Reconstruction of the Face Shape using the Motion Capture System in the Blender Environment

Joanna Smolska[a] and Dariusz Sawicki[b]
Warsaw University of Technology, Warsaw, Poland

Keywords: Motion Capture, Facial Shape, Blender.

Abstract: Motion Capture systems have been significantly improved during the last few years, mainly since they found usage in the entertainment industry and medicine. The main role of Motion Capture systems is to control and record the position of a set of selected points on the scene. Project described in the following article aimed at developing a method of reconstructing the shape of a specific face with the possibility of controlling its movement for the purposes of computer animation. We conducted experiments in a Motion Capture laboratory using Qualisys Miqus M1 and M3 cameras and created a low-poly face model. Moreover, we proposed an algorithm of shape reconstruction using the analysis of the position of a set of points (markers) applied to the surface of the face and Blender software. Finally, we analyzed advantages and disadvantages of both approaches to facial motion capture. Taking into account the publications of recent years, a brief analysis of the trends in the development of facial reconstruction methods has also been carried out.

1 INTRODUCTION

The System of Motion Capture (MoCap) is one of the most important and technically most interesting solutions in twentieth century cinematography (Choosing a performance, 2020, Parent, 2012). Without this technique, films like Lord of the Rings or Avatar could not have been made. MoCap is also used to produce computer games and create virtual reality (Menache, 2011, Kitagawa and Windsor, 2008). On the other hand, Motion Capture is used not only in the entertainment industry. In sport, it allows controlling precisely the player's movements and greatly facilitates training (Kruk and Reijne, 2018, Pueo and Jimenez-Olmedo, 2017). MoCap plays a very important role in medicine, mainly in physical therapy and rehabilitation (Cannell et al., 2018, Zhou and Hu, 2008), but also in psychiatry (Zane, 2019).

The main role of MoCap is to control and record the position of a set of selected points on the scene. In the case of cinematography, games and virtual reality, it allows translating the movement of actors (objects on the stage) into the movement of animated characters (or other objects). In the case of rehabilitation (and other medical applications), the analysis of the position itself serves the proper medical diagnosis or helps in the therapy. In all these cases, the movement of points (or their mutual configuration) is important, not the analysis of the shape of the object. For the surgeon / physiotherapist, it is important how the joint works (to what extent it bends and in what plane), but the details of the limb shape are not so important. In the case of computer graphics, the designer prepares the shape of a fantastic character for an SF movie independently, while MoCap is used to put the character in motion. The procedure is usually similar when the shape of a character's face is to accurately reflect the shape of an actor's face. First, the actor's face is scanned precisely and an appropriate virtual model is built. Then MoCap is used to control the virtual character's movement and express emotions. A good example of such a procedure are the virtual characters in The Matrix (Borshukov, 2003) – for the viewer, the virtual character was indistinguishable from the real one. Generally, systems of Motion Capture are primarily used for motion analysis and control, shape analysis and description is secondary – shape representation is independent of motion control. However, it is possible to show situations when a combination of

[a] https://orcid.org/0000-0001-8534-9241
[b] https://orcid.org/0000-0003-3990-0121

Smolska, J. and Sawicki, D.
Reconstruction of the Face Shape using the Motion Capture System in the Blender Environment.
DOI: 10.5220/0011001700003124
In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) - Volume 1: GRAPP*, pages 310-317
ISBN: 978-989-758-555-5; ISSN: 2184-4321

both of these functions could be convenient. This work is an attempt to use the Motion Capture technique primarily to represent the shape of the face, but also to simultaneously control its movements.

Modern Motion Capture systems allow performing complex tasks related to the recording and analysis of motion. However, facial MoCap still seems to be a difficult task due to subtle changes in shape (expressing emotions) and the fact that the human face is the most known object to man so even the smallest deviations from the expected correctness are immediately noticed. In this project, for the needs of computer animation, the possibility of developing a face model was considered in such a way that it was possible to recreate a specific face – the face of a particular person, and at the same time to control the movement of the face (speech, emotions) using the Motion Capture technique. Such a task should be performed in one programming environment so that there is no need for additional data transferring (and / or conversion). The Blender environment was selected for this task. It is open source and free software for modeling 3D graphics (Blender Foundation, 2021, Blender documentation, 2021). It also has convenient tools for rendering and production of animation.

