International Journal of Advanced Smart Convergence Vol.13 No.3 176-182 (2024)
http://dx.doi.org/10.7236/IJASC.2024.13.3.176

IJASC 24-3-20

Comparative Analysis of Facial Animation Production by Digital Actors – Keyframe Animation and Mobile Capture Animation

Choi, Chul Young

Professor, Major of Visual Animation, Dongseo University, Korea
freechoi@dongseo.ac.kr

Abstract

Looking at the recent game market, classic games released in the past are being re-released with high-quality visuals, and users are generally satisfied. It can be said that the realization of realistic digital actors, which was not possible in the past, is now becoming a reality. Epic Games launched the MetaHuman Creator website in September 2021, allowing anyone to easily create realistic human characters. Since then, the number of animations created using MetaHumans has been increasing. As the characters become more realistic, the movement and expression animations expected by the audience must also be convincingly realized. Until recently, traditional methods were the primary approach for producing realistic character animations. For facial animation, Epic Games introduced an improved method on the Live Link app in 2023, which provides the highest quality among mobile-based techniques. In this context, this paper compares the results of animation produced using both keyframe facial capture and mobile-based capture. After creating an emotional expression animation with four sentences, the results were compared using Unreal Engine. While the facial capture method is more natural and easier to use, the precise and exaggerated expressions possible with the keyframe method cannot be overlooked, suggesting that a hybrid approach using both methods will likely continue for the foreseeable future.

Keywords: *Animation Production, MetaHuman Creator, Studio Library, Maya, Unreal Engine*

1. Introduction

Recently, in the gaming market, remastered games of games that were popular in the past, such as ‘Residence Evil’ or ‘The Last of Us’, continue to be released and are achieving success. Improvements in production technology have made the existing story more tense and immersive by introducing realistic characters, delivering a new game feel. However, compared to the development of visual elements such as realistic characters and backgrounds, no notable technological progress has been made in the character animation production process. Maya, a representative animation production tool, has been established as the main tool for companies for more than 20 years, but recently, other tools are showing high growth. Also, recently, many

Manuscript Received: July. 23. 2024 / Revised: July. 30. 2024 / Accepted: August. 6. 2024
Corresponding Author: freechoi@dongseo.ac.kr
Tel: *** - **** - ****

Professor, Major of Visual Animation, Dongseo University, Korea

videos using A.I. have been exposed on Instagram and production methods have been revealed. Content produced using 'Deepfake' technology shows a completely different method from existing video production methods.[1] In this situation, Epic Games released a user environment in 2019 that allows users to easily apply realistic MetaHuman(MH) characters to Unreal Engine(UE) content.[2] Rather than creating modeling, it provides a way to create the desired character by modifying an existing model. This can be seen as one of the node-based production methods pursued by UE and Blender software. Not only modeling, but also lighting and compositing processes are operated by node base, providing a fast and high-quality workflow. For character animation currently uses keyframe animation and motion capture production methods. Although the level of perfection of each production technology, process, and result is still very high, there is room for improvement in the required production manpower and production efficiency, so new types of animation software are being released recently.[3] And in the case of facial animation, there are still three methods used to express facial expressions and dialogue. The first is facial capture technology using markers with camera, the second is facial capture using a mobile app, and the third is producing keyframe animation using blend shapes.

There was a lot of perception that the second method was of lower quality than the first or third method. However, the method improved by Epic Games in 2023 showed a significant quality upgrade. So, in a situation where face capture is considered an increasingly popular technology, we compared the second and third methods and analyzed the results. To increase the accuracy of the two methods, we created dialogues for four types of emotional expressions, created facial capture animations using keyframe animation and iPhone, and compared the results in UE.

Figure 1. Deepfake Animation(left), [1] AI Animation(Right)[4]

2. Experiments

2.1 Keyframe Animation using the Pose Library

In the case of MH, keyframe animation can be performed using the facial control interface. However, because there are too many controllers to control a realistic face, it is inefficient to create each facial expression by moving the controllers. Therefore, it can be said that it is effective to save the necessary facial expressions in the pose library by moving the controller and then create the desired facial animation using keyframe animation using the saved facial expressions. In order to use the pose library, MH was imported into Maya Software. Once the facial animation is completed, the animation data can be used again as facial animation data in UE's MH in fbx format. To use the pose library, we used 'Studio Library', a Maya plug-in. Table 1 shows the facial expressions and pronunciations required for facial animation. The poses of facial animation can be divided into two types. One is a pose that expresses emotions, and the second is a pose for vocalization. The criteria for distinction are constantly changing and may differ from studio to studio. In the past, the poses were divided into three parts: the area around the eyes, the area around the nose, and the area around the mouth. This method can be said to reduce production efficiency because it requires too many facial expressions from

