International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-14

https://doi.org/10.70917/ijcisim-2025-0221
Article

Inheritance and Innovation Mode of Red Culture in Civic Education of Colleges and Universities in the Digital Era

Yufang Duan[1,*]

[1] Applied Technology College of Soochow University, Suzhou, Jiangsu, 215325, China
* Correspondence author: yfduansuzhou@163.com

Abstract: This paper constructs an immersive red culture learning interactive system, clarifies the architecture and its function of ideological education. With the help of 3D high-precision scene modeling and Leap Motion gesture recognition technology, key methods including vector data acquisition and error correction, hierarchical feature gesture recognition are proposed. Through digital twin, multimodal interaction and adaptive adjustment algorithm, the intelligent and personalized red culture education process is realized. Simulation experiments and system application practices are designed to verify the significant advantages of the proposed system in terms of interactive performance and educational effectiveness. In the 3D high-precision scene construction virtual reality modeling experiment, the proposed method tests the highest scene resolution among the three methods. When the number of scene images is 200, the proposed method still achieves a resolution of 1323.53ppi, which is much higher than the control method. The proposed method also performs optimally in terms of construction time and construction effect, and the scores of all the evaluation dimensions of interactive performance are above 4.0. In teaching practice, the total score of the experimental group class reached 84.26±3.05, far exceeding the 77.12±4.11 of the control group class.

Keywords: civic education; red culture; scene modeling; Leap Motion; learning interaction

1. Introduction

The core values of red culture originated from the profound accumulation of Chinese revolutionary history and socialist construction, which is an ideology, and moreover, a concentrated embodiment of the revolutionary spirit, socialist ideals and communist beliefs [1-2]. Red culture inheritance has a theoretical guidance role in ideological and political education, red culture carries the rich experience and profound lessons of Chinese revolutionary history and socialist construction, through learning and inheritance to help college students to deeply understand the party's struggle and theoretical innovations, and to enhance the identification and confidence in the road of socialism with Chinese characteristics [3-5]. Integrating red culture into the ideological education of colleges and universities is a key initiative to implement the fundamental task of establishing moral education. With the rapid development of the digital era, information dissemination methods and learning modes have undergone profound changes, providing new opportunities for the innovative application of red cultural resources [6-7]. Therefore, the digital improvement of the integration embodiment of red culture and educational carriers in the ideological and political education in colleges and universities will provide a new research direction for the inheritance of red culture [8].

The development of digital intelligence technology provides technological support and innovative

space for the deep integration of red culture dissemination and Civic and political education in colleges and universities [9]. Fu and Ou analyze the logical framework of red culture integration into the Civic and political classroom and study the dissemination path of red cultural resources in Civic and political education driven by digital technology, which plays a crucial role in the enhancement of the teaching quality of the Civic and political education in colleges and universities and the dissemination of red culture[10]. Sun pointed out that the red cultural immersion teaching mode of Civic and Political Education created by using digital technology strengthens students' perceptual cognition, emotional cognition, and practical ability of red culture and Civic and Political content through the narration of typical stories combined with the form of scenarios [11]. Zhao introduced the importance of red culture in civic education, the use of digital platforms to promote the organic combination of civic education and red culture, and enhance the influence of red culture while improving the quality of college students' civic education [12]. Zhang pointed out that the civic classroom is the main carrier of red culture dissemination, so it optimizes the civic education model and builds a virtual practice platform of red culture, realizing the comprehensive, multi-channel integration and "multi-dimensional interaction" of red culture dissemination [13]. It can be seen that the development and application of digital red cultural resources is ushering in an unprecedented development opportunity, which shows great potential in broadening the horizons of ideological and political education and enriching the teaching content [14]. However, behind the booming development, it should also be soberly seen that digital red cultural resources still face many challenges in the depth of excavation, promotion and application practice [15]. These dilemmas not only restrict the full play of the value of red cultural resources, but also hinder the enhancement of the effectiveness of ideological education in colleges and universities.

