Human Factors: The Journal of the Human Factors and Ergonomics Society

http://hfs.sagepub.com/

Choosing Autonomy Modes for Multirobot Search

Michael Lewis, Huadong Wang, Shih Yi Chien, Prasanna Velagapudi, Paul Scerri and Katia Sycara

Human Factors: The Journal of the Human Factors and Ergonomics Society 2010 52: 225 originally published online 11 May 2010

DOI: 10.1177/0018720810366859

SPECIAL ISSUE

Choosing Autonomy Modes for Multirobot Search

Michael Lewis, Huadong Wang, and **Shih Yi Chien,** University of Pittsburgh, Pittsburgh, Pennsylvania, and **Prasanna Velagapudi, Paul Scerri,** and **Katia Sycara,** Carnegie Mellon University, Pittsburgh, Pennsylvania

Objective: The number of robots an operator can supervise increases with the robots' level of autonomy. The reported study investigates multirobot foraging to identify aspects of the task most suitable for automation. **Background:** Many envisioned applications of robotics involve multirobot teams. One of the simplest of these applications is foraging, in which robots are operated independently to explore and discover targets. Depending on levels of autonomy and task, operators have been found able to manage 3 to 12 robots. **Method:** The foraging task can be functionally subdivided into visiting new regions and identifying targets. In the reported experiment, full-task foraging performance was compared with exploration and perceptual search performance for 4-, 8-, and 12-robot teams in a between-groups repeated measures design. **Results:** Operators in the full-task condition could not successfully manage 12 robots, finding only half as many victims as perceptual search operators. Exploration performance was roughly the same in the full-task and exploration conditions, suggesting that performance of this subtask was limiting the number of robots that could be controlled. **Conclusion:** Performance and workload measures indicate that exploration (navigation) tasks are the limiting factor in multirobot foraging. This finding suggests that robot navigation is the best candidate for automation. **Application:** Search tasks, such as foraging or perimeter control, account for many of the near-term applications envisioned for multirobot teams. The results support the choice of task-centered architectures in which the control and coordination of robotic platforms is automated, leaving search and identification of targets to human operators.

INTRODUCTION

Expected applications for multirobot systems (MrS), such as urban search and rescue (USAR) or cooperating uninhabited aerial vehicles, will require close coordination and control between human operator(s) and teams of robots. Controlling multiple robots substantially increases the complexity of the operator's task because attention must constantly be shifted among robots to maintain situation awareness (SA) and to maintain control. When task performance must be tightly coordinated, as, for example, in box pushing (Parker, 2000; Wang & Lewis, 2007) or formation following (Balch & Arkin, 1998), coordination needs to be automated because each action by a robot requires corresponding adjustments by the others, leading to cascading demands. Supervisory control approaches based on selection from among prespecified coordination plans, such as Playbook (Miller & Parasuraman, 2007) or Machinetta (Lewis, Sycara, & Scerri, 2009), appear best suited for these tasks.

There remain, however, a variety of human-robot interaction (HRI) tasks, such as approving targets or extricating robots from impasses, that need to be performed independently with individual robots. For these tasks, difficulty should correspond roughly with the number of robots being controlled. If one robot has the ability to explore area A in time T, for example, then a team of N independent robots should be able to explore an area of NA in time T. If these robots, however, required human input to perform their task, then performance would be limited by the human's ability to supply that

Address correspondence to Mike Lewis, University of Pittsburgh, Department of Information Science, 135 N. Bellefield Ave., Pittsburgh, PA 15260; ml@sis.pitt.edu. *HUMAN FACTORS,* Vol. 52, No. 2, April 2010, pp. 225–233. DOI: 10.1177/0018720810366859.

input. In the neglect tolerance model (Crandall, Goodrich, Olsen, & Nielsen, 2005), this situation is described as a sequence of control episodes in which the operator interacts with the robot for a period, *IT*, needed to raise its performance above some threshold, after which the robot can be neglected for a period, *NT*, until its performance deteriorates to a point at which the operator must again interact with it to restore an acceptable level of performance.

