Rig No-Tradicional para Entornos Virtuales

Caso de Estudio: Rig del Brazo

AUTOR

GABRIEL DE JESÚS SIERRA PICÓN

Trabajo de grado, opción semillero, presentado como requisito para optar al título de:

INGENIERO MULTIMEDIA

Directores:

Carlos ROCHA CASTAÑO

Wilson SARMIENTO MANRIQUE

UNIVERSIDAD MILITAR NUEVA GRANADA

FACULTAD DE INGENIERÍA

PROGRAMA INGENIERÍA MULTIMEDIA

BOGOTÁ, 12 ABRIL 2021

Non-Traditional Rig for Virtual Environment Systems. Case study: Arm rig

Gabriel Sierra Picón, Carlos A. Rocha, Wilson J. Sarmiento

Keywords: Avatar, Rigging, Virtual Environment, Virtual Reality

Summary

This paper presents new design specifications for skeleton rigs based on Virtual Environment problems, taking into account their limitations and requirements, and the relationship between users and the virtual avatar they embody. This approach is an alternative to traditional skeleton rigs used in animation, which may generate undesirable deformations in the mesh since they are usually made with handcrafted animation, and software optimization, in mind. This work includes promissory results in the naturalness of the deformation of the wrist of a digital character, validated with eight experts' subjective evaluation in 3D production.

1 | INTRODUCTION

In the animation world, there have always been techniques to animate dolls and puppets, all the way from Chinese shadow puppets to wooden dolls, where the process of articulating these characters has always been of great importance. In digital world, the process is not very different, animators have also found the need to articulate, these now digital, characters just like their colleagues from past eras did (Xia, et al., 2017; Le Naour, et al., 2019). Animation by skeleton has been a traditional solution to this problem. The process of articulating a character via skeleton is known as rigging (Volonte, et al., 2016). Commercial programs have tools to facilitate this process, but for newbies and veterans alike, this process tends to be very complicated and time-consuming. Additionally, depending on the final purpose or target platform of the character, the rig will have different requirements and limitations, e.g., joint count, type of constraints, and joint location on the skeleton. Several algorithms have tried to optimize this process, but they seldom consider all the differences per final target platform (Vargas Molano, et al., 2019). Other method, is a fully automatic process which builds the entire skeleton by investigating the topological differences on the model, and then refining the final skeleton based on geometric conditions such as the bending angle and thickness of the limbs (Ma & Choi, 2014). This method highlights the importance of how there is no one-size-fits-all for skeleton rigs; each character has its own needs. However, despite how useful an automatic system can be, there is always the need for input from the 3D artist. Noh et al. (2020) proposed a method for automatically building a skeleton rig from raw-scanned 3D models in which the user must give key skeleton annotations for the system to build the final skinned 3D mesh, ready for animation6. The entire process is still art, and automation should never replace the artist's judgment and vision. Most automatic systems and algorithms, at the time of writing of this paper, need a degree of human intervention, after all (Noh, et al., 2020).

To Virtual Environments, such as Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), a traditional solution is to bring the skeleton rigs used in 3D animation; but these, as mentioned above, require the intervention of a 3D artist to get suitable results. Another challenge with Virtual Environments is that the hard part is tracking the body accurately. In some cases, traditional skeleton rigs generate undesirable deformations of the 3D mesh. Some previous work improves the tracking system or inverse kinematic position (Bovet, et al., 2018; Pavllo, et al., 2018); our approach is to propose an alternative skeleton rig design that can generate a softer deformation in some articulations on the mesh by adding an extra bone near the articulation.

This paper presents results obtained from applying our proposal to the hand-forearm area. First, we discuss some previous works related to the topic of this research, then we explain our approach to the problem, briefly showing our workflow. Then we talk about the experiment we conducted, and explain how we implemented and tested it in an arm skeleton of a digital character. We then proceed to talk about the results, and we also discuss highlights seen throughout the experiment. And finally, we discuss what was achieved through this work along with proposed future work to expand upon what we achieved.

