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

Optimization Algorithm for Multimedia Interaction of Computer-Aided Design Models in Virtual Reality Environments

Jin Wang

School of Marxism, Chengdu Polytechnic, Chengdu 610017, China, ince5544@163.com

Corresponding author: Jin Wang, ince5544@163.com

Abstract. The development of virtual reality technology has broadened the path of CAD model presentation and accelerated the development speed of CAD models from two-dimensional plane to three-dimensional space. However, the multimedia interaction of CAD models in a virtual reality environment has problems such as poor interaction sense, high operation complexity, inaccurate 3D sitting, etc. Therefore, this paper introduces the gesture manipulation model and 3D object tracking model in the multimedia interaction system of CAD models so that CAD models in different virtual reality environments can be manipulated by limited gestures to complete more CAD model commands and to enhance the effect of the presentation of CAD models in virtual reality space. model rendering effect in virtual reality space. The experimental results show that in the virtual reality environment, the 3D object tracking model in this paper has better tracking, stability, and timeliness compared with other models, which can effectively improve the accuracy of the CAD model's position matching and enhance its rendering accuracy and effect. In addition, the experimental results show that the gesture operation can reduce the time while ensuring accuracy within 95% compared with the keyboard and mouse operation. Most of the users expressed high recognition of the experience and accuracy of the optimized CAD model's multimedia interaction.

Keywords: Virtual Reality Environment; CAD Model; Multimedia; Interaction Optimization
DOI: https://doi.org/10.14733/cadaps.2024.S25.311-325

1 INTRODUCTION

CAD software can accurately complete the conversion between different figures based on the existing data in a short period to present the CAD model so that the design effect, details, and the degree of completion of the design through a more intuitive and scientific way to present, so CAD models in many areas of the design of the field have been widely used. Hybrid reality technology combines the real world with the virtual world, providing users with a more intuitive and efficient interactive experience. In industrial applications, mixed reality technology not only improves work efficiency but also reduces operational risks, becoming an important force in promoting industrial modernization. In industrial human-machine systems, mixed reality technology enhances the interaction experience between users and machines by providing rich visual, auditory, and tactile feedback. With the help of hybrid reality devices such as head-mounted displays and gesture recognition systems, operators can intuitively view and operate virtual models, achieving seamless integration with machines. This interactive approach not only improves the accuracy and efficiency of operations but also reduces the difficulty of operations, allowing more personnel to participate in complex industrial tasks. By constructing highly realistic virtual scenes, mixed reality technology enables operators to conduct practical operations and simulation training in virtual environments [1].

Although CAD models are constructed based on three-dimensional data of the target object, their previous presentation is mostly done in a two-dimensional plane, which not only affects the overall visual effect of the presentation of the CAD model but also has the problems of inaccurate data and large discrepancies between the three-dimensional model and the actual model. Virtual reality (VR) and digital twin technology are gradually integrating into the field of industrial robots, bringing revolutionary changes to human-computer interaction. The application of these two technologies not only improves the operational efficiency and accuracy of industrial robots but also provides users with a more intuitive and convenient operating experience. Burghardt et al. [2] explored the application and advantages of utilizing virtual reality and digital twins for human-machine interaction in industrial robots. Digital twin technology can collect real-time operational data of industrial robots and provide feedback to users. Users can make real-time adjustments to the robot's operations based on this data to achieve better operational results. By analyzing digital models, potential failures of industrial robots can be predicted and maintenance can be carried out in advance. This approach reduces the probability of malfunctions and improves the reliability and stability of the robot.

The traditional WIMP (Window, Icon, Menu, and Pointer) interaction method has been widely used in the field of CAD assembly modeling, but its limitations are also becoming increasingly prominent. In recent years, immersive natural finger interaction technology has gradually emerged, providing users with a more intuitive and natural interaction experience. Fechter et al. [3] conducted a comparative evaluation of the application of WIMP and immersive natural finger interaction in CAD assembly modeling through user research, in order to provide a reference for the development and application of related technologies. Immersive natural finger interaction is a new human-computer interaction method based on virtual reality technology. Through wearable devices, users can directly use their fingers to perform assembly modeling operations in a virtual environment. This interaction method is more intuitive and natural, which helps to improve user operation efficiency and accuracy. Immersive natural finger interaction has significant advantages over WIMP interaction. When users use immersive natural finger interaction to complete tasks, the time required is significantly less than WIMP interaction. With the development of virtual reality technology, multimedia technology, information communication technology, and human-computer interaction technology, CAD model presentation gradually shifted from a two-dimensional plane to a three-dimensional virtual space. In some application areas, CAD software applications are also more integrated and intelligent, the presentation of CAD models is no longer satisfied with the keyboard and mouse for manipulation, and the operation mode is more inclined to think about the transformation of simplicity and naturalness, to enhance the user's sense of interactivity and experience of the operation of CAD models.