the library, so many facial expressions must be created. Recently, the method of creating basic data by reducing the necessary facial expressions as much as possible and modifying the basic data as needed to create new ones is mainly used.[10] The additional facial expressions in Table 1 are facial expressions created by modifying the basic facial expressions. In the second case of vocalization, the required data varies depending on whether the character is realistic or a cartoon character. In the case of cartoon characters, most expressions are usually possible with just five mouth shapes. In the case of MH, they are slightly more realistic characters, so we tried to express their lines by composing nine mouth shapes.[9] When importing into Maya, if the MH's name is the same, different types of characters can use the same library to create facial expression changes, so it can be said that production efficiency can be increased by configuring the pose library only once.

Table 1. Studio Library Facial Classification Table[9]

Dialouge	A,I	E	F,V	C,D,G,K,R,S,T	L	M,B,P	O	U	W,Q
Shape									
Base Emotion	Smile	Angry	Surprise	Cry	Like	Disappoint-ment	Try Hard	Annoying	
Shape									
Xtra-Emotion	Big Smile	Decide	Satisfied	Focus	Chew	Hopeless			
Shape									

A reference video is also needed to produce facial animation with keyframes. If there is a video recorded for sound, the animator creates the key pose of the facial expression while watching the video. If there is no video, the animator creates the key pose by looking in the mirror and referring to the person's face. The manufacturing process is the same as shown in Figure 2. Data is passed between Maya and Unreal to verify the results. The data format uses fbx format. Since errors may occur depending on the frame rate and data naming structure, the production workflow must proceed after establishing the data standards required for delivery.

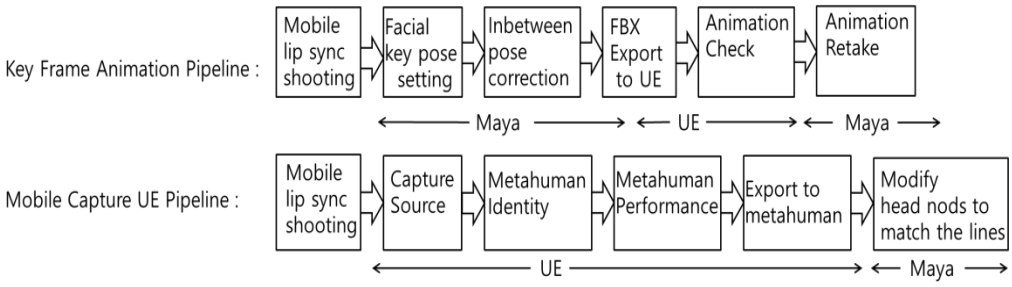


Figure 2. Keyframe and Mobile Facial Capture Animation Production Process

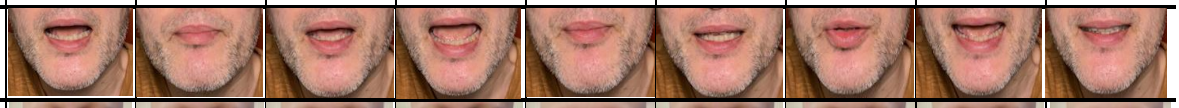
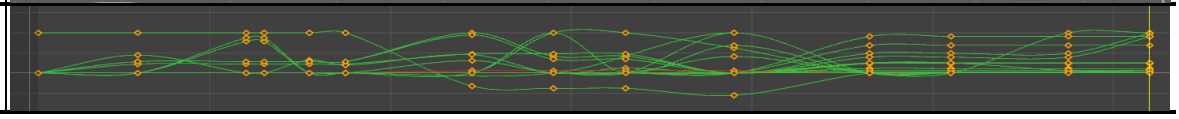
2.2 UE Facial Animation using Mobile Capture

The process of facial capture using video captured in UE is shown in Figure 2. In order to accurately extract the creator's facial expression, a dummy character face is created by basically inputting the creator's data. Using images showing the creator's front and side views and teeth, a dummy model with an accurate facial shape is created, as shown in Figure 3. Because this method is based on the creator's facial shape, it can produce higher synchronization rates than traditional mobile capture results. The dummy character very accurately embodies the creator's facial expressions in the video. Then, if the facial expressions are shown as the final MH instead of the dummy character MH, a high-quality facial expression result is created.

