International Journal of Internet, Broadcasting and Communication Vol.14 No.4 64-79 (2022)
http://dx.doi.org/10.7236/IJIBC.2022.14.4.64

IJIBC 22-4-9

Research on Technology Production in Chinese Virtual Character Industry

Yang Pan , KiHong Kim , JiHui Yan

Doctor, Department of Visual Contents, Dongseo University, China.
Professor, Department of Visual Contents, Dongseo University, Korea.
Professor, Department of VR Contents, Qingdao Huanghai University, China.

pyy8976341@naver.com, khkim@g.dongseo.ac.kr, ikkity@gmail.com

Abstract

The concept of Virtual Character has been developed for a long time with people's demand for cultural and entertainment products such as games, animations, and movies. In recent years, with the rapid development of concepts and industries such as social media, self-media, web3.0, artificial intelligence, virtual reality, and Metaverse, Virtual Character has also expanded new derivative concepts such as Virtual Idol, Virtual YouTuber, and Virtual Digital Human. With the development of technology, people's life is gradually moving towards digitalization and virtualization. At the same time, under the global environment of the new crown epidemic, human social activities are rapidly developing in the direction of network society and online society.

From the perspective of digital media content, this paper studies the production technology of Virtual Character related products in the Chinese market, and analyzes the future development direction and possibility of the Virtual Character industry in combination with new media development directions and technical production methods. Consider and provide reference for the development of combined applications of digital media content industry, Virtual Character and Metaverse industry.

Keywords: *Virtual Character, Virtual Digital Human, Virtual Idol, Virtual YouTuber, Metaverse, Motion Capture, Live 2D, Photo Scan*

1. Introduction

In recent years, China's digital media content industry has developed rapidly, and the digital media market has been greatly expanded. The surge in user demand and the support of technological development have also made Virtual Character related industries develop rapidly. At this stage, the Virtual Character-related markets in China and Japan are the most active, and countries such as the United States and South Korea are also actively accelerating the layout of Virtual Character-related industries, providing a foundation for the combined application of the Metaverse and AI-related industries. With the development of CPU, GPU and

Manuscript Received: September. 2, 2022 / Revised: September. 6, 2022 / Accepted: September. 10, 2022
Corresponding Author: khkim@g.dongseo.ac.kr
Tel: +82-51-320-1704
Professor, Department of Visual Contents, Dongseo University, Korea.

other hardware, the continuous update and iteration of traditional 3D production software, the improvement of game engines and cross-border applications, the digital media content industry has begun to show development bottlenecks in the direction of visual production. The development of Virtual Character related industries and the expansion of market demand have made the digital media content industry gain a new development direction. At the same time, in the Chinese market, stimulated by factors such as self-media, short video platforms, and streaming media platforms, the demand for digital media products has gradually shifted from terminal finished products to front-end and mid-end digital media product production solutions. [1]

As shown in Figure 1. According to the virtual character production technology and its form in the market, a general system framework can be summarized.

