Postprint

Spatial Augmented Reality and Eye Tracking for Evaluating Human Robot Interaction

Elsa Bunz, Ravi Chadalavada, Henrik Andreasson, Robert Krug, Maike Schindler and Achim J. Lilienthal
Örebro University, Sweden
Email: firstname.lastname@oru.se

Abstract—Freely moving autonomous mobile robots may lead to anxiety when operating in workspaces shared with humans. Previous works have given evidence that communicating intentions using Spatial Augmented Reality (SAR) in the shared workspace will make humans more comfortable in the vicinity of robots. In this work, we conducted experiments with the robot projecting various patterns in order to convey its movement intentions during encounters with humans. In these experiments, the trajectories of both humans and robot were recorded with a laser scanner. Human test subjects were also equipped with an eye tracker. We analyzed the eye gaze patterns and the laser scan tracking data in order to understand how the robot's intention communication affects the human movement behavior. Furthermore, we used retrospective recall interviews to aid in identifying the reasons that lead to behavior changes.

I. Introduction

During interaction, humans rely on many implicit and explicit cues to make decisions and to predict the future actions of each other. In Human-Robot Interaction (HRI) it is therefore necessary to express these cues clearly, *i.e.*, the robot must be able to communicate intentions clearly understandable to humans in order to become a reliable, agile and adaptable co-worker. We are interested in intention communication of mobile robots operating in industrial scenarios. Such robots have been commonly using pre-defined paths for navigation. Nevertheless, with the growing demand for increased autonomy, they soon will have to make decisions online and may not always stick to a pre-defined paths in order to increase efficiency. However, such kind of behaviour may lead to uncertainty if the robot doesn't communicate its intentions to the humans sharing its workspace. Our earlier findings regarding a Spatial Augmented Reality (SAR) approach [1], [2] indicate that a mobile robot communicating its future intentions by projecting them onto the shared workspace, was able to improve comfort levels in humans. For certain attributes, this improvement was beyond the levels of human-human interactions. Also, the system encouraged the humans to actively choose safer paths around the robot.

In our previous work [1], Likert scale based questionnaires and trajectory analysis were used to evaluate and understand the influence of the chosen mode of intention communication on the human behavior. In this work, we report a redesigned experimental evaluation in order to provide a deeper understanding of how our system influences the human behavior and how it improves the human-robot interaction. Each subject undergoes 4 trials each with different projections in a
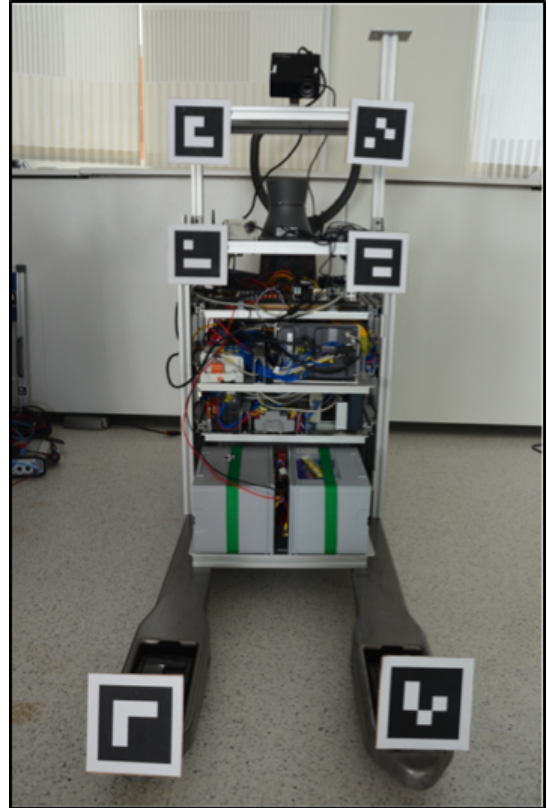

Fig. 1. The platform used for the evaluations.

constrained passage as shown in Fig. 4. The subject wore an eye tracking device during all the trials in order to track the eye gaze and record the world view while interacting with the robot. During the interaction, the subject and the robot were captured using a laser scanner and also a video camera. In order to get further insights about the subjects behavior, upon completion of the interactions with the robot, each subject was interviewed while being shown the recorded video overlaid with their eye gaze data. This helped us to interpret their actions and gather feedback for further development.

The key contribution of our work is the evaluation method and analysis which lets us get deeper insights into how intention communication affects human behavior in a human robot interaction scenario when compared to the likert scale questionnaires and quantitative measurements. To this end, we
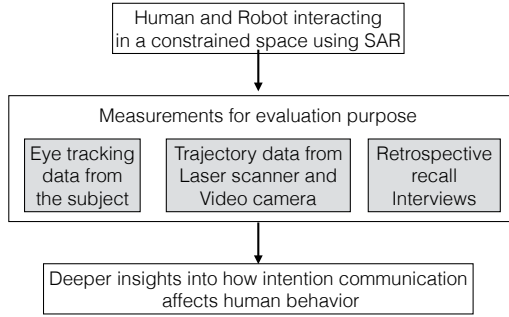
Fig. 2. Outline of the experimental evaluation

carried out interaction experiments as outlined in Fig. 2.

i) First, eye-tracking was used to measure the time spent on paying attention to the projected pattern as well as the robot.

ii) Second, the paths taken by the human subjects in the encounter situation were evaluated to analyze whether a significant difference in the path can be found between experiments.

iii) A recall interview was performed with each participant after the trials were over. These interviews were stimulated with the videos of the eye-tracking overlaid with the gaze in order to acquire the subjective impression and reasoning of the participants.

