

# Depth-Enhanced Mobile Robot Teleguide based on Laser Images

S. Livatino<sup>1</sup> G. Muscato<sup>2</sup> S. Sessa<sup>2</sup> V. Neri<sup>2</sup>

<sup>1</sup> School of Engineering  
and Technology,  
University of Hertfordshire,  
Hatfield, United Kingdom

<sup>2</sup> Dipartimento di  
Ingegneria Elettrica,  
Elettronica e dei Sistemi  
University of Catania,  
Catania, Italy

## ***Abstract***

3D stereoscopic visualization may provide a user with higher comprehension of remote environment in teleoperation when compared to 2D viewing. Works in the literature have addressed the contribution of stereo vision to improve perception of some depth cues often for abstract tasks, and it is hard to find contributions specifically addressing mobile robot teleguide. The authors of this paper have investigated stereoscopic viewing in mobile robot teleguide based on video images in a previous work and pointed out advantages of stereo viewing in this type of application as well as shortcomings inherent to the use of visual sensor, e.g. image transmission delay. The proposed investigation aims at testing mobile robot teleguide based on a different sensor: the laser sensor. The use of laser is expected to solve some problems related to visual sensor while maintaining the advantage of having stereoscopic visualization of a remote environment. A usability evaluation is proposed to assess system performance. The evaluation runs under the same setup of the previous study so to have an experimental outcome comparable to the previous one. The evaluation involves several users and two different 3D visualization technologies. The results show a strong improvement in users' performance when mobile robot teleguide based on laser sensor is (depth-) enhanced by stereo viewing. Some differences are detected between the use of laser and visual sensor which are discussed.

***Key Words*** - Telerobotics, Stereo Vision, 3D Displays, Virtual Reality, Mobile Robotics.

## ***I. INTRODUCTION***

When operating in unknown or hazardous environments, accurate robot navigation is paramount. Errors and collisions must be minimized. Performance in robot teleoperation can be improved by enhancing user's sense of presence in remote environments (telepresence). Vision being the dominant human sensor modality, large attention has been paid to the visualization aspect. Robot teleoperation systems typically rely on 2D displays. These systems suffer of many limitations, e.g. misjudgement of self-motion and spatial localization, limited comprehension of remote ambient layout, object size and shape, etc. The above leads to unwanted collisions during navigation and 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, [2, 3, 4, 5, 6, 10]. The above conclusions can in principle be extended to teleguided robot navigation, where the use of stereo vision is expected to improve navigation performance and driver capabilities [3, 4, 5, 6]. However, it is hard to find work in the literature addressing mobile robot teleguide and the authors' previous work [11] is a quite unique contribution to address stereo viewing on a mobile robot. The experiments presented in [11] demonstrated that stereo viewing can significantly improve user's navigation performance on a number of variables (collisions against objects, mean speed, depth Impression, level of realism, sense of presence).

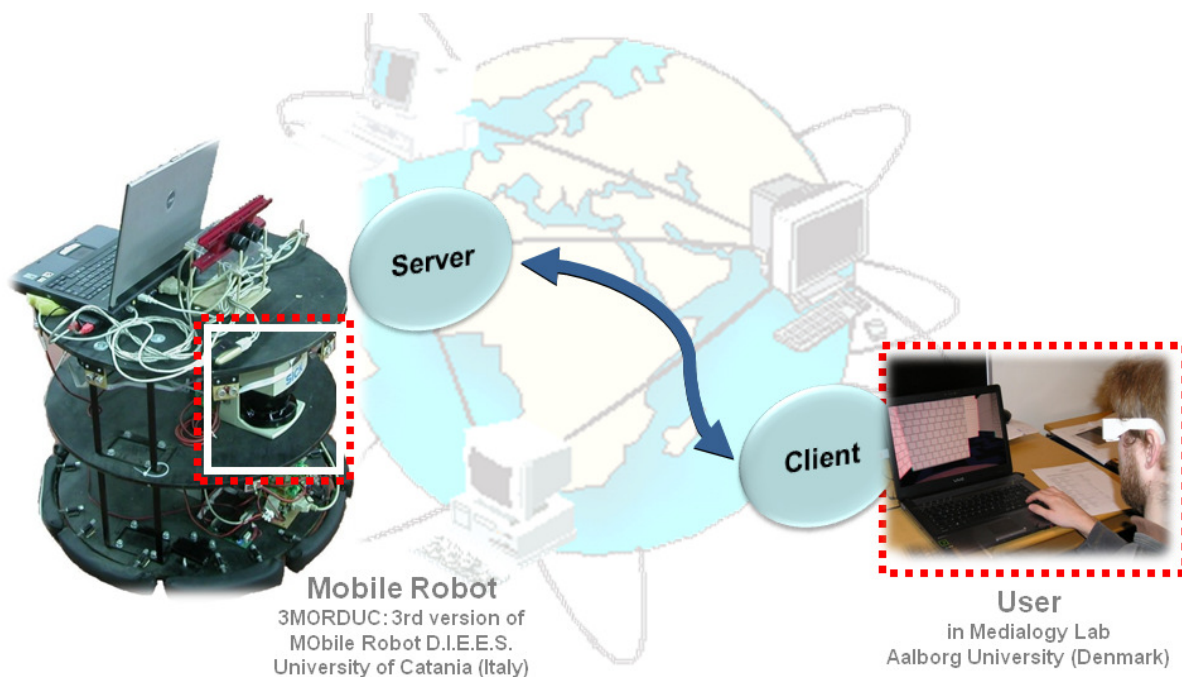
The authors' previous work investigated video-based teleoperation in mobile robot teleguide. The video sensor was considered because it provides rich and highly contrasted information. Therefore, it can largely be used in different types of robot teleguide that need accurate observation and intervention. The rich information provided by a video image may however require a large bandwidth to be transmitted at interactive rates. This often represents a challenge in video-based robot teleoperation, e.g. in case of transmission to distant locations or when the employed medium has limited communication capabilities.

A delay in image transmission is known to affect user-robot interaction performance, e.g. in terms of response time, driving speed, and manoeuvre accuracy. Corde et al. [7] showed that a delay above 1 sec. may lead to a significant decrease in performance. In the authors' previous work, [11], a nearly constant transmission delay of 1 sec. was experienced because of the bandwidth limitation.

An alternative to the use of video technology in robot teleoperation is represented by the use of laser sensor technology, which is proposed in this paper. The figure 1 illustrates the proposed general system setup for a laser-based mobile robot teleguide. The great advantage of adopting laser technology is represented by the possibility of providing real-time feedback to a tele-driving user, even in case of a very narrow communication bandwidth. The disadvantage is represented by the relatively simple description of environment characteristics that a laser-based system can provide when compared to visual sensor. There are many aspects to analyze, compare, compromise, when considering robot teleguide based on video or laser systems. Therefore a usability study is proposed.

The objectives of the proposed investigation are: (1) to assess suitability of stereo viewing in mobile robot teleguide when relying on laser-technology; (2) to analyze the role played by the laser and visual sensors towards increasing navigational accuracy in mobile telerobotic applications.

The proposed experimental setup is identical to that proposed when testing with visual sensor. This allows us to directly compare previous results (based on the use of visual sensor) to new outcome (based on the use of laser sensor).



**Figure 1: A representation of the local-remote system interaction. On the right-hand side the figure illustrates a user who sits in the Medialogy Lab in Denmark in front of a Laptop (or Wall) system. The user wears goggles to obtain 3D visual feedback of the remote environment. On the left-hand side we see our mobile robot equipped with a laser sensor located in the platform front side, responsible for measuring proximity of walls and obstacles surrounding the robot.**

## **A. Laser-based Teleoperation**

In contrast to what typically happens with visual-sensor data, laser data are interpreted by the robotic system before being transmitted and presented to a user. We rely on a laser rangefinder, a type of laser sensor often proposed on mobile robots to assist navigation. This device can be very effective in measuring proximity of walls and obstacles surrounding a robot. It can provide accurate estimate of distance and direction to a detected obstacle.

The accuracy of laser systems has made it suitable for extracting 2D floor maps of a robot workspace. 3D maps can be obtained by combining more sensors readings or a by letting the laser device move.

A 2D floor map of the environment surrounding a robot represents very small information compared to a video image, which can be quickly transmitted over a network. This aspect makes the use of laser very suitable for teleoperation. The provided laser-based information needs however to be conveniently processed and presented to a user in order to be beneficial for teleoperation.

This paper proposes a method that benefits from quick transmission of laser information to a remote user and conveniently presents him/her the sensor data visually, through computer graphics.

