Coordinating (Shared) Perspectives in Robot Assisted Search & Rescue

Martin Voshell
The Ohio State University
Cognitive Systems Engineering Laboratory
Columbus, OH, USA
voshell.2@osu.edu

Augustinus H. J. Oomes
Delft University of Technology
Man-Machine Interaction
Delft, The Netherlands
a.h.j.oomes@tudelft.nl

ABSTRACT

From high fidelity field exercises to disaster response deployments, search and rescue robots are being readily integrated into rescue operations. Previous research has proposed that for such new technology to be successful in an operation the organization architecture needs to support the coordination of shared perspectives between the human team members and the robotic platforms. For this, the robot platforms need to be effective team players in the field of practice. Based on this conceptual model, this paper introduces a novel software interface utilizing virtual position and orientation indicators to alleviate perceptual ambiguities and navigation problems experienced by robot handlers and problem holders. By actively orchestrating and sharing these indicators between handler and operator displays, the interface caters to user expertise and to the natural competency of the human perceptual system. These probes provide a basic tool for aiding robot navigation and way-finding fundamental to effective team coordination and communication in urban search and rescue missions.

Keywords

Urban search and rescue, human-robot interaction, remote perception

INTRODUCTION

Robotic systems have been embraced as an aid to support and speed up rescue and recovery efforts in the urban search and rescue (USAR) domain. The following simplified scenario conveys some of the team coordination and remote perception challenges that must be designed for in a robot assisted search and rescue deployment.

After receiving an operation order, a rescue team is sent to investigate a large pile of debris from a partial building collapse. Upon arrival, the incident commander determines that the area is still too hot from fire damage to allow canine or human search teams to safely enter. Before dedicating rescue efforts, the commander needs to send in a structural specialist to characterize the collapse pattern. In this case, a robot being teleoperated by a human handler at the base camp is to coordinate activity with the structural specialist in order to investigate the site remotely.

The structural specialist needs to be able to gather pertinent information about building materials, types of debris, as well as later revisit specific hazardous areas to continually monitor for condition changes during the rescue and recovery process. In contrast, the handler must maneuver into and around the challenging terrain and avoid obstacles in order to efficiently explore and investigate regions of interest that the structural specialist requests. Many perceptual and platform specific constraints hinder mobility and exploration for the handlers who tend to employ a variety of different navigation strategies to overcome them. One can immediately realize just how disorienting this might be for the structural engineer just `stuck along for the ride' so to speak, when trying to gather information completely separate from that being utilized by the handler.

With these differences in mind, it becomes obvious that simply presenting the same raw video feed to both the handler and to the structural specialist will be of minimal value if information in that feed is not catered to individuals' goals and expertise. Each party has different goals, some interdependent, some highly independent, and the two of them must operate within these constraints to effectively coordinate with one another. In this example, being able to identify the type of structure collapse and locating trapped casualties with certainty are two critical functional roles that require a significant amount of coordination and teamwork between the robot, the handler, and the specialist. To increase the chances of finding survivors, higher level team commanders and incident command must operate in time critical decision making. The response organization that the specialists and robot-handler teams are operating in must be flexible enough to integrate the new types of information access the robot resources provide in such a high tempo mission critical setting. The robot platforms need to fit in the incident command structure, support existing search and operation procedures, and most of all fulfill functional goals.

The need for human and robot coordination in such organizations is traditionally overlooked by system developers. Being an effective team member includes the ability to pick up and adapt to the activities of others in the team to achieve coordinated activity. Decision making in USAR is perceptually triggered and the interfaces need to be able to help operators know when and how to acquire more evidence to act upon. When decoupled from the environment, many basic perceptual and coordination links between the handler and the robot in the remote environment are broken, severely impacting work ability. Based on these challenges faced in the field, we elaborate on an envisioned model of operation, present a novel interface design, and report on a test of this design in a simulated search and rescue event.

