Enhancing Presence in Head-mounted Display Environments by Visual Body Feedback Using Head-mounted Cameras

Gerd Bruder, Frank Steinicke, Kai Rothaus, Klaus Hinrichs
Department of Computer Science
University of Münster
Einsteinstr. 62, 48149 Münster, Germany
{*gerd.bruder, fsteini, rothaus, khh*}*@uni-muenster.de*

***Abstract*—A fully-articulated visual representation of a user in an immersive virtual environment (IVE) can enhance the user's subjective sense of feeling present in the virtual world. Usually this requires the user to wear a full-body motion capture suit to track real-world body movements and to map them to a virtual body model.**

In this paper we present an augmented virtuality approach that allows to incorporate a realistic view of oneself in virtual environments using cameras attached to head mounted displays. The described system can easily be integrated into typical virtual reality setups. Egocentric camera images captured by a video-see-through system are segmented in real-time into foreground, showing parts of the user's body, e. g., her hands or feet, and background. The segmented foreground is then displayed as inset in the user's current view of the virtual world. Thus the user is able to see her physical body in an arbitrary virtual world, including individual characteristics such as skin pigmentation and hairiness.

***Keywords*-Head-mounted displays; augmented virtuality; presence**

I. Introduction

In 1965 Sutherland discussed the concept of an ultimate display and its potential for immersing people in virtual worlds. Three years later the first head-mounted displays (HMDs) were presented that could display a stereoscopic three-dimensional view of virtual environments (VEs) to a user. The displayed VE was updated with respect to the user's head motion, which was detected by simple tracking devices [22]. Today, cameras can be attached to HMDs, which provide real-time images of the physical environment from the user's viewpoint. The images displayed on those so-called *video-see-through HMDs* can be augmented in real-time with digital information. For instance, this allows to display virtual objects in the user's view of the real world, which are updated with respect to position and orientation changes of the user's head. In contrast to such augmented reality setups, in typical HMD setups users are immersed in a virtual world and shielded from physical surroundings. The result is that a user is not able to see her physical body through the HMD. Such situations where users cannot see their own body can be confusing for the user. One approach to provide users with a virtual body, a so-called *avatar*, is to use marker-based tracking systems to detect motions of the user's body, which then can be mapped to a virtual body model. Although this allows to display near-natural virtual bodies, few immersive virtual environments (IVEs) make use of such systems due to expensive hardware and necessary instrumentation of users with motion capture suits.

In most virtual reality (VR) applications intuitive exploration of virtual worlds and natural interaction plays an important role. Therefore, it is highly desirable that users get visual feedback about their body in the VE. During interaction with the virtual world a realistic virtual body serves as egocentric frame of reference, gives cues about the size of objects and enhances estimation of distances as well as space perception in the VE [5][16]. Furthermore, Slater et al. have shown that a user's sense of feeling present in a VE, i. e., the impression of "being there", can be enhanced by a virtual body [18][19]. It has been shown that when the virtual body motion matches the user's physical motion, the user's sense of presence is increased [25]. Hence, it has various benefits to incorporate a realistic visualization of an articulated virtual body in an HMD environment, which can be controlled in real-time by the user's physical movements and viewed from an egocentric perspective.

In this paper we describe how visual body feedback can be provided to users by a simple low-cost approach in order to enhance their VR experience. The described system can easily be integrated into existing HMD setups. In our system cameras attached to the HMD provide real-time images from the user's approximate eye-positions in view direction in the real world. We separate the pixels of the user's real body in these camera images from the background and display these as insets in the rendered view of the virtual world (see Figure 1). Hence, the user is able to see her own body in a virtual environment.

The remainder of this paper is structured as follows. Section II covers related work and background information about virtual bodies and visual identities in VR environments. In Section III we describe the classification and segmentation process of egocentric video images, and how the final view is composed. Section IV describes a preliminary study, which shows that this simple camera-based augmented virtuality approach suffices to increase a user's subjective sense of presence compared to a condition without visual body feedback. Section V concludes the paper and gives an overview about future work.

