

Remote Supervision of an Autonomous Surface Vehicle using Virtual Reality

Mårten Lager^{*,**} Elin A. Topp^{*}

^{*} Department of Computer Science, Lund University, Sweden
(e-mail: {marten.lager, elin.anna.topp}@cs.lth.se)

^{**} Saab Kockums AB, Malmö, Sweden

Abstract: We compared three different Graphical User Interfaces (GUI) that we have designed and implemented to enable human supervision of an Autonomous Surface Vehicle (ASV). Special attention has been paid to provide tools for a safe navigation and giving the user a good overall understanding of the surrounding world while keeping the cognitive load at a low level. Our findings indicate that a GUI in 3D, presented either on a screen or in a Virtual Reality (VR) setting provides several benefits compared to a *Baseline GUI* representing traditional tools.

© 2019, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Teleoperation, Virtual reality, User interfaces, Remote control, Autonomous vehicles, Navigation, Ship control, USV, ASV

1. INTRODUCTION

As the autonomous vehicles are becoming more intelligent, the human's role of being in constant control can be relaxed. In many cases, the semi-autonomous vehicle can plan and execute missions that meet the human needs, while the human can still take control by teleoperating the vehicle, which is useful, e.g., when some situation occurs that the vehicle has not yet been trained for. Many car manufacturers will equip their cars with teleoperation capability, and Levander (2017) sees teleoperation of ships as a key technology in the transferring process towards autonomous ships.

In this study, we are focusing on a small Autonomous Surface Vehicle (ASV) that is being remotely supervised by a human user via a low bandwidth connection. Murphy (2014) believes small ASVs like this are likely to play an important role during future Search and Rescue (SAR) operations at sea. The reason for adding the low bandwidth constraint that prohibits video streams and high-resolution images to be sent, is that we want the Graphical User Interface (GUI) to work far away from land. In these areas, ASVs need to rely on radio communication normally used by ships, as mobile communication has too insufficient coverage, and satellite communication is too costly and has too bulky antennas.

There are many benefits of using an ASV instead of a normal ship. An ASV can be constructed lighter and cheaper. An ASV can also be used when it would be dangerous for a human to operate, e.g., during bad weather at sea. Multiple ASV fleets can be placed in various locations, and during accidents, the closest fleet can be dispatched from a centralized location with teleoperating experts. ASVs and drones can typically be dispatched far more quickly than manned vehicles, as there is no need for waiting on the human crew. Compared to flying drones, ASVs have good endurance and can be on a mission for

many hours, while drones are typically faster and can get a good overview of an area from their high altitude.

Although cars and airplanes are hard to teleoperate due to the constraints that the dynamic traffic situations set on time delays and jitter (see d'Orey et al. (2016) and Neumeier et al. (2018)), traffic at sea is often characterized by more available time for decision making, making it ideal for teleoperation.

During complex situations at sea on manned ships, Porathe (2006) shows that it can be hard even for humans to interpret and match the surrounding environment with the information from the navigation equipment onboard. There are several occasions where the navigators, due to the high cognitive load, have mixed up sea marks, directions or landmarks, which in many cases have led to fatal accidents. This raises the question of how new types of GUIs can support remote supervision of an ASV with limited bandwidth. Of particular interest to us is to see how the user's *situational awareness* and *cognitive load* are affected when using such GUIs in comparison to using traditional ones. We use the term *situational awareness* to describe the ability of the user to assess the situation the vessel is in with respect to surrounding elements like other vessels, seamarks, or shallow areas.

To answer these questions, we propose a 3D-visualisation of the ship's surroundings either on a computer screen or in a Virtual Reality (VR) setup. We base this GUI on ideas from the available research regarding manned ships to increase the situational awareness while maintaining a low cognitive load.

We evaluated the two different versions of the GUI against a baseline GUI, see Figure 1, in a small study with 16 participants, showing that there are significant benefits regarding the mentioned factors.

