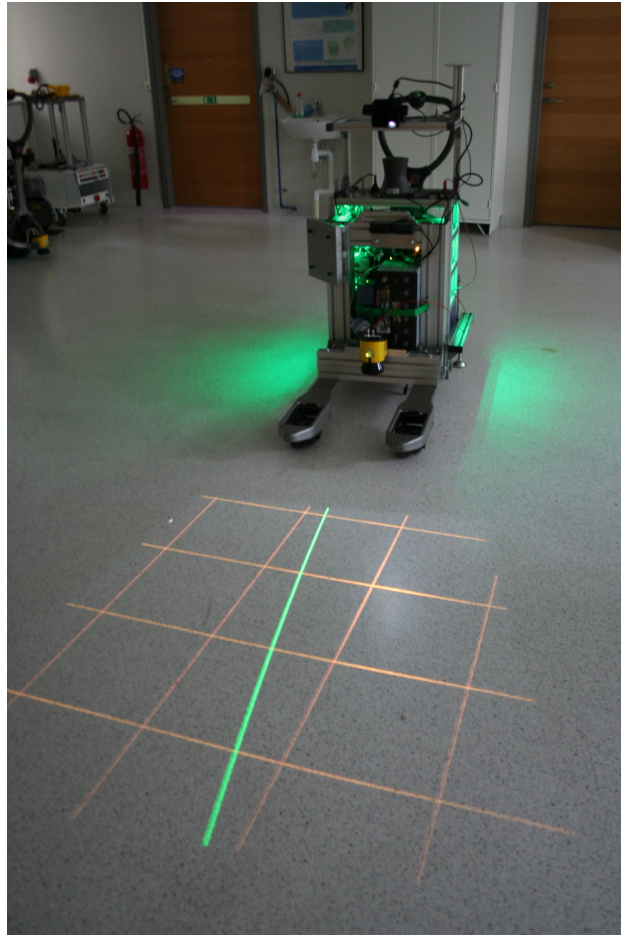




CHALMERS
UNIVERSITY OF TECHNOLOGY



Human Robot Interaction for Autonomous Systems in Industrial Environments

Master's thesis in Systems Control and Mechatronics

RAVI TEJA CHADALAVADA

Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016

MASTER'S THESIS 2016:EX039/2016

Human Robot Interaction for Autonomous Systems in Industrial Environments

A study on robot to human intention communication through
on-board projection on shared floor spaces

RAVI TEJA CHADALAVADA



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016

Human Robot Interaction for Autonomous Systems in
Industrial Environments

A study on robot to human intention communication through on-board projection on shared floor spaces

RAVI TEJA CHADALAVADA

© RAVI TEJA CHADALAVADA, 2016.

Supervisor: Achim J. Lilienthal, MRO Lab, AASS, Örebro University.

Examiner: Knut Åkesson, Department of Signals and Systems.

Master's Thesis 2016:EX039/2016
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Linde CitiTruck AGV projecting its future intentions on the shared floor space using a retrofitted LED projector.

Typeset in L^AT_EX
Printed by Chalmers University of Technology
Gothenburg, Sweden 2016

Human-Robot Interaction for Autonomous Systems in Industrial Environments.

A study on robot to human intention communication through on-board projection on shared floor spaces.

RAVI TEJA CHADALAVADA

Signals and Systems

Chalmers University of Technology

Abstract

The upcoming new generation of autonomous vehicles for transporting materials in industrial environments will be more versatile, flexible and efficient than traditional Automatic Guided Vehicles (AGV), which simply follow pre-defined paths. However, freely navigating vehicles can appear unpredictable to human workers and thus cause stress and render joint use of the available space inefficient. This work addresses the problem of providing information regarding a service robot's intention to humans co-populating the environment. The overall goal is to make humans feel safer and more comfortable, even when they are in close vicinity of the robot. A spatial Augmented Reality (AR) system for robot intention communication by means of projecting proxemic information onto shared floor space is developed on a robotic fork-lift by equipping it with a LED projector. This helps in visualizing internal state information and intents on the shared floors spaces. The robot's ability to communicate its intentions is evaluated in realistic situations where test subjects meet the robotic forklift. A Likert scale-based evaluation which also includes comparisons to human-human intention communication was performed. The results show that already adding simple information, such as the trajectory and the space to be occupied by the robot in the near future, is able to effectively improve human response to the robot. This kind of synergistic human-robot interaction in a work environment is expected to increase the robot's acceptability in the industry.

Keywords: Human robot interaction, intention communication, internal states, Spatial Augmented Reality.

Acknowledgements

I would like to thank Henrik Andreasson, Robert Krug and Achim Lilienthal for their help during the work and the examiner Knut Åkesson. I also would like to thank my family and friends for their constant support.

Ravi Teja Chadalavada, Gothenburg, February 2016

Contents

List of Figures	xii
List of Tables	xiii
1 Introduction	1
2 Theory	3
3 Methods	7
3.1 Proposed Approach	7
3.1.1 Vehicle Platform	7
3.1.2 Projected Pattern	8
3.1.3 Calibration	8
3.2 AR module based interface	11
4 Evaluation	17
4.1 AR based Intention Communication experiments	18
4.1.1 Pilot Experiment 1	18
4.1.2 Pilot Experiment 2	20
4.2 AR based Interface experiments	20
4.2.1 Pilot Experiment 1	21
4.2.2 Pilot Experiment 2	22
5 Results	25
5.1 AR based Intention Communication experiments	25
5.2 AR based Interface experiments	28
6 Conclusion	33

List of Figures

3.1	The platform used for the evaluations: A standard projector (Optoma ML 750) (1) is mounted on a retrofitted Linde CitiTruck AGV (3). The projector is used to project the intention of the vehicle on the ground plane in front of the truck (4). Two SICK S300 scanners are mounted in front (2) and back to ensure safety for human co-workers.	13
3.2	Rendered image to be displayed with the projector. The dark red grid is of size $0.15 \times 0.15 \text{ m}^2$. The green line represents the intended trajectory to be driven and the white lines contain the region that the vehicle needs to occupy in order to traverse the path (green).	14
3.3	Examples of rendered images to be projected with different velocity profiles: (top) faster straight path, (bottom) slower turning path. The slowdown region is depicted in green – if a person enters this region, the robot will slow down. The e-brake region is colored red – violating this zone will cause the robot to stop. The white line is representing the path to be driven.	15
4.1	Photos taken during different stages with different subjects during pilot experiment 2 involving a sharp turn;	18
4.2	Driven paths for the two setups. For the first set of experiments the path was almost straight, whereas the second set of experiments involved a sharper turn.	19

4.3	Pilot experiments setup: (Top) Experiment 1 – the two interacting agents approach each other head-on; (Bottom) Experiment 2 – the two agents meet at a junction-crossing. The sliding door was used to limit a humans field of view while allowing enough maneuvering space for the robot to pass through.	23
4.4	Snapshots taken during different stages with different subjects during the pilot experiments (best viewed in color).	24
5.1	Response of the 13 test subjects to the questionnaire: the improvement in the ratings when the robot’s intentions are projected is evident.	26
5.2	Exemplary trajectories of the robot and a human test subject from pilot experiment 1. Red trajectories represent projection is OFF and blue represents projection is ON.	28
5.3	Experiment results: (Top) Experiment 1; (Bottom) Experiment 2; Robot intention communication allows to approach, and in some cases to exceed, human-human interaction comfort levels.	29
5.4	Experiment results: After the human-human encounter, each of the subjects were asked what they had observed in the other human when they were passing each other and the pie chart depicts the subject’s observations during the encounter with another subject in the human-human experiment.	31
.1	Data collected through the Likert scale ratings after the experiments pertaining to the Section 4.1. AR based Intention Communication experiments	I
.2	Data collected through the Likert scale ratings after the experiments pertaining to the Section 4.2. AR based Interface experiments	II

List of Tables

5.1	Veer-off distance mean and (1-sd) values. Exp 1.1 and Exp 2.1 represent the tasks with projection OFF and Exp 1.2 and Exp 2.2 represent the tasks with projection ON.	27
5.2	Paired sample t-test results comparing the data between projection OFF and projection ON.	27
5.3	Paired sample t-test results for experiment 1 based on Likert scale ratings for each evaluation attribute	30
5.4	Paired sample t-test results for experiment 2 based on Likert scale ratings for each evaluation attribute	31

1

Introduction

Human life comprises a humongous amount of human to human interactions. Future service robotics applications in shared environments will necessitate human-machine interactions. Humans are by nature social animals and the ability to communicate and understand communication has been a fundamental part of the human evolution. Hence, communication plays an important role in the future of human-robot coexistence. In humans, the face is the center of the communication system enabling verbal and non-verbal communications. This communication, done through the human face, controls almost all the social aspects of human-human interactions. Breazeal et al. [1] stress the importance of social interaction capabilities for robots. Thus, when designing robots to assist humans, we tend to make robots resembling humans, often including a face, real or virtual, which would become the center of social human-robot interaction system.

Logistics is an important commercial application of robotics where robots need to accomplish a set of goals in an environment co-populated by human workforce. These kind of robots do not have a face. In our work, we considered such a robot working alongside humans to investigate the questions of how to make humans comfortable around such a robot and how a synergistic human-robot interaction system can work in shared environments.

We consider an Automatic Guided Vehicle(AGV), a robot working in busy environments such as shop- floors and warehouses co-populated by human workforce. AGV's have been providing transportation capabilities in intra-logistic applications for several decades using pre-defined paths for navigation and blinkers for indicating directions.

The upcoming generation of AGV's need to navigate freely without sticking to the pre-defined paths in order to be more flexible, versatile and efficient. However, freely navigating vehicles can appear unpredictable to human workers and thus cause stress and render joint use of the available space inefficient. Human workers are used to collaborate with fellow humans and, in certain scenarios, even do not need to rely on verbal communication. This implicit understanding is key in industrial working conditions and an effective human-robot communication system is absent at the moment, but will play a key role in the future of human-robot coexistence. Our research focus lies on building a communication platform for the robots in logistics scenarios with important commercial applications.

