Mixed-Initiative Interactions for Mobile Robot Search

Curtis W. Nielsen and **David J. Bruemmer** and **Douglas A. Few** and **Miles C. Walton**
Robotic and Human Systems Group
Idaho National Laboratory
{curtis.nielsen, david.bruemmer, douglas.few, miles.walton}@inl.gov

Abstract

At the INL we have been working towards making robots generally easier to use in a variety of tasks through the development of a robot intelligence kernel and an augmented-virtuality 3D interface. A robot with the intelligence kernel and the most recent interface design were demonstrated at the AAAI 2006 robot exhibition and took part in the scavenger hunt activity. Instead of delegating all responsibility for the scavenger hunt to the robot, as is common in traditional AI approaches, we used a mixed-initiative human-robot team to find and identify the objects in the environment. This approach allowed us to identify the objects and, using icons and labels, place them on the digital map that was built by the robot and presented in the 3D interface. Mixed-initiative interactions support the integration of robotic algorithms aspects with human knowledge to make the complete system more robust and capable than only using robots or humans.

Introduction

The robot exhibition workshop at AAAI 2006, provided an excellent opportunity to demonstrate some of the human-robot teaming technology that we have been working on at the Idaho National Laboratory (INL) including collaboration with the Stanford Research Institute (SRI) and Brigham Young University (BYU). The collaborative efforts are focused on bringing together tools for improving the utility of a mobile robot. In particular, the INL has developed a general purpose robot intelligence kernel (Bruemmer *et al.* 2005; Few, Bruemmer, & Walton 2006) that uses a simultaneous localization and mapping (SLAM) algorithm developed by SRI (Konolige 2004) and an augmented virtuality 3D interface originally developed at BYU (Ricks, Nielsen, & Goodrich 2004).

The goal of our research is to improve the usefulness of mobile robots by making them easier for an operator to use. In order to do this, our research is focused on two fronts: First, making the robot more capable of acting in the environment on its own, and second, providing better information about the robot and its environment to the operator. The robot system we use has dynamic levels of autonomy that can be changed depending on the task, the needs of the operator, and the capabilities of the robot. Robot situational awareness information is recorded and abstracted by the robot and presented to the operator via the 3D interface which also provides simplified tasking tools that can be used by the operator to direct the robot.

The development of the RIK and 3D interface has been directed and proved through numerous user-studies in a spiral development process that allowed us to see when particular solutions were more appropriate than other solutions (Nielsen & Goodrich 2006a; 2006b; Bruemmer *et al.* 2005),. One of the key observations from these studies has been the fact that a single level of autonomy or single interface presentation was not always appropriate for every task. Furthermore, when the operator was given the ability to choose the desired level of autonomy, they often met with frustration and subjectively claimed that when choices were limited they felt a higher degree of control (Bruemmer *et al.* 2005). Therefore, although our efforts have focused on the development of multiple levels of dynamic autonomy, when it comes to specific tasks, we look at what aspects of the task are best accomplished by the operator and which ones are best accomplished by the robot and we create a human-robot, mixed-initiative interaction. This approach differs from conventional AI approaches where the end goal is to have the robot perform the complete task.

Our robot was invited to take part in the scavenger hunt portion of the mobile robot exhibition where robots demonstrate their ability to find objects of interest in an environment. Although the approach of our solution is not exactly what was looked for in the AI sense, it does provide some insights into how the teaming of robots and humans can lead to better performance than using robots alone.

In this paper we will discuss the robot intelligence kernel (RIK) used to provide robots with an understanding of their environment and surroundings as well as the ability to dynamically interact with the environment. We then discuss the 3D interface and how it supports the operator's awareness of the environment around the robot. Next, we discuss our approach to mixed-initiative interactions, specifically, how the robot and operator roles can be divided and supported to accomplish tasks. We then show how this approach was used for the scavenger hunt portion of the mobile robot exhibition. The paper concludes with lessons learned and directions of future research.

The Robot

The robot that was used at the AAAI-2006 robot exhibition is an ATRV-mini originally built by iRobot and augmented with the Robot Intelligence Kernel (RIK) developed at the INL (Bruemmer *et al.* 2005; Few, Bruemmer, & Walton 2006). The intelligence kernel is a software architecture that resides on board the robot and serves as the brains of the robot. The RIK is a portable, reconfigurable software architecture that supports a suite of perceptual, behavioral, and cognitive capabilities that can be used across many different robot platforms, environments, and tasks. The RIK has been used for perception, world-modeling, adaptive communication, dynamic tasking, and autonomous behaviors in navigation, search, and detection tasks.