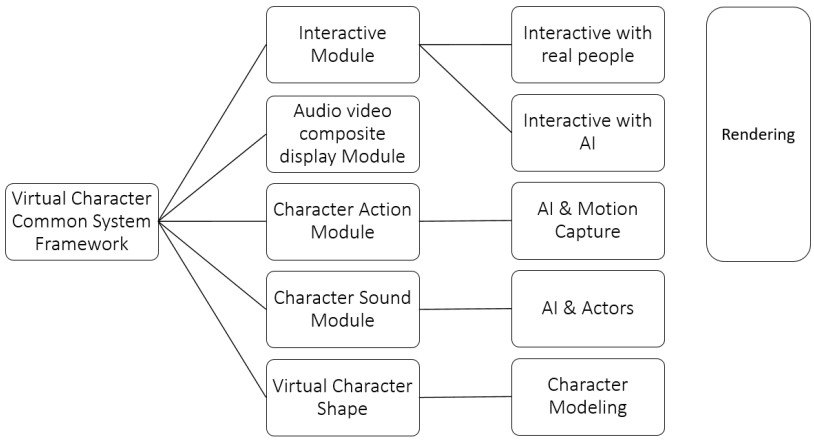

Figure 1.Virtual Character Common System Framework

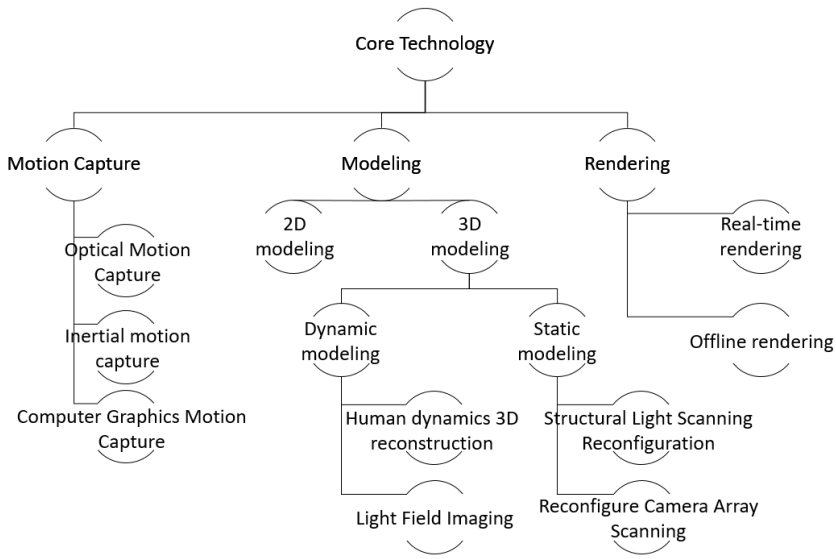

Figure 2.Virtual Character Core Creation Technology Classification

As shown in Figure 2. Behind the general system framework of virtual characters, modeling, AI, motion capture, and rendering are the four key technologies. From the perspective of the digital media content industry, we focus on the research and analysis of motion capture, modeling, and rendering. [2]

2. Research on Virtual Character Motion Capture Technology

2.1 Motion Capture Technology

Commonly used motion capture technologies are divided into mechanical, acoustic, electromagnetic, active optical, passive optical, and gyro-ceremony inertial navigation motion capture in principle. Devices with different principles have their own advantages and disadvantages, which can generally be evaluated from the following aspects: positioning accuracy; real-time; ease of use; range of motion that can be captured; anti-interference; Analyze software connectivity.

With the rapid development of computer software and hardware technology and the improvement of production requirements for digital content products, motion capture has entered a practical stage in many countries, and many companies have successively launched a variety of commercialized motion capture equipment, such as MotionAnalysis, Polhemus, Sega Interactive, MAC, X-Ist, FilmBox, Nokov, etc., are successfully used in virtual reality, gaming, ergonomics research, simulation training, biomechanical research and many more.[3]

From a technical point of view, the essence of motion capture is to measure, track and record the motion trajectory of objects in three-dimensional space. A typical motion capture device generally consists of the following parts:

Sensor

The so-called sensor is a tracking device fixed at a specific part of the moving object, which will provide the motion capture system with the position information of the moving object. Generally, the number of trackers will be determined according to the detail of the capture.

Signal capture device

This equipment will vary depending on the type of motion capture system, which is responsible for the capture of positional signals. For mechanical systems it is a circuit board that captures electrical signals, and for optical motion capture systems it is a high-resolution infrared camera.

Data transmission equipment

Motion capture systems, especially those requiring real-time effects, need to quickly and accurately transmit a large amount of motion data from signal capture devices to computer systems for processing, and data transmission devices are used to complete this work.

Data processing equipment

The data captured by the Motion capture system needs to be corrected, processed and combined with the three-dimensional model to complete the computer animation production work, which requires us to apply data processing software or hardware to complete this work. Whether it is software or hardware, they all use the high-speed computing power of the computer to complete the data processing, so that the 3D model can truly and naturally move.

In the 1970s and 1980s, motion capture began as a photographic image analysis tool in biomechanical research. As the technology matured, the technology began to expand to education, training, sports, computer animation, television, movies, and video games. And other fields. The user is equipped with marker points (Markers) at each joint, and the movements are identified by the changes in the position and angle between the

marker points.

With the wide application of motion capture technology in the digital content industry, most of the current mainstream motion capture devices are optical motion capture devices. With the development of related electronic technologies, gyro-ceremony inertial navigation motion capture devices have begun to show its unique advantages.

Virtual Idol and Virtual YouTuber are a major application of Virtual Character in the field of culture and entertainment. Related Virtual Characters can be divided into human-driven and computing-driven. The human-driven type needs to be driven by capturing real actions or expressions through motion capture devices or cameras, while the computing-driven type can automatically read and parse external input information through an intelligent system. , and generate corresponding voice actions to interact with users.

Common life-driven Virtual Characters need to complete six steps of image design and modeling, modeling and binding, motion capture and driving, rendering, content generation, and interaction. Specifically, the final image of Virtual Character mainly includes four aspects: motion capture, gesture control, lip sync, and expression control.

2.2 Classification of Motion Capture Technology

Optical Motion Capture

Optical uses optical perception to determine the real-time position and orientation of objects. Optical devices mainly include photosensitive devices (receivers), light sources (emitters), and controllers for signal processing. There are various photosensitive devices, such as ordinary cameras, photodiodes, etc. The light source can be ambient light or structured light. In order to prevent the interference of visible light, infrared rays, lasers, etc. are usually used as light sources. Due to the fast propagation speed of light, the most significant advantages of optical devices are fast speed, high update rate and low delay, which are more suitable for occasions with strong real-time performance and work well in a small area. The working principle of infrared passive optical motion capture: a capture space composed of multiple cameras, the near-infrared LED on the camera illuminates the reflective marking points on the target, the camera performs infrared imaging on the marking points, and extracts the two-dimensional information of the marking points. The three-dimensional position information of the marker point is calculated from the spatial data fed back by each camera to the same marker point. The motion capture system will complete the continuous shooting, image storage, analysis, and processing of the performer's movements, and complete the real-time recording of the movement trajectory.