The number of *IT*s that can be fit into the *NT* interval plus 1 places an upper bound on fan-out, the number of robots that can be controlled to a fixed level of efficiency. The term *fan-out* was borrowed by Olsen and Wood (2004) from digital design, where it refers to the number of logic gates a gate can drive. Operators acting in accordance with this model should produce an additive increase in performance with number of robots until saturation is reached. After saturation, the curve should remain flat because the operator must ignore additional robots if interactions are to be of sufficient duration to return the robots to above-threshold performance. In empirical tests, such as those conducted by Olsen and Wood, measures of individual performance (such as area searched per robot) are observed to decline as size increases, causing team performance to follow a decelerating curve as robots are added rather than the linear increase with abrupt transition predicted by the neglect tolerance model.

In either case, the level at which fan-out is reached can be identified by the fan-out plateau (Olsen & Wood, 2004), the point of inflection after which additional robots cease to improve team performance. Finding an appropriate size for robotic teams, therefore, involves a trade-off between the decline in individual performance and improvements for the team at a particular level of automation. The present study investigates components of a search task to identify those components whose automation might most effectively increase fan-out and team performance.

In a foraging task (Cao, Fukunaga, & Kahng, 1997), robots independently search regions for targets of interest. Multirobot foraging was chosen for this experiment because search tasks, such as USAR (Burke, Murphy, Coovert, & Riddle, 2004) or perimeter patrolling (Endo, MacKenzie, & Arkin, 2004), are envisioned to be primary field applications for mobile robots. Teleoperation is currently the preferred mode for deployed robots performing foraging. For example, all 24 types of ground robots tested at a recent National Institute of Standards and Technology field trial to evaluate emergency response robots (Jacoff & Messina, 2007) were teleoperated. Researchers such as Endo et al. (2004), however, have demonstrated that ground robots can be successfully controlled through waypoints (a path designated through a series of points) over the types of uneven terrain expected for foraging tasks. Waypoint control was chosen for this experiment because it is a widely used form of control in experimental robotics and because it requires the lowest level of automation (Wang & Lewis, 2007) compatible with independent control of multiple robots.

Some variant of automation through waypoint control was used in each of the reviewed human-in-the-loop simulation studies investigating performance as a function of the size of mobile robot team (Humphrey, Henk, Sewell, Williams, & Adams, 2007; Olsen & Wood, 2004; Trouvain, Schlick, & Mevert, 2003; Trouvain & Wolf, 2002), with differences arising primarily in behavior after reaching a waypoint. As Olsen and Wood (2004) demonstrate, increased automation can drastically affect fan-out, extending it from two to nine in their study. These studies suggest that for foraging tasks involving waypoint control, the fan-out plateau lies somewhere between four and nine-plus robots, depending on the level of robot autonomy and environmental demands. Many of the tasks envisioned for robot teams, however, require larger numbers. One way to improve fan-out is to increase robot autonomy in ways that reduce the control demands on human operators.

The foraging task can be heuristically decomposed into exploration and perceptual search subtasks corresponding to navigation of an unknown space and searching for targets by inspecting and controlling onboard cameras. The current study investigates the scaling of performance with number of robots for operators performing either the full task or only one of the subtasks (exploration or perceptual search) to identify limiting factors.

For this comparison to be valid, the subtasks must be relatively independent in producing full task performance. The perceptual search task requires the operator to identify a victim from the video window and then locate corresponding features on the laser-generated map to mark the victim's location. Operators in the full-task condition have prior exposure to the situation shown in the camera view and map because the camera view, which allows discrimination among objects and is not limited in range, is typically used to decide where to go, which must then be entered as waypoints on the map.

Endsley and Kaber (1999) have demonstrated that automation of task execution, such as the exploration component of the search task, can lead to out-of-the-loop performance problems and loss of SA, the perception of environmental elements, comprehension of their meaning, and the projection of their status in the near future. In a study particularly relevant to robotic exploration, Peruch, Vercher, and Guthier (1995) determined that self-controlled viewers tended to develop rich survey knowledge more quickly than passive observers. For the current experiment, the accuracy as well as the number of victims located must be considered, because the premise predicts that perceptual-search participants should be less accurate in marking victims than full-task participants, who benefit from active involvement in navigation.