We would like to address this issue by intuitively filling the communication gaps between the human and robot by building a dynamic interface through which the robot can express its intentions and interact in a way that humans can acknowledge the feedback from the robot, thus, establishing an equivalent to human-human understanding. This kind of a mutual understanding with the robot makes the humans feel comfortable and is thereby expected to improve the acceptability of such technologies in industry.

2

Theory

Human-robot interaction has been widely studied for decades, focusing mostly on human intention recognition from the perspective of a robot. However, the contrary situation of robot intention recognition from the perspective of a human has received comparatively little attention so far. A robot communicating its intentions to users in its vicinity allows for a better understanding of the robot while avoiding the unpredictability and communication gap issues. Fong et al. [2] and Breazeal et al. [1] stress the importance of social interaction capabilities for robots which includes overlapping perceptual space, appropriate interaction distance and safety. Researchers like Norman [3], Asada et al. [4], Dautenhahn [5], Bates [6] and Blumberg [7] suggest that a robot's ability to communicate effectively will make it appear more reliable, predictable and transparent to humans. A robot communicating its intentions to users increases predictability and reduces problems caused by communication gaps [1]. In turn, these communication abilities increase a user's willingness towards using the technology and eventually increases the chances of acceptance at the workplace. Usage of AR techniques has proven to be an effective method of enabling robot-human communication. Milgram et al. [8] was one of the first researchers to implement these techniques in tele-robotic control operations which were further developed by Hine et al. [9], Kelly et al. [10], Shayegan et al. [11] and Livatino et al. [12]. The most popular way of integrating AR is via head mounted displays which, for several reasons, is infeasible in industrial environments. Instead, we chose to develop a spatial AR system which projects the robot's intention directly into the real environment.

One of the few applications of using AR for communication purposes

in robotics is to aid human operators in a human-robot co-worker assembly scenario – was recently proposed by Ruther et al. [13]. Daily et al. [14] used head-mounted AR displays for communicating information to humans from large numbers of small-scale robots in a robot swarm to enable situation awareness, monitoring, control for surveillance, reconnaissance, hazard detection and path finding. Fabrizio et al. [15] and Collett et al. [16] used interactive AR to represent a practical, interactive system for visualizing the internal and normally hidden states of the swarm, overlaid in real-time over a live video feed acquired from a fixed camera. This projection of internal states was used for analysis and debugging processes.

With respect to our work, the most relevant developments are done by Matsumaru [17], Florian et al. [18], Lee et al. [19], Park and Kim [20], Costa et al. [21] and Covert et al. [22]. They have developed spatial AR systems to project the intentions of a robot on the shared floor space to enable a user to understand the data and behavior of a robotic assistant, thus providing an opportunity to analyze and potentially optimize the working process. The works in [21], [22] performed tests in a real environment, which showcased encouraging results regarding the usefulness of communicating robot intentions. The authors of [17] introduced a mobile robotic system which presents the scheduled path and basic operation states to the people nearby. Also, they conducted a questionnaire evaluation on 200 people about the direction of motion and the speed of motion only, which indicated that the employed AR system made the robot's intents more intelligible. In [20] the idea of a projector based interface to interact with the robot was proposed.

Covert et al. [22] focused on evaluating the robot's ability to communicate intended movements to a human by asking questions about what the robot's intention might be, while the work in [21] focused more on developing interactive AR interfaces for mobile robots to be used in rehabilitation applications. We have evaluated the mobile robot's ability based on the test subject's reaction in a close to real life situation.

Usage of the Spatial AR system to project the intentions of a robot on the shared floor space, allowed humans to interpret and react to

the behaviors of a robotic agent, thus, providing an opportunity to analyze and potentially optimize the working process. As a further development to this system, we developed a projector based interface to interact with the robot like how Park and Kim [20] implemented.

A significant research effort has been dedicated to generate suitable proxemic robot behaviors including appropriate human avoidance control [23] and a study of spatial distances and orientation of a robot with respect to a human user [24]. Other findings show that experience with robots reduces the discomfort zone [25] and that humans maintain larger distances when a mutual gaze is established [26]. Recently, Dondrup et al. [27] incorporated proxemics into the Qualitative Trajectory Calculus representation.

In our work, we rather focused on generating robot behaviors, by communicating the intentions, which allowed humans to better pre-plan their own motion and to adapt their proxemics by maintaining larger distances from the robot. It is experimentally shown that this adaption leads to natural interactions with comfort levels approaching those of human-human interactions.

The key contributions of this work are an implementation of a spatial augmented reality visualization system for a mobile platform, a projector based interface system to interact with robot and practical evaluations of these systems by mimicking a real life scenarios.

3

Methods

3.1 Proposed Approach

The main objective of this work is to communicate the intention of a fork-lift type vehicle, such as the research prototype depicted in Fig. 3.1, to humans in the vicinity. Ideally, the coverage of the projected floor space should enclose the area around the vehicle and be sufficiently large to allow displaying the intention of the vehicle over a time horizon of several seconds. In the initial evaluation performed in this work, a standard projector was mounted pointing in the direction of the forks as shown in Fig. 3.1. Thus, both the Field of View (FOV) and illumination brightness are limiting factors. For example, to obtain a large enough projection area, the projector was tilted resulting in some non-illuminated area between the vehicle and the projected image on the floor. This is acceptable because even though the robot's path is generated on the fly, the motion of the vehicle is highly predictable in its close vicinity. The projector is connected to an on-board computer which renders images using an available pose estimate of the vehicle's location together with information regarding the current mission.

3.1.1 Vehicle Platform

The mobile base is built upon a manually operated forklift which originally was equipped with motorized forks and a drive wheel only. The forklift has subsequently been retrofitted with a steering mechanism and a commercial AGV control system. The latter is used to interface

the original drive mechanism, as well as the steering servo. To assure safe operation, the vehicle is equipped with two SICK S300 safety laser scanners¹ respectively facing in forward and backward directions.

3.1.2 Projected Pattern

In order to render projection images the GLUT framework is utilized. A common reference frame is used in the rendering of the scene and in the overall architecture [28].

This approach makes it straightforward to draw the common 3D world representation by updating the pose of the projector/virtual camera by using the localization estimate of the AGV and the extrinsic parameters (*i. e.*, the pose of the projector/virtual camera expressed in an AGV-fixed coordinate frame). An example of a rendered image that is used for projection is depicted in Fig. 3.2. The projected red squares remain stationary even when the vehicle is moving.

3.1.3 Calibration

There are two steps in drawing the pattern onto the floor. First, we render the image using the GLUT framework which results in a full screen image. This image looks different depending on where in the virtual world we place the virtual camera. We project the rendered image (full screen) from the virtual camera onto the floor. Another essential part is therefore to determine the parameters of the projector such as its focal length and aspect ratio. Note also that the aspect ratio is dependent on the resolution of the graphics card used to render the image. The goal of the calibration procedure is to be able to consistently place a virtual camera in the GLUT drawing framework such as to generate an image which corresponds to 3D coordinates in the real world when projected on the floor.

The key function of a projector is to display an undistorted image onto a flat surface. Therefore, in this work, we utilize the standard perspec-

¹<http://www.sick.com/>

tive pin-hole camera model [29] to describe the transformation from the image to the projected image in a given reference frame. The standard rendering components available in the OpenGL framework are used to render the image to be projected. The pin hole camera model is described using a camera projection matrix \mathbf{P} which expresses the mapping from a 3D position \mathbf{x} to a 2D image coordinate \mathbf{y} expressed in homogeneous coordinates. The projection matrix is computed as

$$\mathbf{P} = \mathbf{A} (\mathbf{R}|\mathbf{T}) = \begin{pmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} (\mathbf{R}|\mathbf{T}), \quad (3.1)$$

where $f_{x,y}$ are the focal lengths, x_0, y_0 is the center of the projectors coordinates in pixels and \mathbf{R}, \mathbf{T} describe the pose of the projector (rotation \mathbf{R} and position $\mathbf{t} = -\mathbf{R}^T\mathbf{T}$ in the world coordinate frame). Here \mathbf{A} is the matrix of intrinsic parameters and $(\mathbf{R}|\mathbf{T})$ the matrix of extrinsic parameters. $(\mathbf{R}|\mathbf{T})$ represents a block matrix comprising of matrices \mathbf{R} and \mathbf{T} .

In this work we have two projection matrices; the first from the virtual camera \mathbf{P}_C which is used to render the scene and the second representing the projector \mathbf{P}_P . Given a 3D point

$$x_C$$

in OpenGL, a 2D image coordinate \mathbf{y} and the corresponding projected 3D coordinate \mathbf{x}_P , the following relation holds in case of a pin-hole projection model:

$$\mathbf{y} = \mathbf{P}_C \mathbf{x}_C = \mathbf{P}_P \mathbf{x}_P. \quad (3.2)$$

The main goal of the calibration procedure is to find the transition between the vehicle origin (in the real world) and the OpenGL frame to render the image. Given the projection matrices the transition can be computed as:

$$\mathbf{x}_P = \mathbf{P}_P^{-1} \mathbf{P}_C \mathbf{x}_C. \quad (3.3)$$

The main problem lies in obtaining the projection matrices. For the projector matrix \mathbf{P}_P we need to find the extrinsic offset $(\mathbf{R}|\mathbf{T})$ as well as the intrinsic parameters \mathbf{A} . For the virtual camera \mathbf{P}_C finding the intrinsic parameters can easily be done using the parameters in the ray-tracing method with the screen resolution. To determine the

projection matrix for the projector is, however, not straight forward. Since the initial hardware setup only contained the projector and no other sensor which could be used for an automatic calibration procedure, we use a manual calibration approach which essentially works by measuring the projected pattern on the floor. This simple approach could be extended to an automatic procedure.

To simplify the calibration procedure we propose the following approach: firstly, we are not interested in obtaining the projection matrices $\mathbf{P}_C, \mathbf{P}_P$ individually. Secondly, we are also not interested in the factorization of the projection matrices \mathbf{P} into intrinsic \mathbf{A} and extrinsic parameters ($\mathbf{R}|\mathbf{T}$). Instead, our calibration parameters will consist of 7 variables in total; 6 parameters to describe the pose of the virtual camera in the GLUT framework and a scale parameter s , which is used to tune the aspect ratio of the projected image. This scale parameter is directly related to the ratio of the focal lengths f_x/f_y and the focal lengths can be altered by moving the camera back and forth along the viewing axis. The center of projection x_0, y_0 is directly incorporated into the extrinsic parameters (please note that the center of projection only determines where in the image plane the extrinsic parameters refer to).