The software architecture is based on the integration of software algorithms and hardware sensors over four levels of the RIK. The foundation layer of the RIK is the Generic Robot Architecture that provides an object-oriented framework and an application programming interface (API) that allows different robot platforms, sensors, and actuators to interface with the RIK. The second layer is the Generic Robot Abstractions layer which takes data from the first layer and abstracts the data so that it can be used in abstracted algorithms that are designed for generic robot systems and are easily portable to new systems. The third layer is comprised of many simple reactive and deliberative robot behaviors that take, as input, the abstractions from the second layer and provide, as output, commands for the robot to perform. The fourth and final layer provides the "Cognitive Glue" that orchestrates the asynchronous simple behaviors into specific task-based actions. By combining the individual behaviors, the cognitive glue of the final layer supports a suite of meta-level capabilities such as a) real-time map-building and positioning, b) reactive obstacle avoidance and path-planning, c) high-speed waypoint navigation with or without GPS, d) self-status awareness and health monitoring, e) online adaptation to sensor failure, f) real-time change detection, g) human presence detection and tracking, and h) adaptive, multi-modal communication.

The RIK has been installed on a variety of robot platforms including the ATRV-mini, ATRV-Jr, IRobot Packbot, Segue, and specially designed systems from Carnegie Mellon University (CMU). Furthermore versions of the RIK are being used by other HRI teams throughout the community (Garner *et al.* 2004; Everett *et al.* October 26 28 2004; Baker *et al.* October 2004). The ATRV-mini (see Figure 1) was the robot shown during the AAAI mobile robot exhibition. During the exhibition, the robot could be seen wandering around the exhibition floor and sometimes it was wondering autonomously and sometimes the robot was working to get to a particular goal as defined by the user.

The Interface

True human-robot teamwork requires a shared understanding of the environment and task between the operator and the robot. To support a dynamic sharing of task roles and responsibilities between a human and a robot, a 3D interface had been developed through collaborations with Brigham Young University (Ricks, Nielsen, & Goodrich 2004) and extended through current work at the INL. The 3D interface supports an augmented virtuality visualization of information from the RIK that allows the operator to observe the robot's understanding of the environment. Understanding the robot's knowledge of the environment allows the human to anticipate and predict robot behavior and effectively task the robot.

Figure 1: The Atrv-mini used for the exhibition and scavenger hunt.

By utilizing map generated information from the robot and heuristic-based sensor fusion techniques, the interface can present a 3D, virtual representation of its surroundings in real time The 3D interface is based on the ability of the robot to build a map of the environment which, for this research, is done using a consistent pose estimation (CPE) algorithm developed by Stanford Research Institute (Konolige 2004). The map of the environment is presented to the operator on the "floor" of the virtual environment and a model of the robot is presented where the CPE algorithm localized the robot (Ricks, Nielsen, & Goodrich 2004; Nielsen & Goodrich 2006a). The operator views the information from the interface from a perspective tethered to the robot at a position slightly above and behind the robot such that obstacle information on the sides of the robot and behind the robot are visible in the interface. In tasks that require a human's observation of video information, the video data is displayed integrated with the range information to provide the operator with contextual information about where in the environment the visual data is coming from. As the camera is panned around an environment, the video panel in the 3D interface also moves to support the operator's understanding of where the robot is looking in the environment. Figure 2 shows the 3D interface.

This 3D interface has been shown to significantly reduce the time to accomplish navigation specific tasks with mobile robots in comparison to more traditional interfaces. Additionally, operators are able to avoid obstacles better, keep the robot farther from obstacles, and better manipulate sensor equipment such as pan-tilt-zoom cameras (Nielsen & Goodrich 2006a; 2006b). When video information is

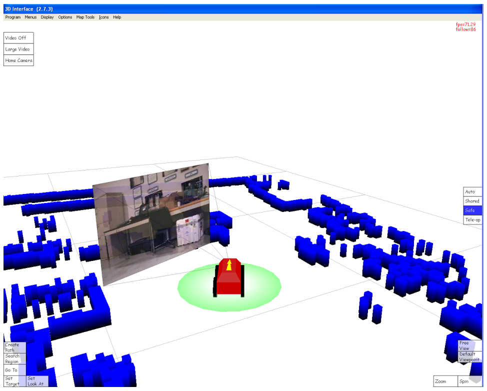


Figure 2: The virtual 3D interface integrates map, video, and robot pose into a single perspective.

not necessary (as is the case with some navigation tasks), This 3D interface has also been shown to utilize between 3,000 and 5,000 times less bandwidth than a video display tasks (Bruemmer *et al.* 2005). This is made possible by providing the operator with simplified map information that illustrates traversable areas of the environment rather than full video streaming.

