Zheng Wang[1]
School of Mechanical and
Aerospace Engineering
Nanyang Technological University,
Singapore

Elias Giannopoulos
EVENT Lab
Facultat de Psicologia
Universitat de Barcelona
Spain

Mel Slater
EVENT Lab
Facultat de Psicologia
Universitat de Barcelona
Spain
and
Institució Catalana de Recerca
i Estudis Avançats (ICREA)
and
Department of Computer Science
University College London

Angelika Peer*
Martin Buss
Institute of Automatic Control
Engineering
Technische Universität München
D-80290 Munich,
Germany

Handshake: Realistic Human-Robot Interaction in Haptic Enhanced Virtual Reality

Abstract

This paper focuses on the development and evaluation of a haptic enhanced virtual reality system which allows a human user to make physical handshakes with a virtual partner through a haptic interface. Multimodal feedback signals are designed to generate the illusion that a handshake with a robotic arm is a handshake with another human. Advanced controllers of the haptic interface are developed to respond to user behaviors online. Techniques to achieve online behavior generation are presented, such as a hidden-Markov-model approach to human interaction strategy estimation. Human-robot handshake experiments were carried out to evaluate the performance of the system. Two different approaches to haptic rendering were compared in experiments: a controller in basic mode with an embedded curve in the robot that disregards the human partner, and an interactive robot controller for online behavior generation. The two approaches were compared with the ground truth of another human driving the robot via teleoperation instead of the controller implementing a virtual partner. In the evaluation results, the human approach is rated to be most human-like, with the interactive controller following closely behind, followed by the controller in basic mode. This paper mainly concentrates on discussing the development of the haptic rendering algorithm for the handshaking system, its integration with visual and haptic cues, and reports about the results of subjective evaluation experiments that were carried out.

1 Introduction

An important goal of haptic research in the context of immersive virtual environments is to introduce physicality into the virtual reality experience. The physicality of haptic interaction, among other factors, is a critical issue to be considered in rendering realistic interaction. For instance, the participant should be able to feel the weight and elasticity of virtual objects as well as to touch and interact with a virtual human partner in a real-life manner.

This paper presents a haptic rendering algorithm that aims at achieving a realistic handshake with a virtual human partner. The human participant grasps a haptic interface, representing the arm of a virtual human partner, and uses

*Correspondence to angelika.peer@tum.de.
[1]Formerly at the Institute of Automatic Control Engineering, Technische Universität München, D80290 Munich, Germany.

Presence, *Vol. 20, No. 4, August 2011*, 371–392

this to perform a handshake with a virtual agent. Advanced controllers for the haptic interface have been developed that implement a partner who adapts online to the human and gives him or her the illusion of taking part in a real human-human handshake.

Handshaking is a common daily activity. In order to simulate this activity through human-robot interaction (HRI), problems such as signal measurement, interaction control, and evaluation must be tackled and solved. Pollard and Zordan (2005) generated handshake animations from a vision system. Kunii and Hashimoto (1995) created the first telehandshake using a simple 1-DOF device, while the recent work of Haans and IJsselsteijn (2006), Bailenson, Yee, Brave, Merget, and Koslow (2007), Gunn, Hutchins, and Adcock (2005), and Tachi, Kawakami, Nii, Watanabe, and Minamizawa (2008) also followed the telepresence route. Haans and IJsselsteijn and Bailenson et al. viewed a telehandshake as a way to express and communicate social and emotional information, while Gunn et al. and Tachi et al. focused on more specific issues of telehandshaking such as time delay, system design, and integration. In Tachi et al., two humans conduct a handshake via a robotic system. In contrast, autonomous robotic handshaking has not been as closely studied until very recent work such as Sato, Hashimoto, and Tsukahara (2007), where the authors take the oscillation synchronization approach to realize human-robot handshaking; and in Y. Yamato, Jindai, and Watanabe (2008), where the authors focused on the approaching and shaking motions of a handshake robot; and in Karniel, Nisky, Avraham, Peles, and Levy-Tzedek (2010), where the authors evaluated different handshake models realized by means of a weighted sum of human and artificial systems. Apart from these very few investigations, the study of autonomous robotic handshake partners in a force/motion interaction context is lacking in the existing literature to the authors' knowledge.

New challenges arise when rendering interactions with a virtual interaction partner compared to interactions with passive environments: (i) the mechanical impedance of a human arm can vary over time, therefore we can no longer assume the mechanical properties of the environment to be constant and independent of the actual state of interaction; (ii) interaction forces between the single participants are measured by a single force/torque sensor mounted at the end-effector of the robot; this allows only the measurement of interaction forces; (iii) the human interaction strategy of how to carry on the interaction is not directly measurable, since the information is part of the internal mental processing of the participant. Bearing in mind these various challenges, two different approaches to design the robot controller can be considered: (a) design a basic controller that carries out the handshake as predefined, ignoring the bilateral nature of human-human interaction and forgetting about the human-related information that is difficult to acquire; or (b) design a complex controller that estimates the necessary human information and uses it to simulate a more interactive virtual partner. Both approaches were adopted for comparison.

In Section 2, two types of haptic rendering algorithms for human-human handshakes are introduced. Section 2.1 presents the basic handshake controller and Section 2.2 discusses the design of the more interactive controller and considers techniques for estimating human interaction strategies. Section 3 focuses on the subjective evaluation of the overall handshake scenario. Finally, Section 4 summarizes the main results and formulates future research directions.

2 Haptic Rendering

2.1 Basic Handshaking Controller

A basic approach to rendering handshakes is to replay prior recorded human-human handshake trajectories on a robot. The robot, controlled by a position controller, uses the recorded trajectory as a reference and follows it faithfully. However, due to the characteristics of the position controller, the robot will try to achieve the desired position at all costs, ignoring the bilateral nature and mutual influence of the single partners that is characteristic for a human-human handshake. This basic handshaking controller resembles an extreme case of a dominant person who does not compromise to his or her partner at all.

In order to increase naturalness of interaction, we introduced compliance by means of a second order impedance model as shown in Equation 1, where M, B, and K are the impedance parameters mass, damping, and stiffness representing the robot arm, f and x denote force and position, and x_0 is the equilibrium position of the virtual partner arm (Wang, Yuan, & Buss, 2008):

$$f(t) = M\ddot{x}(t) + B\dot{x}(t) + K(x(t) - x_0). \quad (1)$$

Constant values were chosen for the mass and damping parameters for the arm; but human arm stiffness was adapted by a changeable stiffness parameter to imitate human arm stiffness. In practice, the contraction of the muscle groups is increased when the human is exerting higher forces. The stiffness of the arm is therefore increased. Inspired by this fact, a time-varying virtual stiffness K consisting of a constant term K_0 in addition to a term proportional to the difference between the actual position $x(t)$ and the equilibrium position x_0, defined as the neutral position of the robot end-effector, was implemented:

$$K = K_0 + \eta(x(t) - x_0). \quad (2)$$

The intuitive explanation of this selection is that the more a participant wants to drive the partner away from the equilibrium, the stiffer the participant's arm should be in order to succeed. In other words, a robot with such a controller is compliant closer to the equilibrium and stiffer when further from the equilibrium. In the evaluation experiments, the robot controlled by this basic controller is used as one of the comparative conditions in haptics. In order to simulate a machine-like handshake, the desired trajectory was chosen to be a repetitive sinusoidal curve instead of one of the recorded trajectories.

The fundamental limitation for the realization of full interactive handshakes makes the basic controller approach clearly different from human-human handshaking, where the arms can provide compliance during interaction, while the participant can select different strategies with respect to adaptation to their partner's form of handshake. Therefore, an interactive handshaking controller has been developed that adapts the robot behavior online to the current selected interaction strategy of the human partner.

2.2 Interactive Hidden-Markov-Model-based Handshaking Controller

The interactive handshake controller, first presented in Wang, Peer, and Buss (2009b), is based on the assumption that humans select between two different strategies when performing handshakes with a partner: either they act passively by following and adapting their behavior as best as possible to the lead of the interaction partner; or they act actively by commanding the handshake trajectory without taking into account the behavior of their partner. Unfortunately, it is not feasible to directly measure human interaction strategies, but they can be estimated from measured behavioral (force and motion) data. Once the currently selected human interaction strategy is estimated, the interactive handshaking controller can be designed: In our specific case the respective opposite role is assigned to the robot. Depending on the personal style of handshaking, humans switch between the two aforementioned strategies while performing a handshake. This again means that the robot needs to continuously estimate the currently selected human interaction strategy to achieve realistic human-robot handshakes.

To implement this interactive handshaking controller on a robotic system, a robot controller consisting of a control, planning, and adaptation module was realized (see Figure 1). These three modules were proposed in Groten (2011) for modeling a partner in a collaborating human-human dyad. While the planning unit decides on the desired trajectory, the control unit implements compliant behavior. Finally, the adaptation unit uses information about the actual estimated human interaction strategy to decide on different adaptation laws that alter the reference trajectory as well as the provided compliance of the robot. Thus, a new double-layered control scheme is proposed consisting of a low-level controller (LLC) that combines a planning unit and a control unit; and a high-level controller (HLC), which is represented by the adaptation unit.

