Implementation and Comfort Assessment of a Robot Behavior

Daniel Jönsson

May 5, 2006
Master's Thesis in Computing Science and Cognitive Science, 20 credits
Internal supervisor at CSC-KTH: Elena Pacchierotti
External supervisor at CS-UmU: Lars-Erik Janlert
Computer Science Examiner: Per Lindström
Cognitive Science Examiner: Timo Mäntylä

Umeå University
Department of Computing Science
SE-901 87 UMEÅ
SWEDEN

Abstract

Robots will become a larger part of our society. They will work along side and interact with people who have no prior experience with robot technology. This demands autonomous robots that fit in and do not cause irritation, stress or fear, but help the human beings they interact with. Social robotic research assumes that people will see robots as human beings, implying that a robot must imitate the behavior of people to the greatest extent.

Research about human behavior is used in order to create a robot behavior which imitates human beings. One specific field in this area is human spatial behavior, addressing how human beings use space when interacting.

A robot behavior, which the master's thesis has extended, is designed to operate in a corridor setting. If the robot meets a person in a corridor, it will maintain a distance to the person in accordance with social rules.

This thesis extends how the robot maintains the distance towards a person by using direction of movement information. The aim of the extension was to make the robot behave more socially. A user study was conducted in order to assess whether the extension made the robot behavior more social.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Robots will become a larger part of our society. They will work alongside and interact with people who have no experience with robotic technology. This demands autonomous robots that fit in and do not cause irritation, stress or fear, but help the human beings they interact with. Social robotic research assumes that people will see robots as human beings, implying that a robot must imitate the behavior of people to the greatest possible extent. Generally, the imitation is achieved by using research results and theory from human psychology. But that is not enough. An implemented robot behavior which is supposed to act "human-like" has to be tested with human beings, to conclude if people find the behavior human-like or not.

But, a large majority of the research consist of very short interaction exposure in a laboratory setting; the longitudinal research in natural settings clearly represents a minority. How people will react to robots in the long run is still unknown; but we will know in the future with, or without research.

This thesis takes part in the social robotic research by extending an existing robot behavior that imitates the human spatial behavior. The existing robot behavior is social in a sense that its behavior is based upon proxemics—the theory of how human beings use distance when interacting with each other. This is necessary as the robot is designed to navigate in a narrow environment occupied by people; the robot must "know" how to maintain distances from human beings in the same way as people do. An example of a narrow environment would be an office corridor; a robot navigating in such corridor would interact more human-like when the robot behavior conforms to the rules of proxemics.

The extension is a robot behavior consisting of a decision module, using direction of movement of an approaching human as an extra parameter. To establish whether or not the extension resulted in a more human-like behavior, subject comfort was assessed repeatedly.

In order to differentiate between the robot behavior including the decision module and the robot behavior not including the decision module, the former will be addressed as "current robot behavior" and the latter "previous robot behavior". In the same manner, this thesis work will be called "current work" whereas the work this thesis is based upon will be called "previous work".

To outline some of the behaviors necessary for a social robot to act human-like in a corridor, a scenario will be used: Two people are walking in a corridor. The first along the corridor and the second across the corridor. They are bound to collide unless one person changes direction or pace. However, in this kind of encounter certain rules can be applied, each with different priority. A rule with a low priority, in most western countries, is to pass the approacher on the right side, i.e. by moving towards the right side. A higher prioritized rule, again in most western countries, is to keep the distance—by all means—towards the approaching person. Now, if the person walking parallel along the corridor is replaced with a social robot, what would the appropriate behavior be then? The same rules apply to the robot as to a human being. The robot would notice a human walking towards the robot's right side. Thus, applying the low priority rule to pass on the right side would be in conflict with the rule with the higher priority to keep the distance—a pass on the right would result in a collision. Therefore a stop would be the most intelligent social behavior on the part of the robot.

The previous robot behavior consist of two different actions which can be taken when meeting a person, i.e., to pass the approaching person on the robot's right side or to stop. This decision of movement action is mostly dependent on which side of the corridor the meeting person is walking—no consideration of the movement direction is taken.

The current robot behavior consists of an information processing (IP) module and a behavior module. The IP module extracts the direction of movement from position data of a walking person. This direction information together with positional data are then used in the behavior module in the decision of how the robot should move. From the scenario above, it is evident that the direction data is crucial when a person crosses the path of the robot.

A user study was conducted in order to answer whether or not the current robot behavior increases the user comfort. The study consisted of exposures where the robot either used the current robot behavior or the previous robot behavior. An attitude questionnaire using Likert-scale was used to assess a subject's comfort level.

As this thesis is interdisciplinary in the sense that it both includes computer science and experimental psychology, the report is divided into two parts; the first of these parts, chapter 2, addresses the implementation, whereas the second part, in chapter 3, addresses the user study.

1.1 Social Robots

How social robots will affect society, how they should behave and where more research is necessary are questions which have been discussed in great detail. In a paper by Butler et al. [5], the authors clearly express an opinion that robots will become part of our society, but that "very little work has been done in trying to understand how people would interact with a robot and what makes them comfortable or uncomfortable, i.e., the aesthetic qualities of the robot's behavior patterns." In addition the authors claim that most robotic research is based on a short term exposure—there is a lack of knowledge regarding the longitudinal effect. A short term exposure can never predict the long term effect of human robot interaction. They also concluded that the best cooperation between a human and a robot can be achieved when the robot imitates human behavior to a large extent. This is also mentioned in [24] where the authors emphasized that a robot has to follow the social rules to be accepted and thus behave as a "human being". Another viewpoint was expressed in [30], that it is also important how robots move—body language is a large part in the perception of a robot as a human.

But why does a robot need to imitate human behavior? The answer is given in [9]: if robots are going to be accepted as a partner to interact with, they need social skills and the ability to recognize social context and conventions. Thus, robots needs social knowledge and acting if they are going to interact with people in their daily lives. However, do we *need* to have robots in our daily lives? In a survey [20], the author describe the idea of an intelligent service robot (ISR), which will have a general use in a domestic setting, assisting elderly and handicapped as well as helping "ordinary" people with their everyday lives. The result of the survey showed a general positive attitude towards the idea of an ISR, saving time and helping elderly as well as disabled. But, a robot realization of this kind requires communication involving interpretation of human intentions; as a 80 year old or a disabled person will probably have a hard time "programming" the ISR. Thus, an ISR need to be socially interactive, and the first question is answered again. In general, this thought was also expressed in [29], when robots will become part of people's daily life, robots will interact closely with people in their everyday environment; implying the need for a robot with the abilities to communicate the way human beings do.

1.2 Distance in Human Interaction

Social robotics research regarding spatial behavior in human robot/interaction is sparse. To have a theoretical framework, psychology theory of human interaction has commonly been applied. A good overview and analysis is given by John R. Aiello in [1] and a more pragmatical view can be found in [28].

One of the research areas in human psychology is human spatial behavior. The research began in the mid 1960s and at the end of 1960, the interest was spurred considerably [1]. A prominent framework of human spatial behavior is proxemics. It was founded by Edward T. Hall, an anthropologist who was greatly inspired by animal studies; but the majority of his research results came from his own reported observation of spatial interaction [16]. The latest definition of proxemics was made by Hall in 1974 [17]: "...the study of man's transactions as he perceives and uses intimate, personal, social and public space in various settings while following out of awareness dictates of cultural paradigms." The definition can be divided into three components: transaction, spatial context labeled "zones", and an unconscious cultural dependency [19]. All the components are of equal importance, but the most interesting component for this thesis is the concept of zones. In a western culture, the zones and ranges have the following general definition [16]:

Range	**Personal Spatial Zone**	**Situation**
0 m–0.45 m	Intimate zone	With a partner or close friends
0.45 m–1.2 m	Personal zone	Conversation with friends
1.2 m–3.6 m	Social zone	Conversation with non-friends
3.6 m+	Public zone	Conversation with an audience

The specific range for each zone is not to be seen as a given boundary. Rather, the transition from one zone to another occurs more gradually. This gradual change has also been formalized by Hall, where each zone contains a near and a far phase [16].

The spatial zones are used in social robotic research as guidelines regarding interaction distance between a robot and people. Usually, the social distance is used between work colleagues and in formal meetings. As a social robot is assumed to be considered as a person, the same distance is assumed to be applicable for human/robot interaction [22].

1.3 Applied Robotic Research with Distance Consideration

Social robotic research has considered spatial behavior in different degrees and with different theoretical frameworks. The theory of proxemics has commonly been used [22, 24, 26, 34], and also, the theory of human territoriality [37]. One example is a robot that enters a group discussion. The robot behavior includes a dynamical adjustment of the distances to the persons in discussion. The relative distances will vary depending on the number of persons in the discussion; four people creates a square, whereas three people creates an equilateral triangle. Also, the robot listens passively by directing its cameras towards the person who is speaking [2]. In a similar common social situation, a robot was programmed to detect a queue and position itself in the queue by recognizing individual silhouettes of human beings. The distance is a critical factor in queuing behavior: Too far away and the robot can be perceived as not standing in the queue; too close and the robot interferes with the personal zone of the person in front [24]. Another study, somewhat unconventional, observed how people perceived a robot communicating with body expressions; the movements of the robot were inspired from classical ballet and were assumed to make movement impressions familiar to human beings [22]. Three studies with the goal of making a robot passage behavior, started with naturalistic observations of how people and a researcher walked passed each other. The purpose was to create an algorithm which described a by-passing trajectory. Then the algorithm was implemented to create an initial robot passage behavior [36]. The second study included further naturalistic observations which resulted in a pass-by algorithm based upon the most common by passing behavior [35]. Finally, the last experiment tested the by pass algorithm created in the previous study. The tests consisted of a robot which passed by both a subject stationary and a subject walking [37]. Another study, on which this thesis is based, was carried out at KTH, CVAP/CAS; it consists of a robot able to navigate in a corridor environment. More importantly, the robot dynamically applies proxemics when passing a person in a corridor by constantly adapting the by pass trajectory from the relative distance of the passing person [26].

1.3.1 Geometrical Form of the Zones

The geometrical shape of the personal spatial zones have been addressed by several studies. The study [26], carried out at KTH, CVAP/CAS, used a model of the social distance shaped as an elliptical region around a person, where the major axis is in parallel with the direction of walking. A video study in a corridor environment indicated that this elliptical model might be correct [7]. Further, an elliptical form of the zones has been mentioned by Hall [16] and in the study of human territoriality, the personal space has been modeled as an oval shape. Also, the same shape was the result of an experiment aimed to ascertain the form of the personal space concerning people standing in a queue [24].

1.4 CVAP/CAS

Much work and the study were carried out at CVAP/CAS (The Computational Vision and Active Perception Laboratory)/(The Center for Autonomous Systems), Kungliga Tekniska Högskolan (KTH), Stockholm, where Henrik I. Christensen is the director.

CAS is a research center which conducts research in autonomous systems including mobile robot systems for service, domestic and field applications. CVAP is a research group at the School for Computer Science and Communication (CSC). The group consists of about 30 people and does research in computer vision and robotics as well as on related problems in geometric modeling and computations.

1.5 Purpose

The purpose of this thesis is to implement and evaluate a robot behavior extension. The extension consists of extracting the direction of movement from the positional data of a moving person, and use that direction information in the decision module. To determine the benefit of the robot behavior extension, a user study will be conducted. The benefit is to be ascertained from subject's comfort level, using a questionnaire. The main hypothesis is that the current robot behavior will be perceived as more comfortable than the previous robot behavior.

1.6 Goals

The goals of the first part of the thesis are to:

– Implement an algorithm that extracts direction of movement from positional data.

– Create a decision module for a corridor environment using direction of movement.

– Integrate the decision module with the existing robot behavior implementation.

And the goals of the second part of the thesis are to:

– Create a user study that determines whether or not the robot behavior extension enhanced the general level of subject comfort.

– Conduct the user study with at least 20 participants.

1.7 Outline of Thesis

Chapter 2 concerns robot behavior implementation, describing the work flow of how the movement decision module implementation was effected. Chapter 3 describes how subject comfort attitude was ascertained by means of a comfort assessment. Chapter 4 presents conclusions and considerations of the work. The acknowledgments and an appendix concludes the thesis.

Chapter 2

Robot Behavior Implementation

Several ideas were discussed regarding what kind of extension could be implemented together with the previous robot behavior. The thesis was to be held within certain limits based upon previous work: the extension would consist of a robot behavior imitating the spatial behavior human beings use in a corridor environment. A final limit was the criterion to include an implementation *and* a user study in the problem of the thesis.

The thesis' aim was thus to enhance the previous work with a spatial robot behavior not yet implemented and make a user study regarding the implementation. To find a suitable behavior to implement, a discussion about the assumptions of the previous work was brought up; extending the previous robot behavior with the consideration of more than one person was an excessively complex and time consuming task—even without a user study; the assumption of constant walking speed does not impose a large limitation to the robot behavior, and additionally simplifies many calculations; but, assuming that a moving person always has a direction of movement parallel to the corridor wall does give rise to problems; and, luckily enough, addressing the task of including direction information in the previous robot behavior seemed feasible.

When a closer look was taken at the chosen problem, it turned out to be a suitable extension for several reasons. First, the position of a moving person was already available from a motion detection implementation. Second, it is a common situation that a person walks across a corridor, e.g., walks from one room to another using the corridor of a office. Third, the problem was appropriate for a user study and an implementation within the time limit.

This chapter begins with a description of the previous work. This work includes the assumptions, the software used implementing the previous robot behavior, an explanation of the previous robot behavior, and a description of the robot. The implementation is thereafter explained with its prerequisites, gait sampling for understanding the data, filter creation because of data noise, the nature of this noise, and finally the algorithm which explains the conversion from positions to directions. The resultant robot behavior using the direction information is then described together with a more visual explanation of the algorithm. The chapter ends with a discussion.

2.1 Previous Work

This thesis' decision module implementation is based upon the work in [26]. The decision module was developed utilizing the software and the robot platform from previous work.

In the following paragraphs, the software and the robot platform are mentioned first. Thereafter, previous work's assumptions and robot behavior are addressed.

2.1.1 Software Platform

A software platform called Player/Stage was used when developing the current and the previous robot behavior. Player is a language and platform independent network server for robot control. Stage simulates a population of mobile robots in a 2D environment.

Together, Player/Stage is an interface which is standardized for steering and controlling robots, with support for the most common type of actuators and sensors. The software is open source and is released under GNU GPL, making it possible for researchers to use it in any way as long as the software developed also is released under GNU GPL. In this way, the code can be reused and ported [33].

2.1.2 Robot Platform

The robot used in this thesis is a PeopleBot from ActivMedia Robotics [8] seen in figure 2.1(b) and 2.1(a). It has several sensors: front sonars 1.1 m above the floor; circumferential sonars 0.2 m above the floor; a web cam which can be pitched and yawed; speech interface; and a SICK LMS laser 0.3 m above the floor.

In the robot behavior extension, only preprocessed motion data were used, originating from the laser. The technical specifications of the laser is a sampling rate of 5Hz, 180^o field of view with 360 samples in the distance range 0.2 m to 8 m. Since the laser is positioned close to floor, the laser scans in lower leg height.

The sampling frequency of the laser could have been 20Hz, but due to an unstable USB interface communicating with the laser, the lower sampling frequency of 5Hz had to be used. Consequences of this lower sampling frequency will be discussed in greater detail in chapter 4.

2.1.3 Assumptions

The previous work was based on the following assumptions:

- The robot operates in a corridor environment.
- Only one person at a time is present in the corridor.
- The person detected is assumed to have constant speed.
- The person detected is assumed to walk parallel with the corridor walls.

2.1.4 Robot Behavior

The previous robot behavior has two basic modes:, obstacle avoidance and person passage, which will be described in the following paragraphs.

To begin with, the robot is positioned in the middle of a corridor. After the robot has localized itself in the environment using a map, it starts to move towards a predefined

(a) Close up against the wall.

(b) Together with the author of this thesis.

Figure 2.1: The robot, a PeopleBot from ActivMedia Robotics, used in this master's thesis.

goal. When the goal is reached, the robot stops. The aim of the robot is thus to reach the predefined goal and avoid (static) obstacles.

When movement is detected, the robot changes behavior mode from obstacle avoidance to person passage. Depending on several metric variables, different passage behaviors come into effect. A basic principle for the robot behavior is not to violate the personal zone of a person. Thus, the robot only moves as long as the relative distance towards the person, is kept in accordance with proxemics. If the person is walking somewhere along the robot's left side and there is enough space to pass at a calculated passage point, given the walking speed of the human, a passage movement is executed. To signal to the person that the robot has seen the individual and wishes to pass, a turn towards the robot's right is started as early as possible.

However, if the lateral distance at a theoretical passage is too narrow, or the walking person is detected too closely, the robot stops. The reason for this behavior is that one of the previous mentioned circumstances could result in a violation of the social zone, which is against the ground rule of the robot behavior. The robot will also stop if the person is walking on the robot's right side of the corridor, as a passage on this side is preferable in the western culture.

After an executed passing or when a movement has disappeared, the robot switches back to obstacle avoidance mode and continues its move towards the goal.

To summarize, the robot's default mode is to move towards a predefined goal while avoiding obstacles. If a person is detected, a dynamic passing behavior is executed with the aim of always passing on the robot's right side of the corridor, or to stop if a pass is inappropriate. When movement no longer is seen by the robot, the mode is switched back to the default.

2.2 Direction of Movement Extraction

This section will first state the prerequisites, describing when the implementation is supposed to work and, implicitly, when it does not. Second is the acquisition and analysis of the positional data described which led to the filter implementation. Finally, the positional data properties are discussed to shed light on the last part, the direction of movement (DoM) extraction algorithm.

2.2.1 Prerequisites

The robot is thought to operate in a corridor environment, where a maximum of one person is walking. Keeping track of one person is easy. Keeping track of many is extremely complex. This topic will be further discussed in section 4.2.

The gait should be "natural", i.e. an individual should walk in the manner preferred when walking in general. This is assumed to imply an approximately constant speed.

Also, the corridor where the robot and a human being interact has to be specified to the decision module as two parameters. The first parameter represents the center line of the corridor and is given as the x *or* y position. The second parameter states the direction of the corridor. Both parameters have to be in the world frame. Together, these parameters are utilized in transforming subject movement positions to a robot frame. The transformation makes the calculation of which side a person is on and of the DoM, relative the robot, more convenient.

2.2.2 Position Data Acquisition and Analysis

The goal was to calculate a DoM, from samples that contained consecutive positions of a person. The collected data originate from a motion detection algorithm implemented in the previous work. The technique behind the motion detection is easy to describe: Two consecutive laser scans are compared, and motion is detected. Due to low vertical height placement of the laser, motion detected consists of the two positions of the legs. The movement of the robot has to be taken into consideration in the computation of the legs' position in the world frame, resulting in a position with an absolute reference. Calculation of DoM is done using consecutive center of masses (CoM) which in turn is calculated from the mean position of the legs. This is described more thoroughly in section 2.2.3

The acquisition of data was accomplished with the author walking in front of the stationary robot. The walking mode was performed with the use of predefined trajectories which were sketched beforehand on paper. The sketches were later used in estimating the performance of a filter, applied as "true" data. Each trajectory was walked at two different speeds: the author's normal walking speed, and a speed the author found distinctively slower. Rough estimates were calculated as being 0.8 m/s and 0.5 m/s for the normal and the slower walking speed, respectively. The slower walking speed was used as it was a probable speed of a subject. The topic of human walking is also mentioned in section 3.2.3. Additionally, it was of interest if walking speed effected the signal to noise ratio and leg occlusion. A faint indication showed that signal to noise ratio decreased and that leg occlusion was unchanged with the slower walking speed compared to the normal.

2.2.3 Filter Implementation

As Matlab was used in data acquisition and analysis, the DoM implementation continued in Matlab.