2 | RELATED WORK

About the process of creating a skeleton rig, many papers have pointed out the hardships and how time-consuming building a skeleton can be for a 3D artist, and thus the importance of having an automatic rigging system to facilitate the process. Pan et al. (2009) propose a method of automatically rigging a character based on its 3D silhouette. The silhouette is detected from two perpendicular projections of the character model, is then used to extract a rough shape of the character's skeleton. They demonstrate the importance of saving time in the character creation pipeline and creating skeleton rigs that adjust to the character being worked on (Pan, et al., 2009).

Other methods, like the one proposed by Ma et al. (2014), is a fully automatic process which builds the entire skeleton by investigating the topological differences on the model, and then refining the final skeleton based on geometric conditions such as the bending angle and thickness of the limbs. This method highlights the importance of how there is no one-size-fits-all for skeleton rigs; each character has its own needs. However, despite how useful an automatic system can be, there is always the need for input from the 3D artist. Noh et al. (2020) proposed a method for automatically building a skeleton rig from raw-scanned 3D models in which the user must give key skeleton annotations for the system to build the final skinned 3D mesh, ready for animation. The entire process is still art, and automation should never replace the artist's judgment and vision. Most automatic systems and algorithms, need a degree of human intervention, after all (Noh, et al., 2020). Ultimately these methods, while important, are only tangentially related to the method we purpose, which is a skeleton rig system that is more adequate for Virtual Environments.

The appearance and structure of the avatars used in Virtual Environments also have a great impact on the behaviors and attitudes of the user. In the Proteus Effect, term coined by Yee & Bailenson (2009), we can observe how users with taller avatars would negotiate in a more aggressive way than users that were given shorter avatars; both height and attractiveness of the avatar affected the behavior of the users. In their second study they note how even these behavioral changes transcend virtual environments on to face-to-face interactions in life. Our self-perception not only affects our behavior, but also how our body reacts physically, as denoted by Hägni et al. (2008) in their study about the galvanic skin response to unexpected threats. Two group of subjects seated in front of a monitor watched virtual arms intercepting virtual balls, until the virtual arms are unexpectedly stabbed by a knife. One group was told to just observe the arms, while the other were told to imagine as if those arms were theirs. The later showed a significantly higher increase in skin conductance compared to the first group. From here we can notice the strong relationship between what the mind considers part of itself, and the reactions and behaviors that take place both at the psychological level and the physical level, and why realism and naturalness will be important factors to consider going forward in the virtual field.

The fact of even having a visual appearance at all affects the user experience (Smith & Neff, 2018). As pointed out by Smith & Neff (2018), in their research where they investigated the communication within virtual experiences, having embodied avatars provides a high level of social presence. They paired some participants with motion tracked avatars, which along with its high level of social presence, presented communication very similar to what happens in face-to-face interactions in the real world. In contrast to the participants not using any avatars at all, which felt their experience was lonelier and with degraded communication (Smith & Neff, 2018).

Besides the visual appearance of oneself, all the process involved in tracking the user’s movements also plays an important role in how users perceive themselves. As latency between the user’s movements and what they see increments, qualitative drops in their experience has been observed (Toothman & Neff, 2019). Another interesting phenomenon that happens when the tracking is not fully in sync, is the Self-Avatar Follower Effect, term coined by Gonzalez-Franco, et al., (2020). In their work they present experimental results on how whenever the virtual avatar’s body is not in sync with the real user’s body, they will try to compensate and unconsciously follow its movements and try to match them.

All these works demonstrate the importance of the avatar people embody in their virtual experiences. Going a level deeper, within the avatar itself, hands are very noteworthy since they are our primary method of interaction with our environment (Pavllo, et al., 2018). A broad range of motion capture devices, that focus on the hands as primary means of motion capture, and control over the virtual avatar, have also become very commonplace on the market (Bovet, et al., 2018). This is why we chose the arms, and in particular the hand-forearm area as our case study.

