www.cad-journal.net Computer-Aided Design AND APPLICATIONS

Animation Design Based on Anatomically Constrained Neural Networks

Zhengrong Wang[1] and Xipeng Ruan[2]

[1]College of Humanities, Xiamen Huaxia University, Xiamen, Fujian 361000, China, wangzr@hxxy.edu.cn
[2]Department of Sewage, Xiamen Water Affairs Center, Xiamen, Fujian 361000, China, xpruan2024@163.com

Corresponding author: Zhengrong Wang, wangzr@hxxy.edu.cn

Abstract. Computer-aided design (CAD) technology, with outstanding modelling and editing capabilities, has shone in multiple fields, such as modern industrial design and architectural design. This study focuses on the combination of CAD and physics engines and their application in animation design and deeply analyzes the practical effects of this combination in teaching visualization. This article improves the physical effects of animation by paying attention to the intelligent design of convolutional neural networks. By intelligently modelling the physics engine of animated works. The introduction of the internet not only significantly improves the overall efficiency of animation design but also allows animation works to reach new heights in visual effects and dynamic sensation. The characters in the animation are more vivid, the scenes are more delicate, and the entire work seems to have life, which can deeply attract the audience's attention. At the level of teaching application, this innovative method also demonstrates enormous potential. Its intuitiveness and interactivity greatly stimulate students' enthusiasm for learning, helping them better understand the core concepts and practical skills of animation design.

Keywords: CAD; Physics Engine; Animation Design; Visual Teaching
DOI: https://doi.org/10.14733/cadaps.2024.S27.274-287

1 INTRODUCTION

The field of character animation has long faced a challenge: how to cleverly transfer a motion style from one animation segment to another clip while ensuring that the motion content of the latter is preserved. The general methods have obvious limitations, as they can only handle style transitions seen during training and are helpless for novel sports styles. To address this challenge, we propose a revolutionary data-driven framework for achieving flexible transfer of sports styles. The proposed framework can cleverly bypass the tedious 3D reconstruction process, directly extract motion styles from videos, and cleverly apply them to 3D input motion [1]. The content code is decoded into output motion through a series of time convolutional layers, while the style code fine-tunes the depth features through time invariant adaptive instance normalization (AdaIN) to give the motion a unique style. Paired motion data is used as the training basis, meaning that movements with the same content need to be executed in different styles. The uniqueness of this framework lies in its ability to learn from unpaired sets of movements with style labels, thus being able to cope with unobserved movement styles during training. This means that we can not only process 3D data but also directly extract styles from videos, greatly expanding the sources and application scope of styles. However, as the complexity of the system increases, especially in multidimensional (such as 2-D and 3-D) environments, the limitations of this method become increasingly prominent. For many engineering students, understanding complex physical phenomena and system operations solely through text and static images is far from enough. This greatly affects the learning experience of engineering students, and may even lead to their inability to effectively master this knowledge. To address this deficiency, Asef and Kalyvas [2] delved into the potential benefits of animation in engineering education. Through animation, students can have a more intuitive view of the internal operation of the system, gain a deeper understanding of the connections between multiple physical equations, and better apply this knowledge to practical engineering. This requires more effort from the academic community in creating animations to visualize subsequent processes and operations. The data shows that after introducing animation teaching methods, the average scores of both modules have significantly improved, while the failure rate has also significantly decreased. With the rapid development of computer hardware technology, its performance has been significantly improved, while the cost has shown a trend of decreasing year by year. In the design and implementation process of virtual reality 3D engines, the frame rate directly affects the user's immersion and experience comfort. Therefore, Bao [3] proposed an innovative rendering queue management method aimed at significantly improving the frame rate of the engine. At the same time, they also conducted in-depth research on key structural techniques in skeletal animation and designed specialized animation controllers to accurately control the playback of animations and interpolation operations of key structures. In the implementation of animation technology, they adopted an object-oriented design method and an emitter regulator particle rendering mode. This method not only optimizes the rendering process but also ensures the smoothness and stability of the image through intelligent resource scheduling.