The aim of the project is to develop a method of reconstructing the shape of a specific face with the possibility of controlling its movement for the purposes of computer animation. The MoCap technique will be used to describe the location of the relevant points in the model. Additionally, it was assumed that the method would be fully integrated with the software in the Blender environment.

2 THE MOTION CAPTURE SYSTEMS AND ITS APPLICATIONS

2.1 The History and State of the Art

The basis for the operation of MoCap systems is rotoscopy – patented by Max Fleischer in 1917 (Bedard, 2020). This technique was used in 1937 by W. Disney in the production of the animated film Snow White and the Seven Dwarfs. The first successful application of contemporary MoCap was done in 1985 by Robert Abel – the film Brilliance (Sexy Robot) (Kitagawa and Windsor, 2008). Among the later productions, the most spectacular ones are worth mentioning: The Matrix – where an advanced character creation system was used for the first time in a way that fragments with virtual actors were inserted into the film imperceptibly to the viewer. The method developed for this film was presented at the most important conference on computer graphics (SIGGRAPH) in 2003 (Borshukov, 2003). The second, important example is the movie Lord of the Rings, where Gollum appears – probably the most famous virtual character performing with real actors. Descriptions of the methods developed to create this character have found a place not only in the history of cinema, but above all in textbooks of computer graphics.

Several technical solutions are used in MoCap systems: mechanical (exo-skeletal), magnetic (electromagnetic), inertial and optical (Menache, 2011, Kitagawa and Windsor, 2008). Due to their advantages, optical systems are the most widely used today and it seems that they are the future of the MoCap technique (Topley and Richards, 2020). Currently conducted research is going in the direction of capturing a motion in an efficient way without the need of markers usage. Such solutions are already available on the market (Nakano et al., 2020).

In the case of Facial MoCap, the set of selected facial points is registered. This set includes reference points that allow to present facial expressions and emotions. Usually it is from a dozen to several dozen reference points, in the form of markers glued to the surface of the face. This technique does not allow eye movement to be recorded. The problem was solved during the research for the film Avatar, by placing cameras mounted on the head very close to the actors' faces. Currently, markerless systems are increasingly used (Choosing a performance, 2020). Especially in the case of face shape analysis for border control purposes. A set of several to a dozen cameras placed in close proximity to the actor's face and specialized software are most often used to reconstruct the shape of the face and the movement at the same time (Kitagawa and Windsor, 2008). The tendency in this case is to build a face model based on a skeletal model – a set of bones corresponding to the anatomical structure of the craniofacial model is built (Parent, 2012, Kitagawa and Windsor, 2008). There are also modern software solutions that add the ability to register and control traffic directly in Blender (Thompson, 2020). Such solutions allow the use of a smartphone camera (Ai Face-Markerless Facial Mocap, 2021, Face Cap - Motion Capture, 2021). There are also similar solutions (Faceware Facial Mocap, 2021), independent of the hardware. These solutions allow registering the location of selected – characteristic points of the face. Thanks to this, it is possible to "transfer" the actor's facial expressions to those of a virtual character. However, this technique does not provide enough information to reconstruct the shape (with details) of the actor's face in a virtual form. This is usually done similarly to the production of the Matrix movie.

It is worth paying attention to the fact that the system sensitivity (and thus the number of analyzed points of face) is directly dependent on the purpose for which facial MoCap is used (Hibbitts, 2020). A small number of points are enough to orient the face (head) in a certain direction, more points to analyze simple facial expressions, even more to analyze emotions. A more sensitive system is needed for speech analysis, the most for full reconstruction (to capture all the subtle details of the shape of the face).

