DOI: 10.1111/cgf.15064
EUROGRAPHICS 2024
A. Aristidou and R. McDonnell
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 43 (2024), Number 2
STAR – State of The Art Report

A Survey on Realistic Virtual Human Animations: Definitions, Features and Evaluations

Rim Rekik[1], Stefanie Wuhrer[1], Ludovic Hoyet[2], Katja Zibrek [2] and Anne-Hélène Olivier[2]

[1] Inria centre at the University Grenoble Alpes, France
[2] Inria, Univ Rennes, CNRS, IRISA, M2S, Rennes, France

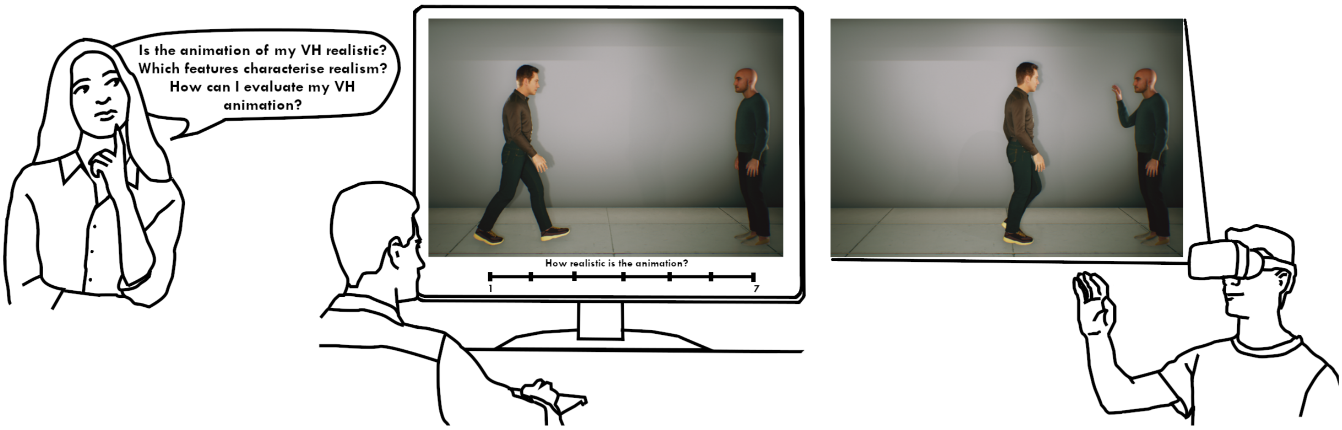

Figure 1: *The central goal of this survey is to explore the aspects of animation realism of virtual humans. We approach this topic by reviewing existing definitions of realism, provide a taxonomy of animation features that influence realism, and discuss existing evaluation methods: objective, subjective and hybrid approaches. Two examples of subjective approaches are depicted in the image above: (left) a user rating realism with a self-report measure, and (right) a user reacting to the virtual human while being immersed in a VR environment).*

Abstract

Generating realistic animated virtual humans is a problem that has been extensively studied with many applications in different types of virtual environments. However, the creation process of such realistic animations is challenging, especially because of the number and variety of influencing factors, that should then be identified and evaluated. In this paper, we attempt to provide a clearer understanding of how the multiple factors that have been studied in the literature impact the level of realism of animated virtual humans, by providing a survey of studies assessing their realism. This includes a review of features that have been manipulated to increase the realism of virtual humans, as well as evaluation approaches that have been developed. As the challenges of evaluating animated virtual humans in a way that agrees with human perception are still active research problems, this survey further identifies important open problems and directions for future research.

CCS Concepts

• ***Computing methodologies*** → ***Animation;***

1. Introduction

Virtual Humans (VHs) are digital representations of humans and present an ever-growing domains in computer graphics and computer vision. Due to recent advancements in immersive technologies [SP18], it is expected the trend to continue. VHs can be used in different domains - users can be presented as avatars in virtual reality (VR) while communicating with other avatars in virtual spaces such as metaverse [WSZ*22], are involved in various applications [CGMB20] such as ecommerce [BJS*22], virtual gaming [AAB*13], visual effects industries and movies [TRL*17], virtual training [Kun06], interactions with virtual doctors [CDFG*22], sports [WDR*22], teaching [BDS23], cultural heritage [MDC18,

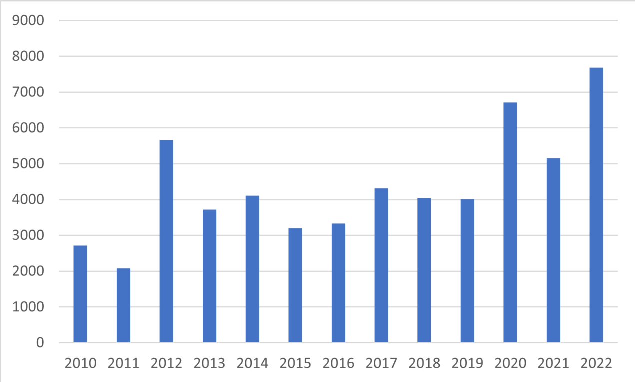

Figure 2: *The figure shows the number of publications per year since 2010 containing keywords "realism" AND "animation" AND ("human" OR "agent" OR "character"). The community's interest in this research topic has increased recently. Source:* `https://app.dimension.ai`*, exported October 30, 2023.*

KPP*21], and virtual try-on for clothing [ZXK*21]. Particularly in non-entertainment applications, VHs which resemble humans in their appearance and behaviour are of high importance. Realism of the VH can help with communicating subtle emotions [ZMM19], while accurate movement can positively affect collaboration [HOB20] and learning [SAS*22].

While being a sought-after feature, realism of VHs is a challenging goal to achieve. First of all, when we speak about realism, we can refer to different aspects of the VHs: their appearance, their movement or behaviour. Each of these aspects presents a particular challenge and combined together can result in further difficulties. For instance, the so called *Uncanny Valley Theory* [Mor70] addresses the problem of a drop in viewers'affinity when the VH's appearance becomes close to human but not quite reaches it. The negative perception or *eeriness* of these VHs increases even more when they are moving. Further research in this area found other aspects of the Uncanny Valley, such as that the affinity is dependent on the 'appeal' of the render style [MBB12] and that mismatch of elements, such as material and shape [ZAJ*15] or behaviour not matching appearance [ZKM18] contribute most to the negative response towards these VHs. Specifically in immersive environments, the most important element was found to be realistic behaviour of the VH [BBBL01], while the appearance was not considered crucial. The underlying theory of the Uncanny Valley itself has been under debate, due to methodological issues [KFMT15, CSJ11, HM10].

Researchers have investigated a wide range of aspects related to the realism of VHs, which is visible from the increasing number of publications on this topic. Compared to year 2010, the number of publications containing a reference to realistic animated virtual humans has almost tripled (see Figure 2) and several surveys over the years attempted to organise this knowledge by focusing on particular aspects of VHs [MTT05, VGS*06, MHLC*22, RWV19, RPA*15]. Our present work is part of this effort, and in our survey we specifically focus on the realism of *animated* VHs. In particular, our goal is to highlight, in the form of a taxonomy, the factors that have been studied in the literature to increase the level of realism in VH animation, and to summarise the various methodological approaches for evaluating the different aspects of such realism. Figure 1 visualizes the goals of this survey. We focus on the aspects of 3D geometrical shape without including render style, material and texture realism in this survey as this would present a separate survey due to the extensiveness of the topic. However, we refer the reader to [ZZM19], which summarises the construction and evaluation of the realism of these aspects in more detail. It is also worth discussing surveys on related topics, and to clarify how this survey complements the previous literature. We can first refer to the three seminal surveys of Magenat-Thalman et al. [MTT05], Vinayagamoorthy et al. [VGS*06] and van Welbergen et al. [VWVBE*10]. Magenat-Thalman et al. [MTT05] discussed the advances as well as the issues and challenges of creating realistic VHs, including three main features which are their appearance (encompassing facial and body shapes, hair, clothes, and textures), their motion and their high-level behaviour. In addition, Vinayagamoorthy et al. [VGS*06] proposed an extensive review focusing on the design of expressive characters, that are able to provide non-verbal communication cues to a user. They defined the main components of non-verbal communication elements and referred to the corresponding existing models. Van Welbergen et al. [VWVBE*10] presented how VHs are modelled (skeletal modeling, physical modeling and biomechanical modeling), the techniques used to animate these VHs (motion editing and physical simulation) and how to control their motion so that VHs are perceived to be natural. This survey was focused on the generation of VH animation without giving the definition of what a realistic VH animation is, and did not detail how the level of its realism is evaluated. Also, They did not include VHs interaction or its evaluation. Similarly, Ursu et al. [UP12] presented a survey on VH assessment without exhaustively focusing on evaluating its realism. Since, several surveys proposed complementary state-of-the-art reports for VH models, such as physics-based [GPEO11], muscle-based [CRPPD17], deep-learning-based [MHLC*22], reinforcement-learning-based [KAK*22] approaches for VH animation, or surveys on 3D morphable face models [EST*20], style in 3D human body motion [RWV19], or on eye-gaze modeling and animation [RPA*15]. While these surveys did not necessarily deal directly with the details of realism and its evaluation, they provided additional insights about the creation of realistic human motions. In particular, it can be noted that Ruhland et al. [RPA*15] discussed the perceptual evaluation of plausibility, effectiveness and the ability to achieve communicative goals. The question of the evaluation of realism was however not at the core of the discussions of these surveys, and to our knowledge this question has never been explicitly addressed in surveys about VHs.

On the topic of motion of VHs, we report studies which have used different approaches to create realistic animations. VHs can be animated using motion capture, where an actor's motion is recorded using sensors or reflective markers, which are placed on the main joints of the body and their position in physical space can be translated into virtual space using infrared cameras (Vicon, Optitrack system) or by using inertial trackers, where the position of the body parts is calculated based on accelerometers, gyroscopes and magnetometers (Xsens). This data is fitted onto a virtual skeleton and retargetted onto the final VH. This process can result in inaccuracies, which can occur at the motion capture phase (occlusions, loss of tracking) or retargetting phase (the skeletons do not match in size, skinning issues when the mesh is added on the skeleton, etc.), reducing the level of realism. Other types of animation, such as procedural and physics-based animation can be used to simulate hair [WBK*07] and clothes [TF88a, TF88b] or blendshapes animation for face reconstruction [JTDP06]) but they highly depend on the ability of an algorithm which simulates the motion to recreate realistic physics. Another way to animate VHs is also by using vision-based motion capture [AC99, Gav99, MG01]. Facial blendshape animation, for example, can be enhanced with video analysis of the facial movements which are translated to a template virtual mesh and retargetted to drive facial blendshapes. Recent advances in computer vision and machine learning approaches have opened up new possibilities to create highly realistic VHs by reconstructing the body shapes and animations using volumetric capture [YZG*21, SHS*23, ZMFB23, ZXB*23]. While the realism of VH animation is constantly improving alongside the evolution of deep learning methods [MHLC*22, MRW*23], the created animations are prone to errors and different animations techniques also create specific issues which reduce the level of realism of VHs. Additionally, since the choice of the technique is related to the application type, there is no consensus nor unified way of evaluating the level of realism. This makes our present survey relevant and timely, since we create a taxonomy of animation features which impact the VH realism and report the various evaluation approaches performed on VHs which have been created over the years using the above mentioned animation techniques. In our survey, we also address the terminology of realism. While there is no formal definition of VH realism, some insights can be found in closely-related fields, in particular in the recent survey of Gonçalves et al. [GCM*22], discussing the effects of realism in immersive virtual experiences at large. More specifically, they first focused on the definitions of the concept of realism available in the literature, and made the distinction between objective and subjective realism. Then, they discussed how the realism of Immersive Virtual Environments (IVEs) can impact users' experiences. The impact was characterized in terms of performance, presence, perceived environment realism, involvement, physiological response, user preference, task satisfaction, virtual agent rating, embodiment, and user behaviour. They also report the factors that can impact individually realism, considering IVEs content and system characteristics, with a short focus on VHs including factors related their behavioural realism, textures or anthropometric fidelity. Interestingly, the authors highlighted the lack of taxonomy to define the factors related to realism and the resulting users'experience in IVEs. Discussions around the concept of perceived realism can also be found in the context of VH applications for human cognition. For instance, authors in neuroimaging research [dBdG15] reviewed whether the perception of VHs differs from the perception of humans, and the potential effect influence of using VHs as stimulus material in social and affective neuroimaging studies. In particular, the authors considered the effect of VH appearance and motion, as well as of expressions of emotions and social interactions. The main variables considered in this review were brain activity measurements, self-reported rating using likert scales, classification speed or uncanny valley stimulus.

To summarise, while the importance of realism is consensually accepted, there is as yet no overview of the way in which realism of VHs is assessed. In this context, our objectives are to review the factors that have been manipulated to increase the realism of VHs as well as on the evaluation approaches that have been developed in the literature. However, given the number of 'technical' components that are involved in the creation of such VHs, we chose in this survey to focus on the components related to the animation of the VHs. This choice enables us to provide a comprehensive review within a tractable space, while simultaneously providing insights that complement the other relevant reviews dedicated to VH motion mentioned above [VWVBE*10, GPEO11, RPA*15, CRPPD17, RWV19, MHLC*22, KAK*22]. We propose a narrative review, where the literature search was mainly performed using google scholar database. Our keywords were "realism" AND "animation" AND "human" OR "agent" OR "character". To describe approaches and theories that are used to study the animation realism, we also extended our search terms to more specific topics such as evaluation methods.

The main contributions of this survey, which are visualized in Figure 3, are therefore:

- A summary of the proposed definitions of animated VH realism depending on the different fields in the literature (presented in Section 2).
- A taxonomy of the different animation features that impact VH realism, considering namely its kinematics, physics, 3D geometrical features as well as potential interactions (presented in Section 3).
- An overview of existing approaches to evaluate the realism of VHs that we divide into 3 categories: objective, subjective and hybrid approaches(presented in Section 4).
- A discussion about future works and challenges related to the design and evaluation of VHs realism (presented in Section 5).

2. Realism of Virtual Humans

In this section, we present the terminology and the definitions related to VH realism that exist in the literature to provide a global view of how researchers define VH realism. We also propose a more precise definition of VH animation realism that excludes texture and colour information.

2.1. Terminology

VH realism has been studied in different research areas, and while a large number of works use the terms *"realism"* or *"realistic"*, various other qualifiers have also been used to describe these concepts. Some authors [RAP08, VHBO14] use the term *"plausibility"* which is defined in the context of immersive environments by Gonçalves et al. [GCM*22] as the events and behaviours that can possibly occur in reality. Others use the term *"believability"* [Hil00, MTT05, KHG*07, APA*14, ZAJ*15], or its synonym *"convincingness"* [ZAJ*15], to qualify not only the physical appearance of the virtual agent but also its emotions, personality and social interactions [NDP10]. However, some researchers [CCZB00, RPE*05, VWVBE*10] use *"naturalness"* as VHs need

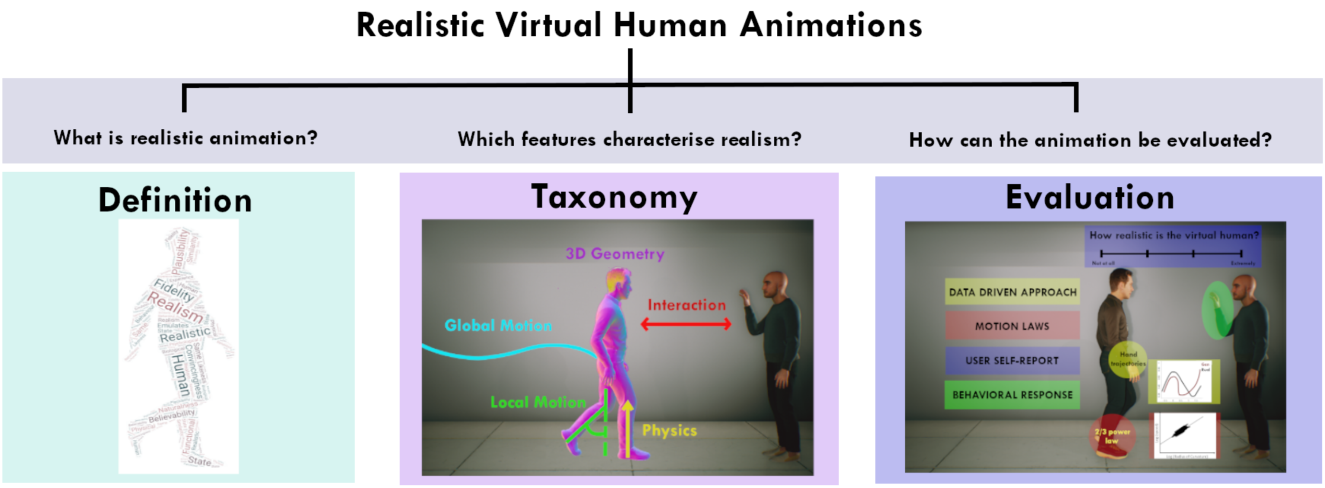

Figure 3: *Overview of the main contributions of this survey. We summarize existing definitions of realistic animation (Section 2), provide a taxonomy of animation features (Section 3), and give an overview of evaluation protocols (Section 4). The bottom of the figure shows the following. Left: a cloud of words used in definitions. Middle: the animation features that can be used to manipulate VH realism, which can be separated into four dimensions: kinematics-based (including global and local ones), physics-based, 3D-geometry-based and interaction-based features. Right: evaluation strategies that have been used to measure realism, and which include data-driven approaches, motion laws, user self-report, and behavioral responses.*

to be perceived to be natural in their different aspects while others [BLB*02, JL21, ZZC*23] opt for the term *"fidelity"*, which was defined by several works [ODGK03, MWT*12, SFJ01, MWT*12, ABS*05] as the degree to which a model or simulation accurately emulates an object's state and behaviour in the real world. More precisely, Alexander et al. [ABS*05] divided the *"fidelity"* of virtual reality experiences into three subcategories: physical simulation which represents the different aspects of an environment including visual, aural, tactile, olfactory and taste stimuli, functional simulation which describes how the operational equipment reacts to the user and psychological fidelity which defines how close the psychological effect of the simulation would be to the same experience in the real-world. Other terms such as *"Human-likeness"* [CSJ11, dBdG15, HRN*20, SLBA22] and *"humanness"* [KMH*19] are also used to indicate the similarity between VHs and real human beings. Finally, several studies focusing on human motion perception [Joh73, GL02, GP03, CFOC05, BS07, SC12, SCI*12, WCC16] have investigated how sensitive humans are to *"biological"* motions. In this context, humans are very sensitive to biological motion, being able to identify gender or emotion from a very simplified visual version of the representation, using Point Light Display for example [ADGY04, ANM*11, PKHS05]. The term *"biological"* was also used to discern that the VH's motion characteristics were based on a model of real human being, in contrast to an artificial one [CHK07, MPM05].

