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

Parametric Design Algorithm Guided by Machine Vision in Animation Scene Design

SiHui Yu[1], Yuzhi Chen[2], Tianyi Guo[3] and Dong Yang[4]

[1]College of Humanities, Xiamen Huaxia University, Fujian, Xiamen 361024, China, ysh@hxxy.edu.cn
[2]College of Information and Smart Electromechanical Engineering, Xiamen Huaxia University, Fujian, Xiamen 361024, China, cyz@hxxy.edu.cn
[3]College of Humanities, Xiamen Huaxia University, Fujian, Xiamen 361024, China, gty@hxxy.edu.cn
[4]College of Design and Art, Xiamen University of Technology, Fujian, Xiamen 361024, China, yangdong@xmut.edu.cn

Corresponding author: SiHui Yu, ysh@hxxy.edu.cn

Abstract. Against the backdrop of rapid progress in digital technology, animation has become an indispensable part of people's lives. CAD parametric design is a design method based on parameter constraints, which can quickly modify and adjust the design by adjusting parameter values. This article introduces machine vision and AI technology into animation design, automatically identifying and extracting key elements in animation scenes and generating scene designs that meet the requirements based on design requirements and constraints. By analyzing the simulation results of different models in animation scene generation, it was found that the proposed model demonstrated significant advantages, achieving lower error rates and higher accuracy. This result validates the enormous potential of DL in the field of animation design while also highlighting its effectiveness and application value in practice. Compared with traditional methods, this algorithm can automatically identify and extract key elements in the scene, greatly reducing the workload of manual adjustment and optimization.

Keywords: Animation Scenes; Machine Vision; CAD; Parameterized Design
DOI: https://doi.org/10.14733/cadaps.2024.S15.261-275

1 INTRODUCTION

Animation has become an indispensable part of people's lives, covering many research fields such as animated feature films, animated advertisements, and games. With the continuous development of computer graphics and virtual reality technology, 3D graphics engines have become the core tool for animation production. In animation scene design, intelligent algorithms are becoming increasingly widespread, providing more possibilities for animation designers. Bao [1] explored the application of intelligent algorithms in the design of 3D graphics engine animation scenes. 3D graphics engine animation scene design refers to the process of using computer graphics technology to create dynamic images with high realism. This process includes multiple stages such as modeling, material mapping, lighting processing, and animation design. In animation scene design, the application of intelligent algorithms can help designers improve work efficiency and enhance animation quality. A genetic algorithm is an optimization algorithm based on the principle of biological evolution, which can optimize the data structure of 3D models and improve rendering efficiency. In animation scene design, genetic algorithms can be used to optimize parameters such as the position, orientation, and size of objects in the scene to achieve the best visual effect. A neural network is an algorithm that simulates the working mode of the human nervous system, which can automatically learn and recognize complex patterns. In animation scene design, neural networks can be used to automatically recognize and predict objects and lighting in the scene, achieving more realistic rendering effects. However, the high cost and time consumption of animation production have become the bottleneck restricting its further development. Computer-aided design (CAD) technology, especially parametric design algorithms, gradually shows its great potential in animation production. Sculpture art is an art form with a profound historical heritage and unique expressive power. In the digital age, the continuous advancement of technology has brought new possibilities to sculpture art. Guo and Wang [2] discussed how to apply machine vision-based CAD parameterization methods to the expression techniques of sculpture art space to improve its creative efficiency and expressive power. Machine vision is a technology that obtains and processes image information by simulating the human visual system. It can recognize various features in an image, such as shape, color, texture, etc., and convert them into data that computers can process. The CAD parameterization method is a technique based on parametric modeling, which achieves accurate representation and efficient optimization of models by representing them as parametric equations or constraint equations. Collect image information of sculpture works through machine vision technology and use image processing techniques to preprocess the images, such as denoising and enhancement, to extract feature information of sculpture works. Using CAD parameterization methods, the preprocessed image information is transformed into parameter equations or constraint equations to achieve efficient representation and optimization of sculpture works. During this process, the expression and details of the sculpture can be adjusted by controlling the parameters.

CAD parametric design is a design method based on parameter constraints, which can quickly modify and adjust the design by adjusting parameter values. In animation scene design, its application is still in the primary stage. Digital calligraphy and painting is an innovative form that combines traditional calligraphy and painting art with modern computer technology. It creates and displays calligraphy and painting works through virtual brushes and digital ink. In digital calligraphy and painting, the construction of virtual brush-shaped scenes is one of the key factors in achieving a realistic calligraphy and painting experience. Huang et al. [3] explored the historical progress, current status, and future development trends of visual research on virtual brush modeling in digital calligraphy and painting. Digital calligraphy and painting have gradually become important branches in the field of computer art. Early digital calligraphy and painting work mainly relied on geometric shapes generated by computer programs and textures generated by algorithms. With the continuous development of computer hardware and software technology, digital calligraphy and painting have made significant progress in visual effects and interactivity. At present, great progress has been made in the construction of virtual brush modeling scenes in digital calligraphy and painting. Some mainstream digital calligraphy and painting software already supports the selection of multiple brush tools and can simulate different types of ink and paper effects. Meanwhile, by capturing and tracking user hand movements in real-time, this software can also achieve a more natural interactive experience. This is mainly because of the complexity and artistry of animation scene design, which makes the traditional CAD parametric design algorithm difficult to meet its needs. Machine vision is an important branch of the AI field, which studies how to make computers "see" the world like people. Traditional CAD tools often have certain limitations when dealing with complex 3D models. To address this issue, 3D reality technology (VR) is combined with CAD to enhance the efficiency and effectiveness of animation design. Jing and Song [4] will explore the application of this combination method in animation design. 3D reality technology is a computer technology that can create and experience virtual worlds. By using devices such as virtual reality helmets and controllers, users can immerse themselves in a computer-generated three-dimensional environment. CAD, on the other hand, is a tool used for design and manufacturing, widely used in industries, construction, and other fields. Introducing 3D reality technology into CAD systems can achieve more intuitive and flexible design and manufacturing. In animation design, the use of 3D reality technology can enable designers to operate and adjust models more intuitively in a virtual environment. For example, designers can use virtual reality helmets and handles to rotate, scale, and move virtual models to observe better and adjust the details of the model. Meanwhile, by utilizing the parametric design function of CAD software, designers can easily make parametric modifications and optimizations to the model. With the continuous growth of deep learning (DL) technology, machine vision has made remarkable progress in image recognition and object detection. These technologies make it possible to automatically identify and extract the key elements in the animation scene.

