Research Article

Task Image Setting of 3D Animation Based on Virtual Reality and Artificial Intelligence

Zhiwei Li

Guangzhou Institute of Technology, Guangzhou, Guangdong 510900, China

Correspondence should be addressed to Zhiwei Li; 1440420711@xs.hnit.edu.cn

Received 22 June 2022; Revised 15 July 2022; Accepted 4 August 2022; Published 7 September 2022

Academic Editor: Yajuan Tang

Copyright © 2022 Zhiwei Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to make the expression and action of animated virtual characters more realistic, this paper proposes a virtual character expression and action system based on 3D animation. The system hardware module is used to complete the collection and processing of image data and human bone data. Then, the positioning of human skeleton points and the mapping relationship between joint points and moving skeleton points are constructed, and finally the virtual character model is constructed. On this basis, this paper completes the correspondence of feature points through facial feature point mapping and completes the face model alignment by aligning the video face with the 3D animation virtual face. The Laplace coordinate recovery model is used to reconstruct the facial expression action, complete the simulation of 3D animation virtual character expression action and realize the design of the virtual character expression action system based on 3D animation. The experimental results show that in the aspect of expression movement, the system in this paper has a better effect than the real-time motion capture technology system and the 2D animation expression movement system with a 95.40% simulation degree. The fidelity of skin texture processing in this system is 97.60%. Conclusion. The designed system can effectively simulate the facial expression in the character image and integrate it into the three-dimensional animation to make the virtual character more vivid. After rendering and skin texture processing on the Unity3D platform of the system, the authenticity of the virtual character is enhanced.

1. Introduction

Virtual reality technology is a new scientific research technology with rapid development in the 21st century. It is based on electronic information technology, completes the integration of virtual reality, establishes a virtual world with simulation and displays it through the three-dimensional model. It is the projection of the real world. Virtual reality technology has the characteristics of immersion, interaction, multiple cognition, and subjectivity [1]. Its important technologies include the establishment of a dynamic solid model, the transformation of three-dimensional graphics, and the transmission of information content. Today, film and television culture are widely used. Many industries, such as culture, education, aerospace, and medicine, have produced excellent practical results and promoted the rapid development of the above industries.

The technicality of 3D animation and virtual reality are two completely different definitions. In order to better create 3D animation, we also need to master the technicality of 3D animation and virtual reality technology [2]. 3D animation is mainly based on static photos. It is necessary to change the general animation into 3D pictures to play. However, due to unilateral transmission, it still cannot give people a very obvious sense of reality and presence. The essence of virtual reality technology is to capture the dynamic information of real people, objects, and environment and then use software to calculate to produce more detailed pictures. Virtual reality technology is based on the database query to develop the relationship between the information content of various factors in the 3D scene, such as matching the position of the sun with the sunlight, the light and shadow of the object will change with the data information of the sunlight part, and the data information of the light and shadow part will also...
change [3]. The figure of the object in the interface will move with the sun, and there is no need to show the relationship between light and shadow through complex description. They are closely related and have the same key significance and application value in 3D animation.

On the basis of this research, this paper proposes a task image setting for 3D animation works based on virtual reality artificial intelligence. The system can not only use Maya software for character modeling but also render the 3D animation environment to make it more realistic; at the same time, it can also simulate the facial expressions in the face images and reconstruct the simulated facial expressions in the face model. The facial expressions and actions of the three-dimensional animation virtual characters are more diverse, and the effect is more realistic.

2. Literature Review

The combination of virtual reality technology and animation has an important practical significance. At present, virtual reality technology is widely used in tourism, medical treatment, education, and other industries. However, the development of virtual reality technology is not mature, and people’s familiarity and technological development cannot meet the needs of these fields. Entertainment is the most suitable field if you want to quickly promote virtual reality technology and get rapid development. VR animation, with its new display mode and 360° panoramic immersive experience, gives the audience a new perspective to observe the virtual world of animation, obtains rich sensory information, and makes the audience easily attractive, moving, and charming dreamlike art world [4, 5]. The application of virtual reality technology makes animation come out of the traditional animation form, changes the audio-visual language of animation, makes the audience participate in animation, and makes the presentation of animation art more diversified.