Through actual running tests, it was found that the frame rate of the system has been significantly improved, and the animation effect has also achieved the expected goals. In modern manufacturing, especially in the field of complex mechanical assembly production, any small error can lead to overall performance degradation or even product failure. The core of this system is an animated robot equipped with two high-precision 2D cameras and an advanced 3D scanner. In order to verify whether the mechanical components meet the design requirements, Ben et al. [4] introduced a 3D CAD model of mechanical assembly as a reference. It has developed an automatic detection system based on computer animation vision. This model accurately describes the shape, size, and positional relationship of each component, providing us with a benchmark for inspection. They developed an inspection solution based on 2D animation image analysis using images captured by 2D cameras. This solution can efficiently identify whether necessary components are missing from the components and whether these components are installed in the correct position. This automatic detection system based on computer animation vision not only improves the accuracy and efficiency of detection but also greatly reduces errors caused by human factors. Facial feature extraction, as the cornerstone of facial analysis, plays a crucial role in many applications. Bodini [5] delved into the latest advances in facial landmark extraction in 2D images and videos, with a particular focus on methods that fully utilize deep learning techniques. Thanks to the rapid development of deep learning technology, the performance of facial landmark extraction methods has also achieved remarkable improvements in field datasets. These datasets not only provide us with rich facial images and video resources but also provide objective standards for evaluating the performance of various methods. Whether it is for subtle facial expression analysis, identity and facial recognition, or even facial animation and 3D facial reconstruction, accurate capture of facial landmarks is an indispensable part. By comparing and analyzing the performance of numerous methods, not only can we gain insight into the advantages and disadvantages of each method, but we can also provide valuable references for future research [6]. By extracting and transforming human pose from input images into abstract pose data structures, we can more accurately capture the action features of characters. In order to improve the accuracy and efficiency of recognition, Ding and Li [7] proposed an innovative solution, which is a real-time 3D pose recognition system based on deep convolutional neural networks. This conversion process not only simplifies data processing but also improves the efficiency of animation generation. It breaks the traditional limitations of monocular 3D pose estimation and greatly improves the feasibility of real-time operation. What's even more exciting is that the system can achieve real-time running speed at an astonishing resolution of 384 frames per second, providing viewers with a smooth and realistic visual experience. The recognition accuracy has been improved by about 4.3%, while the performance has been improved 4-8 times, which fully demonstrates the superiority and practicality of this method. Although current monocular 3D face reconstruction methods can restore fine geometric details, they have some limitations. Other methods are trained on high-quality facial scans and cannot be well extended to wild images. Some methods generate faces that cannot be realistically animated because they do not simulate how wrinkles change with expressions. Feng et al. [8] proposed the first method, which regresses 3D facial shapes and animatable details that are specific to individuals but vary with facial expressions. DECA is learned from field images without paired 3D supervision and achieves state-of-the-art shape reconstruction accuracy on two benchmarks. They introduced a new detail consistency loss that distinguishes individual-specific details from wrinkles related to facial expressions. This kind of unraveling allows us to synthesize realistic, individual-specific wrinkles by controlling facial expression parameters while maintaining invariance to individual details.

Although the combination of CAD and physics engines has obvious advantages in animation design, research on this technology is still in its early stages both domestically and internationally, and its application in the field of education is even rarer. Traditional animation design teaching mainly focuses on the cultivation of theory and hand drawing skills, lacking in-depth teaching and practical aspects of modern animation design technology. Therefore, this study aims to explore animation design methods based on CAD and physics engines and attempt to integrate them into the construction of animation design visualization teaching systems. I hope to introduce these advanced technologies to help students better understand and master the core principles and practical skills of animation design while injecting new vitality and possibilities into the field of animation design education.

This study focuses on exploring the integration of CAD technology and physics engines to revolutionize the animation design process and seeks to integrate these innovative technologies into the animation teaching system. Against the backdrop of the rapid development of artificial intelligence, the field of animation design is gradually expanding its application scope. Especially ACNN technology, which processes a large amount of animation data through deep learning, automatically identifies key features and creates more realistic and vivid animation effects. The introduction of this technology has brought more possibilities to animation design and provided designers with rich, innovative tools. We hope that through this research, students can gain a deeper understanding of the essence of animation design and enhance their practical skills and innovative thinking. Meanwhile, we also hope to inject new vitality into the field of animation design education through this research.

Research highlights:

This study innovatively combines CAD technology with physics engines, providing animation creators with a new creative path and significantly enhancing the realism and dynamic expressiveness of animation.

In our research, we successfully applied ACNN technology to animation design and significantly improved the intelligence level of animation design through the optimization function of ACNN.

The research not only focuses on theoretical exploration but also puts the proposed methods into practice, constructing an innovative visual teaching system for animation design, and providing more diversified means for the teaching of animation design.

2 RELATED WORK

Haouchine et al. [9] have introduced a disruptive interactive image and video editing method that breaks through the limitations of traditional editing methods and provides users with a new creative space in a physically coherent manner. In Calypso, user editing operations are performed directly in 3D space. By using physical simulation, the proxy geometry presented by non-rigid aligned CAD models is fully processed to achieve physics-based operations. After physical simulation processing, these edits are accurately transferred to the target 2D content through the correspondence between shape and image. It introduces an efficient CAD-to-image alignment program that can jointly minimize rigid and non-rigid alignment while maintaining the advanced structure of the input shape. A series of examples were used to demonstrate physics-based editing effects. The physical behaviours in these examples are diverse, and the edited images and videos maintain a high degree of geometric and visual consistency. Hattler and Cheung [10] will V œ RTEX, an abstract animation that originally existed only in two-dimensional planes, endows it with stunning three-dimensional depth and creates an immersive visual experience. Among them, binocular color competition and depth shift are two key technologies which together constitute V œ The core of RTEX. In this unique space, the audience seems to be situated between two worlds, able to feel the depth and infinity of abstract art, as well as touch the reality and concreteness of the real world. By adopting the design of a dual-screen semi-cave automatic virtual environment, a mixed dimension that is neither completely abstract immersion nor pure gallery space has been constructed. By cleverly utilizing these techniques, we have successfully transformed two-dimensional abstract animations into three-dimensional stereoscopic images. In V œ In RTEX, abstraction and concreteness, as well as virtuality and reality, intertwine to form a complex and rich visual world. Enable the audience to feel a strong sense of space and depth while watching.