This paper is organized as follows: In Section 2, an overview of related research is given followed by the GUI

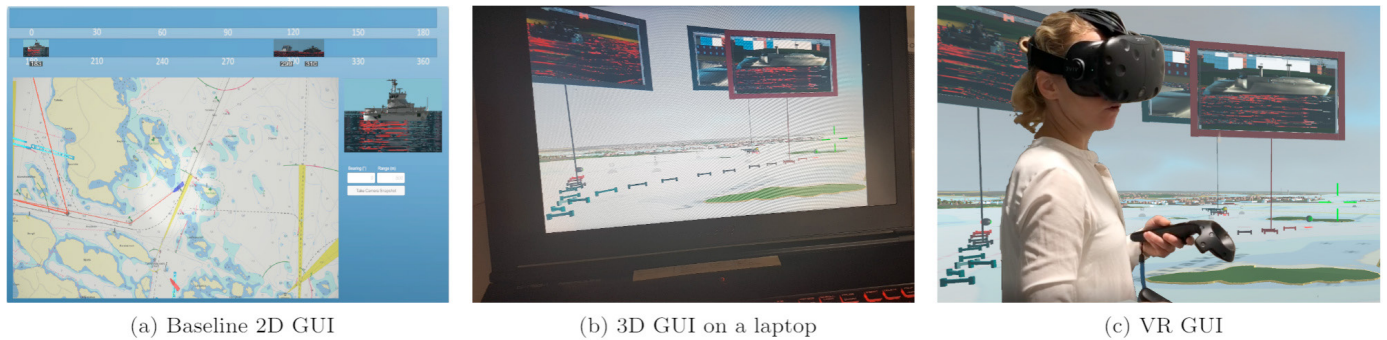


Fig. 1. Three types of GUIs have been developed. (a) is a 2D GUI, that represents a traditional GUI. (b) and (c) are created in 3D, where (b) is presented on a laptop and (c) in VR.

design in Section 3. Then the User Study is presented in Section 4, with related results and discussion in Section 5 and 6. Conclusion is given in Section 7.

2. RELATED WORK

Teleoperation can be used for remote controlling vehicles and robots, e.g., cars (see Neumeier et al. (2018)), drones (see Hedayati et al. (2018)) or ships (see Nava-Balanzar et al. (2017)). While teleoperating a vehicle, it is important to support the human's perception of what the vehicle's sensors detect of the surrounding world. Williams et al. (2018) elaborate on how VR, Augmented Reality (AR) and Mixed Reality can strengthen visualization, and thereby the total communication between the machine and the human, and how various viewpoints, e.g., the ego-centric view can be used. Hedayati et al. (2018) explore how AR can be used for augmenting a drone's field of view for a user collocated with the drone. When not collocated, VR is often a better presentation technique compared to AR. Hosseini and Lienkamp (2016) and Shen et al. (2016) show how VR can enhance the situation awareness when driving a remote car.

When navigating onboard a ship, the traditional way is to use either a paper sea chart or an electronic chart system, showing an abstracted map of, e.g., islands, groundings, depth measurements, and sea marks. Research has shown that it is difficult for humans to match what they see on the chart or Radar to what they see in the real world outside the bridge of the ship. Instead, it is a better approach to visualize a 3D map oriented in a way that matches the user's view of the surrounding world (see Porathe (2007) and Witt (2017)). Our GUI is influenced by these results, as we place the user directly into the 3D world where the surrounding world easily can be matched with the sea chart. We call this ego-centric view *First Person View (FPV)*.

Some research has also investigated how AR can reduce cognitive load when navigating. In these applications, Head Mounted Displays (HMD), normally with see-through technology, augment important information, such as sea lanes and other nearby ships directly in the real-world environment (see Morgère et al. (2014); Mollandsøy and Pedersen (2017); Hugues et al. (2010) and Jaeyong et al. (2016)). In our application, we augment the important information directly in the virtual environment.