First, this paper constructs a theoretical model of immersive red culture learning interaction system based on embodied cognition theory and immersive experience technology. Second, focusing on three-dimensional high-precision scene modeling and gesture interaction recognition technology, this paper discusses the construction strategy of immersive red culture learning and interaction system. A cognitive support system for multimodal interaction is established, and an intelligent scene engine and an open scene editing platform are designed. Finally, the effectiveness of the 3D scene modeling method, the interactive performance of the system and the effectiveness of educational practice are comprehensively evaluated through simulation experiments and practical teaching applications.

2. Artificial Intelligence-driven Optimization Path of Red Culture Learning in Civic Education in Colleges and Universities

Under the background of rapid development of digital and intelligent technology, red culture, as an important resource for ideological and political education in colleges and universities, urgently needs to take advantage of emerging technology to realize the innovation of inheritance and enhance the effectiveness of education.

2.1. Theoretical models

Based on the analysis of design elements and combining the theory of embodied cognition and immersive experience technology, this paper constructs a theoretical model of an interactive system for immersive red culture learning, which is shown in Fig. 1, and is designed to guide the design and development of the system to ensure that it can effectively achieve the educational goals and meet the learning needs of users.

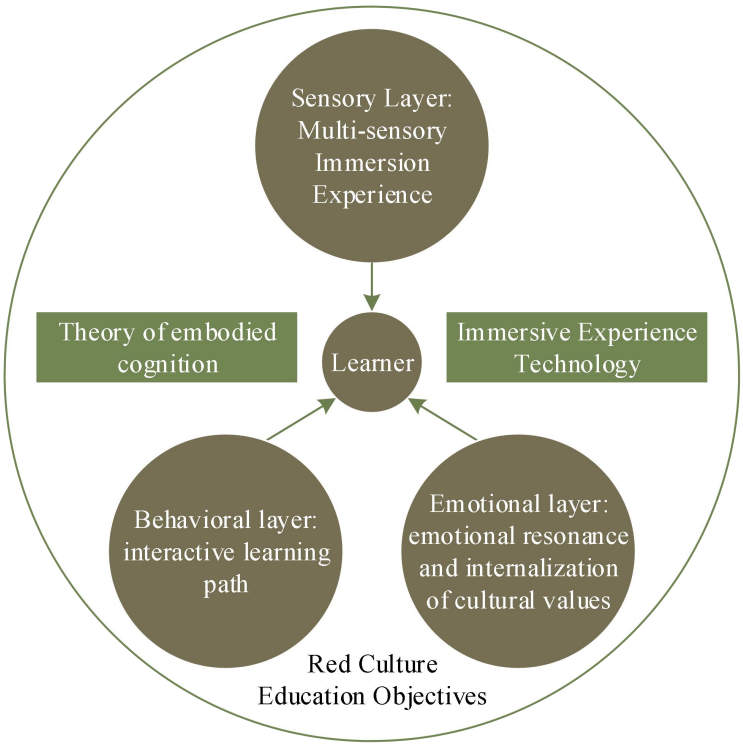


Figure 1. Theoretical model of immersive red culture learning interaction system

The theoretical model constructed in this study focuses on the learner and aims to integrate embodied cognition theory and immersive experience technology to create an immersive learning environment that can stimulate learners' sensory and emotional participation in an all-round way. The design of the model is divided into three levels: multi-sensory immersion experience, interactive learning path, and emotional resonance and internalization of cultural values, which are interrelated to enhance the learners' absorption and understanding of red cultural knowledge and realize the deep-seated goal of red cultural education.

At the sensory level, the model focuses on creating a multi-sensory immersive environment, so that learners can experience rich historical scenes in their audio-visual senses to ensure the authenticity and immersion of the experience.

The behavioral layer focuses on designing an interactive learning path that guides learners to actively explore and participate in the learning process of red culture through task-driven and rewarded feedback. By setting challenging tasks and achievement goals, this layer stimulates learners' endogenous motivation to accomplish the learning goals.

The emotional layer is designed to stimulate learners' emotional resonance, and through the creation of different storylines and substitution of different character perspectives, learners can emotionally connect with the red culture. This emotional resonance helps learners better internalize the values and spiritual connotations of red culture.

Through the guidance of this theoretical model, it can provide a clear research idea and research direction for the subsequent design strategy output and design practice.

