

Subjective Evaluation for 2D Visualization of Data from a 3D Laser Sensor

Patrik Lif^(✉), Gustav Tolt, Håkan Larsson, and Alice Lagebrant

Swedish Defence Research Agency, Linköping, Sweden
{patrik.lif,gustav.tolt,hakan.larsson}@foi.se,
alice.lagebrant@live.se

Abstract. The purpose of this study was to increase understanding of how information from a 3D sensor should be presented in two dimensions for users. Stimuli consisted of video footage where people moved in different patterns and carried out various activities. Seven visualizations were used in a simple- and a complex scene respectively. The subjective ratings and interviews shows that various visualizations highlight different parts of the scene and allow the user to prioritize different information. This means that the choice of display must be connected to the application. In future research these results will be supplemented with objective performance metrics, e.g. response time to detect targets and eye movements. Also, to understand end-users and increase their performance in real settings further task analysis will be conducted.

Keywords: 3D laser sensor · Subjective ratings · Human factors

1 Introduction

To build an effective system in a military setting a number of factors must be taken into account. From an ecological approach [1] and representation design [2] there is a cognitive triad between *domain/environment*, *interface* and *humans/users* that must be considered. From a technical perspective the focus is often on technical solutions, such as sensors and their technical performance. From a military perspective the focus is often on the environment where the system should be used. To develop an effective and user friendly system all these three parts must be taken into account, but considerations regarding sensor type are also important. Human factors put the user with abilities and limitations in focus. In our research we aim for an understanding of the whole picture, but here the focus is on how information from sensors shall be presented to make sure that the user gets good situation awareness [3, 4].

In both civilian and military contexts there is a considerable value to depict the environment from sensor information and in many situations it is important to detect and identify people. In this experiment we focus on data from an advanced 3D camera equipped with a pulsed laser so that each pixel acts as a range finder. Range information has proven to be very useful, e.g. for automatic target recognition at long distances and in difficult lighting conditions. The purpose of this study was to increase the knowledge and understanding of how information from a 3D sensor should be presented in two dimensions (2D) for users. The focus was on subjects' ability to

understand information on the display presentation, not on the sensors *per se* or other technical performance.

2 Method

A within-group design with *three display modes* (distance, intensity, and gated viewing) \times *two pseudo-coloring presentation schemes* (jet and gray) \times *two scenes* (simple and complex) was used. Moreover, an extra visualization of intensity with red marking was added to highlight items moving towards the observer. Participants watched the video sequences and gave their subjective opinions by answering a questionnaire and attending an interview.

2.1 Subjects

The participants in the experiment consisted of twelve subjects (seven women and five men) with average age 21.58 years and range from 19 to 27 years. The requirement to participate was minimum 18 years of age and adequate vision with or without correction, such as glasses or contacts.

2.2 Apparatus

The data used in this experiment came from the 3D imaging laser sensor ASC 3D-FLASH [5]. This sensor is an advanced camera with its own lighting source in form of a pulsed laser. The detector in the camera consists of 128×128 pixels each that acts as a distance meter that gives distance images with an image rate of up to 30 Hz (here: 10 Hz). The number of pixels and the frame rate is much lower than in a normal SLR camera because of the sophisticated electronics in the detector. Apart from a distance assigned to each pixel, there is also an intensity value that corresponds to how much of the emitted laser light is reflected back to the detector.

2.3 Stimuli

Stimuli in the experiment consisted of video footage from a dataset collected during a field trial. Data were collected in a number of scenarios where people moved in different patterns and carried out various activities. From the collected data, movies for one simple and one complex scene were created in MATLAB. In the simple scene, two people walked towards each other, shook hands, passed around each other and went back in the direction they came from. In the complex scene, five people walked irregularly within a limited area, passing a bag between them.

