Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

doi: 10.3850/978-981-18-8071-1_P200-cd

Proposed method for analysis of eye tracking data from unmanned ship operation.

Jon Bernhard Høstmark[1], Gudbrand Strømmen[1], Alf Ove Braseth[2] and Magnhild Kaarstad[2]

[1]*Kongsberg Group, Kongsberg Maritime, Norway.*
Jon.Bernhard.Hostmark@km.kongsberg.com
Gudbrad.strommen@km.kongsberg.com

[2]*Institute for Energy Technology, Norway.*
alf.ove.braseth@ife.no
magnhild.kaarstad@ife.no

Within the maritime domain, there is a focus on applying new technologies to reduce cost, and more recently focusing on environmental sustainability. The use of highly automated and unmanned ships is one such approach. The maritime safety level should not be reduced with the introduction of unmanned remote ship operation and novel technology in the maritime domain. This requires increased knowledge and understanding of how humans perceive information presented through displays in a fully digitalized work environment. Based on this, the paper suggests a method using eye-tracking data to objectively collect and analyze how operators perceive information on displays. The approach is assessed by running scenarios in a simulated environment with navigators and vessel traffic operators. Both information content and the arrangement of information are explored through the approach. The collected eye-tracking data is analyzed and visualized through software and validated against data from participant interviews. This poses a non-intrusive method allowing in-depth post analysis of individual events in a test scenario, without the need to stop and perform e.g., an interview. By using the method, quantitative objective results are obtained, which is valuable for backing up qualitative interview data. The results suggest that the proposed method is promising by enabling quantitative evaluation of visual information accessed by the test participants. Further work should pay attention toward analysis of how individual visual search patterns differ between participants for different test cases. From a safety perspective deeper understanding of multi-asset controls impact on salience and potential tunnel vision is needed.

Keywords: *Keywords*: Maritime, Safety, Autonomy, Operation Centre, Human Factors, Situation Awareness, Eye-tracking.

1. Introduction

In recent years there has been increased focus on the application of autonomy to address sustainability and to reduce the environmental impact in the maritime sector. In addition, there is a continuous drive to reduce the cost for maritime transport. To address these challenges, several companies such as Massterly (Massterly, 2023) and Ocean Infinity vision unmanned vessels (Ocean Infinity, 2023) have started initiatives into operations with unmanned bridge controlled from a ROC (Remote Operation Centre). Cost and emissions are addressed by operating a fleet of vessels at reduced speed with reduced manning compared to conventional manned vessels.

The transition toward autonomy is suggested to be a gradual process, (IMO 2018, Lloyd´s 2016, DNV GL 2018), with classification and legislation undergoing modifications. To maintain a high level of safety (Porathe et al., 2018) at sea when introducing ROC operation, there is a need to expand research on the capabilities of the human in a ROC. Currently limited research exists using objective methods to evaluate information navigators’ access during a voyage.

Scalability of autonomous ships operation depends on manning level and the level of autonomy. Three manning levels is suggested (Rødseth, 2017): manned bridge; unmanned bridge – crew onboard, and unmanned bridge). One way to allow for operation of vessels with unmanned bridges is the use of remote control from a ROC. There is, however, a need to establish a safe and efficient operational concept for MASS. The Autoship project (Colella et al., 2023) discuss opportunities and challenges with MASS (Maritime Autonomous Surface Ships). They state *"However, MASS, which can be classified in the category of complex socio-technical systems, are associated with unprecedented levels of systems complexity as well as multifaceted and unpredictable interactions between the involved subsystems, environment and humans."* The previous research project, MUNIN (MUNIN, 2016) aimed "to develop and verify a concept for an autonomous ship, which is defined as a vessel primarily guided by automated on-board decision systems but controlled by a remote operator in a shore side control station".

