target	2D simulated world	4 UAVs (homogeneous)	Manual, passive, active, and super active
Humphrey et al. (2006): user study of robot team visualization	Robot selection and position identification	2D simulated world	4 x 4 Agents (4 teams of 4)	None (No robot control)
Humphrey et al. (2006): user study of robot team visualization	Search	3D simulated outdoor world (USARSim)	6, 9 UGVs (heterogeneous) 6, 9 UGVs (heterogeneous)	Teleoperation and scripted behaviors
Wang et al. (2007), user studies of cooperating robot teams	Box pushing	3D simulated indoor world (USARSim)	2 UGV (homo/ Heterogeneous)	Manual, waypoint

Trouvain et al. [8] compared the navigation performance of two, four, and eight robots controlled by a single operator in a 2D simulated world. Under conditions of waypoint-based independent robot control and simple robot simulation, the experiment shows that the operator is able to control up to four robots, with the larger number of robots resulting in a higher workload and a stronger impact on the operator's monitoring ability than on the control ability. In a later study [10], the authors added autopilot capability to the robots, upgraded the world simulation to a 3D graph rendering system, and improved the simulated robot to a virtual fail-safe vehicle. The comparison of one, two, and four robots controlled during the exploration task shows that, in single robot control, human intervention improved performance but, when participants shifted to multi-robot (two or four robots) control, they "tend to reactive instead of proactive supervisory control" [10]. Overall performances under this latter condition were worse than those under the condition of full autonomy control. Again, increased human workload was found as the group size increased. In the most recent study [11], researchers used the UGV simulator, USARSim (also used in the current study), to compare robot control behaviors with six and nine independent heterogeneous robots. In this experiment, the high fidelity of the simulator, which introduces realistic SA problems, such as robot collisions, entanglement with objects in the environment, or falls makes the results potentially more generalizable to field robotics. The participants controlled the robots via teleoperation or prescribed behavior with a scalable interface. The results show that a higher number of robots caused a higher workload; however, the increment was less than the ratio of 1:1, which indicates that the operator continued to benefit from control over additional robots at least in the range covered in the study. The experimenters report improving SA with added robots contradicting the common belief that more robots lead to worse SA, however, because only the *attentional demand* scale of the 3D SART was significantly higher, more robots may have required greater effort without actually benefiting SA in other ways.

Multi-robot control appears to impact the human operator's workload in three distinct ways: (1) building and maintaining awareness, (2) making decisions, and (3)