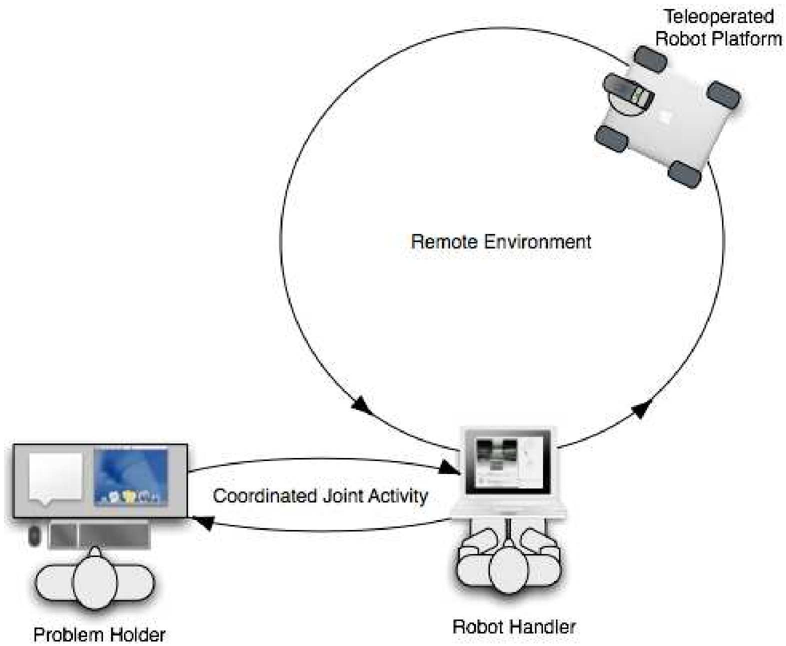

Figure 1. Coordinated joint activity is pivotal toward understanding a remote environment through a robot's sensors. A simplified version of the envisioning model proposed by Woods, et al. (2004), this figure illustrates how our software interface is shared between and catered to the specific Problem Holder and the Handler roles.

Human Robot Coordination

Woods, Tittle, Feil, and Roesler (2004) lay out a basic framework for envisioning the different roles and challenges in coordinated human-robot operations. Based on interactions with Robin Murphy and the Center for Robot Assisted Search and Rescue (CRASAR, University of South Florida, www.crasar.org), the conceptual model hypothesizes that for successful human robot coordination at least two fundamental human roles must be planned for: the robot handler and the problem holder. The handler role essentially plays out as `the driver' of the robot and is typically fulfilled by the system developers who have a better understanding of the robot's capabilities and limits. The robot handler is concerned with the knowledge, practice, and interfaces needed to manage the robotic platform in a physical environment as a resource relative to the challenges the environment poses.

The handler faces many challenges. A basic example from fieldwork indicates that operators spend significantly more time gathering information about the state of the robot and the state of the environment than actually navigating the robot (Burke, Murphy, Coovert, and Riddle, 2004). Another observation is that the mission portion of the task (i.e., searching for a victim in visual clutter, assessing structural conditions) was far beyond the capabilities of the autonomous computer vision and navigation algorithms onboard the robot. As a work around in such cases, the handler role is often fulfilled by more than one person because the navigation and control interfaces are so difficult and poorly designed that it takes at least two people to teleoperate one robot: one operator focusing on navigation in the complex environment while another concentrates on visual search.

The problem holder(s) role is fulfilled by the search and rescue personnel (task force leader, structural, search, canine, medical etc) whose goals lie in ``trying to characterize the search situation and achieve rescue goals" (Casper, 2002). The problem holder role points to the human role(s) responsible for achieving mission goals and the associated knowledge and experience. This again is reflected in how best to support incident commanders in their decision making as well as how to manage the many potential sources of information remote sensor platforms provide to different specialists (e.g., medical director, structural engineer). The problems holders, handler, and robot must be able to work together as effective team members to be able to coordinate their activities and achieve operational goals.

The problem holder utilizes the robot handler/robotic system couple to carry out their intent at a distance. Due to the inherent brittleness of the robot, the handler must operate and monitor to fulfill the context gap and coordinate with the problem holders. We cannot emphasize enough what a critical role the handler plays in coordination and in order to better support how they perform their job, we must address the ambiguities the handler faces inherent in mediated remote perception.

