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

Implementation and Exploration of a Collaborative Platform for Art Creation and Design Based on Multimodal Perception

Dazhuang Wang[1] and Xu Zhao[2]

[1]Department of Fine Arts, Jilin Normal University, Siping 136000, China, wdazhuang1984@163.com
[2]Library, Jilin Normal University, Siping 136000, China,
zhaoxu202304@163.com

Corresponding author: Xu Zhao, zhaoxu202304@163.com

Abstract. The aim of this study is to construct a collaborative platform that integrates art CAD (Computer Aided Design) and VR (Virtual Reality) with multimodal perception. This integration seeks to improve user immersion and interaction during the artistic creation and design process. To accomplish this objective, the article conducts a thorough analysis of user needs and defines the corresponding functions and design goals of the platform. Through architecture design, algorithm implementation, and simulation experiments, this article successfully combines the accuracy of art CAD with the realism of VR and integrates multi-modal perception technology so that users can perceive and interactively design works through multiple senses. The experimental results show that the platform not only provides powerful art CAD function but also supports highly realistic VR display and realizes the integration of visual, auditory, and tactile perception methods. In addition, the platform is easy to use and collaborative and supports multi-person online collaborative editing, commenting, and sharing design works.

Keywords: Art CAD; Virtual Reality; Multi-Modal Perception; Create A Platform Together; User Experience
DOI: https://doi.org/10.14733/cadaps.2024.S28.110-124

1 INTRODUCTION

As science and technology rapidly advance, art CAD technology and VR technology have emerged as crucial instruments in the realm of art design. In applications such as intelligent transportation and video surveillance, precise separation of foreground (FG) and background (BG) is not only crucial for functional implementation but also demonstrates infinite innovative possibilities in the field of art. In terms of artistic expansion, the FG-BG separation technology based on multimodal perception can be applied to various artistic scenes. The term "multimode" here not only refers to different frame sequences in the video but also encompasses various sensor data such as sound, light, temperature, etc. The traditional computer vision-based FG-BG separation technology, although able to meet functional requirements to a certain extent, appears relatively singular in artistic expression. Akilan et al. [1] utilized Long Short Term Memory (LSTM) pipelines and 3D Convolutional Neural Networks (3D CNN) to capture the dynamic features of artworks and also incorporated the concept of multimodal perception. In digital art performances, artists can use this technology to separate the performer's actions from the background environment. It further combines FG-BG separation technology with artistic style transfer technology to achieve the fusion of FG objects in videos with different artistic style backgrounds. The Great Resource Sharing Plan has further accelerated the digitalization process of education, and the digital demands of school education in the digital education environment have become increasingly clear. After preliminary sorting and screening of virtual museums and related educational resources, the Beijing Palace Museum virtual museum platform was selected. In this regard, Bansal et al. [2] discussed and summarized a feasible model framework for integrating virtual museum resources in middle school art appreciation courses, for teachers to flexibly select and use in the future. A network exploration model was adopted for the subsequent integration of teaching resources, and a task-driven teaching activity plan was designed and tested in S School in Nanchang City. Students have significantly improved in situational experience, cultural understanding, resource integration, and learning methods, while only showing a small improvement in image recognition and creative practice. Based on the analysis and practice of the pre - and post-test questionnaire survey results, it can be concluded that the online inquiry teaching model has certain effectiveness in integrating virtual museum resources in art appreciation courses. The iterative evolution of digital technology and the emergence of various public education resources have made the integration of public digital resources into school teaching a new path to compensate for the technological lag in schools. VR technology, on the other hand, enables users to experience and interact in an immersive way by constructing a 3D environment, which greatly enriches the display mode and sense of participation in art design. From a technical perspective, the composition of interactive robot art installations is quite complex and precise. The multi-colour drawing software driven by artificial intelligence will enter an automatic progress system every time it completes the art image drawing. The audience shares their past dreams through specific means, such as voice input, and these imaginative works are then captured and converted into painting instructions by the AI system in the device. KUKA industrial robot, as the executing mechanism, is responsible for drawing according to the instructions of the drawing software. Next is multi-colour drawing software driven by artificial intelligence. This section is responsible for converting the drawing instructions captured by AI into specific drawing actions. In terms of audience interaction design, Canet and Guljajeva [3] cleverly utilized interactive technology to enable the audience to easily participate in the process of artistic creation.

Now, the concept of "new folk art" has been formally and clearly proposed. Network culture in China is like a surging river, absorbing the wisdom of different groups of folk culture. They are completely different from traditional mainstream culture and have become a new cultural ecological style. Cedillo et al. [4] investigated the survival status of the ancient form of art watermark in the new media context (online oil painting community). At present, the online art community is still a new phenomenon and phenomenon, although virtual, it does exist in the real world. It has not been guided and regulated by mainstream consciousness, nor has it been processed and transformed by cultural elites. The online approach has created infinite conditions for the development of new folk culture and has built a new folk culture square. The online art community is a member of this new folk culture square. Because it contains rich consensus and has strong cohesion like grassroots, it also has strong vitality and independence. And be able to approach the profound changes in art brought about by the rapid development of technology more rationally, calmly, and dialectically. Simultaneously, calmly and efficiently utilizing new information media to invest in art education, in order to better promote the inheritance and development of the art industry. In order to deeply address the inherent challenges in the practical teaching of art and design in the classroom, Fan and Li [5] proposed a multi-modal perception-based interactive learning method for artworks. It introduces artificial intelligence technology to evaluate the quality of art and design teaching. In traditional art and design teaching, students often can only learn and evaluate works through a single visual mode. Often relying on the personal experience and subjective judgment of experts or teachers, it is difficult to ensure the accuracy and impartiality of the evaluation. The learning method of multimodal perception emphasizes the use of multiple sensory channels (such as visual, auditory, tactile, etc.) to comprehensively understand and experience artistic works. Using BPNN to build an interactive learning system that can automatically adjust teaching content and methods based on student feedback and performance, thereby achieving personalized teaching. The experimental results indicate that the model used for multimodal perception in evaluating art and design teaching has high accuracy. It can improve students' practical skills. Through multimodal perception and interactive learning, students can gain a deeper understanding and experience of artistic works, thereby continuously improving their skills in practice. The use of artificial intelligence technology for evaluation purposes can also improve the quality of art and design education in universities, and the proposed interactive learning methods have significant benefits for students majoring in art and design.

