Scaling Human-Robot Systems

Prasanna Velagapudi
Robotics Institute
Carnegie Mellon University
pkv+@cs.cmu.edu

Paul Scerri
Robotics Institute
Carnegie Mellon University
pscerri+@cs.cmu.edu

INTRODUCTION

Exciting applications are emerging that involve large, heterogeneous human-robot teams acting in complex environments. Examples include search and rescue [5], disaster response [12], and military applications [4]. Robots are capable of augmentation and force multiplication of human assets, providing superhuman perception, coverage, and mobility, without risking human life.

In these domains, the advantages of robotic teams lie primarily in their ability to cover large areas quickly, by coordinating and parallelizing their efforts. However, when equipped with the latest in multi-spectral imaging, 3D LIDAR, and RF signal analysis, these platforms can also accumulate vast amounts of information to process, reaching gigabits per second [11]. Available communication hardware is ill-equipped to handle such high data rates simultaneously from many sources, and the problem is only exacerbated by power limitations and the overhead of wireless network protocols, especially in remote and military settings [9]. In addition, in many of these domains, robots can be partially or sporadically connected, meaning that information must be buffered and relayed through other robots or other communications hubs. Overall, information from robots is bandwidth-constrained and latency can vary greatly.

Under these conditions, the naïve but common architecture of humans as overarching controllers quickly breaks down. Humans are a necessarily centralized resource, and it is impractical and often impossible to simply pipe all necessary data for them to constantly control and correct autonomous robots. In addition, human attention and workload are tightly constrained resources, and the cognitive costs associated with rapid switching and real-time monitoring weigh heavily on operators, reducing their ability to effectively deal with many robots [7].

On the other hand, fully autonomous teams are deficient in several major areas, making human operators critical and likely to remain so in the near future. Human operators have the advantage of a large database of meta-knowledge to augment their reasoning skills. The ability to make use of this knowledge to perform inference is often necessary to interpret the data collected by robotic agents and convert it into useful hypotheses for the team. For example, in wilderness search and rescue, victims or signs of life can be extremely hard to spot directly, even by humans. Occlusions and unstructured terrain in the visual field limit visibility and add visual noise, e.g. Figure 1. However, humans are exceptionally good at spotting clues that are applicable in context. Patterns left by victims, potential obstacles, anomalous shapes or structures–the ability to notice these things as significant to the task is extremely difficult to codify, but trained humans excel at it.

Figure 1. A wide range of unstructured terrain types complicates the interpretation of UAV video feeds.

The high expense, but necessity, of human operators means that efficient human-robot teams must focus on letting these operators focus on the tasks that they do best, ones that autonomous systems cannot or should not do, while relegating the rest to robot autonomy. It is important to distinguish this from simply giving the autonomy components the “easiest” tasks: certain tasks, such as coordinated navigation, can be quite difficult for humans [16], but have many reasonable autonomous solutions [2].

Tasks relegated to humans face the additional cost of moving information to and from human operators. Unlike autonomous software, humans cannot easily be moved between platforms or decentralized. The information they need to make their decisions must be shipped to them, and their de-

CHI 2009, April 3 - 9, 2009, Boston, MA, USA.

cisions must be propagated back over the network. If humans are to handle more than one robot, this further restricts the types of tasks that humans can reasonably be expected to perform. The information used to do the tasks must be low-bandwidth enough to fit across robot communication channels, and the robots that require the response must be tolerant of the latency of moving the information to the human, having them cognitively process it, and having it sent back out across the network. These restrictions are key drivers in the design of large scale human-robot systems.

TASKS UNDER NETWORK CONSTRAINTS

Traditional models of teleoperation are simply impractical in teams with more robots than humans. Providing live data streams and expecting real-time responses from operators is unrealistic in systems that cannot deliver that information and where operators have to monitor many vehicles. Instead, human tasks need to be described as atomic operations that include the data necessary for sufficient situational awareness and reasoning. Tasks must be tolerant of latencies and dependencies on information must be carefully tracked. This question, of how to recast the domain-specific human tasks in a manner that allows them to be distributable and atomic, is an area for research.

