Extending the Body for Interaction with Reality

Tiare Feuchtner  Jörg Müller

Department of Computer Science, Aarhus University, Aarhus, Denmark

Figure 1: We present an ownership-preserving direct manipulation technique in augmented reality, which allows interaction with remote devices in a ubicomp environment with the help of a long virtual arm. While the user’s real hand is close to the body the virtual arm is of normal length (A) and by simply reaching out the user can make it extend to access remote devices in the room. For instance we allow adjusting the height of a table (B), opening and closing a curtain (C) and adjusting the angle of a tilting surface (D).

ABSTRACT
In this paper, we explore how users can control remote devices with a virtual long arm, while preserving the perception that the artificial arm is actually part of their own body. Instead of using pointing, speech, or a remote control, the users’ arm is extended in augmented reality, allowing access to devices that are out of reach. Thus, we allow users to directly manipulate real-world objects from a distance using their bare hands. A core difficulty we focus on is how to maintain ownership for the unnaturally long virtual arm, which is the strong feeling that one’s limbs are actually part of the own body. Fortunately, what the human brain experiences as being part of the own body is very malleable and we find that during interaction the user’s virtual arm can be stretched to more than twice its real length, without breaking the user’s sense of ownership for the virtual limb.

Author Keywords
Ownership; Augmented Reality; Ubiquitous Computing; Virtual Hand Illusion

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Our rooms and environment are filled with an ever increasing number of interactive devices and embedded computers. These range from permanently installed devices such as automated blinds and smart light bulbs, automatic doors and windows, ventilators and air conditioning, to actuated furniture like height adjusting tables and chairs. How people should best interact with and control this wide range of devices remains one of the big unsolved problems of ubiquitous computing.

The oldest technique is to mount controls on the devices themselves, such as switches, cords, handles etc. Here the main problem is reachability, i.e., the user needs to move around to reach the controls, which are sometimes located in inconvenient locations (e.g., a ceiling-fan). This is improved by mounting controls in some central location, such as light switches arranged at a well accessible position on the wall. However, users still need to access this central location, and it might be difficult to map controls to devices. Another way of solving the reachability problem can be by using remote controls. However, remote controls might not be at hand, and the mapping for multiple devices is once again not trivial.

Recently, more and more devices are controlled by web interfaces or mobile phone apps. Here again the main drawback is that users have to acquire their computer or phone, log in, start the right app, and then select the right device to control. Many other interaction techniques have been proposed, most of which split the problem into device selection and command selection. Speech interfaces might be useful, but can be socially inappropriate or sensitive to noise and often require the user to learn the command vocabulary. Similarly gestural interfaces require learning of a gesture command language.
We propose going back to the original direct manipulation technique of mounting controls on the devices themselves, and solve the problem of reachability directly. We therefore virtually extend the users’ arms in augmented reality (AR) to enable direct manipulation of the controls or even the devices themselves. Furthermore, to create a better experience we aim at preserving the perception that the arm is actually the users own arm (ownership). Rapid advances in AR and gesture tracking hardware make this technique feasible in the future, where lightweight glasses or even contact lenses could provide the visual virtual overlay, and sensing technology could be ubiquitously embedded in our environment.

In this paper we present an ownership preserving direct manipulation technique with a very long arm and describe the iterative development thereof. We investigate how the fact that the hand is connected to the users’ body, the realism of the hand, and visibility of the user’s real arm, contribute to the perception of ownership in an AR application.

To summarize, the contributions of this paper are (1) a study of the body ownership effect (in particular the moving virtual hand illusion) in AR, (2) an exploration of the limits of body ownership in respect to appearance, realism and connectivity of body parts, and (3) a technique for controlling devices in ubicomp environments that are out of reach.

RELATED WORK

3D Selection
In order to interact with a ubicomp environment, users need to select and manipulate objects in space. Most of the work in this area has been performed in virtual reality [2, 4, 8]. 3D selection techniques can be categorized by exocentric metaphors (world-in-miniature and automatic scaling) and egocentric metaphors (virtual hand and virtual pointer techniques) [44]. In regards to the latter, virtual pointing is reported to be more effective, more accurate and less strenuous (e.g., through ray casting), while virtual hand can allow a direct transition to the manipulation task (e.g., with a 3D cursor). Our work falls into the category of the virtual hand metaphor. While virtual hands allow to select locations in 3D space, one fundamental problem is that of reach. If control space (the location of the user’s physical hand) and display space (the location of the virtual hand) coincide, objects beyond the physical reach of the user cannot be selected. The key idea to increase reach is to introduce a flexible mapping between control and display, a control to display (C:D) gain.

