Spatial Settings In Social Navigation Using Deep Q-Learning

Marcell Balogh
Technical Faculty of IT and Design
Department of Architecture, Design, and Media Technology
Aalborg University, Rendsburggade 14 9000 Aalborg, Denmark
mbalog14@student.aau.dk

***Abstract*— The objective of this paper was to investigate whether following from the back or walking side by side is preferred in social navigation. The agent was trained by a trial-and-error learning approach using a Deep Q-Network. The experiment was driven by semi-structured interviews, video based observations and Godspeed surveys. Additionally, the participants were split into two independent groups each of which presented with one of the mentioned spatial settings. The results showed that both settings equally maximize user experience and meet with their requirements. However, following from behind is slightly favored in term of likeability. Moreover, the preference of each settings is highly depending on the context of robot guidance.**

***Index Terms*—Robot-Guidance, Deep Q-Learning, Spatial Settings, Human-Robot Interaction**

I. Introduction

Recently, the use and application of robotics has significantly increased in different fields. Social robotics is one that has captured the attention of many researchers who have proposed various applications such as cooperative or assistive robots. Some researchers have been convinced that human-following could be one of essential way for robots to acquire intelligence. Hence, several human-following robots have recently been developed [1] [2] [3] [4] with the hope that human-following behavior might be efficient for mobile robots in order to train them effectively in unknown and dynamic environments or to cooperate with humans [5]. However, these applications are rather putting too much emphasis on their technical capabilities than social aspects of potential human-robot interactions which would be more important in real life scenarios. In addition, human-following behavior tends to be ineffective in itself and cannot be widely used in social robotics, since it makes robot mobility dependent on the person instead of endowed with the required intelligence. Conversely, this paper addresses the topic of robot tour guidance where human and robot switch places and the human starts following the robots instead and developing an interaction for social navigation where the robot adapts human behavior on human-like basis via trial-and-error learning. This paper, in addition, delineates towards experimental investigation of the hypothesis that different human-robot spatial settings in social navigation task might alter the experience and perception of that interaction, furthermore, make people prefer one than another.

II. Related Work

A. Robot Guidance

In recent years, a fair share of research have been conducted with the aim of developing mobile robots that are able to guide humans to a destination. In studies like [6] [7] [8] [9] various mobile applications were implemented in order to lead people in bounded environments, such as hospitals or museums without neglecting even the social aspect. A paper reports about an autonomous robot interaction with museum visitors where a humanoid robot guided people around several exhibitions while giving explanation of the attractions. The proposal also highlights the practical importance of involving humanoid robot, rather than e.g. a cylinder shaped one, because its physical structure enables it to interact with people by using human-like body movements [9]. Another reference from [10] represents multimodal interaction between humans and a tour guide (humanoid) robot additionally underlines great potential for investigating intuitive natural interaction with humans rather than involved navigation task only. Studies like [11] and [12] focus on how mobile guide-robots should take social acceptance into account when constructing safe path for itself and list major human aware capabilities that a robot could exhibit during navigation. They also emphasize, inter alia, that the robot should ensure safety, and have reliable and effective motion, in addition, the machine should not get too far into the personal space of the human. Consequently, robots should respect personal distances to be more socially acceptable. Hall introduced the science of proxemics, in his book [13], to demonstrate how the use of space can affect personal and business relations, cross-cultural interactions, architecture, city planning, and urban renewal. Figure 1 represents the four different personal distances, introduced and studied by Hall. On top of considering personal distances, another

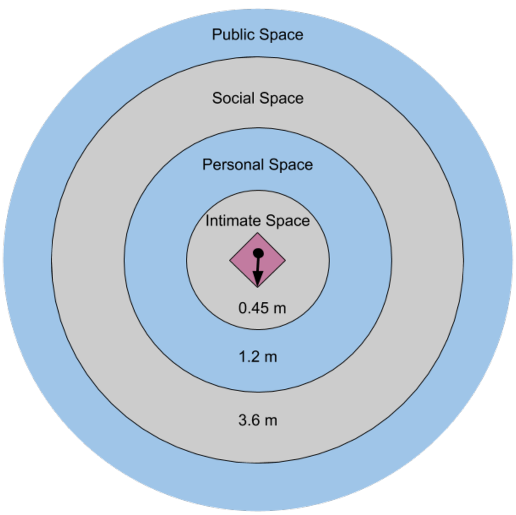


Fig. 1: The four different personal distances defined by Edward T. Hall.

significant matter of interest or importance is the way in which a human is placing themselves in relation to the robot. The study in [14] introduces a preference for walking on the left/right side of the guiding robot not only as following from behind, moreover it also informs that there is a desire to reach the goal quickly with robot guidance. Nevertheless, not all humans track the guiding robots in the same speed, for example elderly [7] or disabled people [15] experience difficulties walking, understandably so leading to a desire for speed adaption. In [16] these crucial aspects (personal zones, switching