The filter implementation began with applying an ad hoc filter to the calculated DoM in order to handle noise. Alas, the results did not live up to expectations. The next attempt was to utilize a finite impulse response (FIR) filter . A FIR-filter gives a weighted mean value from a sampling window of a certain length, where each sample is weighted individually [15].

Two different implementations were considered. In the first, DoM data were computed from the raw position data (CoM) and then filtered with FIR. In the second implementation CoM was filtered prior to calculation of DoM which was not filtered. The comparison was made to ascertain at which stage in direction extraction the FIR-filter was to be applied: on the DoM-data or the CoM-data. The comparison of the two implementations revealed that applying the FIR-filter to CoM data gave the best result.

In spite of the improvements from the FIR filter regarding the signal to noise ratio, there were still problems to solve. Problems involved more than common signal theory could solve—the FIR-filter could not alone handle the noisy data adequately. Several of the ad hoc ideas, used from the start, were brought back for reconsiderations and finally included with alterations.

Before the resultant algorithm is explained, the properties of the positional data will be discussed. This will shed light on the ad hoc filtering design in the algorithm.

Position Data Properties

The raw positional data contain a lot of noise due to different disturbances. Generally, these disturbances can be divided into three categories: missing data, missing leg, and abnormally large positional movements.

The first category is that data seem to disappear in different ways. The laser may fail to detect an object (a leg in this case) depending on its distance and geometry. Moreover, the motion detection algorithm may fail to detect a moving object if the motion is below a certain threshold.

The second category is a missing leg, which occurs due to occlusion. This seems to be due to the relative direction and position between the robot and a moving person; e.g. a person walking in a lateral direction in front of the robot will regularly occlude the most distant leg behind the closest leg. Two missing legs do not occur due to occlusion, but rather from the first category of problems.

The third category is abnormally large positional movements, which can be caused by three different cases. The first case is that a missing leg will result in a positional shift as the CoM-calculation calculates the mean using one raw position instead of two. The second case consist of a slight rotational shift occurring naturally between two consecutive laser readings. This small rotation difference can result in a large distance difference, e.g., when a first laser beam is reflected from an edge of an object and the consecutive laser beam misses the edge and is reflected from another object further away. The third case is when the laser scans a surface of glass which sometimes lets the laser beam pass trough and sometimes reflects it. In the first case it is still possible to track the person's position whereas the tracking is lost in the second and third case.

Algorithm

The calculation of DoM begins with an ad hoc-filter, continues with a FIR-filter and ends with another ad hoc-filter. These three filters will be explained at first and then summarized with an algorithm description.

The initial filter contains two parts. First, CoM is calculated and second, raw positional disturbances are corrected. The CoM calculation creates position data for a moving person, p_i, at a given time, i. p_i is dependent on three parameters: number of legs found, $numLeg_i$, and the position of each leg, $posLeg1_i, posLeg2_i$. The calculation of p_i is in general the mean of the leg positions, $p_i = mean(posLeg1_i, posLeg2_i)$. If only one leg is present, $p_i = posLeg1_i$. When no legs are found, the previous CoM is used, $p_i = p_{i-1}$. The reason behind this needs a digress.

If the positional data for a person is lost in the range of the robot's field of view, there is only one proper assumption: the person is stationary at the previous position. Consequently, the robot will continue to move *as if* a person existed in that position, thus maintaining the social distance in the best possible way. If the robot behavior would have assumed that the person suddenly had disappeared, it would have returned to obstacle avoidance mode and treated a possible stationary person as any other obstacle—not maintaining the social distance.

Also, to be correct, CoM is an estimate of the "real" center of mass. A "more" correct definition of this implementation's CoM that it consist of an intersection of the lower legs in the laser plane. The deviation of CoM and the real center of mass can be seen as negligible considering the resolution of the positional data.

When the CoM has been calculated, the positional movement is evaluated if it is a non-consecutive sample: if two consecutive positions are further than 0.12 m apart,

p_i is defined as *not* being the *real* sample value which followed p_{i-1}; thus, p_i is a non-consecutive sample. If p_i is defined as a non-consecutive sample, and the p_{i-1} calculation was based on at least one leg, p_{i-1} will be replaced. The reason p_{i-1} is replaced and not p_i, is that p_{i-1} is replaced with a value interpolated linearly between the positions p_{i-2} and p_i. An extrapolation of the position (the *real* position) would have been better, but an interpolation is much easier to implement and was assumed to result in practically the same DoM.

An implication of not allowing CoM movements > 0.12 m in 0.2 s (5 Hz sampling frequency), is that the speed during this time window, $0.12/0.2 = 0.6$ m/s, is defined as too fast, even though it is about half the normal walking speed. (See section 3.2.3.) But, with this low threshold, the resulting DoM is smoothed. This double functionality of decreasing the effect of large movements *and* DoM smoothing was not the initial purpose, but was discovered when trying different threshold settings.

When the positional data have passed the ad hoc filter, the FIR-filter is next. A FIR-filter of order n consists of the scalar coefficients $h(n)$. These coefficient values were set with s low-pass filter values, using 0.2 as a cutoff value. The cutoff value set gave the best overall response: not too slow adapting to "normal" directional changes, but also somewhat stiff, removing quick "non-normal" directional changes. The coefficients are multiplied with the input data $p_{[i,i-n]}$ and summed, resulting in a filtered position, pf_i. As the order of the filter is $n = 5$, the last five consecutive positions are used when creating the filtered position.

After the first ad hoc-filter and the FIR-filter, the only step left is the DoM calculation involving the last ad-hoc filter. There are two different ways of calculating DoM: use previous DoM or use current filtered positions, pf. The previous DoM is used when the difference between two consecutive pf are zero. If the difference is not zero, DoM is calculated using $atan2()$ with the relative difference in $pf_{i-1} - pf_i$.

The above explanation of how the DoM calculation is achieved, is summarized with the following algorithm description:

1. Calculate CoM (position), p_i:
   - (a) If $numLeg_i == 2$
     Let $p_i = mean(posLeg1_i, posLeg2_i)$
   - (b) If $numLeg_i == 1$
     Let $p_i = posLeg1_i$.
   - (c) If $numLeg_i == 0$
     Let $p_i = p_{i-1}$.
2. Evaluate if p_{i-1} is a consecutive sampling:
   - (a) If $numLeg_i > 0$ and $|p_i - p_{i-1}| > 0.12$
     Let $p_{i-1} = interpolate(p_i, p_{i-2})$
3. FIR-filter the latest 5 positions, $p_{[i,i-5]}$
   - (a) Let $pf_i = \sum(h_{[1,5]} \times p_{[i,i-5]})$
4. Calculate DoM_i using $pf_i - pf_{i-1}$
   - (a) If $pf_i - pf_{i-1} \neq 0$,
     Let $DoM_i = atan2(pf^x_i - pf^x_{i-1}, pf^y_i - pf^y_{i-1})$
   - (b) Else let $DoM_i = DoM_{i-1}$

2.3 Decision Module Development

The decision module is based upon human spatial behavior in a sense that a pass on the approacher's left side is preferred in western culture. However, the largest part of the decision module was designed through tests and assumed behaviors as no information was found in the literature of human spatial behavior.

The main idea of the decision module is to choose the best movement action based upon the two variables for a walking person: DoM and corridor side. Several different scenarios are possible when the robot, for example, is moving along the corridor and a person suddenly appears: a person moves across the corridor from one room to another in different directions, a person moves out of a room to walk in the same direction as the robot on the left and the right side, et cetera.

Developing the decision module resulted in a division: a movement action decision submodule (MAD) and a movement action filter submodule (MAF). These two submodules will be explained in the following paragraphs.

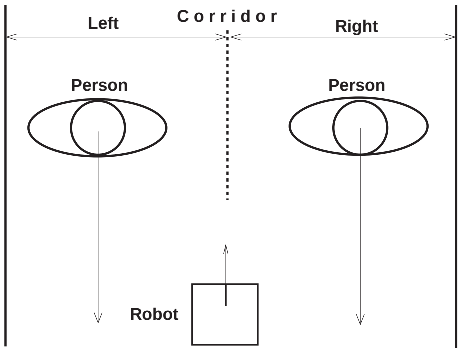
2.3.1 Movement Action Decision

Explaining the MAD will be done in a stepwise fashion, starting with the action decision of the previous robot behavior and ending with the MAD of the current robot behavior.

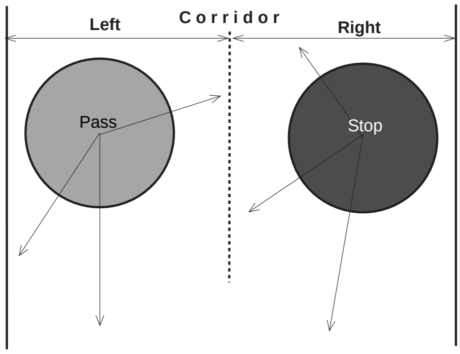
The previous robot behavior assumes that an approaching person is walking parallel to a corridor wall. This is depicted in figure 2.2(a), in an approaching robot's perspective. Additional assumptions are that different movement actions are executed depending on which side the person is moving on. A walk on the right side results in a stop and a walk on the left side results in a pass, irrespective of direction. This independence of direction is depicted using a DoM action diagram (DAD) for each corridor side in figure 2.2(b). A DAD is like a pie chart, where every segment covers an interval of DoM. Consequently, the sum of the segments will always be 360^o. Also, every segment represents which movement action that shall be executed in its directional interval. Each DAD (a circle), representing the previous robot behavior in the figure, contains only one segment; thus, regardless of direction, a movement on the right corridor side will result in the movement action stop and a movement on the left corridor side will result in a pass. This is also shown from the arrows in the figure, coming from the center of each DAD. The arrows represent possible movement directions of a person and the segment under the arrow represents which movement action will be chosen.

The first version of the decision module, of the current work, is depicted as two DADs in figure 2.3(a). As the DADs in the figure show, DoMs toward the right result in a stop and DoMs toward the left result in a pass, irrespective of corridor side. However, when walking parallel to the corridor walls, as the arrows in the figure represent, problems occur due to a natural direction fluctuation in an interval β, shown in figure 2.3(b). The consequence of the direction fluctuation is an ambiguity in the resulting movement action as it alternates between stop and pass.

A common scenario for a robot operating in a corridor setting is to meet a person walking parallel with the corridor walls. Solving the ambiguity problem was therefore of interest. The problem was solved with the addition of an angular interval, α, shown in figure 2.4(a), in which the fluctuations in DoM are ignored. As the figure shows, α is a DoM interval designed to enable walk along a corridor that does not result in an erroneous movement action.

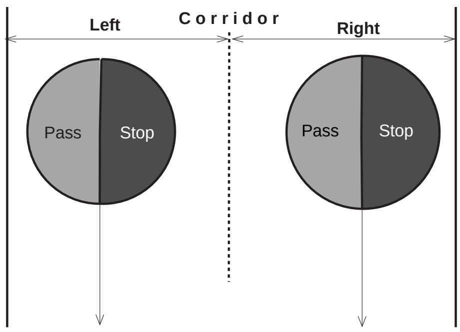


(a) A walk along the corridor.

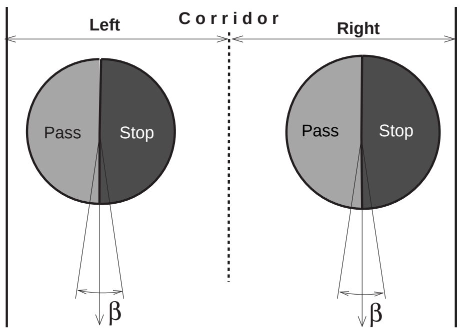


(b) Two DADs showing that irrespective of DoM, the movement action is only dependent on corridor side.

Figure 2.2: Movement action decision of previous robot behavior.

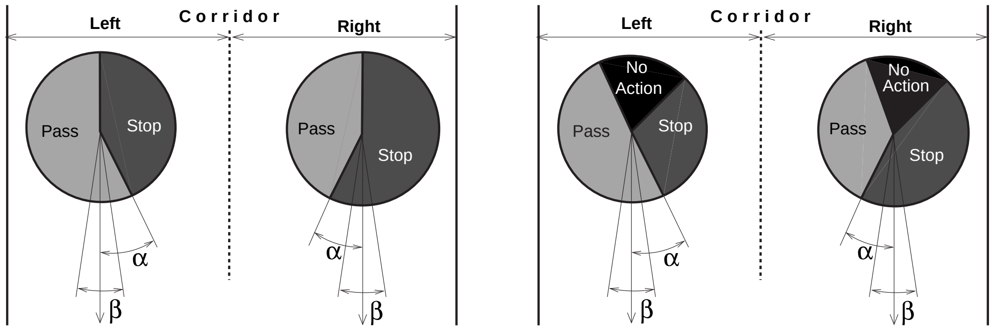


(a)



(b)

Figure 2.3: Initial MAD where a walk along the corridor resulted in a fluctuating movement action due to β.



(a) MAD with the "no action" movement action.

(b) MAD without the "no action" movement action.

Figure 2.4: DADs representing the MAD, taking β into consideration.

From testings, it was shown that $\beta \approx 10^o$; therefore, a fluctuation of 5^o on each side of the main direction was expected. So an α had to be set to meet the criteria $\alpha > \beta/2$. With $\alpha = 13^o$, a margin interval of $13 - \frac{10}{2} = 7^o$ exist and no unwanted movement actions were executed.

Additionally, it was believed that neither pass nor stop should be executed if a person walked in the same direction as the robot, regardless of corridor side. Thus, the "no action" movement action was added as a small segment in both DADs, illustrated in figure 2.4(b). However, the additional movement action was not tested in the user study, but it will be discussed in chapter 4.

2.3.2 Movement Action Filter

In the final implementation of the decision module, a MAF submodule was added. The creation of the MAF arose from two facts: first, fluctuation in the resulting movement action still existed when walking in a direction near the border of movement actions in a DAD; e.g., such as a walk along a corridor towards the robot. Second, the previous robot behavior states: once the movement action stop is executed, a change of movement action is not allowed until movement is out of sensor view.

A consequence of these two facts was that a DoM would occasionally result in a stop, and thus the robot would stop in the middle of a passing maneuver.

To cope with the movement action fluctuations, MAF was created to filter the output from MAD. The output from MAF is a movement action state, initially undefined. The state will only change if three equal and consecutive inputs (ECI) from MAD are received by MAF. Tests revealed that if the number of ECI to MAF was set to three, it would result in a defined movement action state. More than three did not give improvements large enough to motivate the increase in delay and less than three resulted in an erroneous fluctuating state.

A block diagram of the system is shown in figure 2.5. The figure sums up the data flow originating in the raw readings of the laser to the final movement action state by MAF. What can also be seen in the figure is what the decision module consists of.

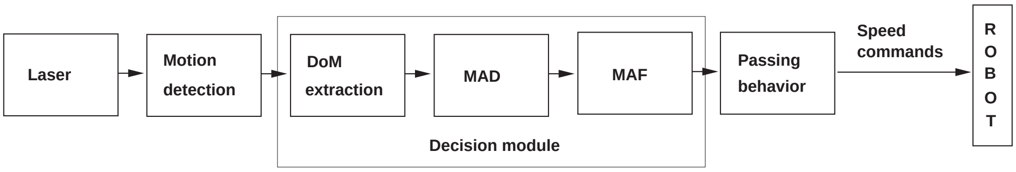


Figure 2.5: The data flow in the current robot behavior from the laser to the speed commands. The decision module is seen in the middle.

2.3.3 Delay

The DoM filter and the MAF in the decision module have sample windows of different sizes. An input to one of these modules will not result in an output until the sample window is filled. A consequence is two types of delays: an initial delay before input results in an output delay and a reactive delay. The initial delay has the effect that the output is undefined until the sample windows are filled. The reactive delay results in a sustained effect from an old sample. The response rate will thus decrease with a large sample window. The current robot behavior has to operate in real-time, and thus general goals have been to minimize any delays and window sizes.

An interesting question: how large is the delay of the decision module? The DoM filter has a sample window of five whereas the MAF has a sample window of three. With the software running at 10 Hz, the total theoretical initial delay is $0.1 \times (5 + 3) = 0.8$ s.

The reactive delay can also be calculated. From the DoM filter, the reactive delay is roughly the same as its initial delay; estimates measured in graphs point to a delay of ~ 0.7 s on average. This is depicted in section 2.4. The reactive delay of the MAF is the time it takes for MAF to change state, the same time to define the state: 0.3 s. In total, the reactive delay is ~ 1.0 s.

To answer the question, the delay of the decision module is the summated delay of the reactive and the initial delay $0.8 + 1.0 = 1.8$ s. However, this delay is not fix, as it will vary depending on the MAF. Thus the decision module can be defined as having a delay of ~ 1.8 s. This delay means that it takes ~ 1.8 s for the decision module to define a movement action state from the instant a moving person is within motion detection range. If the person changes DoM in such way that activation of a different movement state is required, the delay is ~ 1 s.

From tests with the robot a rough estimative delay of ~ 2 s on average was estimated, and is fairly similar to the theoretical delay of ~ 1.8 s. The cause for this small discrepancy between the two delays resides in fluctuations of the input to DAF. If three movement actions, which all are the same, are not received by the DAF in a row, the movement action state will not change.

2.4 Results

First, the results of the DoM filter will be explained by showing how each algorithm part contributed to the final result. Second, the result from the decision module will be explained.

2.4.1 Direction of Movement Extraction

Describing the results of the DoM filter is achieved by dividing it into the four (1–4) algorithm description parts from section 2.2.3. The results are explained in two different ways: incrementally, where the algorithm parts are successively utilized, from a completely incapacitated filter to a fully implemented filter; and comparatively, where combinations of three algorithm parts are compared against all four parts.

Throughout the results, a trajectory depicted in figure 2.6 will be used. The figure shows a top view of the trajectory, depicting a walk back and forth in the form of a rectangle in dotted lines. The figures of the results will also apply the following denotation: rotated crosses denote raw values not manipulated by a filter; dashed lines denote the result from an incapacitated DoM filter; solid lines denote the result of a fully implemented DoM filter; dotted lines denote a "real" or correct value. Additionally, the legends in the figures will apply the following syntax: the word "filter" is followed by four characters representing the utilizing and non-utilizing of a an algorithm part by the number of the algorithm part and an "x", respectively; for example, "filter 12x4" means an incapacitated filter where the third part of the algorithm was not used when the data was filtered, but the other three parts were used.

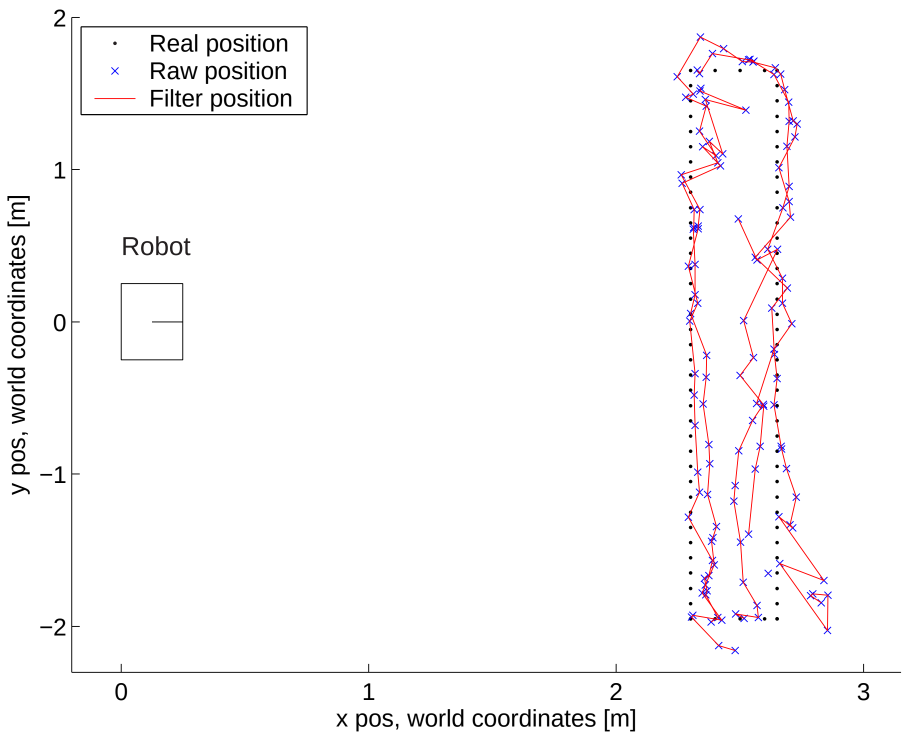


Figure 2.6: The trajectory, from a top view, used throughout the results. The robot is seen on the left, whereas the dots, the points and the line represents the "real" walked trajectory, the raw sampled trajectory and the filter sampled trajectory, respectively.