2.2. Immersive red culture learning interactive system design

2.2.1. 3D high-precision scene building virtual reality modeling

(1) Acquisition of real scene vector data

In this paper, we collect and update the vector data of the real scene in real time through the reference map, which is more intelligent and can automatically acquire the vector data of the real scene. The data source mainly contains three types of vector data: geography, remote sensing image and topology of the real scene. Among them, the geographic vector data is the core of the real scene vector data, while the remote sensing image vector data and topology vector data are convenient for quickly

finding the target area in the process of building the 3D scene. The acquisition process of real scene vector data source is shown in the following equation.

$$\begin{cases} S_1 = \sum(\alpha_1 - \alpha_0)\cdot z(x_1, y_1) \\ S_2 = \sum(\alpha_2 - \alpha_0)\cdot z(x_2, y_2) \\ S_3 = \sum(\alpha_3 - \alpha_0)\cdot z(x_3, y_3) \\ S = S_1 + S_2 + S_3 \end{cases} \tag{1}$$

where S denotes the real scene vector data source; S_1 denotes the real scene geographic vector data; α_1 denotes the vector data containing the geographic features of the real scene; $z(x_1, y_1)$ denotes the geographic vector data node coordinates; S_2 denotes the real remote sensing image vector data of the scene; α_2 denotes the vector data containing the remote sensing image features of the real scene; $z(x_2, y_2)$ denotes the node coordinates of the remote sensing image vector data; S_3 denotes the topological vector data of the real scene; α_3 denotes the vector data containing topological features of the real scene; $z(x_3, y_3)$ denotes topological vector data node coordinates; α_0 denotes useless data. The collected real scene vector data sources are stored in the map service database, and the corresponding functions are configured with reference to the characteristics of each type of vector data, which are used as the data support for the construction of 3D high-precision scene.

(2) Pre-processing vector data

Since there may exist certain offset errors between remote sensing image vector data and geographic vector data, in order to guarantee the accuracy of 3D scene construction, this paper removes the offset errors between remote sensing image vector data and geographic vector data by aligning them. The geographic vector data of the real scene is sliced according to the control points of the same name, and the mapping function between the local geographic vector data and the remote sensing image vector data is used to realize the alignment of the two. The center point of the offset error of the geographic and remote sensing image vector data with the same name is obtained by clustering of the DBSCAN algorithm, and this point is treated as the initial discrete point, so as to realize the slicing operation of the geographic vector data.The DBSCAN algorithm utilizes the closeness of the spatial distribution of the data points to obtain the clusters of each cluster, and it is assumed that the control point of the geographic vector data for the real scene is D_1, and the remote sensing image vector data control point is D_2, then the expression of the geometric error between the two is:

$$\begin{cases} W\cdot x = D_2\cdot x - D_1\cdot y \\ W\cdot y = D_1\cdot x - D_2\cdot y \end{cases} \tag{2}$$

Where $(W\cdot x, W\cdot y)$ denotes the homonymous control points between the geographic vector data of the real scene and the remote sensing image vector data. The error region of each local vector data is obtained by DBSCAN clustering, and then combined with the spatial distribution of vector data points, the data homonymous points with large offset errors are eliminated, and then the remaining homonymous points are used to construct Tyson polygons, which can realize the slicing of the real scene geographic vector data.

The geometric correction of the local geographic vector data of the real scene after the slicing will cause the connecting line to be broken, so in this paper, the local data are spliced by affine transformation, which in turn realizes the alignment of the overall data. A suitable threshold is set, according to which the vector data that are at the intersection of the segmentation lines are filtered out; the local vector data are aligned by using equation (3) to perform the alignment operation:

$$\begin{cases} X_n = \dfrac{X_n^2 - X_n^1}{m-1} * (n-1) + X_n^1 \\ Y_n = \dfrac{Y_n^2 - Y_n^1}{m-1} * (n-1) + Y_n^1 \end{cases} \tag{3}$$

where (X_n^1, Y_n^1) denotes the coordinates of the n th data node of the real scene vector data before alignment; (X_n^2, Y_n^2) denotes the coordinates of the n th data node of the real scene vector data after

node; (X_n, Y_n) denotes the coordinates of the n th data node in the real scene vector data after alignment and smoothing; m denotes the number of real scene vector data nodes. According to equation (3), the local geographic vector data and image vector data of the real scene are aligned, and then the alignment results are spliced to realize the alignment of the overall real scene vector data.

