POTENTIAL AND CHALLENGES OF STEREO AUGMENTED REALITY FOR MOBILE ROBOT TELEOPERATION

Markus Sauer, Frauke Driewer, Manuel Göllnitz, Klaus Schilling

University of Würzburg, Computer Science VII: Robotics and Telematics
Am Hubland, 97074 Würzburg, Germany
{sauer, driewer, schi}@informatik.uni-wuerzburg.de

Abstract: In many mobile robot applications a human operator is still required to control the robot. The design of the user interface is the major element affecting the performance and efficiency of the task execution in the teleoperation scenario. This work investigates an approach to visualize all information gathered by the mobile robot in an integrated fashion by applying stereo augmented reality. Initial experiences and results with this type of interface for navigation tasks are presented.

Keywords: human-machine interface, user interfaces, robot navigation, mobile robots, remote control, man/machine interaction, telerobotics, teleoperation.

1. INTRODUCTION

Today robotics faces the increasing requirement of a paradigm shift from industrial applications with fixed requirements and conditions to everyday, dynamic human environments. In the last decades robotics research has made great advances to operate in such environments with a higher level of autonomy. Nevertheless, there are still many tasks and applications, where the superior perception, interpretation, reasoning capabilities and the experience of human operators are needed. Reasons for this need of human support can be for instance safety issues and/or complex, dynamic real-world scenarios. The most obvious application examples are robots applied in human-robot teams in search and rescue scenarios or robots utilized in space missions. These scenarios include very complex tasks due to e.g. highly dynamic, unknown or hazardous environments, danger for human live, or the risk of a complete mission failure resulting in costs of some million Euros.

Due to these mission crucial issues, mobile robots are often directly teleoperated and drive assistance features like e.g. path planning, path tracking and obstacle avoidance are integrated as support functions for the human operator. The graphical user interface for the robot's operator plays a key role for the overall performance and successful completion of a task or a mission.

A major challenge for the design of these user interfaces is not to overwhelm the operator with a large amount of information from the robot's environment presented side-by-side in various visualizations (e.g. in worst case raw data) without any direct relations. The operator in this case faces in addition to his/her actual task the difficulty that he/she has to mentally integrate all this information in a model about the situation at the remote place of the robot. As a result the operator often only focuses on one element of the graphical user interfaces. Often this is the camera image from the robot, because it delivers the most information about the remote scene in an integrated and intuitive fashion.

This contribution presents an approach how augmented reality (AR) concepts can be applied to visualize and integrate sensor data and additional environment information gathered by the robot directly in the visual presentation – the camera image - of the robot's work environment. Therefore, a stereo video stream from a mobile robot is

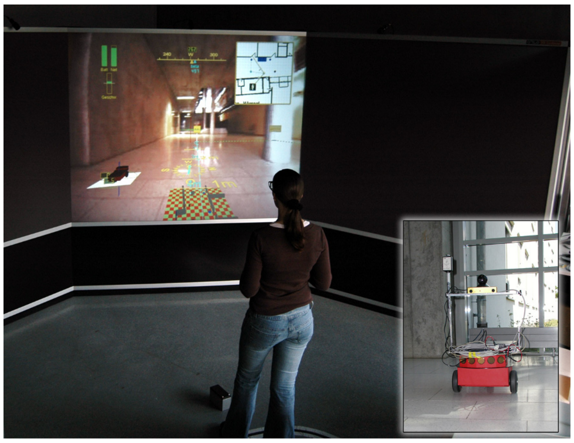

Fig. 1 Implemented AR interface on stereo projection system and teleoperated mobile robot

augmented with correctly aligned virtual three-dimensional representations of sensor data received from the robot. This integrated stereo AR user-interface is present on a large-scale projection system to the robot's operator. Fig. 1 shows this AR interface on the stereo projection system and the utilized mobile robot during tests.

The remainder of this paper describes first the related work and the mentioned, developed exemplary graphical user interface based on augmented reality techniques for mobile robot teleoperation. Furthermore, an overview of the experience, results of the first tests and the yielding potential and challenges of this kind of interfaces is given.