The objective of this study is to delve into the profound amalgamation of art CAD and VR technology, aiming to achieve the digitization, virtualization, and enhanced sensory experience of artistic designs by establishing a collaborative platform. This undertaking strives to introduce a fresh creative atmosphere for artists and designers while enabling the audience to more readily comprehend the allure and design ethos inherent in artistic works. The fruition of this platform is destined to leave a lasting impression on both artistic design and the VR realm, fostering technological advancements in artistic design, bolstering design proficiency and excellence, and ultimately presenting the audience with an array of enriching and multifaceted artistic encounters. This study commences by analyzing the theoretical foundations and technical traits of art CAD, VR, and multimodal perception technology. Subsequently, it deliberates on their potential utilization in art design. Following this, a collaborative platform that amalgamates art CAD, VR, and multimodal perception will be constructed. This platform will facilitate the consolidation of diverse design tools, the formulation of a virtual environment, and the gathering and handling of multimodal sensory data.

In the realm of research methodologies, this article employs multiple avenues, such as literature reviews, case analyses, and experimental validations, to uphold the comprehensiveness and precision of the research. Additionally, it actively engages in exchanges and collaborations with experts and scholars in the field to garner additional research support and resources.

This article is structured as follows: The introduction section outlines the research background, objectives, contents, and methodologies adopted. Next, the literature review section surveys the current research status and development trends in related fields. The third section provides an in-depth description of the design and implementation steps of the co-creation platform. The fourth section presents the algorithm's implementation alongside the corresponding simulation results. The fifth section delves into the practical application and testing of the platform. Lastly, the sixth section offers a concise summary of the research findings and outlines potential future research directions.

2 RELATED WORK

The traditional method of generating artistic glyphs often adopts a multi-stage approach, such as generating shapes first and then adding textures. Introducing local texture refinement loss is the key to improving the quality of synthesized textures. Different decoders simultaneously generate glyph images and texture images, ensuring consistency between shape and texture during the generation process. The loss function ensures that the model pays attention to local details when generating textures, thereby generating more realistic and delicate artistic glyph images. The two encoders in the model are used to decode the representation of content and style, ensuring the generation of multiple content and styles. This not only improves the generation efficiency but also reduces the errors that may be introduced by intermediate steps. The AGIS Net model achieves synchronous transmission of shape and texture styles through a one-stage approach. This means that users can generate rich and diverse artistic glyph images by providing different content (such as different characters) and styles (such as different artistic font styles) [6]. With the rapid development of digital technology, the expression forms and audience experiences of media art are undergoing unprecedented changes. Especially when these technologies are applied to installation art, they can greatly enrich the sense of hierarchy and immersion in the work. The enhancement of this interactivity not only allows the meaning of the work to be more directly conveyed to the audience but also stimulates their sense of participation and creativity, making them common participants in artistic creation. Interactive media art captures and generates corresponding content through sensors, thereby establishing a dynamic relationship with the audience, allowing them to experience and feel the meaning conveyed by the work more deeply. In the process of exploring interactive media art, Jeon [7] noticed its potential challenges and limitations. How to balance the freedom of artistic creation and the operability of technology is also a question that needs continuous exploration. The stability and reliability of technical equipment are crucial for ensuring the smooth operation of the work. By providing real-time interactive feedback, a closer connection is established between the work and the audience.

At the same time, digital art and design education have gradually become a research focus in the field of education. It not only changes the appearance and connotation of modern display design but also provides infinite possibilities for the innovation and development of art and design. In order to better integrate digital technology with art and design, Li [8] combined modern technology with traditional art and design fields and designed a digital art and design system based on big data technology and interactive virtual technology. It can not only provide users with precise data support and rich interaction methods, but also achieve the virtualization, informatization, and digital transformation of art and design. The system aims to provide more accurate data support for art and design through the collection and analysis of big data. In terms of methodology, it delves into key technologies such as big data technology, interactive virtual technology, digital art design, and 3D file formats. This article also introduces the relevant content of support vector machines in algorithms, providing strong technical support for the intelligent recognition and analysis of the system. Meinecke et al. [9] borrowed the popular multimodal perspective to explore the teaching and practice of high school art appreciation, aiming to innovate the classroom teaching form of ordinary high school art appreciation. However, in the new era, teachers are striving to implement the new curriculum concept in their teaching process. Enable students to take the high school art appreciation classroom as the starting point for their love of art. Through systematic learning and practice, their ability to appreciate and discern can be improved. However, in actual teaching, there are still many practical problems that hinder innovation in high school art appreciation classrooms. How to design and carry out specific appreciation activities and flexibly and reasonably apply modern information technology in high school art appreciation classes. By utilizing multimodal composite resources, multimodal network tools, diverse interactive learning, and a combination of multiple evaluation methods, we aim to enrich the high school art appreciation classroom. As researchers of subject teaching, they have explored the common multimodal teaching theories in subject English research that can be applied interdisciplinary to high school art appreciation activities after focusing on the teaching models of the entire subject community.