With the rapid development of virtual reality (VR) technology, its applications in various fields are becoming increasingly widespread. In the field of 3D computer-aided design (CAD) modeling, the introduction of virtual reality technology provides designers with a more intuitive and efficient way of interaction. Especially the application of interactive gestures and voice inspiration greatly improves the efficiency and experience of designers in 3D CAD modeling in virtual environments. Khan and Tunçer [4] explored the application and advantages of interactive gestures and speech inspiration in 3D CAD modeling in the conceptual design of virtual reality environments. The grasping and movement of virtual objects can be achieved through the pinching action of fingers. The rotation action of the palm can achieve the rotation of virtual objects. The sliding action of fingers can achieve scaling of virtual objects, etc. These interactive gestures not only improve the accuracy and efficiency of modeling but also enable designers to immerse themselves more in virtual environments, enhancing the creativity and experience of design. At the same time, many CAD models, in addition to being presented in 3D virtual space, need to be presented in the fusion of virtual space and real space, which requires the introduction of augmented reality technology to accomplish the superposition and supplementation of information in the virtual reality world, so as to achieve the desired visual and sensory effects. In this process, the degree of alignment between virtual space and real space affects the accuracy of the CAD model position, the authenticity of the presentation effect, and the coherence and accuracy of human-computer interaction, so how to present CAD models more accurately and how to improve the sense of multimedia interactive experience of CAD models in virtual reality environment has become one of the most important directions in the development and research of CAD models. As a core component of the immersive VR experience, the design and usability of a 3D user multimedia interaction interface directly determine the interaction effect and experience quality of users. Kharoub et al. [5] explored the design principles and usability optimization methods of 3D user multimedia interaction interfaces for immersive VR, in order to provide valuable references for research and practice in related fields. The interface design should conform to the user's cognitive habits, enabling users to intuitively understand the structure and function of the interface. Reduce user learning costs and improve operational efficiency through reasonable layout, icons, and textual explanations. The interface design should maintain a consistent style, and follow unified visual norms and interaction logic. This helps users form stable psychological expectations during the operation process and reduces the possibility of misoperation. Explore new interaction methods based on the characteristics of VR technology, such as gesture recognition and voice control. This can not only improve the convenience of user operation but also bring users a more natural and realistic interactive experience.

Based on the above problems, this paper introduces gesture manipulation technology and a 3D object tracking algorithm to optimize the CAD model multimedia interaction system in a virtual reality environment and improve the experience of CAD model multimedia human-computer interaction through the scenario-based gesture manipulation model. At the same time, the 3D object tracking algorithm is used to optimize the rendering effect of CAD model multimedia interaction in different virtual reality environments to ensure the richness of the corresponding data sets, the accuracy of the baseline poses of CAD models with different materials, and the diversity of the rendering data.

The innovations of this paper are as follows:

First of all, this paper introduces context-based CAD model gesture manipulation technology in the CAD model multimedia interactive system to realize the same gesture expressing different meanings in different contexts, i.e., to complete as many CAD model commands as possible under the limited number of gestures.

Secondly, 3D object tracking technology can improve the rendering effect of CAD models in virtual reality space, more realistically reflecting the effect of CAD models of different materials, and ensuring the accuracy of the model's baseline posture.

Finally, the combined application of CAD model gesture manipulation technology and 3D object tracking algorithm can effectively optimize the effectiveness and experience of CAD model multimedia interaction and improve the accuracy of human-computer interaction operation.

2 STATUS OF DEVELOPMENT OF CAD MODELS

Kozinets [6] explored new methods and their advantages in service experience research in the context of immersive network virtual reality, augmented reality, and human-computer interaction. By constructing virtual service scenarios, researchers can simulate different service environments and contexts in order to observe and analyze user behavior reactions and experiential feelings under various conditions. Immersive network virtual reality and augmented reality technology allow users to interact in real-time with the virtual environment and receive immediate feedback. This interactive approach enables researchers to understand user needs, thereby optimizing service design more accurately. Through human-computer interaction technology, personalized service content and methods can be recommended to users based on their preferences and behavioral data. This personalized service can enhance user satisfaction and loyalty, and enhance the effectiveness of the service experience. In today's digital age, the application of multimedia interaction design in children's emotional education is becoming increasingly widespread. The concept of social presence, which simulates real-life social scenes through technological means to enhance the effectiveness and quality of children's emotional education, has become a hot research topic. Ma et al. [7] explored the optimization algorithm for multimedia interaction design in children's emotional education based on social presence, in order to provide new research perspectives and practical guidance for related fields. In terms of interaction fluency, optimization algorithms can focus on the interaction response speed between users and the virtual environment. By reducing latency and increasing frame rate, it is possible to ensure timely and accurate feedback on user actions in virtual space, thereby improving the smoothness of interaction.

Immersive virtual reality training based on multiple evaluations, combined with multimedia human-computer interaction technology, not only improves training effectiveness but also enhances user engagement and learning experience. Makransky et al. [8] explored the advantages of multimedia human-computer interaction for immersive virtual reality training based on multiple evaluations. Multiple evaluation refers to the comprehensive and objective evaluation of the performance of trainees through the comprehensive application of various evaluation methods and techniques during the training process. In immersive virtual reality training, multiple evaluations can achieve a precise evaluation of training effectiveness by collecting and analyzing multidimensional information such as user cognitive performance in the virtual environment. Through multimedia human-computer interaction, users can interact with the virtual environment through various methods such as gestures, voice, and eye movements, achieving more natural and intuitive operations. This interactive approach not only enhances user engagement and immersion but also helps to enhance the cognitive ability and operational skills of trainers. Virtual reality (VR) technology has gradually penetrated various fields of manufacturing, bringing revolutionary changes to traditional manufacturing models. Especially in the design of human-machine workspaces, immersive and collaborative human-machine interaction has become an important research direction. Malik et al. [9] explored the application of virtual reality in the design of human-machine workspaces in the manufacturing industry, as well as the advantages and challenges of immersive and collaborative human-machine interaction. Virtual reality technology enables designers and operators to simulate product design and manufacturing processes in virtual spaces by constructing highly realistic virtual environments. Immersive human-computer interaction is one of the important applications of virtual reality technology. In the design of human-machine workspaces in the manufacturing industry, immersive human-machine interaction enables operators to fully immerse themselves in the virtual environment and interact in real time with virtual objects. Electronic learning, with its flexible and convenient characteristics, provides learners with rich and diverse learning resources and methods. However, in the process of e-learning, the standardization and interactivity of multimedia content are crucial for improving learning effectiveness. Mwambe et al. [10] studied automated interactive technologies that support multimedia content specification for e-learning. Standardized multimedia content can provide learners with clear and explicit learning information, reduce learning difficulty, and improve learning outcomes. At the same time, standardized multimedia content also helps to enhance the learning interest and enthusiasm of learners, and enhance their initiative and sustainability in learning. In the process of e-learning, automated interaction technology can intelligently adjust the presentation and interaction methods of multimedia content based on the learning situation and needs of learners, providing a personalized learning experience. This interactive approach not only enhances learner engagement but also provides timely feedback on their learning outcomes, helping learners better grasp knowledge and skills.