2.2 The Experiments in Motion Capture Laboratory – a First and Representative Case Study

The original assumption was to use the Motion Capture system based on the Qualisys camera system with the dedicated Qualisys Track Manager (QTM) software (Qualisys, 2021). In the laboratory at the Faculty of Electrical Engineering of Warsaw University of Technology, there are four Miqus M1 cameras and two Miqus M3 cameras. The cameras differ in terms of technical parameters. M3 cameras allow for obtaining higher image quality than M1 cameras (Resolution: M1 – 1216 × 800, M3 – 1824 × 1088). M3 cameras are characterized by higher fps values (the number of recorded frames per second), which means that the obtained measurement will be smoother in the perception. In addition, Miqus M3 has also been enriched with a "high speed" mode, which allows recording more dynamic movements.

The Miqus M1 and Miqus M3 cameras require a set of markers to identify reference points in space (Figure 1a.). Based on the position of the markers, a "low poly" face model was built – Figures 1b and 1c.

Because there is no direct integration between QTM and Blender software, building a "low poly" model required the development of an additional program to transfer the appropriate coordinates. Additionally, there was a problem of ambiguity in automatic mesh triangularization. For example, the space between the eyebrows looks completely different (Fig 1b versus Fig 1c) depending on the way the points are connected. From the perspective of the triangularization algorithm, both solutions are correct (Shewchuk, 1996). The only effective solution seems to be connecting the relevant points manually. This allows for subtle changes in shape to be taken into account, thanks to the human experience of facial recognition. The main disadvantage of the solution is the limitation of the number of mesh nodes in the "low poly" model due to the lack of the possibility to freely thicken the markers on the face.

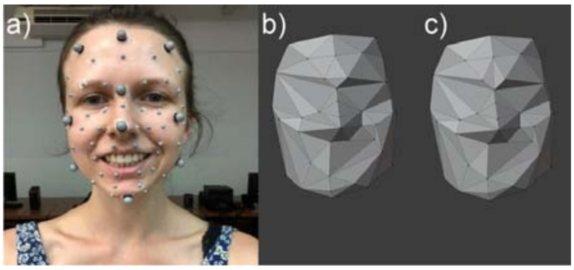

Figure 1: a) Placement of markers on the author's face, b) and c) Built face models based on the analysis of the marker position. A problem of ambiguity in automatic mesh triangularization – the space between the eyebrows.

The use of Miqus M1 and M3 cameras and QTM software does not allow for effective identification of the face shape for the purposes of animation with the application of Blender. The 3D Resolutions of the M1 and M3 cameras at a distance of 10m are respectively: 0.14mm and 0.11mm. The arrangement of the cameras in the laboratory allowed conducting experiments at a distance of about 3 meters. Even so, there were problems due to the fact that these cameras are suited to analyzing the movement of an object rather than the subtle details of a face. In consequence:

- In many cases cameras couldn't differentiate one marker from another – for example two markers were classified as one that changes the position between frames;
- Qualisys track manager didn't detect markers movement correctly – it was creating new markers for each few frames instead of keeping track of one marker (coming back to differentiating markers from each other – distances between them were too small);
- We noticed relatively big measurement errors regarding location in space – facial movements are very subtle, so even small measurement errors result in highly visible animation errors.

The resolution of the cameras and the small size of the markers (diameter 2 mm) allow for full control of the trajectories of the movement. However, identification of the shape of the face with subtle details becomes a very difficult problem. In this way, there is no practical possibility of increasing the mesh density and reproducing the subtle details of the face. Additionally, there is no direct software integration between QTM and Blender.

3 PROPOSED ALGORITHM OF THE SHAPE RECONSTRUCTION

There are known attempts of Motion Capture analysis adapted directly to Blender (Thompson, 2020). Ai Face-Markerless Facial Mocap For Blender (Ai Face-Markerless Facial Mocap, 2021) allows recognizing head movements, eyes closed (blinking) and basic forms of facial expression. The Face Cap - Motion Capture program (Face Cap - Motion Capture, 2021) works in a similar way. Both programs were developed to control the animated character and the movements of the head. They also allow controlling facial expressions and changes in facial expressions to a limited extent. Similar possibilities are offered by independent software (Faceware Facial Mocap, 2021).