In terms of practice, there are significant differences in the application of the educational function of art museums both domestically and internationally. The plan launched by foreign virtual art museum education is "from the inside out, from the bottom up," which means that the designed digital teaching resources of the museum are first aimed at school education and then gradually developed outward, accommodating various social resources [10]. The educational resources follow a development trend of "from the outside to the inside, from top to bottom." The emphasis of public art education in our country on improving the artistic literacy of the people has directly led to the fact that the public art education function of museums has long been mainly focused on propaganda and education. The existing educational resources of virtual art museums are mainly aimed at social groups without segmenting them into age or academic groups. A few teaching resources aimed at schools are also mostly set up for college students [11]. At the same time, when developing teaching resources that match the school curriculum, it is also aimed at K-12 basic stage students, children, and families first, and then extended upwards to continue developing learning resources suitable for groups such as universities, adults, and communities. In summary, when examining the current historical stage of "sensory technology," the aesthetic significance of virtual art museums is self-evident. Nowadays, virtual art museums generally integrate technology into art, creating important conditions for the inheritance of visual culture and image civilization [12].

These technologies can help artists and curators better manage and showcase artworks, improving the efficiency and quality of art exhibitions. The increasingly significant trend of diversification in environmental art design, combined with advanced 3D virtual reality technology, has brought revolutionary changes to modern display design. Wang and Hu [13] explored the integration of modern technology and traditional art, achieving innovative applications of traditional environmental art and thereby promoting the economic and cultural development of cities and society. This indicates that in a virtual environment, art design, fast response, and accurate presentation are the two aspects that users are most concerned about. In the experiment, they used 3D virtual reality technology to immerse users in a comprehensive virtual environment by wearing special glasses. The experimental results show that the weights of fast and accurate indicators are 0.65 and 0.37, respectively. The degree of functionality focuses on the practicality and functionality of virtual environments in art exhibitions, and the degree of use measures the convenience and comfort of users when using virtual environments. In this environment, not only has the visual experience been greatly improved, but the sound is also delivered through speakers or advanced acoustic transmission technology, bringing users a more realistic perception.

The artistic text effect, as a brilliant pearl in design elements, endows the text with visual charm beyond language with its unique contours, rich colours, and delicate textures. Yang et al. [14] learned a large amount of existing art text effect data to enable machines to automatically understand and imitate these effects and apply them to new texts. Due to the complexity of artistic text effects, data collection and annotation require extremely high costs and time. Therefore, existing datasets are often small in scale and difficult to support when training complex models. The creation of such effects often relies on experienced designers who, through exquisite skills and sharp aesthetics, transform texts into pieces of art. This method adopts the architecture of generative adversarial networks and achieves the goal of transferring all art styles within a model by training a generator and a discriminator. Based on summarizing the general design methods of art products, Zhang and Rui [15] adopted a modular decomposition method to implement the art design process on a computer. Research can help promote the teaching philosophy of digital computer graphics and image-assisted art design. It discussed the issue of establishing an art image-assisted design platform system in computer graphics. We have built a computer graphics and image-assisted art teaching platform with digital content innovation as the core. It is using digital means to establish a sensory interactive teaching method, achieving an intuitive experience of the spatial form theory of artistic agricultural products, cultivating students' modelling ability and aesthetic judgment.

3 DESIGN OF ART CAD AND VR CO-CREATION PLATFORM COMBINING MULTI-MODAL PERCEPTION

3.1 Platform Requirements Analysis and Function Definition

Art CAD technology, since its inception, has progressed significantly over the past few decades. Evolving from basic 2D drawing tools to modern 3D modelling and rendering software, it continues to innovate, bolstering the creative processes of artists and designers. Presently, art CAD technology finds widespread application in areas like architectural design, product development, animation production, and beyond. VR technology, on the other hand, enables the creation and experience of virtual worlds. Highly realistic VR presentation: users expect to preview and interact with their own designs in a virtual environment so as to evaluate the design effect more intuitively. Support of multi-modal perception technology: users hope to perceive and interact with the design works through various senses, such as vision, hearing, and touch, so as to enhance the sense of reality and immersion in design. Ease of use and collaboration: the platform needs to be easy to use and support multi-person online collaboration so that team members can participate in the design process together.

Based on the above requirements, this article defines the following main functions for the platform: art CAD design module, providing rich drawing and design tools; VR display module, which supports 3D preview and interaction of design; Multi-modal perception integration module, which realizes the integration of vision, hearing and touch; Collaboration and sharing function, supporting multiple people to edit, comment and share design works online.

3.2 Platform Architecture Design

The overall architecture of the platform is divided into four main layers: data layer, service layer, application layer and interaction layer.

Data layer: responsible for storing and managing all design data, user information and VR scene data.

Service layer: providing core service functions, such as CAD design service, VR rendering service, multimodal perception data processing service, etc.