Control to Display Gain
Poupyrev introduced a C:D gain function that is linear when the hand is close to the body, and increases when the distance of the hand is beyond a certain threshold. In this “Go-Go Interaction Technique” [44] the user controls an abstract hand cursor, which floats in space. The technique increases reach, but not precision. Hindmarsh et al. [20, 25] later applied Go-Go principles in collaborative virtual environments (CVEs), where clicking on distant targets made a humanoid avatar point towards them. The authors found that users had difficulty interpreting other’s pointing gestures and that stretching the pointing arm all the way to the target was helpful. Another class of techniques adjust the C:D gain based on the hand velocity. This is similar to pointer acceleration functions, which follow the same approach for the mapping between mouse and cursor on the desktop [12, 41]. For example, PRISM [21] lowers the C:D gain to enhance precision for object translation and rotation depending on hand speed. Adaptive pointing [37] improves over this approach by simulating absolute pointing behavior. Smoothed pointing [22] presents a further auto-calibrating improvement of the technique.

Distant Reaching in AR and Ubicomp
The need for extending the user’s reach also arises in the context of interaction with interactive tables, large vertical displays, or AR. Some approaches involve using a virtual grabbing tool [1], or laser-style pointing by casting a virtual ray from a touch-pen [42], or the user’s finger [57]. FингAR-tips [10] is an example for a distant reaching technique in both virtual reality (VR) and augmented reality (AR). This interactive urban planning interface allows users to manipulate objects with a virtual representation of their thumb and index fingers. Two interaction spaces are supported: near space (within arm’s reach), and far space. The manipulation technique for far space includes ray based selection and hand gestures that enable direct manipulation, as discussed in the following section.

Direct Manipulation in AR
While Shneiderman’s principles for direct manipulation [47] originally applied to WIMP interfaces, post-WIMP interfaces that are zoomable, include interactive visualizations instead of icons, and widgets instead of menus, are even more natural and direct [6]. Going another step further AR blurs the line between real and virtual so that direct manipulation may not only apply to virtual objects, but just as well to objects with a physical representation in the real world.

For example the HoloDesk [24], a see-through table surface, allows interaction with 3D objects with the bare hands or everyday objects without the burden of wearing a head mounted display (HMD). It is such a nearly tangible interaction between real and virtual objects, which we strive to achieve. However, instead of making users feel like they can touch something virtual with their real hand, we want to make them think they can touch something real with their virtual hand. This may be possible by changing what people perceive as being part of their body.

To summarize, Poupyrev et al. [44] presents a distant reaching technique in 3D but no evaluation thereof. Hindmarsh et al. [20, 25] later applied this technique in desktop-based CVEs where interaction was supported through mouse input. Many other distant reaching techniques involve tools or rays [1, 42, 57], and while some adopt hand shaped pointers [10], these techniques are all evaluated based on performance (e.g., accuracy, task completion time). We propose ownership as another measure by which to evaluate interaction, where we focus on experience instead of performance.
THE PSYCHOLOGICAL CONCEPT OF OWNERSHIP

Body Representations
We have quite a good idea of what we look like when we walk, we can identify our own shadow or mirror image, and we can navigate past obstacles, e.g., walk through a door without bumping into the door post. For these every-day actions we rely on an internal body representation. Research supports that this representation is malleable and is formed and updated continuously [14]. The body representation is frequently divided into two aspects: the body schema and the body image. The body schema is the knowledge of our own body model and can for instance tell us if an object is within reach, or if we have to duck [39]. We use it to plan our movements and it is updated with each change of posture. We can also update our body schema when using tools so that our ‘peripersonal space’, in which we can act, expands. In contrast, our body image can be described as a mental image of our body, e.g., what it looks like from the outside [14].

Self-attribution: Ownership and Agency
Beyond perceiving the world, we also perceive ourselves and the effects we have on our surroundings. We can therefore differentiate between two types of self-attribution: body ownership and agency. Basically we can define body ownership as the feeling of something being part of our own body, e.g., knowing your arm is yours. Agency on the other hand is the feeling of directly causing changes in the environment. With contradicting findings, the correlation of agency and ownership is still unclear [9, 13, 31, 32, 59].

Body ownership illusions (BOIs)
BOIs [34, 43, 56] involve the self-attribution of external objects to the own body, which can be induced through multisensory integration [11, 39], for instance through concurrent visuo-tactile stimulation of a body part. Such illusions may include ownership of a whole artificial body [40, 43, 50], or just an artificial limb [7, 23, 35, 48].