In the reviewed studies, two (Humphrey et al., 2007; Trouvain et al., 2003) extend automation to the exploration subtask, and one (Olsen & Wood, 2004) leaves both tasks to the operator. Fan-out values, however, fall in the same range across these experiments, suggesting that either exploration and perceptual search can be shared without cost or that differences in tasks make direct comparison of fan-out values between these studies invalid. In this study, we test the hypothesis that exploration and perceptual search do in fact interfere with one another.

Hypothesis 1: Exploration and perceptual search subtasks will interfere with one another, making full-task performance poorer than that for some subtask(s).

Because independent subtasks may have different fan-outs, then

1. fan-out is the same for full task and subtasks,
2. both subtasks have larger fan-out than full task, or
3. full task and one subtask have the same fan-out and the other subtask has larger fan-out.

In Case 3, we say that the subtask with the same fan-out as the full task is the *limiting* subtask, because if it were automated, the system would exhibit the larger fan-out of the other subtask. If there were a limiting subtask, it would likely be exploration, because operators can perceptually search only areas they have explored.

Hypothesis 2: Exploration will be a limiting subtask.

Because a larger-fan-out subtask will consume fewer cognitive resources for a given team size, it will be reported as less effortful.

Hypothesis 3: Perceptual search will have lower workload ratings than full task or exploration.

METHOD

Participants

For the study, 45 paid participants, 25 male and 20 female, were recruited from the University of Pittsburgh community and were assigned proportionately to experimental conditions. Their ages ranged from 20 to 33, with mean age 25.18 and standard deviation 2.63. None had prior experience with robot control, although most were frequent computer users. One participant in the exploration condition failed to complete the experiment, and so these data have been excluded from the analysis.

Apparatus

The USARSim (Lewis, Wang, & Hughes, 2007) robotic simulation was used with 4 to 12 simulated unmanned ground vehicles (UGV) performing USAR foraging tasks. USARSim is a high-fidelity simulation of USAR robots and environments developed as a research tool for the study of HRI and multirobot coordination. MrCS (Multi-robot Control System), a multirobot communications and control infrastructure with accompanying

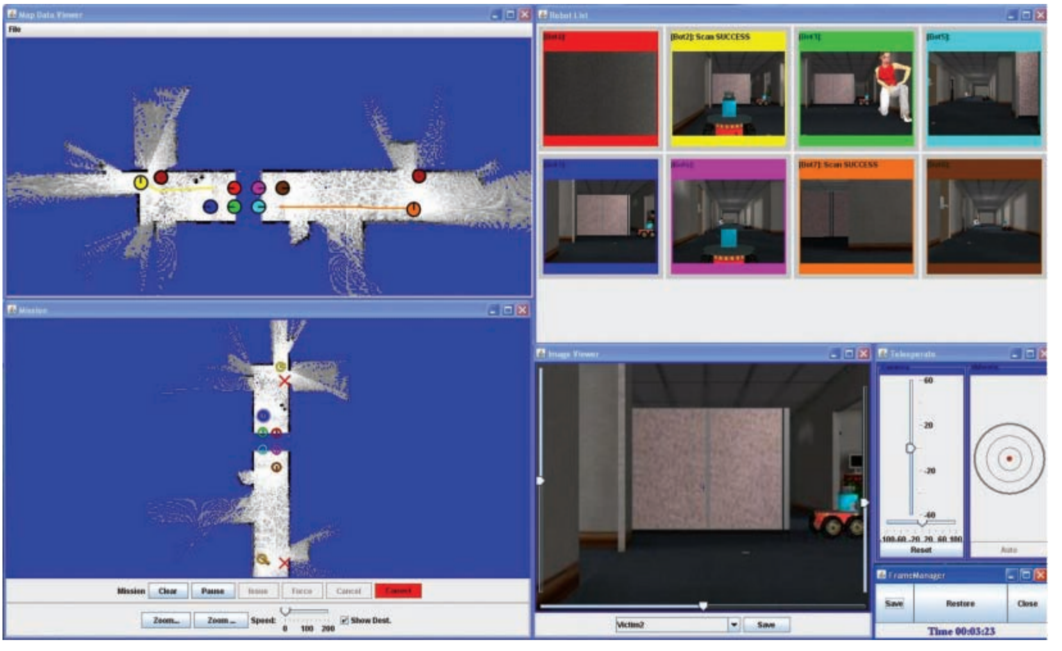

Figure 1. Multi-robot Control System graphical user interface.