II. RELATED WORK

A. Intention Communication

During walking, humans communicate their motion intentions using different types of cues such as gazing or by adapting their trajectories according to a pattern of movement directions [3]. If robots are to operate in human environments, they must adapt to human expectations such that the common human interaction patterns need not to be changed drastically. In this context, several researchers outline the benefits of revealing the intentions of the robot: Takayama et al. [4] claim that if the robot shows forethought before performing a functional action, people will be more likely to see the robot as being more appealing, approachable and sure of its subsequent actions. The ability to predict is stressed by Turnwald et. al. [5] who show that humans are not only reacting but use prediction to plan their motion. This point is substantiated by the consensus in all studies regarding the fact that humans anticipate future motion of other objects and with that also possible collisions. Especially for obstacle avoidance this beforehand usage of information seems to be important.

Lichtenthäler [6] investigated legibility and predictability of robot navigation. Here, robot behavior was defined as legible if a human can infer the next actions, goals and intentions of the robot with high accuracy and confidence and the robot behavior fulfills the expectations of human interaction partner. Similarly, predictability encapsulates the ability to predict the

robot's trajectory. The authors in [6] report various correlations with legibility, namely safety, comfort, surprise, efficiency and perceived value. This stresses the importance of making a robots behavior legible to obtain the desired high values in these categories.

Humans acquire an estimated $85 - 90\%$ of information through their visual system [7]. Researchers like Matsumaru [8], Coovert at al. [9] and also ourselves in prior work [1], [2], used SAR to reveal a robot's future intentions with encouraging results. In this line of thought, May et al. [10] investigated pass-by situations where a human and a robot aim to pass through a corridor trying to circumvent each other given spatial constraints. Hereby they focused on the perceived comfort, the ambiguity of the signal and the preferences of the participants after experiencing the conditions which differed in the way and presence of navigational intent communication. Their results showed that the comfort of the humans encountering the robot can be increased by non-verbal signs.

B. Eye-tracking

The most reported event in eye-tracking data relates to the state when the eye remains still over a period of time lasting anywhere between tens of milliseconds to upto several seconds and this is called a fixation [11]. Just and Carpenter [12] formulated the eye-mind hypothesis [13] which states that there is no relevant delay between what is fixated and what is processed, which is an important assumption to be able to interpret the measured fixations of a human.

Already Yarbus [14] noted that the task is relevant for the gaze behavior. He found that vision seems to be tightly linked to the cognitive goals of the observer. Several studies suggest that the measured movements of the eyes during a task can be used to conclude the distribution of visual attention. The term attention can be defined in terms of its task which is according to Müller and Krummenacher

They state that one of the main functions of attention lays in the selection of perceptive information for the control of behavior. There are several studies supporting the link between selective attention and the planning of saccadic eye movements [15]. The authors of [16], [17] were observing humans using tools and manipulating objects. They found a strong relationship between the executed action and the gaze of the participants. Thus it is possible to use the overt attention, $e.g.$ measured through the fixations, as a sign of the attention. Several psychophysiological and imaging studies furthermore give significant evidence that the relocation of attention is reflected in the fixations [18]. In addition, there are several studies suggesting that the visual system of attention is involved in the planning of the whole motion sequence [15].

By analyzing differences in the scanned paths between novices and experienced drivers, Antonya et al. [19] identified experience as an important factor influencing the gaze. Also Yarbus [14] found that the gaze positions for observing pictures vary a lot depending on the task carried out by the participants. According to recent work on natural tasks [20]

it becomes obvious that the cognitive goals of the observer have an important influence on the distribution of gaze. Furthermore, the gaze priorities seem to be adjusted very quickly to environmental probabilities. To summarize, gaze locations seem to be very tightly linked to the task [21]. There have been several studies looking at gaze patterns during obstacle avoidance tasks, where the participants were moving and needed to avoid obstacles – this situation resembles an encounter with a robot in a corridor. During walking, the information given through the vision is an important source and therefore, according to Hayhoe and Rothkopf [21], observers must learn where and when to look at critical locations during walking while at the same time controlling direction and balance. Thus, the gaze reflects a learned behavior that is adapted depending on the locations that are rated as critical. Patla and Vickers [22] found in their experiments that participants planned stepping over an obstacle before actually reaching the obstacle, so there weren't any fixations on the obstacle when they were stepping over it. Overall, they found these similarities in the gaze pattern over their participants which supports the possibility of analyzing the gaze and comparing it. Considering all these findings, it seems to be valid to draw conclusions about a person's attention from the measurements of the eye movements.