Exploration Specific Developments

The 3D interface is particularly useful in teleoperation tasks where the operator directly controls the movement of the robot through the environment. Furthermore, as mentioned above, the 3D interface has been shown to support the operator's ability to control a pan-tilt-zoom camera. However, it has been difficult to show that the 3D interface actually improves an operator's ability to search for, find, and identify items of interest hidden in the environment. For example, in a study with expert search and rescue personnel, there was little difference in the operator's ability to find victims when using the INL system in comparison to another system (Yanco *et al.* 2006).

One possible reason that the 3D interface has not strongly support search tasks as much as it supported navigation tasks is that navigation tasks required a robot-centric understanding of the robot's environment. Supporting the operator in navigation tasks is facilitated by tethering the operator's perspective of the virtual environment to the robot such that no matter how the robot moves, the environment is always viewed from a position above and behind the robot. The problem is that in search tasks, the operator does not necessarily require a robot-centric view of the environment, but rather an exocentric or environment-centric perspective that would improve his or her understanding of what parts of the environment had already been searched.

To address this reasoning, we provided an option to change the 3D interface such that the operator's perspective was fastened and no longer tethered to the movement of the robot and the video was presented in the top middle of the screen. This, however, brought the challenge of more difficult teleoperation, so, we also focused on supporting the operator's use of higher levels of navigational autonomy in the RIK. Two of the RIK navigational behaviors include waypoints and a go-to option. Traditionally, these navigation behaviors are activated by selecting a button on the interface and setting a series of waypoints for the robot to follow or setting the go-to location at the desired destination in the environment and leaving, the responsibility of finding a path to the destination to the robot. There are circumstances when each approach is better than the other approach however the decision of which one to use was left to the operator. In order to improve the ability of the operator to task the robot we simplified this process and made it so the operator simply drags a ”target” icon to the desired destination. The interface then determines which of the navigational behaviors to request from the robot and the robot performs the requested action. This solution means the operator no longer concerns themselves with the details of how the robot will move, rather they only focus on the end goal. This becomes much like driving the intent of the robot or the goal of the robot rather than pure teleoperation.

Furthermore, since control of the camera is important to an exploration task and camera control from a joystick is often difficult and sluggish, we provided a solution, similar to the navigational control, where the operator drags a ”look-at” icon to the desired place in the environment where they would like the robot camera to look. As the robot moves through the environment, it always attempts to keep the camera oriented towards the desired destination. The combination of these approaches supports the operator's responsibilities to control the robot by minimizing the effort necessary to move the robot and orient the camera. Figure 3 illustrates the interface and icons used for robot control during the scavenger hunt.

Exhibition and Scavenger Hunt

Throughout the exhibition phase of the robot competition, the ATRV-mini could be seen traversing the environment and building a map of its findings. The reactive behaviors of the robot were easily predictable and the robot could be guided by simply walking alongside the robot. Obstruction of a range sensor on a side of the robot would turn the robot in the opposite direction. The 3D interface presented the map information about the environment and provided an intuitive representation of the spatial information that could be seen by interested parties. One of the limitations with the map-building approach was that people standing still were often added to the map as an obstacle. When the people moved, the map still maintained that they were obstacles, even if the robot traversed the place in the environment. This often led to maps that appeared quite cluttered. Fortunately, the robot was able to identify static features such as walls, tables, and equipment that helped the user understand the robot's observations of the environment. When the robot was tasked to autonomously navigate to a particular place in the environment, it uses its internal map to plan how it will get to the goal location. The robot's intent is then displayed in the 3D interface as a set of waypoints to inform the operator. The challenge with the ghost obstacles from moving people is that sometimes the robot would plan inefficient paths through the environment in an attempt to avoid ghost ob-

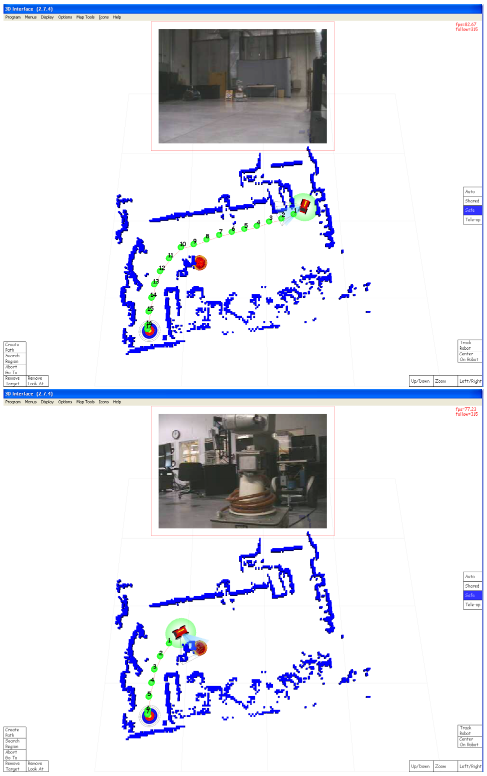


Figure 3: The new 3D interface used for the scavenger hunt.

stacles. Future work will address the issue of removing old signatures in the environment.

