3D Visualization Technologies for Teleguided Robots

Salvatore Livatino
Medialogy Studies
Aalborg University
Copenhagen, Denmark
sal@media.aau.dk

Filippo Privitera
Scuola Superiore di Catania
Catania, Italy
fiprivitera@ssc.unict.it

ABSTRACT

The use of 3D stereoscopic visualization may provide a user with higher comprehension of remote environments in teleoperation when compared to 2D viewing. Works in the literature have demonstrated how stereo vision contributes to improve perception of some depth cues often for abstract tasks, while little can be found about the advantages of stereoscopic visualization in mobile robot teleguide applications. This work investigates stereoscopic robot teleguide under different conditions, including typical navigation scenarios and the use of synthetic and real images. This work also investigates how user performance may vary when employing different display technologies. Results from a set of test trials ran on five virtual reality systems emphasized few aspects which represent a base for further investigation as well as a guide when designing specific systems for telepresence.

Categories and Subject Descriptors: I.3.7 Computer Graphics: Three Dimensional Graphics and Realism. Virtual Reality

General Terms: Human Factors, Verification, Experimentation, Performance.

Keywords: Virtual Reality, Stereo Vision, Telerobotics, Teleoperation, 3D Visualization.

1. INTRODUCTION

The commonly used 2D display systems suffer of many limitations in robot teleoperation. Among which: misjudgment of self-motion and spatial localization, limited comprehension of remote ambient layout and object size and shape, etc. The above leading to unwanted collisions during navigation, as well as long training periods for an operator. An advantageous alternative to traditional 2D (monoscopic) visualization systems is represented by the use of a stereoscopic viewing. In the literature we can find works demonstrating that stereoscopic visualization may provide a user with a higher sense of presence in remote environments because of higher depth perception, leading to higher comprehension of distance, as well as aspects related to it, e.g. ambient layout, obstacles perception, manoeuvre accuracy, etc.

The above conclusions can in principle be extended to teleguided robot navigation. However, it is hard to find works in the recent literature addressing stereoscopic mobile robot teleguide, which motivated the authors to focus on this very important application field. In addition, it is not straightforward how stereo viewing would be an advantage for indoor workspaces where the layout, typically man-made, would be simple and emphasizing monocular depth cues such as perspective, texture gradient, etc.

When analyzing the benefits of stereoscopy, researchers often focus on comparing different depth cues, learning behaviors, etc., but they always run their experimentation trials using one or two specific visualization technologies. A comparison among different virtual reality (VR) facilities is uncommon in the literature, despite some works can be found comparing two different systems, e.g. [3], [9]. Nevertheless, depth perception and task performance may greatly vary for different display technologies, providing a user with different sense of presence and interaction capabilities. In addition, display technologies also differ in cost, portability and accessibility. Different display technologies best fit different application situations. For example, a "light" system, portable and cost-effective, would be required in case of low-range transmission possibility, whereas a larger setup, providing higher immersion, would be more suitable for training purposes.

2. 3D STEREO VISUALIZATION

Several systems have been developed for Teleoperation and VR with different display and interaction possibilities, (e.g. [13], [14]). Large displays for immersive presentations, e.g. Powerwalls, or systems for individual use but allowing for high interaction, e.g. the CAVE, [2], or Head Mounted Display (HMD). Figure 1 shows examples. Different technologies have been developed which confirm the fundamental role of stereoscopic visualization for most VR systems. The basic idea supporting stereoscopic visualization is that this is closer to the way we naturally see the world, which tells us about its great potential in teleoperation. Main approaches to 3D stereo visualization may be classified as:

- *Passive Stereo*. Multiplex images in space and they can be sub-divided in: *Anaglyph* (separation based on color filters); *Polarized* (separation based on polarized filters); *Separated Displays* (separation based on different displays very close to user eye as in HMDs).
- *Active Stereo*. Multiplex images in time typically based on *Shutter Glasses*, (LCD shutter panels in synchronization with the visualization display).
- *Autostereoscopic Stereo*. Separates images based on special reflecting sheets laying on the display, or other methods. Do not require goggles.