The movies were based on the same data but the visualization varied regarding pseudo-coloration as a function of either *intensity* (the amount of received laser light), *distance* or so-called *gated viewing* (GV). Gated viewing means that the camera shutter opens for a very short period of time, so that only laser light corresponding to a

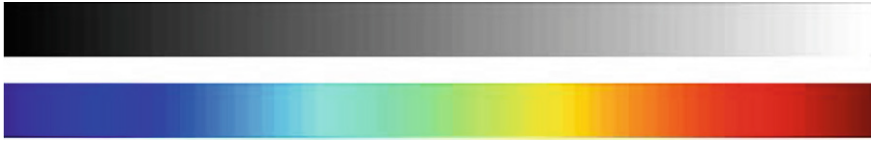


Fig. 1. The colormaps Gray (above) and Jet (below) (Color figure online)

particular range interval is detected. This technique allows for suppression of disturbing elements such as vegetation, rain, snow and fog. By adjusting the range interval so that it does not include the background, objects can also be made to stand out clearly from the background. Strictly speaking, the 3D-FLASH is not a GV system, but since the collected data contain range values typical GV videos could be simulated.

Seven display configurations for each scene were used. The colormaps “Gray” and “Jet” were adopted from MATLAB colormaps [6] (Fig. 1).

Both scenes were visualized regarding distance (Gray and Jet), gated imaging (Gray and Jet) and intensity (Gray, Gray with red marking for objects moving towards the sensor/viewer, and Jet). The seven conditions are hereafter referred to “distance-gray”, “distance-jet”, “GV-gray” and “GV-jet”, “intensity-gray”, “intensity-jet”, and “intensity-red”. Still images from the simple and complex scene are presented in Figs. 2 and 3. The presentation order between simple and complex scene was balanced between the participants and the order of movies within each scene was randomized.

2.4 Procedure

After welcoming the participants individually and briefing them about the experiment (purpose and procedure), they received some written information and had the opportunity to ask questions to the experiment leader.

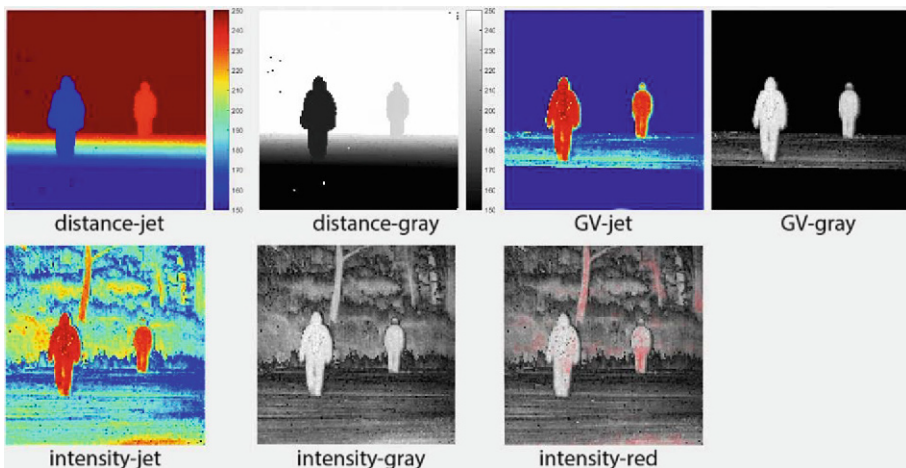


Fig. 2. “Simple scene” conditions (Color figure online)