To perform the calibration, an evenly spaced 2D grid with a fixed size of 0.15×0.15 meters is drawn together with the origin of the common reference frame. A calibrated projector system should then be able to replicate this grid pattern on the floor where the origin of the grid should be at the origin of vehicle located between the two fixed wheels at the forks. The involved manual calibration steps are outlined below:

- roll, pitch
By utilizing a tape measure, the roll and pitch values are adjusted such that lines are parallel. The size is not important, nor that the grid cells are squared.
- scale - s
The aspect ratio of the projected image is adjusted until the grid cells are squared (again, the size is not important).

- Height - z
The height is adjusted such that the grid cells have a size of 0.15×0.15 meters.
- yaw
The heading is adjusted to make the origin of the projected grid be aligned with the direction of the vehicle.
- Position - x, y
The virtual camera position is adjusted such that the origin of the common frame is at the vehicle origin (between the two fixed wheels at the forks).

To give an intuitive movement of the virtual camera for the user in the OpenGL framework during calibration, an orbit type of camera is utilized where the pose is represented using a focal point on the floor $(x, y, 0)$ a distance r and roll, pitch and yaw orientations of the camera. The first step is to use the pitch and roll parameters to adjust the pose to get a pattern with parallel lines on the floor. Next, we use the yaw parameter in order to orient the direction to make the coordinate axes aligned. The third step is to use the distance parameter r and the scale parameter s to obtain projected squares on the floor which are of correct size. Finally, we move the focal point in the (x, y) -plane to get the position of the coordinate system aligned.

The pose of the projector will be computed later on using the global localization estimates of the vehicle and the pose of the projector relative to the reference frame of the vehicle. Therefore, the calibration will be relative to the origin of the vehicle. An automatic calibration procedure is undoubtedly possible but would require additional sensory equipment.

3.2 AR module based interface

Using the data from the sick scanners we define two dynamic regions. One region which will cause the vehicle to slow down (the slowdown speed is set to 0.05 m/s compared to the normal max speed of 0.6 m/s)

and an e-brake region where the forklift will directly stop. The two regions are defined based on the intended velocity profile and the footprint of the vehicle as seen in Fig. 3.3. The slowdown and e-brake regions are defined to be the space the vehicle needs to occupy the next 5 seconds or 2 seconds respectively.

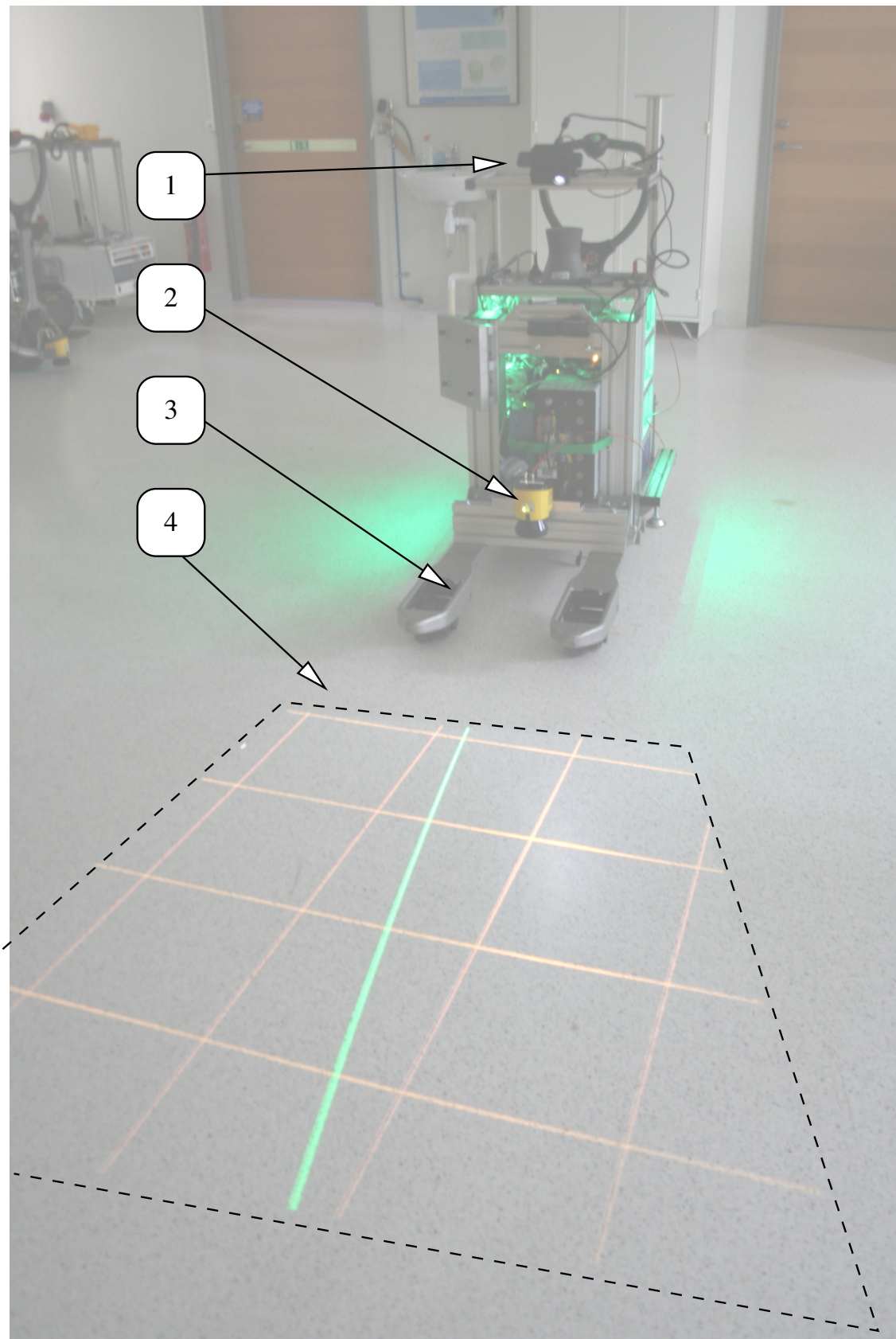


Figure 3.1: The platform used for the evaluations: A standard projector (Optoma ML 750) (1) is mounted on a retrofitted Linde CitiTruck AGV (3). The projector is used to project the intention of the vehicle on the ground plane in front of the truck (4). Two SICK S300 scanners are mounted in front (2) and back to ensure safety for human co-workers.

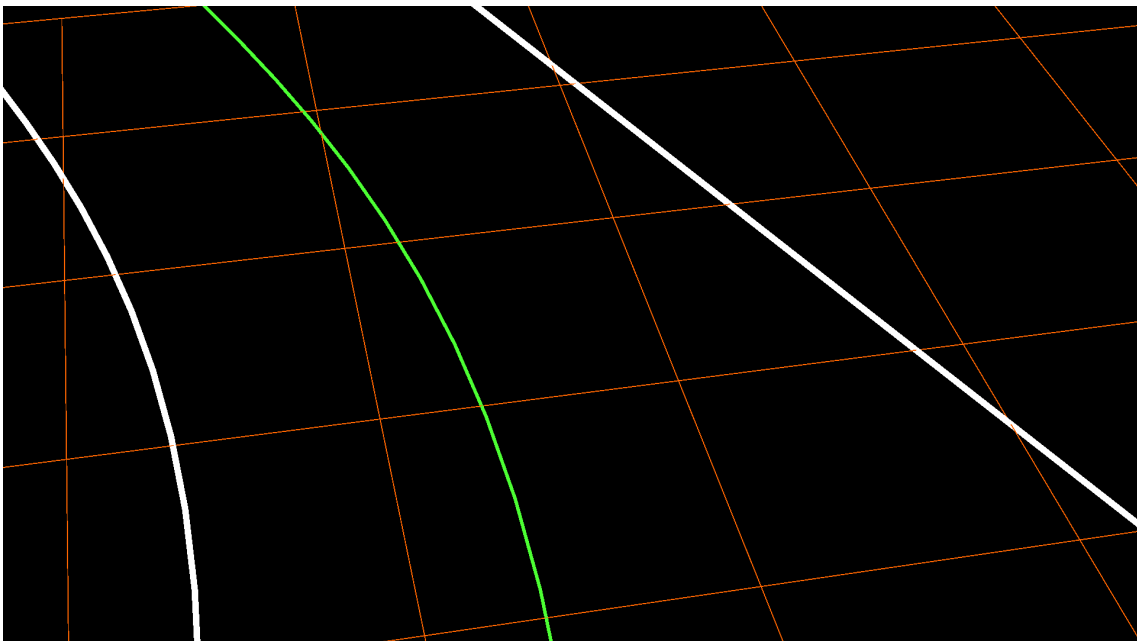


Figure 3.2: Rendered image to be displayed with the projector. The dark red grid is of size $0.15 \times 0.15 \text{ m}^2$. The green line represents the intended trajectory to be driven and the white lines contain the region that the vehicle needs to occupy in order to traverse the path (green).

