# FakeMi: A Fake Mirror System for Avatar Embodiment Studies

Marc Erich Latoschik* Jean-Luc Lugrin† Daniel Roth‡
HCI Group, University of Würzburg

**Figure 1:** *Left three images: Sketches through the design phase of the fake mirror system. Second from left depicts a potential application in a real living scenario. Right image: Actual implementation of the system and setup for typical studies on avatar embodiment. Note that the mirrored avatar looks perspectively correct for the user and is distorted in this picture due to the applied stereoscopy and head-tracking.*

## Abstract

This paper introduces a fake mirror system as a research tool to study the effect of avatar embodiment with non-visually immersive virtual environments. The system combines marker-less face and body tracking to animate the individual avatars seen in a stereoscopic display with a correct perspective projection. The display dimensions match typical dimensions of a real physical mirror and the animated avatars are rendered based on a geometrically correct reflection as expected from a real mirror including correct body and face animations. The first evaluation of the system reveals the high acceptance of the setup as well as a convincing illusion of a real mirror with different types of avatars.

**Keywords:** Avatar, Virtual Body Ownership, Tool, Virtual Mirror

**Concepts:** •**Computing methodologies** → **Virtual reality;** *Mixed / augmented reality;* •**Human-centered computing** → *Empirical studies in HCI;*



## 1 Introduction

Behavioral research, cognitive sciences, and clinical psychology have adopted Virtual Reality (VR) technology as a promising research and therapy tool. The control of comprehensive artificial stimuli covering fundamental human senses to mimic real-world stimuli provides replicable experimental setups. They enable us to investigate correlations between stimuli and behavior and psychological state and state changes and provide controlled manipulations, e.g., for exposure therapies and similar interventions. An interesting correlation is defined by the *Proteus* effect [Yee and Bailenson 2007]. It describes the change of the behavior of an individual caused by the visual and behavioral characteristics of their avatars as digital representation of their bodily appearance in a virtual environment. The Proteus effect relies on the illusion of virtual body ownership (IVBO), the acceptance of the displayed avatar to be a digital representation of one's self. IVBO itself is an extension of the classical rubber hand illusion [Botvinick and Cohen 1998] to the digital domain [IJsselsteijn et al. 2006; Slater et al. 2008].

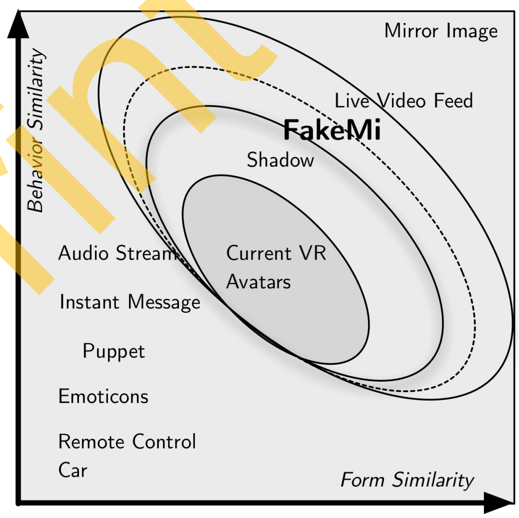


**Figure 2:** *Targeted localization of the proposed metaphor and system following [Bailenson et al. 2008].*

Experimental setups to investigate IVBO and Proteus effects try to cover as much of the behavioral (movement) as well as appearance factors (shape, texture, look) and usually involve the whole body [Lugrin et al. 2015]. Typically, these systems use Head-Mounted Displays (HMDs) or CAVE-like displays [Cruz-Neira et al. 1992] together with marker-based full-body tracking. They often become quite cumbersome and require users to wear additional I/O hardware, e.g., the displays, motion capture suits, or individual retro-

*e-mail:marc.latoschik@uni-wuerzburg.de
†e-mail:jean-luc.lugrin@uni-wuerzburg.de
‡e-mail:daniel.roth@uni-wuerzburg.de


VRST '16,, November 02-04, 2016, Garching bei München, Germany
ISBN: 978-1-4503-4491-3/16/11...$15.00
DOI: http://dx.doi.org/10.1145/2993369.2993399

reflective markers. They are either fully or partially immersive, blocking out much of the real physical space around the user and either completely prevent users to see their real body or they modify the outer appearance caused by the I/O hardware to wear. The effects of avatar embodiment in fully immersive VR systems are well-known (see [Spanlang et al. 2014]). Avatar embodiment in lower immersive mixed-reality environments received less interest, despite a large spectrum of possible applications.