Remote Perception Challenges

Casper (2002) relates many difficulties that hindered rescue workers understanding of the remote environment when exploring rubble piles and void spaces. The typical means of navigation are provided by cameras that are usually fix-mounted on the robot, are small in diameter, have small angular fields of view, operate with many frame dropouts, relay low resolution images, and have poor (if any) color rendering. The robot relays this impoverished video imagery to the handler’s controller software. It is at this critical point where many perceptual ambiguities inherent in the remote setup can arise or even worsen from the crude integration of the direct video sensor feedback within the interface visualization. Trying to understand the remote environment through this very literal soda straw undermines just what the human perceptual system excels at in the natural world.

Seeing through a remote camera is not the same as having a human observer at a scene. Teleoperation is of a mediated perception-action nature and it is the interface where the handler must create a visual understanding based on the constraints of the remote robot agent in the environment. In such cases ambiguities involving object recognition, judgment of scale, and the absolute position of objects in the remote world are abundant. Lighting is usually uncontrolled and platforms equipped with their own illumination may rob color, give false specularity cues, and greatly hinder texture discrimination. As humans we actively sample the distal world with effortless independent coordination between heading and gaze. However, with the fixed camera platforms gaze and heading are neither independent or controllable. When no frame of reference for body awareness is provided there is profound misperception of depth, speed, and scale of obstacles and passages. Many of these ambiguities stem from the impoverished and conflicting cues affecting depth perception and how it relates to size perception and are outlined in figure 2.

Figure 2. Examples of the ambiguities encountered in robot teleoperation. From top left to bottom right. Robots tend to be deployed in vertical voids. It is not always easy maintaining orientation especially with minimal surface and texture information. Top Right illustrates a very basic agent-environment disconnect, the affordance of reachability is absent, the operator has no sense of scale or just where the manipulator arm is, or what can be reached with it. Bottom left, when cameras are fix mounted, traversing stairs constantly pitching upward hinders navigation. Bottom right shows just how much information poor illuminations robs from the scene.

Such visual ambiguities are resultant of poor orchestration between the human operators, the robot and sensor hardware, and the monitoring and control software. For the robot to be useful, it must aid the operator's ability to be able to traverse such debris, find unexplored reachable voids, and relay critical information back from the area. An operator does not need a rich full 3D representation of the pile from our original example, they simply need to be able to operate their robot in a goal-directed manner so as to be able to anticipate their own potential for action given the robot's relationship to conditions in the environment. The handler-robot team needs what Woods et al (2004) term functional presence. Functional presence occurs when the remote observer has sufficient information available to his senses to effectively function as if he were directly perceiving and acting in the remote environment. The handler's interface needs to be more than just a raw video feed and must be enhanced so as to recover what is lost by decoupling the human handler's perceptual system from the environment being explored.

Even as onboard autonomy and sensor hardware mature, the interfaces found in recent fieldable systems still tend to miss the bigger problem. Murphy and Burke (2005) continue to see that operators have difficulty simply integrating the robot's view into their understanding of the search and rescue site. Rescue personnel tended to compensate for this coordination challenge and re-establish common-ground (Clark and Brennan, 1991) by communicating with other team members at the site to gather information that would support their joint activity. This work-around is a valuable artifact of cognitive work and serves as a valuable leverage point emphasizing the need to support such usability into a new interface to support this fundamental cognitive activity.

To aid functional presence and team-robot coordination the interface must help robot handlers and problem holders resolve these perceptual ambiguities in order to reduce navigation errors, enable faster void searching, and support more timely and accurate identification of hazards and victims. Overcoming the ambiguities and catering to these coordination needs in remote perception is critical for creating shared perspectives across actors-in-the-scene and remote specialists effectively enabling them to better work together. The iSearch software architecture we are developing serves as a platform to begin to face these challenges.