In our paper, we use the Virtual Hand Illusion (VHI), which is related to the Rubber Hand Illusion (RHI) as first reported on by Botvinick and Cohen in 1998 [7]. The RHI was originally studied using the classic setup of a plastic hand that the user perceives as part of his body. More recently however this ownership effect is frequently explored in a mediated setup, i.e., inducing of the feeling of ownership for a virtual hand [28, 35, 48, 49, 53], and is referred to as VHI. The original RHI study [7] showed that synchronous visuo-tactile stimulation can lead to a sense of ownership of a rubber hand. However, the illusion does not depend on this specific combination of the visual and tactile sense: It can also be elicited through visuo-proprioceptive (or visuo-motor) integration (without tactile stimulus) [15, 16, 36, 55, 58, 31], or tactile-proprioceptive integration (without vision) [17]. In our system the user integrates sensory information from vision and proprioception.

It is important to point out that some incongruities of stimuli have a stronger disruptive effect on the illusion, than others. To provide a short overview, we categorize the most inhibiting factors as follows:

- **Matching stimuli**: the stimuli should be synchronous (temporal), collocated and have the same orientation (e.g., stroking the index finger synchronously and in the same direction on both the virtual and real hand) [7, 13, 18, 23, 31, 34, 45, 48].
- **Anatomical plausibility**: the artificial limb must be in a plausible posture (position and orientation) with respect to the body [9, 14, 18, 30, 31, 32, 34, 48, 55].
- **Identity**: the limb must have a familiar appearance and shape [23, 34, 43, 45, 48, 55, 54].
- **Connectivity**: the limb should be visibly connected to the body [48, 53].

If the above criteria are sufficiently fulfilled, the illusion has been shown to be quite flexible and robust: It was found that the ownership illusion can apply to supernumerary limbs (e.g., owning a third arm) or unusual limbs [23, 34], such as arms of different skin color [33], different sizes [5], very long arms [35], and even tails [51]. It has been shown that, in a VR setup with tactile feedback, ownership for a passive (non-moving) virtual arm can be preserved, while slowly stretching it up to 4 times the normal arm length [35]. It is however unclear if ownership persists when the arm length is actively controlled by the user and used for interaction.

Two active VHI studies, where the virtual hand can be controlled and ownership is explored during interaction, is the recent work by Lin and Jörg [38] and Argelaguet et al. [3]. The first presents a VHI comparing 6 different hand representations of varying realism, which participants used to fend off spheres flying towards them in VR. The latter compare 3 different hand representations, where participants complete simple pick and place operations in a VR environment. Both confirm that a realistic hand representation positively impacts ownership.

To give a short summary, most VHI studies are non-interactive, i.e., the virtual hand cannot be controlled by the user. Examples for such are the very long arm illusion [35], and also the referenced work exploring the effect of connectivity on ownership [48, 53]. Those VHI studies that do support interaction [3, 38] do not support distant reaching and the realism of the virtual environment is limited.

We argue that controlling real objects in the environment with our AR setup provides a very different experience. We believe that ownership in AR deserves more attention, since a core difficulty lies in bridging the gap between virtuality and reality in a convincing fashion, which is likely to affect our sense of ownership for a virtual limb.

SUPPORTING OWNERSHIP IN INTERACTION

Designing for Multisensory Integration
To create an illusion of ownership for an artificial or virtual limb, it is necessary that a congruent stimulus is received through multiple senses. In our case we combine vision (seeing the virtual arm move) with proprioception (feeling where your actual arm is). These two stimuli agree in most aspects, such as how the hand moves and twists and to some degree
also what it looks like. It has been shown that the unrealistic effect of the arm becoming longer can, under certain circumstances, be tolerated by the brain [35]. Another factor that could disturb ownership is the delay that occurs between moving the actual hand and seeing the motion of the virtual hand. We have measured this delay to be approximately 100 ms which is below the threshold for detecting visuo-motor delays at 150 ms [19, 46].

**Implementation**

Our setup consists of an HMD (Oculus Rift DK2\(^1\)) and we use an OptiTrack\(^2\) system with 10 Flex13 cameras to track the user and the objects in the room. For this purpose retro-reflective markers are attached to the user’s head, shoulders, right elbow and hands (the view from one of the cameras can be seen in Figure 2). To achieve AR, we converted the Oculus into a see-through HMD by attaching a camera to the front, which provides live video from the user’s point of view. The tracking data and camera stream are both channeled into Unity\(^3\), where we have built a “scene” with virtual representations of the interactive devices in the room and a plane onto which the camera image is projected. Furthermore, this scene contains a virtual full body avatar: a 3D model of a young man that is a standard asset in Unity. Wearing the HMD the user can see the real world captured by the camera, as well as overlaid elements from the virtual world. For instance, the user sees the avatar’s virtual arms as if they were his own. By mapping the movement of the user’s hands to the virtual hands the user receives congruent stimuli from the (virtual) limbs he sees and the (real) limbs he feels (visuo-proprioceptive integration).