3. DESIGN

In our study, we focus on a GUI for remote supervision of a small ASV via a limited bandwidth connection which inhibits video or high-resolution images to be transferred. The ASV is expected to be highly autonomous in order to handle a SAR mission, but still assumed to need some human supervision with the ability to take measure if something unexpected happens.

The GUI is developed for a small ASV with a computer capacity and sensor suite comparable to an autonomous car. The postulated sensors and capabilities are:

- Global Positioning System (GPS), or a satellite independent positioning system (see Lager et al. (2017));
- Application for autonomous route steering;
- Camera with 360-degree coverage and zoom;
- Radio communication with a small antenna with a bandwidth of around 10 kbit/s;
- Application for image detection of ships; and
- Application for cropping and compressing images of ships, so that detected ship images can be transferred through the radio communication interface.

3.1 Architectural Overview

Figure 2 shows a summary of the communication interfaces to the GUI. To create the virtual surroundings and corresponding sea chart, the ASV transmits its position. When the camera onboard the ASV detects ships, it transmits all tracks of them every second, as well as a small cropped image every 10 seconds. Some, mainly larger, ships have an Automatic Identification System (AIS) transponder, that transmits, e.g., their identity, position and steering direction. This is also received by the GUI and presented to the user along with the tracks from the camera.

3.2 GUI Design

We have created two GUI versions in 3D to enhance situational awareness while maintaining a low cognitive load. In 3D, a virtual world is perceived in the same way as humans normally perceive the real world. By combining a virtual world with the navigation GUI, our intention has been to give the user a better understanding of the environment, instead of just letting the user look at an electronic chart system. In the virtual environment, objects such as sea marks, surrounding ships, sea lanes,

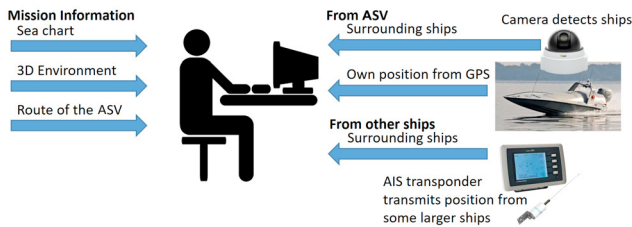


Fig. 2. The GUI uses the position from the ASV's GPS to position the surrounding 3D world and the sea charts. The camera and AIS provide tracks to the GUI.

and routes are augmented for the user. To enhance the immersive experience even further, one of those GUIs are using VR. VR is a technique that presents a fully computer-generated simulated environment for the user, and the user is thereby fully left out from the visual physical world. In Figure 1c, a user supervises the ASV during the user study by using a VR HMD called *HTC Vive*.

The design of the different parts of the GUI has been developed in an iterative approach. First, a prototype was created based on research by e.g. Porathe (2007) about how the cognitive load of the user can be reduced for navigators. After a brief evaluation by navigation experts, an improved version of the application was developed, which has now been tested by a larger user group.

The GUI is implemented in Unity 3D (see Uni (2018)), which is normally used for creating 2D and 3D games. We have used a simulation kernel, received from the shipyard Saab Kockums AB, that simulates own and other ships in a predefined mission. It also creates a 3D replica of the real world, which has been used as a foundation for the implementation of the GUI.

Three different GUIs have been developed; one *Baseline GUI*, representing traditional navigation tools, one 3D tool presented on a laptop, called *3D GUI*, and one 3D tool presented in VR, called *VR GUI*. All these GUI types are presented in Figure 1.

3.3 Baseline - Traditional GUI

Navigators onboard manned ships use an electronic sea chart with north facing upwards. The own ship, positioned by the GPS, as well as other tracks of ships received by the AIS, are visualized directly on the chart. The *Baseline GUI* is created to mimic this design. A navigator normally needs to keep track of how the surrounding real world matches the sea chart, causing a large cognitive load. Because the ship in our application is controlled remotely, this matching is not needed, making it easier for the *Baseline GUI* users.