Virtual reality technology provides artists with a brand-new creative platform and display methods. Multi-source information art painting fusion interactive 3D dynamic scene virtual reality technology aims to integrate multiple information sources into art painting and combine interactive 3D dynamic scenes to create a richer and more vivid art experience. Pan and Deng [11] conducted in-depth research on the application of this technology. Multi-source information art painting refers to the integration of information sources from different channels and forms into painting creation, in order to enrich the content and expression of the picture. These information sources can include various types such as images, videos, audio, text, etc. Through virtual reality technology, artists can present this multi-source information in a three-dimensional form to the audience, making the picture more vivid and three-dimensional. In practical applications, artists can use professional 3D modeling software to convert various information sources into 3D models or textures and embed them into virtual scenes. Wearable devices and augmented reality (AR) technology are gradually integrating into people's daily lives, bringing unprecedented convenience and experience to our lives. The continuous progress of deep learning technology also provides strong technical support for the application of wearable augmented reality. Park et al. [12] explored intelligent interactive task assistance based on deep learning in wearable augmented reality and analyzed its applications and advantages in various fields. Through image recognition technology, the system can recognize the user's gestures and actions, thereby understanding the user's intentions and implementing corresponding operations. Meanwhile, speech recognition technology can recognize the user's voice commands and achieve voice control. In addition, natural language processing technology can help systems understand user text input and achieve more intelligent interaction. Augmented reality technology provides users with a more intuitive and vivid interactive experience by overlaying virtual information in the real world. In the field of multimedia interactive assembly, augmented reality technology can display real-time virtual models, operation prompts, and other information during the assembly process, helping operators better understand and execute assembly tasks. Digital twin technology achieves real-time monitoring, analysis, and optimization of actual systems by constructing virtual models that correspond to them. In the field of multimedia interactive assembly, digital twin technology can build virtual models of the assembly process, simulate the assembly process, predict assembly results, and provide decision support for operators. Specifically, augmented reality technology can present assembly steps, part positions, and other information in the form of three-dimensional images to the operator's field of vision through devices such as head-mounted displays. In this way, operators can have a more intuitive understanding of the assembly process, reduce operational errors, and improve assembly efficiency [13].

The virtual reality technology under multimedia interaction of CAD models not only improves design efficiency but also brings users a brand-new design experience. Sagnier et al. [14] explored the technology acceptance model of virtual reality technology under multimedia interaction of CAD models, in order to provide theoretical support for the promotion and application of this technology. This technology enables users to intuitively observe, operate, and modify CAD models in a virtual environment, thereby improving the accuracy and efficiency of design. The immersion brought by virtual reality technology is an important factor affecting user acceptance. The stronger the immersion, the easier it is for users to immerse themselves in the virtual environment, thereby better understanding and operating CAD models. The use of multimedia interaction methods makes the interaction between users and CAD models more natural and intuitive. The improvement of interactivity helps to enhance user engagement and experience, thereby increasing the acceptance of technology. Human motion recognition technology plays a crucial role in achieving natural interaction between users and the virtual world, enhancing user immersion and experience. Support Vector Machine (SVM), as a powerful classification algorithm, has broad application prospects in the field of human motion recognition. Zhang et al. [15] explored human motion recognition technology and its applications in virtual reality art media interaction environments based on SVM. Support Vector Machine is a machine learning method based on statistical learning theory, which achieves data classification by finding an optimal hyperplane. In human motion recognition, SVM can utilize the features of human motion data for classification, thereby achieving recognition of different motion patterns. By identifying the user's dance movements, real-time interaction and synchronized performance with virtual characters can be achieved. Users can control the dance movements of virtual characters based on their own movements and rhythms, creating unique artistic effects.

In summary, the development of CAD models from two-dimensional gradually changed to three-dimensional, from the plane changed to three-dimensional, in the process of its development, most of the research focused on the study of the effect of three-dimensional presentation of CAD models, the lack of multimedia interaction in the virtual reality environment in the effectiveness of the study of the accuracy and so on, so this paper has a certain practical significance and significance of the application of the study.

3 MULTIMEDIA INTERACTIVE MODELING OF CAD MODELS IN VIRTUAL REALITY ENVIRONMENTS

3.1 Contextualized CAD Model Multimedia Interactive Gesture-Based Operating System

The context-based CAD model is a concrete description of the interaction state between man and machine during the modeling process, which contains three levels: scenario dimensions, scenario items, and scenario values, which can reflect the current user control habits, the CAD model displays state, and the state in which it is located, and provide basic information for the gesture operating system. The relevant commands of the CAD model include generation commands, operation commands, and editing commands, and the relevant commands can be recognized through the geometric model manipulation gestures. CAD model-related commands include the generate command, operate command, and edit command, which can be recognized by geometric model manipulation gestures. According to the temporal characteristics, manipulation gestures can be divided into static gestures, which are generally composed of simple gestures with unchanged shapes in multimedia interactions, and dynamic gestures, which form special shapes or paths in motion trajectories, with the former having no constraints on spatial positions and directions, and the latter having no constraints on the shapes of gestures in the process of motion. Some of the static and dynamic gestures are shown in Figure 1.