2. AUGMENTED REALITY AND MOBILE ROBOT TELEOPERATION

Due to the nowadays available hardware as well as graphic and computing power at reasonable costs AR systems as a new way of building intuitive man-machine interfaces are increasingly of interest. AR systems combine real and computer generated objects correctly aligned in real environments and run interactively and in real-time. Azuma *et al.* (2001) gives an overview of the major characteristics and challenges of an AR system and the major application areas. In the field of robotics AR systems are currently mainly investigated for programming maintenance, service or training tasks. Skourup *et al.* (2004) presents an example, how AR concepts can be utilized for a more effective way of industrial robot programming.

Compared to the industrial robot applications, the requirements for AR for mobile robot teleoperation are in many parts more similar to requirements for portable mixed reality systems for humans in outdoor environments, e.g. the requirements for tracking to register virtual and real objects. The Tinmith project (Piekarski *et al.,* 2002) realized at the University of South Australia is an advanced example in this area. Nevertheless, there are significant differences between the human portable and the teleoperation system, resulting from the remote control scenario. The operator who is not directly in the working environment is always limited by the capabilities of the sensors on the robot and their representations. These limitations significantly reduce the feeling of presence and the situation awareness of the operator, which are key-elements for successful mobile robot teleoperation.

Experience has shown that the challenging application search and rescue is a good domain for the investigation human-robot interaction. In this application area many studies were accomplished to determine requirements for an efficient work between humans and robots. Murphy *et al.* (2005) summarizes the lessons learnt from these studies. Two main findings relevant for the presented work are that building situation awareness is the major bottleneck, not autonomous navigation and the robot should not be considered as something to be controlled, but should be exploited as an active information source

Scholtz *et al.* (2004) derived guidelines for teleoperation interfaces in the field of urban search and rescue (USAR) from an analysis of the observations of the RoboCup USAR competition. Some results of this work like the summarized list of key requirements for an user interface listed in the following also strongly influenced the interface presented here. User interfaces for the remote control of a mobile-robot require:

- a frame of reference to determine position of robot relative to environment
- indicators for the health and status of the robot
- information from multiple sensors presented in an integrated fashion
- automatic presentation of contextually-appropriate information
- as the main feedback of the robot, the camera image, might suffer from disturbances (e.g. communication problems or brightness), it is recommend to supplement the video by other sensors
- the ability to self inspect the robot body for damage or entangled obstacles (not included in this contribution)

Kadous *et al.* (2006) describes one of the recently successful interfaces used in the Robocup USAR competition. Driewer *et al.* (2006) shows an approach, how to use mixed reality for teleoperation of mobile robots in the search and rescue scenario. Nielsen *et al.* (2006) compares a three-dimensional world representation, integrating the video as projection in front of the mobile robot, against side-by-side presentation of map and video as well as two-dimensional maps together with video.

The following sections present an approach how an efficient user interface for mobile robot teleoperation can be implemented utilizing AR-techniques.

3. SYSTEM SETUP

For the test of the AR interface a differential drive Pioneer 1 (see Fig. 1 bottom right) robot is equipped with a stereo camera, ultrasonic sensors and a compass delivering all three orientation angles of the robot. The lenses of the stereo camera were selected for an approx. 100 degree horizontal field of view. This large field of view supports the user in maintaining local situation awareness in the navigation task. The robot's sensor and video data and the commands from the operator to the robot are transferred over a dedicated WLAN link. Thus, network disturbances have not to be considered.

The stereo AR user interface for the teleoperator has been implemented on a large-scale stereo wall with a 100 inch diagonal. The pictures for the left eye and the right eye are polarized orthogonal and displayed merged on the projection screen. The users have to wear special glasses with the same polarization filters to separate the pictures for the left and right eye again and to gain the stereo effect. Fig. 2 shows an overview of the system setup.