The trajectory in figure 2.6 can also be depicted as the two CoM coordinates dependent on time, shown in figure 2.7. Walking was performed at a right angle relative to the robot and thus most of the DoMs should have been 90^o or 270^o, depicted with the two dotted lines in the lower diagram of the figure (2.7). However, this is neither the case for the CoM data nor for the filter data. The figure depicts a filter which only calculates the DoM without any other considerations, i.e., a "filter" that does not utilize any of the algorithm parts. What is clearly seen, is the missing data in the interval around 7 s and 25 s, respectively. Also, the DoM is very often 0^o, which is caused by duplicate samples from the motion detection module. This was discussed earlier in section 2.3.3.

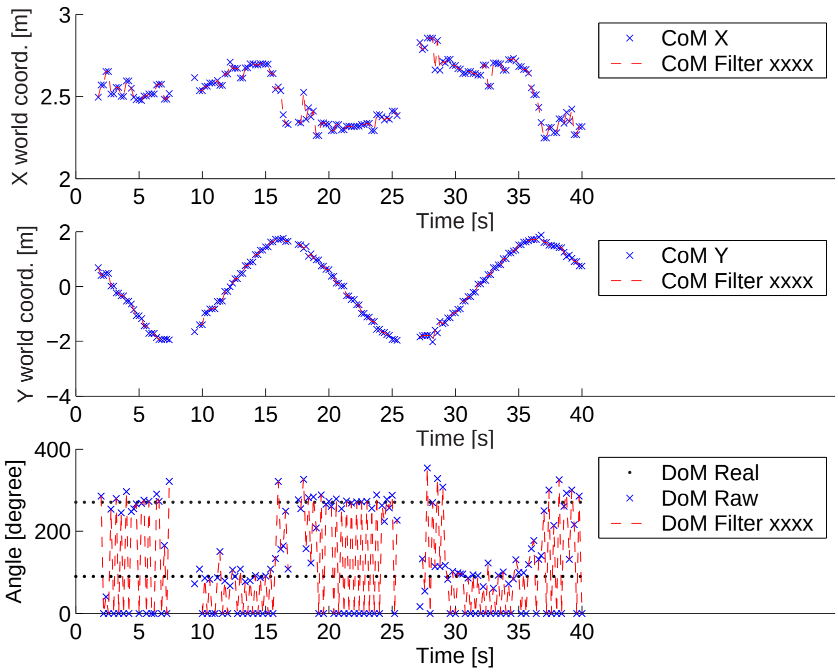


Figure 2.7: Top shows change of the x coordinate in CoM data and CoM filtered data. Middle shows change of the y coordinate in CoM data and CoM filtered data. The bottom shows the DoM from the CoM data, the DoM filter and the "real" DoM. The filter was completely incapacitated shown as "xxxx".

Figure 2.8 depicts the addition of algorithm part 1.c. Comparing the figure with the previous figure 2.7, it is noticeable that the missing positional data is replaced by positional data which do not change over time. As a result of this non-motion the resultant DoMs become zero; they should be undefined, but from the definition of $atan2()$ they are zero.

Continuing, with the addition of algorithm part 2, the result is that of figure 2.9. Now, the biggest change from the previous figure is the decrease in the erroneous DoMs of 0^o. However, several DoMs are still inaccurate in some intervals.

Figure 2.10 shows the contribution of the FIR filter, part 3 in the algorithm description. Nearly all errors are gone, and the outliers are no longer considered, e.g., the cluster of four raw DoMs in the angle interval of 300^o to 400^o and the time interval of 25 s to 30 s. However, a reactive delay is brought in, due to the sample window of the FIR filter. The two uppermost graphs of the figure clearly show this reactive delay as a right shift in the filter CoMs compared to the raw CoMs. This reactive delay was

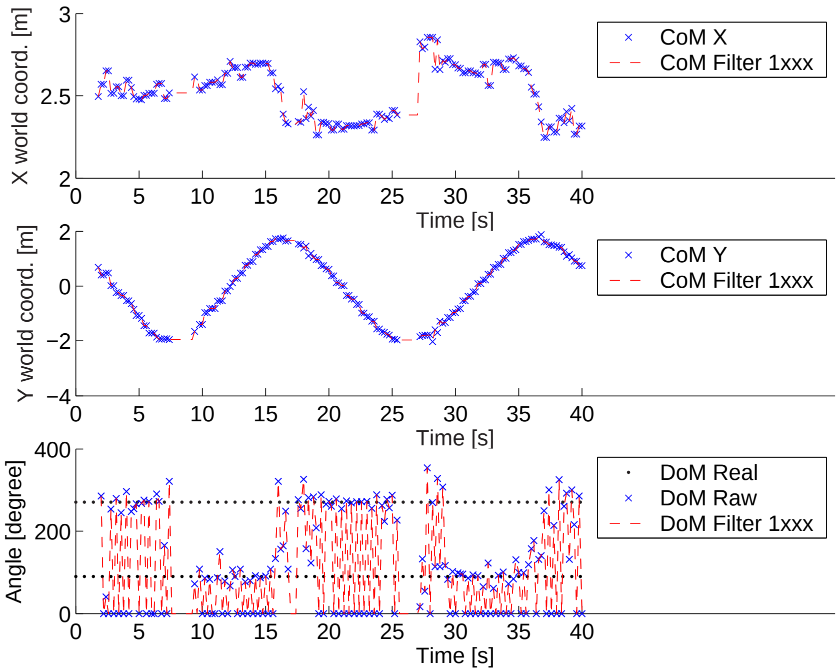


Figure 2.8: Top shows change of the x coordinate in CoM data and CoM filtered data. Middle shows change of the y coordinate in CoM data and CoM filtered data. The bottom shows the DoM from the CoM data, the DoM filter and the "real" DoM. The filter had the first part of the algorithm activated, seen as "1xxx".

mentioned in 2.3.3. The reactive delay can also be seen in the DoM data, although it is more difficult to notice.

Finally, figure 2.11 depicts all algorithm parts of the implemented DoM filter. When comparing the figure with the previous figure, it will be evident that the removal of the DoMs incorrectly set to 0^o are discriminative.

But, what is the contribution of each algorithm part? The answer to this question will be evident from a comparative view seen in figure 2.12. The figure depicts three comparisons between an incapacitated filter (dashed line) and the implemented filter (solid line). The uppermost graph's incapacitated filter lacks part 2 of the algorithm description; the middle graph's incapacitated filter lacks part 3 of the filter and the last graph's incapacitated filter lacks part 4.

It is clear that the middle diagram in the figure, depicting the absence of the FIR filter, has the largest amount of deviated dashed curves; the FIR filter is thus the most important part of the DoM filter. Second best is part 2 of the algorithm 2, shown in the uppermost diagram; and the smallest addition to the final result is given by the final fourth algorithm part.

2.4.2 Decision Module

The final result of how the decision module worked are shown in figure 2.13 and 2.14. Each figure is divided into three subfigures, representing three different time steps, and

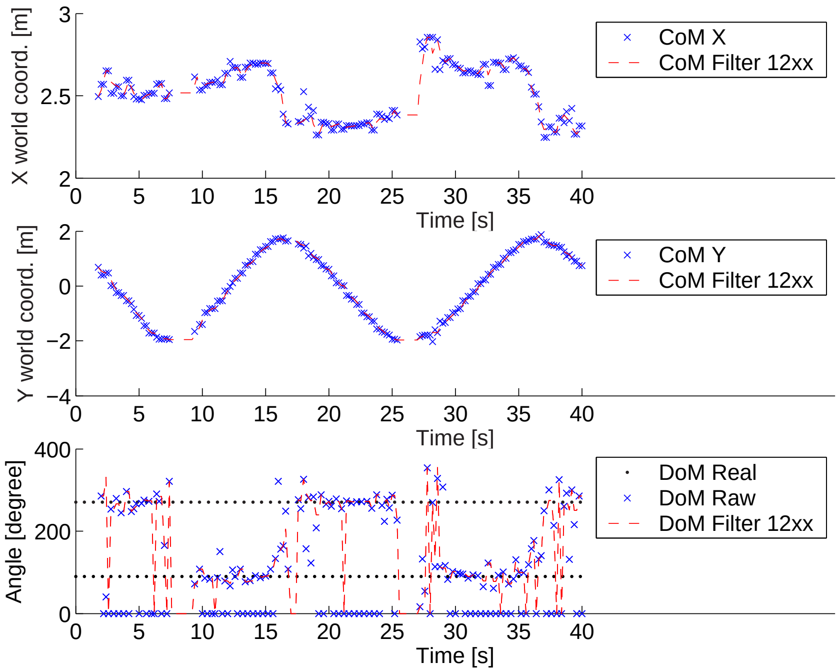


Figure 2.9: Top shows change of the x coordinate in CoM data and CoM filtered data. Middle shows change of the y coordinate in CoM data and CoM filtered data. The bottom shows the DoM from the CoM data, the DoM filter and the "real" DoM. The filter had the first and the second part of the algorithm activated, seen as "12xx".

each show four different positions. The circles and the crosses represent the position of the previos robot behavior and the current robot behavior, respectively. The rotated crosses and the stars represent an approaching person's position when exposed to the previous robot behavior and the current robot behavior, respectively. The time between each positional difference for both the person and the robot is 0.5 s.

The subfigures furthest to the left in figures 2.13 and 2.14 show the positions before the decision module has defined a state. The central subfigures clearly show that the current robot behavior detetects the DoM and that the previous robot behavior does not, as movement actions carried out are opposite one another; in the central subfigure of figure 2.13, it is clear that the previous robot behavior stops whereas the current robot behavior performs a pass; in the central subfigure of figure 2.14, the previous robot behavior starts a pass when the approaching person is very close, whereas the current robot behavior stops. The subfigures furthest to the right depict the resulting positional difference between the robot behaviors. In both figures, the robot behaviors are four positions apart, which is approximatly a 2 s delay.

2.5 Discussion

This section covers the discussion about the implementation; chapter 3, that addresses the user study, will have its own discussion.

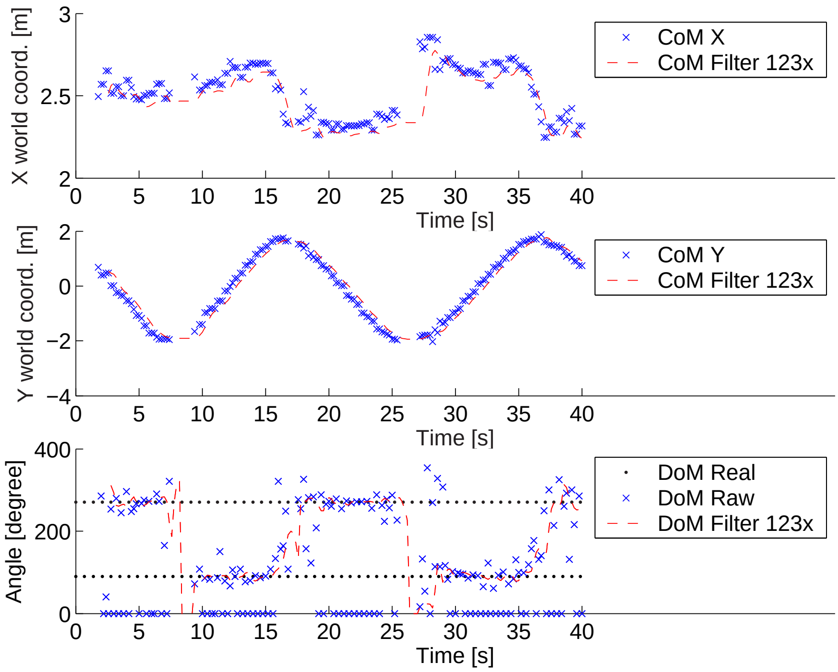


Figure 2.10: Top shows change of the x coordinate in CoM data and CoM filtered data. Middle shows change of the y coordinate in CoM data and CoM filtered data. The bottom shows the DoM from the CoM data, the DoM filter and the "real" DoM. The filter had the first three parts of the algorithm activated, seen as "123x".

2.5.1 Equipment Changes

If the laser's sampling frequency had been as high as possible, 20 Hz, a lot of things could have been done differently.

For example, the total delay from the decision module would have decreased to a quarter, on average 0.5 s; this time is small, perhaps close to the reaction time a human being has, and would have improved the decision. Additionally, the increased sampling frequency would probably have increased the accuracy of the DoM filter, even though the sampling window was of the same size. This is because the DoM range that has to be considered is smaller and because of the need of accuracy. In other words, a high sampling frequency will only sample smaller changes in DoM, whereas a lower sampling frequency has to deal with both small and large changes in DoM.

Another approach with a 20 Hz sampling frequency, would be to double the buffers and thus half the delay to 1 s. The result from the DoM filters with larger sample windows can be seen in figure 2.15. The solid lines in the figure's graphs display the results from the final DoM implementation. The dashed lines show, from top to bottom, a FIR window size of 7, 9 and 11, respectively. The reactive delay clearly increases as the sample window increases, but there are also benefits. The last graph shows an implementation with a very stable DoM nearly perfectly on top of the dotted lines representing the "true" DoM.

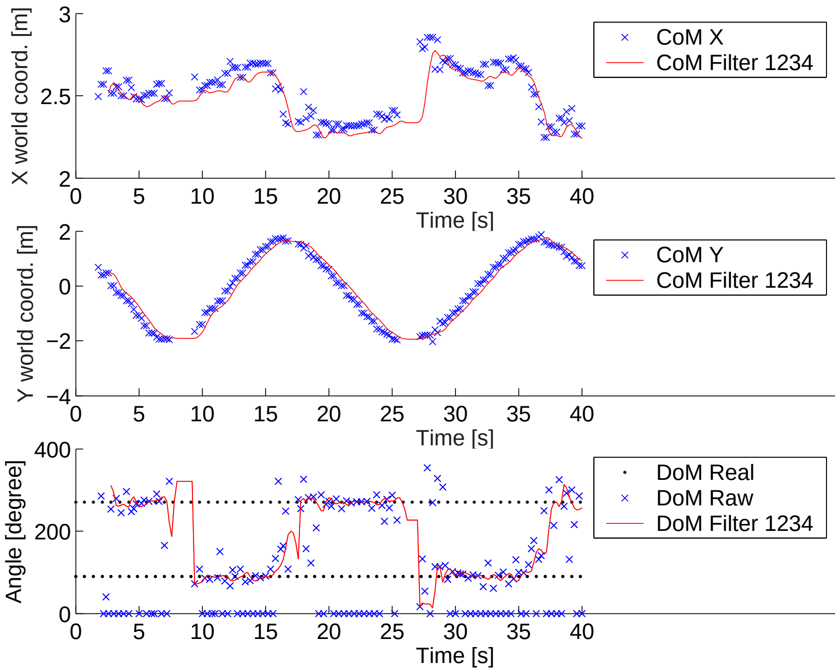


Figure 2.11: Top shows change of the x coordinate in CoM data and CoM filtered data. Middle shows change of the y coordinate in CoM data and CoM filtered data. The bottom shows the DoM from the CoM data, the DoM filter and the "real" DoM. The filter had all part of the algorithm activated, seen as "1234".

If the time scale was divided by four, the reactive delay would not have been a practical drawback compared to the beneficial stability.

2.5.2 Filter Techniques

FIR is not the only filter that can be applied to a noisy signal. Several filters exist, but there is one filter in particular that would really do the job well. This is the Kalman filter. An idea to use the Kalman filter was discussed, but as the scope of this thesis was partly an implementation and partly a user study, the Kalman idea was abandoned in preference to the FIR filter, which was easier to implement.

The Kalman filter idea involved the concept of filtering each leg position with a Kalman filter. Filtering each leg with a lower sampling frequency was assumed to be too hard—the sampling frequency had to be 20 Hz or higher. This sampling would involve multiple targets, which is much harder than tracking a single target. Sampling the legs involves two distances: the distance moved by each leg, and the distance between the legs. The former is variable in comparison with the latter which can be assumed to be constant. With a low sampling frequency, the variable distances and the static distance can become approximately equal. But with a higher sampling frequency, the variable distances are much smaller than the static distance. Loss of samples always occurs and with the higher sampling frequency the probability of mixing up the legs due to leg occlusion is minimized.

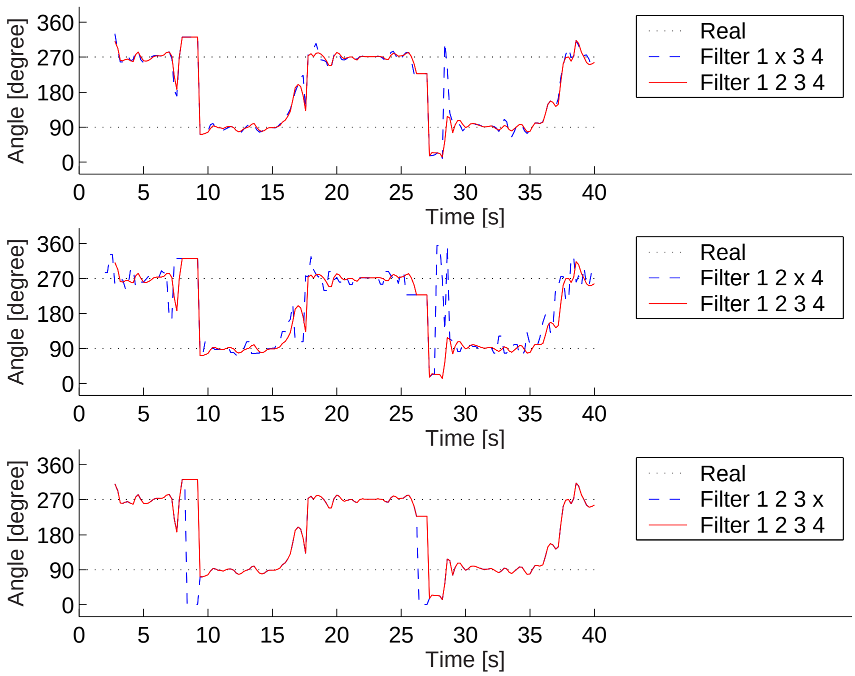


Figure 2.12: Comparison what each filter part adds to the overall result. The dashed line shows an incapacitated and a solid line shows the final implemented DoM filter.

2.5.3 Walking Style

During the analysis of position samples it was noted that a walking gait that felt "unnatural", i.e. walking at a pace that felt too fast or too slow, affected the resultant DoM to a great extent. It was speculated that unnatural walking brings uncertainty into the movement and thus creates a larger sway than is natural. The speculation led to the idea that a subject should walk as natural as possible in the user study, without constraints on the walking speed. This topic is brought up in greater detail in chapter 3.

