



Toward Virtual Reality-based Evaluation of Robot Navigation among People

Fabien Grzeskowiak, Marie Babel, Julien Bruneau, Julien Pettré

► To cite this version:

Fabien Grzeskowiak, Marie Babel, Julien Bruneau, Julien Pettré. Toward Virtual Reality-based Evaluation of Robot Navigation among People. VR 2020 - 27th IEEE Conference on Virtual Reality and 3D User Interfaces, Mar 2020, Atlanta, United States. hal-02473081

HAL Id: hal-02473081

<https://hal.science/hal-02473081>

Submitted on 11 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward Virtual Reality-based Evaluation of Robot Navigation among People

Fabien Grzeskowiak*

Inria Rennes Bretagne Atlantique

Marie Babel†

Inria Rennes Bretagne Atlantique

Julien Bruneau

Inria Rennes Bretagne Atlantique

Julien Pettre‡

Inria Rennes Bretagne Atlantique

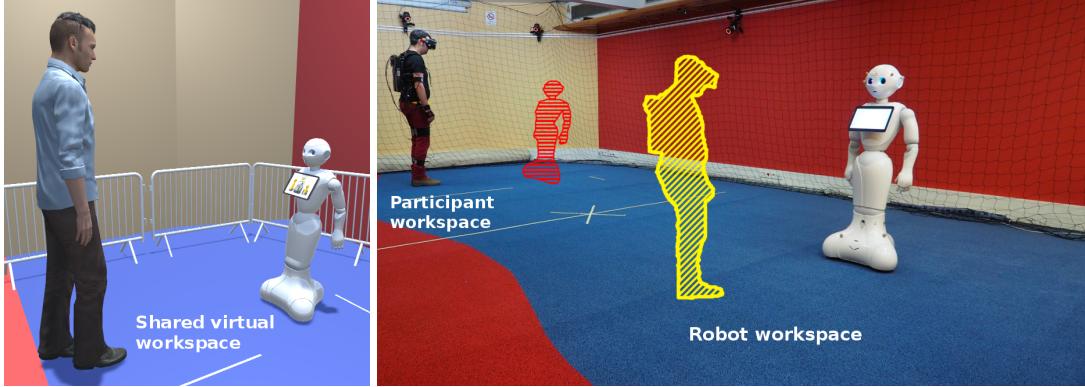


Figure 1: *Left:* the avatars of a real robot and of a real human are face to face in a virtual shared environment. *Right:* Avatars' motion are respectively steered by tracking a real robot and a real human who are physically separated to prevent collision.

ABSTRACT

This paper explores the use of Virtual Reality (VR) to study human-robot interactions during navigation tasks by both immersing a user and a robot in a shared virtual spaces. VR combines the advantages of being safe (as robots and humans interacting by the means of VR but can physically be in remote places) and ecological (realistic environments are perceived by the robot and the human, and natural behaviors can be observed). Nevertheless, VR can introduce perceptual biases in the interaction and affect in some ways the observed behaviors, which can be problematic when used to acquire experimental data. In our case, not only human perception is concerned, but also the one of the robot which requires to be simulated to perceive the VR world. Thus, the contribution of this paper is twofold. It first provides a technical solution to perform human robot interactions in navigation tasks through VR: we describe how we combine motion tracking, VR devices, as well as robot sensors simulation algorithms to immerse together a human and a robot in a shared virtual space. We then assess a simple interaction task that we replicate in real and in virtual conditions to perform a first estimation of the importance of the biases introduced by the use of VR on both a Human and a robot. Our conclusions are in favor of using VR to study human-robot interactions, and we are developing directions for future work.

Keywords: Human Robot Interaction, Motion Tracking, Virtual Reality, Robots simulation

1 INTRODUCTION

This work stands at the interface between Virtual Reality and Robotics. The presence of mobile robots in places shared with humans has become a reality. This fact raises the question of the potential risks caused by the navigation techniques used by these robots in these places, which are dynamic and partially unpredictable by nature. Ideally, to ensure an essential level of safety, all possible configurations of interaction between humans and robots should be evaluated. This is tractable through a theoretical approach, but evaluating them empirically is difficult in practice. There are technical issues such as controlling the numerous interaction parameters, but also ethical ones as some of the most interesting configurations to study should present potential risks of collision between humans and robots.

Simulation, i.e., using algorithms to replicate the environment and the robot behaviors, partially mitigates this problem because it offers the possibility of sampling and standardizing a very large number of configurations to study how interactions unroll without the need to have human participants. However, it raises the question of the level of realism of the simulation, and the transposition of its results to real cases. In particular, the modeling of human behaviour largely remains an open problem.

In this work, our idea is to try to reconcile the advantages of simulation tools (effective and safe) with those of empirical tests involving real robots and humans (realistic conditions).

To do this, we use Virtual Reality (VR) to immerse robots and humans in the same virtual workspace in which they can interact with each other. In reality, they evolve in physically separated workspaces, as illustrated in Fig. 1, which removes any risk of the robot hurting the human.

The idea of using VR in the study of human-robot interactions is not new as discussed in Sect. 2. However, to our knowledge, this is the first time that, a mobile robot is immersed together with a human user to allow a natural interaction. Indeed, the novelty of our work lies in the fact that the sensory inputs of both the human and the

*e-mail: fabien.grzeskowiak@inria.fr

†e-mail: marie.babel@irisa.fr

‡e-mail: julien.pettre@inria.fr

robot are simulated, preserving all the mechanisms of decision and action for each of them. For the robot, this means that it responds to the situation using the on-board electronics, algorithms, control and motors: its real capacities are preserved and replicated in the virtual world by motion capturing the resulting motion.

Our objective is also to explore the ability of such platform to generate realistic situations of interaction, i.e., verifying that the simulation of sensory input we generate for the robot and the human are of enough quality to allow a robot and a human to exhibit interaction behaviors which are similar to reality. This is a classical problem in VR when studying the immersion of a human, but this raises novel question concerning the immersion of a robot. Such a goal is however a long term one, we propose in this work a pilot study to explore whether we face the previously observed perception biases on the side of the human user, and we perform a first evaluation of biases which concern the robot sensing.

Our contributions are twofold:

- We provide the technical description of a VR platform. The design of the platform is specifically adapted to the study of human-robot interactions in navigation tasks.
- We perform a pilot evaluation of this platform, mainly to assess the perception of robot motion in the virtual environment by humans and vice-versa. To this end, we observe the influence of this mutual perception on the interaction.

Our first contribution is described in Sect. 3. Our second contribution is presented in Sect. 4. We set a robot and a human in a simple situation of interaction, face-to-face, and give the instruction successively to the robot and the human to preserve a distance between them while the other is moving forward and backward. We compare how this simple task is performed in real and VR conditions. This way, we assess the perception of motion and distances in VR and compare it to the real case. Our results are discussed in Sect. 4.4, before Conclusion.