We were also invited to use our robot in the scavenger hunt portion of the robot exhibition. The system was not specifically designed to perform in the scavenger hunt activity and our system did not meet all the requirements for the scavenger hunt however, organizers were interested in a practical demonstration of the available technology. The field of mobile robots using artificial intelligence has long been interested in designing algorithms and robots that can perform a task similar to that of a robot and this was the main goal of the scavenger hunt task. Conference organizers placed numerous bright colored objects throughout the environment and participants in the scavenger hunt were then tasked to find these objects. Bonus points were given to correct identification and the ability to mark the objects on a map generated by the robot.

In the scavenger hunt, we successfully identified 5 objects (a Winnie the Pooh bear, two tennis balls, a blow up toy, and a pail). Since the robot built a map of the environment, the operator was able to place the items within the digital map of the environment and record their location with labels and icons for future reference. The division of labor was such that the robot performed the map-building and movement from place to place and the operator handled control of the camera, the placement of navigational goals for the robot, identification of objects, and the iconic representation of the items in the digital map. While this approach was definitely not congruent with traditional AI approaches, it does demonstrate that the human and the robot could sufficiently perform the task when working together.

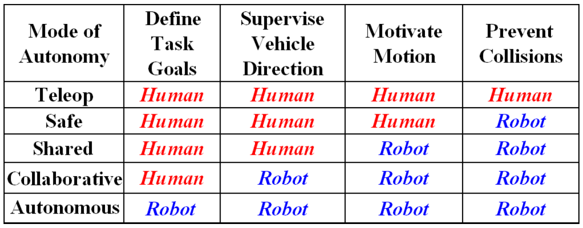
Mixed-Initiative Interactions

It could be said that of course our robot performed well, it had a human identifying objects. The purpose of our involvement in the study was not to demonstrate any technology that would work because the simplest solution would be to just send a human looking for the objects. After all, a human would be able to search in many sneaky places where a robot would not know to look. The purpose of the scavenger hunt was to demonstrate the current state of the art in mobile robot search technology. While our contribution is not in the field of algorithms or sensors, it is in the field of human-robot interactions.

Previous user studies and anecdotal observations have illustrated that a human operator and robots have different sets of strengths and weaknesses. To capitalize on the strengths of the team members, requires the orchestration of the interaction between the human and the robot and may change depending on the task. For example, in navigation tasks, we have found that the robot tends to be more proficient and precise than the human. In search tasks with a sensor that can specifically identify things of interest, again, the robot is more proficient than the operator. However, in a search task where there is not a sensor for identifying specific items of interest (especially in video), then it might be best to allow the operator to perform the identification of objects.

It is important to note that even when searching the video for the objects of interest, navigational aspects of AI are particularly helpful because they can reduce the cognitive requirements on the operator. For example, in the scavenger hunt task, the operator only had to specify the goal position for the robot and desired look-at position of the camera. The robot was then responsible to move the robot and the camera. This left the operator with time to monitor the video and determine if anything of interest is found. Approaches that have not used intelligent robots for search tasks have demonstrated that the cognitive responsibilities on the operator are quite demanding and often the operator misses important information in the environment (Casper & Murphy June 2003; Burke *et al.* 2004).

The balance between human and robot responsibilities is often referred to as mixed-initiative interactions. Traditionally, this has meant that both humans and robots are viewed as peers and they work together to solve problems or tasks that they are unsure about how to solve themselves. Often this problem solving takes place as a dialog between interested parties in which each participant reasons about available information and they come to a solution. This approach is especially applicable in domains where the possible intentions of the operator are varied and unpredictable (e.g. Microsoft word, excel) (Horvitz 1999). However, in domains where the task and responses are more predictable, we have found that it may be beneficial to define the task in terms of human and robot responsibilities. Then the lines are drawn, and the robot knows when the operator should take initiative and the operator is limited in their possible actions when the



Mode of Autonomy	Define Task Goals	Supervise Vehicle Direction	Motivate Motion	Prevent Collisions
Teleop	*Human*	*Human*	*Human*	*Human*
Safe	*Human*	*Human*	*Human*	*Robot*
Shared	*Human*	*Human*	*Robot*	*Robot*
Collaborative	*Human*	*Robot*	*Robot*	*Robot*
Autonomous	*Robot*	*Robot*	*Robot*	*Robot*

Figure 4: Mixed-initiative responsibilities.

robot should take initiative. A mixed initiative chart that delineates the responsibilities of the human and robot for the different modes of autonomy that have been developed for the RIK is shown in Figure 4.