Figure 1: VR facilities at Aalborg University VR Media Lab and Medialogy Copenhagen. Top-row from left: 160deg. Panorama; 6-sided CAVE; 1-sided CAVE. Bottom-row from left: Small Powerwall with projectors and filters; HMD; 3D Laptop, 3D Desktop.

Different stereoscopic approaches can be used coupled to different display systems. The latter being responsible for the degree of immersion, interactivity, isolation from the surroundings, etc. Among main components: *Display Size*, from tiny HMD monitors to large 360deg panoramic screens; *Display Structure*, e.g. flat, curved, table-like, cubic shaped, head mounted; *Projection Modality*, LCD/CRT monitors, front/back projected screens; *Image Quality*, e.g. resolution, brightness, contrast, color range, refresh rate.

The literature works investigating the benefits of stereoscopy can be classified as either application specific, or abstract tasks with general performance criteria, [3]. In literature test trials often deal with assessing the role of most dominant depth cues, e.g. interposition, binocular disparity, movement parallax, [11], and their consequence to user adaptation to new context (user learning capabilities). The parameters through which assess stereoscopy benefits typically are: item difficulty and user experience, accuracy and performance speed, [11], [4]. Test variables altered during experiments include: changes in monocular cues, texture type, relative distance, etc., other than stereoscopic versus monoscopic visualization. Everybody seems to agree that stereoscopic visualization presents the necessary information in a more natural way, which facilitates all human-machine interactions [4]. In particular, stereoscopy improves: comprehension and appreciation of presented visual input, perception of structure in visually complex scenes, spatial localization, motion judgement, concentration on different depth planes, perception of surface materials. The main drawback, which have yet prevented large application, is that users are called to make some sacrifices, [12]. A stereo view may be hard to "get right" at first attempt, hardware may cause crosstalk, misalignment, image distortion, and all this may cause eye strain, double images perception, depth distortion.

Most of the benefits of stereoscopy may affect robot teleguide. Among the conclusions gathered from the literature: "most telemanipulation tasks require operators to have a good sense of the relative locations of objects in remote world", [4]; "stereopsis gives better impression of telepresence and of 3D layout", [6]; "binocular disparity and movement parallax are important contributors to depth perception", [11]; "a robot in a dangerous environment can be controlled more carefully and quickly when the controller has a stereoscopic view", [1].

3. ROBOT TELEGUIDE AND 3D VISUALIZATION TECHNOLOGIES

It is proposed to investigate benefits of stereoscopic viewing based on the analysis of few factors typically described as predominant, (e.g. [9]), i.e. depth relationships and motion perception. Two categories of user studies are proposed:

- *Aptitude Tests*. To assess user's ability in estimating egocentric distance and self-motion when using stereoscopic visualization under static conditions or passive (computer controlled) motion. Proposed trials concern "Egocentric Distance", (the user stands in front of a corridor while he/she is asked to estimate the egocentric distance to the far-end wall-plane); and "Self-Motion", (the user is driven along a corridor while is asked to estimate robot speed).

- *Interactive Tests*. To assess user's ability in estimating relative and egocentric distance when using stereoscopic visualization under dynamic conditions or user controlled motion. Proposed trials concern "Collision Avoidance", (the user drives along a narrow corridor avoiding making collisions against the walls); and "Access Width". (the user is asked to estimate access width of visible doorways).

The outcome of the above experimentation is in terms of: *measurement accuracy*, (directly provided by the user as answers to questionnaires, or inferred based on number of collisions registered); and *task-completion time*, (recorded in some trials).