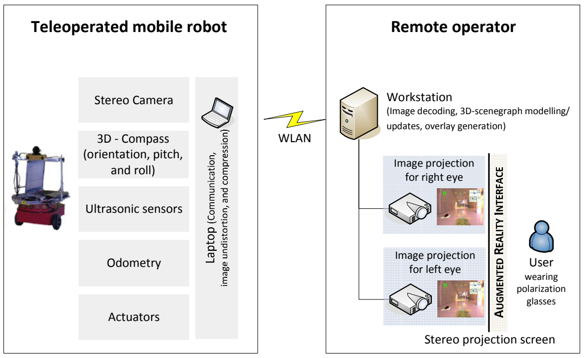

Fig. 2 System setup

The extrinsic (position and orientation of the two camera heads with respect to robot's reference point) and intrinsic (projection properties of camera heads) stereo camera parameters are determined by a calibration process before the interface is used. The extrinsic parameters are used for modeling the virtual stereo camera position and orientation in the AR interface. The intrinsic parameters are used to calculate undistorted camera images on the robot. This undistortion process is primarily needed, because the lenses with a large field of view introduce a large radial distortion on the images. The undistorted images together with the utilization of the intrinsic parameters to model the virtual cameras' projections on the user interface side are necessary to realize a well registered AR overlay.

4. THE AR-STEREO INTERFACE

The main design criteria this user interface was secure and fast navigation of a mobile robot in an unknown environment during an exploration and search task. The operator should be able to understand and memorize its local surrounding fast, detect obstacles, determine the next target point, how to go there and immediately detect if robot is approaching a dangerous situation. One important example especially in unknown environments is the danger that the robots overturns, which will eventually lead to a loss of the robot.
The main component of the teleoperation interface is the stereo video from the robot. It represents the work environment of the teleoperated mobile-robot. This stereo camera view is therefore the base element as background of the AR interface. All other data presented to the operator are visualized as three dimensional virtual objects. This information is integrated, and therefore correctly registered and overlaid, into the stereo camera image.

For the first test interface the following information is shown in an integrated fashion in the AR interface:

- Orientation of the mobile robot
- Pitch and roll angle of the robot
- Status (e.g. battery, network link) of the robot.
- Distance measurements from ultrasonic senors, corresponding to possible obstacles in the sensor field of view of the robot
- Distance reference
- Planned path and waypoints and navigation support to reach the next waypoint
- Maps as global frame of reference for the operator with robots current position

Waypoint system. The interface includes a waypoint management system. The user can plan a path by selecting waypoints in a 2D map. These waypoints are then placed in the correct three-dimensional position in the field of view of the robot. Additionally the path connecting the waypoints and the distance to the next waypoint is visualized. The yellow signs in Fig. 3 represent these waypoints. Waypoint "Alpha" is the next planed waypoint.

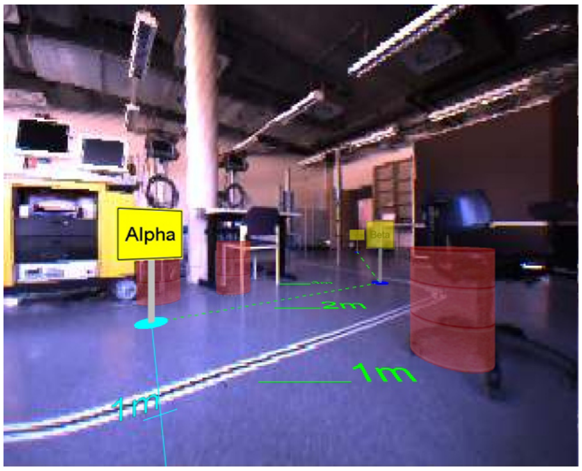

Fig. 3 Waypoint system, highlighted obstacles and distance reference components

Orientation visualization. The orientation of the robot as state element of the robot can be visualized (Fig. 4 from left to right) as classical display, as Head-Up Display (HUD) element comparable to those in modern aircrafts and as compass rose projected on the floor in front of the robot. The displays of compass also visualize the direction and the deviation to the direction to the next waypoint. The compass element is mainly relevant for navigation purposes.

