Scientific and Social Research, 2025, Volume 7, Issue 7
https://ojs.bbwpublisher.com/index.php/SSR
Online ISSN: 2981-9946
Print ISSN: 2661-4332

Research on facial Micro-Expressions of Digital Humans on AIGC

Mingming Xu*

City Graduate School, City University of Malaysia, Petaling Jaya 46100, Malaysia

**Author to whom correspondence should be addressed.*

Abstract: Digital-human facial Micro-expression generation suit constructs the film CG's basic model. This can be used to upgrade the basic CG digital-human to an intelligent digital-human. Digital humans lack the reality in micro-expressions, so the effect is lost during application in film or interactive games. This paper introduced the face plus-in of a soft mod manipulator, accurately calculated the expression feature points by using the blend shape technology, and proposed a method to produce CG digital human micro-expressions using markerless point capture. The study showed that this micro-expression technology can be applied to realistic digital humans. Finally, this technical method of micro-expression production can achieve a satisfactory in-depth experience effect for the audience in virtual studio platforms, CG films, and interactive games. This study proves the feasibility and importance of this technology.

Keywords: Digital human; AIGC; Micro-expression; Non-marker point; Bland Shape

Online publication: August 12, 2025

1. Introduction

People's faces have complex structures that can produce complex and subtle changes in expression, and each tiny change can reflect the character's emotional changes and state of mind. How to rebuild the face of 3D characters has always been the research hotspot of the computer vision and graphics field. Nowadays, the facial expression of digital humans, except for games, movies, social contact, etc., also involves the technical issue of network security. In the future, following the development of metaverse technology, digital technologies will be expanded [1]. This paper points out the two types of research: On the one hand, controllers in software such as Maya or Unity3D are used to generate expression feature point parameters, which are adjusted to manipulate the micro-expressions of the model. On the other hand, the artificial intelligence generative component (AIGC) drives the model motion, which is very important for the ordinary digital-human gradual change to intelligent digital-human.

In the generated technology of micro-expressions for digital human faces, the data capture of faces is of vital importance. Data-capture modelling is a technique for loading motion data onto a model of an object's face.

It allows the object to display expressions according to the loaded motion data. In the traditional method of data capture that gauges points, the more frequent use of encapsulated, commercially available software is used to generate production campaign data. Example: Motion Builder, 3DsMax, Maya, etc., but the motion data generated after needs the edition and adjustment for artificial [2]. CG characters have difficulty in generating a natural face. Based on the technology of face generation of Blend Shape, the research suggests a method for the motion of the model of an unmarked point. Intelligent expression effects for digital avatars through AIGC empowerment.

2. Culture review and prospect

Digital human includes reality, digital human, and virtual digital human. Virtual digital human technology covers character modelling design, face 3D generation, character movement face capture, audio and video synthesis, display and interaction, etc., especially focusing on the character's head of the virtual digital human expression in the virtual world has a broad application prospect. In 2017, the State Council issued the New Generation Artificial Intelligence Development Plan clearly defined goals and a roadmap for the development of AI in a number of areas [3]. With the rapid development of artificial intelligence and big data, cloud computing and digital human technology have become the new industry's new wind direction. Especially in the entertainment industry, the rapid rise of virtual idols, virtual anchors in the Chinese market, has become a sought-after object for young people. The emergence of virtual idols such as Luo Tianyi has contributed to the expansion of the movie and television industry.

Abroad has already researched the field for many years. There are research organizations that try to achieve similar functionality with different technological routes. For example, in scholar Siddiqui's study, he presented a face-generated method for digital humans, called FExGAN-Meta. It works great for digital people to generate videos. Siddiqui's research results indicated that FExGAN-Meta stably classified the image face as simple and complex [4]. Similarly, in Zou K's study, which presented a specific micro-expressions technology, A series of 3D Model sequences were generated from the 3D model using the diffusion model, including the information which the alteration of the facial points by characters [5]. Such technology has achieved wide application in the fields of movies, music, and games. The virtual protagonists in the American movie Battle Angel Alita, Virtual singer Hatsune Miku, etc., because of their vivid facial expressions, have a large number of fans around the world.