The iSearch Interface Design

Our interface design looks to overcome many of the significant perceptual ambiguities handlers face as well as provide a structure for coordinating multiple perspectives between handlers and problem holders as a remote exploration plays out. Our previous research has looked at reducing keyhole effects that hinder remote functional presence in virtual USAR environments (Voshell, Woods, Phillips, 2005) and the current research uses a physical robot platform to test new interface in the real world (see Figure 3). The resulting robot control software, dubbed `iSearch', is under development to address these many HRI/HRC challenges.

Figure 3. The robotic sensor platform dubbed iRescue can be controlled either wirelessly or tethered by any machine running the iSearch control software. In this case, the control software is utilizing a live video sensor feed along with a dynamic shared map representation.

The robot and handler team must coordinate with problem holders to accomplish their goals. The basic interface layout for the handler and problem holder(s) is the same and consists of two viewports. The pair of viewports varies in functionality based on user role: the problem holder and handler share similar information catered to their specific needs. Each interface viewport represents a different point of view, a perspective, and more importantly control over point of view is delineated relative to the needs and goals of the handler and problem holders.

Wickens and Prevett (1995) demonstrated that local navigation tasks (such as obstacle avoidance) are maintained by knowledge in relation to an egocentric frame of reference. For this, we provide a robot sensor navigation viewport with support for multiple arrangements of video and sensor overlays. The teams must also accomplish many global spatial awareness tasks (path finding, waypoint planning) which are maintained by information relative to an exocentric frame of reference. A second viewport consists of a jointly created shared map for such exocentric tasks. Péruch, Pailhous, and Deutsch (1986) showed that the redundancy of information between the egocentric and exocentric views is crucial toward resolving discrepancies between the different frames of reference. The ability to take different points of view between the two displays, the easier the coordinated task at hand will be.

The interface components of the handler-robot display and the problem holder-handler display are designed to support two key requirements necessary for joint cognitive systems by making the monitored processes and robot agents observable and directable (Christofferson & Woods, 2002). This is present in the navigation view where the handler views feedback from the robot, and is just as relevant in the problem holder's viewport containing feedback from the actions of the handler-robot team.