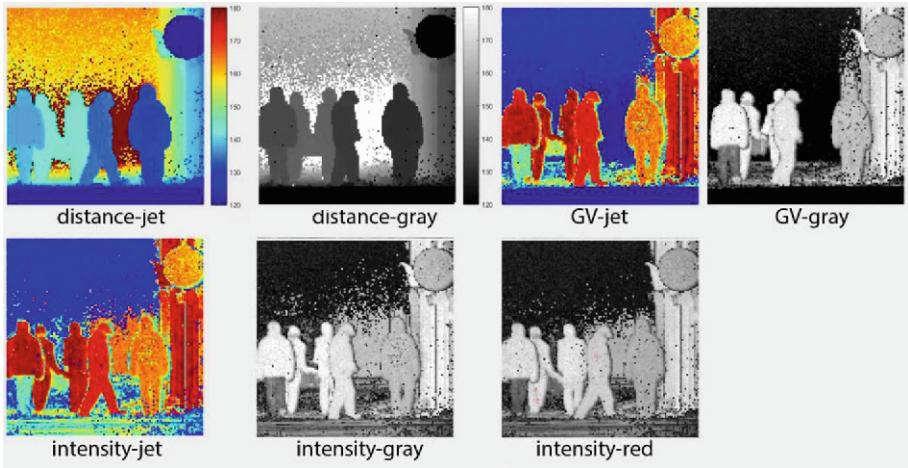


Fig. 3. “Complex scene” conditions (Color figure online)

Then an introduction was given to make the participants familiar with the situation and the test material. They were introduced to the different types of visualization and then received about ten minutes training, where all types of visualizations to be used in the trial and the survey questions were explained. The participants were informed to focus on the visualizations with different display modes, type of scene, and colormap.

After each scene, subjective information were collected from the participants in the form of a questionnaire. Each question was answered using a seven-point scale, seven being equivalent to the best possible results and one representing the worst case. When the participant had seen all scenes and responded to the related questionnaire, a semi-structured interview was conducted to evaluate participants understanding of the display configurations, e.g. perception of color, distance and direction.

3 Results

The results include statistical analysis of data from the surveys and summarized information from interviews. The data from surveys were analyzed first with a two-way ANOVA [7] with type of visualization (7 types as described in Figs. 2 and 3) and type of scene (2 types as in Figs. 2 and 3) as factors. This was followed by a three-way analysis of variance to analyze the main and interaction effects of type of visualization (distance, intensity, and GV), colormap (gray and jet), and type of scene (simple and complex). In the later analysis the visualization intensity-red were excluded. A Post Hoc test was conducted with Tukey’s Honest Significance Test [8]. Only the most important results are presented here. Information from the interview are here presented summarized, highlighting only the most important and frequent answers.

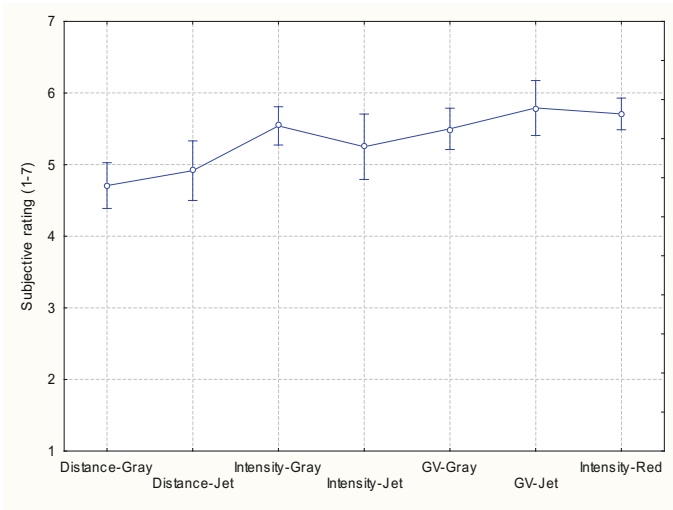


Fig. 4. Mean and standard error of mean for the seven display configurations

3.1 Survey

Here we present data from seven questions (translated from Swedish), five (question 1–5) about the visualization and two about the image quality (question 6–7).

Question 1: How easy/difficult was it to understand what happened in the scene?

A two-way analysis of variance showed that there were a significant main effect for type of scene, $F(1,11) = 5.22, p < .05$. The participants perceived it harder to see which person moved against the observer in the complex scene than in the simple one. There were also significant main effect for type of display, $F(6,66) = 4.63, p < .001$ (Fig. 4). Tuckey Post Hoc test shows that display distance-gray were rated lower than intensity-gray, GV-jet, and intensity-red ($p < .05$). Also, distance-jet were rated lower than GV-jet ($p < .05$).

