Proceedings of the
2009 IFAC Workshop on Networked Robotics
Golden, CO, USA, October 6-8, 2009

Towards a Predictive Mixed Reality User Interface for Mobile Robot Teleoperation

Markus Sauer * **Martin Hess** ** **Klaus Schilling** **

* *Center for Telematics, Würzburg, Germany (e-mail: markus.sauer@telematik-zentrum.de).*
** *Department of Computer Science VII, University of Würzburg, Am Hubland, Würzburg, Germany (e-mail: {hess,schi}@informatik.uni-wuerzburg.de).*

Abstract: Lack of situation awareness significantly decreases the performance in missions where a mobile robot is operated by a human from remote. The user interface is a key influencing element for situation awareness of the human operator. The information from the remote site is limited to what the robot's sensors can provide. In addition, this information is in general only delivered with a certain - maybe varying - communication delay. Predictive displays provide a promising approach to cope with these problems. In order to increase the situation awareness for the human operator in teleoperation scenarios predictive user interfaces can be used to achieve an artificial excoentric view.
This work presents an approach how a predictive mixed reality user interface can be realized with the help of motion control theory. The human operator commands a virtual robot projected into the camera image delivered to the human operator from the physical robot. Hereby a trajectory for the real physical robot is generated, which is executed by the physical robot after a certain time. Combined with mixed reality technologies an artificial, exocentric view of the mobile robot is achieved which leads to a short time predictive user interface for mobile-robot teleoperation.

Keywords: teleoperation, human-machine interface, mobile robots, telematics, control

1. INTRODUCTION

Recent advances in robotics enable more and more autonomous behaviors for mobile robots in real world environments. Nevertheless, there are a lot of applications for mobile robots like e.g. search and rescue, where in many cases a human is still required to operate the robot from remote. In a teleoperation system the human is included in the control loop of the robot. Sheridan (1992) provides with his supervisory control model a theory for this kind of systems. Assistive (semi-) autonomous behaviors, drive assistance systems and the user interface for the human operator are major influencing aspects for the performance of a teleoperation system for mobile robots.

The importance of these elements for a teleoperation system is mainly caused by the fact that the understanding of the human operator about the robot's situation and its surrounding is limited to the information he/she receives from the sensors of the robot and how things are preprocessed for and presented to the human with the user interface. This aspect from the human factors area is often summarized as the task to maintain situation awareness (Endsley (2000)). When (semi-)autonomous behaviors and drive assistance systems are included in the teleoperation system also the issue to maintain common ground (Stubbs et al. (2007)) is a major design aspect. Common ground means in this case that the human operator understands what the operated robot does and why it is behaving like this. In teleoperation systems, where this issue is not considered in the user interface, often the human operator gets confused by the robot's behavior and loses the feeling of control, such that he/she prefers the (semi-)autonomous behaviors to be disabled.

Another important aspect, when realizing user interfaces for teleoperation of vehicles, is the fact that different frames of reference introduced by the user interface cause the need of mental rotations by the human operator (Wickens et al. (2005)), such that the overall performance of the system might be reduced. Many user interfaces for mobile-robot teleoperation neglect this very important aspect. Mixed reality (Azuma et al. (2001)) as technology enables to realize integrated displays, where many things can be displayed in correct spatial relation, such that the need for mental rotations for the human operator is minimized.

Nielsen et al. (2007) suggest an user interface, where such type of integrated display is realized with augmented virtuality and an exocentric viewpoint of the operated robot. The environment is modeled as a virtual three dimensional world and the received camera image is projected on a plane in front of the robot. The results of the performed user tests in this work supports the assumption that the exocentric viewpoint (slightly behind and above the operated vehicle) improves the situational awareness of the human operator like it was also stated in Wickens et al. (2005). The exocentric view in this work leads to the fact that the virtual three-dimensional objects are no longer correctly aligned with the real objects seen in the two-dimensional camera image.

Another augmented virtuality user interface designed for the operation of multiple robots was proposed in Driewer et al. (2007). A pure augmented reality user interface was realized in Sauer et al. (2007), where it could be shown that also the egocentric viewpoint with the correct spatial alignment of user interface elements also leads to a performance increase.