The research topic of this article, namely "CAD parameterized design algorithm guided by machine vision in animation scene design," is proposed in this context. The research objective of this article is to develop a machine vision-based CAD parametric design algorithm that can automatically recognize and extract key elements in animation scenes and achieve rapid modification and adjustment of these elements through parametric design. We hope to address the limitations of traditional CAD parametric design algorithms in animation scene design by introducing machine vision technology, thereby improving rendering efficiency and reducing animation production costs. An animation scene design model based on the above algorithm was constructed in the article, which can automatically generate scene designs that meet the requirements according to design requirements and constraints. This study has important theoretical value in promoting the advancement of animation production technology and also plays a positive role in promoting the sustainable growth of the animation production industry. In web design, animation design is an important component that can enhance the visual effect of web pages and improve the user experience. In web animation design, the dynamic effect of visual communication is one of the key factors affecting the animation effect. Kiu et al. [5] explored the dynamic effects of visual communication in web animation design and its implementation using CAD technology. In web animation design, the dynamic effect of visual communication refers to the dynamic changes and expressions of visual elements achieved through the design and control of animation. This dynamic effect can enhance the visual effect of web pages and improve the user experience. For example, animation design can create smooth transitions, interactive animations, and deformation effects on-page elements, thereby attracting user attention and enhancing their impression of the webpage. CAD technology refers to computer-aided design technology, which has been widely applied in various fields, including web design. In web animation design, CAD technology can achieve precise design and efficient production of animations. For example, using CAD software can accurately control the frame rate, timeline, and motion curves of animations, thereby achieving precise control and optimization of animations.

This article mainly studies the application of machine vision and CAD in animation design, which includes the following innovations:

(a) Traditional CAD parametric design algorithms mainly rely on manual input and adjustment of parameters, while this article introduces machine vision and AI technology into animation scene design to automatically recognize and extract key elements.

(b) This article verifies the feasibility of a CAD parametric design algorithm based on machine vision through experiments and proves that this algorithm can improve the efficiency of animation scene design.

(c) The research methods and results of this article can not only be applied to animation scene design but also provide useful reference and inspiration for fields such as film special effects production, game design, virtual reality, etc.

Firstly, this article discusses the basic theories of CAD design methods and machine vision. Then, based on the specific requirements of animation scene design, a new machine vision-based CAD parametric design algorithm will be proposed, and it will be described and implemented in detail. Finally, the superiority of the proposed method is verified through experiments, providing a new and efficient solution for animation scene design.

2 THEORETICAL BASIS

With the continuous development of virtual reality technology, the design of 3D models for animated characters has become an important application field. The 3D model design method for multi-visual animated characters based on virtual reality technology can present the images and actions of animated characters more realistically and vividly, improving the visual experience and viewing effect of the audience. Li et al. [6] introduced a 3D model design method for multi-visual animated characters based on virtual reality technology and its application. It is necessary to use virtual reality technology to establish a three-dimensional model of animated characters. In the modeling process, it is necessary to consider details such as the proportions of the characters, muscle lines, clothing, and props in order to provide a high-quality model foundation for subsequent animation production. After establishing the 3D model, it is necessary to perform texture mapping and material mapping on the model. Texture mapping is the process of mapping a two-dimensional image onto a three-dimensional model, making the surface of the model richer and more realistic. Material mapping is the process of improving the visual effect of a model by setting surface material properties such as color, gloss, texture, etc. In order to achieve dynamic effects of animated characters, it is necessary to bind bones to the model and use motion simulation techniques to achieve the actions of the characters. In the process of bone binding, it is necessary to consider the proportion and structure of the character in order to set the connection relationship of the bones correctly. Meanwhile, through motion simulation technology, accurate control and adjustment of character movements can be achieved. With the rapid development of the digital age, animation scene design has become an important creative field. Although advances in computer graphics and artificial intelligence technology have brought higher efficiency and better visual effects to animation scene design, achieving diverse interactive recommendations for animation scenes remains a challenging issue. Li et al. [7] explored how to use deep learning techniques combined with multi-view visualization to achieve diverse interactive recommendations for animation scenes. Deep learning is an artificial neural network technique that can process and learn from large amounts of data. In computer graphics, deep learning has been widely applied in fields such as image recognition, object modeling, and scene rendering. Multi-view visualization is a technique that presents data from multiple perspectives and viewpoints, which can help people understand and analyze data more comprehensively. Users can interact with multi-view visualization interfaces to explore and evaluate different animation scene design solutions. Deep learning models can learn user interaction behavior and preferences and generate personalized animation scene recommendation schemes.