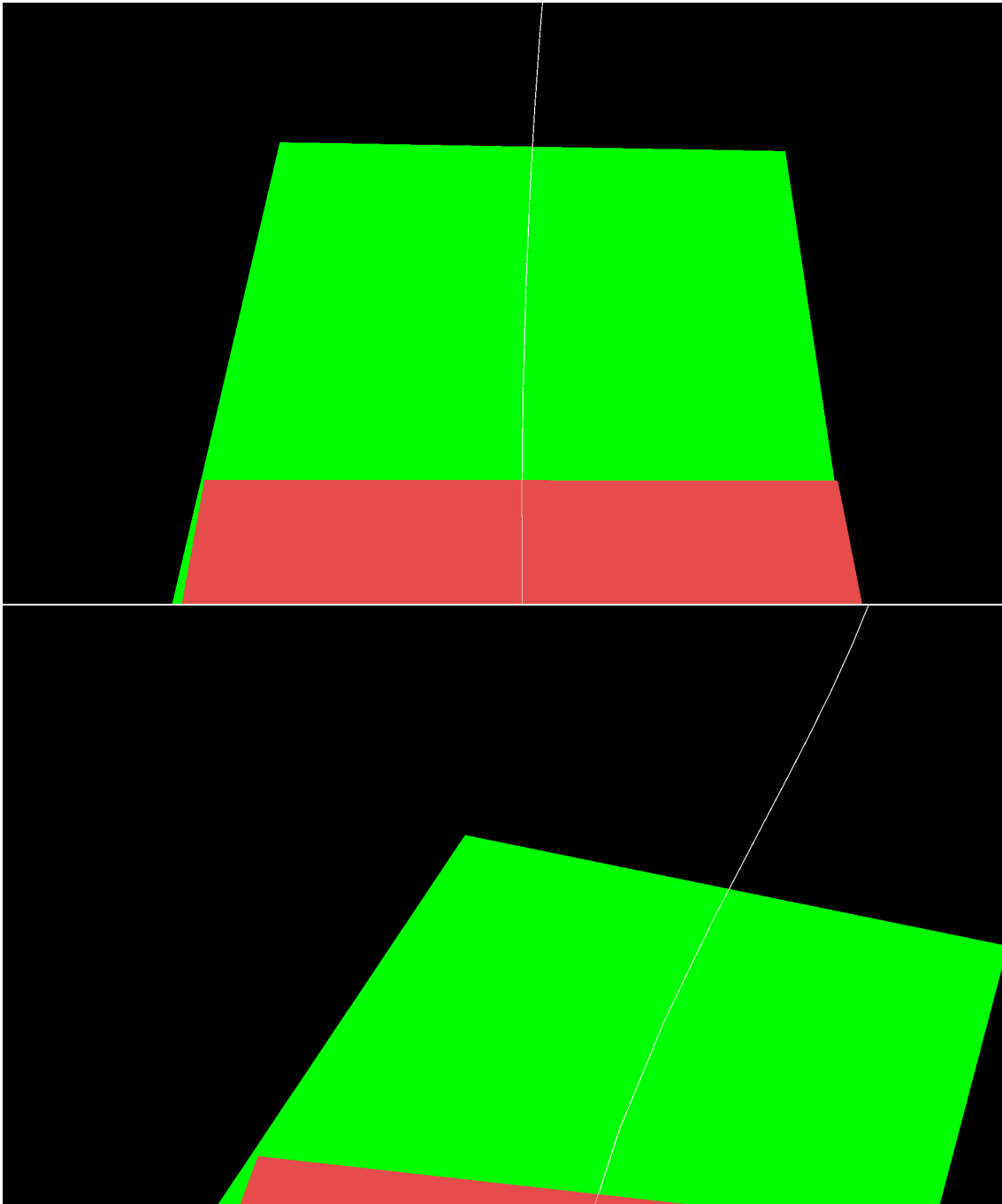


Figure 3.3: Examples of rendered images to be projected with different velocity profiles: (top) faster straight path, (bottom) slower turning path. The slowdown region is depicted in green – if a person enters this region, the robot will slow down. The e-brake region is colored red – violating this zone will cause the robot to stop. The white line is representing the path to be driven.

4

Evaluation

In order to understand how useful this technology can be to establish a sustainable human-robot interaction, our aim is to determine quantitatively how humans react to the robot's intentions projected on the shared floor space and how well do they interact with them, pilot experiments are divided into two main parts. The first set of experiments are called AR based Intention communication experiments, in which the robot projected the future intentions and the subjects couldn't interact with the projections. The second set of pilot experiments are called AR based Interface experiments, the projection patterns were modified when compared to the first set of experiments and this time the subjects were able to interact with the robot using the projections on the floor. In each of these experiments, a chosen set of key attributes were selected for measuring and comparison purposes. These attributes were chosen based on a study of human factors and a literature review. The chosen key attributes along with their respective measured abilities are given in the following table.

Attribute	Measured ability
Communication	convey information to humans
Reliability	encourage trust in humans
Predictability	humans comfortable around the robot
Transparency	intentionally share the information
Situation awareness	convey necessary information corresponding to the current situation

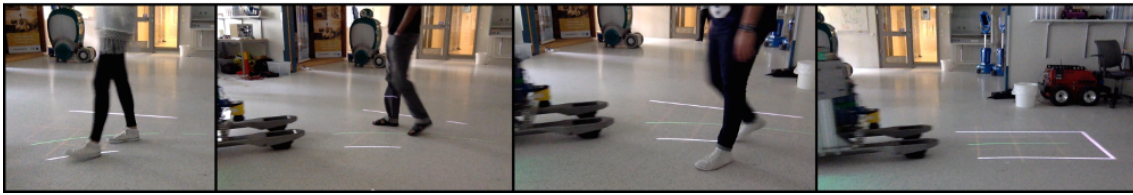


Figure 4.1: Photos taken during different stages with different subjects during pilot experiment 2 involving a sharp turn;

4.1 AR based Intention Communication experiments

AR based intention communication experiments were further divided into two pilot experiments designed around real world scenarios to test the key attributes that contribute to a synergistic robot-human work environment. In each pilot experiment, as soon as the robot starts moving, the test subject was asked to start walking towards the robot until no longer comfortable with the approaching robot. Every test subject was later asked to rate their experience with the robot on a scale of 1 to 7 with respect to the chosen key attributes. A total number of 13 subjects were chosen from a wide spectrum of backgrounds and ages, such as students, social workers, socio-economists, administration workers, researchers and engineers. Only two of them had some experience with robots but not in particular with the robot employed in the experiments.

The obtained data was used to measure the level of human reactions. Necessary safety precautions were taken during all the pilot experiments and all the test subjects were informed about the potential risks and how to behave in safety critical situations. The maximum velocity of the vehicle was limited to 0.6 m/s during all evaluations.

4.1.1 Pilot Experiment 1

Pilot experiment 1 essentially constituted a chicken game and was subdivided into two parts. In pilot experiment 1.1 the robot moved in a straight line without projecting its intentions, while the test subject was asked to walk in a straight line towards the robot and to veer off

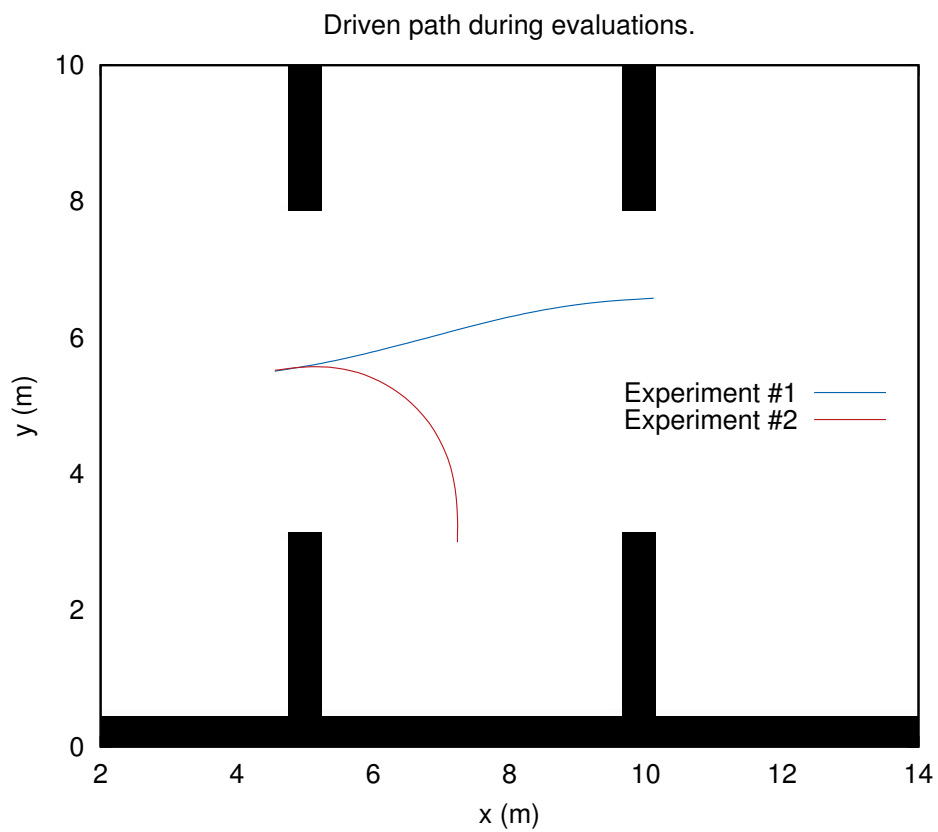


Figure 4.2: Driven paths for the two setups. For the first set of experiments the path was almost straight, whereas the second set of experiments involved a sharper turn.

her path when no longer comfortable with the approaching robot.

Pilot experiment 1.2 is the same as above with the addition that the robot projected its intentions onto the shared floor space.

4.1.2 Pilot Experiment 2

Pilot experiment 2 is sub-divided into two parts as well. In pilot experiment 2.1 the robot makes a sharp turn without projecting its intentions. The test subject was asked to initially walk towards the robot in a straight line and, after the robot initiated its turn, to veer off in the opposite direction.

Again, pilot experiment 2.2 is the same as above, with the addition of the robot indicating its intentions.

The paths driven by the robot in the two respective experiment sets are illustrated in Fig. 4.2. An exemplary test run sequence is shown in Fig. 4.1. In all experimental test runs the projector was first switched off before switching it on in the corresponding second run.