Figure 3. Identity Process for Mobile Capture

The face capture image in Table 2 is the result of facial expression retargeting to the face of the main actor that will be shown on the screen, rather than a dummy character based on the creator's face. Comparing the facial expression images, it seems that the facial expressions created using the keyframe method make the character's emotions easier to read, and when comparing the curves in the graph editor, the facial expression capture data appears to create more complex and natural facial expression changes.

Table 2. Facial Capture Data and Keyframe Data Comparison

Frame	12	24	31	49	58	66	78	93	115
Live Shot									
Facial Capture									
Keyframe Animation									
Facial Capture Curve									
Keyframe Animation Curve									

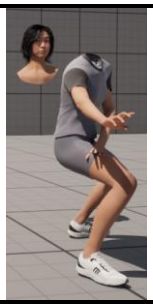
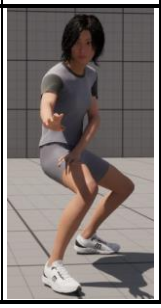
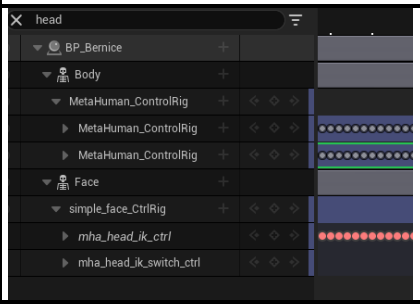
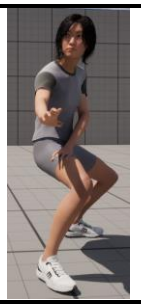
2.3 Sync Facial Animations and Movements

Facial animation is not completed by implementing facial expressions. This is because facial movements are expressed depending on the level of emotion or intensity of tone during a conversation, so if this is ignored,

the character's movements may become awkward, so ignoring this can make the character's movements awkward. So, in the case of keyframe animation, head movement is added while looking at a reference image or a mirror to create natural head movement according to vocalization. In the case of mobile capture, the head rotation is also recorded, so using this can further increase the efficiency of the production process.

At this point, problems may arise. This is because head rotation values are applied to mobile capture, and head rotation values may also be present in motion capture or when animating the body with key frames. In this case, the head moves separately from the body. There are two ways to solve this problem: using a UE Blueprint and solving it with copy and paste head rotation data in a level sequence. The solution in Level Sequence is to bake the body and facial capture layers, make the keyframes visible, and then change the head rotation data from mobile capture to the head rotation data from body animation. The level sequence solution is a more intuitive way to do it. Since head rotation data present in the body animation may also be required, the most appropriate solution is to add an additional layer of body animation and apply the head rotation data from the mobile capture data to it rather than deleting the head rotation data from the body.

Table 3. Synchronize Process Head Rotation According to Vocalization

	A	B	C	D
Synchronize Process				

A : Both head and body have head rotation information
B : When head rotation information is turned off
C : Create an additive layer for the body and move the head rotation information
D : Head rotation information captured on the mobile is shared with the body's head rotation information

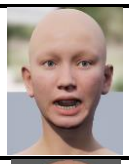
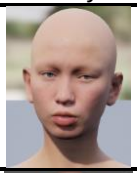
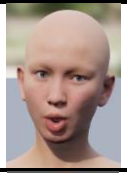
As a result, in order to implement the natural facial expressions and dialogue of realistic characters such as MH, it is easier to use mobile capture data than keyframe data and can produce natural results, but it does not express emotions as accurately as keyframe animation. This is an area where the process of producing keyframe animation is still considered necessary. For a more accurate test, we attempted to compare longer sentences with different emotions.

3. Results

Table 4 shows lines expressing four different emotions produced using keyframe animation and mobile capture, and a table comparing the scenes corresponding to each word. Looking at the table, we can see that, in general, the results created with keyframe animation are easier to read facial expressions. This is because exaggerated emotions can be seen compared to the results of applying mobile capture data. As can be seen from the keyframe animation curve in Table 2, keyframe animation has the advantage of applying only necessary and accurate facial expressions and making emotional expressions accurate and remembered for a long time through moving holds of important facial expressions. However, mobile capture data can be said to have consistent quality and more accurate mouth shapes compared to keyframe animation. In Table 2, the

curve of the mobile capture data continues to change, which means that the facial expression is continuously changing. Therefore, when comparing the two methods using video rather than images, the facial capture method delivers the feeling of MH being more alive because the mouth shape and facial expression matching the sound are very important for a live-action character. If it were a cartoon character, the opposite result might have been obtained.