Taking into account the analyzed small number of characteristic points of the face (and their arrangement), it should be stated that none of these programs is suitable for the reconstruction of the face shape – for presenting its subtle surface details. To solve this problem, we developed an algorithm for the reconstruction of the face shape using the analysis of the position of a set of points (markers) applied to the surface of the face. The proposed Algorithm is as follows:

```
A1.  Create a generic 3D model (Blender
     + Face Builder).
A2.  Import the video into Blender.
A3.  Pin specific areas (based on7
     reference photos – figure 2).
A4.  Make a reconstruction in Blender
     (move the markers to the 3D
     projection
     area).
A5.  Match the model to the set of
     markers of the real face (figure 3).
A6.  Create bones. Combine them with the
     mesh of the fitted model.
A7.  Add a sphere of influence for each
     bone.
A8.  Set the bones so that in each frame
     they copy the position of a given
     marker, ignoring the value in the Z
     axis (identification of the
     position based on the image from one
     camera is in fact a 2D analysis).
A9.  Go to face animation.
```

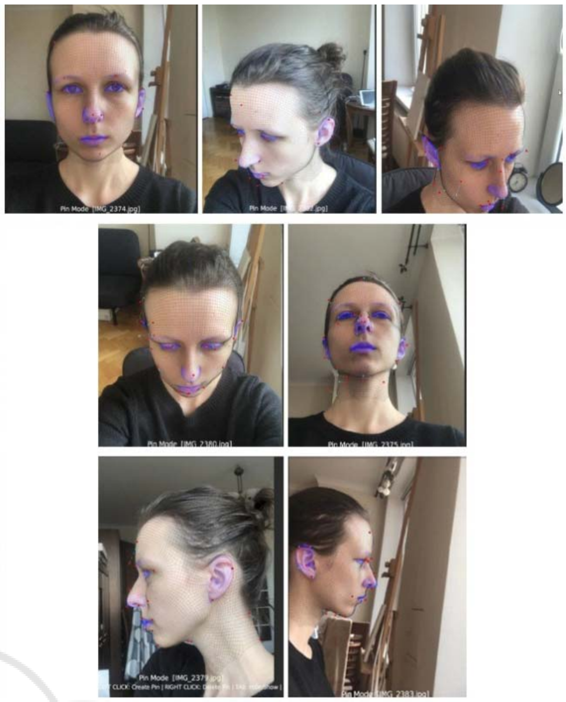

Figure 2: The set of 7 reference photos.

The 3D model was created using Blender software and the Face Builder tool. This tool allows creating a generic (universal) face model, and then modifying the resulting model in order to represent a specific face. The basic head model is presented in Figure 4. We took 7 photos of the face from different angles (Figure 2). These photos served as references for modifying the model (pinning process – matching the model to the photos). The generic model is a universal (averaged) model and practically does not coincide with any particular face. In the pinning process, it is worth starting with the most characteristic points, such as the nose or mouth, highlighted in blue in the figure.

Figure 3: Face with markers painted on the skin.

In Figure 4 selected successive stages of the implementation of the facial reconstruction algorithm are presented.

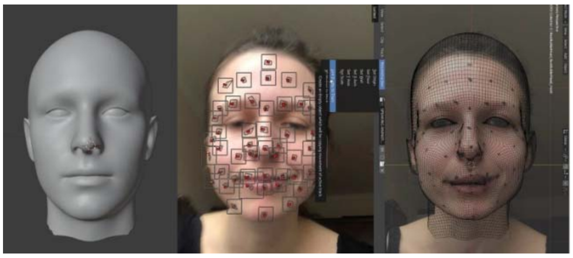

Figure 4: The basic head model (left) and selected successive stages of the face reconstruction.