Application layer: contains specific functional modules, such as CAD design module, VR display module, etc. These modules realize their respective functions by calling the services of the service layer.

Interaction layer: handling user interaction, capturing their inputs, and delivering pertinent feedback are among its responsibilities. This layer is also responsible for the integration of multi-modal sensing technology, such as displaying virtual scenes through a VR helmet and receiving user's operation instructions through gloves or a stylus.

4 ALGORITHM IMPLEMENTATION AND SIMULATION EXPERIMENT

4.1 Algorithm Implementation

3D modelling algorithm: 3D modelling using art CAD technology is one of the core functions of the platform. In this article, a modelling algorithm based on boundary representation (B-rep) is adopted, and an accurate 3D model is constructed by defining the geometric information and topological structure of the model. Basic structure definition of the algorithm:

The point $P = x,y,z$ and the line L are defined by the parametric equations of the two points:

$$L: \begin{cases} x\ t = x_0 + t\ x_1 - x_0 \\ y\ t = y_0 + t\ y_1 - y_0 \\ z\ t = z_0 + t\ z_1 - z_0 \end{cases} \tag{1}$$

Where $t \in [0,1]$. In the concrete implementation, the B-rep algorithm first describes the shape of the model through basic geometric elements such as points, lines and surfaces. The relationship between these elements, namely the topological structure, is strictly defined, which ensures logical consistency within the model.

For a 3D model whose surface geometry information can be given by a parametric equation, its volume V can be obtained by the following integration:

$$V = \iiint_S \det\left(\frac{\partial \vec{r}}{\partial u}, \frac{\partial \vec{r}}{\partial v}, \vec{n}\right) dA \tag{2}$$

Where $\vec{r}\ u,v$ is the parametric equation representing the model surface, $\det$ is the determinant, $\vec{n}$ is the unit normal vector of the point on the surface S, and dA is the area element on the surface S. The surface area A of the model can be obtained by summing the surface areas of each independent surface:

$$A = \sum \text{Superficial area}_{\text{i}} \tag{3}$$

Each $\text{Superficial area}_{\text{i}}$ can be obtained by edge matching and curvature calculation of the parametric equation of the surface. In this way, designers can easily modify any part of the model, and the whole model will be automatically updated to maintain its geometric and topological integrity. This not only improves the efficiency of modelling but also greatly enhances the accuracy and editability of the model. The platform adopts a convolutional neural network model to realize art feature detection (Figure 1).

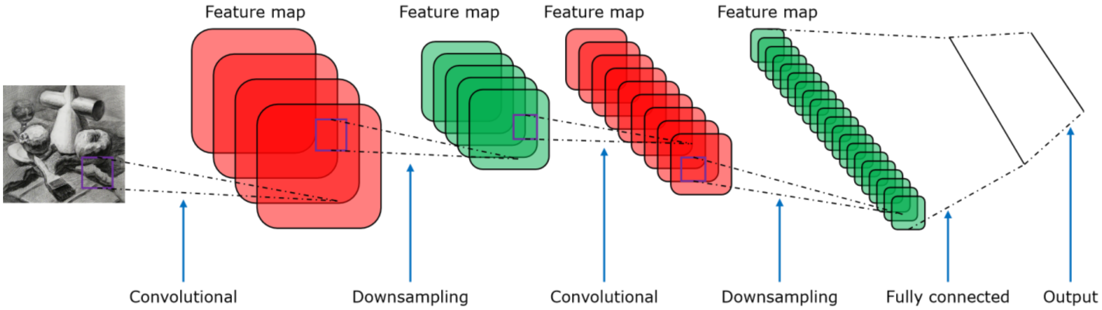


Figure 1: Feature detection model.

VR rendering algorithm: In order to truly show how the design works in the virtual environment, this article uses a rendering algorithm based on ray tracing. The algorithm can simulate the path of light and generate realistic light and shadow effects, thus improving the visual realism of VR scenes. Let $\vec{r}$ be the light from the viewpoint, $\vec{o}$ be a vertex of the object in the scene, and $\vec{d}$ be the vector from $\vec{o}$ to $\vec{r}$, then the intersection P of the light and the object can be calculated by the following formula:

$$P = \vec{o} + \frac{\vec{r} \cdot \vec{d}}{\left\|\vec{d}\right\|^2} \vec{d} \tag{4}$$

In the process of rendering, the ray tracing algorithm will trace the rays from the viewpoint and calculate their interaction with the objects in the scene. This includes complex optical phenomena such as light reflection, refraction, and shadow projection. The calculation formula for light reflection is as follows:

$$R = 2\vec{N} \cdot \ \vec{r} - \vec{P} \ \ \vec{r} - \vec{P} \ + \vec{r} \tag{5}$$

Where N is the normal vector at the intersection P and R is the direction vector of reflected light.

The calculation formula for light refraction is as follows:

$$\vec{j} = \frac{\vec{i}}{n_1} + \frac{n_2 - n_1 \ \vec{o}}{n_1 n_2} \sqrt{n_1^2 - \ \vec{i} \cdot \vec{o} \ ^2} \tag{6}$$

Where $\vec{r}$ is the direction vector of the incident light, $\vec{o}$ is the normal vector at the intersection P, $\vec{j}$ is the direction vector of refracted light, and $\vec{i}$ is the refractive index of two media. In order to calculate the shadow produced by light $\vec{r}$, the following formula can be used:

$$\text{Shadow} = \begin{cases} 0 \ \textit{if} \ \vec{s} \text{ such that } \vec{s} \in \text{Lightsource and } \vec{s} \to \vec{r} \text{ is not obstructed} \\ 1 \text{ Otherwise} \end{cases} \tag{7}$$

Where $\vec{s}$ is the position of the light source and $\vec{s} \to \vec{r}$ represents the path from the light source to the light $\vec{r}$? Through the above calculation, the algorithm can generate highly realistic lighting and shadow effects, which makes the VR scene more realistic visually and provides users with an immersive feeling.