4.2 AR based Interface experiments

For the AR based interface experiments, the projected pattern included the future trajectory, a green region and a red region. The green region indicates the safe zone to walk around the robot and the red region indicates the danger zone. The green and red region areas dynamically change their areas depending on the velocity of the robot and the position of the human. Human subjects position can be determined using the laser scanner and if the human is in the green region, robot slows down and as the robot slows down, the green area starts to shrink indicating the robot is coming closer. If a human is detected in the red region, robot takes an emergency stop. In order to quantitatively evaluate the designed system, the two pilot experiments illustrated in Fig. 4.3 were designed around real-world scenarios. These scenarios were carefully chosen from a variety of everyday encounters

humans experience on a daily basis and thus are relevant for future applications in service robotics. The first pilot experiment consists of an encounter in a corridor or aisle, while the second pilot experiment represents a junction crossing situation. For evaluation, we chose 14 subjects from various backgrounds and age groups. A briefing was given where the functionality of the AR module was explained and safety precautions were outlined. After the experiments, the subjects were asked to rate their experience on a Likert scale against the key attributes mentioned earlier. Likert scale is a psychometric response scale primarily used in questionnaires to obtain participant's preferences. In our evaluation, we used a 0-7 Likert scale, where 0 represents a poor rating and 7 represents an excellent rating.

4.2.1 Pilot Experiment 1

Pilot experiment 1 was divided into three parts. In the first part, two randomly chosen subjects were asked to walk towards each other and veer off in a natural way like how they would do in normal everyday situations. This part of the experiment was designed in order to create a benchmark for the evaluations and to prepare the subjects for the follow-up experiments with the robot. This way of evaluation is expected to bring in more originality to the ratings as the subject interacts with the robot immediately after the interaction with a human in an exactly same situation. Subjects were also asked to pay attention to how they were communicating to each other before veering off and later were questioned about their observations.

In the second part, one of the subjects was replaced with a robot. During the experiment, human and robot moved towards each other until the robot took a slight pre-defined turn to the right as depicted in Fig. 4.3. In the third part of the experiment, a projector display was added as shown in Fig. 4.4 and the experiment was repeated. The second subject, which was replaced at the beginning, was brought back for the experiments after the first subject finished the third part of the pilot experiment 2.

4.2.2 Pilot Experiment 2

The basic experimental procedure was the same as in the previously described pilot experiment. However, this time the interacting agents needed to cross paths at a junction. The experiment was set up such that both parties would arrive at the junction at approximately the same time. In cases where one agent was a robot, its corresponding path was hard-coded to be straight (see Fig. 4.3).

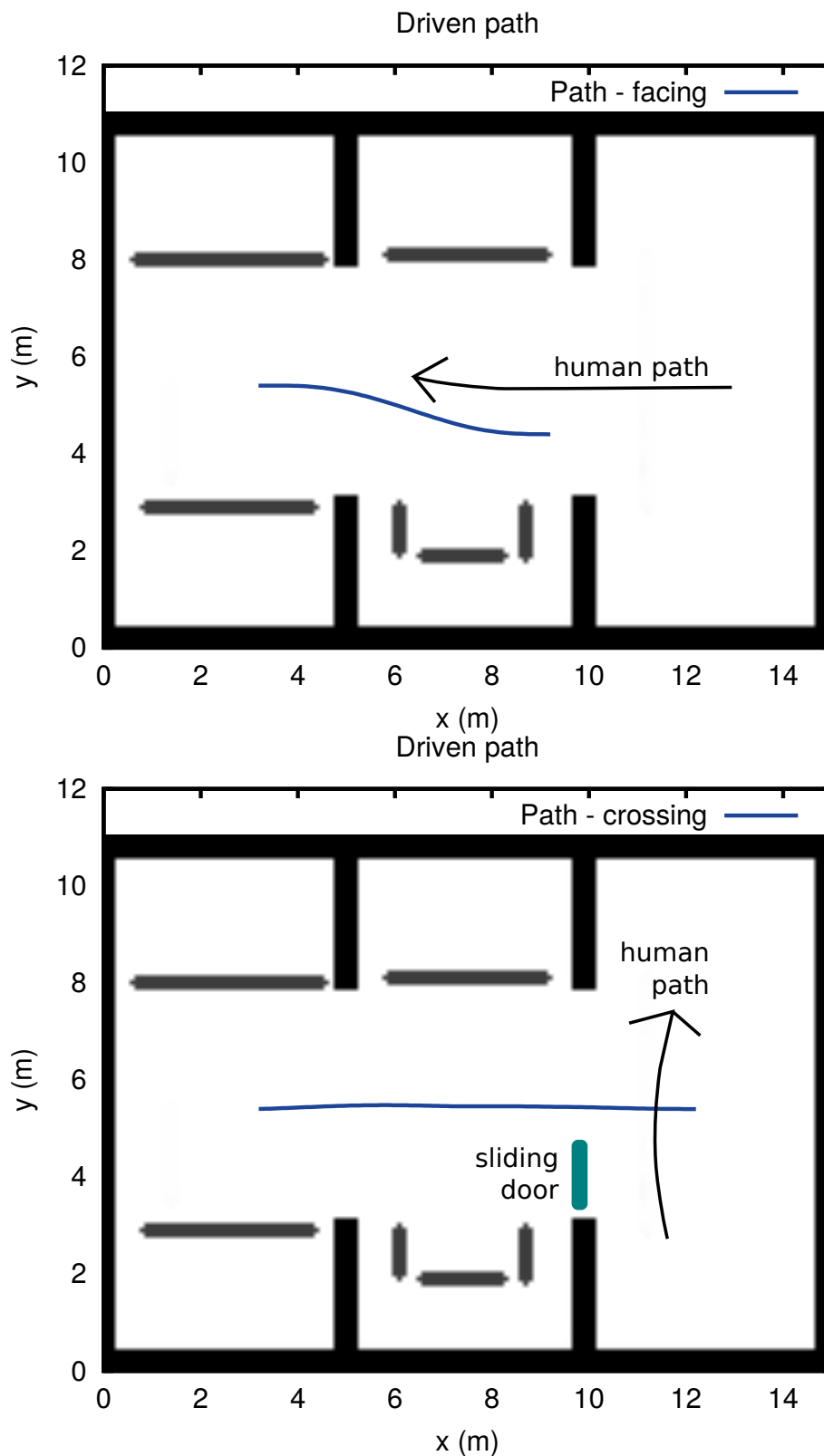


Figure 4.3: Pilot experiments setup: (Top) Experiment 1 – the two interacting agents approach each other head-on; (Bottom) Experiment 2 – the two agents meet at a junction-crossing. The sliding door was used to limit a humans field of view while allowing enough maneuvering space for the robot to pass through.

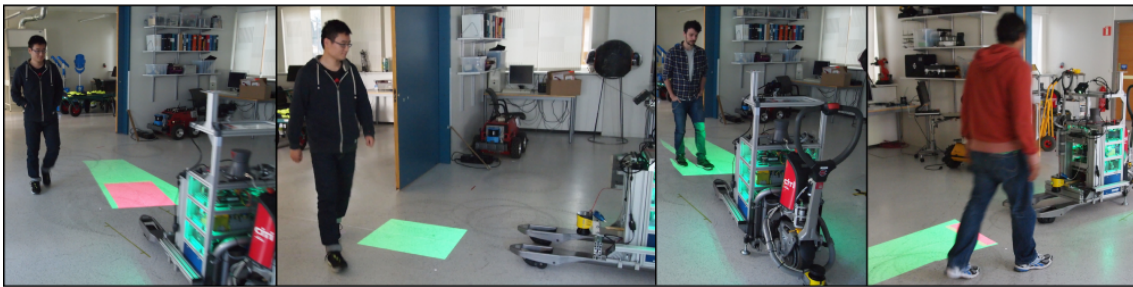


Figure 4.4: Snapshots taken during different stages with different subjects during the pilot experiments (best viewed in color).

5

Results

The Likert scale ratings given by all the subjects were averaged for each attribute and for every task of the pilot experiments. In both pilot experiments, a significant change in the human reaction was apparent when the robot projected its intentions on the shared floor space. This is in strong agreement with the hypothesis that expressing the essential states of the robot is important for a natural human-robot interaction to take place.

5.1 AR based Intention Communication experiments

When the robot projected its intentions in the pilot experiment 1.2, there was an average increase of 53% in user ratings compared to pilot experiment 1.1, in which the robot did not convey any intentions. Of the attributes considered, communication, predictability and transparency are the most vital components for the acceptability of a robot technology into a human-robot work environment and they achieved significant increases with communication at 81% and predictability as well as transparency at 62%.

For the pilot experiment 2, which is a somewhat more complex and practical scenario, there was a larger increase in the average ratings than in the previous pilot experiment. When the robot projected its intentions in pilot experiment 2.2, there was a 65% increase in the average user rating over all the considered attributes. Here, communication, predictability and transparency achieved over 90% rise.