C. Retrospective Interviews

Quantitative measurements, as discussed above, provide us with data to measure the changes and possible influences on the human behavior. However, the reasons behind the observed behavior cannot be inferred unambiguously. Therefore, sometimes retrospective interviews are used in addition to the eye-tracking to find explanations for the gaze patterns and to get a deeper insight in the underlying processes. Hereby often videos with the overlaid gaze are shown to the participants and they are asked to think aloud or explain their gaze patterns. Hansen [23] investigated the question of whether participants really remember their eye movements or resort to guesses when they are asked to explain them. To this end, he showed the participants recordings of someone else's eye movements. As the participants were able to detect the error, he concluded that humans indeed remember their own eye movements. Guan et al. [24] supported this hypothesis with their finding that humans look at objects in the same order as they afterwards also say they do. Thus, using retrospective interviews seems to be an appropriate method to deepen the understanding of the eye-tracking data and the laser tracking data.

III. Methodology

The robot used for the interaction scenario is the fork-lift type vehicle depicted in Fig. 1. It was built using a manually operated forklift which originally was equipped with motorized forks and a drive wheel. The platform was retrofitted with a steering mechanism and a commercial AGV control system. Two SICK S300 safety laser scanners ensure a safe operation forwards and backwards. A projector is mounted on the robot to project a pattern in front of the forks. For more
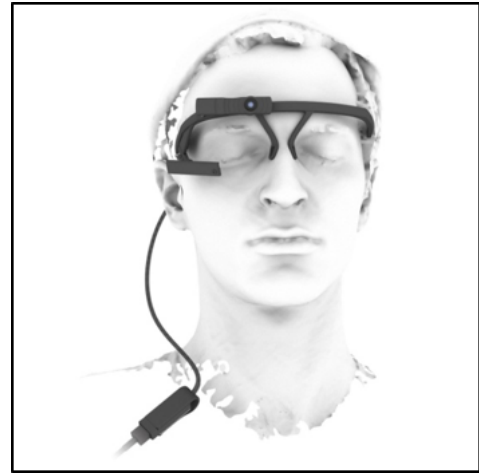

Fig. 3. The eyetracker used for the evaluations developed by Pupil Labs [25]

detailed information about how the projection was generated see [1]. Fiducial markers were attached to the robot in different places in order to define the areas of interest, thus, enabling an automatic categorization of the detected eye-gaze fixations.

For the acquiring of the eye-tracking data we used an eyetracker from Pupil labs [25], which is a mobile eye-tracking headset (see Fig. 3). It is equipped with a high speed world camera with a resolution of 1920×1080 for a framerate of 30 fps and two infrared spectrum eye cameras with a resolution of 640×480 and a framerate of 120 fps for each eye. Scene capturing was done using the open source software Pupil Capture, for categorization and analysis the open source software Pupil Player was used which was developed by Kassner et al. [25].

During the experiment, to track the path of the subject and the mobile robot during the encounter, a SICK LMS 500 laser scanner was used and was also recorded with a stationary video camera. The camera was facing the robot and thus recording the participant from behind. Furthermore, the screen of the computer used for playing the videos with the gaze overlay was captured together with the audio during the interview.

IV. Experimental Design

The experiment was designed to compare three different projections as shown in Fig. 5 and a case without projection, thus resulting in four trials per participant. Hereby pattern A and B were supposed to convey the future trajectory, with A being equivalent to the projection already used in previous experiments reported in [1]. Pattern A depicts the future trajectory over a time horizon while in pattern B the information is more compressed using an arrow pointing along the instantaneous movement direction. The arrow was chosen for several reasons. Bertamini et al. [26] provide evidence that angles attract attention while Bar and Neta [27] suggest that the human brain can detect sharp features very fast as this can help to signal potential danger. Larson et al. [28] showed
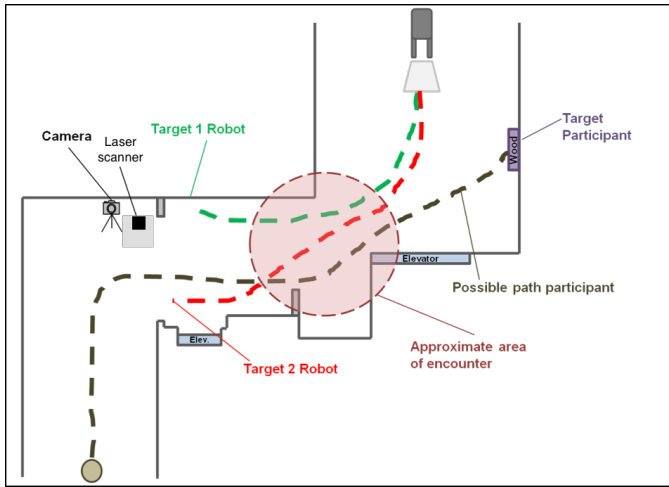
Fig. 4. The experimental setup used for the evaluations