Multi-modal sensing fusion algorithm: In order to fuse data from different sensing modes, this article designs a fusion algorithm based on a weighted average. The algorithm can dynamically assign weights according to the confidence of each perceptual mode so as to get more accurate perceptual results.

$$C_i = f\ D_i, R_i, S_i \tag{8}$$

$$W_i = \frac{C_i}{\sum_{j=1}^{N} C_j} \tag{9}$$

Where N is the total number of perceptual modes. Then, the weights are dynamically assigned according to these confidence levels to ensure that more reliable modes occupy a larger proportion of the fusion results. Accordingly, the algorithm can generate a perception result that integrates all modal information, thus improving the accuracy and robustness of perception.

Collision detection and optimization algorithm: In the VR environment, in order to ensure that the interaction between users and virtual objects is authentic, this article adopts a collision detection algorithm based on a bounding box. The bounding box algorithm simplifies the calculation of collision detection by creating a compact bounding volume (usually a simple geometric shape, such as a cube or sphere) for virtual objects. When the bounding boxes of two objects intersect, the algorithm will further carry out accurate collision detection.

Let P_1 and P_2 be the accurate collision detection points of two virtual objects and N_1 and N_2 are the normal vectors of the two objects at the collision points, then the accurate collision detection can be expressed by the following formula:

$$N_1 * \ P_2 - P_1 \ = 0 \tag{10}$$

If the above equation holds, it means that collisions occur at points P_1 and P_2.

Let F be the collision feedback force, m be the mass of the interactive object, v be the velocity of the interactive object and c be a constant related to the properties of the object, then the collision feedback force F can be expressed as:

$$F = -c * \ v - v_0 \tag{11}$$

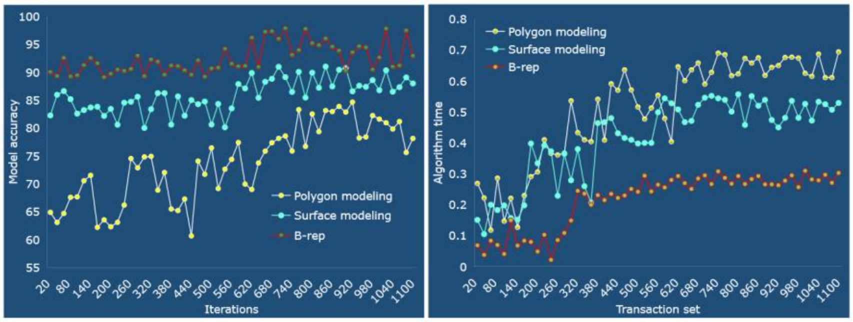
Where v_0 is the speed of the interactive object before the collision and c is a constant for controlling the strength and direction of feedback.

The algorithm can swiftly determine collisions between virtual objects and provide interactive feedback based on the outcomes. This not only enhances detection speed but also guarantees precise collision detection. Furthermore, by refining the bounding box's dimensions and shape, we can boost the algorithm's efficiency, ultimately delivering a smoother, more authentic VR interactive experience for users.

4.2 Algorithm Simulation Experiment

To validate the performance of the introduced algorithm, the section outlines the following simulation experiments:

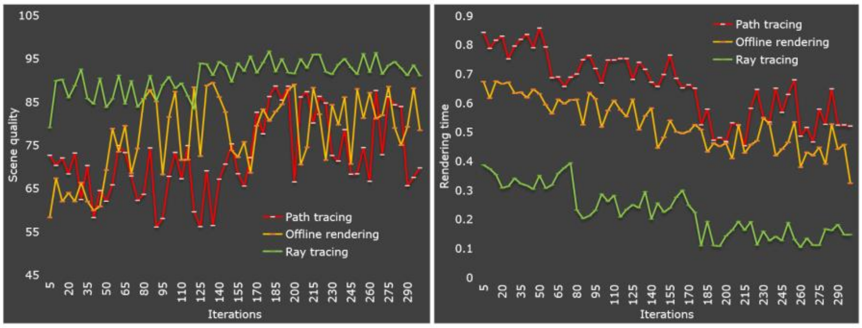
3D Modeling Assessment: Through a comparative analysis of the precision and efficiency of models produced by various modelling algorithms, the superiority of our chosen 3D modelling technique is demonstrated, with results illustrated in Figure 2.



a. Model accuracy b. Model efficiency

Figure 2: Accuracy and efficiency of models generated by different modelling algorithms.