Scaling control

We have studied the effect of the number of controlled robots on performance of an urban search and rescue (USAR) task using the high fidelity USARSim robotic simulation environment [3]. Participants controlled either 4, 8, or 12 robots, both dictating the robots paths and controlling their cameras to search for victims. The results established a significant drop in human performance between controlling 8 and 12 robots (Figure 2). Task performance increased in going from four to eight controlled robots but deteriorated in moving from eight to twelve while workload measures increased monotonically with number of robots. Overall, performance per robot decreased with increases in team size. These results are consistent with earlier studies suggesting a limit of between 8-12 robots for direct human control. Our results demonstrate that these findings generalize to a more realistic setting and complex task [14].

Task decomposition

Building off of the scaling control work, we have also investigated the effects of decomposing the USAR problem into two subtasks. In the perceptual search task, participants search for victims by controlling cameras mounted on robots following autonomous paths. In the exploration task, participants directed the team of robots in order to explore as wide an area as possible. By decomposing the search and rescue task into exploration and perceptual search subtasks it is possible to individually assess their scaling characteristics in order to provide a basis for tentative task allocations among humans and automation for controlling larger robot teams. Our data suggest that there is some cost of concurrence for performing the exploration and perceptual search tasks, as individually, performance in these tasks continued to increase between 8 and 12 robots, while performance in both tasks together caused performance to drop between 8

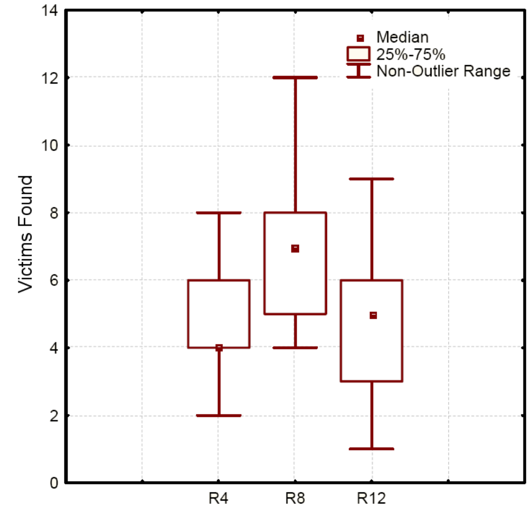

Figure 2. In a simulated search and rescue task, team performance decreased as number of robots increased from 8 to 12.

and 12 robots, echoing the results of the scaling control experiments. Interestingly, the workload ratings for perceptual search were significantly lower than for exploration or the full task, suggesting that operators may have reserve capacity for monitoring additional robots in this type of perceptual task. This suggests that focusing autonomy on exploration tasks may allow human operators to handle larger team sizes [16].

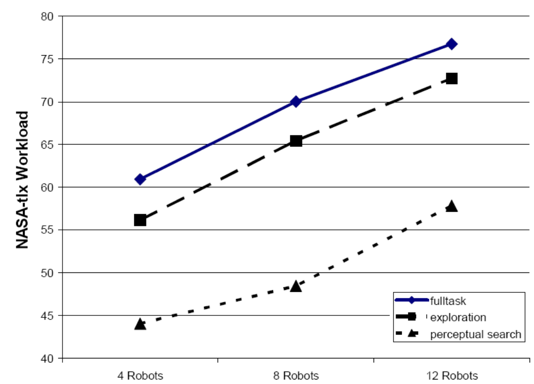

Figure 3. Workload measurements suggest operators may be able to handle more robots while only dealing with perceptual search than while only dealing with exploration in simulated search and rescue.

Asynchronous imagery

Camera guided teleoperation has long been the preferred mode for controlling remote robots, with other modes such as asynchronous control only used when unavoidable. We evaluated the usefulness of asynchronous operation for a multi-robot search task. Because controlling multiple robots places additional demands on the operator, we hypothesized that removing the forced pace for reviewing camera video might reduce workload and improve performance. Furthermore, using still images with no latency requirements greatly minimized network constraints. Participants operated four robot teams performing a simulated urban search and rescue (USAR)