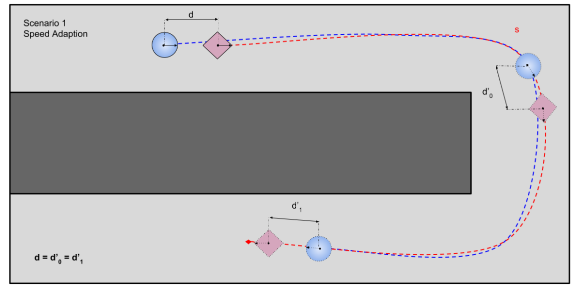


Fig. 2: Speed adaption (human is represented by blue circle while the robot with red square).

dispositions and speed adaption) are all assessed and it devises a solution in the form of a so-called switching strategy which is based on a bilateral interaction between a human and a mobile robot. This control strategy enables the robot to guide a human while it follows a specified path and maintains the desired distance between them (Figure 2). Furthermore, it detects human intentions such as leaving the task and changing the interaction zones (Figure 3). Another study [17], with similar

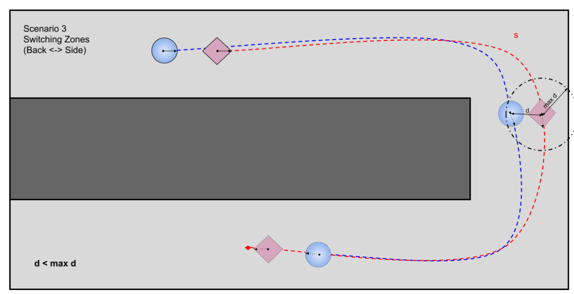


Fig. 3: Switching zones from back to side and vice versa (human is represented by blue circle while the robot with red square).

considerations in mind, implemented a mobile robot with *"Hierarchical Mixed Observability Markov Decision Processes"* that indeed emphasized speed adaption, robot dis-positioning behavior relative to humans and human's task leaving affinity. However, it also touched upon the proactive sideways of the application as the robot could arbitrarily choose to change pace by decelerating or accelerating itself.

B. Research Questions

Many of the above mentioned studies developed and presented assistive mobile robot applications that can be crucial in social robotics, especially in robot guidance. These papers describe social aspects both in the sense of robot behavior and desired physical form of the ideal social mobile robot, while in addition looking for the ideal solution that makes robot guidance more socially acceptable and easier for people with reduced mobility. Some studies pay attention to spatial behaviors such as e.g. understanding the effects of body orientation and gaze in group conversations, but very small amount have made spatial settings in social navigation a focus of attention. In addition, the research area is lacking in studies of dynamic spatial settings. There is also a deficiency in regards to what people think, feel and perceive socially about their spatial settings (orientation relative to a mobile robot) in human-robot interaction and since such reciprocal actions require a certain flexibility and intelligence, this study aims to investigate:

1) *"how social navigation behavior can be learned by a robot without being explicitly programmed"*
2) and *"how such learned behavior affects user experience"*.

Thereby, expected to gain a better understanding of how mobile robots guide a person to a destination in a socially acceptable way and improve the user experience by taking into account human preferred spatial settings: following from the back or one of the sides.

III. Deep Q-learning

The human brain is able to generate guidance strategies even in unknown environments and simultaneously control the essential body kinetics regarding to another person's behavior, and additionally to improve performance over multiple trials. Such capabilities are challenging to implement in autonomous systems. In order to endow Pepper robot (subsection IV-A) with human learning capabilities for social navigation behavior, reinforcement learning (RL) was applied. RL is a type of machine learning where an agent learns how to behave in an environment by interacting with it and receiving certain feedback as rewards or penalties for performing particular actions, which gives the agent a policy (π).

A. State and Action Sets

In this project, the environment is a linear pathway with neither dynamic nor non-dynamic obstacles. The length of this corridor was specified by the range of WIFI router in meters, that approximately is 20m. In fact, the building where the experiment has taken place, has WIFI coverage on all of its premises, however, due to security and authorization issues, a separated and individual WIFI router has been set up for wireless communication with the agent. That has affected and limited the range of the experimental space.

The agent, correspondingly to the given task which is guiding a person to a destination, has the capability to play two actions: move forward and stop. There are two actions only due to the speed limitation of the agent (described in subsection IV-A2), since the low speed by default obstructs to have an appropriate speed interval or continuous output space.

There are two scenarios or spatial settings: a subject follow the agent from behind (scenario A) or from the desired side (scenario B). The state of the robot at each time step is defined in its velocity (max/min), distance (m) between a person and the robot itself, and by the orientation (number of segments of the laser scanner and its length in meters) of the person relative to the robot. On that basis, the agent must be rewarded if it keeps the appropriate distance and orientation, else penalized, with other words, negatively rewarded.