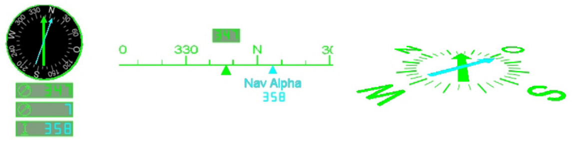

Fig. 4 Different implemented visualizations of mobile robot's orientation

Artificial Horizons. The artificial horizon can be visualized (Fig. 5 from left to right) as classical display, as lines in the center of the screen comparable to HUDs of modern aircrafts, as a mobile robot model and as a 3D visualization composed of two 3-sided Polygons. These visualizations of the robot's pitch and roll angle are mainly important for a safe teleoperation of the robot. They should help the user to prevent the loss of the robot by overturning.

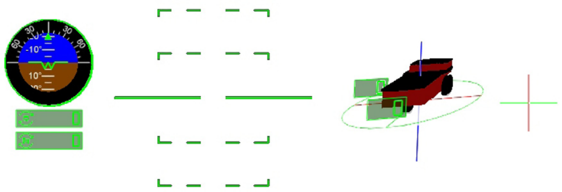

Fig. 5 Different implemented visualizations of pitch and roll angle of the robot

Maps. As a frame of reference with respect to the environment a two-dimensional (Fig. 6 top right) and a three-dimensional map (Fig. 6 bottom center) have been implemented. They include a priori known structure, show the robot's position and field of view, range sensor information and the planned waypoints/path.

Distance information. A distance reference component is projected on the floor in front of the robot (Fig. 3 center, green lines and digits). This should support distance estimations of the operator. The distance measurements from the ultrasonic sensors are utilized to highlight obstacles (Fig. 3 red barrels) in the near surrounding of the robot.

Status information and warnings. The network, battery status and the current velocity is presented by bar charts to the user (Fig. 6 top left). If values reach critical states, e.g. proximity of obstacles or low battery, warnings are shown in the center of the interface.

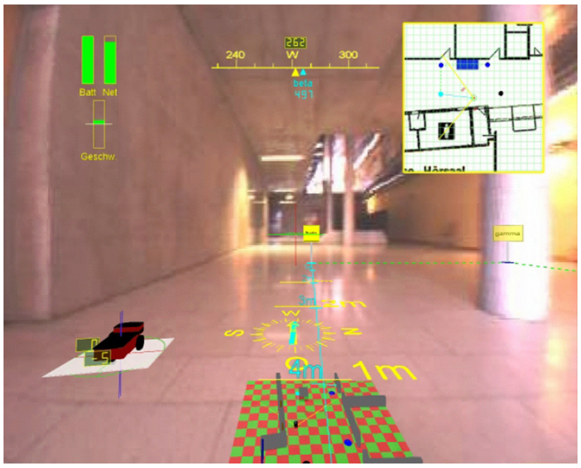

Fig. 6 AR interface with selected components

Fig. 6 shows the complete user interface with selected components. The system was designed in a way that the user can fade in and out the different possible graphical components according to the current needs and own preferences. The user can even fade out all the augmentations to have the full stereo video without any occlusions.

5. CHALLENGES, TESTS AND POTENTIAL

5.1 Technical challenges and results

The implementation of such a stereo AR teleoperation system is very challenging, because of the large amount of technical parameters, which can significantly influence the system performance.

The first parameters to optimize are the background video characteristics: frame rate, resolution, and delay. With the given constraints from the setup (WLAN bandwidth, available computing power for compression on robot) a stereo frame rate of more than 20 fps, a resolution of 320 x 240 (for the video, the projection has 1280x1024) could be reached. These parameters turned out to be adequate for the navigation task. Nevertheless, if details in the environment shall be identified a higher resolution is desirable.

The accuracy of the camera calibration, i.e. the registration of the virtual objects and the overlay, needs to be very high. The applied calibration methods proved to be suitable. The localization of the robot in the global frame of reference needs to be improved. Currently, it is only based on robot's odometry and therefore there is by-and-by an increasing position error between the overlayed global virtual elements (e.g. waypoint, planned path) and their position in the real world.