user interface developed for experiments in multirobot control and RoboCup competition (Balakirsky et al., 2007), was used in these experiments. MrCS provides facilities for starting and controlling robots in the simulation, displaying camera and laser range finder output, and supporting interrobot 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 and 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). Because the maps show only the regions within the laser's range, the operator needs to use the camera view to choose where to go and the map to mark the waypoints.

Experimental Task

A large USAR environment previously used in the 2006 RoboCup Rescue Virtual Robots competition (Balakirsky et al., 2007) was selected for use in the experiment. The environment was a mazelike hall with many rooms and obstacles, such as chairs, desks, cabinets, and bricks. Victims, unobscured human figures, were evenly distributed within the environment. A simpler environment was used for training.

The experiment followed a between-groups repeated measures design with number (4, 8, 12) of robots defining the repeated measure and control regime defining the groups. Participants in the full-task condition performed the complete USAR task. In the subtask conditions, they performed variants of the USAR task requiring only exploration or perceptual search. Because of the large differences in difficulty between controlling small (4) and large (12) teams, we chose a fixed ordering to prepare participants to perform the more difficult tasks at a level more nearly approaching their asymptote. The comparisons between full- and part-task groups remain valid because they are between "equally trained" participants at each team size. The estimate of fan-out remains appropriate because although participants controlling larger teams have more practice, the decline or leveling off of performance still indicates a resource limitation for which training should have a countervailing effect.

The analysis requires that the three conditions be equivalent in terms of the area searched by participants. In the full-task condition, operators used waypoint control to explore an office-like environment. When victims were detected by the onboard cameras, the robot was stopped and the operator marked the victim on the map and returned to exploration. Equating the exploration subtask was relatively straightforward. Operators were given the instruction to explore as large an area as possible with coverage judged by the extent of the laser range finder–generated map.

Developing an equivalent perceptual search condition was more complicated. The perceptual search operator's task resembles that of the payload operator for an unmanned aerial vehicle. or a passenger in a train, in which there is no control over the platform's trajectory but the operator can only pan and tilt the cameras to find targets. The targets the operator has an opportunity to acquire, however, depend on the trajectories taken by the robots. If an autonomous path planner were used, the robots would explore continuously and, depending on the algorithm, might cover a wider area than when operated by a human (where pauses typically occur after arrival at a waypoint). If human-generated trajectories are taken from the full-task condition, however, they would contain additional pauses at locations where victims were found and marked, providing an undesired cue. Instead, we have chosen to use trajectories from the exploration condition, because they should contain pauses associated with waypoint arrival but not those associated with identifying and marking victims. As a final adjustment, operators in the perceptual search condition must be able to pause their robots to identify and mark the victims they discover. A typical trajectory was selected for each team size for use in the perceptual search condition.

As noted, trajectories used in the perceptual search condition (labeled *per*) were selected from exploration participant data to match performance found in the full-task condition. *Potential victims* refers to victims located within the laser range finder–generated map. Because the range finder has a 180° field of view, whereas the robot's camera is limited to a realistic 45°, a victim may lie within the map without ever having appeared on a robot's camera.

Procedure

After collecting demographic data, the participant read standard instructions on how to control robots via MrCS. In the following 20-min training session, participants in the full-task and exploration conditions practiced control operations. Participants in the full-task and perceptual search conditions were encouraged to find and mark at least one victim in the training environment with the guidance of the experimenter. Participants then began three testing sessions (15 min each) in which they performed the search task using 4, 8, and finally 12 robots. After each task, the participants were asked to complete workload ratings on the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988).