A virtual reality predictive display was introduced in the ROTEX project (Hirzinger et al. (1994)) and proved to be advantageous. A predicted virtual representation of a robot manipulator was used to operate a physical manipulator in space over a longer time delay.

In Sugimoto et al. (2005) augmented reality is applied for mobile robot teleoperation combined with an exocentric viewpoint. The authors augmented past images which were stored in a database during the robot's movement with a virtual model at the current position of their tele-operated mobile robot. Their approach also supports the hypothesis that an exocentric view of the human operating the robot leads to a better situation awareness. The physical placement of the camera at the required position in order to achieve this exocentric view would increase the size of the robot to an undesirable size. With the system in Sugimoto et al. (2005) it is possible to gain this exocentric view without increasing the robot size. Nevertheless, with their approach the authors loose the important live characteristic of the camera image received from the robot. If there is a change in the environment or something unexpected happens, this can not be recognized because the system presents only past images.

In the work presented here an user interface is proposed, where the advantages of the exocentric viewpoint are combined with the correct spatial alignment of all user interface elements in an augmented reality user interface with a live camera image. The human operator operates a virtual robot which is projected correctly registered into the live view from a camera placed on a physical mobile robot. By commanding the virtual robot a reference trajectory is generated which is followed by the physical robot after a certain selectable time (cf. Figure 1). This approach has various important advantages. The applied techniques from motion control and augmented reality lead to a teleoperation user interface with an exocentric view. Wrong planned paths and movements can be canceled before execution due to the predictive nature of the system. Additionally the system gets more robust with respect to delay in the complete teleoperation chain.

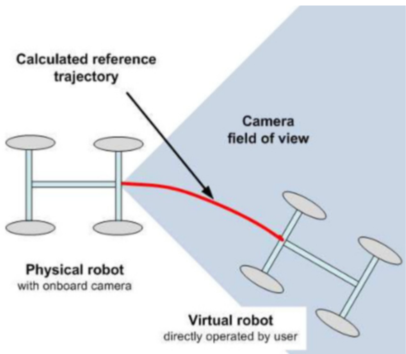

Fig. 1. Principle of proposed predictive user interface.

The remainder of this work presents the concepts of this predictive mixed reality system, the underlying motion control theory and the implemented demonstration system. Finally the conclusion and an outline of the planned future work is given.

2. MIXED REALITY PREDICTIVE DISPLAY

When designing a teleoperation system and its components the different timings and delays in the system need to be considered. Figure 2 gives an overview of a typical teleoperation setup. First of all there is a certain time for each sensor i for data acquisition and pre-processing (Δt_{ai}). Then there is a time which is needed to communicate and to present the data to the human operator (Δt_{com}) and finally there is a delay caused by the time a human needs to percept and react to a certain presentation of information (Δt_p). All these delays can cause significant disturbances in the closed loop system with the human as controller.

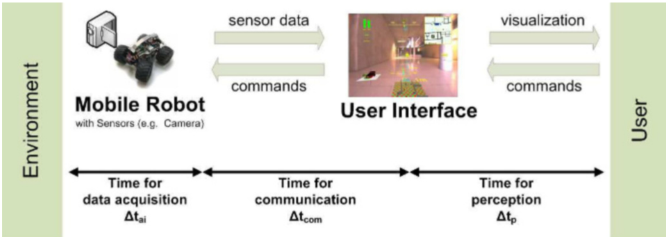

Fig. 2. Tele-operation chain with most significant delays.

Thus, here a concept of a predictive mixed reality display is proposed where these delays can be neglected up to a certain limit. As described before the human operator operates a virtual robot in the camera image which is retrieved from a physical robot with a certain delay. By operation of the virtual mobile robot with a certain prediction horizon of $\Delta t_{L,F}$ a reference trajectory is generated which the physical robot subsequently follows. The human operator gets instantaneous feedback to his/her commands through the virtual model of the robot without disturbances caused by delays in the system.