that a triangle with a downward-pointing vertex is recognized more rapidly than the identical shape with an upward-pointing vertex. Also the work in [29] used arrows to indicate the intention of their robot and concluded that their system is intelligible. Furthermore, people are used to arrows indicating directions as in everyday life these are vastly used. So, to sum up, using an arrow to communicate the future path of a robot seems to be a good choice, as due to its angled v-shaped top it might attract the attention and might be detected faster than other symbols. Furthermore, it has already been used successfully and people already have a conception about the meaning of an arrow. This facilitates the understanding of the pattern and thus helps the humans to understand the intention of the robot faster. To see whether the projection of a pattern makes a difference, pattern C was the projection of white area. Finally, in condition D there was no projection at all.

In our experiment 17 persons participated (8 female, age: M = 28.8, SD = 6.08). Out of these, 13 stated Swedish as their nationality. The participants had a wide variety of backgrounds with four persons stating that they have experience with robots. Four persons reported being left handed, the remaining twelve stated to be right handed. For each condition, every participant had one encounter with the robot resulting in a total of four trials per participant. To avoid measuring learning effects, the order of the conditions per participant was varied according to a balanced latin square thus resulting in four different groups (ABDC, BCAD, CDBA, DACB).

The setting of the experiment is shown in Fig. 4. The task of the participants was to reach a wooden object placed at the end of the corridor. The setting was chosen due to its spatial constraint which resulted in a tight encounter of the participant and the robot. Therefore, it was necessary to decide at which side to pass the robot. Furthermore, the path of the robot was varied randomly between two different paths, in order to reduce the possibility of a learning effect. The

experimental procedure was as follows: first the participants were greeted, told about the experiment and asked to fill out a general questionnaire as well as to sign the consent form. Afterwards the standing robot without projection was shown to the participants in order to familiarize them with the platform. This was done in an attempt to make the first trial more comparable to the following trials. As a next step, the task was explained to the participants. They were told that after setup of the eye tracker, their task will be to reach and pick up the wooden object and that during this task they will encounter the robot which is on its way to accomplish a task as well. Finally, they were informed that they were to repeat this task four times and that the robot's paths and behaviors will vary across the trials.

After the participant confirmed that he/she understood the task, the eye-tracking goggles were set up. Here, the eye cameras needed to be adjusted such that the pupil was robustly detected. Furthermore, the eye tracker needed to be calibrated using manual marker calibration [25]. Then, the recording was started, a laptop used to record the eye-tracking data was put into a backpack carried by the participant who started to walk at a given signal corresponding to the click sound of the release of the break of the forklift. After the participant had picked up the wooden object and put it back to its place he/she was told to come back and the recording was stopped by the experimenter. Meanwhile, the robot drove back to its start position. This procedure was repeated four times. After all trials commenced, the participant was told that he/she could take off the eye-tracking headset and was then ask to come in another room for a stimulated-recall interview. Hereby the videos with the gaze overlay were shown to the participants and they were asked to explain their gaze and to comment on the experiment. The whole experiment took approximately 40 minutes, with the instruction and set up part taking about $10 - 15$ minutes, the trials taking about 15 minutes and the interview taking $10 - 15$ minutes.

V. DATA ACQUISITION AND EVALUATION

A. Eye-tracking Data

The fixations were extracted using the Pupil Player software, which uses a dispersion algorithm. The fixations were extracted following the recommendation of Blignaut [30] to use the values suggested by Holmqvist et al. [11]: a minimum duration threshold of 150 ms and a dispersion threshold of $1°$.