RESULTS

Data were analyzed using a repeated-measures ANOVA comparing full-task performance with that of the subtasks. When measures were inappropriate for some subtask (perceptual search operators do not navigate, exploration operators do not find or mark victims), comparisons are pairwise rather than tripartite. Results are reported on the basis of an alpha of .05. Overall, full-task participants were successful in searching the environment at all team sizes, finding as many as 12 victims on a trial. Because operators sometimes marked victims twice or became confused about the in-focus robot's location and marked victims in a wrong room, victims were treated as found only if they were located within 2 m of a marked location. When multiple markings were found, only the most accurate was retained. Participants were not told of this scoring procedure, so indiscriminate marking was not encouraged.

For victims marked within 2 m in the full-task condition, the average number of victims found was 4.8, $\sigma = 1.57$, with the use of 4 robots; 7.06, $\sigma = 2.25$, with 8 robots; but dropped back to 4.73, $\sigma = 1.98$, with 12 robots. Participants in the perceptual search condition, by contrast, were significantly more successful, between-groups $F(1, 28) = 27.4$, $p < .001$, finding 6.93 (σ = 1.22), 8.2 (σ = 0.86), and 8.33 (σ = 0.62) victims, respectively. As shown in Figure 2, full-task performance peaked for full-task participants at 8 robots, then declined, resulting in a significant *N* Robot × Task interaction, within-groups $F(2, 56) = 8.45$, $\eta^2_p = .23$, $p = .001$. Adjusted pairwise comparisons found differences between 4 and 8 robots ($p < .001$) and between 8 and 12 robots ($p = .006$), but the difference between 4 and 12 ($p = .08$) did not reach significance.

Participants in the perceptual search condition were significantly more precise than those in the full-task condition.

Comparing accuracy in marking victims directly, root mean square error shows that participants in the perceptual search condition were significantly more precise, between-groups $F(1, 28) = 38.37$, $p < .001$, than full-task participants, although accuracy in both groups declined for an increasing number of robots, within-groups $F(1.44, 40.43) = 41.837$, $p < .001$, Greenhouse-Geisser corrected.

As Figure 3 shows, coverage for the full-task and exploration conditions was very similar at 4 and 8 robots but diverged at 12 robots, between-groups $F(1, 27) = 11.43$, $p = .002$; Within Groups × *N* Robots $F(2, 54) = 4.15$, $p = .021$. Post hoc Bonferroni corrected pairwise comparisons showed increases in area covered between 4 and 8 robots ($p < .001$) but no difference between 8 and 12 for full-task and perceptual search participants. Paired *t* tests found significant differences in coverage, favoring exploration participants at 4 robots (exploration, $\bar{X} = 173.8$, $\sigma = 52.2$; full task, $\bar{X} = 129.2$, $\sigma = 48.48$), $t(15) = 4.25$, $p = .024$, and 12 robots (exploration, $\bar{X} = 377.8.8$, $\sigma = 48$; full task, $\bar{X} = 285.8$, $\sigma = 60$), $t(15) = –3.62$, $p = .001$, but no difference at 8 robots (exploration, $\bar{X} = 348.9$, $\sigma = 73.9$; full task, $\bar{X} = 327.6$, $\sigma = 55.4$), showing that full-task and perceptual search participants had very similar performance up to 8 robots and that perceptual search participants maintained performance at 12 robots whereas those in the full-task condition declined.

As Figure 4 shows, subjective workload as reported on the overall scale of the NASA-TLX increased monotonically in all groups,

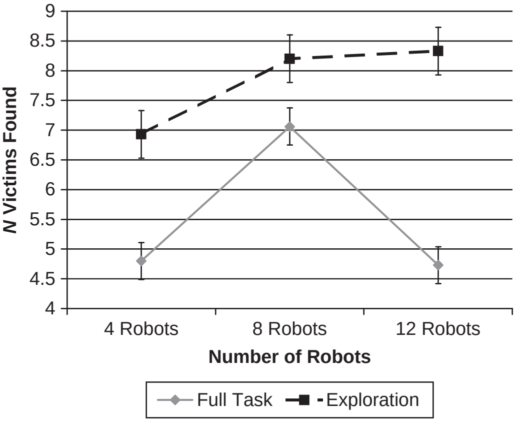

Figure 2. Victims found as a function of N robots (2 m). Bars show standard errors.

