The final authenticated version is available online at https://doi.org/10.1109/TVCG.2019.2912607.

A Survey of Full-Body Motion Reconstruction in Immersive Virtual Reality Applications

Polona Caserman, Augusto Garcia-Agundez, and Stefan Göbel

Abstract—Due to recent advances in virtual reality (VR) technology, the development of immersive VR applications that track body motions and visualize a full-body avatar is attracting increasing research interest. This paper reviews related research to gather and to critically analyze recent improvements regarding the potential of full-body motion reconstruction in VR applications. We conducted a systematic literature search, matching VR and full-body tracking related keywords on IEEE Xplore, PubMed, ACM, and Scopus. Fifty-three publications were included and assigned in three groups: studies using markerless and marker-based motion tracking systems as well as systems using inertial measurement units. All analyzed research publications track the motions of the user wearing a head-mounted display and visualize a full-body avatar. The analysis confirmed that a full-body avatar can enhance the sense of embodiment and can improve the immersion within the VR. The results indicated that the Kinect device is still the most frequently used sensor (27 out of 53). Furthermore, there is a trend to track the movements of multiple users simultaneously. Many studies that enable multiplayer mode in VR use marker-based systems (7 out of 17) because they are much more robust and can accurately track full-body movements of multiple users in real-time.

Index Terms—

✦

1 INTRODUCTION

FULL-BODY motion reconstruction in virtual reality (VR) applications is necessary to enable natural interaction and a much higher level of immersion [1], [2]. VR applications aim to increase the feeling of presence in VR (sense of being there) by tracking the full-body movements of the user and applying them to an avatar [3], [4]. Similarly, the work of Slater and Wilbur [5] shows that immersion requires a virtual body whereas presence requires that the user identifies with that virtual body. In other words, for a high sense of presence, the users must recognize the movements of the virtual body as their movements. Many recent research publications have added a growing base of evidence to support the use of motion capture technologies in VR applications. Previous studies show the benefits of having a full-body avatar in a virtual environment [6], [7]. A realistic looking avatar evokes a significantly higher acceptance of the virtual body [8]. Recent studies have demonstrated that immersive VR can be used to induce the illusion of ownership over a virtual body, e.g., to reduce the implicit racial bias against dark-skinned people [9] or to reproduce the experience of the world as a child [10]. The progress of real-time full-body tracking has also proven to be advantageous to perform collaborative tasks between human and robot. [11].

In this survey, we aim to gather and critically analyze recent improvements regarding the potential of full-body motion reconstruction in VR applications. First, we give an overview of recent studies by discussing the advantages and disadvantages of different motion capture systems and Head-Mounted-Display (HMD) technologies. Then, we compare different avatar forms (realistic avatar vs. stick avatar and full-body vs. partial-body representation) and perspectives (first-person vs. third-person view). We discuss the VR related key findings, such as end-to-end latency, cybersickness, and level of immersion. Finally, we analyze the papers in terms of the number of players (single vs. multiplayer) and system evaluation.

The need for full-body motion reconstruction is already evident in a single-player experience [12]. When a user is immersed in VR, she is not able to see the real environment and particularly not the real body. Especially in multiplayer games, it is essential to reconstruct full-body motions in real-time [3]. Otherwise, the users cannot see the movements of other users and therefore cannot interact with each other. Desai et al. [13] note that the visualization of full-body avatars in multiplayer applications improves the VR experience. Due to a lack of user's motion data, current VR applications typically do not represent the user's body [14]. In particular, many VR applications show floating hands or even only VR controllers. Recent work reveals that there is a significantly higher sense of presence for the full-body avatar over the hands-only representation [6], [15], although there is some contradictory evidence [16]. When VR applications do not represent body movements to the user, there is no visual sense that the user is entirely in the virtual environment.

Recent advances in motion capture technologies provide various possibilities for full-body tracking in real-time to reconstruct avatars with respect to the user. Different motion capture technologies have their advantages and disadvantages regarding accuracy, robustness, latency, and complexity. A markerless system requires no additional sensors attached to the user's body [3]. The human motions are analyzed from the RGB-D camera, such as Microsoft Kinect. Body tracking with the Kinect device has been possible for

---

- *P. Caserman, A. Garcia-Agundez, and S. Göbel are with the Multimedia Communications Lab, Technische Universität Darmstadt, 64283 Darmstadt, Germany, E-mail: firstname.lastname@kom.tu-darmstadt.de*

Manuscript received October 22th, 2018; accepted April 18th, 2019.

DOI: https://doi.org/10.1109/TVCG.2019.2912607

some time now. However, the potential of this technology could not be fully exploited in VR. The data obtained from this sensor are often affected by jitter [17], inconsistent tracking [17], [18], cannot handle occlusion (some body parts are not visible for the sensor) [14], [19]–[21], and suffer from high latency [12], [22], [23]. The tracking depends on the position of the device and Field-of-View (FoV), making it difficult to see and to track all body parts. The Kinect can extract the skeleton of the user accurately when the user is facing the sensor and struggles when the user is only visible from the side [20]. These challenges are even more obvious in room-scale VR applications. To track users regardless of their orientation, multiple Kinect devices are required. Previous work presented fusion strategies to combine multiple Kinect devices to either track a single user [13], [24] or multiple users [4], [17].

More precise motion capture technology is based on multiple markers and cameras. A marker-based motion capture system either requires attachment of small retro-reflective markers to a bodysuit which are then tracked by infrared (IR) cameras [1], [8], [9], [25] or to use LED markers attached directly to the body [26]. Therefore, markers or sensors must be tied to the key points of the body to track the movements of the bone joints over time. Such a system is very accurate [1], [17]; however, it can cause discomfort since the user must wear a tight suit [17], [27] and it can be very expensive [9], [17].

Alternatively, an inertial measurement unit (IMU), which typically consists of a gyroscope, an accelerometer, and a magnetometer, can be easily bound to the user's body [21]. Such a system can track body movements in real-time [28] and is inexpensive [29], [30]; however, IMUs usually suffer from drift [28], [30].

To experience an immersive VR, the user must wear a novel HMD, such as the Oculus Rift CV1[1] or HTC Vive[2]. Both HMDs provide a high-definition resolution of 2160×1200 pixels, split between each eye with a wide FoV of 110° and a high refresh rate of up to 90 Hz. Users can experience the VR while seated (in a chair) or by physically walking around. External tracking allows the user to walk around freely. Recent studies already provided significant evidence that the room-scale VR configuration leads to higher immersion [31].

When motions are tracked in real-time, it is possible to reconstruct a full-body avatar. The users wearing a VR headset can then view an avatar from a first- or third-person view. A study by Bodenheimer and Qiang [7] shows that different forms of virtual avatars alter the user's perception and attitude. Realistic representations of full-body avatars can increase the sense of embodiment [8]–[10]. Representing a full-body avatar has a significant influence on the behavior of the user and can increase the sense of presence [15], [22]. Eye tracking will further enhance the sense of presence in a virtual world shared between two users [32]. Moreover, tracking of lips can also be essential to improve the level of embodiment [33].

The paper is organized as follows. In Section 2 we describe the method of selecting papers, tracking and representing a full-body avatar in VR. In Section 3, we present the results of research publications which meet our requirements and discuss them in Section 4. We conclude our paper in Section 5.

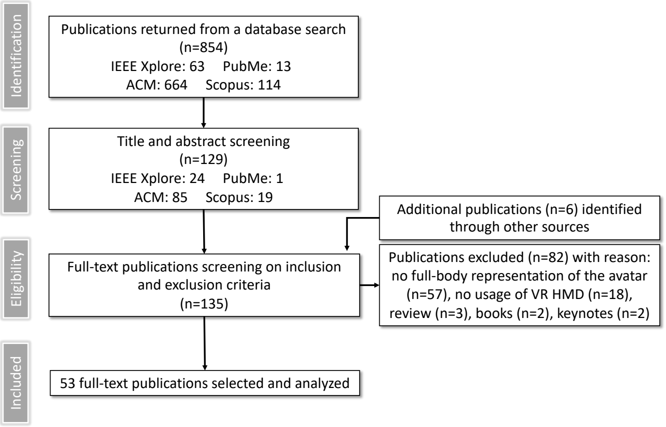

Fig. 1. PRISMA [34] flowchart of the results from the literature search.

2 METHODS

We searched for publications matching the following keywords: *virtual reality* AND (*full-body tracking* OR *motion tracking*) AND (*head mounted display* OR *headset*), published since January 1st, 2013 (release of the Oculus DK1[3]). The search was performed on February 7th, 2018 and again on September 29th, 2018. Inclusion criteria were: a) the application reconstructs full-body movements; b) the system tracks body movements; c) the application transforms the user's motions onto a virtual avatar; d) research publications were published in the past six years; e) research publications are written in English. Exclusion criteria were: a) the application represents only a part of the body; b) the application does not utilize a VR HMD; c) reviews; d) keynotes; e) books:

- **IEEE Xplore**: 63 publications, of which we excluded 49 publications due to the described criteria above, leaving us with 14 publications
- **PubMed**: 13 publications, only one additional publication meets our criteria
- **ACM**: 664 publications, 26 additional non-duplicate publications using the same criteria
- **Scopus**: 114 publications, six additional non-duplicate publications using the same criteria

Figure 1 shows an overview of the literature search. First, we reduced our corpus to a total of 47 research publications. Then, we identified six more publications through other sources that meet our requirements. Finally, we proceeded to classify the total number of 53 publications in the following groups:

1) **Markerless Motion Capture Systems**: Publications using motion capture systems based on time of flight, such as the Microsoft Kinect[4].

---

[1] https://www.oculus.com, last visited on October 9th, 2018

[2] https://www.vive.com, last visited on October 9th, 2018

[3] https://www.kickstarter.com/projects/1523379957/oculus-rift-step-into-the-game/posts/440293, last visited on October 9th, 2018

[4] https://support.xbox.com/en-GB/browse/xbox-360/accessories/Kinect, last visited on October 9th, 2018

2) **Marker-based Motion Capture Systems**: Publications using motion capture suits, such as OptiTrack[5] or Vicon[6]. We also include publications which use VR devices, such as HTC Vive tracker and controller, to track movements of the user and thus to reconstruct a full-body avatar.
3) **Motion Capture Systems using IMUs**: Publications using IMU sensors, such as Perception Neuron[7], PrioVR[8] or Xsens Inertial Motion Capture system[9].

TABLE 1
Classification of research publications, included in this survey.

Motion Capture System	**Number of Papers**
Markerless	29
Marker-based	17
IMU	7
Total	53

3 RESULTS

In the first step, we studied the full-body tracking methods used in each publication to classify them in groups. We noticed that there is a strong focus on using markerless motion capture systems (29 out of 53) in comparison with marker-based systems (17 out of 53) or approaches using IMUs (7 out of 53). Table 1 shows the results of the paper analysis.

3.1 Markerless Motion Capture Systems

Many research publications use a markerless motion capture system to reconstruct full-body motions in VR. Body pose estimation can be done using a single RGB-D camera, such as the Kinect sensor, e.g., to personalize the furniture by posing and acting [35], to virtually climb ladders [36], for cycling-based exergames [23], [37], [38], for collaborative manufacturing tasks between human and robot [18], as a training simulation in a classroom [39], as an exposure therapy for acrophobia (fear of heights) [22], to create role-playing games [40], for engineering and construction [2], for task planning [19], for posture recognition [41], and even for risk perception analysis [42].

Games that are restricted to a single camera limit the possible in-game experience regarding user interactions [13]. Occlusion can cause discrepancies in tracking when some body parts are not visible by the Kinect sensor [14], [21]. Another disadvantage of the Kinect is that it can only track the full-body of users when they are standing at a certain distance in front of it [43]. Reus et al. [42] propose a solution to gently force the user to keep their body aligned with a motion sensor fixed in the real world. The Kinect does not perform well, especially when the user is not facing the sensor [19], [20]. To overcome this challenge, some studies demonstrated the usage of multiple Kinect devices to either track a single user [13], [24] or to create multiplayer VR experiences [4], [17].

In multiplayer VR applications, users see their full-body avatar and also the avatar of other users. Among others, several studies focus on attempting to develop multiplayer games by including users without a HMD into the VR experience. For example, Chou et al. [40] developed a multiplayer game, where one user with the VR headset is playing a giant. The Kinect device tracks only the motions of this player. Users without a HMD join forces to defeat the giant through different platforms. A user controlling a catapult aims at the giant and launches projectiles through smartphone sensors such as a gyroscope and an accelerometer. In another multiplayer setting, a teacher can train classroom management skills [39]. Similarly, the Kinect device tracks only the body movements of a trainee wearing a VR headset. An instructor controls the simulation of a virtual classroom with 24 virtual agents via a desktop application.

Because the Kinect device cannot recognize hand gestures properly, developers often use this device in combination with a Leap Motion[10] controller [2], [17], [19], [44]. Leap Motion contains a pair of IR cameras and can recognize the user's hand in the camera's FoV. The Kinect sensor can be used to detect rough body posture (legs, upper body, arms) and the Leap Motion sensor to detect hand gestures [2]. Research by Leoncini et al. [17] suggests to apply sensor fusion of multiple Kinect devices and Leap Motion controllers to enable full-body tracking with direct hand manipulation. A Leap Motion controller can also be replaced by a Myo armband[11], to recognize hand gestures by tracking electrical activities of the muscles [15]. In addition to the Kinect device, one research publication uses a dance pad for navigation [19]. Table 2 presents details of each study using a markerless motion capture system.