The implementation of the AR interface as stereo system introduced further challenges and requirements. Some of these requirements only appeared after different test persons used the system. One example of these partially very subjective requirements is accuracy of the alignment of the two camera heads. While some users could compensate several pixels of deviation in the vertical alignment of the two camera heads, others totally lost the stereo effect. Therefore, this deviation was also measured in the calibration process and the corrected images are shown to the user. Moreover, it is difficult to see very close objects. Humans normally slightly adjust their line-of-sight for each eye dependent on the distance to the object he/she focuses on, but the parallel cameras are fixed with respect to each other. This is a limitation of the current setup as the alignment cannot be adapted to the current distance to the focused object.

Various problems which only occur when using the system in stereo could be identified. For instance, using the HUD visualization of the artificial horizon in the center of interface as it is common in HUD displays of aircrafts and in mono interfaces, significantly disturbed the stereo impression of the operator and is therefore not applicable in this type of stereo AR interface.

In addition, a function for the operator to adjust the brightness of the background stereo video to the environment conditions (e.g. contrast, light) during remote control was implemented. This very helpful feature was a result of the already gained experience with this type of interface during the implementation phase. It is now possible to adapt the brightness of video to the current requirements and situation.

5.2 AR interface user tests

All user tests were performed with a principal test setup as described in section 3. The test participants were volunteers, mainly students with different study background. The operator was located in the room with the stereo projection wall without seeing or hearing the robot. Furthermore, they were not allowed to see the robot's environment before the test. However, the test participants knew the building before it was prepared for the experiments (e.g. with obstacles, objects to find …). The remotely controlled robot was videotaped for later evaluation. The participants were asked to fill in questionnaires and rate the different approaches and features. All experiments were performed with nine test participants. In the following an extract of some experiments and result is given.

At first various experiments, where the participants had to travel along a defined path (specified by narrow numbered passages), with different user interface configurations were performed. One of the important results in these tests was that the interface with stereo projection is only slightly better than that with only mono projection. A reason might have been that most of the test participants knew the building beforehand and therefore they had a model of distances, sizes and relations in this area in mind.

In the second experiment the test participants had to follow a virtual path defined with the waypoint system of the interface. These virtual path and the waypoint symbols were augmented to the stereo video, like e.g. in Fig. 3. All test participants appreciated this way of navigation support.

After these two experiments the test participants were asked how useful the different elements of the user interface are. Fig. 7 shows the results of this part of the questionnaire. Almost all elements were rated as useful, especially the waypoint system. The only element below the average of 3.5 was with an average of 3.22 the obstacle high lightening. From the comments of the test participants two reason could be derived. At first, this type of visualization hides too much of the camera image though the barrels are partially transparent. Secondly, most of the test participants did not remember that they can fade out and in the different interface components. This obstacle high lightening was in design only thought to be used and faded in, if required, e.g. when there are bad lightening conditions.

The third experiment was related to a search task. The test participants had to navigate through an area covered with obstacles and find certain objects. Two interfaces were used for this test. The interface presented here and an interface based on a virtual model of the environment together with a small video image from the robot (AV interface). The AV interface was designed to coordinate multiple robots with supervisory control and the focus in the design laid on global situational awareness. As main performance indicator for this experiment, the time until completion of the task was taken. For the AV user interface the average completion time was 1212 seconds (100%) and for the AR 464 seconds (38.2%). The hypothesis that the average is equal could be rejected with 99.8% using the Kruskal-Wallis test. This significant difference in the completion time validates the assumed better local situational awareness for the AR interface.

In the final discussion with the test participants about the AR interface, the received opinion was that there were no suggestions for further user interface

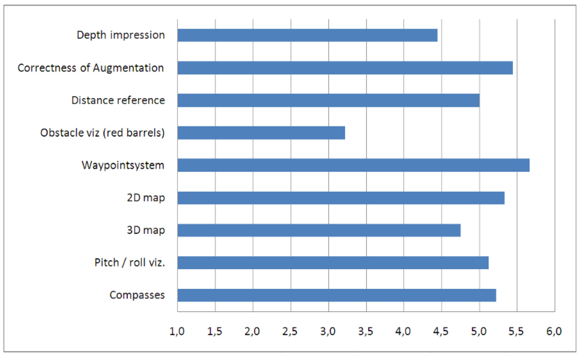

Fig. 7 Average rating of usefulness by test particpants (1 means useless/bad; 6 very useful/very good)

components. Most desirable would be a higher resolution of the stereo video or a second high resolution camera with a small field of view to inspect certain regions of interest in the large stereo video. The general opinion about the AR interface was that is suitable for this task. Deriving from their comments most of the test participants did not take consciously advantage of the stereo presentation. The reason here might be that stereo view is natural for humans and that the experiment scenarios were too easy to really take advantage of the direct depth impression. Nevertheless this is also a good indicator for the elaborate implementation of the stereo video and the overlay for the AR interface, because there were no disturbances for the test participants, regarding stereo.