Figure 3 shows the timing of the different system events at time t. At time t the (old) sensor data (e.g. camera image, position, distance measurements) from time $t-\Delta t_1$ is presented to the human. The human generates through operation of the virtual robot a reference trajectory for $t+\Delta t_3$, such that a total time difference between physical and virtual operated robot of $\Delta t_{L,F} = (\Delta t_3 + \Delta t_c + \max t_{ai})$ results. Δt_3 denotes an additional prediction time parameter which can be selected during the design phase of the interface. It should be larger than the time for perception Δt_p ($\approx$ 100ms). The whole concept is applicable as long as the virtual operated robot stays in the field of view of the delayed camera image, what also determines the limit for the prediction time. In principle the different sensors on the robot need an additional time synchronization which was neglected for this work.

The different components and the data-flow in between these components to realize this concept are shown in Figure 4. Based on the received pose data from the mobile robot $p(t-\Delta t_2)$ and the joystick position $J(t)$, the motion

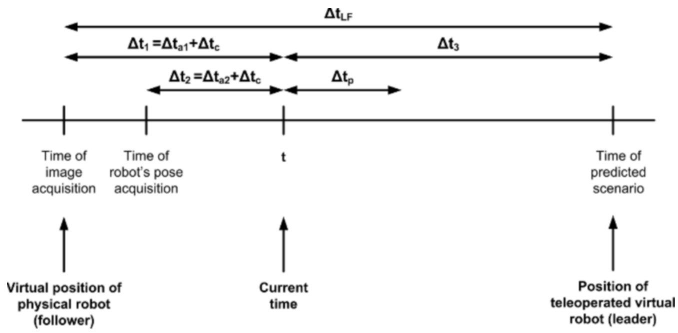

Fig. 3. Timing of different system events for the predictive mixed reality system.

control module calculates the position and orientation for the virtual operated robot $p(t + \Delta t_3)$ and the virtual representation of the physical robot $p(t - \Delta t_1)$. In addition the control inputs $u(t)$ for the mobile robot are generated. Together with the intrinsic and extrinsic calibration data, this leads to the integrated user interface visualization of the camera image I, the virtual operated robot VR_L and the model of the physical model VR_F.

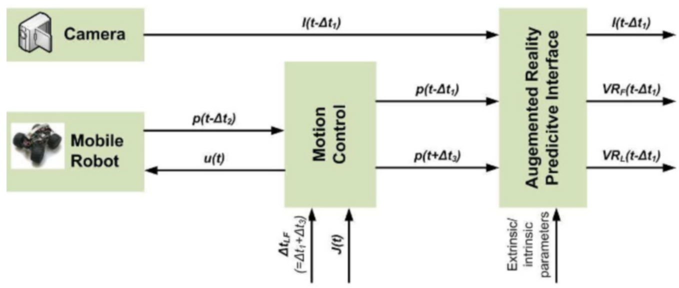

Fig. 4. Overview of dataflow for the user interface.

For the system proposed here, a robot with Ackermann steering was chosen, because the kinematic constraints of a robot of this type make it easier and more natural for the user to stay in the limited field of view with a fixed mounted camera on the physical robot. To overcome this limitation in future a pan tilt unit for the camera on the robot or other robot configurations (e.g. differential drive) can be used.
To make this kind of predictive interface, a robust reference trajectory generation and trajectory tracking is needed, that will be described in chapter 4.

3. MIXED REALITY

Due to the enormous increase in available computational and graphical power of todays computers, mixed reality systems get more and more applicable (Azuma et al. (2001)). In this work from the spectrum of mixed reality interface an augmented reality interface was implemented. As it can also be used as an augmented virtuality interface, in the remainder of this paper the term mixed reality will be used which covers both.

Mixed reality user interfaces are especially interesting for mobile robot tele-operation because they enable a correct spatial alignment of the different sensor data from robot with a camera image representing a very rich information for the human from the real remote environment. Nevertheless, they raise a lot of technological challenges, e.g. calibration of the camera image with the virtual world, implementing an overlay of virtual and camera information which runs in real-time, occlusion handling, etc.

For the demonstrator developed in this work the intrinsic parameters of the camera (projective properties of the camera) and the extrinsic parameters were determined by standard photogrammetric methods with a chessboard as reference. For the presented concept, the extrinsic calibration denotes the identification of the relative position and orientation with respect to the robot's reference point.

