Overlaying Graphics to Augment a Telerobotic Stereo-vision System

Syed M.S. Islam
School of Dentistry
The University of Western
Nedlands, Australia
shams@csse.uwa.edu.au

Mayez A. Al-Mouhamed
Department of Computer Engineering
KFUPM
Dhahran, Saudi Arabia
mayez@kfupm.edu.sa

Talal Mousa Alkharobi
Department of Computer Engineering
KFUPM
Dhahran, Saudi Arabia
talalkh@kfupm.edu.sa

Seyed M. Buhari
Faculty of Science
Universiti Brunei Darussalam
Brunei Darussalam
mibuhari@yahoo.com

***Abstract*— A telerobotic stereo-vision system is used to extend operator's eye-hand motion co-ordination to a distance and has been used in many applications. One of the most critical problems of such system is communication delay that leads to teleoperation instability. The problem can be minimized by using Augmented Reality (AR) concepts of superimposing virtual objects onto the real video image of the workspace and slave robot so that the operator can see the effects of his hand motion on the manipulated object. This facilitates tele-operation by reducing the number of trial and error in positioning the object. In addition, AR improves task safety and reduces network interactions. In this paper, a detailed model to augment a given stereo-vision system with AR is presented. At first, a six DOF 3D graphical arm is designed and then superimposed onto the video image using camera calibration and registration methods. Motion activation algorithms are developed and interfaces are designed to facilitate task simulations. The system is implemented using Microsoft .NET with Visual C# and DirectX 9 to augment a stereovision system comprising of a PUMA-560 robot operating over a LAN.**

Keywords-augmented reality; telerobotics; stereo-vision; computer graphics.

I. INTRODUCTION

Telerobotics is a multidisciplinary area that aims at extending eye-hand motion co-ordination through a distance using real-time wired or wireless computer networks. In a tele-robotic stereo-vision system, a master arm is used to describe hand motion, slave arm to reproduce the motion, and stereo vision and force feedback to enhance human depth perception and senses. Now-a-days, such system has been used for operating in deep space, underwater, nano and micro scales, carrying out surgery inside the patient body and in many other scaled down or scaled up or hazardous situations.

One of the most critical problems of the conventional telerobotic stereo-vision system is the communication delay [1]. Due to this problem, a tele-operator sometimes has to go for a move-n-wait strategy. Thus, it leads to tele-operation instability. Various techniques such as predictive feedbacks, supervisory control and most recently Augmented Reality (AR) have been proposed to solve the problems [1, 11, 12].

AR is a system that combines real and virtual environment and which is interactive in real time and registered in 3-D Area based stereo. A comprehensive survey on AR is made in [2] exploring number of applications including medical visualization, maintenance & repair, annotation, entertainment and military aircraft navigation & targeting, interior design and many more. In robotics, AR lies between tele-presence (completely real) and Virtual reality (completely synthetic) and between manual teleportation and autonomous robotics [1].

AR allows developing and validating a teleoperation plan in the client machine by superimposing some virtual objects onto the real video image of the slave robot working in the remote space. The plan substitutes frequent low-level interactions between the user and the remote site by sending only the finalized data and thus, reduces real-time network interactions. It also increases the task safety.

Proper registration of real world i.e. robot workspace scene data with the graphics image data is required to apply AR on a telerobotic stereovision system. Registration refers to the proper alignment of the virtual object with the real world. Two kinds of registrations are: static and dynamic [4]. In the static registration user and the objects in environment remain still. It is done at initialization with the help of a human operator. The dynamic registration is done while the viewpoint starts moving to automatically update the registration data.

Accurate camera calibration technique is to be used to align graphic camera co-ordinate to the real video camera co-ordinates. Camera calibration is the establishment of the projection from the 3D world co-ordinates to the 2D image co-ordinates by finding the intrinsic and extrinsic camera parameters [3]. Intrinsic parameters includes optical and electronic properties of a camera, such as focal length, lens distortion coefficients, image center, scaling factors of the pixel array in both directions. While the extrinsic parameters are the pose estimation (rotation and translation) of the camera system relative to a user-defined 3D world coordinate frame [4].