It is also important to mention that these qualifiers might be relative to the application context, where the important features of the realism might vary depending on the task to be done, the interaction to be performed with the VH or the locations of the VHs with respect to the user. For example, adding eye-motions indicating the side of avoidance in a head-on collision avoidance task is important for the participants to avoid a virtual pedestrian [NHH09], while it has no effect on the collision avoidance behavior in a 90° crossing task [LPB*18]. This can also be illustrated by several research work in the context of games or simulation, where the major characters are expected to be more realistic (e.g., performing finer collision avoidance manoeuvres within a crowd) than the supporting cast [DHOO05, KDC*08], which are less visible to the user.

2.2. Existing Definitions

While there is currently no consensus in the literature on the exact definition of VH realism, several authors attempted to define the minimum requirements for creating realistic VHs. First, Bailenson et al. [BB04] described the VH globally as the *"perceptible digital representation"* of a human being. Then, Magnenat-Thalmann et al. [MTT05] stated that *"three levels of modeling should be considered to create these believable VHs: 1) realistic appearance modeling, 2) realistic, smooth and flexible motion modeling, and 3) realistic high-level behaviours modeling."* Hertzmann et al. [HOP09] extended this definition to include the simulation of *"motions that have personality and individual characteristics"*, which according to them *"adds whole new dimensions of complexity"*.

Blascovich et al. [BLB*02] separated photographic and behavioural realism, where *"photographic realism refers merely to the photographically realistic appearance of virtual humans and objects"*, while *"behavioural realism refers to the extent to which virtual humans and other objects behave like their counterparts in the physical world"*. For them, *"behavioural realism is more important than photographic realism in terms of social influence, devising compelling, behaviourally realistic, human-like charac-*

ters". Vinayagamoorthy et al. [VGS*06] also suggested that a VH's behaviour influences realism more than its visual aspect, and claimed that expressive VHs are perceived as believable because they *"reflect the simulated situation more like the user would expect them to"*. They further stated that *"virtual characters are plausible within the context of the scenario, (meaning) that the virtual humans must be perceived by the user to be an authentic part of that scenario."*

Realism can also be approached from a physical point of view. For instance, Cruz et al. [CRPPD17] defined realism as *"the visual degree of similarity between the actions, motions and responses of virtual characters with those of their real counterparts, at both the dynamic and kinematic level"*. Interestingly, Geijtenbeek et al. [GPEO11] pointed out that even though the motion of physics-based VHs may be *"physically accurate, they may not exhibit natural properties of real-world motion, such as gait symmetry, passive knee usage or passive arm swing, and may therefore not be perceived as natural."*

Focusing on human faces, Wallraven et al. [WBCB08] proposed to define realism from the viewpoint of human perception, therefore choosing *"the human visual system rather than physical accuracy as (...) gold standard."* In their work, the question of realism was addressed by raising several questions, such as "*What is needed in order to produce perceptually realistic images?"*, *"How do we know when to stop?"*, *"When is realism 'realistic enough'?"* Such a definition therefore links the realism of VHs with the human visual system as the goal of *"computer graphics (is to) produce images that humans look at and have to appreciate."* In contrast, Ren et al [RPE*05] approached this problem *"based on the assumption that the evaluation of naturalness is not intrinsically a subjective criterion imposed by the human observer but is, instead, an objective measure imposed by the data as a whole."* While focusing on the naturalness of VH motions only, their idea was that *"motions that we have seen repeatedly are judged natural, whereas (synthetic) motions that happen very rarely are not."*

2.3. Proposed definition of VH animation realism

After summarizing the existing definitions for realistic VHs in the literature, we propose a specific definition that can be used in vision, graphics, cognitive sciences as well as other fields: VH animation realism is the degree to which the digital representation enables the emulation of the human in its global trajectory, body parts motion (actions or expressions), 3D shape and interactions within a virtual environment (surroundings, other VHs or human users). This mimicry can vary depending on the application while being outside the uncanny valley region.

3. Taxonomy of Animation Features Influencing Virtual Human Realism

In this section, we present a detailed taxonomy of the features that authors found to impact the realism of VH animations. As mentioned in Section 1, our survey is interested in exploring features related to the motion of VHs without focusing on its rendering style including its texture, skin colour or material. We propose to divide these features into four dimensions, illustrated in the middle of Figure 3: 1) kinematics-based features that include global and local motion features , 2) physics-based features, 3) 3D-geometry-based features in relation to motion, and 4) interaction-based features. Table 1 summarises the relevant works presented in the following sections which studied features impacting the realism of VH animation (column 3).

3.1. Kinematics-based Features

Here, we start by discussing approaches that rely on kinematics-based features for VH realism. We subdivide them into features impacting the realism of the VH global motion (Section 3.1.1), and those considering the VH local motions (Section 3.1.2). Following Lynch et al. [LKM*18], we define global features as the VH features that impact the realism of the VH's global trajectory and local features as the VH features related to the motion of the VH's specific body parts. This definition is consistent with Fiset et al. [FLM20], where the local movement was related to limb motion and the global movement to whole-body motion.

3.1.1. Global Features

To generate realistic animations of VH, some researchers manipulated features that impact the quality of their global body motion.

Authors have focused on the similarity between successive body poses, which is leveraged to create natural transitions between existing animation clips [AF02, KGP02, LCR*02]. To generate realistic animations, the general idea is that transitions between animation clips should occur at the time of highest similarity. Such similarities are usually measured based on joint relative differences, joint positions [AF02, KGP02, TLKS08] or angles [LCR*02], joint velocities [AF02, KGP02, LCR*02], or even higher-order derivatives. The more recent *Motion Matching* approach [BC15, HKPP20] is based on the similar principle, even though it generally uses more complex features representing the current animation state globally (i.e., usually a combination of 2D or 3D positions, directions, velocities, etc. expressing either locally or globally different elements of the animation), in order to search in a large database the animation best fitting a given context, which is then blended with the current animation. However, the duration of the transition also influences its realism, as blend duration impacts the perception of the animation [WB04]. In particular, Wang et al. [WB04] demonstrated that transitions with an optimal blend duration (calculated based on geodesic distance or joint velocity differences) are perceptually preferred to generic fixed-duration transitions. Apart from similarity, multiple global features can affect the perceived realism of VH animation. For instance, McDonnell et al. [MNO07] found that the perceived smoothness of walking animations is impacted by factors including linear velocity, motion complexity and group size (i.e. the number of characters in the background). More specifically, they observed that lower intensity movements require fewer pose updates, that characters moving faster across the screen need more frequent updates, and more energetic bodily movements necessitate more poses to look smooth. Furthermore, Turnwald et al. [TEW15] focused on evaluating the perceived similarity in global trajectory while comparing walking motions with real human ones and found that huge path deviation with high curvature, stopping motion, and changing accelerations result into higher differences.

3.1.2. Local Features

Kinematics features related to realistic animation were also addressed considering the local features of the VH animation .

We can first refer to studies interested in the features influencing the realism of facial expressions. Wallraven et al. [WBCB08] found that *"rigid head motion"* and *"natural internal motion"* (i.e., mix of motion capture's natural timing with highly detailed 3D scans of neutral and peak expressions) were crucial to ensure realistic animations, evaluated by the recognition, intensity, sincerity and typicality of the expressions. Both *"rigid head motion"* and *"natural internal motion"* showed similar effects on recognition, but not for judging intensity, sincerity, and typicality where *"natural internal motion"* provided more information. The authors also found that motion blurring impacts the intensity of the expressions while eyes and teeth addition significantly increase recognition, intensity and sincerity. In a similar vein, Hodgins et al. [HJO*10] studied the saliency of different anomalies in animated VHs, including both facial and body anomalies, and highlighted the importance of considering multiple dimensions in modeling the Uncanny Valley. They showed that facial anomalies were the most salient, compared to body animation anomalies, especially when they resembled medical conditions. The presence or absence of facial animation in the vignettes affected how the user respond to questions about emotional impact. Also, eye-tracking data revealed that participants focused on VH faces for a significant period of time. This supports the notion that faces attract the most attention and play a crucial role in the perception of animated characters. More recently, Tessier et al. [TGRJ19] were interested in the facial expression of pain, especially considering the effect of the order of 3 actions units characterizing such an animation, namely B) brow lowering, N) nose wrinkling + upper lip raising and O) orbit tightening + eyelid closure. Results showed that the triggering order of such actions units impacts differently the perceived realism, perceived pain intensity or perceived unpleasantness, showing that not only the identification of motion features but also their temporal sequence is fundamental to design realistic animations.

Other works were interested in the local features related to trunk and limbs motion. Treal et al. [TJM20, TJJ*21] were interested in the effect of adding natural idle trunk motion mimicking human postural oscillations (due to dynamic balance and breathing movements) on the perception of a VH pain. They showed that combining trunk motion and pain facial expressions in comparison to facial expression only, allows to increase not only the perceived level of pain intensity as well as the perceived believability of the VH animation [TJM20], but also the level of empathy of a human observer towards the VH [TJJ*21]. In the specific case of walking motions, McDonnell et al. [MNO07] found that the perceived smoothness is also impacted by cycle rate (i.e., cyclical speed of limbs). In particular, characters with a high cycle rate need more frequent updates to look smooth. Niay et al. [NOZ*20] investigated an invariant characteristic of gait at the lower limb level which can be a good candidate to introduce variety in walking animations while preserving realism. Authors proposed a perceptual experiment based on the Walk Ratio variable (i.e., the step length to step frequency ratio), which was proved to distinguish each individual and to be the same across different walking speeds. Authors asked participants to identify the combination of step length and step frequency that they perceived as the most natural. Results showed that participants were able to perceive and identify the actual natural Walk Ratio of virtual characters, which motion were derived from motion capture data of real walkers. In addition, they were also sensitive to gender in motion, since the walk ratio perceived as the most natural differed between male and female virtual humans. Authors conclude therefore on the potential of such a variable to generate realistic walking animation. Such 'rhythmic' local features are however not necessarily restricted to walking motions, and can be leveraged for other types of motions as well. For instance, Su [Su16] demonstrated that viewers can perceive the rhythm of dance movements (namely Charleston and Balboa) and extract different metrical periodicities both from local features (limb or trunk tempo) or from the global tempo of the dance motion, similar to how one would perceive rhythm in music. Overall, this suggest that the realism of rhythmic movements might be achieved by replicating the kinematic patterns that humans naturally exhibit when moving.

Other authors focused on hand motion features. For example, Jörg et al. [JHO10] studied the perceived realism of finger motions by assessing user's perceptions of errors in synchronization between finger and body motions, as well as using different finger animation methods. Their results indicate that whether changes in finger motion are noticeable or not depends on the type of action performed, but that overall viewers are highly sensitive to synchronization errors. Especially, such errors can impact the interpretation of a scene even if the perceived quality of the animation is not affected. Their results also suggest that finger animations created using motion capture are usually preferred to other approaches (i.e., keyframed, random, or no animation), even though several following works demonstrated that animation perceived as very similar to full motion capture can be obtained by capturing only a subset of the finger joints [HRMO12, WJZ13]. More recently, Hirose et al. [HNH*20] proposed to identify which characteristics of movement is relevant to define kinematic human likeness. They focused specifically on jerk, which is the third time derivative of position and were interested in evaluating the effect of kinematic jerk on the affinity of users for a moving artificial agent, which was represented by a robotic hand. From motion captures of a real user performing a simple 1D grasping motion, they created stimuli by minimizing jerk or adding noise to the original signal. Participants were presented those stimuli both using a laptop or a HMD. On overall, authors found a negative correlation between movement preference and jerk levels. On the laptop, participants preferred first the original motion which integrates very small jerky kinematics, second the minimum jerk stimulus, followed by the jerkier stimuli. In VR, participants preferred a motion slightly jerkier than the original one, second the original motion, followed by stimuli with increasing jerk values. Authors discussed the specificity of the robotic hand representation, that might have affected the preference of kinematics smoothness. However, this study opens perspective on the use of human motion invariant to generate realistic VH animations.

Finally, some authors have investigated how body movement and gestures are perceptually linked in order to identify the important features that should be preserved for creating synthetic realistic animations, where body motion and gestures can come from different motion capture. Luo et al. [LN12] have shown that merging whole body motion with gestures coming from a different motion source generates less perceptually realistic motions, and that viewers' sensitivity to motion realism is related to the variance of motion energy.

3.2. Physics-based Features

In this section, we detail animation features that are based on physics to enhance the realism of VHs. As physics-based simulation is a whole field of research on its own, we suggest the reader to refer to specific surveys for more details about the approaches [VWVBE*10, GPEO11, CRPPD17].

Physics-based animation of VHs was traditionally performed using torque-based methods. In such approaches, specific motion controllers are typically designed to handle the animation of VHs by producing joint torques, that indirectly control the pose of a physics-based character [GPEO11]. However, such approaches originally produced overly-stiff (almost robotic like) animations, which therefore led to the progressive integration of human-like mechanisms. For instance, Wang et al. [WFH09] proposed to optimize the parameters of a physics-based controller to generate natural walking patterns animation. The realism of the simulation was objectively evaluated by comparing the resulting walking patterns with motion capture data. Despite some differences, authors highlighted the importance of the relative magnitude of lower-body joint torque, of minimizing angular momentum as well as stabilizing the head.

To improve the overall realism of physics-based animations, muscle-based approaches then started to include muscles to more finely control joint actuation, in a more similar way to how humans work [CRPPD17]. Realism of muscle-based approaches depend of several parameters, at the core of which usually lies a parametric muscle representation (such as the commonly used Hill muscle model [Hil38]), its activation dynamics (i.e., a non linear relationship between neural excitation and muscle activation), or muscle activation profiles. To benefit from these biological elements, Wang et al. [WHDK12] proposed to augment traditional torque-based approaches by including Hill-type musculotendon units in the physical humanoid model, in order to produce more natural walking and running controllers. More specifically, they used these musculotendon units to have torques for the most important degrees of freedom of their humain model (i.e., sagittal plane hip, knee, and ankle degrees-of-freedom), while simultaneously defining biologically-motivated control functions, where authors minimized an effort term based on the rate of energy expenditure. Authors evaluated the realism of their simulation by comparing the resulting kinematics of the simulated motion with real data in the sagittal plane for lower limbs joint angle for running and walking at different speeds. Furthermore, Geijtenbeek et al. [GVDB-VBE10] evaluated the realism of such musculoskeletal models and found that motion dynamics (i.e., walking or jumping) and speed should be synchronised in order to have natural muscular VH motions. In the specific case of muscle-based approaches, Geijtenbeek et al. [GVDPVDS13] drive the motion of the entire body, while optimizing for the best muscle routing geometry.