These projects focus on the challenge of moving operating a single vessel from the ROC. However, limited research has been done for the operational concept where when one ROC operator is responsible for more than one vessel, a multi-asset control situation. Some similarities might be found comparing with the task of a VTS (Vessel Traffic Service) operator. Today VTS operators monitor and supervise the safe passage of vessels above 50m in areas having high traffic density. However, the responsibility of the VTS operator does not cover monitoring of leisure crafts. Further IMO adopted resolution A.578(14) in 1985(IMO, 1985). The Guidelines states. *"2.1.5 Care should be taken that VTS operations do not encroach upon the master's responsibility for the safe navigation of his vessel, or disturb the traditional relationship between master and pilot."* The responsibility of a ROC operator includes safe navigation that extends to the duties of the ship master. The International Convention for the Safety of Life at Sea (SOLAS), 1974, chapter V describes the general obligation of the master (SOLAS, 1980).

Using a simulated environment to explore challenges related to MASS operation challenges and opportunities is relevant, but there is a need to establish effective ways to collect data. One promising technology is to use eye tracking to provide a quantitative comparison of objective data and to validate the method using interviews. Additionally, eye tracking data can provide a method to compare the visual vigilance of individual operators. Based on this, the paper suggests a method analyzing eye-tracking data to visualize objective findings when operating MASS.

To ensure repeatability the test was performed using a tailor-made simulator based on findings from LOAS user study (Kaarstad et al. 2021).

The purpose of using eye-tracking data in this second study is to help quantify the actual time spent by participants on different information elements in the visual information displays.

The method has the potential to provide a quantitative comparison of objective data validated through interviews and workload measurement.

Situational awareness (SA) is central when focusing on safe operation and effective display interaction design (Endsley, 2012.) As a framework to assess the method and give momentum to the data collection, we explore two areas through the proposed approach using of eye-tracking.

- Which information in the land-based operation centre displays is most important?
- Which work display concept do the operator prefer for specific events and increased stress levels in the scenarios, bridge view vs. birds' perspective?

The first research question is inspired by the first level of Situation Awareness (Endsley, 2012): does the operator have the necessary information to understand the situation? The second question is inspired by the second SA level, being able to comprehend data into a meaningful picture through a suitable display concept. The questions are explored through the proposed method in a user study using a simulated maritime environment with navigators and vessel traffic operators.

Strategic thinking and effective decision making are essential skills to maintain operational safety during MASS operation. Boyd's Observe-Orient-Decide-Act OODA loop (Necesse, 2020) provide a framework for evaluation of the participants shift in focus over the test scenario.

The research presented in this paper is performed within a larger project financed by the Norwegian Research council of Norway through the project Land-based Operation of Autonomous Ships (LOAS). Participants of the project are Kongsberg Maritime, IFE and NTNU. The project objective is to develop and test interaction solutions for a ROC, ensuring safe and efficient supervision of autonomous ships.

This paper is organized as follows, the next chapter presents background information from an earlier LOAS study using eye-tracking data, then this paper´s study explain setup, participants and research into eye-tracking and analysis. Next, statistics and results from eye-tracking analysis are presented, followed by a discussion and topics for further work.

2 The proposed method based on earlier studies

The first section summarizes the research findings from previous LOAS user studies. The second section describes the participants, scenario and task assigned to the test participants. Last section explains the user study procedure and data collection.

This paper´s research concept builds further on the findings in the previous studies (Kaarstad et al., 2021; Braseth et al., 2022). The previous study found that a "bird´s" perspective, seeing the maritime situation through a large map coverage was preferred when supervising more than one ship.

The method developed in this study used eye tracking data collected through Pupil Labs Invisible (Pupil-Labs, 2023) eye-tracking glasses and software. The following is developed on basis on earlier research:

- A scene analysis software implemented in MATLAB (MATLAB, 2023) has been developed to identify screen and information elements the participants focus on during the scenarios.

- A statistical software has been developed to analyze the screen and information used by each participant.

The developed Software-tool (SW) enable eye tracking data collected in the head reference frame Fig.1 to be projected onto the reference frame of the screens, that is the global reference frame.

Fig 1. View from eye-tracking glasses and camera used by participants. Image recognition is used to identify markers and display information overlay.

An example of a resulting eye tracking point cloud in the global reference frame is show in Fig. 2. By comparing Fig.1 and Fig.2 it can be seen that the global frame is extended beyond the visual field of the camera. Data analysis and statistics in this paper is based on the global space time eye point cloud from each participant.