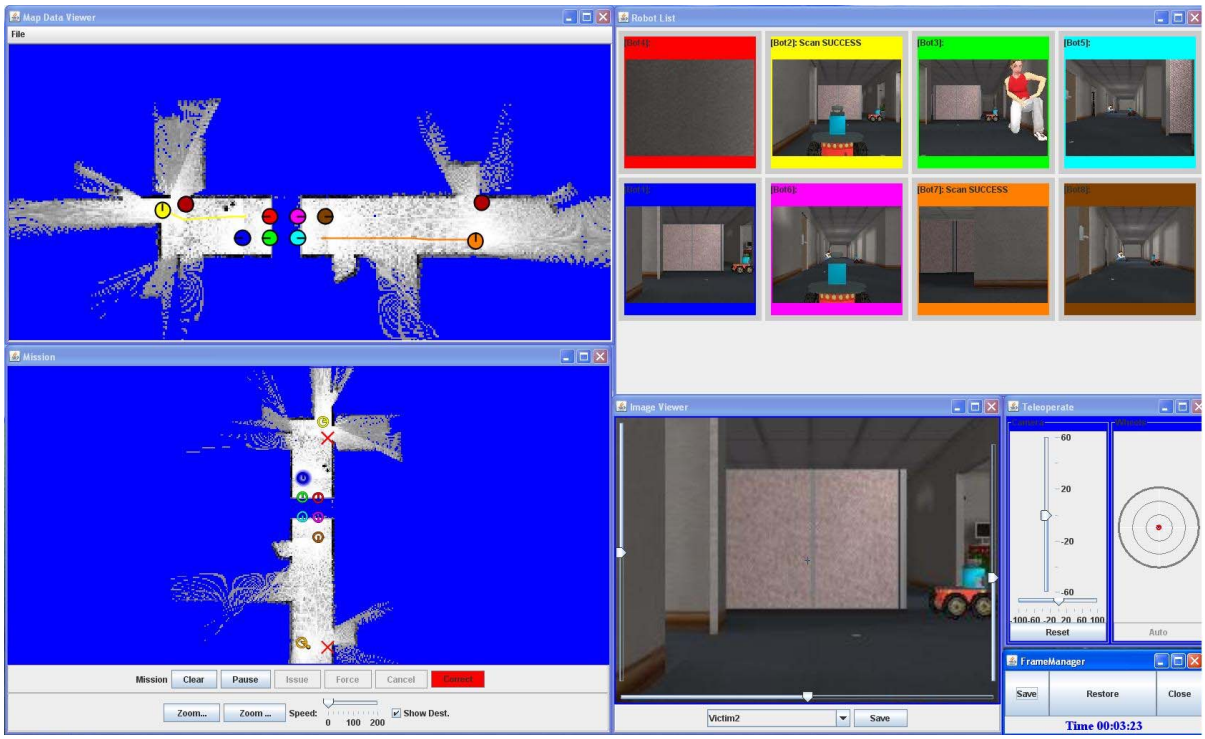

Figure 1. GUI for multi-robot control

controlling the system. Increasing the autonomy level in robotic system, whether providing decision support or individual robot autonomy, allows us to shift the decision-making and robot control workload from the human to the robotic system. On the other hand, increased robot autonomy may cause an increase in perception and decision-making workloads. Thus, there is a trade-off between the autonomy level of the robotic system and the level of human intervention.

The present study investigates the effect of the number of controlled robots on performance of an urban search and rescue (USAR) task using a realistic simulation. The USAR task imposes high mental workload in all three ways by requiring high SA to detect victims and mark them on the map, frequent decisions in switching between robots, and both waypoint control and teleoperation of individual robots. By covering a wider range of team sizes than [11] we were able to bound the number of robots that could be effectively controlled at the USAR task. Subsequent experiments are planned to distinguish the impacts of N of robots on SA (perceptual search with autonomous navigation) and navigation alone. The results of these studies are intended to aid system designers in allocating functions statically or dynamically between the operator and robot autonomy.

USARSIM AND MRCS

The reported experiment was performed using the USARSim robotic simulation with 4-12 simulated UGVs performing Urban Search and Rescue (USAR) tasks. USARSim is a high-fidelity simulation of urban search and rescue (USAR) robots and environments we developed as a research tool for the study of HRI and multi-robot coordination. USARSim supports HRI in ways lower fidelity simulations cannot by accurately rendering user interface elements (particularly camera video), accurately representing robot automation and behavior, and accurately representing the remote environment that links the operator's awareness with the robot's behaviors. USARSim can be downloaded from www.sourceforge.net/projects/usarsim and serves as the basis for the Virtual Robots Competition of the RoboCup Rescue League. The current version of USARSim includes detailed replicas of NIST USAR Arenas, as well as large-scale indoor and outdoor hypothetical disaster scenarios, and a large outdoor area along the Chesapeake Bay. USARSim complements these maps with high fidelity models of commercial (pioneer P2-DX, P2-AT, iRobot's ATRV Jr., Foster-Miller's Talon, and Telerob's Telemax) and experimental (PER from CMU, Zerg from University of Freiburg, Kurt 3D from University of Osnabruk) robots, including a snake (Soyu from Tohoku University), air (Air-robot helicopter) and a large Ackerman-steered UGV (Hummer) and sensor models for many common robotic sensing packages. USARSim uses Epic Games' UnrealEngine2 [12] to provide a high fidelity simulator at low cost. Validation studies showing agreement for a variety of feature extraction techniques between USARSim images and camera video are reported in Carpin et al. [13]. Other sensors including sonar and audio are also accurately modeled.