The three-way analysis of variance showed a significant main effect of type of display $F(2,22) = 9.2202, p < .005$. Tukey’s Post Hoc test showed that display distance were rated lower than displays for intensity and GV ($p < .05$). Also, there was a tendency to significant interaction effect $F(2,22) = 3.3971, p = .052$. The post hoc test showed that display distance were rated lower than display intensity and GV in the complex scene ($p < .05$), while all display configuration in the simple scene were rated equal ($p > .05$).

Question 2: How easy/difficult was it to see the different directions that people were moving in?

A two-way analysis of variance showed that there was a significant main effect for type display, $F(6,66) = 4.97, p < .001$ (Fig. 5). According to Tukey’s post-hoc analysis, participants rated the display intensity-red display higher than intensity-gray, intensity-jet and GV-gray ($p < .05$).

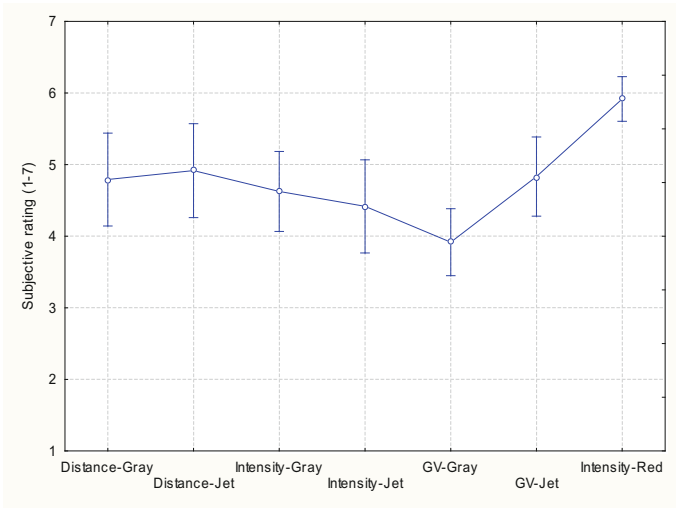


Fig. 5. Mean and standard error of mean for the main effect of type of display

The three-way analysis of variance showed that there was a significant interaction effect between type display and colormap, $F(2,22) = 4.27$, $p < .005$. According to Tukey's post-hoc analysis, participants rated the display GV-jet higher than GV-gray ($p < .05$), and no significant differences due to colormap for distance- and intensity displays.

Question 3: How easy/difficult was it to see which person walked against you?

A two-way analysis of variance showed that there was a significant main effect for type display, $F(6,66) = 9.77$, $p < .001$. According to Tukey's post-hoc analysis, participants rated the display intensity-red higher than all other displays ($p < .05$). The three-way analysis of variance showed that there was a significant main effect of display type, $F(2,22) = 3.81$, $p < .05$, where display GV was rated lower than display distance. There was also a significant main effect of colormap, $F(1,11) = 9.61$, $p < .05$, and Tukey Post Hoc test showed that colormap gray were rated lower than jet ($p < .05$).

Question 4: How good/bad was your experience of the visualization concerning estimation of distance?

A two-way analysis of variance showed that there was a significant main effect for type display, $F(6,66) = 9.13$, $p < .001$, see Fig. 6. According to Tukey's post-hoc analysis, participants rated the display distance gray and distance jet higher than all other displays ($p < .05$).

The three-way analysis of variance showed that there was a significant main effect of display type, $F(2,22) = 16.51$, $p < .001$, where display distance was rated higher than display intensity and GV ($p < .05$).