2.2.2. Hierarchical Gesture Recognition Based on Leap Motion Multi-Feature Fusion

(1) Leap Motion gesture recognition principle

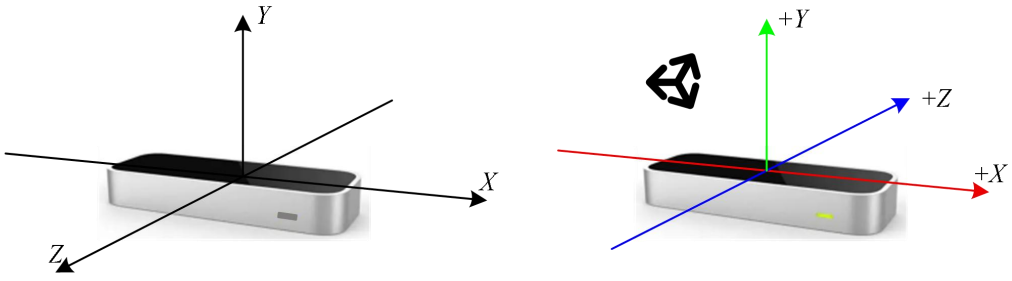
Leap Motion is a high-precision somatosensory device introduced by a company to capture hand movements, with built-in dual cameras, four LEDs and infrared filters for recognizing hand geometry information. Two 1.5 megapixel cameras simulate human eyes to process the scene, capturing hand movements in real 3D space from different angles, capturing left and right visual images of the user's gestures, and generating depth images through stereoscopic calculations. An infrared filter filters natural light into infrared light, which returns the location and direction of movement of the hand and detects hand-specific features through the light field formed by the reflective properties of the object in response to the infrared illumination.The field of view of the Leap Motion is an inverted quadratic prism formed by the location of the device, between 30 mm and 800 mm upwards. The detection system uses frame as the basic unit of data acquisition, acquires hand data at a high frame rate of 200 f/s, and stores each acquired frame in a Leap Frame instance.Leap Motion is simple to use, supports two-handed interaction, and tracks the complete hand object, including localized objects such as the hand, arm, and fingers, within the detection range, which meets the needs of the dissertation research.

The thesis uses Unity engine as the development platform to construct a digital garden scene, conduct gesture recognition research based on Leap Motion, and realize landscape interaction in the virtual scene.The Leap Motion system adopts a right-handed right-angle coordinate system, with millimeters as unit, the center of the device as the origin, the x -axis and z -axis are located in the horizontal plane, and the y -axis is vertically upward. And Unity uses a left-handed coordinate system, in meters, with the z -axis in the opposite direction. The coordinate system conversion and mapping process of Leap Motion is shown in Fig. 2, so it is necessary to map the positional gesture of the real hand captured by Leap Motion to the virtual hand in unity, and to perform the conversion of coordinates and directions between the two based on Eq. (4).

$$\begin{cases} (x, y, z) = k * (x', y', z') * T \\ c = c' * T \end{cases} \tag{4}$$

where $(x', y', z'), c'$ is the hand finger coordinates and orientation captured by Leap Motion, $(x, y, z), c$ is the coordinates and orientation transformed to Unity coordinates, k is the scaling factor, and T is the transformation matrix:

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{5}$$



(a) Leap Motion coordinate system (b) Unity coordinate system transformation

Figure 2. Leap Motion coordinate system transformation and mapping

(2) Improved Hierarchical Gesture Recognition Approach

Combining multiple features for gesture recognition has the defects of large amount of data and complex calculation. To solve this problem, a feature layering strategy is introduced into gesture recognition, where gesture features are categorized by layer to reduce the variety of gestures recognized at each layer. For some gestures, instead of matching all the features of the gesture to be tested with each template, the features are dispersed to each layer, and the gestures are recognized step by step through the layered features. For example, if a gesture can be recognized in the first layer, there is no need to carry out the matching calculation of other features in the next layer, and it directly jumps to the comparison of inter-similarity values, which reduces the amount of computation and improves the recognition speed. Next, using the gestures defined in the template library as a reference, we consider how to divide the features into reasonable layers.

