Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Omnidirectional Robotic Telepresence through Augmented Virtuality for Increased Situation Awareness in Hazardous Environments

Daniel Johansson
MSE Weibull AB
Älmhult, Sweden
daniel.johansson@mseab.se

Leo J. de Vin
Centre for Intelligent Automation
University of Skövde
Skövde, Sweden

***Abstract*—This paper proposes a novel low-cost robotic telepresence approach to situation awareness, initially aimed for hazardous environments. The robot supports omnidirectional movement, wide field of vision, haptic feedback and binaural sound. It is controlled through an augmented virtuality environment with an intuitive position displacement scheme that supports physical mobility. The operator thereby can conduct work away from danger whilst retaining situation awareness of the real environment.**

***Keywords*—telepresence, augmented virtuality, omnidirectional movement, haptic feedback, binaural.**

I. Introduction

It is vital for humans to have a high level of situation awareness in a great number of application areas. One such area is work within hazardous environments, where humans are exposed to increased levels of danger. Awareness of changes in an often dynamic environment is crucial for the safety of both the person conducting the work as well as for persons working or living in close proximity. To decrease the level of risk for personal injury, a versatile telepresence robot would enable the user to be situated away from potential harm while still retaining vital sensory capabilities.

Extensive work has been done within the areas of robotic telepresence and haptic feedback. In [4] for example, a system is presented in which VR is utilized to enhance the feeling of presence as the user remotely controls parts of an advanced humanoid robot. Commercial robotic telepresence products, such as [13] and [14], tend to provide good performance but price can limit their use in cost sensitive applications.

This paper presents a complete robotic telepresence system that is a versatile and cost-effective platform focused towards remote presence. Through the use of an accompanying control environment, the system is able to accurately match human movement and manipulation through natural interaction.

This work was supported by the Knowledge Foundation and MSE Weibull AB and is part of the graduate school: Intelligent Systems for Robotics, Automation and Process Control (RAP) at Örebro University in Sweden.

II. System Design

To retain as much of the user's normal level of control and situation awareness, the robot and its interaction system must match a set of human features. This set of factors leads to a number of canonical system requirements, listed below:

- Untethered omnidirectional movement capability; a human can normally translate and rotate freely in any direction on the ground plane.
- Wide and high-resolution field of vision; the human peripheral vision system is used both for object detection and balance.
- Interaction with haptic feedback; a human is able to interact with objects through hands and arms, relying on the sense of touch for muscle torque and pressure adjustments.
- Hearing; directional sound apprehension can contribute greatly to awareness in a dynamic environment.
- Low intersensory latency; humans exist and utilize their senses in real-time and noticeable delays in the robot's sensory or propulsion systems can be a safety threat and also break the feeling of presence.

To complement the robot, a user interface is required that facilitates the utilization of the robot's capabilities. In [1]-[3] an immersive augmented virtuality simulator is presented by the authors that uses a three meter in diameter blue cylinder, software based chroma keying [11] and a video see-through head mounted display (VSTHMD). The chroma keying algorithm is applied to the VST camera images to digitally render the blue color transparent, which are subsequently overlaid on a computer-generated virtual environment. The VSTHMD, shown in Fig. 1, is custom-made to enable a high field of vision of 110 x 55 degrees, at a low cost. Real-time camera lens distortion reduction is performed on the graphical processing unit through Bezier patches, displayed in Fig. 2, by which the camera image is warped. Furthermore, wireless connectivity and 6 DOF tracking, combined with natural position displacement allows users of the simulator to physically move untethered. This form of simulator technology

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

Figure 1. Body-mounted simulator hardware.