Our objective is then to provide a non-visually immersive VR system which implements a *fake mirror* metaphor (see Figure 1) as a tool to explore the pscho-physiological effects of avatar embodiment [Spanlang et al. 2014]. The proposed system is named FakeMi (Fake Mirror). It aims at extending the choices of technology currently available as defined by the spectrum for the classification and representations of humans in physical and digital space by [Bailenson et al. 2008] as illustrated in Figure 2.

*Fake mirrors* could offer novel perspectives for psychological studies as well as for entertainment (e.g, video games or interactive storytelling), therapeutical (e.g. exposure therapies), or even for sociological or commercial applications (e.g. public displays for advertising or information). Here, we focus on the requirements and ability of such a platform to elicit a convincing mirror illusion with current technology limitations as initially proposed in [Lugrin et al. 2016]. We investigated the impact of three main factors on the potential elicited sense of agency:

1. Low-level physical immersion
2. Avatar realism and user resemblance
3. Motion and face tracking latency and fidelity

### 1.1 System Characteristics

We introduce a non-visually immersive VR system faking a real mirror. Its main characteristics are:

1. It allows users to *come as they are*, i.e., the physical self is not artificially augmented with sensors and devices, therefore preparation and rigging time are largely reduced.
2. It embeds a virtual mirror screen in the physical environment, thereby anchoring the virtual self in the real environment enabling users to constantly compare their virtual and real appearance.
3. It includes marker-less body and face tracking, the latter one being an important aspect of a faithful user representation, identity acceptance, and an equally important cue when it comes to interpersonal communication [Roth et al. 2015].
4. It incorporates individualized avatars approximating the users' real physical appearance using a scan of the users' heads and faces.

## 2 System Design

Figure 1 illustrates three design sketches and the final implementation of the FakeMi system. We investigated potential hardware and software solutions and tools given the following requirements: *R1)* low intrusiveness, to reduce equipping time, let users *come as they are*, and increase the users' comfort. In order to foster a wider adoption the system should incorporate tools being *R2)* low cost. We assume that a combined external (camera-based) sensor for face and body tracking is not sufficient, thus our approach demands *R3)* mutual sensor compatibility. To counteract further preparation time we require *R4)* fast user calibration. Additional requirements involve *R5)* accurate tracking in a given range of operation and finally *R6)* integration in state-of-the-art graphics or simulation engines (such as Unity3D or Unreal engines, which are both popular within the VR research and professional communities, as well as capable of providing high-quality avatar animation and appearance). The overall architecture is depicted in Figure 3.

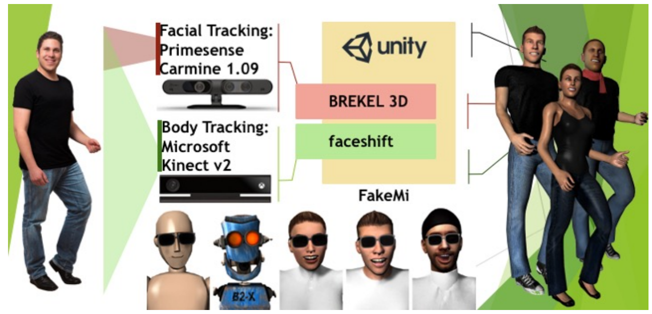


**Figure 3:** *FakeMi - A low cost and marker-less fake mirror system integrating both real-time face and full-body avatar embodiment in VR following [Lugrin et al. 2016].*

### 2.1 Tracking of Facial Expressions

We analyzed *Pro Face 2*, *Performer Suite*, *Faceshift Studio* and *FacePlus* as solutions for facial tracking. We tested and analyzed critical features to fulfill the requirements for each system using simple demo scenes. Table 1 illustrates our overall comparison criteria and results. We found Faceshift Studio to robustly and adequately track the user's facial expression, due to the user template generation and the utilization of both RGB and depth data. The system can also replicate a faithful 3D model of user's head after a short training phase. This choice led to the selection of Unity® as visualization engine since no plugin exists for Unreal®. Marker-based tracking systems or head-mounted cameras have not been considered since they are too invasive (*R1*).