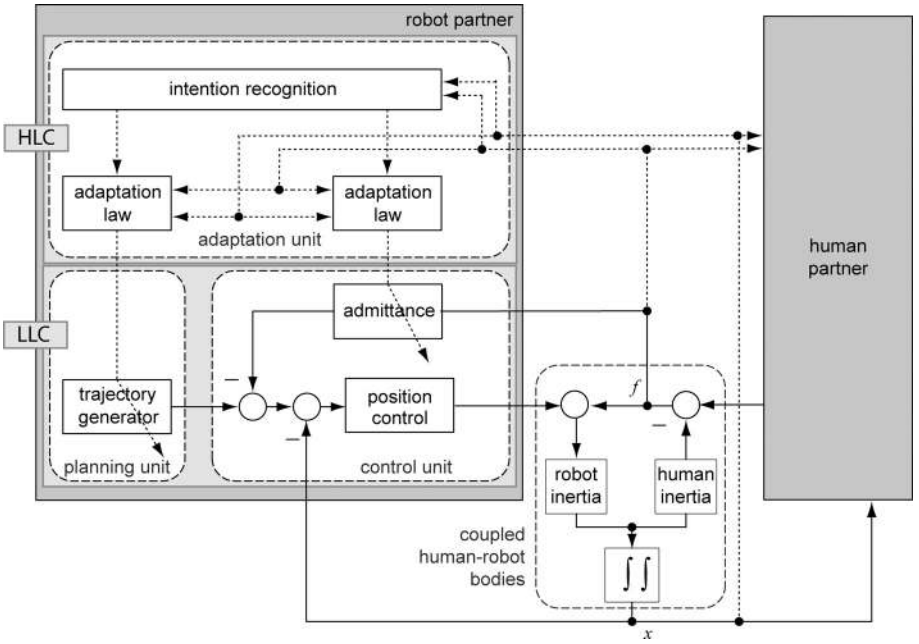


Figure 1. *Implementation of human-human handshakes using an advanced, interactive handshaking controller consisting of a low-level controller (LLC) that combines planning and control unit and a high-level controller (HLC) represented by the adaptation unit.*

2.2.1 Human Behavior Model. In order to adapt to human behavior online, information about the currently selected interaction strategy must be available. Unfortunately, interaction strategies are not directly measurable and thus need to be measured by indirect means. In order to obtain a valid estimation, the selection of input data is crucial. In haptic human-robot interaction, force and position data directly result from the interaction. Hence, any phenomenon observed in either signal can be the consequence of either the robot or the human. For this reason, motion or force signals alone cannot be adopted to estimate the currently selected human interaction strategy, but instead need to be further processed to remove the influence of the robot. In order to overcome this problem, we introduce a human behavior model and estimate parameters of this model from the observable haptic interaction data. The estimated parameters are finally used as input for the human intention estimator described below.

For the basic handshaking controller, the human was assumed to be passively following the robot, which means that the human was modeled by a passive impedance represented by mass, damping, and stiffness of the human arm without further excitation signals to the system. However, for an active human, the desired human trajectory becomes an additional input to the coupled human/robot system. Consequently, the old human model no longer applies and thus a new human behavior model that implements the three units proposed in Groten (2011) is assumed:

1. The human is modeled by a position-controlled arm with a trajectory planner and an adaptation module that adjusts the compliance of the arm as well as the reference trajectory according to the actual estimated interaction strategy and the currently observed haptic data.
2. The human behaves as a collaborative partner. In other words, the human planner adapts the reference trajectory based on the actual interaction status. It does not matter whether the decision of the human is to follow or to change the current trajectory; the decision is made based on actual measurements.

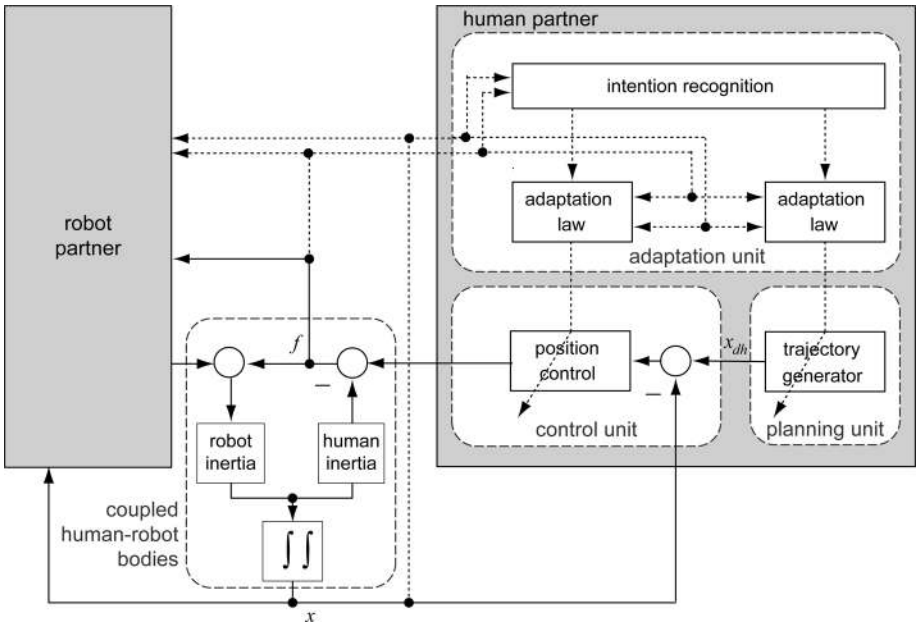


Figure 2. *Human behavior model.* $\boldsymbol{x_{dh}}$ *is the desired trajectory generated by the human,* $\boldsymbol{x}$ *is the actual human/robot position, and* $\boldsymbol{f}$ *is the actual interaction force. The parameters of human arm stiffness and damping represented by the position controller are variable in time.*

An illustration of the human behavior model is given in Figure 2. This model allows the human to change the arm impedance represented by the position controller as well as the desired trajectory with respect to time. The assumption that the human will behave reasonably is a natural approach, since it is expected that a human plans the behavior of the next step based on the information gathered about the current step. Further information needed to generate the trajectory is integrated inside the planner block. This makes the current position and force signals the only inputs to the human model, hence the input and output signals of the human block are all known to the robot.

The model in Figure 2 is expected to be time-varying and nonlinear. However, in practice the robot needs an easily identifiable model to estimate the currently selected interaction strategy. Hence, the following approach is taken to linearize the human behavior model:

1. A linear differential equation was used to represent the relationship between position input and force output signals.
2. The differential equations were limited to the second order, as shown in Equation 3, where f and x are the position input and force output signals, h_2, h_1, and h_0 are the three parameters of the differential equation, each denoted as a human behavior parameter (HBP), representing the current human behavior that determines the force output based on the current position input. Since the adaptation unit is part of the human behavior model, the HBPs are time-varying.

$$f = h_2(t)\ddot{x} + h_1(t)\dot{x} + h_0(t)x. \tag{3}$$

The HBP set (h_2, h_1, h_0) is similar to the impedance parameter set of a passive human in the sense that they are both relationships between force and position signals. However, HBPs also take into account the influence of the human planning and adaptation unit, and thus are not necessarily equivalent to impedance parameters. In the remainder of this paper, impedance parameters denote the parameters in the admittance filter of the robot controller, while HBPs denote the estimated behavior parameters of the human.

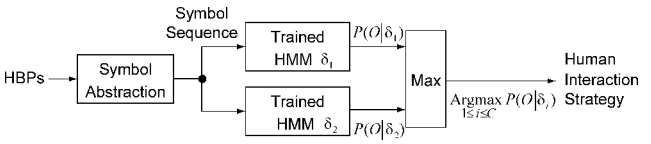


Figure 3. *Estimation of human interaction strategies using HMM-based estimator: HMM* δ_1 *stays for the active and HMM* δ_2 *for the passive state.*

A fast-converging forgetting factor Least Squares Estimation (ffLSE) algorithm is employed to estimate HBPs from measured force and motion data. In Wang, Peer, and Buss (2009a), this algorithm was used to estimate mechanical impedance parameters. Although HBPs are newly proposed in this paper, their estimation can be carried out in a manner similar to impedance parameters. Detailed discussion and validation experiments related to the ffLSE algorithm can be found in Wang et al.

2.2.2 Human Intention Estimation. As human intention is not directly measurable, a special estimator is needed, which is part of the HLC. In our case, we adopt pattern recognition algorithms which process the aforementioned HBPs and determine probabilities for the active or passive interaction strategy. A detailed scheme of the implemented human intention estimator, the interaction strategy estimator, is shown in Figure 3. Using measured force and motion data, an online parameter estimator identifies human behavior parameters, abstracts them into symbols, and feeds them into a discrete Hidden-Markov-Model (HMM)-based intention recognition module which outputs an estimate of the current human intention, that is, interaction strategy. Two HMMs are defined for the estimator, reflecting the two opposite roles of active and passive. Active indicates that the human is trying to lead the handshake, while passive means the human is trying to follow the lead of the robot.

An HMM has an underlying Markovian stochastic process that is not observable directly, but only through another stochastic process with respect to a certain probability of observation. This manages the relationship of hidden human interaction strategies and observable actions. While each human interaction strategy is modeled by an HMM, the hidden states; of the respective HMM encode the hidden mechanism that describes how the human generates the observed action. Therefore, for the estimation of n human mental states, there are n different HMMs, each having σ_n hidden states; and the values of σ_n are determined by the training of the specific HMM. The estimation algorithm then decides which HMM is currently the best fit for the given sensory information.

This type of HMM estimator has been used for speech, handwriting, and other types of pattern recognition since the 1980s (see Rabiner & Juang, 1986). Applications have been reported in human motion recognition by using image sequences, as in J. Yamato, Ohya, and Ishii (1992) and haptic signals, as in Takeda, Kosuge, and Hirata (2005). The application in Takeda et al. is to estimate human dancing steps using force signals measured by a force sensor mounted on the robot. In Avizzano (2007), HMMs are used to classify human behaviors based on haptic measurements when drawing on a 2D plan. In Calinon, Evrard, Gribovskaya, Billard, and Kheddar (2009), HMMs with continuous force and velocity inputs are employed in recognizing human interaction strategies in a joint object carrying task.