(a) (b) (c)

Figure 1. Virtual body in an augmented virtuality scenario: (a) user in a video-see-through HMD setup; cameras attached to an eMagin 3DVisor Z800 HMD at approximate eye positions. (b) Camera images showing the user's hands in the real world and (c) the user's virtual view with the segmented hands.

II. Related Work

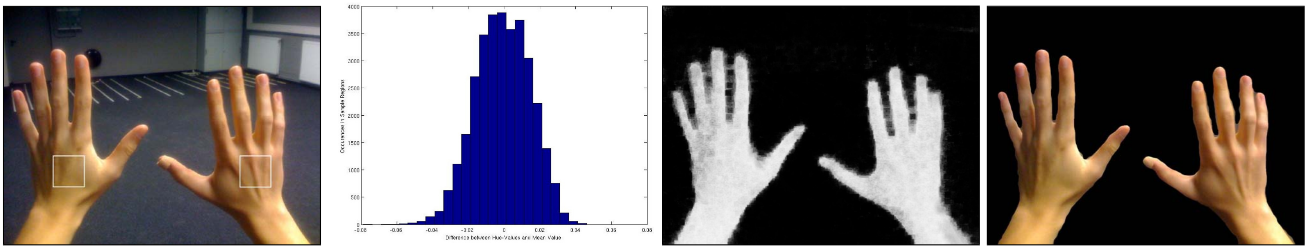
Augmentation of real as well as virtual worlds with digital or analog objects is in the focus of many research groups. Technology that superimposes views to the real world with computer-generated images is called *augmented reality*, whereas the enhancement of virtual worlds using real-world data is called *augmented virtuality*. The term *mixed reality* [13] encompasses both augmented reality as well as augmented virtuality [7][15]. Main issues of mixed reality environments are consistency of virtual and real geometry, time and illumination [7][14]. In augmented virtuality environments the predominant virtual space is enhanced with real-world objects, which may be used to dynamically interact with the VE in real-time. Usually this interaction is not mutual, i. e., real objects are used to interact with virtual objects, but virtual objects give no physical feedback. One popular application for augmented virtuality environments are teleconferencing systems, which integrate two-dimensional real-world camera images into a shared VE so that the users see realistic views of each other in real-time, including body pose, gestures and facial expressions [13].

Real-time camera images taken from objects in the real world can also be used to adapt and texture 3D objects in a VE, which may result in visually faithful virtual objects. Furthermore, certain multi-camera setups allow capturing of real object shapes, which then can be transferred to a VE and textured with images obtained from the cameras [1][6][11]. A study conducted by Lok et al. showed a slight, but not significant, increase in presence when users were presented with realistically textured and visually faithful virtual bodies, which were reconstructed in real-time from camera images, opposed to a condition with generic avatars [12]. Those setups have the advantage that they do not require additional user instrumentation. However, most approaches to reconstruct object shapes and to texture virtual objects using real-time camera images require several cameras and static perspectives, i. e., the setups cannot easily be moved around with respect to a user's position and orientation.

In general, researchers have shown that visualization of avatars can increase a user's sense of presence, i. e., if a virtual body is displayed and if a user has some association with that body, it is likely that the user will believe that she is in that location [17][19]. Besides impact on the sense of presence, it has been shown that a virtual body can affect a user's behavior [27]. For instance, virtual human models for VR applications have been reported to have an impact on social interaction [24]. Furthermore, virtual bodies have an influence on VR interaction tasks [5][16][18] and may increase task performance [12].