The Discrete Element Method (DEM), as a numerical tool for studying abstract animation, has long been widely recognized for its excellent performance. In recent years, with the rapid development of abstract animation physics engine technology, its potential in simulating real particles has gradually emerged. The physics engine was originally designed for video games, to simulate complex physical and mechanical processes in the real world to create a realistic gaming experience. Moreover, significant breakthroughs have been made in computational efficiency, gradually expanding it from the entertainment field to the vast field of scientific research. This approach requires extremely high computational resources when simulating a large number of particles, which limits its application in more complex scenarios. Traditional DEM often uses spherical clusters to approximate the true particle shape during the simulation process. He et al. [11] delved into the modelling methods of abstract animation physics engines in simulating real particles. Through carefully designed particle representation methods, physics engines can more accurately capture the physical properties of particles such as shape, size, and mass. The contact model is responsible for simulating the interactions between particles, including collisions, friction, and other processes, making the entire simulation process more realistic. Ho et al. [12] are committed to exploring the application of 6-DoF (six degrees of freedom) highly immersive virtual reality technology in the field of stereo spatial mapping. In order to comprehensively and accurately evaluate the impact of virtual reality technology, they especially invited 111 students from the university's digital media system to participate in a special 3D VR painting experiment. On this platform, students have the opportunity to experience the unparalleled highly immersive experience brought by HTC VIVE. In this experiment, students will use Google Tilt Brush to create and transform their imagination into three-dimensional paintings through virtual reality technology. It focuses on evaluating the impact of individual perceptual spatial ability and how this influence further affects the motivation for 3D software learning. In order to ensure the scientificity and accuracy of the research, they used a 5-component scale based on the ARCS learning motivation model to systematically collect and analyze various data of students during the experimental process. The analysis results showed that the significance level of the selected factors reached $p=0.000<0.05$, which fully proves that our analysis results are effective and reliable.

Developers have long used animation engines to create a virtual world for players to explore. Unity3D is not only renowned in the field of game development, but its powerful visualization capabilities and flexibility also provide us with the possibility of integrating real-world data into the virtual world. To further enhance the authenticity of the virtual world, Laksono and Aditya [13] utilized the Mapbox for Unity plugin. In addition, through Mapbox for Unity based on OpenStreetMap (OSM) data, they also added road and place name layers to the virtual world, making the entire scene more vivid and realistic. By utilizing the first-person view of the scene, users can freely roam every corner of the virtual world, just like in the real world. Users can explore scenes through first-person views of both ground and air, providing an unprecedented immersive experience. In the end, they successfully achieved multi-platform 3D visualization, able to display detailed 3D models in LOD3, and provided an intuitive and user-friendly user interface. At the same time, in order to achieve a more unique flying experience, they also provide drone views, allowing users to overlook the entire scene from the air and experience a different visual feast. This mixed-use of data from different sources not only ensures the granularity of the scene but also effectively controls the amount of data and computational load.

In the vast field of computer graphics, creating lifelike facial animations has always been a fascinating and challenging task. Because geometry-based animation methods often struggle to achieve precise facial deformation, especially when dealing with complex areas such as the mouth and eyes, it is difficult to generate realistic rendering effects. In order to overcome these difficulties, image-based animation technology has emerged. Dynamic texture sequences require seamless connectivity, which is not always achievable in practice and often leads to visual artifacts that affect the overall quality of the animation. Paier et al. [14] proposed an innovative hybrid animation framework that fully utilizes the latest advances in deep learning. This technology greatly enhances the realism of animation by using dynamic textures to capture subtle expressions and small actions that are difficult to interpret in geometry. Intended to provide users with an interactive animation engine that enables them to edit facial expressions simply and intuitively. Parkhmenko et al. [15] are committed to exploring the rich connotations of modern screen culture, particularly in the theoretical understanding of the development of 3D animation Emergent. Among them, the uniqueness of animated character design and its impact on the aesthetic image of the screen works through audio-visual imagery are the top priorities of this article's research. As the soul of animated works, the design of their audio-visual images directly affects the audience's feelings and cognition. They combined excellent animated film cases from both domestic and foreign sources to compare and analyze the types of ensembles in different technologies. These concepts have their own unique meanings in animation character design, and their relationships are particularly unique. The deductive method was used to distinguish the components of artistic images in screen works from the perspective of theme space content. Through in-depth analysis of these concepts, we can more accurately grasp the core elements of animation character design.

3 THE INTEGRATION AND APPLICATION OF ACNN IN ANIMATION DESIGN

The uniqueness of ACNN lies in its introduction of an attention mechanism, which enables the model to adaptively focus on the most critical information, like humans when processing information while ignoring irrelevant or secondary details. These feature information are the essence of image content and the key to subsequent processing. Next is the fully connected layer, which plays a role in filtering and integrating key features. These components work together, making the entire model perform well in handling complex image tasks. To better understand this structure, you can refer to Figure 1 for a visual experience.

In deep learning, convolution operations connect input units with hidden units through local connections to facilitate the processing of specific tasks. This means that each hidden unit is only connected to a subset of the input unit. Due to the three-dimensional properties of convolutional neurons, which include depth information, a series of filters are needed to process depth information. Each filter is specifically designed to train a specific depth level, and the depth of the output unit matches the number of filters.