The method employed in this work is extended from the method described in Yang, Xu, and Chen (1997). The input data is changed from gesture paths to haptic data, in this case the HBPs. The number of hidden states in each HMM is set to four. The method can be formulated as follows:

Given $S = \{S_n\}, n = 1, 2, \ldots, N$, state S_n being the nth hidden state, and $O = O_1, O_2, \ldots, O_N$ the observed symbol sequence, choose the best matching HMM from δ_i, $i = 1, 2, \ldots, C$; that is, calculate $P(O|\delta_i)$ for each HMM δ_i and select δ_{c^*}, where

$$c^* = \arg\max(P(O|\delta_i)). \tag{4}$$

Given the observation sequence $O = O_1, O_2, \ldots, O_N$ and the HMM δ_i, the problem is how to evaluate $P(O|\delta_i)$, the probability that the observation sequence was generated by HMM δ_i. This probability can be calculated by using the forward-backward algorithm, as shown in Rabiner and Juang (1986). The HMM with the highest likelihood is selected as the recognition result.

Table 1. *Generated HMM Observable Symbols*

h_2	h_1	h_0	Symbol
Low	Low	Low	1
Low	Low	High	2
Low	High	Low	3
Low	High	High	4
High	Low	Low	5
High	Low	High	6
High	High	Low	7
High	High	High	8

HBPs are preferred to the interaction force signal as input to the HMM estimator, for the reason that the interaction force contains the influences of both the robot and the human. For instance, a high measured interaction force can result from a strong robot acting against human arm inertia in the case of a passive human, or from an active human who is trying to change the current style of shaking by either stiffening his or her arm or imposing a certain trajectory. In other words, high force measurements do not necessarily mean that it was the human who was overactive. Therefore, it is not sufficient to estimate human interaction strategies from the force measurement alone. The HBPs, however, consider the interaction trajectory and force signals at the same time. In case of high interaction forces resulting from a strong robot shaking the hand of a passive human, both position and force values would be high, resulting in moderate HBP estimations for h_0 and h_1; if the force results from an active human, the arm impedance or the imposed force would be high, while the position deviation would be small, which will be reflected in the HBP estimations as well. Hence HBPs are the better choice in representing human interaction strategies.

For a given HBP sequence, eight symbols are used in the abstraction as shown in Table 1. The thresholds for low and high are set by heuristics: 1 for the h_2 coefficient, 50 for h_1, and 500 for h_0.

Performance validation tests were carried out for the HMM estimator. Detailed discussions are given in Section 2.3.

Table 2. *Averaged Duration and Amplitude of Human-Human Handshakes*

Position	Min	Max	Mean
Duration (s)	0.712	1.816	1.01
Amplitude (m)	—	0.161	0.098

2.2.3 Trajectory and Impedance Parameter Adaptation. The trajectory and impedance parameter adaptation modules are the last modules in the HLC. The following information is available to generate the reference trajectory: force and motion (position, velocity, and acceleration) measurements, identified HBPs, and the estimated human interaction strategy. In addition, history information of the above data can be stored if necessary. The task of the trajectory generator is to generate a path on the basis of certain criteria which are to be defined in this section.

There are many possible solutions for the selection of criteria, since handshaking contains a large amount of intercultural, intersubjective, and even intertrial variations. Therefore, instead of defining one standard handshake, here the approach is taken to find boundaries for recreating a handshake trajectory that feels human-like to the user. In our case study of handshaking (see Wang et al., 2008), 1800 recordings were made on human-human handshaking. The averaged duration and amplitude of the position trajectories are shown in Table 2. The average handshake lasts about 1 s, with a peak amplitude of about 0.1 m.

The following strategy is used to implement active and passive robot behavior. When the human is in a passive state, the robot tries to take the lead. Impedance parameters in the admittance filter are set to high values, and no modification is made to the current reference trajectory; subsequently, the robot goes on as planned. When the human is active, the robot tries to follow the human's lead. The impedance parameters are set to low values, and the commanded trajectory is modified such that the robot synchronizes to the human motions, which requires a continuous adaptation of the reference trajectory.

A first trajectory adaptation algorithm was proposed and implemented in Wang et al. (2009b). In the current work, a refined trajectory adaptation strategy is proposed and implemented. The main issue of the trajectory adaptation algorithm in Wang et al. is dividing each handshake cycle into four segments and updating the planned trajectory only at the beginning of each segment. Therefore, no matter how much effort the human applied to change the current trajectory during the segment, the effect only appeared when the next segment started. This delay of a quarter of a cycle left an irresponsive impression on the participants in the experiment.

To fix this latency issue, updating the trajectory at each time instance is desired. The strategy of trajectory adaptation is not unique; selected strategies clearly reflect different personalities of a robotic partner during handshaking. Here a general frame of online trajectory adaptation is proposed, with more possible tunable parameters for future customization.

1. A handshake cycle is still divided into segments. Each starts at peak/valley points (A) and ends at equilibrium points (E), or starts at E and ends at A, as shown in Figure 4. However, this segmentation is no longer used for trajectory updating, but to keep a sinusoidal form of the overall shaking trajectory.
2. As suggested by the human-human handshaking results in Table 2, a typical handshake lasts for 1 s. Therefore, each segment is initialized to be of time length $D_1 = 250$ ms. A count t is introduced to note how much time passed since the beginning of the current segment, $0 < t \leq D_n$.
3. At each time instance t, D_n is updated according to the interaction force and current position. The current segment can be immediately terminated once a significant force opposite to the current direction of motion is detected, which suggests the human wants to lead. The current t is set to D_{n+1}.
4. At each time instance, the destination amplitude A_{n+1} or E_{n+1} of the current segment is updated by Equations 5 and 6 depending on the actual segment:

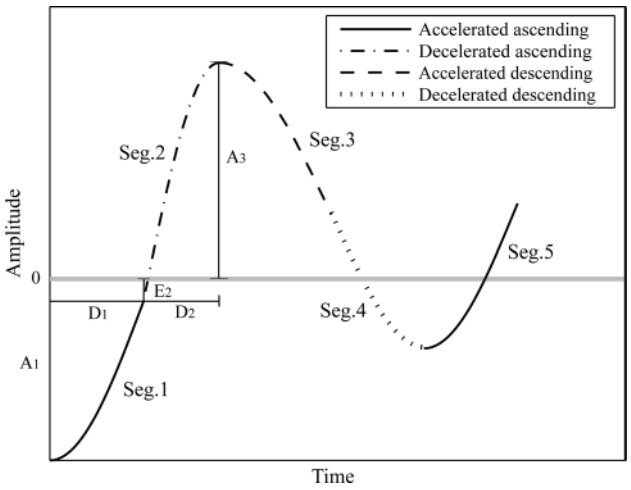


Figure 4. *Refined segmentation of the shaking trajectory.*

$$E_{n+1} = -dir \cdot \beta_1(\beta_2 + f_t)\sin\left(\frac{\pi}{2}\left(\frac{t}{D_n} + 3\right)\right) + x_t. \quad (5)$$

$$A_{n+1} = dir \cdot \beta_1(\beta_2 + f_t)\left(1 - \sin\frac{\pi t}{2D_n}\right) + x_t. \quad (6)$$

In both cases *dir* is the current direction of motion, $dir = 1$ for ascending, $dir = -1$ for descending. f_t and x_t are the current measured force and position, respectively. β_1 and β_2 are tunable parameters determining the amount of adaptation to the interaction force. The sine function is introduced to take into consideration the feasible amount of change applied to the desired amplitude: the closer the current t is to the end of a segment, the less the allowed modification term will be, in order to avoid exceeding the speed limitations of the robot.

5. With the updated D_n, A_{n+1}, and E_{n+1}, together with the measured current position, velocity, and acceleration, a fifth order polynomial can be calculated connecting the current and desired position smoothly (with continuous position and velocity).

The refined trajectory adaptation algorithm shares identical input and output signals with the original algorithm. The performance of the refined trajectory adaptation algorithm is superior to the one proposed

in Wang et al. (2009b), providing improved response speed and a wider range of adaptability. The result of the interactive controller in Section 2.3 was achieved with the refined trajectory adaptation algorithm. The results of the interactive controller using the original trajectory adaptation algorithm can be found in one of our earlier publications (see Wang et al.).

2.2.4 Extended Degrees of Operation. In early implementations, the arm motion for handshaking was modeled only along the vertical direction. Hence, when implemented in the robot, the other directions were fixed, with the human user only able to move the robot along the vertical translational DOF.

A human-human handshake is not subject to such a 1-DOF constraint. To remove it, and hence improve natural interaction, the system needs to be extended from 1-DOF to multiple 1 DOF.

The force/torque sensor mounted on the robot end-effector is capable of measuring 6 DOF information, namely three translational forces and three rotational torques. All calculations are made in Cartesian space. A virtual impedance model is introduced in each of the 5 DOF, namely two other translational DOF and three rotational DOF. The impedance models enable the user to move the robot end-effector along that direction while maintaining stability. The robot is passive along the other 5 DOF, with the reference trajectory generated by the virtual engine only applied to the vertical direction. Distinctive impedance parameters are set empirically for each DOF.

2.3 Objective Performance Evaluation and Results

Validation tests with a small number of participants were carried out prior to the subjective evaluation experiments to assess the performance of each module of the system.