\(^1\)https://www3.oculus.com/en-us/dk2/
\(^2\)http://optitrack.com/
\(^3\)https://unity3d.com/

**Room layout and interactive devices**

To simulate the actuated and sensing environment of the future, an electric curtain and two actuated tables were modified with microcontrollers, allowing them to be controlled from the computer. The layout of the room can be observed in Figure 2. The user sits on a stool in the center of the tracking area with a large motorized desk (adjustable in height) located in front of him. The desk is placed out of reach with its front edge 110 cm away from the user’s position (i.e., the stool). An electric curtain is mounted on the trusses just behind the table, 220 cm away from the user’s position. Slightly to the user’s left and about 150 cm away there is a smaller flip-table that can be tilted from horizontal to an angle of about 45°. Markers are placed on the curtain, the desk and the flip-table in order to track their orientation and position. The actuated devices (desk, flip-table and curtain) all have virtual representations in Unity, which are not rendered but simply serve as occluders to the virtual hand when the user reaches underneath or behind them. As a further depth cue we project a shadow from the virtual hand onto these virtual surfaces. Correctly placed in front of the plane with the camera image, this gives the user the impression of actually reaching into the real world, for instance above or underneath the desk, as can be seen in Figure 5. Furthermore, we implemented collision detection preventing the hand representation from passing through the virtual representations of real world objects, giving the user the perception of actually pushing against solid objects.

---

**Figure 2:** Optical tracking of the user and the actuated objects in the room is achieved with retro-reflective markers. This picture shows the view of one of the OptiTrack cameras.

**Figure 4:** The arm-stretch function was implemented based on the Go-Go Interaction Technique. The graph shows the function adjusted to a participant’s arm length, with the exponent \( p = 4 \) and a maximum virtual arm length of \( R_v = 5 \).

**Extending the virtual arm**

Our arm-stretch function is based on the Go-Go-Interaction technique [44] by which the user’s virtual arm is stretched non-linearly in relation to the extension of his real arm. While the hand is located close to the body its position is mapped directly to the virtual hand (\( R_r \) defines the distance from hand to shoulder). But once the user extends his arm beyond a certain threshold (\( R_r > D \)), the virtual arm is stretched along a
vector drawn from the user’s right shoulder to the right hand, so that the virtual hand is located at a distance

\[ R_v = R_r + k \times (R_r - D) \]

from the user’s body [44]. After some testing, we found \( p = 4 \) to be a good value, and both \( D \) and \( k \) are calculated thus to achieve the maximum length of the virtual arm \( R_v = 5 \). The function is shown in Figure 4. That means that when the user’s real arm is fully extended, the virtual hand appears to be 5 meters away from the user’s shoulder.

Apart from the non-linear stretching of the virtual arm, the virtual hand follows the user’s hand motions closely. This allows the user to directly and naturally interact with objects in the environment that are out of his physical reach.

Three such objects are the above mentioned desk, curtain, and flip-table. Reaching out to the top of the desk with the virtual hand, the user can “push” it down making it move down, and likewise he can raise it again by reaching underneath it, “pushing” it up. The curtain behaves similarly: It has a handle fastened on one side and by “pushing” this handle to the left or right, the user can either open or close the curtain. To adjust the angle of the tilting-table, the user can again just lightly push against the surface of the table to make it rotate in the desired direction.

Masking of the real hand

To enhance ownership, the user’s real right arm is made “invisible” by editing each frame of the video. For this purpose the user wears a green sleeve which covers his hand and lower arm (Figure 3, A). Then chroma keying is applied to detect the relevant area in the camera image. After dilating, filtering and blurring we obtain a binary mask (as shown in Figure 3, B), which is then filled with background-colored pixels (Figure 3, C). A detailed account of the implementation is made in the following section of this paper.

DESIGN PROCESS

Our setup and direct manipulation technique were iteratively built and developed involving repeated testing with multiple users and a pilot study (7 participants). During this process, several types of hand representations with varying degree of realism were evaluated. Details about the evaluation methods used in the pilot study can be found in the Evaluation section of this paper, and the most important learnings from this process are described as follows.