3 | OUR APPROACH

This work proposes design changes for the classic VR skeleton rig (in which the elbow joint connects directly to the wrist joint), for anthropomorphic characters, specifically the arms. Skeleton rigs for video games, real-time tracking, and audiovisual productions often have different needs and requirements to be met within each one's environment. Video game rigs, being the simplest between the previously mentioned ones, only require a basic skeleton structure and are often limited in joint count depending on the target platform's hardware capabilities and development engine limitations. On the other hand, audiovisual productions require complex rig systems, so the animators do not feel limited at conveying all the emotions and actions through the characters they are trying to give life. These characters often also require clothing and fur simulations to go with the rigged system, which may have variations depending on if realistic or cartoon-styled animations are the goal. Finally, real-time tracking skeletons, such as the ones used for Virtual Environments, require a more precise bone structure on the limbs, so the movement feels natural to the end-user, and an accurate Inverse Kinematics system to reduce the number of points that need to be tracked in real-time. One of the biggest offenders in unnaturalness is the twist between the wrists and the forearm, which cannot be made to look natural unless design changes are taken into account.

Consequently, a problem is that the skeleton rig currently used in all these cases does not consider the limitations and requirements of each of the other applications. We propose modifications to the traditional human arm rig to get the best skin deformation results when the wrist rotates while following a tracking system. One of the design challenges we had was to solve is the twist between the wrist and the forearm for real-time tracking skeleton rigs. Our solution was to add an extra joint, as a child of the elbow joint, and have it rest between the middle of the wrist and the elbow joint. Then we linked this new joint's rotation to be exactly one half of the wrist joint's rotation. It has certain limitations at extreme angles, but it looks much more natural than the classic skeleton arm of only the elbow and wrist joint. See Figure 1.

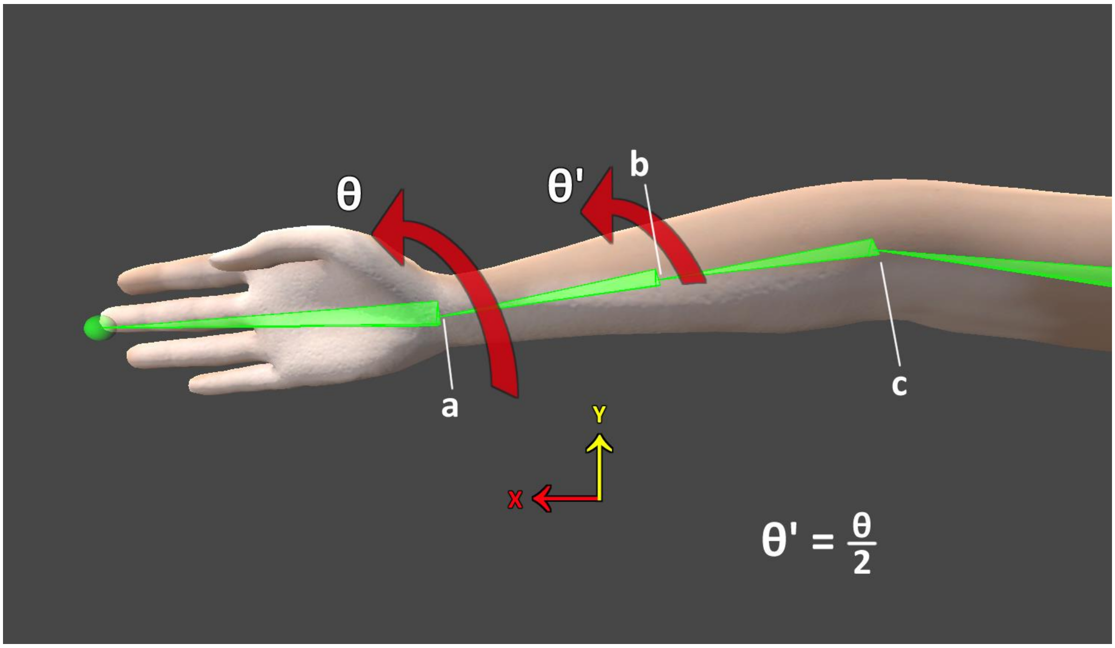

Figure 1. Our proposed rig design. Wrist joint (a), new joint addition (b), elbow joint (c). Where θ represents the degrees rotated on its local X axis by the wrist joint, and θ' represents the degrees rotated on its local X axis by the new joint addition, which will always have half of the rotational value of the former.