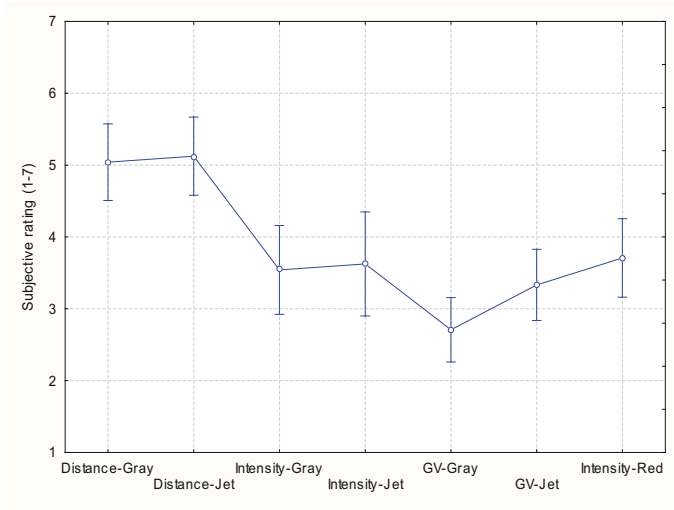


Fig. 6. Mean and standard error of mean for main effect of type of display

Question 5: How did you perceive the risk of confusing different people in the scene?

A two-way analysis of variance showed that there was a significant main effect for type of scene, $F(1,11) = 8.29, p < .05$. According to Tukey’s post-hoc analysis, participants rated the complex scene lower than the simple scene ($p < .05$). There was also a significant interaction effect between type of scene and type of display $F(6,66) = 3,1943, p < .01$, see Fig. 7. Tukey Post Hoc test showed rated display distance-jet and intensity-jet lower in complex- than in simple scene ($p < .001$). Also in complex scene distance-jet were rated lower than display GV-jet and intensity-red ($p < .05$).

The three-way analysis of variance showed that there was a significant main effect for type of scene, $F(1,11) = 8.44, p < .05$, the risk to confuse subject between each other was higher in the complex- than in the simple scene ($p < .05$). There was also a significant interaction effect between type of scene and type of colormap, $F(1,11) = 6.91, p < .05$ and a three-way interaction effect between type of scene, display and colormap, $F(2,22) = 3.20, p < .05$. Tukey Post Hoc test show that display distance-jet and intensity-jet in the complex scene was rated lower than the other displays ($p < .05$). There was no differences between displays coded in gray ($p < .05$), and no differences between displays coded in jet for simple scene ($p < .05$).

Question 6: How good/bad was your experience of the contrast between different persons?

There was no significant differences in the analysis of variance ($p > .05$).

Question 7: How did you experience the image noise?

The three-way analysis of variance showed a significant main effect for type of display, $F(2,22) = 4.27, p < .05$, noise was perceived as more annoying with display

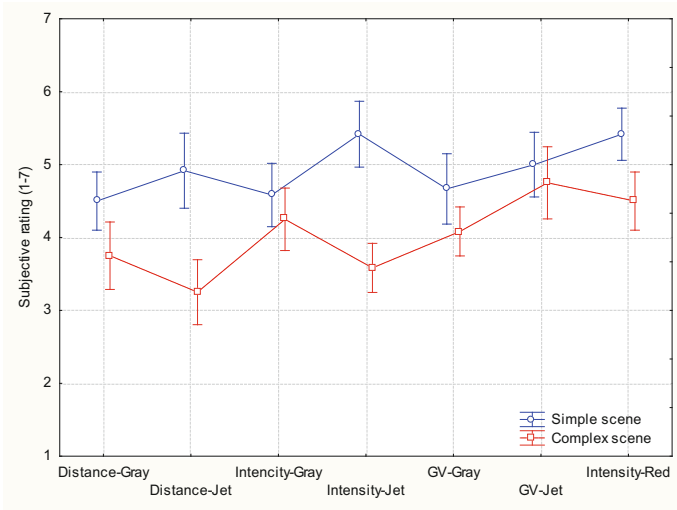


Fig. 7. Mean and standard error of mean for the interaction effect of type of display and type of scene.

intensity compared to GV ($p < .05$). A tendency to main effect was also found for colormap, $F(1,11) = 4.84$, $p = .050$, where Jet was perceived as more disturbing than Gray.