According to the gesture recognition from rough to detailed, the above gesture features are classified into three layers, the first layer feature is the palm displacement direction D_m, which distinguishes between static gestures and dynamic gestures by Δ_p, the second layer feature is the number of fingers N_f, and the third layer feature is the palm direction O_{palm}, the projection distance of neighboring fingers D_{fi} and angle A_{fi}, and joint rotation angle A_{bi}. Using D_m, Δ_p as the first layer of recognition can distinguish the gestures into static and dynamic categories and identify the static bS or dynamic bD labels for the type of motion. For example, when there are two dynamic gestures in the template gesture library, the range of the target gesture is narrowed down to two dynamic template gestures, and it is only necessary to calculate the Euclidean distance between the gesture to be tested X and the two template gestures Y, and then determine whether it is the target gesture or an irrelevant gesture based on the similarity computation; the second layer chooses N_f as the classification basis because the number of fingers is an obvious and important feature to distinguish the gesture, and the detailed expression of the gesture is also based on the fingers, counting the stretched fingers and identifying the corresponding number of fingers labeled $bF_i(\mathrm{i}=0,1,\ldots,5)$; The third layer is to make fine-grained distinction between the gestures to realize the recognition of gestures with higher similarity, and finally only X is calculated by matching with Y which has the same motion label and number of fingers label. The complete layered recognition process is shown in Fig. 3. Each layer can contribute to the gesture recognition, and for some gestures, it is not necessary to go through the complete recognition process, and the recognition result of the middle layer is the final recognition result.

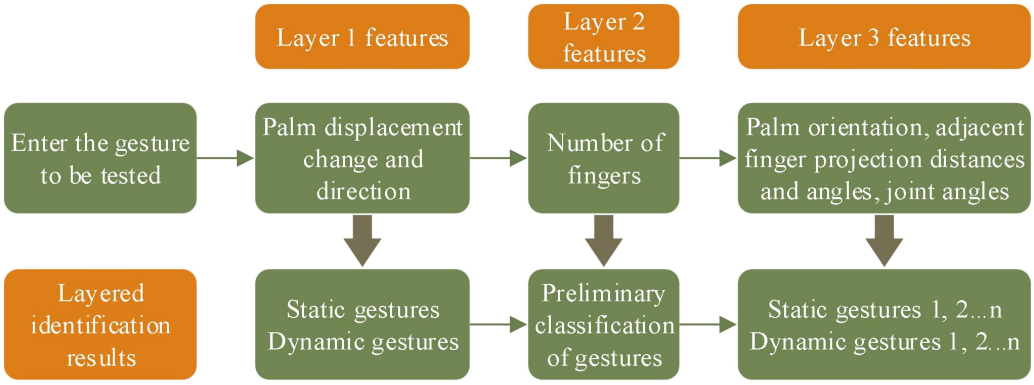


Figure 3. Hierarchical recognition results of gesture features

The gesture matching algorithm based on the geometric features of the hand skeleton calculates the Euclidean distance between the feature pairs based on the features extracted from the input gesture and the template gesture as shown in Eq. (6) so as to measure the degree of matching between the two, and the template gesture with the smallest distance is taken as the target gesture to mark the success of the matching.

$$\| X,Y\| = \sqrt{\sum_{i=1}^{n}\left(x_i - y_i\right)^2} \tag{6}$$

where $\| X,Y\|$ denotes the Euclidean distance between the input gesture X and the template gesture Y, and x_i and y_i denote their i-dimensional features, respectively.