6. CONCLUSION AND FUTURE WORK

The users tests carried out with the implement system validated the great potential of the applied AR techniques. It could be shown, that there is a significant increase in performance and situational awareness of the teleoperator. These initial tests promise a great potential for taking advantage of the AR user interface and its specific components in more challenging scenarios and environments. The stereo AR user interface approach presented here seems to be especially promising for tasks where a fast navigation and a good estimation of distances by the operator is required.

In the future work this system will be transferred to a faster, more agile robot, and the test will be performed in a more challenging unstructured three-dimensional outdoor environment. This also leads to higher technical requirements for the AR interface, e.g. the registration and calibration process. Therefore the future improvements of the presented system will include a more dynamic calibration and registration procedure which minimizes the overlay errors and respectively the possibly occurring confusion of the operator. The performed tests under these more challenging conditions will prove how far the AR approach can outperform other approaches for mobile robot teleoperation.

Another important element of the future work is the improvement and more detailed test of the single components of the user interface. This includes also the adding of further sensors for environment perception and their visualizations to support the increase of the situational awareness.

Finally, the possibility to adjust the viewpoint and respectively the frame of reference of the operator will be researched. Thomas *et al.* (1999) studied this for different tasks with a virtual reality simulation. The main application here was the coordination of different entities. It needs to be investigated, if the results can be transferred and adapted to the teleoperation task. For the presented approach here, this could be a chance to improve the performance and situational awareness with a larger virtual field of view than the real stereo video field of view.

ACKNOWLEDGMENTS

The authors want to thank the test participants for their contributions to this work.

REFERENCES

Azuma, R., Y. Baillot, R. Behringer, S. Feiner, S. Julier and B. MacIntyre (2001). Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications,* **21**, 34-47.

Driewer, F., M. Sauer and K. Schilling (2006) Mixed Reality for Teleoperation of Mobile Robots in Search and Rescue Scenarios, In: *Proceedings 12th IFAC Symposium on Information Control Problems in Manufacturing (INCOM)* Vol II, St. Etienne, France.

Kadous M. W., R. Sheh and C. Sammut (2006). Effective user interface design for rescue robotics. In: *Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. Salt Lake City.

Murphy, R. and J. L. Burke (2005). Up from the Rubble: Lessons Learned about HRI from Search and Rescue. In: *Proceedings of the 49th Annual Meetings of the Human Factors and Ergonomics Society*.

Nielsen, C. W. and M: A. Goodrich (2006). Comparing the usefulness of video and map information in navigation tasks. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human - Robot interaction,* Salt Lake City.

Piekarski, W. and B.H. Thomas (2002). The Tinmith System - Demonstrating New Techniques for Mobile Augmented Reality Modelling, In: *Proceedings 3rd Australian User Interfaces Conference*, Melbourne.

Scholtz, J., J. Young, J. L. Drury and H.A. Yanco (2004). Evaluation of Human-Robot Interaction Awareness in Search and Rescue, In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, New Orleans.

Skourup, C., T. Pettersen and J. Pretlove (2004). Augmented Reality System for Robot Programming. In: *Proceedings. IFAC/IFIP/ IFORS IEA Symposium Analysis, Design and Evaluation of Human-Machine Systems*, Atlanta.

Thomas, L.C., C.D. Wickens and J. Merlo (1999). Immersion and Battlefield Visualisation: Frame of Reference Effects on Navigation Tasks and Cognitive Tunneling. *Technical Report ARL-99-3 Aviation Research Lab - University of Illinois,* Illinois.