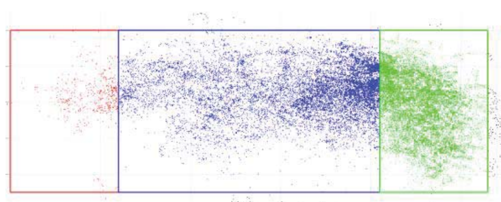

Fig. 2. Eye tracking point cloud in the global frame. The points are classified as map (red points on left screen, blue for center screen) or video (green on right screen).

The research questions from the previous chapter are addressed through analysis of data from a simulator study performed with eight participants of experienced navigators and vessel traffic operators. They are tasked to safely navigate three autonomous cargo ships simultaneously crossing the Oslo fjord. The control room setup designs use map centric layouts seeing the maritime environment from a distance (birds view), with additional on-ship video feed support. The study uses objective data driven techniques collecting eye-tracking data, processed by image recognition software, and evaluated by repeatable performance matrixes to minimize the risk of evaluation bias.

2.1 Study 2 participants

The user second study was performed at Kongsberg Norcontrol over a two-week period. All eight (one female, seven male) participants were educated navigators with experience from bridge watch. Four of the participants worked at a major ferry company, four worked at the Norwegian Vessel Traffic Service operators. The VTS operators had former background as navigators.

The participants' average age was 42.6 years, Average experience working in the maritime domain was 22 years, ranging from 8 to 40 years.

2.2 Scenario

A multi asset scenario was designed and recorded in advance of the testing, ensuring repeatability. The scenario mirrored a normal crossing of the Oslo-fjord with realistic commercial and leisure traffic complexity. Participants were told to monitor the safe passage of three unmanned ships (own ships) simultaneously crossing between Moss and Horten. The scenario duration was approximately 12 minutes where the participants supervised the 5.3 nm crossing.

Fig. 3. Operational area the port of Moss in the top righ corner, Horten to the left. TSS indicated by dotted lines north-south in the center of the area.

Own vessels were crossing the Traffic Separation Scheme of the Oslo fjord with larger ships traveling North or South. The scenario was set to daylight conditions with light wind and waves. Hazy conditions meant that visibility varied from less than one nautical mile to a few nautical miles, stimulating participants to make use of instruments and digital tools. Own ships were equipped with functionality for automatic docking and undocking. This was used when leaving or arriving at the port of Moss next to Bastø Fosen ferry lane.

All participants had local knowledge from sailing area, the operational area with voyage plan in blue is shown in Fig. 3

Four situations in the scenario needed high attention. In situation 3 it would be expected that the participants would intervene or take manual command of own vessel.

1. Time: 1:30 ASKO 1 leaving port of Moss
2. Time: 6:42 ASKO 1 with leisure craft on starboard side at a shallow angle towards Moss. Visible in map view, leisure craft not visible in onboard camera view.
3. Time: 8:43(T1) -11:09(T2) ASKO 2 leisure craft from port visible in map view, not visible in onboard camera view.
4. Time: 12:06(T3) ASKO 1 collide with dock in port of Moss. ASKO 1 does not reduce speed hitting the dock sailing at 10.0 kt.

Situations 3 and 4 are analyzed in further detail and resulting findings are presented later in this paper. The numbers in Fig. 4 show the location of high attention situations.

2.3 Test environment design layout

The user study (Kaarstad et. al, 2021; Braseth et. al, 2022) found that a "bird´s" perspective, seeing the maritime situation through a large map coverage is preferred when supervising more than one ship. Fig. 4. show the test environment with 3 screens.

Fig. 4. Test environment with participant to the left and observer to the right in the picture.

This paper featured an updated design based on findings and workshop feedback from two navigators. Detailed view of screen information is shown in Fig 5. The left screen presents a "zoom out view" (projection) providing overview and long timeframe. The center screen shows the main operational area map (comprehension) with added information detail for medium timeframe. Video feed from each vessel bridge and instrument information was presented on the right screen (perception) for the immediate timeframe. The video feed was based on a commercial 3D vessel simulator from Kongsberg Digital.