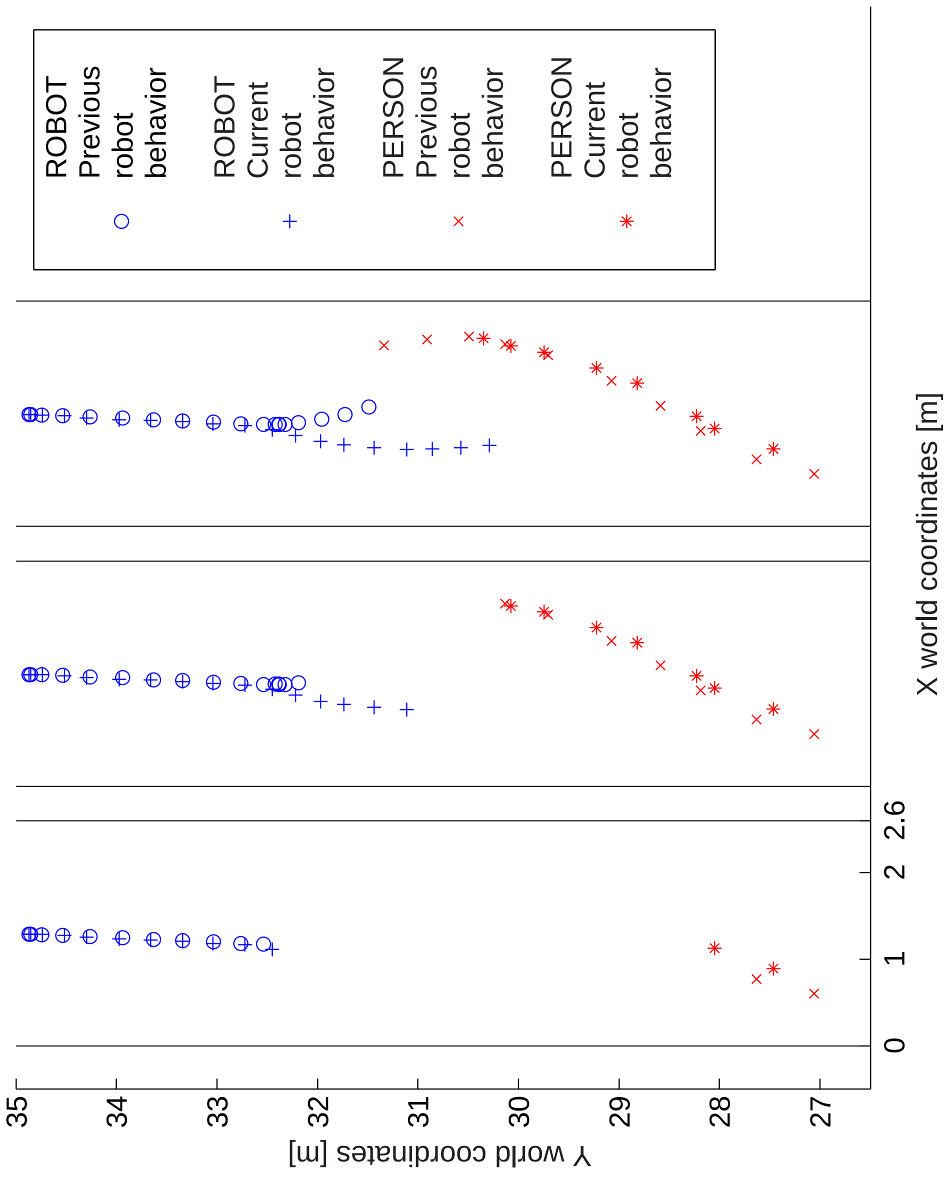


Figure 2.13: The positional data of the robot and a subject with both robot behaviors at the right to left walking path. The three scatterplots show the positions at different time steps. It is evident from the middle figure that the two robot behaviors perform the opposite movement actions.

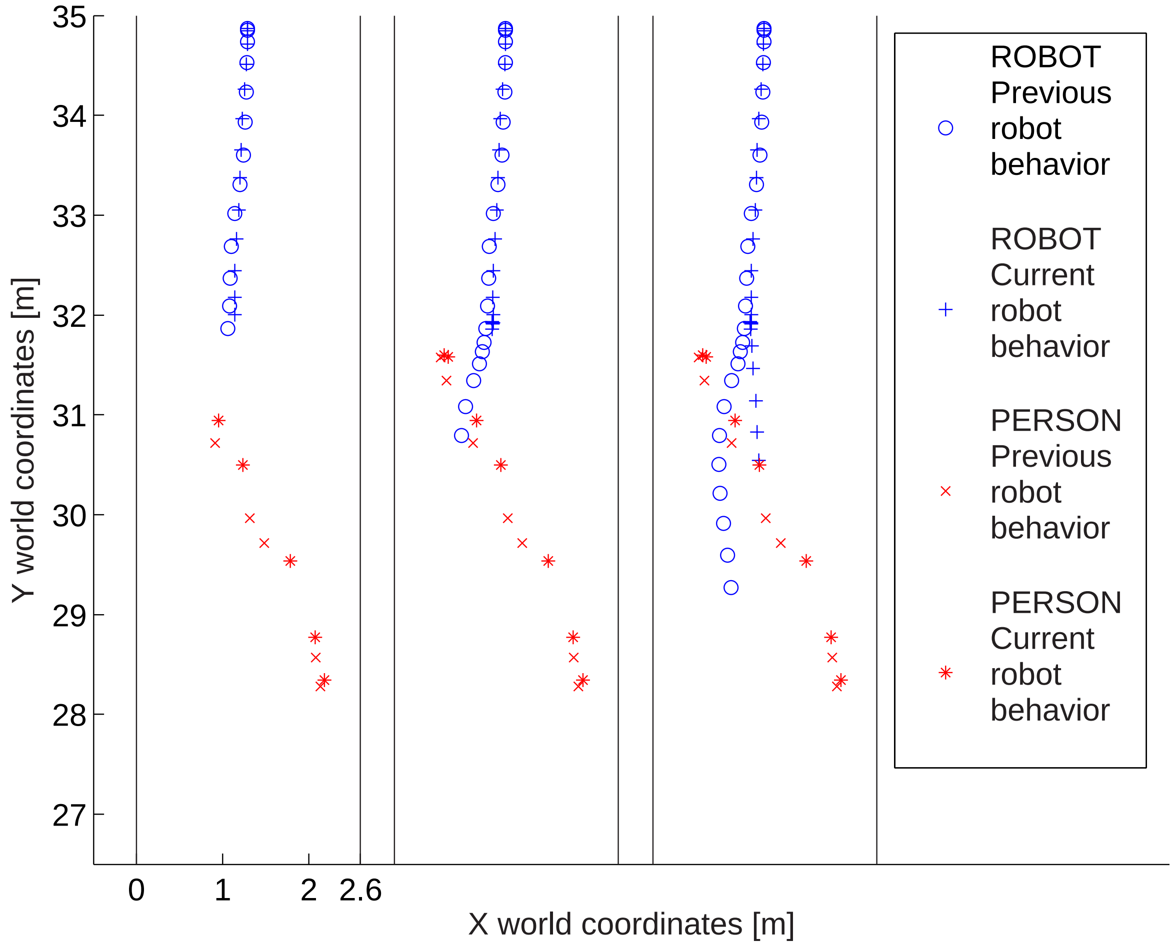


Figure 2.14: The positional data of the robot and a subject with both robot behaviors at the left to right walking path. The three scatterplots show the positions at different time steps. It is evident from the middle figure that the two robot behaviors perform the opposite movement actions.

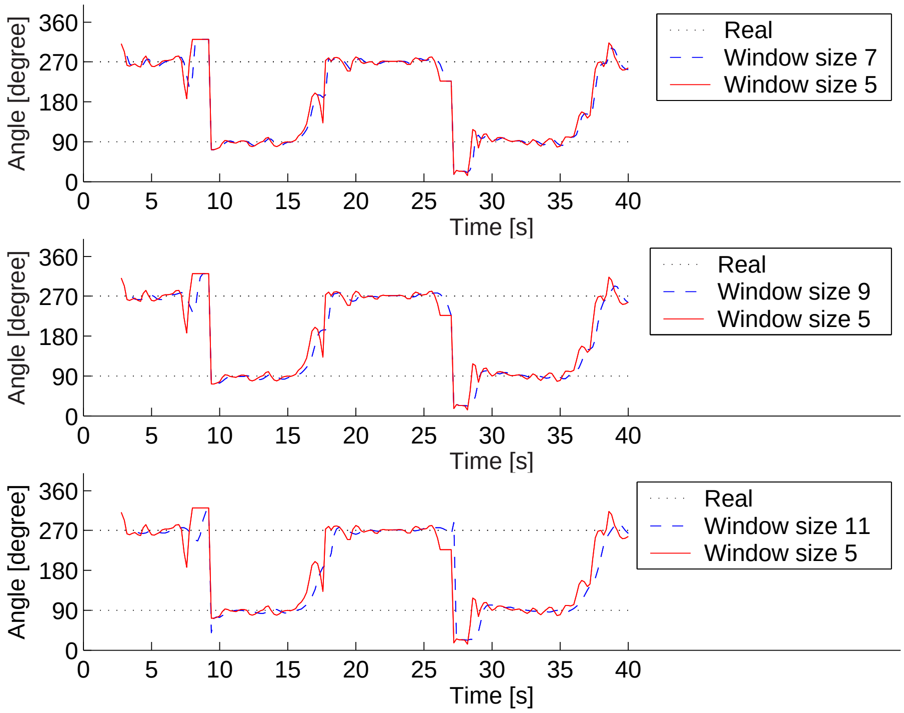


Figure 2.15: Comparison between different window sizes of the FIR filter. The dashed line represents a filter with a window size different from the finally implemented DoM filter.

Chapter 3

Comfort Attitude towards Robot Exposure

The goal of the user study is to investigate the possible human comfort difference between a robot behavior using and not using the decision module.

To fulfill the goal, certain techniques and methods were applied. A study setup was designed to create the largest perceivable difference between the two robot behaviors. This design was achieved with the use of two different walking paths, starting on one side of a corridor and ending on the other side. Subjects were directed to walk both walking paths by given oral instructions to expose the users to the difference between the previous and the current robot behavior. Both paths included a change of direction at crucial positions, and would thus favor the current robot behavior according to the main hypothesis: current robot behavior will be perceived as more comfortable than the previous robot behavior.

With the use of Likert-type scale statements, subject comfort was assessd. By means of repeated measures, inter subject comfort difference was calculated.

Combining the robot behaviors and the walking paths results in four different by-passing interactions. Additionally, each passing interaction results in one of two types of movement actions from the robot: stop or pass. These four passing interactions are labeled PI_{1-4} and are shown in table 3.1 with their respective movement action in parentheses.

This chapter describes how the study setup was taken from a theoretical level to the final setup. Then a description of the creation of the comfort assessment together with

Table 3.1: The resulting four different passing interactions PI_{1-4} from the combination of the current and the previous robot behaviors and walking paths. The side change is seen in the robot's perspective.

		Walking Path	
		Right → Left	**Left → Right**
Robot behavior	**Previous**	PI_1 (stop)	PI_2 (pass)
	Current	PI_3 (pass)	PI_4 (stop)

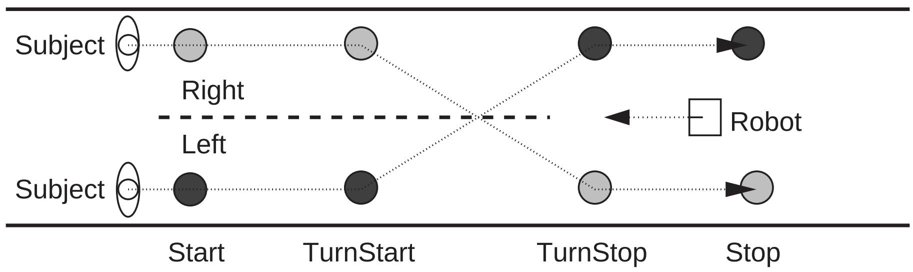


Figure 3.1: A top view of the study setup. The solid lines are the walls of the corridor and the dashed line in the middle indicates the border between the right and the left side of the corridor. The circles depicts locations where walking changes take place. The dotted lines depict the two walking paths and the movement direction of the robot.

the statistical method and hypothesis follows. Finally, the results and a discussion are presented.

3.1 Design of the Study Setup

In general, a practical study setup should, among many things, enable the collection of data in line with the goal of the study. Prior to the practical study setup, a theoretical study setup was created, and verified via simulation, providing a platform for the practical study setup. The theoretical study will be described in this section and the simulation will be described in section 3.3.

From the goal of this study to determine whether a subject comfort difference exists or not, the design of the theoretical study setup was focused on variables that were assumed to enable the greatest perceivable discrimination. As the "input" difference between the current and the previous robot behavior is in the consideration of direction, a setup that focused on direction of movement was designed.

An outline of a setup which enables a focus on the direction of movement is depicted in figure 3.1. The figure shows the starting positions for a subject, *Start* on left and right side of the corridor. The starting position for the robot is also shown in the middle of the corridor. Dotted lines are drawn from both starting positions, *Start*. These lines represent two walking paths along which the subject is to walk: One from left to right and one from right to left. The aim of the walking paths' design is an exposure—for the robot—of a subject who changes side.

Each dotted line has four filled circles representing where a change of speed or direction takes place. Additionally, each line/walking path is divided into three distinct segments. The first segment, *Start–TurnStart*, is for the subject to accelerate to normal walking speed. The second segment, *TurnStart–TurnStop*, is the part of the walking path where the robot is exposed to a subject's change of side. The last segment, *TurnStop–Stop*, is where the subject continues on the other corridor side, opposite from where the individual started.

To summarize, the walking paths will enable the robot to perceive a subject who changes direction and thus different spatial robot behaviors will be executed which the subject, in turn, is to perceive. However, other considerations than perception of spatial robot behavior differences have to be taken into account for a human being. The nature of these considerations is given in the paragraphs that follow.

3.1.1 Study Setup Considerations When Using Proxemics

How should a setup involving proxemics be designed for the most reliable and valid data collection? To begin with, Hall mentions that the personal zones and spatial behaviors are in proxemics seen as outside the realm of consciousness [16]. This has a theoretical impact in research where distances in an interaction should not be brought up to awareness.

Validity problems originating from the choice of research method have been discussed in a review of human spatial behavior [1]. Three different categories of method techniques exist in the research field of human spatial behavior: simulation method, laboratory method and interactional method.

The first method, simulation, is conducted by asking subject *how they think they would have behaved* if interaction had occurred, e.g., a subject expresses the spatial behavior of a hypothetical situation.

In laboratory methods, subjects are asked to distance themselves *as if* a real interaction was occurring, though no interaction actually takes place.

The final method, interactional, can be divided into an unobtrusive and an obtrusive interactional method. An example of the unobtrusive interactional method is unobtrusive observation of people in actual everyday interaction—subjects are unaware of the fact that they are being observed. The obtrusive interactional method is where subjects are asked to engage themselves in an interaction in a laboratory setting, such as the study described in this thesis.

It has been observed that studies performed using a simulation or a laboratory method have a low correlation with real life studies; the results may be valid per se, but lack external validity [1]. Thus, it is of great importance that a subject is unaware of the distance considerations in a study involving proxemics. Also, the same type of study should include to the largest extent "normal" interaction. This topic is also mentioned in section 3.5

3.2 Setting Theoretical Study Setup Variables

The two walking paths will expose the robot to a subject who changes direction. In a similar manner, a subject will be exposed twice to each robot behavior. However, it is of importance that the subject *perceives* a difference between the robot behaviors. But the perception of the difference between the robot behaviors does not necessarily imply a difference in comfort level.

To maximize the probability that a subject perceives a difference, the most influential variables have to be used. To find these variables, a good starting point was the diagonal segment, $TurnStart$–$TurnStop$; it is when a subject enters this segment that the current robot behavior can react differently than the previous robot behavior.

Which variables have the largest influence to the diagonal segment? Foremost it is time—a change of direction has to be detected *before* a change of side has taken place. Otherwise the current robot behavior will react similarly to the previous robot behavior. Secondly, time is causally dependent on speed and distance; slower walking speed and a longer diagonal segment length enables increased time to detect a direction change.

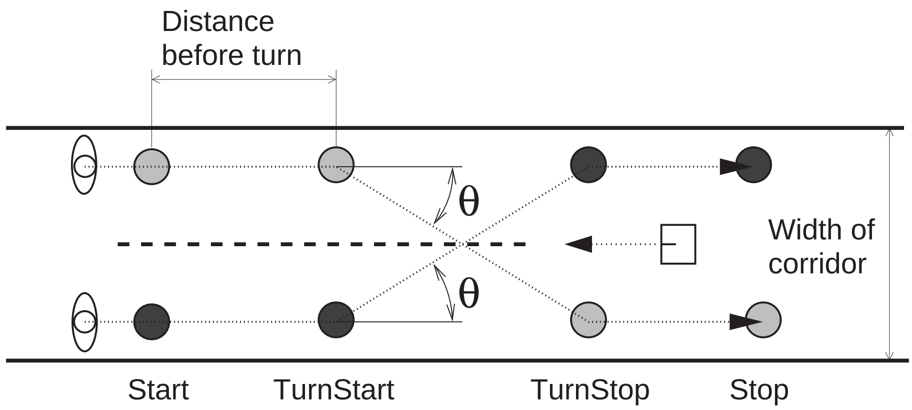


Figure 3.2: A top view of the study setup. The variables which are important in the setup are: θ, "Width" (of corridor) and "Distance before turn".

Thus, the vital task is to create a study setup which enables the current robot behavior to perceive the change of direction as quickly as possible. Also, if the subject changes side too late, the distance to the robot might be too small. This would result in a stop by the robot, due to the risk of violating the subject's personal space.

However, the time variable is limited by four constant factors. First, the motion detection module starts to detect motion at a distance of $\sim$6 m. Second, when a human being is within the motion range, there is a 2 s delay before the movement action state is defined. This delay is due to the decision module. Third, the average speed of the robot is 0.5 m/s. Fourth, the speed of the subject has to be considered.

Within this frame of limiting factors, there are three variables that have the greatest influence on the time before a change of side is detected: the direction of movement change, θ, the corridor width and the length of the first segment, *Start–StartTurn*. These three variables are shown in figure 3.2 and will be discussed and evaluated in the following paragraphs.

3.2.1 Angle of Side Change

The angle of side change, θ, is probably the most important factor in the setup. A decrease in the angle will increase the length of the second segment and thus increase the time available for the robot to define the movement action state. This can be seen in figure 3.3. Albeit, a too small θ does not work due to two factors, α and β, as described in section 2.3.1.

From previous definitions of α and β, a definition of a (DoM) margin interval was made, seen as γ in figure 3.4(a). The margin interval exists for the same reason as α: to enable a walk to be carried out *along* a corridor without causing the execution of an unwanted movement action. Thus, the same marginal interval should exist for θ to enable a corridor to be crossed without an unwanted movement action; and as β is always present, irrespective of DoM, the result can be seen in figure 3.4(b). Consequently, the DoM along the corridor and the DoM θ are equidistant from the movement action border, α. Empirical tests showed that when $\theta = 25^o$, the movement action state was stable.