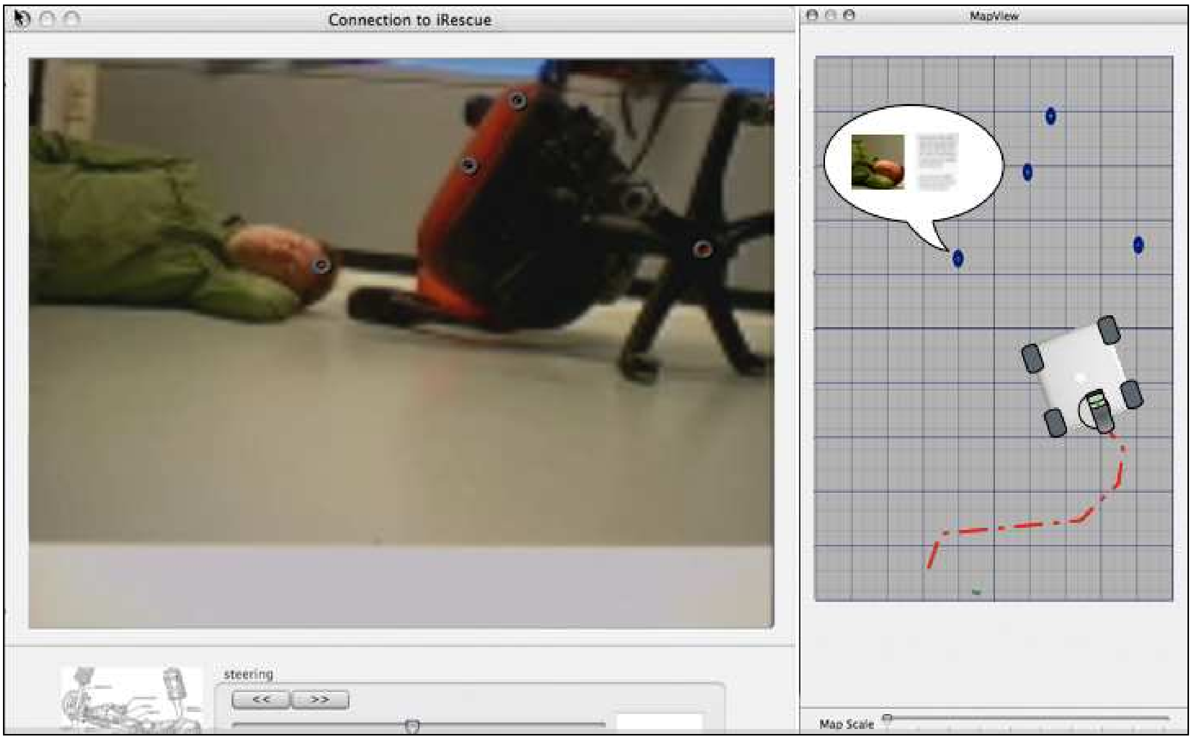

Figure 4. iSearch Application with sensor and map viewport. This is indicative of the robot handlers interface. Combined with on-board range finding, the handler actively positions orientation markers to mark-up the sensor feed. This information is related to the dynamic map where the handler may further tag groups of figures as well as receive input from problem holders viewing the created map.

Robot Sensor Navigation Viewport

The robot sensor navigation viewport displays remote sensor information, enables the viewer control over point of view when available, and gives 3D tools to mark-up the 2D video feed.

CRASAR's findings have shown just how important having multiple points of view in a scene can be. Since there is no way for a designer to guarantee that they have picked the right point of view for any task a priori, the viewport allows different users to vary and control different points of view (themes first considered as Viewtracks, Tittle, Roesler and Woods (2001)). Points of view and camera position and track are made explicit in the design and interface of the 2-D and 3-D representations.

One effective way to support inter-display navigation in the computer medium is to include landmarks in the artificial data field (Woods and Watts, 1997). Taking a very literal interpretation of this, and given the navigation through mediated physical exploration challenges we are facing, we introduce virtual landmarks in the display to project into the physical world. Inspired by the work of Jan Koenderink's use of gauge figures to discern pictorial depth information from 2D scenes (Koenderink, 1998), we are looking at discerning valuable 3D information through shared sets of oriented gauge figures, as well as the implications that groups of these figures serve as reference landmarks for higher level navigation and coordination tasks.

Shared Map Viewport

The second viewport serves as a navigation tool for the creation, displaying, and sharing of experiential maps between problem holders and handlers. Observable status at a glance cues (Woods, Watts, 1998) are available relaying the robot's current position and orientation along with a trace of the robots path history. The map is continuously updated by telemetry data from the robot and is complemented with the positioned landmarks from the handlers and other problem holders. As sets of these landmarked gauge figures and defined map areas are formed, access to higher level critical operation information such as ‘you are here’ information, landmark pathing, and escape routes, can be considered in context to operational constraints, (i.e. battery power, signal strength) amongst the multiple groups.

Such a map may not be immediately useful to a handler preoccupied with mobility problems, however enabling the ability for another member to quickly extract survey knowledge directly from that map makes a powerful navigation support tool. Even when spatial knowledge is well developed and functional, it still has a tendency to be error prone and subject to ambiguities- for this we rely upon the onboard autonomy and sensor data to overcome discrepancies in the relative positions between the many obstacles and areas the robot and handler explore.

Simulated Test

An early version of the iSearch software was tested in a simulated search and rescue robot competition at TU Delft (Figure 5.). Four teams competed with different robot hardware and control software. Each robot was controlled by a handler and monitored by a navigator. The robot handler controlled the remote robot platform in a separate room from the actual exploration and the navigator was responsible to help the handler locate victims, hazards, and build a map of the explored space. This competition served as first test of our coordinated perspectives interface within iSearch.

