AUGMENTED REALITY INTERFACE FOR FREE TELEOPERATION

Vincent LE LIGEOUR Samir OTMANE Malik MALLEM

Laboratoire Systèmes Complexes
Université d'Evry Val d'Essonne
40, rue du Pelvoux, 91000 Evry, France

Abstract: Augmented Reality (AR) can provide to a Human Operator (HO) a real help to achieve complex tasks such as remote control of robots and cooperative teleassistance. Using appropriate augmentations, the HO can interact faster, safer and easier with the remote real world. This paper presents a new Man Machine Interface (MMI) allowing high-level telewok, using augmented reality technologies. An example of telework is presented in this paper as a teleoperation of the remote 4 Degree of Freedom (DoF) robot. An important feature of this interface is based on the use of two new modes devoted for HO perception and interaction. In a new perception mode, a wide screen with stereo glasses are used to allow stereoscopic perception of the remote real environment. The new interaction mode is given thanks to ART* optical tracking system that allows HO to interact freely with the remote 4 DoF robot. An other important work presented in this paper is a distributed software architecture developed with three client/server modules (robot control, stereoscopic video feedback and tracking) and using appropriate methods (communication protocols, image compressions etc.). This new interface is tested and compared with ARITI interface (acronym of Augmented Reality Interface for Teleoperation via Internet) and some preliminary results are presented and discussed.

* ART : Advanced Realtime Tracking

Keywords: Telerobotics, teleoperation, tracking systems, distributed system, stereo vision, virtual reality

1. INTRODUCTION

Early Computer Assisted Teleoperation (CAT) architectures, (Espiau, 1986)(Hasegawa, 1991), provided the operator with many kind of assistances. Yet, the tendency of these assistances was mainly focused on the operator interface improvement. Indeed diverse assistance strategies and informations concerning tasks are proposed and the robot commands where a function of operator actions and the adopted strategy (such as freezing some robot degrees of freedom) i.e. the robot is still a slave of the master control. Nowadays the CAT architectures benefit from a considerable trump: virtual reality which sets high standards in the human-machine interface (Burdea and Coiffet, 1994). VR was also investigated to solve the time delay problem namely through predictive displays (augmented reality systems) and different teleprogramming architectures. There is a great amount of ways and purposes in using Augmented Reality (AR) for teleoperation such as, for instance, sensory feedback improvement, etc. The basic control of task execution is through superimposing, on the online video feedback, the corresponding robot and its surroundings virtual model

(Mallem *et al.*, 1992)(Milgram *et al.*, 1993)(Otmane *et al.*, 2000).

During the last years, Augmented Reality (AR) has been proved to be useful to improve performances of a human operator while this one is performing complex tasks. The LSC (Laboratoire Systèmes complexes) has worked on helps to provide to the operator in the field of teleoperation of a remote robot which led to the ARITI system (*Nasa space telerobotics program web site*, n.d.)(Otmane *et al.*, 2000)(Otmane *et al.*, 2002). But during this time, technology has evolved and provides us new devices and techniques allowing us to develop new ways to perceive and interact with remote environments and to allow teleassistance and collaboration of remote operators.

In the second section of this paper, the ARITI interface is presented with different kinds of the HO assistance. A new LSC EVR@ [1] platform is described in the third section with different devices concerning the new perception and interaction modes. The fourth section is dedicated to the software architecture developed and methods implemented to achieve the desired perception and interaction approaches. The preliminary experiments and results are presented in the last section.

2. THE ARITI SYSTEM

ARITI (Otmane *et al.*, 2000) is a Client/Server system, which allows visualizing and controlling a 4 DoF robot by using any remote computer. It is aimed to enhance HO capabilities of achieving complex telerobotic operations. One has noticed that, in a teleoperation activity, the HO reacts faster when he is given three kinds of help:

- Helps devoted to the environment perception
- Helps devoted to the robot control.
- Helps devoted to check the execution of the actions.

The ARITI Java web applet [Fig. 1] is aimed to provide the maximum help to the HO without giving useless help decreasing user skills.