In the sixth point of the algorithm, bones were attached to the facial model. For further convenience of work, all the bones have been combined into one armature. Its size has been reduced while maintaining the position of the point from which each bone emerges – the position of the appropriate marker in the first frame of the recording. At this stage, it is possible to manually modify the shape of the face model (Blender pose mode) by "grabbing" a given bone. Before grabbing bones we needed to add weights to them, so they have an influence on the models's shape.

Additionally, it is worth paying attention to the fact that the proposed algorithm, after shape reconstruction, directly gives the possibility of further work in animation. From a formal point of view, it could be said, in this case, about a disadvantage of the algorithm – it is difficult to distinguish the moment when the reconstruction works are completed and when the movement control starts. However, from a practical point of view, this is a huge advantage – it is fully integrated into the animation process. And at the same time, it has been implemented directly in the Blender software. In this way, the proposed solution meets the assumptions of the project.

The algorithm uses a video recording (in our case it was done using the camera in the iPhone 8). Dots painted on the skin of the face with a black eye pencil were used as markers. This approach does not require sticking markers (as in the Qualisys system). In addition, the painted dots are always smaller than the sticky markers (even in the smallest version). Thanks to this it is possible to easily add points to the face, if it is necessary to correct the shape.

4 RESULTS

The implementation of the proposed algorithm for the reconstruction of the face shape and control of its position allowed for real-time animation. This means that the facial expressions recorded by the smartphone's camera are converted into the movements of the virtual animated face in Blender. Examples of images (freeze-frames) of a virtual face are presented in Figure 5. On the Internet, a short film is available showing the animation of a virtual face coordinated with the author's facial expressions (https://youtu.be/j7Spse-psEU).

Figure 5: Examples of animation frames.

5 DISCUSSION

Two methods / sets of tools were used in the research described in the article: set of Qualisys cameras with integrated algorithm and author's algorithm with smartphone camera. When comparing the author's methods to a professional Qualisys solution, it is worth paying attention to several aspects.

- The proposed solution works in 2D (as opposed to 3D Qualisys cameras). The consequence of this is the need to calibrate the shape of the face and "pin" the appropriate regions at the preliminary stage;
- All the software was developed independently – it was not possible to use a ready-made, Qualisys system. However, the work is greatly facilitated by well-prepared libraries supporting the work of Motion Capture in the Blender environment;
- A significant advantage of the proposed solution is the possibility of using practically any number of face markers. In the Qualisys system, the diameter of the smallest marker limits their number. This is important for precise shape mapping. On the other hand, one may consider how the large number of markers in the 2D registration affects the accuracy of the position measurement and the precision of facial reconstruction. Such an analysis is foreseen in the future.

The authors' solution turned out to be more useful for the application considered in the project. The professional MoCap system has great advantages, and above all, it is universal when it comes to various applications. However, in a specific task, a "tailor-made" solution may prove to be more useful.

Table 1: Summary of the properties of the various methods.