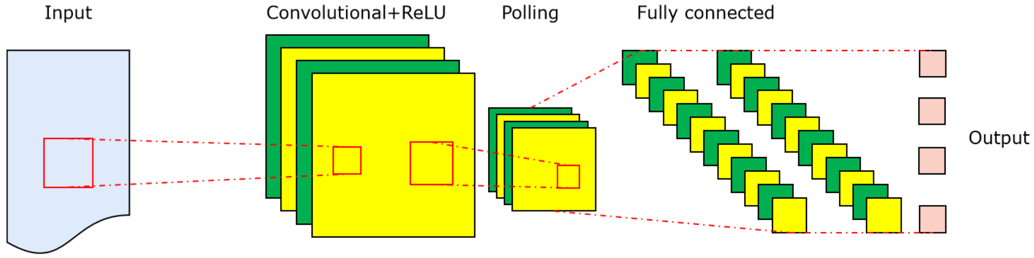

Figure 1: CNN structure.

On the other hand, pooling is a process of spatial or feature fusion and simplification, which plays a key role in reducing model computational complexity, preventing overfitting, and enhancing the robustness of the network to position, size, scaling, and nonlinear deformations. Usually, pooling operations involve downsampling samples, which involves traversing the entire feature map through a fixed-size sliding window, selecting a specific value in each window as the output result, and finally combining these values into a brand-new feature map. Figure 2 shows in detail the operational steps of maximum pooling.

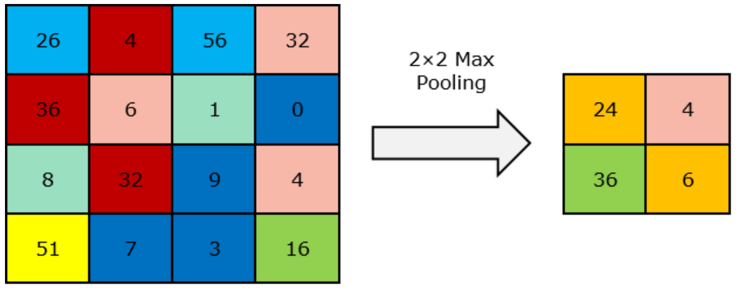

Figure 2: Maximum pooling.

The fully connected layer, also known as the standard neural network hierarchy, is characterized by the fact that each neuron establishes a connection with all neurons in the previous layer. If we mark a fully connected layer as the l th layer and set its weight matrix as W^l with a bias term of b^l, the operation of that layer can be summarized as follows: each neuron receives the output of all neurons in the previous layer, adds a weighted sum to the weight matrix W^l, adds a bias b^l, and then passes it to the activation function for processing, finally obtaining the output result of that layer:

$$Z_j^l = f\ W^l X^{l-1} + b^l \tag{1}$$

Capturing local and global features is crucial in animation design. However, traditional CNNs may be limited in recognizing these features, which can affect accuracy.

In order to enhance the accuracy of feature recognition, ACNN introduces a unique attention module that can intelligently weigh the input features. This attention module has the ability to learn, calculate the importance of each feature, and generate a weight vector based on it. After convolution processing, this weight vector will adjust the importance of each part of the feature map, making the model more focused on features that are beneficial to the task.

The advantages of ACNN lie in several major features:

Adaptability: The attention module can adaptively learn and adjust the level of attention to features based on actual tasks, thereby exhibiting higher flexibility in handling diverse tasks.

Interpretability: By observing the weight vectors generated by the attention module, we can gain insight into the model's level of attention to different features when processing input data, which increases the model's interpretability.

Performance optimization: Thanks to the introduction of the attention mechanism, ACNN performs better in handling complex tasks, especially when dealing with input data containing a large amount of redundant information, its performance is particularly outstanding.

The free fusion and simulation process of the entire animation features adopts the Euler grid method, which includes three stages: data processing, simulation, and rendering. Please refer to the flowchart in Figure 3 for details.

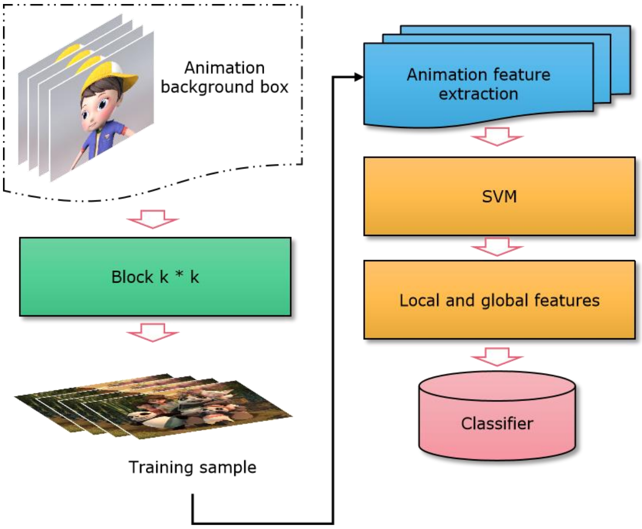

Figure 3: Simulation process.

Firstly, a linear model is used to determine the training of the fusion coefficients of the linear model. During the training process, the following exponential loss function was selected as the optimization objective:

$$L(y,f)=e^{-yf(x)} \tag{2}$$

In this loss function, set $y\in\{-1,1\}$, and f represents the output result of the classifier. By minimizing this loss function, the optimal linear model fusion coefficient can be obtained.