Three fourths of the display is devoted to the perception of the telerobotic environment. This perception is based on a combination of Augmented Reality (AR) and a 3D interactive graphics.

In order to help the HO to control the robot and achieve a complex telemanipulation task, an interactive assistance is given to him. It is the interactive intervention of Virtual Fixtures (VF) in the operation area [Fig. 2]. These VF (Otmane *et al.*, April 16-20, 00c)appear and disappear as the robot's peg comes close or gets away from the objects to be manipulated. The type of virtual fixture appears according to the kind of tasks that had already been achieved.

[1] a new LSC technological platform for telework in Virtual and Augmented Reality Environments

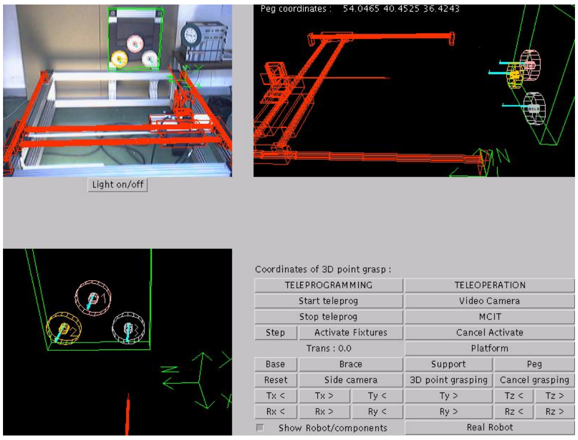

Fig. 1. ARITI Man Machine Interface

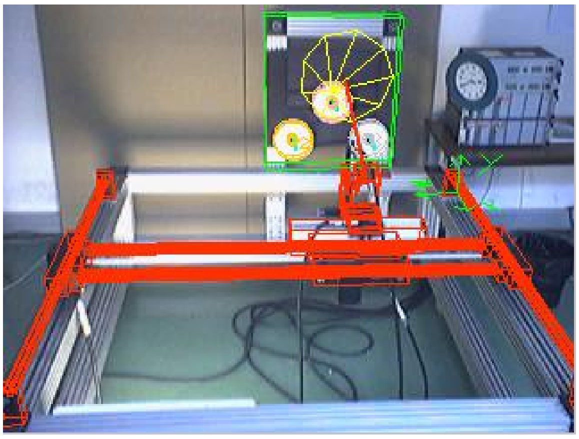

Fig. 2. Virtual fixture enhanced on the video feedback

3. THE NEW PLATFORM : EVR@

3.1 The platform

The Evr@ platform is the new AR system of the LSC. The aim is to provide new ways of perception and interaction through a set of new technologies.

3.2 The tracking system

To give a real time interaction between the virtual robot and the remote robot, an ART tracking system is used. This system allows real time detection of position and orientation of the interaction device called Flystick [Fig. 3].

Cameras worked in the near infrared light spectrum. An infrared light flash illuminates the measurement volume periodically. The system uses retro-reflecting markers for marking the object to be tracked. The procedure giving tracking information consists in 4 steps:

- Measurement volume is illuminated by an IR flash and camera images are taken.
- Two-step calculation unit within the camera recognizes the markers and calculates the marker positions in image coordinates (2D) with high accuracy.

Fig. 3. A.R.T. flystick

- 2D marker coordinates of the single tracking cameras are handed over to the central server via Ethernet.
- 3 DoF marker position and 6 DoF target/body position and orientation are computed by the server

The Flystick also provides 8 buttons and the operator can now interact with the system through this new way of interaction and gives positions, orientations and orders to the system.

3.3 *The Display system*

To really emerge the operator into the system, a large screen [Fig. 4] (3,20x2,40 m) provides stereoscopic video and augmentation of the reality. The DLP technology used in the projector can project a clear images in normal illuminating environment, so the operator can work in a non-dark chamber.