Yang and Song comprehensively introduced the concept, characteristics, short development history, three-dimensional interactive tools, and interactive technology based on natural skills of virtual reality technology [4]. Umeda and others analyzed the performance technology of virtual reality and then elaborated on the significance of studying virtual reality for vision, hearing, force/touch, and smell/taste [6]. Research on various forms of interaction will bring changes to the way of human-computer interaction now, even in the future. Zhang and Chen regarded virtual reality as a core relationship between humans and images and analyzed the interaction and symbiosis between humans and images from a macro perspective [7]. Taking “immersion” and “illusion” as clues, this paper traces back the historical space of virtual reality fantasy from the frescoes in mystery villa and discusses the relationship between virtual art and viewer’s perception and consciousness.

In order to better express the emotional changes of virtual characters in 3D animation, this paper designs a virtual character expression and action system based on 3D animation. On the basis of basic hardware, the system realizes the design of the 3D animation virtual character expression and action system through facial feature point mapping of character image, alignment of video face and virtual face, and facial expression and action reconstruction in Laplace coordinates.

3. Research Methods

3.1. Design of the 3D Animation Virtual Character Expression and Action Simulation System

In order to make the expression and action of 3D animation virtual characters more natural and lifelike, it is necessary to complete the design of the 3D animation virtual character expression and action simulation system under a certain system development environment [8]. The system development environment required here is as follows: a Socket-AM2+ processor with 16 GB memory and an NVIDIA GeForce GT 520 m graphics card, equipped with Kinect for windows, USB interface, and power supply device. Windows 7 is selected as the operating system, Unity3D 5.5.2 is used as the graphics engine, and the system design is completed in the visual studio 2014 development environment through C language. Under the above system development environment, a 3D animation virtual character expression action system is designed, as shown in Figure 1.

In Figure 1, the system is composed of virtual character model building module, data acquisition module, bone data acquisition and an update module, and model joint point binding module. Firstly, the construction of a virtual scene should be completed in the Unity3D platform, and then the collection of human expression and action information should be completed through the Kinect device [9, 10]. On this basis, the virtual character model in 3D animation should be manipulated to complete the simulation of expression and action.

1. Data acquisition and a processing module. Kinect equipment includes infrared projector, RGB camera, infrared camera, tilt control motor, and other components. The speckle data is calculated by the infrared camera, and the depth image of the visible range is established. In order for Kinect to collect the image information from the best angle of view, it can be realized by controlling the angle of equipment through the angle control motor. The data stream collected by Kinect is transmitted through a USB cable. In order to meet the high-power requirements of Kinect, an independent power supply is required. In order to collect RGB data streams, color cameras are usually used to collect RGB data streams. The BCM43694 chip produced by a company is used to process the data stream collected by Kinect, and the depth image and RGB image information can be output through USB. The BCM43694 chip not only reduces the data computation of the host but also reduces the configuration requirements of Kinect for the host.

2. Bone data acquisition and an update module. The collection of human bone data information is mainly completed through Kinect equipment. Kinect
equipment first recognizes the human contour and then identifies the corresponding part of the human body through the human contour [11]. Based on the recognition of this part of the human body, it completes the positioning of human bone points and the updating of bone points.

(3) Model joint point binding module. Through the bone data acquisition and an update module, the position information of human bone points in three-dimensional coordinates can be obtained. In order to control the animation virtual character model through Kinect data, the mapping relationship between the model joint points and the motion control nodes need to be built [12]; then the model established by Maya software is imported into Unity3D, and 20 mapping relationships between model joint points and Kinect bone points are completed to realize the binding of model joint points.