3. Research methodology

3.1. A summary of the face of a virtual digital human

In the virtual digital human in the movie, the generation of micro-expressions of the digital human mainly relies on the head and facial motion capture of the real person, such as high-resolution cameras, expression capture systems, etc. James Cameron directed the sci-fi masterpiece "Avatar: The Way of Water" on the use of advanced digital human technology, the fictional world of humans and the Na'vi fusion, and for the actors to add a CG digital character with rich and delicate expressions, to provide the audience with a realistic visual feast [6].

This method is overly dependent on expensive equipment and professional performance. How to use cutting-edge artificial intelligence technology to break through these limitations, the low-cost realization of high-definition realistic digital human micro-expressions has quickly become the most popular virtual digital human technology, and also the combination of it and artificial intelligence technology to explore the future [7]. Because virtual digital humans are distinct from the marking capture algorithms of real human faces in movies, the method of traditional produce is to capture real face data by using marker points first, then directly match the model by motion data, but

this method has many problems, such as, many data require extensive corrections; expressions are poorly generated, etc., but the workflow proposed in this paper can solve the problem of generating facial micro-expressions for digital human characters [8]. While the digital human facial expression capture is completed, it is then published on the digital human platform to generate digital virtual objects based on the animations and sounds it adds.

3.2. Subtitle technology implementation process

3.2.1. Dilb library and Soft Mod Manipulator

The first step is to create facial controllable points on the 3D model. This can be done with the Soft Mod Manipulator facial expression generation plugin. It works seamlessly with Maya software to generate unmarked controllable points, which can then be used to control changes in the model's surface by binding them to the model. Dilb library is a more classic face feature detection library; its face feature point detection is fast and accurate [9]. With Dilb's face detection method, 68 feature points can be detected. In this paper, the traditional detection points are captured by capturing the parts of the model where important expressions occur, and then the Soft Mod Manipulator is applied to capture the points of these parts, and these feature points are transformed into capture points for the model. According to Dilb library's detection, the human face has some important points (**Figure 1**). Points such as those around the eyes, points in the middle of the brow, points on the sides and around the corners of the mouth, points in the cheek area, and points in the nose area are all points that need to be captured.

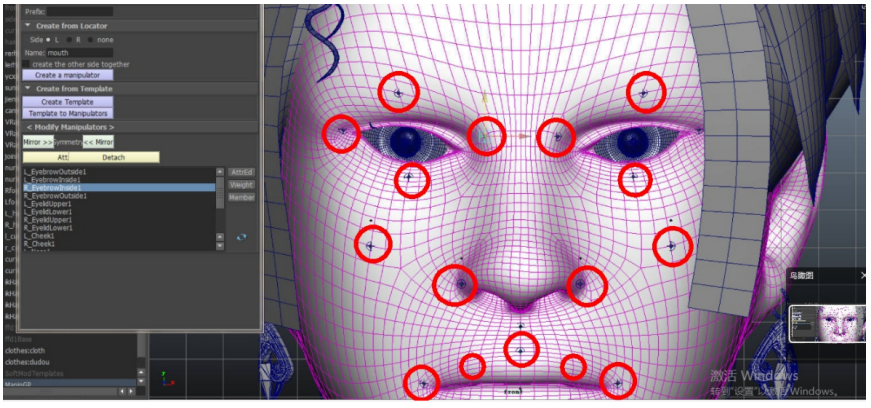

Figure 1. The character's main capture point of micro-expression

According to these points detected by Dilb, using the Soft Mod Manipulator plug-in to bind the above feature points of the character, the capture points of the face can be generated, and the parameters of these points can be adjusted and controlled separately in the panel, and finally, the successfully generated micro-expression (**Figure 2**).