As shown in Figure 3. Optical motion capture is a method that uses multiple high-speed cameras to detect target feature points from different angles, and combines algorithms to complete motion capture. It can be divided into non-marked point type and marked point type. The marker point method captures the suit marks on the person through the camera installed on the venue to obtain the person's position and motion information; the non-marker method directly detects the human shape and extracts the general motion trajectory map of the joints. The advantage of the optical motion capture system is that it has high accuracy, many people can be present at the same time, and accurate positions can be obtained, but it requires a large number of cameras, and requires high operating costs and the number of professionals in the early stage.

Figure 3.Optical motion capture schematic (Data source: Ubisoft La Forge studio)

Representative manufacturers of optical motion capture systems are VICON and OptiTrack:

1) OptiTrack as a full-body motion capture system, the price is slightly lower than VICON; the whole set includes motion capture camera, motion capture software and other accessories, using any head display (HMD) or virtual simulation system (CAVE), you can get extremely low Latency and very smooth tracking effects. The main users in the industry are the Vtuber virtual girl group MaRiNaSu.

2) VICON's optical motion capture system is widely used in medical, sports and other fields. The VICON system was first used in animation production for film and television, and was the world's first optical motion capture system for animation production. In 2017, Vicon acquired IMeasureU, the founder of inertial sensing technology and a manufacturer of motion measuring instruments. In the future, it may combine optical and inertial motion capture technology to continuously enhance motion measurement capabilities. The Japanese virtual idol company Rainbow Club introduced the VICON system in October 2019 to enhance the 3D effect.

As shown in Figure 4. In the process of studying optical motion capture, we found that the fatal disadvantage of optical motion capture is the occlusion of the sensor. Before using the motion capture device, it is necessary to debug the device to ensure the correct operation of the device. There are some differences between different brands of motion capture devices. This article takes the optical motion capture device OptiTrack as an example, using a 16-camera surround solution, and the software used is MOTIVE 2.0.1 Final that is matched with the motion capture device.

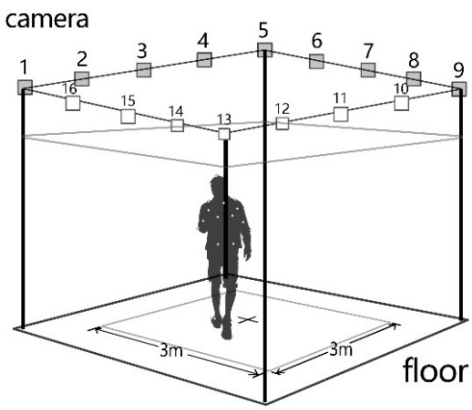

Figure 4.Optical motion capture device OptiTrack schematic diagram

Gyro Ritual Inertial Navigation Motion Capture

Gyro ritual inertial navigation motion capture, referred to as gyro ritual motion capture. The principle of gyroscopic motion capture equipment and optical motion capture equipment is different. The motion capture device of gyroscopic ceremony senses its own movement through the sensors on the actor, and then sends the data to the computer. The computer will calculate the actor's movement according to the data of each sensor. Action data.

Through inertial navigation sensor AHRS (Attitude Reference System), IMU (Inertial Measurement Unit), the performer's motion acceleration, azimuth, tilt angle and other characteristics are measured. Not affected by environmental interference, not afraid of occlusion. High capture accuracy and high sampling speed of 1000 times per second or more. Due to the use of highly integrated chips and modules, it is small in size, small in size, light in weight and cost-effective. The sensor is worn on the performer's head, or through 17 sensors to form a data suit, which is connected to the host through USB cable, Bluetooth, 2.4Gzh DSSS wireless, etc., respectively, can track head and whole body movements, and display complete movements in real time.

As shown in Figure 5. The gyroscopic ritual motion capture device used in this paper is XSENS, with a total of 17 sensors. The data transmission equipment of this device uses ASUS Wi-Fi network equipment. This device is powered by Wi-Fi 5 technology (802.11ac).

Inertial motion capture is the real-time measurement of the movement of the main skeletal parts of the human body through inertial sensors worn by each joint of the character. The inertial solution has the characteristics of high cost performance, not limited by the venue and light, and easy to use, but it cannot accurately obtain the absolute position of each sensor, and it is difficult to perform motion capture for multiple people at the same time, and the inertial sensor will accumulate errors, which will cause excessive use time. Misplaced, unable to track human movements for a long time.

Figure 5.Inertial motion capture diagram

Representative manufacturers of inertial motion capture systems include XSens and Noitom:

1) MVN Studio of XSens: The inertial motion capture solution is widely used in Vtuber and animation production. It has the advantages of short preparation time, easy operation, long use time, high cost performance, and is not limited by light and space, the feature of supporting multiple people to perform motion capture at the same time.

2) Perception Neuron, produced by the domestic manufacturer Noitom, was used in the field of virtual idols earlier and has a certain price advantage.

Computer Graphics Visual Capture

Facial expression capture is also a part of motion capture technology, which refers to the process of using mechanical devices, cameras and other equipment to record human facial expressions and movements and convert them into a series of parameter data. In the early days, facial capture was mainly accomplished by rotating a mechanical device, but with the development of technology, the capture device has been iterated to an optical device in the past ten years.