## **B. Stereoscopic Viewing and Displays**

The performance in mobile robot teleguide is affected by the capability of a user to estimate: spatial localization, spatial configuration, depth relationships, motion perception, and action control, [11]. The possibility for stereoscopic visualization influences some of these factors to a different extent, depending on available space and budget, type of robot platform and sensor data, as well as chosen approach for stereo viewing and visual display type.

Different types of display are today available, and they can be characterized by display size and structure, projection technology, image quality and observation condition. Different displays technologies have also been developed for generating 3D stereoscopic visualization, [4]. 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. Furthermore, stereoscopy can increase user's involvement and immersion, due to the increased level of depth awareness, and this leads to a more accurate action performance and environment comprehension. There are several works in the literature that focus on stereoscopic visualization. These can be classified as application oriented user studies, or abstract tasks and content with general performance

criteria, [2]. The parameters through which to assess stereoscopy benefits typically are: item difficulty and user experience, accuracy and performance speed, [9].

Stereoscopic visualization is claimed to improve comprehension and appreciation of presented visual input (perception of scene structures and surfaces, object motion, etc.), and to facilitate human-machine interaction [1, 3, 4]. Most of the benefits of stereo viewing may improve robot teleguide, however, the users' performance may be challenged by eye strain, double images perception, depth distortion, etc., [10].

The proposed investigation strategy is presented in next section. It follows the experimental design (section III), and the results analysis (sections IV and V). Some final remarks conclude the paper (section VI).

## ***II. PROPOSED INVESTIGATION***

The two main objectives of the proposed investigation are:

- 1) **Performance of Laser-Sensor and Stereo-Viewing.** To assess suitability of stereo viewing in mobile robot teleguide when relying on laser-technology.
- 2) **Comparison Laser-Sensor and Visual-Sensor.** To compare performance of robot teleguide based on laser sensor against that based on visual sensor evaluated in previous experiments.

### **A. Performance of Laser-Sensor and Stereo-Viewing**

The stereo visualization has demonstrated its great potential in improving performance of mobile robot teleguide when using video images. It is therefore proposed a system that visually presents laser-based measurements to a tele-driving user.

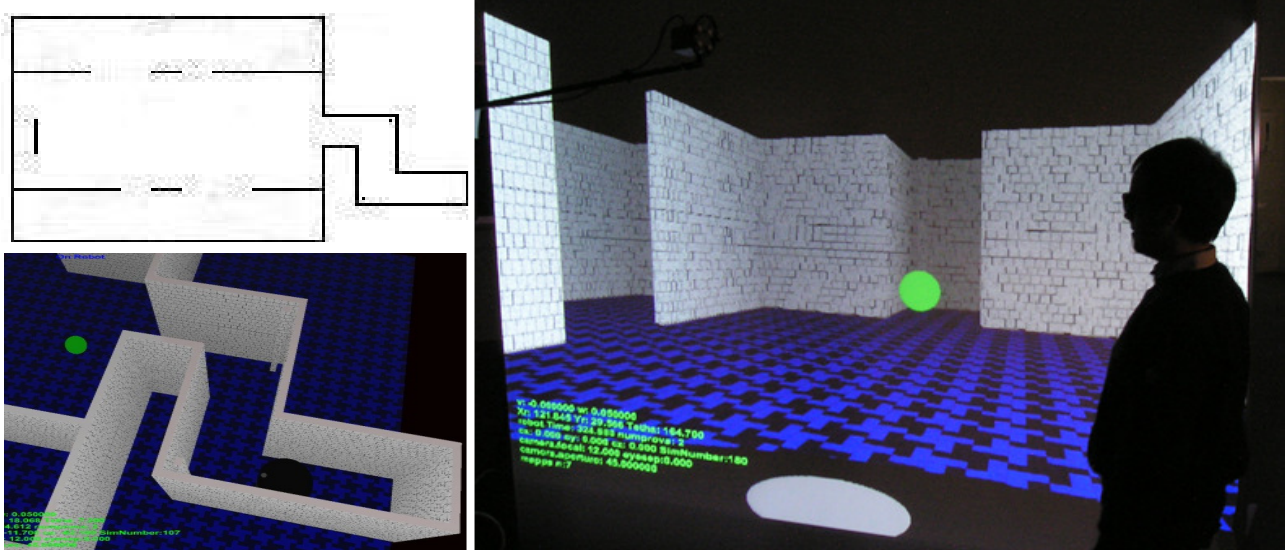
An additional challenge for the proposed stereoscopic visualization is that our visual representation of the environments is rich of strong monocular depth cues. The binocular depth cues are therefore less needed to comprehend depth relationships in the visualized sceneries. This makes nevertheless more meaningful any detected performance improvement under stereo-viewing conditions.

The system is designed to allow a tele-driving user to examine proximity measurements in a way that it is: real-time, visual, and intuitive. In particular:

1. **Real-time.** The information that will be presented to a tele-driving user will correspond to current situation on the remote site. This represents a main advantage compared to the use of video technology. Users can achieve a better perception of robot position and orientation and they can manoeuvre the robot more skilfully. Users will be able to drive faster and make rapid decisions.
2. **Visual.** Users can exploit the advantages of a visual representation and the option of having stereo viewing.
3. **Intuitive.** The visual information needs to be presented in a way that is comprehensible and of easy catch. This allows for prompt users' reaction and real-time transmission to their commands to control the robot.

The progress of last years algorithms in environment-map reconstruction based on laser measurements, allows us today to reliably construct in real-time 2D-maps of robot surrounding workspace. A reconstructed 2D floor-map can be represented as a 2D image, e.g. a black and white image where black pixels describe detected obstacles and white pixels free space. The figure 2 includes an example of a 2D map. This representation has the advantage of being "light", being contained in few Kbytes of information (which can further be reduced by applying image compression). The constructed 2D-map can be processed onboard the robot in real-time and it can quickly be transmitted through a network connection like the Internet. This allows for real-time communication between the robot and the teleoperator's site.

A 3D representation of the observed map can be extrapolated from the 2D floor map by elevating wall lines and obstacle posts. Current front-views of robot workspace can then be generated and visualized on the user's screen by using graphical software. The 3D map building and visualization can be performed in real-time and these operations can be executed on the teleoperator's computer. The figure 2 illustrates the process of building up the 3D map.



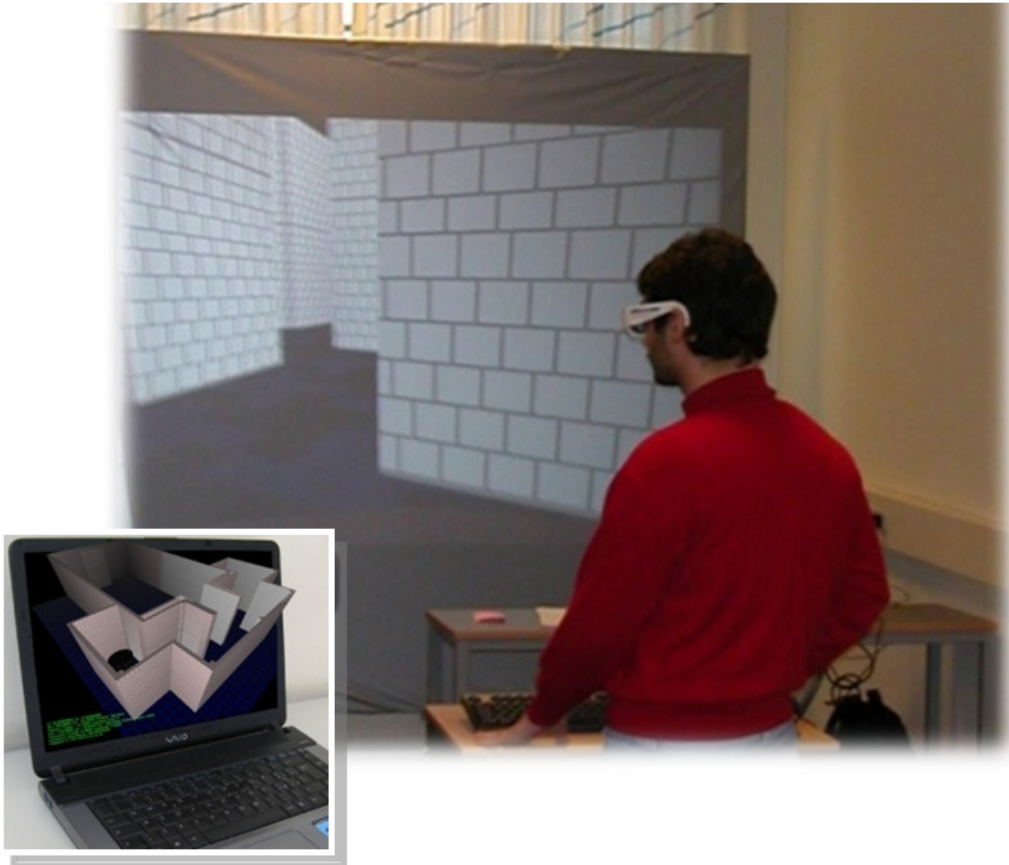
**Figure 2: The process of generating 3D graphical environment views from laser range information. The top-left image shows a 2D floor-map generated by the laser sensor. The bottom-left image shows a 3D extrapolation of a portion of it. The right-image shows a portion of the workspace visible to a user during navigation.**

Two different 3D visualization facilities are proposed in our investigation to evaluate performance on systems with different characteristics, cost and application context. The aim is to gain insight into the problem and to understand on what system, and to what extent, are display type and stereo viewing beneficial.