2.3.1 HMM Estimator. It is not straightforward to define a controlled environment to collect training data for the HMMs, since the actual interaction strategy is practically hidden. In order to obtain training data where the human was either always active or always passive, the participants were instructed to act either uniformly actively or passively throughout a trial. Training data were obtained from four participants. They were instructed to either always take the lead or always follow the lead of the robot. One hundred groups of training data were recorded for each participant, that is, 50 for active and 50 for passive. The results of training using different data sets are shown in Figure 5. For 50 groups of training data (taken from the same participant), a sequence length of 100 ms achieved an accuracy rate of more than 90%. Further increasing the length of the symbol sequence would further delay the response of the system to human behavior changes, while the benefit of increasing estimation accuracy is marginal. Therefore, the combination of 50 groups of training data and a symbol sequence length of 100 ms is implemented in the robot controller.

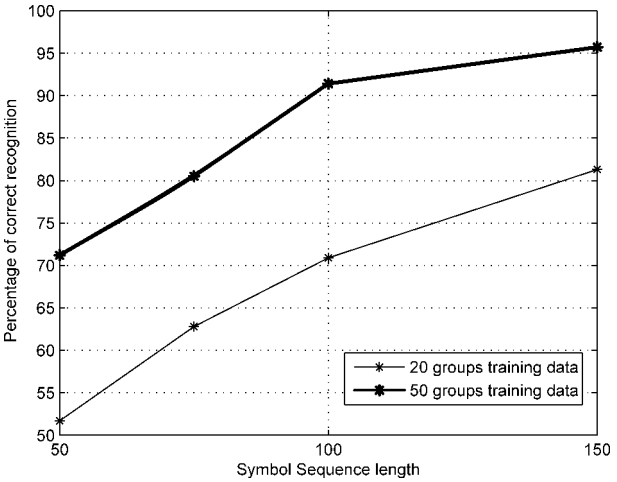


Figure 5. *HMM training results comparison for different symbol sequence lengths. A symbol sequence of length 100, corresponding to 100 ms in time, provides an accuracy rate of more than 90% for 50 groups of training data.*

The two trained HMMs for passive and active are implemented into the HLC which executes in real time (1 kHz). The input sequence length of 100 ms is employed.

Instructing participants about their handshake behavior is a necessary compromise, given that information about the actual interaction strategy is not available to the robot. However, naturally this will degrade the

confidence level of the training results, since accurate validation is not possible. Fortunately, the lack of information about the current human interaction strategy can be compensated for by employing a large amount of training data. First of all, each of the four participants was trained for a long session before the actual training data for HMMs was recorded. Second, in Figure 5 the accuracy results are defined as the percentage of time that the output of the interaction strategy estimator is correct. Training with 50 groups of data from the same participant provided the best result; where using the evaluation recordings made with the other three participants, for more than 90% of the time, the estimator gave correct interaction strategy estimations.

2.3.2 Overall System Performance. After testing the individual modules, the HLC was integrated. Human-robot experiments were carried out using both the basic controller and the interactive controller. The participants were different from those in the training group, and the results are shown in Figure 6 and Figure 7. For the basic controller, compliance was only provided by the virtual impedance model of the robot, while for the interactive controller the robot could synchronize to the human. Another observation can be made from the force signals, that for the basic controller, the human needs a relatively large force to drive the robot along the measured trajectory; while for the interactive controller, much lower forces are needed since the robot can detect the human trying to lead and hence adapts its reference trajectory to follow.

It is worth mentioning that the trajectories in Figure 6 and Figure 7 are significantly longer than the general human-human handshake length of around 1 s. They were recorded with naïve subjects handshaking with the robot. Note that a handshake with a robotic partner generally takes longer than with a real human partner. In Figure 7, the participant released the hand at around 5 s, while the trajectory generator was set to stop after 2 s of the last nonzero force measurement. Therefore, with the estimated human interaction strategy being passive in the last 2 s, the robot took over the lead and followed a rather standard trajectory as predefined in the trajectory generator, until it stopped at around 7 s.

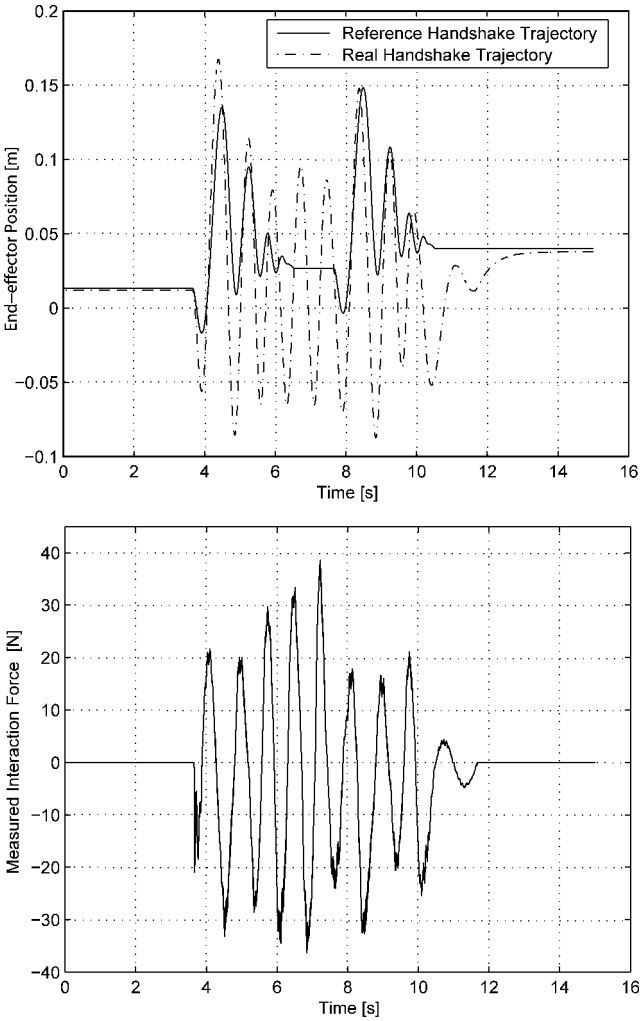


Figure 6. *Experimental results of human-robot handshaking using the basic controller. Upper: position; lower: force. The reference trajectory is not changed according to human input. The human keeps applying large forces to drive the robot.*

The 2 s waiting period before stopping a handshake is a practical measure, since the robot does not know when the human releases the hand. This issue can be easily solved by adding a sensorized robotic hand onto the robotic arm, which measures human grip force and hence obtains the information of when the human hand has been released. Practically, however, after releasing the hand, the participant will not feel any of the robot motion, hence this issue does not affect the experimental results.

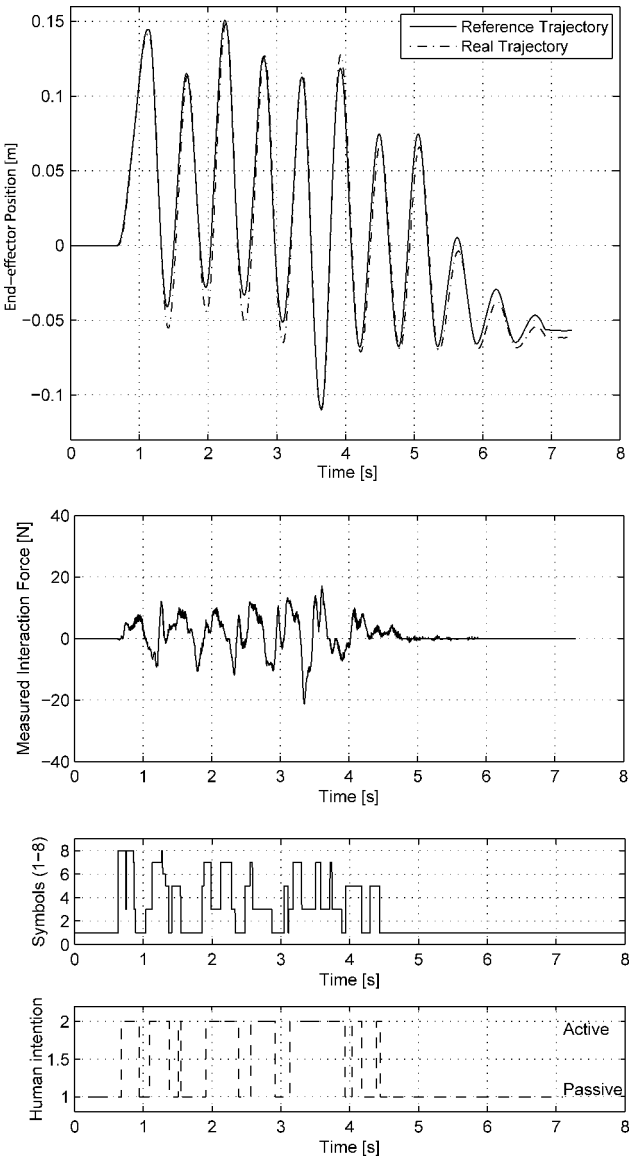


Figure 7. *Experimental results of human-robot handshaking using the interactive controller. Upper to lower: position, force, symbols, estimated human interaction strategy. Reference trajectory is modified according to the human input. Force decreases as the robot synchronizes with the human.*

Figure 8 shows the estimated results during the same handshake as in Figure 7. The first observation from Figure 8 is the occasionally negative HBP estimates. The stability of the system is not affected by the negative estimations, as they are not directly involved in the position control loop. The separation of large and small values,

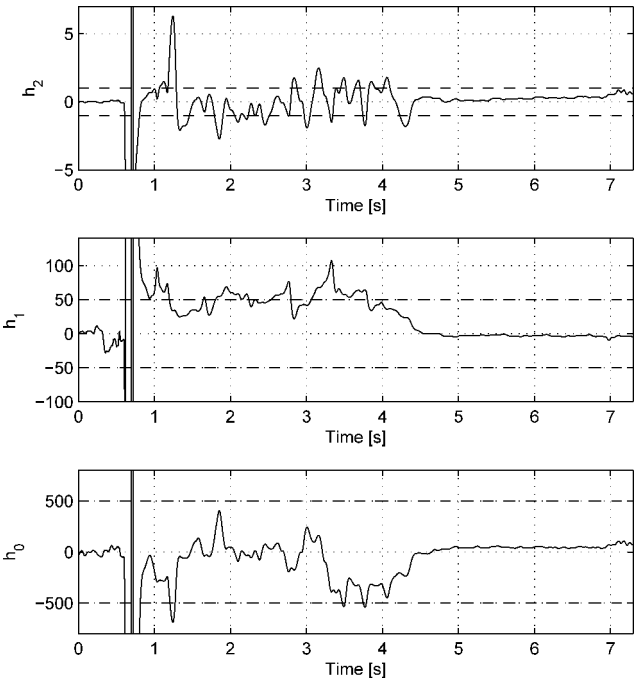


Figure 8. *Estimated HBPs of human-robot handshaking using the interactive controller.*

however, should take into account the possible negative values and use absolute values of each parameter. The thresholds are shown in Figure 8 as dashed lines.