The above decision rules combined with the hierarchical strategy can be used to roughly recognize

the input gesture and categorize the input gesture into one of the categories. However, when the input gesture does not belong to any of the known sample definitions, the classifier will also categorize the gesture into one of the categories, which deviates from the ideal recognition result and reduces the accuracy of gesture recognition. To avoid this, a threshold E is set to describe the similarity of the gesture among the third layer features, comparing the magnitude of the third layer feature dt of this input gesture with E, and the value of E is computed by Eq. (7) in advance through experiments:

$$E = \frac{1}{C}\left(d_1 + d_2 + d_3 + \cdots + d_C\right) \tag{7}$$

Where C represents the number of known gesture sample categories, the value of $C = 7, d_i (i = 1,2,3 \ldots c)$ in this method is taken as after obtaining the experimental data, 20 inputs are randomly sampled for each of seven known sample gestures, and the average distance between the third-level features of the 20 inputs and the third-level features of the corresponding templates is calculated.

The input gestures are classified into known template categories and unknown template categories by a combination of template matching and similarity interval values. When the distance of the third layer feature dt is smaller than E, the recognition result can be correctly output; when dt is larger than E, it means that the gesture is an unknown template gesture, and the subsequent task is not performed, and the recognition flow is shown in Fig. 4.

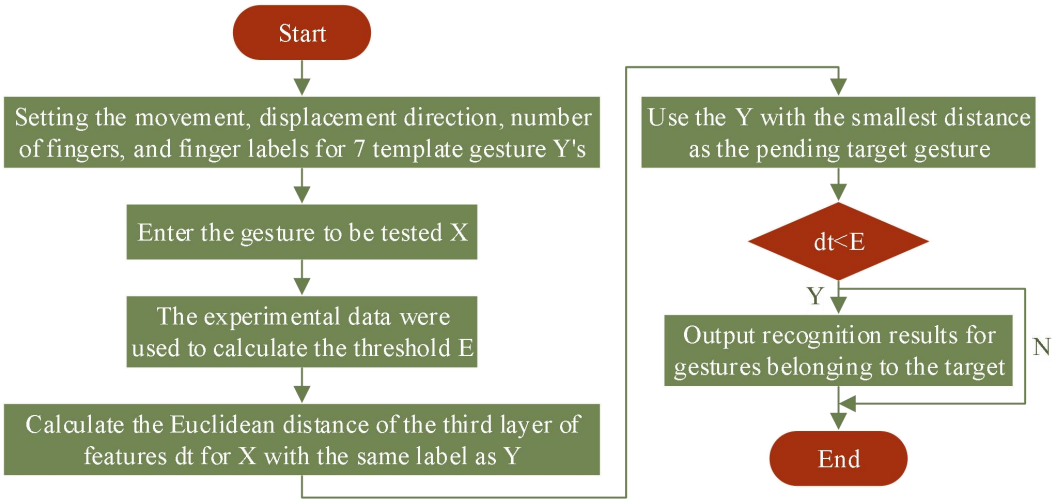


Figure 4. Recognition process

First set the motion, displacement direction, number of fingers and finger labels of the to-be-tested gesture and compare them with the template gesture labels, then calculate the Euclidean distance between the to-be-tested gesture and the template gesture with all the labels being the same in the third layer of the feature, label the template gesture with the smallest distance as a to-be-determined target gesture, and those with a distance less than the similarity threshold belong to the target gesture, and output the recognition results.

2.3. Optimization Path of Artificial Intelligence Driven Education Model

The construction of the intelligent scene engine is a key link in the artificial intelligence technology-enabled red culture ideological education, the essence of this optimization path is to transform the historical logic, practical logic and theoretical logic embedded in the red culture into a perceptible, responsive and evolvable construction through technological means, which requires a breakthrough in the technical limitations of unidirectional presentation, and the establishment of a multimodal interactive cognitive support system. Specific implementation needs to be promoted in three steps: first, build a digital Lison base of the historical scene, using high-precision three-dimensional scanning and semantic modeling technology, the physical space of the revolutionary sites, cultural relics and appliances into an editable digital asset library; second, design multi-level

interaction response rules, so that the learner's gestures, voice questions and other behaviors can be triggered by the real-time feedback of the elements of the scene, for example, touching the virtual telegraph For example, touching the virtual telegraph can trigger the Morse code sending task; finally, the introduction of adaptive adjustment algorithms to dynamically adjust the complexity of the scene according to the cognitive trajectory of the learner - when the system detects that there are obstacles to the user's understanding of the concept of “strategic shift”, it can automatically insert the battle When the system detects that the user has obstacles in understanding the concept of “strategic shift”, it can automatically insert a battlefield projection module to help establish spatial cognition through visualization of troop deployment.