The high attention situations are added in Fig. 5 to give an impression of location and workload during the test.

Fig. 5. The visual information from map and onboard video making up the control room environment. Main screen 75", left and right 43" @ reading distance 1.25 [m].

2.4 User studies procedure

The study was conducted in a simulator setting. Participants were welcomed and informed about the project and about the study. Participants were told their role as master and that their task during the study was to supervise the safe crossing over the Oslo-fjord for three ships. Further witch data we were going to collect was explained. As the scenarios were prerecorded, it was not possible to take action or intervene with the autonomous ships during the test. Instead, they were asked to verbalize their observations and any actions they would have taken on a real voyage. If found necessary, they were told to verbalize if they would like to take manual control, disabling autonomous functionality. The participants were informed about how to communicate on VHF with other vessels for clarifications. VHF communication was simulated by participants calling and instructor answering the VHF call. All participants were briefed about the design layout for them to familiarize themselves with the setup. The location of own ships on map and video feed was explained prior to the tests.

Test scenarios were started after participants signed an informed consent and filled in a background questionnaire. Questionnaires were answered after the scenario, followed by a semi-structured interview.

2.5 Data collection

The study collected audio and video data for each participant as they supervised the autonomous ships. The location of recording equipment was behind the participant as shown in Fig. 4.

Additionally head mounted eye-tracking glasses with camera recording were used. Fig. 1 shows a video frame from the eye tracking camera augmented with identification points from the developed analysis SW. Eye tracking dataset for participant five and six of was unusable and has been omitted from the analysis.

During the tests the participants were observed by the researchers, taking notes for the designed high attention events. After completing the scenarios participants filled in questionnaires on situation understand and workload. For workload evaluation, the NASA Task Load Index (NASA-TLX) was used (Hart et al., 1988). Finally, a semi-structured interview was performed with each participant.

3. The resulting findings

Based on quantitative eye tracking data presented in section 2 and interviews with participants, we outline the following findings for simultaneous supervision of three autonomous or highly automated ships.

3.1 Eye tracking findings

The results of the eye tracking analysis indicate map to be the preferred visual information source when supervising three ships. Fig. 6 shows the percentage of focus on map or video over the duration of the scenario. All test subjects prefer map view the first minute, after this test subject one, three, four and eight switch focus to video, indicated by graphs falling. Five minutes into the scenario all test subjects had stabilized within an individual band of 5-7% until time T3. The final ratio between preference for map or video varies significantly with each participant. Five of the six test subjects stay within 60-90% band, averaging over 2/3 of time spent focusing on the map. Participant four deviates from the group with a slight 55% preference for video.

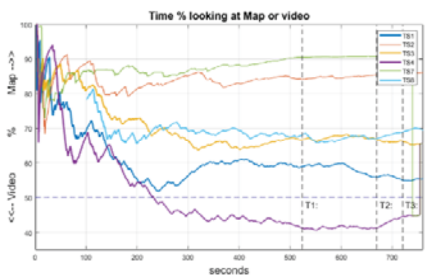

Fig. 6 shows the percentage of map or video preference plotted against time on the x-axis over the scenario duration.

For the high attention situation data from test subject two and seven has been omitted due to poor data quality after time T1. T1 is the time when Asko 2 changes to collision course with John. Fig. 7 visualizes the relative change in preference of map or video, graphs has been aligned at T1. At time T2 all four participants had comparable scores. After this point in time all test subjects start to focus on the video indicated by falling graphs. At time T3 John does an evasive maneuver, three of the four test subjects then shift focus to the map view, while the last test subject focus 50/50 between map and video.

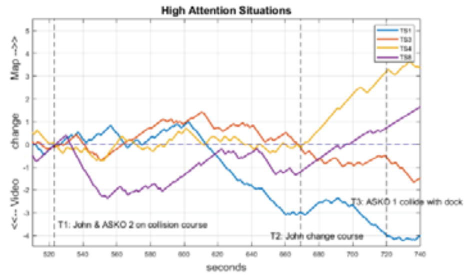

Fig. 7 compares the relative change of map or video preference plotted against time on the x-axis over high attention situation.