Because the Kinect does not require to attach additional sensors on the user's body, many researchers use this technology to develop exergames. Hoang et al. [45] present a VR system for posture or movement training, such as yoga or dance. The movements of both, the instructor and student are captured through skeletal tracking by the Kinect sensor and represented by a stick figure. The student can see both avatars superimposed. In other words, the student can mimic the movements of the instructor from a first- or third-person perspective. Tanaka and Hirakawa [46] developed a similar exergame, where the player trains in a virtual gym. A virtual trainer gives advice and encouragement during the exercises to keep the player motivated to do exercises effectively.

Further existing exergames used an exercise bike in combination with a Kinect and a VR headset. Bolton et al. [37] developed a VR cycling-based exergame inspired by the arcade game PaperBoy. Shaw et al. [23] developed a similar exergame to increase user motivation to work out on

---

[5] http://optitrack.com, last visited on October 9th, 2018
[6] https://vicon.com, last visited on October 9th, 2018
[7] https://neuronmocap.com/, last visited on October 9th, 2018
[8] https://yostlabs.com/priovr/, last visited on October 9th, 2018
[9] https://www.xsens.com/functions/human-motion-measurement/, last visited on October 9th, 2018
[10] https://www.leapmotion.com, last visited on October 9th, 2018
[11] https://support.getmyo.com/, last visited on January 23st, 2018

TABLE 2
Summary of the research publications using a markerless motion capture system.

Author	HMD	Motion Capture System	Avatar	Synchronizing	Nr. of Players	View	Latency
Charles [20]	Oculus Rift DK2	Kinect	Unknown	Full-body	Single	First	Zero latency
Bolton et al. [37]	Oculus Rift DK1	Kinect	Unknown	Upper body	Single	First	/
Shaw et al. [23]	Oculus Rift DK1	Kinect v1	Sphere	Head	Single	First	Not responsive enough
Matsas et al. [18]	Oculus Rift DK2	Kinect v1	Realistic avatar	Full-body	Single	First	/
Collingwoode-Williams et al. [33]	Oculus Rift CV1	Kinect v1	Realistic avatar	Upper body, lips	Single	First	/
Yan et al. [47]	Oculus Rift DK1	Kinect v1	Realistic avatar vs. stick figure	Full-body	Single	Third	/
Greuter and Roberts [43]	Oculus Rift DK1	Kinect v2	Unknown	Full-body	Single	First	/
Hoang et al. [45]	Oculus Rift DK2	Kinect v2	Stick figure	Full-body	Multiplayer	First and third	/
Lee et al. [35]	Oculus Rift DK2	Kinect v2	Stick figure	Full-body	Single	First	/
Latoschik, Lugrin et al. [39]	Oculus Rift DK2	Kinect v2	Realistic avatar	Full-body	Multiplayer (HMD + desktop)	First	73 ms
Chou et al. [40]	Oculus Rift DK2	Kinect v2	Realistic avatar	Upper body of one player	Multiplayer (HMD + desktop)	First	/
Kondo et al. [16]	Oculus Rift DK2	Kinect v2	Realistic avatar vs. feet and hands	Full-body	Single	Third	80 ms
Liu et al. [41]	HTC Vive	Kinect v2	Realistic avatar	Full-body	Single	First	/
Tanaka and Hirakaqa [46]	Oculus Rift DK2	Kinect v2	Stick figure	Full-body	Single	First	/
Takala and Matveinen [36]	Oculus Rift DK1	Kinect, PS Move controller, Razer Hydra	Realistic avatar	Full-body	Single	First	/
Schäfer et al. [22]	Oculus Rift DK1	Kinect v1, PS SixAxis controller	Realistic avatar	Full-body	Single	First	/
Reus et al. [42]	Sensics zSight HMD	Kinect v1, Arduino	Realistic avatar	Full-body	Single	First	/
Mendes et al. [15]	Oculus Rift DK2	Kinect v2, Myo armband	Realistic avatar vs. simplified hands	Full-body	Single	First	/
Hilfert and König [2]	Oculus Rift DK2	Kinect v2, Leap Motion	Humanoid avatar	Full-body + hand gestures	Single	First	Maintaining 75 FPS
Czesak et al. [44]	Oculus Rift DK2	Kinect v2, Leap Motion	Realistic avatar	Full-body + hand gestures	Single	First	/
Lin et al. [19]	Oculus Rift CV1	Kinect v2, Leap Motion, Dance Pad	Humanoid avatar	Full-body + hand gestures	Single	First	/
Tuveri et al. [38]	Oculus Rift DK2	Kinect v2, Raspberry PI	Realistic avatar	Full-body	Single	First	/
Friðriksson et al. [14]	HTC Vive	Kinect v2, HTC Vive controllers	Humanoid avatar	Full-body	Multiplayer	First	Minimize latency by predicting user's motions
Sra and Schmandt [4]	Oculus Rift DK2	Multiple Kinect v1	Humanoid avatar	Full-body	Multiplayer	First	/
Leoncini et al. [17]	HTC Vive	Multiple Kinect v2, Leap Motion	Stick figure	Full-body + hand gestures	Multiplayer	First	/
Desai et al. [13]	Oculus Rift	Multiple Kinect v2 (12)	Realistic avatar	Full-body	Single	First	<50 ms
Otto et al. [24]	Oculus Rift DK2	Multiple Kinect v2, WiiMote	Humanoid avatar	Full-body	Single	First	/
Rhodin et al. [27]	Oculus Rift DK2	Fisheye camera attached to a HMD	Stick figure	Full-body	Single	First	10 -15 FPS
Headleand et al. [48]	Oculus Rift DK2	Leap Motion, xBox controller	Humanoid avatar	Arms (the person is sitting in a wheelchair)	Single	First	/

TABLE 3
Summary of the research publications using a marker-based motion capture system.

Author	HMD	Motion Capture System	Avatar	Synchronizing	Nr. of Players	View	Latency
Chagué and Charbonnier [49]	Oculus Rift DK2	Vicon	Realistic avatar	Full-body	Multiplayer	First	250 Hz for tracking
Bodenheimer and Fu [7]	NVIS nVisor SX	Vicon (8 cameras + 6 markers)	Realistic avatar vs. stick figure	Full-body	Single	First	/
Young et al. [6]	Oculus Rift DK2	Vicon (8 cameras + 6 markers)	Realistic avatar vs. floating hands	Full-body	Multiplayer	First	/
Thomas et al. [50]	Oculus Rift DK2	Vicon (10 cameras + 15 markers)	Stick figure	Full-body	Single	First	39 ms
Egeberg et al. [51]	Oculus Rift DK2	OptiTrack (22 cameras + 5 markers)	Realistic avatar (with wings)	Upper body	Single	First	/
Kasahara et al. [1]	Oculus Rift DK2	OptiTrack (8 cameras + 37 markers)	Stick figure	Full-body	Multiplayer	First	$\approx$ 70 ms
Latoschik, Roth et al. [8]	Oculus Rift CV1	OptiTrack	Realistic avatar vs. wooden mannequin	Full-body	Single	First	$<$ 150 ms
Schmidt et al. [52]	Oculus Rift DK1	OptiTrack (3 markers)	Realistic avatar	Full-body	Single	First	$<$ 100 ms
Banakou et al. [10]	Unknown HMD	OptiTrack (34 markers)	Realistic avatar (adult vs. child)	Full-body	Single	First	/
Peck et al. [9]	NVIS nVisor SX	OptiTrack (12 cameras)	Realistic avatar	Full-body	Single	First	100 Hz
Bourdin et al. [25]	NVIS nVisor SX	OptiTrack (12 cameras + 34 markers)	Realistic avatar	Full-body	Multiplayer	First	/
Borland [53]	Unknown HMD	OptiTrack (12 cameras)	Unknown	Full-body	Single	First	/
Debarba et al. [26]	Oculus Rift DK2	PhaseSpace ImpulseX2 (14 cameras + 4 LED markers) , PS Move controller	Realistic avatar	Upper body	Single	First	/
Jiang et al. [3]	HTC Vive	HTC Lighthouse sensors (2 Vive controllers)	Realistic avatar	Full-body	Multiplayer	First	7 ms
Seele et al. [32]	Oculus Rift DK2, HTC Vive	HTC Lighthouse sensors (2 Vive Controller)	Realistic avatar	Upper body, eyes	Multiplayer	First	/
Loviska et al. [54]	Oculus Rift DK2	Location tracking system, Leap Motion	Minecraft avatar	Upper body	Multiplayer	First	Noticeable delay
Tan et al. [11]	Oculus Rift DK2	Unknown	Unknown	Full-body	Single	First	/

TABLE 4
Summary of the research publications using IMUs.

Author	HMD	Motion Capture System	Avatar	Synchronizing	Nr. of Players	View	Latency
Eubanks et al. [21]	Oculus Rift DK1	IMU sensors (17), Nintendo Wii Remote (2)	Unknown	Full-body	Single	/	/
Podkosova et al. [28]	Oculus Rift DK2	PrioVR (11 IMUs)	Realistic avatar	Full-body	Multiplayer	/	60 FPS
Malleson et al. [30]	Oculus Rift CV1	Perception Neuron	Personalized avatar	Full-body	Single	/	/
Johnson et al. [29]	Samsung Gear VR	Perception Neuron (32 IMUs)	Unknown	Full-body	Multiplayer	/	$>$ 300 ms
Melo et al. [55]	Oculus Rift DK2	A tracker and a sensor on the bike	Realistic avatar	Full-body (only when the feet are on the pedals and hands on the handlebar)	Single	First	/
Caserman et al. [56]	Oculus Rift DK2	Embedded inertial sensors of a HMD	Realistic avatar	Full-body (only when walking on treadmill and grabbing the handle)	Single	First	/
Kishore et al. [57]	NVIS nVisor SX	Xsens	Robot	Full-body	Multiplayer (HMD + Robot)	First	/

the bike. The game contains a very simple avatar with two spheres (one for the head and one for the body). However, not all movements of the virtual body match the movements of the real body. The game developed by Bolton et al. [37] only tracks the player's upper body to detect when the user throws a newspaper. Both research publications combining these two hardware components (a Kinect sensor and a HMD) report some issues. Due to the limitations of the Kinect sensor, the player has to lean too far in either direction, possibly causing a too high safety risk [37]. The players were also frustrated due to the low framerate and non-instantaneous movement reactions [23]. Tuveri et al. [38] developed a similar game and used a Kinect device to track the user's movements. Additionally, the authors mounted a Raspberry PI to the exercise bike to sense the user's cycling speed.

Due to low FoV of the Kinect device, other researchers attempted to develop new body tracking technologies. Rhodin et al. [27] present an egocentric markerless motion capture system. Their system estimates the full-body skeleton pose from a pair of fisheye cameras attached to the VR headset. However, their solution reaches only 10 to 15 FPS and is therefore not capable of full-body tracking in real-time. For more responsive and accurate tracking, Takala and Matveinen [36] demonstrate that the Kinect device can be used together with a PS Move and a Razer Hydra controller. Other recent studies try to minimize latency by interpreting and adaptively predicting the movements of the player [14]. By analyzing the user's posture and movements, it is possible to understand what the user is doing to decide how to represent or adapt the actions inside the virtual environment.

3.2 Marker-based Motion Capture Systems

A marker-based motion capture system utilizes retro-reflective markers attached to the bodysuit. These markers are then tracked by multiple IR cameras to provide skeleton data of a human body [9], [25]. LED markers can also be directly attached to the body [26]. The motion capture system reconstructs then the skeleton from observed marker positions and several cameras, usually eight or more. The movements of the user can be determined with millimeter precision [1]. The cameras are typically placed around a scene to cover the area in which the user will move approximately. The user must place the markers carefully to prevent occlusion problems, i.e., in areas that users are unlikely to touch [49]. Table 3 presents a detailed description of each study using a marker-based motion capture system.

Since a marker-based motion capture system is capable of real-time full-body tracking of multiple users, many authors used this system to develop VR multiplayer applications or games, e.g., to explore cultural heritage [49], to control the physical condition of athletes and dancers [1], and to create a VR shooting game [3]. Further multiplayer experiences include a co-located social VR game [32] or allow one person to visit a distant location to interact with local people there [25]. Furthermore, Young et al. [6] present a VR experience where one user gives another user a high five while the users can either see a full-body avatar or only hands. Moreover, Chagué and Charbonnier [49] combine the real and virtual world by placing optical markers on the real-world objects to bring them into the virtual environment. For example, a simple cardboard box in the real world can be represented as an Egyptian chest in the virtual world. Hence, multiple users can easily interact with the 3D virtual objects and with other users by using the sense of touch. Thus, in all these VR applications, multiple users can share the same virtual environment and can see each other's full-body avatar.