The two proposed visual displays are:

- **Laptop.** This display uses LCD technology and it has a relatively small display size, typically up to 19 inches, with high resolution.
- **Wall.** This display is typically composed by a projector and a screen with a size up to several meters. Our system is front projected.

The figure 3 shows the visualization systems used in our tests.



**Figure 3: The visualization systems used in our tests: the Laptop (left) and the Wall (right).**

The two proposed approaches to stereo viewing are:

- **Colored Anaglyph.** This approach is very economic, easy to produce and very portable. However, it has poor colour reproduction and it often generates crosstalk which affects precision and viewing comfort.
- **Polarized Filters.** This approach nicely reproduce colours, has nearly no crosstalk, and it is very comfortable to a viewer. However, it requires a more complex and expensive setup and it is less portable than Anaglyph

## **B. Comparison Laser-Sensor and Visual-Sensor**

The comparison between laser and visual sensor is relevant because of the different nature of the information provided may affect different systems and users' behaviours. The comparison therefore gives us insight on the role played by different behaviours on teleoperation performance. It also gives us indications



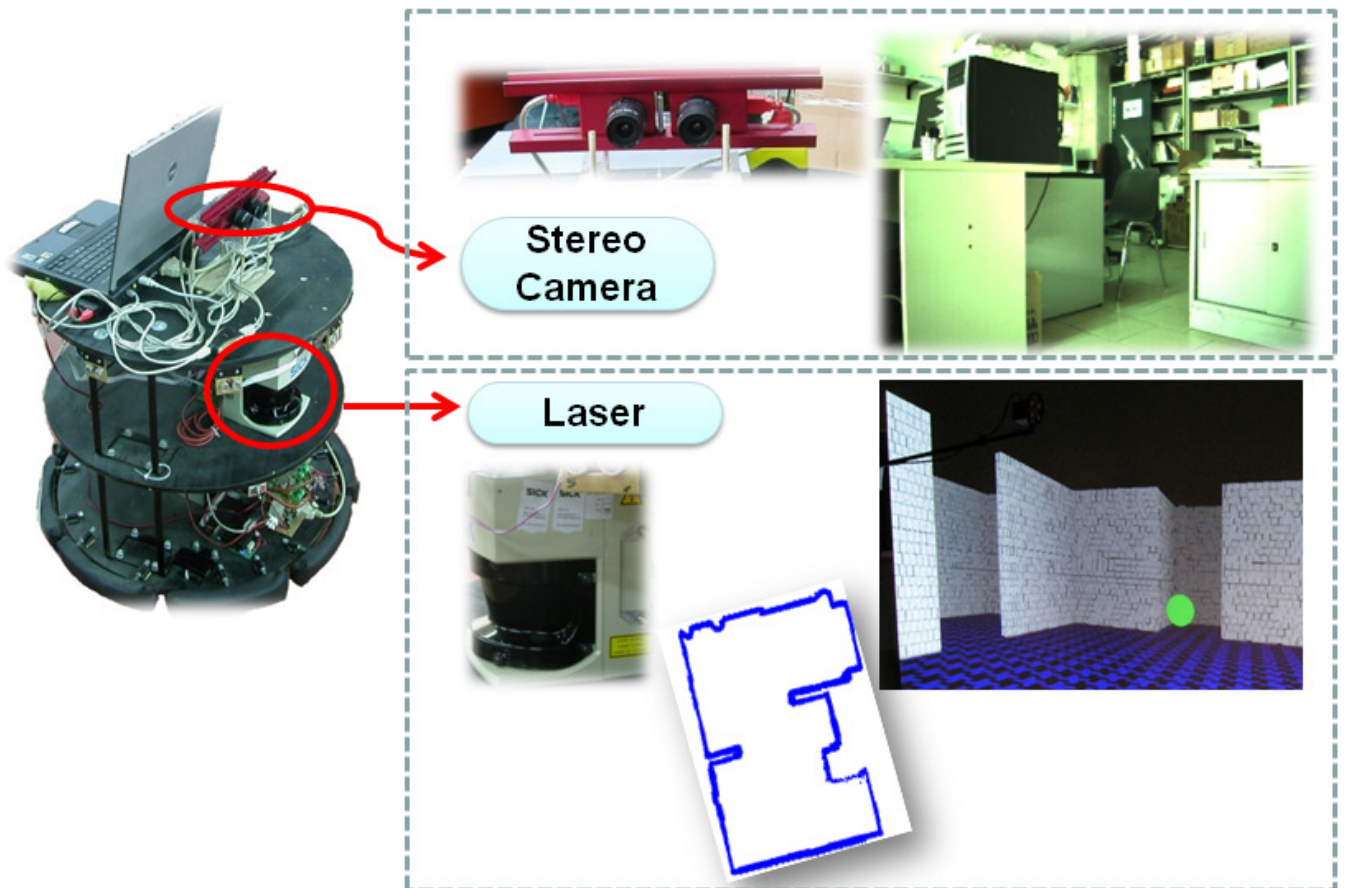
for future developments of teleguide systems based on multiple sensors and augmented reality visualization.

The system and users' behaviours are expected to be affected by:

- **Information Amount:** Our laser sensor provides a smaller amount of information than the visual sensor. The information only consists of distance measures on a specific horizontal plane (the one at the same height of the laser device). The visual sensor provides instead a much richer photo-like information of the workspace. Nevertheless, laser measures are very precise (with errors of the order of millimetres).
- **Visualization Detail:** The 2D map synthesized from laser measurements is made visual and intuitive by generating 3D front-views of robot workspace (based on estimate of current robot position). The "laser images" can so be observed the same way as video images. The laser images show however a lower level of detail because the represented environment features only correspond to an extension of a floor map. There may be therefore a substantial approximation in visualized images. Furthermore, our laser-based visual representation typically shows only few planar surfaces. This may have consequences on obstacle perception and their visual estimation. The figures 2 and 3 show examples.
- **Action Response:** When relying on laser sensor, the image of the remote environment is presented to a tele-driving user, is visualized in real-time. Users can therefore respond in real-time and observe in real-time the effect of their response. This behaviour is very different from that occurring in video-based teleoperation.

For the proposed evaluation we keep the same experimental setup proposed when testing with visual sensor. This way the expected outcome is directly comparable to previous experiments.

Our comparative study looks at differences and similarities between laser and video in the proposed robot teleguide. Our study also looks at specific differences associated with different display types and approaches to stereo. The illustrations in figure 4 summarize the different components of a laser and visual sensors based teleguide systems.



**Figure 4: The figure summarizes the different components and an example of visual result of a laser and visual sensors based teleguide systems.**

### ***III. EXPERIMENTAL DESIGN***

The proposed evaluation aims at detecting the overall usability of the proposed system. The purpose is to obtain tangible proof of user's navigation skills and remote environment comprehension, under different circumstances. The research question involves the following three aspects:

- **Mono versus Stereo.** What are the main characteristics and advantages of using stereoscopic visualization in mobile robot teleguide in terms of navigation skills and remote environment comprehension?
- **Anaglyph Laptop versus Polarized Wall.** How may the characteristics and advantages associated to stereoscopic viewing vary for different approaches of stereo and display systems?
- **Laser Sensor versus Visual Sensor.** What are main performance differences between laser sensor and on visual sensor in mobile robot teleguide?

The usability study is designed according to recommendation gathered from the literature and authors' experience and previous work on evaluation of VR applications, [8]. The study is a within-subjects evaluation with 12 participants in case of the first objective (Performance of Laser-Sensor and Stereo-Viewing) and a between-subjects evaluation with 24 participants in the second objective (Comparison Laser-Sensor and Visual-Sensor).