In [1], a telerobotic AR system that implements 3D view in the monitors is discussed. It uses position tracking through mechanical and optical sensors for camera calibration. Gomez et al. [5] developed a virtual environment for teleoperation using OpenGL. Herve et al. [4] proposed a model based camera calibration and registration for AR where camera moves. Abdullah et al. [3] described a pinhole camera model

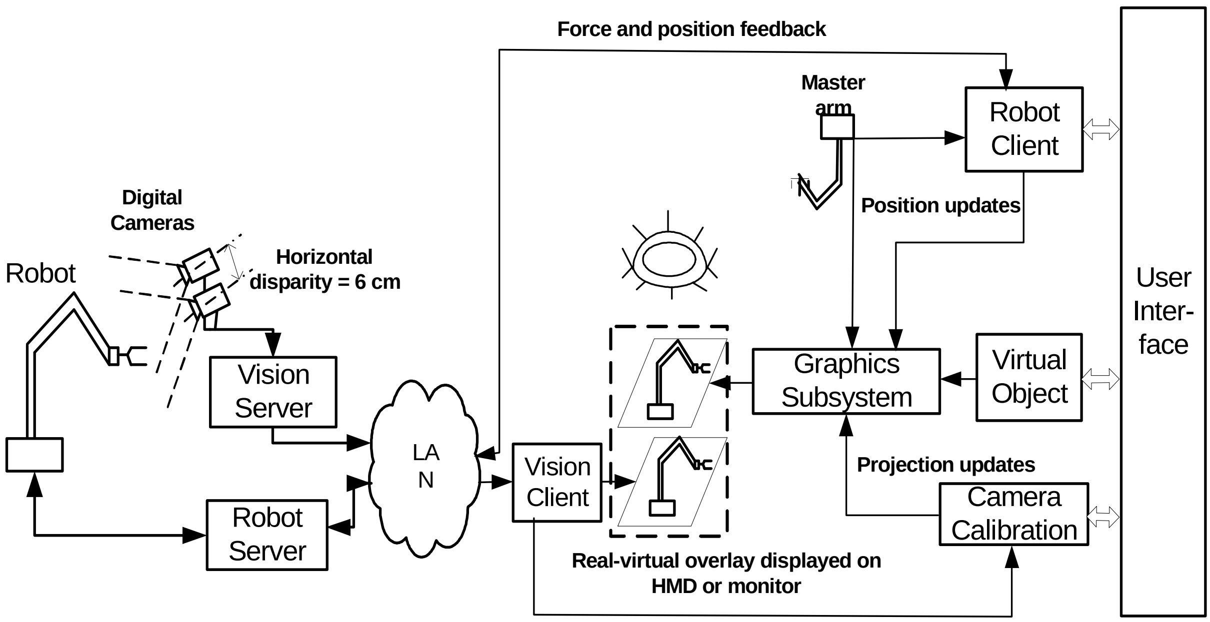

Fig. 1. System design for superimposition of graphic robot onto the real video image of the slave robot.

with weak perspective projection, correction for radial distortion of lens and display on monitors for telerobotic AR applications. Heikkila [6] proposed a geometric camera calibration using circular control points that gives calibration accuracy up to 1/50 th of a pixel size. Marin et al. [7] proposed an Internet-based high level telerobotic framework using Java3D and CORBA. Al-Mouhamed et al. [8] developed multi-threaded distributed telerobotic framework over LAN with video frame transfer rate of 17 fps. They used a camera calibration technique based on fiducial frame of reference for drawing a small red ball in the most current position of the gripper to show the AR effect. Livatino et at. [11] superimposed and aligned graphical objects representing proximity measurements to video information in a real telerobotic system with stereoscopic viewing.

In this paper, we will present a detail strategy for augmenting a telerobotic stereo-vision system with graphical overlays. A mathematical model of robot arm is developed. The arm is then drawn and animated at first in a stand-alone graphic system and then, seamlessly overlaid on the stereo-video received at the client side to facilitate graphical telemanipulation.

The organization of the paper is as follows. In section 2, the design strategy for the proposed model is described. Implementation and performance evaluation are discussed in section 3 and conclusion is drawn in section 4.

II. Design Strategy

A hierarchical design approach is adopted for the augmentation (see [10] for detail). At first, a mathematical model of the robot is derived. Then, choosing the appropriate graphics software the graphic robot arm is designed.

Algorithms are developed for movements and manipulation of the graphics using user interface. Graphics sub-system is then interfaced with the stereo-vision system. Camera calibration and image registration techniques are used to superimpose the graphic model onto real video image. GUI's are then developed for augmented tele-manipulation. Lastly, the full integration is tested. Overall system design is shown in fig. 1.