It is not always necessary to track all extremities to reconstruct full-body movements. Egeberg et al. [51] developed a game in which a player binds markers only on both hands, shoulders, and hip. Thus, the player can freely move most of the body parts, apart from the feet. As long as the player follows the rules of the game, the full-body avatar remains synchronized with the player. Likewise, in the application developed by Debarba et al. [26], the user has to stay seated and is not allowed to move the torso and legs.

Some attempts have been made to reconstruct body movements with the HTC Vive controllers. Seele et al. [32] investigate the impact of avatar eye movement on the perceived quality in a social multiplayer VR application. To track hand movements, both users have to hold two VR controllers each. Thus, the framework tracks only the movements of the upper body. However, the full-body is still synchronized because the users are sitting across a table and do not move their legs. In addition, Jiang et al. [3] introduce a real-time motion reconstruction and recognition method by using the position and orientation of the HMD and two controllers. The upper body reconstruction algorithm is based on Inverse Kinematics (IK) and the lower body is based on animation blending. A machine learning approach (neural network) detects the movements very accurately. The researchers show that their method can reconstruct various full-body motions, even though the lower body animation does not always correspond to the reality, e.g., when the player is walking it was sometimes not possible to determine which foot is in front.

Other recent studies have tried to combine different motion capture devices, markerless as well as marker-based. Loviska et al. [54] use the Oculus Rift DK2 headset, a Leap Motion controller, and an additional camera. The user has to wear this camera which can track a set of markers. Their approach allows multiple users to control a single avatar. However, even though they represent a full-body avatar, only the motions of the user's torso and fingers are tracked and reconstructed.

3.3 Motion Capture Systems using IMUs

An IMU typically consists of an accelerometer (measures acceleration), a gyroscope (measures orientation), and a magnetometer (measures magnetism) [21]. Different tracking systems, such as PrioVR, Perception Neuron or Xsens utilize IMUs, which can be attached to the users' body to track their movements. PrioVR is a body-worn suit with multiple inertial sensors and two additional handheld controllers. Perception Neuron provides a Hi5 VR glove with integrated IMU sensors on the fingers to enable gesture tracking.[12] For positional and orientation tracking of the hand, the user

[12] https://hi5vrglove.com, last visited on March 12rd, 2019

has to fasten a Vive tracker to the wrist. Thus, Perception Neuron combines tracking with IMUs and marker-based technology to track the hand and fingers accurately. Many research publications show that IMUs can provide accurate tracking of the upper and lower body [21], [28]–[30], [57]. Table 4 presents the key findings of each study using IMUs to track body motions.

Integrated sensors of a HMD can also be used to develop a real-time step detector [56]. The step detection algorithm can synchronize the user's feet while walking or running on a treadmill. However, the hands are only synchronized when the user grabs the handle on the treadmill. In another previous work, the authors equipped an ergometer with a tracker and a sensor to capture steering and braking information [55]. Thus, the players' movements are synchronized when they are sitting on the bicycle seat with the feet on the pedals or standing on the floor. Regardless of whether the players are sitting or standing, their hands always have to be on the handlebar. To increase the feeling of presence, the authors also simulate different senses, e.g., audio as well as smell.

Recent works show that motion capture data can also be mapped to a robot, instead of an avatar. Kishore et al. [57] transfer the full-body motion data of a user onto the limbs of a humanoid robot located at a distant position. Thus, one person wearing a HMD and multiple IMUs can interact with the people through the robot in real-time. Since the user wearing a HMD can see the people, a simple interaction such as a hand-shake is possible.

4 DISCUSSION

This paper reviews immersive VR applications developed over the past six years that use different tracking technologies to reconstruct full-body movements. In the following sections, we will summarize key findings, including the strength of evidence for each important outcome. These results should help other researchers gain a deeper understanding of the tracking technology and explore the design space of immersive VR.

4.1 Motion Capture Technologies

In recent years, many advancements in motion tracking technologies have been developed. The developers use these technologies to create immersive VR applications. Figure 2 summarizes the number of studies using different tracking hardware. The markerless motion capture system, in particular, the Kinect device, tends to be the most commonly used (29 out of 53). The second most used motion capture technology is marker-based (17 out of 53), followed by IMUs (7 out of 53).

State of the art research reveals that many studies using a markerless system combine different hardware to enable natural full-body interactions as well as recognition of hand gestures. Many studies use the Kinect as a stand-alone device (n = 14) or in combination with other devices (n = 9), such as Leap Motion [2], [19], [44], Myo armband [15], and other controllers [14], [22], [36], [38], [42]. A few studies even combine multiple Kinect devices (n = 4) to either track a single user [13], [24] or to track multiple users simultaneously [4], [17]. Thus, almost half of the studies (13 out of 27) combine the Kinect sensor with at least one additional tracking device. Conversely, studies using a marker-based motion tracking technology usually do not utilize additional devices. The majority of the studies (13 out of 17) use motion capture suits whereas two studies reconstruct an avatar with only Vive controllers [3], [32]. Furthermore, the studies using IMUs (6 out of 7) typically attach the sensors directly to the body [21], [28]–[30], [57] or the objects, such as a bike [55]. Only one study uses no other sensors except for the embedded inertial sensors of a HMD to reconstruct the movements [56].

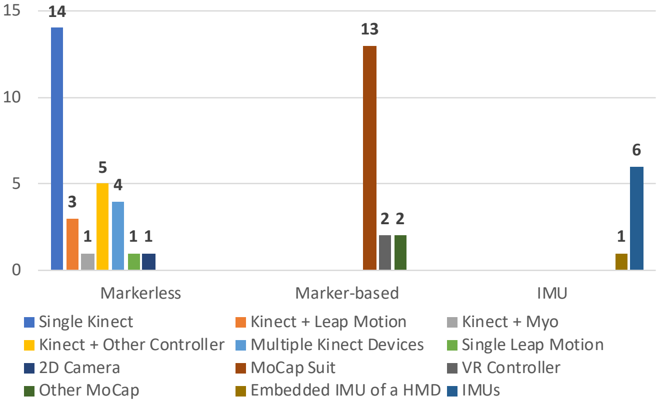

Fig. 2. Overview of tracking technology.

4.1.1 Markerless Motion Capture Systems

Many research publications take advantage of a low-cost solution, such as the Kinect device (27 out of 29). The main drawback of the Kinect device is that it is stable only when the user is facing the sensor or turned away from the sensor [14], [19]–[21]. Because one of the main features of VR is the ability to allow the users to look and move in all directions, the Kinect is not suitable for many immersive VR experiences. When the user wearing a headset is facing to the side, the sensor cannot track all body parts reliably. The sensor cannot handle occlusion when the limbs are out of sight or when one user stands between the sensor and another user. Research by Eubanks et al. [21] suggests using multiple Kinect cameras for a full 360° tracing. However, with numerous devices, tracking problems can occur due to multiple IR sources for depth recognition. Nevertheless, the Kinect device is suitable for applications where the users look only in one direction and as long as they remain in the FoV of the camera, e.g., for flight simulators or for interpersonal skills training.

Additionally, the disadvantage of the Kinect device is that the motion data are inaccurate [42]. In comparison to Kinect v1, the next generation is more precise and suffers less from noise [22]. Unfortunately, the Kinect v2 still has similar tracking issues, including inconsistent tracking [18], jitter problems, and unreliable data [17]. Hoang et al. [45] also found a significant error rate using the Kinect v2 for lower body tracking. These negative characteristics of the Kinect do not get better when combined with a VR headset.

On the contrary, the resulting problems become much more apparent. Due to tracking issues with the Kinect

TABLE 5
Summary of pro and contra for each motion capture system.

Motion Capture System	Pro	Contra	Possible Application Szenario
Markerless	Inexpensive [2], [22], [43], No additional sensors have to be attached to the user's body [3] Simple setup [42], [43]	Cannot handle occlusion [14], [19]–[21] High latency [12], [22], [23] Not precise [42] Inconsistent tracking [17], [18] Jitter Problems [14], [17]	Suitable if the user only looks in one direction and remains in the FoV of the camera, e.g., for flight simulators or for interpersonal skills training.
Marker-based	Real-Time [58], [59] High accuracy [1], [17]	Expensive [9], [17], [21] Intrusive, uncomfortable [17], [27]	Suitable to support a room-scale setup where the player can walk around freely. In addition, motion capture suits with retro-reflective markers are very useful when multiple users must be tracked simultaneously.
IMU	Real-time [28] Inexpensive [29], [30]	Drifting problems [28], [30] No global position [21] Intrusive, uncomfortable [17], [27] Long setup time [21], [27]	Suitable when position data, acceleration, and velocity measurements are needed, e.g., motion recognition for training purposes or rehabilitation.

device, the avatar limbs will jitter, or the user will see movements delayed. In this case, the full-body avatar will not correspond to the user's motions. When the user notices these inaccuracies from the first-person perspective, this can decrease immersion and, in particular, reduce the sense of body ownership. Table 5 summarizes the advantages and disadvantages of the Kinect device.

As previously mentioned, developers often combine the Kinect device with the Leap Motion controller to support natural hand-based gestural interactions. Leap Motion can be mounted in two positions to track hand gestures. It can be positioned directly under the hands; this position works well for desktop applications but is not well suited for immersive VR applications. The users wearing a HMD cannot see the device and will not be able to keep the hands in the tracking area. The users can also mount the Leap Motion controller on the HMD.[13] To track the fingers correctly, the users must therefore continuously look at their hands. Since the users know what the hands are doing without looking at them, they will eventually look away. At this moment, the Leap Motion controller cannot track the fingers anymore. Even though lost tracking is not crucial for motion reconstruction (when the users do not look at the hands, the hands also do not need to be animated), this can be unacceptable when interacting with game objects (the users want to grab and move an object without continually looking at the hands).

4.1.2 Marker-based Motion Capture Systems

Marker-based motion tracking solutions can track full-body motions very accurately and reliably [1], [17]. The OptiTrack system provides a skeletal data stream of the human body with high precision at up to 360 FPS [58]. The solution using multiple Vicon cameras is capable of capturing movements in real-time at 120 Hz with a 16-megapixel resolution [59]. Other solutions achieve an even faster frame rate, however, with reduced resolution. Full-body motion reconstruction requires multiple cameras, usually eight or more. The user has to wear a tight suit with retro-reflective markers. The skeletal 3D model is then created using the data collected by the system. Finally, according to the achieved tracking data, the motions of the user can be mapped to an avatar. Table 5 summarizes the advantages and disadvantages of marker-based motion tracking systems.

Motion capture suits with retro-reflective markers are indeed accurate and very fast. However, the tight suits often create discomfort [17], [27] and are expensive [9], [17], [21]. Previous work has shown that it is not always necessary to track all bone joints of the human body to reconstruct full-body motions. Jiang et al. [3], as well as Seele et al. [32], suggest only to track the hands with HTC Vive controllers. The system can then estimate the skeleton of the upper body by solving the IK problem. Such a solution is very suitable when the number of attached devices should be reduced to increase the user's comfort.

As already discussed in Section 4.1.1, markerless motion capture technologies (e.g., time of flight cameras) often cannot handle occlusion. On the contrary, a technology using multiple markers and cameras provides good tracking results even under partial occlusions. Thus, the user can freely move in a room. Such motion capture systems can be used to support a room-scale setup. A marker-based system is especially suitable when developing multiplayer VR experiences. Due to a large number of markers and cameras, multiple users will not disturb the tracking of one another. Indeed, the state of the art research also revealed that most VR applications enabling multiplayer modus also use marker-based systems (see Section 4.4).

4.1.3 Systems using IMUs

A system that uses IMUs is easy to set up compared to suit-based motion capture systems. Such a system eliminates the need for additional cameras. The user must only fasten multiple sensors on the body. Although IMUs can track motions in real-time [28] and at low-cost [29], [30], the research publications use such a system for full-body motion reconstruction least frequently (7 out of 53).

IMU sensors often suffer from drift and have to be calibrated many times [28], [30]. Such a system does not

---

[13] https://store-us.leapmotion.com/products/universal-vr-mount-pre-order, last visited on October 9th, 2018

TABLE 6
Results of the research publications representing different avatar forms.

Author	Avatar Form	Result
Yan et al. [47]	Realistic avatar vs. stick figure	The users prefer a real human shape over a skeleton model.
Latoschik, Roth et al. [8]	Realistic avatar vs. wooden mannequin	The users rather accept a realistic avatar which invokes significantly higher body ownership.
Bodenheimer and Fu [7]	Realistic avatar vs. stick figure	The researchers found no significant difference between the avatar forms. Nevertheless, the representation of an avatar is essential.
Banakou et al. [10]	Realistic avatar (adult vs. child)	Depending on whether the users have a child or adult avatar, not only their perception but also their behavior changes.

report or measure any global position and cannot be used to track any of the user's physical locomotion or movements through the real world [21]. Table 5 summarizes the advantages and disadvantages of motion capture systems using IMUs.

IMUs are particularly suitable when, in addition to the position data, acceleration or velocity measurements also are needed. Thus, researchers can use such a system for accurate motion recognition, e.g., for step detection [56]. Another possible application scenario could be a serious game for training purposes or rehabilitation where motions or actions should be recognized.