The results shown in Figure 7 and Figure 8 support the notion that HBPs are more suitable for human interaction strategy estimation than the measured force signal, since force measurements contain information exerted by both the human and the robot. Between 1.6–1.9 s in Figure 7, for example, the measured force spike is applied by the robot driving the human, while the human arm is actually loose and acting only as additional inertia as can be observed in Figure 8. Therefore, despite the high force measurement, the estimated interaction strategy is passive.

On the other hand, the estimation results are also affected by the threshold values in HBPs. If in the aforementioned example the threshold of h_1 were set to 20 instead of 50, the resulting symbol between 1.6 and 1.9 s would have been 3 instead of 0, and the resulting human interaction strategy would have been different. Varying the threshold in HBPs can be interpreted to some extent as changing the personality of the virtual

robot partner, where higher thresholds suggest a more dominant role, while lower values result in the robot seeming more gentle and adaptive.

3 Subjective Evaluation with Human Participants

In order to evaluate the overall handshake system, a subjective evaluation experiment was designed where the ground truth approach was employed. This approach compared the actions of a person within a multisensory virtual environment to the actual behavior of people in a similar situation in physical reality. In order to achieve that, an experiment was designed where subjects performed handshakes with the robotic system, as well as with a human, and their evaluation scores for the handshakes performed provided the basis for the evaluation of this robotic system. This follows the strategy for the measurement of presence in virtual environments suggested in Sanchez-Vives and Slater (2005), which views presence as the extent to which responses in a virtual environment are similar to what would be observed or expected in a similar real environment. In the present experiment we directly compare the responses of people to the robot handshake with their responses to the human handshake.

3.1 Experimental Design

3.1.1 Scenario. In order to simulate a scenario where several handshakes would be performed, we chose a cocktail party, as in Giannopoulos, Wang, Peer, Buss, and Slater (2010). Here we additionally introduced a visual representation of the cocktail party along with virtual characters. The participant was immersed in this cocktail party as a virtual character standing behind a bar and interacting with the other virtual characters. To make the environment and the scenario more realistic, a real table was placed in the laboratory in a 1:1 mapping with the virtual bar seen on the virtual environment, so that whenever the subject touched the virtual bar, which he or she would be standing behind, there would be a corresponding haptic feedback from the real table. The sound recordings used were made from a real cocktail party. The virtual environment, with the visual realism of lighting and textures, the realistic sound, as well as the haptic consistency of the virtual bar, created an immersive experience to the participant from the onset of the experiment.

Once the participant became immersed within the virtual environment, a virtual character entered the room and walked toward the participant. The virtual character stopped in front of the participant and greeted him or her with a simple sentence and reached out his or her hand to receive a handshake, which was expected in the greeting context. When the participant reached out for the virtual hand, he or she found the robot hand in the real world which was mapped 1:1 to the location of the virtual character's hand in the virtual environment. Furthermore, the participant's hand was tracked by the data glove and he or she could fully control the position of his or her own virtual hand. After completing a handshake, the virtual character turned around and left the room. This procedure was repeated with 18 distinct virtual characters, which once completed, concluded the experiment.

There were two experimental factors. The first was the main experimental condition of interest, which we call robot type, whereas the second factor was the gender of the avatar that shook hands with the subject. There were three robot types: basic robot, interactive robot, and human driven. In the basic robot condition, the robot was controlled by a position controller with a fixed sinusoidal reference trajectory, as described in Section 2.1. In the interactive robot condition, the controller with the updated trajectory generator was employed, as introduced in Section 2.2; while in the human-driven condition, the second robotic arm was employed by an experimenter to deliver a handshake to the participant via teleoperation. No comparative conditions were intentionally introduced in vision and sound during the actual experiment.

3.2 Response Variables

3.2.1 Questionnaires. Prior to the experiment, the subjects completed a preexperiment questionnaire

providing demographic information about sex, age, prior experience on virtual reality, level of knowledge of computing, and whether they had experienced any similar experiments before (as the participants were students from Technische Universität München, there was a possibility that they had participated in previous studies involving the robot performing the handshakes).

At the end of the experimental session, the subjects were required to complete a postexperiment questionnaire that assessed various aspects of presence. The complete set of questions is given in the Appendix.

3.2.2 Handshake Score (HS). The main response variable studied for the evaluation was provided by the subjects throughout the experiment. Namely, immediately after performing each handshake, the subjects were required to call out a score between 1 and 10 representing the degree of belief that the handshake had been exchanged with a human. A value of 1 meant that it had definitely not been with a human, and 10 definitely with a human.

3.2.3 Number of Correct Classifications (NCC). In order to examine the impact of factors such as age, gender of the subjects, and their relationship with the presence scores, a new response variable was constructed: the number of correct classifications (NCC) out of the 18 choices made by the participants. This classification was defined as follows: for a given response by the participant, a variable x was classified as *true* if:

- The response was between 1 and 3 for the basic robot.
- The response was between 4 and 7 for the interactive robot.
- The response was between 8 and 10 for the human-driven robot.

The total number of *true x*s for every participant defined the NCC of that participant.

3.2.4 Hypotheses. Our hypothesis was that on average, HS would be greater for the interactive robot than for the basic robot. A very good result would be that on the average, HS would not be significantly different between the interactive robot and the human-driven condition. The overall hypothesis was therefore

$$\text{Mean(HS(human-driven))}$$

$$= \text{Mean(HS(interactive robot))} > \text{Mean(HS(basic robot))}.$$

Regarding the effect of gender of the avatar, this was an exploratory issue to examine whether there was any impact of avatar gender on the response, as well as any possible interaction between gender and robot type.

Figure 9. *Experimental setup.*

3.3 Materials and Methods

3.3.1 Equipment. For the purposes of this experiment, two ViSHaRD10 robotic arms were used: one robotic arm was used as a handshaking device, while the other was used as the haptic interface in order to enable the participants to shake hands remotely with another human. Two identical 6-DOF force/torque sensors integrated within each of the robotic arms were used to measure the interaction forces exerted between the users. Similar to the objective performance evaluation experiments, the robotic arm used by the participant was mounted on a platform and a dummy hand made of rubber was mounted on this robotic arm (see Figure 9).

The participants wore an nVISOR SX 60 stereo head-mounted display (HMD; NVIS, 2011) for the display of the visual scenario, and a pair of sound-isolating headphones for the sound inside the virtual world, which was

also used for the purpose of isolating the motor noise of the robot.

In order to map the body and head motions into the virtual system, Intersense IS 900 motion trackers (Intersense, 2011) were placed on the head, back, and wrist of the participant. An Immersion data glove (CyberGlove Systems, 2011) was used to measure the participant's hand movement in order for it to be reflected by his or her avatar.

In order to validate the performance of the developed robot control algorithms, a human handshake partner was used as a comparative condition during the evaluation experiment. Although it is possible to use a specially designed end-effector that two humans can hold, as shown in Wang, Lu, Peer, and Buss (2010), here a different approach was used: teleoperation. When using the two-sided end-effector as formerly proposed in Wang et al., each handshake partner held on to one end of the end-effector, but only one force sensor was used, providing only one combined force measurement for both sides. As a consequence, it was not possible to distinguish between the forces applied by each side. To achieve separate force measurements, in this work, the two humans were separated by using a teleoperation setup. This setup consisted of two robotic arms, each equipped with one force sensor mounted at the end-effector, such that the force applied by either side could be accurately measured. In addition, a universal virtual impedance was employed to govern the force/position relationship at both sides. Force measurements were exchanged bilaterally to calculate the corresponding desired position for each robot (see Peer & Buss, 2008).

3.3.2 Visual Rendering. In order to provide multimodal feedback during handshaking, a virtual cocktail party scenario was created (see Figure 10), which shows the view of the participant into the virtual world. An extended visual scene was developed based on the prototype reported in Wang et al. (2010). In comparison with recently reported similar systems (Giachritsis, Barrio, Ferre, Wing, & Ortego, 2009; Spanlang, Fröhlich, Fernandez, Antley, & Slater, 2007), the new visual scenario generates full body human animations online with

Figure 10. *Participant view of the virtual system in the experiment.*

Figure 11. *Comparison of virtual character decorations: (a) before decoration, (b) after decoration.*

the input data from the haptic subsystem as well as from the motion trackers placed on the human participant. At the same time, high-fidelity facial details are maintained to provide realistic interaction experiences for the participants.

Building on the visual rendering algorithm developed in Wang et al. (2010), facial details of the virtual characters were significantly improved by importing external textures to the face model to restore lost information due to importing the character from development software to real-time rendering software. Figure 11 shows a comparison between the initial virtual character and the one after decoration. Significant improvements can be observed for the eye, eyebrow, lips, hair, and so on.