In this context, Cave Automatic Virtual Environments (CAVEs) benefit from a user's implicit ability to see her body, since the VE is displayed on the walls of a room so that the user's real body is visible in the foreground with the VE in the background. This gives the user the feeling of standing with her own body in a virtual world, which only leads to confusion in situations when virtual objects are supposed to be visualized closer to the viewer than her body. HMD setups on the other hand have to display a rendered virtual body on the HMD in order to give a user feedback about her body pose. Since the beginning of the 90s, researchers have proposed concepts to incorporate full static or dynamic virtual bodies in HMD environments [2][4]. Direct control of a full virtual body usually requires a user to attach marker patches to her body or wear a motion capture suit that contains markers or sensors to detect body movements [16]. As a result of time-consuming instrumentation, latency and/or expensive hardware, few applications make use of more than a virtual hand model, which can be controlled with a tracked input device, wired glove or through markers attached to a user's hand.

Since it is desired to reduce the user's instrumentation in real walking HMD setups, markers or sensors for position and orientation tracking, headphones etc. are usually attached to the HMD itself. We propose to attach one or two cameras to the HMD approximately at the user's eye positions, so that egocentric camera images can be used to present a user's real body in a VE similar to a CAVE setup. Various algorithms have been proposed that allow to segment a user's body as foreground from the background in dynamic camera images. For this purpose we use a simple skin color detection, which has proven to be a



(a) (b) (c) (d)

Figure 2. Classification and segmentation process to obtain a virtual body: (a) back of the right and left hand covered by rectangular training regions, (b) centralized empirical observation set, (c) region confidence map representing plausibility of pixels to be part of the skin, and (d) identified skin pixels.

useful and robust cue for image segmentation and tracking purposes [3][9][23]. In this paper we present a simple implementation that allows to segment camera images in real-time based on skin color detection and also allows to segment a user's lower body parts from the usually uniformly colored floor of a laboratory (see Figure 6). This simple segmentation process proved applicable in our VR laboratory, which provides sufficient color differences between the user in the foreground and the walls and floor of the laboratory in the background. Furthermore, we have examined if such egocentric body visualization, which gives a user visual feedback about her own body in the virtual world, suffices to increase the user's sense of presence.

III. Obtaining a Virtual Body

In this section we describe how we display a virtual body using real-time egocentric camera images. We explain in detail how we identify a user's extremities in these images, e. g., the user's hands and feet, and how we transfer this real-world information to the user's view of the virtual world. Therefore, no additional instrumentation of the user is required despite one or two cameras attached to the HMD, approximately at the user's eye positions. Since usually the field of view (FoV) of a camera is larger than the FoV of an HMD, we crop the camera images such that images from the HMD and camera match. When the user moves her hand in the view of one of the cameras, the hand is segmented from the rest of the camera image and displayed as inset in the rendered view of the virtual world on the HMD (see Figures 1). Hence, the user is able to see her body in the virtual world when her extremities in the real world come into her view.

A. Video-see-through Augmented Virtuality Setup

We use an eMagin 3DVisor Z800 (800×600, 60 Hz, $40°$ diagonal FoV) for visual presentation. Since this HMD does not feature native video-see-through support, we mounted two USB cameras with a resolution of 640×480 pixels and an update rate of 30 frames per second in front of the HMD. These cameras are used to capture egocentric real-time views of the real world. An infrared LED is positioned on top of the HMD. We track the position of this LED in the room with an active optical tracking system (WorldViz' Precise Position Tracking), which provides sub-millimeter precision and sub-centimeter accuracy. The update rate is 60 Hz providing real-time positional data of the active marker. For three degrees of freedom orientation tracking we use an InterSense InertiaCube2 with an update rate of 180 Hz. The InertiaCube is also mounted on top of the HMD. The virtual world with visual body feedback is rendered on an Intel computer with dual-core processors, 4 GB of main memory and an *n*Vidia GeForce 8800 GTX.

B. Classification and Segmentation

The input images for the segmentation process are provided by the cameras attached to the HMD. As mentioned above, the cameras show the real world from the position and orientation of the user in the physical HMD setup, while head movements are used to render the VE according to tracked motions. An example for an image captured by a camera attached to the HMD is shown in Figure 2 (a).