4.2 Avatar Representation

4.2.1 Avatar Form

A small number of studies (n = 4) represent various forms of avatars. The results of studies that evaluate the effects of different avatar representations are detailed in Table 6. The researchers found that a realistic full-body avatar increases the sense of embodiment [8], [33], whereby skin tone can also be important [9].

According to Latoschik, Roth et al. [8] the user prefers to see a realistic virtual body, rather than a wooden mannequin. In their study, each user became an individual, gender-matched avatar created from 3D scans. Since the scanned avatars look very human-like, they evoke a significantly higher acceptance of the virtual body. Yan et al. [47] complement these results. The researchers show that the user prefers a real human shape over a skeleton model. Although Bodenheimer and Fu [7] also evaluated the effect of the presence of different avatar forms, they found no significant difference between the stick figure and full-body avatar. However, they confirm that the representation of an avatar is nevertheless important. Furthermore, Banakou et al. [10] discovered that the virtual body size could affect the perception of the object. The researchers found that when users experience the virtual world as a child avatar, they will overestimate the virtual object sizes.

Other recent studies have attempted to show advantages of a system for rapid acquisition of animated, full-body avatars [30]. The users are first scanned so that a personalized avatar can be created. A high texture fidelity can influence the user's ability to recognize an avatar [60]. Other recent studies evaluated the impact of having a personalized avatar in a video game (and not in an immersive VR game). The results suggest that having such an avatar improves the subjective experience [61]. Even though the users feel more connected and engaged, there is no significant effect on how the user performs. The work of Wauck et al. [62] further asserts that the appearance of the avatar does not affect the user's performance or subjective experience. The authors themselves suggest that immersive VR can increase the player's sense of presence; especially when the player can control the avatar directly via body movements rather than by keyboard.

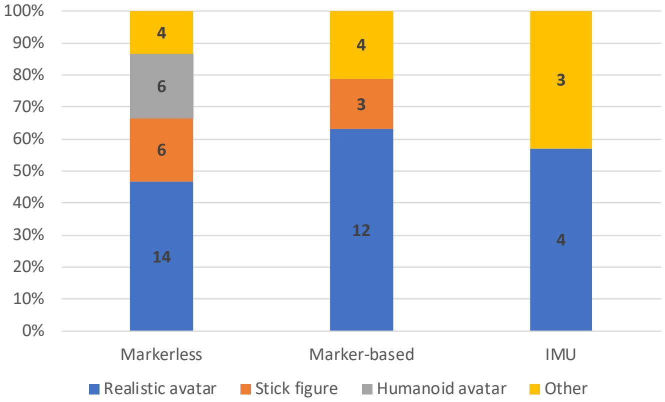

Fig. 3. Comparison of different avatar forms regarding tracking technology. Note that one markerless study and three marker-based studies represent two different avatar forms.

The analysis has yielded some interesting results, revealing that more than a half of the studies using a marker-based motion capture system (12 out of 17) or IMUs (4 out of 7) represent a realistic avatar. Conversely, studies using a markerless motion tracking system often show a stick figure (n = 6) or a humanoid avatar (n = 6). Because the Kinect SDK already provides a stick figure, the developers like to use them. Besides, the Leap Motion includes some pre-made hand models for the Unity game engine, e.g., a robot hand, a hand presented with cubes, and a human hand with realistic skin texture.[14] In addition, movements of a user can also be mapped to a humanoid robot and not to an avatar [57]. Figure 3 summarizes the avatar representation, regarding the different tracking technology.

The state of the art research suggests that there is a focus on using a realistic avatar (n = 30), rather than a stick figure (n = 9) or even a humanoid avatar (n = 6). Since there are some contradictory results regarding the avatar form, further evaluations are required. Future research should also focus on advanced sensors, to measure the sense of presence, e.g., using biosensors such as EEG and ECG, measuring skin temperature, eye-blink rate, pulse.

4.2.2 Full-Body vs. Partial-Body Representation

Among others, several studies focus on attempting to evaluate the effect of different avatar representations, i.e., full-

---

[14] https://developer-archive.leapmotion.com/documentation/v2/unity/unity/Unity_Hand_Assets.html, last visited on October 11th, 2018

TABLE 7
The results of research publications comparing a full-body avatar with a partial-body representation.

Author	Comparing	Result
Mendes et al. [15]	Full-body avatar vs. simplified hands	Full-body avatar improves the sense of presence; however, simplified hands-only representation is more effective in performing tasks.
Young et al. [6]	Full-body avatar vs. floating hands	Results suggest a significantly higher presence for the full-body avatar over hands-only representation.
Schäfer et al. [22]	Full-body avatar vs. no avatar	Body tracking and representation of a full-body avatar increase the sense of presence.
Kondo et al. [16]	Full-body avatar vs. feet and hands	Visual hands and feet are sufficient to induce illusory body ownership. This effect is as strong as using a full-body avatar.

body avatar and hands-only representation. Table 7 summarizes the results of the research publications comparing a full-body avatar with a partial-body representation.

Some of the studies included in this survey suggest that a representation of a full-body avatar in contrast to hands-only will significantly improve the sense of presence within the VR [6], [15]. Makled et al. [63] argue that full-body animations are much more important than head animations. Conversely, research by Kondo et al. [16] suggests that a full-body avatar may not be necessary. The authors state that representing only hands and feet can be sufficient to induce body ownership illusion. Mendes et al. [15] also note that the simplified hands-only representation is more effective in performing tasks. Because these results are contradictory, more research needs to be carried out in this area. In particular, more evaluations and comparative studies investigating the effect of full-body avatars in immersive VR are necessary.

It is already well known that the sense of agency and illusion of body ownership will be increased when the users can see a virtual body from a first-person perspective [9], [10], [51], [64]. The virtual movements must match the actual physical movements. Williams et al. [33] observe that tracking and visualizing the movements of the lips also can significantly improve the level of embodiment. Debarba et al. [26] state that bigger FoV (which causes to see the virtual body more often) does not affect the sense of agency. However, when the movements of an avatar are either shown in the past (high latency) or future (prediction), the user is not able to identify with the virtual avatar [1].

All VR applications included in this survey track the body movements and represent these motions through an avatar or a robot. Most studies track only gross body movements to match the position and orientation of the extremities. However, some recent studies track not only arms and feet but also hand gestures. VR applications that use a markerless motion capture system focus on visualizing a full-body avatar with the individual fingers [2], [17], [19], [44]. If the hand gestures should be recognized, one can primarily use a markerless system such as a Leap Motion controller. Another possibility would be to use data gloves (see also Section 4.3.2). Additionally, some attempts have been made to detect eye movements using a HMD with an integrated eye tracker [32].

It is not always necessary to track the full-body movements to reconstruct a full-body avatar. Some studies even limit the motions of some body parts. For example, in some games, the players can only freely move the upper body while their feet must remain at the same position during the entire gameplay [40], [51]. In addition, as described by Headleand et al. [48], in their study the user sits in a wheelchair and can move the hands. Because the users cannot move their legs, the system only tracks their upper body. The users can use the Xbox controller to steer the wheelchair and a Leap Motion controller to interact naturally with the virtual environment. Although only some of the body movements are tracked, a full-body avatar is still synchronized with the user. Despite these body movement limitations, the feeling of the embodiment is not affected.

4.3 Level of Immersion

Immersive VR is characterized by interaction between the user and the virtual environment. Slater and Wilbur [5] define immersion as the extent to which a system can deliver an illusion of reality to the user. Thus, an ideal system would display in all sensory systems.

Only a few studies stimulated multiple human senses, such as a sense of touch and smell. Table 8 classifies the different feedback modalities used in the included studies, e.g., visual, auditory, tactile, and olfactory. All VR applications included in this survey provided the user with a sense of agency by visualizing a full-body avatar. In other words, all studies provided synchronous visual-tactile stimulation over the body parts. As shown in Figure 4, four studies using markerless tracking technology include sound effects. Studies using a marker-based motion capture system often add auditory (n = 2), tactile (n = 5) or auditory and tactile (n = 2) modalities. Only one study using IMUs includes the sense of smell as feedback modality in their system.

To improve the sense of presence in a virtual environment, the applications should also stimulate other human senses. There is a considerable potential to improve immersion in VR, e.g., using air fans for wind simulations and using smell dispenser to stimulate different flavors (sweet, sour, bitter). Various sensors can be used to stimulate the mediation of the temperature (warm, cold) or vibration sensors to stimulate pain.

4.3.1 Auditory

Auditory output (apart from visual, which is included by all research publications in this survey) is the most common (n = 11) way to increase immersion. To provide realistic 3D audio feedback of the environment, the developers usually use headphones. The modern HMDs, such as Oculus Rift CV1 and HTC Vive Pro already have integrated 360° spatial

TABLE 8
Level of immersion.

Feedback Modality	
Auditory	Headphones [2], [8], [22], [28], [38], [51], [55], [57]
	Surround sound system [25]
	Unknown [18], [50]
Tactile	Passive feedback [28], [32], [49]
	Vibrotactile devices [25], [51]
	Force Feedback [52]
	Interacting with other users [1], [6]
Olfactory	Smell dispenser [55]

sound, which makes it easier for the developers to deliver immersive audio. One study uses a surround sound system to play transmitted audio streams from other users, who were not in the same room [25]. Other studies use headphones for audio instructions [8], [22], to enable voice communication between multiple users [25], [28], [57] or for sound effects to enable a more realistic impression [2].

Most VR applications focus on providing sound effects. Noise canceling headphones lower the attention focused on the real environment and thus leads to a higher presence [22], [55]. Developers should consider using sounds and music to achieve a higher level of immersion. Especially ambient sound (e.g., sound of the wind, birds chirp) can give the user a sensation of being part of the virtual environment.

4.3.2 Tactile

The sense of touch can be stimulated using real physical objects (passive haptics). Chagué and Charbonnier [49] demonstrate how users can naturally interact with virtual objects while touching them in the real world. Therefore, the researchers place optical markers on the objects to track them. The player can interact with the objects in the virtual environment while touching the object in the real world. For realistic tactile feedback, the two objects (the virtual and the real one) must have a similar size and surface.

Furthermore, vibrotactile devices fastened to the body can be used to create haptic feedback, e.g., either to generate vibrotactile feedback when one user virtually touches another user [25] or to allow the user to physically feel when the virtual objects hit them [51]. Vibrotactile solutions include different VR devices, such as HTC Vive controller [32]. When motion reconstruction is accurate enough, the users do not need additional devices at all. For example, in co-located multiplayer VR experiences the users can naturally interact with each other [1], [6]. They can touch the virtual body and thus other users.

For more realistic haptic feedback, Schmidt et al. [52] developed special shoes to experience an actual force when walking up and down the stairs. To enable interaction with virtual objects, recently many commercial haptic gloves have been proposed, such as AvatarVR[15], VRgluv[16], HaptX[17], and Plexus[18]. Other commercial solutions directly attach small devices on the fingertips to interact and manipulate VR objects.[19] Furthermore, commercial solutions for full-body haptics were recently introduced, such as Teslasuit.[20]

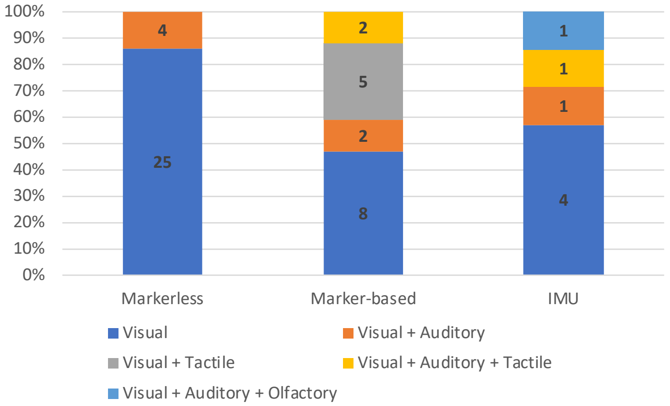

Fig. 4. Feedback modality.

Even though many commercial solutions providing a realistic sense of touch exist, only a few studies focus on using this haptic technology (n = 8). Tactile feedback can significantly increase the level of physical possession and agency. Hence, novel haptic devices can increase the feeling of presence in VR. Egeberg et al. [51] report that visuotactile feedback (user can simultaneously see and feel touch) enhances ownership significantly. Conversely, the results of the study by Bourdin et al. [25] suggest that tactile intervention has no impact on the sensation of body ownership, the presence or behavior. Because these results are contradictory, more research needs to be done in this area. Due to recent improvements in haptic technology, we hope that further studies will investigate the effect of touch in VR.

4.3.3 Olfactory

Only one study, which developed a cycling-based exergame, stimulated user's vision, sense of hearing, and sense of smell. Melo et al. [55] increased the feeling of presence by using a smell dispenser so that the user can smell lavender when passing the fields. Recently, Feelreal presented a sensory mask which can simulate hundreds of smells to immerse the player into a virtual world.[21]

4.4 Number of Players

There is a growing trend of creating true multiplayer VR applications. A true multiplayer application tracks the body movements of multiple users simultaneously. All these users wear a HMD and can see each other. In other multiplayer VR applications, only one user can view the virtual environment via a HMD. Other users without a HMD (non-HMD users) can influence the application from outside.