Validation data showing close agreement in detection of walls and associated Hough transforms for a simulated Hokuyo laser range finder are described in [14]. The current UnrealEngine2 integrates MathEngine's Karma physics engine [15] to support high fidelity rigid body simulation. Validation studies showing close agreement in behavior between USARSim models and real robots being modeled are reported in [16,17,18,19,20].

MrCS (Multi-robot Control System), a multirobot communications and control infrastructure with accompanying user interface developed for experiments in multirobot control and RoboCup competition [21] was used. MrCS provides facilities for starting and controlling robots in the simulation, displaying camera and laser output, and supporting inter-robot communication through Machinetta a distributed mutiagent system. Figure 1 shows the elements of the MrCS. The operator selects the robot to be controlled from the colored thumbnails at the top of the screen. To view more of the selected scene shown in the large video window the operator uses pan/tilt sliders to control the camera. Robots are tasked by assigning waypoints on a heading-up map on the Mission Panel (bottom right) or through a teleoperation widget (bottom left). The current locations and paths of the robots are shown on the Map Data Viewer (middle left).

II. Method

A large USAR environment previously used in the 2006 RoboCup Rescue Virtual Robots competition [21] was selected for use in the experiment. The environment was a maze like hall with many rooms and obstacles, such as chairs, desks, cabinets, and bricks. Victims were evenly distributed within the environment. A second simpler environment was used for training. The experiment followed a repeated measures design with participants searching for victims using 4, 8, and finally 12 robots. Robot starting points were varied over the three trials. Because our primary concern was with changes in performance as N robots increased, trials were presented in a fixed order. This design confounding learning effects and starting points with N of robots was adopted because the randomly selected starting points were sufficiently comparable not to bias results and any learning effect would attenuate rather than accentuate the expected decrements.

Participants

15 paid participants, 8 male and 7 female students were recruited from the University of Pittsburgh community. None had prior experience with robot control although most were frequent computer users.

Procedure

After collecting demographic data the participant read standard instructions on how to control robots via MrCS. In the following 20 minute training session, the participant practiced control operations and tried to find at least one victim in the training environment under the guidance of the experimenter. Participants then began three testing sessions (15 minute each) in which they performed the search task using 4, 8, and finally 12 robots. After each task, the participants were asked to complete the NASA-TLX workload survey.

III. Results and Discussion

Overall participants were successful in searching the environment in all conditions finding as many as 12 victims on a trial. The average number of victims found was 4.80 using 4 robots, 7.06 for 8 robots, but only 4.73 when using 12 robots. A paired t-test shows that in the 8 robots condition (R8) participants explored larger regions, $t(15) = -10.44$, $p = 0.000$, while finding more victims, $t(15) = -3.201$, $p = 0.003$, than using a 4 robot team (R4). On the other hand, adding four addition robots degraded performance with participants in the 8 robot condition (R8) exploring larger regions, $t(15) = -1.19$, $p = 0.059$, as well as finding more victims, $t(15) = -3.014$, $p = 0.005$, than they did using a 12 robot team (R12).

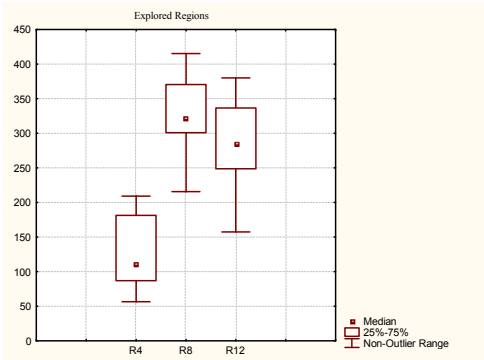

Figure 2. Explored Regions

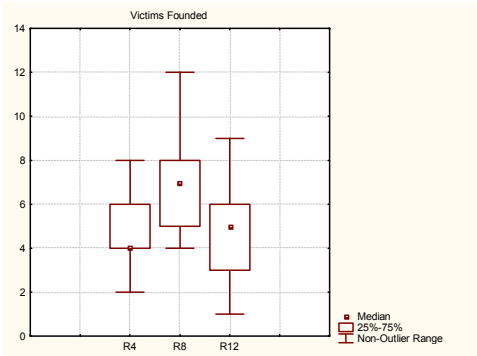

Figure 3. Victims Found

Figure 4 shows that as the number of robots is increased, fewer victims were found per robot. This measure should remain constant if robots were controlled to the same level of task at between 8-12. This can be seen for both the number of effectiveness. However, these differences were not

significant.