A. *Developing the Model of the Graphic Arm*

At first, mathematical model is derived for the PUMA-560 which is a robot arm having 6 revolute joints. Each of the links connected with the is attached with a frame of reference to describe its position vector (with respect to a fixed frame and its orientation matrix (M_0^n) as functi the joint vector (θ) and total number of joints, n. The direct geometric is expressed as,

$$\Gamma(\theta) = O_0O_{n,0}(\theta),\ M_0^n(\theta) \qquad (1)$$

Where the orientation matrix is the product of all orientation matrices from first to last joint:

$$M_0^n = M_0^1 . M_1^2 . M_2^3 ... M_{n-1}^n \qquad (2)$$

And the position vector of the arm tip is the sum of all the link Vectors from first to last link:

$$O_0O_{n,0} = O_0O_{n-1,0} + M_0^n . O_{n-1}O_{n,n} \qquad (3)$$

Inverse geometric model of the arm is used to compute the joint vector θ based on the knowledge of the position vector of the arm tip X_{new} and its orientation matrix M_{new}. The new value of the joint vector is used for controlling the slave arm using the new position and orientation of the end effector:

$$\theta = G^{-1}(X_{new}, M_{new}) \qquad (4)$$

A flowchart for movement of graphic arm in Cartesian space is shown in fig. 2. The details of the model can be found in [9]. This model provides the skeleton of the graphic arm.

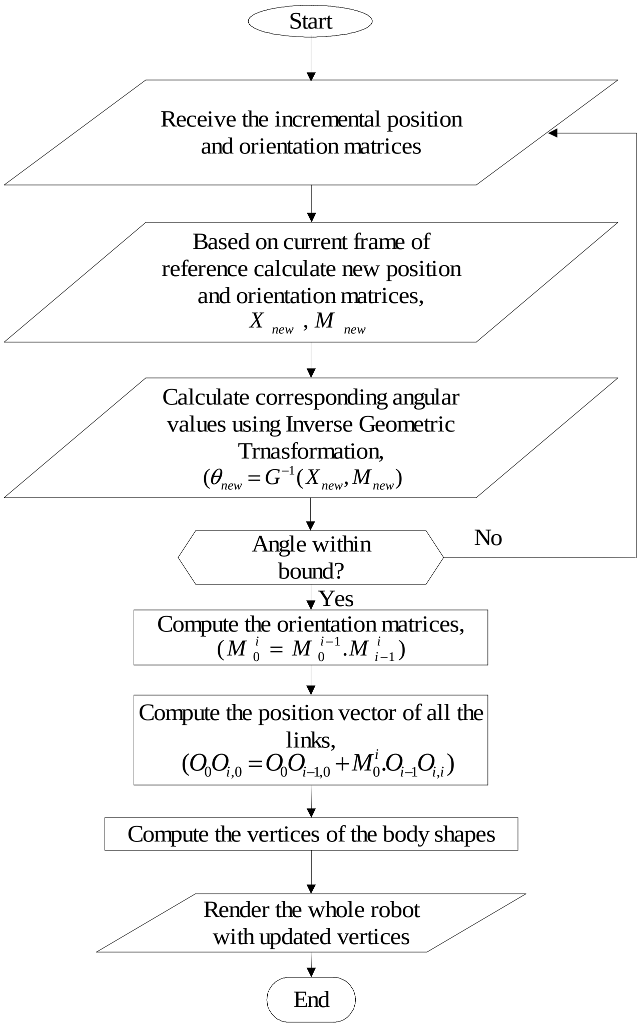

Fig. 2. Flowchart for movement of graphic arm in Cartesian space.

Cylindrical body shapes are then attached to the links [10] to have a solid view.

B. Design of Data Structure

Each link of the arm is modeled as a different object with properties representing the link; see [10] for detail. Thus, geometric data of each link is kept apart from the other link. The configuration data such as number of links and number of segments in cylinder are also kept in separate data file.