Each participant is asked to tele-drive a remotely located mobile robot on both the proposed facilities (Laptop and Wall systems), using both stereoscopic and monoscopic visualization. This results in 4 navigation trials per participant. The approaches for questionnaires and activities schedule follow the same recommendations given in [11]. The study considers quantitative and qualitative evaluations, and it includes the same evaluation measurements and subjective parameters as in [11]. The evaluation measurements are: Collision Rate, Collision Number, Obstacle Distance, Completion Time, Path Length, and Mean Speed. The subjective parameters are: Depth Impression, Suitability to Application, Viewing Comfort, Level of Realism, and Sense of Presence. The acquired data follows the same approach proposed in [11] for the statistical and graphical evaluation.

The experiment involved facilities on different sites: local and remote. The remote site is the location where the robot operated. This was the Robotics laboratory at the DIEES, University of Catania, Italy. The local site is the location where the user (tele-) operated. This was the Medialogy lab at the Aalborg University in Copenhagen, Denmark. The figure 1 illustrates local and remote systems.

Similarly to previous experiments is the indoor environment and camera setup. However, this time the cameras are virtual so are the images, (referred as "laser images"). The test setup is the same for what concerns: robotic and laser systems, visualization systems, network connection, test organization and procedure. The setup is different in the following aspects:

- 1) **Map Building and Graphical Rendering:** The laser measurements are processed by the on board PC (Mobile AMD *Athlon* 796MHz, 512MB RAM) before being transmitted through the Internet. Users observe on their screen views of the 3D model generated by a graphical simulator built in *C++* language using the *OpenGL* graphic libraries.

- 2) **Participants:** The target population, composed by participants with varying background and none or medium experience with virtual reality devices, has an age that ranges between 23 and 40, with an average age of 26.2.

#### ***IV. RESULTS ANALYSIS:***

##### ***PERFORMANCE OF LASER-SENSOR AND STEREO-VIEWING***

The results of the experimentation are shown in figures 7 and 8 for the descriptive statistics and tables 1 and 2 for the inferential statistics. We measure statistical significance of results by estimating the Analysis of Variance (ANOVA). In particular a two-way ANOVA is applied to measure the effect of “Stereo-Mono” and “Laptop-Wall”, on each of the dependent variables, (the quantitative evaluation measurements and qualitative subjective parameters). We set to 0.05 the p-value to determine whether the result is judged statistically significant.

The results for the first objective are presented and commented as in our previous work, [11], to facilitate a comparison among those two investigations. A comparison is nevertheless specifically addressed in a systematic manner (second objective), supported by a statistical analysis which is presented in next section. In this section the results are presented according to the proposed research questions.

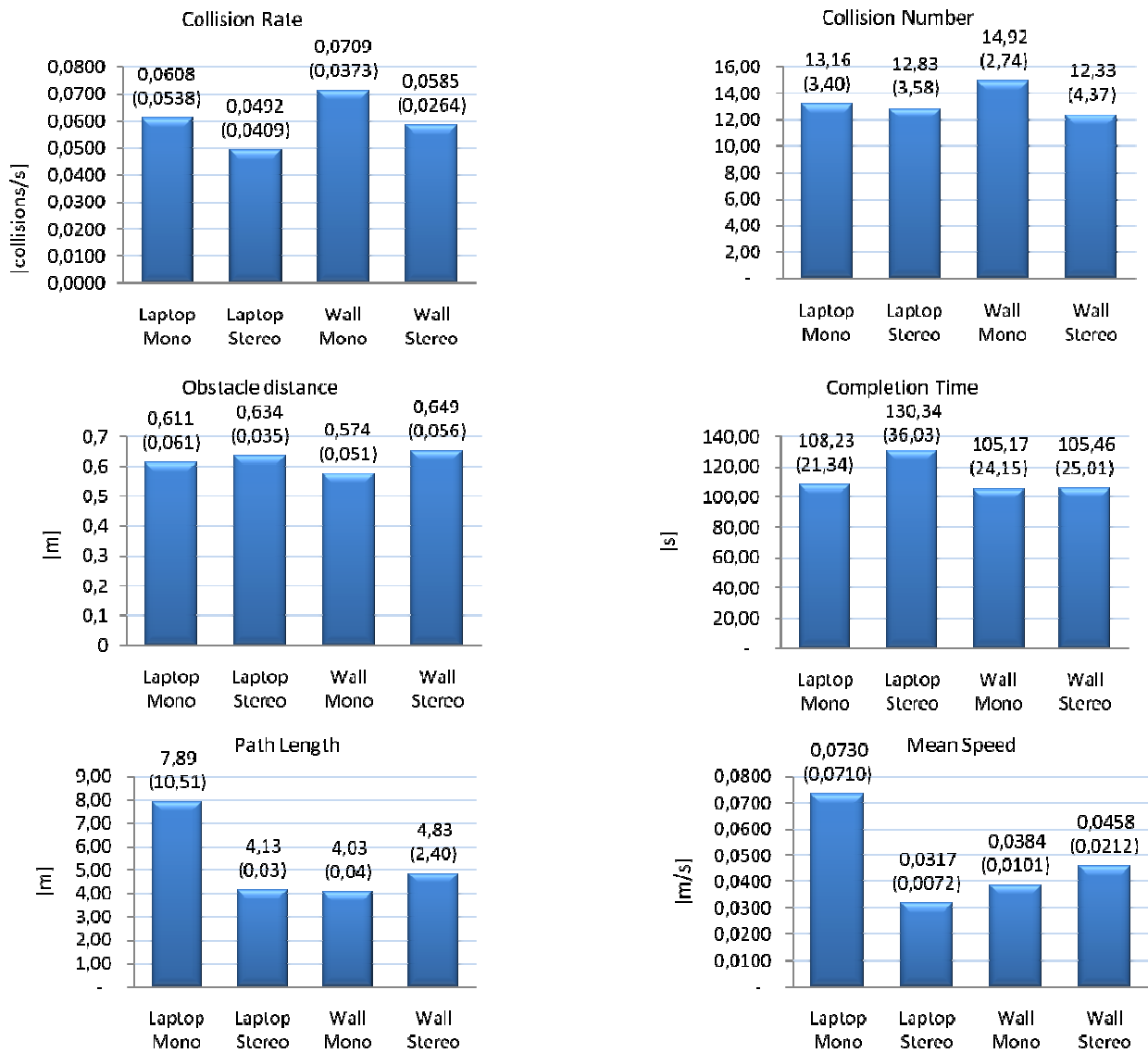
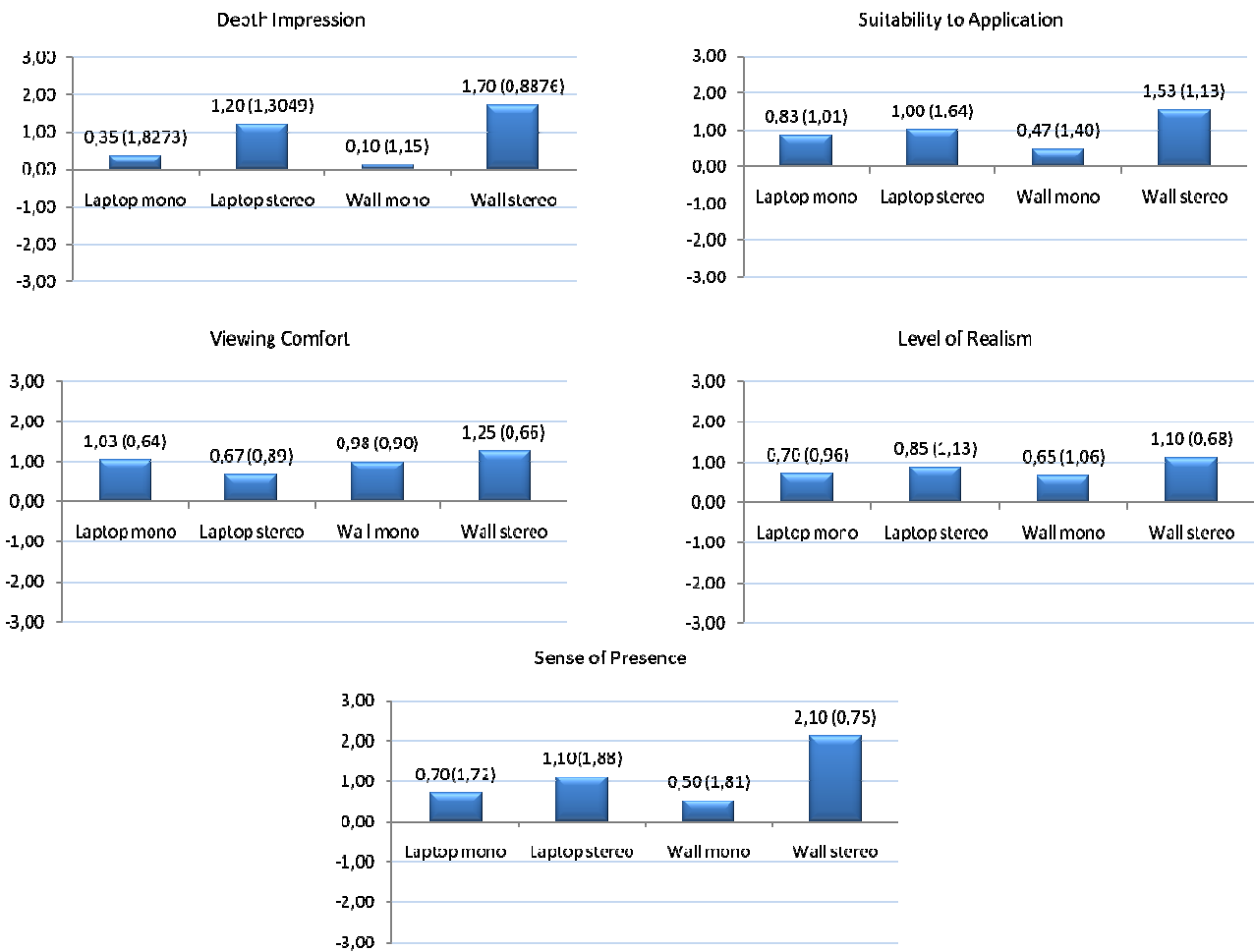