B. Reward Function

RL is based on the idea of the reward (R) hypothesis, that means to have the best behavior, the agent selects actions with the goal of maximizing the expected sum of reward, where outcomes received in the far future are discounted compared with outcomes received more immediately (discount factor = γ). Consequently, cautious reward design (advised definition of evaluation metrics of taking an action in a given state) is crucial to avoid convergence to a local optima or a positive feedback loop. Since the task is to navigate a person to a destination, the robot must know whether that person moves or stops (with a faster robot could be the velocity of that person) in both scenarios. In scenario A, the agent is negatively rewarded if its relative distance to the follower human is greater than the determined threshold

($T = 1m$) while its velocity is at its maximum, else positively rewarded. The threshold value corresponds to the range of human personal space. In scenario B, the agent is positively rewarded if the robot cannot detect anything within the threshold behind its back but the tactile sensors on its head are activated.

C. The Algorithm

RL agents can be divided into two classes, model-based and model-free. This implementation focuses on the model-free approach using off-policy Temporal Differences (TD) (Eq.2) control algorithm known as Q-learning, defined by

$$Q_t(s,a) := Q_{t-1}(s,a) + \alpha TD_t(s,a) \tag{1}$$

where α is the probability of taking random actions (exploration) while 1-α is probability of taking actions from learned experience (exploitation).

$$TD(s,a) = R_{t+1} + \gamma \max_a Q(s_{t+1},a) - Q_{t-1}(s,a) \tag{2}$$

The advantage of model-free agents is that it does not require a model. Generally there is no complete model of the environment, so the effectiveness of model-based algorithms is dependent on their ability to estimate the model well. Since Pepper's sensory (with laser scanners and sonar) is not able to precisely represent its surroundings and the environment is stochastic, model-free RL was chosen. However, the disadvantage of these algorithms is that more trial-and-error experience is required to turn the values into good estimates of future consequences. Thus, even though they are less computationally expensive, it takes more time to learn the optimal policy for these algorithms. An alternative for TD is called Monte Carlo approach, however, it is an offline method, which performs an update only after the episode is finished and since the defined task is continuous, TD was chosen. An issue with Q-learning is that the combination of all possible configurations stored in the Q-table. This would make the table huge and impossible to fit in todays computer memory, if action and state spaces are enormous. Fortunately, there is a solution for this: replacing the Q-table with a neural network, which would tell the agent what the optimal action is in each state.

Deep Q-Learning consists of combining Q-Learning to an Artificial Neural Network (ANN). Inputs are encoded vectors, each one defining a state (in this case 32 sensor inputs) of the environment. These inputs go to an ANN, where the output is the action to play. More precisely, in this implementation, there are two possible actions. The output layer of the neural network is comprised of two output neurons, each one corresponding to the Q-values of each action played in the current state. Then the action played is the one associated with the output neuron that returned by the softmax method. Hence, in each state the prediction is the Q-value:

$$Q(a_t, s_t) \tag{3}$$

where a is chosen by softmax (Eq.7), the target is:

$$R_t + \gamma \max_a Q(a, s_{t+1}) \tag{4}$$

and the loss error is:

$$Loss = \frac{1}{2} R_t + \gamma \max_a (Q(a, s_{t+1})) Q(a_t, s_t)^2 \tag{5}$$

Then this loss error is backpropagated into the network, and the weights are updated according to how much they contributed to the error. The problem with this is that s_t is most of the time very correlated with s_{t+1}, therefore the network is not learning much. This is improved if, instead of considering only one previous transition, the last m transitions are considered where m is a large number. This pack of the last m transitions is what is called the Experience Replay (ER). Then from this ER some random batches B of transitions M are taken to make the updates. In summery, the Q-network is trained with the samples randomly drawn from the replay memory to minimize the correlations between samples.

D. Learning Performance

The agent was trained in the same manner for both scenarios: let the robot drive itself on the designated path and reward or penalize for moving forward or stopping while a human follow it from the back or one of the sides. The training was terminated when the agent reached the maximum average reward. The agent's hyper-parameter values are shown in Table I. Evidently, the training time of the two scenarios are not identical due to the difference in the number of sensor inputs. Since in scenario B, the agent was

Algorithm 1 Deep Q-Network

Initialization:

1) For all couples of actions a and states s, the Q-values are initialized to 0:

$$\forall_a \in A, s \in S, Q_0(a, s) = 0 \tag{6}$$

2) The ER is initialized to an empty list M.
3) Start in the initial state s_0. Play a random action and reach the first state s_1.