Optronic systems at sea often present a 360 view, split along two or four stripes with 180 degrees or 90 degrees each (see Maltese et al. (2010)). As a compliment, normally an enlargement of one camera view can be seen. These features are also available in *Baseline GUI*. Because the received images have low resolution, we have been able to fit everything including the sea chart on one screen without compromising too much with the size of the small images. When the camera on the ASV detects a ship, the large

image in Figure 1a is shown at the same time as the image is placed in the correct location on the lower 180-degree stripe in the top of the GUI. At the same time, it indicates directly with a marker in the chart which area that has been photographed.

The users can manually take photos as well, by entering a bearing and a range. The camera onboard the ASV then takes a photo and transmits the compressed image when there is available bandwidth.

3.4 3D GUI and VR GUI

Our assumption is that we can create an easy-to-use GUI which provides a good situational awareness by:

- Creating the GUI in 3D, and present it in VR.
- Providing different views of the surrounding environment, optimized for various situations.
- Augment objects directly in the 3D world.

Our hypotheses are that a user operating a GUI built by these foundations will have a better overall understanding of the situation, and will observe potential dangerous situations earlier.

The GUIs seen in Figure 1b and 1c, named *3D GUI* and *VR GUI*, share most of the design and have three different views; *FPV*, *Tethered View* (TeV) and *Exo-Centric View* (ECV), see Figure 3. Each view has its own benefits so that the views complement each other.

First-Person View The *FPV*, see Figure 3a, visualizes the world from the ASV's (ego) perspective, hence it simulates what the surrounding environment looks like. A camera has good bearing accuracy but poor range accuracy. From the *FPV*, the range is of lower importance, hence information from passive sensors such as cameras is well visualized in this view. By capturing real-world images of landmarks and comparing the bearings to the sea chart, it will be obvious if the current GPS position diverges from the correct position, as the landmark bearings will not match the chart.

Tethered View The *TeV* is created by a virtual camera hanging above the ASV, viewing the ASV from above, providing a bird's eye view of the ASV in its environment, see Figure 3b.

Exo-Centric View The *ECV*, see Figure 3c, is used for presenting a north-up facing sea chart for the user where the own ship is located in the middle, much resembling the *Baseline GUI*. The *ECV* has been implemented as a room with a large sea chart in front of the user. In the *3D GUI*, the user can zoom in and out, and in the *VR GUI*, the user can walk around freely in the room and look at other parts of the sea chart. The strength of this view is that the user can get an overview of the situation and plan a long way ahead. It is also easy to get an understanding of if the own route is well positioned according to the sea chart so that it does not pass any groundings.

3.5 The differences between the 3D GUI and the VR GUI

The main difference between the *3D GUI* and the *VR GUI* is how to interact with the GUI, see Figure 4. In the