The simplest approach to segment the user's foreground body from the background in the camera images would be to use a green box approach. Using this procedure the background is uniformly colored with a color that ideally does not occur in the foreground body parts. Since all pixels of that color belong to the background, segmentation of the user's body is a simple task. However, typical VR laboratories do not provide a uniformly colored background and therefore segmentation of body parts is a more challenging task [10]. In the following we describe how to realize an algorithm based on skin detection.

The first step is to transform the image into hue $H \in [0, 1]$, saturation $S \in [0, 1]$ and intensity values $V \in [0, 1]$. In Figure 2 (b) the distribution of the hue values of the skin pixels is shown. Almost all pixels in the training region have a difference of ± 0.04 to the mean color. This observation suggests that the HSV color space is suitable for skin detection [3][23].

We divided the task into two phases, which is common practice in supervised pattern recognition: a training phase and a classification task. To account for different skin colors, the first phase is to train the skin classifier. The user is asked to move her hands, so that the backs of the right and left hand cover the white, square training regions, which we highlighted during this phase for the user (see Figure 2 (a)). The hue values of the pixels within these regions are taken as observation set to compute the mean skin color μ and the standard deviation σ. Figure 2 (b) shows the centralized empirical observation set. In each session this training phase has only to be done once.

The classification phase is realized in three steps. In the first step, the hue values are computed and centralized by $H' = H - \mu$. Since the hue is defined over a cyclic domain, we can assume that H' is mapped into the cyclic domain $[-0.5, 0.5]$. A slight smoothing using a Gaussian filter kernel completes the preprocessing step. In the second step, we estimate for each pixel a value which represents the plausibility of the pixel to be part of a skin region (see confidence map in Figure 2 (c)).

The plausibility value $P(p)$ of a pixel p depends on the centralized hue value $H'(p)$ and the hue contrast $C(P)$, which is the difference of the minimal and maximal H' in a 21×21 neighborhood of p:

$$P(p) = (1 - 2\,|H'(p)|) \cdot (1 - C(p))^2$$

In contrast to [8], where the authors construct a Bayes classifier and assume certain distributions for non-skin pixels, in our approach we only model the class of skin pixels and hence cannot define a statistical classifier. As illustrated in Figure 2 (c), a white-colored pixel corresponds to a high plausibility that the pixel is part of the skin, a black-colored pixel corresponds to a low plausibility. The plausibility function as well as the parameters and thresholds are derived from experiments.

In the third step, segmentation of the skin pixels is performed. We first smooth the plausibility map P. All pixels with a plausibility value larger than $1 - 25\sigma$ are taken as skin pixel candidates. The resulting binary image contains some holes in skin regions and some wrong classified skin regions in the background. Both are reduced by using an extension of the median filter technique, i. e., by counting the number of skin pixel candidates in a 7×7 neighborhood of p. A pixel is redefined as skin pixel candidate if and only if there are at least 13 skin pixel candidates in its neighborhood. We repeat this procedure by increasing the lower bound to 17, 21, and 25. Subsequently, we proceed with non-skin pixels in an analogous way until we obtain a foreground image which shows only the skin.

The approach can be extended to detect regions of the user's body that are covered with cloth, for example, using texture-based classification approaches [9]. However, in the current implementation we focused on skin segmentation.

To be able to display the user's legs and feet in the virtual world, we segment the lower part of the user's body using a slightly modified approach. Instead of segmenting the pixels belonging to the user's lower body when the user looks down towards her feet, we segment the pixels belonging to the floor of the laboratory. Since usually the color of the laboratory floor is nearly uniform, pixels belonging to the floor can easily be segmented (see Figure 6 (b)). We classify all pixels not belonging to the laboratory floor as foreground pixels. Currently, the non-optimized implementation of the segmentation process runs at 14 frames per second on average.

(a) (b)

Figure 3. Composition of pixels corresponding to a user's arm and hand with a view to a VE (a) without and (b) with brightness adaptation.