At the Aalborg University we have a large variety of state-of-the-art VR facilities, which represents a formidable testing ground for the proposed investigation (Figure 1 shows the VR facilities). In particular, for our investigation we have considered:

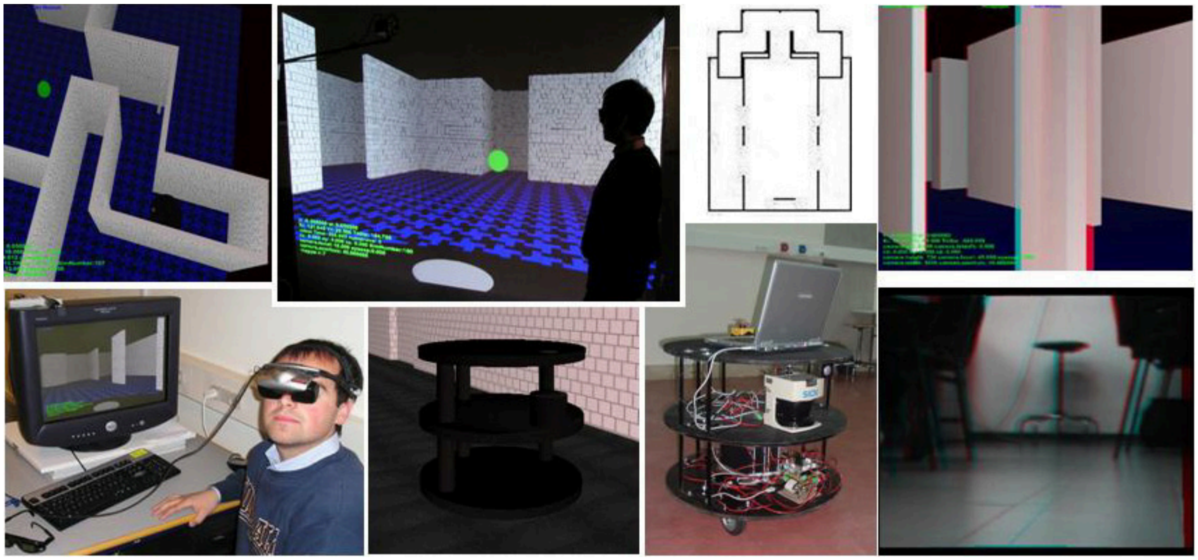


Figure 2: The simulated workspace. Top-row: 3 images on 1-sided CAVE and workspace map. Bottom-row from left: a moment during comparison testing of 3D Desktop and HMD; simulated and real robot (the Morduc system developed at DIEES, University of Catania, Italy); real stereo sequence snapshot. Right column represents 2 anaglyph images: graphical simulation and real capture.

- *3D Laptop*. Medium Notebook, 15in high-res display. Passive anaglyph.
- *3D Desktop*. 21in CRT high-res monitor. Passive anaglyph / active shutters.
- *Small Powerwall*. Front projected 1.5m x 1.5m silver screen, 2 high-res projectors. Passive stereo with polarized filters.
- *1-sided CAVE*. Rear-projected 2.5m x 2.5m screen, 1 low-res high-freq projector. Passive anaglyph / active shutters, [10].
- *Head Mounted Display*. 2x0.59in OLED LCDs 800x600. Separated displays.

Users test the systems observing both simulated and real stereoscopic images. The simulated images rendered based on a graphical model constructed as if extracted from an elevation map generated by an onboard laser rangefinder. The real images by a stereo camera setup for mobile robots expected onboard. We assess systems capabilities for different display technologies asking test users to report about their experience through questionnaires. In particular, questions are grouped into 5 judgement categories: *adequacy to application*, *realism*, *immersion*, *3D impression*, *viewing comfort*.

4. TESTING

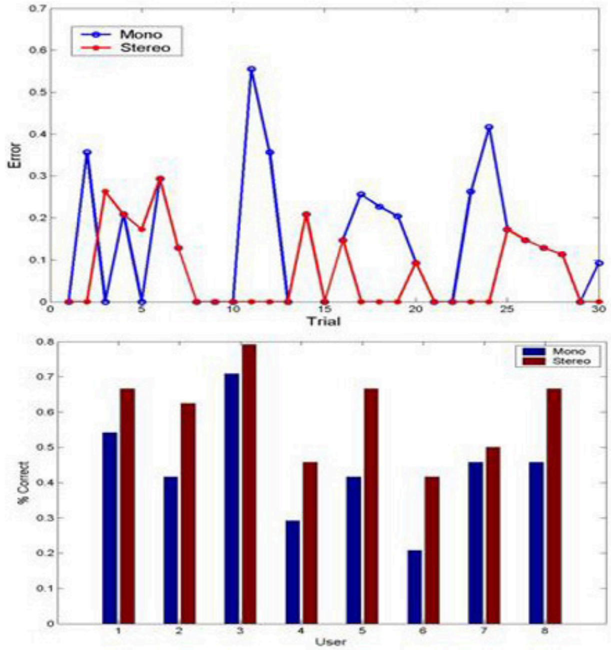
Figure 2 shows the simulated workspace. During the experimentation we altered: *Illumination*, *Texture*, *Planes depth*, *Access width*, *Robot speed*. Concerning testing with real stereoscopic images, the users were asked to observe 4 different stereo videos on the VR facilities. The cameras provided low resolution images (but in color) suitable for low-bandwidth data transmission. Figure 2 bottom-right shows a stereo snapshot from a recorded video.