4 | EXPERIMENT

To test our hypothesis, we recorded motion tracked sequences of users moving their arms a total of six times, to then recreate these sequences using both our rig and the classic VR rig, inside a real time video game engine. Then we captured these real time rendered recreations into video clips, both in shaded and wireframe view.

Afterwards, a subjective evaluation with eight experts in a wide range of fields that include, but were not limited to, 3D Modelling, 3D Animation, Character Rigging, and traditional 2D animation, was made with the purpose of analyzing sequences of videos comprised of motion tracked data being represented by the 3D models of our case study.

Each video was composed by four clips; the classic VR rig and our proposed rig simultaneously, first in shaded view, and then both in wireframe view. Six videos were made in total, one for each motion tracked sequence, so the experts could observe the difference each rig made on different movement sets.

4.1 Implementation

In order to validate our proposal, we applied both our proposed rig and the classic VR rig to a freely available model made with MakeHuman, using a Geodesic Voxel bind method in Autodesk Maya. Then, the model was exported to Unity 2020.1.8f1, via FBX file format. Motion tracking data was recorded into log text files, which were then converted into Unity's animation file format, via a Python script, for further display within the engine. The script took the coordinates of the wrist and the middle finger of each hand and calculate the rotation between them, relative to the rest of the arm, and then convert these Euclidean rotation values to their equivalent in quaternions, which Unity requires for its animation file format. A viewport camera was placed so that both arm models would be centered in the same place in each shot, ensuring the only difference visible was a product of the effects each rig has on every model. For the movement of the arms, the wrist's movement and rotation was handled by the motion tracked data we imported into Unity, as mentioned previously, while the rest of the arm had Inverse Kinematics handled by Unity's Animation Rigging module.

The first set of video clips displayed both set of arms, one with the classic VR rig and the other with our proposed rig, rendered with Unity's Standard PBR shader, using only a diffuse, roughness, and normal map, so the effects of the wrist twisting on the skin could be more noticeable. A point light and a checkered background were also added to further enhance the effect. Later, the second set of video clips displayed both sets of arms with a wireframe shader, using a freely available wireframe shader downloaded from Unity's Asset Store. For the wireframe clips, the checkered background was replaced with a solid grey background. See Figure 2 and 3.

Finally, we proceeded to make a poll. For each video, participants were asked to fill a form in which they rated, on a scale from 1 to 5, if they agreed or not, that there were notable differences between each clip; where 1 stands for "I do not agree" and 5 stands for "I fully agree". Participants were then be asked to pick which clip looked more natural, or alternatively, say neither felt natural, or that they both looked the same. Our pool of eight participants was comprised of experts with a minimum of three years of professional work in their respective fields; 3D animation, render, traditional animation, motion graphics, and character rigging.