Figure 7: Bar graphs illustrating mean value and standard deviation (in brackets) for the quantitative variables.



**Figure 8: Bar graphs illustrating mean value and standard deviation (in brackets) for the qualitative variables. The qualitative data were gathered through questionnaires where the participants provided their opinions by assigning values which ranged between +3 (best performance) and -3 (worst performance).**

Table 1. The results of two-way ANOVA for the quantitative measurements.									
	Collision Rate				Collision Number				
	SS	DoF	F	P	SS	DoF	F	P	
<b>Mono-Stereo</b>	0.00942	1	6.15	0.017	<b>Mono-Stereo</b>	25.521	1	1.95	0.1694
<b>Laptop-Wall</b>	0.00674	1	4.4	0.0418	<b>Laptop-Wall</b>	0.021	1	0	0.9683
<b>Interaction</b>	0.00203	1	1.32	0.2564	<b>Interaction</b>	7.521	1	0.58	0.4522
	Obstacle distance				Completion Time				
	SS	DoF	F	P	SS	DoF	F	P	
<b>Mono-Stereo</b>	18230.5	1	5.99	0.0185	<b>Mono-Stereo</b>	1504.2	1	2.09	0.1557
<b>Laptop-Wall</b>	6224.6	1	2.04	0.1598	<b>Laptop-Wall</b>	4631	1	6.42	0.0149
<b>Interaction</b>	662	1	0.22	0.6433	<b>Interaction</b>	9.1	1	0.01	0.9112
	Path Length				Mean Speed				
	SS	DoF	F	P	SS	DoF	F	P	
<b>Mono-Stereo</b>	26.45	1	0.9	0.3482	<b>Mono-Stereo</b>	0.002	1	1.47	0.2325
<b>Laptop-Wall</b>	29.91	1	1.02	0.3187	<b>Laptop-Wall</b>	0.00016	1	0.11	0.7446
<b>Interaction</b>	50.05	1	1.7	0.1988	<b>Interaction</b>	0.00248	1	1.69	0.2008

Table 1: The results of 2-way ANOVA for the quantitative measurements. Rows show values for the independent variables (Mono-Stereo, Laptop-Wall), their interaction, and error. Columns show the sum of squares (SS), the degrees of freedom (df), the F statistic and the p-value.

Table 2. The results of two-way ANOVA for the qualitative measurements.									
	Depth Impression				Suitability to Application				
	SS	DoF	F	P	SS	DoF	F	P	
<b>Mono-Stereo</b>	27.755	1	15.18	0.0003	<b>Mono-Stereo</b>	6.0208	1	3.33	0.0748
<b>Laptop-Wall</b>	0.255	1	0.14	0.7105	<b>Laptop-Wall</b>	2.2245	1	1.23	0.2734
<b>Interaction</b>	0.422	1	0.23	0.6333	<b>Interaction</b>	0.1875	1	0.1	0.749
	Viewing Comfort				Level of Realism				
	SS	DoF	F	P	SS	DoF	F	P	
<b>Mono-Stereo</b>	0.0133	1	0.02	0.8894	<b>Mono-Stereo</b>	4.0833	1	3.95	0.0531
<b>Laptop-Wall</b>	0.2133	1	0.31	0.5788	<b>Laptop-Wall</b>	0.0208	1	0.02	0.8877
<b>Interaction</b>	0.12	1	0.18	0.6769	<b>Interaction</b>	0.1875	1	0.18	0.6722
	Sense of Presence								
	SS	DoF	F	P					
<b>Mono-Stereo</b>	15.188	1	5.47	0.024					
<b>Laptop-Wall</b>	2.521	1	0.91	0.346					
<b>Interaction</b>	1.021	1	0.37	0.5475					

Table 2: The results of 2-way ANOVA for the qualitative measurements. Rows show values for the independent variables (Mono-Stereo, Laptop-Wall), their interaction, and error. Columns show the sum of squares (SS), the degrees of freedom (df), the F statistic and the p-value.

## **A. Mono versus Stereo**

### **1) Collision**

Under stereoscopic visualization users perform significantly better in terms of Collision Rate. The ANOVA shows a main effect of stereo viewing on the number of collisions per time unit:  $F=6.15$  and  $p=0.017$ . The improvement when comparing mean values is similar on both facilities. This is 18.3% in average.

Both Collision Rate and Collision Number are higher in case of monoscopic visualization, both as mean value and in most users' trials. This supports the expectation, based on the literature, that the higher sense of depth provided by stereo viewing may improve driving accuracy.

### **2) Obstacle Distance**

Under stereoscopic visualization users perform significantly better in terms of Obstacle Distance. The ANOVA has  $F=5.99$  and  $p=0.0185$ . The improvement when comparing mean values is higher on the larger screen: 11.5%.

### **3) Completion Time**

There is no significant difference in Completion Time between mono and stereo viewing. Nevertheless, we have observed that the time employed for a trial is greater in stereo visualization in most of the trials. The test participants have commented that the greater depth impression and sense of presence provided by stereoscopic viewing, make a user spending a longer time in looking around the environment and avoid collisions.

### **4) Path Length**

There is no significant difference in Path Length. The users show different behaviours on the facilities under mono and stereo conditions. In the Laptop we have a reduction of path length in mean values of 48% under stereo viewing conditions. An increase of length in mean values is instead observed in the Wall under the same viewing conditions. Generally, the path is more accurate and well balanced in stereo viewing, which justifies the above mentioned significant improvement in the Obstacle Distance measurement.



### 5) Mean Speed

There is no significant difference in Mean Speed. The results show opposite trends. Users drive faster on Laptop in mono viewing. This is probably one of the causes for more collisions with this facility and configuration.

### 6) Depth Impression

Most of the users had no doubts that Depth Impression is higher in case of stereo visualization. The result from the ANOVA shows a main effect of stereo viewing:  $F=15.18$  and  $p=0.0003$ . This result is expected and agrees with the literature.

### 7) Suitability to Application

The Suitability to Application ANOVA shows a tendency to significant ( $F=3.33$  and  $p=0.0748$ ). Most of the users found stereoscopic visualization more adequate for the assigned teleguide task. We notice an improvement of 69.3% on mean values in case of polarized stereo. Anaglyph stereo penalizes the final result, (only 17% improvement).

### 8) Viewing Comfort

There is no significant difference in Viewing Comfort between stereo and mono visualization and we observe opposite trends in mean values. This result contradicts the general assumption of stereo viewing being "painful" compared to mono. Stereo viewing is even considered more comfortable than mono in the Polarized Wall. The higher sense of comfort on the Wall system is claimed to be obtained by a stronger depth impression in stereo. Our conclusion is that the low discomfort of polarized filters is underestimated as effect of the strong depth enhancement provided in the Polarized Wall.

### 9) Level of Realism

The synthetic images generated from laser data and visualized by the graphic simulator show simple and planar environment features. This affects the perceived level of visual realism. All users find nevertheless that stereo visualization provides more realism than mono viewing. The result from the ANOVA shows a tendency to significant ( $F=3.95$  and  $p=0.0531$ ). The mean values show an improvement of 17.6% on Laptop and 40.9% on Wall.