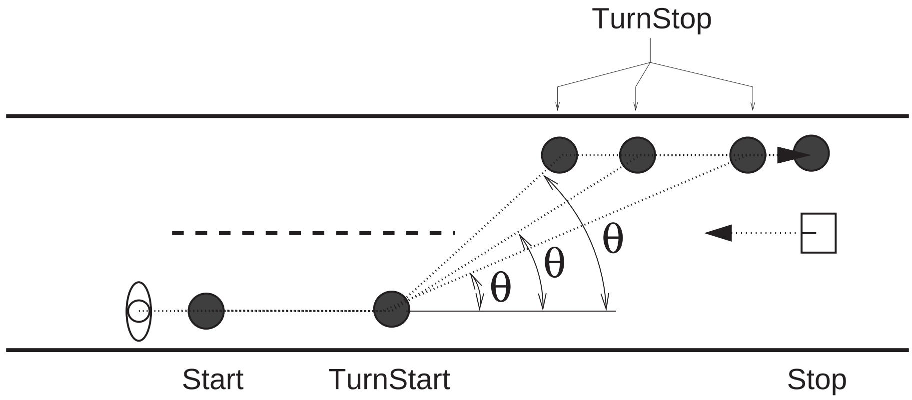
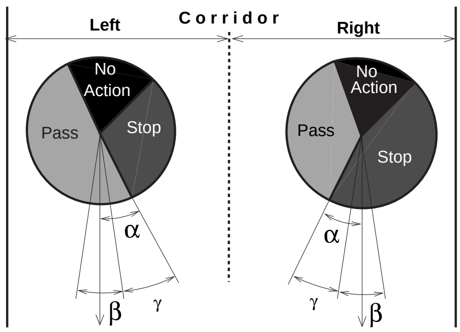
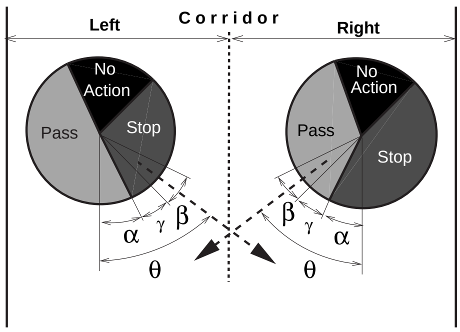


Figure 3.3: A top view of the study setup. A smaller θ will increase the length of the diagonal segment and vice versa.



(a) γ is also used on the other side of α in the creation of θ.



(b) The DADs in the decision module with the additional labeled γ as the marginal interval.

Figure 3.4: How θ was chosen from the setup of the DADs.

3.2.2 Corridor Width

It was early decided that the main study would be conducted in a corridor setting. As there were several corridors to choose from around the area of CVAP/CAS, a criterion of the corridor width had to be established.

Together with θ, the corridor width is also an important variable. It takes shorter time to change side in a narrow corridor than a wider corridor, and a longer time increases the robot's possibility to establish the movement action state prior to the change of corridor side. With the assumption that people tend to walk ~ 0.5 m from the wall, in a 2 m wide corridor, with $\theta = 25^o$, the length of the second segment, *TurnStart–TurnStop*, is $\frac{2-0.5\times2}{sin(25)} \approx 2$ m; a 3 m wide corridor has a second segment length of $\frac{3-0.5\times2}{sin(25)} \approx 5$ m.

However, an excessively wide corridor can be seen more as a room, not complying with the initial setting of a corridor environment. It was thus decided to only consider corridors with a maximum width of 3 m.

3.2.3 Length of First Segment

Before the length of the first segment is calculated, an assumed human walking speed has to be decided. Studies from pedestrians indicate a walking speed of 1.2 m/s, but also that "walking speeds are influenced by environmental [...] characteristics" [21]. Given the assumption that a person would not walk as fast in an unknown environment, an approximated normal walking speed of 1m/s was assumed. Together, the first mentioned walking speed can be seen as a worst case speed, whereas the latter speed can be seen as a more normal speed. However, the normal walking speed will be used in calculations as this will be the most common case.

As is mentioned above, the aim for the first segment is to let the subject accelerate to a normal walking speed. Additionally, it is of importance that a *defined* state from the decision module is available *before* the subject enters the second segment and changes direction. From the assumed walking speed of 1.0m/s and the delay of 2 s of the decision module, the distance walked is $1.0 \times 2 \approx 2$ m. Furthermore, a shorter distance of 1 m is needed to accelerate to normal walking speed, resulting in $2 + 1 \approx 3$ m as the length of the first segment.

3.2.4 Resulting Theoretical Study Setup

When the tree variables, length of first segment, corridor width, and θ, were set, four distances could be calculated and therewith finish the theoretical study setup. These four distances are labeled $D1$–$D4$ and are depicted in figure 3.5 together with the three variables.

Before using the theoretical study setup in the pilot study, the setup was tested utilizing the simulation software Stage.

3.3 Simulation of Study Setup

A simulation was performed using the Stage software. A screen shot of a simulation is shown in figure 3.6. The figure shows a top view of a map representing the corridor where the study took place. This virtual corridor contains two blobs, depicted in the figure. The blob furthest to the left is a simulation of the PeopleBot robot used in the

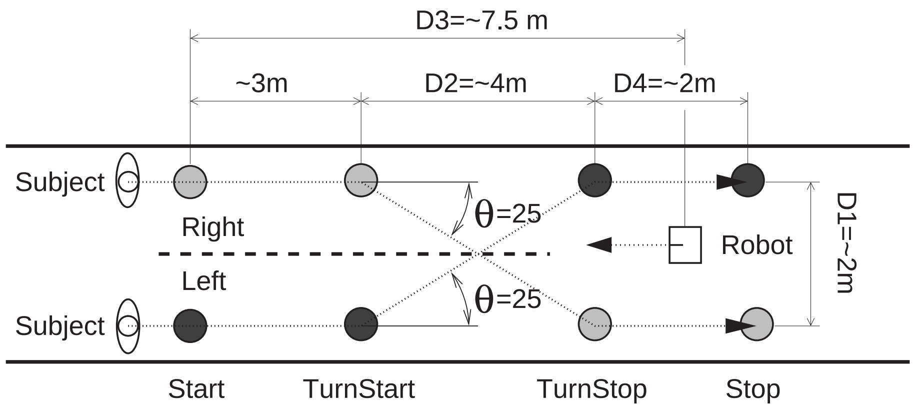


Figure 3.5: The theoretical study setup.

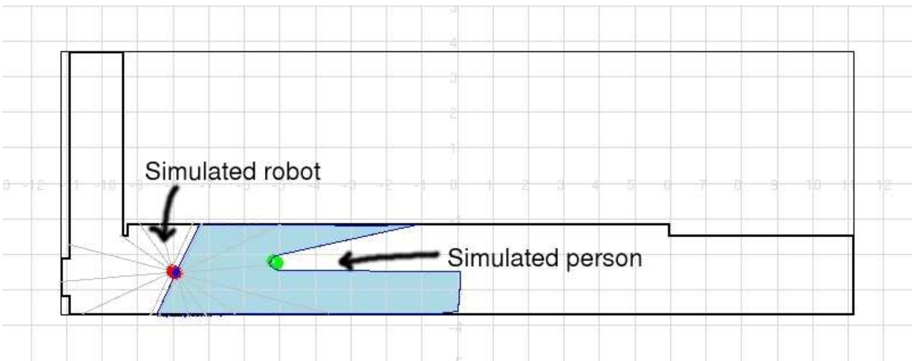


Figure 3.6: A screen shot from the simulation program Stage. The blob on the left is a simulated robot, whereas the blob on the right is a simulated robot used to simulate a person, programmed to move as a subject would move in the user study

user study, and the blob furthest to the right is a simulated robot used to simulate a moving person; no simulation of human walking was thus made.

The simulations were conducted in order to verify the theoretical study setup. Several parameters were altered and tested: different walking speeds of the human, different distances before entering the second segment, different θ et cetera. The results from the simulation showed that the theoretical setup could be used with little alterations, such as the distance between the starting positions of a subject and the robot. This subject is more thoroughly considered in section 3.4.

3.4 Creation of Practical Study

The creation of the practical study began with testing the setup. The setup test was conducted with a pilot study. Then, the statement evaluation study followed, with the aim of refining the comfort assessment questionnaire. Finally, the main study was conducted using the results from the evaluation study.

3.4.1 Conducted Studies

The first study to be conducted was the pilot study. The aim of the pilot study was to discover practical design flaws in the theoretical study setup, both regarding the setup and also in the design of the comfort assessments. The pilot study revealed many flaws, which the theoretical study setup did not cover. The participants of the study consisted of researchers at CVAP/CAS and people randomly asked if they could participate.

The second study was the statement evaluation study, which consisted of the evaluation of ten statements using a questionnaire, seen in appendix A.3, in order to get a validity measurement. The result of the validity measurement was then used to extract the six most valid statements, resulting in a questionnaire seen in appendix A.2. A description of statements and how they were chosen are mentioned in greater detail in section **??**.

The statement evaluation study was conducted using five robotic researchers from CVAP/CAS. They all had significant prior experience with robots and were defined as being "experts" in this field and thus good at evaluating robotic research. The setup was the same as the practical study setup, with the aim of exposing each subject to all conditions. After the study, the participants were given the questionnaire seen in appendix A.3.

The main study is described in great detail in section 3.5.

3.5 Practical Study Setup

The main study will first be generally described and thereafter parts of the main study will be described in greater detail.

The study began by giving a subject a questionnaire, seen in appendix A.1. The questionnaire asked the participant to divulge gender, age, highest eduction, current occupation and height. Upon completion of the questionnaire, five different robot exposures was experienced by the subject. Each robot exposure was followed by the comfort questionnaire, seen in appendix A.1.

The aim of the first robot exposure was to get the subject familiarized with the robot. The remaining four exposures consisted of four different combinations of the passing interactions, PI_{1-4} (defined in table 3.1).

Only oral instructions were given to subjects, with the exception of the introduction written in the first questionnaire. The subjects were told to start walking when the robot started to move. They were also told how to walk and which objects to take along each walking path. The walking path from right to left, in the robot's perspective, was to be performed by first walking towards a glove attached on the right wall; the glove was to be taken as the subject walked along, whereafter the subject turned directly and walked towards the blue telephone on the other side of the corridor. This object was also to be picked up and finally, the subject would continue to walk on the left side of the corridor and place both objects in front of a battery charger. The walking path from left to right consisted of a similar setup; take a glove attached to left wall while walking and directly turn and walk towards a folder on the other side of the corridor; finally, the subject continued on the right side of the corridor and placed both objects on a chair.

The following paragraphs first describe how the points of change were defined, objects taken, and how they and other aspects altered the form and size measurements of the walking paths. Then a description how habituation effects changed the number of robot exposures for a subject. Last, the oral instructions are described.

3.5.1 Participants

Data collected initially from 29 participants were sampled. However, no data from six of these participants was used. In total, four participants passed *behind* the robot instead of in front of it during one passing interaction. Also, the robot was once initially misplaced, resulting in a faulty passing interaction. Finally, one participant did not state Swedish as the mother tongue, which was a prerequisite to be included in the result.

An initial idea of gathering an heterogeneous group of participants for the study was soon replaced by the idea of an homogeneous sample. The reason for this was a lack of subjects to test an heterogeneous sample compared to the availability of subjects from an homogeneous population. Thus, another four participants were removed from the remaining 23, resulting in 19 participants consisting of 9 females and 10 males with an average age of 24.0 ± 3.14 years.

The final sample consisted only of participants that were university students. The amount of participants that had prior experience with a robot similiar to Sony's AIBO dog was 15.8% (3/19), but none of them had experienced an autonomous robot like the one used in the study. Thus, none of the participants had had any prior robot interaction experience which could have biased the results.

3.5.2 Points of Change

The points of change in the theoretical study setup, $Start$, $TurnStart$, $TurnStart$, and $Stop$ on both sides, were represented by five different objects, which are described below.

$Start$, consisted of red and green tape stripes on the floor on the robot's right and left side, respectively.

To motivate a subject's $TurnStart$ point to be closer to the wall, and thus increasing the time before change of side, $TurnStart$ consisted of two black gloves attached to both walls on each corridor side. The tip of the gloves were placed on a vertical height dependent on a subject's height: 200 cm, 215 cm and 230 cm above the floor when a subject had the height $< 175\,\text{cm}$, $175\,\text{cm}$–$185\,\text{cm}$, and $> 185\,\text{cm}$, respectively. These three height placements were needed to make the distance relative to the wall at $StartTurn$ as consistent as possible; with the first, non-height dependent, placement of the glove, tall people reached the glove with the arm in a horizontal alignment; short people had to stretch one of their arms up wards, thus resulting in a vertical alignment of the arm. Consequently, tall people were closer to the middle of the corridor when entering the second segment of a walking path, compared to short people. However, the three height placements solved the problem.

$TurnStop$ consisted of a blue telephone and a black green folder on the robot's right and left side, respectively.

$Stop$ consisted of a wooden chair and the place in front of the black battery charger for the robot on the robot's right and left side, respectively. The objects placed in corridor used in the study, can be seen in figure 3.7

The objects were chosen for a contemplated practical and theoretical reason. The practical reason consists of an easiness in discrimination, both visually and a semantically. Visual discrimination existed as the objects composed a large contrast towards the pale color of the floor and the walls. The semantic discrimination came from the fact that the objects had very little in common, except the gloves; thus, the probability of mixing up the objects was small, except for the gloves. Taken together, the large discriminability also made the oral instructions to be more easily understood.

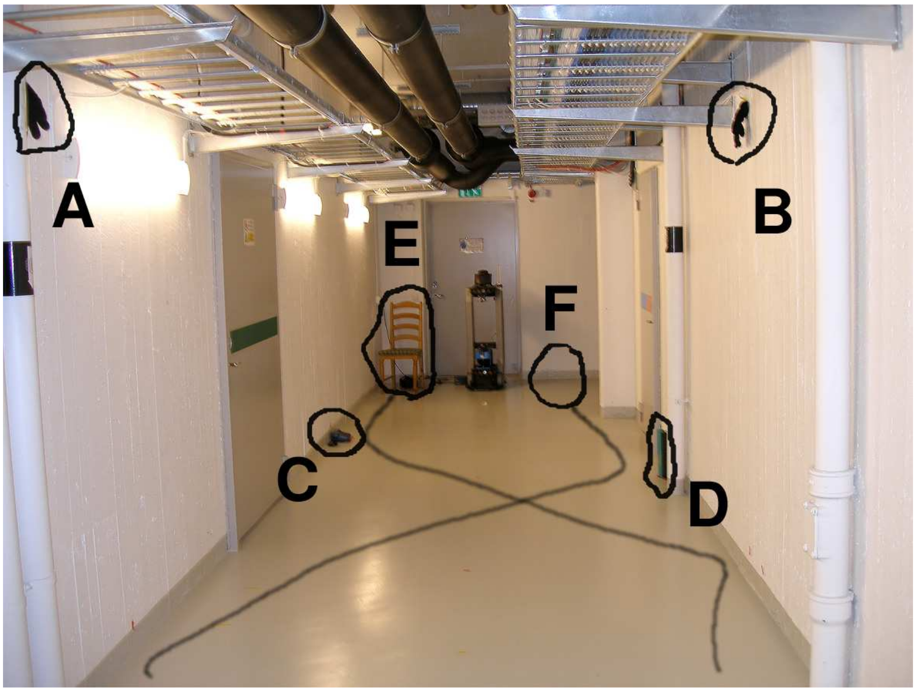


Figure 3.7: The view of the corridor seen by a subject at the first exposure. The places of change are encircled and marked with letters: A, left glove; B, right glove; C, telephone; D, folder; E, chair; F, place in front of black box (not seen in the picture). Also, the walking paths are outlined as airbrushed markings on the floor.

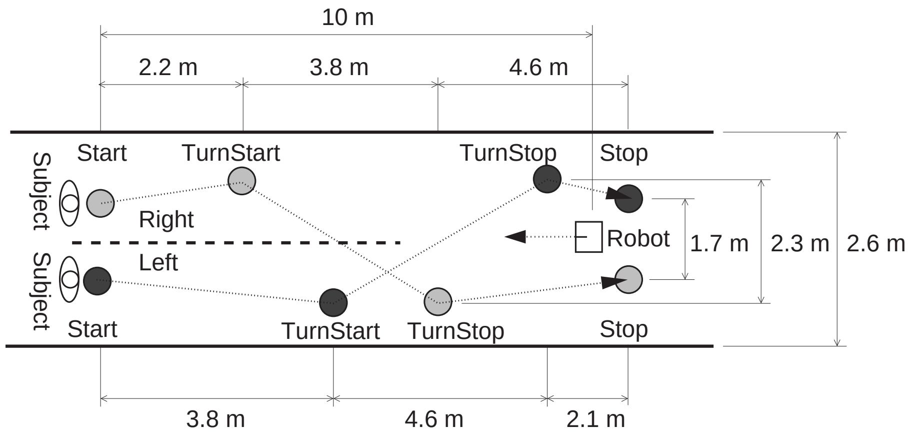


Figure 3.8: A top view of the final study setup.

The theoretical reasons for the use of common objects, were to achieve unawareness of the laboratory setting and to increase the naturalness of the interaction. These being prerequisites in obtaining a valid study. Subjects were told to take the objects along each walking path, while walking, and finally place them on and in front another object. It is assumed that a taking of objects imitates the "natural behavior" of people in an everyday situation; we seldom wander around during several seconds and wonder where the next foot is to be placed. Instead, it is assumed that people constantly have a goal, like "I want to walk to the table" or "I want to place the apple on the table".

3.5.3 Walking Paths

The symmetrical walking paths of the final setup was changed in three ways, resulting in asymmetrical walking paths depicted in figure 3.8.

The first change applied to both walking paths as the gloves at *StartTurn* were attached to the wall at a height that forced a subject to stretch an arm up wards in order to take a glove. This resulted in a shorter distance from the wall for *StartTurn* than *Start*, as evident from figure 3.8.

The second change applied to the walking path from left to right. Subjects in the pilot study did not experience or perceive any difference, which was the aim, when the left to right walking path in figure 3.5 was used. Therefore, *StartTurn* was moved in order to get a subject closer to the robot while walking the second segment; this was shown to increase the experienced difference between the robot behaviors. However, *TurnStart* of the right to left walking path could not be moved closer; as the distance between an approaching person and the robot would thus be too small to execute a pass for the current robot behavior.

The third and last change consisted of all the set distances in figure 3.5. They were somewhat accurate for the right walking path, seen in the figure, but was changed in the left walking path as previously explained. Additionally, two major length changes for both paths caused resizing. The first was the distance between the robot's starting position and the subject's starting position. This had to be increased due to the non-violation distance the robot needs in order to execute the movement action pass. The corridor width also changed as the corridor chosen to use for the studies had a width of

2.6 m (compared to 3 m). The choice was made as the corridor lay in the same building as the CVAP/CAS lab and was rarely used by any people. Due to these changes, all other distances had to be changed.

3.5.4 Habituation

In the beginning of the pilot study, there were four robot passing interactions, resulting in four robot exposures. From the comfort scores and discussions with subjects, it was revealed that the first robot exposure was perceived as much different, compared to the subsequent robot exposures. Subjects said that they were "on their guard the first time", as they "did not know what to expect." But, after the first exposure, subjects "felt more comfortable as they knew what to expect." Thus, a familiarization interaction was added, with the aim of habituating subjects. Consequently, the familiarization interaction increased the number of robot exposures to five. This is mentioned in greater detail in section 3.7.1. Also, with the familiarization interaction, the procedures of passing interaction and answering the questionnaire were practiced by the subject to some extent.

3.5.5 Oral Instructions

Oral instructions were used as the author gained by experience that a usage of *only* written instruction increases the probability of trial failure. With oral instruction, the experimenter can "see" whether a subject has understood the instructions or not to a greater extent than with the usage of written instructions. Yet, oral instructions does inevitably lead to trial inconsistency. To minimize the downside of oral instructions, they were formalized in order to enhance consistency. With this study, formalization was acheived by virtue of writing down the instruction and then memorizing them.

The experimenter used five sets of instructions: two instructions how to take the objects along each walking path; two instructions to urge a subject to practice each walking path prior to a robot exposure; and one instruction how to walk in the first trial where a subject got familiarized with the robot.

3.6 Comfort Assessment

To assess a subject's comfort attitude in a passing interaction, a comfort questionnaire was filled out after each passing interaction exposure. The questionnaire was written in Swedish and only subjects having Swedish as their mother tongue were used as participants. People who do not have Swedish as their mother tongue are believed to confound a test result too much from their interpretation of Swedish. Such a confounding factor is impossible to avoid, but is minimized with this type of language criterion

Several types of scaling techniques can be used for a questionnaire where comfort attitude is to be assessed. It has been shown that Likert-type attitude scales are good in measuring attitude and are also one of the most popular attitude assessment methods [3]. Others also report Likert-type scales to be a common scale [11, 25] and it has been used previously to assess the level of comfort for a subject interacting with a robot [5]. Therefore, Likert-type scaling was chosen as the comfort assessment scale to use.