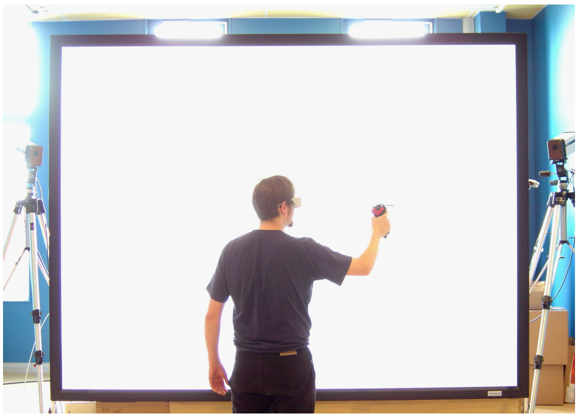

Fig. 4. Screen of the EVR@ platform

3.4 *The frame grabbing part*

And finally to make the system flexible and simple, we can acquire video with different kind of devices :

- Sony compact cameras providing light and efficient system, but needing a frame grabber card.
- Logitech USB webcams providing colored images and a non material dependent system, but providing blurriest images.

4. MIXING NEW WAYS OF PERCEPTION AND INTERACTION

The first challenge is to find a way to combine the different servers :

- The tracking server, emitting frame containing flystick positions, orientations and buttons states.
- The ARITI command server, providing robot position and controlling the 4 DoF
- The cameras/webcams server providing stereo image flow
- The display server, using the developed client integrating the different network flows and displaying the results on the wide screen.

The integration is done via a TCP and UDP IP network layer [Fig. 5]. Each server is a multi client server, so that data can be sent on different clients easily and even be used by different systems thanks to an high level protocol.

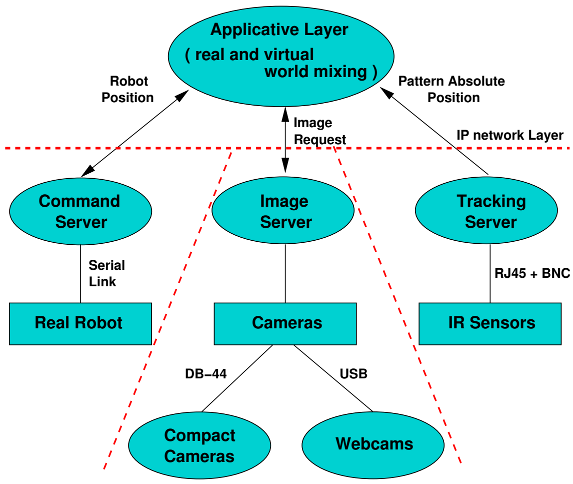

Fig. 5. Global system architecture

4.1 *The tracking client/server*

The tracking server sends UDP frames containing 3D (postion), 6D (3D+orientation) and 6D Flystick (6D+button state) data. Each frame also contains a timestamp a unique ID, a confidence factor and an object identification. Decoding algorithm was designed to fill objects and freely manipulate these data.

Acquiring and decoding process is done in a different process to ensure optimum performance. The protocol enables to start and stop tracking with a single UDP frame.

4.2 *The image client/server*

The image client-server protocol is a complete set of orders that can transmit stereo images in different formats:

- some based on compression : The server can provide 3 different compression methods. A none compression (sending raw image data) very fast and with no encoding delay) very useful on fast Local Area Network (LAN). The second one is the ZIP compression, very powerful and non destructive compression, but on an other hand it is slow to encode (needing all raw image reading before compressing it) and larger than the third encoding way, the JPEG format. The JPEG format encode data by lines, it's very fast and produce smaller buffers than ZIP compression and can be adjusted by the compression level. But this compression is a destructive one that cannot fit the next point.
- the other ones based on evolution : two different evolution methods were tested. Due to the JPEG destructive ability, this simple method cannot be combined. So only the ZIP method is valuable. The first is the XOR method, fast and simple, but very sensitive to small local image variations, and the second is the difference. This second method has a problem : it increase the size of the buffer (encoding **signed** difference) but produces a more compressed buffer (smaller values).