The key to the realization of the technology lies in the establishment of an open scenario editing platform, thus allowing educators to independently configure teaching elements in practice. Teachers can combine different historical scenarios through drag-and-drop interface and set up branching task triggering conditions, for example, embedding the collaborative task of “draft resolution drafting” in the scenario of “Zunyi Conference”, in which learners need to complete the processes of literature review, point of view debates, text revision, etc. by dividing up the work into roles. Learners are required to complete the process of literature review, opinion debate, and text revision through the division of roles. This design transforms technical tools into carriers for educational creation, evolving the historical scene from a static display to a dynamic cognitive laboratory.

3. Analysis of the application of inheritance and innovation mode of red culture in the ideological education of colleges and universities

3.1. Analysis of the effectiveness of virtual reality modeling for 3D high-precision scene building

In order to verify the overall effectiveness of the above 3D interactive scene virtual reality modeling methods, MATLAB simulation software is used for experimental analysis. The 3D high-precision scene building virtual reality modeling method (the proposed method), a hybrid multi-view 3D reconstruction method based on scene graph segmentation (Method A), and an improved 3D reconstruction method for moving spherical panoramic images (Method B) were used for testing, respectively. The images used for the experiments are from the Database dataset, which contains 74 indoor image categories, totaling 16,364 images. Two hundred different scene images were extracted from this dataset, and the 3D interactive scene construction effects of the three methods were analyzed under this condition.

In the process of 3D interactive scene construction, the construction performance of the construction methods is the key to detect the advantages and disadvantages of the methods. When the proposed method, method A and method B are used for 3D interactive scene construction, the scene reconstruction performance of the three methods is tested in terms of resolution and construction time.

3.1.1. Reconstructing Scene Resolution Tests

During the 3D scene construction process, the resolution of the constructed scene is the key to test the reconstruction performance of the reconstruction method. When constructing 3D scenes, the higher the resolution of the constructed scene, the better the constructing performance of the constructing method, and vice versa, the worse it is. When the proposed method, method A and method B are used to carry out 3D interactive scene construction, the scene resolution of the above three methods is tested, and the test results are shown in Figure 5. With the increase of the number of scene images, the scene resolution detected by the three methods decreased to different degrees. The proposed method has the highest scene resolution among the three methods, and when the number of scene images is 200, the resolution of the proposed method still reaches 1323.53ppi, which is much higher than that of the control method. The experiment proves that the proposed method has better performance in 3D scene construction, and the constructed scene is more realistic, which is conducive to improving the effect of human-computer interaction experience.