Facial expression capture solutions mainly include image recognition and iPhone depth camera recognition:

1) Image recognition: The actor's expression is captured by the camera and imported into the virtual idol's facial expression. The more commonly used software is Faceware. Founded in 2001, Faceware is a hardware and software company that has won two Oscar nominations for Science and Technology and a Lumiere Award for Best AR Experience for its facial capture solutions. In addition, 7 of the top 10 games in 2017 use Faceware's solutions. At present, Faceware users have expanded to more than 60 countries around the world, mainly involving 3A game development, animation film and television production, virtual reality, augmented reality and other fields.

2) IPhone depth camera recognition: It mainly tracks and detects the 3D information of the person through the depth camera of the iPhone, records the facial expression of the person, and then converts it into the expression of the avatar. Both Unreal and Unity have implemented this method as open source, and such micro-expressions can be mass-produced, further reducing the difficulty of production.

3. Research on Virtual Character Modeling Technology

3.1 2D two-dimensional class virtual character modeling technology

In the application of Virtual Characters such as Virtual YouTuber and Virtual Idol, traditional 2D art style and two-dimensional, Japanese art style Virtual Character occupies a large part of the market. Different from traditional 2D plane technology, based on 2D animation and computer graphics, 2D characters using 2D technology and 2D art have also developed movable and controllable character models. Not only can you design and make 2D character models according to your needs, but you can also use facial capture, motion capture and other technologies to achieve real-time interactive functions. As shown in Figure 6. This technology is Live2D technology, referred to as L2D. [4]

Figure 6.Schematic diagram of L2D technology

"Live2D" is a graphics rendering technology used in video games, developed by Japanese Cybernoids. Through a series of continuous images and character modeling to generate a 2D image that resembles a 3D model, it is very useful for an animation-based adventure game. The disadvantage is that Live 2D characters cannot turn around a lot. The developers are trying to Allows the technology to display 360-degree images.

Live2Dized characters can interact in a number of different ways. Since its release in 2008, it has been adopted by numerous software (games). In addition to the home game system, Live2D can also be used on Android (Android), IOS, Play Staion, Play Staion Vita, Nintendo 3Ds and other home game devices. At the same time, we are also preparing to develop the corresponding Unity version.

The Graphics Principle and Realization of Live2D Animation Engine

Live2D Euclid technology is still in the development stage. The technology that has been applied is Live2D Cubism.

Live2D Cubism. This is completely different from 3DCG rendering 2D results, which is not more advanced. As shown in Figure 7. Animations can be seen in Live2D, which are achieved through two-dimensional translation, rotation, and deformation operations. So Live2D Cubism is indeed a veritable 2D technology without using any 3D technology.

But Live2D Cubism does have improvements and innovations compared to previous 2D animation software, and it also absorbs some of the ideas of 3D animation software. For example, the model is managed in a tree structure, and operations on the upper control grid will affect all other parts below that grid.

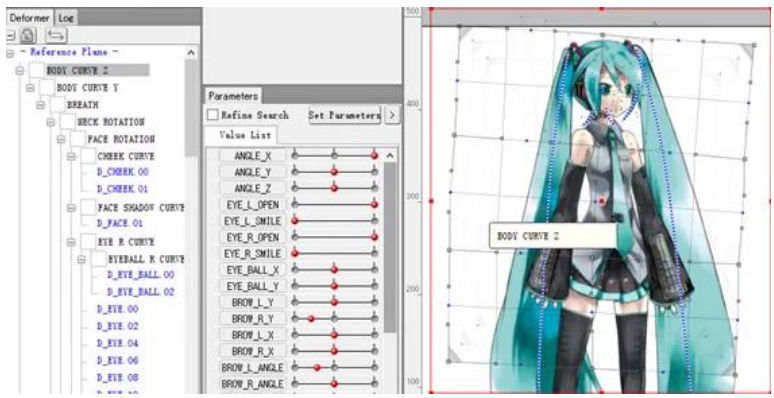

Figure 7.L2D character model Action realization principle diagram

To sum up, the advantages of Live2D Cubism over pure 3D animation are low threshold, low cost, and good effect (the effect of non-real rendering of NPR).

3.2 3D realistic virtual character modeling technology

The first step in creating a digital character requires pre-image design and modeling. 2D digital characters require image design such as original paintings, while 3D digital characters require additional use of 3D modeling technology to generate digital images, the information dimension increases, and the required amount of calculation is larger, whether it is based on IP or real-life design, face and body modeling is required.

At present, 3D modeling technology mainly includes static scanning modeling and dynamic modeling. Static scanning modeling is still the mainstream, among which camera array scanning and reconstruction are developing rapidly. Currently, millisecond-level high-speed photo scanning can be achieved (high-performance camera array accuracy To meet the needs of virtual character scanning and reconstruction, replacing structured light scanning reconstruction has become the current mainstream method of character modeling. Compared with static reconstruction technology, dynamic light field reconstruction can not only reconstruct the geometric model of the character, but also obtain the dynamic character model data at one time, and reproduce the light and shadow effects of the human body from different perspectives with high quality,

become the key development direction of virtual character modeling.