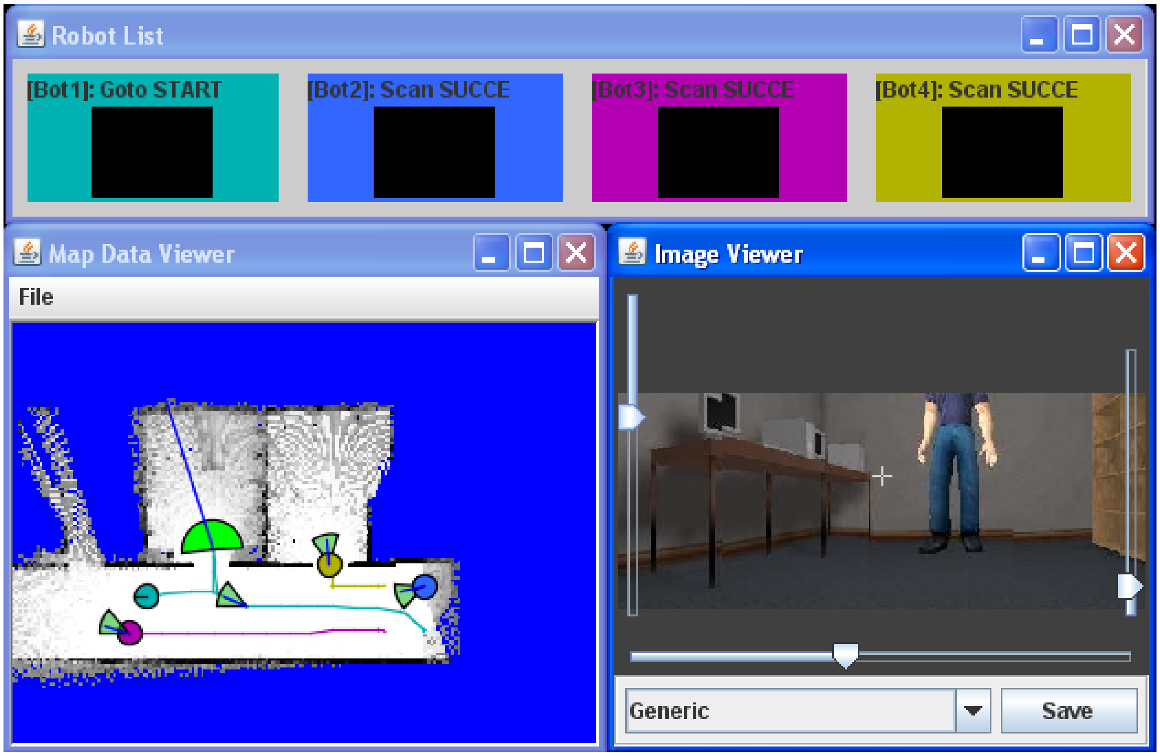

Figure 4. Asynchronous map-registered imagery was shown to yield comparable performance to streaming video in a simulated search and rescue task. Here the large semicircle on the map at left shows the position and field of view of the image on the right. The smaller wedges indicate other images, and the circles are robots exploring the map.

task using either conventional streaming video and a map interface or an experimental interface without streaming video but with the ability to store panoramic images in the map interface to be viewed at leisure. Results showed that search performance was somewhat better using the conventional interface, however, ancillary measures suggested that the asynchronous interface succeeded in reducing temporal demands for switching between robots [15].

HUMAN INFORMATION NEEDS

As team size increases, it becomes increasingly necessary for robots to be selective in the information shared with human operators. Decentralized information sharing methods such as BDDF [8] and distributed particle filters [10] combine fusion with Bayesian frameworks to reduce computational load, while decision-theoretic approaches [6] attempt to reason about the utility of propagating information to particular robots. However, for human operators, it is unclear what the utility of pieces of information are, and fusing information prematurely can hide important features of the source data that humans may be able to identify as anomalous. Thus, there remains the question of determining if humans need particular pieces of information, both before and after data fusion.

Cognitive modeling

Building off of our USAR work, we are currently adapting our multi-robot control system so that we can redirect operator control information to and from ACT-R cognitive models [1]. This will enable us to build models from operation data captured during human user studies, and test those models by having them feed operator decisions back into the control system. These cognitive models have the potential to predict subtle scaling and workload effects that previously would not be discernible from user study data.

Environmental factors

As robots become smaller and more portable, it becomes increasingly common to see them and their operators deployed directly in the field. In the field, a number of external cues are available that are not present for remote teleoperators. Operators can often see or hear the robot itself or landmarks that appear in feedback from the robot. One of our current areas of research is studying the effects of field deployment on operator performance, specifically in the wilderness search and rescue domain. Operators are tasked with identifying visual targets in video feeds from UAV flights both in the field and in controlled lab settings. Differences in performance may identify significantly helpful or harmful environmental effects.