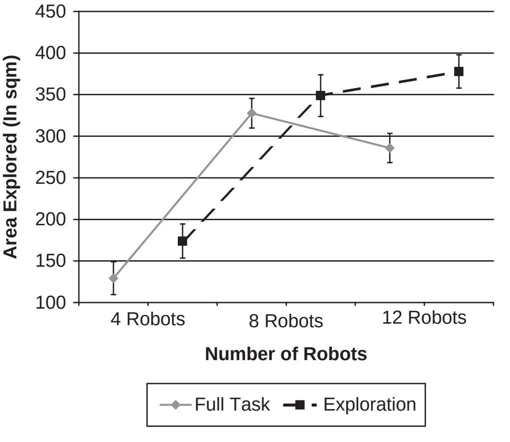

Figure 3. Area explored as a function of N robots. Bars show standard errors.

within-participants $F(1.55, 63.35) = 27.347$, $p = .0001$, but differed between groups, between-groups $F(2, 41) = 9.554$, $p = .001$. Post hoc Tukey HSD tests show perceptual search workload ($\bar{X} = 62.44$, $\sigma = 12.5$) to be significantly lower than either the exploration ($\bar{X} = 64.79$, $\sigma = .12.5$, $p = .008$) or full-task ($\bar{X} = 69.244$, $\sigma = 12.5$, $p < .001$) conditions. Switching focus among robots increased with team size, within-participants $F(2, 82) = 21.856$, $p < .001$, with pairwise comparisons showing an increase between 4 and 8 robots ($p < .001$) but none between 8 and 12.

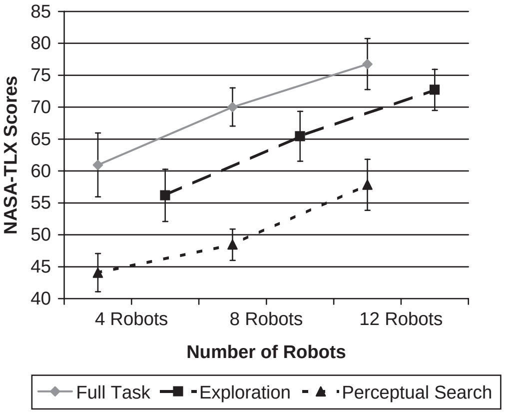

Figure 4. Workload as a function of N robots. Bars show standard errors. NASA-TLX = NASA task load index.

DISCUSSION

The purpose of this experiment was to examine the tasks that go into controlling foraging robots to identify subtasks that might benefit most from increased automation. Rather than looking for cognitive tasks that go into controlling robot teams (Crandall & Cummings, 2007; Olsen & Wood, 2004), we have adopted a naive model of the process by identifying two isolatable activities (each consisting of many cognitive tasks) that could be separated for an experiment and potentially automated.

Our data suggest that there is some cost of concurrence for performing the exploration and perceptual search tasks, given that both victims found (perceptual search) and area explored (exploration) were performed better alone than in the full-task condition. In particular, for both measures, performance at the subtasks was maintained between 8 and 12 robots, whereas both measures for full-task performance decreased. An examination of Figures 2 and 3, however, shows the difference to be more pronounced for finding victims (perceptual search) than for area explored (exploration). In fact, Figure 3 shows very similar exploration performance in exploration and full-task conditions that diverges only at 12 robots, when full-task operators exceed their capacity. Although there is no direct way to compare improvements in perceptual search and exploration, examination of Figures 2 and 3 and comparison of η^2_p values (accounting for variance) for the N Robots × Subtask interaction shows that improvements in perceptual search ($\eta^2_p = .35$) contribute more to victims found than improvements in exploration ($\eta^2_p = .13$) do to area explored.