In the process of steganalysis, as steganalysis noise is generated by embedding secret information into the original signal, its intensity is usually much weaker than the information carried by the signal itself. Therefore, if the ReLU activation function is used in the first few layers of the neural network, it may lead to the loss of negative range information, and there may be too much carrier content information mixed in the positive range. In addition, when the ReLU function approaches its saturation point, the gradient of the activation unit approaches zero, which increases the possibility of gradient disappearance.

To address these issues, a linear truncation activation unit TLU is introduced in the first convolutional layer of the neural network:

$$TLU(x)=\begin{cases}-T, & x<-T\\ x, & -T\le x\le T\\ T, & x>T\end{cases} \tag{3}$$

In the field of animation production, the application advantages of ACNN are mainly reflected in the following aspects:

Capture dynamic elements: Due to the large number of dynamic elements in animation, such as character actions and scene changes, ACNN, with its excellent feature extraction function, can accurately capture these elements, providing strong data support for animation creators.

Identifying core action points: When creating an animation, identifying key action nodes is crucial for maintaining the smoothness of the animation. The attention mechanism of ACNN enables it to automatically identify key action frames and features, assisting animators in more accurately grasping the character's action transitions.

Drawing on and combining styles: ACNN can also learn and integrate multiple animation styles, providing animation creators with richer creative tools and inspiration, thereby promoting innovative thinking and artistic inspiration.

In addition, ACNN consists of two main networks: the first is the network used for image segmentation; Next is the 3DResNeXt classification network. For the detailed structure of ACNN, please refer to Figure 4.

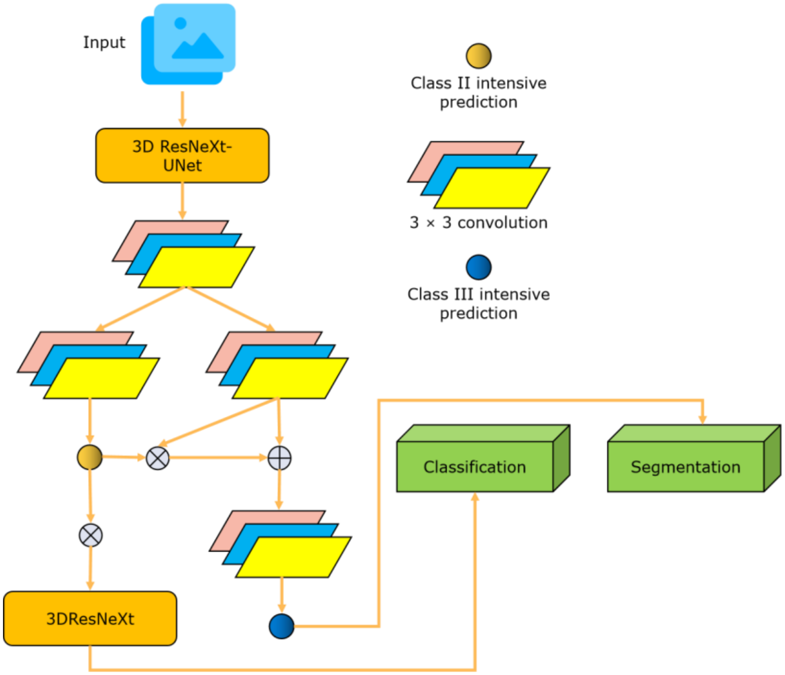

Figure 4: ACNN structure.

When introducing ACNN in animation design, the primary task is to build an ACNN model that is compatible with animation data. This requires converting the original animation data into a format that ACNN can process, including converting animation frames into image sequences and performing normalization operations. In the process of building the ACNN model, basic components such as convolutional layers, pooling layers, and fully connected layers are applied. The convolutional layer is mainly responsible for extracting features from input data, while the pooling layer plays a role in reducing the dimensionality of the data. After the convolutional layer, we embedded an attention module to adjust the weight allocation of the feature map.

The core idea of STFT is to use a fixed-length window function as a truncation tool to sequentially extract small segments of the original time-domain signal. Next, Fourier transform is used to analyze these small segments of signals, in order to obtain local spectral information near specific time points. Its mathematical expression is as follows:

$$STFT_f = t,w = \int_{-\infty}^{\infty} x\ t\ g\ t-\tau\ e^{-jwt}dt \tag{4}$$

The time-domain signal $x\ t$ is processed through the window function $g\ t-\tau$, which centers around τ.

Convolutional and down-sampling operations significantly reduce the complexity of the model and reduce the required parameters. This article chooses the mean down-sampling method, which calculates the average value of pixels within the scanning window and sets this value as the pixel value at the corresponding position of the down-sampling layer feature map. The specific mathematical formula is as follows:

$$y = \overline{w\sum D_{zk}x_i + b_{sub}} \tag{5}$$

During the downsampling process, w, b_{sub} represents the weight and bias term of the downsampling kernel, which acts on the local region D_{zk} of the feature map. In this region D_{zk}, there are multiple neurons whose values are represented by x_i.

Choose ReLU (Linear Rectification Function), which is widely used in deep learning, as the activation function, and its mathematical formula is as follows:

$$f_j = \max\ 0, x_j^l \tag{6}$$

f_j is the output result processed by the activation function.