(4) Virtual character model building module. It is mainly divided into several main parts: the establishment of a character model, the construction of a human skeleton, the calculation of skin binding weight, and map rendering. The character model is called the human patch model, which is mainly composed of polygon mesh. The design of mesh should conform to the human structure, and the setting of vertices should follow the principle of setting more vertices at joints. In the process of character modeling with Maya software, the animation module, dynamics tools, and rendering tools in the software are mainly used. First, the initial pose of the model is set, and a set of a virtual character skeleton structure is designed with Maya software to complete the establishment of the model skeleton and realize the construction of the virtual character model.

3.2. Extracting Facial Expression Features. The extraction of facial expression features is a prerequisite for the realization of facial expression migration algorithm. Only when the corresponding facial expression features are extracted, the feature points of virtual animation facial expressions can be marked to establish the corresponding three-dimensional model. There are many algorithms that can realize face feature extraction, among which AAM (active shape model) and ASM (active shape model) are the most classical and widely used [13]. The active appearance model (AAM) is an image segmentation algorithm based on the active appearance model. The method can be divided into two parts. The first part is the training part, which is to construct an active appearance model of a thing. This part requires a training set, its function is to let the program remember the shape features and appearance features of the image to be cut; the second part is the image segmentation stage, this part is to find the features of the thing collected in the training set in the image to be segmented and to find the contours and appearances of object images that are similar to the training set, and then segment them from the entire image. AAM establishes the corresponding data model by analyzing the texture data and shape data of the human face. The specific algorithm steps are shown in Figure 2.

The main disadvantage of AAM is that the selection of initial values will have a great impact on the final results [14]. If the initial value is improperly selected, the number of iterations will increase, and it is easy to lead to local extremum. ASM realizes the migration and construction
of facial expressions through the training of sample sets and corresponding search. The schematic diagram of the creation of corresponding local features is shown in Figure 3.

3.3. Research on the Expression Algorithm of 3D Animation Virtual Characters. Figure 4 shows the design flow of the expression and action system of 3D animation virtual characters [15].

It can be seen from Figure 2 that the three-dimensional model design and the expression action capture system are the two main parts of the production of three-dimensional animation virtual characters and scenes. After the production of three-dimensional animation virtual characters and scenes, they are sent to the three-dimensional animation virtual character expression action platform, which sends them to local/network animation virtual users through display output [16]. The process of 3D animation virtual character and scene production includes feature point mapping, face model alignment, and facial expression animation, which jointly promote the completion of 3D animation virtual character expression action. Figure 5 shows the structure diagram of virtual character expression action implementation.

3.3.1. Mapping Based on Facial Feature Points. Facial expressions and movements, such as opening mouth, closing eyes, and smiling, can be called flexible movements. Five key points are selected on the face of the character, and these points are regarded as the selected feature points of the head posture change. Set the key point coordinates of the human face in the first frame image as $w_0 = (a^0, b^0, c^0, 1)$, then the key point coordinates of the human face in the nth frame can be expressed as $w_n = (a^n, b^n, c^n, 1)$, so the formula can be integrated into the following formula [17]:

$$w_n^s = (w_0^s \cdot Q),$$  \hspace{1cm} (1)

where $Q$ is used to describe the relationship between the five key points on the face, so the pure expression action data can be expressed by the following formula:

$$w_n^e = (w_n^s \cdot Q^{-1}) - w_0^s.$$  \hspace{1cm} (2)

At this time, the posture data and pure expression action data are separated first, and then the pure expression action data are mapped on the 3D animation model. Therefore, the mapping function formula to be established can be expressed as the following formula:

$$w_n^e = f(w_n^e).$$  \hspace{1cm} (3)

Equation (3) represents the pure expression action data of the 3D animation model at the frame.

3.3.2. Alignment of the Video Face and the 3D Animation Virtual Face. The video face model mainly obtains the two-dimensional image of the face through the Kinect device and adds the Candide-3 grid topology model based on the two-dimensional image to complete the construction of the video face model. The 3D animation virtual face model mainly integrates FaceGen and OpenGL to complete model
In terms of model construction, the two models are not built in the same way, so there are differences in the size and proportion of these models. Before completing the 3D virtual face animation, the video face must be aligned with the 3D animated virtual face [18]. The alignment process can be described as follows: firstly, the external contour of the video face and the 3D animated virtual face are scaled and aligned; then the 3D animation virtual face is deformed to promote the 3D animation virtual face to be more similar in shape; finally, the facial features and eyebrows of 3D animation face are aligned to make it more realistic.