At each time t $\geq$ 1:

1) With probability α select a random action a_t otherwise play the action a_t, where a_t is the action with the highest preference in each state is given the highest probability of being selected from the exponential softmax distribution:

$$a_t \simeq \pi(a|s) = \frac{\exp(Q(s,a))^\tau}{\sum_{a'} \exp(Q(s,a'))^\tau}, \tau \geq 0 \tag{7}$$

2) Get the reward r_t = R(a_t, s_t)
3) Get into the next state s_{t+1}, where s_{t+1} is a random draw from the T(a_t,s_t,.) distribution.
4) Append the transition (s_t, a_t, r_t, s_{t+1}) into M.
5) Take a random batch $B \subset M$ of transitions.
6) For each transition (s_{tB}, a_{tB}, r_{tB}, s_{tB+1}) of the random batch B:
   - Get the prediction (Eq.3).
   - Get the target (Eq.4).
   - Get the loss (Eq.5).
7) Backpropagate the loss error and update the weights according to how much they contributed to the error.

fed with 30 times more distance measurements than in scenario A, the environment represented by input sensory became more stochastic, hence, complicating and decelerating the agent to learn the optimal policy. Figure 4 and 5 show the learning process of scenario A and B respectively.

E. Discussion

Even though Q-Learning performs well, other RL algorithms could be tried out and used for this problem for the sake of comparison. In addition, learning time could be improved by either using imitation learning methods where the model learns to imitate the actions of an expert (human) or as another option is to train the model in simulations. Simulated environment could give an ability to shape a more optimal reward function, however, the so-called reality gap between simulated and real-world environments often causes problems for algorithms to induce all the learnable behaviours. An alternative solution could be the utilization of motion capture where the reward system could precisely determine spatial arrangement and relation of the person or even group of people relative to agent. In addition, extracting and recognizing human legs pattern from the raw laser sensor signal could notably improve the performance of current implementation.

Hyper-parameter	Value
Discount Factor (γ)	0.9
Exploration Factor (α)	0.1
Learning rate	0.001
No. Input Units	32
No. Hidden Layers	30
No. Output Units	2
Size of ER Memory	100000
Temperature Parameter (τ)	100
Rewards	+1/-1

TABLE I: Hyper-parameter values of the Deep Q-Network

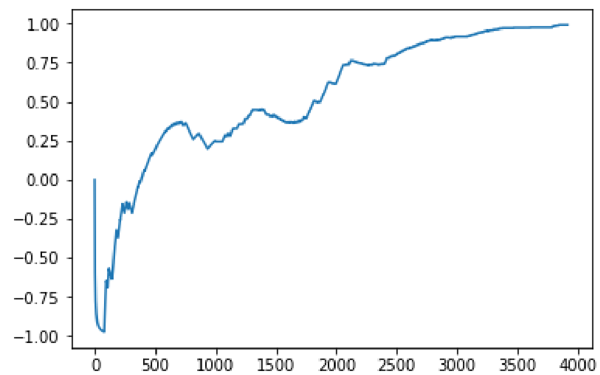


Fig. 4: The learning performance of scenario A, where the y axis stands for the average reward R and the x axis represents time step t. ($\approx$ 15 minutes)

IV. Methods and Materials

A. Robot Platform

1) Appearance: Pepper (Figure 6), with a height of 1.2 meters and a total of 20 degrees of freedom, is a semi-humanoid robot that was made especially

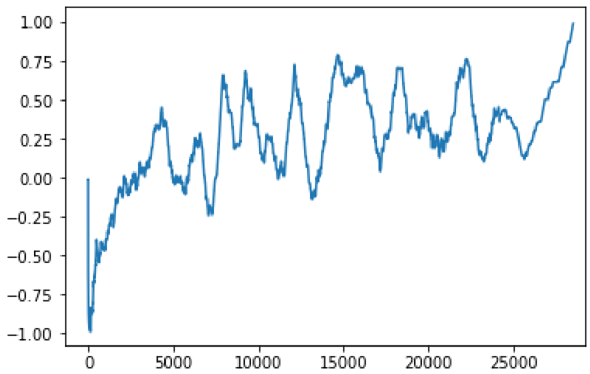


Fig. 5: The learning performance of scenario B, where the y axis stands for the average reward R and the x axis represents time step t. ($\approx$ 110 minutes)

for human-robot interaction. It was designed to be socially acceptable and have friendly looks. Reasonably, its appearance was highly influenced by a phenomenon so-called uncanny valley which describes relation between the human likeness of a robot, and humans affinity for it [18]. Therefore, Pepper is an appropriate choice for social navigation which supports all findings are discussed in the subsection II-A at the page 1.

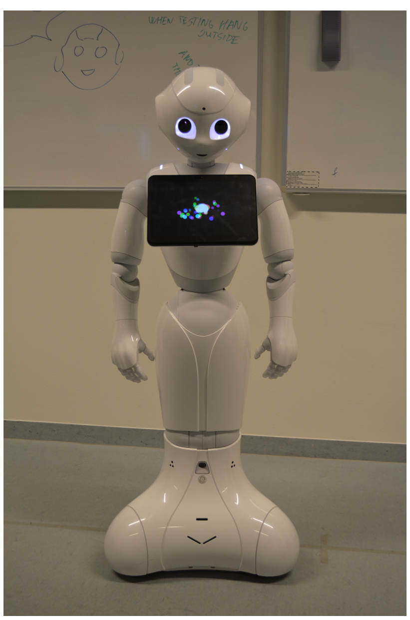


Fig. 6: Pepper, the human-shaped robot from Soft-Bank Robotics.