Due to the setting of the experiment the fiducial markers that were attached to the robot were occluded for a substantial time of the experiment in which the projection was already visible. Thus it was not possible to use an automated categorization of the fixations but it had to be done manually. Therefore each fixation that was categorized as such by the dispersion algorithm was classified according to the place where it was measured. To this end, the videos with the gaze overlay were analyzed and it was decided whether a fixation either was on an Area of Interest (AOI) on the robot (AOI-R), on the area where the projection was or would be in the no projection
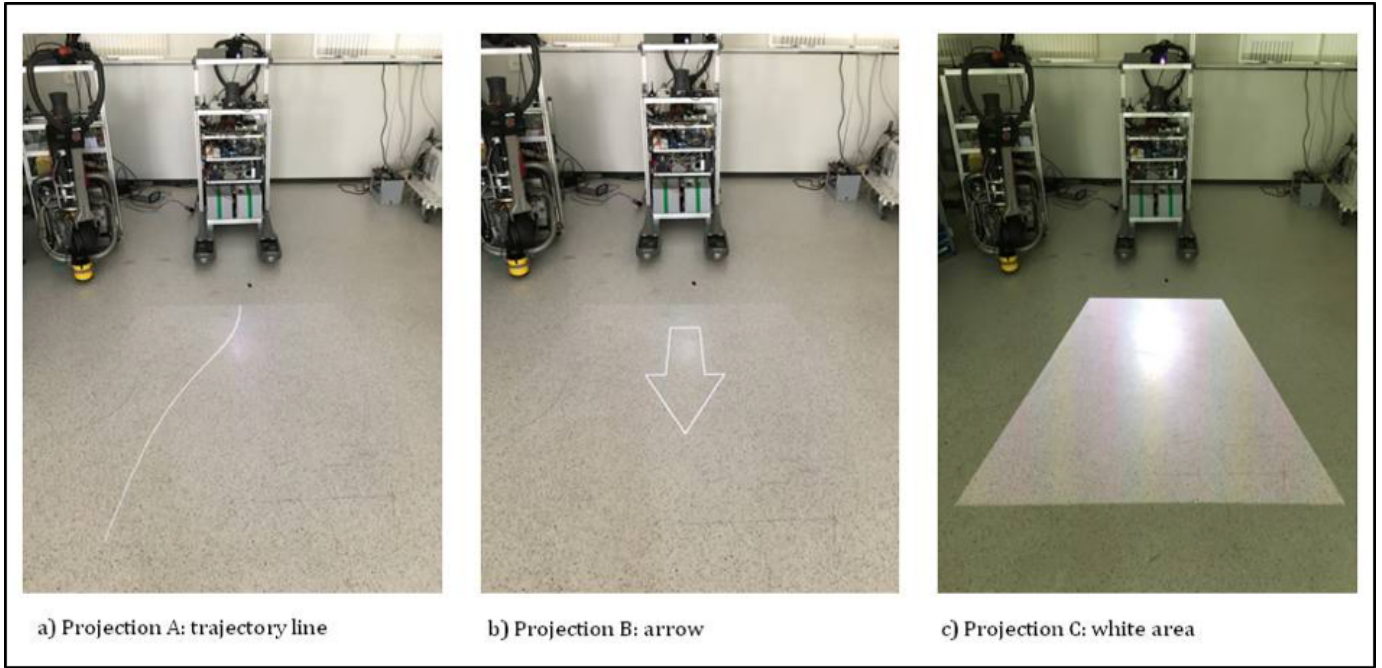
Fig. 5. The different intention communication modes: In total 4 types of intention communication were tested, three of which are depicted above. The forth mode is without projection.

condition (AOI-P) or somewhere else (see Fig. 6). All fixations that belonged to somewhere else were ignored.

The classified fixations were then used to calculate our dependent variables for each trial: the total fixation duration and average fixation duration of all fixations in an AOI. To further investigate the gaze pattern of the participant and whether it changed depending on the patterns or the number of trials done, another dependent variable was computed: the number of times the AOI was fixated first over all trials for both AOI. Here, the idea was that whether the projection area or the robot is first fixated might change due to the projected pattern or also due to the number of trials the participant already completed. If this would be the case, it would mean that the gaze pattern of the participant, at the beginning of the encounter, was influenced.

First, in order to determine whether the AOI and the pattern have an influence on the dependent variables, an analysis of variance (ANOVA) was performed. The goal of running the ANOVA was to see whether the projected pattern, as well as the location (robot vs. projection area), have an influence on the total and average duration of fixations. This was done to answer the question whether the projection of patterns influences the gaze behavior of the participants.

To make sure that the observed significant difference was not caused by learning effects or an insufficient counterbalancing for each dependent variable, there were two more ANOVA performed. In one, the number of the trial for the participant $(1 - 4)$ was used as a within factor, to make sure that the position of the trial did not have an influence on the dependent variable. To test whether the counterbalancing was successful,

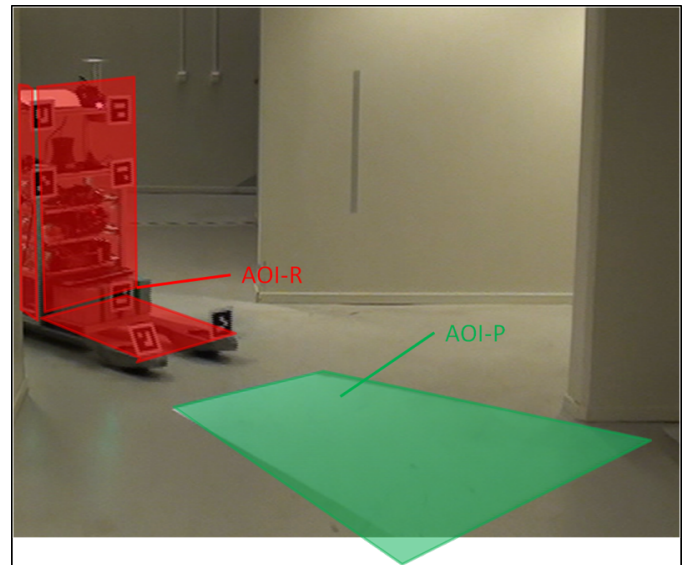Fig. 6. The defined areas of interest: AOI-R represents robot, AOI-P represent the projected surface

each participant was assigned to a group $(1 - 4)$ depending on which order of patterns he/she had in the experiment. This factor was used as a between factor in the ANOVA with the pattern and the AOI being a within factor. To analyze the number of first fixations on the two AOI a Chi-squared test was conducted.