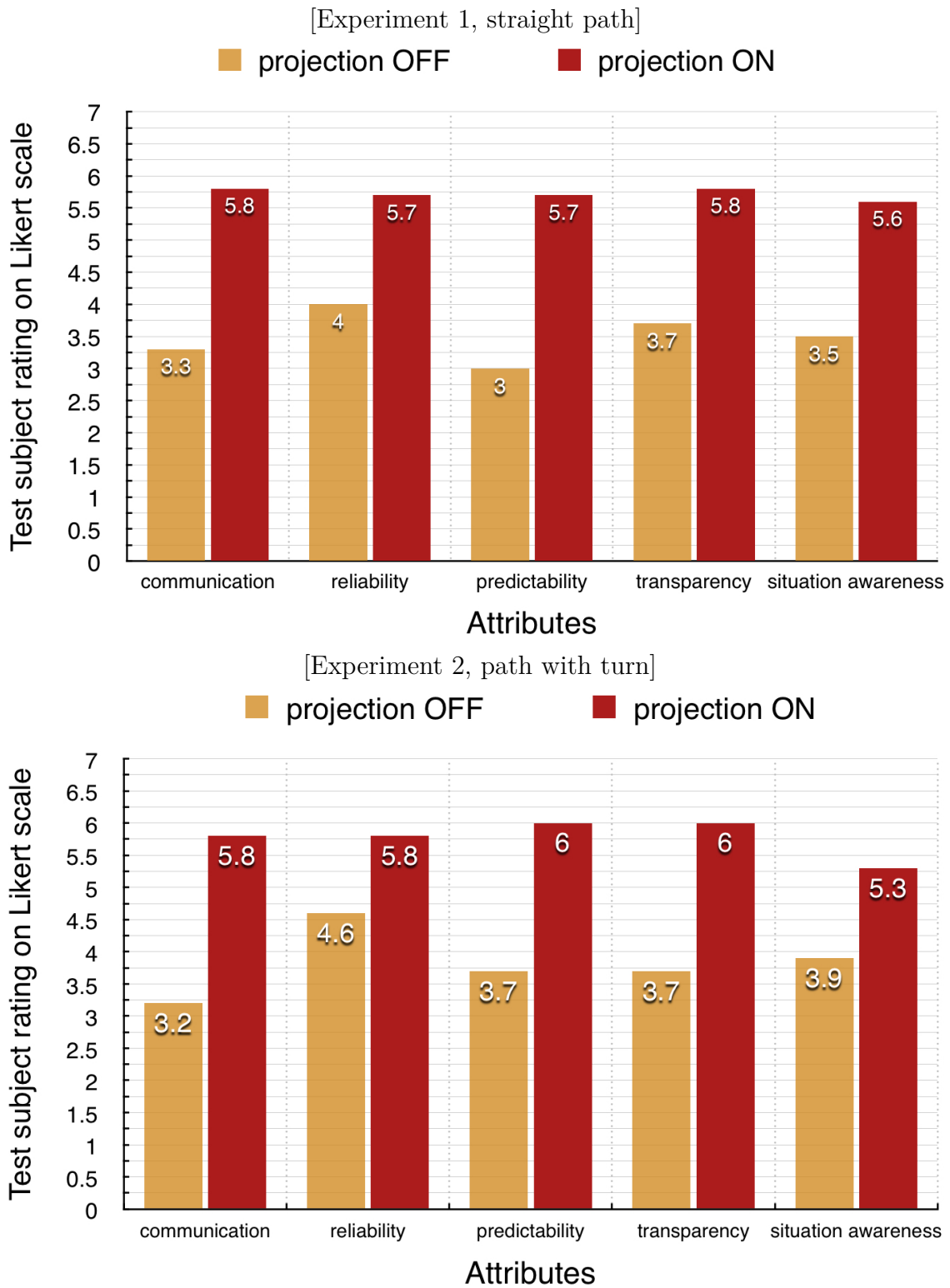


Figure 5.1: Response of the 13 test subjects to the questionnaire: the improvement in the ratings when the robot's intentions are projected is evident.

Table 5.1: Veer-off distance mean and (1-sd) values. Exp 1.1 and Exp 2.1 represent the tasks with projection OFF and Exp 1.2 and Exp 2.2 represent the tasks with projection ON.

	Exp 1.1	Exp 1.2	Exp 2.1	Exp 2.2
$\bar{d} [m]$	1.40 ± 0.45	2.01 ± 0.79	1.45 ± 0.33	1.81 ± 0.58

Table 5.2: Paired sample t-test results comparing the data between projection OFF and projection ON.

	$\overline{\Delta d} [m]$	t	p	$c_i [m]$
Exp. 1	0.68 ± 0.78	3.17	0.008	[0.21, 1.15]
Exp. 2	0.37 ± 0.37	3.55	0.004	[0.14, 0.59]

The results summarized in Fig. 5.1 indicate that the communication system installed on the robot to project its intentions have been a valuable utility for humans in the presence of the robot. This supports our hypothesis that a robot exhibiting its internal states is an asset for the technology’s acceptance at shared work scenarios and can aid in achieving harmonious work environments.

In addition to the subjective questionnaires, the subject’s trajectory during the experiment was recorded using the laser scanner of the robot and subsequently analyzed. During the pilot experiments, the point where the subject starts to veer-off from the robot’s intended path was identified and the distance d between this point and the robot was measured. Intuitively, one would expect the test subjects to approach closer to the robot in case the projection is enabled. However, the obtained results proved the contrary as shown in Table 5.1 which summarizes the mean and 1-STD of the distance values for the four experiment sets. A possible explanation for this could be that if humans are aware of future intentions of the robot, they are able to plan their path ahead as well which is beneficial in applied scenarios. A look at the trajectories extracted from the laser scanner data, corroborates this point. When the projection is ON, subjects had planned their path in advance and had a comfortable encounter with the robot instead of a hasty deviation as exemplary shown in Fig. .2. This observation relates to the positive experiences the test subjects

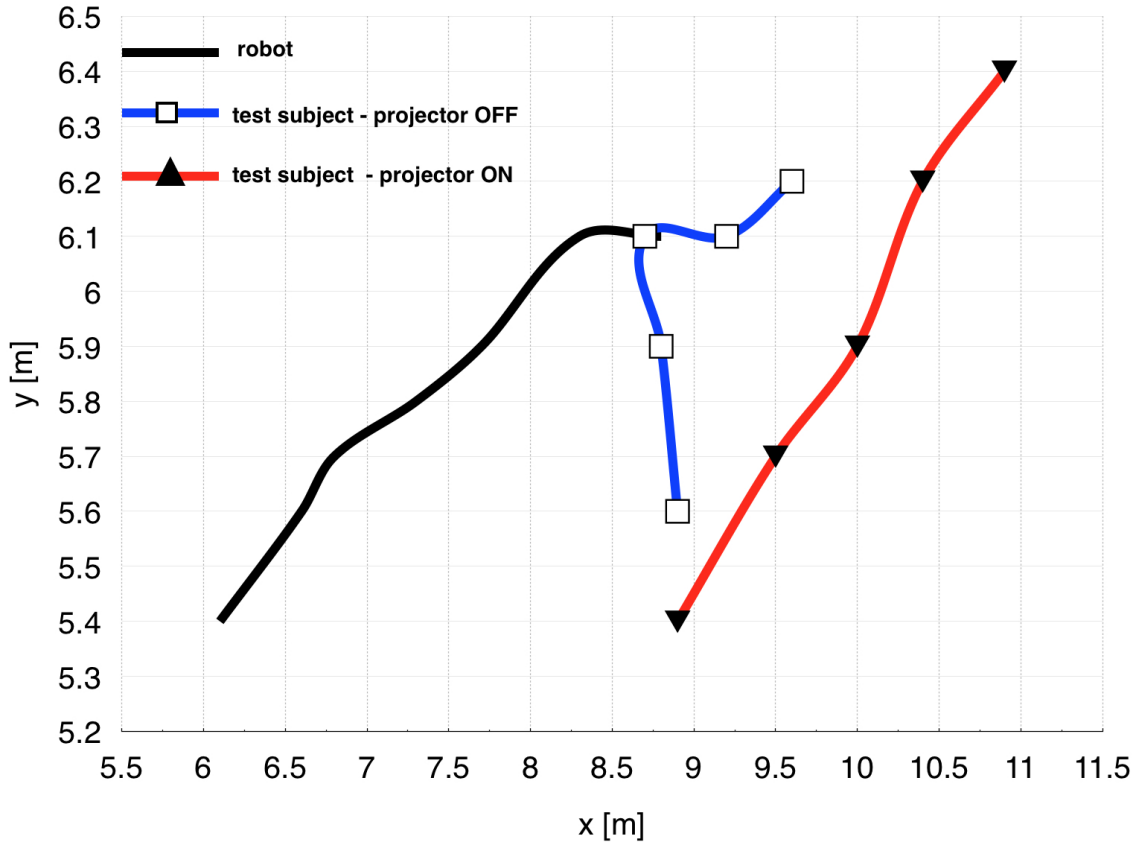


Figure 5.2: Exemplary trajectories of the robot and a human test subject from pilot experiment 1. Red trajectories represent projection is OFF and blue represents projection is ON.

had when the projection was enabled as argued previously. It is worth noting that also the distance variance increased with the projection enabled as the test subjects adopted varying reaction behaviors. To ascertain the statistical significance we conducted a paired sample t-test with a significance level of $\alpha = 0.05$. The results are summarized in Table 5.2 where $|\Delta d|$ denotes the mean and 1-STD of the veer-off distance differences between experiments conducted with the projection turned on and off respectively, t indicates the test statistics, p denotes the p-value and c_i describes the corresponding confidence intervals.

5.2 AR based Interface experiments

Here, our aim is to evaluate how humans react to the robot's intentions projected on the shared floor space and to compare these reactions to