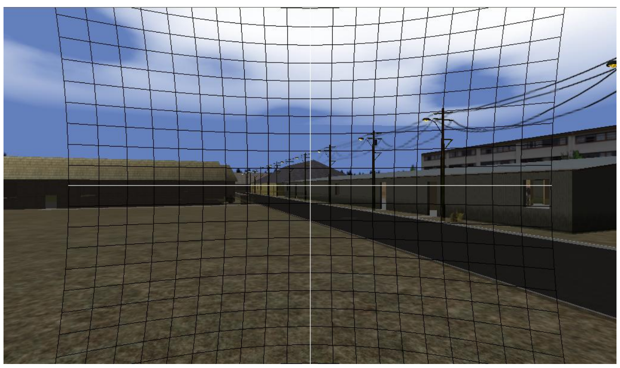

Figure 2. An example of a virtual environment with Bezier grid overlay.

allows mixed environments to be created that facilitates the users to view themselves, physical items and simultaneously other users within a virtual environment.

When the simulator is connected to a camera-equipped robot which is designed according to the previous requirements, the simulator's virtual environment can be substituted for the robot's camera images. The human operator is then able to control the robot by physical movements as if she or he was in the robot's environment. Furthermore, the robot camera motion can be coordinated with the operator's head movements while the operator's arms control the robot arms. These abilities combined with the fact that the operator, through chroma keying, can see herself or himself in the robot's environment, could increase the situation awareness of the operator. Augmented cognition is also possible with the system, by overlaying additional computer generated visual information.

This form of system design creates a human-in-the-loop regulator, [5] and [6], where one part of the loop is proportionally controlled directly by the user and another is indirectly controlled by a PID regulator.

The structure of the paper is as follows. Firstly, an overview of the robot implementation is given. This is followed by an explanation of the system's method of position displacement, an evaluation of the system and ends with concluding remarks and methods for improvement.

III. Robot Implementation

A. *Untethered Omnidirectional Movement*

In relation to the first of the system requirements given in section II, the robot must be able to follow the human operator and therefore be able to simultaneously translate and rotate on the ground plane. To provide the necessary speed and acceleration, the robot has four geared DC motors, as in [7] and [12], which are mounted 90 degrees apart. Mounted on each motor is a wheel equipped with free-turning rollers at the periphery, as shown in Fig 3. The wheels thereby can provide traction in the direction normal to the motor axis, whilst it can slide freely in the motor axis direction.

A motor pair, (i.e. two motors mounted 180 degrees apart), is used to translate the robot in one dimension. The other pair is used for perpendicular translation, thereby achieving two-dimensional movement. To rotate the robot, all motors rotate in the same direction. This arrangement allows the robot to match human omnidirectional movement. Even though three motors would be sufficient to achieve the necessary motion, four motors enable more efficient translation and a higher maximum velocity, since the motor pairs are mounted orthogonally to each other. In addition, the use of four motors provides the robot with increased mechanical stability.

The simulator's tracking produces 6DOF position information, which is measured relative to defined marker patterns on the user. For robot movements, a marker pattern is situated on the back of the operator's vest containing the body-mounted simulator hardware.

Untethered robot control is gained through an ultra-wideband (UWB) USB hub, to which the motor and servo controllers are connected. No calculations are required to be made onboard, so the operation of the platform is both rugged and low-powered.

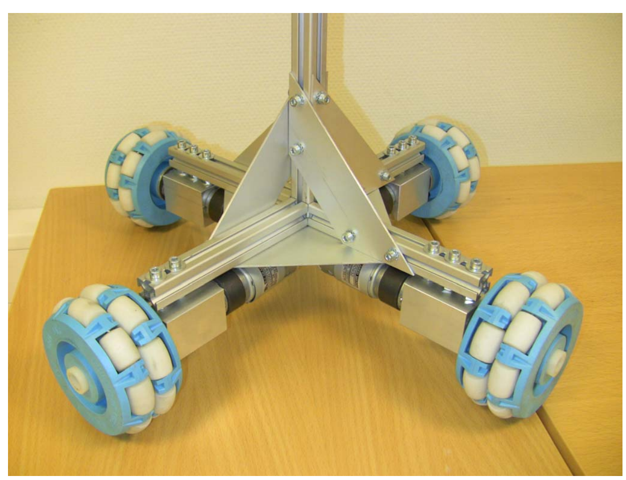

Figure 3. Omnidirectional platform.

An OceanServer tilt and deviation compensated, digital 3-axis compass is utilized to correlate the simulator's coordinate system with that of the robot.