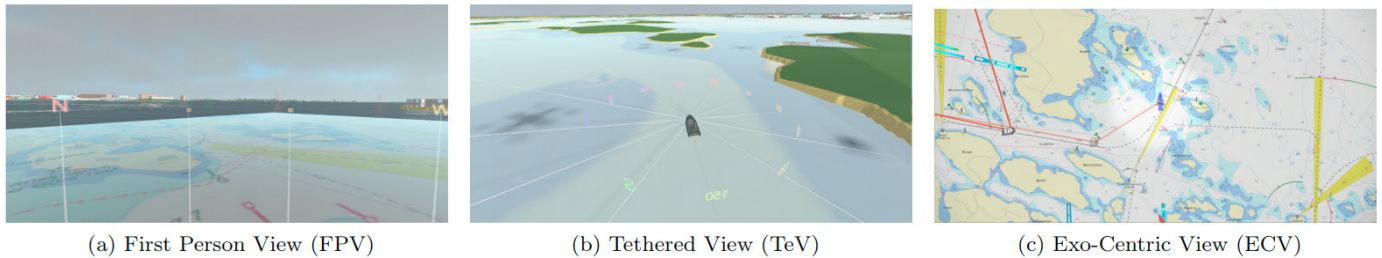


Fig. 3. a) *FPV*: In the *FPV* the virtual user is placed onboard the ASV. The center of the sea chart is positioned in the user's location and is oriented to be consistent with the surrounding world. b) *TeV*: The own ASV is situated in front of the virtual user, as the user was paragliding after the boat. The dark blue in the sea chart indicates a dangerous area with a bottom depth less than 3m. In *TeV*, the user can easily get a feel for the own ASV's size compared to the passage between the shallow areas. c) *ECV*: This view provides a large sea chart for the user, centered in the ASV's position.

3D GUI, the switching between the various views and the zooming is done with keyboard buttons, and photos are taken by pointing and clicking with the mouse. In the *VR GUI* on the other hand, the associated HTC Vive hand controllers are used for the same interaction.

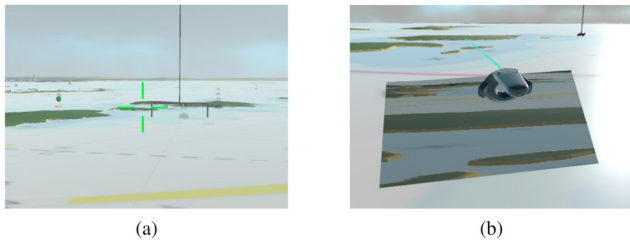


Fig. 4. a) In the *3D GUI*, photos are taken with a mouse click while pointing the green hair cross. b) In the *VR GUI*, photos are taken by pointing the green ray from the hand controller and then pressing a button. The image also shows a zoom-window below the hand controller, which enlarges what the hand controller is pointing towards.

3.6 Features to reduce the cognitive load

Several features have been implemented to support navigation and situational awareness, while still limiting the cognitive load:

- A correctly oriented sea chart surrounds the user in *FPV*, see Figure 3a.
- A sea chart is presented instead of the water in *TeV*.
- A rainbow-colored compass is shown in *FPV* and *TeV*, to guide the user with directions.
- Sea marks are augmented in *FPV* and *TeV*.
- Indications of surrounding ships are visualized in all views.
- Routes and waypoints are visualized in all views.

4. USER STUDY

We evaluated our implementations with 16 participants, recruited mainly from Lund University and the shipyard Saab Kockums AB, in 50-minute long trial sessions (recorded on video) with the task and scenarios described below. We had the ethics board of our university check the study setup and were informed, that no supervision by

the board or formal approval was needed to conduct the study. The participants were informed of the possibility to withdraw at any time and agreed upon the use of the recordings and other data for research purposes.

The user should, after an introduction phase based on written instructions and a slide show, use two different GUIs, either the *Baseline GUI* and the *3D GUI* (on screen), or the *Baseline GUI* and the *VR GUI* to supervise a ship that was passing through an archipelago on a predefined route in two different scenarios. The task was to observe potential dangerous situations and report them as soon as they were detected, and to keep track of the closest nearby ships. Also, they were asked to take pictures of surrounding islands when possible, assuming that we could measure their cognitive load by getting an idea of how much spare time (and mental capacity) they had to handle this secondary task. The participants were told that the safety tasks were most important, and the photo task was least important. A two minutes introduction to each GUI was given before each test.

From analyzing the video recordings of the user study, four final score values were given for each run (objective results) for *Collision Observations*, *Grounding Observations*, *Situational Awareness* (closest ships correctly identified), and *Photos Taken* (Cognitive Load). The scores were computed as percentages of the respective possible values. After the experiments, the GUIs were evaluated subjectively by the participants by answering the following three questions for each GUI on a scale of 1-10 (10 was best):

- (1) Do you feel that you had a good overall picture of the situation? (Situational awareness rating)
- (2) Do you think that the GUI tool was easy to use?
- (3) If you had practiced 100 hours on this GUI, do you think it would be best for the task?