3.3.3. Facial Expression Motion Reconstruction Based on Laplace Coordinates. The facial features and the facial contour of the 3D animation virtual face model need to be realized by deformation, but the restoration of facial expression action in the initial frame needs to be completed by facial expression action reconstruction. Here, the Laplacian coordinate restoration model is used to reconstruct facial expression, but the translation invariance of differential coordinates will make it impossible for Laplacian coordinate geometry to obtain Cartesian coordinates through reverse calculation [19]. Therefore, the matrix coordinate transformation matrices $M$ and $M'$ should be regarded as singular.
matrices, and the Laplace operator is represented by the weight \( p_{ij} \), and the new vertex \( v_j \) is obtained to promote the vector \( k \) to transform the mesh. From the Laplace operator of the new vertex, the rank of the \( M \) matrix is \( \text{rank}(M) = (n - u) \), and \( u \) represents the number of meshes. Then the Cartesian coordinate formula can be expressed as the following formula:

\[
R' \mathbf{M}(v_j) = \sum_{j \in \mathbf{N}(i)} p_{ij}(v_i - v_j)
\]

\[
= \sum_{j \in \mathbf{N}(i)} p_{ij}[(v_i + k) - (v_j + k)]
\]

\[
= \sum_{j \in \mathbf{N}(i)} p_{ij}(v_i - v_j)
\]

\[
= \mathbf{M}(v_j).
\]

However, the 3D animation virtual face model has certain complexity, so it is necessary to set multiple vertex coordinates to make the face model more realistic. In order to obtain the unique solution of the Cartesian coordinates of the face model [20], set a 3D animation virtual face model \( O \) with connectivity, and set the grid vertex number and the index value to constrain geometry \( F(1, 2, \ldots, u) \), and \( v_j \in f_j, j \in F \); then the constraint condition of the unique solution of the Cartesian coordinates can be expressed as the following formula:

\[
\left( \frac{\mathbf{M}}{p_{Iu\alpha}^0} \right) x = \begin{pmatrix} \alpha^{(x)} \\ p_{C^u} \end{pmatrix},
\]

where \( I \) represents the identity matrix. The position constraint points are usually completed by the least square method. Assuming \( \mathbf{M} = (\mathbf{M}/p_{Iu\alpha}^0) x \), the unknown formula can be obtained as the following formula:

\[
\hat{x} = \arg \min_x \left( \| Mx - \alpha^{(x)} \|^2 + \sum_{j \in \mathbf{N}(i)} v_j^2 x_i - c_j^2 \right).
\]

Assume that the weight \( P \) is 1, so the constraint factors of each vertex are consistent, and the change of weight will not change the animation effect. The unique solution of the linear equation is obtained by the least square method [21]. The least square method assumes that the \( S \) matrix is \( (n + u) \times u \), and describes the matrix with the following formula:

\[
\hat{x} = \left( \mathbf{M}^T \mathbf{M} \right)^{-1} \mathbf{M}^T g,
\]

where \( g = (\beta, \omega_1, \cdots, \omega_m)^T \). Then the linear equation obtained by integrating (7) is as follows:

\[
\left( \mathbf{M}^T \mathbf{M} \right) x = \mathbf{M}^T \beta.
\]

Matrix \( \mathbf{M} \) can be regarded as a scarf matrix. Since \( S = \mathbf{M}^T \mathbf{M} \), formula (9) can be obtained after decomposing matrix \( S \) as the following:

\[
S = \mathbf{R}^T \mathbf{R}.
\]

According to the formula, the matrix \( \mathbf{R} \) can be obtained by decomposing the \( S \) matrix. In the decomposition process, the values of unknown components \( x, y, \) and \( z \) in the network model can be obtained. Therefore, the formula can be expressed as the following formula:

\[
\mathbf{R}^T \delta = \mathbf{M} \alpha^{(x)},
\]

\[
\mathbf{R} x = \delta.
\]

In the whole process of facial expression action reconstruction, the complex operation is simplified, the Laplace coordinates are distributed in the 3D animation virtual face model by the least square method to avoid local distortion, and multiple constraints are set in the calculation of Cartesian coordinates, which will enhance the authenticity of the 3D animation virtual face model.