2) Mobility: Pepper is a holonomic mobile robot that can move easily to any direction using its three independent omnidirectional wheels that function together to generate motion. However, this capability can cause unpredictable motion of the robot when it goes forward in the form of turning slightly instead of following a straight trajectory. In addition, although the robot accelerates up to 3 km/h, the normal human walking speed is defined about 5 km/h, which is slightly exceeding Pepper's velocity. The low speed resulting in difficulty and unpleasantness while following the robot is discussed in subsection V-C at the page 9.

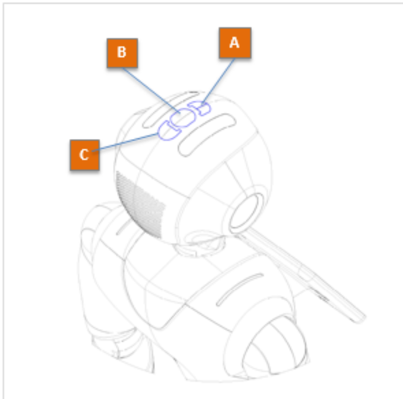
3) Perception: The robot has many highly valued sensors and actuators, those are well listed in Peppers documentation on its website [19]. In this project, the utilized sensors were: one rearward sonar, three lasers of which pointing horizontally in the robots surrounding area and three tactile sensors on the top of the head. Each laser and the sonar has a field-of-view of 60 degrees, and are non-overlapping, creating blind spots. Additionally, each laser provides only 15 points, making them insufcient for reflecting the surroundings precisely. Figure 7 illustrates the used sensor technology.

B. Qualitative Methods

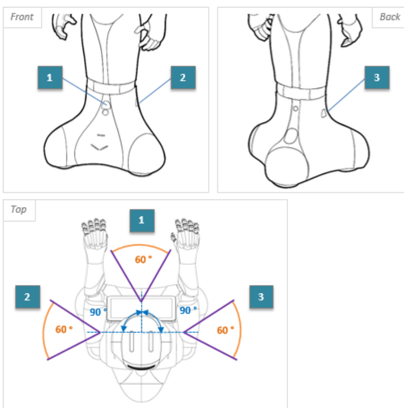
This research used both quantitative and qualitative methods. However, this subsection presents observatory techniques that attempt to explore *"meanings, concepts definitions, characteristics, metaphors, symbols, and description of things"* [20] via observation and semi-structured interviews. The goal of the observation was to understand people's spatial behavior with and perception of robots as a social phenomenon by using qualitative data collection methods. Ethnographic research uses camera to collect empirical data as a row form and the analysis of which is an interpretation of the meanings and functions of human actions and mainly takes the form of verbal descriptions or explanations [21]. Hence, despite the fact that the experiment was conducted in a controlled manner, Ethnographic camera was used as a tool to record the human-robot interactions and Grounded Theory to identify and integrate categories of meaning from the video data.

1) Ethnographic camera: From the "ethnographic toolbox for design", the method *"shadowing"* [22] was chosen for evaluating the performance of the developed system and to record the behavior patterns of the participants in connection with the robot. *"Shadowing"* produces material on the details of what the participant is doing and where his/her attention moves, in addition, it allows the test subject to control what can be videotaped.

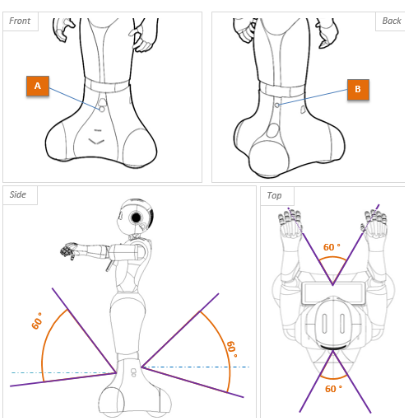
The position, point of interest and angle of the camera can influence if and how attention is drawn to the camera and video recording. Hence, "what" and "how" it is filmed depending on the role of video and its influence. For example, a test subject and his/her activity can be observed with the camera either as a proverbial *"fly on the wall"* or, on the other hand, enable the informants to reflect on their own practice and how that act might alter according to the camera (*"fly in the eye"*) [23]. Although, the choice of these concepts is often a dilemma, in this research, there is no difficult choice to be made between the above mentioned two opposed alternatives, since both concepts were applied in different role during the experiment. One action camera (Figure 8 represents the perspective of Cam1) amounted to robot represented the *"fly in the eye"* while a normal video camera (Cam2) used for *"shadowing"* embodied the *"fly on the wall"* metaphor.



(a) Head tactile sensors.



(b) Laser sensors.