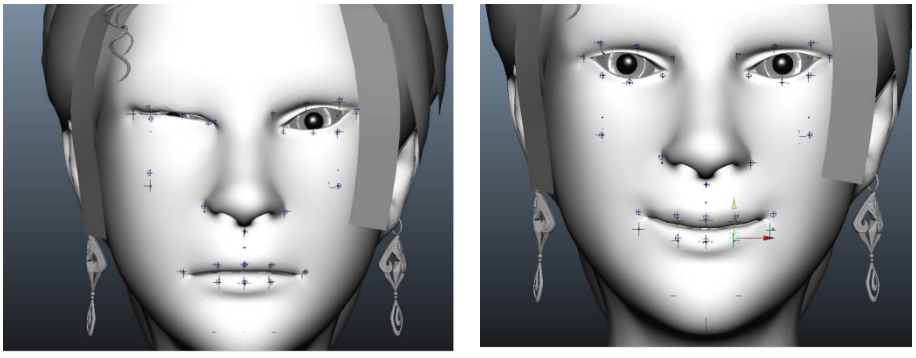

Figure 2. Dilb Detection, a micro-expression created by Soft Mod Manipulator

3.2.2. Bland Shape and merge micro-expression

In the production process, it is worth noting that each expression, such as the left eye closed, requires the corresponding application of Blend Shape to generate a separate expression, and then the model with the expression will be individually copied out to a new object model. The producer needs to generate as many emoticons as there are objects to be copied out. Eventually, the many model expressions generated will be centrally managed in Blend Shape. On the side, every face's key point is only one. The chart below divides and categorizes each key point and outlines 7 Deformation categories. Any combination of all these deformation points basically covers the face of the entire digital person (**Table 1**).

Table 1. Key for each facial expression

	Part	Expression	Key point	Key point name
A	Mouth	Open-mouth	57	Open-mouth
		Move-mouth	53/45	Move-L-R-mouth
		Pull-down	55/56	Pull-L-R-down
B	Left-eyes	Open-eyes	43,44	Open-eyes
		Move-eyes	42/45	Move-up-do-eyes
C	Right-eyes	Open-eyes	36,37	Open-eyes
		Move-eyes	38/35	Move-up-do-eyes
D	Nose	Up=down-move	32	Up=down-move
E	Face	Move-face	33/31	Move-up-do-face
F	Left-brow	Eyebrow-inner	21	Eyebrow-inner
G	Right-brow	Eyebrow-inner	19	Eyebrow-inner

3.2.3. The calculation method of key weights

Finally researcher utilizes the key weights algorithm to calculate the precise moving position of capture points. This study proposes that the preferred project is to import the Unity 3D from model data in Maya first, subsequently select a basis of Blend Shape in Unity 3D, and calculate the complete change value. After all key points are calculated in this way. The producer needs to adjust these points before the expression changes in order to gain the oriental coordinate of the key point, which has been named E_ori(n).

Example:

D(n)= E(n) - E_ori(n).

This method of weight calculation is based on the fact that the head does not move at all, and the digital human expression needs to be recorded into a video after it is generated, so the bias caused by the head movement cannot be avoided. Here is another way of calculating coordinates introduced, as a consideration of the problem, which is the distance and stability between facial points. The researcher needs to find one relatively constant points that generate the kinds of expression and is used as a reference point E0. Starting the standard phase, the coordinates of each point are E_ori. The difference D_ori of each point with respect to E0 is:

D_or(n)=E_ori(n)- E0_ori(n).

During the run phase, the variable of the difference of each point with respect to E0 is:

D(n)=E(n)-E0- D_or(n).

Of course, the above calculations are based on the case where the facial expression is basically unchanged or the change is not obvious. If the 2D coordinates are changed based on a video-like image, the positions of the key points will change if the expression of the character's face changes drastically, in which case they will have to be recalculated [10]. The above three chapters are the process of micro-expression creation using the innovative method of Blend Shape expression fusion and key weights combination.

3.3. The process of traditional digital human into intelligent digital human

In the previous chapter, the researcher investigated micro-expression generation techniques for CG characters, which can subsequently be utilized with the convenience of the AIGC generative style to convert the produced traditional digital human into an intelligent digital human. The platforms of intelligent digital humans included Niverse Avatar, Meta Human Creator for Unreal Engine, DI-D, etc. Take the DI-D platform as an example, loading one digital human, and generating digital human animations across modalities through forms such as inputting expressions, recording sounds, and Cross-textual. Digital human expressions become very rich as the AI takes control of them (**Figure 3**).