B. Vision System

The operator receives visual information from a camera on the robot. Given that high resolution images are required in order to provide for the necessary level of detail when using a wide field of vision a HDMI camera with full HD resolution (1920x1080) is employed. The camera is equipped with a wide angle lens to match the VSTHMD. Novel wireless HD links are utilized to transmit the camera images. At the moment UWB is used as it is available, but other technologies not yet on the market can provide a non-line-of-sight transmission. Until then, to maximize transmission integrity, the receiver is placed as close as possible to the working area of the robot. The signal is then relayed through an HDMI cable to the simulator. There it is captured, altered and sent wirelessly to the operator.

The camera is mounted on a pan and tilt platform controlled by hobby servos. A marker pattern on the VSTHMD is used to control the platform to match head movements. Thus, head movements are independent of robot rotation. This is to enable the operator to work with a stationary robot body and arms, whilst allowing the movement of the head for a better overview.

C. Interaction and Haptics

The robot employs two 5DOF servo-controlled arms equipped with grippers. The arm servos position demands are calculated based on the position of marker patterns on the operator's controllers.

The joints of the wrists, as well as of those of the shoulders, are calculated directly from the tracking system and current joint positions. Only the two servos on each arm that are used for combined forward and backward movement of the wrist requires inverse kinematics calculations. The problem is therefore reduced to 2DOF in 2-dimensional input space. However, to avoid unnecessary calculations at execution time, forward kinematics is used to create precalculated look-up tables so that the joints and corresponding 2-dimensional position values are deduced. The servos used on the robot have a resolution of approximately 0.1 degrees, which limits the maximum size of the data arrays. More computational power is therefore left over for propulsion, image acquisition and chroma keying.

To manipulate the opening and closing of the grippers, two Bluetooth controllers, Wii Remotes, are employed. One button controls the opening of the gripper, whilst another controls the closing.

The gripper claws are equipped with thin film force sensors. These provide linear feedback of the force detected between claws and objects. To relay the amount of force detected by the force sensors to the operator, the vibrators within the Wii Remotes are used. This forms a low-cost implementation of interaction and haptic feedback when compared to, for example [8], [9] and [10]. However, depending on the application requirements and budget, the system could be fitted with more complex equipment.

Other omnidirectional platforms, such as hexapods, could be used if this would be beneficial for the application, and interaction and haptics could then be integrated in the platform's legs.

Fig. 4 shows an example of the user's view within the simulator. The image illustrates how the left robot arm is controlled by the user. Also visible are the tracking markers situated on the controller. The transparency of the overlaid image can be regulated in 256 levels. This is to ensure that the user's hands do not obstruct crucial areas within the robot's view.

D. Binaural Sound

Binaural sound is used to allow positional sound to be forwarded to the operator. The camera is mounted inside a model of a human head equipped with external ears. A microphone is mounted at each ear canal entrance and the signal is routed to the operator unaltered. This setup enables the use of the natural human positional sound system. However, the positional information can only be approximated without an exact copy of the operators head and ears.

E. Latency

Only real-time capable hardware is used in the simulator to minimize the latency. Virtually no latency can be accepted in the wireless vision interfaces. A hardware accelerated HD capture card and a high performance multi-core processor are also utilized.

The joints of the robot are controlled by high speed and low-cost servos and controllers. The gears attached to the robot propulsion motors are matched to provide velocities which are comparable to human walking speed.

F. Hardware

A block diagram over the entire system is shown in Fig. 5.

Figure 4. An example of the user's view through the video see-though head mounted display.