To get a deeper understanding of the visual focus for the high attention situation, the visual eye point cloud for the test subjects from time T1 to T2 has been analyzed. Fig.7 and Fig.8 show the two extremes of the point cloud shapes. Test subject tree, Fig.7 has a concentrated blue point cloud around the location of John Asko2 vessels. This concentration of visual focus was seen in five of the six participants. Test subject one deviates with a large less focused point cloud fig.8.

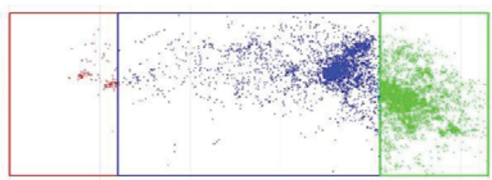

Fig.7. High attention situation point cloud in the global frame participant three. The points are classified as map (red points on left screen, blue for center screen) or video (green on right screen).

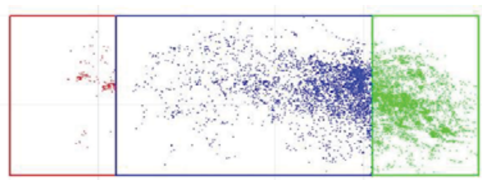

Fig.8 High attention situation point cloud in the global frame participant one. The points are classified as map (red points on left screen, blue for center screen) or video (green on right screen).

3.2 Interview findings

The participants reported back the map display to be easier to read than normal ECDIS (Electronic

Chat Display and Information System) on ships. Regarding available information in the map participant reported sufficient and clearly visible information.

Assigned task during the test was not hindered by the eye tracking glasses. All but one participant forgot about the eye tracking glasses during the test. The one who noticed the glasses reported the strength was different to his normal glasses.

4. Discussion and further work

In this paper we have proposed a new approach to study the operation of a fleet of autonomous ships in a simulated environment by combining the use of eye-tracking and software. The purpose is to establish an objective method to quantify and evaluate visual focus. We have applied the method on a case with supervision of three ships. For large parts of the study, the data show map centric view (birds' perspective) to be the preferred display interface. However, as stress levels intensify, on-ship video feed becomes the preferred visual information source. Using the research questions as cases for the method, it was found that five of the six participants analyzed used map as the most frequent source of information. This finding aligns with the findings in interviews used as validation of the method. Additionally, the method enables deeper understanding of participant visual focus, showing all participants favored the map in the beginning of the test. This aligns with Boyd's Observe-Orient-Decide-Act OODA loop. That is participants favored the map view to orient in the navigational area, before accessing detailed information in the video feed from each vessel. Significant variation was found for final preference for map or video, this suggests that participants may have different strategies for obtaining visual information. Further studies should seek to quantify the influence of individual factors for visual information preference and resulting SA.

A high attention scenario was studied to evaluate the methods ability to detect change in focus by the participants. The findings show strong indication of the method's ability to detect shift of focus in stressful situations. Careful examination of Fig. 7 indicates that all participants detect and focus on the collision course of ASKO 2 in the timeframe 550-565 seconds. At time ~600 seconds all participants have primary focus on the video feed. This aligns well with the qualitative findings in the interview as all participants reported looking at video (bridge view) to confirm the situation visible in the map display. In addition, most of the participants tried to call John on VHF to clarify the situation during the test. The high attention situation is resolved when the leisure craft John change course at time ~670. A clear shift in focus from video to map is visible for three of the four participants, after the situation is resolved. This is shown in Fig. 7 at time ~670 seconds by rising graphs towards map preference.

The tendency to prefer video feed is greater in the high attention situation, than for normal operation, this might indicate misplaced salience and attention tunneling (Endsley et al., 2012). Further research should investigate if the proposed method can be used to detect SA daemons such as attention tunneling. Real-time monitoring of visual perception may also be helpful providing ability to use machine-initiated cues to attract attention of an operator, without adding unnecessary visual cluttering.

Further study into the ability of one operator to safely monitor more than one vessel has been addressed in the user study LOAS and the results from this study will be presented in separate papers.