2 RELATED WORK

This section provides an overview of the work related to our objectives in the areas of human-robot interaction, VR for robotics as well as VR for human behavioral studies.

2.1 Human-Robot Interaction

Service robotics is an expanding field [31, 48] and people have a growing trust and interest in robotics, and more especially in mobile robotics, i.e. robots that have a moving base. This makes human-aware navigation an active topic [22] with many emerging solutions to adapt robots behaviour to the proximity of humans. Such studies fall in the scope of *Human-Robot Motion* (HRM) defined in [33] as a sub-domain of Human-Robot Interaction (HRI) and focused on mobility aspects. One goal of HRM is to avoid the robot states that can lead to inevitable collisions with a human [18]. The difficulty of this exercise lies within the limits of the robot capabilities to sense and move as well as the predictability of human behaviors [17]. This is why sets of rules and heuristics were proposed to guide the design of robots [10, 42]. Evaluating the capabilities of robots to follow those guidelines is absolutely required before operating them among humans.

The standard work-flow generally includes a validation in simulation, followed by real experiments. For example, [6, 13, 40] used various algorithms to simulate human behaviors during interactions with robots. The use of simulators raises the question of their realism including, for example, the realism of the employed physics engines [21]. Nevertheless, the interest of simulators is undeniable. In particular, the trend is toward using machine learning approaches which are data-intensive, a need that simulator can, at least, partially satisfy. For example, CARLA [14] is becoming popular in the task

of creating photo-realistic environments for robot training and evaluation. However, CARLA is dedicated to autonomous driving cars, a field where interactions with humans are limited since pedestrians are almost absent of roads and their behaviors is essentially driven by traffic regulations.

In service robotics, close interactions between humans and robots are the main issues simulators have to face. They generally include simplistic human models. For instance, Chi *et al.* [6] simulate humans with a simple point moving according to a predefined trajectory and the human-robot interaction is ignored. Real experiments remain then the only valid method to study interactions, at the cost of possible risks for the involved participants. To limit this risk, [27] managed to conduct experiments in real conditions with a lightweight robot moving in a crowded environment.

Positioning. The established practices to evaluate the capabilities of a robot to safely move in proximity with humans are based on two methods: simulation-based and experiment-based evaluations. The level of realism of simulators limits the nature and the usefulness of conclusions. The latter method, experiment-based, raises ethical questions when it comes to evaluating the safety levels of robots, since it requires to set robots and humans in situations of possible danger. This statement justifies the need for an alternative method that can combine the benefit of simulation and real experiment, that is explored in this paper. The means of VR can let real robots and real humans performing an interaction in shared virtual space, whilst they remain physically separated, and thus, without risk of collisions between them. The previous usage of VR in the field of robotics, that also study human behaviors and interactions, is addressed in the next section.

2.2 Virtual Reality for robotics and behavioral experiments

VR provides a unique way to control and manipulate human perception and, for this reason, has been used to study human behaviour [38]. Beyond control of perception, the use of VR is also interesting to exactly reproduce situations over several participants, to facilitate data measurement and recording, or to avoid exposing participants at risks, when dangerous situations for them are studied. For example, VR has been used with post-stroke patients [2, 9] or to study emergency evacuation [28].

It is also used to study human locomotion [7, 37, 44] as well as their interactions with other humans [4, 26, 29, 41] or with their environment [25]. In the latter, Lin *et al.* propose a virtual scene for humans to interact with and suggest that the use cases can be extended to robotics applications. Such application is concretely explored in [20], where VR is used as an interface for industrial collaborative robot control, and in [11], where it is used to teach users how to drive a power wheelchair in safe conditions.

Nevertheless, VR might introduce some experimental biases. For instance, [43] suggests that a real robot is more appealing for humans than a simulated one displayed on a screen. Also, perception in VR is biased. It is known that for instance depth perception [3] is affected, and that objects are perceived closer than they are.

Then, comparing same tasks performed in VR and in real conditions is required to estimate those biases and avoid misinterpretations of experimental results. Such comparisons have been done for many different studies: locomotion toward a goal [1, 8], collision avoidance [1, 5, 16, 30], personal space [19, 28], etc. Also, HRI experiments in VR in [23] explore the bias when the human only is immersed in VR. All these studies show that, while there is some quantitative differences, such as slower walking speed or bigger personal space, humans display similar patterns and behaviors in VR.

Positioning. Our objective is to further explore the use of VR for experimenting with robots. To the best of our knowledge, this is the first time that a real robot is immersed in VR in order to make it capable of operating in a real or a virtual environment in the

same way. In the closest work to our [23], the robot and the human positions are tracked based on HTC vive trackers for them to share their mutual position, the robot is not perceiving the human position through its sensors, and the tracking is a low dimensional one (the posture of the human and of the robot are ignored). The robot has direct access to its externally tracked position, thus not dealing with the robot sensing and processing capacities issues. In our case, this requires in particular to simulate the robot's perception of its virtual environment. In comparison with previous works, we suggest using VR to safely experiment the robot navigation capabilities among humans, while keeping a real robot and real humans in the loop. We then describe in the next Section a platform that fulfills the requirements for such studies. Especially, we decide to immerse both a real robot and a real human in a shared virtual environment, instead of simulated ones, in order to reach a higher degree of realism than simulation. However, because the perception of distance and of motion are crucial in such tasks, our paper first proposes to perform a simple case of interactions in both VR and real world, and to estimate possible biases introduced by the use of VR.

3 HUMAN-ROBOT INTERACTIONS THROUGH VIRTUAL REALITY

In this section, we provide a detailed technical description of our VR platform to study human-robot interactions.

3.1 Overview

Our technical objective, illustrated by Fig. 2, is to create a shared virtual workspace (VW) where the avatar of a human and a robot can interact through VR. We need an accurate representation of the human and its limb in virtual reality, first to give the human a better embodiment feeling, and second to simulate the robot sensors in the VW more realistically. In the same way, the virtual robot should move its joints by mimicking the real robot, so as to provide relevant visual cues to the human. To do so, they both directly control the motion of their avatar using motion capture. The human perceives the VW including the robot motion and its own avatar motion by graphically rendering the VW and displaying it through a Head Mounted Display (HMD). The Robot perceives the VW and the human avatar through virtual sensors that bypass the real ones.

3.2 Equipment

Here is the list of the main components of the platform, which are also represented in Fig. 2:

- A game engine: we used *Unity3D* [46]. The game engine generates the VW and provides tools and plugins to manage it and to link all the different components.
- A mobile robot: we here used *Pepper* [32]. It is a humanoid robot with an omni-directional mobile base built by *Softbank Robotics* to interact with humans. It embeds a processor *Intel Atom E3845*, and is equipped with Lidar, ultrasound and infrared proximity sensors, RGB cameras and RGB-D cameras to detect the surrounding obstacles. It is programmed to stop if it appears to be too close to an obstacle. The stop distance depends on the speed of the robot (0.30 m for a speed of 0.35 m/s, see [12]).
- A root position tracking system: we used the *Vicon* [39]. It is able to track multiple rigid objects using infrared cameras and reflective markers. It comes with a software, *Tracker 3*, which can process data in 1.5ms.
- A body movement tracking system: we used the *Xsens* suit [47] [36]. The suit is composed of 18 Inertial Measurement Units (IMU) connected together, to estimate the body pose.