Notice that in this chart, except for fully autonomous, the human always has the responsibility for defining the task goals. Even in traditional AI solutions, defining the task goal will likely remain the responsibility of the operator for some time.

Although it may sound that traditional AI and human-robot interactions are separate approaches, the two are actually related and only differ by maturity of the technology. The reason some mixed human-robot solutions are more effective than traditional AI approaches is that the AI approach has not, as of yet, fully solved the problem. As solutions with AI improve, then the human-robot interactions will also change up to the point where the AI solution can complete the task without human input. In the interim, a workable solution is to use the best that AI has to offer and augment that with knowledge from the human operator.

Conclusion

In this paper we discuss the Robot Intelligence Kernel and the 3D interface as they were used at the AAAI 2006 Mobile Robot Exhibition and Scavenger Hunt. Some of the differences between our approach and other participants are discussed, namely, that we use the human to identify objects of interest in the environment where the other approaches use the robot.

Although the goal of the conference was to demonstrate algorithms that support artificial intelligence without human input, our approach demonstrates the field of mixed-initiative interactions where the robot and human are responsible for accomplishing different and well-defined aspects of the task. Artificial Intelligence is a worthy goal that is pursued by many roboticists, however, when the AI algorithms are not robust enough to work in many conditions, a mixed-initiative approach, that utilizes as much of AI as possible but also incorporates human-knowledge is a solution that can make the human-robot system more robust and capable than only using robots or humans. As AI algorithms and technologies mature, the interactions between humans and robots to solve tasks will change until the solution is fully autonomous and the human is no longer needed for the task.

In the future we plan to continue to explore methods of making unmanned vehicles more capable and providing sufficient information to the operator such that human-robot interactions are facilitated and can accomplish challenging tasks.

References

Baker, M.; Casey, R.; Keyes, B.; and Yanco, H. A. October 2004. Improved interfaces for human-robot interaction in urban search and rescue. In *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, 2960–2965.

Bruemmer, D. J.; Marble, J. L.; Few, D. A.; Boring, R. L.; Walton, M. C.; and Nielsen, C. W. 2005. Shared understanding for collaborative control. *IEEE Transactions on Systems, Man, and Cybernetics—Part A* 35(4):494–504.

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.

Casper, J., and Murphy, R. R. June, 2003. Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center. *IEEE Transactions on Systems, Man, and Cybernetics Part B* 33(3):367–385.

Everett, H. R.; Pacis, E. B.; Kogut, G.; Farrington, N.; and Khurana, S. October 26-28, 2004. Towards a warfighter's associate: Eliminating the operator control unit. In *SPIE Proc. 5609: Mobile Robots XVII.*

Few, D. A.; Bruemmer, D. J.; and Walton, M. C. 2006. Improved human-robot teaming through facilitated initiative. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication*.

Garner, J.; Smart, W. D.; Bennett, K.; Brummer, D. J.; Few, D. A.; and Roman, C. M. 2004. The remote exploration project: A collaborative outreach approach to robotics education. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 2, 1826–1830.

Horvitz, E. 1999. Uncertainty, action, and interaction: In pursuit of mixed-initiative computing. *Intelligent Systems* 14(5):17–20.

Konolige, K. 2004. Large-scale map-making. In *Proceedings of the National Conference on AI (AAAI).*

Nielsen, C. W., and Goodrich, M. A. 2006a. Comparing the usefulness of video and map information in navigation tasks. In *Proceedings of the 2006 Human-Robot Interaction Conference*.

Nielsen, C. W., and Goodrich, M. A. 2006b. Testing the usefulness of a pan-tilt-zoom (PTZ) camera in human-robot interactions. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*.

Ricks, B. W.; Nielsen, C. W.; and Goodrich, M. A. 2004. Ecological displays for robot interaction: A new perspective. In *International Conference on Intelligent Robots and Systems IEEE/RSJ*.

Yanco, H. A.; Baker, M.; Casey, R.; Keyes, B.; Thoren, P.; Drury, J. L.; Few, D. A.; Nielsen, C. W.; and Bruemmer, D. J. 2006. Analysis of human-robot interaction for urban search and rescue. In *Proceedings of IEEE International Workshop on Safety, Security, and Rescue Robotics*.