(c) Sonar sensors.

Fig. 7: Type of sensors were used as inputs for Deep Q-Net. Source: http://doc.aldebaran.com/2-5/family/peppertechnical/indexpep.html

Fig. 8: The *"fly in the eye"* perspective of Cam1.

The awareness of being filmed and observed with Cam1 did not influence the participants due to the simplicity of their task, following the robot until it reaches its destination. However, it enabled their spatial behavior and even facial expressions to being observed more specifically and closely. On the other hand, Cam2 could roll the film quietly, without notice and capture from other point of view the equally important practice of the test subjects. The Ethnographically Informed Design methodology could have been better used if the context of the experiment was less controlled, however, its methods and tools gave a structured ability to record human-robot interaction progressively.

2) Semi-Structured Interview: By taking into account the ideas and opinions produced by the test subjects or occurring suddenly in their mind before and after human-robot interaction, interviews were conducted with each test subjects to explore their thoughts. More precisely semi-structured interviews which were guided by a script of six open questions but interesting issues were explored in more depth with following sub-questions.

C. Quantitative Methods

In order to compare and test statistically the hypothesis of the two said interaction (side walk vs. following from behind), a measurable quantitative method was applied. Godspeed survey is standardized measurement tool for human-robot interaction, more precisely, for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Trough these five categories, it consists of 28 survey items in total each attached with a semantic differential scale in the format of a typical five-level Likert item, within the higher score on the scale is the better.

V. Experiment

This paper discusses an experiment to study the relationship of humans and robots in two spatial settings in a term of social navigation. The objective of the experiment is to test whether from the back (scenario A) or walking side by side (scenario B) is preferred to follow a robot and how these settings might affect people's perception and experience of the interaction. This paper presents the participants and procedures for the experiment, the experiment's results, and an analysis of those results.

A. Participants and Ethics

In this research, there was no specific target group, thus the informants were randomly recruited adults from the age 20 to 40. In both scenario A and B, 15-15 test subjects participated respectably, of which 100% was male. The scenarios were tested on two distinct group of participants. Ethical issues associated with recording people and their speech, hence, permissions need to be obtained. Even if the participants gave their consent, there is no moral right for the researcher to use the captured or photographed data unrestrictedly. Thus, anonymity of the actor and speaker was ensured.

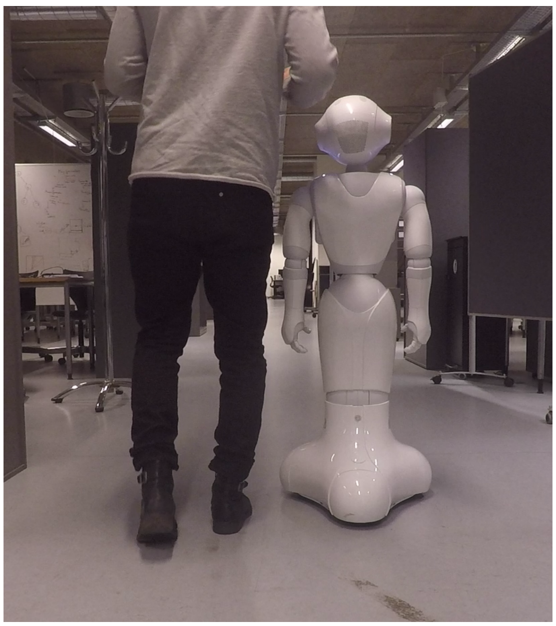
B. Procedure

Two independent tests were conducted to evaluate the defined spatial settings. The idea behind this was to test whether scenario A or B would actually maximize the users experience and preference of social navigation. In both cases, the participants were uninformed so-called "blind" about the experiment, with the only knowledge of being participating in human-robot interaction. Firstly, they were asked to sign the consent form and then answer three questions related to their previous or existing experiences with robots, about their expectation of the interaction and their conceptualization of robot guidance. Secondly, they were asked to approach and interact with Pepper:

1) *Scenario A:* Following Pepper from the back, apart from that, there is no other criterion.
2) *Scenario B:* Following Pepper from one of the sides, apart from that, there is no other criterion.

The *"no other criterion"* means they were allowed to change their speed or even stop for a while then start off again. The test was done when the robot reached the goal of min 18 to max 20 meters from the start point. After the interaction, firstly the participants were asked to fill out the Godspeed survey then further questions were investigated such as which of the four personal distances that they perceived between themselves and Pepper during the interaction, opinion about the overall speed of the robot and lastly about the experiment generally in itself. All participants were videotaped in a fashion described in subsection IV-B1 for later analysis.

(a) Scenario A.



(b) Scenario B.

Fig. 9: Snapshots of the two defined spatial settings from the experiment (Cam2).