5. Limitations

The method presented in this paper only evaluates what visual information participants spend most time on, not the value of the information to maintain a high level of SA. We acknowledge that factors such as poor readability or poor visual design may increase the time participants spend obtaining information from a display. We expect that we and other researchers will modify, refine and further develop the method seeking an objective quantitative method to evaluate SA.

On average each participant dataset contains 18.000 image frames and 144.000 eye-tracking data points, however the sample size of six datasets does not pose a sufficient statistical background to firmly conclude findings related to the research questions in this study.

Acknowledgement

We would like to thank the navigators and vessel traffic operators who participated in our simulator study on their spare time or in a busy working day. We would also like to acknowledge the shipping company that forwarded our request and generously let us use their simulator in our study. The study is part of a larger research project financed by the Norwegian Research council of Norway through Land-based Operation of Autonomous Ships

References

Braseth, A.O., Kaarstad, M., Høstmark, J.B., Strømmen, G. (2022). Supervising Autonomous Ships – A Simulator Study with Navigators and Vessel Traffic Supervisors. ESREL 2022

Colella, M. M., Azarko, A., Rødseth, Ø. J., Nordahl, H., Theotokatos, G., Ruud, T., Helgås, V. (2023). The Autoship project: Speeding-up the transition towards a next generation of autonomous ships in the EU, project article, downloaded 28.03.23: https://www.autoship-project.eu/wp-content/uploads/2023/03/AUTOSHIP-Ebook.pdf

Endsley M.R., Bolté B., Jones D.G. (2012). *Designing for Situation Awareness*, An Approach to User-Centered Design, Taylor & Francis, London & New York.

Hart, S.G., Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Hancock, P.A., Meshkati, N. (eds.). Human Mental Workload. Advances in Psychology. Vol. 52. Amsterdam: North Holland. pp. 139–183.

IMO (1985) RESOLUTION A.578(14) Guidelines for vessel traffic services, Downloaded 23.03.2023 https://wwwcdn.imo.org/localresources/en/KnowledgeCentre/IndexofIMOResolutions/AssemblyDocuments/A.578(14).pdf

Kaarstad, M., Braaseth, A.O., Strange, E., Høstmark, J.B. (2021). Toward safe and efficient operation of autonomous ships from a land-based center. ESREL-2021

Massterly (2023), web page accessed 29.04.2023, https://www.massterly.com/news-1

MATLAB (2023), web page accessed 29.04.2023, https://www.mathworks.com/products/matlab.html

MUNIN (2016) Maritime Unmanned Navigation through Intelligence in Networks, web page downloaded 28.03.23: http://www.unmanned-ship.org/munin/

Necesse (2020) Vol. 5, Issue 1, 142-165 Sjøkrigsskolen, web page accessed 29.04.2023, https://fhs.brage.unit.no/fhs-xmlui/bitstream/handle/11250/2683228/Boyds%20OODA%20Loop%20Necesse%20vol%205%20nr%201.pdf?sequence=1&isAllowed=y

Ocean Infinity (2023), web page accessed 29.04.2023, https://oceaninfinity.com/ourtechnology/

Porathe, T., Hoem, Å., Rødseth, Ø., Fjørtoft, K., Johnsen, S.O. (2018). "At least as safe as manned shipping? Autonomous shipping, safety and 'human error,'" Safety and Reliability – Safe Societies in a Changing World, Haugen et al. (Eds), Tyler & Francis Group, London, UK (2020)

Pupil-Labs (2023) Pupil Invisible, web page accessed 29.04.2023, https://treaties.un.org/doc/Publication/UNTS/Volume%201184/volume-1184-I-18961-English.pdf

SOLAS, (1980) Treaties and international agreements Registered on 30 June 1980 No. 18961, Downloaded 20.04.2023: https://treaties.un.org/doc/Publication/UNTS/Volume%201184/volume-1184-I-18961-English.pdf

Rødseth, Ø. J. (2017) Definitions for Autonomous Merchant Ships. NFAS Norwegian Forum for autonomous ships. DOI:10.13140/RG.2.2.22209.17760