From the calibration data, a calibrated projection plane for the camera image could be generated in the virtual world of the user interface. With the extrinsic parameters of the camera calibration this projection plane could be placed correctly on the virtual representation of the physical robot. Hereby, an augmented reality overlay can be realized if the viewpoint is in the virtual representation of the physical camera position and looking in direction of the optical axis of the physical camera. Figure 5 shows a screenshot illustrating the modeling of the different elements in the Java3D scenegraph. The occlusion problem is not yet considered in this work.

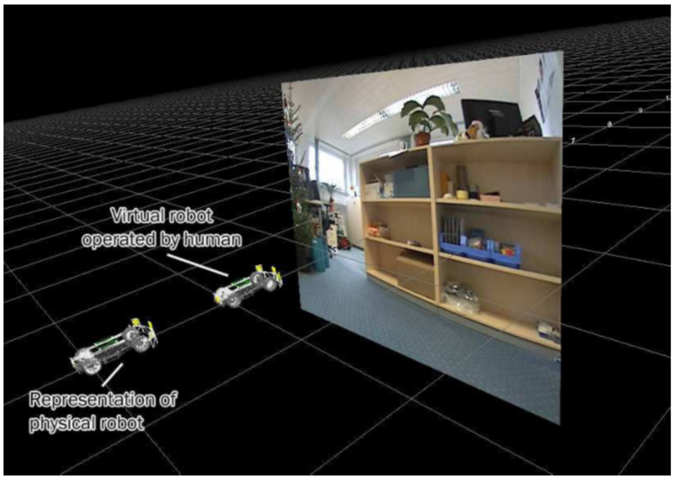

Fig. 5. Sideview (viewpoint not on the physical camera axis) of the realized three-dimensional system showing the modeling of the camera image, the physical robot, and the operated virtual robot.

4. MOTION CONTROL

This section describes how the motion control of the virtual and the real vehicle are realized. First, a method to achieve robust tracking of a feasible reference trajectory is outlined. After this it is clarified how such a reference trajectory could be generated from the joystick inputs of the operator.

4.1 Trajectory Tracking

For the approach to work successful it is essential that the real vehicle follows the trajectory of the virtual one with an error as small as possible. Since the output of the mobile robots actuators depends highly on the current charging level of the batteries, it is not sufficient to utilize open-loop control, i.e. to apply the same control inputs as applied to the model of the robots virtual counterpart. The idea of our approach is to extract a trajectory from the joystick inputs and apply a feedback control to track it while compensating uncertainties.

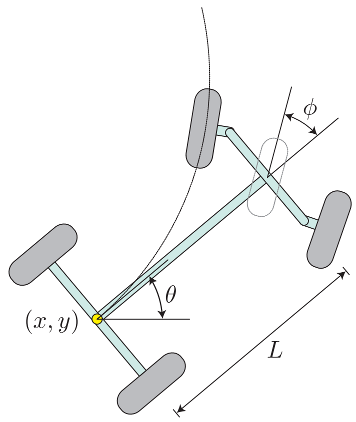

Fig. 6. Illustration of the car-like robot model.

The kinematics of a rear-wheel driven car-like robot (see Fig. 6) in the plane that is actuated by forward and steering velocity can be described with the following coupled nonlinear differential equations:

$$\begin{aligned} \dot{x} &= u_1 \cos\theta, \\ \dot{y} &= u_1 \sin\theta, \\ \dot{\theta} &= \frac{\tan\phi}{L} u_1, \\ \dot{\phi} &= u_2. \end{aligned} \tag{1}$$

The state vector of the vehicle consists of the cartesian coordinates $[x, y]$, and its orientation θ measured against the x-axis. The control vector consists of the vehicles speed u_1 and steering velocity u_2, and L denotes the distance between front and rear axle (LaValle (2006)).

During motion the orientation vector of the robot corresponds to the tangent in the current position on its trajectory. Therefore, a reference trajectory for a car-like robot is fully defined by its coordinates $[x_d(t), y_d(t)]$.