3.6.1 Likert-type Scale

A Likert-type setup contains several statements, *items*, which are clearly positive or negative, such as "I think that robots are something good" or "Robots should be banned". The answer to each item is given by choosing a degree of agreement on the *response scale* which contains a number of *response categories*, like "Disagree", "Disagree somewhat", "Neutral", "Agree somewhat", and "Agree". Each agreement is associated with an *item score* and all item scores are summated to obtain a subject's *test score* [3].

There is given below an example of a Likert-type setup with two items and five response categories where a high score indicates a subject who is positive towards robots and a low score indicates a subject who is negative towards robots:

Statement 1	"I think that robots are something good"				
Score 1	1	2	3	4	5
Answer 1	Disagree	Disagree somewhat	Neither disagree nor agree	Agree somewhat	Agree

Statement 2	"Robots should be banned"				
Score 2	5	4	3	2	1
Answer 2	Disagree	Disagree somewhat	Neither disagree nor agree	Agree somewhat	Agree

As can be seen in the example, "Score 2" and "Score 1" have the opposite grading as their statements are positive and negative, respectively. This means that the summated test score will indicate how negative or positive a subject is towards the statement's topic. If "Score 1" and "Score 2" switched places, a high test score would indicate that a subject had a negative attitude towards robots.

When creating a questionnaire using Likert-type scaling, the number of scale items and response categories have to be set, and be in accordance with two factors. The first is the repeated measures concept and the second is the subject's fatigue. An example of the latter is when the subject has to answer too many items. The implication of these two factors on Likert scaling are discussed below.

Generally, the data in statistical calculations have to come from an interval scale, or better, a ratio scale. Albeit, a Likert-type scale item answer belongs to an ordinal scale. However, test scores, the summation of the item scores, from Likert-type scales have been found by Likert and others to conform to normal distributed data [3]. Also, the use of Likert-type scale data as interval data is commonplace for practical reasons, even though it has been criticized [10]. But there are some constraints which has to be satisfied before the Likert-type setup can be treated as interval. The constraints involve both the number of response categories and the number of items.

The resulting Likert scale statements can be seen both in appendix A.3 as the original statements and in appendix A.2 as the final statements.

Number of Response Categories

The literature reveals the usage of several different numbers of response categories; the most common number of categories is five [11, 32], but if Likert scales are to be treated as interval level data, seven categories are preferable [4, 10, 11], or compared to ten categires, are the second best alternetive [6]. Also, in a paper by Pemberton (1933) cited in [23], seven categories got the maximum reliability.

Thus, the scientific research points to the use of seven response categories or more. A larger number of response categories could increase the reliability in theory, but "On the other hand, if the scale has too many choices, then the subject will be indifferent between two or more consecutive choices that represent a single internal level. Being forced to select only a single choice rather than a range, the subject randomly assigns the response to one of the choices within this range" [23].

To summarize, a large number of response categories would increase the reliability and the conformation towards normal distributed test score data, but *too* many categories would cause subject fatigue. Therefore seven response categories are chosen—the minimum required for making use of the data as normal distributed data.

Response categories in Swedish, in [27]:

– (5) Stämmer absolut (Definitely correct)

– (4) Stämmer (Correct)

– (3) Vet ej (Don't know)

– (2) Stämmer inte (Incorrect)

– (1) Stämmer absolut inte (Definitely incorrect)

were used to create the Swedish response categories in this study:

– (7) Stämmer absolut (Definitely correct)

– (6)

– (5)

– (4) Neutral (Neutral)

– (3)

– (2)

– (1) Stämmer absolut inte (Definitely incorrect)

The response categories 2, 3, 5 and 6 are not explicitly associated with an agreement and were left out deliberately to make the scale construction. The middle alternative was changed, as the author found "Neutral" better than "Don't know".

It should be pointed out that the number in the parentheses are part of both response scales and should not be confused with item scores; incorporating the item score in a mixture of positive and negative statements could bias the subject.

Number of Items

As with the number of response categories, the number of scale items is also dependent on whether the test score can be used in statistical calculation assuming normal distributed data. A span of 4 to 40 items can be used [11] but others [3] find 6 to 30 items better. The same argument is used here as in the above paragraph in the decision of the number of scale items; the minimum of six items were chosen as it was believed that subject fatigue probability would be a confounding factor above six items.

The statements was made with the intention of measuring the level of comfort that was perceived during the interaction with the robot. Also, the statements implicitly assume that the robot behaves and is seen as a human being.

The comfort questionnaire with the comfort statements is available in appendix A.2.

3.7 Statistical Method

An early decision in the design of a user study resides in the method to be used. Several factors effect the choice of method. In this particular study, one of the strongest factors was the limitation of the amount of data that could be collected. This is due to the fact that the robot is battery powered and used by several researchers at CVAP/CAS; a study session could not take the whole day.

Given these limits, a suitable statistical method had to be chosen. To begin with, factorial design is in general better than a one-way design: Greater generalizability of the results, interaction effects can be calculated, and, above all, it requires fewer participants for the same degree of statistical power [18]. However, an even better design is the concept of repeated measures. Besides having all three mentioned benefits of factorial design compared to a one-way design, it requires fewer participants. With the current setup of table 3.1, four groups are needed for a factorial design, whereas a repeated-measure design only needs one group.

This maximized use of participants was the main argument in using repeated-measure design compared to factorial design. Additionally, there are other considerations in comparing factorial design with repeated-measure design, and these will be discussed below.

3.7.1 Repeated Measures

In general, one of the main advantages of the repeated measures concept is that it enables the reduction of the overall variability by using a common subject pool for all treatments. In this case, a treatment corresponds to a passing interaction. Additionally, the subject difference can be removed from the error term, making the error part independent from treatment to treatment. This independence is sometimes expressed as "*partialing out* effects that cause the dependency." To put it in another way, subjects will respond differently to different treatments creating a variance in the score of each treatment. This variability, and thus error term, can however be divided into an inter subject variation and an intra subject variation. With this division of the variability, the unwanted inter-subject variation can be partialed out from the total variation, leaving only the variation within the subjects as an error term. Consequently, the error term is decreased and therefore the relative power is increased [18].

In the current study with repeated measures, each subject will experience all four passing interactions; thus, the test scores can be compared between subjects. This type of comparison is harder in a factorial design for two reasons.

Firstly, four passing interactions would result in four passing interaction groups; and each group has to be statistically equal to the other groups in order to be compared. Secondly, a subject who is biased with a negative attitude towards robot would grade all robot behaviors with lower scores than a unbiased subject.

These two problems do not apply to repeated measures for two reasons. Firstly, inter subject comparison removes the need for group equality—the subject is the group. Instead of creating treatment groups, repeated measures makes each subject a group. Secondly, the *difference within* a subject's scores, compared to the *difference between* four groups, tell the relative difference for each subject independently whether there is a positive or negative bias towards robots.

To summarize the benefits of the repeated measures concept in this user study: Only one group is used as it is compared internally, and relative scoring maximizes the result of a difference between passage interaction exposures.

Until now, only the benefits of repeated measures have been mentioned. As with all things, there are also disadvantages with repeated measures. One problem is that a subject has to be exposed to all passing interactions—it takes approximately four times longer per subject compared to a factorial design. A consequence of this is the increased probability of subject fatigue. This is why the number of passage interactions exposed to a subject is minimized to four. If a subject tires, the risk of withdrawal, unnatural performance, et cetera, increases. This would induce a withdrawal bias [31] and other biases, thus jeopardizing the validity of the study.

Other problems are the sequence effects and practice effects [18,31]. The first effect is due to the fact that each passing interaction will influence the perception of subsequent interactions. A passing interaction could put a subject in a negative mode whereas another subject could react positively. The succeeding interaction's evaluation could thus be biased by the sequence effect.

The practice effect consists of learning in a trial which affects the outcome of subsequent trials. This is not applicable with the current setup as it explicitly does not involve learning. However, an effect in close resemblance to the practice effect is habituation—response decrease with repeated presentations of a stimulus [13]—which holds for this kind of study. Habituation in the context of this study, was mentioned in greater detail in section 3.5.4.

From the pilot study it was shown that a subject was often "nervous" at the first passing interaction compared to subsequent interactions. Thus, the habituation was very quick but a distinct difference in score, independent of the passing interaction, was always present between the first two exposures.

Problems remedies

Luckily, there are remedies which reduce the sequence effect and the effect of habituation. Regarding the sequence effect one solution is counterbalancing, i.e. balancing the order of treatments among subjects [31]. The best counterbalancing is to use all possible combinations of passing interactions. With four different interactions, the resulting number of combinations are $4! = 24$. Thus, at least 24 subjects would be needed to counterbalance the sequence effect. As this kind of combinatorial setup is somewhat impractical, a trade off between counterbalancing and sequence effect has to be made; a practical method often applied when the number of treatments is greater than three, is Latin square [31]. In table 3.2 the Latin square counterbalancing is shown, requiring only four combinations, which simplifies the setup. Because of its practical benefit, Latin square was used to decrease the sequence effect in the main study.

The habituation was dealt with by adding an familiarization interaction, with the aim of making a subject familiar with the robot and thus remove the initial tension seen in the pilot study. This interaction was conducted before the other passing interactions.

Table 3.2: The exposure order using Latin square.

	Order of robot exposures			
First group of subjects	PI_1	PI_2	PI_3	PI_4
Second group of subjects	PI_2	PI_3	PI_4	PI_1
Third group of subjects	PI_3	PI_4	PI_1	PI_2
Fourth group of subjects	PI_4	PI_1	PI_2	PI_3

Table 3.3: The exposure order using Latin square together with a first familiarization interaction (FI).

	Order of robot exposures				
First group of subjects	FI	PI_1	PI_2	PI_3	PI_4
Second group of subjects	FI	PI_2	PI_3	PI_4	PI_1
Third group of subjects	FI	PI_3	PI_4	PI_1	PI_2
Fourth group of subjects	FI	PI_4	PI_1	PI_2	PI_3

Together with Latin square, the resultant order of the interaction exposures are shown in table 3.3.

3.8 Hypotheses

It is hypothesized that subjects will find the current robot behavior more comfortable as it will conform to a human proxemic behavior to a greater extent than the previous robot behavior. Thus, a main effect is expected to exist of the robot behavior. This is the main hypothesis. Moreover, a subject will have a shorter distance in relation to the robot when walking from left to right rather than vice versa. Thus, the left to right walking path should be experienced as less comfortable than the right to left walking path. Also, an interaction effect is expected with a simple effect on both walking paths.

3.9 Results

Throughout the results, the term "test score" will refer to the summated comfort-assessment score. The scale used to assess the comfort attitude is reversed. This means that a low score indicates that a subject found a passing interaction comfortable whereas a high score indicates that a subject found a passing interaction uncomfortable. Moreover, the interval of the comfort score is from 6 to 24. All calculations were performed using Matlab and its statistical toolbox.

The results consist of two parts. A first part describes the calculations of the prerequisites needed, verifying that the data is "tenable" for statistical analysis. A second part consists of data analysis to ascertain whether or not the hypotheses were met.

Table 3.4: Cronbach's α for each passing interaction's comfort evaluation and the mean value of Cronbach's α

	PI_1	PI_2	PI_3	PI_4	**mean(PI_{1-4})**
Chronbach's α	0.8678	0.7784	0.8711	0.8386	0.8390

3.9.1 Prerequisites

When utilizing Likert scales, a test should be performed, such as Cronbach's α, to ascertain the scale's inter-correlation [11]. Cronbach's α is an internal consistency reliability estimation, which can be interpreted in many ways. One interpretation, given the context of the user study, would be the amount of the variance, in percent, of the Likert scale statements that would explain a hypothetical true scale composed of all possible statements [12].

Four passing interactions evaluated with the same questionnaire result in four independent test score lists for each passing interaction. Thus, four individual Cronbach's α values and the mean of all four values were calculated, shown in table 3.4. From [14], citing George and Mallery (2003), the following rules of thumb are used in the interpretation of the results: "≥ 0.9—Excellent, ≥ 0.8—Good ≥ 0.7—Acceptable ≥ 0.6—Questionable ≥ 0.5—Poor ≤ 0.5—Unacceptable".

The results are overall "good", except the second passing interaction which is "acceptable", although the mean Cronbach coefficient is "good".

The repeated measures method is based upon assumptions in order for the F ratios to actually follow the F distribution. A within-subject repeated measures design has three assumptions that have to be met: normality of data, homogeneity and sphericity of the $\hat{\sum}$ (covariance matrix) [18].

The normality of data is an assumption which means that a set of data values conforms to the normal distribution. In this user study, the data consist of the test score lists for each passing interaction. Testing if the test scores conformed to normality was done using the Lilliefors Goodness-of-fit test, with the null hypothesis that all four sets of data have normal distributions. The Lilliefors-statistics with $\alpha = 0.05$ can be seen in table 3.5. Interpretation of the table is done by looking at the first row, which should contain the p-values. But these values are unavailable when the Lilliefors-statistics (lstat) is outside the Lilliefors table, and replaced by NaN (Not a number). The reason for the unavailable p-values is that the Lilliefors calculations in matlab are based on a table of pre-calculated values. However, to find out if the null hypothesis is rejected or not, the final two rows can be used: if the lstat is below the critical value (CV), the null hypothesis is not rejected. Thus, the last two rows clearly show that all four test score lists have a normal distribution.

To test homogeneity of variance, Levene's test for equality of variance was used. In context of the user study, it tests whether the test scores from each passing interaction varies to a degree equally with the walking path and the robot behavior. If the test score would vary more with one independent variable than with another, the variance would have been heterogeneous. The result of Levene's test with $\alpha = 0.05$ was that the assumption of homogeneity of variance was met with $p = 0.8467$ ($F = 0.2702, df1 = 3, df2 = 72$).

The last assumption to fulfill was statistical sphericity, which means that the diagonal elements of the covariance matrix (shown in table 3.6) are comparatively equal and

Table 3.5: Lilliefors-statistics testing if the test scores from each passing interaction conform to the normal distribution. If the lstat is less than the CV, the null hypothesis of conforming to a normal distribution is not rejected. The statistics clearly show that the scores from each passing interaction has a normal distribution.

	PI_1	PI_2	PI_3	PI_4
P-value	NaN	NaN	NaN	NaN
lstat	0.1614*	0.1751*	0.1430*	0.1210*
CV	0.1950	0.1950	0.1950	0.1950

$^*p < 0.05$

Table 3.6: The covariance matrix of the passing interactions PI_{1-4}

$\hat{\sum} =$

	PI_1	PI_2	PI_3	PI_4
PI_1	38.2749	27.2105	14.7749	13.5468
PI_2	27.2105	30.3977	11.3216	9.0760
PI_3	14.7749	11.3216	45.2749	30.7690
PI_4	13.5468	9.0760	30.7690	35.695

the off-diagonal elements are also comparatively equal [18]. With a visual inspection of the covariance matrix, the assumption of sphericity does not seem to be met as the off-diagonal elements vary between 9 and 30. A statistical test, however, is more reliable than a visual inspection. Bartlett's test for sphericity was calculated using the same data as Levene's test with $\alpha = 0.05$. The result was in accordance with the visual inspection—the assumption of sphericity was not tenable. (Approximation to $\chi^2 = 32.6452, df = 6, p = 0.0001$.) Thus, not all of the three assumptions were met.

However, adjustment techniques can be applied when this happens, as statistical tests regarding the assumption of sphericity often rejects the null hypothesis (that the data have sphericity). "The Greenhouse and Geisser and Huynh and Feldt adjustments to the degrees of freedom appear to do an adequate job of correcting [overcoming] for problems with the sphericity when testing for overall main effects or interactions." [18]. This adjustment consists of changing the degrees of freedom in the F-statistics used in repeated measures. A correction is performed by multiplying each degree of freedom with ϵ, thus $(number\ of\ walking\ paths - 1) \times \epsilon$, $(number\ of\ robot\ behaviors - 1) \times \epsilon$. However, the ϵ must first be estimated with $\hat{\epsilon}$ and $\tilde{\epsilon}$. From calculations, both $\hat{\epsilon} = 0.6012$ and $\tilde{\epsilon} = 0.6641$ were in the neighborhood of 0.75, suggesting the use of $\tilde{\epsilon}$ as an estimate of ϵ. The unadjusted degrees of freedom, $(1,\ 1)$, are thus replaced with $(0.6641,\ 0.6641)$.

3.9.2 Analysis

Having adjusted the degrees of freedom, the repeated measures F-statistics with $\alpha = 0.05$ from the comfort attitude assessment, is seen in table 3.7. The table shows a significant effect on the walking path, an almost significant effect on the robot behavior, and a non-significant interaction effect. From table 3.8, displaying each passing interaction's mean of the test score, and the standard deviation, it is clear that the right to left walking path was experienced as more comfortable than the left to right walking path, which is

Table 3.7: The results from the comfort assessment. Only walking path was significant

Source	SS	df	MS	F	p
Robot behavior	24.3289	0.6641	36.6358	4.6302	0.0572
Walking path	174.0132	0.6641	262.0383	6.2329	0.0327*
Walking path× Robot behavior	18.32	0.6641	27.1252	3.0376	0.1068

$^*p < 0.05$

Table 3.8: Mean test score and standard deviation for each passing interaction. Nota bene, the comfort score is reverse: a low score indicates high comfort and vice versa.

	PI_1	PI_2	PI_3	PI_4
Mean	12.9474	12.7895	16.9474	14.8421
Standard deviation	6.1867	5.5134	6.7287	5.9746

in line with secondary hypthesis. Nota bene: the test score scale is reversed, and thus a low test score indicates a high degree of comfort (towards an interaction with the robot) and vice versa. The main effect of the walking path is also shown in the interaction figure 3.10.

Additionally, the data are depicted with a waisted box plot in figure 3.9. Each box in the figure shows statistics for the test scores from each passing interaction. The narrow part of a box represents the mean test score. The horizontal lines delimiting the box are the lower and upper quartile values, corresponding to the distribution of half the sample. The remaining half of the data's extent are depicted with the whiskers from each box (dashed lines).

The main effect on the walking path could be caused by several things. Were there sequence effects? This question can be answered by comparing the mean difference scores from each subject group with each other (seen in table 3.3). A difference score is calculated as the test score difference between the robot behaviors from the same walking path. Thus, a subject in a subject group contributes with two difference scores.

Comparison between several mean values is preferably perfomred using ANOVA. A prerequisite is that the relative scores have normal distribution. A Lilliefors test with $\alpha = 0.05$ was conducted on all four groups, with the results shown in table 3.9. The results show that the fourth subject group's difference test scores did not conform to normal distribution. Thus, a nonparametric version of the ANOVA test should be used instead. A suitable test is the Kruskal-Wallis test with the null hypothesis that the samples are different. With $\alpha = 0.05$, $p = 0.3028$, and thus the difference scores' differences among the subject groups were not statistically significant. This is also shown in the boxplot in figure 3.11.