The combination of computer-aided design (CAD) and machine vision technology provides designers with new creative tools. Especially for ink elements with strong artistic and cultural connotations, CAD parameterization based on machine vision and forced perspective interactive projection mapping technology can better achieve their simulation and application in virtual environments. Liang and Kim [8] discussed the practice and application of this technology in the field of technical art. Machine vision technology can achieve precise analysis and understanding of images through image processing and pattern recognition algorithms. In CAD systems, automatic recognition and parameterization of ink elements can be achieved through machine vision technology. Firstly, through image processing techniques, feature extraction and edge detection are performed on ink elements, transforming them into geometric shapes that computers can recognize. Then, through parameterization techniques, these geometric shapes are represented as parameterized models for subsequent simulation and rendering. In the application of ink elements, through forced perspective interactive projection mapping technology, dynamic projection transformations can be performed on ink elements based on the user's perspective and position, making their presentation in virtual environments more realistic and vivid. At the same time, users can manipulate and adjust ink elements through interactive means, achieving real-time modification of their shape, size, position, and other attributes. With the development of technology and the arrival of the digital age, computer-aided design (CAD) has become an important tool in various fields. In landscape design, especially for Hlai ethnic villages with rich cultural heritage, the application of CAD parametric design provides designers with more possibilities. Liu et al. [9] explored how to apply machine vision-based CAD parametric design to the visualization of animated scenes in Hlai ethnic village landscape design in order to improve design efficiency and expressiveness. Machine vision is a technology that obtains and processes image information by simulating the human visual system. It can recognize various features in an image, such as shape, color, texture, etc., and convert them into data that computers can process. CAD parametric design is a technique based on parametric modeling, which achieves accurate representation and efficient optimization of models by representing them as parametric or constraint equations. Collect image information of the Hlai ethnic village through machine vision technology and use image processing techniques to preprocess the image, such as denoising and enhancement, to extract the village's characteristic elements and spatial relationships. After obtaining the parameterized model, computer graphics technology can be used to render and output the landscape, presenting animated scene effects with Hlai ethnic characteristics. For example, diverse landscape expression techniques can be achieved by simulating different lighting conditions and material effects.

In the field of art, blind artists often have unique and creative ways of perceiving and expressing the world. However, due to visual deficiencies, they often face challenges in understanding and shaping scenes. With the advancement of technology, animation and virtual reality (VR) have provided new possibilities for blind art practitioners to compensate for the lack of visual information and enhance scene perception. McSwan et al. [10] explored the application and potential of animation and virtual reality in the scene perception experience of blind art practitioners. Animation can express emotions and artistic conception through visual elements such as shape, color, lines, etc. These elements are endowed with unique meanings and symbols in the animation, providing blind artists with a new perspective on understanding the world. For example, through color changes in animations, natural phenomena such as day and night alternations and seasonal cycles can be conveyed, helping artists perceive and understand the flow of time. On the other hand, animation can enhance scene perception through auditory elements such as sound and sound effects. Sound can provide a sense of space, direction, and depth, helping blind artists better understand and perceive scenes in animation. For example, through changes in sound effects, spatial relationships such as distance, height, and size can be conveyed, allowing artists to perceive the levels and details of the scene. Automatic stereoscopic displays with LCD (liquid crystal display) are changing the way people watch 3D animations due to their unique advantages. Mori and Bao [11] discussed the working principle, advantages, and how to apply this new display device to watch 3D animations based on the Moir é effect. The automatic stereoscopic display with LCD utilizes the Moir é effect theory. This theory suggests that when two structurally similar objects move relative at a certain distance and speed, they produce a visual effect that makes people feel the third dimension. In an automatic stereoscopic display with an LCD, the pixels on the LCD screen flicker in a specific way while being synchronized with the special glasses worn, allowing the left and right eyes to see different pixels, thus forming a three-dimensional visual effect in the brain. Automatic stereoscopic displays with LCD have many advantages. Firstly, it does not require special glasses or other external devices, and users only need to wear the accompanying special glasses to enjoy stereo-visual effects. Secondly, LCD screens have the characteristics of high definition and resolution, which can provide users with an excellent picture quality experience. In addition, due to the use of automatic adjustment, this monitor can adjust the screen in real-time according to the position of the user's head and eyes, providing a more comfortable viewing experience.

In animated narrative short films, unique identity transformation experiences can be created through clever plot design and visual presentation. This experience often goes beyond simple role changes, touching on an individual's inner world and self-awareness. Najafi [12] uses 3D motion capture technology as an example to explore the heuristic exploration of identity transformation in animated narrative short films and how to endow animated characters with richer emotions and the inner world through the replacement of self-continuity. 3D motion capture technology is a real-time recording and analysis of object motion trajectories, widely used in the field of animation production. Through this technology, animators can capture the motion data of real actors or objects and convert it into digital models, achieving realistic animation effects. In animated narrative short films, 3D motion capture technology provides greater flexibility for character performance, providing more possibilities for the development of the story plot and the transformation of character identity. In animated narrative short films, the replacement of self-continuity refers to the character maintaining a certain internal continuity and consistency during the process of identity transformation in order to achieve the emotional expression and inner world display of the character. This kind of substitution can be physical, such as changing appearance or body shape; It can also be psychological, such as changing personality or behavior. In the field of animation production, scene design is a crucial aspect that can directly affect the overall effect and quality of animation. To improve the efficiency and effectiveness of scene design, we can use multiple CAD systems to generate parameterized animation motion scene designs. Sadler et al. [13] introduced how to use multiple CAD systems to achieve this goal. Using multiple CAD software can generate multiple design solutions simultaneously, and designers can choose and optimize according to their needs. Different CAD software has different characteristics and functions, and designers can use these characteristics to improve the effectiveness and quality of scene design. After the scene model is established, animation needs to be added to it. During this process, multiple CAD software can be used to design and adjust animation effects. For example, kinematic and dynamic analysis software can be used to simulate and analyze the actions and expressions of a character. Finally, it is necessary to render and post-process the scene. During this process, multiple CAD software can be used to optimize and adjust the effects and quality of the scene. For example, image processing software can be used to adjust parameters such as color, contrast, and brightness.