In their work, Jiang et al. [JVWDGL19] argued that *"models based on muscle-actuation are able to impose physiologically realistic constraints and energetic cost on the (...) torque trajectories, leading to (...) human-like movements that reflect the mechanics of human anatomy"*. However, as such realism *"comes at the expense of complex modeling efforts and costly computation of simulation"*, they proposed an approach to transform an optimal control problem formulated in the *muscle-actuation* space to an equivalent problem in the *joint-actuation* space. More specifically, their method transforms constant bounds on muscle activations to nonlinear, state-dependent torque limits in the joint-actuation space, based on a transformation of the metabolic energy function into a nonlinear function of joint torques. This transformation enables the consideration of energetic costs in the animation process, making the generated motions more biologically plausible by the use of more natural and realistic torque patterns in the resulting animations.

Finally, other works on the physical realism of VH motions focused on evaluating the features that impact their perceived realism. In particular, Reitsma et al. [RP03] studied the features impacting the realism of ballistic motions. More specifically, they evaluated user sensitivity to physical errors in VH animations, by introducing errors in the velocity and in the gravitational force of VH jumping motions. They found that viewers detect the errors in horizontal velocity easier than the ones added in vertical velocity and that added accelerations were easier to notice than added decelerations. Also, participants were more sensitive to errors in effective gravity, with a lower gravity being easier to detect than a higher gravity. In a following work [RAP08], they demonstrated that such user sensitivities vary depending on whether ballistic motions are displayed on a realistic human character or an abstract ball. More precisely, viewers were more sensitive to gravity errors displayed on the human character, compared to gravity errors displayed on the ball shape, even though the authors claim that other factors may partially account for such difference, e.g., distance of character, or user interpretation size of the abstract object size.

3.3. 3D-geometry-based Features

In this section, we discuss the 3D-geometry-based features influencing the realism of the VH animation. We remind the reader that we do not however discuss features impacting the realism of the rendering style of the VHs.

Early works exploring 3D-geometry-based features showed that the geometric model used for displaying VH animations impacts the perception of the animation [HOT98]. They found that viewers were better able to observe changes in the motion of VHs when using a polygonal model, even though visually limited by the resources of the time, than they were with a simpler stick figure model. Similarly, Mcdonnell et al. [MJM*09] found that motion information alone is enough to recognize the VH emotion however adding shape information increases its intensity and the recognition accuracy. On the contrary, Chaminade et al. [CHK07] showed that motion is considered to be more biological (coming from a real person) when displayed in more impoverished formats. Using

VHs with various degrees of anthropomorphism (from Point Light Displays, through abstract to realistic character), they asked participants to categorize whether the motion was biological or artificial, and showed that participant's response bias towards 'biological' decreased with characters' anthropomorphism. However, with the recent leap in the quality of reconstructed 3D VH, it is important to understand how these new levels of visual realism can impact the perception of VHs. For instance, recent results with more visually realistic VHs suggest that visual realism affects empathic responses to VHs [HZCM23]. Subramanyams et al. [SLVC20] investigated also in the geometrical features impacting the realism of 3D point-cloud VHs. They found that perception pattern distortion on the body and clothes, visual artifacts (i.e., blurriness) and level of detail and resolution of facial features can influence the user perception of the animation. For facial animations, Wallraven et al. [WBCB08] studied the correlation between geometrical features and perceived facial expressions realism. They found that degrading face geometry (shape and texture blurring) degrades a number of recognizable charactertistics, including recogition and intensity of emotions, therefore leading to less realistic VH faces. More recently, Mcdonnell et al. [MZCD21] focused in their study on perceptual importance of particular blendshapes for facial animation and found that some were more salient then others. However, they did not mention if manipulating blendshapes would make the facial expression more realistic. In addition, the increase of complexity also influences the capacity of using realistic 3D VHs for interactive social applications (e.g., social virtual reality).

Several authors [KHT16, KMH*19, RSM*23] were interested in studying the consistency between the shape and the motion of VH animation. For instance, Kluver et al. [KHT16] studied the effect of internal consistency, which refers to the coherence or harmony between different traits or features within a stimulus. In the context of their study, it specifically refers to the match or alignment between the anthropometry (body shape and proportions) and kinematics (movement patterns) of a moving body, i.e., the degree to which these two aspects of biological motion are consistent or in agreement with each other. While not evaluating realism, their results show a significant link between internal consistency and attractiveness, and is therefore potential relevant for creating realistic VH motions. This question was further explored by Kenny et al. [KMH*19], who studied how the dense representation of a VH (e.g., geometry of the articulated structure and distribution of skinning weights over that structure) could influence the kinematics of his/her movements. Particularly, they examined the effect of internal inconsistencies introduced by retargeting motion from one person onto the shape of another. They found that inconsistencies between shape and motion led to lower attractiveness ratings for 'inconsistent' walkers (i.e., with motion and shape from different actors) compared to 'consistent' walkers' (i.e., with motion and shape from the same actor). Shape-motion inconsistency affected the perceived weight and thrown distance of objects, suggesting that the visual system assimilates shape and motion information to alter the perception of action outcomes. The effect of inconsistencies however varied depending on the action observed, with viewers' detection of inconsistencies between shape and motion being typically close to chance level for transitive actions (i.e., actions involving an object) [KMH*19,RSM*23], but above chance level for walking animations [RSM*23]. Overall, these results suggest that detection rate is typically low, despite slight differences between male and female animations.

Another geometrical feature that was studied in the literature is the impact of the sex of the VH on its animation [MJH*09, HRZ*13, TBM*20]. Mcdonnell et al. [MJH*09] evaluated how motions as well as body shapes could impact the perceived sex of VHs, and showed that both have an effect on sex judgement. More specifically, female walks were consistently rated as female for all body shapes, while the neutral walks were rated according to the sex of the VH. In addition, the 3D geometry of the VH will tend to influence our perception of the character's sex, especially for motions with little male or female characteristics. This suggests that the body shape also influences the perception of the character's sex, which highlights the importance of considering both factors when creating realistic virtual characters. Hoyet et al. [HRZ*13] investigated how distinctive the motions of different actors are, when displayed on the same virtual body. The authors found that distinctiveness of motions can be impacted by different gaits and by sex, with female motions tending to be more distinctive than male motion. They also found that asian participants could recognize the motions of caucasian actors as good as or better than their european counterparts, particularly for dancing males, that have been rated more attractive. Thaler et al. [TBM*20] found that, for the specific case of walking motions, sexual dimorphism results in attributing different biological and personality traits to male and female VHs. Specifically, sexual dimorphism in walking was more important for female attractiveness, while increased vertical motion was important for male attractiveness. Interestingly, the opposite was true for perceived confidence. These findings have important implications for applications using animated VHs, as inferences made from the character's shape and motion could influence user behaviour.

3.4. Interaction-based Features

In this section, we detail the studies that were interested in the realism of VHs during interaction tasks. We first consider tasks where the VH interacts with the physical non-human environment, including interactions with objects and second, tasks involving non verbal interactions between the VH and other VHs or users. In this context, the interaction-based features include both the global characteristics of the whole body-trajectory (e.g., whole body trajectory adaptations to navigate an environment without collisions), and the local characteristics of the motion during such situations of interaction. In addition, let us note that our survey only focuses on the realism of VHs and did not include all the studies related to avatar realism during social interactions, which of course are an important and active research area (e.g., [FBH*22,RBB*22,HOB20,LRG*17,GSV*03]).

As highlighted in the recent paper of Hassan et al. [HGW*23], synthesizing realistic character animation implies to consider interactions between a VH and its surroundings and is still a challenging and open research question. To this end, these authors have proposed a method which learns scene interactions by imitating reference motion data from an unstructured large motion data set in order to generate "natural and life-like" animations of interaction. This is achieved using an adversarial framework composed of a discriminator, which distinguishes simulated motion data from reference data, and a policy, which controls the character's motion to maximize a reward. Interactions with the environment are modeled by conditioning the discriminator and the policy on the human motion and the environment. While not directly evaluating the nature of the parameters that ensure the realism of VH interaction animation, they demonstrated the success of their approach considering carrying, sitting on or lying down on different objects. In particular, they highlighted the success rate of their method, being able to generate such motions even when applying physical perturbation, and interestingly its ability to generate motions with different styles. Other authors have considered the case of throwing animations by manipulating the characteristics of overarm or underarm motion capture throwing motions and evaluating observers sensitivity to those changes [VHBO14, YO22]. The main objective was to identify which characteristics of captured throwing motions can be edited while preserving their plausibility. Vicovaro et al. [VHBO14] proposed 4 studies where they manipulated 1) the speed of release of the ball 2) the speed of release of the ball and the motion of the thrower while preserving the angle of release 3) the angle of release of the ball, and 4) the representation of the thrower, which was not anymore a VH but a mechanical throwing device. Their results showed that for studies 1 and 2, shortened underarm throws were perceived unnatural, even when modifying the thrower motion. In study 3, results showed that the manipulation of the angle of release helped to improve the perceived plausibility of shortened underarm throws. In study 4, interestingly, the sensitivity to modification of speed release showed results in contrast with a VH thrower, highlighting the specificity of biological motion. As a continuation of those studies, Yamac and O'Sullivan [YO22] manipulated the point of release of a ball, under different viewing conditions (near/far and side/front). Participants were asked to indicate whether or not the motion was modified. Authors showed that observers were asymmetrically sensitive to modifications of point of release depending on the type of throwing. Early release was mainly detected in overarm throws while late release was more noticeable in underarm throws. Those results taken together showed how subtle is the human eye to human motion editing during interactions with objects, and such animation process should then be carefully conducted to guarantee the preservation of perceived realism.

Another very important feature related to the design of realistic VHs animation is their capacity to interact during social interaction. Here, we are interested in the animation of VHs during social interactions through non-verbal behaviour, that can be described through proximity (including interpersonal distance and touch), gaze behaviour, posture and gestures [HRS08]. In their survey, Pan and Hamilton [PH18] discussed the relevance of VR to study human social interaction and proposed a landscape of virtual interactions depending on the level of graphical realism and the level of interactivity between participant and computer system. They specifically highlighted the future challenges related to the design of fully responsive VHs that are able to deal with a wide range of interaction situations. In this context, authors argued that the biggest challenge would be to design a VH that can successfully pass a Turing test, and this can only be done if researcher from complementary fields of research (psychology, VR, AI) work together on this project. In line with Pan and Hamilton's paper [PH18], Kyrlitsias and Michael-Grigoriou [KMG22] proposed a survey about social interactions with VHs and avatars in immersive environments. One section of the survey is dedicated to the factors influencing social interaction with VHs. In this context, authors highlighted the importance of visual realism, behavioural realism, the uncanny valley concept and the self-representation of the user. Closely related to our survey on the topic of VHs animation realism, authors defined behavioral realism as the "extent to which a VH behaves in the way an actual person would behave", including "verbal and non-verbal behavior, responsiveness and interactivity with the environment and the user". Authors report that behavioral realism has important implications to guarantee the level of social presence and mutual awareness during the interaction. Authors conclude that more research on VHs behaviour would be beneficial to increase their "social potential". Referring to the theory of agency proposed by Blascovich [BLB*02], authors suggest research perspective to the design of VHs that are plausible, intelligent and interactive.

Let us first consider the interaction between a user and one single VH. While our survey is primarily focused on the animation of VHs, it is worth to mention that several authors have considered the effect of the visual appearance on social interactions with VHs. For example, a study by Bailenson et al. [BSH*05] studied the effect of VH's appearance and behaviour on the user's feeling of social presence (copresence) with them and found that the impact of behavioral realism on social presence cannot be estimated without taking into account the appearance of the VH. Similarly, Zibrek et al. [ZMM19] manipulated the render style realism of a VH (photorealistic, simple and sketch styles), standing in front of a user and interacting through verbal and non verbal communication including eye-contact in different emotional scenario (sad, happy and unfriendly). Their results showed that the render style realism did not influence the level of comfort with the VH, as well as the level of social presence, highlighting the importance of VH behaviour rather than appearance in the interaction. However, they found that participants emotional response differed between the photorealistic condition and the stylised ones, where participants felt more concerned in the happy scenario and less concerned in the sad one. Mousas et al. [MKR*21] also evaluated the effect of visual appearance in a collision avoidance task, where a user has to avoid while walking a static obstacle that can be a mannequin, human, cartoon, robot or a zombie. Their results did not show significant differences regarding collision avoidance behaviour between the mannequin, human or cartoon conditions, which can be considered as 3 levels of visual realism of a VH. The main effect was more related to the emotional reaction (i.e., emotional reactivity and contagion) induced by the obstacle, where zombie obstacle was associated with an aversive reaction and a stronger collision avoidance manoeuvre from the user. In a recent study, Nelson and al. [NKAM22] have manipulated the rendering style of the VH (realistic, toon, creepy, scary, and robot) considering the same collision avoidance scenario. Results showed similar collision avoidance behaviours between realistic and toon characters, despite a higher level of empathy towards the realistic character. In line with Mousas et al. [MKR*21], it seems that the emotional reaction induced by the character has a main effect because the avoidance behaviour was stronger for the scary condition.

Regarding behavioural realism of non-verbal interactions, authors took interest in the addition of realistic local body motions to improve the interaction with a user. For example, Aburumman et al. [AGWH22] have studied the effect of VH head nodding motion in a face-to-face interaction using a cartoon-style VH. They designed a study, where users had to perform a virtual maze task while being able to ask help from 2 VHs, animated with eye-blinking, facial expressions, and changing of gaze but with 2 different head nodding behaviour: nodding disabled, or animated with a human-like head nodding motion Their results showed benefits of natural nodding on the interaction, where users rated more positively the nodding VH, and approached more often this VH to ask for help in the maze, revealing higher values of implicit trust. Study by Cafaro et al. [CVB16] examining greeting encounters in VR found that adding nonverbal cues such as smiling, gazing, and proximity during user–agent encounters have effect on user's perception of agent's extraversion and affiliation, and increase the likelihood and frequency of further encounters. Jovane et al. [JRZ*23] have investigated a new method for realistic automatic motion editing based on human visual motion features. The method relies on motion-warping units expressing the motion and position of the VH relative to the point of view of the user with which the VH interacts. They evaluated their method with a user study, where a VH aimed at catching the attention of the user through a waving motion. While their method was not subjectively rated as significantly more realistic than a simple method with a straightforward orientation of the whole body of the agent towards the target, they showed a better rate of detection of the waving motions directed towards the user. This opens perspectives towards the design of VH with attention-aware non-verbal communication capabilities. In relation to visual attention with medical applications, Volonte et al. [VRDB18] investigated the interaction with a virtual patient in a conversational task, in front of a computer screen, where the virtual patient 1) CA: was realistically animated with all the communicative features including head and body movements, dynamic eye gaze and mutual gaze with the user, interactive motion with the environment, conversational facial expression or spoken audio response, 2) NCA: was animated without conversational motions and showed only idle motion, interaction with the environment but not with the user, and audio 3) NA: was not animated and only the audio is played. Results showed that visual attention of the user, evaluated through their gaze behaviour, was higher in conditions CA and NCA comparison to condition NA. Contrary to their expectations, the visual attention was the same in conversational and non conversational situations. Finally, we can mention studies that were interested in the animation of voluntary physical interactions between individuals such as pushing someone. Hoyet et al. [HMO12] were interested in evaluating the effect of mismatches between a physical contact, i.e., a VH pushing another VH, and the subsequent reaction on the perceived realism of such a causality interaction. Their results showed the importance of matching the source force strength and orientation and the reaction, as well as its timing especially for strong push, in order to ensure the plausibility of the interaction.

Other authors have also considered more complex situations when several VHs, namely a crowd, interact together. Cabrero-Daniel et al. have recently designed a quality function metric to characterize trajectories of the crowd [CDMH*21]. Authors defined quality as the level of perceived realism of a trajectory. To this end, they have developed an approach which first relies on crowd simulation and animation experts who allow to identify and evaluate the relevant trajectory features that are related to trajectory quality perception. Second, they designed the quality function which is based on a weighted combination of local, global and interaction variables. Third, they validated their quality metric through a user study, where they showed good correlation between the quality scores and the level of realism rated by participants. Considering the interaction between a crowd of VHs and a user, Kyriakou et al. [KPC17] were interested in identifying the factors that contribute to increase the plausibility of a crowd simulation as well as the level of user's experience. Their results showed that providing the crowd with the ability to interact with the user through collision avoidance as well as basic social behaviour such as gaze or waving, contributes to a perceived higher level of realism of the simulation and higher level of presence. Interestingly, enabling such interactions also induces behavioural adaptations from the user, showing a more social attitude with respect to the VHs. At the local body motion level, Hoyet et al. [HOKP16] have evaluated the effect of adding local shoulder rotation to the collision avoidance behaviour of VH. They conducted two perceptual studies, involving 1) 2 VHs and 2) a crowd of VHs. Their results showed that adding simple secondary shoulder motion to the VHs collision avoidance behaviour increases the perceived naturalness of the situation while hiding the perceived residual collisions. Local body animation characteristics have also been considered by Molina et al. [MRP21]. Authors showed that adding animation variety on top of the simulated locomotor trajectory not only increases the overall perceived realism of the simulated crowd but also the perceived realism of the locomotion trajectory and the animation. An interesting result was that these conclusions applied only when the agents were rendered as VHs and not bots. Finally, let us note that some other factors can also influence the perceived level of realism of a crowd, such as the position and orientation of the VHs, context of the scene and its related crowd behaviour as well as the view point of the camera that allows to visualize the crowd [EPO08, EPO11].