The importance of visual depth cues

During early tests it became evident that depth cues were necessary to effectively create the impression of reaching “into” reality with the virtual arm. For this purpose the real world scene was remodeled in Unity to provide realistic collision, occlusion, and shadows. The latter two effects can be observed in Figure 5, where the hand is cut off at the edge of the table, and thus appears to be reaching underneath it, or seems to hover above the table projecting a shadow onto it. The addition of these effects was found to reduce the impression of just waving the arm around in front of a flat display, and effectively provided a sense of reaching into 3D space instead.

Comparison of inpainting approaches

During the pilot study participants indicated that they found it distracting to be able to see both their real arm and the virtual arm at the same time. They mentioned that this made it more difficult to focus on interacting with the virtual arm. Consequently we removed the real arm by masking the area and replacing it with pixels that are similar to the background color. To fill the masked area three different functions were compared, starting with the OpenCV inpaint\(^4\) function using the Navier-Stokes based method, and the variation thereof by Telea [52]. For performance reasons and to reduce flickering

---

\(^4\)http://docs.opencv.org/2.4/modules/photo/doc/inpainting.html

Figure 5: To provide a better sense of interaction with a 3D environment, we implemented depth cues such as shadows (A) and occlusion (B) of the virtual hand.
in the masked area we then implemented a third approach, which we find better suited for inpainting video frames. Here the arm-region is simply filled with old pixel values, which are remembered from before the arm entered the image, thus we call it Memory Inpainting.

**Defining the direction of the arm-stretch**

A non-trivial aspect for the implementation of the arm stretch was to determine in which direction the users arm should be extended (i.e., along which vector the virtual hand should be displaced) to provide the best sense of control. Several alternatives were tested with the displacement vector originating either at the user’s head position, the sternum, the right shoulder, or the elbow. Placement at the head would allow pointing in the line of sight, but it also causes the extended arm to move when the user moves his head to look around. Placement at the sternum is more stable in that respect, however it causes the virtual hand to drift closer when the right arm is stretched to the left and further away when the arm is fully extended to the right. Stretching the arm along a vector originating from the elbow results in a very clumsy, long forearm that proved difficult to control. Finally, drawing the vector from the user’s shoulder position makes the hand follow a path that feels fairly natural. We found this to be best suited for our direct manipulation technique.

**EVALUATION**

We conducted a user study to evaluate our technique and to explore the impact of realism and body-connectivity on the feeling of ownership for the virtual hand representation. Furthermore, we wanted to determine if hiding the real arm is essential for preserving ownership of a virtual arm in AR.

**Experimental Design**

We evaluated the four conditions arm (C1), hand (C2), abstract-hand (C3), and arm w/o inpainting (C4), which we counterbalanced using a Latin Square. The conditions mainly differ in the visual connectivity of the virtual hand to the user’s body, and the hand’s degree of realism and similarity to the user’s actual hand. In C1 the user could see a whole virtual arm and hand replacing his real limb (see Figure 6, top left). When the user extended his arm to reach for a distant object, the virtual arm was stretched further leading to a thinning of the limb giving it a rubbery appearance. In C2 the virtual arm was removed and the user could only see the virtual hand floating in the air (Figure 6, top right). In C3 this virtual hand was replaced by a hand pointer similar to the one shown in “The Go-Go-Interaction Technique” [44] (Figure 6, bottom left). C4 was identical to C1, with the difference that inpainting was deactivated and both the real and virtual arm were simultaneously visible (Figure 6, bottom right). Furthermore, all users started with a baseline in which the virtual arm was of normal length (arm-stretch deactivated), as shown in Figure 7.

**Questionnaire**

In each condition the users were asked to answer a questionnaire (see Table 1) with statements about ownership (Q1-Q4), agency (Q6-Q7,Q10), and other effects related to the direct manipulation technique. The greatest part of the questionnaire is based on previous work [7, 31, 35], since it has been shown to be a reliable indicator for ownership. We added Q4 to explore if ownership could shift over time and Q8 to find out if seeing both the real and the virtual arm (C4) is distracting. Further, we added Q11-13 to explore in how far ownership involved the expectation of haptic stimuli and Q14 to gather subjective impressions of the experience. All of the statements, except for Q4 and Q14, were to be rated on a 5 point Likert scale (1 “strongly disagree” - 5 “Strongly agree”). After all trials were concluded, the users were asked to indicate their preference of condition. Furthermore, to explore lasting effects of the arm stretch, they were asked if they perceived their right arm to be longer than their left.
During the experiment there were times when...