5. RESULTS

The collected data from the user experiments have been summarized in the objective and subjective results below. The interpretation of the results is done in Section 6.

5.1 Objective Results

For the objective results from the user experiments, we found that both the *3D GUI* and the *VR GUI* gave

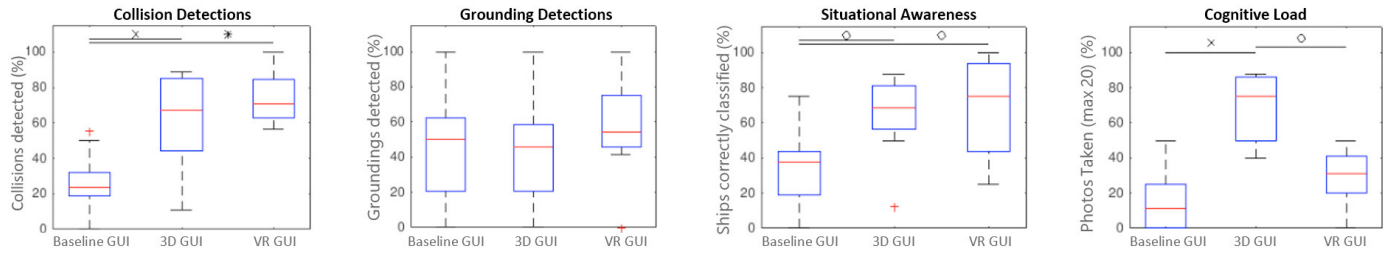


Fig. 5. Objective results show that the situational awareness as well as the ability to detect collisions have been improved in both the 3D GUI and the VR GUI. The users of the 3D GUI have also managed to take more photos, which could indicate a lower cognitive load compared to the Baseline GUI and the VR GUI. (o), (x) and (*) denote comparisons with $p < 0.05$, $p < 0.01$ and $p < 0.001$ respectively.

significantly better results regarding the collision detection and situational awareness than the Baseline GUI, while there was no major difference regarding the detection of groundings. The 3D GUI was significantly better than both the Baseline GUI and the VR GUI regarding the possibility to take photos, where the results for the VR GUI are somewhat inconclusive. Figure 5 summarizes the objective results, along with the p-values showing if there was a significant difference, computed in a series of one-tailed t-tests. The mean values are presented in Table 1. Even though it was a quite small user study, power tests ($\alpha=0.05$, $\text{power}>0.80$) have shown that there were enough participants to support the significant results.

Table 1. Objective Results Summary

	Baseline	3D GUI	VR GUI
Collision Detections	26.8%	61.5%	74.0%
Grounding Detections	42.2%	43.8%	56.3%
Situational Awareness	33.6%	64.1%	68.8%
Photos Taken	15.9%	82.8%	29.4%

5.2 Subjective Results

For the subjective results from the user evaluation, we found that the users experienced a significant benefit of the VR GUI compared to the Baseline GUI, regarding having a good Situational Awareness. The users also experienced that both the VR GUI and the 3D GUI were more Easy to Use. The users expected the VR GUI to be the best and the 3D GUI to be the next best tool, for an expert user with many hours of training.

Figure 6 summarizes the subjective results, along with the p-values, computed in a series of one-tailed t-tests. The mean values are presented in Table 2. The power tests have shown that there were enough users to support the significant results with a p-value less than 0.01, but not the two results with a p-value between 0.01 and 0.05.

Table 2. Subjective Results Summary

	Baseline	3D GUI	VR GUI
(1) Situational Awareness	5.8	7.5	8.0
(2) Ease of use	5.3	7.3	7.9
(3) Best tool	4.3	8.4	9.6

6. DISCUSSION

Our results show, that the users had better situational awareness when using 3D or VR, which can be seen in

Collision Detection and (objective) Situational Awareness. They also said they experienced this in the question regarding the Situational Awareness and Ease of Use in the user evaluation. The VR GUI got a higher score than the 3D GUI on all these four metrics, and the 3D GUI, in turn, got a higher score on all these than the Baseline GUI, which is also reflected in the user ratings of the best tool for the task (given more training hours than they had experienced themselves). We think the good score for the VR GUI (and 3D GUI) has to do with the fact that the perception in the GUI much resembles how a human normally perceives the world.