In section 3, our objective was to propose a taxonomy of the animation features influencing VH realism. We then reviewed illustrative papers according to 4 dimensions, related to kinematics, physics, 3D geometry and interaction. These features can be used to evaluate the quality of VH animation quantitatively through objective approaches or as a factor to change the perceived realism by human subjects through subjective methods. In the next section, we propose an overview of these evaluation methods for VH animation realism.

4. Evaluation methods

The organisation of this section is inspired by the survey of Gonçalves et al. [GCM*22]. Authors considered the realism of virtual environments according to two dimensions: 1) objective realism that depends on the similarity of the stimuli (hardware and software) provided by the virtual environment and the ones in real word independently of the user's perception, and 2) subjective realism, that depends on how the users perceive the environment. To review the approaches that have been developed to evaluate the realism of VH motion, we extend this classification considering 1) the objective approaches, 2) the subjective approaches and 3) the hybrid approaches, which are perception-based computational methods. We detail each category below. Figure 3 (right) provides a visual overview of the first two categories, which include data-driven approaches and motion laws for objective approaches, and user self-report and behavioral studies for subjective approaches. The relevant papers which use the evaluation methods described in this section, are summarised in Table 1 (column 4) alongside their animation feature classification.

4.1. Objective approaches

In this part, we present some computational methods that were used in order to evaluate the realism of VH animation. Realism is objectively evaluated using quantitative metrics without the user's opinion in the evaluation loop.

4.1.1. Data driven approaches

The first type of objective approaches can be characterised as a data-driven, where researchers compared one or more features of VH animation with those derived from real recorded human data.

In this category, the main principle is to evaluate similarity between generated VH animation and motion capture information. One solution was to compute the L_2 distance between either corresponding joints (e.g., [MBR17]) or corresponding vertices of densely sampled meshes (e.g., [MRW*23]), possibly after aligning the two sequences using the Dynamic Time Warping (DTW) [Vin68]. Later, Coskun et al. [CTC*18] proposed a novel deep metric that aims to capture the semantic relationships across human motions and achieved significant improvements in measuring similarity compared to traditional approaches based on DTW and L_2 computations. It is an attention-based recurrent neural network that can cluster similar human motions while taking into account not only high level information but also contextual information. Authors found that combining the Maximum Mean Discrepancy (MMD) loss [GBR*12] within a triplet [SKP15] learning paradigm allows for a better learning of an embedding space of different motion categories. Incorporating MMD loss guarantees better separation between the distributions of motion samples from different categories and enforces the clustering of the distributions of similar motions in the same location of the embedding space.

Other metrics were also proposed to evaluate the realism based on motion similarity with real data. At the global trajectory level, Brogan et al. [BJ03] proposed 3 quantitative error metrics for measuring the accuracy of steering models: distance error, area error, and speed error that measure the mean distance, area and speed between the real path and the predicted path. At the local-body motion level, Khan et al. [KHSG17] introduced a motion quality metric for quantitatively evaluating the realism of synthesized full body motions. It is composed of the Inter-joint-variation and the motion energy. Inter-joint-variation relates to the correctness of skeletal motion by evaluating if a generated motion has the same inter-joint relationships as captured motion, while motion energy was defined as the average amount of motion in a sequence leveraging the fact that motion sequences with faster dynamics will have more energy than the ones with slower dynamics. Shen et al. [SDK21] presented another local-body motion level approach to evaluate the realism of hand trajectories. They considered geometric micro metrics (e.g., curvature, aspect, vicinity curliness, linearity, and slope) to evaluate the low-level features of the trajectories. These metrics measure different aspects of the trajectory shape and behaviour, providing insights into how closely the generated trajectories resemble real human hand motion. The authors presented also the KL divergence between the micro metrics per trajectory as a metric to assess the similarity between the generated trajectories and real user traces.

Some researchers also focused on the evaluation of the realism of VH animation by considering its interaction behaviour. For instance, Wang et al. [WOO16] proposed *"a new semantic-level"* crowd evaluation data-driven metric that aims to compare simulated crowds with real-world ones. Instead of focusing on low-level features (e.g.,individual trajectories) or global features (e.g.,crowd densities), the metric seeks to uncover latent path patterns that exist in both real and simulated data. They introduced a Stochastic Variational Dual Hierarchical Dirichlet Process (SV-DHDP) model to compute the fidelity between the extracted patterns and the ones of real crowd data. To extract these patterns, the authors used unsupervised clustering by non-parametric bayesian inference. When considering interactions with the environment, Hassan et al. [HGW*23] used the success rate of their method, defined as the percentage of results where the desired task was completed successfully, to evaluate their physically-simulated characters method.

Not directly aiming at evaluating the realism of the VH animation, but of interest for such an evaluation, other researchers investigated human motion similarities based on Laban Movement Analysis (LMA) [LU71]. They provided an alternative manner for comparing global motion features by considering the four categories proposed by the LMA : Body, Effort, Shape, Space. The Effort component is further divided into four subcategories (namely Space, Weight, Time, Flow), which have been used to quantify VH animation. For instance, Durupinar et al. [Dur21] studied the perception of human motion similarity based on LMA and demonstrated, using a user study, that not all the LMA effort factors are *"significantly discernible from the baseline motion, where all the efforts are zero"*. They found that "Flow", which ensures that motions are continuing or constant, was statistically the most significant when comparing two motions. This approach has also been used by Aristidou et al. [ASC*15] in the context of folk dance in order to compare and evaluate motions, not only by considering posture but also style.

4.1.2. Human motion laws

Another objective way to know if a VH animation is realistic or not is to evaluate whether it respects human motion and interaction laws.

First, at the kinematic level, one can refer to several motion invariants highlighted by studies from human movement sciences or neurosciences such as: the *"two-third power law"* between velocity and curvature [VT82, VKDB01] which expresses the strong covariation between movement and geometry of the trajectory, the *"minimum jerk"* [VF95, PHA*07], associated with maximum smoothness of generated trajectories or the *"Fitts' law"* [Fit54], often associated with the concept of speed-accuracy trade-off motion, which relates time and movement amplitude, as a function of accuracy requested by the task. Authors then proposed to evaluate whether the resulting global motion and animations conform with such human motion laws. For instance, Gibet et al. [GLM01] used *"Fitt's law"* and *"two-third power law"* to evaluate their simulated hand-arm movements communication gestures, with applications for sign language gestures. Bogaers et al. [BYV20] used smoothness metrics for evaluating expressive music gestures of a VH playing piano. To quantify motion smoothness, Balasubramanian et al. [BMCB11] proposed the *"spectral arc-length"* metric. It relies on the frequency components of the movement's speed profile (i.e.,changes of speed over time), represented using the Fourier magnitude spectrum. Closely related to this motion invariant but sharing the data-driven comparison, let us also mention the study of Delbosc et al. [DOS*23] which aimed at generating synchronized and believable facial non-verbal animations for conversational VH. Authors proposed to evaluate their resulting animation in comparison to ground-truth data both using a distance metrics based on DTW but also comparing jerk results as a good indicator of motion naturalness. Finally, at the kinematics level, we can mention the study of Shen et al. [SDK21] focusing on hand motions. In order to have accurate human-like hand trajectories, authors proposed to define a minimum trajectory error to be tolerated, allowing the generated trace to be recognized as human one. The smallest tolerance length was measured using a simulated gesture keyboard recognizer.

The evaluation of VH animation realism based on motion laws has also been investigated from a physical point of view. Geijtenbeek et al. [GVDBVBE10] presented an evaluation approach for VH physical realism using musculoskeletal model simulations. It was composed by 2 metrics: The first metric *"dynamics error measure"* was used to measure if the VH animation respects the Newton-Euler laws of motion. This involves assessing the consistency between linear and angular momentum changes and external forces (e.g., gravity and contact forces). The second metric *"muscle error measure"* was defined as *"the amount of force a muscle must produce on top of its maximum capacity"*. Both evaluation metrics allow for assessing the muscle forces estimated by a musculoskeletal model, providing insights into the adherence to physical laws and the realistic portrayal of muscle forces.

4.2. Subjective Approaches

Including human perception in the evaluation of VH realism is a common practice of evaluation studies in computer graphics. Perceptual evaluation is beneficial since certain objective metrics may result in implementation of computationally heavy methods when designing VHs but their absence would not be detrimental to the perceived realism from the point of view of the observer. In addition, some objective measures may not be perceptually sound. For example, increasing the level of changes to the facial structure of facial animation is not linear to the perceived change [CCB01, MZCD21], therefore, it is important to take user evaluation into account. This evaluation can be done by conducting perceptual user studies, where the users need to respond to questions about the realism aspects of VHs they have observed. We can also study the behaviour of the users when they observe or even interact with the VHs while being immersed with them in virtual environments.

4.2.1. Self-report studies

Self-report studies can be designed to evaluate the realism of VHs. Those explicit user studies are the ones where the user is directly asked about the aspects of VH's realism. They are generally based on questionnaires which can be analysed quantitatively or qualitatively. Researchers show the experimental stimuli to the participants on a display, that can be a screen [HOKP16,RAP08] or an immersive setup [ZMM19, SLVC20], and ask them to respond to one or a series of questions. These questions are typically custom made by the researchers to correspond to the experimental task and the research questions, although some attempts to create standardised questionnaires related to various aspects of VH realism have been proposed, such as the Godspeed questionnaire [BKCZ09, HM10] or the social presence questionnaires [BHB03, BBBL01].

User experiments are typically conducted in the laboratory to ensure the same controlled conditions for all participants. In order to collect a larger and more diverse sample of participants, some researchers use crowdsourcing platforms such as the mechanical turk [AVR21] or Prolific [PS18] platforms, however, these are not considered as controlled settings as the participants can be using various devices with different display settings, resolution, and be conducted in distracting environments.

The simplest way to explicitly evaluate VH realism is to present one stimulus to participants and ask them to rate the perceived realism on a scale. This psychometric tool can be in a form of a Likert scale [Lik32], where the participants rate their level of agreement with an affirmation such as "The VH animation is realistic.", the minimum value of the scale being a strong disagreement and the maximum value being a strong agreement. There are many variations of this scale, varying in the number of possible answers which can be an even or an odd number, this latter authorizing a central neutral answer. For example, Reitsma et al. [RAP08] used a Likert scale from 0 representing the user being most confident that the stimulus has been modified to 9 representing the user being most confident the stimulus was not modified (the shown animation is the reference version). Hoyet et al. [HOKP16] asked participants to rate the degree of naturalness of an interaction between 2 virtual walkers performing a head-on collision avoidance task, using a scale ranging from 1 (not very natural) to 7 (very natural). A similar scale is a semantic differentiation scale, where the stimuli is evaluated using bipolar adjectives, such as "Artificial"-"Natural" [HM17]. Others user studies presented one stimulus to be evaluated but opted for binary responses for questions such as whether the animation was natural, not modified, unchanged (no error) or changed (error) and modified [VHBO14, RP03, EPO08, RAP08], *"the trajectory of the ball was correct or incorrect with respect to the preparatory motion"* [VHBO14], the *"animation smooth or jerky?"* [MNO07], and the pedestrians formation was *"real or synthetically generated"* [EPO08].

Some other researchers evaluated the realism of the VH stimulus implicitly by comparing it to a reference animation. The principle is then to present two stimuli to participants, where one is animated using reference data (e.g., a motion capture) and the other one is either a modified version or a newly generated animation, and asked the participants about the similarity between the two versions. In some user studies [JADJ22, HTMS04, Dur21, MRW*23], the participants were asked to rate the level of similarity between the stimuli using a Likert scale such as *"Choose an option: exactly same, very similar, moderately similar, similar, slightly similar, not similar"* [Dur21]. In other user studies [HOT98, JHO10, WB04, RSM*23, HRVDP04, TLKS08], the participants were asked to respond to one or more binary questions such as whether a *"pair of two motions were the same or different"* [HOT98], *"in which clip was the motion of better quality?"* [JHO10], *"whether the first or second motion of a pair was more natural"* [WB04], *"which of two presentations of the arm contained a change"* [HRVDP04], *"judge whether two given postures appear similar or not"* [TLKS08], and *"which animation was the hybrid?"* [RSM*23]. Other researchers [PMKO09, CZXL09, VCH*21] showed multiple stimuli and asked the participants to pick the two motions that they felt were most similar to each other, or to choose the animation that looks closest to the reference one.

Other researchers designed user studies to implicitly evaluate the realism of VHs by asking about specific animation features. Wallraven et al. [WBCB08] used a 1 to 7 Likert scale to evaluate the intensity, sincerity and the typicality of facial expressions. To evaluate the impact of shape-motion consistency on visual realism of VH animation, Hoyet et al. [HRZ*13] asked their participants to rate the degree to which the motions were probably made by the same actor. Su et al. [Su16] asked their participants to choose the tempo in the movement that felt natural to synchronize with.

4.2.2. Behavioural user studies

In contrast to self-report studies, where users are aware of what is being measured, behavioural user responses are more indirect and can provide additional information about the perceived realism of VHs. Behaviour of users can be examined systematically with measures such as eye-tracking [HJO*10, RKB*18, VRDB18] and response time [CSJ11]. For example, eye-gaze can spot saliency of different areas on the VH's body, which could signal an anomaly that causes difficulty in visual processing [CPJ*13]. It has been used to study the uncanny valley effect but most of these studies focused on limited representations of VHs (static images, only faces without hair, etc.).

Immersive environments, such as VR, give a unique opportunity to study human behaviour, since the user's whole body can be tracked while interacting with the VH, providing a rich source of data. For example, in the study of Zibrek et al. [ZMM19] the interpersonal distances between the participants and VHs were measured in VR (see Figure 4) to study the effects of render style realism of VHs. The proximity metric was computed, which represents the minimum distance between two people in an interaction. The connection between minimum distance and realism perception is motivated by the idea that the users will exhibit social behaviour when they will be presented with a VH whom they perceive to be "alive", i.e. they experience copresence [BSH*05]. As a behavioural metric, it can be expressed as Euclidean distance of the current camera (user) position and the position of central mass of the virtual character in the virtual space (see image *a*, Figure 4). Avoidance behaviour, where the VH is an obstacle in the environment and the user needs to avoid it to reach a goal, can generate a full walking trajectory dataset from the positions of the user through time (see image *b*, Figure 4). These trajectories can include minimum passing distance and average distance from the clearance area, as well as a calculation of deviation point or the point in the trajectory in front of the VH which marks the onset of user's deviation from a straight path [BMC21]. The avoidance behaviour was used to study differences between real and virtual obstacles [SOB*15].

Behavioural methods were often employed in virtual social interaction studies to study behavioural realism of VHs. The idea behind this approach is that realistic behaviour of VHs will result in social influence, prompting the users to respond with natural social behaviour [Bla02, HBS03]. The creation of realistic interactive autonomous VHs is still a challenging pursuit, however, attempts have been made to examine the behaviour of VHs which will induce favourable impressions in the user. A study by Cafaro et al. [CVB16] explored this by observing how likely it was for users to interact with the agent again after being exposed to it and how often they would interact with it. Similar measure of the length and frequency of interaction was used by Volonte et al. [VRDB18] when the user was interacting with a virtual patient and when measuring implicit trust in the study of Aburumman et al. [AGWH22]. Behavioural realism was also studied using collaborative immersive environments, where multiple users were embodied in virtual avatars. Here, the importance is on accurate avatar representation (visual, motion mapping) which can aid or hinder users' communication. While this is a special category of VHs which we do not specifically cover in this survey, the following measures could potentially be used with autonomous VHs as well. Herrara et al. [HOB20] tracked head and hand motions of users as they were communicating with other users in a collaborative task in VR. A higher number of head and hand rotations towards other users in VR was indicative of realistic behaviour and the use of nonverbal cues to aid communication between users, embodied as avatars. Another sign of social influence is nonverbal synchrony, where multiple users exhibit similar behaviours at similar times. It is defined as the correlation between the extent of bodily movements of the two participants, with higher correlation scores indicating higher synchrony [OKKHB20]. A higher number of such occurrences through time were indicative of higher perceived and behavioural realism [HMD*23, OKKHB20].