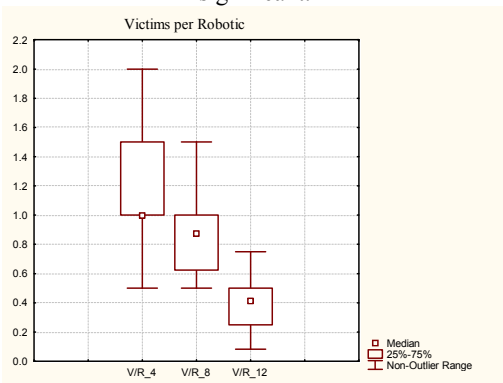

Figure 4. Victims Found per Robot

With increasing numbers of robots operators tended to neglect some robots either entirely or after an initial movement as shown in Table 2. A paired t-test indicates that participants neglected more robots in the 12 robot condition, t (15) = -1.922, p = 0.064, than under 4 robots team (R4). More robots were neglected after an initial move in the 8 robot (R8) condition t (15) = -2.092, p = 0.046, than for 4 robots (R4); and still more comparing a 12 robot team (R12) to the 8 robot (R8) condition t (15) = -3.761, p = 0.001.

Table 2. Neglected Robots

Number of Robots	R4	R8	R12
Totally	0.00	0.13	0.80
After the Initial Move	0.00	0.33	1.87

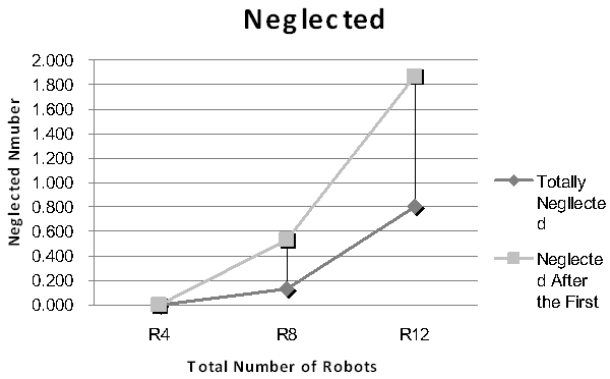

Figure 5. Neglected Analysis

As in earlier studies we found a positive relation between the number of times the operator switched between robots and the victims that were found. Higher switching rates are an indicator of shortened ITs or more efficient use of NT to service additional robots and hence should improve team performance. Figure 6 shows the number of switches observed under each of the three conditions. There were significant differences in number of switches between robots for the 4 robot and 8 robot conditions (t= -2.914, P<0.007) and the 4 robot and 12 robot conditions (t= -2.620, P<0.014). Similar results were found for numbers of missions (waypoint assignments) between the 4 robot and 8 robot condition (t= -3.079, P<0.005) and the 4 robot and 12 robot condition (t= -2.118, P<0.043).