Xsens ensure real-time, reliable and accurate human motion analysis, and the suit is easy to set up and calibrate.

- A robot movement tracking system: we used Pepper's own *odometry sensors*.
- A VR system: we used the *Fove* [45]. It embeds a WQHD OLED (2560 X 1440) screen, which specifies a frame rate of 70fps, and a field of view of up to 100 degrees. The *Fove* tracking system combines IMU sensor and infrared-based tracking. It embeds gaze tracking possibilities even though they are not used in our current solution.

3.3 Workspaces and Virtual Environment

As shown in Fig. 1, there are three different workspaces used by the system. The robot and the human have their own workspace in the real world, and thus remain physically separated. All their actions, physically performed in their respective workspaces, are translated to the VW (see Sect. 3.3). They also both perceive the VW, instead of their respective workspaces (see Sect. 3.5).

The VW is a 3D scene that can be modeled and visualized using *Unity*. Fig. 2 shows the VW designed specifically for the experiment presented in Sect. 4. To prevent participants and robots to collide with physical obstacles, those in their own workspace should be accurately positioned in the VW (including the walls).

3.4 Sensing Human and Robot Actions

In order to replicate the actions of both the human and the robot in the VW, we need the full body motion. To do so, the human motion is tracked using the Vicon system and the *Xsens* suit while the robot's motion is tracked using the Vicon system and its own odometry sensors, as described in figure 2.

In order to track an object with the *Vicon* system, we need to paste reflective markers all around the object to be detected by the cameras. Then we associate them in the software, creating a specific shape for this object. This system is sensitive to symmetries in the shape, and requires to have enough markers to be visible by the cameras in order to detect the object. The robot has 5 reflective markers on its base and 5 on its head while 7 markers were put on the *Fove* to track the human. These markers can be seen on the *Fove* in Fig. 2. The Vicon system is subject to tracking loss which can be caused by occlusions of markers, networking latency, bad lightning condition... Even if these losses are very brief, as we require real time and precise motion capture, they can lead to unrealistic behavior of the avatars. This can break immersion for the human participant, include huge biases or, even worst, lead to increase cybersickness ([34]). For this reason, a second tracking system is used for the robot and the human.

For the robot, the Vicon tracking system is combined to the robot odometry sensors. The *odometry sensors* estimate the global position of the robot in the environment, and the joints position relatively to the global position. The *odometry sensor* of the robot's base is less precise than the Vicon tracking (or the joint's sensors) due to friction but does not suffer from data loss. In the end, the robot movement is estimated using the data from the odometry sensors and its global position is periodically calibrated using the Vicon data to prevent any drift on the global position. This estimation is only used to replicate the real robot pose in the virtual environment, however, the robot estimates his pose and global position using its own sensors only, in order to keep the imperfections of the robot odometry. In other words, the accurate positional information of the robot is required to simulate correctly sensed data according to the sensor position configuration.

The participant is equipped with a motion tracking suit *Xsens*. The *Xsens* suit estimates the global position of the participant in the environment, and the limb positions relatively to the global position. The embedded IMU accumulates errors over time and the estimated

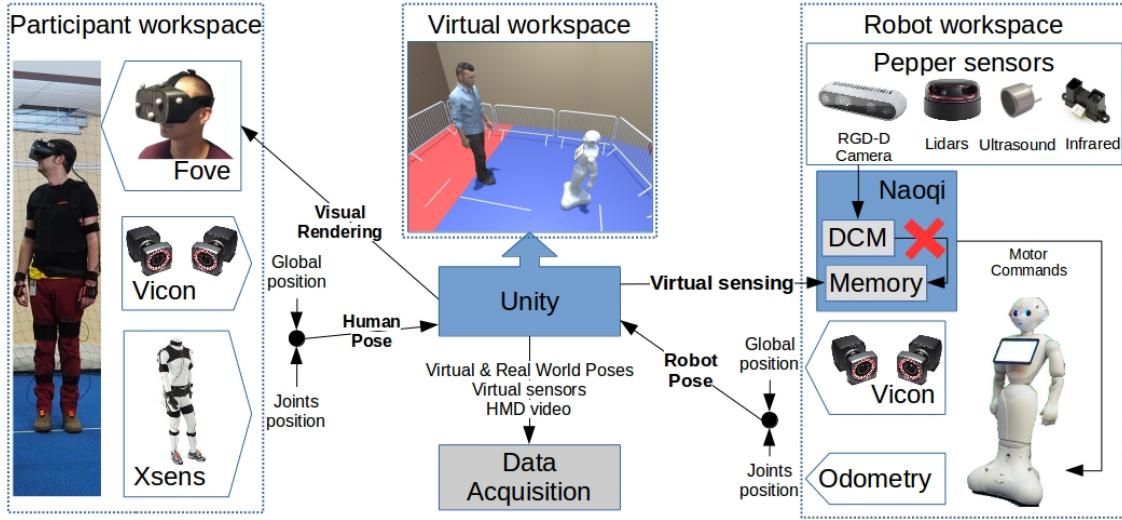


Figure 2: System overview. A 3D virtual scene, the virtual workspace, is replicated from the 2 separated real workspaces in which the human and the robot move. Participants are embodied in an avatar in the virtual workspace thanks to the combination of the tracking systems. The participant perceives the virtual environment thanks to a Fove HMD. The robot *Pepper* can move in the robot workspace. A virtual robot mimics the real one in the virtual environment thanks to the tracking system and its sensors. The virtual robot sensors replace the real robot sensors in the control process of the real robot.

global position of the suit is subject to a noticeable drift after a few minutes of use. The limb positions do not suffer from this drift as they are located relatively to the global position. Same as for the robot, a periodic calibration of the *Xsens* suit is done with the *Vicon* system to prevent any drift on the global position.

3.5 Perception of the Virtual Environment

Human perception The human perceives the VW mostly through vision and tactile sensation. Vision is managed thanks to a head-mounted display (HMD), in our case the *Fove*. While the human avatar is controlled by another tracking system (see Sect. 3.3), the camera rotation, through which the human perceives the VW, is controlled by IMU sensors of the *Fove*. Using the *Fove* sensors allows to be more precise in the calibration and reduces latency issues. This improves the experience of the human and decreases cybersickness. In order to provide the human a greater sense of presence in VR, the 3D environment is accurately replicating the real environment and is superimposed over it. This way, when the human gets to virtually touch a virtual wall, he/she also touch the real wall.