With the rapid development of technology, machine vision and CAD technology have been widely applied in various fields, including animation scene design. The CAD parametric design algorithm based on machine vision can convert visual information into analyzable data, providing more possibilities for animation scene design. Shan and Wang [14] discussed the application of this algorithm in animation scene design and how it can improve design efficiency and creative quality. Machine vision is a technology that obtains and processes image information by simulating the human visual system. It can recognize various features in an image, such as shape, color, texture, etc., and convert them into data that computers can process. CAD parametric design algorithm is a technology based on parametric modeling, which achieves accurate representation and efficient optimization of models by representing them as parametric equations or constraint equations. In animation scene design, CAD parametric design algorithms based on machine vision can combine the advantages of both to achieve efficient design and precise control of animation scenes. Through machine vision technology, designers can extract useful information from a large number of images and convert it into data that computers can process. Through CAD parametric design algorithms, designers can efficiently process and optimize these data, achieving the design and presentation of complex animation scenes. Chinese ink painting is a unique art form, with its concise lines and elegant colors expressing unique aesthetic taste and cultural connotations. In recent years, with the development of computer graphics, CAD parameterization methods based on machine vision have been widely applied in 3D model drawing. Yan et al. [15] proposed a non-realistic 3D mountain model drawing method based on machine vision and CAD parameterization for Chinese ink painting style, aiming to combine the artistic style of Chinese ink painting with modern computer technology to create a unique artistic charm of 3D mountain models. Machine vision is a technology that obtains and processes image information by simulating the human visual system. It can recognize various features in an image, such as shape, color, texture, etc., and convert them into data that can be processed by computers. The CAD parameterization method is a technique based on parametric modeling, which achieves accurate representation and efficient optimization of models by representing them as parametric equations or constraint equations. After obtaining the parameterized model, non-realistic rendering was performed using the Chinese ink painting style. Specifically, 3D mountain models with Chinese ink painting characteristics can be created by simulating the stroke effects, color matching, and rendering methods of Chinese ink painting.

3 MACHINE VISION GUIDED CAD PARAMETRIC DESIGN ALGORITHM

3.1 Special Requirements of Animation Scene Design

Machine vision is an important branch of the field of AI, which studies how to enable computers to "see" the world like humans and automatically recognize and understand objects in the real world. In this study, machine vision technology was used to automatically recognize and extract key elements in animation scenes, providing basic data for subsequent parametric design. CAD parametric design is a design method based on parameter constraints, which can quickly modify and adjust the design by adjusting parameter values. In this article, the CAD parametric design algorithm is used to realize the rapid modification and adjustment of key elements in an animation scene. By adjusting these parameters, the elements in the scene can be quickly modified and adjusted to meet different design requirements. In addition, the rule-based method is used to realize the automatic adjustment and optimization of parameters. Animation scene design involves multiple aspects, such as scene layout, lighting, and material settings. This article focuses on how to improve the efficiency of animation scene design through machine vision and CAD parametric design techniques. This method can automatically extract key elements in the scene and quickly modify these elements through CAD parametric design algorithms. In addition to the theoretical foundations mentioned above, the research also involves multiple technologies and tools. For example, the DL framework TensorFlow was used in the study to train convolutional neural networks (CNN); Using CAD software AutoCAD to achieve the rendering of parameterized design models; OpenCV, an image processing library, was used for image processing and pattern recognition. Animation is not just about presenting visuals but, more importantly, conveying emotions and atmosphere. Scene design needs to accurately create a specific atmosphere according to the needs of the plot. Designers need to convey these emotions through elements such as color, light and shadow, and composition. The scene is not just the background of the character; it should have an interactive relationship with the character. Designers need to ensure that the scene provides sufficient performance space for the characters and echoes their personalities and behavior. Although animation is fictional, it needs to give people a sense of realism so that the audience can immerse themselves in it. This means that animation design needs to be as realistic and detailed as possible for every detail in the scene.

Unlike static image design, animated scenes need to consider dynamic factors. Designers need to anticipate the presentation effects of the scene at different angles, lighting, and time points, and ensure that it remains aesthetically pleasing in motion. Scene design needs to be closely integrated with the plot, using visual elements to drive the growth of the story and reveal the character's inner world.

3.2 Extraction of Key Elements in the Scene

A deep learning (DL) based image segmentation algorithm was used in the study to classify and recognize scene images at the pixel level. By training a large amount of scene image data, the algorithm learns the features and patterns of different scene elements. In practical applications, the first step is to use a camera or scanner to obtain the 3D model data of the scene and convert it into a 2D image format. Then, input these images into the image segmentation algorithm to automatically recognize and extract key elements from the scene. These key elements will be labeled with different colors or labels for easy processing and modification in the future.

When representing a 3D solid model with polygons, triangles are usually used as the most basic polygon units, and the 3D model can be represented as a set of triangles. In Euclidean geometric space, triangles represent the simplest and most accurate way to represent 2D regions, while points can only be represented as particles. A triangular patch can represent a 2D region in 3D space, but when using points to describe it, at least three points are required. This means that compared to triangles, point groups of the same size have a weaker ability to describe 3D entities. Figure 1 illustrates that when using point representation models, the required quantity of points is much greater than the quantity of patches.