It was shown e.g. in De Luca et al. (1998) that the car-like model (1) can be input-output linearized in $[x, y]$ by means of dynamic feedback linearization. As a result one receives the nonlinear dynamic feedback controller

$$\begin{aligned} u_1 &= \xi_1, \\ u_2 &= \frac{-3\xi_2 \cos^2\phi \tan\phi}{\xi_1} + \frac{Lcos^2\phi}{\xi_1^2}(r_2 \cos\theta - r_1 \sin\theta), \\ \dot{\xi}_1 &= \xi_2, \\ \dot{\xi}_2 &= \frac{\xi_1^3 \tan^2\phi}{L^2} + r_1 \cos\theta + r_2 \sin\theta, \end{aligned} \tag{2}$$

and the original system is transformed into the equivalent linear system

$$\dddot{z} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \tag{3}$$

under the state transformation

$$\begin{aligned} z &= \begin{bmatrix} x \\ y \end{bmatrix}, \\ \dot{z} &= \begin{bmatrix} \xi_1 \cos\theta \\ \xi_1 \sin\theta \end{bmatrix}, \\ \ddot{z} &= \begin{bmatrix} -\xi_1^2 \tan\phi \sin\theta / L + \xi_2 \cos\theta \\ \xi_1^2 \tan\phi \cos\theta / L + \xi_2 \sin\theta \end{bmatrix}. \end{aligned} \tag{4}$$

With this controller it is easy to achieve trajectory tracking by applying a global stabilizing feedback to the linear system:

$$\begin{aligned} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} \dddot{x}_d \\ \dddot{y}_d \end{bmatrix} + k_a \left(\begin{bmatrix} \ddot{x}_d \\ \ddot{y}_d \end{bmatrix} - \ddot{z} \right) \\ + k_v \left(\begin{bmatrix} \dot{x}_d \\ \dot{y}_d \end{bmatrix} - \dot{z} \right) + k_p \left(\begin{bmatrix} x_d \\ y_d \end{bmatrix} - z \right) \end{aligned} \tag{5}$$

where the feedback gains have to be chosen such that the polynomial

$$\lambda^3 + k_a \lambda^2 + k_v \lambda + k_p \tag{6}$$

is Hurwitz. For our application only the relative positioning of the real vehicle to the virtual one is important. Hence the position of the vehicle $[x(t), y(t)]$ which is needed for the feedback can be obtained by means of dead reckoning. For this purpose two wheel encoders and a gyroscope were utilized in the hardware implementation.

4.2 Reference Trajectory Generation

In order to tele-operate the vehicle with the described controller, a suitable method that generates a feasible reference trajectory from the operator's joystick inputs is needed.

The main problem is that in general the dynamics of the vehicles actuators are too slow to follow high frequency joystick commands accurately. To avoid this, our implementation takes values from the joystick in a predefined time interval and smoothly interpolates the values in between. In the experiment we limited the maximum driving speed of the vehicle to $0.6\,m/s$ and it turned out that in this case, a sampling rate of half a second is an adequate tradeoff between smoothness and sluggishness. Since the steering angle servo easily can change its setting from maximum left to maximum right and vice versa during that interval we choose to fill the gaps with cubic splines that have zero first and second derivative in the end points. Fig. 7 shows a typical profile for the generated desired steering angle.

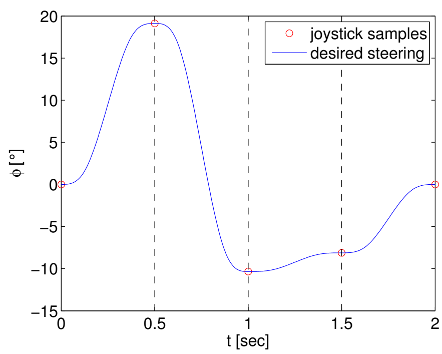

Fig. 7. Typical profile for the generated desired steering angle.