C. Acquisition and 3D Visualization of Real Stereo-image

A multi-threaded distributed telerobotic framework developed in [8, 9] is used for acquiring video image of the worksite. In the server PC, Vision Server is used to capture image by two digital cameras and then relayed over network. On the client side, after detecting and making connection with the server, pictures are received. Page flipping technique is used for 3D visualization and HMD is used as display device

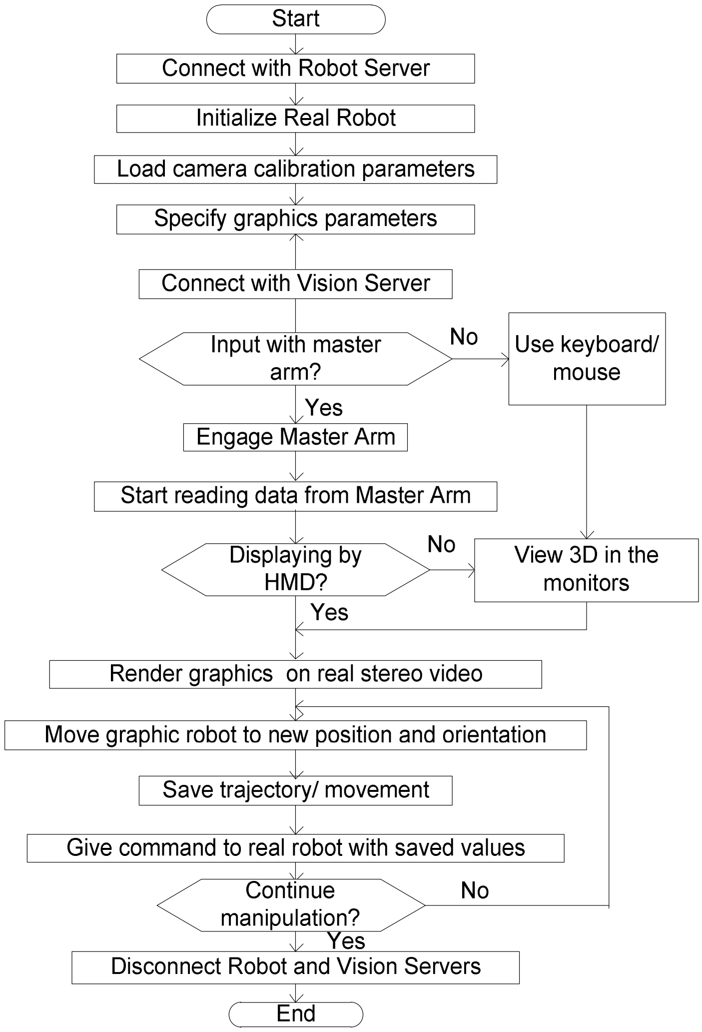

Fig. 3. Overall client system flowchart showing steps for graphical tele-manipulation.

that gives the effect of 3D depth perception.

D. Superimposition of Graphic Arm on to the Real Image

Heikkilas camera calibration [6] method is used to find out the intrinsic and extrinsic parameters of the cameras used. Graphical camera used to display the graphical arm is also configured accordingly for proper superimposition. During each page flipping operation, real image acquired from the network video stream is copied in a surface, say *frontSurf* while the drawing of the current image on graphics screen is in progress in the another surface, say *augSurf*. *AugSurf* is copied to the primary surface which is used to display the contents. Left and right camera image are displayed on two different view-ports on the monitors or HMD for 3D views.

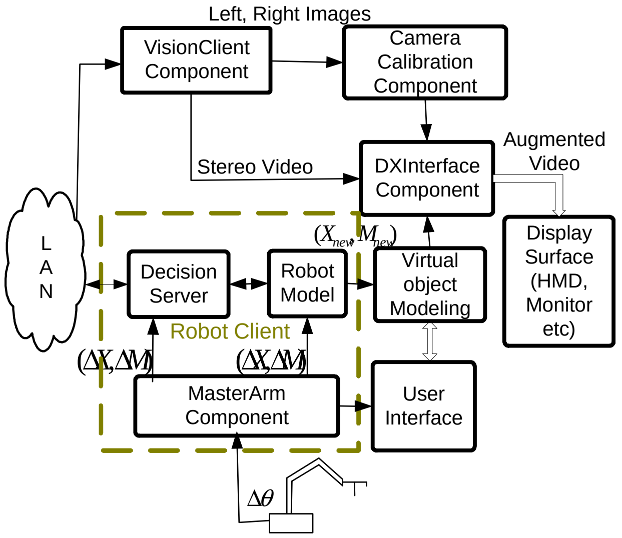

Fig. 4. The basic building blocks of the client system.