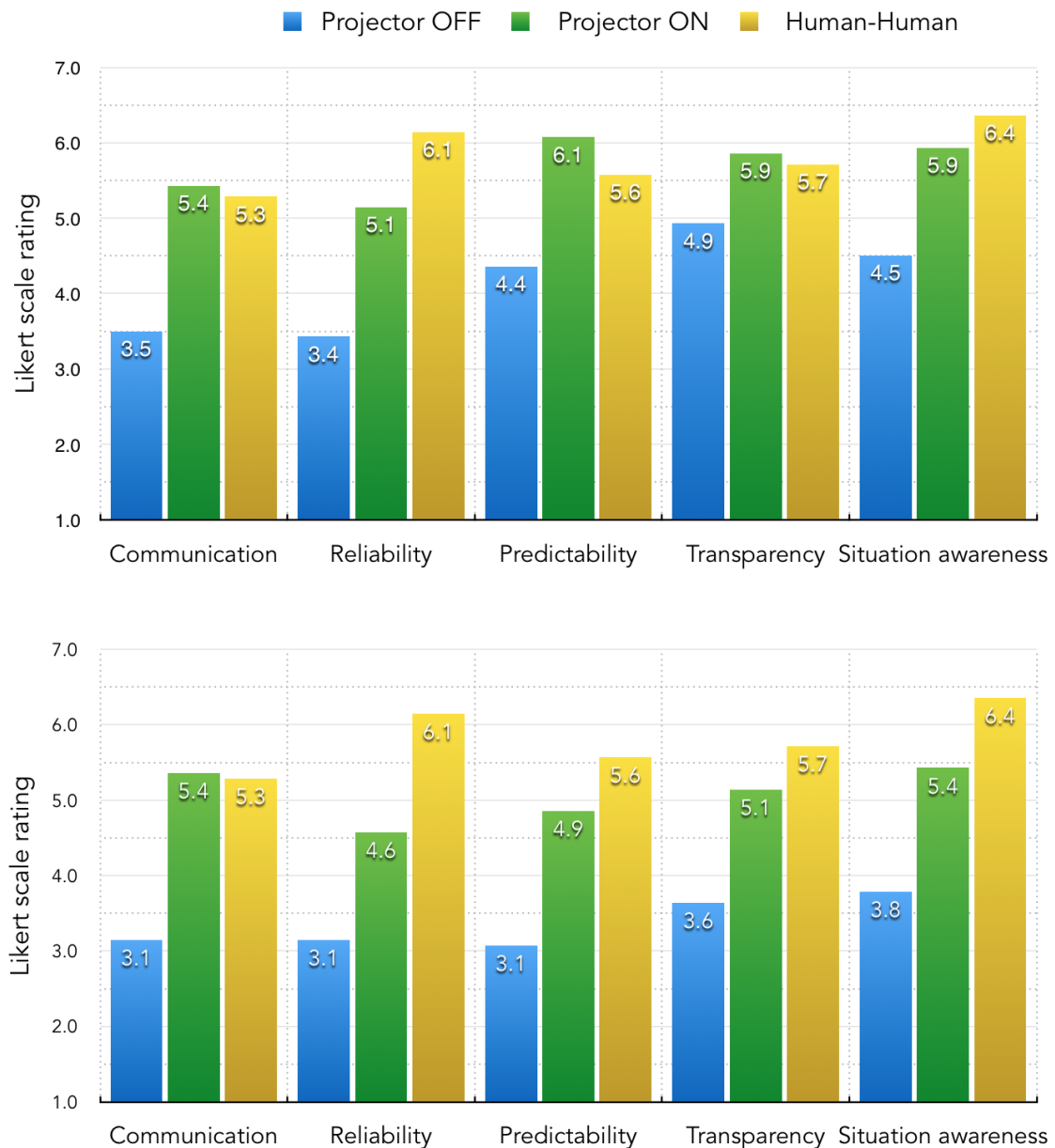


Figure 5.3: Experiment results: (Top) Experiment 1; (Bottom) Experiment 2; Robot intention communication allows to approach, and in some cases to exceed, human-human interaction comfort levels.

corresponding human-human encounters.

A breakdown of the results is presented in Fig. 5.3. Our intention communications system significantly increases the scores for all abilities. It is evident that the system allows for interactions approaching and, in case of the communication attribute, even exceeding human-human interaction comfort levels. The biggest discrepancy is in the reliabil-

Table 5.3: Paired sample t-test results for experiment 1 based on Likert scale ratings for each evaluation attribute

	$\overline{\Delta L}$	t	p	c_i
Communication	0.14 ± 0.62	0.23	0.82	$[-1.48, 1.19]$
Reliability	-1.0 ± 0.42	-2.38	0.03	$[0.09, 1.91]$
Predictability	0.5 ± 0.34	1.45	0.17	$[-1.25, 0.24]$
Transparency	0.14 ± 0.36	0.40	0.70	$[-0.92, 0.64]$
Sit. awareness	-0.43 ± 0.2	-2.12	0.05	$[-0.01, 0.87]$

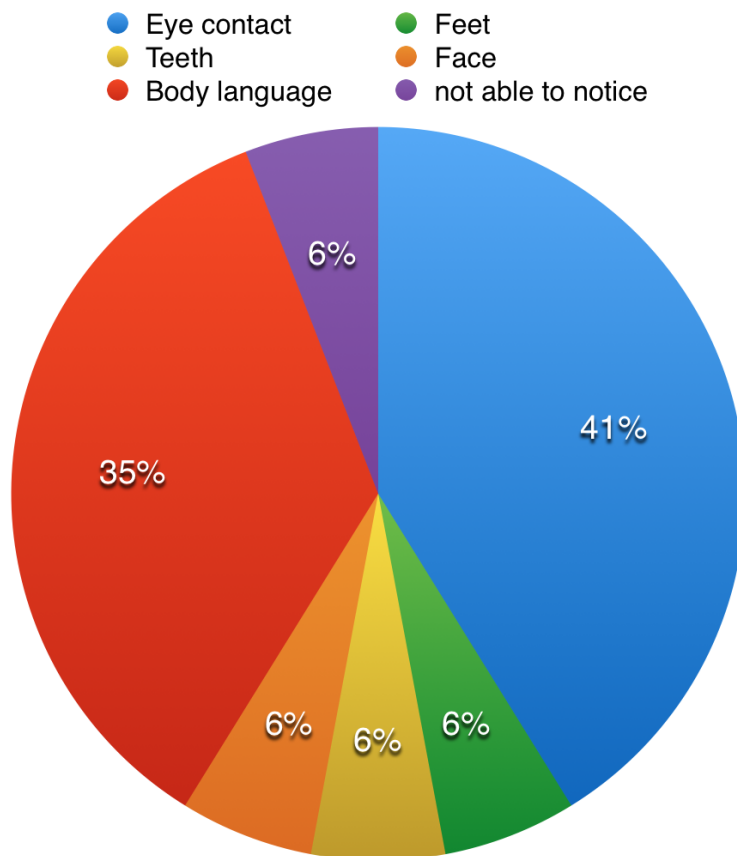
ity measure which indicates that an element of hesitation towards the robot remains. In general, the results are better for Experiment 1, where the robot communication system outperforms human-human interaction also in Predictability and Transparency. We attribute this to the fact that, due to the straight approach, the robot’s proxemic data is visualized to the human over a longer time-span than in Experiment 2).

To ascertain the statistical significance we conducted paired sample t-test for each evaluation attribute with a significance level of $\alpha = 0.05$. The results for both experiments are summarized in Table 5.3 and Table 5.4 respectively. A t-test is a statistical examination of two population means. A paired sample t-test examines whether two samples are different and is commonly used when the variances of two normal distributions are unknown and when an experiment uses a small sample size. In the tables, ΔL denotes the mean and (1-sd) of the corresponding Likert rating between human-robot interaction experiments with enabled projection and human-human interaction experiments. Furthermore, t indicates the test statistics, p denotes the p -value and c_i describes the corresponding confidence intervals.

In addition, we questioned the subjects of the human-human interaction experiments about their observations during the encounter. The result is visualized in Fig. 5.4 and highlights the importance of anthropomorphic features such as gaze and body language. We see this as a further validation of the usefulness of our intention recognition

Table 5.4: Paired sample t-test results for experiment 2 based on Likert scale ratings for each evaluation attribute

	$\overline{\Delta L}$	t	p	c_i
Communication	0.07 ± 0.64	0.11	0.91	$[-0.87, 0.01]$
Reliability	-1.57 ± 0.45	-3.47	0.0	$[-2.55, -0.59]$
Predictability	-0.71 ± 0.34	2.11	0.05	$[-0.02, 1.45]$
Transparency	-0.57 ± 0.36	-1.59	0.14	$[-0.20, 1.35]$
Sit. awareness	-0.93 ± 0.16	-5.64	0.0	$[0.57, 1.28]$

**Figure 5.4:** Experiment results: After the human-human encounter, each of the subjects were asked what they had observed in the other human when they were passing each other and the pie chart depicts the subject's observations during the encounter with another subject in the human-human experiment.

system, which allows a human to comfortably interact with a faceless mobile robot as discussed above.

6

Conclusion

The aim of this work was to evaluate how humans react in terms of their proxemic behavior to intentions projected by a robotic vehicle. Although the number of subjects used in the experiments is relatively small, the results presented in this work indicate an improvement of all evaluation criteria when the robot's intention is visible. However, the most inspiring result lies in the comparison to human-human intention communication. We have shown, that the presented approach almost performs as well in terms of providing the necessary information to convey the current situation and thus allows a human to naturally interact with a faceless mobile robot.

In this work the main focus has been to present the robot's intention to the humans, rather than determining the intentions of the humans. A natural next step (apart from estimating humans intention) is to evaluate how the robot could convey not only its own, but also the estimated intentions of its co-workers and how this would influence the response to the system.

Future work will mainly focus on evaluating what needs to be projected and evaluating the system in an industrial environment upon installing the suitable hardware and conducting experiments over a larger number of subjects. Furthermore, we are planning to implement the presented AR system for human-robot communication in an industrial environment, by augmenting it with the capability to project person specific information and provide an intuitive way to interact with the robot. The key technology with respect to the implementation of our approach is the projection system. For the presented implementation, a standard LED projector was used. We have

plans to experiment with a combination of other technologies such as pico-projectors, laser projectors and holographic projectors.

Acknowledgements

This work has partly been supported by the Swedish Knowledge Foundation under contract number 20140220 (AIR) and the European Commission under contract number FP7-ICT-600877 (SPENCER).

Bibliography

- [1] C. Breazeal, A. Edsinger, P. M. Fitzpatrick, and B. Scassellati, “Active vision for sociable robots.” *smcA*, vol. 31, no. 5, pp. 443–453, 2001.
- [2] T. Fong, I. Nourbakhsh, and K. Dautenhahn, “A survey of socially interactive robots,” *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.
- [3] D. A. Norman, *The design of everyday things*. Basic Books, 2002.
- [4] H. Asada, M. Branicky, C. Carignan, H. Christensen, R. Fearing, W. Hamel, J. Hollerbach, S. LaValle, M. Mason, B. Nelson, G. Pratt, A. Requicha, B. Ruddy, M. Sitti, G. Sukhatme, R. Tedrake, R. Voyles, and M. Zhang, “A roadmap for US robotics: From internet to robotics,” *Computing Community Consortium on Emerging Technologies and Trends*, 2009.
- [5] K. Dautenhahn, “Socially intelligent robots: dimensions of human–robot interaction,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1480, pp. 679–704, 2007.
- [6] J. Bates, “The role of emotion in believable agents,” *Communications of the ACM*, vol. 37, no. 7, pp. 122–125, 1994.
- [7] B. M. Blumberg, “Old tricks, new dogs: ethology and interactive creatures,” Ph.D. dissertation, Massachusetts Institute of Technology, 1996.