3.2 Interviews

During the semi-structured interview all the visualizations were presented and the participants was instructed to discuss from a number of selected focus areas: scene understanding, the color scale impact, distance and 3D perspective as well as noise and overall quality. In display mode distance-gray the low image detail level made it hard to understand how people moved and interacted with each other. There was an obvious risk to confuse individuals in certain situations (when they were at the same distance), but they stood out clearly from the background. In display mode distance-jet the color scale provided more accurate and precise distance and direction of the assessment than the gray scale which affects the understanding positively. Display mode intensity-gray gives good details and sharp contours and makes objects, people and the background clearly different from each other. Display mode intensity-jet was visually demanding and affected the reference system, mental workload and overall understanding very negatively. The colors, details and noise make the picture chaotic and the interpretation problematic. In display mode intensity-red the understanding of the scene was very high because of detail richness, and this improvement is without noise or workload is affected negatively. The visualization thus becomes very sophisticated and easy to use.

4 Discussion and Summary

This study shows that the various visualizations highlight different parts of the scene and allow the user to prioritize different information. This means that the choice of display must be connected to the application.

Display intensity were considered to be better for tracking people, seeing details and understanding the overall scene. Red marked direction provides further understanding of movement patterns and make the users more confident in their judgments. GV was considered useful mainly when focus was on the individuals in the scene and not on understanding the environment. Coding displays with distance was difficult to use in real-time as it requires a divided attention between the scale and the film sequence. It would therefore be better to use a pause function and then do more qualified assessments from a distance view. Grayscale is generally considered easier to use and perceived as easier to the eyes. Since the noise tends to be less disturbing in grayscale this visualization should be used when there are many impressions, e.g. detailed backgrounds. Presentation with color required more training and initially perceived as more demanding. However, color demonstrated strengths in being more accurate and sensitive. Generally speaking, direction and distance were perceived as easier to determine, but noise became more disturbing. The environment is used as reference to get an understanding of the scene and movement of people. The perception of people's direction and estimation of distances are negatively affected when the environment is absent in the visualization. On the other hand, the absence of disturbing background makes it easier to focus on people.

In this experiment we focused on what users thought of the different visualizations. In the future this will be supplemented with objective performance metrics. Examples of such measures are response time to detect targets, time to solve a task and measurement of eye movements. The basis for our research is to understand the end-users and increase their performance in real settings. With task analysis we can get even better understanding of user needs, and thereby tailor visualizations for specific users and tasks. These results are important to better understand how information from 3D sensors shall be presented for users, e.g. military personal on the ground and unmanned aerial operators.

References

1. Flach, J., Hancock, P.: An ecological approach to human-machine systems. In: Proceedings of the Human Factors and Ergonomics 36th Annual Meeting, pp. 1056–1058, Santa Monica, CA, USA (1992)
2. Woods, D.D.: Towards a theoretical base for representation design in the computer medium: ecological perception and aiding human cognition. In: Flach, J., Hancock, P., Caird, K., Vicente, K. (eds.) *An Ecological Approach to Human Machine Systems I: A Global Perspective*, pp. 157–188. Erlbaum, Hillsdale (1995)
3. Endsley, M.: Measurement of situation awareness in dynamic systems. *Hum. Factors* **37**(1), 65–84 (1995)

4. Endsley, M.: Toward a theory of situational awareness in dynamic systems. *Hum. Factors* **37** (1), 32–64 (1995)
5. ASC. Portable 3D Flash Lidar Camera Kit. <http://www.advancedscientificconcepts.com/products/older-products/portable.html> (2016). Accessed 11 Feb 2016
6. MATLAB. Colormap (2015). <http://se.mathworks.com/help/matlab/ref/colormap.html>
7. Hays, W.L.: *Statistics*. Harcourt Brace College Publishers, Fort Worth (1994)
8. Greene, J., D'Oliveira, M.: *Learning to Use Statistical Tests in Psychology*. Open University Press, Milton Keynes, Philadelphia (1982)