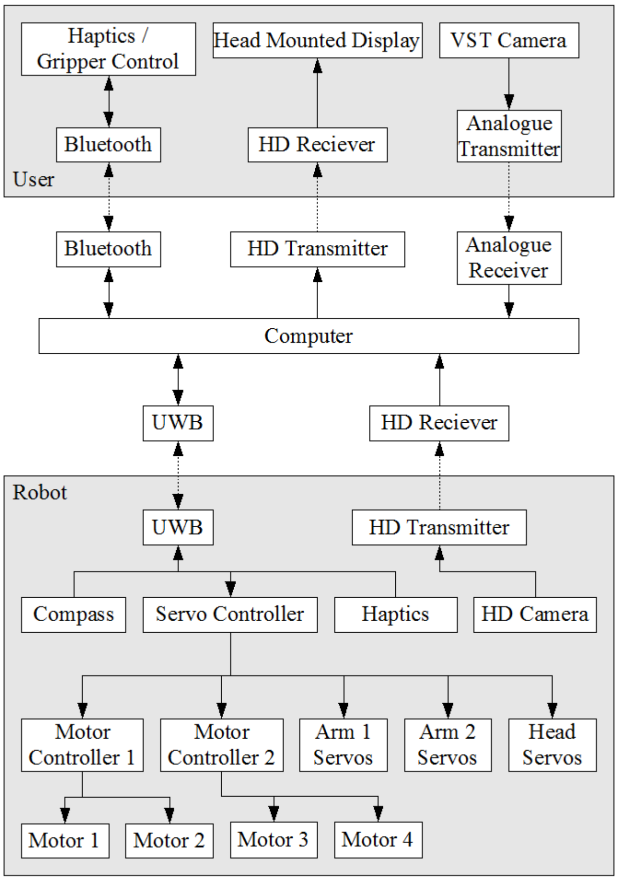

Figure 5. Hardware schematics

IV. CONTROL

A. Position Displacement

Fig. 6 is a plan view of the physical simulator environment from above, where the position displacement principle is illustrated. The circle represented by the radius r_o shows the physical outer boundary of the area in which the user is allowed to move. "A" and "B" represent the coordinate systems of the robot and user respectively. The angular offset between the two coordinate systems, labeled α_{pv}, is due to the rotational angle of the user. This angle is a process variable that is read from the compass onboard the robot. The angular set point, α_{sp}, is read from the tracking system and represents the user's current heading. The difference between α_{sp} and α_{pv} originates in the robot's failure to maintain correlation with the user due to for example fast user rotation or wheel slippage. This error value is used by a rotational velocity PID regulator. The innermost circle, represented by r_i, is a virtual border which confines the area of purely absolute correlation of user and robot translation direction and velocity. If the user is outside this border, a continuous robot translation is added to the absolute translation. The direction of this continuous translation depends on the current user position relative to the origin, while the velocity depends on the current distance from r_i. Examples of absolute and continuous translation are illustrated by "C" and "D" respectively. The continuous translation is added to enable robot translation in an area larger than the user environment. Mathematical formulas relating each of the quantities described above are given in the following section. The direction and velocity of the resulting robot translation is represented by "E". The radius of r_o is here instead the maximum robot velocity, v_{max}.

Using this position displacement principle, the robot's position and orientation can be controlled without any other input than physical user movement. A result is a user that is physically active, which can increase the feeling of awareness compared to a desktop setup.

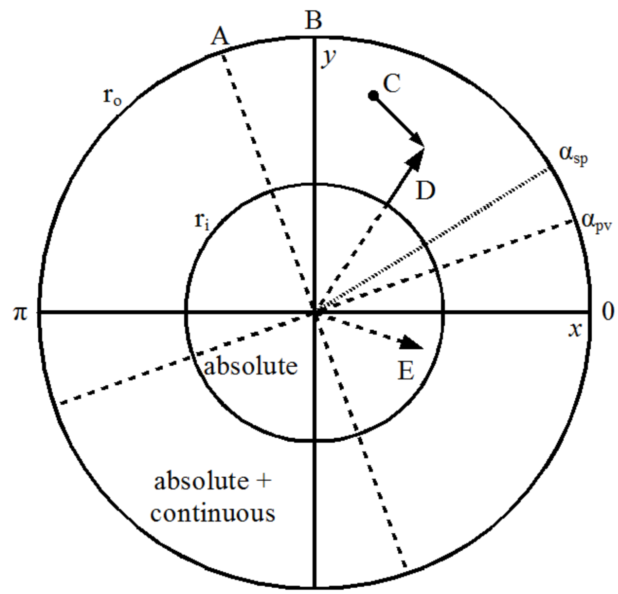

Figure 6. Robot position displacement.