The handler and the navigator were co-located on different continents (Delft, the Netherlands, and Saratoga Springs, New York). The navigator in New York shared a real time video feed with the handler in Delft. To achieve common ground in the activity, the setup enabled the handler to point at objects in the video feed from the robot. The navigator could then also see the robot’s video as well as what the handler was pointing at. They jointly built a map of the explored area to use for navigation, and as an artifact at the end of the competition. A combination of high quality optics and the powerful coordination supported by iSearch won the competition.

Though the arena was set up to imitate the structural mess after a building collapse, the light conditions were unrealistic compared to the real-world examples in Figure 2. Our present prototype can not handle dark environments; we plan to improve our robot and its interface in the next phase of development and also make the experimental circumstances more and more realistic.

CONCLUSION

A USAR incident could occur anywhere there is potential for casualties to be trapped under collapsing structure. To be prepared for such unexpected events, robotics systems developers must invest in equally robust platforms to be able to respond effectively in the face of such uncertainty. To do this, they need to acknowledge the fundamental need for coordination between the roles that must be accounted for in human robot operations.

The iSearch software interface we have proposed is one such attempt to support coordination by catering to handler-robot work interaction and catering to the problem holder as a supervisory controller within the system. The handler and the robot actively build a model of the remote scene by rapidly probing and tagging the remote environment with virtual orientation indicators in the navigation window. This information is then translated to the exocentric shared map viewport of both the handler and other problem holders. Problem holders then have the ability to monitor, mark up, and redirect the handler based on these two sets of information and as new information comes to light. The interface synchronizes these shared models of the world across the parties involved and allows them to detect and repair discrepancies in operation.

Figure 5. In the simulated search and rescue task, the robot operator communicates with an off-site navigator over an early version of the iSearch software to jointly build a map of the environment while exploring the disaster area for victims.

ACKNOWLEDGMENTS

This work was carried out while Voshell was a visiting researcher in Delft from August to December 2005. Financial support from the Ohio State University and the Delft University of Technology is gratefully acknowledged. We thank David Woods for feedback and support.

REFERENCES

1. Burke, J. L., Murphy, R. R., Coovert, M. D., and Riddle, D. L. (2004). Moonlight in Miami- a field study of human robot interaction in the context of an urban search and rescue disaster response training exercise. *Human Computer Interaction*, *19*, 85-116.
2. Casper, J. (2002). *Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center*. Unpublished master's thesis, University of South Florida, Dept. Computer Science and Eng., Tampa, FL.
3. Christofferson, K., and Woods, D. D. (2002). How to make automated systems team players. In E. Salas (Ed.) *Advances in human performance and cognitive engineering research* (p. 2-12). Stamford, CT: JAI Pres/Elsevier.
4. Clark, H. H., and Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasly (Eds.), *Perspectives on socially shared cognition.* Washington: D.C.: American Psychological Association.
5. Koenderink, J. J. (1998) Pictorial Relief, *Phil. Trans. R. Soc. Lond.* A, 356, 1071-1086.
6. Murphy, R. R., and Burke, J. L. (2005). Up from the rubble: Lessons learned about HRI from search and rescue. In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, Sep. 26-30 2005. Human Factors and Ergonomics Society.

7. Péruch, P. Pailhous, J., and Deutsch, C. (1986). How do we locate ourselves on a map: a method for analyzing self-location processes. *Acta Psychologica*, 61, 71-88.
8. Tittle, J. S., Roesler, A., and Woods, D. D. (2002). The remote perception problem. In *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.
9. Voshell, M., Woods, D. D., and Phillips, F. (2005). Overcoming the keyhole in human-robot coordination: simulation and evaluation. In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, Sep. 26-30 2005. Human Factors and Ergonomics Society.
10. Wickens, C. D. and Prevett, T. T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied.* 1(2), 110-135.
11. Woods, D. D., Tittle, J., Feil, M., and Roesler, A. (2004) Envisioning human-robot coordination in future operation. *IEEE Transactions on Systems, Man, and Cybernetics- Partc C, 34*(2), 210-218.
12. Woods, D. D. and Watts, J. C. (1997). How not to have to navigate through too many displays. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human computer interaction* (2 ed., p. 617-650). Amsterdam, The Netherlands: Elsevier Science.