**Table 1:** *Comparison of facial tracking software solutions.*

| | ProFace2 | Performer | Faceshift | FacePlus |
|---|---|---|---|---|
| Robust Tracking | × | ✓ | ✓ | ✓ |
| Accurate Tracking | × | ✓ | ✓ | ✓ |
| Approximate Price | $150 | $10000 | $1500 | $1500 |
| Tracking without Training | ✓ | × | × | ✓ |
| Manual Training Process | × | ✓ | ✓ | × |
| Webcam video | × | ✓ | × | ✓ |
| Depth sensing device | ✓ | × | ✓ | × |
| Head Mesh Creation | × | × | ✓ | × |
| Integrated in Unity3D | ✓ | ✓ | ✓ | ✓ |
| Integrated in Unreal | × | ✓ | × | × |

### 2.2 Tracking of Body Movements

Markerless body tracking systems compatible with Unity® and utilizing the *Kinect v2* can be divided into three types: i) *Unity plugins* (e.g., *Kinect v2 with MS-SDK*), ii) *VR Middleware* (e.g., *MiddleVR*), and iii) *Standalone applications with a Unity plugin* (e.g., *Brekel Pro Body*). The comparison is depicted in Table 2. Brekel Pro Body v2 is our final choice due to its elaborated smoothing techniques and the ready-to-use plugin integration. It allows an extensive control of the applied smoothing and the resulting latency via the user interface and an instant mapping of tracked user movements to the avatars.

**Table 2:** *Comparison of body tracking software solutions.*

| | Pro Body v2 | MiddleVR | Unity Plugins |
|---|---|---|---|
| Tracking without Training | ✓ | ✓ | ✓ |
| Training Process available | ✓ | × | × |
| Smoothened Tracking Data | ✓ | × | × |
| Network Capability | ✓ | ✓ | × |

### 2.3 System Integration

Following our initial requirements and comparative tests, we selected Pro Body (V 2.280, 64bit) and Faceshift Studio (V 2015.1.02) as the target software solutions. Both deliver state-of-the-art tracking quality (*R5*), fast user calibration (*R4*) and model integration into Unity® (V 5.1.1) (*R3, R6*). A fully-working prototype has been developed. The system's general architecture from Figure 3 is implemented by dedicated software components depicted in Figure 4. Three main component are responsible for the management of (1) the virtual characters, including face and body tracking and the control of the experimental conditions, (2) the views, i.e., the head-tracked stereoscopic rendering in combination with MiddleVR, and (3) the virtual environment with the mirrored room and the virtual version of the real mirror.

The system drives Faceshift Studio with the Carmine 1.09 and Pro Body with the *Kinect v2* (both software and hardware sensors fulfill *R1*, *R2*, *R3*, *R5* and *R6*). Both sensor devices work well in combination: the Kinect v2 utilizes time-of-flight to measure depth [Sarbolandi et al. 2015] and the Carmine 1.09 uses structured light [Vongbunyong and Chen 2015]. As depicted in Figure 3, FakeMi is capable of projecting both user's facial expression and body movement to a multitude of different avatars in real-time with a very short calibration and equipment phase (*R4*).

The FakeMi system includes a Unity® module providing a generic avatar, animated by the combined input from both tracking systems. It also provides a high-level API to customize and calibrate avatar's dimension to individual body sizes and proportions. In addition, our generic avatar permits to rapidly import character models produced by popular 3D modeling softwares, such as *Autodesk Character Generator* or *Poser*. One of the most interesting features is truly the large spectrum of avatar types possible, and their interchangeability; from low realistic (e.g., mannequin avatar) to moderately realistic (e.g., human male or female avatar) to highly realistic (e.g., custom head mesh and texture) avatars. FakeMi allows developers to quickly import new avatars as well as to replace their head by the scanned user's head model from Faceshift Studio, after a short modeling phase of approximately 15-20 minutes.

The optimal operating range of the system is represented by a volume of approximately 1.3 m length $\times$1.3 m width $\times$2 m height *(R5)*, which suitable for a large variety of VR applications and experiments settings, especially for virtual mirror configurations.

The system uses an standard consumer display of type LG 55UB850V used in portrait orientation with an outer dimension of 123 cm (h) x 73 cm (w) and a native resolution of 3840 x 2160 pixels. The screen borders were covered by a wooden panel to hide any screen details and to mimic a real physical mirror. The display was mounted approximately 80 cm above the floor (see Figure 1 right). The screen is theoretically capable to deliver a 1920 x 1080 pixel full HD 3D resolution which was fed from the rendering computer to the display. Stereoscopy was achieved by polarization filters modified to match the portrait orientation. As typical for similar consumer devices, the screen induces a notable input lag between 66 ms and 84 ms increasing the overall end-to-end latency.