These results support automation of path planning and navigation versus efforts to improve automation for target recognition and cueing, provided that the technical challenges are comparable. Because robots searched comparable areas in the full-task and perceptual search conditions, the larger number of victims found in the perceptual search condition represents a more thorough search and indicates that full-task participants were missing victims they should have found. Because avoiding missed targets is crucial to many foraging tasks, such as demining or search and rescue, this advantage may be more important than other performance gains, such as widening the search area.

The finding of distinctly lower workload ratings for perceptual search is also encouraging in that it suggests that there may remain reserve capacity to monitor additional robots or perform other tasks. The top of the curve for perceptual search participants between 8 and 12 robots may have been artificially low because of a ceiling on the number of targets that could be observed given the limitations in trajectories. Although suboptimal trajectories were needed to match the trajectories of other experimental conditions, an MrS involving autonomous exploration could avoid the pauses and overlaps of our user-generated trajectories.

The concern that the passive nature of the perceptual search task might lead to reduced SA was not borne out by these data. In fact, accuracy in marking victims, our best measure of operators' SA, was significantly better for perceptual search operators. We believe that the conventional advantages active navigation offers to SA (Peruch et al., 1995) may be negated in multirobot control because of the operator's need to constantly switch between robots. The disorientation attendant to these switches and difficulties in keeping track of landmarks across constantly shifting viewpoints may make memory of past trajectories of little use in assembling a picture of the environment. Under these conditions, the additional load imposed by planning paths for multiple robots appears to overwhelm any advantage to SA the activity might have offered.

An additional benefit from automating exploration would be the shift from platform-centric to network-centric (Alberts, Garstka, & Stein, 1999) control. Given that the objective of foraging is to search a prescribed area, navigation is only instrumental. Because the operator must continually shift between viewpoints anyway, automating the task of navigating lawnmower, successive sweeps through rows or columns of a grid, or other search routes frees the operator to devote greater attention to the camera views being obtained.

The image of an MrS operator sitting in front of a bank of screens like a security guard monitoring surveillance cameras raises concerns about SA. Operators might be able to detect targets but be unable to relate these detections to the overall search. This was not the case in our data, however, because victims could be scored only by locating them on a map, and perceptual search participants were best at this task. These results are encouraging for the future of large-scale MrS in foraging because they suggest that fan-out could be increased by autonomous exploration. Because foraging robots can be controlled relatively independently, very large teams could be formed simply by the addition of operators, each with larger fan-out.

Other forms of automation, such as self-diagnosis or self-reflection, could lead to alarm-based control strategies that might extend fan-out even farther. Although our results do not directly address more complex cooperative tasks, such as assembly or construction, our participants' ability to maintain SA without direct control over exploration behaviors indicates that they may be able to successfully monitor and direct large teams of indirectly controlled cooperating robots as well.

ACKNOWLEDGMENTS

This research has been sponsored in part by Air Force Office of Scientific Research FA9550-07-1-0039, Air Force Office of Scientific Research FA9620-01-1-0542, and Office of Naval Research Grant N00014-03-1-0580.

REFERENCES

Alberts, D., Garstka, J., & Stein, F. (1999). *Network centric warfare: Developing and leveraging information superiority*. DoD C4ISR Cooperative Research Program, Washington, DC. Retrieved from http://www.dodccrp.org/files/Alberts_NCW.pdf

Balakirsky, S., Carpin, S., Kleiner, A., Lewis, M., Visser, A., Wang, J., & Zipara, V. (2007). Toward hetereogeneous robot teams for disaster mitigation: Results and performance metrics from RoboCup Rescue. *Journal of Field Robotics*, *24*, 943–967.

Balch, T., & Arkin, R. (1998) Behavior-based formation control for multi-robot teams. *IEEE Transactions on Robotics and Automation*, *14*, 926–939.

Burke, J., Murphy, R., Coovert, M., & Riddle, D. (2004). Moonlight in Miami: An ethnographic study of human-robot interaction in USAR. *Human-Computer Interaction*, *19*(1/2), 85–116.