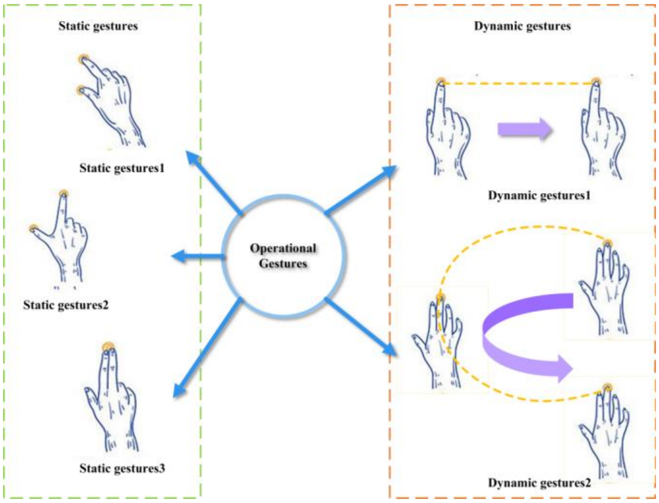

Figure 1: Partial static gesture and dynamic gesture.

The CAD model for gesture recognition is implemented by the HMM algorithm for gesture observation based on the observation of a sequence of vectors, where all types of states are described by a probability density distribution. The HMM contains five variables, as shown in Equation (1) for the description of the hidden state variables:

$$Y = \{y_1, y_1, ..., y_N\} \tag{1}$$

The observed state variables are described as shown in Equation (2):

$$G = \{g_1, g_1, ..., g_M\} \tag{2}$$

If an observation is made at the moment t, the observation is expressed as $O_t = \{g_1, g_1, ..., g_M\}$.

The state transfer matrix variables are described as:

$$\begin{cases} A = \{a_{ij}\} \\ a_{ij} = P(q_{t+1} = Y_j \middle| q_t = Y_i) \\ \sum_{j=1}^{L} a_{ij} = 1 \; 1 \le i \;\; j \le N \end{cases} \tag{3}$$

Where the actual state t is currently in is denoted as q_t, $t = 1,2,\ldots$.

The probability transfer matrix representation of the observations is shown in (4):

$$\begin{cases} B = \{b_j(k)\} \\ b_j(k) = P(v_k, t \middle| q_t = Y_j) \; 1 \le j \le N \;\; 1 \le k \le M \end{cases} \tag{4}$$

The initial probability distribution variable is denoted as:

$$\begin{cases} \pi = \{\pi_n\} \\ \pi_n = P(q_n = Y_n) 1 \le n \le N \end{cases} \tag{5}$$

HMM-based gesture recognition requires first defining the sequence of gesture observations and determining the corresponding learning model parameters during the gesture modeling process. The gesture recognition module calculates the conditional probability of the observed sequence based on the corresponding parameters to generate the conditional probability of the sequence, and the similar values obtained are the results of the most similar gesture types.

In the virtual reality environment, CAD model gestures are based on the combination of basic elements of different situations to constitute a specific interaction of the scenario interaction specific instructions, in order to improve the accuracy of the instruction flexibility mapping relationship, this paper adopts the decision tree algorithm based on information entropy and information gain to obtain the corresponding mapping model.

Let the number of sample categories in the dataset C be M, the probability of the data contained in the category with sequence number i over all data is denoted as p_i, and the corresponding entropy calculation formula is shown in (6):

$$info(C) = -\sum_{i=1}^{M} p_i lb(p_i) \tag{6}$$

If the dataset is categorized according to a certain feature it will reduce its uncertainty, and the entropy obtained at this time will also be reduced, let the feature be T, the number of categories it contains is recorded N, and its calculation formula is shown in (7):

$$info_T(C) = -\sum_{j=1}^{M} \frac{\left|C_j\right|}{C} \times info(C_j) \tag{7}$$

This gives the difference in entropy reduction before and after categorization as shown in equation (8):

$$Gain(T) = info(C) - info_T(C) \tag{8}$$

The result obtained with the above equation is the information gain.

A schematic diagram of the CAD model multimedia interactive gesture operating system is shown in Figure 2.

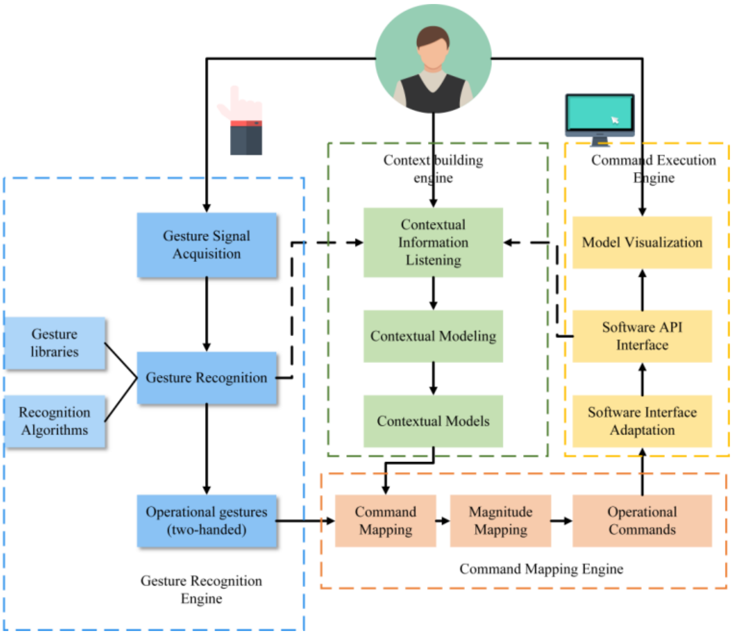

Figure 2: Schematic diagram of CAD model multimedia interactive gesture operating system.