To reduce data flows (10 Mbits network between client and server) the JPEG compression method were used.

Used images are stereoscopic RGB pictures with a resolution of 640x480 pixels (the stereoscopic raw size is then 1843,2 Ko).

For a zip image, the average time of compression and send is 131.05 ms (standard deviance 1.33) and the size average 519.72 ko (standard deviance 1.54).

The JPEG compression is the best in terms of compression; this is due to the line per line compression in opposite to the one block ZIP compression. There is now one problem to the JPEG compression, this is a destructive compression.

In fact two different servers was design to grab frame both on compact Sony cameras and color webcams.

In the compact Sony's cameras, pictures are grabbed simultaneously via the frame-grabber giving two interlaced gray-scaled pictures. In fact those buffers can be compressed as an RGB buffer. The size is smaller (quite the same picture) but the quality is degraded (echo images on other channels).

In the colored webcams, images are grabbed via a common USB output webcams and acquiring two 640x480 YUV buffers. These buffers are converted into RGB buffers (bigger one) on the server to ensure client-server protocol compatibility and JPEG compression.

4.3 *The command client/server*

The command client - server is based on a TCP protocol. At the first connection the server sends the limit values beyond the robot cannot go, then the client can send different orders:

- Setting positions in different ways (Setting all coordinates or just updating one)
- Resetting the robot position
- Getting the current of the 4 DoF robot (useful for updating a supervisor client)

To ensure correct commands order, the server uses FIFO (First In First Out) commands and mutexed read and write on the serial port connected on the robot.

The following order are part of the protocol [Tab. 1] :

Order	**Argument**	**Answer**	**Effect**
SETPOS	1 Axis (X, Y, Z, W) and 1 position	None	Change a position and send it to the robot
UPDATE	None	None	Update server stored positions from robot
GETALLPOS	None	4 positions (X, Y, Z, W)	Get server stored positions
GETREALPOS	None	4 positions (X, Y, Z, W)	Update server positions and send them
RESET	None	None	Place the robot in (0, 0, 0, 0) position
GETALLLIMITS	None	8 limit positions (1 for each axis)	Limits are stored in the server and disable position in dangerous position

Table 1. Command protocol

If an order failed then "KCA" is sent back otherwise "ACK" is sent. All the limits stored in the server are checked twice. First the clients disable this position and avoid user to send them. But if a problem occurs (network problem, other client, ...) the server check the value just before sending them to the serial robot interface. Due to the robot interface (serial) no "tracking" could be implied. A tracking method could enable real time robot tracking with every movements of the operator. The protocol already make it available but it cannot be used due to the serial interface.

4.4 Man Machine Interface

The important part of this new way of controlling the 4 DoF is to find the new MMI. In fact we cannot control and visualize a robot in a small web applet with the keyboard and in a large screen with a totally free flystick. The flystick is also a 6 DoF interface and the robot a 4 DoF, we had to choose the degrees used to control the robot.

The first challenge is to find a way to manage the flystick as a new way of interaction. 3 different methods were tested :

- Totally absolute control, giving the same position and orientation of the flystick. It can seem the perfect way, but in fact it is the worse. It tires the operator and this interaction is not simple. The two-tested solutions were: controlling the 2 translations with the flystick translation and the angles with the 2 angles, and controlling the translations the same way and on of the translation with the TOP-BOTTOM translation.
- Totally relative control, it is most useful, but the angle sum can produce strange effects (for example pointing forward and the resulted effect is pointing left or right).
- Both absolute and relative controlling. It is the method used. The two translations are controlled relatively to enable a move less position (on a chair for example) and the two rotations are controlled absolutely. The result is that the robot is controlled like a sword or a foil.