---

[15] https://www.neurodigital.es/, last visited on October12th, 2018
[16] https://vrgluv.com, last visited on October 12th, 2018
[17] https://haptx.com, last visited on October 12th, 2018
[18] http://plexus.im, last visited on January 30th, 2019
[19] https://www.gotouchvr.com/, last visited on October 12th, 2018
[20] https://teslasuit.io, last visited on April 26th, 2018
[21] https://feelreal.com, last visited on October 12th, 2018

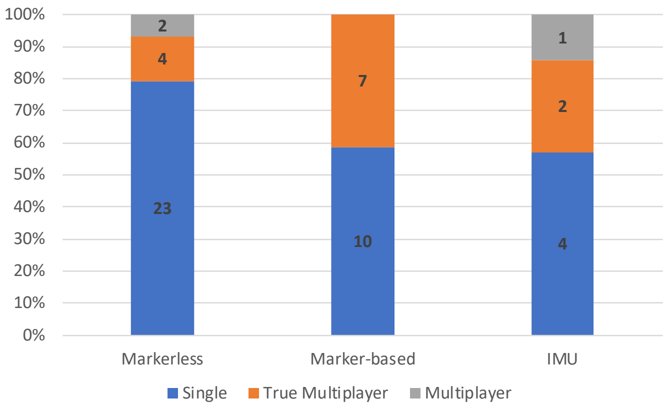

Fig. 5. Overview of the research publications enabling multiplayer mode. In a true multiplayer VR application, multiple users are wearing a HMD. The application tracks the body movements of all these users. In other multiplayer applications, only one user is wearing a HMD and the non-HMD users can influence the VR application from outside.

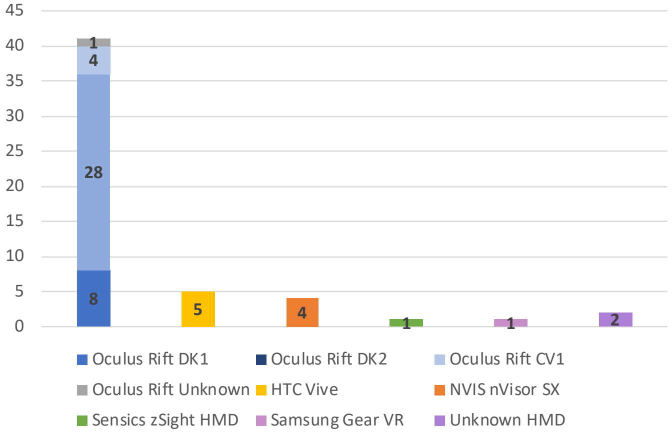

Fig. 6. The Oculus Rift is the most frequently used HMD, followed by HTC Vive and NVIS nVisor SX.

4.4.1 *True Multiplayer Applications*

One-fourth (13 out of 53) of research publications included in this survey developed a true multiplayer VR application. As shown in Figure 5, almost half of the studies using a marker-based motion capture system developed a multiplayer VR experience (7 out of 17). Only four studies using a markerless motion capture system [4], [14], [17], [45] and two using IMUs [28], [29] developed a collaborative virtual environment. Most multiplayer applications (n = 9) allow co-located users to interact with each other in a shared virtual environment [1], [4], [6], [17], [28], [29], [32], [49], [54]. In some minor multiplayer VR applications (n = 2), the users are located in separate rooms [25], [45].

When developing true multiplayer VR games, it is essential to reconstruct the full-body motions in real-time. The players wearing a HMD cannot see the real world and, in particular, not the other players. Therefore, it is necessary to track the full-body motions of all players to visualize virtual avatars. By presenting a full-body avatar to each player, a team can communicate and interact much better with each other, e.g., through the recognition of different gestures [3] or eye-tracking [32].

One of the main challenges when developing multiplayer games is latency. The VR system must distribute a large amount of tracking data of each user while keeping perceived latency as low as possible. When latency is noticeable, it can negatively affect the gameplay experience [29]. The results of our analysis suggest that many multiplayer VR applications are using a marker-based system (7 out of 13) because such a system provides tracking data in real-time and is very accurate. Thus, developers should consider using marker-based systems when developing true multiplayer games to track multiple users simultaneously.

4.4.2 *Multiplayer Applications that Include non-HMD Users*

Only a small number of applications (3 out of 53) include non-HMD users into the VR experience. Such applications create an entirely different VR experience. Two studies developed a multiplayer application, where one user can view the virtual environment via HMD. The motion tracking system tracks only the body movements of this one player. Other users can interact with the player through a smartphone [40] or a desktop application [39]. Thus, these users do not wear a HMD and their movements are not tracked.

One further study developed a framework where a user wearing a HMD can interact with people at a distant location [57]. The system tracks body movements of this one user and transfers them onto the limbs of a robot. The user can see the people via the HMD and can thus interact with them via the robot.

4.5 Head-Mounted Display Technology

Figure 6 summarizes the utilization of different HMDs. Note that one research publication used two different HMDs, Oculus Rift DK2 and HTC Vive. We classified the involved headsets into three categories:

- Oculus Rift (n = 41): including DK1, DK2, and CV1
- HTC Vive (n = 5)
- Other: NVIS nVisor SX (n = 4), Sensics zSight HMD (n = 1) and Samsung Gear VR (n = 1). Two research publications did not specify which HMD they used in their research.

Oculus Rift DK2 is by far the most frequently used HMD (n = 28). However, it does not support 360° tracking, has limited tracking space and no motion controller. Modern HMDs, such as Oculus Rift CV1 and HTC Vive are used less frequently. Both HMDs have the same technical specifications and rely on an embedded IR tracking system to enable positional and rotational tracking. HTC Vive requires two Lighthouse sensors to track the Vive devices in a big room up to 15 × 15 feet whereas the Oculus Rift CV1 requires even three sensors to provide a 360° tracking.[22] However, Oculus Rift CV1 and HTC Vive were both first released in 2016. Because the majority of the studies were published in 2016, both HMDs were not always yet available.

One of the main advantages of the HTC Vive is room-scale tracking. This feature is particularly beneficial for

[22] https://www.digitaltrends.com/virtual-reality/oculus-rift-vs-htc-vive/, last visited on October 9th, 2018

TABLE 9
Summary of the research publications using a wireless solution.

Author	Nr. of Players	Main Findings
Greuter and Roberts [43]	Single	Backpack with a tablet computer
Yan et al. [47]	Single	Wearable transmitting system
Schmidt et al. [52]	Single	Backpack with a laptop (carried on the front)
Eubanks et al. [21]	Single	Backpack with a laptop; theoretically infinite tracking area
Chagué and Charbonnier [49]	Multiplayer	Backpack with a laptop
Podkosova et al. [28]	Multiplayer	Backpack with a laptop; huge tracking area ($200\,\mathrm{m}^2$)
Leoncini et al. [17]	Multiplayer	Client-server network
Johnson et al. [29]	Multiplayer	Client-server network; mobile HMD

creating immersive VR experiences, where the user can walk around to explore the virtual environment freely. Additionally, users can use VR controllers to interact with game objects. As already previously mentioned, only two studies [3], [32] used Vive controllers to reconstruct the body motions and to visualize a full-body avatar. To the best of our knowledge, no research publication uses an Oculus Touch controller to enable full-body tracking.

4.5.1 Wireless vs. Wired

One of the biggest limitations of the HMDs are the attached cables that break the sense of presence and restrict the freedom of movement. The two most commonly used HMDs in this survey, Oculus Rift and HTC Vive, are both wired. Especially in co-located multiplayer games, the cables attached to the various devices affect health and safety. The users will eventually stumble over cables belonging to the HMD they are wearing. More likely, they will stumble over cables belonging to the other users. In particular, the cable attached to the HMD, can also break the sense of presence and significantly limit the freedom of movement [49]. One option is to attach the cables to the ceiling so that there are no cables on the floor. However, such a structure can be complicated.

A small number of studies (n = 8) suggest wireless solutions to enable VR experiences without disturbing cables. As it can be seen in Table 9, several studies used a backpack, where a laptop or tablet computer is connected to a VR headset [21], [43], [47], [52]. Multiplayer applications additionally have to distribute the tracking data to the server and other clients over a wireless network [28], [49]. With a mobile HMD, such as Samsung Gear VR, a wireless solution is ensured [29].

Wireless solutions significantly reduce trip hazards for the players [43]. For example, Yan et al. [47] show that wireless solutions can be used to practice dance movements in VR. With a wireless solution, the developers can thus increase the freedom of movement. Wireless solutions are especially advantageous to enable collaborative experiences in a large virtual environment [28]. Half of the studies using a wireless solution (4 out of 8) are also a multiplayer VR application.

The analysis has yielded some interesting results, revealing that many multiplayer VR applications use a wireless solution. Hence, the results suggest that there is a trend of developing wireless solutions, mainly when supporting multiplayer mode. The developers should consider using a wireless solution when they aim to achieve a 360° free movement. Recent advances in VR technologies have already contributed to the advancement of HMDs. For example, HTC Vive recently released wireless adapters that can be used for Vive and Vive Pro HMDs.[23] Likewise, Oculus announced in 2019 a new HMD (Oculus Quest), that is standalone.[24] This improvement in VR technology makes it easier for developers to create multiplayer VR games. We believe that in the future more VR games will be developed where the user can freely move around without disturbing cables.

4.6 First vs. Third-Person Perspective

HMDs can show multiple viewpoints to the user, including a first- or third-person perspective. On the one side, from the first-person perspective, the users can look down and see their virtual body as they would in the real world. Previous work already validated that the first-person perspective enhances the sense of embodiment [9], [10], [33]. On the other side, from the third-person perspective, the users will see the full-body avatar which is positioned somewhere in the virtual world [47]. As in common non-immersive virtual environments, the avatar can be placed in front of the viewer, so that the user can observe it from behind. The third-person perspective can also reduce the probability that cybersickness will occur [65]. However, Maselli et al. [66] recently found that looking at the avatar from the third-perspective will inhibit the illusion of body ownership.

The majority of studies included in this survey (46 out of 53) visualize the full-body avatar only from the first-person perspective. Two further studies represented the avatar from the third-person perspective. Yan et al. [47] developed a dance-based exergame in which the dancers can observe themselves from external self-image. In another study, the users can see the avatar 2 m in front of and facing away from them in a virtual room [16].

One further publication evaluated the effect of both perspectives. Hoang et al. [45] show the significant advantage of the first-person over the third-person perspective in VR. In this study, the users with the HMD must imitate a virtual instructor performing different poses, which they can see either from the first- or third-person view. Body postures are either prerecorded or shown by an expert in real-time. Thus, the users can see through a HMD their movements superimposed with the movements of the instructor. However, additional results of the study indicate that the mean completion time is significantly higher for the users who can see the avatar from the first- compared to the third-person view. Thus, viewing the avatar from the first-person view

---

[23] https://www.vive.com/us/wireless-adapter/, last visited on January 22th, 2018

[24] https://www.oculus.com/quest/, last visited on January 22th, 2019

takes more time. The authors also found that the first-person perspective does not receive high preference ranking, compared to more traditional methods such as video calls (e.g., Skype). This result could have been specific to the tasks or the poor quality of the Kinect tracking technology.

Thomas et al. [50] conducted a similar study comparing different perspectives. The findings suggest that the tasks will be performed differently when the user sees the avatar from a first- or third-person perspective. Here, the user can view the avatar from the first-person when wearing a HMD and from the third-person when watching a 3-D TV. Especially when the movements of the player cannot be translated into the VR accurately and in real-time, the third-person perspective can be beneficial.

These results suggest that during learning, it is not always necessary to create a first-person experience to be effective. In a 2016 study, the outcomes nevertheless showed that the users prefer a first-person over the third-person perspective [45]. Desai et al. [13] have expressed a similar view. In their study, however, the researchers investigated an effect of visual quality between the first-person view on a HMD compared to the third-person view on a 3-D TV.

The analysis regarding the perspective shows that there is a strong focus on visualizing the avatar from the first-person view when developing immersive VR experiences. However, according to [45] and [47], the third-person view can be beneficial for VR applications where users have to learn a particular movement. The users can then see their movements in a virtual mirror or from behind. For other VR applications, especially those requiring interaction between users, a first-person view should be used. The state of the art research also revealed that in most multiplayer VR applications the users could view their virtual body from the first-person view and can thus naturally interact with other users.

4.7 Latency

End-to-end latency refers to the total time between the physical movement and the update of the corresponding image that reflects that movement [67]. In other words, latency indicates the time between a user's motion and the time when this motion is visible on the HMD. High latency of the HMD can cause cybersickness [68]. Additionally, a significant delay between a physical movement and an output image can negatively affect the sense of presence and will thus reduce immersion [69].

Only one-third (18 out of 53) of the research publications measured end-to-end latency or specified at least the FPS at which the VR application runs. Most studies measured a very high value of 39 ms and above. Only one study could reconstruct the full-body motions in 7 ms [3]. For a real-time application and especially for VR, the delay should not exceed 20 ms [70].

Previous work has already shown that high latency will reduce the sense of body ownership. In the study conducted by Kasahara et al. [1], movements of the avatar body were intentionally represented in the past or future. The authors make clear that body ownership disconnects when the users perceive the discrepancy between their body and the virtual body. The users notice a delay between real and virtual movements at approximately 30 ms. When the virtual movements are presented 25 to 100 ms in the future, the users perceive their virtual avatar as lighter. In both cases (movements presented in the past or future), the sense of ownership could not be preserved.