#### 10) **Sense of Presence**

Most of the users believe that stereo visualization enhances presence in the observed remote environment. The ANOVA has  $F=5.4$  and  $p=0.024$ . The improvement in mean values is 36.4% on Laptop and 69% on Wall.

### **B. Anaglyph Laptop versus Polarized Wall**

#### 1) **Collision**

Users perform significantly better in the Laptop system in terms of Collision Rate. The ANOVA has  $F=4.4$  and  $p=0.0418$ . The improvement when comparing mean values is 15%. The Collision Number ANOVA shows no significant difference between the two systems. The effect of stereoscopic visualization compared to monoscopic is analogous on both facilities, with stereo viewing performing better in mean values.

#### 2) **Obstacle Distance**

There is no significant difference between the two systems and the improvement when comparing mean values is only 1.7%. It is the mono-stereo viewing condition that makes a more relevant contribution on this measurement rather than the facility.

#### 3) **Completion Time**

Users perform significantly better in the Wall system. The ANOVA has  $F=6.42$  and  $p=0.0149$ . The improvement in mean value is 11.7%. Most of the participants argued that the faster performance is due to the higher sense of presence given by the larger screen. The higher presence enhances driver's confidence. Therefore a smaller time is employed to complete a trial.

#### 4) **Path Length**

There is no significant difference in Path Length. Nevertheless, most of the users operating on the Wall system ran along paths 23.6% shorter in mean value. The mean values show different trend in mono and stereo performance on the two facilities

## **5) Mean Speed**

There is no significant difference in Mean Speed. The slower mean speeds are typically detected on the Wall. The mean values show different patterns for mono-stereo performance on the two facilities, which seems to be the consequence of the similar pattern in Path Length.

## **6) Depth Impression**

There is no significant difference between the two facilities. This confirms that the role played by the stereoscopic visualization is more relevant than the change of facilities. Both on Laptop and Wall the results show very similar trends. The improvement when driving under stereo-viewing conditions is 71% on the Laptop and 94% on the Wall. The results show that even on a Laptop system a very high 3D impression can be perceived. A result confirmed in the literature, [6].

## **7) Suitability to Application**

There is no significant difference between the two systems. Looking at the mean value, we can only observe that users in mean value believe that a large visualization screen is more suitable to mobile robot teleguide under stereo visualization. The larger screen should be considered more suitable according to the literature, [2], because our robot teleguide is a "looking-out" task (i.e. where the user views the world from inside-out as in our case), which require users to use their peripheral vision more than in "looking-in" tasks (e.g. small object manipulation). This is not the case shown in mean value of the Wall mono. Based on user's comments, the reason seems to be that the Laptop system is much appreciated as low-cost and portable facility.

## **8) Viewing Comfort**

There is no significant difference between the two systems. However, the mean values best result is perceived in case of the Wall in stereo viewing. This result is expected and it confirms the benefit of front-projection and polarized filters which provide limited eye-strain and crosstalk, and great colour reproduction. The benefits are so appreciated to make most users believe that the Wall in stereo is more comfortable than the Wall in mono. An opposite trend in mean values is detected for the Laptop facility. Here the passive Anaglyph technology (Laptop stereo) strongly affects viewing comfort and it calls for high brightness to mitigate viewer discomfort.

#### 9) **Level of Realism**

There is no significant difference between the two systems. The mean values of Level of Realism show the same trends on the two facilities with stereo viewing better performing.

#### 10) **Sense of Presence**

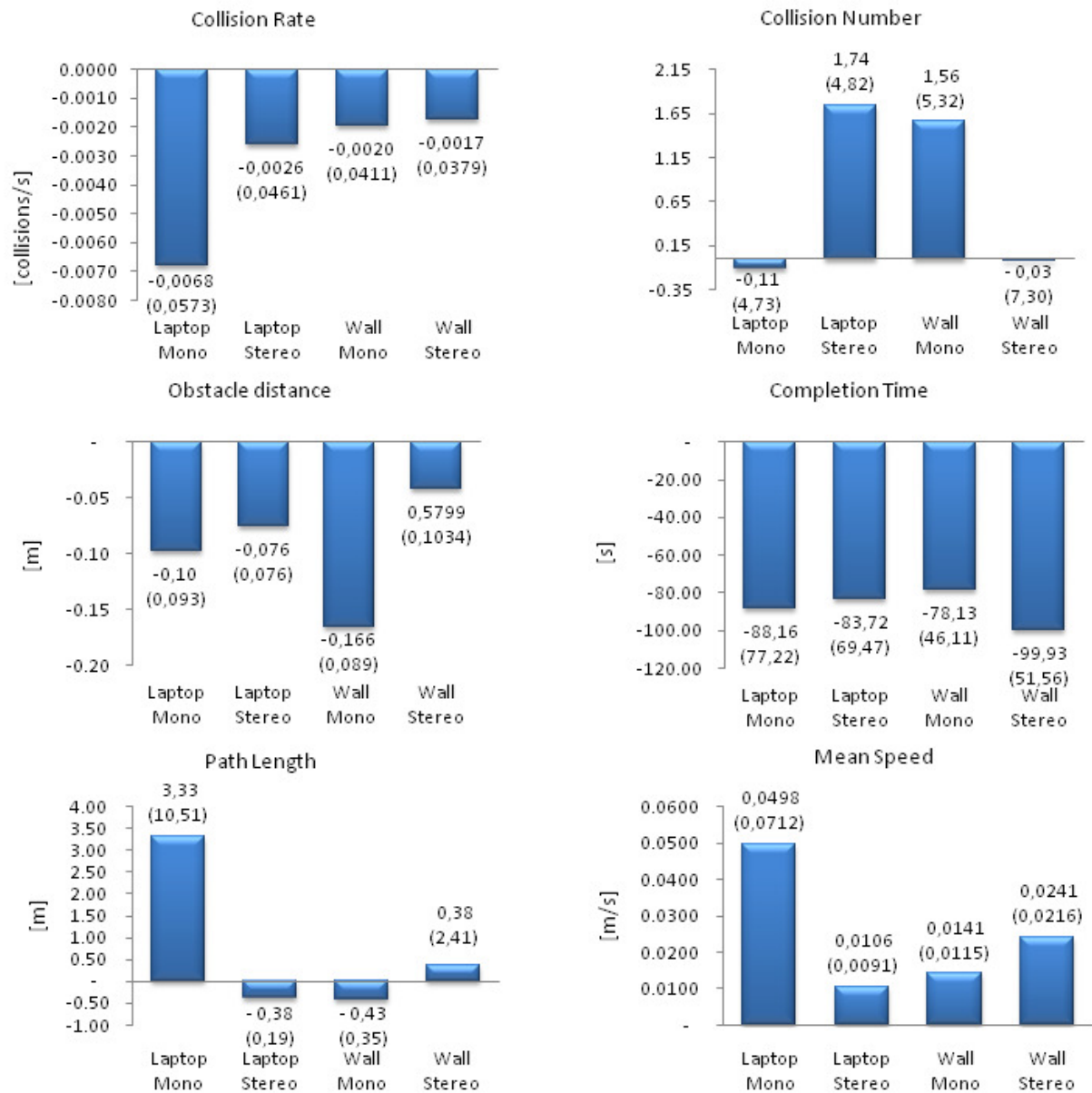
There is no significant difference between the two systems. The mean values show the same trend on both the facilities with Sense of Presence higher under stereo visualization. The improvement under stereo viewing is higher in mean value for the Wall system (76%) than the Laptop (36%).