4. Result Analysis

Taking the human expression and action image database provided by a City Science and Technology Co., Ltd. as the experimental object, the images in the image database are 10000 human expression and action images taken by 2000 people, and multiple groups of facial expression and action images are randomly selected from the image database [22].

4.1. Simulation and Analysis of Expression and Action of 3D Animation Virtual Characters. Using the system designed in this paper, the 3D animation virtual character expression and action simulation is completed for the image.

Through this system, we can effectively simulate the facial expression and action in the character image and integrate the simulated expression and action into the three-dimensional animation to complete the expression and action addition of the virtual character, making the virtual character more vivid.

4.2. Simulation Performance Comparison. In order to verify the practicability of this system, the real-time motion capture technology system, the 2D animation expression action system, and this system are used to complete the expression action simulation of several groups of images in the image library, respectively, and the simulation results are compared. Although the real-time motion capture system can well simulate the rigid motion of the head, the facial expression is not effectively simulated. The 2D animation expression action system has basically completed the simulation of facial expression and action, but the virtual characters are not realistic enough because the texture and background decoration are not completed in the Unity3D platform. The three-dimensional animation virtual characters’ expressions and actions simulated by the system in this paper are better than those simulated by the other two systems. They not only effectively express the expression changes of human face but also make the virtual characters more realistic through the rendering of the Unity3D
platform and the processing of skin texture. The results of expression motion simulation and skin texture simulation fidelity of the three systems are shown in Figure 6.

It can be seen from Figure 6 that in terms of facial expressions, the system in this paper has a better effect than the real-time motion capture technology system and the 2D animation expression action system with a simulation degree of 95.40%; the fidelity of the skin texture processing of the system in this paper is 97.60%. The designed system can effectively simulate the facial expressions and actions of the human face in the image and integrate it into the three-dimensional animation to make the virtual character more vivid; after rendering and skin texture processing on the unity3d platform, the authenticity of the virtual character is enhanced.

5. Conclusion

In view of the shortcomings of stiff expression and inflexible action in the traditional 3D animation character design system based on evolutionary computation, this paper designs a virtual character expression and action system based on 3D animation. The system can not only use Maya software for character modeling but also use animation modules, dynamics tools, and rendering tools in the software to render the 3D animation environment and make it more realistic; at the same time, it can also simulate the expression action in the face image, and reconstruct the simulated facial expression action in the face model, so that the expression action of the 3D animation virtual character is more diversified and the effect is more real.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] C. Mousas, D. Kao, A. Koilias, and B. Rekabdar, “Evaluating virtual reality locomotion interfaces on collision avoidance task with a virtual character,” The Visual Computer, vol. 37, pp. 1–17, 2021.
[2] A. A. Chubarov, D. V. Tikhomirova, A. V. Shursova, N. O. Veselov, and A. V. Samsonovich, “Virtual listener: a Turing-like test for behavioral believability,” Procedia Computer Science, vol. 169, pp. 892–899, 2020.
[3] N. O. Veselov, A. A. Chubarov, M. N. Roshchin, L. M. Marder, S. M. Segal, and A. V. Samsonovich, “Emotional bica for non-player characters: new empirical data,” Procedia Computer Science, vol. 169, pp. 412–422, 2020.
[4] Y. Jing and Y. Song, “Application of 3d reality technology combined with cad in animation modeling design,” Computer-Aided Design and Applications, vol. 18, no. S3, pp. 164–175, 2020.
[5] M. Liao, B. Song, S. Long, M. He, C. Yao, and X. Bai, “Synthtext3d: synthesizing scene text images from 3d virtual worlds,” Science China Information Sciences, vol. 63, no. 2, Article ID 120105, 2020.
[6] K. Umeda, K. Kobayashi, T. Minato, and H. Yamada, “Atomic-scale three-dimensional local solvation structures of ionic liquids,” The Journal of Physical Chemistry Letters, vol. 11, no. 4, pp. 1343–1348, 2020.
[7] S. Zhang and F. Chen, “Enhancing data-driven algorithms for human pose estimation and action recognition through simulation,” IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 9, pp. 3990–3999, 2020.
[8] Y. Zhi, C. Zhou, T. Li, S. Liu, and Y. Jin, “Action unit analysis enhanced facial expression recognition by deep neural network evolution,” Neurocomputing, vol. 425, pp. 135–148, 2021.
[9] D. Ludl, T. Guilde, and C. Curio, “Enhancing data-driven algorithms for human pose estimation and action recognition through simulation,” IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 9, pp. 3990–3999, 2020.
[10] H. Kim, J. Ben-Othman, L. Mokdad, and K. Lim, “CONTVERB: continuous virtual emotion recognition using replaceable barriers for intelligent emotion-based iot services and applications,” IEEE Network, vol. 34, no. 99, pp. 1–7, 2020.
[11] S. Cao, Y. Yao, and G. An, "E2-capsule neural networks for facial expression recognition using au-aware attention," *IET Image Processing*, vol. 14, no. 11, pp. 2417–2424, 2020.