B. Motor Control

To enable simultaneous translation and rotation in any direction, the control of the four propulsion motors is implemented by the following steps, and updated for each sample:

1) Absolute Components

"C" in Fig. 6 shows that the user has moved from the start to the end of the arrow within one update sample, written as *dx/dt* and *dy/dt*. The absolute velocity component for each motor pair is calculated by projecting the velocities into the correct quadrant in the robot's coordinate system. Equation (1) shows how the absolute velocity is calculated for horizontal robot translation. The horizontal contribution is multiplied with the total velocity. Forward/backward translation uses the same equation, although with sine instead of cosine.

$$v_{xa}=\cos\left(2\arctan\left(\frac{\frac{dy}{dt}}{\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}+\frac{dx}{dt}}\right)-\alpha_{PV}\right)*\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2} \quad (1)$$

2) Continuous Components

The continuous component, calculated as in equation (2), depends only on positional information of the user. By analogy to equation (1), the horizontal component is calculated by using cosine and the forward/backward component using sine. The output velocity is calculated using the distance that the user lies beyond r_i and the robot's maximum translational velocity.

$$v_{xc}=\cos\left(2\arctan\left(\frac{y}{\sqrt{x^2+y^2}+x}\right)-\alpha_{PV}\right)*\left(\frac{\sqrt{x^2+y^2}-r_i}{r_o-r_i}\right)*v_{max} \quad (2)$$

3) Rotational Component

The rotational velocity PID regulator, given in equation (3), generates a correctional velocity at each update sample. This velocity is common to all motors and added to each motor's calculated total velocity. The maximum physical rotational error is $\pm\pi$, and the PID output is scaled accordingly as described below.

$$v_r=\left(K_p e(t)+K_i\int_0^t e(\tau)d\tau+K_d\frac{de(t)}{dt}\right)*\frac{v_{max}}{\pi} \quad (3)$$

4) Scaling

The final motor velocities are calculated as shown in equation (4), which represents the horizontally moving motor pair. In case where any of the motor's velocity exceeds v_{max}, all velocities are downscaled correspondingly. Equation (5) calculates the scaled output to the motors based on the highest velocity, $v_{highest}$, among all four motors.

$$v_{x1,2}=v_{xa}+v_{xc}\pm v_r \quad (4)$$

$$v_{x,y\,scaled}=\frac{v_{max}}{|v_{highest}|}*v_{x,y} \quad (5)$$

5) Regulator

Fig. 7 shows an illustration of the system's regulator functions. The user is the human-in-the-loop as the robot relays visual and haptic information to be processed and reacted to directly by the user. The compass is only used for alignment of the user and robot coordinate systems. The user is responsible for correcting translational deflection, due to, for example surface inclination or uneven surface resistance. If the wheels start to slip the user will notice this through changes in the images and audio from the robot and then be able to react. Although the motors can be made to compensate for any incline, using the pitch and roll data available from the compass and/or angular rate counters on the wheels, this has not been implemented. The reason for this is, that the intention is to make the user aware of the robot's surrounding environment, by forcing the user to compensate for this motion.

With the exception of image and audio feed to the user, the update rate of the regulator loop is primarily limited by the compass. The compass can provide a refresh rate of maximum 40Hz with an accuracy of 1°.

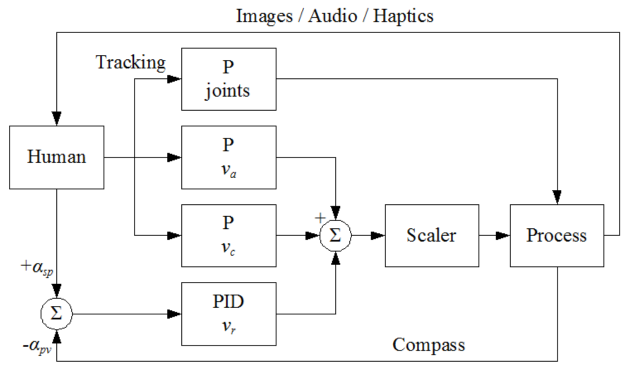

Figure 7. Regulator schematics

V. Evaluation

Early evaluation of the system implies that situation awareness and presence is improved when compared to a desktop configuration. The system can be extended to enable, for example, lifting, assembly and manipulation of real equipment within hazardous environments.