Robot perception First, The real robot perceives the VW based on simulated sensors. Therefore, we need to be able to bypass the sensors embedded into the robot, which might not be possible with a commercial robot. Typically, for the robot *Pepper*, the framework *Naoqi* [12] is used to manipulate the robot. The low level controller of this robot is called DCM, part of *Naoqi*, and handles electronic devices according to commands, then updates the memory with the sensors values. The DCM does not allow us to shut down the real sensors, so we solved this problem by working at the memory level, overwriting the real robot sensor memory as soon as the DCM updated its values.

Sensors Our virtual sensors have to give the same output as real ones. To achieve this, we need to have good knowledge of the robot sensors, by using the data-sheet or testing the sensors in many situations. Our simulated sensors offer the possibility to tune the

usual parameters according to the real sensor properties, such as maximum and minimum distance of detection, maximum incidence angle required for detection, field of view, update rate, bit-precision... In order to provide more realistic sensors simulation, we have the possibility to add various types of noise to any type of sensors. We implemented a Gaussian noise generator, an offset generator and a peak generator, which are configurable and can be added to the output of any kind of sensors described below.

Robot sensor data are generated thanks to the 3D elements of the VW: the scene, and the human avatar. We implemented simulated infrared (IR), ultrasound (US) proximity sensors, 2D laser (LIDAR) sensors, and depth camera (RGB-D), which are the sensors that *Pepper* uses (Fig. 2). The simulated sensors implementation uses the capabilities of *Unity* in terms of physics engine and GPU computation.

For the US, IR, and Lidar, the simulated sensors cast a predefined number of rays in a given detection zone. These rays are generated by the physics engine which returns a collision point when a ray hits a collider, i.e. a component which defines the shape of an obstacle for the purpose of physical collisions. The rays are cast evenly in the detection zone, according to the predefined number of rays. For a more realistic result, we propose to simulate the fact that such sensors fail to detect obstacles if the incidence angle is too high, by giving the possibility to specify a maximum incidence angle. If the ray hits an object in its path, and the hit angle is smaller than this value, then the object is detected and the distance between the sensor and the hitting point is recorded.

US sensors have specific detection zone, which we propose to approximate by a teardrop function centered on the position and the angle of the sensor in the VW. The simulated US sensor returns the smallest observed value from the set of rays. We propose to approximate the detection zone of a IR sensor by a cone centered on the sensor reference in the VW. The system supervisor can configure the maximum distance of detection, the angle of the cone and the level of sampling. We propose to simulate Lidar sensor by casting rays according to a partial sphere. The system supervisor is able

to configure the angular resolution, the starting and ending angles, as well as the minimum and maximum range of the LIDAR. RGB cameras are simulated using simple images extracted from Unity. The depth component of the RGB-D camera is simulated using Unity shaders, which are usually in charge of rendering textures. We use them to generate a depth texture, that gives us a gray scale image where each pixel value corresponds to the depth value. Also, the simulated camera is highly configurable (*e.g.* field of view, resolution...).

4 EVALUATION

4.1 Objectives and Hypotheses

We intend to use our VR platform to study human-robot interactions. To this end, we seek to evaluate the ability of our platform to conduct these studies. In particular, we want to explore robots and humans behaviors by comparing a simple interaction performed both in VR and in reality. Our experiment is however a pilot one: we do not intend to explore novel aspects of perception and behavior in VR. In this new kind of situation of interactions between a robot and a human, we verify that the human perception biases are similar to what was described in the literature, and we report our observations concerning the robot sensing.

We first want to evaluate distant interactions happening in our platform. Our evaluation is based on the experiment of [15], which addresses the question of interpersonal coordination when a pair of participants walk simultaneously. In our case, we replace one of the humans by a robot. We set the robot and the human face-to-face. We give the instruction, alternatively to the robot and to the human, to be the leader, and to perform a back and forth motion. The other one is the follower, and has to keep a constant distance with the leader. Our hypotheses are:

- **H1:** In virtual and real conditions, the trajectories of the participant and the robot will be highly correlated over time.
- **H2:** The virtual condition will induce a higher reaction time from the robot and the participant than the real condition.
- **H3:** In the virtual condition, the human speed will be lower than in real condition, as it is a commonly observed fact [16].

4.2 Experiment Description

Population For this pilot study, 7 unpaid participants, recruited via internal mailing lists among students and staff, volunteered for the experiment. (2F, 5M; age: avg.= $27, 2 \pm 3, 89$, min=21, max=42). They were all naive to the purpose of the experiment, had normal or correct-to-normal vision, and gave written and informed consent. The study conformed to the declaration of Helsinki, and was approved by the local ethical committee. Because of tracking issues that appeared for 2 participants, only the data from 5 participants were finally used in this paper.

Conditions The tasks, detailed below, were performed under two conditions: real and virtual. The condition in which a participant started was chosen randomly.

- **Real condition:** this is our baseline situation. The robot and the human share the same physical workspace. Note that participant head pose was tracked by wearing glasses with reflective markers.
- **Virtual condition:** The participant is immersed with the robot in a virtual copy of the participant workspace. They are physically separated, in their own respective real workspaces, as shown in Fig. 1.

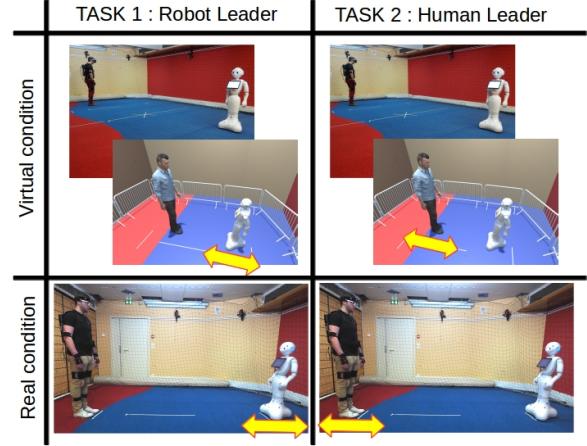


Figure 3: Tasks and Conditions: The leader moved forward and backward with a sinusoidal movement. The follower tried to stay in front of the follower and to maintain a constant interpersonal distance between the leader and the follower. The leader was either the robot (task 1) or the participant (task 2), and the task was performed in VR or in real condition

4.3 Experiment Description

Population For this pilot study, 7 unpaid participants, recruited via internal mailing lists among students and staff, volunteered for the experiment. (2F, 5M; age: avg.= $27, 2 \pm 3, 89$, min=21, max=42). They were all naive to the purpose of the experiment, had normal or correct-to-normal vision, and gave written and informed consent. The study conformed to the declaration of Helsinki, and was approved by the local ethical committee. Because of tracking issues that appeared for 2 participants, only the data from 5 participants were finally used in this paper.