Before we go to the next section, we should also mention physiological measures, such as heart rate, skin conductance and brain signals. While there is there is a lack of studies using these measures to explore realism of virtual stimuli [GCM*22], some studies used them to investigate the response to VHs. For example, Sterna et al. [SCIC*23] used heart rate and skin conductance response to explore realism of behavior (different amounts and type of movement) and interactability (talking or silent) of virtual agents. The authors found that a high behavioral realism led to an increase in the skin conductance response but not heart rate, while the level of interactability did not influence the physiological measures.

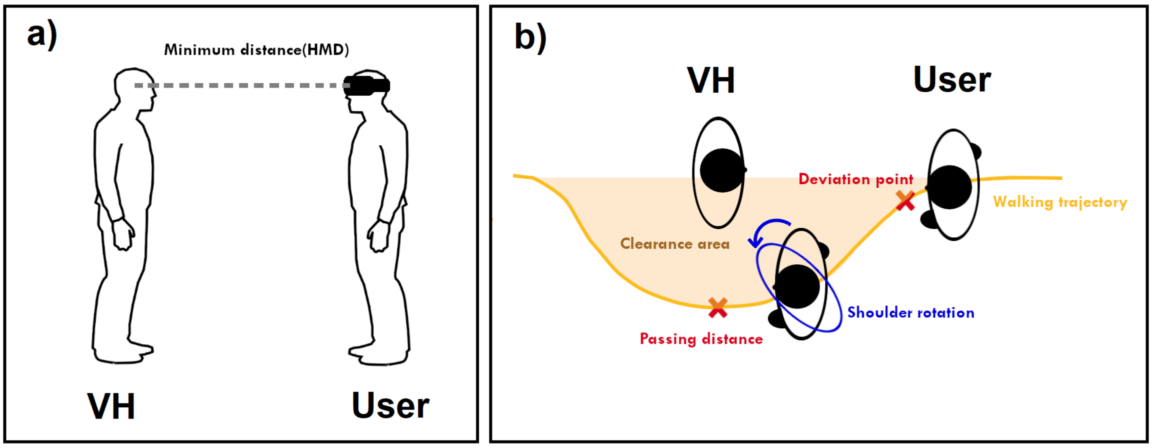

Figure 4: *Examples of the proximity behavioural measure to evaluate the VH realism in immersive environments: a) minimum distance between the user's HMD and the central mass of the VH; b) avoidance behaviour of the user when passing the VH and examples of the metrics of proximity: deviation point, passing distance, clearance area, and shoulder rotation.*

4.3. Hybrid approaches

In this part, we present the different evaluation approaches that correlates quantitative metrics with the human perception of VH realism. Since VH realism can be assessed by comparing VH motion to a reference solution, we present first the works that investigated whether similarity metrics correlate with human perception. Then, we focus on works that were interested specifically in quantifying the perceived realism of VH motion.

Some works such as [PMKO09, HTMS04, TLKS08, CZXL09] investigated the process of defining analytical metrics based on existing methods and optimizing their parameters using the results of a perceptual experiment. Pravzak et al. [PMKO09] found that, for computing the similarity between two motion frames (or poses), the sigmoidal distance function emulates the best the human perception with the manhattan metric being in the second place and the classical squared euclidean metric being the worst. Harada et al. [HTMS04] investigated the optimisation of several quantitative metrics based on human attention to specific body parts more than others. They optimized 3 similarity evaluation metrics (normal distance, minimum mirror-body distance and minimum mirror-region distance) each represented by 4 ways of describing pose (3-dimensional joint angle representation, 3-dimensional position representation, 2-dimensional position representation and 2-dimensional joint angle representation) using both a user study and an exponential simulated annealing algorithm. They found that optimised normal distance with pose described by 3-d position representation best matches human perception of motion similarity. Tang et al. [TLKS08] stated that the proposed *"joint relative distance"* emulates the human perception of VH motion similarity better than the traditional euclidean distance. It is a weighted average distance to compute joint proximities with weights being optimized using users' answers to the question *"Are they similar"*. Similarly, Chen et al. [CZXL09] presented the *"Relational Geometric Distance"*, which is a weighted distance between joints (or bones) that was also optimised using a user study. The responses of the questionnaire were used to label data and therefore training the Adaboost algorithm for the weights optimisation. Lastly, in the frame of crowd simulation trajectory evaluation with respect to real-world data, Guy et al. [GVDBL*12] proposed an Entropy Metric, where the crowd is described as a complex system, and compared this metric with the results of a user study where participants were asked to rate the perceived similarity between a simulated and a real stimulus, using four different levels of entropy for the simulation. Entropy metric quantifies the magnitude of the prediction error: the lower the value of entropy, the smaller the distribution error and the better the similarity with a reference data-set. Authors showed that the simulated stimuli with a lower value of entropy were consistently perceived with higher similarity to the real-data stimuli. In contrast, some studies (e.g., [TEW15, Dur21, JADJ22]) did not found any correlation between perceptual similarity and tested objective metrics. Turnwald et al. [TEW15] studied the relationship between quantitative similarity measures and human perception of differences in motions for human trajectories. They compared between different time series analysis methods to compute the similarity between trajectories: the average Euclidean distance, the DWT distance, and the Longest Common Subsequence(LCSS) method [DGM97]. They found that comparing the derivatives of the path and velocity profiles after aligning the sequences using DTW method, best reproduces human perception of similarity between walking motions. However, none of the similarity measures evaluated in the study particularly matched human perception. Similarly, other works such as [Dur21] and [JADJ22] found that DTW and mean metric scores, respectively, do not generalize to the human perception. These various findings enhance that there is no yet unified quantitative metric that mimics human perception of motion similarity.

Other authors [RP03, RPE*05, LN12, CDMH*21] explicitly studied the correlation between user perception of VH realism and objective metrics to quantify its level. Reitsma et al. [RP03] presented a ballistic error metric for human jumping motions that was optimized using the results of a user study. It is a quantitative metric that detects incorrect gravity capture errors (horizaontal errors and vertical errors) during the flight phase of such motions and was constrained by the thresholds identified based on users' opinions. [RPE*05] explored the possibility of developing a measure to quantify the naturalness of human motion using statistical models. The authors investigated the performance of different statistical models such as of gaussians, hidden markov models, and switching linear dynamic systems, on motion capture data. They then assessed the results of these models by conducting a user study that has as objective to know if a motion is natural or unnatural. They even combined statistical models into a hierarchical ensemble model to capture the naturalness of different parts of the body. However, they found that while their approach is successful, human observers perform better in identifying natural VH motion. Another way to evaluate VH realism was to assess its motion-gesture consistency. For instance, Luo et al. [LN12] investigated a metric to predict if the body motion and the gestures are aligned. They found, based on users' responses, that the mahalanobis distance on energy variance is the best in terms of perceptual realism. Focusing on VHs interaction, Cabrero-Daniel et al. [CDMH*21] presented a data-driven Quality Function (QF) to evaluate the perceived realism of simulated crowd trajectories. The metric is a weighted average of trajectory features (i.e., which are individual, interaction and global features) that were identified by animation and crowd simulation experts. The proposed QF has as objective to penalize deviations between the features of the evaluated trajectory and a golden set of reference values obtained from real data and it was validated by non-expert users' opinions.

5. Discussion and future challenges

While the importance of realism is generally acknowledged, there was still a lack of a general overview of how the realism of VH animation is assessed. In this context, the main objectives of our survey were to review the factors that have been manipulated to enhance the realism of VHs, as well as the evaluation approaches that have been developed in the literature.

First, it was crucial to provide a clear and precise definition of the concept of realism. We reviewed the definitions that have already been proposed in the literature and highlighted the wide range of terminologies surrounding this term. We then proposed to define the realism of VH animation as *"the degree to which the digital representation enables the emulation of the human in its global trajectory, body parts motion (actions or expressions), 3D shape and interactions within a virtual environment (surroundings, other VHs or human users). This mimicry can vary depending on the application while being outside the uncanny valley region."*.

Second, we provided a detailed taxonomy of the features that were considered to design realistic VH animations in different tasks. We proposed to consider dimensions that reflect the complexity of the design of VH animation processes, and identified 4 main categories of features, namely: 1) kinematics-based features which are related to full body (e.g., global trajectory and motion transitions) and local body motions (e.g., face, eyes, limbs, trunk or hands motions.), 2) physics-based features which are linked to the mechanical and muscular properties of human body and the laws of physics, 3) 3D-geometry-based features, which characterize the geometric models as well as the consistency between the shape and the motion and 4) interactions-based features which characterize the relation between the VH animation and its surroundings that can be the physical environment or the other VHs or users. We did not mention that aspect in our survey, but it is important to note here that there is an increasing interest in the VH literature in the animation of accessories and clothes, which are often derived from physics-based or 3D-geometry-based approaches. However, while there are different approaches for generating realistic clothed VHs in motion using physical simulation [TF88a, TF88b] or deep learning methods [RPC*10, SOC19, GBH23, XPB*21], there is little research related to the evaluation of the level of realism of such cloth animations [AOGT15, BX16].

Third, we showed how authors proposed to assess VH realism specifically by considering 3 main approaches that are either 1) objective, 2) subjective or 3) hybrid. In the context of evaluating VH animation realism, where the applications are user-oriented, those hybrid approaches appear to be promising since they combine objective metrics with users' perception and behaviours. While not having systematically reviewed the overall characteristics of the user studies conducted to evaluate VH animation realism, it is worth to note here that authors always report the age and the gender of their participants but provide less often other socio-demographics characteristics, such as their level of expertise (e.g., in human animation or relevant fields). However, such characteristics could be associated with specific attention mechanisms to the resulting animation that might affect participants' answers and behaviours [ZKM18]. We also want to point out that the majority of the user studies, usually conducted in laboratories, recruited young and healthy adults, i.e., typically students (e.g., age was between 18 and 26 in the work of Kenny et al. [KMH*19]). One can then question the generalisation of the results to more diverse population, which is seldom considered in our community. We therefore encourage our research community to engage discussions on the relevant user characteristics to be considered to evaluate VH realism, as well as to collaborate in setting up large studies between research groups from different countries to ensure cultural diversity (e.g., [HRZ*13, WVE*22]).

Recent work has opened up avenues for future research in evaluating the realism of VH animation. In particular, we suggest that the next challenge would be to design a generalized multi-factorial metric of realism. Even though such a metric would more than likely be application-dependent, we believe that exploring such a direction of research would be valuable for advancing our understanding of VH animation realism. In addition to the studies already mentioned in this survey, such as the one of Cabrero-Daniel et al. [CDMH*21] where authors proposed a quality metric of pedestrians trajectories in a crowd, we can mention here other recent studies from other fields of research tending in this direction, such as recent objective quality assessment metrics for volumetric videos represented as textured meshes [NDD*23, MNGL23]. We can also refer to the realism metric for generated LiDAR point clouds proposed by Triess et al. [TRPZ22] or the deep learning-based perceptual metric for trees modeling introduced by Polasek et al. [PHBČ21]. Inspired by these works and the surge and high-performance of deep learning approaches in VH animation modeling, we believe that investigating a unified DL-based metric for VH realism would be particularly interesting, with the goal of automatically quantifying perceptual realism along the dimensions mentioned above (Section 3). In the process of building such realism metric, we also believe that the task has to be considered. Indeed, as described in Section 3, animation features are task-specific (e.g., walking [PMKO09, NOZ*20], dancing [TLKS08, Su16]). A possible way to consider the specificity of the task is to use the concept of motion complexity [YLYD10, PGB14, SLCK17] which helps describing the nature of the motion. Then, it would be of interest to relate the complexity of motion with the level of realism. In addition, the conditions of viewing (e.g., the camera viewpoint) should also be integrated into the metrics since they can affect the user perception of realism [EPO11].

Based on the review of the literature presented in this survey, we also suggest another direction for research which is related to the identification of the features influencing the most the perceived realism of the VH animations. As shown in Table 1, there is no integrative approach that evaluates the various aspects of VH animation with respect to human perception of realism. Instead, each study focuses on a particular feature, taking advantage of the need to know if there is ultimately a weighting of these parameters to ensure VH animation realism. We can draw inspiration from previous works which have explored which body parts were the most salient when evaluating a motion [HTMS04, TLKS08, PMKO09]. Other aspects such as the ones related to the properties of motion revealed by Laban Motion Analysis can also be included since Durupinar et al. [Dur21] showed that the Flow element is more perceptually discernible than other LMA effort factors when comparing to unmodified motions. We believe that these perceptual features could be promising features for explicitly estimating VH realism, and could be considered for the creation of a level of details function and help to define the weight of the features to be considered when creating a realism metric.

Finally, we acknowledge that we did not conduct a systematic review and thus our narrative review might show selection bias. Nevertheless, we would like to highlight that our goal was to understand the main features and evaluations that have been conducted by authors about the realism of VHs animations, and can serve as a starting point for researchers interested in this topic.

To conclude, we have attempted in this survey to provide a global vision around the question of the realism of animated virtual humans, in particular through a study of existing definitions and of the multiple factors that have been used in the literature to increase or study realism of VHs. We believe that further advances on the remaining challenges around this topic will be achieved through multi-disciplinary collaborative research between Computer Graphics, Biomechanics and Neurosciences researchers, in order to analyse and identify human motion invariants as well as to propose novel evaluation methodologies.

6. Acknowledgement

This work was funded by the French National Research Agency (ANR) 3DMOVE - 19-CE23-0013.

Study	Situational context or Movement type	Global Kinematics-based features	Local Kinematics-based features	Physics-based features	3D geometrical features	Interaction-based features	Details	Objective Evaluation	Subjective Evaluation	Hybrid Evaluation	Details
[HOT98]	running motion				✓		shape information		✓		user self-report
[GLM01]	communication gestures		✓				motion of hands and arms	✓			Human motion laws
[RP03]	ballistic motion			✓			physical errors (velocity error, acceleration and deceleration addition)		✓		user self-report
[WB04]	walking, standing to walking, and running to turning	✓					duration of transition between animation clips		✓		user self-report
[BSH*05]	Observing and approaching an agent					✓	different visual representations and head movements of the agent (static, random, mimic, recorded)		✓		user self-report and behavioral response
[MNO07]	kungfu kick, jumping jack, walking and jogging	✓					linear velocity, motion complexity (+number of characters in the background)		✓		user self-report
[CHK07]	running motion				✓		degrees of anthromorphism		✓		user self-report
[WBCB08]	facial motion		✓		✓		animation technique, presence of teeth and eyes, motion blur face geometry (shape and texture blurring)		✓		user self-report
[RAP08]	ballistic motion			✓			gravity errors		✓		user self-report
[WFH09]	walking motion			✓			relative magnitude of lower-body joint torque, angular momentum, head stabilization	✓			data-driven
[MJH*09]	motions w/wo little male or female characteristics				✓		VH sex		✓		user self-report
[MJM*09]	emotions				✓		shape information		✓		user self-report
[GVDB-VBE10]	labor-intensive motions (e.g. jumping or knee bending) and relaxed motions (e.g. waving)			✓			total amount of external force, moment, and excess muscle force	✓			human motion law
[JHO10]	counting, drinking from a bottle, pointing, and snapping		✓				errors in synchronization between finger and body motion, method of animation		✓		user self-report
[HJO*10]	VH motion anomalies		✓				anomalies of face or body or both		✓		user self-report
[LN12]	"random" body motion and full body motion with more posture gesture mergers		✓				merging whole body motion with gestures		✓		user self-report
[HRMO12]	finger motion		✓				method of animation		✓		user self-report
[WHDK12]	walking and running			✓			inclusion of Hill-type musculotendon units in traditional torque-based approaches	✓			data-driven
[HMO12]	pushing motion					✓	w/wo mismatch between the pushing motion and the reaction		✓		user self-report
[GVD-PVDS13]	locomotion with speed variation, steering, uneven terrain and external perturbations			✓			muscle routing geometry	✓			data-driven
[WJZ13]	finger motion		✓				hand features (principal components of hand motion sequence)	✓			data-driven
[HRZ*13]	walking, jogging and dancing motion				✓		different gaits and by sex		✓		user self-report
[VHBO14]	throwing motion					✓	speed of release, thrower motion, angle of release (+representation of the thrower)		✓		user self-report
[TEW15]	walking motion	✓					path deviation, stopping motion and changing motion accelerations		✓		user self-report
[Su16]	dancing motion		✓				metrical periodicities from limb or trunk tempo or global tempo		✓		user self-report
[KHT16]	walking motion				✓		alignment between the anthropometry (body shape and proportions) and kinematics (movement patterns)		✓		user self-report
[CVB16]	greeting encounters					✓	smiling, gazing, proximity		✓		user self-report and behavioral response