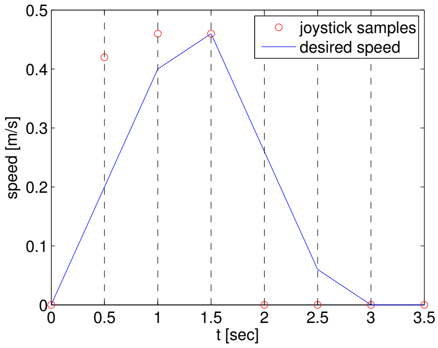

Fig. 8. Typical profile for the generated desired speed. Acceleration is bounded to $0.4\,m/s^2$.

In contrast to the steering angle, the speed of the robot cannot change so fast. Therefore it is essential to put an upper bound on the absolute value of the robots acceleration. Even though the resulting curves level of continuity is higher by using cubic splines to fill the gaps, for bigger speed changes it leads to undesired oscillations along the movement direction of the vehicle. This is avoided by simply using linear approximation between the sampling points (see Fig. 8), which on the other hand leads to non-differentiability in most of the sampling points and consequently to tracking errors from a mathematical point of view. But in the hardware implementation these errors are not measurable due to its small value and the exponential convergence of the controller.

With the obtained values for desired speed and steering angle it is now possible to numerically integrate the kinematic model (1) to obtain the values for the desired position $[x_d(t), y_d(t)]$ which are used to update the actual position of the virtual vehicle. For the real robot to follow the trajectory of the virtual one with a certain temporal offset, we calculate the first three time derivatives of the desired position by means of numerical differentiation and feed them into the controller (5) after a fixed time interval. If the discrete time of the implementation accidentally hits one of the non-differentiable points we just take a value an instant later.

5. DEMONSTRATION SYSTEM AND FIRST RESULTS

For the implemented system some simplifications have been assumed to quickly reach a first proof of concept system implementation. It was assumed that $\Delta t_{a1} = \Delta t_{a2}$ what means the robots pose and image were taken at the same time. This assumption caused no problems in the system setup used for this implementation. All gathering of sensor data and communication are fast enough such that no negative effect could be recognized due to this assumption.

In the implemented system a four wheel indoor MERLIN robot with Ackermann steering and onboard PC104 was used. The communication was realized over WLAN. As camera a standard webcam with a resolution of 640x480 pixels was used. The software has been written in C,C++ and Java. The three dimensional user interface was implemented in Java3D.

Occlusion handling for the augmented reality overlay was not considered in this implementation. As these incorrect occlusions of the real world by the operated virtual robot, will cause significant disturbance to the user, this needs to be addressed in future work to make such kind of predictive augmented reality interface really applicable. These expected occlusion problems could already be investigated by first test runs where the human operators were asked to drive around a sharp corner.

The first tests with different people has already provided interesting feedback and proved the potential of the concept. It was mentioned that they had the impression to really operate the virtual robot in the real world and did not think anymore about the real physical robot. The difference between the input commands of the virtual robot issued by human and the change of the camera position according to the delayed position change through the movement of the real robot did not disturb significantly in open areas. This is a very good indicator that our overlay, reference trajectory generation and tracking works well. Future performance and situational awareness tests have to prove the advantages of the concept quantitatively. Figure 9 shows the user interface as it is seen by the human operator. The virtual robot which is commanded by the human and used to generate the trajectory of the physical robot, is here overlayed spatially correct aligned in the middle of the camera image.

Fig. 9. A screenshot from the predictive user interface where the virtual robot is operated.

In addition, the way how the whole user interface is modeled in 3D, allows for any other viewpoint of the generated world (cf. Figure 5) in order to investigate any spatial relations of gathered data. This is especially interesting if other additional information is integrated into the system (e.g. laser measurements, map information, position of other robots...).

6. CONCLUSIONS AND FUTURE WORK

In this work a concept to reach a short-term predictive mixed reality interface and the first implementation were presented. Due to the implemented prototype we expect good performance results for future experiments with test persons. The first trials showed that it is possible to hide short time-delays during the tele-operation control loop from the human operator.