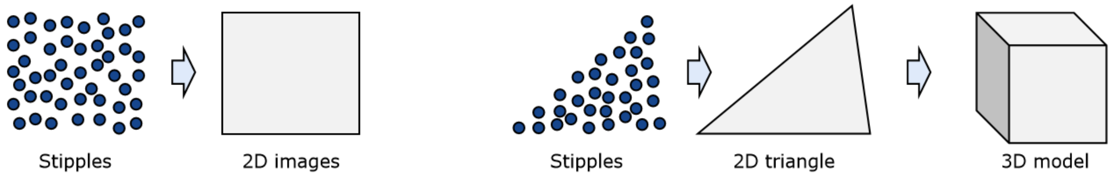

Figure 1: The difference between using point sets to represent images and 3D graphics.

Firstly, it is needed to prepare an annotated animation scene dataset., Contains various types and styles of animated scene images, and each scene image has corresponding labels or annotation information. After the model construction is completed, a pre-trained CNN model is used for feature extraction, and the animated scene images are input into the CNN model. The output of the convolutional layer is obtained through forward propagation as the feature representation. The extracted feature representations can be used as input data to train a classifier or other machine learning algorithms for learning specific tasks in animation scenes. The trained and optimized model can be used in practical animation scene recognition and analysis tasks. The corresponding hole-pooling operation is shown in Figure 2.

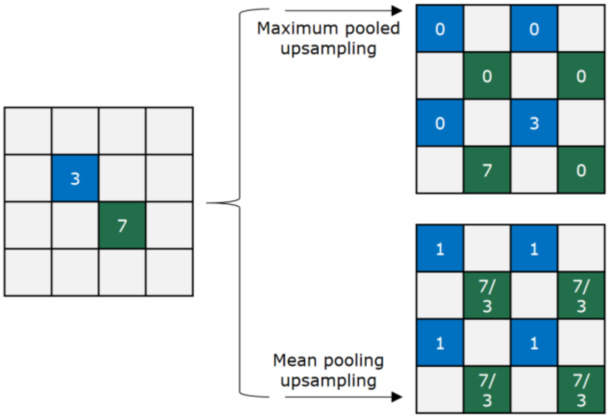

Figure 2: Cavity pool reverse propagation.

After extracting key elements, use CAD parametric design algorithms to quickly modify these elements. Firstly, based on the extracted key element information, construct a parameterized model of the scene. This model includes the shape, size, position, and other parameters of all elements in the scene, as well as their relationships and constraints. Then, use the parametric design tools provided by CAD software to adjust and optimize these parameters. Each adjustment is immediately reflected in the rendering effect of the scene, allowing the design to preview and adjust the scene's effect in real time. In addition to linear SVM, there is also non-linear SVM, which is proposed to solve the problem of not being able to use linear plane segmentation feature vectors, as shown in Figure 3.

The kurtosis distribution of coefficients obtained from the basis function responses of different styles of test images and static scene images is shown in Figure 4. The kurtosis distribution of coefficients obtained by using the basis function response can effectively distinguish different styles of animation scenes. When using test images with the same style as static scene animation works, the maximum kurtosis value is obtained. This means that the response between the same style of animation scene and the base function is the strongest. Therefore, the coefficients obtained from the response are the sparsest, meaning only a few coefficients have significant values, while most of the other coefficients are close to zero. For other styles of test images, the kurtosis value is lower, indicating that there are differences in style between them and static scene works. This difference may be caused by differences in color, texture, layout, or other visual elements.

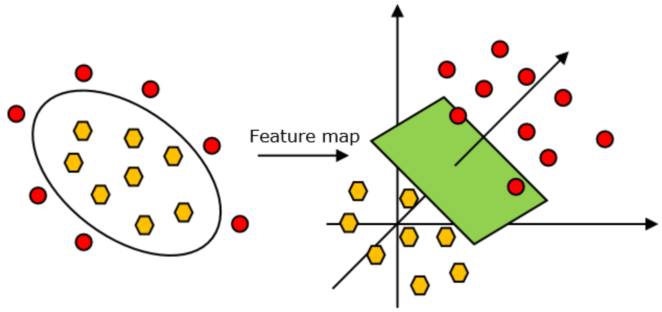

Figure 3: Principle of SVM nonlinear classifier.

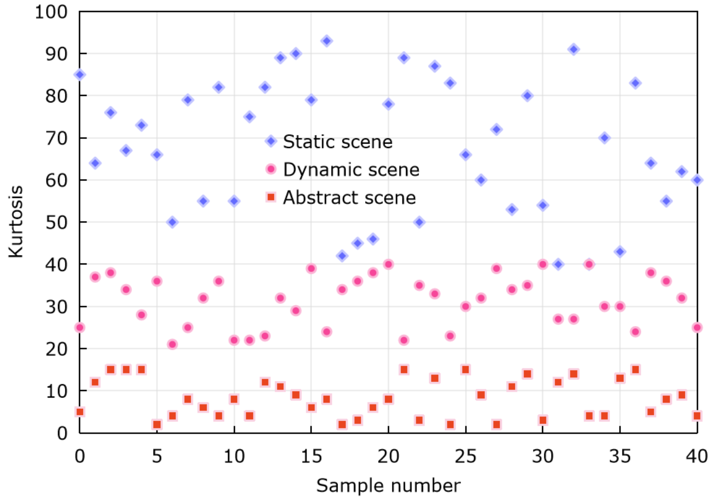

Figure 4: Kurtosis distribution diagram.

From a practical application perspective, this method provides an effective tool for style classification and recognition in animation scenes. For example, in the animation production process, this method is used to ensure that the style of the scene matches expectations; Alternatively, in animation archiving and retrieval, this method can be used to identify and classify different scene styles quickly.