C. Results and Discussion

The interviews underlined that the participants had no previous experience with physical robots at all apart from robotic toys. Furthermore, the majority expected some kind of verbal communication from human-robot interaction. People's mental model of robot guidance were clear and they interpreted easily the task. In order to specify statistically which of the two spatial settings maximized user experience and preference the following null-hypothesis was stated: H_0 = *There is no significant difference how people experience and perceive scenario A and B in terms of anthropomorphism, animacy, likeability, perceived intelligence and safety*. Hence, the 28 items of Godspeed survey were categorically compared between the two test groups (Table II) which led to that result: there is no significant difference how people experience and perceive scenario A and B in terms of anthropomorphism, animacy, perceived intelligence and safety (Table III). Hence the null-hypothesis is incorrect and cannot be rejected. However, in one case the null-hypotheses could be rejected and underlined that following from the back is more likeable than scenario B. Likeability was measured by how nice, pleasant, kind, friendly and likable the participants found the robot during interaction in each scenario. In regards to the interviews, the majority of participants described their spatial arrangements corresponding to the measures of the four different personal distances, introduced by Hall. For scenario A, 10 out of 15 participants found their positions in relation to the robot "social", while in scenario B, 13 participants reported it as "personal". About 70% of the interviewees of both scenarios indicated that Pepper's speed is slow while others

Groups	Obs.	Items	μ	$\sum$
Anthropomorphism				
A	75	5	2.96	222
B	75	5	2.893	217
Animacy				
A	90	6	3.01	271
B	90	6	2.93	264
Likeability				
A	75	5	3.893	292
B	75	5	3.64	273
Perceived Intelligence				
A	75	5	3.386	254
B	75	5	3.32	249
Perceived Safety (beginning)				
A	45	3	3.53	159
B	45	3	3.77	170
Perceived Safety (end)				
A	45	3	4.08	184
B	45	3	3.86	174

TABLE II: Group statistics represented by the number of observation based on the number of survey items, the mean and the sum of scores given in semantic differential scale in the format of a five-level Likert item.

Normality (Shapiro-Wilk)	Test	p-value
Anthropomorphism		
p-value = 1.459e-09	Wilcoxon	0.4087
Animacy		
p-value = 4.405e-10	Wilcoxon	0.4911
Likeability		
p-value = 4.524e-12	Wilcoxon	0.03077
Perceived Intelligence		
p-value = 6.334e-10	Wilcoxon	0.8874
Perceived Safety (beginning)		
p-value = 1.247e-06	Wilcoxon	0.216
Perceived Safety (end)		
p-value = 8.819e-10	Wilcoxon	0.3381

TABLE III: Test Statistics (confidence level = 0.95). The Shapiro-Wilk normality test tells whether the distribution of the sample is significantly different from a normal distribution (p-value<0.05) or not (p-value>0.05) and thereby help for determining the appropriate type of statistical test. Wilcoxon rank sum test (aka Mann-Whitney U-test) is used when is asked to compare the means of two groups that do not follow a normal distribution where p-value<0.05 rejects statistical equality of the means.

commonly replied: *"...it was slower than my walking speed but that is fine."* or *"...2-4 times slower...but in an exhibition I would not walk that fast."*. Results of the interviews highlighted that the preference of spatial settings in social navigation is highly dependent on the context. For examples, following from back, in a task of leading a person from A to B in an unknown environment for the subject, is preferred because of a better visibility of the followed trajectory set by the robot. On the other hand, walking side by side tends to be chosen as a preference for guidance involving verbal communication such as museum or exhibition tour guides. The video tapes documented that people only focused on the robot, their posture and gaze was constantly on Pepper excluding everything else. Half of the participants found the agent responsive enough while the other half reported latency. This could be also seen in the recordings, for example, when latency confuses and makes the test subjects uncertain in their movements. At the base of the videos, this latency ranged from about 0 to 1.5 seconds. The results might be biased by the small sample size, limited time of interaction and parameters (short length and lack of curve) of the test path not mentioning the exclusive male attendance.

VI. Conclusion

In this study the main question was whether a type of spatial settings (side or back) does affect the experience of people in relation with spatial navigation in human-robot interaction. Additionally there was a challenge to develop an agent that adapts human behavior on human-like basis via trial-and-error learning to conduct the experiment. Although the agent could generalize an optimal policy to determine when to stop or move, the current algorithm with human leg pattern recognition feature could be more accurate and reliable. The significant findings show that spatial settings are highly contextually dependent and the resulted preference for spatial settings was too heterogeneous, thus statistically there is no significant difference in people's experience and perception of their spatial relation to a robot. However, there is a preference in term of likeability for scenario A. In the future, the experiment could be repeated with the involvement of more test subjects, test out other evaluation metric than Godspeed and design further scenarios such as for example the task leaving affinity. Further experiments could be designed in a less controlled manner, where the agents trained to interpret human intention of changing his/her orientation, additionally switching the control of interaction.