3.2 Optimized Model for Multimedia Interaction of CAD Models Based on 3D Object Tracking

The application of 3D registration can make the CAD static model in the virtual reality environment accurately localized in the 3D registration needs to transform the coordinates of the model accordingly, set the origin of the coordinate system as the model's identification coordinates, the positive direction is recorded as the direction of the horizontal axis, and the relationship between the world coordinates and identification coordinates is coincident. The relative position and direction between the viewer and the camera are known, and the model's coordinates are transformed. The transformation relationship between the world coordinate system, and the camera coordinate system can be expressed by the rotation matrix and translation vector, i.e., set the chi-square coordinate system of a certain point under the two coordinate systems $(X_w,Y_w,Z_w,1)^T,(X_c,Y_c,Z_c,1)^T$, and the transformation relationship is shown in Equation (9):

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} R & W \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} = C \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \tag{9}$$

The orthogonal 3×3 matrix is denoted as R, which is the rotation matrix between the camera and marker coordinates, and the 3D translation vector is denoted as W, which is the translation matrix between the two.

Let a point in the camera coordinate system be $P(x_c,y_c)$, its projection in the screen coordinate system is denoted as $P'(m,n)$ the image pixel coordinates are (m,n) , and the physical size of each pixel in the direction of the x and y axes are dx,dy, respectively, as shown in Equation (10) for the transformed relational equations between the camera and screen coordinate systems:

$$Z_c\begin{bmatrix}m\\n\\1\end{bmatrix}=\begin{bmatrix}a_x&0&m_0&0\\0&a_y&n_0&0\\0&0&1&0\end{bmatrix}\begin{bmatrix}X_c\\Y_c\\Z_c\\1\end{bmatrix}=P\begin{bmatrix}X_c\\Y_c\\Z_c\\1\end{bmatrix} \tag{10}$$

Among others, $a_x=f/dx, a_y=f/dy$.

This can be obtained after derivation:

$$Z_c\begin{bmatrix}m\\n\\1\end{bmatrix}=PC\begin{bmatrix}X_m\\Y_m\\Z_m\\1\end{bmatrix}=T\begin{bmatrix}X_m\\Y_m\\Z_m\\1\end{bmatrix} \tag{11}$$

In multimedia interaction, the virtual reality environment in which the CAD model is located has more uncertain information, and the dynamics and interactivity of the scene are stronger, which requires the introduction of 3D object tracking technology based on early 3D registration to improve the problem of alignment between the virtual space and the real space. In this paper, we choose to realize tracking of 3D object feature points based on PTAM, and its principle is based on the feature recognition results in the object image combined with keyframe technology to realize the tracking of the object, image feature recognition can effectively avoid the loss of the tracking target, while the keyframe technology can effectively deal with the problem of object jitter in terms of rendering the object. In this algorithm, all the sample keyframes in the dataset will be divided into subsets of smaller size, and each camera image and subset are realized to match, the constructed subset is shown in (12):

$$\begin{cases}S_1=\{k_1,k_2,\ldots,k_j\}\\S_2=\{k_{j+1},k_{j+2},\ldots,k_{2j}\}\\\vdots\\S_{N/f}=\{k_{f-con\frac{N}{f}+1},\ldots,k_N\}\end{cases} \tag{12}$$

where the keyframes are denoted as k_j , where the number of keyframes capable of frame rate processing is denoted as f .

In order to realize the comprehensive evaluation of the algorithm, this paper chooses three metrics to evaluate the algorithm. The first one is the average edge distance (AED) of the 3D object model, which is a measure of the distance of the edge contour under the baseline pose and the estimated pose, and the number of image sequences is set to be i , and its calculation formula is shown in (13):

$$AED=\frac{1}{C}\sum_{k=0}^{C-1}\left\|\pi(K(R_iX_k+t_i))-\pi(K(R_i^gX_h+t_i^g))\right\|_2 \tag{13}$$

Where the rotation matrix of the object in camera coordinates obtained by the frame algorithm with sequence number i is R_i and the corresponding translation vector is t_i , the rotation matrix of the camera in the world coordinate system is denoted as R_i^g and the corresponding translation vector is denoted as t_i^g . The 3D contour point of the current pose is denoted as X_k , the 3D contour point with the closest distance between the reference pose and the 2D projection point of the current pose is denoted as X_h , the number of 3D contour points of the current pose is denoted as C , the camera's internal parameter matrix is denoted as K, and the non-chiral coordinates of the chi-square coordinate mapping are denoted as π.

This is followed by the average surface distance, which is calculated as shown in (14):

$$ASD = \frac{1}{V}\sum_{k=0}^{V-1}\left\|(R_i X_k + t_i) - (R_i^g X_h + t_i^g)\right\|_2 \tag{14}$$

where the number of 3D points in the 3D object model with sequence number k is denoted as X_k and the number of all its vertices is denoted as V.

Finally, there is the reinitialization rate on each sequence, calculated as shown in (15):

$$RIR = \frac{R_n}{F_n} \tag{15}$$

where the number of reinitializations is denoted as R_n and the corresponding number of sequence frames is denoted as F_n.