4.7.1 Latency of Markerless Motion Capture Systems

Because the Kinect device runs at 30 FPS[25], tracking with this device results in higher end-to-end latency compared to expensive motion capture suits. This low framerate is only poorly compatible with VR applications which require to run at 90 FPS.[26] The latency of the Kinect v2 in combination with a VR headset has been reported to be even higher, between 60 and 80 ms.[27] In general, combining a large sensing space with high accuracy leads to substantially higher latency. For example, end-to-end latency of Oculus Rift and Leap Motion is as high as 125 ms, whereas end-to-end latency of Oculus Rift and Kinect is 169 ms [12]. The VR latency should ideally not be more than 11.11 ms (90 Hz) since this would satisfy the refresh rate of the HTC Vive and Oculus CV1. Using the Kinect sensor will eventually result in non-instantaneous responses, which leads to negative feedback about the controls [23]. The low framerate will not only deteriorate the VR experience but will also eventually cause cybersickness [22], [23].

Due to its high latency, the Kinect device is often not sufficient for full-body reconstruction in immersive VR. Even though the most research publications using the Kinect device claim that their application can track and visualize the full-body avatar in real-time and with only a low delay, the fact that the Kinect runs at 30 FPS cannot confirm this assumption. The latency of the Kinect is way too high to satisfy the requirements of the immersive VR. As already discussed before, the VR applications need to maintain 90 FPS. Tracking with Kinect can still be suitable, when used together with another device, such as HTC Vive controllers. These controllers are much more responsive and can be used to track the hands. Using sensor fusion strategies to combine full-body tracking along with hand-tracking can be a solution.

4.7.2 Latency of Marker-based Motion Capture Systems

Other research publications use marker-based motion capture systems with accurate position and orientation tracking. The commercial systems, such as OptiTrack and Vicon promise imperceptible latency of below 10 ms [58], [59]. However, most authors using motion capture suits nevertheless report a higher end-to-end delay, between 39 and 150 ms [1], [8], [50], [52]. In other words, most applications that use expensive marker-based systems achieved different results, most of them above the ideal value. Only one publication, using two Vive controllers to reconstruct full-body motions reported latency of below 7 ms [3].

[25] http://www.imaginativeuniversal.com/blog/2014/03/05/quick-reference-kinect-1-vs-kinect-2/, last visited on October 10th, 2018

[26] https://www.newegg.com/vr/guides/find-out-if-your-pc-can-handle-vr.html, last visited on October 10th, 2018

[27] https://www.reddit.com/r/oculus/comments/2au31o/kinect2_latency_is_6080ms_video_included/, last visited on October 10th, 2018

TABLE 10
Summary of research publications reporting the level of cybersickness.

Author	**Reported cybersickness**	**Condition**
Shaw et al. [23]	Increased (severe nausea)	Sitting, driving a bicycle
Debarba et al. [26]	No	Sitting
Schäfer et al. [22]	Low	Standing (teleporting)
Latoschik, Lugrin et al. [39]	Unknown	Standing in place
Eubanks et al. [21]	Low (general discomfort and fullness of head)	Standing (natural walking)

The analysis on tracking in VR revealed that marker-based motion capture systems are very well suited to reconstruct motions in real-time. These systems, in particular, can be used for multiplayer VR experiences. When multiple users have to interact with each other, it is essential that the users perceive little to no delay. Since latency is crucial for immersive VR experiences, it is essential to keep this aspect in mind. The developers should consider the time at which the motion occurs until the corresponding movement is mapped onto the avatar and displayed on the HMD. Research by Jiang et al. [3] suggests reducing latency by pre-processing sensor data. Likewise, Friðriksson et al. [14] propose using motion prediction algorithms to minimize delays. In the future, developers and researchers should focus on proposing more strategies to reduce latency in immersive VR applications.

4.7.3 Latency of Systems using IMUs

Only one study using IMUs measured end-to-end latency [29]. According to the authors, end-to-end latency of approximately 300 ms was reported as noticeable and had a negative impact on the gaming experience. The results suggest that a solution using a smartphone-based HMD and Perception Neuron motion capture system is not sufficient for real-time full-body motion reconstruction in VR.

In a study by Podkosova et al. [28], the authors did not explicitly measure latency; they nevertheless specified the rate at which the avatars were synchronized between the users (60 FPS). The researchers conclude that the interactions between the user and virtual objects did not reveal noticeable latency between haptic feedback and graphics.

4.8 Cybersickness

A minority (5 out of 53) studies included in this survey focused on evaluating cybersickness. Only one study reported increased symptoms in which one participant had to abort the evaluation due to severe nausea [23]. The evaluation results of cybersickness are detailed in Table 10.

4.8.1 Sitting Condition

Increased cybersickness was reported in a cycling-based exergame in which the player was cycling down ramps or falling down pits [23]. When driving uphill, the players mentioned no discomfort. According to the researchers, this is because the players feel in control when driving upwards. Additionally, the authors report too high latency and note that the full-body motion controls in VR are not responsive enough. Garcia et al. [68] argue that the high latency of the HMD can cause cybersickness. Because the authors use Oculus Rift DK1 (60 Hz) and Kinect v1 (30 FPS), this can be the reason for the high latency and thus the increase in cybersickness.

In contrast, Debarba et al. [26] report no cybersickness. In this study, the participants were sitting. Compared to [23], the VR experience in [26] did not force any movement upon the user. The researchers evaluated the impact of FoV (106° or 180° vertical FoV) on cybersickness and found no significant difference [26]. Thus, bigger vertical FoV does not increase the probability of cybersickness.

4.8.2 Standing Condition

None of the research publications allowing the participants to stand and walk around reported increased cybersickness. Two studies reported only low cybersickness [21], [22]. The developers can alleviate cybersickness by allowing the user to walk in the physical space naturally (see also Section 4.8.3). If the users in VR can see and feel the same motion, the probability that cybersickness occurs is very low. In addition, teleportation can also reduce the probability of cybersickness [22]. When teleporting, the users can move by pointing to some location in the virtual environment. The users can select the desired position by pressing a button and immediately teleport there when the button is released.

4.8.3 Reducing cybersickness

Capturing of user's motions and representing a full-body avatar not only enhance immersion but also reduce cybersickness [65]. Maintaining low latency will diminish the probability of discomfort caused by cybersickness. Cybersickness can also be reduced when the VR game is played from the third-person perspective [65]. From the third-person perspective, the user wearing a HMD can cover a large virtual environment with a much smaller movement in the physical space. Because the user does not necessarily have to move much, the probability that cybersickness will occur is much smaller. For example, for the low sensory conflict condition, the developers of a very popular VR game Lucky's Tale[28] represent the character from the third-person perspective. Additionally, some attempts have been made to add a virtual nose to reduce cybersickness.[29]

VR experiences in which the users remain seated, however, their virtual position moves, are more likely to induce nausea symptoms. When the movement is forced upon the user, the physical and virtual movements do not match. This problem is also known as *the sensory conflict theory* [71]. The cybersickness will especially occur when the user does not use head movements to rotate the virtual scene, but some input device, e.g., a keyboard, a mouse or a controller. One possibility is to allow the user to stand and to walk naturally. When the users can see and feel the same movement, the probability that cybersickness occurs is much lower.

---

[28] https://www.oculus.com/experiences/rift/909129545868758/, last visited on October 10th, 2018

[29] https://www.wired.com/2015/04/reduce-vr-sickness-just-add-virtual-nose/, last visited on October 10th, 2018

TABLE 11
Summary of the evaluation studies. Abbreviations: *Tech.* Technical evaluation, *Usab.* Usability test, *UX* User Experience

Author	**Evaluation Method**		**Number of Subjects**
Charles [20]	Tech.	The accuracy of the Kinect device	
Headleand et al. [48]	Tech.	Validation study	3 (experts)
Hilfert and König [2]	Tech.	Performance test	
Leoncini et al. [17]	Tech.	Validation study of the proposed approach	
Rhodin et al. [27]	Tech.	Visibility test of body parts for different camera angle views (Kinect vs. GoPro vs. fisheye camera), runtime measurements (FPS); detection and pose accuracy	8
Reus et al. [42]	Tech.	Evaluating drifting and head-body misalignment	
Jiang et al. [3]	Tech.	Latency evaluation; the accuracy of a natural action recognition algorithm based on a neural network to reconstruct a lower body	3
Thomas et al. [50]	Tech.	Technical evaluation	17
Podkosova et al. [28]	Tech.	Evaluation of jitter, update rate, network performance	
Malleson et al. [30]	Tech.	Measuring processing time to create a personalized avatar	12
Johnson et al. [29]	Tech.	Measuring tracking error for the proposed technique and end-to-end latency	
Caserman et al. [56]	Tech.	The accuracy of a step detector to synchronize feet with the virtual avatar while walking on a treadmill	2
Matsas et al. [18]	Usab.	Efficiency evaluation	30
Lee et al. [35]	Usab.	Usability test (e.g., how intuitive and easily the user can personalize furniture using gestures and voice commands)	10
Mendes et al. [15]	Usab.	Effectiveness of the avatar representation (full-body avatar vs. simplified hands)	24
Otto et al. [24]	Usab.	Usability test	22 (experts)
Eubanks et al. [21]	Usab.	Usability test (evaluating cybersickness)	4 (men)
Tuveri et al. [38]	Usab.	The effectiveness of gamification techniques in immersive VR environments for fitness	22
Chou et al. [40]	UX	Gameplay evaluation	9
Collingwoode-Williang et al. [33]	UX	Evaluating the sense of agency and body ownership	40
Desai et al. [13]	UX	User experience (e.g., gameplay experience) evaluation	25
Egeberg et al. [51]	UX	Evaluating the sense of agency and body ownership as well as the gameplay experience	30
Bodenheimer and Fu [7]	UX	Evaluation of the effect and presence of different avatar forms (realistic avatar vs. stick figure)	18
Seele et al. [32]	UX	Evaluation of the sense of presence	42
Banakou et al. [10]	UX	Evaluating the sense of agency and body ownership (adult vs. child avatar)	32
Peck et al. [9]	UX	Evaluating the sense of agency and body ownership	60 (women)
Bourdin et al. [25]	UX	Body ownership study	44
Kondo et al. [16]	UX	Body ownership study (full-body avatar vs. feet and hands)	50
Kishore et al. [57]	UX	User experience (e.g., experience with the robot) evaluation	6
Yan et al. [47]	Usab., UX	The effectiveness of the training (compared to the regular training); user experience evaluation (e.g., realistic avatar vs. stick figure)	8
Schäfer et al. [22]	Usab., UX	Usability test (effect of the body tracking with Kinect, evaluating cybersickness); evaluation of the sense of presence (full-body avatar vs. no avatar)	42
Melo et al. [55]	Usab., UX	Evaluating the impact of the body position on the usability of VR environments (e.g., measuring completion time); user experience evaluation (e.g., satisfaction)	47
Kasahara et al. [1]	Tech., UX	End-to-end latency evaluation; evaluating the sense of agency and body ownership	29
Latoschik, Roth et al. [8]	Tech., UX	End-to-end latency evaluation; body ownership study (comparing different avatar representations, realistic avatar vs. wooden mannequin)	20
Schmidt et al. [52]	Tech., UX	Validation test (also latency evaluation); gameplay evaluation (e.g., how enjoyable is the VR experience, how realistic is the proposed method)	12
Hoang et al. [45]	Tech., UX	Performance evaluation regarding posture accuracy and user's preference (first-person vs. third-person view)	23
Shaw et al. [23]	Tech., Usab.	A technical evaluation comparing four tracking methods; evaluating cybersickness	
Debarba et al. [26]	Tech., Usab., UX	A technical evaluation comparing different FoV (e.g., time and precision); evaluating the impact of FoV on cybersickness; evaluating the sense of agency and body ownership	6
Young et al. [6]	Tech., Usab., UX	Accuracy measurement (e.g., offset distance); effect of the avatar representation (full-body avatar vs. hands-only); evaluation of the sense of presence	16
Latoschik, Lugrin et al. [39]	Tech., Usab., UX	End-to-end latency evaluation; usability test (e.g., intuitive use and task load, performance, cybersickness); evaluation of the sense of presence	

Another possibility to reduce cybersickness is to use real-time motion capture technology. In future work, the researchers should also focus on evaluating the correlation between cybersickness and personal factors, such as age, gender, and experience playing 3D (immersive) video games.

4.9 System Evaluation

The included publications are classified into four groups, based on their evaluation status: a) no evaluation (n = 13); b) technical evaluation (n = 12); c) usability test (n = 6), and d) user experience evaluation (n = 11). Three studies evaluated usability and user experience. Four further studies conducted a technical and a user experience evaluation. One study performed a technical evaluation and a usability test whereas three studies conducted all. Table 11 illustrates the systems evaluation status in detail.