### ***V. RESULTS ANALYSIS:***

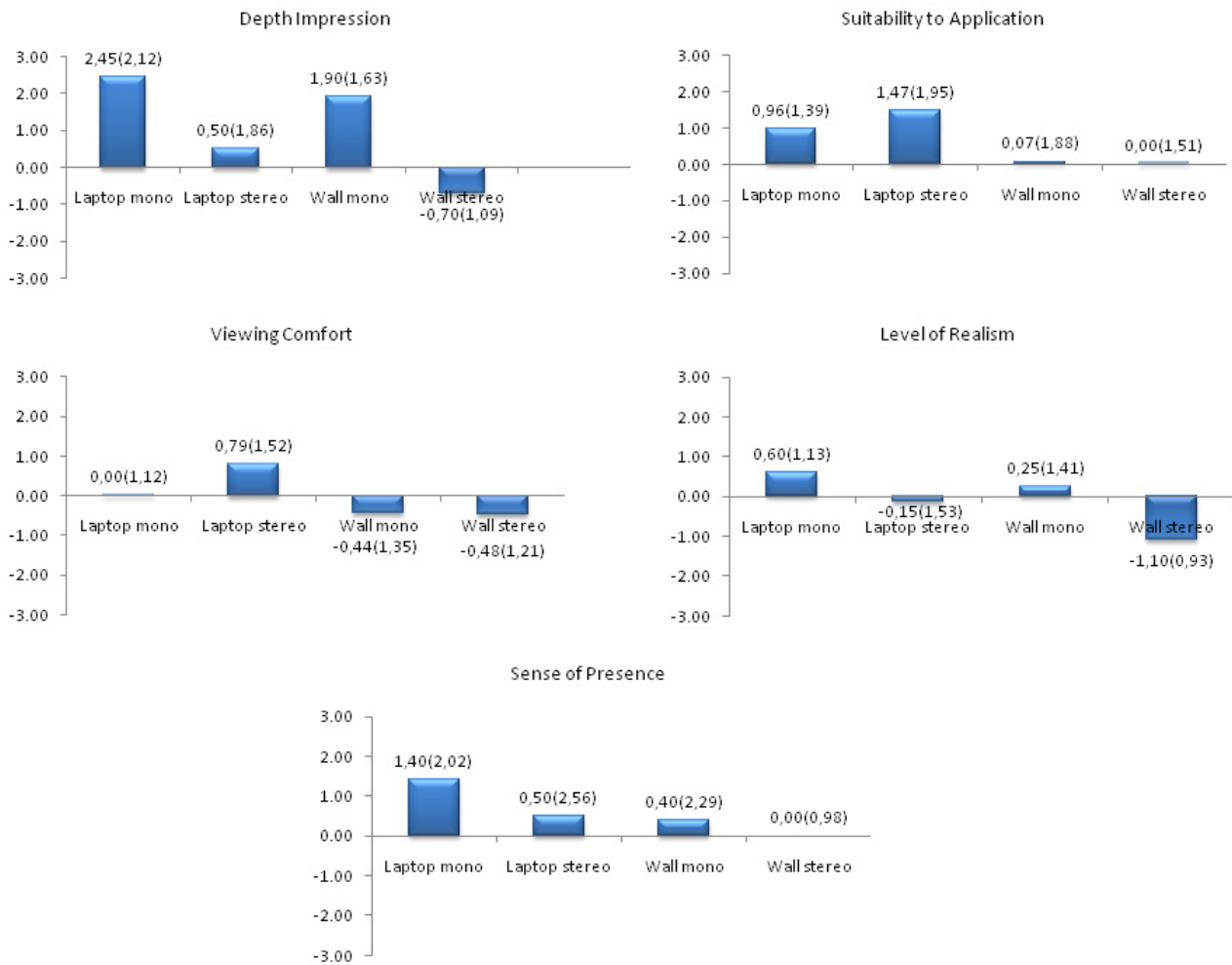
#### ***COMPARISON LASER SENSOR – VISUAL SENSORS***

The figures 9 and 10 show descriptive statistics. In particular they show the difference between mean values that were estimated for the video and laser -based robot teleguide. The tables 3 and 4 show inferential statistics. As in case of the first objective we measure statistical significance of results by estimating the ANOVA. In this case a two-way ANOVA is applied to measure the effect of “Mono-Stereo” and “Laser-Video” on each of the dependent variables. Both data from video and laser trials are considered.

In this section the results are commented for each quantitative and qualitative parameter.



**Figure 9: Bar graphs illustrating difference in mean values (and standard deviation in brackets) for the quantitative variables of laser and video based teleguide.**



**Figure 10: Bar graphs illustrating differences in mean values (and standard deviation in brackets) for the qualitative variables of laser and video based robot teleguide. The qualitative data were gathered through questionnaires where the participants provided their opinions by assigning values which ranged between +3 (best performance) and -3 (worst performance).**

Table 3. The results of two-way ANOVA for the quantitative measurements.									
	Collision Rate					Collision Number			
	SS	DoF	F	P		SS	DoF	F	P
<b>Mono-Stereo</b>	0.01026	3	1.9	0.1336	<b>Mono-Stereo</b>	66.67	3	1.45	0.2335
<b>Laser-Video</b>	0.00048	1	0.2	0.604	<b>Laser-Video</b>	39.56	1	2.59	0.1116
<b>Interaction</b>	0.00506	3	0.95	0.4227	<b>Interaction</b>	47.76	3	1.04	0.3789
	Obstacle distance					Completion Time			
	SS	DoF	F	P		SS	DoF	F	P
<b>Mono-Stereo</b>	1430.1	3	0.07	0.9739	<b>Mono-Stereo</b>	8469.2	3	0.73	0.5376
<b>Laser-Video</b>	31693.5	1	4.9	0.0296	<b>Laser-Video</b>	1508.1	1	0.39	0.5344
<b>Interaction</b>	8399.1	3	0.43	0.7298	<b>Interaction</b>	621.3	3	0.05	0.9836
	Path Length					Mean Speed			
	SS	DoF	F	P		SS	DoF	F	P
<b>Mono-Stereo</b>	26.45	1	0.9	0.3482	<b>Mono-Stereo</b>	74.05	3	1.44	0.2379
<b>Laser-Video</b>	29.91	1	1.02	0.3187	<b>Laser-Video</b>	3.99	1	0.23	0.6311
<b>Interaction</b>	50.05	1	1.7	0.1988	<b>Interaction</b>	62.67	3	1.22	0.309

Table 3: The results of 2-way ANOVA for the quantitative measurements. Rows show values for the independent variables (Mono-Stereo, Laser-Video), their interaction, and error. Columns show the sum of squares (SS), the degrees of freedom (df), the F statistic and the p-value.

Table 4. The results of two-way ANOVA for the qualitative measurements.									
	Depth Impression					Suitability to Application			
	SS	DoF	F	P		SS	DoF	F	P
<b>Mono-Stereo</b>	107.75	3	21.11	0	<b>Mono-Stereo</b>	27.771	3	5.68	0.0014
<b>Laser-Video</b>	2.557	1	1.5	0.2238	<b>Laser-Video</b>	2.122	1	1.3	0.2573
<b>Interaction</b>	2.693	3	0.53	0.6645	<b>Interaction</b>	1.095	3	0.22	0.8796
	Viewing Comfort					Level of Realism			
	SS	DoF	F	P		SS	DoF	F	P
<b>Mono-Stereo</b>	20.1353	3	8.29	0.0001	<b>Mono-Stereo</b>	29.327	3	10.79	0
<b>Laser-Video</b>	0.997	1	1.23	0.2706	<b>Laser-video</b>	0.48	1	0.53	0.4687
<b>Interaction</b>	0.5861	3	0.24	0.8673	<b>Interaction</b>	0.349	3	0.11646	0.9429
	Sense of Presence								
	SS	DoF	F	P		SS	DoF	F	P
<b>Mono-Stereo</b>	100.986	3	14.26	0					
<b>Laser-Video</b>	0.48	1	0.2	0.6532					
<b>Interaction</b>	1.122	3	0.16	0.9239					

Table 4: The results of 2-way ANOVA for qualitative measurements. Rows show values for the independent variables (Mono-Stereo, Laser-Video), their interaction, and error. Columns show the sum of squares (SS), the degrees of freedom (df), the F statistic and the p-value.

### 1) **Collision**

Under stereoscopic visualization users perform significantly better in terms of Collision Rate both on case of laser and visual sensor. The ANOVAs show similar values for “F” and “p”. The mean values show same trends on both facilities with users performing better in stereo-viewing conditions. It is therefore very clear that stereo viewing plays a more dominant role than the different image type and system behaviours.

The differences between mean values on the Collision Number are relatively small. However, if we consider that users tele-driving on laser images employ less time to complete a trial, we can conclude that the real-time response copes for the lack of image quality because we keep an approximately same number of collisions.

For what concern differences among the visualization facilities, the Laptop performs significantly better on Collision Rate both for video and laser images, (the ANOVA “p” value is lower in case of video-images). As for the Collision Number the improvement in Laptop performance has a tendency to significant in case of video-images ( $F=3.32$  and  $p=0.0757$ ), and there is not significant difference in case of laser-images.

### 2) **Obstacle Distance**

The Obstacle Distance is the quantitative measurement that shows the largest result discrepancy (between laser and video-images trials). Users perform significantly better on stereo-viewing conditions but only in case of laser-images. Looking at the visualization facility, we find that users perform significantly better on the Laptop, but only on video-images.

When considering all laser and video -based trials we note that users perform significantly better on video-images (keeping robot farer from obstacles). The ANOVA has  $F=4.9$  and  $p=0.0296$ .