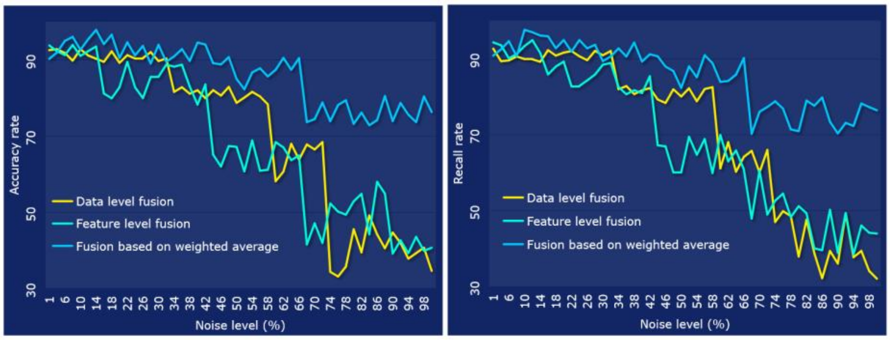
VR rendering experiment: Under the same hardware conditions, compare the scene quality and rendering speed generated by different rendering algorithms and verify the efficiency of our adopted rendering algorithm. The result is shown in Figure 3.



a. Scene quality b. Rendering speed

Figure 3: Scene quality and rendering speed generated by different rendering algorithms.

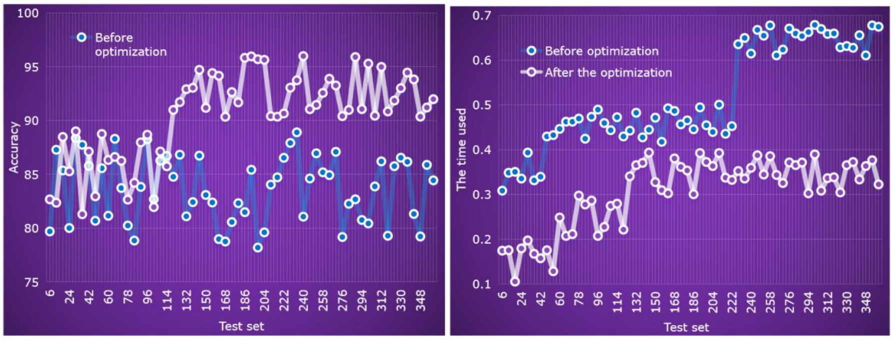
Multi-modal sensing fusion experiment: By simulating the data input of different sensing modes, the accuracy and robustness of the multi-modal sensing fusion algorithm are verified. The performance of the algorithm under different noise interferences will be evaluated, and the results are shown in Figure 4.



a. Accuracy rate b. Recall rate

Figure 4: Performance of the algorithm under different noise interferences.

Collision detection and optimization experiment: simulate the interaction process between users and virtual objects in a virtual environment, and test the accuracy and real-time performance of collision detection algorithm. Furthermore, we will compare the performance improvement of the algorithm before and after optimization, and the result is shown in Figure 5.



a. Accuracy of algorithm b. Real-time performance of algorithm

Figure 5: Accuracy and real-time performance of collision detection algorithm.

The results of the simulation experiment and its analysis are as follows:

Experimental results of 3D modelling: By comparing the experimental data, this article finds that the 3D modelling algorithm adopted is superior to other comparison algorithms in model accuracy and modelling efficiency. This is due to the accurate geometric representation and efficient topology construction method in this algorithm.

Experimental results of VR rendering: The experimental results indicate that the rendering algorithm employed in this article boasts swift rendering speeds without compromising the quality of the rendered output. This proves the effectiveness of the ray tracing algorithm in VR rendering.

Experimental results of multi-modal sensing fusion: By simulating the data input of different sensing modes, this article finds that the multi-modal sensing fusion algorithm can accurately fuse the information from different modes and improve the accuracy and robustness of sensing. Even under noise interference, the algorithm can still maintain good performance.

Experimental results of collision detection and optimization: The experimental results show that the collision detection algorithm adopted in this article can accurately detect collisions between virtual objects and provide timely interactive feedback. After optimization, the algorithm improves the real-time performance while maintaining accuracy and provides users with a smoother VR interactive experience.

5 PLATFORM IMPLEMENTATION AND TESTING

5.1 Platform Development Environment and Tool Selection

Considering the complexity of the platform and the integration requirements of multi-modal sensing technology, this article adopts C++ as the main programming language because it has excellent performance and flexibility in the underlying system programming and graphics rendering. Furthermore, we also use Python to deal with some advanced functions, such as data analysis and user interface design, because Python is widely used in the fields of data processing and machine learning, and its syntax is concise and easy to understand, which can speed up development.

In terms of development tools, this article chooses Visual Studio as the main integrated development environment because it supports multiple programming languages and provides powerful debugging and testing tools. Furthermore, Unity 3D is used as the development environment of the VR module. Unity 3D is a professional game and VR development platform, which provides rich graphics rendering and physical engine functions, which can well meet our needs in VR display.

5.2 Platform Function Implementation Details

The function realization of the platform is mainly divided into the following parts:

The Art CAD Design Module utilizes an open-source CAD library and incorporates a user-defined data structure and algorithm to achieve precise 2D and 3D graphic renderings. Through an intuitive interface, users can design graphics and view real-time previews of their designs.

For the VR Display Module, we've leveraged the Unity 3D engine to craft a highly engaging VR setting. Designs from the CAD module can be effortlessly transferred to this virtual world, allowing users to freely explore and engage with their creations in a virtual space.

The Multi-modal Perception Integration Module integrates sensors and devices like VR helmets and gloves. Through specialized interfaces and data processing methods, it enables the seamless fusion of visual, auditory, and tactile sensations. This offers users an immersive experience akin to being in a genuine design setting.