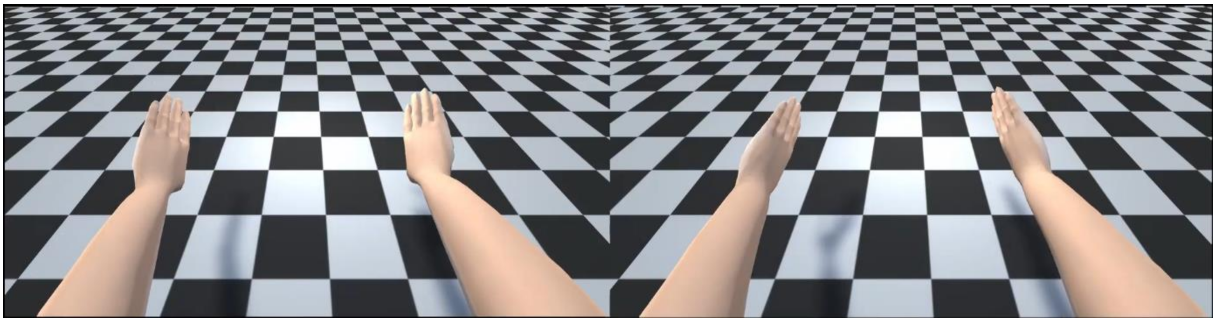

Figure 2. Our proposed rig (Left), Classic VR rig (Right). Shaded view. It can be noted how when the wrists twist, that section of the arms shrinks on the classic VR rig, but retains its size and looks more natural on our proposed rig.

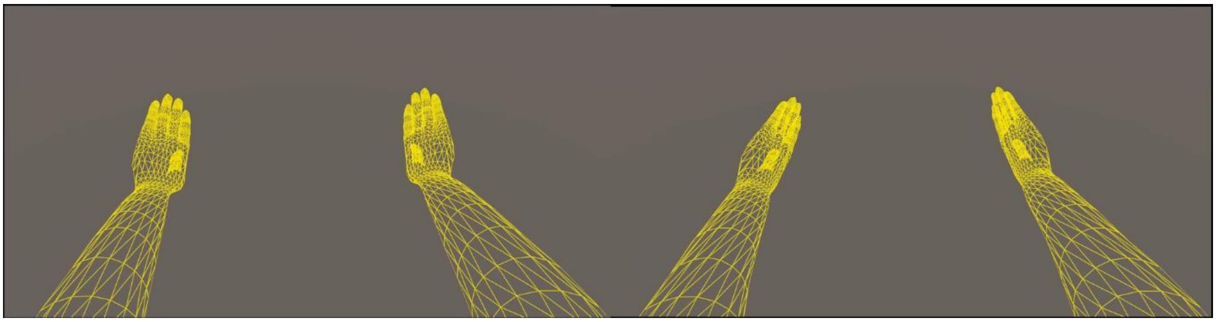

Figure 3. Our proposed rig (Left), Classic VR rig (Right). Wireframe view.

4.2 Results

Out of the six tests we presented, we observed that on average, most of the experts agree that our rig allows the model to have a more natural deformation when the wrists twist, closer to what they would expect to see in real life with their own arms, in contrast with the classic rig. The worst-case scenario, out of all of the polled answers, coincides with a motion tracked sequence that had little wrist rotation involved; 50% of the participants answered they couldn’t see any notable differences at all between both rigs. In contrast, our best-case scenario reported an overwhelming majority of answers, 62.5% of the participants, favoring our proposed rig, coinciding with a motion tracked sequence that had ample wrist rotations. See Figures 4 through 6 for a more detailed breakdown of the results.

From these results we can tell there is a relationship between the magnitude of the wrist rotation angle and how noticeable the difference of naturalness between rigs becomes.

These results give us an idea of how we can structure design changes on a skeleton rig so that it can solve reoccurring problems in animation, depending on the platforms, limitations, and requirements; and eventually, making future automatic implementations of these rigs. Now it is possible to establish an implementation algorithm with its accompanying tests with end-users to validate this that has been exposed.

“There are striking differences between the two video clips”				
Strongly Disagree	**Disagree**	**Borderline Agree**	**Agree**	**Strongly Agree**
3	1	2	2	0
3	0	1	3	1
2	3	0	2	1
2	2	1	3	0
3	1	3	1	0
1	2	3	1	1

Figure 4. *Results of the first question of the poll. Participants were asked if they agreed or not, that striking differences were observed between our proposed rig and the classic VR rig.*

"Of the two video clips, which one would you consider looks more natural and adjusted to reality?"			
Clip A	**Clip B**	**They look the same**	**Neither**
3*	0	2	2
0	5*	1	2
1	4*	1	2
3*	1	1	3
2*	0	4	2
1	5*	2	1

Figure 5. Results of the second question of the poll. Participants were asked to tell which clip looked more natural, between our proposed rig and the classic VR rig. The asterisk (*) denotes which one of the two was our proposed rig in each particular instance.