A high fidelity virtual hand model of 24 DOF (see Cobos, Ferre, Ortego, & Sanchez-Uran, 2008) was used to represent the hands of virtual characters. The hand model is capable of real-time hand animation with the

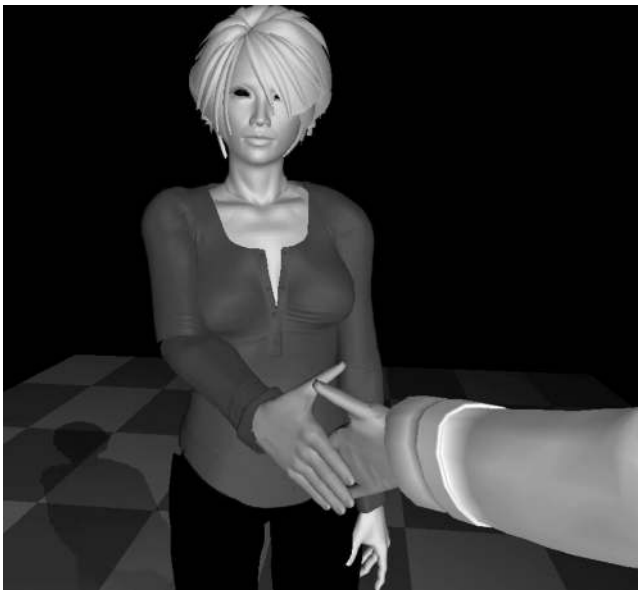


Figure 12. *The first virtual environment used in training sessions.*

input data measured from a data glove worn by the participant. The location and orientation of the hand in the virtual world was registered with the location and orientation of the robot end-effector in the real world. An unactuated rubber hand was mounted on the robot arm, such that after fine calibration, the participant could reach out and grab this hand at the same location of the virtual hand as in the virtual world.

Eighteen virtual characters were created with different clothing and facial details. They were all used in the evaluation experiment as virtual handshake partners. An additional virtual character was created as the avatar or virtual representation of the human participant. Since the avatar's head was always behind the virtual camera, it never appeared in the view. In order to minimize computational load, the avatar's head was not designed.

Two virtual environments were designed, one simplified for use during training sessions, the other for the main experiment. The first environment was kept simple, consisting of an empty space with a floor with black and white mosaic bricks, with one fixed virtual character together with the virtual representation of the participant. A single spotlight was used to light the virtual world, as shown in Figure 12. This environment was used during training sessions. The aim was to familiarize

Figure 13. *The rendered result of the second virtual environment.*

the participant with the virtual reality without experiencing the main experimental environment. Therefore, an empty space was provided instead of the barroom scenario, with a fully functional virtual character as the handshake partner who happened to be in a fixed location.

The second virtual environment was much more complex than the training environment; in this case, the virtual world represented a cocktail bar. A number of virtual characters were programmed to approach the human participant in order to shake hands, while the actual robot location was fixed in the physical world, meaning the participant was required to stand in the same place the virtual characters came over to greet him or her and then leave again. Figure 13 shows the design of the virtual bar environment, consisting of the following elements:

1. There was a virtual room equal to the physical room in which the virtual reality was embedded, and which was designed to be a bar.
2. Two doors were placed on each side of the room, so that characters could enter and leave the room.
3. In the real world, the participant stood on a 1×1 m platform. A table was placed in front of the participant such that the participant could reach the robot over the table while preventing an accidental walking away from the platform.
4. In accordance with the table in the real world, a virtual table of the same height was placed in the virtual world. The participant, similar to a bartender, was hence constrained by this table, being

able to see and shake the hand, but not to actually walk into the virtual world.
5. To address the ambient noise of the bar, a number of additional virtual characters were placed around the room. However, to minimize processing time, these characters did not move.
6. The lighting condition in the room was set to be dim, with rotating lights of time-varying colors.

3.3.3 Sound Rendering. In order to display plausible sounds that are typical for a cocktail bar, a sound rendering algorithm was developed, which was able to replay background sound as well as the real-time conversations of virtual characters.

For sound rendering, one background sound clip was prerecorded with a number of virtual character dialogue clips as add-ons to it, which were mixed in online and were triggered by events. During the experiment, the background noise started to play first. When the robot was ready, the experimenter controlling the events hit one key that triggered the playback of an animation that moves the character's mouth and the playback of a corresponding dialogue at the same time. In order to fit mouth motion with the dialogue, synchronization of auditory and visual cues was required.

In order to achieve mouth synchronization in the virtual characters, Voice-O-Matic (2011) was used to generate mouth motion information out of voice clips. Once this information was available, it was mapped to key frames in visual rendering, while keeping the timing information. For each dialogue clip, such a sequence was created off-line. An audio library FMOD (2011) then read in this information and switched key frames to generate the talking animation that was synchronized with the voice.

3.4 Experimental Procedure

Twenty-one participants were recruited (10 male, 11 female). They were full-time students from Technische Universität München. Each of them signed an informed consent for their participation to the experiment. Each participant took around 30 min to finish the experiment, and was given €5 as compensation.

The experiment consisted of three parts: preexperiment briefing, the training session, and the main experiment. The participants were not allowed to see the robot throughout the experiment, avoiding the appearance of the robot which could interfere with the immersion into the VR environment. A trained experimenter guided the participants through the entire procedure.

1. **Preexperiment Briefing.** Once the participant arrived at the site, the trained experimenter introduced the experiment and handed all the necessary documentation to be filled out (consent form, preexperiment questionnaire) outside the experiment room, thus preventing the participant from viewing the robot prior to the conclusion of the experiment. Upon completion of the documentation, the participant was blindfolded and guided into the experiment room again to prevent visual contact with the robot. The participant was then requested to put on the necessary devices (HMD along with backpack and data glove) and run a set of required calibrations.
2. **The Training Session.** The participant performed six handshakes of the basic condition. At this stage, the participant saw a female virtual character in front of him or her with her hand protruding as if awaiting a handshake. The participant was then asked to reach for the hand of the virtual character and perform a handshake. At this point, the experimenter explained to the participant about the rating procedure, that after each handshake a score from 1 to 10 needed to be provided verbally by the participant, about how close the handshake felt to shaking hands with a real person, where a score of 1 meant that the handshake was definitely not performed by a human, while a score of 10 that the handshake definitely was performed by a human. The participant then continued with the training session for another six handshakes, this time giving a score after each one. Throughout this training session, the handshakes were with the basic robot and the interactive robot, presented in a random order.

3. **The Main Experiment.** This session consisted of 18 handshakes, matching the number of different avatars that had been created. There was an arbitrarily determined sequence of male and female avatars, but it was the same sequence used for all subjects. Specifically it was:

   M-F-M-F-F-M-F-F-M-F-M-M-F-F-M-M-F-M

   At this stage, the third condition was introduced, namely the human-driven condition, where the handshake was performed between the participant and an experimenter in a remote location via teleoperation of the robotic arm. Each of the three conditions in comparison were repeated six times in random order. The visual environment presented to the participant at this stage was the cocktail barroom, which had been unknown to the participant up to that moment. Therefore, initially the participant was given some time to look around and get accustomed to the environment as well as his or her virtual representation and get familiar with moving his or her virtual hand. A prerecorded cocktail party background sound, with people chatting and background music, was played back. At intervals of 15 s a virtual character entered the room, walked toward the participant, performed a handshake, and then left the room. After each handshake, the participant was supposed to rate the handshake performed (as practiced throughout the training session), according to their experience. No further interaction between the expert and the participant was expected, as the participant was fully instructed by the virtual events occurring in the virtual world.
4. **Postexperiment.** After the experiment, the participant was led out of the experiment room to fill out the postexperiment questionnaire and provide further comments.

3.5 Statistical Analysis

This was a two-factor within-groups design. In order to account for the within-group effect, this was analyzed as a 3-way ANOVA (Robot Type, Gender, and Subject) with interactions.

Table 3. *Mean and Standard Errors of Handshake Scores for Each Robot Type*

Robot type	Mean	*SE*
Basic robot	3.0	0.17
Interactive robot	5.3	0.17
Human-driven	6.8	0.21

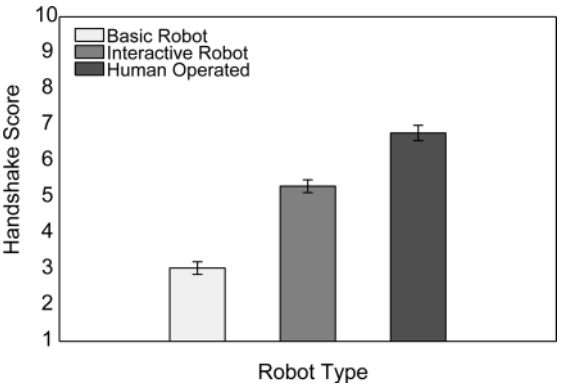


Figure 14. *Mean and standard errors of the handshake scores by robot type.*

NCC can be considered as a binomial random variable with $n = 18$, and hence logistic regression can be carried out on the explanatory variables such as gender, age, game playing, and so on.

Additionally we can consider Classification as Human (CH), which is a score of how often they felt they were shaking hands with a real human. This is defined as the number of times out of 18 they gave a score of ≥ 8. Correlations of CH with the presence-related scores from the questionnaires could be used as a consistency check, since we should expect, for example, that the more often the participants had felt that they were shaking hands with a real person, the greater should be the CH.