Conditions The tasks, detailed below, were performed under two conditions: real and virtual. The condition in which a participant started was chosen randomly.

- **Real condition:** this is our baseline situation. The robot and the human share the same physical workspace. Note that participant head pose was tracked by wearing glasses with reflective markers.
- **Virtual condition:** The participant is immersed with the robot in a virtual copy of the participant workspace. They are physically separated, in their own respective real workspaces, as shown in Fig. 1.

Tasks We propose 2 interaction tasks inspired by [15]. We define the interpersonal distance as the euclidean distance between the participant and the robot. The participants were asked to move in coordination with the robot, and always started by the task 1.

- **Task 1:** The robot leads the interaction. The participants were asked to stay in front of the robot and to maintain the starting distance (1.5 m) with the robot between the robot and themselves. We applied a sinusoidal velocity command to the robot with a frequency of 0.15 Hz and an amplitude of 0.12 m/s, making it moving forward and backward, as shown in Fig. 3, during one minute.
- **Task 2:** The participant leads the interaction and was asked to have a similar behaviour as the robot in the task 1, as shown in Fig. 3. The participant had to move forward and backward in a periodic movement while facing the robot at all time.

The robot, using its real or virtual sensors, tried to maintain a constant interpersonal distance of 1.5 m during one minute, and tried to stay face to face with the participant. Human detection was done with the real-time object detection neural network tiny-YOLO [35]. The obtained bounding box was used for control in orientation. The distance to the human was given by proximity sensors (US and LIDAR).

Experimental design The environment, shown in Fig. 1, was an empty space shaped by a 8mx5m rectangle, surrounded by walls and a protective net. The room was separated in two 4mx5m squares with a floor marking, delimiting the robot and the participant workspace. Participants started either in virtual condition or in real condition, alternatively. At the beginning of the virtual condition, we suggested to the participants to move around and feel the correspondence between the real walls of the room and the virtual walls of the VW. After a training of 1 minute of the task, participants performed two trials of 1 minute each of the experimental tasks for each of these conditions. In total, each participant performed 8 trials of 1 minute (2 Conditions \times 2 Tasks \times 2 repetitions). A re-calibration of the tracking system was done between each trial.

Collected Data During the experiment, the participant and the robot trajectories were recorded, as well as the robot sensors. Data acquisition used the Vicon system in real condition, and the Xsens suit and the robot inner odometry in virtual condition. The data were acquired with a frequency of 120Hz. Note that since the network communication relies on *ROS* which does not guarantee hard real-time constraints, this frequency is not precisely respected by the system.

4.4 Results

We sampled the data with a fixed frequency of 100Hz and filtered the data with a linear forward-backward filter, on each of the 60 seconds long records. We choose to limit the analysis according to the preferential direction of the interaction. We temporally normalized the data: we removed the 5 first seconds of each record which correspond to the time the robot needs to be fully stabilized in task 2. The statistics take into account all the values of each trials (5 participants per task, with 2x55 seconds of recording for each condition and each task).

Similarly to [15], we propose to analyze the similarities and delays between the leader and follower positions (which correspond to the robot or the human depending on the task) thanks to a partial cross-correlation over the time for each trial (i.e., performed on a moving time window). Our time window is 10s long, and is moved by steps of 2.5ms. Fig. 4 detail our computation over 1 example trial. More precisely, we estimate the variation of the Pearson correlation coefficient in time based on normalized cross correlation, as well as the variation of the delay between the two signals. This process results into both a Pearson coefficient function as well as a delay function over time. Their distribution is plotted for each participant and condition in Fig. 5.

We use a Lilliefors test [24] on the cross-correlation scores and delays, which rejected the normality hypothesis for each condition and task ($p < 0.05$ for each test). We then use a Wilcoxon signed rank test between the two conditions, for each task and for each of those variable. Those tests rejected the null hypothesis of zero median ($p < 0.05$), except for the cross correlation data on task 1 ($p = 0.4620$).

For each task and condition, we computed the interpersonal distance, at each time step for each trial of each participant, which has a value of [avg:1.39; std:0.05] m and [avg:1.03; std:0.06] m for the task 1 in real condition and virtual condition respectively, and [avg:1.73; std:0.17] m and [avg:1.74; std:0.12] m for the task 2 in real condition and virtual condition respectively.

We also computed the velocity of both the robot and the participant, summarized in the Table 1.

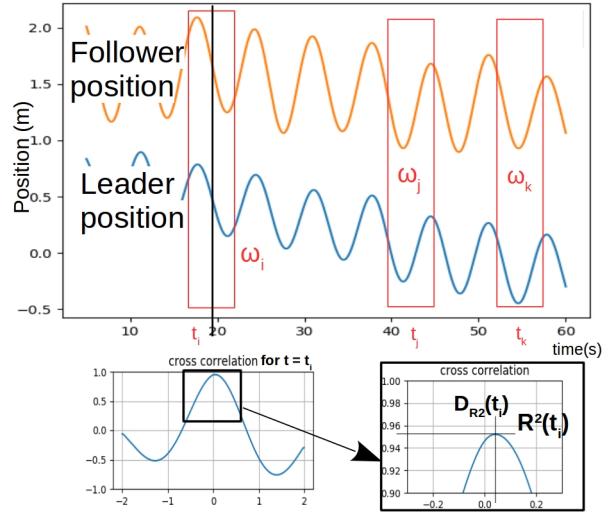


Figure 4: Computation of cross correlation and delay over time example trial: for a given time t_i , we define a time window ω_i centered on t_i . We perform a normalized cross correlation between the leader position and the follower position. The delay D_{R2} for t_i and the correlation R^2 for t_i are the ordinate and abscissa respectively of the point of maximum value of the function given by the normalized cross correlation. The process is repeated for each time step of the trial, which gives us a set of values displayed on Fig. 5.

	Robot velocity		Human velocity	
	Real	VR	Real	VR
T1	-0.006±0.012 -0.336; 0.349	0.008±0.034 -0.432; 0.400	-0.002±0.013 -0.336; 0.369	0.007±0.024 -0.472; 0.477
T2	0.006±0.086 -0.422; 0.373	0±0.059 -0.563; 0.496	0.001±0.045; -0.270; 0.363	0.005±0.052 -0.421; 0.360

Table 1: Mean and standard deviation (first line of a task), as well as amplitude (second line of a task), of the velocities of the robot and the human for each condition and each task.