Publication	Algorithm / method	Accuracy of representation	Speed of work. Ease of implementation	Hardware / software requirements	Integration with Blender / other environment
(Paysan et al., 2009)	3D morphable models	High accuracy with textural shape of the whole face.	Complicated 3D model of the face.	Device independent method	The need to develop an appropriate library
(Perakis et al. 2012)	Localization of a basic set of landmark points from face. Usage of geometrical dependencies (shape index and spin image) for encoding the coordinates of points.	Relatively high accuracy but a small number of face points	Relatively long time (1 – 6 s) of localization of the landmarks for each facial scan..	Device independent method	The need to develop an appropriate library
(Kazemi and Sullivan 2014)	Localization of a basic set of landmark points from face. Usage of an ensemble of regression trees	Relatively high accuracy but a small number of face points	Real time, very fast method	Device independent method	The need to develop an appropriate library
(Ichim et al. 2015)	Morphable model and multi-view geometrical analysis	High accuracy with textural shape of the whole face	Real time, very fast method	Any hand-held video input (smartphone)	The need to develop an appropriate library
(Zhu et al., 2018)	Regression-based face alignment. Usage of regresion tree.	High accuracy.	Real-time. Complicated method.	Device independent method	The need to develop an appropriate library
(Valle et al., 2019)	Coarse-to-fine cascade of ensembles of regression trees. Usage of Convolutional Neural Network.	High accuracy. Only contour – e.g. no points on the cheeks.	Real-time. Advanced training stage of the CNN	Device independent method	The need to develop an appropriate library
(Thompson, 2020, Face Cap - Motion Capture, 2021)	MoCap of a basic set of landmark points from face.	Relatively the small set of landmark points	Real time. Easy to use.	iPhone, iPad	Fully integrated with Blender
(Park and Kim 2021,)	Combination of two methods for landmark points (coordinate regression and heatmap regression), usage of Attentional Combination Network	Very high accuracy. The ability to adjust the resolution. Only contour – e.g. no points on the cheeks.	Real-time. Complicated method.	Device independent method	The need to develop an appropriate library
(Ai Face-Markerless Facial Mocap, 2021)	MoCap of a face contour, eye contour, lip contour (max 24 apple ARkit shapekeys).	Relatively the small set of landmark points. Only contour – e.g. no points on the cheeks.	Real time. Easy to use.	Any camera	Integrated with Blender, integrated with Unity3d
(Faceware Facial Mocap, 2021)	MoCap of a face contour, eye contour, lip contour.	Relatively high accuracy (adjustable resolution). Only contour – e.g. no points on the cheeks.	Real time. Easy to use.	Dedicated to iClone software	Dedicated to iClone software, no other integration possible
(Luo et al., 2021)	Nonlinear morphable model and multi-view geometrical analysis using Generative Adversarial Network	Relatively high accuracy depends on training stage.	The realization time depends on the number of iterations in the testing stage - indirectly on accuracy	Device independent method	The need to develop an appropriate library
(Li et al., 2021)	3D morphable model and multi-view topological analysis	High accuracy depending on the mesh size (number of vertices).	Real-time. Complicated method	Device independent method	The need to develop an appropriate library
Our method	Localization of a set of landmark points from face. Usage of geometrical dependencies.	High accuracy, shape of the whole face. The ability to adjust the resolution.	Real-time animation.	Any camera, smartphone	Fully integrated with Blender

Professional MoCap systems rarely give the possibility of direct integration with other / different software. It is also worth paying attention to the high price of such a system.

It would have been interesting to see how well the proposed solution works. A good confirmation of the correctness of the developed algorithm and its implementation is the comparison of the photo of the face and the generic face and the matched (reconstructed) face. For comparison, face settings other than those used in the reference photos set (viewed from different angles) were selected. An example of such a comparison is presented in Figure 6. You can see the compliance of the shape of the profile line (forehead – nose – mouth – chin) of the fitted face compared to the photo of the real face.

At this stage of the project implementation, the virtual character lacks details such as the image of the eyes or hair. The color of the skin should also be properly selected. Such elements will be added in the next stage of the project.

It is also worth comparing the proposed solution with other modern methods of recognizing the shape of the face. Even, if they are not directly related to the MoCap system. The properties of various methods are summarized in the Table 1. A comparison of publications from recent years shows a certain trend in the development of methods. It seems that currently the most popular methods are simply those that use neural networks (convolutional networks). However, methods based on geometric analysis – although less frequently used – also bring good results (Ichim et al. 2015, Li et al., 2021). In addition, many modern solutions allow only (or primarily) to analyse a set of selected (and basic) landmark points. Although, there are works (Graham et al., 2013) in which it is proposed to recognize the subtle details of the shape of any part of the face. Using morphing, normal map and albedo / texture map gives very good results (Ichim et al., 2015, Luo et al., 2021). The tendency to use these methods is also visible in the most recent works (Li et al., 2021).