Utility-based information sharing

Finally, we are exploring methods of information sharing amenable to the dynamic, heterogeneous information that humans require. We focus on simple, probabilistic routing strategies that depend only on a mathematical sense of utility, which we hope to bridge with our cognitive modeling work. Our results to date show that simple stochastic strategies using some easily obtainable probabilistic information about the team can achieve reasonable overall belief sharing performance. Specifically, by collectively estimating the value of a piece of information, the team can make more efficient use of its communication resources [13].

CONCLUSION

It is important to address these issues in human-robot interaction before they become limiting factors in modern multi-robot platforms. Already, many autonomous systems are limited by the need for human supervision, and the trends in research and industry are heading towards larger and more complex teams. By eliminating existing reasons for human supervision and consolidating and decoupling the remainder from real-time network constraints, the need for skilled human operators can be greatly reduced, opening the floodgates for practical, cost-effective applications of multi-robotic technology across a wide range of domains.

REFERENCES

1. J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Quin. An integrated theory of mind. *Psychological Review*, 111:1036–1060, 2004.
2. W. Burgard, M. Moors, C. Stachniss, and F. Schneider. Coordinated multi-robot exploration. *IEEE Transactions on Robotics*, 21:376–386, 2005.
3. S. Carpin, M. Lewis, J. Wang, S. Balakirsky, and C. Scrapper. USARSim: a robot simulator for research and education. In *Proc. of the 2007 IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2007.
4. L. Chaimowicz and V. Kumar. Aerial shepherds: Coordination among uavs and swarms of robots. In *7th International Symposium on Distributed Autonomous Robotic Systems*, 2004.
5. J. L. Drury, J. Richer, N. Rackliffe, and M. A. Goodrich. Comparing situation awareness for two unmanned aerial vehicle human interface approaches. In *Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics*, 2006.
6. C. V. Goldman and S. Zilberstein. Optimizing information exchange in cooperative multi-agent systems. In *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 137–144, New York, NY, USA, 2003. ACM.
7. M. Goodrich, M. Quigley, and K. Cosenzo. Task switching and multi-robot teams. *Multi-robot Systems: From Swarms to Intelligent Automata*, 2005.
8. A. Makarenko, A. Brooks, S. Williams, H. Durrant-Whyte, and B. Grocholsky. An architecture for decentralized active sensor networks. In *IEEE International Conference on Robotics and Automation (ICRA'04)*, New Orleans, LA, USA, 2004.
9. C. D. Moffatt and R. D. Lilley. High-data-rate, network radio for wireless communications in military environments. *Military Communications Conference, 2007. MILCOM 2007. IEEE*, pages 1–7, Oct. 2007.
10. M. Rosencrantz, G. Gordon, and S. Thrun. Decentralized sensor fusion with distributed particle filters. In *Proceedings of the Conference on Uncertainty in AI (UAI)*, 2003.
11. B. Ross, J. Bares, D. Stager, L. Jackel, and M. Perschbacher. An Advanced Teleoperation Testbed. In *Field and Service Robotics: Results of the 6th International Conference*. Springer, 2008.
12. N. Schurr, J. Marecki, M. Tambe, P. Scerri, J. Levis, and N. Kasinadhuni. The future of disaster response: Humans working with multiagent teams using DEFACTO. In *AAAI Spring Symposium on Homeland Security*, 2005.
13. P. Velagapudi, O. Prokopyev, K. Sycara, and P. Scerri. Maintaining shared belief in a large multiagent team. In *Proceedings of Tenth International Conference on Information Fusion*, 2007.
14. P. Velagapudi, P. Scerri, K. Sycara, H. Wang, M. Lewis, and J. Wang. Scaling effects in multi-robot control. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 2121–2126, Sept. 2008.
15. P. Velagapudi, J. Wang, H. Wang, P. Scerri, M. Lewis, and K. Sycara. Synchronous vs. asynchronous video in multi-robot search. *Advances in Computer-Human Interaction, 2008 First International Conference on*, pages 224–229, Feb. 2008.
16. H. Wang, P. Velagapudi, P. Scerri, M. Lewis, and K. Sycara. How Search and its Subtasks Scale in N Robots. *To appear in ACM SIGCHI/SIGART Human-Robot Interaction*, 2009.