The attention mechanism is essentially a similarity evaluation method. When the similarity between the input and the target state is high, the weight of the input will also increase accordingly, reflecting that the current output is more dependent on this input. In this article, an attention mechanism is constructed by adding a fully connected layer after global pooling. Assuming there is a set of feature vectors $h_i\ i=1,2,\cdots,m$, in order to integrate the information of these k feature vectors to generate a vector h^*, a weighted average method is adopted, namely:

$$h^* = \sum_{i=1}^{k} a_i h_i \tag{7}$$

The key to the attention mechanism lies in the precise allocation of weight a_i to generate an appropriate a_i.

The fully connected layer serves as a key bridge between feature fusion and classification recognition. It receives inputs composed of each element of the feature information matrix and maps these scattered features to the space marked by the samples as the output of this layer. The dimension of this output is equal to the number of target categories. Subsequently, the Softmax layer processes the output of the fully connected layer, converting the values into probabilities for the corresponding categories. Finally, by selecting the maximum value in the output vector of the Softmax layer, the predicted target category can be determined:

$$p_n = \frac{e^{x_m}}{\sum_n e^{x_n}} \tag{8}$$

Assuming x is the input to the Softmax function, x_m is the m th element in x. After Softmax processing, p_n represents the probability distribution of the element for all categories, where p_n reflects the probability that the predicted result belongs to the GGth category.

SVM is a model suitable for binary classification, which can handle both linear classification problems and non-linear classification scenarios. When dealing with nonlinear inputs, SVM maps these inputs to a high-dimensional feature space and finds the best hyperplane in this high-dimensional space for classification. In this study, the radial basis function (RBF) was used as the kernel function:

$$K(x_i,x_j) = \exp\left(-g\left\|x_i - x_j\right\|^2\right) \tag{9}$$

In the above formula, the parameters of the kernel function are represented by g. In the model training process of this article, the optimal parameters of SVM are found through grid partitioning, where the penalty coefficient is set $C = 6, g = 0.006$.

In order to process animated images, N sampling points were set and the original signal components within M frequency bands were evenly divided into K equal parts in time, resulting in l time-frequency small blocks. In order to further improve accuracy, each time interval is further subdivided into P equal parts. Finally, the energy of all component time-frequency blocks was defined for subsequent analysis and processing:

$$e_{il} = \sum_{m=1}^{p} \left|a_{il}(m)\Delta t\right| \tag{10}$$

In the calculation process, $a_{il}(m)$ represents the energy value corresponding to the m data points of the subband i in a specific time-frequency block l, while Δt representing the length of time occupied by each equal segment.

4 EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Analysis of Experimental Results

The significant reduction in modelling time indicates that in a fast-paced animation production environment, this algorithm can significantly improve work efficiency, enabling animation producers to complete model construction more quickly.

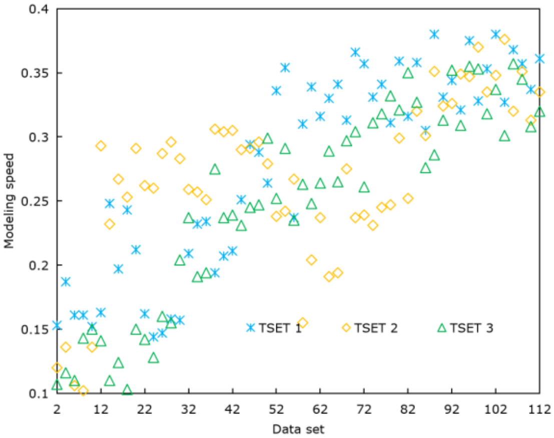

Figure 5: Modeling speed.

Figure 5 shows the modeling speed under different test results. This algorithm has successfully achieved the goal of fast modelling by simplifying the computing process, improving data access and processing methods, and combining cutting-edge technologies such as parallel computing.

In order to comprehensively evaluate rendering speed, multiple animation scenes were carefully selected for testing, each with different complexity and characteristics. When selecting these scenes, consideration was given to the diverse lighting environment, complex shadow effects, and fine texture mapping. Figure 6 vividly illustrates the speed advantage of the real-time rendering algorithm introduced in this article in animation production. Whether dealing with simple or complex scenes, this algorithm has demonstrated excellent rendering speed.

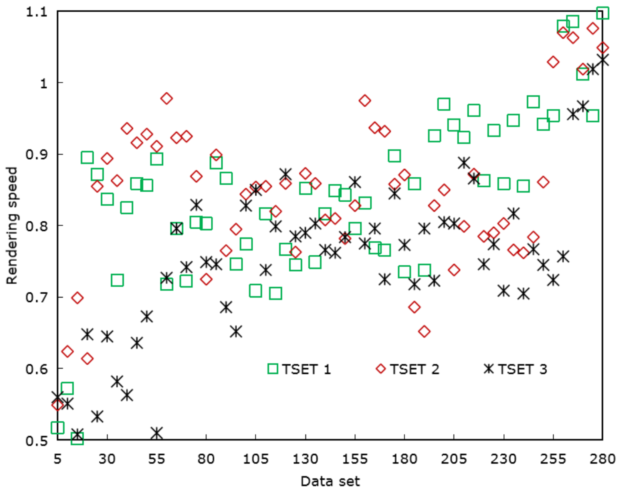

Figure 6: Rendering speed.