Collaboration and sharing function: The platform supports multi-person online collaborative editing. This article designs an efficient real-time synchronization mechanism to ensure that multiple users can edit the same design project at the same time without conflict. Users can also share their design works on social media or online platforms to communicate and discuss with others.

5.3 Platform Testing and Performance Evaluation

After the platform is developed, this section conducts comprehensive tests to ensure its stability and performance. The test mainly includes the following aspects:

Unit test: write detailed test cases for each module of the platform to ensure that each function can work as expected. Integration testing: testing the interaction and dependence between modules to ensure that data can be transmitted and processed correctly between different modules. Compatibility test: tests were conducted on different operating systems and devices to ensure that the platform has good compatibility. The above test results are shown in Table 1.

Test type	*Test module/environment*	*Number of test cases*	*Pass quantity*	*Number of failures*	*Passing rate*	*Remarks*
Unit testing	Module a	25	25	0	100%	All functions are as expected
	Module b	30	29	1	96.7%	Boundary condition test case failed
	Module c	20	20	0	100%	Pass all
Integration testing	Module A interacts with module B.	15	15	0	100%	Data transmission and processing are correct
	Module b interacts with c	12	12	0	100%	Data transmission and processing are

						correct
Compatibility test	Windows 10	10	10	0	100%	Good compatibility
	macOS Catalina	10	10	0	100%	Good compatibility
	Android 10	10	9	1	90 %	There is a small problem with one equipment model
	iOS 14	10	10	0	100%	Good compatibility

Table 1: Platform detailed test results.

Performance test: This section simulates the scene of online editing by multiple users at the same time and tests the response time and resource consumption of the platform. The results are shown in Figures 6 and 7.

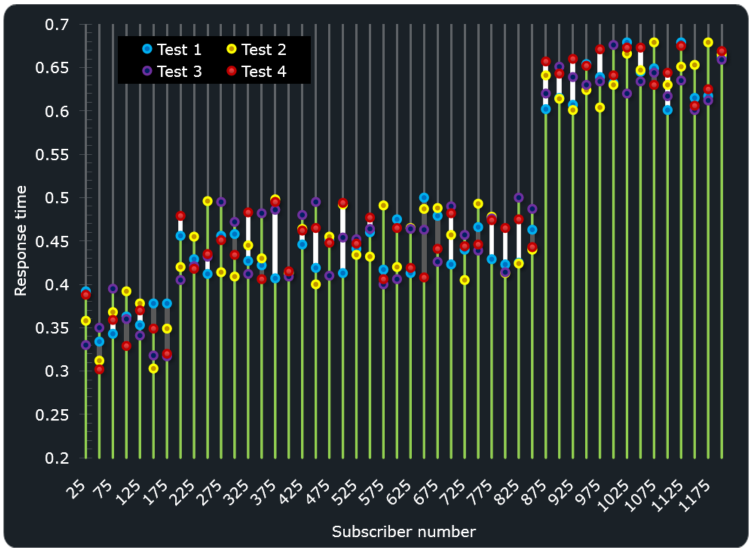


Figure 6: Response time of the platform.

The results show that the response time of the platform is kept within a reasonable range even when multiple users edit at the same time, which is very important to ensure the user experience. Furthermore, the platform performs well in resource consumption, and can reasonably allocate and utilize system resources under different loads to ensure the stable operation of the platform.

User acceptance test: invite real users to test and collect their feedback and suggestions in order to further optimize the user experience. The results are shown in Table 2.
Note: The satisfaction rating range is 1-10, with 10 being very satisfied and one being very dissatisfied.

Generally speaking, users have a high rating on the platform. Through the above series of tests, this article verifies the stability and performance of the co-creation platform and makes full preparations for the official release.

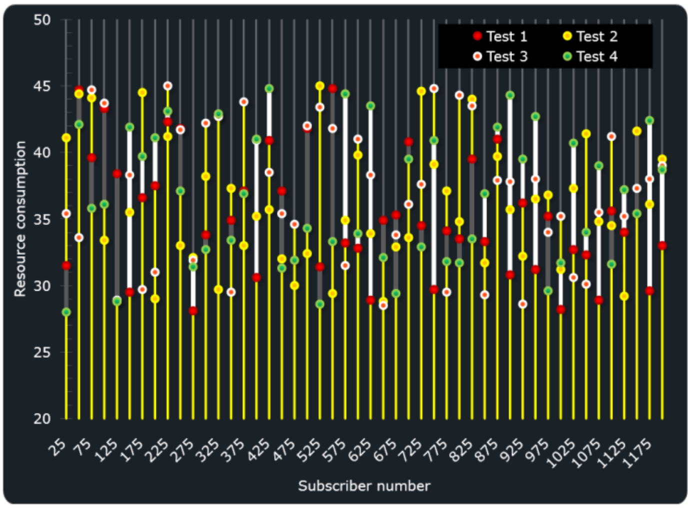


Figure 7: Resource consumption of the platform.

User number	*Test module*	*Number of test cases*	*Satisfaction score (1-10)*	*Main feedback and suggestions*
User 001	Module A	5	8.9	The operation process is smooth, but the interface design can be optimized.
User 002	Module B	4	8.4	In some operations, the response time is long and performance needs to be optimized.
User 003	Module C	3	9.4	The function is complete, and the user experience is good. It is suggested to add operation instructions.
User 004	Module D	6	7.9	Some functions are complicated to operate, and the process needs to be simplified.
User 005	Overall experience	-	9.1	The overall experience of the platform is good, but I hope to add more personalization options.