3.3 Automated Generation of Animation Scenes

In order to achieve an automated generation of animation scenes, the results obtained from feature extraction and learning through CNN can be utilized. In the process of automated animation scene generation, scene layout refers to how to arrange the positions and relationships of different elements in the scene to form a reasonable and attractive spatial structure. Object placement refers to how to place various objects in a scene to make it more vivid and rich. To achieve these two tasks, use rule-based or learning-based methods. Rule-based methods can lay out scenes and place objects based on predefined rules and constraints. Learning-based methods can teach how to lay out scenes and place objects from a large amount of training data. We can model the task of scene layout and object placement as an optimization problem and use gradient descent or other optimization algorithms to solve the optimal solution of the problem. In this process, the results obtained through feature extraction and learning using CNN are used as input features to improve the performance of the model.

The feature of the animation image is extracted by feature functions f_A and f_B, and the output feature vectors are u and v respectively. Then, in the part of the pooling function P, the bilinear convolution neural network model adopts the bilinear pooling method to bilinear combine the output feature vectors:

$$bilinear\ f_A, f_B = f_A \otimes f_B = u^T v \tag{1}$$

That is, the feature vector u and the feature vector v are subjected to the outer product operation to obtain the bilinear feature $bilinear\ f_A, f_B$. On this basis, in order to merge all the bilinear features in the image, the bilinear features at each position are accumulated and summed:

$$\phi\ f_A, f_B = \sum_{d=1}^{D} bilinear\ f_A, f_B = \sum_{d=1}^{D} u^T v \tag{2}$$

Where D is the one-dimensional length of the feature map; Then, the vector $\phi\ f_A, f_B$ is normalized and L_2 L_2 regularized:

$$y = sign\ \phi\ f_A, f_B \sqrt{\left|\phi\ f_A, f_B\right|} \tag{3}$$

$$z = \frac{y}{\|y\|_2} \tag{4}$$

Finally, the obtained vector z is input into the classification function C for recognition and classification, and the classifier selects the Softmax function.

In the process of automatically generating animated scenes, lighting, and rendering are two key steps. Lighting refers to how to set the light sources and lighting conditions in a scene to create a specific atmosphere and visual effect. Rendering refers to how computer graphics techniques are used to convert a scene from a 3D model to a 2D image or video. Physics-based lighting models can simulate lighting effects in the real world, making scenes more realistic. Efficient rendering algorithms can improve rendering speed, and reduce production time and cost while ensuring image quality. The error is defined as:

$$\cos t\ u, v = \|u - v\| \times \max_{f \in T_u} \left\{ \min_{n \in T_{uv}} \ 1 - f \cdot normal \cdot n \cdot normal \ \div 2 \right\} \tag{4}$$

$\|u - v\|$ is the distance between two u, v points, T_u is all triangular surfaces adjacent to u point, and T_{uv} is all triangular surfaces with uv as edges.

The eigenvector corresponding to the smallest eigenvalue M can be regarded as the normal vector of point P_i.

$$M = \sum_{i=1}^{K} P_i - P' \ \ P_i - P' \ ^T / K \tag{5}$$

The angle between the new normal N_{new} of the triangle and the normal N_{ini} in the original mesh is also smaller than the threshold ε:

$$\arccos\ N_{new} \cdot N_{ini} \ \le \varepsilon \tag{6}$$

The texture distribution set of the fuzzy regional animation image is calculated as follows:

$$w\ i,j\ = \frac{1}{Z\ i}\exp\left(-\frac{d\ i,j}{h^2}\right) \tag{8}$$

Where $Z\ i$ are the first-order and second-order texture distribution operators. The relevant parameters are replaced and converted into:

$$W^{'} = \frac{1}{2} f\ x^{'}, y^{'}, z^{'}\ + E \tag{9}$$

Where: $x^{'}, y^{'}, z^{'}$ is the 3D coordinate with visual constraint.

4 MODEL TESTING AND ANALYSIS

4.1 Simulation Test Results

In order to achieve the goal of parameterized design of animation scenes, this study prepared a large-scale animation scene dataset. This dataset contains various types and styles of animation scenes, as well as relevant descriptions or label information for each scene. These data can be used to train models to recognize and understand the description or requirements of animation scenes, as well as generate corresponding animation scenes. This study uses error rate and accuracy as the main performance indicators. The error rate is used to evaluate the degree of difference between the generated scene and the real scene of the model, while the accuracy rate is used to evaluate the model's ability to recognize test samples. Through comparative experiments, it is expected that the animation scene generation method proposed in this article can demonstrate better performance in terms of error rate and accuracy, verifying its effectiveness and application value in the field of animation design.

The results in Figure 5 show the error rate changes during the training stage of the classical CNN model, the classical GAN model, and the model proposed in this article. As the quantity of training sessions increases, the error rates of all three network models show a decreasing trend. This indicates that these models gradually learn effective representations and mapping relationships of the data during the training process. In contrast, the training error rate of the model in this article is the lowest and ultimately stabilizes at around 0.05.

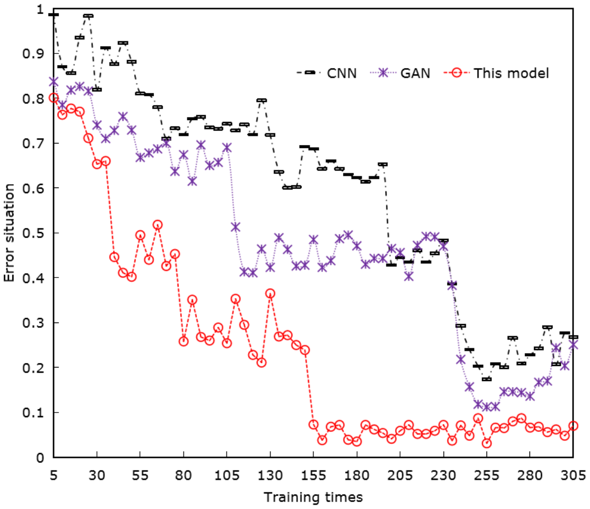

Figure 5: Training effect diagram of different models.