And to interact with the client both 8 buttons were used to control the robot:

- the trigger: used to take control and the robot and sending orders on the release
- a reset button (resetting robot position)
- two buttons to increase and decrease translation scale
- a button to initialize peg orientation
- a button to validate the target release of ARITI system
- a button to switch full-screen/windowed style
- a button to end the application

The virtual fixtures to grab object were also integrated in the new MMI. Volumetric object were used to augment volume perception. The AR was used to give additional informations. So when the robot is far from the target the stereoscopic video is full-screened and the stereoscopic virtual model is paste in the right-hand corner of the screen as shows in [Fig. 6].

When the robot is near the object and the virtual fixtures self activated the stereoscopic virtual model of the robot is now full-screened [Fig. 7] and the video

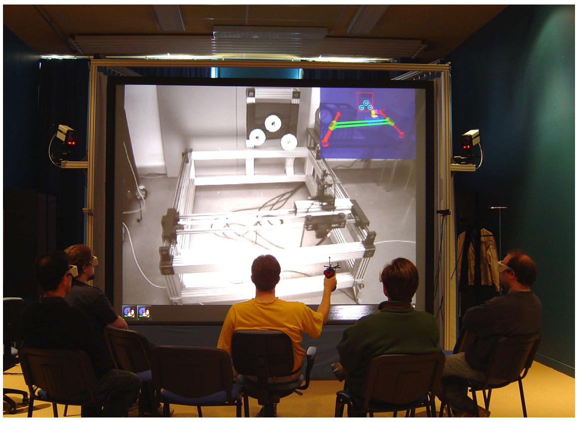

Fig. 6. Screen of the EVR@ platform - Fullscreen video

stereo is reduced, because not providing any useful information.

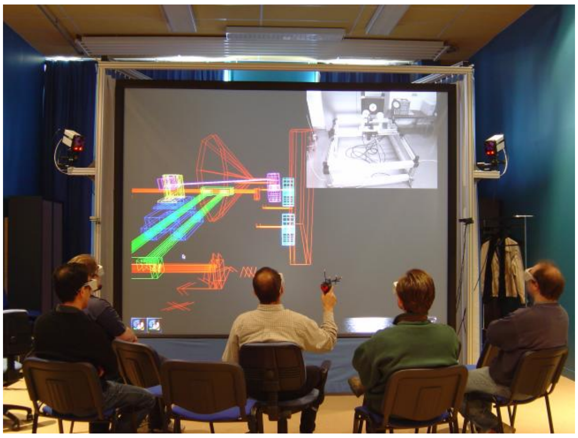

Fig. 7. Screen of the EVR@ platform - Fullscreen virtual model

Finally a third perception method was used : superimpose AR and VR [Fig. 8]. The registration was done using an interactive least square minimisation. To help users during the dangerous task of picking or replacing an object, a virtual camera was placed near the target. When the HO activate the fixture, it also activate the virtual camera and a visual help is placed in the left top hand corner.

5. EXPERIMENTS AND RESULTS

To evaluate this new way of perception and interaction, some tests were done with different telemanipulation tasks and with different user skills. People that have already manipulate ARITI through the web and having already used the flystick, and others totally beginners. The manipulation was, with full fixtures activates, to pick an object, to remove it safely from its support and to finally replace it safely on a support. The time of each task was evaluate. The tests are achieved for advanced users. In a series of 10 users and each of them making six tests on each interface (ARITI web interface and the new proposed MMI

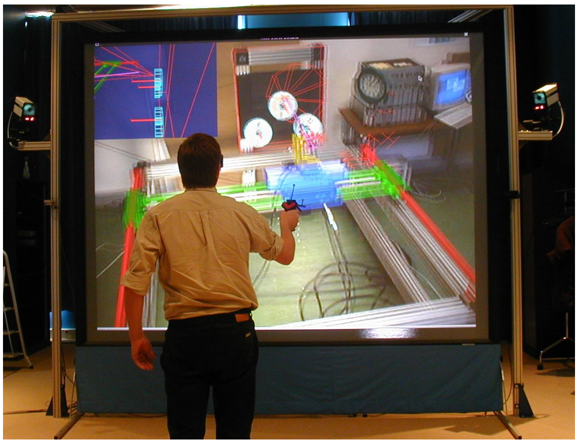

Fig. 8. Screen of the EVR@ platform - Superimpose with virtual zoom

interface). We can note that the average time with the new MMI interface (EVR@) is a third better (in time) than the web ARITI interface. The average time with the flystick is 34.64s (standard deviance of 3,4s) and with the web interface 47,55s (standard deviance of 3,81s). The first tests with beginners seems to demonstrate that the Evr@ interface is more intuitive than the web one; In fact the learning is better with the flystick than the mouse and keyboard. These first tests already show that a beginner user is faster with the flystick than an advanced user with the web interface.