B. Laserscan Data

The data that was acquired through the laser scanner needed to be processed and filtered before an analysis was possible. Out of the 17 participants the path data of 14 could be used for the analysis, as for two participants the measurement was not complete and for one participant the quality of the data was not good enough to extract the path reliably. An important point about the path data is, that due to the different walking speeds of the participants the encounter with the robot was slightly different in every trial and thus the comparison of the path was done relative to the time before the encounter with the robot. Using standard ROS tools, the laser data was extracted in form of the trajectories of the human and the robot and then used to compute the dependent variables for the statistical analysis.

The dependent variables used for the evaluation of the laser scan data are the average speed and the maximum deviation from the average x-value in the three seconds before the encounter as well as the minimal distance between human and robot. These were chosen to consider different aspects of the path and the human behavior. First the average speed in the short interval before the encounter could vary between the different conditions due to a different perception the participants could get from the robot, or also due to a varying certainty on what to do. This certainty could also be reflected in the shape of the path, which should be possible to see in the maximum deviation from the average x-value. For the participant the target was located approximately straight ahead in y-direction. As people prefer to take straight paths, the preferred way without the robot as an obstacle would therefore be a straight line, with coordinates only varying in y- and only to a small extent in x- direction. If the participant now has to veer off or adjusts his/her path, the maximum deviation of the x-coordinate from the average coordinates in that period of time reflects the straight path deviation. Finally, the minimal distance shows how much space the human kept between himself and the robot. Variations in this quantity depending on the projected path could reflect a different perception the human has of the robot.

VI. Results and Discussion

A. Eye-tracking

The two-way repeated-measures ANOVA for the total fixation duration revealed a main effect of the AOI, $F(1, 16) = 28.68$, p ¡ .001 as well as an interaction effect between AOI and pattern, $F(3, 48) = 4.25$, p = .010, which means that the projected pattern and AOI had an influence on the total fixation duration and average fixation duration. Both of these were also found in the two-way repeated-measures ANOVA for the average fixation duration (main effect AOI: $F(1, 16) = 11.74$, p = .003 and the interaction effect between AOI and pattern: $F(3, 48) = 3.62$, p = .041).

There were additionally two ANOVA computed for both dependent variables. These were first a two-way repeated-measures ANOVA with the number of the trial as an independent variable, to determine whether there was a learning

effect independent of the order of the projections. The second was a three-way mixed ANOVA with AOI and pattern as a within-factor and the group the participant belonged to due to the counterbalancing as a between factor. This was used to control whether the counterbalancing was effective. For both dependent variables there was no significant main or interaction effect of the trial number. Furthermore, in the three-way mixed ANOVA no significant interaction between group and AOI or pattern was found for both dependent variables. Thus, the statistical analysis verifies the effectiveness of the counterbalancing and no learning effect was measured.

As the interaction effect cannot be interpreted directly, due to the fact that four different patterns were compared, post-hoc tests for this effect were necessary to determine where the significance stems from. Paired t-tests were performed and to control for the multiple comparisons that were made, the p-values were adjusted using the method described by [31] For the post-hoc test the data was divided into two datasets, one for each AOI. Then, paired t-tests to compare the different patterns with each other were performed for each AOI and for both dependent variables. For AOI-R, the area of interest on the robot, the three projections were compared with each other. This was done to investigate how different projection patterns impact the gazes directed towards the robot. The post-hoc t-test showed that the total fixation duration spent by the subjects on projection A is significantly longer than in the case of projection C, $i.e.$, the subjects looked at the robot significantly longer when its trajectory was displayed opposed to only displaying an white area. There was no significant difference found regarding the average fixation duration.

A possible explanation for this results could be that the white light was very visible and drew a lot of attention, whereas the line segment visualizing the trajectory was not that visible. Actually, in the retrospective interviews, several participants reported that they saw the line segment very late or not at all. One participant even reported that she felt fooled because she saw the line only shortly before she had to pass the robot. Thus, it is very likely that in the line projection mode participants only used the robot itself as an indication to determine where to go and how to react. As proposed by [11], the distribution of attention between different targets can be measured by the total fixation duration over an AOI and thus a difference here might reflect a change in the distribution of attention between the two AOI's. In our case, during white area projection mode (C) the fixations on the robot were shorter than during trajectory projection mode (A) which might indicate that the attention was shifted from the robot to the projection area.

Regarding AOI-P, the area of interest on the projection, all six possible pairwise t-test comparisons were carried out per dependent variable. The post-hoc tests revealed that the total fixation duration spent on projection D (no projection) is significantly lower than that of projections B and C.

It can be clearly seen that the white light projection is the one that caused the greatest change, as it changed the behavior for both dependent variables. An explanation for this

effect is offered by the visibility of the patterns. From all participants, only two did report in the interview that they had not seen the white light, 6 reported that they had not seen the arrow and 10 could not remember seeing the trajectory line. Thus, the visibility, or at least the attention-drawing abilities of the projection, might vary and explain why the projections differed in the fixations that were spent on them. This is also might explain why projection C differed most from the other projections.