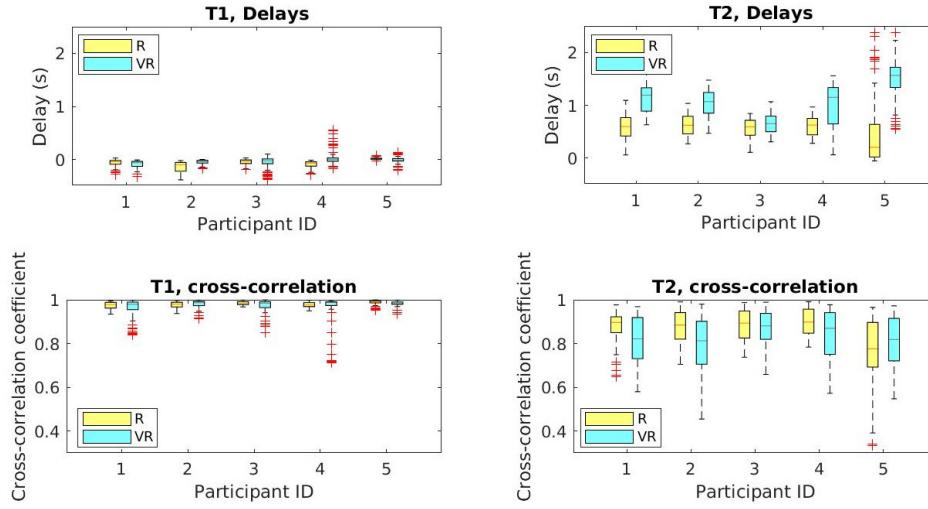


Figure 5: Box plots of delays and cross-correlations for each participant, for each task and each condition

4.5 Analysis

We here discuss the results shown in Sect. 4.4.

First we can notice a difference in speed from the robot between the real and the virtual conditions. The robot goes faster in VR. This is due to the error of its odometry sensors which is explained in Sect. 3.3.

Influences of virtual reality in the Human Robot interaction
Results show clear similarities between real and virtual conditions. There are however quantitative differences between real and virtual conditions.

First, the Wilcoxon signed rank test on the delays between the two conditions rejects the hypothesis of equality. Also, we can observe that the delays on Fig. 5 are slightly higher in VR than in real condition on task 2. We suppose that the system described in Fig. 2 has a network latency. Thus our hypothesis **H2** is verified.

The delays on task 1 are close to 0 for both conditions, which means that the participant is able to realize in time that the robot changed its velocity. Since the participants were able to hear the robot moving, we suppose that in VR they sometimes anticipated the movement of the robot, which explains the positive values for task 1 for participant 4. In [15], where the leader and follower are both humans, the author suggests that instead of anticipating, followers tend to be more reactive. In our case, when the robot was the follower, it was also reactive.

However, when the robot was the leader, it moved fairly slowly and with a predictable movement: the human could anticipate the robot movement efficiently instead of being reactive. Contrary to [15] who observes that velocities are higher when the leader walks backward and the follower forward, we do not observe this in our situation. We believe that the small maximal speed is in cause here, and it was easy for humans to follow the robot. However, it might be due to the way the robot is being controlled: for higher velocities, different controllers might have different influence on the forward and backward speeds of both the robot and the human, as leader or follower. Also, we notice that the average interpersonal distance during task 1 shows that the variable is smaller in virtual reality than in real condition. It is possible that humans experience a fear of collision in the real scenario that is mitigated in VR. Also, when the robot was leading, the security distance (see Sect. 3) between the robot and the human was never reached: the human seems to have a bigger personal space than our robot. Therefore, humans were almost never surprised by brutal stops, which reinforce the idea that human could anticipate the robot movements.

The Fig. 5 shows a high score of cross-correlation for each condition for task 1, and the cross-correlation data for task 1 passed the Wilcoxon signed rank test. This verifies our hypothesis **H1** for this task, but not the second task. We explain the lower cross-correlation scores for task 2 with the fact that the participants had different pace between each other, while the robot was always moving at the same speed in task 1 with each participant. Also, the participants were attentive to the robot's movements during the task 2: they were waiting for the robot because it was moving slowly, while the robot in task 1 was always in movement. We observe this behavior in both conditions, therefore this bias is not due to virtual reality.

Also, the participant speed for task 1 and 2 is zero centered, which differs from [15], and has higher maximum values in VR, which contradicts our hypothesis **H3** and [16]. However, the robot speed overestimated his speed in VR, so we cannot clearly assesses the bias here. For the second task, the small velocity of the robot (<0.35 m/s) constrained the participant speed in both conditions, as explained in the previous paragraph. We believe that the default walking speed is smaller in VR but in our experiment, participants do not even reach their default walking speed. The hypothesis **H3** is then wrong in this specific case.

Finally, in [15], the trajectories were correlated, thus subjects were totally involved in a bidirectional interaction with mutual influences. We observe this when the experience involves a robot: the human and the robot were imitating each other by simply watching. Even with such a simple controller for the robot, the whole interaction was preserved thanks to the human.

5 CONCLUSION

In this paper, we present a new VR platform designed to study interactions between robots and humans. Its main originality is to immerse both a human and a robot in a shared virtual environment. Both the robot sensing as well as the human visual perception are synthetic, but the robot and the human are able to do actions as in real conditions. This is a safe method to study human-robot interactions, and it outperforms the level of realism that can be reached in a pure simulation environment. In particular, the human response to the robot behavior would be difficult to predict based on only algorithms.

In this work, we are concerned by the perception biases introduced by the use of VR technologies. Their effect on the interactions in our pilot study, a simple coordination task performed between a human and a robot, are limited. Thus, it seems that VR could be used to

perform HRI experiments, while the quantitative results should be interpreted with care, since slight differences can be found compared to real experiments.

We believe that our platform fills the gap between simulation and experiments in real conditions. It offers more realistic conditions than simulation, at the cost of setting up more hardware and recruiting participants which takes time, which means less data than if we use simulation. Also, it offers more control, reproducibility of the experiments, and safety than real condition experiments, at the cost of using a complex set up which generate less realistic data than real condition experiments.

We believe that our platform offers new capabilities for VR experiments on HRI, allowing us to directly develop new methodologies to study HRI. We believe that it is absolutely necessary to provide a detailed description of our platform, so in this paper, we designed a simple pilot experiment to introduce the platform as we don't want important technical aspects to be eluded. Much more ambitious experiments are underway, which rely on the description in this paper.

It seems that latency, induced by the system described in Sect. 3, might be a major source of disturbance for the participant, and can cause cybersickness. The frequency of configurations updates in the VR world for both the human and the robot was 20.6Hz during the presented experiment.

Another limitation is that we used the robot odometry to estimate its motion during the experiment, which is subject to drifts and overestimation of velocity. To solve this issue, we would like to use a Kalman filter using both data from the *Vicon* tracker and the odometry to have a more precise tracking without any losses. Also, the system described in Sect. 3 is complex and expensive. There are some constraints, such as a limited area of tracking, limited real time capabilities. Each subsystem is subject to potential flaws, and has system requirements that make the whole platform difficult to maintain.