The CNN model is a commonly used method for image classification and feature extraction, which can effectively learn local features and spatial structures in images. However, in the task of generating animation scenes, in addition to local features, it is also needed to consider the global structure and semantic information of the scene. Therefore, CNN models may not be able to capture this global and semantic information well. The GAN model is a generative model that can learn potential representations of data distribution and generate new data samples. In the task of generating animation scenes, it is not only needed to generate the image itself but also to ensure that the generated image has a specific style and content. This article combines the advantages of CNN and GAN and introduces new technologies and methods to solve the challenges in animation scene generation tasks. This article also uses rule-based or learning-based methods for scene layout and object placement and uses physics-based lighting models and efficient rendering algorithms for lighting and rendering. These technologies can capture global and semantic information of animation scenes, and control the style and content of generated images.

Carry out experiments individually on the CNN model, GAN model, and our novel model using the test samples. Figure 6 illustrates the error occurrence of the model within the test sample, while Figure 7 demonstrates the model's precision in the test sample.

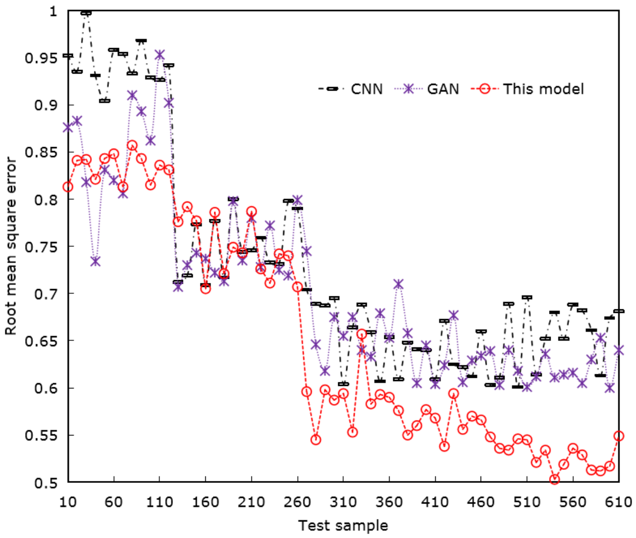

Figure 6: Error of the model in the test sample.

Among the test samples, the CNN model exhibited a considerable error rate, eventually leveling off at an approximate value of 0.68. The GAN model's error also plateaued, settling around 0.63. However, the technique presented in this article displayed the lowest error and reached stability at around 0.54, indicating its superior performance in generating animation scenes.

Within the test sample, the CNN model's accuracy performance falls short of expectations, hovering around 80%. This deficiency stems from the inherent limitations of CNN models in animation scene generation tasks, leading to constrained recognition capabilities for intricate scenes and subsequently impacting accuracy. On the other hand, the GAN model boasts a higher accuracy rate compared to the CNN model, at approximately 85%. This advantage can be attributed to the GAN model's proficiency in generating novel data samples and its enhanced ability to discern data distribution within the test samples, thereby elevating accuracy. Notably, the model proposed in this article achieves optimal accuracy, surpassing 90%. This exceptional result underscores the superior performance of the method introduced in this article for animation scene generation tasks. By adeptly capturing the global and semantic nuances of scenes, this method generates exceptional-quality scenes with distinctive styles and rich content.

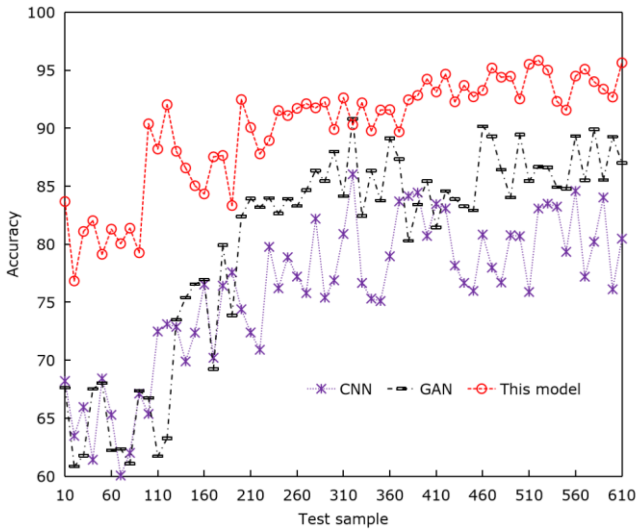

Figure 7: Accuracy of the model in the test sample.

4.2 Results Discussion and Analysis

Comparative experiments conducted on the traditional CNN model, conventional GAN model, and our novel model in animation scene generation tasks lead to the following inferences:

In terms of error rate, the CNN model has a relatively high error, the GAN model has a moderate error, and the method proposed in this article has the lowest error. This indicates that the method proposed in this article has better performance in animation scene generation tasks and can reduce error rates. In terms of accuracy, the GAN model has a higher accuracy than the CNN model, while the accuracy of our model is the best, reaching over 90%. This further confirms the superiority of the method proposed in this article in the task of generating animation scenes.

According to the results, the method proposed in this article performs the best in the task of generating animation scenes. Therefore, when designing animations, priority should be given to using the model proposed in this article or other learning-based generative models to assist the design process. By automating the generation of animation scenes, designers can greatly reduce their manual drawing and adjustment workload. This is particularly important for animation production projects that require rapid iteration and mass production. Automated generation technology can help designers quickly generate scenes and objects of various styles, thereby providing designers with more inspiration and possibilities. By combining DL and computer graphics technology, automated generation technology can generate scenes and objects with high quality and rich details.