E. *Design of GUI and Graphic Tele-manipulation*

The server side GUI, developed by Al-Mouhamed et al. [8], provides a user the facility of connection and initialization of the real robot. Any client (master station) may connect to the server (slave station) through the Internet. The client will establish connection and starts the Augmented Reality (AR) functions. Here, an HMD is used to display the stereo images which are streamed from the server robot. In addition, a graphic robot is superimposed on the real stereo image. The user has control of a number of tele-operation functions like space scaling and shift master arm to flexible position and map to slave arm in addition to control the start and end of a connectivity session.

The client side GUI is designed to take user's input from master arm, keyboard or mouse to perform these functionalities.

As shown in fig. 3, for graphical tele-manipulation, at first the Real Robot Server is connected and robot is initialized. The camera calibration parameters are then loaded to the program. Graphics parameters are also specified. Then, the Vision Server is connected. A user then can interact with the graphic arm and manipulate real robot by wearing the HMD.

III. Implementation and Performance Evaluation

The proposed system is implemented on a Microsoft .NET framework using MS Visual *C#*.NET, MS Visual C++ and MS DirectX 9. Direct3D is used as a 3D graphics API. Both the server and client systems of the given telerobotic stereo-vision system are running on PCs having 2-GHz Intel P4 processors with 1GB DRAM and 512 KB cache memory and connected to a campus network by using a 100 Mbps NIC card.

The server PC is interfaced to two Sony Handycam digital cameras using a 400 mbps FireWire PCI (IEEE-1394) card. The performance of the system is evaluated against the rendering time and accuracy. Data is taken by averaging over 1000 samples.

A. *Basic Components of the Client System*

The basic building blocks implemented at the client side to augment given stereo-vision system is shown in fig. 4. DirectX Interface module is the core module of the client application for displaying the augmented view. It takes input from virtual object modeling module, camera calibration module video client and output the augmented video. Thus, it synchronies real and virtual data, make projection on video surface and performs the page flipping for the use with HMD. The master arm component computes an increment in the Cartesian position and orientation of the master arm, which is due to operator motion. The above increments are forwarded to the slave arm through the network. The motion parameter transfer is subject to transmission and protocol delays. For this, the proposed approach consists of locally applying the motion increment on the overlay image of manipulated object. This allows the operator to see the effects of his hand motion on the manipulated object within the latest stereo image.

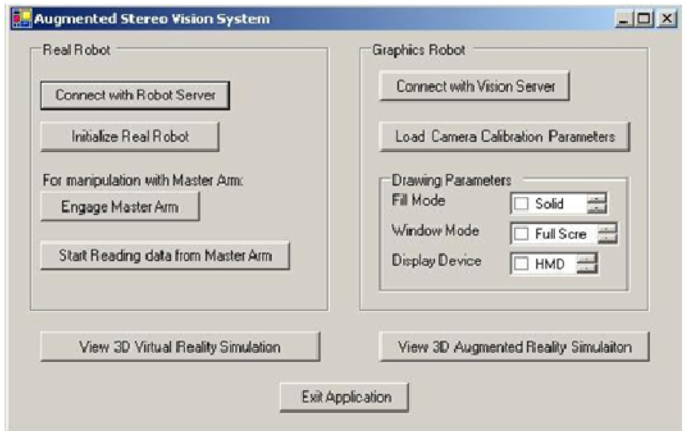

Fig. 5. Main GUI at the client side.

B. *User Interfaces and Augmented Manipulation*

Client side GUI's helps a user to tele-manipulate. Fig. 5 shows the main user interface to be used for setting some configuration parameters and making connection with real robot and the graphic robot system. It also let the user to open the simulation interface for manipulation in stereo-view. The simulation window interface facilitates user in changing the graphic robot position and orientation using master arm, mouse or keyboard and to see the effect on the real stereo-video of the workspace. Once finalized with the trajectory, commands can be issued to perform the specific task with the slave robot.

C. *Solid and Wire-frame Views of the Graphic Arm*

Both wire-frame and solid models of the graphic arm (with some simplified assumption) are drawn (see fig. 5). Rendering times for arm with 50 segments in each cylinder representing the body for these two views are recorded as 64.09 ms

&64.283 ms respectively. When the arm is overlaid on real video the times found were 105.37ms and 109.62ms.