PhotoScan Technology - A photo-based 3D model building method

1) The technical advantages of photo modeling are reflected in:

A) The shape, size and location of the subject are not restricted when shooting;

B) The learning threshold is low, and beginners without modeling experience can also operate;

C) Relying on the relevant software processing system, you can choose fully automatic or semi-automatic modeling methods, and can also process multiple 3D modeling projects in batches, which greatly reduces the workload of modelers;

D) The requirements for hardware equipment are low, using ordinary digital cameras, even mobile phones can be used as modeling input devices, and modeling equipment only needs a high-performance workstation. The above advantages are destined that the photo modeling method will become an important development direction in the field of 3D modeling in the future, and can be widely used in aerial remote sensing measurement, 3D landscape map production, 3D printing, archaeology, film and television production and other fields. [5]

2) Principle of photo modeling technology

The entity pictures captured by input devices such as cameras are two-dimensional images, and the three-dimensional modeling based on photos needs to analyze multiple two-dimensional photos obtained, and extract the three-dimensional information of the object from them.

First, the camera should be calibrated, and the internal parameters of the camera, such as focal length, optical center coordinates, etc., should be reversed from the obtained photos; external parameters, the shooting direction and position information of the camera in the actual three-dimensional space.

3) Photo modeling process

Image acquisition:

Refers to the digital image acquisition process of the modeling target, which is the starting point of photo modeling.

Image feature matching:

This process includes the identification, extraction, analysis and matching of image features. Regarding feature extraction, it is the process of transforming the original sample features of high-dimensional space into low-dimensional space features to describe.

Camera calibration:

This process is to obtain camera information. According to the parameter properties of calibration, camera calibration can be divided into internal calibration and external calibration. Internal calibration refers to the camera's internal geometric optical parameters, such as focus, focal length, lens distortion, and entrance pupil; external calibration refers to the camera's spatial transformation parameters, such as position and rotation information.

Sparse point cloud:

After the image is subjected to feature matching and camera calibration, the image of the sparse point cloud can be obtained, and the external calibration process of the camera is carried out simultaneously with this.

Spatial dense point cloud:

After getting more detailed scene information, you can use this to perform multi-view stereo reconstruction. The number of photos taken, shooting quality, lighting conditions, the complexity of the subject itself, and the target material are all key factors that affect the quality of dense point clouds.

Meshing:

There are many meshing algorithms, and the most used is the mesh model generated based on the triangular surface structure. After meshing, the obtained model should continue to be optimized. On the one hand, the holes are repaired. Because of the loss of some features of the modeling target, the obtained point cloud is sparse and of low quality, and some holes will be generated after meshing; On the one hand, the number of triangles obtained after meshing is huge, which is not practical, and the number of polygons needs to be reduced to improve efficiency.

Mesh topology optimization:

After the mesh is repaired and simplified above, it is necessary to manually re-topologize the modeling target to optimize the structure of the polygons.

UV processing:

The UV coordinate system is the texture mapping coordinate system of the polygonal model, and the 3D model establishes a one-to-one mapping relationship with the 2D image through UV coordinates. For any polygonal model, the UV coordinate information can theoretically have countless solutions. These solutions will not affect the structure of the model itself, but will affect the
Affects the precision and rendering of the final texture map. The goal of UV processing is to establish a coordinate system that is flat and relaxed, easy to identify, scientific and reasonable in layout, and can effectively utilize UV space.

Texture optimization:

The texture is the surface map of the 3D model. Generally speaking, the establishment of the texture is usually after the establishment of the UV coordinates, but for the 3D reconstruction technology based on the image, after the dense point cloud is generated, the texture information is mapped through the image space. can be obtained, but the UV coordinates have not been obtained at this time. The obtained texture information is a spatial texture information that can neither be edited nor mapped onto a flat image. Therefore, this texture needs to be optimized into a traditional texture map for use.

Epic Unreal Engine Metahuman

In early 2021, Epic announced the Metahuman Creator tool. The product is based on a pre-made high-quality face material library, allowing users to quickly generate virtual humans by means of automatic mixing and manual adjustment. Contains full rigs that can be used for animation in Unreal Engine projects.
Features: Efficient template mixing technology, which can quickly get a new face after combining multiple basic faces, with delicate micro-expression animation; cloud rendering, making the production process lightweight, high-quality and convenient.

On June 9, 2022, MetaHuman launched a new version. While still in Early Access, this release not only provides new capabilities for MetaHuman Creator, a cloud-based application that enables anyone to create photorealistic hair and clothing digital humans in minutes, but also Unreal Engine brings an exciting new

feature - Mesh to MetaHuman, and support for the new character rigging, animation and physics features in Unreal Engine 5.

The Mesh to MetaHuman feature will use the textured mesh you created through scanning, sculpting, or traditional modeling tools as a starting point, combined with the body type you selected in the MetaHuman options, and use the automatic landmark tracking in Unreal Engine 5 to template the MetaHuman topology to match the mesh. This template is then submitted to the cloud and matched to the most suitable MetaHuman in our database. Next, this mesh is used to drive the face rig while applying the differences from the original mesh.

After converting a mesh to MetaHuman, you can download it immediately, or open it in MetaHuman Creator, play its animation there, see it come to life immediately, and then continue tweaking and refining it.

While elements like hair and skin textures need to be remade or applied in MetaHuman Creator or your tool of choice, with the Mesh to MetaHuman feature, you can get fully rigged from your own unique Static Mesh in just a few minutes. Definite digital human.