In addition, the robot velocity was limited to a maximum (0.35 m/s), which limited the scope of the results. Some participants reported that they heard the real robot moving around while they were in virtual reality and used this sound of the real robot moving in its own workspace, to synchronize their movement in VR. Some participants were also disturbed by their avatar, which did not perfectly fit their own body in terms of look. Moreover, the HMD we used has a limited field of view which impacts the participants immersion. Concerning the pilot experiment, the number of participants is low (5). However, it is enough to make a proof of concept and to identify ways of improvement. Finally, the tasks put velocities constraints on the participants, and lack of variation as the robot sinusoidal command is always the same during the whole experiment. This limits the variety of interactions, thus the scope of the results.

Future work

In the short-term, we must take into account the issues raised by this study to properly validate the platform. To do so, we must first deal with implementation issues. Preliminary optimization (post-experiment) showed that the frequency of data acquisition can reach at least 30Hz (20.6Hz during the experiments). Concerning the robot noise, the long-term solution will implement a second room equipped with a tracking system, synchronized with the tracking of the first room, in order to carry out the experiments in two separate rooms. For the short-term, the participant will be equipped with an active noise-canceling headset to cover the surrounding sound (or to relocate the sound source at the virtual robot position). We are also interested in improving the setup with bigger tracking room. We would like to enhance the participant perception thanks to a HMD with a bigger field of view. Also, we would like include new sensations to the perception, using wearable haptics. In the long term, we consider adding Xsens systems to enable the participation

of multiple users around the robot.

Our pilot study, despite its limitations, already highlight that the perception biases from the robot point of view are not quite of the same nature as the ones from the participant point of view. Once the technical and practical issues are solved, we will be able to conduct a series of experiments to evaluate remaining biases considering new interactions tasks. In particular, we should study the cases in which the robot is performing a motion using a combination of different sinusoids with different frequencies and phase shifts, as well as the case where the human performs a totally free movement, in order to compare more deeply with [15].

In addition, giving more freedom to the human will allow us to study the interaction with more metrics such as path efficiency, shoulder orientation, or jerkiness in the movement. Also, by design, the participant is safe from contacts with the robot and the platform allows to bypass robot security rules and to avoid classical issues in human-robot interactions such as the freezing robot problem.

Hence, we want to use this platform in HRI studies that present risks of collisions. For instance, we want to study human motion behavior before collisions with a robot, and we want to compare different navigation algorithms by reporting on virtual collisions. Moreover, we are highly interested in crossing scenarios (with or without collision-avoidance) and study risks of collision with this setup. Also, we would like to perform experiments with different robots, for instance with a smart power wheelchair. Also, this platform could be used to validate learning based algorithms which have been trained in simulation, and become the bridge between simulation and reality. Finally, generally speaking, we plan to improve the immersion provided by the platform, validate its usage in HRI study as it evolves, and use it in scenarios that present risks of collisions in order to study the human behavior in the presence of a robot, as well as evaluating navigation algorithms according to collisions related metrics.

ACKNOWLEDGEMENT

This work is funded by the Crowdbot Project, a European Union Horizon 2020 research and innovation project in the area of robotics, information and communication technology.

REFERENCES

- [1] P. Agethen, V. S. Sekar, F. Gaisbauer, T. Pfeiffer, M. Otto, and E. Rukzio. Behavior analysis of human locomotion in the real world and virtual reality for the manufacturing industry. *ACM Transactions on Applied Perception (TAP)*, 15(3):20, 2018.
- [2] G. Aravind, A. Darekar, J. Fung, and A. Lamontagne. Virtual reality-based navigation task to reveal obstacle avoidance performance in individuals with visuospatial neglect. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(2):179–188, 2015.
- [3] C. Armbrster, M. Wolter, T. Kuhlen, W. Spijkers, and B. Fimm. Depth Perception in Virtual Reality: Distance Estimations in Peri- and Extrapersonal Space. *CyberPsychology & Behavior*, 11(1):9–15, Feb. 2008. doi: 10.1089/cpb.2007.9935
- [4] J. Bruneau, A.-H. Olivier, and J. Pettré. Going through, going around: A study on individual avoidance of groups. *IEEE transactions on visualization and computer graphics*, 21(4):520–528, 2015.
- [5] M. A. Bühler and A. Lamontagne. Circumvention of pedestrians while walking in virtual and physical environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018.
- [6] W. Chi, H. Kono, Y. Tamura, A. Yamashita, H. Asama, and M. Q.-H. Meng. A human-friendly robot navigation algorithm using the risk-RRT approach. In 2016 IEEE International Conference on Real-time Computing and Robotics (RCAR), pp. 227–232. IEEE, Angkor Wat, Cambodia, June 2016. doi: 10.1109/RCAR.2016.7784030
- [7] Y.-h. Chou, R. C. Wagenaar, E. Saltzman, J. E. Giphart, D. Young, R. Davidsdottir, and A. Cronin-Golomb. Effects of optic flow speed and lateral flow asymmetry on locomotion in younger and older adults: a virtual reality study. *Journals of Gerontology: Series B*, 64(2):222–231, 2009.