As the first trials proved that this user interface is a very promising approach, in future a lot of extensions might be realized. Currently the camera is fixed on the physical robot, thus, to not confuse the human operator the virtual robot movement and the reference trajectory is limited to the field of view the camera. This limitation can be removed by the application of a pan and tilt unit, which allows to keep the virtual robot in the field of view of the human operator independent from the specific field of view of the camera. A different approach would be to use mobile robots with different kinematics (e.g. differential drive). In order to utilize the user interface with other robot configurations the motion control and the robot models inside the motion control component need to be adjusted. Another aspect for further investigation would be an intuitive way for the human operator to pause, start and reset the execution of the delayed commands. This would enable a much more comfortable way to use the predictive user interface for mobile robot tele-operation.

Important from the augmented reality point of view will be to integrate an occlusion handling with the help of range sensors mounted on the mobile robot. This will help to significantly reduce confusion of the human operator which might raise due to incorrect overlays with respect to obstacles. In a perfect overlay they should occlude the virtual robot and are currently occluded by the virtual robot, because the virtual robot is simply rendered on the camera image without checking the occlusion conditions.

Other points for future work are the integration of predictive assistance systems. For instance it is possible integrate an obstacle avoidance for the virtual operated robot in combination with a laser range finder on the physical robot. If a possible collision of the virtual robot is detected according to the range data, the execution of the delayed commands can be stopped. Also some additional local autonomy functions on the mobile robot (e.g. local obstacle avoidance) will make the concept more robust against disturbances.

Currently this concept is limited to flat surfaces. Therefore it would be helpful to integrate a physical model of the used mobile robot for the reference trajectory calculation of the virtual robot. Together with a local elevation map generation in the area of the moving virtual robot, this would help to bring this kind of user interface also to unstructured or real three-dimensional worlds. Another additional application area of this type of short time mixed reality predictive user interface would be to use it for step wise planning of manipulator movements before execution of the actual movement for instance for bomb disposal tasks.

ACKNOWLEDGEMENTS

This work was partially supported by the Elitenetwork of Bavaria through the program "Identification, Optimization and Control with Applications in Modern Technologies".

REFERENCES

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6), 34–47.

De Luca, A., Oriolo, G., and Samson, C. (1998). Feedback control of a nonholonomic car-like robot. In J.P. Laumond (ed.), *Planning robot motion*. Springer-Verlag.

Driewer, F., Sauer, M., and Schilling, K. (2007). Design and evaluation of a teleoperation interface for heterogeneous human-robot teams. In *Proc. 10th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems (HMS), Seoul, Korea*.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M.R. Endsley and D.J. Garland (eds.), *Situation awareness analysis and measurement*, chapter 1, 3–26. Lawrence Erlbaum Associates. ID002.

Hirzinger, G., Landzettel, K., and Fagerer, C. (1994). Telerobotics with large time delays-the rotex experience. In *Proc. IEEE/RSJ/GI International Conference on Intelligent Robots and Systems '94. IROS '94*, volume 1, 571–578.

LaValle, S.M. (2006). *Planning Algorithms*. Cambridge University Press.

Nielsen, C.W., Goodrich, M.A., and Ricks, B. (2007). Ecological interfaces for improving mobile robot teleoperation. *IEEE Transactions on Robotics*, 23(5), 927 – 941.

Sauer, M., Driewer, F., and Schilling, K. (2007). Potential and challenges of stereo augmented reality for mobile robot teleoperation. In *Proc. 10th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems (HMS), Seoul, Korea*.

Sheridan, T.B. (1992). *Telerobotics, automation, and human supervisory control*. MIT Press, Cambridge, MA, USA.

Stubbs, K., Hinds, P.J., and Wettergreen, D. (2007). Autonomy and common ground in human-robot interaction: A field study. *IEEE Intelligent Systems*, 22(2), 42–50.

Sugimoto, M., Kagotani, G., Nii, H., Shiroma, N., Inami, M., and Matsuno, F. (2005). Time follower's vision: A teleoperation interface with past images. *IEEE Computer Graphics and Applications*, 25(1), 54–63.

Wickens, C.D., Vincow, M., and Yeh, M. (2005). *Handbook of visual spatial thinking*, chapter Design Applications of Visual Spatial Thinking: The Importance of Frame of Reference, 383 – 425. Cambridge University Press. ID001.