4 EXPERIMENTAL RESULTS OF MULTIMEDIA INTERACTIVE MODEL OPTIMIZATION OF CAD MODELS IN VIRTUAL REALITY ENVIRONMENT

In this paper, gesture operation technology and 3D object tracking technology are introduced into the CAD model multimedia interaction system to achieve the purpose of multimedia interaction model optimization in terms of gesture recognition, interaction, and dynamic position matching of CAD model in a virtual reality environment. In order to verify the performance of interactive gesture operation of CAD model in a virtual reality environment, this paper compares the gesture operation with the keyboard and mouse operation in the experiment, in which the same gesture will be operated in ten rounds, and the results are shown in Figure 3. From the results in the figure, it can be seen that in the ten rounds of gesture task experiments, both gesture and keyboard and mouse task correct rates have some small fluctuations, the difference between the two correct rates is not large. From the ten rounds of the correct rate, the average can be seen, although the gesture correct rate is slightly lower than the correct rate of the keyboard and mouse, the gap between the two is not large, and there is no obvious variability. From the results of the single round of experiments, it can be seen that there is still a possibility that the gesture correct rate is higher than the keystroke correct rate. This indicates that the gesture operation can effectively replace the keyboard and mouse in the CAD model interactive operation, and after a certain amount of training, the correct rate of gesture will have a certain degree of improvement can reach the level of the correct rate of the keyboard and mouse, and even reach the correct rate of the correct rate over the correct rate of the state of the keyboard and mouse manipulation.

A comparison of the CAD model command accuracy between gesture manipulation and keyboard and mouse manipulation is shown in Fig. 4, where Figure 4 (a) shows the manipulation time of both when the manipulation accuracy is less than or equal to 95%. From the results, it can be seen that when the control accuracy is below 95%, the gesture control time is significantly lower than the keyboard and mouse control time in all three control states. Figure 4 (b) shows the final control accuracy results of the two control modes, and the results show that the control accuracy of the keyboard and mouse under the three command states is higher than that of the gesture control, but the difference between the two is within 1%. This shows that in the virtual reality environment, the gesture manipulation CAD model takes a relatively shorter time to reach the 95% accuracy requirement, i.e., it reaches the corresponding commands in a faster time. In terms of the final accuracy, although the precision of the keyboard and mouse control is higher, which is more suitable to be used in the operation of CAD models with higher fine requirements, the final accuracy of gesture control and keyboard and mouse control is similar. Gesture operation can meet the multimedia interaction requirements of CAD models in the virtual reality environment. It will shorten the completion time of the commands and improve the interaction experience.

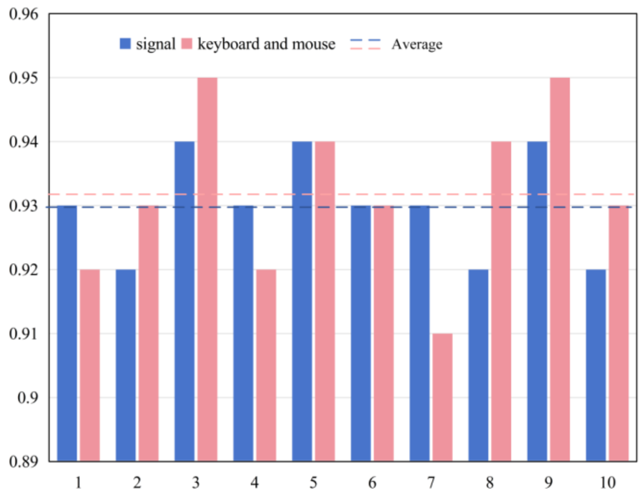

Figure 3: Results of ten rounds of correctness comparison between gesture manipulation and keypad manipulation.

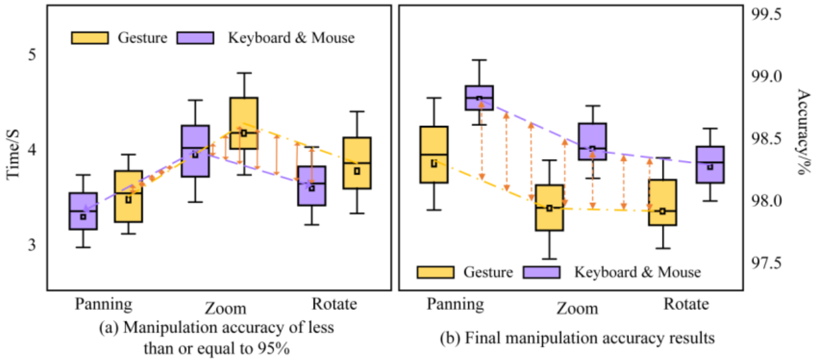

Figure 4: Comparison of CAD model command accuracy between gesture manipulation and keystroke manipulation.

In order to verify the performance of 3D tracking algorithms for CAD model position matching in a virtual reality environment, four other algorithms are selected for comparison in this paper, and the results are shown in Figure 5. From the results in the figure, it can be seen that overall the five algorithms have a certain consistency in the trend of the AED and ASD error accuracy curves, and the results of the two metrics of this paper's model have the best performance among all the algorithms. The model algorithm of this paper can extract feature information of more regions based on the object feature information, realize the optimization of the edge region, show a good feature extraction effect, and therefore show better tracking results. However, in terms of three-dimensional space, the model algorithm of this paper is not the best performance, and the algorithm of model 2 has a slight advantage, but it still has a large potential for improvement.

A comparison of the average speeds of five 3D object tracking algorithm models in the same dataset is shown in Figure 6, where BD is the base dataset, HFL denotes severe illumination variations in the augmented dataset in virtual reality environment, HNO denotes severe noise in the augmented dataset, HMB denotes severe motion blur in the augmented dataset, and OCC denotes the augmented dataset occlusion dataset.