VII. Acknowledgments

I would like to thank Victor Schmuck who gave technical support to Pepper robot, hence, helped me get results of better quality. I am also grateful for my professor Matthias Rehm who supervised me through the whole project. Additionally I would like to acknowledge the equipment support of Aalborg University in form of an actual Pepper robot.

References

[1] N. Hirai and H. Mizoguchi, "Visual tracking of human back and shoulder for person following robot," Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003), Kobe, Japan, 2003, pp. 527-532 vol.1.

[2] Shinichi Okusako Shigeyuki Sakane "Human Tracking with a Mobile Robot using a Laser Range-Finder" Journal of RSJ vol. 24 no. 5 pp. 605-613 2006.

[3] Kazuyuki Morioka, Yudai Oinaga, Yuichi Nakamura, "Control of Human-Following Robot Based on Cooperative Positioning with an Intelligent Space", IEEJ Transactions on Electronics Information and Systems, vol. 131, no. 5, pp. 1050-1058, 2011.

[4] Kazuyuki Morioka, Joo-Ho Lee, Hideki Hashimoto, "Human-Following Mobile Robot in a Distributed Intelligent Sensor Network", IEEE Transactions on Industrial Electronics, vol. 51, no. 1, pp. 229-237, 2004.

[5] F. Hoshino and K. Morioka, "Human following robot based on control of particle distribution with integrated range sensors," 2011 IEEE/SICE International Symposium on System Integration (SII), Kyoto, 2011, pp. 212-217.

[6] J. F. Engelberger. Health-care robotics goes commercial: The Helpmate experience. Robotica, 11(6):517523, Nov-Dec 1993.

[7] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, and V. Verma. Experiences with a mobile robotic guide for the elderly. In Proceedings of the National Conference of Articial Intelligence (AAAI 02), pages 587592, Edmonton,AB, July 2002.

[8] W. Burgard, A. B. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. The interactive museum tourguide robot. In Proc.of the Fifteenth National Conference on Articial Intelligence, Madison, Wi, 1998.

[9] Shiomi, M., Kanda, T., Ishiguro, H., Hagita, N.: Interactive Humanoid Robots for a Science Museum. IEEE Intell. Syst. 22, 25-32 (2007)

[10] Faber, Felix., et al.: The Humanoid Museum Tour Guide Robotinho. The 18th IEEE International Symposium on Robot and Human Interactive Communication. pp. 891-896 (2009)

[11] Sisbot, E. A., Marin-Urias, L.F., Alami, R., and Simeon, T.: A Human Aware Mobile Robot Motion Planner. IEEE Transactions on Robotics, 23(5):874-883. (2007)

[12] Kurse, T., Pandey, A. K., Alami, R., and Kirsch, A.: Human-aware robot navigation: A survey. Robotics and Autonomous Systems, 61(12):1726-1743. (2013)

[13] Hall, Edward T.: The Hidden Dimension. Anchor Books. ISBN 0-385-08476-5. (1966)

[14] Henry, P., Vollmer, C., Ferris, B., and Fox, D.: Learning to navigate through crowed environments. In 2010 IEEE International Conference on Robotics and Automation, pp. 981-986 (2010)

[15] Feng, C., Azenkot, S., and Cakmak, M. (2015). Designing a Robot Guide for Blind People in Indoor Environments. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts (HRI'15 Extended Abstracts). ACM, New York, NY, USA, 107-108

[16] LeicaJuan, P., Toibero, M., Roberti., and Carelli, R., (2015). Switched Control to Robot-Human Bilateral Interaction for Guiding People. Journal of Intelligent Robotic Systems. SPRINGER. pp 73-93

[17] Fiore M., Khambhaita H., Milliez G., Alami R. (2015) An Adaptive and Proactive Human-Aware Robot Guide. In: Tapus A., Andr E., Martin JC., Ferland F., Ammi M. (eds) Social Robotics. ICSR 2015. Lecture Notes in Computer Science, vol 9388. Springer, Cham

[18] M. Mori. (1970). "The uncanny valley" Energy vol. 7 no. 4 pp. 33-35

[19] Aldebaran Documentation What's New in NAOqi 2.4.3? Overview - NAO Software 1.14.5 Documentation, Retrieved Dec 03, 2018 from http://doc.aldebaran.com/2-8/family/peppertechnical/indexpep.html

[20] Berg, Bruce Lawrence; Lune, Howard. (2008). Qualitative Research Methods for the Social Sciences (8th ed.). Boston. p. 3. ISBN 9780205809387. Scott Reeves, Ayelet Kuper, and Brian David Hodges.

[21] Brian A. Hoey. What is Ethnography? Retrieved Nov 29, 2018 from http://brianhoey.com/research/ethnography/

[22] Ylirisku, S. (2007). Designing with video: Focusing the user-centred design process, S.l.: SPRINGER. p. 65- 66

[23] Ylirisku, S. (2007). Designing with video: Focusing the user-centred design process, S.l.: SPRINGER. p. 49