Compared with traditional methods, this algorithm can automatically identify and extract key elements in the scene, greatly reducing the workload of manual adjustment and optimization. Moreover, the algorithm can also be quickly modified and adjusted according to the requirements of directors and designers, ensuring the scalability of the design. The algorithm can also achieve real-time preview and adjustment of the scene, making it more convenient for designers to create.

5 CONCLUSION

CAD parametric design is a design method based on parameter constraints, which can quickly modify and adjust the design by adjusting parameter values. In animation scene design, its application is still in its early stages. The research objective of this article is to develop a machine vision-based CAD parametric design algorithm that can automatically recognize and extract key elements in animation scenes, and achieve rapid modification and adjustment of these elements through parametric design. According to the results, the method proposed in this article performs the best in the task of generating animation scenes. Therefore, when designing animations, priority should be given to using the model proposed in this article or other learning-based generative models to assist the design process. By automating the generation of animation scenes, designers can greatly reduce their manual drawing and adjustment workload. This is particularly important for animation production projects that require rapid iteration and mass production. Automated generation technology can help designers quickly generate scenes and objects of various styles, thereby providing designers with more inspiration and possibilities.

In the future, more DL technologies can be combined with animation scene generation tasks, such as reinforcement learning, transfer learning, etc., to further improve the diversity of generated scenes. In order to improve the generalization ability of the model, it is possible to try using a larger-scale animation scene dataset for training and study how to achieve better performance on a limited dataset.

Sihui Yu, https://orcid.org/0009-0000-4884-709X
Yuzhi Chen, https://orcid.org/0009-0000-4149-8122
Tianyi Guo, https://orcid.org/0009-0008-3413-5145
Dong Yang, https://orcid.org/0000-0003-2465-8555

REFERENCES

[1] Bao, W.: The Application of intelligent algorithms in the animation design of 3D graphics engines, International Journal of Gaming and Computer-Mediated Simulations, 13(2), 2021, 26-37. https://doi.org/10.4018/IJGCMS.2021040103
[2] Guo, S.; Wang, B.: Application of computer aided modeling design in the expression techniques of sculpture art space, Computer-Aided Design and Applications, 19(S3), 2021, 1-12. https://doi.org/10.14733/cadaps.2022.S3.1-12
[3] Huang, L.; Hou, Z.-X.; Zhao, Y.-H.; Zhang, D.-J.: Research progress on and prospects for virtual brush modeling in digital calligraphy and painting, Frontiers of Information Technology & Electronic Engineering, 20(10), 2019, 1307-1321. https://doi.org/10.1631/FITEE.1900195
[4] Jing, Y.; Song, Y.: Application of 3D reality technology combined with CAD in animation modeling design, Computer-Aided Design and Applications, 18(S3), 2020, 164-175. https://doi.org/10.14733/cadaps.2021.S3.164-175
[5] Kiu, Z.: The Dynamic effect of visual communication in web design and its technical realization, Computer-Aided Design and Applications, 17(S2), 2020, 78-88. https://doi.org/10.14733/cadaps.2020.S2.78-88
[6] Li, L.; Zhu, W.; Hu, H.: Multivisual animation character 3d model design method based on VR Technology, Complexity, 2021(4), 2021, 1-12. https://doi.org/10.1155/2021/9988803
[7] Li, Y.; Qi, Y.; Shi, Y.; Chen, Q.; Cao, N.; Chen, S.: Diverse interaction recommendation for public users exploring multi-view visualization using deep learning, IEEE Transactions on Visualization and Computer Graphics, 29(1), 2022, 95-105. https://doi.org/10.1109/TVCG.2022.3209461
[8] Liang, K.; Kim, H.-G.: Compulsory perspective interaction projection mapping with elements of ink paintings, TECHART: Journal of Arts and Imaging Science, 7(3), 2020, 9-16. https://doi.org/10.15323/techart.2020.8.7.3.9
[9] Liu, J.; Wu, X.; Zhang, Y.; Wang, L.: Visualization system of Hlai ethnic village landscape design based on machine learning, Soft Computing, 27(14), 2023, 10001-10011. https://doi.org/10.1007/s00500-023-08196-8
[10] McSwan, A.: Exploring animation and virtual reality to represent the perceptual-experiences of art-practitioners with sight-loss, The Design Journal, 24(2), 2021, 315-324. https://doi.org/10.1080/14606925.2021.1877237

[11] Mori, S.; Bao, Y.: Autostereoscopic display with LCD for viewing a 3-D animation based on the moiré effect, OSA Continuum, 3(2), 2020, 224-235. https://doi.org/10.1364/OSAC.383279
[12] Najafi, H.: Displacement of self-continuity: An heuristic inquiry into identity transition in a 3D motion-capture-based animated narrative short film, Animation Practice, Process & Production, 8(1), 2020, 165-188. https://doi.org/10.1386/ap3_00010_1
[13] Sadler, J.-E.; Day, R.-D.; Bronson, P.-G.: A neutral xml design framework for generating parametric parts in multiple CAD systems, Computer-Aided Design and Applications, 2019(5), 2019, 16. https://doi.org/10.14733/cadaps.2019.923-935
[14] Shan, F.; Wang, Y.: Animation design based on 3D visual communication technology, Scientific Programming, 2022(11), 2022, 1-11. https://doi.org/10.1155/2022/6461538
[15] Yan, M.; Wang, J.; Shen, Y.; Lv, C.: A non-photorealistic rendering method based on Chinese ink and wash painting style for 3D mountain models, Heritage Science, 10(1), 2022, 1-15. https://doi.org/10.1186/s40494-022-00825-z