The dominant system latency originates in the acquisition of the robot camera images, which have a latency of approximately 20ms before they are presented to the user. The minor part of this is due to the relatively large amount of data that has to be transferred to the video memory for each frame. The major part is generated by the camera itself, which requires time to output each image. This problem can be addressed by using an alternative camera with lower latency.

VI. Conclusions

It is shown that a relatively low-cost robotic telepresence system can be built that allows a high amount of situation awareness and with several of the human capabilities transferred. The system supports intuitive and natural position displacement and correlated arm movements. The operator uses physical movements to control the robot instead of common desktop interaction or joysticks.

The system could be improved in a number of ways, depending on, for example, haptic and physical needs of the application. Examples of improvements would be stereoscopic vision and the use of manipulator arms with more joints and higher lifting capacity. Another example is control of the optical zoom of the camera, which can be implemented through multi-modal input.

Future work will include a thorough evaluation of usability and presence. The aim is also to further develop the body-mounted hardware to increase comfort and ease of use.

ACKNOWLEDGMENT

The authors gratefully acknowledge the advice of Professor Sillitoe, Department of Engineering and Technology, University of Wolverhampton, U.K., in the preparation of this paper.

REFERENCES

[1] D. Johansson and L. J. de Vin, “An augmented virtuality simulator with an intuitive interface: concept, design and implementation,” Proceedings of Virtual Reality International Conference VRIC 09, Laval, France, 2009.

[2] D. Johansson and L. J. de Vin, “A low cost video see-through head mounted display for increased situation awareness in an augmented environment,” Proceedings of Intuition 2008, Turin, Italy, 2008.

[3] D. Johansson and L. J. de Vin, “Design and development of an augmented environment with high user mobility for training purposes,” Proceedings of Mechatronics’08, Limerick, Ireland, 2008.

[4] Y. H. Seo, H.Y. Park, T. Han and H.S. Yang, "Wearable telepresence system using multi-modal communication with humanoid robot," Proceedings of ICAT 2003, Tokyo, Japan, 2003.

[5] J. Tang, Q. Zhao and R. Yang, “Stability control for a walking-chair robot with human in the loop,” International Journal of Advanced Robotic Systems, Vol. 6, No. 1, 2009, pp. 47-52.

[6] R. Stanciu and P. Y. Oh, “Feedforward control for human-in-the-loop camera systems,” Proceedings of the 2004 IEEE International Conference on Robotics and Automation, New Orleans, LA, USA, 2004.

[7] R. Rojas and A. G. Forster, “Holonomic control of a robot with omnidirectional drive,” Künstliche Intelligenz, Vol. 20, Nr. 2, 2006.

[8] C. Glover, B. Russell, A. White, M. Miller and A. Stoytchev., "An effective and intuitive control interface for remote robot teleoperation with complete haptic feedback," Proceedings of the 2009 Emerging Technologies Conference (ETC), Ames, IA, USA, 2009.

[9] A. Shiele and G. Visentin, “The ESA human arm exoskeleton for space robotics telepresence,” Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics and Automation in Space: i-SAIRAS 2003, NARA, Japan, 2003.

[10] R. J. Stone, “Haptic feedback: a potted history, from telepresence to virtual reality,” Proceedings of the Workshop on Haptic Human-Computer Interaction, Glasgow, UK, 2000.

[11] F. vd Bergh and V. Lalioti, “Software chroma-keying in immersive virtual environments,” South African Computer Journal, No. 24, 1999, pp. 155-162.

[12] http://seniord.ece.iastate.edu/may0835/robot_chassis.html, visited on 25/06/09.

[13] http://anybots.com, visited on 25/06/09.

[14] http://robodynamics.com, visited on 25/06/09.