Q1  "I felt as if the virtual hand were part of my body."
Q2  "I felt as if the virtual hand were my own hand."
Q3  "it felt as if I had more than one right hand."
Q4* "If you felt like you had more than one right hand, was that simultaneous, or alternating?"
Q5  "the virtual hand began to look like my real hand."
Q6  "when I moved my real hand I expected the virtual hand to move in the same way."
Q7  "I adjusted the movement of my real hand according to the movement of the virtual hand."
Q8  "During the task I concentrated more on the virtual hand than on my real hand."
Q9  "it felt as if my real arm were becoming longer."
Q10 "I felt as if I was causing the movement of the desk, table and curtain."
Q11 "I was expecting to feel the desk when the virtual hand 'touched' it."
Q12 "I was expecting to feel the flip-table when the virtual hand 'touched' it."
Q13 "I was expecting to feel the curtain when the virtual hand 'touched' it."
Q14 "Describe the interaction in 3 words."

Table 1: Participants were asked to respond to this questionnaire after each trial (C1-4 and CB). Questions 9 to 14 were omitted in the baseline. All questions, apart from Q4 and Q14, were to be answered by indicating agreement on a 5-point Likert scale (1 strongly disagree, 5 strongly agree). Q4 was only asked if Q3 was rated >= 3.

Participants and apparatus
We performed our user study with a total of 15 participants (4 female), aged 25 to 53 (median 29). Of these, 11 participants claimed to have a computer science-related occupation, and 7 indicated to have experienced AR or VR on more than 3 occasions.

As was described earlier, the apparatus consisted of a see-through HMD that provided a view of the room augmented with virtual content, and motion tracking to allow reaching out with a virtual arm to control remote devices.

Procedure
Participants were outfitted with retro-reflective markers before entering the room and were then led into the tracking area with their eyes closed. They were only allowed to open their eyes once they had put on the HMD, so that they could only inspect their surroundings through the head mounted camera. In the beginning of the study each user was asked to extend his right arm and the virtual avatar was rescaled to match the user’s arm length (measured from the shoulder to the right hand). A large part of the virtual body was made invisible, so that the user could only see the virtual hand representation (depending on the condition).

The participants were then asked to familiarize themselves with their virtual right arm by observing it move in accordance to their own hand movements. To establish a baseline with a virtual arm of normal length the participants were then instructed to alternately tap three virtual cubes that were displayed in mid-air within comfortable reach of their right hand (see Figure 7). The cubes changed color when they were touched and collision detection prevented the hand from passing through, giving the user the impression of solid objects. After 90 seconds of interaction a shortened version of the questionnaire evaluated the user's sense of ownership for the virtual arm.

Thereafter the arm-stretch was activated and the user was shortly introduced to the long virtual arm and how it could be used to control the devices. The participants were allowed to practice interaction with the desk, curtain and flip-table, and then followed four trials with the conditions C1-C4 (Figure 6) in counterbalanced order. In every trial the participants were asked to interact with each of the devices, adjusting them as they pleased, resulting in at least 90 seconds of interaction time for each condition. Each trial was again concluded by answering the questionnaire.

RESULTS
Of the 15 study participants, 3 were excluded from further analysis. One person was disregarded due to technical difficulties during the study, and two were omitted on the basis that they did not give a positive rating (>= 3) to any of the ownership statements in the baseline. Since ownership is very subjective, great interpersonal differences may exist [14] and some people are less susceptible to the illusion. In previous work participants have been categorized into “responders” and “rejecters” [31], and it is not uncommon to disregard people of the latter group [18]. Since our aim is to explore what effect our different hand representations have on the ownership illusion, we are only interested in the remaining 12 responders. Our main findings are described as follows.

Evidence of ownership based on questionnaire
The questionnaire responses indicate ownership in the arm condition (C1 with median of agreement: 4 for Q1, 3 for Q2), with similar values as in the baseline (with median of agreement: 4 for both Q1 and Q2) (Figure 8, left and center). We performed a Kruskal-Wallis test to compare the conditions and found significantly stronger agreement to Q1 in the baseline, than in the abstract-hand condition (C3) (p = 0.0084). No further significance was found for Q1 (C1 vs. C3: p = 0.0535, C4 vs. baseline: p = 0.0845).

The pattern of stronger ownership indications in C1 than in C2-C4 is also weakly reflected by the ratings for Q9 (felt as if real arm was becoming longer) with median of agreement of 3.5 (Figure 8, right). No noticeable ownership occurred for any of the other conditions (C2-C4). It is clear however, that the abstract-hand (C3) was perceived as the most artificial since it got the lowest ratings.