The Chi-squared test performed between the two AOI had a significant difference between the numbers of first fixations. In general, AOI-R was fixated first more often than the AOI-P. However, in the case of projection C, it was the opposite. So the projection of the pattern (at least in the case of C) does change the gaze pattern and thus it might help the participant if useful information is displayed in an easily understandable way. The Chi-squared test with respect to the trial number for each participant had no significance, which means that the gaze pattern at the first fixation doesn't change depending on the number of trials the subject already participated in.

B. Path Data

For the path data, we carried out an analysis similar to the previously described eye-tracking data. The one-way ANOVA with the projection pattern as independent variable did not reveal any statistically significant main effect for the respective pattern independent of the choice of dependent variable. However, the one-way ANOVA with the trial number as independent variable showed a significant main effect on the minimal distance ($F(3, 39) = 3.16$, $p = .035$), as well as on the maximal deviation of the mean x-coordinate in the 3 seconds before the encounter ($F(3, 39) = 3.52$, p = .024). For both of these variables, a two-way mixed ANOVA to control for the counterbalancing did not show a significant interaction between the trial number and the group.

The projection pattern had no significant main effect on any of the dependent variables. However, there was a significant effect of the trial number on the minimal distance and maximal deviation from the average x-coordinate in the 3 seconds before the encounter. For these main effects post-hoc tests were computed which, after correction, were not significant. However, maximal deviation showed a trend to significance for the comparison between trials 2 and 3. The maximal deviation in the second trial is on average smaller than in the third trial.

The Chi-squared test showed a significant result regarding the total number of subjects evading the robot on the right side compared to the total number choosing the left side independent of the projection pattern or the trial number. However, there was no significant result if the data was further divided into the patterns or the trial number. There was a visible trend for the comparison of the second to the forth trial regarding the minimal distance to the robot. On average, this distance is higher in the second trial than in the fourth trial. Although there was no significant difference found for the speed and the veer-off distance, the average value of these variables might still be interesting to interpret, *e. g.*, in terms

of proxemics theory [32] for the distance. Over all participants, the computed average of the mean speed in the last 3 seconds before the encounter with the robot was 1.09 m/s, the average veer-off distance was 2.11 m, whereas the average minimal distance was 0.79 m.

VII. CONCLUSION

In this work we investigated robot intention communication via projections in industrial scenarios. To this end, we conducted experiments using eye tracking and recorded path data to verify how intention communication effects the human behavior. Retrospective recall interviews were conducted to check the results obtained from the quantitative data. The results respectively obtained from the eye tracking and the path data differ in various aspects. In the chosen experiment scenario, the projection pattern had an influence on the gaze of the participants which differed for the different projections. However, the projection did not influence the path of the participants. The question here is whether the participants did not need the information of the pattern to decide on the path or if the communicated information was unsuitable. As many participants reported that they did not see the arrow or the line projections, it is more probable that they could not interpret the information in a meaningful way, the white projection probably did not give enough information. This highlights the importance of designing the projected patterns in a way that they are clearly visible and easy to interpret. With our combination of eye tracking and path data analysis, it becomes evident that the projection did influence the participants attention but not enough to change their walking behavior in the chosen scenario. This is in contrast to our earlier results which could be due to a change in the design of experiments, which is a more constrained space that affected the exposure time of the subjects to the projections and change in the visibility conditions due to the varying lighting conditions. Thus, the design of the patterns needs to be improved to increase their benefit. Also, we noticed that retrospective interviews give valuable insights as well. Without these interviews, there would be no way to find out which participant had noticed which pattern, as a gaze on the pattern does not automatically mean that it was processed as well. To summarize, the used measurements yielded different results. Regarding the eye tracking data, it was found that the projections changed the gaze patterns with respect to the first fixation, as well as for duration of the fixations on the areas of interest. Projecting a white rectangle evoked the largest changes in human movement and gaze behavior. Possible explanations for these results were found through the conducted retrospective interviews.

References

[1] R. T. Chadalavada, H. Andreasson, R. Krug, and A. J. Lilienthal, "Thats on my mind! robot to human intention communication through on-board projection on shared floor space," in *Proc. of the European Conference on Mobile Robots*, 2015, pp. 1–6.

[2] ——, "Empirical evaluation of human trust in an expressive mobile robot," in *Robotics Science and Systems (RSS), Workshop on Social trust in Autonomous Robots*, 2016.

[3] A. Watanabe, T. Ikeda, Y. Morales, K. Shinozawa, T. Miyashita, and N. Hagita, "Communicating robotic navigational intentions," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2015, pp. 5763–5769.

[4] L. Takayama, D. Dooley, and W. Ju, "Expressing thought: improving robot readability with animation principles," in *Proceedings of the 6th international conference on Human-robot interaction*. ACM, 2011, pp. 69–76.