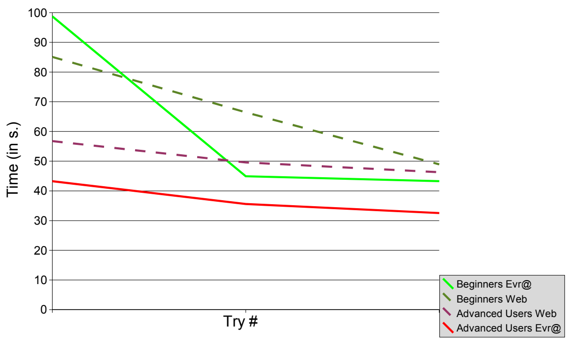

Fig. 9. Comparison of time achievement task between the ARITI and EVR@ interfaces (Web results are dashed)

In the figure [Fig. 9] we can see that even if the new interface is harder to use in the first try, the operator can improve his skill in the second attempt and can do faster with the Evr@ interface (with the Flystick) than with the ARITI web interface.

6. CONCLUSION AND PERSPECTIVES

In this paper the new MMI for high level teleoperation is presented with a developed architecture tacking a count of the two new approaches of perception and interaction. The next steps in AR software development will be to reduce the logical glue (used to glue the different network clients in the main application). Using small pattern and fast inter-pattern communication method could achieve that. Then networked communication could be used totally freely by user and enhanceability to build user-sized distributed application. An other step will be to add a scripting method to configure Input/Output and robot description. A complete XML robot was already designed but not yet implemented, that could lead to a complete robot architecture free program, able to manipulate 4 DoF and 6DoF robot.

REFERENCES

Burdea, G. and P. Coiffet (1994). *Virtual Reality technology*. John Wiley et Sons, Inc Eds, New York.

Espiau, B. (1986). Téléopération avancée : Le projet de téléopération avancée. Technical report. Saclay, France.

Hasegawa, T (1991). Model-based telerobot system manipulation skill. pp. 499–506.

Mallem, M., F. Chavand and E. Colle (1992). Computer-assited visual perception in teleoperated robotics. In: *Robotica* (J. Rose, Ed.). Vol. 10. pp. 93–103. Combgidge university, England.

Milgram, P., S. Zhai and D. Drasic (1993). Application of augmented reality for human-robot communication. In: *IROS: Int. Conf. on Inteligent Robot and Systems*. Yokohama, Japan. pp. 1467–1476.

Nasa space telerobotics program web site (n.d.). http://ranier.oact.hq.nasa.gov/telerobotics page/.

Otmane, S., M. Mallem, A. Kheddar and F. Chavand. (2000). Ariti: an augmented reality interface for teleopeation on the internet. In: *Advanced Simulation Technologies Conference (ASTC2000)*. Wyndham City Center Hotel, Washington, D.C., USA. pp. 254–261.

Otmane, S., M. Mallem, A. Kheddar and F. Chavand (April 16-20, 00c). Active virtual guide as an apparatus for augmented reality based telemanipulation system on the internet. In: *IEEE Computer Society - 33rd Annual Simulation Symposium ANSS2000*. Wyndham City Center Hotel, Washington, D.C., USA. pp. 185–191.

Otmane, S., N.Khezami and M. Mallem (2002). Distributed predictive display for collaborative teleoperation on the internet. In: *4th Virtual Reality International Conference (VRIC2002)*. Laval, France. pp. 111–116.