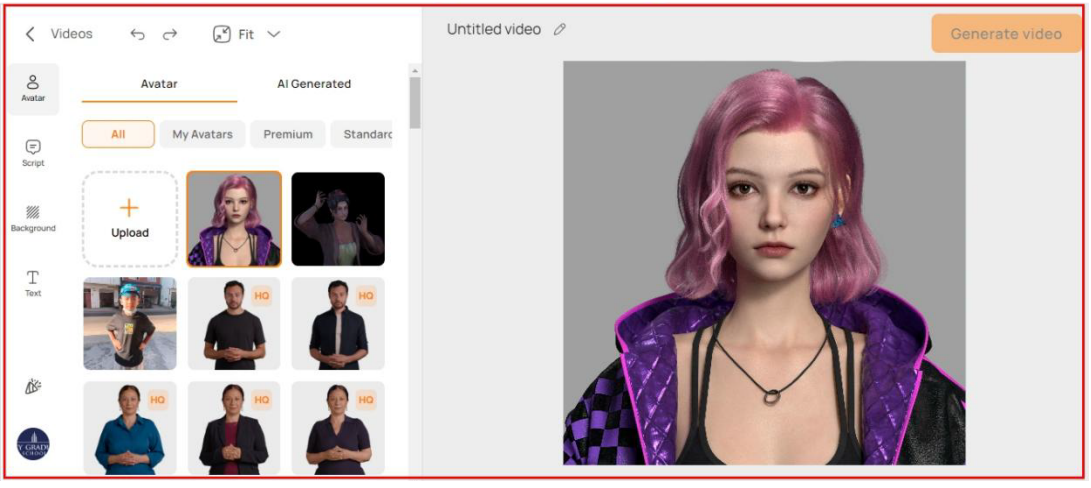

Figure 3. Intelligent digital human micro-expression generation sketch (Work source: Racing Games)

4. Conclusion

In this paper, the author proposes a complete production scheme for the generation of micro-expressions for digital humans from three aspects, namely, Blend Shape expression generation for digital humans, key weights computation, and AIGC-enabled digital humans. It is hoped that it will be helpful and inspiring to those who study computer vision and the creation of film special effects art. Recently, the development of AIGC has had a huge impact on cross-media image generation and video design. As the digital human technology of previous is continuously developed as well, besides the micro-expression technology in this paper, intelligent digital humans also involve motion capture, voice recording, virtual interaction, etc. In the end, this paper holds that only by constantly exploring new technologies and discovering new ideas can digital human technology be continuously updated and better serve humanity and society.

Acknowledgments

This research was conducted and completed under the supervision of Dr. Fauzi Naeim during my PHD studies, and I would like to thank the professor for his guidance.

Disclosure statement

The author declares no conflict of interest.

References

[1] Wang S, 2018, A Review of 3D Face Expression Acquisition and Reconstruction Techniques. Journal of System Simulation, 30(7): 2423–2424.

[2] Miao Y, 2016, Expression Data-Driven 3D Face Modeling Approach. Journal of Changchun University of Science and Technology, 39(2): 112.

[3] CCTV, 2024, Empowering Future Industries, the Central Economic Work Conference Laid out this Action. Wuhan Science: Technology and Innovation Bureau, 12.

[4] Siddiqui JR, Fexgan-meta, 2022, Facial Expression Generation with Meta Humans. arXiv preprint arXiv, 2203.05975.

[5] Zou K, Fai SS, Yu B, et al., 2023, 4D Facial Expression Diffusion Model. ACM Transactions on Multimedia Computing, Communications and Applications, 21(1): 1–21.

[6] You G, Chan Z, 2022, Research Article Motion Capture Technology and Its Applications in Film and Television Animation. Advances in Multimedia, 2022(6392168): 1–10. https://doi.org/10.1155/2022/6392168.

[7] Zhang MY, 2023, Application of Performance Motion Capture Technology in Film and Television Performance Animation. Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation (ISCCCA-13).

[8] Daněček R, Chhatre K, Tripathi S, et al., 2023, Emotional Speech-driven Animation with Content-emotion Disentanglement: SIGGRAPH Asia 2023 Conference Papers, 1–13.

[9] Xie Y, 2023, Overview of AIGC-empowered Digital People in Movies. Modern Film Technology, 2023(10): 34–35.

[10] Leiker D, Gyllen AR, Eldesouky I, et al., 2023, Generative AI for Learning: Investigating the Potential of Learning Videos with Synthetic Virtual Instructors. International Conference on Artificial Intelligence in Education. Springer Nature Switzerland, Cham, 523–529.