Since Quixel Bridge is now integrated into the Unreal Editor, you can add Metauman to your Unreal Engine 5 project with just a few clicks. This all combined allows us to get more realistic real-time animated characters in less time. [6]

4. Research on Virtual Character Rendering Technology

After the virtual digital character animation is obtained through motion capture, it is necessary to render lights, hair, clothing, etc. through the engine. With the upgrade of hardware and algorithms, the authenticity and real-time performance of rendering have been significantly improved.

For example, several 3D engines currently on the market, such as Unity 3D from Unity Technology and UE 4 from Epic Games, all include PBR (Physically Based Rendering) technology, better simulation of light and increased realism of renderings.

In terms of content production software, Motion Builder is one of the important real-time 3D animation production software. It provides users with real-time animation tools and innovative HIK (Human Inverse Kinetics) character technology, which is convenient for users to quickly create various 3D virtual characters. Based on the real-time 3D engine, the rendering speed is greatly optimized. The software also provides support for motion capture devices. [7]

Computer graphics rendering styles are mainly divided into two types: Physically Based Rendering (PBR) and Non-Photorealistic Rendering (Non-Photorealistic Rendering).

4.1 Non-Photorealistic Rendering (NPR)

In NPR, due to the prevalence of two-dimensional wind in the mobile game market, Japanese cartoon style has become more common in recent years. Two-dimensional cartoon rendering is a kind of NPR and a branch of stylized rendering. When setting up a project to make a game, first of all, it is necessary to determine the overall picture style, whether it is more realistic or more stylized, and if it is more stylized, what is the degree of stylization. There are many types of stylized rendering, commonly found in: ink style, pencil style, cartoon style, cartoon style can be divided into American animation style and Japanese two-dimensional style, what we usually call two-dimensional cartoon rendering refers to this Japanese rendering style.

The American cartoon style is relatively continuous in color, with gradient colors. The coloring style depends largely on the tones defined by the artist, while exaggeration and deformation are often used in the shadows and highlights. The Japanese cartoon style often has more realistic character shapes. But in terms of coloring, it tends to be large blocks of solid color, and some borders of light and dark. Although such

classifications are not clearly demarcated, they are easy to describe.

Stroke: Graphical Definition of Outlines

There are different stroke implementation methods for different rendering styles. Graphical definition of stroke in 2D cartoon rendering. In general, the stroke algorithm can be divided into four categories according to the classification of "Real-Time Rendering":

The first category is Shading Normal Contour Edges, the second category is Procedural Geometry Silhouetting, the third category is Edge Detection by Image Processing, and the fourth category is Geometric Contour Edge Detection.

Shading Normal Contour Edges

Principle: For objects rendered by the camera, the edge normals are always perpendicular or approximately perpendicular to the camera's viewing angle. Using the angle between the model normal vector and the viewing vector, the closer the angle is to the vertical, the closer it is to the contour line. At the same time, we can also control the thickness of the stroke through a threshold parameter.

The principle of this stroke algorithm is very similar to the implementation of edge light. Unlike rim light, which illuminates the edges of objects, this stroke algorithm darkens the edges of objects. A disadvantage of this algorithm is that the thickness of the stroke depends on the curvature of the object's surface. When the curvature change is not uniform at the edges of the object, the thickness of the stroke cannot be controlled.

A more serious problem is that when the rendered object is far from the camera, the normal of the edge of the object may not be close to 90 degrees with the camera's viewing angle due to the pixel accuracy. As a result, the strokes may become thinner or even disappear.

Procedural Geometry Silhouetting

Principle: Double pass rendering, the first pass renders the front of the object, the second pass renders the back of the object, and makes the outline visible. There are several different ways to implement the geometric stroke of the program: Z-Bias method, Triangle fattening method, Triangle Shell method (triangle face back expansion method) and so on.

Advantages: independent processing of all vertices, fast rendering, and controllable line width.

Disadvantages: As shown in Figure 8. Non-shared fixed-point rendering will form gaps, and the model enlargement effect is obvious. There are three solutions, forcing vertices in the same position to have the same normal orientation; creating additional mesh structure at these contours (increasing the density of the mesh), controlling the width of the contour lines according to the distance to the Camera-; z- The method of bias is to draw the back face, but without expanding it, but offset the Z value of the back vertex a little bit forward, so that some parts of the back face are displayed to form a stroke effect.

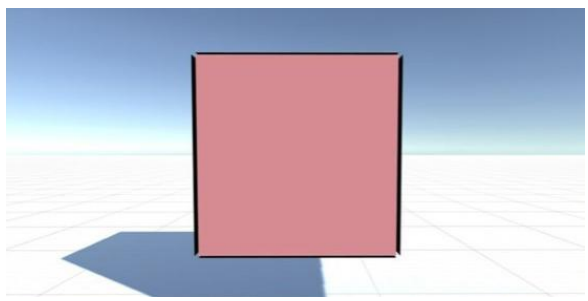

Figure 8.Schematic diagram of contour calculation rendering

The remoteness of the back part may conflict with the original model in depth and occlude the model. There are two solutions, set the Z-offset for the backfaces, so that the contour lines are buried in the adjacent faces; modify the normals of the expansion of the backfaces to flatten the contour lines.