The reason why the rendering speed of this algorithm is so fast is mainly because it fully utilizes the capabilities of current computer hardware. Especially by using GPU for parallel computing, it can handle multiple rendering tasks simultaneously, significantly improving overall efficiency. The advantage of GPU parallel computing is that it can efficiently process large amounts of data, which makes it very suitable for computationally intensive tasks such as graphics rendering. Through parallel processing, rendering tasks that originally required sequential execution can now be performed simultaneously, greatly reducing the overall rendering time.

The data shown in Figures 7 and 8 clearly demonstrate the trend of real-time animation production in the key indicators of smoothness and clarity as the number of iterations increases in the animation production method that combines CAD and physics engines. Fluency is one of the important criteria for evaluating the quality of animation, which is related to whether the animation can bring a coherent and natural visual experience to the audience. The data in Figure 7 shows that as the number of iterations increases, the smoothness of the animation also continues to improve.

Cutting-edge rendering algorithms can efficiently calculate the image for each frame, ensuring smooth and unobstructed animation playback without any lag or delay. By utilizing parallel processing technology, our model can handle multiple tasks simultaneously, greatly improving the rendering efficiency of animations and ensuring smooth playback at high frame rates. High frame rate refers to the ability to present more frames per unit of time, which directly enhances the coherence of the animation and brings a smoother visual experience to the audience.

The clarity of animation is crucial for the presentation of details, directly affecting the overall visual experience of the audience. According to the data in Figure 8, it can be clearly seen that as the number of iterations increases, the clarity of the animation also steadily improves. High-end graphics processing technology can more accurately simulate the transmission and reflection of light, thereby creating more lifelike images.

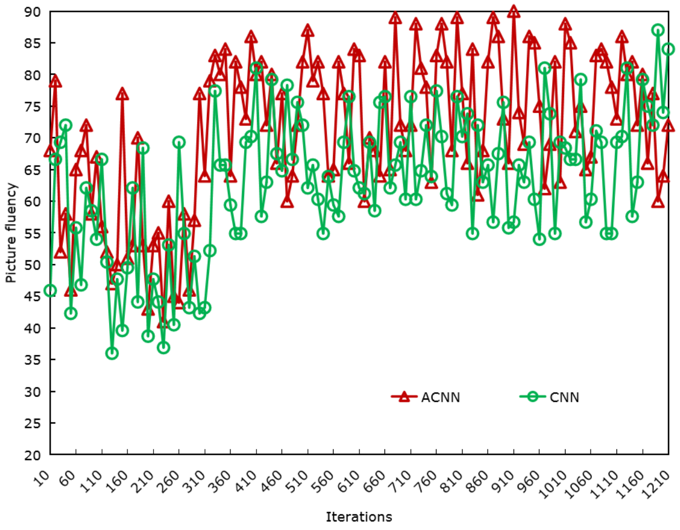

Figure 7: Analysis of animation fluency.

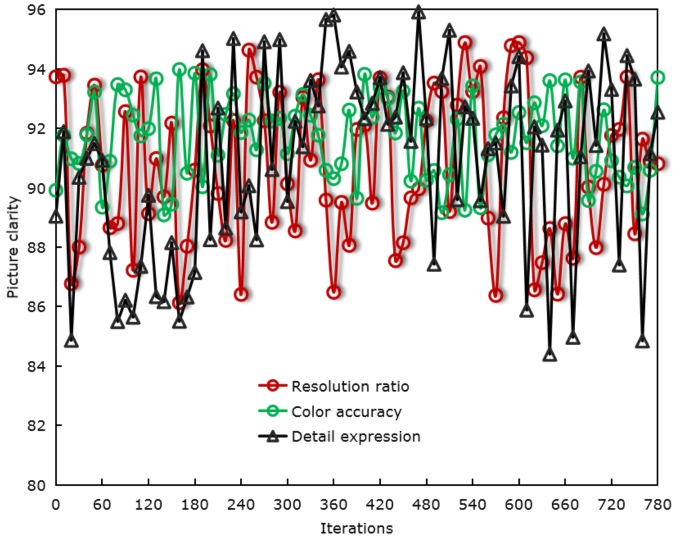

Figure 8: Animation clarity.

The animation design and production mode combines CAD and physics engines, achieving continuous improvement in animation fluency and clarity through efficient rendering algorithms, parallel processing methods, and cutting-edge graphics processing technology. This progress not only enhances the overall quality of animation but also provides viewers with better visual enjoyment.

4.2 DISCUSSION

This study combines the precise modelling ability of CAD technology with the simulation function of physics engines for the real world, providing animation designers with an efficient tool to help them create animation works that combine artistic and physical realism more quickly.

In traditional animation design education, students usually need to master the basic rules and skills of animation through extensive hand-drawing practice and practice. Traditional animation design methods are often cumbersome and time-consuming, and students not only face enormous pressure during the learning process but also find it difficult to ensure the systematic and accurate knowledge they have learned. The visual operation interface of the software allows students to observe their design results in real-time and make timely adjustments and optimizations. Through CAD software, students can easily construct three-dimensional models of animated characters and scenes, achieving precise size adjustments and detailed refinement. In the process of animation design, students need to consider the influence of physical factors such as gravity, friction, and collision to make the animation more realistic.