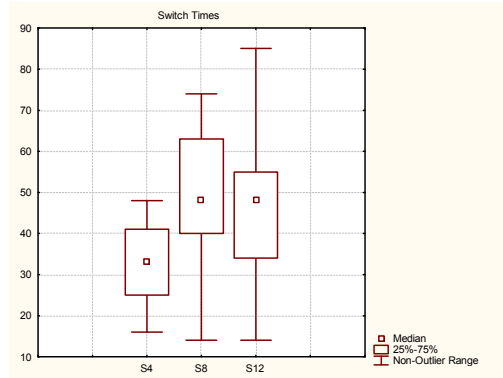

Figure 6. Number of Switches

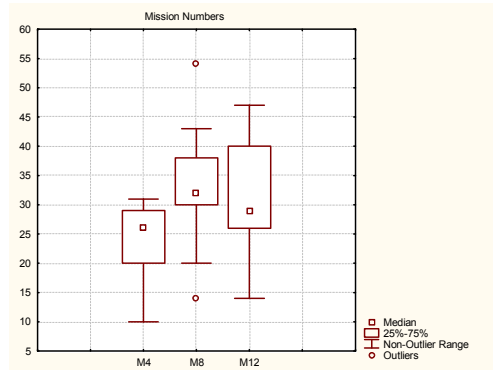

Figure 7. Mission Numbers

Table 3 Fan-out based on observed NT

Table 3 shows Fan-out for the three conditions estimated by

	IT	NT	Fan-out
R4	0.230079	0.769921	4.402437
R8	0.115556	0.884444	9.116028
R12	0.079661	0.920339	13.46839

$$Fan-out = \frac{NT}{IT} + 1$$

as proposed in [4].

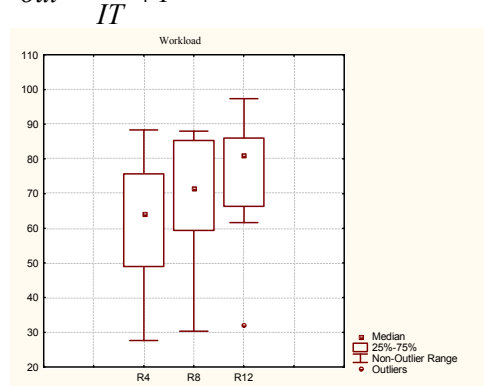

Figure 8. NASA-TLX Measurement of Mental workload

The result of the workload assessment indicates that workload increased with increasing numbers of robots to be victims found and the regions explored which improve between the four robot and eight robot conditions but decline

again between 8 and 12 robots. Determining Fan-out empirically as in [5] the Fan-out plateau (point at which performance is no longer improved by adding additional robots) lies somewhere within this region. The point at which operator capabilities become saturated can be estimated more closely by observing the number of robots that are either completely neglected or neglected after the first move. This number is approximately 2.7 for the 12 robot condition suggesting that the actual limit for this experiment is approximately 9 robots. An examination of the number of switches between robots supports this estimate because the number of switches is essentially the same in both the 8 and 12 robot conditions. This means that operators have reached their limit for interactions and are neglecting the robots for slightly longer times to accommodate the additional robot they are actively controlling. Similar conclusions can be drawn from the relation between number of missions (waypoint assignments) and regions explored. There is an increase in assigned missions between the 4 and 8 robot conditions that is accompanied by a substantial increase in the area explored. The decrease in mission numbers in the 12 robot condition is likewise reflected in a decline of the explored regions.

Although we do not have a direct measure of robot effectiveness such as AT (active time during which robot is moving) used in [5] using the estimate proposed in [4] Fan-out neatly parallels the number of robots operators were assigned to control. This suggests operators were using a satisficing strategy in which they attempted to distribute their attention approximately equally among the robots. This resulted in a lowering of the accepted standard of performance in order to accommodate the additional robots. This conjecture is borne out by examining per-robot performance which declines steadily across the three conditions.

With this initial experiment we are establishing a baseline for exploring the effects of robot team size on human performance. Even these early results suggest that the navigation component of the operator's task and the perceptual tasks involved in search may be somewhat differently impacted by increases in team size. We hope to be able to use such results to guide system designers in allocating functions statically or dynamically between the operator and robot (team) autonomy.

ACKNOWLEDGMENT

This research has been sponsored in part by AFOSR FA9550-07-1-0039, AFOSR FA9620-01-1-0542, ONR Grant N00014-03-1-0580 and L3-Communications (4500257512)

REFERENCES

[1] B. Gerkey and M. Mataric. A formal framework for the study of task allocation in multi-robot systems. International Journal of Robotics Research, 23(9):939–954, 2004.

[2] Miller, C. Modeling human workload limitations on multiple UAV control, Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, 526-527, 2004.