C. Compositing Process

In this section we describe how we compose the final view from the segmented camera images and a rendered image showing a virtual world. Since the virtual view is updated with respect to tracked physical head motions, including head pitch, yaw and roll, the virtual and physical camera orientations are almost identical. Thus, body parts segmented from the real camera images, which are displayed as insets in the view of the virtual world, appear spatially stable.

We make use of alpha blending in order to display the segmented body in the rendered view of the virtual world. We mask all foreground pixels in the segmented camera images with an alpha value of 1 and all background pixels with a value of 0. During the composition step background pixels from the camera images are neglected and pixels with an alpha value of 1, i. e., regions showing the user's body, replace the corresponding pixels from the image showing the virtual world. Since the real and virtual images may have different resolutions, the real camera image can be scaled and then blended over the virtual image.

Since we do not have information about the depth of pixels in the real-world camera image, we blend the images regardless of depth. This provides correct views as long as the user's hands do not penetrate objects displayed in the VE. In the current implementation we do not extract depth information from the (stereo) camera images or an additional depth camera. However, with this information a depth composite of the real and virtual images could be performed, allowing closer virtual objects to occlude body parts that are farther away. Incorporating depth information would allow to implement three-dimensional interaction of the user's body with the VE.

D. Image Adaptation

As described above, the segmented camera images showing a user's body are displayed as insets in the view of the virtual world. Global illumination effects like shadows and reflections caused by the user's virtual body as described in [21] are not considered. However, in order to adapt a camera image of the user's body to a virtual world,

(a)

(b)

Figure 4. (a) Photo of the real VR laboratory and (b) view of the virtual replica model used in the experiment.

we adjust the brightness of the pixels that correspond to the body to match the brightness of the VE. We compute the brightness of the current view of the virtual world by means of averaging the intensity values V for all pixels. Then we brighten or darken the image showing the virtual body such that the average of the intensity values for each pixel that is part of the body matches the average intensity of the VE. As illustrated in Figure 3, this procedure results in a more realistic visualization of the virtual body. Figure 3 (a) shows a user's arm and hand without brightness adaptation. Figure 3 (b) shows the same scene, but the pixels corresponding to the user's body have been adapted to the average brightness of the VE.

IV. Preliminary Evaluation

In this section we describe a simple experiment, which we conducted to evaluate whether the ability to see one's own body, as segmented from egocentric camera images, in a VE leads to an increased subjective sense of presence. We collected subjective self-estimation of the participants' level of presence using the Slater-Usoh-Steed (SUS) presence questionnaire [26].

A. Experiment Design

The visual stimulus consisted of a virtual replica of our real-world laboratory, i. e., a VE that was modeled and textured to match our physical HMD setup one-to-one (see Figures 4 (a) and (b)). We mapped the tracked orientation and position of a subject in the laboratory to the same position in the virtual replica. We compared the two conditions without (condition C1) and with (condition C2) visualization of the subject's body in the rendered view of the virtual replica. Since the subjects knew the VE from the real world, task instructions and training took only a few minutes.

Setup: We performed all experiments in a $10m \times 7m$ laboratory room. The hardware setup in the experiment consisted of an eMagin Z800 HMD with attached cameras as described in Section III-A. The HMD and cameras were connected via $10m$ cables to the rendering computer. For the experiment we used images of only one head mounted camera and displayed the visual body monoscopically to speed up the update rate of the entire system.

The virtual scene was rendered using OpenGL and our own software with which the system maintained a frame rate of $10-30$ frames per second in both conditions C1 without and C2 with displayed virtual body. Prior to the experiment the subjects trained the skin segmentation process as described in Section III-B, which was supervised by the experimenter to ensure optimal results. The entire training process requires only a few seconds.

Participants: Seven male subjects (age $25-31$, $\varnothing : 28$) participated in the experiment. Subjects were students of computer science or professionals with experience using HMDs. All had normal or corrected to normal vision. All subjects were volunteers and naive to the experimental conditions. Two subjects obtained class credit for their participation. The total time per subject took about 30 minutes, including pre-questionnaire, training, experiment, post-questionnaires, and debriefing. Subjects were allowed to take breaks at any time.