3.6 Results

3.6.1 Handshake Scores. The means and standard errors of the 126 observations for each robot type are shown in Table 3 and Figure 14. The 3-way ANOVA showed a highly significant difference between the means of Robot Type ($p = .0000$), but not for

Gender ($p = .68$). Moreover there was no significant interaction between Robot Type and Gender ($p = .54$) and the 3-way interaction term was also not significant ($p = .09$). However, the Jarque-Bera test rejected the assumption of normally distributed errors ($p = .004$). Investigation of the residual errors showed that there was one outlier, and when this was removed the ANOVA satisfied the assumption of normality with Jarque-Bera ($p = .16$). Qualitatively, the results do not change with this outlier removed. The difference between the main effects of Robot Type remain highly significant ($p = .0000$), there is no difference between the Gender types ($p = .97$), no interaction effect ($p = .32$) and the full 3-way interaction is not significant ($p = .13$).

A multiple comparisons test of the differences between the main effects shows that

$$\text{Mean(RT(basic robot))} < \text{Mean(RT(interactive robot))} < \text{Mean(RT(human-driven))},$$

even at an overall level of significance $P < 1.0 \times 10^{-15}$.

Hence, although the evidence is extremely strong that the interactive robot was considered more like a real handshake than the basic robot, it is equally strong that it could be reliably distinguished from the real human handshake.

Table 4 shows the handshake score frequency distributions across the three types of interface. It is clear that although on the whole each of the three types of interface were correctly recognized, there were classification errors. For example, the human operator was scored as a robot (handshake score ≤ 2) 11 out of 126 times, and there was even one classification of the basic robot as human (a score of 9).

Figure 15 provides a view of the distribution that is easier to understand, where the scores have been grouped into successive pairs. The basic robot distribution is approximately a reverse J-shaped distribution with the mode at scores (1,2), and the human is almost J-shaped with the mode at (7,8) which has a frequency just slightly higher than that for (9,10). However, the interactive robot distribution is more symmetrical about the modal scores of (5,6), with some skew toward the higher scores of (7,8). Using Kolmogorov-Smirnov tests, the three distributions are highly significantly different with the basic different from the interactive ($p < 2.05 \times 10^{-12}$) and the interactive different from the human ($p < 4.0 \times 10^{-6}$).

Table 4. *Frequency Table of Handshake Score by Type of Robot*

Handshake score	Basic robot	Interactive robot	Human-driven	Total
1	39	3	4	46
2	23	12	7	42
3	14	11	4	29
4	20	12	6	38
5	17	30	14	61
6	6	19	12	37
7	3	21	22	46
8	3	13	19	35
9	1	5	28	34
10	0	0	10	10
Total	126	126	126	378

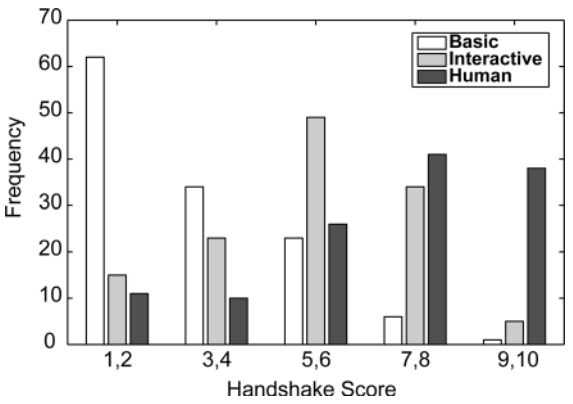


Figure 15. *Frequency distribution of the handshake scores grouped in successive pairs.*

3.6.2 Individual Differences. Next, we consider whether the results could have been influenced by the characteristics of the participant—their age, gender, previous experience with computers, with virtual reality in general, their computer game playing,

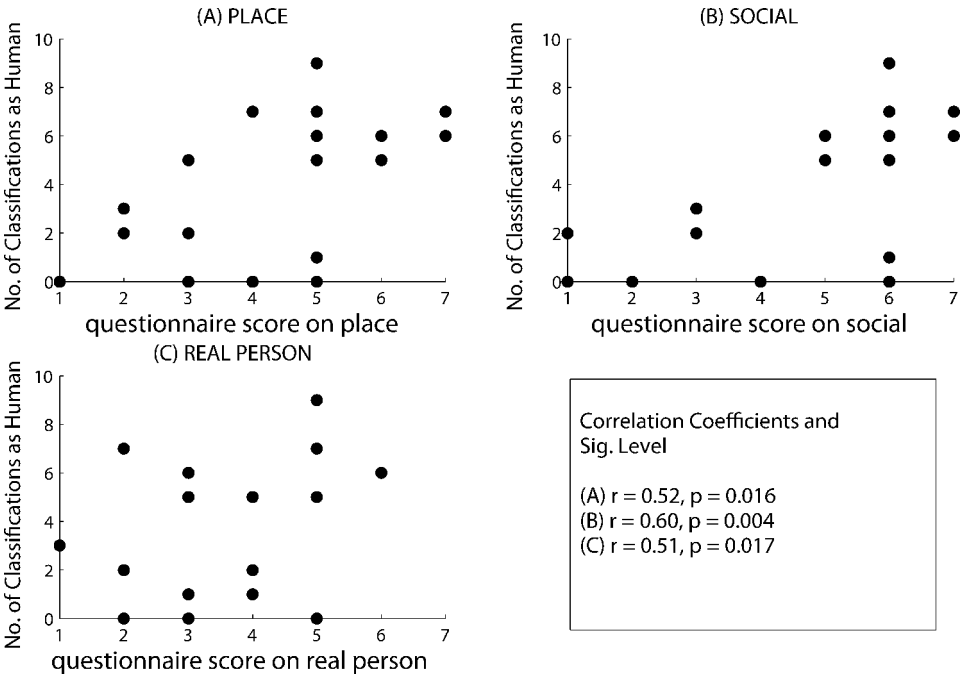


Figure 16. *Scatter diagrams for CH on three presence-related questions.*

and whether they had either taken part in any previous experiment, or seen the robot before. For this we use the constructed variable NCC. The mean of NCC is 10.2 with *SD* 3.3. Logistic regression of NCC on these possible explanatory variables derived from the preexperiment questionnaire found no significant associations at all.

3.6.3 Consistency Checks. Finally for consistency checks we use the variable CH and examine how well this correlated with various presence-related questions in the postexperiment questionnaire, in particular how much they felt to be in a place (place), how much they felt that they were socializing with real people (social), and how much they felt they were shaking hands with a real person (real person).

Figure 16 shows the scatter diagrams and gives the corresponding Pearson correlation coefficients for CH on the scores for these three questions. It is clear that generally people gave handshake scores that were consistent with their later subjective evaluations in the sense that the more they classified the handshake as human the greater the likelihood that they would later say that they felt that they had been in a place that was a social situation where they were interacting with real people.

3.7 Discussion

The experiment demonstrates a fundamental advantage of the interactive robot over the basic robot, in that participants were more likely to rate it higher on the scale toward being like a human handshake. However, it is also the case that there was a great discrepancy between the interactive robot and the actual human handshake on this same score.

A possible cause can be found from the comments given by participants after the experiment, where many mentioned that the robot controller at times failed to stop a handshake at the right time. Currently, due the fact that there is no hand with movements that are directly robot controlled, the robot controller can only decide when to stop a handshake by the measured arm interaction force, as discussed in Section 2.3.2. In other words, for the current setup, the interactive robot could only stop handshaking after the human partici-

pant stopped first. This sometimes resulted in lingering handshakes and hence lower presence scores for the interactive robot.

4 Conclusions and Future Work

We have described an interactive controller designed to adapt the reference trajectory of the low-level controller for human-robot handshaking according to the actual interaction. HMMs with haptic inputs were employed for human interaction strategy estimation. A fast online parameter identification method was used to provide estimations of human behavior parameters. An HMM-based estimator then estimated the current human interaction strategy according to the identified human behavior parameters.

Multimodal virtual scenarios were created and integrated with haptics. Human-robot experiments were carried out to evaluate the performance of the overall system. Finally, the robot operating in interactive mode was scored much more human-like than the robot in basic mode.

The interactive controller can be improved and generalized in several ways. The system can be extended to generate a variety of other types of handshakes, since active and passive do not necessarily cover all human interaction strategies. The interactive controller was designed with a high potential for generalization, and the number of HMMs and the number of states within each HMM can both be modified to meet the need of specific tasks; continuous HMMs could be studied when considering HBPs as HMM inputs. HMMs could also be considered for motion generation. Currently, the participant grips an unactuated rubber hand, which could be replaced by an actuated robotic hand in the future, which would provide the participant with improved overall experience. The fields of application could be generalized, for instance to rehabilitation, where robotic assistance can be provided with the functionality of adapting to human interaction strategies, as well as sensorimotor skill training, where robotic experts can guide human trainees.

Acknowledgments

The work described in this paper was funded by the projects IMMERSENCE of the 6th Framework Program of the European Union, FET, Presence Initiative, contract number IST-2006-027141 and the project BEAMING of the 7th Framework Program of the European Union, contract number ICT-2010-248620.

The authors would like to thank Y. Gao, J. Lu, and Q. Liang, who helped in implementing the system as well as carrying out the experiment.

The authors would also like to thank the anonymous reviewers for their valuable comments and suggestions.

References

Avizzano, C. (2007). Cognitive human-computer communication by means of haptic interfaces. In *Proceedings of the 16th International Conference on Robot & Human Interactive Communication*, 75–80.

Bailenson, J., Yee, N., Brave, S., Merget, D., & Koslow, D. (2007). Virtual interpersonal touch: Expressing and recognizing emotions through haptic devices. *Human-Computer Interaction*, *22*(3), 325–353.