However, visual inspection of the figure indicates a similarity between group one and two, and group three and four. From table 3.3 together with table 3.1, it is clear that the first and second groups began with the right to left walking path, and the third and fourth groups began with the left to right walking path. In order to establish whether or not there is a factor that confounds the main effect on the walking path, the difference scores from the first two subject groups were compared with the two last

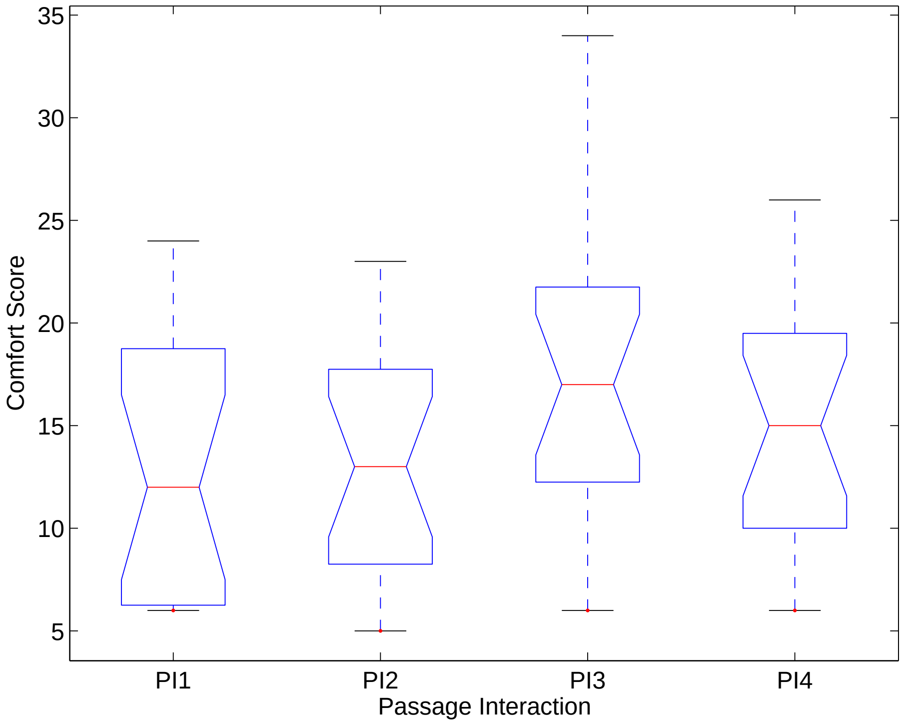


Figure 3.9: A box plot of the four passing interactions.

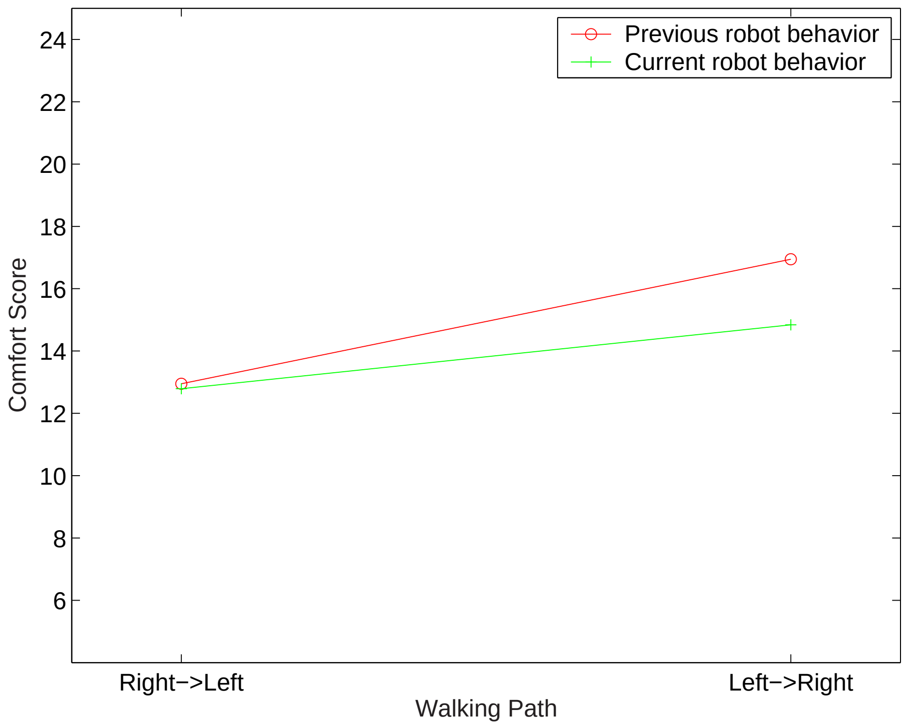


Figure 3.10: The robot behavior×walking path interaction. It should be noted that the minimum comfort score is 6 and the maximum is 42. The results in table 3.7 show that the difference in walking path is significant, and that people preferred the right→left walking path rather than the left→right walking path.

Table 3.9: Lilliefors-statistics establishes whether or not the relative robot behavior test score for each subject group conforms to the normal distribution. The result in the fourth group did not have a normal distribution.

	PI_1	PI_2	PI_3	PI_4
p-value	NaN	NaN	NaN	0.0379
lstat	0.2028*	0.1690*	0.2095*	0.2990
CV	0.2850	0.2490	0.2580	0.2850

$^*p < 0.05$

Table 3.10: Lilliefors-statistics for both the women's and the men's groups.

	Women's Passing Interaction				**Men's Passing Interaction**			
	PI_1	PI_2	PI_3	PI_4	PI_1	PI_2	PI_3	PI_4
p-value	0.0947	NaN	0.1870	0.1805	0.1585	NaN	NaN	NaN
lstat	0.2513	0.3262	0.2256	0.2269	0.2225	0.0947*	0.2146*	0.1483*
CV	0.2710	0.2710	0.2710	0.2710	0.2580	0.2580	0.2580	0.2580

$^*p < 0.05$

subject groups. A Lilliefors test with $\alpha = 0.05$ resulted in that only the difference scores of the last two groups were of normal distribution, and thus a one-tailed Mann-Whitney test was conducted instead of a t-test; the null hypothesis, that the subject group one and two were together smaller than subject group three and four together, could not be rejected due to $p = 0.9661$.

Gender differences were also evaluated. With the attempt to make a $2\times2\times2$ mixed repeated measures analysis where gender was the inter group variable, a Lilliefors test was first conducted with $\alpha = 0.05$ including all test scores of the men and the women. The Lilliefors-statistics is shown in table 3.10 and displays that the women's second passing interaction's test scores were normal distributed.

Therefore a simple approach was tried to compare the groups—a Mann-Whitney test was conducted. The result was $p = 0.0014$, and thus the null hypothesis that there was no difference between the genders' test scores was rejected. As the Mann-Whitney U-value was half the size for the women compared with the men ($U = 416.5 < 1023$), a conclusion was drawn that women found the robot more comfortable.

As the main effect on the walking path apparently does not originate from a sequence effect, the most probably cause is from the setup. The right to left walking path was deliberately designed to have a larger interaction distance than the left to right walking path, as was explained in 3.5.3. Additionally, this is depicted in the figures 2.13 and 2.14.

3.10 Discussion

The discussion is divided into two parts, consisting of a result discussion and a general discussion. The former involve speculations regarding the causes of the hypotheses

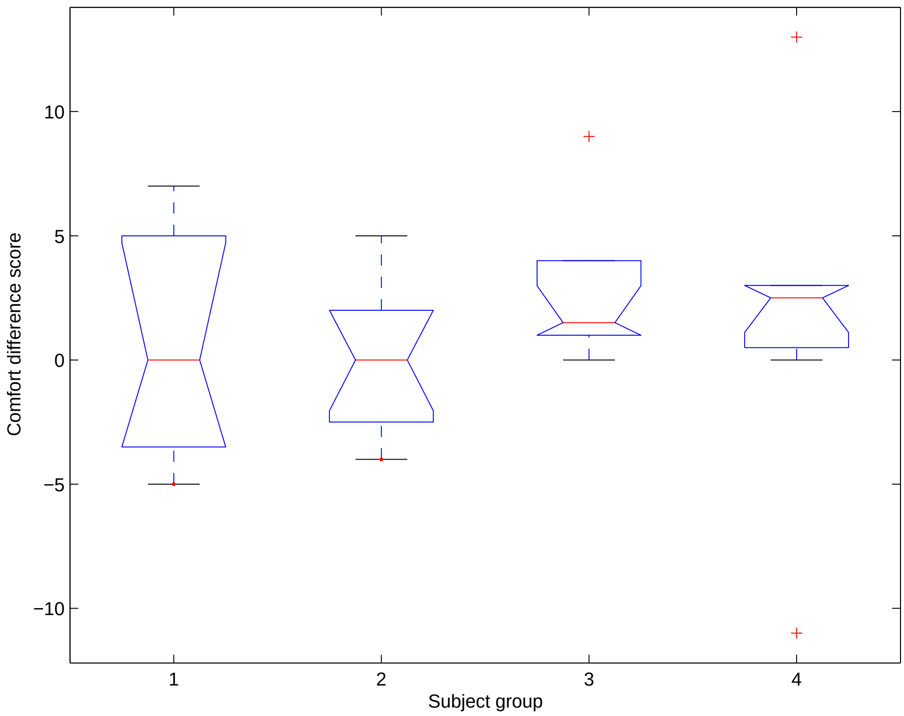


Figure 3.11: Boxplot of a Kruskal-Wallis test, comparing the difference score between the different subject groups shown in table 3.3. The results showed that no difference existed between the groups.

results and connection to the literature. The latter discusses in general terms results that were not part of the hypotheses, also with connection to the literature.

3.10.1 Results

Why was there no main effect on the robot behavior? There are several possible answers to that question.

The sequence effect does not seem to be the answer as there were no significant difference between the subject groups. However, a sequence effect must have existed, but its effect was negligible.

Another answer to said question could be the gender difference found, i.e. that the women in general had lower comfort scores than the men. From the results displayed in table 3.10, it was of interest to observe the data distribution for both genders with a normplot. In the figures 3.12(a) and 3.12(b) the women's test scores and the men's test scores are normplotted, respectively. A normplot is a way of illustrating how close a set of data are to the normal distribution, which is shown as a diagonal line in both figures. The more the data (plusses) are aligned with the diagonal line, the greater the tendency of the data to lean towards a normal distribution.

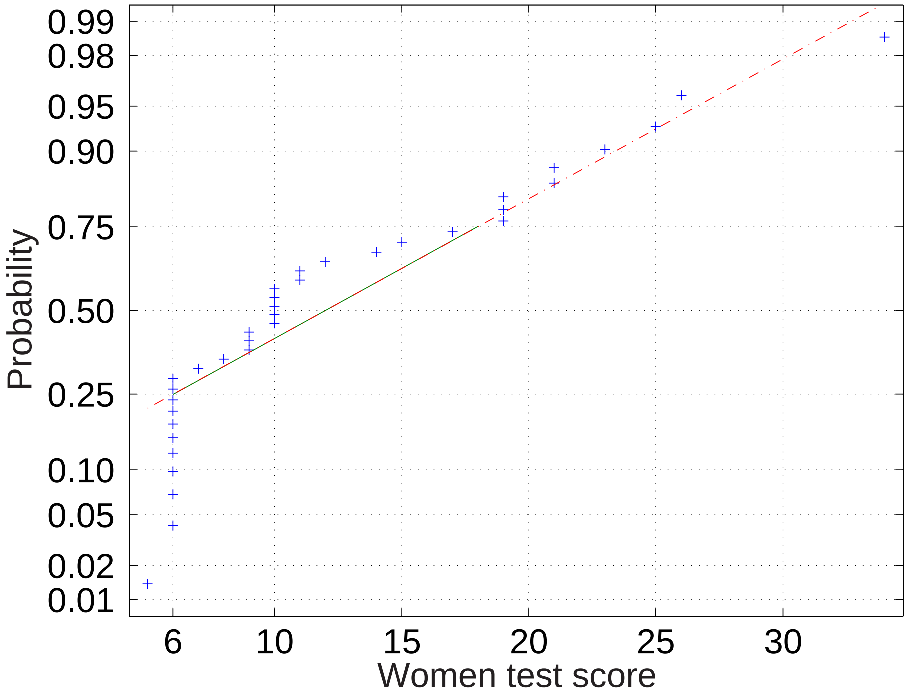
Looking at the data of the women, there seems to be a floor effect. Several test scores are placed at the bottom of the comfort scale (6) compared to the men's test score, which are spread in the interval from 12 to 22. Could a floor effect of the women thus have resulted in the non-significant robot behavior?

Before the question is answered, an analysis of the floor effect tendency is performed. A floor effect does not originate from the comfort questions, given the assumption that men and women perceive language relatively similarly; no tendency of floor effect is found in the normplot of the men, and therefore a possible explanation regarding the indicative floor effect is that women experienced the robot as being more comfortable then men. Thus, the question is rephrased: Did a gender difference in perceived comfort, result in the non-significant robot behavior?

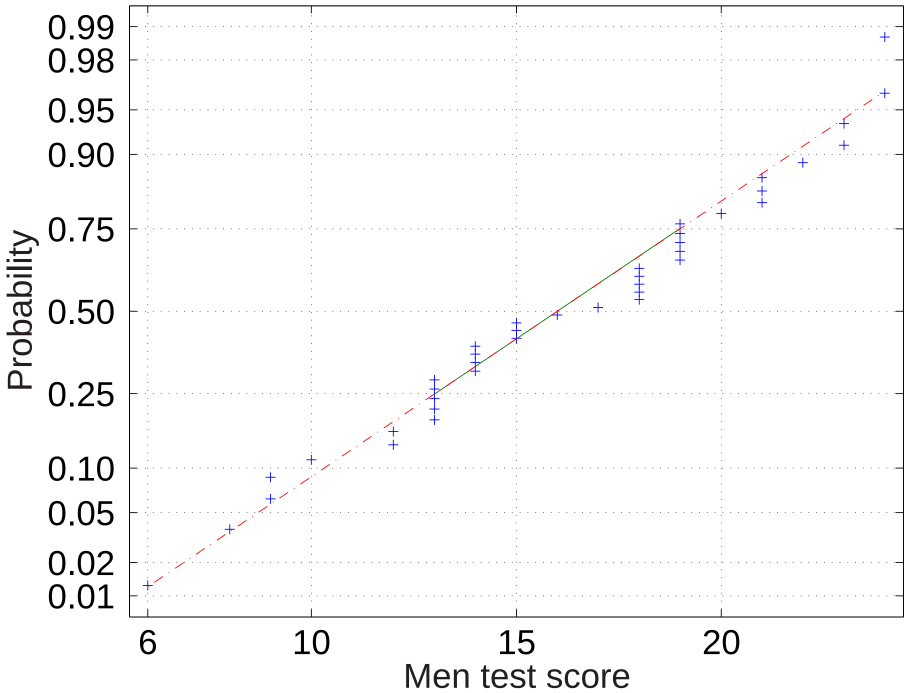
A gender dependency has been found in spatial behavior research. The spatial behavior differs depending on the genders taking part in an interaction; the only conclusive results are that a male-male interaction uses more space than a female-male and female-female interaction. The order of the last two setups has not yet been established. The reason men use more space than women is explained by the fact that men are taller than women and proxemic behavior is relative to body size [16]. This explanation is strengthened, as it has been shown that the interaction distance increases with age through adolescence [1].

With the assumption that men and women walked both walking paths with approximately the same variance, the mean distances towards the robot were constant in comparison of the genders. Thus, with the theory of spatial behavior research, one conclusion could be that women found the robot more comfortable as they experienced the difference in interaction distance as negligible; and this would explain the indicative floor effect in terms of proxemics. The interesting question is then: Did the *men* experience a difference in robot behavior as they did not have a floor effect and with the assumption that they had the same mean interaction distances as the women?

A repeated measures calculation from the men's test score, answers that question. The men's test scores were of normal distribution, and the covariance matrix did have sphericity and homogeneity. With $\alpha = 0.05$ the F-statistics for the men are shown in table 3.11. The results for the men alone are the same as those for men and women



(a) The normplot for the women's data.



(b) The normplot for the men's data.

Figure 3.12: Normplot: the alignment of plusses with the diagonal line show the tendency toward normal distribution.

Table 3.11: The men's results from the comfort assessment. Only walking path was significant.

Source	SS	df	MS	F	p
Robot behavior	32.4000	1	32.4000	3.6633	0.0879
Walking path	324.9000	1	324.9000	5.5899	0.0423*
Walking path×Robot behavior	8.1000	1	8.1000	0.0040	0.9509

$^{*}p < 0.05$

together: The only significant main effect is the walking path. Thus, the answer to the latter question is no. The men did not experience the robot behavior differently from the women. The gender difference in both theory and test score results does not therefore explain the non-significant robot behavior result.

Could this non-significance be explained by how subjects experienced the gender of the robot? As the difference in space usage between female-male and female-female interaction is inconclusive, perceiving the robot as being a female would have explained the non-significance. Survey and interview results do, however, point to the fact that people in general find the gender of a robot as being neutral. The hypothesis of a female robot would thus not hold.

Inconsistencies in robot movements were observed by the author. Could they have "ruined" a possible significant effect on robot behavior? The expected movement action for each passing interaction is shown in table 3.1. The movement actions are opposite one another with regard to each robot behavior for both walking paths. However, this was not the result for all passing interactions. In a few cases where the subject walked fast, and consequently had walked across the corridor too fast, the current robot behavior reacted in almost the same way as the previous. However, a subject eliciting the "wrong" movement action from the current robot behavior, did also elicit a somewhat "different" movement from the previous robot behavior. Alas, the difference between the robot behaviors was smaller than it should have been. But as these inconsistencies only were approximately four, they probably did not "ruin" the study.

Another explanation regarding the non-significance on the robot behavior is also because of inconsistencies, but not to setup design or subject issues, but from the actual inconsistency of the robot behavior. Even though the same input in form of a corridor is received by the robot, the resulting movement output may sometimes be different. This difference is thus rare, but it should not be disregarded. How big the inconsistencies of the robot behavior were and how it affected the result is however hard to tell. But a little difference would have significantly affected the robot behavior, as it was just above 0.05, and this little difference could have been on the "wrong" side of the p-value in this study. More participants and a method of measuring the inconsistency in the robot would have explained if the difference was on the wrong side or not. Also, a part of this difference might have come from fast walking subjects.

Finally, to give a possible answer to the question that was asked in the beginning. The difference between the walking paths and in the robot behavior is clearly evident in figures 2.13 and 2.14. But, simply, it might not have been perceived by the subjects. This could be a very plausible explanation regarding the non-significant robot behavior effect.

3.10.2 General

In the study's interaction, the subject took two objects while walking along each walking path. If the person paid too much attention in taking the objects, the person might have looked more at the objects to take, than the surroundings. This could have made the spatial interaction between the subject and the robot to be an unconscious reaction, which is theoretically beneficial. However, it might also have lessened the difference experienced from the robot behaviors if the subject paid too little attention to the robot.

A survey, consisting of answers from 134 people living in Sweden, focused on attitudes towards intelligent service robots. The results showed that there are several factors which people are consciously aware of. One factor is the size of the robot which was to have a height from 100 cm to 150 cm [20]. The robot in this study had a height of 130 cm, and was assumed not to "scare" people by it's height, inducing a bias. However, the size of the robot relative to subject was speculated to be a factor of comfort scoring: A small person will see the robot as less comfortable. With the null hypothesis that height and test score were uncorrelated, the correlation coefficient r was calculated with $\alpha = 0.05$. The result was $p = 0.0002$, and thus the correlation $r = 0.4138$ is valid. However, the percentage of variation, $r^2 = 0.1712$, is very low, and does not strongly support the speculation that the height of people influences a test score.

Proxemics is culturally bounded, meaning that different cultures have different spatial behaviors [16]. There are some findings from questionnaire data indicating "that individuals from different cultures did not attach the *same meaning* to the same components of proxemic behavior." However, much research regarding cultural differences of proxemic behavior are non-interactional, making their validity uncertain [1]. Consequently, there is an indication that the results from this user study are only valid in a Swedish culture setting. This implies that the same setup could have resulted in floor effects with a culture where interactional space is less than in Swedish culture. Spatial behavior research defines culture as either being a contact culture or a non-contact culture, where the former uses less interactional space than the latter. Northern Europeans, Asians and North Americans are said to have non-contact cultures. Arabs, Latin Americans and Southern Europeans are said to have contact cultures [1].