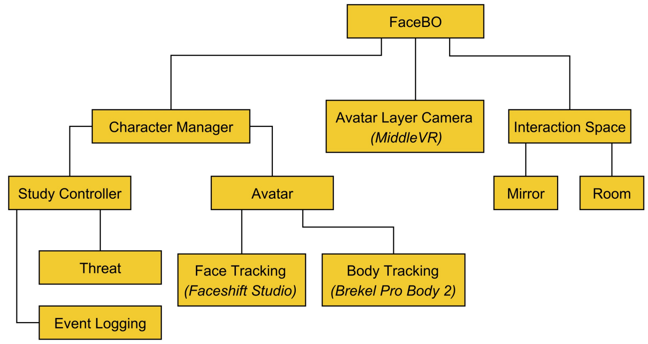


**Figure 4:** *Overview of the main software components implementing the FakeMi system.*

## 3 Evaluation

To evaluate the system, participants were asked to interact with the system as if they were looking at a real mirror. The basic setup of the study is depicted in Figure 1 right. The study used a within-subject design consisting of five conditions of one minute each. We assumed that uncanny valley effects could happen in our study when using human avatars [Lugrin et al. 2015]. Thus, we used avatars of different human resemblance and added another condition with an individualized avatar initially generated using a face scan. Condition one to four were using a randomized order of four avatars, i.e., a wooden mannequin, a robot, a generic human (male or female, depending on the participants gender), and an individualized avatar as illustrated in the bottom row in Figure 3 (here just the faces are depicted but complete body shapes were used).

The user task ensures participants to constantly interact with their real and their virtual body. Prerecorded oral instructions were played-back asking participants to perform three randomized types of actions:

1. Look at a specific body part with minimal body movement (e.g. right foot, left upper arm or belly).

2. Make facial expression (e.g., anger, disgust, fear, joy, sadness, and surprise) based on Paul Ekmans' primary emotions.

3. Move body part (e.g., turn head up, raise right arm).

Users had three seconds to execute each instruction. A sound informed participants to look back onto the screen and stop the current task. After another three seconds the next instruction was given. This was repeated as long as time was left during the one minute sessions.

### 3.1 Measures

The study included the following questionnaires and measures (questions using a 7-point Likert scale were originally translated in German):

1. **Avatar's agency**: using the agency's factors (i.e., *myMovement*, *bodyControlEnjoyment, controlMovements, causeMovements*) of an extended post-experimental IVBO questionnaire based on [Slater et al. 2010; Gonzalez-Franco et al. 2010; Kalckert and Ehrsson 2012; Lugrin et al. 2015].

2. **Frame counting**: To measure end-to-end latency based on frame counting [He et al. 2000] comparable to [Lugrin et al. 2015] using a high speed-camera running at 480 Hz.

### 3.2 Results

The results of the agency's factors are illustrated in Figure 5. High values for all the scores indicate that participants had a stronger feeling that the movements they observed seemed to be their movements. These results did hold true for any of the chosen avatars regardless of their composition and look. Hence, the successful elicitation of agency seems to be provided by the system characteristics, i.e., a sufficiently convincing visio-motor synchrony, correct perspective projection and hence an effective mirror illusion. These results were achieved despite a borderline end-to-end latency of approximately $150 - 200$ ms which is hardly within the necessary threshold ($\leq 150$ ms) for real-time interactions.

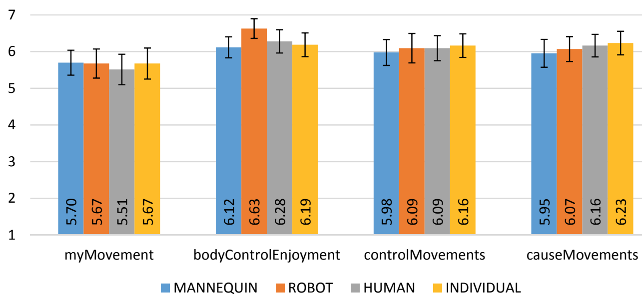


**Figure 5:** *The results from the IVBO questionnaire's agency factor.*

### 3.3 Conclusion and Future Work

This paper presented a new apparatus enabling further research on avatar embodiment effects. The FakeMi system implements a fake mirror metaphor. It targets the significance and impact of the degree of immersion while providing a combined body and face tracking with the application of individualized avatars generated from face scans. The proposed system is indeed on the *low end* on the immersion scale, in combination with the support of individualized face scans and face animations it considerably differs from previous work and existing platforms for avatar embodiment research, typically in immersive setting. Our first results are promising and support the initial goals, specifically the system's capability to elicit a convincing mirror illusion despite current technological limitations, and hence to open-up novel research perspectives.