Table 2: User acceptance test results and feedback.

6 CONCLUSIONS

In this study, a co-creation platform of art CAD and VR combined with multimodal perception was successfully constructed. Through in-depth user demand analysis, the function and design goal of the platform is defined, and a highly interactive and realistic VR environment is designed. In terms of algorithm implementation, this article innovatively combines 3D modelling, VR rendering and multimodal perception technology, providing users with an immersive creative experience. The results show that these algorithms are effective and the platform is practical. Users can design, preview, interact and collaborate with many people on this platform.

Although some achievements have been made in this article, there are still limitations in the research. For example; The integration of multi-modal sensing technology needs to be further improved, especially in improving sensing accuracy and response speed; The performance and stability of the platform need to be continuously optimized in practical applications to cope with the larger-scale concurrent operation and data transmission of users; The friendliness and ease of use of the user interface also need to be iteratively improved according to user feedback. In the future, we will continue to explore the deeper application of multi-modal sensing technology in the co-creation platform, such as improving the intelligent level of user perception through AI algorithms. Furthermore, we will pay attention to the latest developments in VR technology and integrate new technologies and methods into the platform in time to enhance the user experience.

7 ACKNOWLEDGEMENT

Jilin Normal University Doctoral Research Launch Project "Calibration and Precision Guidance of Audience Participation in Major Emergencies of Network Public Opinion" (Grant No. 2023042).

Dazhuang Wang, https://orcid.org/0009-0005-5270-6392
Xu Zhao, https://orcid.org/0000-0002-0810-5590

REFERENCES

[1] Akilan, T.; Wu, Q.-J.; Safaei, A.; Huo, J.; Yang, Y.: A 3D CNN-LSTM-based image-to-image foreground segmentation, IEEE Transactions on Intelligent Transportation Systems, 21(3), 2019, 959-971. https://doi.org/10.1109/TITS.2019.2900426
[2] Bansal, M.; Mishra, A.; Sharma, A.: Multiple scaling Fuzzy-PSO watermarking scheme for gray-scale and colored images, Multimedia Tools and Applications, 81(11), 2022, 15219-15248. https://doi.org/10.1007/s11042-022-12526-7
[3] Canet, S.-M.; Guljajeva, V.: Dream painter: exploring creative possibilities of ai-aided speech-to-image synthesis in the interactive art context, Proceedings of the ACM on Computer Graphics and Interactive Techniques, 5(4), 2022, 1-11. https://doi.org/10.1145/3533386
[4] Cedillo, H.-M.; Cedillo, H.-A.; Garcia, U.-F.-J.: Improving dft-based image watermarking using particle swarm optimization algorithm, Mathematics, 9(15), 2021, 1795. https://doi.org/10.3390/math9151795
[5] Fan, X.; Li, J.: Artificial intelligence-driven interactive learning methods for enhancing art and design education in higher institutions, Applied Artificial Intelligence, 37(1), 2023, 2225907. https://doi.org/10.1080/08839514.2023.2225907
[6] Gao, Y.; Guo, Y.; Lian, Z.; Tang, Y.; Xiao, J.: Artistic glyph image synthesis via one-stage few-shot learning, ACM Transactions on Graphics (TOG), 38(6), 2019, 1-12. https://doi.org/10.1145/3355089.3356574
[7] Jeon, M.: Hands-on real-time interactive art using particle system, TECHART: Journal of Arts and Imaging Science, 8(4), 2021, 5-8. https://doi.org/10.15323/techart.2021.11.8.4.5
[8] Li, L.: Digital art design system based on big data technology and interactive virtual technology, IEEE Consumer Electronics Magazine, 12(2), 2021, 49-55. https://doi.org/10.1109/MCE.2021.3133267
[9] Meinecke, C.; Hall, C.; Jänicke, S.: Towards enhancing virtual museums by contextualizing art through interactive visualizations, ACM Journal on Computing and Cultural Heritage, 15(4), 2022, 1-26. https://doi.org/10.1145/3527619
[10] Ornes, S.: Virtual reality helps shape and respond to artistic impulses, Proceedings of the National Academy of Sciences, 119(31), 2022, e2210323119. https://doi.org/10.1073/pnas.2210323119
[11] Pallasena, R.-K.; Sharma, M.; Krishnaswamy, V.: A study of interaction, visual canvas, and immersion in AR design: a DSR approach, AIS Transactions on Human-Computer Interaction, 14(3), 2022, 390-425. https://doi.org/10.17705/1thci.00173

[12] Wang, X.; Cai, L.: Application of artificial intelligence in 6G internet of things communication in interactive installation art, International Journal of Grid and Utility Computing, 13(2-3), 2022, 195-203. https://doi.org/10.1504/IJGUC.2022.124401
[13] Wang, Y.; Hu, X.-B.: Three-dimensional virtual VR technology in environmental art design, International Journal of Communication Systems, 35(5), 2022, e4736. https://doi.org/10.1002/dac.4736
[14] Yang, S.; Wang, W.; Liu, J.: Te141k: artistic text benchmark for text effect transfer, IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(10), 2020, 3709-3723. https://doi.org/10.1109/TPAMI.2020.2983697
[15] Zhang, B.; Rui, Z.: Application analysis of computer graphics and image-aided design in art design teaching, Computer-Aided Design and Applications, 18(S4), 2021, 13-24. https://doi.org/10.14733/cadaps.2021.S4.13-24