Experiment: We instructed the subjects to perform a simple task: We attached a Wii remote controller to each wall of the VR laboratory at a fixed position. The Wii remote controllers were also visible in the virtual replica model of the real laboratory. The subjects' task was to walk through the laboratory to such a Wii remote controller whenever its virtual representation started blinking. When the subjects pressed a button on the corresponding real-world Wii remote controller with their hand it stopped blinking and a different Wii remote controller blinked to indicate the new target location. The subjects had to click a total of ten buttons to complete the task. Both experimental conditions without (condition C1) and with visualization (condition C2) of a virtual body were tested in separate, randomized sessions. After both sessions for conditions C1 and C2, subjects had to fill out SUS presence questionnaires.

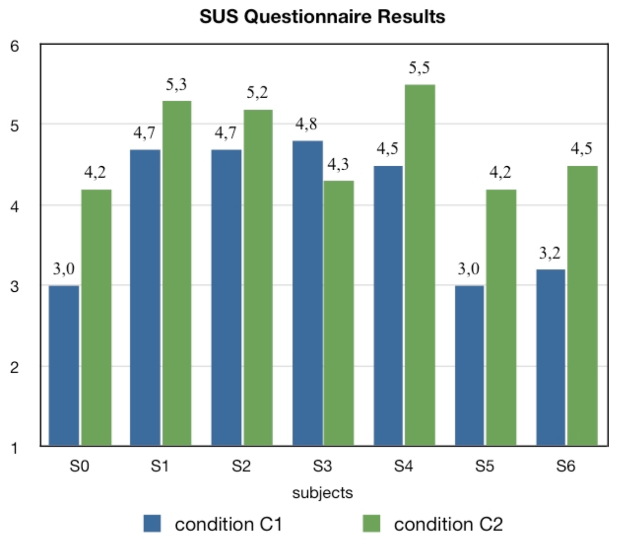


Figure 5. Individual SUS scores for condition C1 without and condition C2 with virtual body.

B. Results

The individual mean SUS scores for conditions C1 and C2 are shown in Figure 5. Averaged over all subjects the mean SUS score in condition C1 was 3.99, and 4.74 in condition C2. The mean SUS score in condition C2 was approximately 19% higher than in condition C1 without virtual body. The increase is statistically significant ($\rho < 0.05$). The results conform with comments of the subjects in an informal debriefing session after the experiment. Only one subject remarked that due to latency of the camera-based setup he felt less present in condition C2. The results motivate that visualization of a virtual body indeed can increase a user's level of presence.

The results of the SUS questionnaire generally show low presence scores. Some subjects informally remarked that the low update rate limited their sense of presence. Other subjects stated that they did not entirely trust the tracking system and the mapping from real to virtual room coordinates, and that they had less fear of colliding with a wall when they were able to see their own hands in front of them.

Most subjects reported informally that they were faster at clicking a button on one of the Wiimote devices when they were able to see their hand, even despite the low update rate. Furthermore, subjects reported that they felt safer when they touched the Wii remote controllers in condition C2, which is pointed out by the comment of one subject:

> *"It was definitely easier for me to judge distances and find objects when I was able to see my hands."*

We have performed further questionnaires in order to determine the subjects' level of simulator sickness. The subjects had to fill out Kennedy's simulator sickness questionnaire (SSQ) before and after the experiments. The Pre-SSQ score averages to 8.23 and the Post-SSQ score to 15.71 with an average increase of 7.48. No subject showed strong symptoms of simulator sickness. There was no significant difference in the sensitivity to simulator sickness between both conditions.