Study	Situational context or Movement type	Global Kinematics-based features	Local Kinematics-based features	Physics-based features	3D geometrical features	Interaction-based features	Details	Objective Evaluation	Subjective Evaluation	Hybrid Evaluation	Details
[HOKP16]	shoulder motion during crowd interaction					✓	adding local shoulder rotation to the collision avoidance behaviour		✓		user self-report
[KPC17]	crowd motion					✓	collision avoidance ability, gazing, waving		✓		user self-report, behavioral response
[VRDB18]	conversation with a virtual patient					✓	w/wo conversational motion features (dynamic eye-gaze, mutual gaze, facial expression, lip-sync, head and body motion)		✓		behavioral response
[TGRJ19]	facial expression of pain		✓				order of action units (temporal sequence)		✓		user self-report
[JVWDGL19]	jumping, swinging, walking and running			✓			transformation of the metabolic energy function into a nonlinear function of joint torques	✓			data-driven
[KMH*19]	pushing, throwing, and lifting objects				✓		geometry of the articulated structure and distribution of skinning weights over that structure		✓		user self-report
[ZMM19]	standing body motion with gestures, facial motion					✓	emotional monologue, gaze interaction with the user		✓		behavioral response, user self-report
[TJM20]	trunk and limbs motions for VH pain		✓				combining trunk motion and pain facial expressions		✓		user self-report
[NOZ*20]	walking motion		✓				walk ratio		✓		user self-report
[HNH*20]	hand motion		✓				kinematic jerk		✓		user self-report
[SLVC20]	full-body motion				✓		perception pattern distortion on the body and clothes visual artifacts (i.e., blurriness) and level of detail, and resolution of facial features		✓		user self-report
[TBM*20]	walking motion				✓		sexual dismorphism		✓		user self-report
[TJJ*21]	facial pain expression		✓				combining trunk motion and facial motion		✓		user self-report
[MZCD21]	facial expression				✓		particular blendshapes for facial animation		✓		user self-report
[CDMH*21]	crowd motion	✓				✓	validation of a quality metric based on expert report (walking speed, trajectory length, flickering direction, time to collision or closest approach, local density, personal space overlap...)			✓	user self-report, data-driven approach
[MRP21]	crowd motion					✓	animation variety		✓		user self-report
[SDK21]	hand motion trajectories		✓				balance between variation and accuracy og generated trajectory	✓			data-driven
[AGWH22]	face-to-face interaction during a maze task					✓	VH head nodding motion		✓		user self-report, behavioral response
[YO22]	throwing motion					✓	point of release (+viewing condition)		✓		user self-report
[JRZ*23]	catching attention through waving motion					✓	viewpoint-dependent motion warping units (visibility, spatial extent)		✓		user self-report, behavioral response
[DOS*23]	conversational VH					✓	head movement, gaze orientation, facial expressions			✓	user-self report, human motion laws, data-driven approach
[RSM*23]	transitive actions and walking motions				✓		observed action		✓		user self-report
[SCIC*23]	passive or active viewing of a VH					✓	body motion, eye contact, meaningful questions		✓		user self-report, behavioral response

Table 1: *A chronological table summarizing works presented in this survey that were interested in the realism of VH animations. The third column outlines the features manipulated by authors to influence the realism of VH animation (Section 3) and the fourth column reports the different categories of evaluation used by authors to assess the effect of those manipulations (Section 4). Please note that in this table, we included the studies cited in this survey that contribute both to the manipulation of specific features of realism and their evaluation.*

References

[AAB*13] ANDERSON K., ANDRÉ E., BAUR T., BERNARDINI S., CHOLLET M., CHRYSSAFIDOU E., DAMIAN I., ENNIS C., EGGES A., GEBHARD P., ET AL.: The tardis framework: intelligent virtual agents for social coaching in job interviews. In *International conference on advances in computer entertainment technology* (2013), Springer, pp. 476–491. 1

[ABS*05] ALEXANDER A. L., BRUNYÉ T., SIDMAN J., WEIL S. A., ET AL.: From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games. *DARWARS Training Impact Group 5* (2005), 1–14. 4

[AC99] AGGARWAL J. K., CAI Q.: Human motion analysis: A review. *Computer vision and image understanding 73*, 3 (1999), 428–440. 3

[ADGY04] ATKINSON A. P., DITTRICH W. H., GEMMELL A. J., YOUNG A. W.: Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception 33*, 6 (2004), 717–746. 4

[AF02] ARIKAN O., FORSYTH D. A.: Interactive motion generation from examples. *ACM Trans. Graph. 21*, 3 (2002), 483–490. doi:10.1145/566654.566606. 5

[AGWH22] ABURUMMAN N., GILLIES M., WARD J. A., HAMILTON A. F. D. C.: Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies 164* (2022), 102819. 10, 13, 18

[ANM*11] ALAERTS K., NACKAERTS E., MEYNS P., SWINNEN S. P., WENDEROTH N.: Action and emotion recognition from point light displays: an investigation of gender differences. *PloS one 6*, 6 (2011), e20989. 4

[AOGT15] ALIAGA C., O'SULLIVAN C., GUTIERREZ D., TAMSTORF R.: Sackcloth or silk? the impact of appearance vs dynamics on the perception of animated cloth. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception* (2015), SAP '15, p. 41–46. doi:10.1145/2804408.2804412. 15

[APA*14] ASENSIO J. M. L., PERALTA J., ARRABALES R., BEDIA M. G., CORTEZ P., PEÑA A. L.: Artificial intelligence approaches for the generation and assessment of believable human-like behaviour in virtual characters. *Expert Systems with Applications 41*, 16 (2014), 7281–7290. 3

[ASC*15] ARISTIDOU A., STAVRAKIS E., CHARALAMBOUS P., CHRYSANTHOU Y., HIMONA S. L.: Folk dance evaluation using laban movement analysis. *Journal on Computing and Cultural Heritage (JOCCH) 8*, 4 (2015), 1–19. 11

[AVR21] AGUINIS H., VILLAMOR I., RAMANI R. S.: Mturk research: Review and recommendations. *Journal of Management 47*, 4 (2021), 823–837. 12

[BB04] BAILENSON J. N., BLASCOVICH J.: Avatars. *Encyclopedia of human-computer interaction 1* (2004), 64–6. 4

[BBBL01] BAILENSON J. N., BLASCOVICH J., BEALL A. C., LOOMIS J. M.: Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments 10*, 6 (2001), 583–598. 2, 12

[BC15] BÜTTNER M., CLAVET S.: Motion matching-the road to next gen animation. *Proc. of Nucl. ai 1*, 2015 (2015), 2. 5

[BDS23] BERG C., DIEKER L., SCOLAVINO R.: Using a virtual avatar teaching simulation and an evidence-based teacher observation tool: A synergistic combination for teacher preparation. *Education Sciences 13*, 7 (2023), 744. 1

[BHB03] BIOCCA F., HARMS C., BURGOON J. K.: Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments 12*, 5 (2003), 456–480. 12

[BJ03] BROGAN D. C., JOHNSON N. L.: Realistic human walking paths. In *Proceedings 11th IEEE International Workshop on Program Comprehension* (2003), IEEE, pp. 94–101. 11

[BJS*22] BILLEWAR S. R., JADHAV K., SRIRAM V., ARUN D. A., MOHD ABDUL S., GULATI K., BHASIN D. N. K. K.: The rise of 3d e-commerce: the online shopping gets real with virtual reality and augmented reality during covid-19. *World Journal of Engineering 19*, 2 (2022), 244–253. 1

[BKCZ09] BARTNECK C., KULIĆ D., CROFT E., ZOGHBI S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics 1* (2009), 71–81. 12

[Bla02] BLASCOVICH J.: Social influence within immersive virtual environments. In *The social life of avatars: Presence and interaction in shared virtual environments*. Springer, 2002, pp. 127–145. 13

[BLB*02] BLASCOVICH J., LOOMIS J., BEALL A. C., SWINTH K. R., HOYT C. L., BAILENSON J. N.: Immersive virtual environment technology as a methodological tool for social psychology. *Psychological inquiry 13*, 2 (2002), 103–124. 4, 9

[BMC21] BOURGAIZE S. M., MCFADYEN B. J., CINELLI M. E.: Collision avoidance behaviours when circumventing people of different sizes in various positions and locations. *Journal of Motor Behavior 53*, 2 (2021), 166–175. 13

[BMCB11] BALASUBRAMANIAN S., MELENDEZ-CALDERON A., BURDET E.: A robust and sensitive metric for quantifying movement smoothness. *IEEE transactions on biomedical engineering 59*, 8 (2011), 2126–2136. 12

[BS07] BLAKE R., SHIFFRAR M.: Perception of human motion. *Annu. Rev. Psychol. 58* (2007), 47–73. 4

[BSH*05] BAILENSON J. N., SWINTH K., HOYT C., PERSKY S., DIMOV A., BLASCOVICH J.: The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence 14*, 4 (2005), 379–393. 9, 13, 17

[BX16] BI W., XIAO B.: Perceptual constancy of mechanical properties of cloth under variation of external forces. In *Proceedings of the ACM Symposium on Applied Perception* (2016), SAP '16, p. 19–23. doi:10.1145/2931002.2931016. 15

[BYV20] BOGAERS A., YUMAK Z., VOLK A.: Music-driven animation generation of expressive musical gestures. In *Companion Publication of the 2020 International Conference on Multimodal Interaction* (2020), pp. 22–26. 12

[CCB01] CAMPANELLA S., CHRYSOCHOOS A., BRUYER R.: Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual cognition 8*, 2 (2001), 237–262. 12

[CCZB00] CHI D., COSTA M., ZHAO L., BADLER N.: The emote model for effort and shape. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (2000), pp. 173–182. 3

[CDFG*22] COOKS E. J., DUKE K. A., FLOOD-GRADY E., VILARO M. J., GHOSH R., PARKER N., TE P., GEORGE T. J., LOK B. C., WILLIAMS M., ET AL.: Can virtual human clinicians help close the gap in colorectal cancer screening for rural adults in the united states? the influence of rural identity on perceptions of virtual human clinicians. *Preventive Medicine Reports 30* (2022), 102034. 1

[CDMH*21] CABRERO-DANIEL B., MARQUES R., HOYET L., PETTRÉ J., BLAT J.: A perceptually-validated metric for crowd trajectory quality evaluation. *Proceedings of the ACM on Computer Graphics and Interactive Techniques 4*, 3 (2021), 1–18. 10, 14, 15, 18

[CFOC05] CHAMINADE T., FRANKLIN D. W., OZTOP E., CHENG G.: Motor interference between humans and humanoid robots: Effect of biological and artificial motion. In *Proceedings. The 4th International Conference on Development and Learning, 2005* (2005), IEEE, pp. 96–101. 4

[CGMB20] Carrozzino M. A., Galdieri R., Machidon O. M., Bergamasco M.: Do virtual humans dream of digital sheep? *IEEE Computer Graphics and Applications 40*, 4 (2020), 71–83. 1

[CHK07] Chaminade T., Hodgins J., Kawato M.: Anthropomorphism influences perception of computer-animated characters' actions. *Social cognitive and affective neuroscience 2*, 3 (2007), 206–216. 4, 7, 17

[CPJ*13] Cheetham M., Pavlovic I., Jordan N., Suter P., Jancke L.: Category processing and the human likeness dimension of the uncanny valley hypothesis: eye-tracking data. *Frontiers in psychology 4* (2013), 108. 13

[CRPPD17] Cruz Ruiz A., Pontonnier C., Pronost N., Dumont G.: Muscle-based control for character animation. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 122–147. 2, 3, 5, 7

[CSJ11] Cheetham M., Suter P., Jäncke L.: The human likeness dimension of the "uncanny valley hypothesis": behavioral and functional mri findings. *Frontiers in human neuroscience 5* (2011), 126. 2, 4, 13

[CTC*18] Coskun H., Tan D. J., Conjeti S., Navab N., Tombari F.: Human motion analysis with deep metric learning. In *Proceedings of the European conference on computer vision (ECCV)* (2018), pp. 667–683. 11

[CVB16] Cafaro A., Vilhjálmsson H. H., Bickmore T.: First impressions in human–agent virtual encounters. *ACM Transactions on Computer-Human Interaction (TOCHI) 23*, 4 (2016), 1–40. 10, 13, 17

[CZXL09] Chen C., Zhuang Y., Xiao J., Liang Z.: Perceptual 3d pose distance estimation by boosting relational geometric features. *Computer Animation and Virtual Worlds 20*, 2-3 (2009), 267–277. 13, 14

[dBdG15] de Borst A. W., de Gelder B.: Is it the real deal? perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in psychology 6* (2015), 576. 3, 4

[DGM97] Das G., Gunopulos D., Mannila H.: Finding similar time series. In *European Symposium on Principles of Data Mining and Knowledge Discovery* (1997), Springer, pp. 88–100. 14

[DHOO05] Dobbyn S., Hamill J., O'Conor K., O'Sullivan C.: Geopostors: A real-time geometry / impostor crowd rendering system. In *Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games* (2005), I3D '05, p. 95–102. doi:10.1145/1053427.1053443. 4

[DOS*23] Delbosc A., Ochs M., Sabouret N., Ravenet B., Ayache S.: Towards the generation of synchronized and believable non-verbal facial behaviors of a talking virtual agent. In *International Cconference on Multimodal Interaction*. 2023, pp. 228–237. 12, 18

[Dur21] Durupinar F.: Perception of human motion similarity based on laban movement analysis. In *ACM Symposium on Applied Perception 2021* (2021), pp. 1–7. 11, 13, 14, 16

[EPO08] Ennis C., Peters C., O'Sullivan C.: Perceptual evaluation of position and orientation context rules for pedestrian formations. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization* (2008), pp. 75–82. 10, 12

[EPO11] Ennis C., Peters C., O'sullivan C.: Perceptual effects of scene context and viewpoint for virtual pedestrian crowds. *ACM Transactions on Applied Perception (TAP) 8*, 2 (2011), 1–22. 10, 16

[EST*20] Egger B., Smith W. A., Tewari A., Wuhrer S., Zollhoefer M., Beeler T., Bernard F., Bolkart T., Kortylewski A., Romdhani S., et al.: 3d morphable face models—past, present, and future. *ACM Transactions on Graphics (ToG) 39*, 5 (2020), 1–38. 2

[FBH*22] Fraser A. D., Branson I., Hollett R. C., Speelman C. P., Rogers S. L.: Expressiveness of real-time motion captured avatars influences perceived animation realism and perceived quality of social interaction in virtual reality. *Frontiers in Virtual Reality 3* (2022), 981400. 8

[Fit54] Fitts P. M.: The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology 47*, 6 (1954), 381. 12

[FLM20] Fiset F., Lamontagne A., McFadyen B. J.: Limb movements of another pedestrian affect crossing distance but not path planning during virtual over ground circumvention. *Neuroscience letters 736* (2020), 135278. 5

[Gav99] Gavrila D. M.: The visual analysis of human movement: A survey. *Computer vision and image understanding 73*, 1 (1999), 82–98. 3

[GBH23] Grigorev A., Black M. J., Hilliges O.: Hood: Hierarchical graphs for generalized modelling of clothing dynamics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 16965–16974. 15

[GBR*12] Gretton A., Borgwardt K. M., Rasch M. J., Schölkopf B., Smola A.: A kernel two-sample test. *The Journal of Machine Learning Research 13*, 1 (2012), 723–773. 11

[GCM*22] Gonçalves G., Coelho H., Monteiro P., Melo M., Bessa M.: Systematic review of comparative studies of the impact of realism in immersive virtual experiences. *ACM Computing Surveys 55*, 6 (2022), 1–36. 3, 10, 13

[GL02] Giese M., Lappe M.: Measurement of generalization fields for the recognition of biological motion. *Vision research 42*, 15 (2002), 1847–1858. 4

[GLM01] Gibet S., Lebourque T., Marteau P.-F.: High-level specification and animation of communicative gestures. *Journal of Visual Languages & Computing 12*, 6 (2001), 657–687. 12, 17

[GP03] Giese M. A., Poggio T.: Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience 4*, 3 (2003), 179–192. 4

[GPEO11] Geijtenbeek T., Pronost N., Egges A., Overmars M. H.: Interactive character animation using simulated physics. *Eurographics (State of the Art Reports)* (2011), 127–149. 2, 3, 5, 7

[GSV*03] Garau M., Slater M., Vinayagamoorthy V., Brogni A., Steed A., Sasse M. A.: The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2003), pp. 529–536. 8

[GVDBL*12] Guy S. J., Van Den Berg J., Liu W., Lau R., Lin M. C., Manocha D.: A statistical similarity measure for aggregate crowd dynamics. *ACM Transactions on Graphics (TOG) 31*, 6 (2012), 1–11. 14

[GVDBVBE10] Geijtenbeek T., Van Den Bogert A., Van Basten B., Egges A.: Evaluating the physical realism of character animations using musculoskeletal models. In *International Conference on Motion in Games* (2010), Springer, pp. 11–22. 7, 12, 17

[GVDPVDS13] Geijtenbeek T., Van De Panne M., Van Der Stappen A. F.: Flexible muscle-based locomotion for bipedal creatures. *ACM Transactions on Graphics (TOG) 32*, 6 (2013), 1–11. 7, 17