Figure 6: Comparison in pairs (for specific head positions): photo of the face and reconstruction of the face for the purposes of animation. Minor differences result from the difficulty of repeating the identical position in the real comparison and the difficulty of choosing the right lighting.

On the other hand, older works (Paysan et al., 2009) make it possible to analyze the shape of the entire face surface, but with a complex 3D model of the face. It seems that in this context our method is a compromise. It uses traditional geometric analysis (we do not use any neural network), but allows analyzing the shape of the entire face with a given (limited) resolution. An additional application advantage (from the point of view of our project goal) is full integration with Blender software.

6 SUMMARY

In the work, the implementation of facial animation and control of facial expressions based on the recorded video was presented. The developed reconstruction algorithm was implemented directly in the Blender environment. This approach does not require any additional data conversion. The real face image was registered with the use of the iPhone 8 smartphone camera. The project confirmed the possibility of obtaining good results without spending large funds – neither expensive MoCap cameras nor specialized software were used.

An important advantage of using a smartphone and a set of libraries in the Blender system is the reduction of hardware requirements and lower costs. Work on the project will be continued. We hope that in the end we will be able to achieve photorealistic quality of the animation using the solution described here.

REFERENCES

Ai Face-Markerless Facial Mocap For Blender. (2021). Available at: https://blendermarket.com/products/ai-face-markerless-facial-mocap-for-blender. (Accessed 1 September 2021) .

Bedard M. (2020). What is Rotoscope Animation? The Process Explained. September 20, 2020. https://www.studiobinder.com/blog/what-is-rotoscope-animation-definition/. (Accessed 10 Septenber 2021).

Blender documentation. (2016). https://docs.blender.org/api/blender_python_api_2_77_release/. (Accessed 11 September 2021).

Blender Foundation. https://www.blender.org/about/. (Accessed 1 September 2021).

Borshukov G., Lewis J.P. (2003). Realistic Human Face Rendering for ”The Matrix Reloaded”, *SIGGRAPH 2003 Sketch #2* .

Cannell J., Jovic E., Rathjen A., Lane K., Tyson A.M., Callisaya M.L., Smith S.T., Ahuja K.D., Bird M.L. (2018). The efficacy of interactive, motion capture-based rehabilitation on functional outcomes in an inpatient stroke population: a randomized controlled trial. *Clin Rehabil*. 32(2), 191-200. https://doi.org/10.1177/0269215517720790.

Choosing a performance capture system for real-time mocap, (2020). White Paper. https://www.unrealengine.com/en-US/tech-blog/choosing-a-performance-capture-system-for-real-time-mocap. (Accessed 10 Sept 2021).

Face Cap - Motion Capture. Available at: https://apps.apple.com/us/app/face-cap/id1373155478. (Accessed 1 Sept 2021).

Faceware Facial Mocap. https://mocap.reallusion.com/iclone-faceware-mocap/. (Accessed 1 Sept 2021).

Graham P., Tunwattanapong B., Busch J., Yu X., Jones A., Debevec P., Ghosh A. (2013). Measurement-Based Synthesis of Facial Microgeometry. *EUROGRAPHICS 2013*. Vol. 32 (2013), Number 2

Hibbitts D. (2020). Choosing a real-time performance capture system. Jan. 29 2020, available at. https://www.unrealengine.com/en-US/tech-blog/choosing-a-performance-capture-system-for-real-time-mocap. (Accessed 15 Sept 2021).

Ichim A.E., Bouaziz S., Pauly M. (2015). Dynamic 3D avatar creation from hand-held video input. *ACM Transactions on Graphics* (ToG) 34 (4). 1-14. https://doi.org/10.1145/2766974.

Kazemi V., and Sullivan J. (2014). One millisecond face alignment with an ensemble of regression trees. In *Proc. of the IEEE conference on computer vision and pattern recognition*. 2014. https://doi.org/10.1109/CVPR.2014.241

Kitagawa M., Windsor B. (2008). *MoCap for Artists: Workflow and Techniques for Motion Capture*. Elsevier/Focal Press, Burlington.