[3] Cummings, M. and Guerlain, S. An interactive decision support tool for real-time in-flight replanning of autonomous vehicles, AIAA Unmanned Unlimited Systems, Technologies, and Operations, 2004.

[4] J. W. Crandall, M. A. Goodrich, D. R. Olsen, and C. W. Nielsen. Validating human-robot interaction schemes in multitasking environments. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 35(4):438–449, 2005.

[5] D.R Olsen and S.B. Wood, *Fan-out: measuring human control of multiple robots*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2004, ACM Press: Vienna, Austria. p. 231-238.

[6] C. Nielsen, M. Goodrich and J. Crandall. Experiments in human-robot teams. In *Proceedings of the 2002 NRL Workshop on Multi-Robot Systems*, October 2003

[7] R. Parasuraman, S. Galster, P. Squire, H. Furukawa and C. Miller. A flexible delegation-type interface enhances system performance in human supervision of multiple robots: Empirical studies with roboflag. *Transactions on Systems, Man, and Cybernetics, Part A*, 35(4):

[8] B. Trouvain and H. Wolf. Evaluation of multi-robot control and monitoring performance. In *Proceedings of the 2002 IEEE Int. Workshop on Robot and Human Interactive Communication*, pages 111–116, September 2002.

[9] R.R. Murphy and J.L. Burke, Up from the Rubble: Lessons Learned about HRI from Search and Rescue, in *Proceedings of the 49th Annual Meetings of the Human Factors and Ergonomics* Society. 2005: Orlando.

[10] B. Trouvain, C. Schlick, and M. Mevert, Comparison of a map- vs. camera-based user interface in a multi-robot navigation task, in *Proceedings of the 2003 International Conference on Robotics and Automation*. 2003. p. 3224-3231.

[11] C.M. Humphrey, C. Henk, G. Sewell, B. Williams, J. A. Adams. *Evaluating a scaleable Multiple Robot Interface based on the UTFARSim Platform*. 2006, Human-Machine Teaming Laboratory.

[12] (UE 2) UnrealEngine2, http://udn.epicgames.com/Two/rsrc/Two/KarmaReference/KarmaUserGuide.pdf, accessed February 5, 2008.

[13] S. Carpin, T. Stoyanov, Y. Nevatia, M. Lewis and J. Wang. Quantitative assessments of USARSim accuracy". *Proceedings of PerMIS 2006*

[14] S. Carpin, J. Wang, M. Lewis, A. Birk and A. Jacoff. High fidelity tools for rescue robotics: Results and perspectives, *Robocup 2005 Symposium, 2005.*

[15] Mathengine, *MathEngine Karma User Guide*, http://udn.epicgames.com/Two/KarmaReference/KarmaUserGuide.pdf, accessed May 3, 2005.

[16] S. Carpin, M. Lewis, J. Wang, S. Balakirsky, C. Scrapper. (2006b). Bridging the gap between simulation and reality in urban search and rescue. *Robocup 2006: Robot Soccer World Cup X*, Springer, Lecture Notes in Artificial Intelligence, 2006.

[17] M. Lewis, S. Hughes, J. Wang, M. Koes, and S. Carpin. Validating USARsim for use in HRI research, *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, 2005, 457-461

[18] C. Pepper, S. Balakirsky, and C. Scrapper. Robot Simulation Physics Validation, *Proceedings of PerMIS'07*, 2007.

[19] B. Taylor, S. Balakirsky, E. Messina and R. Quinn. Design and Validation of a Whegs Robot in USARSim, *Proceedings of PerMIS'07, 2007*.

[20] M. Zaratti, M. Fratarcangeli and L. Iocchi. A 3D Simulator of Multiple Legged Robots based on USARSim. *Robocup 2006: Robot Soccer World Cup X*, Springer, LNAI, 2006.

[21] S. Balakirsky, S. Carpin, A. Kleiner, M. Lewis, A. Visser, J. Wang and V. Zipara. Toward hetereogeneous robot teams for disaster mitigation: Results and performance metrics from RoboCup Rescue, *Journal of Field Robotics*, 2007