"Of the two video clips, which one would you consider looks more natural and adjusted to reality?"				
	Our Proposed Rig	**Classic VR Rig**	**They look the same**	**Neither**
Min	2	0	1	1
Max	5	1	4	3
Average	3.666	0.5	1.83	2.0

Figure 6. Results of the second question of the poll. This table condenses the results of Figure 5, showing the highest, lowest and average amount of answers in each category.

5 | CONCLUSION

This research explores the hypothesis that, upon adding an additional joint between the wrist and the elbow, in a humanoid rig, can lead to more lifelike deformation along the arm when the wrist rotates. For testing this hypothesis, we recorded motion tracking data from live subjects. Then we implemented both our rig proposal and the classic rig commonly used in VR applications, into the Unity game engine, so we could replicate the motion tracked animations with both rigs.

An experiment was conducted in which we gathered several experts that specialize in related fields; from traditional animation, to 3D animation and rigging. The experts were then shown six different videos comprised of various clips representing both rigs playing the motion tracked sequences, both in shaded and wireframe view. From here, the experts were prompted to answer a series of questions regarding if they could perceive any difference between both rigs, and the naturalness of the deformations. Results showed that a majority of the experts, on average, leaned to our proposed rig in terms of naturalness across all tests. We saw a direct relation between the magnitude of the wrist's rotation angle and how noticeable the differences between rigs were. We only focused on the hand-forearm area, but in reality the elbow, the upper arm and all the way to the shoulder, have structural considerations that should be taken into account as they may affect the naturalness of the movement of the rest of the arm. These should be given attention for future works, and to further enhance the realism of motion tracked arms in virtual environment settings. The visual appearance of one's own avatar in a VR experience, and how accurately it moves when compared to reality, has impacts both on how immersed and how we behave within the virtual experience. Thus, we believe that our rig proposal could be adapted and streamlined into existing VR pipelines, and improve the naturalness of the virtual experience.

Acknowledgments

We would like to thank the Office of the Dean of Research of Nueva Granada Military University for their funding of this work, which is a result of the undergraduate scientific initiation project PIC-ING-3331. We would also like to thank the experts that participated in the poll we conducted. And finally, Jesús Sierra for reviewing this paper.

Conflict of interest

The authors declare no potential conflict of interests.

SUPPORTING INFORMATION

The following supporting information is available as part of the online article:

References

Bailenson, J. N., Yee, N., Merget, D. & Schroeder, R., 2006. The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoperators and Virtual Environments,* 8, Volume 15, p. 359–372. https://doi.org/10.1162/pres.15.4.359

Baran, I. & Popovi, J., 2007. *Automatic rigging and animation of 3D characters.* New York, New York, USA, ACM Press, p. 72. https://doi.org/10.1145/1275808.1276467

Bhatti, Z. & Shah, A., 2012. *Widget based automated rigging of bipedal character with custom manipulators.* New York, New York, USA, ACM Press, p. 337–340. https://doi.org/10.1145/2407516.2407593

Bovet, S. et al., 2018. The Critical Role of Self-Contact for Embodiment in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics,* Volumen 24, p. 1428–1436. https://doi.org/10.1109/TVCG.2018.2794658

Callejas Cuervo, M., Díaz, G. M. & Ruíz-Olaya, A. F., 2015. Integration of emerging motion capture technologies and videogames for human upper-limb telerehabilitation: A systematic review. *DYNA,* 2, Volumen 82, p. 68–75. https://doi.org/10.15446/dyna.v82n189.42066