The sample size, especially in the studies evaluating either user experience or usability, ranges from 2 to 50, whereas one study [9] involved even 60 subjects. As one can see in Table 11, almost all studies exceed the minimal sample size. Virzi [72], for example, points out that 80% of the usability problems will be detected with four or five participants. Faulkner [73] suggests that at least ten subjects are needed, and with 20 subjects the percentage of problems revealed will be increased to 95%.

The analysis in this survey also reveals that only two studies evaluated with experts: one study did a usability test [24] and another study did a technical evaluation [48]. Other studies were conducted with students whereas one study involved only women [9] and another study only men [21].

4.9.1 Usability evaluation

The usability evaluation was conducted with regard to the following aspects: intuitiveness, effectiveness, and efficiency. Lee et al. [35] evaluated how intuitive and easily the user can use gestures (tracked by a Kinect device) and voice commands to personalize furniture. Latoschik, Lugrin et al. [39] also evaluated intuitive use of their system. Further studies evaluated the effectiveness of the VR-based training compared to the regular training [47], the effect of the therapy using a markerless motion tracking technology [22], and the effectiveness of gamification techniques in immersive VR for fitness [38]. Matsas et al. [18] evaluated efficiency, whereby the users had to follow the moving virtual hand, which defines the movement pattern of the task. Furthermore, two studies [6], [15] evaluated the effect of the avatar representation by comparing a full-body avatar and hands-only representation (see also Section 4.2). In a study conducted by Melo et al. [55], the researchers evaluated the impact of the body position on the usability of VR environments. Additional studies evaluated cybersickness [21]–[23], [26], [39] (see also Section 4.8).

4.9.2 Technical Evaluation

In the technical evaluation, a few studies measured end-to-end latency [1], [3], [8], [29], [39] (see also Section 4.7). Further studies evaluated the accuracy of body tracking, e.g., the accuracy of the step detector [56], the accuracy of a natural action recognition algorithm [3], the accuracy of the Kinect device [20], the posture accuracy [45], and the accuracy of high fiving while either seeing the full-body or only hands [6]. One study evaluated the effects of FoV, i.e., how big a camera angle view should be to see the full-body [27]. In another study, the researchers compared different FoVs to understand how this affects the relation between the virtual body and environment [26]. Further works validated the design of the shoe which allows the user to experience physical elevation [52], the fidelity of the virtual wheelchair's functionality [48] or the proposed approach of multiple Kinect devices [17].

4.9.3 User Experience

When evaluating user experience, many studies (n = 9) evaluated the sense of agency and body ownership. The majority of the studies represent a realistic avatar [9], [10], [16], [25], [26], [33], [51]. One study compares the representation of a realistic avatar over a wooden mannequin [8] whereas another study [1] shows only a stick figure.

Some research publications also focus on the sense of presence [6], [32] or even analyze whether different avatar forms (gender-matched vs. a simple skeleton avatar) affect judgment [7]. One study analyzes which viewpoint is preferred by the user (comparing first- and third-person view) [45]. Furthermore, some attempts have been made to investigate the level of immersion [22], [39] or the experience between a human and a humanoid robot [57]. Many studies developing VR games evaluated gameplay experiences [13], [40], [47], [51], [55].

The state of the art research revealed that most research publications evaluate user experience (n = 21). Some of them compared different avatar forms (see Section 4.2.1), whereas others evaluated the effect of a full-body avatar over a hands-only representation (see Section 4.2.2). However, none of them evaluated different tracking technologies. A direct comparison between a motion capture suit such as Vicon or OptiTrack and a Kinect device could provide further useful information regarding body ownership or sense of presence.

5 CONCLUSION

In this paper, we have reviewed the state of the art research on tracking motions of the user to reconstruct a full-body avatar. The research activity in this particular area has significantly improved in recent years. We included 53 studies that met our criteria and discussed the potential of full-body motion reconstruction in VR applications. Our findings indicate that even though it is known that markerless motion capture systems, such as the Kinect device, suffer from high-latency, noise in skeleton tracking, and cannot handle occlusion, the majority of studies still use such a system. The second most common technology to track the movements in real-time, and with high accuracy, is a marker-based motion capture system, such as OptiTrack and Vicon, followed by systems using IMUs. The results also show that many studies combine different tracking technologies. Almost half of the studies using a Kinect device utilize at least one additional body tracking device or a controller, such as Motion Leap or Myo armband. Conversely, studies using marker-based systems usually utilize no additional devices to enable natural interactions with the virtual environment.

The analysis shows that most users prefer a realistic or a humanoid avatar, unlike a stick figure. It has been confirmed by many studies that body tracking and representation of a full-body avatar increase the sense of presence. A small number of studies evaluated the effect of a full-body avatar and hands-only representation. Compared to hands-only representation, the full-body avatar induces a significantly higher sense of presence [6], [15], although there is some contradictory evidence [16]. We hope that in the future more VR applications will reconstruct full-body movements to demonstrate the benefits of using motion capture technologies in VR applications.

The analysis also revealed some interesting results regarding latency. Unfortunately, most studies do not measure end-to-end latency of their VR application; and those that did mostly reported too high delay. When developing VR applications, the developers should thoroughly consider the total delay from the time motion occurs, to the time the results of that motion are displayed on a HMD. High end-to-end latency not only causes cybersickness but will also break the sense of presence. Thus, the developers should always try to ensure a high frame rate and low HMD latency.

Apart from the fact that high latency is one of the most significant limitations of most VR applications, the cables attached to the various devices break the sense of presence and limit the freedom of movement. With wireless HMD technologies such as HTC Vive (Pro) with a wireless adapter and Oculus Quest, new possibilities are arising. The novel HMDs do not limit mobility anymore, and will probably ensure better natural locomotion. We believe that by creating more powerful VR wireless solutions, the number of multiplayer VR applications will increase.

Some studies attempted to increase immersion by stimulating additional senses, such as the sense of touch or smell. However, none of the studies show a full-body avatar and use novel haptic gloves. Hence, there is a considerable potential to improve VR experiences. By using a motion capture system and advanced haptic technology, the VR applications not only reconstruct a full-body avatar but also bring a realistic touch. Such an application creates the illusion of touching a real object or another player. We hope that this survey will encourage researchers and developers in the future to reconstruct full-body movements and provide multi-modal feedback to enhance VR experiences.

REFERENCES

[1] S. Kasahara, K. Konno, R. Owaki, T. Nishi, A. Takeshita, T. Ito, S. Kasuga, and J. Ushiba, "Malleable Embodiment: Changing Sense of Embodiment by Spatial-Temporal Deformation of Virtual Human Body," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: ACM, 2017, pp. 6438–6448.

[2] T. Hilfert and M. König, "Low-Cost Virtual Reality Environment for Engineering and Construction," *Visualization in Engineering*, vol. 4, no. 1, p. 2, Jan 2016.

[3] F. Jiang, X. Yang, and L. Feng, "Real-time Full-body Motion Reconstruction and Recognition for Off-the-shelf VR Devices," in *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1*, ser. VRCAI '16. New York, NY, USA: ACM, 2016, pp. 309–318.

[4] M. Sra and C. Schmandt, "MetaSpace: Full-body Tracking for Immersive Multiperson Virtual Reality," in *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, ser. UIST '15 Adjunct. New York, NY, USA: ACM, 2015, pp. 47–48.

[5] M. Slater and S. Wilbur, "A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 6, pp. 603–616, 1997.

[6] M. K. Young, J. J. Rieser, and B. Bodenheimer, "Dyadic Interactions with Avatars in Immersive Virtual Environments: High Fiving," in *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, ser. SAP '15. New York, NY, USA: ACM, 2015, pp. 119–126.

[7] B. Bodenheimer and Q. Fu, "The Effect of Avatar Model in Stepping off a Ledge in an Immersive Virtual Environment," in *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, ser. SAP '15. New York, NY, USA: ACM, 2015, pp. 115–118.

[8] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch, "The Effect of Avatar Realism in Immersive Social Virtual Realities," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '17. New York, NY, USA: ACM, 2017, pp. 39:1–39:10.

[9] T. C. Peck, S. Seinfeld, S. M. Aglioti, and M. Slater, "Putting Yourself in the Skin of a Black Avatar Reduces Implicit Racial Bias," *Consciousness and Cognition*, vol. 22, no. 3, pp. 779 – 787, 2013.

[10] D. Banakou, R. Groten, and M. Slater, "Illusory Ownership of a Virtual Child Body Causes Overestimation of Object Sizes and Implicit Attitude Changes," *Proceedings of the National Academy of Sciences*, vol. 110, no. 31, pp. 12 846–12 851, 2013.

[11] J. T. C. Tan, Y. Mizuchi, Y. Hagiwara, and T. Inamura, "Representation of Embodied Collaborative Behaviors in Cyber-Physical Human-Robot Interaction with Immersive User Interfaces," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '18. New York, NY, USA: ACM, 2018, pp. 251–252.

[12] J. Botev and S. Rothkugel, "High-Precision Gestural Input for Immersive Large-Scale Distributed Virtual Environments," in *Proceedings of the 9th Workshop on Massively Multiuser Virtual Environments*, ser. MMVE'17. New York, NY, USA: ACM, 2017, pp. 7–11.

[13] K. Desai, S. Raghuraman, R. Jin, and B. Prabhakaran, "QoE Studies on Interactive 3D Tele-Immersion," in *2017 IEEE International Symposium on Multimedia (ISM)*, Dec 2017, pp. 130–137.

[14] F. A. Friðriksson, H. S. Kristjánsson, D. A. Sigurðsson, D. Thue, and H. H. Vilhjálmsson, "Become your Avatar: Fast Skeletal Reconstruction from Sparse Data for Fully-tracked VR," in *Proceedings of the 26th international conference on artificial reality and telexistence and the 21st eurographics symposium on virtual environments: posters and demos. Eurographics Assoc. Arkansas, Little Rock*, 2016.

[15] D. Mendes, D. Medeiros, M. Sousa, R. Ferreira, A. Raposo, A. Ferreira, and J. Jorge, "Mid-Air Modeling with Boolean Operations in VR," in *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, March 2017, pp. 154–157.

[16] R. Kondo, M. Sugimoto, K. Minamizawa, T. Hoshi, M. Inami, and M. Kitazaki, "Illusory Body Ownership of an Invisible Body Interpolated Between Virtual Hands and Feet via Visual-Motor Synchronicity," *Scientific Reports*, vol. 8, no. 1, 2018.

[17] P. Leoncini, B. Sikorski, V. Baraniello, F. Martone, C. Luongo, and M. Guida, "Multiple NUI Device Approach to Full Body Tracking for Collaborative Virtual Environments," in *Augmented Reality, Virtual Reality, and Computer Graphics*, L. T. De Paolis, P. Bourdot, and A. Mongelli, Eds. Cham: Springer International Publishing, 2017, pp. 131–147.

[18] E. Matsas, G.-C. Vosniakos, and D. Batras, "Modelling Simple Human-robot Collaborative Manufacturing Tasks in Interactive Virtual Environments," in *Proceedings of the 2016 Virtual Reality International Conference*, ser. VRIC '16. New York, NY, USA: ACM, 2016, pp. 20:1–20:4.

[19] J. Lin, X. Guo, J. Shao, C. Jiang, Y. Zhu, and S.-C. Zhu, "A Virtual Reality Platform for Dynamic Human-scene Interaction," in *SIGGRAPH ASIA 2016 Virtual Reality Meets Physical Reality: Modelling and Simulating Virtual Humans and Environments*, ser. SA '16. New York, NY, USA: ACM, 2016, pp. 11:1–11:4.

[20] S. Charles, "Real-Time Human Movement Mapping to a Virtual Environment," in *2016 IEEE Region 10 Symposium (TENSYMP)*, May 2016, pp. 150–154.

[21] J. C. Eubanks, C. Lai, and R. P. Mcmahan, "Portable Virtual Reality: Inertial Measurements and Biomechanics," in *2015 IEEE 1st Workshop on Everyday Virtual Reality (WEVR)*, March 2015, pp. 1–4.

[22] P. Schäfer, M. Koller, J. Diemer, and G. Meixner, "Development and Evaluation of a Virtual Reality-System with integrated Tracking of Extremities under the Aspect of Acrophobia," in *2015 SAI Intelligent Systems Conference (IntelliSys)*, Nov 2015, pp. 408–417.

[23] L. A. Shaw, B. C. Wünsche, C. Lutteroth, S. Marks, and R. Callies, "Challenges in Virtual Reality Exergame Design," *Conferences in Research and Practice in Information Technology Series*, vol. 162, pp. 61–68, 2015.

[24] M. Otto, M. Prieur, P. Agethen, and E. Rukzio, "Dual Reality for Production Verification Workshops: A Comprehensive Set of Virtual Methods," *Procedia CIRP*, vol. 44, pp. 38–43, 2016.

[25] P. Bourdin, J. M. T. Sanahuja, C. C. Moya, P. Haggard, and M. Slater, "Persuading People in a Remote Destination to Sing by Beaming There," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '13. New York, NY, USA: ACM, 2013, pp. 123–132.

[26] H. G. Debarba, S. Perrin, B. Herbelin, and R. Boulic, "Embodied Interaction Using Non-planar Projections in Immersive Virtual Reality," in *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '15. New York, NY, USA: ACM, 2015, pp. 125–128.