Regarding Grounding Detections and Photos taken (our assumed indicator for the cognitive load), our results are somewhat inconclusive, which we attribute to the relative unfamiliarity with the GUI as the users simply did not manage to switch into the optimal view (ECV) for this task reliably enough. We also observed a “fun factor” of taking photos of ships instead of islands in the VR setting, significantly reducing the scores for the VR GUI users.

Of the sixteen users in the user study, six persons considered themselves to be computer gamers (three used 3D GUI, and three used VR GUI). These persons scored better in the experiments in general, indicating that the performance of the GUIs is increased with training. Table 3 shows a summary for each of the GUI of how much better scores the gamers had compared to the persons that were not gamers. As can be seen, the gamers performed particularly well in the VR GUI. This finding goes in line with the user evaluation where the users suggested that the VR GUI and 3D GUI would be a better expert tool after a long training.

Table 3. How much better gamers performed compared to users that were not gamers.

	Baseline	3D GUI	VR GUI
Collision Detection	14.8%	6.9%	38.4%
Grounding Detection	88.6%	20.0%	93.9%
Situational Awareness	-19.5%	40.7%	52.2%

7. CONCLUSION

We investigated, how 3D and VR approaches could support the remote operation of an ASV with a low bandwidth connection, by comparing respective GUIs with a Baseline GUI following the currently applied interfaces in such contexts. Our findings show, that both the 3D and VR approaches outperform the traditional approach

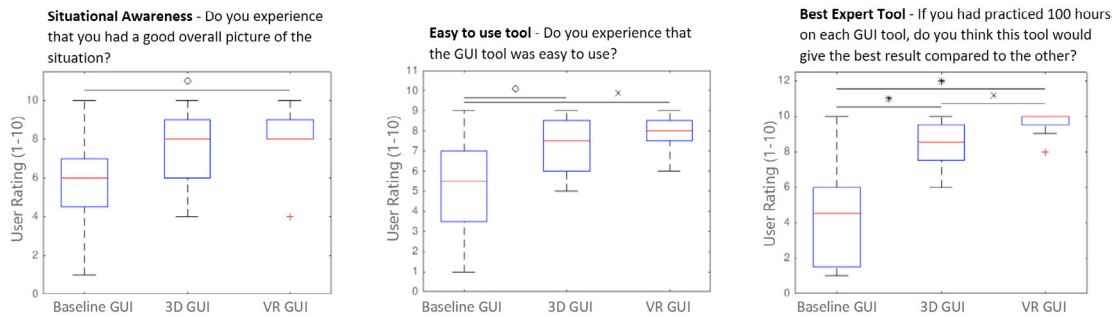


Fig. 6. The evaluation score for *Situation Awareness*, *Easy to Use* and *Expert Tool* have been improved for the users of the *3D GUI*, and improved even more for the users of the *VR GUI*. (o), (x) and (*) denote comparisons with $p < 0.05$, $p < 0.01$ and $p < 0.001$ respectively.

significantly. We found the *3D GUI* and *VR GUI* users to be better to react to potentially dangerous situations compared to the *Baseline GUI* users, and they could keep track of the surroundings more accurately. They also reported that they expected the *3D GUI*, and especially the *VR GUI*, to be the best tool of the three choices for an expert user with many hours of training.

As our investigations so far only have covered the supervision of a simulated ASV, we see the next step to be the integration of functionalities to control an actual ASV. This might potentially also mean looking at the integration of further information sources into the interfaces, which might entail new aspects for the user(s) to handle when interacting with the ASV.

ACKNOWLEDGEMENTS

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by Knut and Alice Wallenberg Foundation. The simulation tool was provided by Saab Kockums AB.