- [8] G. Cirio, A.-H. Olivier, M. Marchal, and J. Pettré. Kinematic evaluation of virtual walking trajectories. *IEEE transactions on visualization and computer graphics*, 19(4):671–680, 2013.
- [9] A. Darekar, A. Lamontagne, and J. Fung. Dynamic clearance measure to evaluate locomotor and perceptuo-motor strategies used for obstacle circumvention in a virtual environment. *Human Movement Science*, 40:359–371, 2015.
- [10] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi. An atlas of physical human-robot interaction. *Mechanism and Machine Theory*, 43(3):253–270, Mar. 2008. doi: 10.1016/j.mechmachtheory.2007.03.003
- [11] L. Devigne, M. Babel, F. Nouviale, V. K. Narayanan, F. Pasteau, and P. Gallien. Design of an immersive simulator for assisted power wheelchair driving. In *2017 International Conference on Rehabilitation Robotics (ICORR)*, pp. 995–1000. IEEE, London, July 2017. doi: 10.1109/ICORR.2017.8009379
- [12] P. documentation. <http://doc.aldebaran.com/2-5/naoqi/sensors/dcm.html>.
- [13] Dongqing Shi, E. G. Collins Jr, B. Goldiez, A. Donate, Xiuwen Liu, and D. Dunlap. Human-aware robot motion planning with velocity constraints. In *2008 International Symposium on Collaborative Technologies and Systems*, pp. 490–497. IEEE, Irvine, CA, USA, May 2008. doi: 10.1109/CTS.2008.4543969
- [14] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun. CARLA: An Open Urban Driving Simulator. *arXiv:1711.03938 [cs]*, Nov. 2017. arXiv: 1711.03938.
- [15] T. Ducourant, S. Vieilledent, Y. Kerlirzin, and A. Berthoz. Timing and distance characteristics of interpersonal coordination during locomotion. *Neuroscience Letters*, 389:6–11, 2005. doi: 10.1016/j.neulet.2005.06.052
- [16] P. W. Fink, P. S. Foo, and W. H. Warren. Obstacle avoidance during walking in real and virtual environments. *ACM Transactions on Applied Perception*, 4(1):2–es, Jan. 2007. doi: 10.1145/1227134.1227136
- [17] T. Fraichard. Motion Safety with People: an Open Problem. *IEEE Int. Conf. on Robotics and Automation (ICRA)*, p. 3, 2015.
- [18] T. Fraichard and H. Asama. Inevitable collision states: a step towards safer robots? *Advanced Robotics*, 18(10):1001–1024, Jan. 2004. doi: 10.1163/1568553042674662
- [19] M. Gérin-Lajoie, C. L. Richards, J. Fung, and B. J. McFadyen. Characteristics of personal space during obstacle circumvention in physical and virtual environments. *Gait & posture*, 27(2):239–247, 2008.
- [20] F. Hernoux, E. Nyiri, and O. Gibaru. Virtual reality for improving safety and collaborative control of industrial robots. In *Proceedings of the 2015 Virtual Reality International Conference on ZZZ - VRIC '15*, pp. 1–6. ACM Press, Laval, France, 2015. doi: 10.1145/2806173.2806197
- [21] S. Ivaldi, V. Padois, and F. Nori. Tools for dynamics simulation of robots: a survey based on user feedback. *arXiv:1402.7050 [cs]*, Feb. 2014. arXiv: 1402.7050.
- [22] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems*, 61(12):1726–1743, Dec. 2013. doi: 10.1016/j.robot.2013.05.007
- [23] R. Li, M. van Almkerk, S. van Waveren, E. Carter, and I. Leite. Comparing human-robot proxemics between virtual reality and the real world. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 431–439. IEEE, 2019.
- [24] H. W. Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, 62(318):399–402, 1967.
- [25] J. Lin, X. Guo, J. Shao, C. Jiang, Y. Zhu, and S.-C. Zhu. A virtual reality platform for dynamic human-scene interaction. In *SIGGRAPH ASIA 2016 Virtual Reality Meets Physical Reality: Modelling and Simulating Virtual Humans and Environments*, SA '16, pp. 11:1–11:4. ACM, New York, NY, USA, 2016. doi: 10.1145/2992138.2992144
- [26] S. Lynch, J. Pettré, J. Bruneau, R. Kulpa, A. Cretual, and A.-H. Olivier. Effect of virtual human gaze behaviour during an orthogonal collision avoidance walking task. In *IEEE Virtual Reality*, 2018.
- [27] C. I. Mavrogiannis, A. M. Hutchinson, J. Macdonald, P. Alves-Oliveira, and R. A. Knepper. Effects of Distinct Robot Navigation Strategies on Human Behavior in a Crowded Environment. p. 10.
- [28] M. Moussaïd, M. Kapadia, T. Thrash, R. W. Sumner, M. Gross, D. Helbing, and C. Hölscher. Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of The Royal Society Interface*, 13(122):20160414, 2016.
- [29] L. Nummenmaa, J. Hyn, and J. K. Hietanen. I'll walk this way: Eyes reveal the direction of locomotion and make passersby look and go the other way. *Psychological Science*, 20(12):1454–1458, 2009.
- [30] A. H. Olivier, J. Bruneau, R. Kulpa, and J. Pettré. Walking with virtual people: Evaluation of locomotion interfaces in dynamic environments. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2017. doi: 10.1109/TVCG.2017.2714665
- [31] L. Palopoli, A. Argyros, J. Birchbauer, A. Colombo, D. Fontanelli, A. Legay, A. Garulli, A. Giannitrapani, D. Macii, F. Moro, P. Nazemzadeh, P. Padeleris, R. Passerone, G. Poier, D. Prattichizzo, T. Rizano, L. Rizzon, S. Scheggi, and S. Sedwards. Navigation assistance and guidance of older adults across complex public spaces: the DALI approach. *Intelligent Service Robotics*, 8(2):77–92, Apr. 2015. doi: 10.1007/s11370-015-0169-y
- [32] A. K. Pandey and R. Gelin. A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics Automation Magazine*, PP:1–1, 07 2018. doi: 10.1109/MRA.2018.2833157
- [33] R. Paulin, T. Fraichard, and P. Reignier. Human-Robot Motion: Taking Human Attention into Account. *IROS - IEEE*, p. 6, 2018.
- [34] L. Rebenitsch and C. Owen. Review on cybersickness in applications and visual displays. *Virtual Reality*, 20(2):101–125, June 2016. doi: 10.1007/s10055-016-0285-9
- [35] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You Only Look Once: Unified, Real-Time Object Detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788. IEEE, Las Vegas, NV, USA, June 2016. doi: 10.1109/CVPR.2016.91
- [36] M. Schepers, M. Giuberti, and G. Bellusci. Xsens mvn: Consistent tracking of human motion using inertial sensing. 03 2018. doi: 10.13140/RG.2.2.22099.07205
- [37] M. A. Sheik-Nainar and D. B. Kaber. The utility of a virtual reality locomotion interface for studying gait behavior. *Human factors*, 49(4):696–709, 2007.
- [38] M. J. Tarr and W. H. Warren. Virtual reality in behavioral neuroscience and beyond. *nature neuroscience*, 5:1089, 2002.
- [39] V. tracking system website. www.vicon.com.
- [40] P. Trautman, J. Ma, R. M. Murray, and A. Krause. Robot navigation in dense human crowds: Statistical models and experimental studies of human-robot cooperation. *The International Journal of Robotics Research*, 34(3):335–356, Mar. 2015. doi: 10.1177/0278364914557874
- [41] K. Varma, S. J. Guy, and V. Interrante. Assessing the relevance of eye gaze patterns during collision avoidance in virtual reality. 2017.
- [42] M. Vasic and A. Billard. Safety issues in human-robot interactions. In *2013 IEEE International Conference on Robotics and Automation*, pp. 197–204. IEEE, Karlsruhe, Germany, May 2013. doi: 10.1109/ICRA.2013.6630576
- [43] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric. Embodiment and Human-Robot Interaction: A Task-Based Perspective. In *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 872–877, Aug. 2007. doi: 10.1109/ROMAN.2007.4415207
- [44] W. H. Warren Jr, B. A. Kay, W. D. Zosh, A. P. Duchon, and S. Sahuc. Optic flow is used to control human walking. *Nature neuroscience*, 4(2):213, 2001.
- [45] F. H. website. www.getfove.com.
- [46] U. website. www.unity.com.
- [47] X. website. www.xsens.com.
- [48] J. Wirtz, P. G. Patterson, W. H. Kunz, T. Gruber, V. N. Lu, S. Paluch, and A. Martins. Brave new world: service robots in the frontline. *Journal of Service Management*, 29(5):907–931, Oct. 2018. doi: 10.1108/JOSM-04-2018-0119