[HBS03] Hoyt C. L., Blascovich J., Swinth K. R.: Social inhibition in immersive virtual environments. *Presence 12*, 2 (2003), 183–195. 13

[HGW*23] Hassan M., Guo Y., Wang T., Black M., Fidler S., Peng X. B.: Synthesizing physical character-scene interactions. *arXiv preprint arXiv:2302.00883* (2023). 8, 11

[Hil38] Hill A. V.: The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B-Biological Sciences 126*, 843 (1938), 136–195. 7

[Hil00] Hill R.: Perceptual attention in virtual humans: Toward realistic and believable gaze behaviors. In *Proceedings of the AAAI Fall Symposium on Simulating Human Agents* (2000), AAAI Press Menlo Park, pp. 46–52. 3

[HJO*10] HODGINS J., JÖRG S., O'SULLIVAN C., PARK S. I., MAHLER M.: The saliency of anomalies in animated human characters. *ACM Transactions on Applied Perception (TAP) 7*, 4 (2010), 1–14. 6, 13, 17

[HKPP20] HOLDEN D., KANOUN O., PEREPICHKA M., POPA T.: Learned motion matching. *ACM Trans. Graph. 39*, 4 (2020). doi:10.1145/3386569.3392440. 5

[HM10] HO C.-C., MACDORMAN K. F.: Revisiting the uncanny valley theory: Developing and validating an alternative to the godspeed indices. *Computers in Human Behavior 26*, 6 (2010), 1508–1518. 2, 12

[HM17] HO C.-C., MACDORMAN K. F.: Measuring the uncanny valley effect: Refinements to indices for perceived humanness, attractiveness, and eeriness. *International Journal of Social Robotics 9* (2017), 129–139. 12

[HMD*23] HAN E., MILLER M. R., DEVEAUX C., JUN H., NOWAK K. L., HANCOCK J. T., RAM N., BAILENSON J. N.: People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication 28*, 2 (2023), zmac031. 13

[HMO12] HOYET L., MCDONNELL R., O'SULLIVAN C.: Push it real: Perceiving causality in virtual interactions. *ACM Trans. Graph. 31*, 4 (2012). doi:10.1145/2185520.2185586. 10, 17

[HNH*20] HIROSE J., NISHIKAWA A., HORIBA Y., INUI S., PATAKY T. C.: Integrated jerk as an indicator of affinity for artificial agent kinematics: laptop and virtual reality experiments involving index finger motion during two-digit grasping. *PeerJ 8* (2020), e9843. 6, 18

[HOB20] HERRERA F., OH S. Y., BAILENSON J. N.: Effect of behavioral realism on social interactions inside collaborative virtual environments. *Presence 27*, 2 (2020), 163–182. 2, 8, 13

[HOKP16] HOYET L., OLIVIER A.-H., KULPA R., PETTRÉ J.: Perceptual effect of shoulder motions on crowd animations. *ACM Trans. Graph. 35*, 4 (2016). doi:10.1145/2897824.2925931. 10, 12, 18

[HOP09] HERTZMANN A., O'SULLIVAN C., PERLIN K.: Realistic human body movement for emotional expressiveness. In *ACM SIGGRAPH 2009 Courses*. 2009, pp. 1–27. 4

[HOT98] HODGINS J. K., O'BRIEN J. F., TUMBLIN J.: Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics 4*, 4 (1998), 307–316. 7, 13, 17

[HRMO12] HOYET L., RYALL K., MCDONNELL R., O'SULLIVAN C.: Sleight of hand: Perception of finger motion from reduced marker sets. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (2012), I3D '12, p. 79–86. doi:10.1145/2159616.2159630. 6, 17

[HRN*20] HOPPE M., ROSSMY B., NEUMANN D. P., STREUBER S., SCHMIDT A., MACHULLA T.-K.: A human touch: Social touch increases the perceived human-likeness of agents in virtual reality. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (2020), pp. 1–11. 4

[HRS08] HARRIGAN J., ROSENTHAL R., SCHERER K.: *New handbook of methods in nonverbal behavior research*. Oxford University Press, 2008. 9

[HRVDP04] HARRISON J., RENSINK R. A., VAN DE PANNE M.: Obscuring length changes during animated motion. *ACM Transactions on Graphics (TOG) 23*, 3 (2004), 569–573. 13

[HRZ*13] HOYET L., RYALL K., ZIBREK K., PARK H., LEE J., HODGINS J., O'SULLIVAN C.: Evaluating the distinctiveness and attractiveness of human motions on realistic virtual bodies. *ACM Transactions on Graphics (TOG) 32*, 6 (2013), 1–11. 8, 13, 15, 17

[HTMS04] HARADA T., TAOKA S., MORI T., SATO T.: Quantitative evaluation method for pose and motion similarity based on human perception. In *4th IEEE/RAS International Conference on Humanoid Robots, 2004.* (2004), vol. 1, IEEE, pp. 494–512. 13, 14, 16

[HZCM23] HIGGINS D., ZHAN Y., COWAN B. R., MCDONNELL R.: Investigating the effect of visual realism on empathic responses to emotionally expressive virtual humans. In *ACM Symposium on Applied Perception 2023* (2023), pp. 1–7. 8

[JADJ22] JUSTICE J., ADKINS A., DONG T., JÖRG S.: Do we measure what we perceive? comparison of perceptual and computed differences between hand animations. In *ACM SIGGRAPH 2022 Posters*. 2022, pp. 1–2. 13, 14

[JHO10] JÖRG S., HODGINS J., O'SULLIVAN C.: The perception of finger motions. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (2010), pp. 129–133. 6, 13, 17

[JL21] JUNG S., LINDEMAN R. W.: Perspective: Does realism improve presence in vr? suggesting a model and metric for vr experience evaluation. *Frontiers in virtual reality 2* (2021), 693327. 4

[Joh73] JOHANSSON G.: Visual perception of biological motion and a model for its analysis. *Perception & psychophysics 14* (1973), 201–211. 4

[JRZ*23] JOVANE A., RAIMBAUD P., ZIBREK K., PACCHIEROTTI C., CHRISTIE M., HOYET L., OLIVIER A.-H., PETTRÉ J.: Warping character animations using visual motion features. *Computers & Graphics 110* (2023), 38–48. 10, 18

[JTDP06] JOSHI P., TIEN W. C., DESBRUN M., PIGHIN F.: Learning controls for blend shape based realistic facial animation. In *ACM Siggraph 2006 Courses*. 2006, pp. 17–es. 3

[JVWDGL19] JIANG Y., VAN WOUWE T., DE GROOTE F., LIU C. K.: Synthesis of biologically realistic human motion using joint torque actuation. *ACM Transactions On Graphics (TOG) 38*, 4 (2019), 1–12. 7, 18

[KAK*22] KWIATKOWSKI A., ALVARADO E., KALOGEITON V., LIU C. K., PETTRÉ J., VAN DE PANNE M., CANI M.-P.: A survey on reinforcement learning methods in character animation. In *Computer Graphics Forum* (2022), vol. 41, Wiley Online Library, pp. 613–639. 2, 3

[KDC*08] KAVAN L., DOBBYN S., COLLINS S., ŽÁRA J., O'SULLIVAN C.: Polypostors: 2d polygonal impostors for 3d crowds. In *Proceedings of the 2008 Symposium on Interactive 3D Graphics and Games* (2008), I3D '08, p. 149–155. doi:10.1145/1342250.1342273. 4

[KFMT15] KÄTSYRI J., FÖRGER K., MÄKÄRÄINEN M., TAKALA T.: A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in psychology 6* (2015), 390. 2

[KGP02] KOVAR L., GLEICHER M., PIGHIN F.: Motion graphs. *ACM Trans. Graph. 21*, 3 (2002), 473–482. doi:10.1145/566654.566605. 5

[KHG*07] KENNY P., HARTHOLT A., GRATCH J., SWARTOUT W., TRAUM D., MARSELLA S., PIEPOL D.: Building interactive virtual humans for training environments. In *Proceedings of i/itsec* (2007), vol. 174, pp. 911–916. 3

[KHSG17] KHAN M., HUTTUNEN H., SUOMINEN O., GOTCHEV A.: Recurrent neural networks for modeling motion capture data. *spine 10* (2017), 200. 11

[KHT16] KLÜVER M., HECHT H., TROJE N. F.: Internal consistency predicts attractiveness in biological motion walkers. *Evolution and Human Behavior 37*, 1 (2016), 40–46. 8, 17

[KMG22] KYRLITSIAS C., MICHAEL-GRIGORIOU D.: Social interaction with agents and avatars in immersive virtual environments: A survey. *Frontiers in Virtual Reality 2* (2022), 786665. 9

[KMH*19] KENNY S., MAHMOOD N., HONDA C., BLACK M. J., TROJE N. F.: Perceptual effects of inconsistency in human animations. *ACM Transactions on Applied Perception (TAP) 16*, 1 (2019), 1–18. 4, 8, 15, 18

[KPC17] KYRIAKOU M., PAN X., CHRYSANTHOU Y.: Interaction with virtual crowd in immersive and semi-immersive virtual reality systems. *Computer Animation and Virtual Worlds 28*, 5 (2017), e1729. 10, 18

[KPP*21] KARUZAKI E., PARTARAKIS N., PATSIOURAS N., ZIDIANAKIS E., KATZOURAKIS A., PATTAKOS A., KAPLANIDI D., BAKA E., CADI N., MAGNENAT-THALMANN N., ET AL.: Realistic virtual humans for cultural heritage applications. *Heritage 4*, 4 (2021), 4148–4171. 1

[Kun06] KUNKLER K.: The role of medical simulation: an overview. *The International Journal of Medical Robotics and Computer Assisted Surgery 2*, 3 (2006), 203–210. 1

[LCR*02] LEE J., CHAI J., REITSMA P. S. A., HODGINS J. K., POLLARD N. S.: Interactive control of avatars animated with human motion data. *ACM Trans. Graph. 21*, 3 (2002), 491–500. doi:10.1145/566654.566607. 5

[Lik32] LIKERT R.: A technique for the measurement of attitudes. *Archives of psychology* (1932). 12

[LKM*18] LYNCH S. D., KULPA R., MEERHOFF L. A., PETTRÉ J., CRÉTUAL A., OLIVIER A.-H.: Collision avoidance behavior between walkers: global and local motion cues. *IEEE transactions on visualization and computer graphics 24*, 7 (2018), 2078–2088. 5

[LN12] LUO P., NEFF M.: A perceptual study of the relationship between posture and gesture for virtual characters. In *Motion in Games: 5th International Conference, MIG 2012, Rennes, France, November 15-17, 2012. Proceedings 5* (2012), Springer, pp. 254–265. 7, 14, 15, 17

[LPB*18] LYNCH S. D., PETTRÉ J., BRUNEAU J., KULPA R., CRÉTUAL A., OLIVIER A.-H.: Effect of virtual human gaze behaviour during an orthogonal collision avoidance walking task. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2018), IEEE, pp. 136–142. 4

[LRG*17] LATOSCHIK M. E., ROTH D., GALL D., ACHENBACH J., WALTEMATE T., BOTSCH M.: The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM symposium on virtual reality software and technology* (2017), pp. 1–10. 8

[LU71] LABAN R., ULLMANN L.: The mastery of movement. 11

[MBB12] MCDONNELL R., BREIDT M., BÜLTHOFF H. H.: Render me real? investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics (TOG) 31*, 4 (2012), 1–11. 2

[MBR17] MARTINEZ J., BLACK M. J., ROMERO J.: On human motion prediction using recurrent neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 2891–2900. 11

[MDC18] MACHIDON O. M., DUGULEANA M., CARROZZINO M.: Virtual humans in cultural heritage ict applications: A review. *Journal of Cultural Heritage 33* (2018), 249–260. 1

[MG01] MOESLUND T. B., GRANUM E.: A survey of computer vision-based human motion capture. *Computer vision and image understanding 81*, 3 (2001), 231–268. 3

[MHLC*22] MOUROT L., HOYET L., LE CLERC F., SCHNITZLER F., HELLIER P.: A survey on deep learning for skeleton-based human animation. In *Computer Graphics Forum* (2022), vol. 41, Wiley Online Library, pp. 122–157. 2, 3

[MJH*09] MCDONNELL R., JÖRG S., HODGINS J. K., NEWELL F., O'SULLIVAN C.: Evaluating the effect of motion and body shape on the perceived sex of virtual characters. *ACM Transactions on Applied Perception (TAP) 5*, 4 (2009), 1–14. 8, 17

[MJM*09] MCDONNELL R., JÖRG S., MCHUGH J., NEWELL F. N., O'SULLIVAN C.: Investigating the role of body shape on the perception of emotion. *ACM Transactions on Applied Perception (TAP) 6*, 3 (2009), 1–11. 7, 17

[MKR*21] MOUSAS C., KOILIAS A., REKABDAR B., KAO D., ANASTASLOU D.: Toward understanding the effects of virtual character appearance on avoidance movement behavior. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)* (2021), IEEE, pp. 40–49. 9

[MNGL23] MARVIE J.-E., NEHMÉ Y., GRAZIOSI D., LAVOUÉ G.: Crafting the mpeg metrics for objective and perceptual quality assessment of volumetric videos. *Quality and User Experience 8*, 1 (2023), 4. 15

[MNO07] MCDONNELL R., NEWELL F., O'SULLIVAN C.: Smooth movers: perceptually guided human motion simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation* (2007), pp. 259–269. 5, 12, 17

[Mor70] MORI M.: The uncanny valley: the original essay by masahiro mori. *IEEE Spectrum 6* (1970). 2

[MPM05] MORRIS J. P., PELPHREY K. A., MCCARTHY G.: Regional brain activation evoked when approaching a virtual human on a virtual walk. *Journal of cognitive neuroscience 17*, 11 (2005), 1744–1752. 4

[MRP21] MOLINA E., RÍOS A., PELECHANO N.: The impact of animations in the perception of a simulated crowd. In *Advances in Computer Graphics: 38th Computer Graphics International Conference, CGI 2021, Virtual Event, September 6–10, 2021, Proceedings 38* (2021), Springer, pp. 25–38. 10, 18

[MRW*23] MARSOT M., REKIK R., WUHRER S., FRANCO J.-S., OLIVIER A.-H.: Correspondence-free online human motion retargeting. *arXiv preprint arXiv:2302.00556* (2023). 3, 11, 13

[MTT05] MAGNENAT-THALMANN N., THALMANN D.: Virtual humans: thirty years of research, what next? *The Visual Computer 21* (2005), 997–1015. 2, 3, 4

[MWT*12] MEYER G. F., WONG L. T., TIMSON E., PERFECT P., WHITE M. D.: Objective fidelity evaluation in multisensory virtual environments: auditory cue fidelity in flight simulation. 4

[MZCD21] MCDONNELL R., ZIBREK K., CARRIGAN E., DAHYOT R.: Model for predicting perception of facial action unit activation using virtual humans. *Computers & Graphics 100* (2021), 81–92. 8, 12, 18

[NDD*23] NEHMÉ Y., DELANOY J., DUPONT F., FARRUGIA J.-P., LE CALLET P., LAVOUÉ G.: Textured mesh quality assessment: Large-scale dataset and deep learning-based quality metric. *ACM Transactions on Graphics 42*, 3 (2023), 1–20. 15

[NDP10] NIEWIADOMSKI R., DEMEURE V., PELACHAUD C.: Warmth, competence, believability and virtual agents. In *Intelligent Virtual Agents: 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings 10* (2010), Springer, pp. 272–285. 3

[NHH09] NUMMENMAA L., HYÖNÄ J., HIETANEN J. K.: I'll walk this way: Eyes reveal the direction of locomotion and make passersby look and go the other way. *Psychological science 20*, 12 (2009), 1454–1458. 4

[NKAM22] NELSON M. G., KOILIAS A., ANAGNOSTOPOULOS C.-N., MOUSAS C.: Effects of rendering styles of a virtual character on avoidance movement behavior. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2022), IEEE, pp. 594–599. 9

[NOZ*20] NIAY B., OLIVIER A.-H., ZIBREK K., PETTRÉ J., HOYET L.: Walk ratio: Perception of an invariant parameter of human walk on virtual characters. In *ACM Symposium on Applied Perception 2020* (2020), SAP '20. doi:10.1145/3385955.3407926. 6, 16, 18

[ODGK03] O'SULLIVAN C., DINGLIANA J., GIANG T., KAISER M. K.: Evaluating the visual fidelity of physically based animations. *ACM Transactions on Graphics (TOG) 22*, 3 (2003), 527–536. 4

[OKKHB20] OH KRUZIC C., KRUZIC D., HERRERA F., BAILENSON J.: Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific reports 10*, 1 (2020), 20626. 13