Gonzalez-Franco, M. et al., 2020. The Self-Avatar Follower Effect in Virtual Reality. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR),* p. 18–25. https://doi.org/10.1109/vr46266.2020.00019

Hägni, K. et al., 2008. Observing virtual arms that you imagine are yours increases the galvanic skin response to an unexpected threat. *PLoS ONE,* Volumen 3, p. 1–6. https://doi.org/10.1371/journal.pone.0003082

Le Naour, T., Courty, N. & Gibet, S., 2019. Skeletal mesh animation driven by few positional constraints. *Computer Animation and Virtual Worlds,* 5.Volumen 30. https://doi.org/10.1002/cav.1900

Ma, J. & Choi, S., 2014. Kinematic skeleton extraction from 3D articulated models. *Computer-Aided Design,* 1, Volumen 46, p. 221–226. https://doi.org/10.1016/j.cad.2013.08.036

Mathee, H. & Haux, B., 2016. *Portable real-time character rigs for virtual reality experiences.* New York, NY, USA, Association for Computing Machinery, Inc, p. 5–6. https://doi.org/10.1145/2947688.2947694

Noh, S.-T., Takahashi, K., Adachi, M. & Igarashi, T., 2020. Shape refinement and rigging of raw-scanned 3D volume by a user-specified skeleton. *Computers & Graphics,* 4, Volumen 87, p. 80–88. https://doi.org/10.1016/j.cag.2020.01.009

Pan, J. et al., 2009. Automatic rigging for animation characters with 3D silhouette. *Computer Animation and Virtual Worlds,* 6, Volumen 20, p. 121–131. https://doi.org/10.1002/cav.284

Pantuwong, N. & Sugimoto, M., 2011. *A fully automatic rigging algorithm for 3D character animation.* New York, New York, USA, ACM Press, p. 1. https://doi.org/10.1145/2073304.2073337

Pavllo, D., Porssut, T., Herbelin, B. & Boulic, R., 2018. *Real-Time Finger Tracking Using Active Motion Capture: A Neural Network Approach Robust to Occlusions.* New York, NY, USA, Association for Computing Machinery. https://doi.org/10.1145/3274247.3274501

Smith, H. J. & Neff, M., 2018. Communication behavior in Embodied virtual reality. *Conference on Human Factors in Computing Systems - Proceedings,* Volumen 2018-April, p. 1–12. https://doi.org/10.1145/3173574.3173863

Toothman, N. & Neff, M., 2019. The impact of avatar tracking errors on user experience in VR. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings,* p. 756–766. https://doi.org/10.1109/VR.2019.8798108

Vargas Molano, J. S., Díaz, G. M. & Sarmiento, W. J., 2019. Parametric Facial Animation for Affective Interaction Workflow for Avatar Retargeting. *Electronic Notes in Theoretical Computer Science,* 5, Volumen 343, p. 73–88. https://doi.org/10.1016/j.entcs.2019.04.011

Volonte, M. et al., 2016. Effects of Virtual Human Appearance Fidelity on Emotion Contagion in Affective Inter-Personal Simulations. *IEEE Transactions on Visualization and Computer Graphics,* Volumen 2626, p. 1–1. https://doi.org/10.1109/TVCG.2016.2518158

Wu, D., Weng, D. & Xue, S., 2016. Virtual Reality System as an affective medium to induce specific emotion: A validation study. *Electronic Imaging,* 2, Volumen 2016, p. 1–6. https://doi.org/10.2352/ISSN.2470-1173.2016.4.ERVR-419

Xia, S. et al., 2017. A Survey on Human Performance Capture and Animation. *Journal of Computer Science and Technology,* 5, Volumen 32, p. 536–554. https://doi.org/10.1007/s11390-017-1742-y

Yee, N., Bailenson, J. N. & Ducheneaut, N., 2009. *The proteus effect: Implications of transformed digital self-representation on online and offline behavior.* s.l.:s.n. https://doi.org/10.1177/0093650208330254