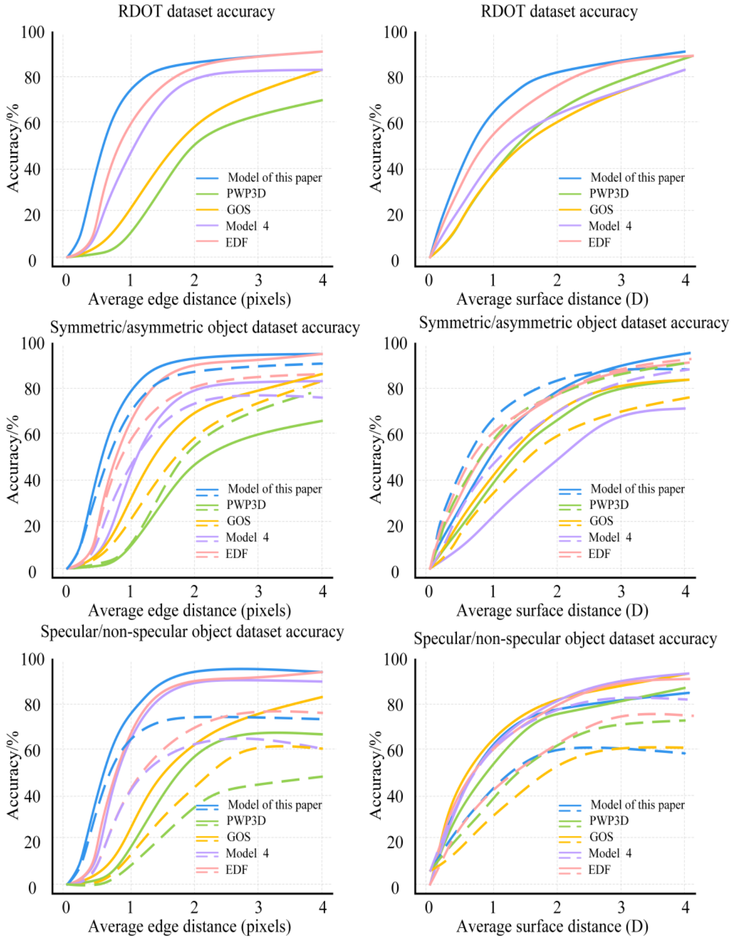

Figure 5: Comparison results of model performance metrics of five 3D object tracking algorithms.

The results show that the model in this paper exhibits the best temporal performance in different types of dataset environments, and the performance performance stability is relatively high. This indicates that the multimedia interaction model of the CAD model in this paper has improved the accuracy and efficiency of bit-position matching in both virtual and real spaces and can provide a better environment for users to use.

In order to further verify the optimization effect of the multimedia interaction model of CAD model in a virtual reality environment, 120 users were randomly selected for interaction experience in this paper, and the result statistics are shown in Figure 7. The results in the figure show that on the whole, most users approve of the optimized multimedia interaction model of CAD model, in which the convenience of gesture control has the highest degree of approval, but the degree of approval in terms of accuracy is relatively low. In addition, more than 95% of the users showed high recognition of the accuracy of the dynamic display of CAD models, and in the three dynamic displays of CAD model panning, zooming, and rotating, except for the relatively low recognition of zooming, the recognition of the other two was higher than 92%. This indicates that the introduction of gesture manipulation technology and 3D object tracking technology in CAD models effectively improves the interactive performance and experience of CAD model multimedia and can improve the performance of the model application in a virtual reality environment.

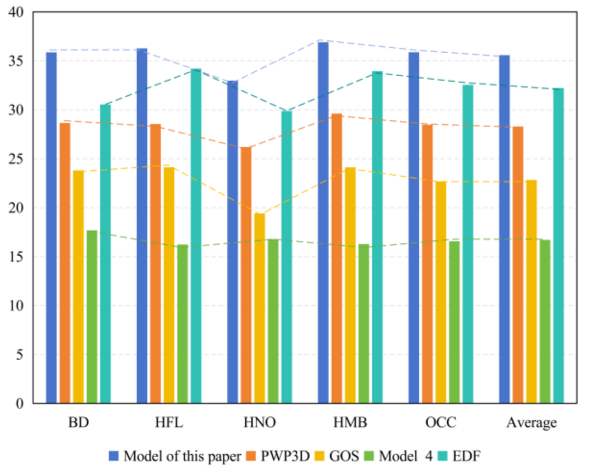

Figure 6: Comparison of the average speed of five 3D object tracking algorithm models in the same dataset.

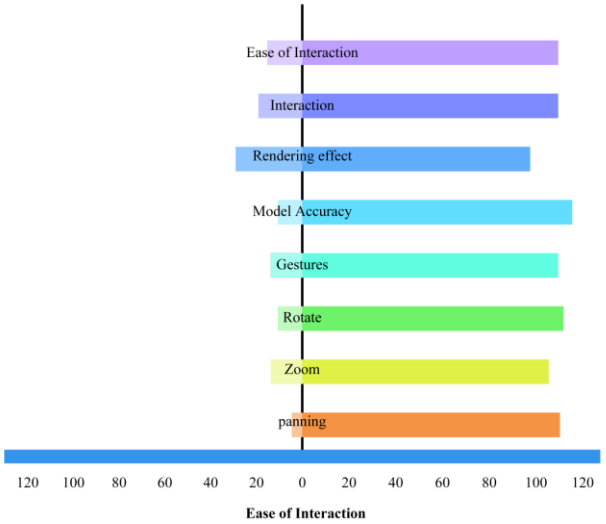

Figure 7: Optimized experience results of a multimedia interactive model of CAD model in a virtual reality environment.