When asked if they felt like they owned more than one hand (Q3) most participants indicated agreement in C4 where both their real arm and the virtual arm were visible (median of agreement: 4, see Figure 10, left). This is quite interesting, because in this condition they did not actually indicate ownership based on Q1 (median of agreement: 2), meaning that they felt like they owned a third hand, which was not however part of their body. Here a Kruskal-Wallis test shows a significant difference between the abstract-hand (C3) and the arm w/o inpainting (C4) with p = 0.0419. There is no significant difference between C4 and the baseline (p = 0.0711), or any of the other conditions.
Experiencing an elongated arm
As mentioned earlier, participants slightly agreed that they felt like their real arm was becoming longer (Q9) in C1 (arm) with a median of 3.5, not however in all other conditions (C2-C4).
Interestingly however, the interaction throughout all four trials appears to cause a lasting impression of having an elongated arm: After completing all trials we asked the users to stretch out both arms straight in front of them and indicate if they felt that their right arm was longer (P1) and could reach further (P2) than their left arm. Most participants agreed to both of these statements (median of agreement: 4 for both P1 and P2), as shown in the right boxplot of Figure 10.

Agency ratings indicate strong sense of control
Agency was strong across all conditions with a median of agreement ratings between 4.5 and 5 for both Q6 (expecting virtual hand to move in the same way as real hand, see Figure 9, left) and Q10 (feeling of causing movement of devices, see middle boxplot in Figure 9). The participants also indicated that they regulated the movement of their real hand based on the movements of the virtual hand (e.g., adjusting their hand movement to the speed of the moving table that the virtual hand was pressing against), with median of agreement for Q7 between 4 and 5 for all conditions (Figure 9, right).

Strongest focus of attention on virtual hand
In regards to Q8 some participants indicated that their attention switched back and forth between the real and the virtual hand, and in particular in C4 they found it distracting to see the real hand as well and notice the discrepancies. But overall the majority claimed to have focused more strongly on the virtual hand during interaction, with a median of agreement for Q8 between 4 and 5 across all conditions (Figure 10, center).

The long arm is the preferred hand representation
After all trials were completed, we asked the participants to compare the long arm (C1, C4) with both the hand (C2) and the abstract-hand (C3) and indicate their preference for each pair. While each hand representation found some supporters, the majority of all users (67%) preferred the arm to the abstract-hand. The reasons include that having the arm made the virtual hand feel more connected to the body, it made navigation easier and was a helpful depth cue. Arguments against it were that it was pretending to be something it was not (the
term ‘uncanny valley’ was mentioned). When asked about their preference between arm (C1) and hand (C2) opinions were split with half preferring one and half the other. Arguments for the arm were the same as mentioned earlier and the most common argument against it was that it blocked the user’s view of the target. Furthermore we asked about their preference concerning seeing their real arm during interaction (arm w/o inpainting, C4), or hiding it through inpainting (C1-C3). Two thirds of all users (67%) preferred not seeing their own arm. Even though our Memory Inpainting creates a blurry area in the image, which still reveals the real hand position, most participants found it less distracting and more immersive than when seeing their real arm at the same time. On the other hand, supporters of having their real arm visible appreciated it as a reference and claimed that it helped them navigate the room with the virtual hand.

**Informal description of the interaction**

After each trial at the end of the questionnaire, the participants were asked to describe the interaction with three short terms or phrases (Q14). We categorized these descriptions and list the most common in Table 2, along with the number of occurrences.

We also asked informally if it was clear to the participants how to interact, if they found it easy, and if the devices reacted as they intended. All of these were confirmed, showing that the participants were able to effectively interact with the system without extensive training.

**DISCUSSION**

**Study Results**

Our user study showed that ownership for a virtual hand in AR can be preserved during interaction with remote devices, if the virtual hand representation is sufficiently realistic and there is a visual connection to the user’s body. We also find that seeing the real hand as well as the virtual hand during interaction disrupts the ownership illusion, thus it is important to hide it from view. In our arm condition these requirements were all fulfilled and, in comparison to the other conditions, participants reported highest ownership for this long virtual arm. The long arm was furthermore appreciated because it helped the participants navigate the environment and judge the depth of their reach. However, users also remarked on it occluding their view.

**Applications**

While it remains to be evaluated if direct manipulation with an extended body is more efficient than using other techniques, it certainly provides a very different experience. We believe that it is precisely this experience that is the main strength of our technique. Obvious applications include control of the environment for first-time users, and specially also elderly or disabled users who may have less control over their own body than others. This limits their feeling of empowerment and also the perception of their empowerment by others. This empowerment might be increased by the ability to directly manipulate the own environment through a virtual long arm. In the future we envision our technique to be integrated in ordinary glasses and the normal environment of users. Thus, it could be used on a daily basis.