Calinon, S., Evrard, P., Gribovskaya, E., Billard, A., & Kheddar, A. (2009). Learning collaborative manipulation tasks by demonstration using a haptic interface. In *Proceedings of the 14th International Conference on Advanced Robotics (ICAR)*, http://infoscience.epfl.ch/record/138466/files/ICAR2009_Paper194.pdf

Cobos, S., Ferre, M., Ortego, J., & Sanchez-Uran, M. (2008). Simplified hand configuration for object manipulation. *Proceedings of EuroHaptics 2008*, *Lecture Notes in Computer Science*, Vol. 5024 (pp. 730–735). Berlin: Springer.

CyberGlove Systems. (2011). CyberGlove II Data Glove. http://www.cyberglovesystems.com/products/cyberglove-ii/overview

FMOD. (2011). Audio library. http://www.fmod.org

Giachritsis, C., Barrio, J., Ferre, M., Wing, A., & Ortego, J. (2009). Evaluation of weight perception during unimanual and bimanual manipulation of virtual objects. In *Proceedings of the Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems: World Haptics*, 629–634.

Giannopoulos, E., Wang, Z., Peer, A., Buss, M., & Slater, M. (2010). Comparison of people's responses to real and virtual handshakes within a virtual environment. *Brain Research Bulletin*, *85*(5), 276–282.

Groten, R. (2011). *Haptic human-robot collaboration: How to learn from human dyads*. PhD thesis, Technische Universität München.

Gunn, C., Hutchins, M., & Adcock, M. (2005). Combating latency in haptic collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *14*(3), 313–328.

Haans, A., & IJsselsteijn, W. (2006). Mediated social touch: A review of current research and future directions. *Virtual Reality*, *9*(2), 149–159.

Intersense. (2011). Intersense IS900. http://www.intersense.com/pages/20/14

Karniel, A., Nisky, I., Avraham, G., Peles, B.-C., & Levy-Tzedek, S. (2010). A Turing-like handshake test for motor intelligence. In A. M. Kappers, J. B. van Erp, W. M. B. Tiest, & F. C. van der Helm (Eds.), *Haptics: Generating and Perceiving Tangible Sensations, Part I*, *Lecture Notes in Computer Science*, Vol. 6192 (pp. 197–204). Berlin: Springer.

Kunii, Y., & Hashimoto, H. (1995). Tele-handshake using handshake device. In *Proceedings of the 21st IEEE International Conference on Industrial Electronics, Control, and Instrumentation*, Vol. 1 (pp. 179–182).

NVIS. (2011). HMD nVisor SX60. http://www.nvisinc.com/product.php?id=16

Peer, A., & Buss, M. (2008). Robust stability analysis of a bilateral teleoperation system using the parameter space approach. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2350–2356.

Pollard, N., & Zordan, V. (2005). Physically based grasping control from example. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 311–318.

Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, *3*(1), 4–16.

Sanchez-Vives, M., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339.

Sato, T., Hashimoto, M., & Tsukahara, M. (2007). Synchronization based control using online design of dynamics and its application to human-robot interaction. In *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, 652–657.

Spanlang, B., Fröhlich, T., Fernandez, V., Antley, A., & Slater, M. (2007). The making of a presence experiment: Responses to virtual fire. In *Annual International Workshop on Presence*, 303–307.

Tachi, S., Kawakami, N., Nii, H., Watanabe, K., & Minamizawa, K. (2008). TELEsarPHONE: Mutual telexistence master slave communication system based on retro-reflective projection technology. *SICE Journal of Control, Measurement, and System Integration*, *1*(5), 335–344.

Takeda, T., Kosuge, K., & Hirata, Y. (2005). HMM-based dance step estimation for dance partner robot-MS DanceR. In *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)*, 3245–3250.

Voice-O-Matic. (2011). http://www.voice-o-matic.com/

Wang, Z., Lu, J., Peer, A., & Buss, M. (2010). Influence of vision and haptics on plausibility of social interaction in virtual reality scenarios. In *Proceedings of EuroHaptics 2010*, 172–177.

Wang, Z., Peer, A., & Buss, M. (2009a). Fast online impedance estimation for robot control. In *Proceedings of the IEEE International Conference on Mechatronics*, 1–6.

Wang, Z., Peer, A., & Buss, M. (2009b). An HMM approach to realistic haptic human-robot interaction. In *Proceedings of the Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems: World Haptics*, 374–379.

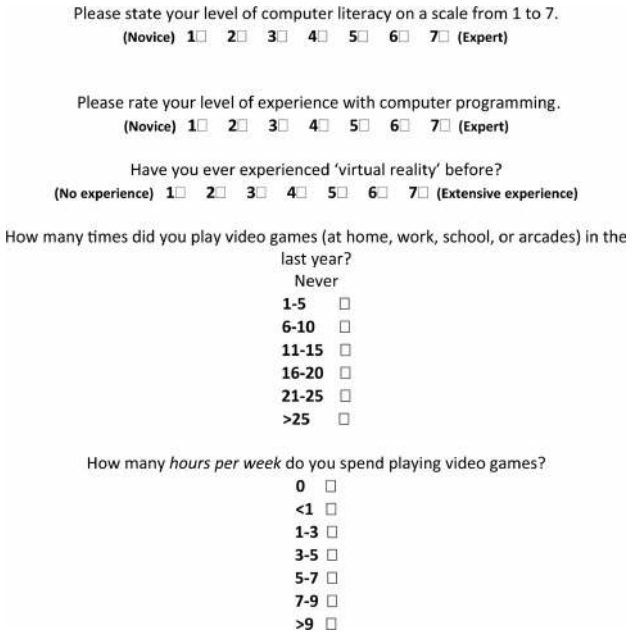
Wang, Z., Yuan, J., & Buss, M. (2008). Modelling of human haptic skill: A framework and preliminary results. In *Proceedings of the 17th IFAC World Congress*, 14761–14766.

Yamato, J., Ohya, J., & Ishii, K. (1992). Recognizing human action in time-sequential images using hidden Markov model. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 379–385.

Yamato, Y., Jindai, M., & Watanabe, T. (2008). Development of a shake-motion leading model for human-robot handshaking. In *Proceedings of the SICE Annual Conference 2008*, 502–507.

Yang, J., Xu, Y., & Chen, C. (1997). Human action learning via hidden Markov model. In *IEEE Transactions on Systems, Man, and Cybernetics*, *27*(1), 34–44.

Appendix: Preexperiment and Postexperiment Questionnaires



Please state your level of computer literacy on a scale from 1 to 7.
(Novice) 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ (Expert)

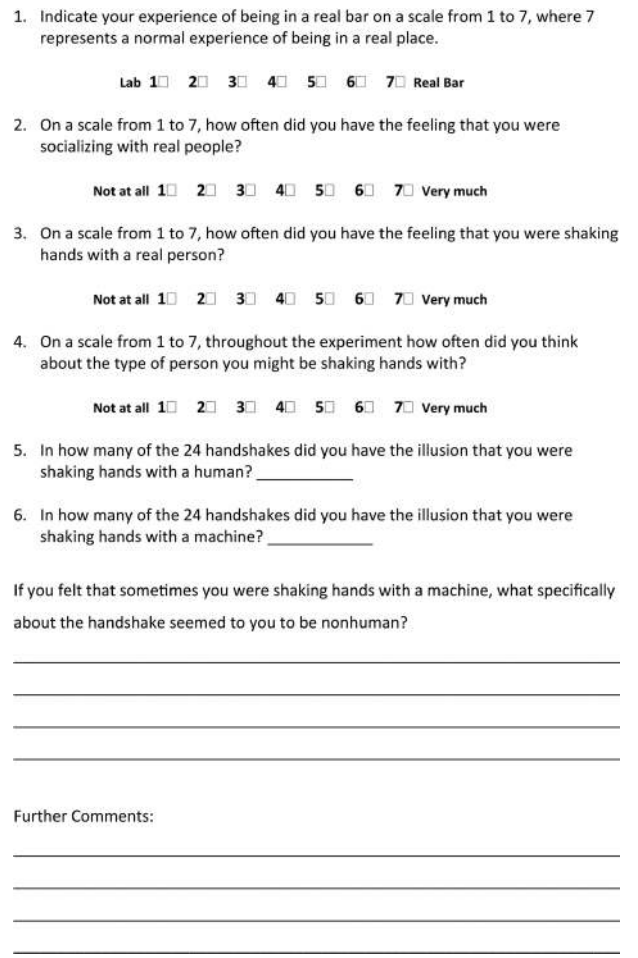
Please rate your level of experience with computer programming.
(Novice) 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ (Expert)

Have you ever experienced 'virtual reality' before?
(No experience) 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ (Extensive experience)

How many times did you play video games (at home, work, school, or arcades) in the last year?
Never
1-5 ☐
6-10 ☐
11-15 ☐
16-20 ☐
21-25 ☐
>25 ☐

How many *hours per week* do you spend playing video games?
0 ☐
<1 ☐
1-3 ☐
3-5 ☐
5-7 ☐
7-9 ☐
>9 ☐

Figure 17. *Preexperiment questionnaire.*



1. Indicate your experience of being in a real bar on a scale from 1 to 7, where 7 represents a normal experience of being in a real place.

   Lab 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ Real Bar

2. On a scale from 1 to 7, how often did you have the feeling that you were socializing with real people?

   Not at all 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ Very much

3. On a scale from 1 to 7, how often did you have the feeling that you were shaking hands with a real person?

   Not at all 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ Very much

4. On a scale from 1 to 7, throughout the experiment how often did you think about the type of person you might be shaking hands with?

   Not at all 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ Very much

5. In how many of the 24 handshakes did you have the illusion that you were shaking hands with a human? _____

6. In how many of the 24 handshakes did you have the illusion that you were shaking hands with a machine? _____

If you felt that sometimes you were shaking hands with a machine, what specifically about the handshake seemed to you to be nonhuman?

Further Comments:

Figure 18. *Postexperiment questionnaire.*