V. Conclusion

In this paper we have presented a simple VR-based approach to incorporate a realistic visual representation of oneself in virtual worlds, which makes use of low-cost camera hardware and software. In contrast to current approaches, such a virtual body can be integrated easily into existing HMD setups without additional user instrumentation such as several tracked markers or motion capture suits. The described augmented virtuality system makes use of real-time images captured by cameras that are attached to an HMD. We have shown how a user's body can be segmented from these egocentric camera images and how we merge this body information with the user's current view of a virtual world. We presented a simple computer vision algorithm, which allows real-time skin segmentation in dynamic camera images, as well as segmentation of a user's legs, feet etc. Furthermore, we evaluated the system in terms of the user's subjective sense of feeling present in a virtual world displayed on the HMD. The results show that using this simple camera-based approach subjects reported a higher sense of presence.

In the future we will extract depth information from stereo camera images or an additional depth camera, which allows more effective segmentation of a user's body in the camera images. Depth information further allows to compute occlusion of the user's real body with objects displayed in the virtual world, as well as body-based interaction with the VE. One limitation of the proposed camera-based approach is that only interaction within the current view of the user is possible. We will evaluate in how far this restricts interaction in practical situations or if this could be enhanced with additional hardware such as more head mounted cameras.

We have tested the setup in a real-walking environment [20] showing the Grand Canyon Skywalk (see Figure 6). Test users were really excited about the possibility to perform a virtual walk on this attraction. They were focused on their feet while walking over the glass-bottomed horseshoe and getting visual body feedback in the VE (as shown in Figures 6 (b) and (c)).

We believe that the described approach is a valuable alternative to full-body tracking in situations where users cannot be instrumented with additional hardware. The system allows an enhanced sense of presence and is based on easily replaceable and upgradeable hardware. In the future we will evaluate how visualization of such an augmented virtuality body affects space cognition, which is important for applications such as architectural exploration, and performance in typical VR tasks.

Figure 6. (a) Picture of the Grand Canyon Skywalk, which is a giant, glass-bottomed horseshoe that juts out 65 feet from the edge of a cliff and that is 3,800 feet above the canyon floor, (b) camera image showing the user's lower body parts, i. e., legs and feet, and (c) virtual view of a digital skywalk with visual body feedback.

REFERENCES

[1] J. Allard, C. Ménier, B. Raffin, E. Boyer, and F. Faure. Grimage: Markerless 3d interactions. In *SIGGRAPH (Emerging Technologies)*, 2007.

[2] N. Badler, C. Phillips, and B. Webber. *Simulating Humans: Computer Graphics Animation and Control*. Oxford University Press, 1993.

[3] J. Brand and J. Mason. A comparative assessment of three approaches to pixellevel human skin-detection. In *In Proceedings of the International Conference on Pattern Recognition*, volume 1, pages 1056–1059, 2000.

[4] V. Buchmann, T. Nilsen, and M. Billinghurst. Interaction with partially transparent hands and objects. In *ACM Proceedings of the Sixth Australasian Conference on User interface*, volume 40, pages 17–20, 2005.

[5] M. Draper. Exploring the influence of a virtual body on spatial awareness. Master's thesis, Seattle, WA 98195, USA, 1995.

[6] M. Gross, S. Würmlin, M. Naef, E. Lamboray, C. Spagno, A. Kunz, E. Koller-Meier, T. Svoboda, L. V. Gool, S. Lang, K. Strehlke, A. V. Moere, and O. Staadt. Blue-c: A spatially immersive display and 3d video portal for telepresence. In 819-827, editor, *ACM Transactions on Graphics (TOG)*, volume 22, 2003.

[7] K. Jacobs and C. Loscos. Classification of Illumination Methods for Mixed Reality. Technical report, Eurographics 2004, 2004.

[8] M. J. Jones and J. Rehg. Statistical color models with application to skin detection. In *International Journal of Computer Vision*, volume 46, pages 81 – 96, 2002.

[9] P. Kakumanu, S. Makrogiannis, and N. Bourbakis. A survey of skin-color modeling and detection methods. *Pattern Recognition*, 40(3):1106–1122, 2003.