After some pilot studies, 15 users were asked to run the Aptitude and the Interactive trials. An error-rate index was calculated based on user answers and ground truth [5]. The Aptitude trials consisted of 2 test trials (distance and motion). In the Interactive trials users were asked to drive through a path which combined the ”Collision Avoidance” and ”Access Width” tests. Figure 2 shows a general workspace map for the Interactive trials.

Under stereoscopic visualization users perform better the Aptitude ”Egocentric Distance” tests and the Interactive ”Access Width” tests. Variance analysis with repeated measures showed a main effect of stereo viewing on percentage of correct answers: $F=5.38$ and $p=0.0388$ (egocentric distance); $F=5.33$ and $p=0.0368$ (access width). Figure 3 shows accuracy of a typical run, and the percentage of correct answers for a representative pool of users.

The highest improvement was obtained on the 3D Desktop. We believe this may due to the lower 3D impression sometime provided in the CAVE for positive parallax, (far-end wall-planes appeared ”compressed”). The benefits of stereoscopy when estimating robot self-motion were instead not significant. This seemed to agree with the theory of Hubona et al., [7], with motion saturating the visual system so that stereopsis would not be relevant. The Interactive ”Collision Avoidance” trials did not provide significant results.

The results of the comparative tests are shown in figure 3. We can observe that users believe that larger visualization screens provide higher Realism and Adequacy of depth cues in robot teleguide than other VR systems. This goes along with Demiralp et al. considerations , [3], telling that ”looking-out” tasks (i.e. where the user views the world from inside-out as in our case), require users to use more their peripheral vision. Larger screens provide higher Immersion, (as expected). Interestingly, the sense of immersion drops in case of passive anaglyph, mostly justified by eye-strain arising from rear-projection (screen alters colors causing high crosstalk). It may surprise the reader that most users claim a higher 3D Impression with 3D Desktop rather than with the CAVE. In particular, this concerned the behind display impression (positive parallax). Confirmation that 3D Desktop perceived 3D Impression can be large, can be found in the work of Jones et al., [8]. In our case the range of perceived depth represents a large workspace portion for small screens than for larger. The highest Comfort judgement is assigned to the Small Powerwall, as confirmation of the benefits of front-projection and polarized filters.



Synthetic Images		Adequacy	Realism	Immersion	3D Impression	Comfort
3D Laptop Console	Anaglyph	4.2	2.5	2.5	5.8	2.5
3D Desktop Console	Anaglyph	4.2	2.5	3.3	6.7	3.3
	Shutters	5	3.3	4.2	7.5	5
1-sided CAVE	Anaglyph	5.8	5	3.3	6.7 / 5.8	1.6
	Shutters	6.7	5.8	7.5	7.5 / 5.8	6.7
HMD	Separated	5	4.2	5	6.7 / 5.8	5.8
Small Powerwall	Polarized	6.2	5.8	6.7	7.5	7.5

Video Images		Adequacy	Realism	Immersion	3D Impression	Comfort
3D Laptop Console	Anaglyph	1.7	3.3	2.5	5.8	2.5
3D Desktop Console	Anaglyph	2.5	2.5	4.2	3.3	2.5
	Shutters	5	4.2	4.2	6.7 / 7.5	5
1-sided CAVE	Anaglyph	5.8	5.8	4.2	5.8 / 5	8.3
	Shutters	6.7	6.7	6.7	7.5 / 5.8	6.7
HMD	Separated	5	5	5	6.7	6.7

Low (1-4) Medium (5-7) High (8-10)

Figure 3: Left: inaccuracy of a typical run from the Aptitude Egocentric Distance tests (top) and percentage of correct answers for a representative pool of test users for the Interactive Access Width (bottom). Right: results for the comparative tests with synthetic images (top) and stereoscopic videos (bottom). Cells with 2 values represent judgment for negative (left) and positive (right) parallax.