[12] M. V. Swati and G. S. Baghel, "Systematic design study of the elliptical sheet beam electron gun for terahertz TWT applications," *IEEE Transactions on Plasma Science*, vol. 48, no. 11, pp. 3854–3861, 2020.

[13] A. A. Abdulhameed and A. M. I. Said, "Systematic design of short-span segmental beams reinforced by CFRP plates," *Key Engineering Materials*, vol. 857, pp. 130–137, 2020.

[14] D. Cocchi, L. Raimondi, T. M. Brugo, and A. Zucchelli, "A systematic material-oriented design approach for lightweight components and the CFRP motor wheel case study," *International Journal of Advanced Manufacturing Technology*, vol. 109, no. 7-8, pp. 2133–2153, 2020.

[15] X. Zhan, Z. H. Mu, R. Kumar, and M. Shabaz, "Research on speed sensor fusion of urban rail transit train speed ranging based on deep learning," *Nonlinear Engineering*, vol. 10, no. 1, pp. 363–373, 2021.

[16] H. Li, H. Yang, and K. Chen, "Feature point extraction and tracking based on a local adaptive threshold," *IEEE Access*, vol. 8, Article ID 44325, 2020.

[17] X. Jiang, J. Yu, and J. Jiang, "Research on feature point generation and matching method optimization in image matching algorithm," *Wireless Networks*, vol. 2021, pp. 1–14, 2021.

[18] X. Lu and Y. Zhang, "Human body flexibility fitness test based on image edge detection and feature point extraction," *Soft Computing*, vol. 24, no. 12, pp. 8673–8683, 2020.

[19] N. Yuvaraj, K. Srishti, G. Dhiman et al., "Nature-inspired-based approach for automated cyberbullying classification on multimedia social networking," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–12, Article ID 6644652, 2021.

[20] P. Ajay, B. Nagaraj, B. M. Pillai, J. Suthakorn, and M. Bradha, "Intelligent ecofriendly transport management system based on iot in urban areas," *Environment, Development and Sustainability*, vol. 2022, no. 3, pp. 1–8, 2022.

[21] J. Liu, X. Liu, J. Chen, X. Li, T. Ma, and F. Zhong, "Investigation of ZrMnFe/sepiolite catalysts on toluene degradation in a one-stage plasma-catalysis system," *Catalysts*, vol. 11, no. 7, 828 pages, 2021.

[22] R. Huang, P. Yan, and X. Yang, "Knowledge map visualization of technology hotspots and development trends in China’s textile manufacturing industry," *IET Collaborative Intelligent Manufacturing*, vol. 3, no. 3, pp. 243–251, 2021.