In future work we will improve the system in terms of tracking quality (end-to-end latency, tracking volume, tracking resolution) and avatar appearance (photorealistic rendering). A reduced latency is critical to support a wider range of interactions without inducing unwanted effects. A straight forward goal is the exchange of the utilized display type. Finger tracking certainly has to be improved since the current solution lacks robustness and reliability. A potential obstacle is the unclear future of the currently used Faceshift software after its recent acquisition by Apple, but we are already investigating promising alternatives with our research partners as well as novel commercial ones (e.g., *Intel RealSense, Faceware Live*).

## Acknowledgements

The authors like to thank David Zilch for his extensive support during the design and implementation of the system and user study.

## References


BAILENSON, J. N., YEE, N., BLASCOVICH, J., AND GUADAGNO, R. E. 2008. Transformed social interaction in mediated interpersonal communication. *Mediated interpersonal communication*, 77–99.

BOTVINICK, M., AND COHEN, J. 1998. Rubber hands' feel'touch that eyes see. *Nature 391*, 6669, 756–756.

CRUZ-NEIRA, C., SANDIN, D. J., DEFANTI, T. A., KENYON, R. V., AND HART, J. C. 1992. The cave: audio visual experience automatic virtual environment. *Commun. ACM 35*, 6, 64–72.

GONZALEZ-FRANCO, M., PEREZ-MARCOS, D., SPANLANG, B., AND SLATER, M. 2010. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE virtual reality conference (VR)*, IEEE, 111–114.

HE, D., LIU, F., PAPE, D., DAWE, G., AND SANDIN, D. 2000. Video-based measurement of system latency. In *International Immersive Projection Technology Workshop*.

IJSSELSTEIJN, W. A., DE KORT, Y. A. W., AND HAANS, A. 2006. Is this my hand i see before me? the rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators and Virtual Environments 15*, 4, 455–464.

KALCKERT, A., AND EHRSSON, H. H. 2012. Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in human neuroscience 6*.

LUGRIN, J.-L., LATT, J., AND LATOSCHIK, M. E. 2015. Anthropomorphism and illusion of virtual body ownership. In *International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments*.

LUGRIN, J.-L., ZILCH, D., ROTH, D., BENTE, G., AND LATOSCHIK, M. E. 2016. Facebo: Real-time face and body tracking for faithful avatar synthesis. In *in Proceedings of the 23rd IEEE Virtual Reality (IEEE VR) conference*.

ROTH, D., LATOSCHIK, M. E., VOGELEY, K., AND BENTE, G. 2015. Hybrid Avatar-Agent Technology - A Conceptual Step Towards Mediated "Social" Virtual Reality and its Respective Challenges. *i-com Journal of Interactive Media 14*, 2.

SARBOLANDI, H., LEFLOCH, D., AND KOLB, A. 2015. Kinect range sensing: Structured-light versus time-of-flight kinect. *Computer Vision and Image Understanding 139*, 1 – 20.

SLATER, M., PEREZ-MARCOS, D., EHRSSON, H., AND SÁNCHEZ-VIVES, M. V. 2008. Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience, 2008, vol. 2, num. 6*.

SLATER, M., SPANLANG, B., SANCHEZ-VIVES, M. V., AND BLANKE, O. 2010. First person experience of body transfer in virtual reality. *PloS one 5*, 5, e10564.

SPANLANG, B., NORMAND, J.-M., BORLAND, D., KILTENI, K., GIANNOPOULOS, E., POMÉS, A., GONZÁLEZ-FRANCO, M., PEREZ-MARCOS, D., ARROYO-PALACIOS, J., MUNCUNILL, X. N., ET AL. 2014. How to build an embodiment lab: achieving body representation illusions in virtual reality. *Frontiers in Robotics and AI 1*, 9.

VONGBUNYONG, S., AND CHEN, W. H. 2015. *Disassembly Automation: Automated Systems with Cognitive Abilities*, 1 ed. Springer.

YEE, N., AND BAILENSON, J. 2007. The proteus effect: The effect of transformed self-representation on behavior. *Human communication research 33*, 3, 271–290.