Cao, Y., Fukunaga, A., & Kahng, A. (1997). Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots*, *4*, 1–23.

Crandall, J., & Cummings, M. (2007). Developing performance metrics for the supervisory control of multiple robots. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction* (pp. 25–32). New York: Association for Computing Machinery.

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

Endo, Y., MacKenzie, D., & Arkin, R. (2004). Usability evaluation of high-level user assistance for robot mission specification. *IEEE Transactions on Systems, Man, and Cybernetics*, *34*, 168–180.

Endsley, M., & Kaber, D. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*, 462–492.

Hart, S., & Staveland, L. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, Netherlands: Elsevier.

Humphrey, C., Henk, G., Sewell, B., Williams, B., & Adams, J. (2007). Assessing the scalability of a multiple robot interface. In *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robotic Interaction* (pp. 239–246). New York: Association for Computing Machinery.

Jacoff, A. S., & Messina, E. R. (2007). Urban search and rescue robot performance standards: Progress update. In G. Gerhart, D. Gage, & C. Shoemaker (Eds.), *Proceedings of SPIE Unmanned Systems Technology IX* (Vol. 6561, No. 65611L). Bellingham, WA: SPIE.

Lewis, M., Sycara, K., & Scerri, P. (2009). Scaling up wide area search munition teams. *IEEE Intelligent Systems*, *24*(3), 10–13.

Lewis, M., Wang, J., & Hughes, S. (2007). USARSim: Simulation for the study of human-robot interaction. *Journal of Cognitive Engineering and Decision Making*, *1*(1), 98–120.

Miller, C., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors*, *49*, 57–75.

Olsen, D., & Wood, S. (2004). Fan-out: Measuring human control of multiple robots. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 231–238). New York: Association for Computing Machinery.

Parker, L. (2000). Lifelong adaptation in heterogeneous multi-robot teams: Response to continual variation in individual robot performance. *Autonomous Robots*, *8*, 239–267.

Peruch, P., Vercher, J., & Guthier, G. (1995). Acquisition of spatial knowledge through visual exploration of simulated environments. *Ecological Psychology*, *7*(1), 1–20.

Trouvain, B., Schlick, C., & Mevert, M. (2003). 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* (pp. 3224–3231). Piscataway, NJ: IEEE.

Trouvain, B., & Wolf, H. (2002). Evaluation of multi-robot control and monitoring performance. In *Proceedings of the 2002 IEEE International Workshop on Robot and Human Interactive Communication* (pp. 111–116). Piscataway, NJ: IEEE.

Wang, J., & Lewis, M. (2007). Assessing coordination overhead in control of robot teams. In *Proceedings of 2007 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 2645–2649). Piscataway, NJ: IEEE.

Michael Lewis is a professor in the School of Information Sciences at the University of Pittsburgh in Pittsburgh, Pennsylvania. He received his PhD in psychology from Georgia Institute of Technology in Atlanta in 1986.

Huadong Wang is a doctoral student in the School of Information Sciences at the University of Pittsburgh in Pittsburgh, Pennsylvania. He received his ME in industrial engineering from Tsinghua University in Beijing, People's Republic of China, in 2006.

Shih Yi Chien is a doctoral student in the School of Information Sciences at the University of Pittsburgh. He received his MSIS in information sciences from the University of Pittsburgh in Pittsburgh, Pennsylvania, in 2009.

Prasanna Velagapudi is a doctoral student at the Robotics Institute at Carnegie Mellon University in Pittsburgh, Pennsylvania. He received his BE in computer science and engineering from Carnegie Mellon University in 2005.

Paul Scerri is a systems scientist at Carnegie Mellon University in Pittsburgh, Pennsylvania. He received his PhD in computer science from Linkoping University in Linkoping, Sweden, in 1999.

Katia Sycara is a professor at the Robotics Institute at Carnegie Mellon University in Pittsburgh, Pennsylvania. She received her PhD in computer science from Georgia Institute of Technology in Atlanta in 1987.

Date received: April 30, 2009
Date accepted: January 25, 2010