Edge Detection by Image Processing

The edge detection method based on image processing is generally realized by screen post-processing. Usually used in deferred rendering projects, because this method requires depth information, normal information and object ID information in G-Buffer. After obtaining the depth information, normal information and object ID information in the screen space, the edge detection operator is used to detect the difference between the depth value, normal value and object ID value between two adjacent pixels. Situation to find and draw the stroke. The advantage of this type of method is that the strokes have a consistent line width, but the disadvantage is that it requires additional normal and depth information. However, in the popular deferred rendering framework in recent years, normals and depths are already part of G-Buffer, so they often do not require additional information for drawing normals and depths.

Shading

Cel Shading

Cel Shading imitates the traditional Japanese anime celluloid style, and its distinctive features are: the colors are bright and blocky, and the relationship between light and dark is obvious. As shown in Figure 9. Therefore, the basic idea of Cel Shading is to reduce the multi-color level to a low level, reducing the richness of the color level, so as to achieve an effect similar to manual coloring.

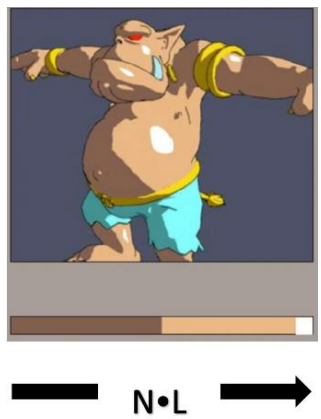

Figure 9.Cel Shading transition diagram

Tone Based Shading

The biggest difference between Tone Based Shading and Cel Shading is that there is no obvious color block distinction, and the color scale obtained by interpolation is continuous.

Figure 10.Tone Based Shading transition diagram

As shown in Figure 10. miHoYo's cartoon rendering style is the effect of using four layers of Tone Based Shading layer by layer. It's worth noting that diffuse ramp maps and specular ramps are a bit like pre-integrated LUT maps for real skin rendering. Guess to use this to show the subsurface scattering effect of cartoon character skin.

4.2 Physically Based Rendering (PBR)

The advancement of PBR (Physically Based Rendering) rendering technology and the emergence of new rendering technologies such as heavy lighting have made digital human skin textures realistic, breaking through the uncanny valley effect. The Uncanny Valley effect was proposed by Japanese robotics expert Masahiro Mori, who believes that people's affinity for robots increases with the increase of its simulation level, but when it reaches a higher critical point, the affinity will suddenly fall into at the bottom, negative psychology such as rejection, fear and confusion arise. The uncanny valley effect of digital human is mainly brought about by the difference between digital human appearance, expressions and actions and real people. The key to the realism of appearance is the realism of skin material. Whether it is plastic or waxy, it will bring discomfort to human beings.

Before the emergence of PBR technology, limited to the development of related software and hardware, all 3D rendering engines focused more on achieving 3D effects, which were not satisfactory in terms of realism. PBR is a collection of rendering technologies based on the simulation of imaging laws in the real physical world. Its key lies in the micro-surface model and energy conservation calculation. By reflecting more realistically the strength of reflected and refracted light on the surface of the model, the rendering effect breaks through plastic feel. Several common 3D engines, such as UnrealEngine 4, CryEngine 3, Unity 3D 5, have their own PBR implementations.

The physical composition of human skin is very complex, and its surface and interior are composed of very complex substances. Real skin contains many details: pores, villi, acne, moles, and oil. The oily layer on the skin surface mainly contributes to the reflective part of skin light (about 6% of the light is reflected), while the epidermis and dermis below the oily layer mainly contribute to the subsurface scattering (about 94% of the light is scattered). Although skin composition is very complex, researchers in the field of graphics rendering use simplified thinking to model skin as several layers.

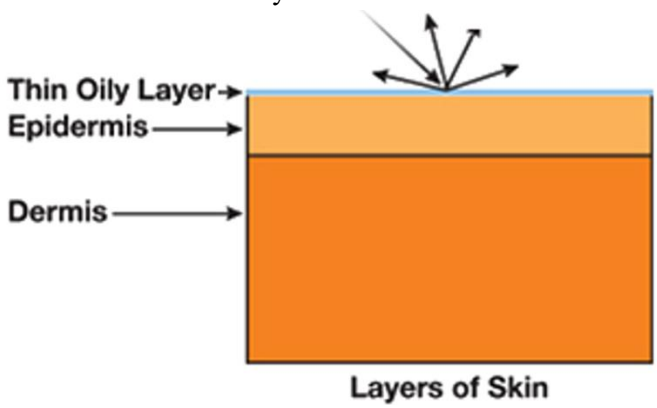

Figure 11.Rendering of human skin

As shown in Figure 11. Thin Oily Layer: Simulates the specular reflection of the skin. Epidermis: Contributing layer for modeling subsurface scattering. Dermis: Contributing layers for modeling subsurface scattering.