Kruk van der, E., Reijne M.M. (2018). Accuracy of human motion capture systems for sport applications; state-of-the-art review. *European Journal of Sport Science*. 18(6), 806-819. https://doi.org/10.1080/17461391.2018.1463397

Li T., Liu S., Bolkart T., Liu J., Li H., ZhaoY. (2021). Topologically consistent multi-view face inference using volumetric sampling. In *Proc. of the IEEE International Conference on Computer Vision 2021 (ICCV 2021)*. 3824-3834. Available at: https://arxiv.org/abs/2110.02948. (Accessed 26 December 2021).

Luo H., Hu L., Nagano K., Wang Z., Kung H-W., Xu Q., Wie L., Li H. (2021). Normalized avatar synthesis using stylegan and perceptual refinement. In *Proc. of the 34th IEEE International Conference on Computer Vision and Pattern 2021 (CVPR 2021)*. https://doi.org/10.1109/CVPR46437.2021.01149.

Menache A. (2011). *Understanding Motion Capture for Computer Animation*. 2nd edition. Morgan Kaufmann, Burlington.

Nakano N., Sakura T., Ueda K., Omura L., Kimura A., Iino Y., Fukashiro S., Yoshioka S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in Sports and Active Living*. 2(May), 50 pages (2020). https://doi.org/10.3389/fspor.2020.00050

Parent R. (2012). *Computer Animation: Algorithms and Techniques*. 3rd Edition. Morgan Kauf-mann, Burlington.

Park H. and Kim D. (2021). ACN: Occlusion-tolerant face alignment by attentional combination of heterogeneous regression networks. *Pattern Recognition* 114 (2021) 107761.

Paysan P., Knothe R., Amberg B., Romdhani S., Vetter T. (2009). A 3D face model for pose and illumination invariant face recognition. In *Proc. of the Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*. https://doi.org/10.1109/AVSS.2009.58

Perakis P., Passalis G., Theoharis T., Kakadiaris I.A. (2012). 3D facial landmark detection under large yaw and expression variations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35.7 (2012): 1552-1564

Pueo B., Jimenez-Olmedo J.M. (2017). Application of motion capture technology for sport per-formance analysis. *Retos*, 32, 241-247 . https://doi.org/10.47197/retos.v0i32.56072

Qualisys. https://www.qualisys.com/. (Accessed 11 September 2021).

Shewchuk J.R. (1996). Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In: Lin, M.C. and Manocha, D., (eds), Applied Computational Geometry: Towards Geometric Engineering, *LNCS*, vol. 1148, pp. 203–222. Springer, Berlin.

Thompson B. March 31, (2020). How To Use Facial Motion Capture With Blender 2.8. Available at: https://www.actionvfx.com/blog/how-to-use-facial-motion-capture-with-blender-2-8. (Accessed 1 Sept 2021).

Topley M., Richards J.G. (2020). A comparison of currently available optoelectronic motion capture systems. *Journal of Biomechanics*, 106 (June), 109820. https://doi.org/10.1016/j.jbiomech.2020.109820.

Valle R., Buenaposada J.M., Valdés A., Baumela L. (2019). Face alignment using a 3D deeply-initialized ensemble of regression trees. *Computer Vision and Image Understanding*. 189 (2019) 102846.

Zane E. (2019). Motion-Capture Patterns of Voluntarily Mimicked Dynamic Facial Expressions in Children and Adolescents With and Without ASD. *Journal of Autism and Developmental Disorders*. 49(3), 1062–1079. https://doi.org/10.1007/s10803-018-3811-7

Zhou H., Hu H. (2008). Human motion tracking for rehabilitation—A survey. *Biomedical Signal Processing and Control*. 3(1), 1-18 (2008). https://doi.org/10.1016/j.bspc.2007.09.001

Zhu H., Sheng B., Shao Z., Hao Y., Hou X., Ma L. (2018). Better initialization for regression-based face alignment. *Computers & Graphics*. 70 (2018) 261–269.