5 CONCLUSIONS

CAD model based on the existing three-dimensional object data to build a three-dimensional model can be shown in the two-dimensional plane, from the visual better presentation of the model design effect, so it has been widely used in many fields. With the development of virtual reality technology and information technology, CAD models are also gradually developed from two-dimensional to three-dimensional presentation, but due to the complexity of the operation of different multimedia interactions, inaccurate three-dimensional spatial presentation effects, data loss, and other issues, the CAD model in the virtual reality environment of the multimedia interaction experience is poor. Therefore, this paper introduces gesture operation technology and 3D object tracking technology in the multimedia interaction system of CAD models to strengthen the interactivity and augmented reality of the model and improve the accuracy of the model's baseline sitting matching in the virtual reality environment. The experimental results show that the gesture manipulation technology shortens the realization time of CAD model commands in a virtual reality environment within 95% accuracy standard, and the overall accuracy is less than 1% than that of keyboard and mouse operation. This indicates that gesture operation technology can enhance the interactivity of CAD model multimedia in a virtual reality environment while ensuring certain correctness and accuracy. In addition, through comparison experiments, it can be seen that this paper's three-dimensional object tracking technology and other models, compared with its different types of dataset environments in different types of performance of the time and stability of the best, the overall tracking effect is optimal, but in the three-dimensional space there is still a greater potential for the development of the CAD model in the virtual reality environment can improve the CAD model in the position of the matching accuracy. According to the results of the user survey on model optimization, most users believe that gesture operation technology and 3D object tracking technology effectively improve the stability and accuracy of CAD models in virtual reality space, improve multimedia interactivity and experience, and can be applied in virtual reality space in different contexts.

Jin Wang, https://orcid.org/0009-0006-2523-2790

REFERENCES

[1] Bruzzone, A.-G.; Fancello, G.; Daga, M.; Leban, B.; Massei, M.: Mixed reality for industrial applications: interactions in human-machine system and modeling in an immersive virtual environment, International Journal of Simulation and Process Modelling, 14(2), 165-177. https://doi.org/10.1504/IJSPM.2019.099910
[2] Burghardt, A.; Szybicki, D.; Gierlak, P.; Kurc, K.; Pietruś, P.; Cygan, R.: Programming of industrial robots using virtual reality and digital twins, Applied Sciences, 10(2), 2020, 486. https://doi.org/10.3390/app10020486
[3] Fechter, M.; Schleich, B.; Wartzack, S.: Comparative evaluation of WIMP and immersive natural finger interaction: A user study on CAD assembly modeling, Virtual Reality, 26(1), 2022, 143-158. https://doi.org/10.1007/s10055-021-00543-0
[4] Khan, S.; Tunçer, B.: Gesture and speech elicitation for 3D CAD modeling in conceptual design, Automation in Construction, 106(1), 102847. https://doi.org/10.1016/j.autcon.2019.102847
[5] Kharoub, H.; Lataifeh, M.; Ahmed, N.: 3D user interface design and usability for immersive VR, Applied sciences, 9(22), 2019, 4861. https://doi.org/10.3390/app9224861
[6] Kozinets, R.-V.: Immersive netnography: a novel method for service experience research in virtual reality, augmented reality and metaverse contexts, Journal of Service Management, 34(1), 2023, 100-125. https://doi.org/10.1108/JOSM-12-2021-0481
[7] Ma, X.; Cao, C.; Sun, Y.: Research on multimedia interaction design of children's emotional education based on social presence, Journal of Computer-Aided Design & Computer Graphics, 32(7), 2020, 1087-1094. https://doi.org/10.3724/SP.J.1089.2020.18352.z25
[8] Makransky, G.; Borre, G.-S.; Mayer, R.-E.: Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments, Journal of Computer Assisted Learning, 35(6), 2019, 691-707. https://doi.org/10.1111/jcal.12375
[9] Malik, A.-A.; Masood, T.; Bilberg, A.: Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace, International Journal of Computer Integrated Manufacturing, 33(1), 2020, 22-37. https://doi.org/10.1080/0951192X.2019.1690685
[10] Mwambe, O.-O.; Tan, P.-X.; Kamioka, E.: Endogenous eye blinking rate to support human automation interaction for e-learning multimedia content specification, Education Sciences, 11(2), 2021, 49. https://doi.org/10.3390/educsci11020049

[11] Pan, J.; Deng, X.: Multi-source information art painting fusion interactive 3d dynamic scene virtual reality technology application research, International Journal of Communication Systems, 35(5), 2022, e4705. https://doi.org/10.1002/dac.4705
[12] Park, K.-B.; Kim, M.; Choi, S.-H.; Lee, J.-Y.: Deep learning-based smart task assistance in wearable augmented reality, Robotics and Computer-Integrated Manufacturing, 63(1), 2020, 101887. https://doi.org/10.1016/j.rcim.2019.101887
[13] Qiu, C.; Zhou, S.; Liu, Z.; Gao, Q.; Tan, J.: Digital assembly technology based on augmented reality and digital twins: a review, Virtual Reality & Intelligent Hardware, 1(6), 2019, 597-610. https://doi.org/10.1016/j.vrih.2019.10.002
[14] Sagnier, C.; Loup, E.-E.; Lourdeaux, D.; Thouvenin, I.; Valléry, G.: User acceptance of virtual reality: an extended technology acceptance model, International Journal of Human–Computer Interaction, 36(11), 2020, 993-1007. https://doi.org/10.1080/10447318.2019.1708612
[15] Zhang, F.; Wu, T.-Y.; Pan, J.-S.; Ding, G.; Li, Z.: Human motion recognition based on SVM in VR art media interaction environment, Human-Centric Computing and Information Sciences, 9(1), 2019, 40. https://doi.org/10.1186/s13673-019-0203-8