[10] Y. Kojima, Y. Yasumuro, H. Sasaki, I. Kanaya, O. Oshiro, T. Kuroda, S. Manabe, and K. Chihara. Hand manipulation of virtual objects in wearable augmented reality. In *Proceedings of International Conference on Virtual Systems and Multimedia*, pages 463–469, 2001.

[11] B. Lok. Online model reconstruction for interactive virtual environments. In *Symposium on Interactive 3D Graphics*, pages 69–72, 2001.

[12] B. Lok. *Interacting With Dynamic Real Objects in Virtual Environments*. PhD thesis, University of North Carolina at Chapel Hill, 2002.

[13] H. Maeda, T. Tanikawa, J. Yamashita, K. Hirota, and M. Hirose. Real world video avatar: Transmission and presentation of human figure. In *Proceedings of the IEEE Virtual Reality 2004*, page 237, Washington, DC, USA, 2004. IEEE Computer Society.

[14] P. Milgram, D. Drascic, J. Grodski, A. Restogi, S. Zhai, and C. Zhou. Merging Real and Virtual Worlds. In *Proceedings of IMAGINA'95*, 1995.

[15] P. Milgram and F. Kishino. A Taxonomy of Mixed Reality Visual Displays. In *IEICE Transactions on Information and Systems, Special issue on Networked Reality*, 1994.

[16] B. Mohler, H. Bülthoff, W. Thompson, and S. Creem-Regehr. A full-body avatar improves egocentric distance judgments in an immersive virtual environment. In *oceedings of the 5th Symposium on Applied Perception in Graphics and Visualization (APGV 08))*, pages 194–197, ACM Press.

[17] M. Slater, A. Steed, J. McCarthy, and F. Maringelli. The influence of body movement on subjective presence in virtual environments. *Human Factors*, 40(3):469–477, 1998.

[18] M. Slater, M. Usoh, and A. Steed. Depth of presence in immersive virtual environments. In *Presence: Teleoperators and Virtual Environments*, volume 3, pages 130–144, 1994.

[19] M. Slater, M. Usoh, and A. Steed. Taking steps: The influence of a walking metaphor on presence in virtual reality. In *ACM Transactions on Computer-Human Interaction (TOCHI)*, volume 2, pages 201–219, 1995.

[20] F. Steinicke, G. Bruder, L. Kohli, J. Jerald, and K. Hinrichs. Taxonomy and implementation of redirection techniques for ubiquitous passive haptic feedback. In *Cyberworlds*, pages 217–223. IEEE Press, 2008.

[21] F. Steinicke, K. Hinrichs, and T. Ropinski. Virtual Reflections and Virtual Shadows in Mixed Reality Environments. In *10th International Conference on Human-Computer Interaction (INTERACT2005)*, pages 1018–1021, 2005.

[22] I. Sutherland. The Ultimate Display. In *Proceedings of IFIP Congress 2*, pages 506–509, 1965.

[23] J. Terrilon, M. Shirazi, H. Fukamachi, and S. Akamatsu. Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *In Proceedings of the International Conference on Face and Gesture Recognition*, pages 54–61, 2000.

[24] N. M. Thalmann and D. Thalmann. The artificial life of synthetic actors. *Transactions of the Institute of Electronics, Information and Communication Engineers D-II*, J76D-II(8):1506–14, 1993. MIRALab, Geneva Univ., Switzerland.

[25] M. Usoh, K. Arthur, M. Whitton, R. Bastos, A. Steed, F. Brooks, and M. Slater. The visual cliff revisited: A virtual presence study on locomotion. In *International Workshop on Presence*, 2006.

[26] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionaires in reality. *Presence: Teleoperator in Virtual Environments*, 9(5):497–503, 1999.

[27] N. Yee. *The Proteus Effect: Behavioral Modification via Transformations of Digital Self-Representations*. PhD thesis, Stanford University, 2007.