D. Speed and Accuracy of Graphics Rendition

Speed of graphics rendering is measured for different environments [8]. Refresh rate of 253.59 and 11.384 fps are recorded for with or without superimposition respectively. Significant accuracy was also achieved in matching of graphics with real video image.

IV. Conclusion

AN AR-BASED TELEROBOTIC SYSTEM IS PROPOSED IN THIS PAPER TO AUGMENT A CLIENT-SERVER TELEROBOTICS SYSTEM OPERATING OVER THE INTERNET.

Transmission and protocol delay causes instability in the tel-operations system which makes complicate the operator task. An augmented reality scenario has been proposed for verlaying the user motion corrections on a graphic representation of the manipulated object and transmits these corrections to the slave arm. This overcomes the delay problem because the operator can see the effects of his motion on graphical representation of the object. It is shown that using hardware accelerated graphics rendition through Direct3D provides excellent refresh rate of the output screen. A flexible and generalized data structure suitable for telerobotic visualization is used. A computer vision-based calibrations method is used that provides accuracy up to 1/50 of pixel size. A user-friendly graphical user interface is developed for simple manipulation in the telerobotic AR system. It can be used as a base framework for further research on virtual and augmented telerobotic manipulations.

ACKNOWLEDGMENT

The authors would like to acknowledge and give thanks to King Fahd University of Petroleum and Minerals for its support of providing required computational and hardware facilities towards accomplishment of this research work.

REFERENCES

[1] A. Rastogi, P. Milgram and D. Drascic. Telerobotic Control with Stereoscopic Augmented Reality, *SPIE*, Feb. 1996, Vol. 2653: Stereoscopic Displays and Virtual Reality Systems III, pp. 135-146.

[2] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, A survey of augmented reality, Computer Graphics and Applications, *IEEE*, 2001, Vol. 21(6), pp. 34–47.

[3] J. Abdullah and K. Martinez, Camera self-calibration for the artoolkit, *The First IEEE International Workshop on Augmented Reality Toolkit*, Sept. 2002, pp. 5.

[4] J.-Y. Herve, C. Duchesne and V. Pradines, Dynamic registration for augmented reality in telerobotics applications, *IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 2, October 2000, pp 1348 – 1353.

[5] S. R. Gomez, J. Carro, E. Moreda, J.J. Escribano, and C. Cerrada, A design srategy for low cost experimental telerobotics platforms. *Proc. IEEE International Workshop on Robot and Human Interaction*, Pisa, Italy, Sept. 1999, pp. 273 – 278.

[6] M. Yorozu, K. Oka Hirano, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface." *IEEE Transl. J. Magn. Japan*, Vol. 2, , August 1987, pp.740-741.

[7] R. Marin, P.J. Sanz, and J.S. Sanchez, A very high level interface to teleoperate a robot via web including augmented reality. *Proc. IEEE International Conference on Robotics and Automation*, 2002 ICRA'02, May 2002, Vol. 3, pp. 2725 – 2730

[8] M. Al-Mouhamed, O. Toker, and A. Iqbal, A Multi-Threaded Distributed Telerobotic Framework, *IEEE T. on Mechatronics*, Vol 11, No 5, Oct., 2006, pp. 558-566.

[9] M. Al-Mouhamed, O. Toker, and N. Merah, Design of an intelligent telerobotic system, *Technical Report AT-20-80*, King Abdulaziz City for Science and Technology, Kingdom of Saudi Arabia, 2004.

[10] S.M.S. Islam, M. Al-Mouhamed, S. M. Buhari, and T. Al-Kharobi, A Hierarchical Design Scheme for Application of Augmented Reality in a Telerobotic Stereo-Vision System, Proceedings of the Saudi 18th National Computer Conference, Riyadh, Saudi Arabia, March 26-29, 2006, pp 399-412.

[11] S. Livatino, G. Muscato, D. De Tommaso, and M. Macaluso, Augmented reality stereoscopic visualization for intuitive robot teleguide. *Proc. IEEE International Symposium on Industrial Electronics*, 2010, pp 2828 – 2833.

[12] Sangwoo Lee, Hoiju Jung, Hyunjong Song and Shinsuk Park, Development of immersive Development of immersive augmented reality interface system for construction robotic system augmented reality interface system for construction robotic system, *Proc. International Conference on Control Automation and Systems (ICCAS), 2010*, pp 2309 – 2314.