The real videos were in general best appreciated in the CAVE, then on the HMD, and then on the 3D Desktop with active stereo. A higher Realism was felt by the users compared with synthetic images (except in case of Anaglyph stereo), while Adequacy was generally worse. The Immersion category gets the same scores obtained for synthetic images but scores for passive anaglyph are higher (probably a consequence of higher realism). 3D Impression is generally worse. The viewing comfort is typically worse too. It is best for HMD and CAVE with shutters and almost unacceptable for CAVE with anaglyph. Figure 3 shows average scores for the different judgment categories.

5. CONCLUSION

The proposed work investigated the role of 3D stereoscopic visualization in applications related to mobile robot teleguide. Results from a set of test trials ran on five different VR systems emphasized few differences which represented a base for further investigation. The stereoscopic viewing improves performance in case of estimation of egocentric and related distance, while it does not show significant improvements in case of self-motion perception compared to other depth cues. A main purpose of the proposed investigation was also the comparison of different systems, which characteristics were described in terms of adequacy to application, realism, immersion, 3D impression, and viewing comfort. We hope that our comments on those aspects can support system design for specific applications in Telerobotics (to be considered in relation to cost and portability). Future investigation will concern with a deeper analysis of the Interactive tests and the extension of our tests to large VR displays such as Panoramas, Powerwalls, and the 6-sided CAVE.

6. REFERENCES

[1] O. Bimber et. al. The ultimate display: What will it be? In *SIGGRAPH'05*.

[2] C. Cruz-Neira, D. Sandin, T. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *SIGGRAPH93*.

[3] C. Demiralp, et al. Cave and fishtank VR displays: A qualitative and quantitative comparison. *IEEE Trans. on Visualization and Computer Graphics*, 2006. 12(3).

[4] D. Drascic. Skill acquisition and task performance in teleoperation using monoscopic and stereoscopic video remote viewing. In *Proc. 35th Human Factors Society*, 1991.

[5] A. Gaggioli, R. Breining. Perception and cognition in immersive vr. *Commun. Through Virtual Technology*, 2001.

[6] A. Geiser. Ergonomische grundlagen fur das raumsehen mit 3d anzeigen. Technical report, ETH Zurich, 1994. Nr.10656.

[7] G. Hubona, et al. The effects of motion and stereopsis on 3D visualization. *Human-Computer Studies*, 1997. 47.

[8] G. Jones, D. Lee, N. Holliman, D. Ezra. Perceived depth in stereoscopic images. In *Human Factors Society*, 2000.

[9] D. Kasik, J. Troy, S. Amorosi, M. Murray, S. Swamy. Evaluating graphics displays for complex 3d models. *IEEE Computer Graphics and Applications*, 2002. 22.

[10] S. Livatino et al. Designing a virtual reality game for the cave. In *Eurographics IC*, 2006.

[11] U. Naeplin, M. Menozzi. Can movement parallax compensate lacking stereopsis in spatial explorative tasks? *Elsevier DISPLAYS*, 2006. 22, 157-164.

[12] I. Sexton, P. Surman. Stereoscopic and autostereoscopic display systems. *IEEE Signal Processing Magazine*, 1999.

[13] T. Boult. DOVE: Dolphin Omni-directional Video Equipment. *IC on Robotics and Automation*, 2000.

[14] R. Ott, M. Gutierrez, D. Thalmann, F. Vexo. Advanced VR Technologies for Surveillance and Security. *IC on VR Continuum and Its Applications (VRCIA)*, 2006.