[27] H. Rhodin, C. Richardt, D. Casas, E. Insafutdinov, M. Shafiei, H.-P. Seidel, B. Schiele, and C. Theobalt, "EgoCap: Egocentric Markerless Motion Capture with Two Fisheye Cameras," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 162:1–162:11, Nov. 2016.

[28] I. Podkosova, K. Vasylevska, C. Schoenauer, E. Vonach, P. Fikar, E. Bronederk, and H. Kaufmann, "Immersivedeck: a Large-Scale Wireless VR System for Multiple Users," in *2016 IEEE 9th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS)*, March 2016, pp. 1–7.

[29] M. Johnson, I. Humer, B. Zimmerman, J. Shallow, L. Tahai, and K. Pietroszek, "Low-Cost Latency Compensation in Motion Tracking for Smartphone-Based Head Mounted Display," in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, ser. AVI '16. New York, NY, USA: ACM, 2016, pp. 316–317.

[30] C. Malleson, M. Kosek, M. Klaudiny, I. Huerta, J. C. Bazin, A. Sorkine-Hornung, M. Mine, and K. Mitchell, "Rapid One-Shot Acquisition of Dynamic VR Avatars," in *2017 IEEE Virtual Reality (VR)*, March 2017, pp. 131–140.

[31] R. Shewaga, A. Uribe-Quevedo, B. Kapralos, and F. Alam, "A Comparison of Seated and Room-Scale Virtual Reality in a Serious Game for Epidural Preparation," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–14, 2017.

[32] S. Seele, S. Misztal, H. Buhler, R. Herpers, and J. Schild, "Here's Looking At You Anyway!: How Important is Realistic Gaze Behavior in Co-located Social Virtual Reality Games?" in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '17. New York, NY, USA: ACM, 2017, pp. 531–540.

[33] T. Collingwoode-Williams, M. Gillies, C. McCall, and X. Pan, "The Effect of Lip and Arm Synchronization on Embodiment: A Pilot Study," in *2017 IEEE Virtual Reality (VR)*, March 2017, pp. 253–254.

[34] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," *International Journal of Surgery*, vol. 8, no. 5, pp. 336–341, 2010.

[35] B. Lee, M. Cho, J. Min, and D. Saakes, "Posing and Acting As Input for Personalizing Furniture," in *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, ser. NordiCHI '16. New York, NY, USA: ACM, 2016, pp. 44:1–44:10.

[36] T. M. Takala and M. Matveinen, "Full Body Interaction in Virtual Reality with Affordable Hardware," in *2014 IEEE Virtual Reality (VR)*, March 2014, pp. 157–157.

[37] J. Bolton, M. Lambert, D. Lirette, and B. Unsworth, "PaperDude: A Virtual Reality Cycling Exergame," in *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '14. New York, NY, USA: ACM, 2014, pp. 475–478.

[38] E. Tuveri, L. Macis, F. Sorrentino, L. D. Spano, and R. Scateni, "Fitmersive Games: Fitness Gamification Through Immersive VR," in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, ser. AVI '16. New York, NY, USA: ACM, 2016, pp. 212–215.

[39] M. E. Latoschik, J.-L. Lugrin, M. Habel, D. Roth, C. Seufert, and S. Grafe, "Breaking Bad Behavior: Immersive Training of Class Room Management," in *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology*, ser. VRST '16. New York, NY, USA: ACM, 2016, pp. 317–318.

[40] K.-T. Chou, M.-C. Hsiu, and C. Wang, "Fighting Gulliver: An Experiment with Cross-Platform Players Fighting a Body-Controlled Giant," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '15. New York, NY, USA: ACM, 2015, pp. 65–68.

[41] X. Liu, X. Feng, S. Pan, J. Peng, and X. Zhao, "Skeleton Tracking Based on Kinect Camera and the Application in Virtual Reality System," in *Proceedings of the 4th International Conference on Virtual Reality*, ser. ICVR 2018. New York, NY, USA: ACM, 2018, pp. 21–25.

[42] V. Reus, M. Mello, L. Nedel, and A. Maciel, "Correcting User's Head and Body Orientation Using a Comfort Pose Function," in *2013 XV Symposium on Virtual and Augmented Reality*, May 2013, pp. 167–174.

[43] S. Greuter and D. J. Roberts, "SpaceWalk: Movement and Interaction in Virtual Space with Commodity Hardware," in *Proceedings of the 2014 Conference on Interactive Entertainment*, ser. IE2014. New York, NY, USA: ACM, 2014, pp. 1:1–1:7.

[44] K. Czesak, R. Mohedano, P. Carballeira, J. Cabrera, and N. Garcia, "Fusion of Pose and Head Tracking Data for Immersive Mixed-Reality Application Development," in *2016 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, July 2016, pp. 1–4.

[45] T. N. Hoang, M. Reinoso, F. Vetere, and E. Tanin, "Onebody: Remote Posture Guidance System Using First Person View in Virtual Environment," in *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, ser. NordiCHI '16. New York, NY, USA: ACM, 2016, pp. 25:1–25:10.

[46] Y. Tanaka and M. Hirakawa, "Efficient Strength Training in a Virtual World," in *2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, May 2016, pp. 1–2.

[47] S. Yan, G. Ding, Z. Guan, N. Sun, H. Li, and L. Zhang, "OutsideMe: Augmenting Dancer's External Self-Image by Using A Mixed Reality System," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '15. New York, NY, USA: ACM, 2015, pp. 965–970.

[48] C. J. Headleand, T. Day, S. R. Pop, P. D. B. Ritsos, and N. W. John, "A Cost-Effective Virtual Environment for Simulating and Training Powered Wheelchairs Manoeuvres," *Medicine Meets Virtual Reality 22: NextMed/MMVR22*, vol. 220, p. 134, 2016.

[49] S. Chagué and C. Charbonnier, "Real Virtuality: A Multi-user Immersive Platform Connecting Real and Virtual Worlds," in *Proceedings of the 2016 Virtual Reality International Conference*, ser. VRIC '16. New York, NY, USA: ACM, 2016, pp. 4:1–4:3.

[50] J. S. Thomas, C. R. France, S. T. Leitkam, M. E. Applegate, P. E. Pidcoe, and S. Walkowski, "Effects of Real-World Versus Virtual Environments on Joint Excursions in Full-Body Reaching Tasks," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 4, pp. 1–8, 2016.

[51] M. C. S. Egeberg, S. L. R. Lind, S. Serubugo, D. Skantarova, and M. Kraus, "Extending the Human Body in Virtual Reality: Effect of Sensory Feedback on Agency and Ownership of Virtual Wings," in *Proceedings of the 2016 Virtual Reality International Conference*, ser. VRIC '16. New York, NY, USA: ACM, 2016, pp. 1–4.

[52] D. Schmidt, R. Kovacs, V. Mehta, U. Umapathi, S. Köhler, L.-P. Cheng, and P. Baudisch, "Level-Ups: Motorized Stilts That Simulate Stair Steps in Virtual Reality," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '15. New York, NY, USA: ACM, 2015, pp. 359–362.

[53] D. Borland, "Integrating Head and Full-Body Tracking for Embodiment in Virtual Characters," in *2013 IEEE Virtual Reality (VR)*, March 2013, pp. 81–82.

[54] M. Loviska, O. Krause, H. A. Engelbrecht, J. B. Nel, G. Schiele, A. Burger, S. Schmeißer, C. Cichiwskyj, L. Calvet, C. Griwodz, and P. Halvorsen, "Immersed Gaming in Minecraft," in *Proceedings of the 7th International Conference on Multimedia Systems*, ser. MMSys '16. New York, NY, USA: ACM, 2016, pp. 32:1–32:4.

[55] M. Melo, T. Rocha, L. Barbosa, and M. Bessa, "The Impact of Body Position on the Usability of Multisensory Virtual Environments: Case Study of a Virtual Bicycle," in *Proceedings of the 7th International Conference on Software Development and Technologies for*

Enhancing Accessibility and Fighting Info-exclusion, ser. DSAI 2016. New York, NY, USA: ACM, 2016, pp. 20–24.

[56] P. Caserman, P. Krabbe, J. Wojtusch, and O. von Stryk, "Real-Time Step Detection Using the Integrated Sensors of a Head-Mounted Display," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2016, pp. 003 510–003 515.

[57] S. Kishore, X. Navarro, E. Dominguez, N. de la Pena, and M. Slater, "Beaming into the News: A System for and Case Study of Tele-Immersive Journalism," *IEEE Computer Graphics and Applications*, pp. 1–1, 2016.

[58] OptiTrack. (2018) Prime 41. Last visited on October 9th, 2018. [Online]. Available: http://optitrack.com/products/prime-41/

[59] Vicon. (2018) Compare Cameras. Last visited on October 9th, 2018. [Online]. Available: https://www.vicon.com/visualization

[60] J. Thomas, M. Azmandian, S. Grunwald, D. Le, D. Krum, S.-H. Kang, and E. S. Rosenberg, "Effects of Personalized Avatar Texture Fidelity on Identity Recognition in Virtual Reality," in *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, R. W. Lindeman, G. Bruder, and D. Iwai, Eds. The Eurographics Association, 2017.

[61] G. Lucas, E. Szablowski, J. Gratch, A. Feng, T. Huang, J. Boberg, and A. Shapiro, "The Effect of Operating a Virtual Doppleganger in a 3D Simulation," in *Proceedings of the 9th International Conference on Motion in Games*, ser. MIG '16. New York, NY, USA: ACM, 2016, pp. 167–174.

[62] H. Wauck, G. Lucas, A. Shapiro, A. Feng, J. Boberg, and J. Gratch, "Analyzing the Effect of Avatar Self-Similarity on Men and Women in a Search and Rescue Game," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: ACM, 2018, pp. 485:1–485:12.

[63] E. Makled, Y. Abdelrahman, N. Mokhtar, V. Schwind, S. Abdennadher, and A. Schmidt, "I Like to Move It: Investigating the Effect of Head and Body Movement of Avatars in VR on User's Perception," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '18. New York, NY, USA: ACM, 2018, pp. LBW099:1–LBW099:6.

[64] K. Kilteni, R. Groten, and M. Slater, "The Sense of Embodiment in Virtual Reality," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 373–387, 2012.

[65] M. Sra, "Asymmetric Design Approach and Collision Avoidance Techniques For Room-scale Multiplayer Virtual Reality," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ser. UIST '16 Adjunct. New York, NY, USA: ACM, 2016, pp. 29–32.

[66] A. Maselli and M. Slater, "The Building Blocks of the Full Body Ownership Illusion," *Frontiers in Human Neuroscience*, vol. 7, p. 83, 2013.

[67] M. Mine and G. Bishop, "Just-in-time Pixels," *University of North Carolina at Chapel Hill Technical Report TR93-005*, 1993.

[68] A. Garcia-Agundez, A. Westmeier, P. Caserman, R. Konrad, and S. Göbel, "An Evaluation of Extrapolation and Filtering Techniques in Head Tracking for Virtual Environments to Reduce Cybersickness," in *Serious Games*, M. Alcañiz, S. Göbel, M. Ma, M. Fradinho Oliveira, J. Baalsrud Hauge, and T. Marsh, Eds. Cham: Springer International Publishing, 2017, pp. 203–211.

[69] M. Meehan, S. Razzaque, M. C. Whitton, and F. P. Brooks, "Effect of Latency on Presence in Stressful Virtual Environments," in *IEEE Virtual Reality, 2003. Proceedings.*, March 2003, pp. 141–148.

[70] K. Raaen and I. Kjellmo, "Measuring Latency in Virtual Reality Systems," in *Entertainment Computing - ICEC 2015*, K. Chorianopoulos, M. Divitini, J. Baalsrud Hauge, L. Jaccheri, and R. Malaka, Eds. Cham: Springer International Publishing, 2015, pp. 457–462.

[71] J. J. LaViola, Jr., "A Discussion of Cybersickness in Virtual Environments," *SIGCHI Bull.*, vol. 32, no. 1, pp. 47–56, Jan. 2000.

[72] R. A. Virzi, "Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?" *Human factors*, vol. 34, no. 4, pp. 457–468, 1992.

[73] L. Faulkner, "Beyond the five-user assumption: Benefits of increased sample sizes in usability testing," *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 3, pp. 379–383, Aug 2003.

Polona Caserman is a PhD student in the Serious Games group at TU Darmstadt since July 2017. Her research areas are full-body tracking and motion reconstruction in immersive VR-based serious games, particularly accuracy issues and latency reduction. Her interests include motion recognition using only a small number of HTC Vive controllers and trackers.

Augusto Garcia-Agundez is a PhD student in the Serious Games group at TU Darmstadt since July 2015. He specializes in games for health and rehabilitation, particularly the measurement of the effect of serious games. His research interests include controllers and sensors for exergames (Leap Motion, Wii Balance Board, Microsoft Kinect) as well as virtual reality and cybersickness.

Stefan Göbel is an assistant professor and academic councilor, lecturer and head of the Serious Games group at the Multimedia Communications Lab of the TU Darmstadt. His research topics cover standardization, authoring, personalized and adaptive games as well as effect measurement of serious games. He also provides lectures and practical courses in Serious Games.