- [8] P. Milgram, S. Zhai, D. Drascic, and J. J. Grodski, “Applications of augmented reality for human-robot communication,” in *Intelligent Robots and Systems’ 93, IROS’93. Proceedings of the 1993 IEEE/RSJ International Conference on*, vol. 3. IEEE, 1993, pp. 1467–1472.
- [9] B. Hine, P. Hontalas, T. Fong, L. Piguet, E. Nygren, and A. Kline, “Vevi: A virtual environment teleoperations interface for planetary exploration,” SAE Technical Paper, Tech. Rep., 1995.
- [10] A. Kelly, N. Chan, H. Herman, D. Huber, R. Meyers, P. Rander, R. Warner, J. Ziglar, and E. Capstick, “Real-time photorealistic virtualized reality interface for remote mobile robot control,” vol. 30, no. 3, pp. 384–404, 2011.
- [11] S. Omidshafiei, A.-A. Agha-Mohammadi, Y. F. Chen, N. K. Ure, J. P. How, J. Vian, and R. Surati, “Mar-cps: Measurable augmented reality for prototyping cyber-physical systems,” in *AIAA Infotech@ Aerospace Conference*, 2015.
- [12] S. Livatino, F. Banno, and G. Muscato, “3-d integration of robot vision and laser data with semiautomatic calibration in augmented reality stereoscopic visual interface,” *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 69–77, 2012.
- [13] S. R  ther, T. Hermann, M. Mracek, S. Kopp, and J. Steil, “An assistance system for guiding workers in central sterilization supply departments,” in *Proc. of the International Conference on Pervasive Technologies Related to Assistive Environments*, 2013, p. 3.
- [14] M. Daily, Y. Cho, K. Martin, and D. Payton, “World embedded interfaces for human-robot interaction,” in *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*. IEEE, 2003, pp. 6–14.
- [15] F. Ghiringhelli, J. Guzzi, G. A. Di Caro, V. Caglioti, L. M. Gambardella, and A. Giusti, “Interactive augmented reality for understanding and analyzing multi-robot systems,” 2014, pp. 1195–

1201.

- [16] T. H. Collett and B. A. MacDonald, “Augmented reality visualisation for player,” in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE, 2006, pp. 3954–3959.
- [17] T. Matsumaru, “Mobile robot with preliminary-announcement and display function of forthcoming motion using projection equipment,” in *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*. IEEE, 2006, pp. 443–450.
- [18] F. Leutert, C. Herrmann, and K. Schilling, “A spatial augmented reality system for intuitive display of robotic data,” in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2013, pp. 179–180.
- [19] J.-H. Lee, J. Kim, and H. Kim, “A note on hybrid control of robotic spatial augmented reality,” in *Ubiquitous Robots and Ambient Intelligence (URAI), 2011 8th International Conference on*. IEEE, 2011, pp. 621–626.
- [20] J. Park and G. J. Kim, “Robots with projectors: an alternative to anthropomorphic hri,” in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM, 2009, pp. 221–222.
- [21] N. Costa and A. Arsenio, “Augmented reality behind the wheel - human interactive assistance by mobile robots,” in *Proc. of the International Conference on Automation, Robotics and Applications*, 2015, pp. 63–69.
- [22] M. D. Covert, T. Lee, I. Shindeev, 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.
- [23] E. Pacchierotti, H. I. Christensen, and P. Jensfelt, “Human-robot

- embodied interaction in hallway settings: a pilot user study,” in *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*. IEEE, 2005, pp. 164–171.
- [24] H. Hüttenrauch, K. S. Eklundh, A. Green, and E. A. Topp, “Investigating spatial relationships in human-robot interaction,” in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 5052–5059.
- [25] L. Takayama and C. Pantofaru, “Influences on proxemic behaviors in human-robot interaction,” in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 5495–5502.
- [26] J. Mumm and B. Mutlu, “Human-robot proxemics: physical and psychological distancing in human-robot interaction,” in *Proceedings of the 6th international conference on Human-robot interaction*. ACM, 2011, pp. 331–338.
- [27] C. Dondrup, N. Bellotto, and M. Hanheide, “Social distance augmented qualitative trajectory calculus for human-robot spatial interaction,” in *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*. IEEE, 2014, pp. 519–524.
- [28] H. Andreasson, J. Saarinen, M. Cirillo, T. Stoyanov, and A. J. Lilienthal, “Fast, continuous state path smoothing to improve navigation accuracy,” in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 662–669.
- [29] D. A. Forsyth and J. Ponce, *Computer Vision: A modern approach*. Pearson Publishing, 2003.

Testing (Scale 1-7)

Pilot Experiment 1	Subject Background	Computer Engg	Computer Engg	Computer Engg	Social worker	Researcher	Admin	Admin	Lab Engineer	Researcher	Researcher	Researcher	Researcher	Researcher	Researcher	Sum	Average
Communication	No projector	3	2	2	1	5	3	4	7	3	5	1	3	5	5	46	3.2857142857
	With projector	6	6	6	3	6	6	7	7	6	7	5	5	6	6	81	5.7857142857
Dependability & Reliability	No projector	5	5.5	5	5.5	4.5	5.5	4	7	3	7	2	3.5	3	65.5	4.6785714285	
	With projector	5	6.5	6	4.5	5.5	6	6	7	5	7	6	5	6	80.5	5.75	
Predictability	No projector	6	2	1	3	3	6	5	7	4	4	1	2	5	52	3.7142857142	
	With projector	7	5	6	4	4	5	6	7	5	7	7	7	7	84	6	
Transparency	No projector	3	4	4	4	3	5	4	7	2	4	1	2	6	52	3.7142857142	
	With projector	7	6	6	5	5	6	5	7	5	6	6	6	7	84	6	
Situation Awareness	No projector	4	4	4	4	3	2	5	4	7	3	6	3	3	55	3.9285714285	
	With projector	4	5	5	5	3	5	7	7	4	6	6	5	6	75	5.3571428571	
Pilot Experiment 2	Subject Background	Computer Engg	Computer Engg	Computer Engg	Social worker	Researcher	Admin	Admin	Lab Engineer	Researcher	Researcher	Researcher	Researcher	Researcher	Sum	Average	
																	Computer Engg
Communication	No projector	3	2	2	5	1	5	2	4	7	3	4	1	3	5	47	3.3571428571
	With projector	6	6	6	6	4	5	6	7	7	5	7	5	6	6	81	5.7857142857
Dependability & Reliability	No projector	5	5.5	2	5	3.5	4	3	4	7	2	7	2	3.5	3	56.5	4.0387142857
	With projector	5.5	6.5	6	5	5	5	6	6	7	4	7	5	6	79	5.6428571428	
Predictability	No projector	5	2	2	3	3	2	5	3	4	2	3	1	2	5	42	3
	With projector	7	5	6	5	4	4	5	7	7	5	6	5	7	7	80	5.7142857142
Transparency	No projector	4	4	3	5	5	4	4	4	4	2	4	1	2	6	52	3.7142857142
	With projector	6	6	5	6	4	6	5	7	7	5	5	6	6	7	81	5.7857142857
Situation Awareness	No projector	4	4	2	4	4	3	4	4	4	3	6	2	3	3	50	3.5714285714
	With projector	4	5	6	5	4	6	7	7	7	5	6	6	5	6	78	5.5714285714

Figure .1: Data collected through the Likert scale ratings after the experiments pertaining to the Section 4.1. AR based Intention Communication experiments

HR2018 (PKU) experiment results

Subtype	1		2		3		4		5		6		7		8		9		10		11		12		13		14		
	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	
Human-Human																													
Communication	4	1	4	1	6	0	1	0	6	0	6	0	6	0	7	0	6	0	5	0	7	0	7	0	7	0	6	0	5.2857142857
Reliability	7	6	6	2	7	6	7	6	7	6	7	6	7	6	7	6	7	6	4	7	6	7	6	7	6	7	6	6.1428571428	
Predictability	5	5	5	3	5	5	5	3	5	5	5	5	5	5	7	6	5	5	6	6	7	6	7	6	7	6	3	5.5714285714	
Transparency	5	5	5	5	6	4	7	6	6	6	4	7	6	7	7	6	6	6	6	6	7	6	7	6	7	6	3	5.7142857142	
Situation awareness	5	6	5	6	6	6	6	6	6	6	7	6	7	6	7	6	6	6	7	6	7	6	7	6	7	6	6	6.3771428571	
Human-robot 1																													5.8142857142
Communication	6	0	4	1	6	0	0	0	1	5	0	6	0	6	1	7	2	2	5	2	6	2	7	3	4	3	4	5.4	
Reliability	5	2	3	1	6	0	1	1	5	3	6	5	7	6	6	7	2	4	2	4	2	7	5	7	6	2	5.1		
Predictability	5	2	5	1	7	2	5	3	5	2	5	4	7	6	7	6	6	6	6	6	7	7	7	7	6	6	2	6.1	
Transparency	5	3	5	4	7	2	6	6	5	5	2	4	7	7	7	7	3	6	6	6	7	7	7	7	6	2	5.9		
Situation awareness	6	0	5	2	5	2	5	5	5	5	3	7	5	7	6	6	7	5	5	7	6	7	7	7	4	2	5.9		
Human-robot 2																													4.1
Communication	6	0	4	1	6	0	0	6	1	5	0	7	5	6	4	6	5	2	2	6	3	6	5	6	5	2	5.4		
Reliability	4	0	3	1	6	0	2	2	4	4	5	5	7	5	5	5	5	2	2	5	3	5	4	5	6	5	2	4.6	
Predictability	4	1	5	1	5	0	2	2	5	3	5	5	3	5	6	3	6	3	6	7	6	6	3	6	5	2	4.9		
Transparency	4	0	5	4	5	0	5	5	5	0	5	3	5	7	3	5	7	3	6	7	6	6	4	6	5	2	5.1		
Situation awareness	4	0	5	2	5	1	5	5	5	5	3	6	4	6	4	7	6	5	5	7	6	5	4	6	5	4	5.4		
																													3.4

Figure .2: Data collected through the Likert scale ratings after the experiments pertaining to the Section 4.2. AR based Interface experiments