[PGB14] PENG Z., GENEWEIN T., BRAUN D. A.: Assessing randomness and complexity in human motion trajectories through analysis of symbolic sequences. *Frontiers in human neuroscience 8* (2014), 168. 16

[PH18] PAN X., HAMILTON A. F. D. C.: Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology 109*, 3 (2018), 395–417. 9

[PHA*07] PHAM Q.-C., HICHEUR H., ARECHAVALETA G., LAUMOND J.-P., BERTHOZ A.: The formation of trajectories during goal-oriented locomotion in humans. ii. a maximum smoothness model. *European Journal of Neuroscience 26*, 8 (2007), 2391–2403. 12

[PHBČ21] POLASEK T., HRUSA D., BENES B., ČADÍK M.: Ictree: automatic perceptual metrics for tree models. *ACM Transactions on Graphics (TOG) 40*, 6 (2021), 1–15. 15

[PKHS05] POLLICK F. E., KAY J. W., HEIM K., STRINGER R.: Gender recognition from point-light walkers. *Journal of Experimental Psychology: Human Perception and Performance 31*, 6 (2005), 1247. 4

[PMKO09] PRAŽÁK M., MCDONNELL R., KAVAN L., O'SULLIVAN C.: A perception based metric for comparing human locomotion. *Eurographics Ireland* (2009). 13, 14, 16

[PS18] PALAN S., SCHITTER C.: Prolific. ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance 17* (2018), 22–27. 12

[RAP08] REITSMA P. S., ANDREWS J., POLLARD N. S.: Effect of character animacy and preparatory motion on perceptual magnitude of errors in ballistic motion. In *Computer Graphics Forum* (2008), vol. 27, Wiley Online Library, pp. 201–210. 3, 7, 12, 17

[RBB*22] ROGERS S. L., BROADBENT R., BROWN J., FRASER A., SPEELMAN C. P.: Realistic motion avatars are the future for social interaction in virtual reality. *Frontiers in Virtual Reality 2* (2022), 750729. 8

[RKB*18] ROTH D., KULLMANN P., BENTE G., GALL D., LATOSCHIK M. E.: Effects of hybrid and synthetic social gaze in avatar-mediated interactions. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2018), IEEE, pp. 103–108. 13

[RP03] REITSMA P. S., POLLARD N. S.: Perceptual metrics for character animation: sensitivity to errors in ballistic motion. In *ACM SIGGRAPH 2003 Papers*. 2003, pp. 537–542. 7, 12, 14, 17

[RPA*15] RUHLAND K., PETERS C. E., ANDRIST S., BADLER J. B., BADLER N. I., GLEICHER M., MUTLU B., MCDONNELL R.: A review of eye gaze in virtual agents, social robotics and hci: Behaviour generation, user interaction and perception. In *Computer graphics forum* (2015), vol. 34, Wiley Online Library, pp. 299–326. 2, 3

[RPC*10] ROHMER D., POPA T., CANI M.-P., HAHMANN S., SHEFFER A.: Animation wrinkling: augmenting coarse cloth simulations with realistic-looking wrinkles. *ACM Transactions on Graphics (ToG) 29*, 6 (2010), 1–8. 15

[RPE*05] REN L., PATRICK A., EFROS A. A., HODGINS J. K., REHG J. M.: A data-driven approach to quantifying natural human motion. *ACM Transactions on Graphics (TOG) 24*, 3 (2005), 1090–1097. 3, 5, 14, 15

[RSM*23] RUSSELL J., SABOOR S., MAKKAR M., BARUA A., THALER A.: Detection of inconsistency between shape and motion in realistic female and male animation. *STEM Fellowship Journal 8*, 1 (2023), 17–23. 8, 13, 18

[RWV19] RIBET S., WANNOUS H., VANDEBORRE J.-P.: Survey on style in 3d human body motion: Taxonomy, data, recognition and its applications. *IEEE Transactions on Affective Computing 12*, 4 (2019), 928–948. 2, 3

[SAS*22] STUART J., AUL K., STEPHEN A., BUMBACH M. D., LOK B.: The effect of virtual human rendering style on user perceptions of visual cues. *Frontiers in Virtual Reality 3* (2022), 864676. 2

[SC12] STOUT D., CHAMINADE T.: Stone tools, language and the brain in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences 367*, 1585 (2012), 75–87. 4

[SCI*12] SAYGIN A. P., CHAMINADE T., ISHIGURO H., DRIVER J., FRITH C.: The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social cognitive and affective neuroscience 7*, 4 (2012), 413–422. 4

[SCIC*23] STERNA R., CYBULSKI A., IGRAS-CYBULSKA M., PILARCZYK J., SEGIET N., KUNIECKI M.: How behavioral, photographic, and interactional realism influence the sense of co-presence in vr. an investigation with psychophysiological measurement. *International Journal of Human–Computer Interaction* (2023), 1–16. 13, 18

[SDK21] SHEN J., DUDLEY J., KRISTENSSON P. O.: Simulating realistic human motion trajectories of mid-air gesture typing. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2021), IEEE, pp. 393–402. 11, 12, 18

[SFJ01] SCHRICKER B. C., FRANCESCHINI R. W., JOHNSON T. C.: Fidelity evaluation framework. In *Proceedings. 34th Annual Simulation Symposium* (2001), IEEE, pp. 109–116. 4

[SHS*23] SHETTY A., HABERMANN M., SUN G., LUVIZON D., GOLYANIK V., THEOBALT C.: Holoported characters: Real-time free-viewpoint rendering of humans from sparse rgb cameras. *arXiv preprint arXiv:2312.07423* (2023). 3

[SKP15] SCHROFF F., KALENICHENKO D., PHILBIN J.: Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 815–823. 11

[SLBA22] STEIN J.-P., LINDA BREVES P., ANDERS N.: Parasocial interactions with real and virtual influencers: The role of perceived similarity and human-likeness. *New Media & Society* (2022), 14614448221102900. 4

[SLCK17] SUH I. H., LEE S. H., CHO N. J., KWON W. Y.: Measuring motion significance and motion complexity. *Information Sciences 388* (2017), 84–98. 16

[SLVC20] SUBRAMANYAM S., LI J., VIOLA I., CESAR P.: Comparing the quality of highly realistic digital humans in 3dof and 6dof: A volumetric video case study. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2020), IEEE, pp. 127–136. 8, 12, 18

[SOB*15] SANZ F. A., OLIVIER A.-H., BRUDER G., PETTRÉ J., LÉCUYER A.: Virtual proxemics: Locomotion in the presence of obstacles in large immersive projection environments. In *2015 IEEE virtual reality (vr)* (2015), IEEE, pp. 75–80. 13

[SOC19] SANTESTEBAN I., OTADUY M. A., CASAS D.: Learning-based animation of clothing for virtual try-on. In *Computer Graphics Forum* (2019), vol. 38, Wiley Online Library, pp. 355–366. 15

[SP18] SUH A., PROPHET J.: The state of immersive technology research: A literature analysis. *Computers in Human behavior 86* (2018), 77–90. 1

[Su16] SU Y.-H.: Visual tuning and metrical perception of realistic point-light dance movements. *Scientific reports 6*, 1 (2016), 22774. 6, 13, 16, 17

[TBM*20] THALER A., BIEG A., MAHMOOD N., BLACK M. J., MOHLER B. J., TROJE N. F.: Attractiveness and confidence in walking style of male and female virtual characters. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (2020), IEEE, pp. 678–679. 8, 18

[TEW15] TURNWALD A., EGER S., WOLLHERR D.: Investigating similarity measures for locomotor trajectories based on the human perception of differences in motions. In *2015 IEEE International Workshop on Advanced Robotics and its Social Impacts (ARSO)* (2015), IEEE, pp. 1–6. 5, 14, 17

[TF88a] TERZOPOULOS D., FLEISCHER K.: Deformable models. *The visual computer 4*, 6 (1988), 306–331. 3, 15

[TF88b] TERZOPOULOS D., FLEISCHER K.: Modeling inelastic deformation: viscolelasticity, plasticity, fracture. In *Proceedings of the 15th annual conference on Computer graphics and interactive techniques* (1988), pp. 269–278. 3, 15

[TGRJ19] Tessier M.-H., Gingras C., Robitaille N., Jackson P. L.: Toward dynamic pain expressions in avatars: perceived realism and pain level of different action unit orders. *Computers in Human Behavior 96* (2019), 95–109. 6, 18

[TJJ*21] Treal T., Jackson P. L., Jeuvrey J., Vignais N., Meugnot A.: Natural human postural oscillations enhance the empathic response to a facial pain expression in a virtual character. *Scientific Reports 11*, 1 (2021), 12493. 6, 18

[TJM20] Treal T., Jackson P. L., Meugnot A.: Combining trunk movement and facial expression enhances the perceived intensity and believability of an avatar's pain expression. *Computers in Human Behavior 112* (2020), 106451. 6, 18

[TLKS08] Tang J. K., Leung H., Komura T., Shum H. P.: Emulating human perception of motion similarity. *Computer Animation and Virtual Worlds 19*, 3-4 (2008), 211–221. 5, 13, 14, 16

[TRL*17] Turner N., Reeves M., Letteri J., Lemmon D., Barrett D.: Weta digital vfx: War for the planet of the apes. In *ACM SIGGRAPH 2017 Computer Animation Festival* (2017), SIGGRAPH '17, p. 27. doi:10.1145/3078280.3114817. 1

[TRPZ22] Triess L. T., Rist C. B., Peter D., Zöllner J. M.: A realism metric for generated lidar point clouds. *International Journal of Computer Vision 130*, 12 (2022), 2962–2979. 15

[UP12] Ursu E. A., Pronost N.: A literature review on virtual character assessment. *Utrecht University* (2012). 2

[VCH*21] Villegas R., Ceylan D., Hertzmann A., Yang J., Saito J.: Contact-aware retargeting of skinned motion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 9720–9729. 13

[VF95] Viviani P., Flash T.: Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance 21*, 1 (1995), 32. 12

[VGS*06] Vinayagamoorthy V., Gillies M., Steed A., Tanguy E., Pan X., Loscos C., Slater M.: Building expression into virtual characters. 2, 5

[VHBO14] Vicovaro M., Hoyet L., Burigana L., O'sullivan C.: Perceptual evaluation of motion editing for realistic throwing animations. *ACM Transactions on Applied Perception (TAP) 11*, 2 (2014), 1–23. 3, 9, 12, 17

[Vin68] Vintsyuk T. K.: Speech discrimination by dynamic programming. *Cybernetics 4*, 1 (1968), 52–57. 11

[VKDB01] Vieilledent S., Kerlirzin Y., Dalbera S., Berthoz A.: Relationship between velocity and curvature of a human locomotor trajectory. *Neuroscience letters 305*, 1 (2001), 65–69. 11

[VRDB18] Volonte M., Robb A., Duchowski A. T., Babu S. V.: Empirical evaluation of virtual human conversational and affective animations on visual attention in inter-personal simulations. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)* (2018), IEEE, pp. 25–32. 10, 13, 18

[VT82] Viviani P., Terzuolo C.: Trajectory determines movement dynamics. *Neuroscience 7*, 2 (1982), 431–437. 11

[VWVBE*10] Van Welbergen H., Van Basten B. J., Egges A., Ruttkay Z. M., Overmars M. H.: Real time animation of virtual humans: a trade-off between naturalness and control. In *Computer Graphics Forum* (2010), vol. 29, Wiley Online Library, pp. 2530–2554. 2, 3, 7

[WB04] Wang J., Bodenheimer B.: Computing the duration of motion transitions: an empirical approach. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation* (2004), pp. 335–344. 5, 13, 17

[WBCB08] Wallraven C., Breidt M., Cunningham D. W., Bülthoff H. H.: Evaluating the perceptual realism of animated facial expressions. *ACM Transactions on Applied Perception (TAP) 4*, 4 (2008), 1–20. 5, 6, 8, 13, 17

[WBK*07] Ward K., Bertails F., Kim T.-Y., Marschner S. R., Cani M.-P., Lin M. C.: A survey on hair modeling: Styling, simulation, and rendering. *IEEE transactions on visualization and computer graphics 13*, 2 (2007), 213–234. 3

[WCC16] Wykowska A., Chaminade T., Cheng G.: Embodied artificial agents for understanding human social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences 371*, 1693 (2016), 20150375. 4

[WDR*22] Witte K., Droste M., Ritter Y., Emmermacher P., Masik S., Bürger D., Petri K.: Sports training in virtual reality to improve response behavior in karate kumite with transfer to real world. *Frontiers in Virtual Reality 3* (2022), 903021. 1

[WFH09] Wang J. M., Fleet D. J., Hertzmann A.: Optimizing walking controllers. *ACM Trans. Graph. 28*, 5 (2009). doi:10.1145/1618452.1618514. 7, 17

[WHDK12] Wang J. M., Hamner S. R., Delp S. L., Koltun V.: Optimizing locomotion controllers using biologically-based actuators and objectives. *ACM Transactions on Graphics (TOG) 31*, 4 (2012), 1–11. 7, 17

[WJZ13] Wheatland N., Jörg S., Zordan V.: Automatic hand-over animation using principle component analysis. In *Proceedings of Motion on Games* (2013), MIG '13, p. 197–202. doi:10.1145/2522628.2522656. 6, 17

[WOO16] Wang H., Ondřej J., O'Sullivan C.: Trending paths: A new semantic-level metric for comparing simulated and real crowd data. *IEEE transactions on visualization and computer graphics 23*, 5 (2016), 1454–1464. 11

[WSZ*22] Wang Y., Su Z., Zhang N., Xing R., Liu D., Luan T. H., Shen X.: A survey on metaverse: Fundamentals, security, and privacy. *IEEE Communications Surveys & Tutorials* (2022). 1

[WVE*22] Wang C.-C., Volonte M., Ebrahimi E., Liu K.-Y., Wong S.-K., Babu S. V.: An evaluation of native versus foreign communicative interactions on users' behavioral reactions towards affective virtual crowds. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2022), pp. 340–349. doi:10.1109/VR51125.2022.00053. 15

[XPB*21] Xiang D., Prada F., Bagautdinov T., Xu W., Dong Y., Wen H., Hodgins J., Wu C.: Modeling clothing as a separate layer for an animatable human avatar. *ACM Trans. Graph. 40*, 6 (2021). doi:10.1145/3478513.3480545. 15

[YLYD10] Yang Y., Leung H., Yue L., Deng L.: Evaluating human motion complexity based on un-correlation and non-smoothness. In *Advances in Multimedia Information Processing-PCM 2010: 11th Pacific Rim Conference on Multimedia, Shanghai, China, September 2010, Proceedings, Part II 11* (2010), Springer, pp. 538–548. 16

[YO22] Yamac G., O'Sullivan C.: Fauxthrow: Exploring the effects of incorrect point of release in throwing motions. In *ACM Symposium on Applied Perception 2022* (2022), pp. 1–5. 9, 18

[YZG*21] Yu T., Zheng Z., Guo K., Liu P., Dai Q., Liu Y.: Function4d: Real-time human volumetric capture from very sparse consumer rgbd sensors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2021), pp. 5746–5756. 3

[ZAJ*15] Zell E., Aliaga C., Jarabo A., Zibrek K., Gutierrez D., McDonnell R., Botsch M.: To stylize or not to stylize? the effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics (TOG) 34*, 6 (2015), 1–12. 2, 3

[ZKM18] Zibrek K., Kokkinara E., McDonnell R.: The effect of realistic appearance of virtual characters in immersive environments-does the character's personality play a role? *IEEE transactions on visualization and computer graphics 24*, 4 (2018), 1681–1690. 2, 15

[ZMFB23] Zhou B., Meng D., Franco J.-S., Boyer E.: Human body shape completion with implicit shape and flow learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 12901–12911. 3

[ZMM19] ZIBREK K., MARTIN S., MCDONNELL R.: Is photorealism important for perception of expressive virtual humans in virtual reality? *ACM Transactions on Applied Perception (TAP) 16*, 3 (2019), 1–19. 2, 9, 12, 13, 18

[ZXB*23] ZINS P., XU Y., BOYER E., WUHRER S., TUNG T.: Multi-view reconstruction using signed ray distance functions (srdf). In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 16696–16706. 3

[ZXK*21] ZHAO F., XIE Z., KAMPFFMEYER M., DONG H., HAN S., ZHENG T., ZHANG T., LIANG X.: M3d-vton: A monocular-to-3d virtual try-on network. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 13239–13249. 2

[ZZC*23] ZHANG L., ZHAO Z., CONG X., ZHANG Q., GU S., GAO Y., ZHENG R., YANG W., XU L., YU J.: Hack: Learning a parametric head and neck model for high-fidelity animation. *ACM Transactions on Graphics (TOG) 42*, 4 (2023), 1–20. 4

[ZZM19] ZELL E., ZIBREK K., MCDONNELL R.: Perception of virtual characters. In *ACM Siggraph 2019 Courses*. 2019, pp. 1–17. 2