5 CONCLUSIONS

This study not only delves into the animation design method that combines CAD and physics engines but also verifies its practical application effect in animation creation and visualization teaching through a series of experiments. In the vast world of animation creation, CAD software has become a sharp sword in the hands of designers due to its powerful functions and flexibility. As we further integrate CAD technology with physics engines, the boundaries of animation design are further expanded, bringing infinite possibilities. It can quickly build and modify complex 3D models; whether it is the precise portrayal of characters or the delicate depiction of scenes, CAD can complete them with astonishing efficiency. The combination of the two has brought the visual and dynamic experience of animated works to unprecedented heights. The precise modelling ability of CAD allows designers to construct their own animated world freely. Through an intuitive operating interface and powerful technical support, students can more efficiently master the core skills of animation design while also gaining a deeper understanding of the principles and processes of animation production. The introduction of this method not only brings more powerful creative tools to animation designers but also injects new vitality and energy into the field of animation design teaching. It makes animation design teaching more intuitive, efficient, and scientific, laying a solid foundation for cultivating more outstanding animation design talents.

Zhengrong Wang, https://orcid.org/0009-0009-0501-6128
Xipeng Ruan, https://orcid.org/0009-0007-5088-1730

REFERENCES

[1] Aberman, K.; Weng, Y.; Lischinski, D.; Cohen, O.-D.; Chen, B.: Unpaired motion style transfer from video to animation, ACM Transactions on Graphics (TOG), 39(4), 2020, 64:1-64:12. https://doi.org/10.1145/3386569.3392469
[2] Asef, P.; Kalyvas, C.: Computer-aided teaching using animations for engineering curricula: A case study for automotive engineering modules, IEEE Transactions on Education, 65(2), 2021, 141-149. https://doi.org/10.1109/TE.2021.3100471
[3] Bao, W.: The application of intelligent algorithms in the animation design of 3D graphics engines, International Journal of Gaming and Computer-Mediated Simulations, 2021, 13(2), 1-12. https://doi.org/10.4018/IJGCMS.2021040103
[4] Ben, A.-H.; Jovančević, I.; Orteu, J.-J.; Brèthes, L.: Automatic inspection of aeronautical mechanical assemblies by matching the 3D CAD model and real 2D images, Journal of Imaging, 5(10), 2019, 81. https://doi.org/10.3390/jimaging5100081
[5] Bodini, M.: A review of facial landmark extraction in 2D images and videos using deep learning, Big Data and Cognitive Computing, 3(1), 2019, 14. https://doi.org/10.3390/bdcc3010014
[6] Dey, U.; Cheruvu, S.-K.: A web-based integrated GUI for 3D modeling, kinematic study, and control of robotic manipulators, Computer Applications in Engineering Education, 28(4), 2020, 1028-1040. https://doi.org/10.1002/cae.22282

[7] Ding, W.; Li, W.: High speed and accuracy of animation 3d pose recognition based on an improved deep convolution neural network, Applied Sciences, 13(13), 2023, 7566. https://doi.org/10.3390/app13137566

[8] Feng, Y.; Feng, H.; Black, M.-J.; Bolkart, T.: Learning an animatable detailed 3D face model from in-the-wild images, ACM Transactions on Graphics (ToG), 40(4), 2021, 1-13. https://doi.org/10.1145/3450626.3459936

[9] Haouchine, N.; Roy, F.; Courtecuisse, H.; Nießner, M.; Cotin, S.: Calipso: physics-based image and video editing through CAD model proxies, The Visual Computer, 36(1), 2020, 211-226. https://doi.org/10.1007/s00371-018-1600-0

[10] Hattler, M.; Cheung, T.-M.-C.: Into the Vœrtex: Case study of a stereoscopic abstract animation installation, Animation Practice, Process & Production, 10(1), 2021, 51-71. https://doi.org/10.1386/ap3_000026_1

[11] He, H.; Zheng, J.; Schaefer, V.-R.: Simulating shearing behavior of realistic granular soils using a physics engine, Granular Matter, 23(3), 2021, 56. https://doi.org/10.1007/s10035-021-01122-5

[12] Ho, L.-H.; Sun, H.; Tsai, T.-H.: Research on 3D painting in virtual reality to improve students' motivation of 3D animation learning, Sustainability, 11(6), 2019, 1605. https://doi.org/10.3390/su11061605

[13] Laksono, D.; Aditya, T.: Utilizing a game engine for interactive 3D topographic data visualization, ISPRS International Journal of Geo-Information, 8(8), 2019, 361. https://doi.org/10.3390/ijgi8080361

[14] Paier, W.; Hilsmann, A.; Eisert, P.: Interactive facial animation with deep neural networks, IET Computer Vision, 14(6), 2020, 359-369. https://doi.org/10.1049/iet-cvi.2019.0790

[15] Parkhomenko, Y.-A.; Lugovtsev, A.-Y.: The role of object design in the creation of animation imagery, art analysis experience, Observatory of Culture, 16(4), 2019, 374-385. https://doi.org/10.25281/2072-3156-2019-16-4-374-385