<table>
<thead>
<tr>
<th>arm (C1)</th>
<th>hand (C2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>elastic (5)</td>
<td>realistic (5)</td>
</tr>
<tr>
<td>weird (3)</td>
<td>disconnected (4)</td>
</tr>
<tr>
<td>depth cue (3)</td>
<td>hand-like (3)</td>
</tr>
<tr>
<td>natural (2)</td>
<td>natural (2)</td>
</tr>
<tr>
<td>easy (2)</td>
<td>powerful (2)</td>
</tr>
<tr>
<td>abstract-hand (C3)</td>
<td>arm w/o inpainting (C4)</td>
</tr>
<tr>
<td>disconnected (6)</td>
<td>natural (4)</td>
</tr>
<tr>
<td>mouse cursor (6)</td>
<td>easy (4)</td>
</tr>
<tr>
<td>unrealistic (5)</td>
<td>reference point (4)</td>
</tr>
<tr>
<td>difficult (5)</td>
<td>distracting (3)</td>
</tr>
</tbody>
</table>

Table 2: This table lists the most common categories of descriptions given for each condition. The number of occurrences is indicated next to each tag.
Figure 11: As a step beyond command languages, direct manipulation interfaces allow users to directly manipulate a representation of data. In the first generation, with Graphical User Interfaces (GUIs), the screen provided the border between the physical and digital worlds. Tangible User Interfaces (TUIs) and Radical Atoms provide a physical representation of the digital world, thereby moving this border into the physical world. Body extension goes beyond these paradigms by moving the border between the physical and the digital into the user’s own body.

Relationship to Body Implants
Another area that proposes to “extend the body” for interaction is that of implantable electronics [26]. In this field electronic devices are implanted in the human body, e.g., under the skin, to facilitate interaction. Interestingly, while implanted electronics physically become part of the body, they do not necessarily become so from the perspective of the human brain. The brain might well perceive the implanted devices as foreign bodies, if they fail to become part of the body schema. This would presumably happen if the devices fail to provide synchronous sensory feedback to events such as touch and pressure. In contrast, our approach does not manipulate the physical body at all, but merely the brain’s impression of what is part of the body. Whether interfaces are part of the physical body or part of what the brain experiences as the body does not need to correlate.

Body Extension as an Interaction Paradigm
Current interaction paradigms include command languages, agents, and direct manipulation with and without tools. Many current techniques can be subsumed under the “command language” paradigm. Gestural interaction, for example, often consists of a command language of gestures, which are recognized by the computer and execute discrete events. Voice commands and buttons (whether virtual or physical) fall into the same category. The main drawback of commands is that the command vocabulary is limited and needs to be learned by users. Because of the “open mic” problem, the command recognition usually needs to be activated. Target selection, command selection, and command parameterization are usually separate steps.

Virtual (or physical) agents have been proposed to solve the problem of having to learn the command vocabulary. These can range from an invisible voice that can control the room to a mobile robot. Users can converse with these agents in a dialogue. Main drawbacks of agents are that interaction might be cumbersome and socially awkward.

Direct manipulation [47] has been proposed to overcome many shortcomings of command languages, in particular the learning of the command set, thereby producing less cognitive load [27]. Further, direct manipulation interfaces allow users to directly manipulate a representation of data. Instrumental interaction [6] proposes that we can use physical and virtual instruments (tools) to overcome limitations of our own body.

Graphical User Interfaces (GUIs) provided the first instantiation of direct manipulation interfaces, and the screen provided the border between the physical and digital worlds. Tangible User Interfaces (TUIs) and Radical Atoms [29] extend this concept by providing a physical representation of the digital world. Thus, they move the border between physical and digital into the physical world. Body extension goes beyond these paradigms by moving the border between the physical and the digital into the user’s own body (see Figure 11).

Combining this extended body with tools is a very interesting direction for future research. Compared to command languages and agents, being able to use the own body with “super powers” might strengthen empowerment, independence, and well-being. This might be particularly beneficial for otherwise less-able users. We believe this might become a very interesting alternative to having to learn commands or interacting with “agents”, e.g., talking to the curtain and asking it to open.

CONCLUSION
We have presented a technique where users can directly manipulate devices in their environment through a virtual extension of their real arm. We have shown that the users’ perception of ownership for a long virtual arm can be preserved, if it looks realistic and appears connected to their physical body. Also the visual removal of the users’ real arm aids the illusion. We believe that the virtual extension of the users body is an interesting alternative to gestural or voice commands and virtual agents.

REFERENCES
   I-grabber: Expanding physical reach in a large-display tabletop environment through the use of a virtual grabber. In Proceedings of the ACM International


42. Petkova, V. I., and Ehrsson, H. H. If i were you: perceptual illusion of body swapping. *PloS one 3*, 12 (2008), e3832.