[5] A. Turnwald, D. Althoff, D. Wollherr, and M. Buss, "Understanding human avoidance behavior: interaction-aware decision making based on game theory," *International Journal of Social Robotics*, vol. 8, no. 2, pp. 331–351, 2016.

[6] C. Lichtenthäler, "Legibility of robot behavior: Investigating legibility of robot navigation in human-robot path crossing scenarios," Ph.D. dissertation, München, Technische Universität München, Diss., 2014, 2014.

[7] R. Mangold, *Informationspsychologie: Wahrnehmen und Gestalten in der Medienwelt*. Springer-Verlag, 2015.

[8] K. Matsumaru, "Mobile robot with preliminary-announcement and display function of forthcoming motion using projection equipment," in *Proc. of the IEEE International Symposium on Robot and Human Interactive Communication*, 2006, pp. 443–450.

[9] M. D. Coovert, T. Lee, I. Shindev, and Y. Sun, "Spatial augmented reality as a method for a mobile robot to communicate intended movement," *Computers in Human Behavior*, vol. 34, pp. 241–248, 2014.

[10] A. D. May, C. Dondrup, and M. Hanheide, "Show me your moves! conveying navigation intention of a mobile robot to humans," in *Proc. of the European Conference on Mobile Robots*. IEEE, 2015, pp. 1–6.

[11] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[12] M. A. Just and P. A. Carpenter, "Eye fixations and cognitive processes," *Cognitive psychology*, vol. 8, no. 4, pp. 441–480, 1976.

[13] ——, "A theory of reading: from eye fixations to comprehension." *Psychological review*, vol. 87, no. 4, p. 329, 1980.

[14] A. L. Yarbus, *Eye movements during perception of complex objects*. Springer, 1967.

[15] D. Baldauf and H. Deubel, "Attentional landscapes in reaching and grasping," *Vision research*, vol. 50, no. 11, pp. 999–1013, 2010.

[16] M. M. Hayhoe, A. Shrivastava, R. Mruczek, and J. B. Pelz, "Visual memory and motor planning in a natural task," *Journal of vision*, vol. 3, no. 1, pp. 6–6, 2003.

[17] N. Mennie, M. Hayhoe, and B. Sullivan, "Look-ahead fixations: anticipatory eye movements in natural tasks," *Experimental Brain Research*, vol. 179, no. 3, pp. 427–442, 2007.

[18] J. M. Findlay and I. D. Gilchrist, *Active vision: The psychology of looking and seeing*. Oxford University Press, 2003, no. 37.

[19] C. Antonya, F. Barbuceanu, Z. Rusák, D. Talaba, S. Butnariu, and H. Erdélyi, "Obstacle avoidance în simulated environment using eye tracking technologies," in *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2009, pp. 1581–1590.

[20] J. Jovancevic-Misic and M. Hayhoe, "Adaptive gaze control in natural environments," *The Journal of Neuroscience*, vol. 29, no. 19, pp. 6234–6238, 2009.

[21] M. M. Hayhoe and C. A. Rothkopf, "Vision in the natural world," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 2, no. 2, pp. 158–166, 2011.

[22] A. E. Patla and J. N. Vickers, "Where and when do we look as we approach and step over an obstacle in the travel path?" *Neuroreport*, vol. 8, no. 17, pp. 3661–3665, 1997.

[23] J. P. Hansen, "The use of eye mark recordings to support verbal retrospection in software testing," *Acta Psychologica*, vol. 76, no. 1, pp. 31–49, 1991.

[24] Z. Guan, S. Lee, E. Cuddihy, and J. Ramey, "The validity of the stimulated retrospective think-aloud method as measured by eye tracking," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 1253–1262.

[25] M. Kassner, W. Patera, and A. Bulling, "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction," in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication*. ACM, 2014, pp. 1151–1160.

[26] M. Bertamini, L. Palumbo, T. N. Gheorghes, and M. Galatsidas, "Do observers like curvature or do they dislike angularity?" *British Journal of Psychology*, vol. 107, no. 1, pp. 154–178, 2016.

[27] M. Bar and M. Neta, "Visual elements of subjective preference modulate amygdala activation," *Neuropsychologia*, vol. 45, no. 10, pp. 2191–2200, 2007.

[28] C. L. Larson, J. Aronoff, I. C. Sarinopoulos, and D. C. Zhu, "Recognizing threat: A simple geometric shape activates neural circuitry for threat detection," *Journal of cognitive neuroscience*, vol. 21, no. 8, pp. 1523–1535, 2009.

[29] D. Matsui, T. Minato, K. F. MacDorman, and H. Ishiguro, "Generating natural motion in an android by mapping human motion," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2005, pp. 3301–3308.

[30] P. Blignaut, "Fixation identification: The optimum threshold for a dispersion algorithm," *Attention, Perception, & Psychophysics*, vol. 71, no. 4, pp. 881–895, 2009.

[31] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian journal of statistics*, pp. 65–70, 1979.

[32] E. T. Hall, *The hidden dimension*. Doubleday & Co, 1966.