### 3) **Completion Time**

The users drive slower in mean value under stereo visualization conditions both in case of laser and video images. The performance on Laptop is significantly slower only in case of laser-images.

An interesting outcome is to observe that users always employ less time to complete a trial in case of laser-images. This seems to be the immediate consequence of having real-time feedback. Most interestingly, we can observe that in case of laser-images the number of collisions is comparable to those detected when using video-images. We can conclude that despite a lower image-quality and the more approximated



environment representation, the real-time performance provided by a laser-based teleguide allows for faster completion time of the assigned task while keeping the same driving-accuracy as with video-images.

#### **4) Path Length**

There is no significant difference or relevant trend in Path Length on any of the proposed research questions. It can only be observed that the longer paths in mean value are those related to users operating on the Laptop under mono-viewing condition.

#### **5) Mean Speed**

The improvement in Mean Speed under monoscopic viewing conditions has a tendency to significant in case of video-images while there is not significant difference in case of laser-images. The slower speed under stereo condition is the consequence of a higher Completion Time.

#### **6) Depth Impression**

Most of the users had no doubts that Depth Impression is higher under stereo visualization conditions both in case of laser and video images. Stereoscopic viewing performs significantly better on both types of images. If we consider the results on stereo-viewing facilities only (both for laser and video images), users performs significantly better on the Wall facility. The ANOVA has  $F=11.99$  and  $p=0.0013$ .

#### **7) Suitability to Application**

The improvement of the Suitability to Application parameter in case of stereo viewing shows a tendency to significant only in case of laser-image. The ANOVA has  $F=3.33$  and  $p=0.0748$ . Nevertheless, if we consider results for both laser and video images the improvement of stereo viewing becomes statistical significant. The ANOVA has  $F=5.68$  and  $p=0.0014$ .

If we consider the results on stereo-viewing facilities only (both for laser and video images), users performs significantly better on the Wall facility. The ANOVA has  $F=12.61$  and  $p=0.001$ . This result is mostly due to the very low performance of Anaglyph stereo for video-images. Therefore we can conclude that the Anaglyph stereo on Laptop is better tolerated on laser-images than video-images.

## 8) Viewing Comfort

The improvement of stereo visualization in Viewing Comfort when considering both laser and video images is statistically significant. The ANOVA has  $F=8.29$  and  $p=0.0001$ . Both on laser and video images, stereo and mono visualization show opposite trends in mean values for the two facilities.

If we consider the results on stereo-viewing facilities only (both for laser and video images), users performs significantly better on the Wall facility. The ANOVA has  $F=19.11$  and  $p=0.0001$

## 9) Level of Realism

Stereoscopic viewing performs significantly better with both laser and video images. As expected the best result is for video images. The improvement of stereo visualization in Level of Realism when considering both laser and video images is statistically significant. The ANOVA has  $F=10.79$  and  $p=0$ .

If we consider the results on stereo-viewing facilities only (both for laser and video images), users performs significantly better on the Wall facility. The ANOVA has  $F=11.25$  and  $p=0.0018$ .

## 10) Sense of Presence

Stereoscopic viewing performs significantly better with both laser and video images. The best result is for video images. The improvement of stereo visualization in Sense of Presence when considering both laser and video images is statistically significant. The ANOVA has  $F=14.29$  and  $p=0$

If we consider the results on stereo-viewing facilities only (both for laser and video images), users performs significantly better on the Wall facility. The ANOVA has  $F=15.82$  and  $p=0.0003$ .

## ***VI. CONCLUSION***

This work investigated the role of 3D stereoscopic visualization in laser-based mobile robot teleguide. Two different visualization systems were considered. A main aim was to experimentally demonstrate the performance enhancement in mobile robot teleoperation when using laser-based stereoscopic visualization. Furthermore, the advantage of binocular stereo viewing was challenged by a visual representation rich of strong monocular depth cues.

A usability evaluation was proposed to assess system performance. The evaluation involved several users and two different working sites located approximately 3,000 km apart.

The use of laser sensor was proposed as alternative to the use of visual sensor previously experimented. A main aim was therefore also to compare performance of mobile robot teleguide based on laser sensor against that based on visual sensor evaluated in previous experiments.

The results were evaluated according to the proposed research questions. This involved three factors: monoscopic versus stereoscopic visualization, laptop system versus wall system, and laser-based images versus video images. The three factors were evaluated against different quantitative variables (collision rate, collision number, obstacle distance, completion time, path length, mean speed) and qualitative variables (depth impression, suitability to application, viewing comfort, level of realism, sense of presence).

The result of the evaluation on the stereo-mono factor indicated that 3D visual feedback leads to fewer collisions and a safer driving than 2D feedback therefore is recommended for future applications. The number of collisions per time unit was significantly smaller when driving in stereo and the mean of minimum distance to obstacles was significantly higher when driving in stereo. A statistically significant improvement of performance of 3D visual feedback was also detected for the variables depth impression and sense of presence, (while it was detected a tendency to significant for the suitability to application and level of realism variables). The other variable did not lead to significant results on this factor.

The results of the evaluation on the laptop-wall factor indicated significantly better performance on the laptop in terms of collision rate and on the wall in terms of completion time. No statistically significant results were obtained for the other variables.

The results of the comparative evaluation which included also the results of the previous experiments based on visual sensor, indicated significantly better performance on the obstacle distance variable (laser-video factor) and on all qualitative variables (mono-stereo factor).

The Interaction between the factors was never statistically significant.

We observed that in laser-based teleguide the real-time response copes for the lack of image quality. We also showed that users always employed less time to complete a trial while making approximately the same number of collisions

Further studies are under development with the aim of combining laser and video technology and augmented reality visualization to assist mobile robot teleguide. Further visualization systems are also being considered.

We expect that 3D visualization will soon become very popular in telerobotic application and it will spread on different application contexts as well, e.g. interactive television, cinema, and computer games.

## ***Acknowledgments***

Dr. Sessa research is supported by the Japan Society for the Promotion of Science (JSPC) postdoctoral fellowship for Foreigner Researchers FY2008.

## ***REFERENCES***

- [1] M. Bocker, D. Runde, L. Muhlback, "On the Reproduction of Motion Parallax in Videocommunications". In 39th Human Factors Society, 1995.
- [2] C. Demiralp, C. Jackson, D. Karelitz, S. Zhang, D. Laidlaw. "Cave and Fishtank Virtual-Reality Displays: A Qualitative and Quantitative Comparison". IEEE Transactions on Visualization and Computer Graphics, Vol.12, issue 3, 2006.
- [3] D. Drascic, "Skill Acquisition and Task Performance in Teleoperation using Monoscopic and Stereoscopic Video Remote Viewing". In proc. 35th Human Factors Society, 1991.
- [4] M. Ferre, R. Aracil and M. Navas, "Stereoscopic Video Images for Telerobotic Applications", Journal of Robotic Systems, vol.22, issue 3, pp. 131-146, 2005.
- [5] G. Hubona, G. Shirah, D. Fout, "The Effects of Motion and Stereopsis on Three-Dimensional Visualization". Int. Journal of Human-Computer Studies, Vol. 47, 1997.
- [6] G. Jones, D. Lee, N. Holliman, D. Ezra, "Perceived Depth in Stereoscopic Images". In proc. 44th Human Factors Society, 2000.
- [7] L.J. Corde, C.R. Caringnan, B.R. Sullivan, D.L. Akin, T. Hunt, R. Cohen, "Effects of Time Delay on Telerobotic Control of Neural Buoyancy", IEEE Proceedings of Int. Conference on Robotics and Automation (ICRA), pp 2874-2879, Washington, USA, 2002
- [8] S. Livatino, C. Koeffel. "Handbook for Evaluation Studies in Virtual Reality". IEEE Int. Conf. in Virtual Environments, Human-Computer Interface and Measurement Systems (VECIMS), Ostuni, Italy, 2007
- [9] U. Naepflin, M. Menozzi, "Can Movement Parallax Compensate Lacking Stereopsis in Spatial Explorative Tasks?". Elsevier DISPLAYS, 2006.
- [10] I. Sexton, P. Surman, "Stereoscopic and Autostereoscopic Display Systems". IEEE Signal Processing Magazine, 1999.

- [11] S. Livatino, G. Muscato, C. Koeffel, S. Sessa, C. Arena, A. Pennisi, D. Di Mauro, E. Malkondu. "Mobile Robotic Teleguide Based on Video Images". IEEE Robotics and Automation Magazine. Vol. 14. No. 4. 2008
- [12] M. Ferre, R. Aracil, M. Sanchez-Uran. "Stereoscopic Human Interfaces". IEEE Robotics and Automation Magazine. Vol. 14. No. 4. 2008