There are other aspects that influence the interactional space, one of which is personality. But as with the cultural aspect of proxemics, research regarding the personalitie's influence on proxemics is beset with a validity problem as most research has been conducted using simulation methods. Anyhow, results indicate that self-esteem and locus of control are two important factors. People that have lower self-esteem need to be further apart when interacting with on another. This is explained by the fact that people who feel more confident about themselves seem to be more inclined to approach others more closely and to allow others to approach them more closely. People who feel that they have control of a situation will allow others to approach them more closely or vice versa. Also, they react negatively with respect to more remote distances [1].

The personalities of the respective participants were unknown. However, some subjects were notably nervous prior to the first interaction, and it can be assumed that people felt that they had no control of the robot. Thus, the test scores for the familiarization interaction should have been generally higher, if people felt uncomfortable about the first trial compared to the trials that followed. Since all the test scores were found to have normal distributions in a Lilliefors test with $\alpha = 0.05$ ($p = 0.1072$, $lstat = 0.1776$, $CV = 0.1950$), four one tailed t-tests ($\alpha = 0.05$), with the null hypothesis that the familiarization interaction was equal in comfort to all passing

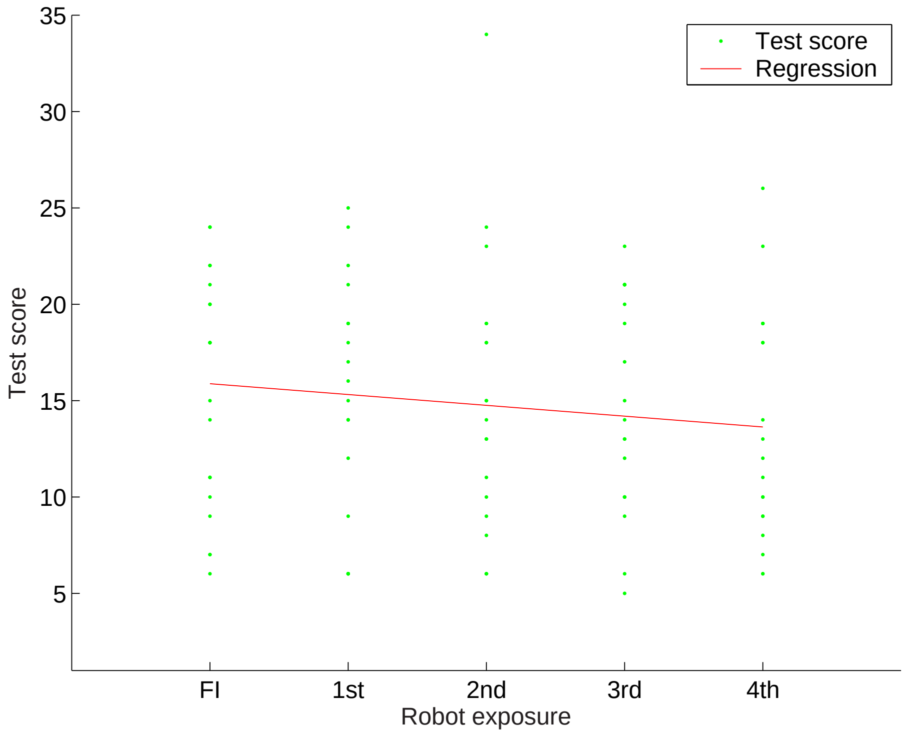


Figure 3.13: The test scores for each robot exposure with a robust regression line showing the tendency for the test scores as number of robot exposures increases.

interaction, were conducted. The t-tests showed that none of passing interactions was experienced as being different from the familiarization interaction in terms of assessed comfort ($p = 0.0931$, $p = 0.0702$, $p = 0.7342$, $p = 0.3446$). Does this imply that the inclusion of a familiarization interaction was unnecessary, i.e. that the habituation effect did not exist? Probably not; the t-test only indicates that the differences between the passing interactions were not measurable. Additionally, it is the author's strong opinion from observations that participants in general behaved differently when comparing the first exposure with the second.

But if habituation existed, there should have been a general tendency towards lower test scores with increasing number of robot exposures experience by participants. A robust regression was performed in Matlab using all five robot exposures. The result can be seen in figure 3.13 where the regression line has a gradient of $k = -0.4456$. As there is a clear negative tendency with increased number of robot exposures, this would imply two things. Firstly, the habituation existed throughout all five exposures. Secondly, there is support for the observations that subjects behaved differently between the first two robot exposures.

Chapter 4

Conclusions

What are the conclusions of this thesis? The answer to this question resides in the extent to which its goals have been fulfilled.

This thesis consists of a robot behavior implementation and its evaluation in a user study.

The behavior implementation includes an algorithm that extracted direction of motion from positional data. Additionally, the implementation showed good results considering the simple techniques utilized; limitations of the FIR filter were overcomed with the creation of ad hoc filtering techniques.

The decision module was implemented, and integrated with the previous robot behavior, and was found to work effectively in a corridor environment with respect to the conducted study. Extensive tests were performed both in simulation and with the robot.

A user study was designed and conducted to assess the subject comfort during the robot exposures in a corridor passing interaction. The design of the study setup was based upon a pre-study, and thus the setup was refined to a degree that yielded valid and reliable data. Moreover, the statements for the comfort assessment were based on a pre study, resulting in very reliable answers in the user study.

One of the hypotheses was rejected, although the data were both valid and reliable. This would imply that the data collection and setup were properly performed and designed, respectively. The probable reason behind the rejection of the robot behavior hypothesis was that there is no difference in experienced comfort between the robot behaviors. Although difference might exist, this study did not find one in terms of comfort attitude.

The user study did originally have 29 participants, even though the available time with the robot was limited by several factors. Moreover, the goal of at least 20 participants was almost reached with 19 participants.

4.1 Restrictions

This thesis work has dealt with a lot of problems. Most of which were solved in one way or another. Some problems, however, were not solved, but before these are mentioned, an explanation regarding robotic research with user studies needs to be described. Robotics by itself is a complex research field. It involves physics, mathematics, signal theory, computing science, to name but a few research areas. With the addition of a user study of social robotics, the complexity is increased by a power of two. The additional application of statistics and psychology in making a robot behave and move in a pre-specified way is in itself almost work enough for a master's thesis.

4.1.1 Implementation

The DoM filter was observed to give a fluctuating result to some people. Probably this was because the author structured the DoM filter solely on the basis of his walking data.

The delay of ~ 2 s from the decision module is very large. This has the effect that the implementation does not work in "normal" settings, i.e., a more narrower corridor than the one used in the experiment, and other direction-change-angles larger than θ. Additionally, no objective measurement was created to enable a formal method of measuring the accuracy of the filter to be acheived. However, it was not in the concept of the thesis to develop the best technique of extracting the DoM.

4.1.2 User Study

Many subjects were not included. Better planning should have been utilized. With more participants, the main effect regarding robot behavior would have been more certain.

The direction of change, θ, can be seen as somewhat unnatural. This was noticed in the pilot study and the main study when subjects seemed to interpret the instruction "Also, walk towards the telephone and pick that up" as "Also, walk towards the other side of the corridor and pick up the telephone"; the result was a directional change greater than θ.

4.2 Limitations

When the robot initially was tested in a corridor 2.3m in width, the delay of the decision module was too large to use the direction information, i.e. a person will cross the center of the corridor before the movement action state is defined. This wil result in current robot behavior that is identical to previous robot behavior. Thus, the functionality of the current implementation is limited to wider corridors and to people walking in a direction similiar to a small θ.

Since people cross corridors in several different ways, there is more to consider than the direction of movement originating from the legs. More importantly, though, is that people, seen from a raw data view, swing their direction back and forth. The aimed direction is somewhere between the extremeties. However, the "true" direction of movement is hard to establish while still using a filter that is sensitive enough to tell when a human being changes direction slightly.

4.3 Future Work

Several things can be done to improve movement extraction. Some points have been mentioned earlier while some have not. As described in the chapter 2 discussion, a Kalman filter can be implemented to filter each leg's position, instead of filtering the CoM. Subjective measures other than comfort attitude can be used, such as "intelligent behavior", "human like" combined with other types of attitude scale assessments such as Thurstone or Guttman scales.

An even bigger improvement would be to create a new scale, specifically appropriate when pertaining to social robotic research. The scale could be structured with the aid of several techniques:

– Creation of a standardized scale for use when assessing how a robot is experienced by a subject. This scale can be "true" in a sense that it is based on physiological stress indicators, such as heart beating, perspiration et cetera.

– Normalization of an attitude score with different measures, such as the minimum or the average distance when passing the robot.

– Longer habituation period for a subject than one period of interaction exposure.

– Use of an heterogeneous group instead of an homogeneous.

– Create a model of the gait of an human being to anticiapte the movement of a person.

– Above all, test all different DoMs in each DAD, i.e. use more scenarios that are possible in a corridor environment.

The previous work is not yet able to deal with more than one person, but it will. An interesting implementation would be a decision module that receives multiple movement-data and returns a priority list of suitable directions for the robot to move, in accordance with proxemics.

Chapter 5

Acknowledgments

Big thanks to my supervisor Elena Pacchierotti for all her patience, help, and explanations. Also to Ali and Babak for explaining and discussing signal theory with a novice and for being moral boosters. My internal supervisor is to be thanked for his comments and email discussion about how to proceed and his comments on this report. Christian gave very good input on all sorts of topics regarding my work. Daniel, for sharpening a picture and encoding of the data. Also, Thomas and Ken for their proof reading and laughter.

Finally a big thanks to my family and friends who gave me a lot of moral support during my monologues of my troubles. Additionally, a big thank to Alan who helped me with my English in this report and Leo for all discussions about statistics, its practice and, above all, how to look at things from a different perspective. Also, without the music from www.di.fm, this thesis would never have been finished.

References

[1] J. R. Aiello. Human Spatial Behaviour. In D. Stokels and I. Altman, editors, *Handbook of Environmental Psychology*. John Wiley & Sons, New York, NY, 1987.

[2] P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, and H. I. Christensen. Navigation for human-robot interaction tasks. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 1894–1900, April 2004.

[3] W. E. Arnold, J. C. McCroskey, and S. V. O. Prichard. The likert-type scale. *Today's Speech*, (15):31–33, 1967.

[4] N. J. Birkett. Selecting the number of response categories for a likert-type scale. In *Survey Research Methods Section*, volume XII of *COGNITIVE ASPECTS OF SURVEY METHODOLOGY*. American Statistical Association, 1986.

[5] J. T. Butler and A. Agah. Psychological Effects of Behavior Patterns of a Mobile Personal Robot. *Autonomous Robots*, 10(2):185–202, March 2001.

[6] P. C. C. and C. A. M. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 1(104):1–15, March 2000.

[7] D. Chen, J. Yang, and H. D. Wactlar. Towards automatic analysis of social interaction patterns in a nursing home environment from video. In *6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, volume Proc of ACM MultiMedia 2004, pages 283–290, New York, NY, October 2004.

[8] J. Dietsch. The human interface robot. `http://www.activrobots.com/ROBOTS/peoplebot.html`, 2005. ActivMedia's home page [Online; accessed 02-Oct-2005].

[9] T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3):143–166, March 2003.

[10] G. D. Garson. Data levels and measurement — pa 765 statnotes: an online textbook. `http://www2.chass.ncsu.edu/garson/pa765/datalevl.htm`, 2005. [Online; accessed 22-Nov-2005].

[11] G. D. Garson. Scales and standard measures — pa 765 statnotes, an online textbook. `http://www2.chass.ncsu.edu/garson/pa765/standard.htm`, 2005. [Online; accessed 22-Nov-2005].

[12] G. D. Garson. Reliability analysis — pa 765 statnotes, an online textbook. `http://www2.chass.ncsu.edu/garson/PA765/reliab.htm`, 2006. [Online; accessed 03-Feb-2006].

[13] M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun. *Cognitive Neuroscience, The Biology of the Mind.* W. W. Norton & Company, New York, 2nd edition, 2002. page 315.

[14] J. A. Gliem and R. R. Gliem. Calculating, interpreting and reporting cronbach's alpha reliability coefficient for likert-type scales. In *Midwest Research to Practice Conference in Adult, Continuing, and Community Education.*, pages 82–88, Columbus, Ohio, Ochtober 2003.

[15] G. R. Griffin. Finite impulse response faq. `http://www.dspguru.com/info/faqs/firfaq.htm`, 2005. [Online; accessed 14-Nov-2005], The Finite Impulse Response FAQ is published on Iowegian International Corporation's web site (`http://www.dspguru.com/.net/.org`).

[16] E. Hall. *The Hidden Dimension.* Doubleday & Company, New York, anchor books edition 1969 edition, 1966.

[17] E. T. Hall. *Handbook for Proxemic Research.* Society for the anthropology of visual communication, Washington, 1974. p. 2.

[18] D. C. Howell. *Statistical Methods for Psychology.* Duxbury, Thomson Learning, 5th edition, 2002.

[19] W. J. Ickinger. Proxemics research — improving our understanding of human behavior through research with a behavioral game methodology. `http://sharktown.com/proxemics/intro.html`, 2006. [Online; accessed 18-Jan-2006].

[20] Z. Khan. Attitudes towards intelligent service robots. Report IPLab-154, TRITA-NA-P9821. Stockholm Royal Institute of Technology, 1998.

[21] R. L. Knoblauch, M. T. Pietrucha, and M. Nitzburg. Field studies of pedestrian walkning speed and start-up time. *Transportation research record 1538, Transportation Research Board*, pages 27–38, 1996.

[22] H. Mizoguchi, T. Sato, K. Takagi, M. Nakao, and Y. Hatamura. Realization of expressive mobile robot. In *Proceedings of the 1997 IEEE International Conference on Robotics and Automation*, pages 581–586, Albuquerque, NM, April 1997.

[23] J. Munshi. A method for constructing likert scales — research report. `http://www.munshi.4t.com/papers/likert.html`, 1990. [Online; accessed 18-Dec-2005].

[24] Y. Nakauchi and R. Simmons. A social robot that stands in line. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 1, pages 357–364, October 2000.

[25] J. C. Nunnally. *Psychometric Theory,.* McGraw-Hill Book Company, New York, 3rd edition, 1994.

[26] E. Pacchierotti, H. I. Christensen, and P. Jensfelt. Human-robot embodied interaction in hallway settings: a pilot user study. In *Proc. of the IEEE Int. Workshop on Robot and Human Interactive Communication, to appear*, Nashville, TN, August 2005.

[27] N. Santavirta. Introduction to statistical analysis — biomaterial graduate school, the 2nd annual meeting of the national biomaterials and tissue engineering graduate school. `http://www.helsinki.fi/science/bgs/1.ppt`, 2005. [Online; accessed 23-Nov-2005].

[28] L. Scott-Webber. *In Sync: Environmental Behavior Research and the Design of Learning Spaces.*, chapter 2nd, pages 9–25. Society for College and University Planning, Ann Arbor, MI, 2004.

[29] K. Severinson-Eklundh, A. Green, and H. Hüttenrauch. Social and Collaborative Aspects of Interaction with a Service Robot. *Robotics and Autonomous Systems*, 42(3):223–234, March 2003.

[30] S. Shibata and H. Inooka. Psychological evaluations of robot motions. *Int. Journal of Industrial Ergonomics*, (21):483–494, 1998.

[31] T. Triggs and S. Moss. Between-subject versus repeated-measures designs — research design and analysis (rda) unit. `http://www.med.monash.edu.au/spppm/research/rda/Between%20subject%20vs%%20repeated%20measures.htm`, 2002. [Online; accessed 18-Dec-2005] Professor Tom Triggers and Dr Simon Moss are in part responsible for the RDA.

[32] W. M. K. Trochim. Likert scaling — the web center for social research methods. `http://www.socialresearchmethods.net/kb/scallik.htm`, 2005. [Online; accessed 22-Nov-2005].

[33] R. Vaughan, B. Gerkey, and A. Howard. On device abstraction for portable, reusable robot code. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 2121–2127, Las Vegas, NV, Oct. 2003.

[34] M. L. Walters, K. Dautenhahn, R. te Boekhorst, K. L. Koay, C. Kaouri, S. Woods, C. Nehaniv, D. Lee, and I. Werry. The influcence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. In *Int. Workshop on Robots and Human Interactive Communication*. IEEE, 2005.

[35] M. Yoda and Y. Shiota. Analysis of human avoidance motion for application to robot. In *Proc. of the IEEE Int. Workshop on Robot and Human Interactive Communication*, pages 65–70, November 1996.

[36] M. Yoda and Y. Shiota. Basic study on avoidance motions for human behaviors. In *Proc. of the IEEE Int. Workshop on Robot and Human Interactive Communication*, pages 318–322, 1996.

[37] M. Yoda and Y. Shiota. The mobile robot which passes a man. In *Proc. of the IEEE Int. Workshop on Robot and Human Interactive Communication*, pages 112–117, September 1997.

Appendix A

Appendix

A.1 Questionnaire

Den studie du snart ska vara med i är helt frivillig och kan avslutas när som helst. Syftet med studien är att studera hur vi uppfattar robotar till vardags. Studien omfattar fem omgångar där du var gång ska gå mot roboten. Totalt tar försöket 10 minuter.

Innan försöket börjar behöver du svara på nedanstående enkla frågor. Svaren kommer endast att visas i sammanställd form där enskilda individers svar ej går att utröna. Efter försöket kommer du som tack för hjälpen att få kaffe, te eller chokladdryck.

Har du några frågor? Tveka inte att ställa dem till Daniel!

Kön:

Längd:

Ålder:

Modersmål:

Sysselsättning:

Högsta utbildningsnivå:

Tidigare kontakt med robotar:

Lämna detta papper till Daniel när du är klar. Tack!

A.2 Statements

Vänligen besvara i vilken grad påståendena stämmer för senaste gången du och roboten gick i riktning mot varandra:

Påstående	**Stämmer absolut** (1)	(2)	(3)	**Neutral** (4)	(5)	(6)	**Stämmer absolut *inte*** (7)
Jag kände mig bekväm när jag och roboten rörde oss i korridoren							
Jag kände mig negligerad av robten när vi rörde oss i korridoren							
Jag kände mig lugn när jag och roboten rörde oss i korridoren							
Jag kände mig osäker med roboten när jag gick i korridoren							
Jag kände mig obekväm när jag mötte roboten i korridoren							
Jag upplevde att roboten tog hänsyn till mig när jag och roboten rörde oss i korridoren							

A.3 Statements Evaluation

Syftet med försöket du nyss varit med om har varit att mäta din personliga *känsla* av bekvämlighet/trevnad (comfort) vid robotexponeringarna. Tänk dig nu en försöksperson som *inte* har någon tidigare erfarenhet av robotar. Besvara i vilken grad du tycker att påståendena du bedömde nyss, mäter "comfort" för robotexponeringen du nylingen upplevde. Läs helst föregående mening en gång till och kom ihåg att påståendena kommer besvaras av personer utan roboterfarenhet. Detta är alltså en utvärdering av utvärderingarna.

Påstående	**Stämmer absolut** (1)	(2)	(3)	**Neutral** (4)	(5)	(6)	**Stämmer absolut *inte*** (7)
Jag kände mig bekväm när jag och roboten rörde oss i korridoreni							
lJag kände mig besvärad när jag mötte roboten i korridoren							
Jag kände mig negligerad av roboten när vi rörde oss i korridoren							
Jag upplevde att roboten var bufflig när jag gick i korridoren							
Jag kände mig lugn när jag och roboten rörde oss i korridoren							
Jag kände mig osäker med roboten när jag gick i korridoren							
Jag uppfattade roboten som artig när den och jag rörde oss i korridoren							
Jag upplevde att roboten tog hänsyn till mig när jag och roboten rörde oss i korridoren							
Jag var trygg med roboten när vi rörde oss i korridoren							
Jag kände mig obekväm när jag mötte roboten i korridoren							