There are many methods of skin rendering, UE uses Separable Subsurface Scattering, also called SSSS or

4S. Jorge and others from Blizzard first showed the rendering of SSSS in the speech "Next-Generation Character Rendering" at GDC2013, and formally proposed "Separable Subsurface Scattering" through a paper in 2015. It is approximated by two passes of horizontal and vertical convolution, and the efficiency is further improved. This is the mainstream technology currently used in movies and games.

In the PBR process, the real normal + specular map is used to express the texture of the object under the condition of light irradiation, and the size of the map is increased; the scene performance in the game will be more in line with the physical rules, and the calculation of lighting will be more realistic. The goal is physically based rendering, which is a game-changing technology for visual development today. The main difference between PRB and the traditional workflow is the change in the production process of the texture; the change of the information contained in the texture; the Shader coloring scheme for the texture in line with the PBR workflow.

Breakthroughs in real-time rendering technology help realistic digital humans achieve real-time interaction, and the scope of application is rapidly expanding. Real-time rendering refers to the real-time calculation and output of graphics data, and each frame is an image calculated for the actual ambient light source, camera position and material parameters at that time.

Compared with offline rendering, real-time rendering faces greater challenges. First, the rendering time is short. Real-time rendering needs to render at least 30 frames per second, that is, to complete one frame of image rendering within 33 milliseconds, while offline rendering can take hours or even longer to render one frame of image; Rendering is limited by timeliness requirements, and computing resources generally cannot be adjusted in time, while offline rendering is less limited by timeliness, and more computing resources can be temporarily allocated. Early real-time rendering can only choose highly abstract and simplified rendering algorithms, sacrificing picture quality.

With the improvement of hardware capabilities and breakthroughs in algorithms, the rendering speed, the realism of the rendering effect, and the resolution of the rendered picture have all been greatly improved. In the real-time rendering of virtual characters, it has been possible to mix the real with the fake. In 2016, the real-time-driven virtual characters jointly developed by EpicGames and 3Lateral, Cubic Motion, Ninja Theory and other companies were demonstrated at the Siggraph conference that year, and they successfully rendered high-quality virtual characters in real-time in a consumer-grade hardware environment.

5. Conclusion

In general, in addition to artificial intelligence technology, the core technologies of virtual characters are modeling, action, and rendering. In modeling related technologies, whether it is a two-dimensional 2D art style, a 3D cartoon art style, or a 3D realistic art style or a hyperreal art style, different art styles correspond to different user groups and different application modes. Style has its own characteristics and necessity of existence. Modeling technologies of different art styles will develop together. There should be no so-called pros and cons between virtual character models with different art styles. Different art expressions and different technical production methods are worthy of respect. Worth further study and research.

In motion-related technologies, whether it is optical motion capture or inertial motion capture, corresponding to different product requirements also has its own characteristics and necessity of existence. According to the different needs of the virtual character products produced, choosing the corresponding motion capture equipment or using them in combination can greatly improve the efficiency and quality of product production. At the same time, the simplification of motion capture technology and the control of the cost of related hardware equipment are also the most worthy of research at present when facing users who need to create products related to virtual characters. The popularization of simple motion capture technology and

equipment will also bring huge and good influence and development to the virtual character industry.

In the technical part of rendering, with the advancement of technology, engines such as Unity and UE can render virtual characters with different art styles, and at the same time, they can also adjust the scene, lighting, materials, etc. The most important thing is that engine rendering can achieve almost real-time rendering and present excellent visual effects compared to traditional CG rendering technology. It is believed that with the development of industry-integrated applications such as virtual characters and Metaverse, engine-based real-time rendering plug-ins will make the basic production of related products faster and more convenient.

References

[1] Wang,Y.Q.(2022,May).A Brief Analysis of the Development of Chinese Virtual Idol Industry Empowered by 5G+ Motion Capture Technology——Taking the Virtual Idol Group A-SOUL as an Example. In Journal of Physics: Conference Series (Vol. 2278, No. 1, p. 012011). IOP Publishing.

[2] Kong, R., Qi, Z.,&Zhao, S. (2021, December). Difference Between Virtual Idols and Traditional Entertainment from Technical Perspectives. In 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021) (pp. 344-349). Atlantis Press.
DOI: https://doi.org/10.2991/assehr.k.211209.058

[3] Saida, K., & Foudil, C. (2020). Motions Planning for Virtual Character. Computación y Sistemas, 24(1), 317-324.
DOI: https://doi.org/10.13053/cys-24-1-2932

[4] Ma Wenyi, Jiang Lili & Fan Yanwen. (2022). Design and Implementation of Social Software Based on Live2D+Mlkit. Modern Information Technology (11), 37-40.

[5] Gao Huaqian. (2019). Research and Application of Photo Modeling Technology. Modern Film Technology (02), 42-46.

[6] Pan, Y., Kim, K., Lee, J., Sang, Y., & Cheon, J. (2022). Research on the Application of Digital Human Production Based on Photoscan Realistic Head 3D Scanning and Unreal Engine MetaHuman Technology in the Metaverse. The International Journal of Advanced Smart Convergence, 11(3), 102-118.
DOI: https://doi.org/10.7236/IJASC.2022.11.3.102

[7] Yu Chenping.(2021). Research on the Design and Development of the Virtual Idolization of Game Characters (Master's Thesis, Zhejiang Sci-Tech University).