REFERENCES

- (2018). Unity 3d. URL <https://unity3d.com/>. Last accessed 29 January 2018.
- d'Orey, P.M., Hosseini, A., Azevedo, J., Diermeyer, F., Ferreira, M., and Lienkamp, M. (2016). Hail-a-drone: Enabling teleoperated taxi fleets. In *Intelligent Vehicles Symposium (IV), 2016 IEEE*, 774–781. IEEE.
- Hedayati, H., Walker, M., and Szafir, D. (2018). Improving collocated robot teleoperation with augmented reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 78–86. ACM.
- Hosseini, A. and Lienkamp, M. (2016). Enhancing telepresence during the teleoperation of road vehicles using hmd-based mixed reality. In *Intelligent Vehicles Symposium (IV), 2016 IEEE*, 1366–1373. IEEE.
- Hugues, O., Cieutat, J.M., and Guitton, P. (2010). An experimental augmented reality platform for assisted maritime navigation. In *Proceedings of the 1st Augmented Human International Conference*, 12. ACM.
- Jaeyong, O., Park, S., and Kwon, O.S. (2016). Advanced navigation aids system based on augmented reality. *International Journal of e-Navigation and Maritime Economy*, 5, 21–31.
- Lager, M., Topp, E.A., and Malec, J. (2017). Underwater terrain navigation using standard sea charts and magnetic field maps. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2017)*. IEEE. Doi: 10.1109/MFI.2017.8170410.
- Levander, O. (2017). Autonomous ships on the high seas. *IEEE Spectrum*, 54(2), 26–31.
- Maltese, D., Deyla, O., Vernet, G., Preux, C., Hilt, G., and Nougues, P.O. (2010). New generation of naval IRST: Example of EOMS NG. In *Infrared Technology and Applications XXXVI*, volume 7660, 766004. International Society for Optics and Photonics.
- Mollandsøy, M.K. and Pedersen, P.H. (2017). *Augmented Reality og Head Mounted Display-navigatørens verktøy i en teknologisk fremtid*. B.S. thesis, Sjøkrigsskolen.
- Morgère, J.C., Diguët, J.P., and Laurent, J. (2014). Electronic navigational chart generator for a marine mobile augmented reality system. In *Oceans-St. John's, 2014*, 1–9. IEEE.
- Murphy, R. (2014). Disaster Robotics. Intelligent robotics and autonomous agents series. *The MIT Press*.
- Nava-Balanzar, L., Sanchez-Gaytán, J., Fonseca-Navarro, F., Salgado-Jimenez, T., Gómez-Espinosa, A., and Ramirez-Martinez, A. (2017). Towards teleoperation and automatic control features of an unmanned surface vessel-rov system: Preliminary results. In *International Conference on Informatics in Control, Automation and Robotics*.
- Neumeier, S., Gay, N., Dannheim, C., and Facchi, C. (2018). On the way to autonomous vehicles teleoperated driving. In *AmE 2018-Automotive meets Electronics; 9th GMM-Symposium*, 1–6. VDE.
- Porathe, T. (2006). *3-D nautical charts and safe navigation*. Ph.D. thesis, Institutionen för Innovation, Design och Produktutveckling, Mälardalen University.
- Porathe, T. (2007). User-centered map design. In *Usability Professionals Association Conference*.
- Shen, X., Chong, Z.J., Pendleton, S., Fu, G.M.J., Qin, B., Frazzoli, E., and Ang, M.H. (2016). Teleoperation of on-road vehicles via immersive telepresence using off-the-shelf components. In *Intelligent Autonomous Systems 13*, 1419–1433. Springer.
- Williams, T., Tran, N., Rands, J., and Dantam, N.T. (2018). Augmented, mixed, and virtual reality enabling of robot deixis. In *the International Conference on Virtual, Augmented and Mixed Reality*.
- Witt, S. (2017). Technologies used for navigation. *Novel Interaction Techniques for Oceangoings*, University of Passau.