Table 4. Mobile Capture Data and Keyframe Data Emotional Expression Comparison

Emotion	Dialogue	I'm	very	happy	Today.
Joy	Mobile capture				
	Keyframe				
Emotion	Dialogue	Why	are you	doing	this to me?
Angry	Mobile capture				
	Keyframe				
Emotion	Dialogue	I don't	want to	talk to you	right now!
Sorrow	Mobile capture				
	Keyframe				
Emotion	Dialogue	This is the	best	day of	my life!
Pleasure	Mobile capture				
	Keyframe				

As a result, the more realistic the shape of the character is, the more easily and naturally the audience can accept the face capture method using a mobile phone rather than the key frame method of facial expression using blend shapes. Nevertheless, because keyframe animation has the advantage of being able to implement accurate and exaggerated emotional expressions, there still seems to be room for its inclusion in the production process.

4. Conclusion

Keyframe animation using the pose library has been widely used as a method of creating non-realistic facial expressions of characters. However, as the number of content using realistic characters such as MH increases, the need for facial capture is increasing. In this situation, the upgrade of Epic Games' mobile capture method shows usability as good as key frame animation with increased accuracy and quality. We created and compared four emotional expression situations using keyframe animation and mobile capture. The results showed that mobile capture data produces accurate and natural mouth movements. Keyframe animation has the advantage of allowing larger and more exaggerated expressions of emotion, but the disadvantage is that the mouth movements are somewhat inaccurate and the quality is unstable. This can become an element that hinders the audience watching the content from drawing attention and concentrating. Methods that leverage mobile capture data can produce reliable and consistent quality. However, because the level of expression expressed by the actor is different, the quality of the result is bound to vary greatly depending on the actor, which can be said to be a major disadvantage. In order to compensate for these shortcomings, it is believed that keyframe animation processing using the pose library is still necessary to clearly convey emotions.

References

[1] Deepfake Singularity, *https://www.youtube.com/watch?v=oxXpB9pSETo*
[2] MetaHuman Creator, *https:// MetaHuman.unrealengine.com*
[3] Cascadeur, *https://www.youtube.com/@Cascadeur*
[4] A.I. animation, *https://www.youtube.com/watch?v=OY48vtnYzD0*
[5] Gu Jie, J.W. Hwang, and C.Y. Choi, "Case Study of Short Animation with Facial Capture Technology Using Mobile," International Journal of Internet, Broadcasting and Communication, Vol. 12, No. 3, pp. 56-63, Sep. 2020. DOI: http://dx.doi.org/10.7236/IJIBC.2020.12.3.56
[6] Yangyang He, C.Y. Choi, "A Study of Facial Expression of Digital Character with Muscle Simulation System," International Journal of Advanced Smart Convergence, Vol.8, No.2, pp.162-169, June 2019. DOI: http://dx.doi.org/10.7236/IJASC.2019.8.2.162
[7] Gu Jie and C.Y. Choi, "Case Study of Motion Capture Data Processing for Actual Digital Actor Motion Implementation," The Korean Journal of Animation, Vol. 14, No. 2, pp. 23-41, June 2018. DOI: 10.51467/ ASKO.2018.06.14.2.23
[8] C.Y. Choi, K.D. Choi, " Production Technology and Method of Rig System for Facial Capture of Digital Creatures ", The Korean Journal of Animation, Vol. 8, No. 2, pp.60~78, June 2012.
[9] Choi, Chul Young, "Case Study of Animation Production using MetaHuman," International Journal of Advanced Smart Convergence, Vol.11, No.3, pp.150-156, Sep. 2022. DOI: http://dx.doi.org/10.7236/IJASC.2022.11.3.150
[10] C.Y. Choi, "Study of Character Animation to improve Production Efficiency," International Journal of Advanced Smart Convergence, Vol. 9, No. 2, pp. 179-184, June 2020. DOI: http://dx.doi.org/10.7236/IJASC.2020.9.2.179
[11] J.W. Song and J.Y. Noh, "Direct Retargeting Method from Facial Capture Data to Facial Rig," J. of KCGS, Vol. 22, No. 2, pp.11-19, June 2016.
[12] Noor Adibah Najihah Mat Noor, N. M. Suaib, " Review on 3D Facial Animation Techniques," International journal of engineering and technology, Vol.7, No.4.44, pp.81-187, Dec. 2018. DOI: https://doi.org/10.14419/ijet.v7i4.44.26980