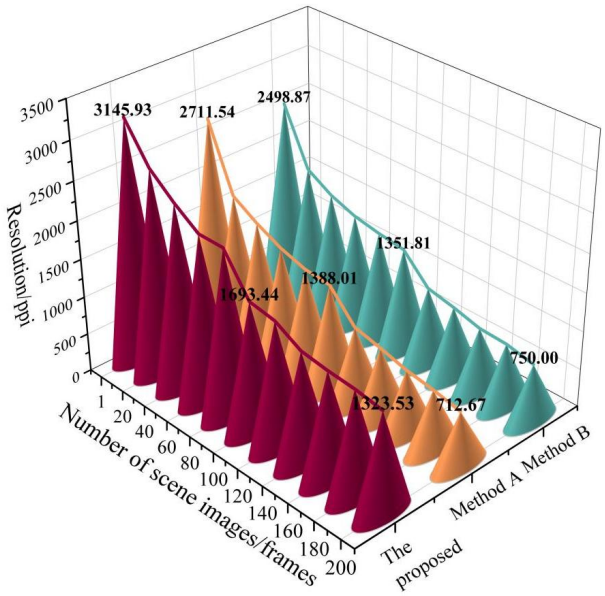


Figure 5. Test results of scene resolution for different methods

3.1.2. Build time testing

The build time can accurately reflect the build performance during 3D scene construction. When the proposed method, method A and method B are used to carry out 3D interactive scene construction, the construction times of the three methods are tested, and the test results are shown in Table 1. With the increasing number of scene images, the construction time of the scene detected by the three methods shows different degrees of growth The proposed method maintains the lowest construction time for all the number of images, which is 11.43ms for 1 frame and 56.92ms for 200 frames, which is significantly lower than that of Methods A and B. Especially when the number of images is large, such as 200 frames, the proposed method saves 62.35ms compared to Method A and 12.35ms compared to Method B. The construction time for Method A and Method B is significantly lower. Especially when the number of images is high, such as 200 frames, the proposed method saves 62.35 ms compared with method A and 12.11 ms compared with method B, which indicates that it still has the ability to build the 3D scene efficiently under high load. The proposed method analyzes the user's 3D action response state before 3D scene construction, so the method has a short construction time during scene construction.

Table 1. Test results of construction time

Number of scene images/frames	Scene construction time/ms		
	The proposed	Method A	Method B
1	11.43	16.55	14.32
20	14.15	23.47	19.04
40	18.64	28.31	26.45
60	20.38	39.26	30.29
80	26.16	48.33	33.28
100	30.29	64.12	35.42
120	34.72	75.61	38.11
140	38.28	89.04	47.33
160	46.17	94.22	53.27
180	54.28	109.56	63.41
200	56.92	119.27	69.03

3.1.3. Actual build results

The virtual reality scene construction method has gained a simulation application in a university's Civic and Political Education communication platform. The project is aimed at upgrading the immersive learning interaction system as a whole, with the goal of improving students' interactivity and spatial immersion. The comparison results between the original system and this study are shown in Table 2. In terms of frame rate stability, the new method achieves an average of 71.81 FPS, which is a

significant improvement compared to the original system's 39.76 FPS. The action recognition accuracy is improved to 92.25% on average, especially reaching 94.22% in region 2, indicating that the system is robust. The delayed response time decreases from 140.48ms to 90.48ms on average in the original system, which significantly compresses the length of the response chain. The semantic region matching rate is improved from 80.77% to 94.78%, with a maximum of 96.25% in region 4, indicating that the system has strong semantic constraints. In terms of resource occupancy, the improved scheme achieves resource scheduling convergence while improving the rendering quality. The mean value is controlled at 56.69%, effectively compressing the redundant computation paths. The rendering anomaly rate is reduced from 2.95% to 0.54%, which can support the stable output of immersive interaction.

Table 2. Comparison results of the original system and this study

Evaluation indicator	Zone 1	Zone 2	Zone 3	Zone 4	Mean value
Frame rate stability(FPS)	36.27	43.18	40.36	39.21	39.76
	67.45	74.92	70.51	74.36	71.81
Accuracy rate of action recognition(%)	68.38	75.58	75.22	74.13	73.33
	91.26	94.22	90.35	93.17	92.25
Delay response time(ms)	162.91	133.45	132.71	132.86	140.48
	96.25	95.12	84.29	86.27	90.48
Semantic area matching rate(%)	83.44	79.05	85.12	75.47	80.77
	94.15	95.26	93.46	96.25	94.78
Resource occupancy rate(%)	42.48	37.33	45.54	42.58	41.98
	56.22	60.15	53.28	57.11	56.69
Rendering error rate(%)	3.04	2.94	3.25	2.56	2.95
	0.55	0.41	0.58	0.62	0.54

3.2. Evaluation of Interactive Systems in Virtual Simulation Teaching and Learning

3.2.1. Cognitive tasks

In order to ensure that the subjects were able to properly understand and comply with the task requirements during the task execution, the subjects received only a brief introduction to the process and task assignment before performing the standardized test tasks, and were all using the system for the first time. In this test evaluation, 25 subjects completed two cognitive tasks. These tasks required subjects to interact with the virtual reality environment through natural gestures in order to achieve specific goals. The results showed that only three subjects were able to successfully complete both tasks at once. This suggests that natural gesture interaction is challenging in virtual reality environments.5 subjects completed the test after failing one attempt at the task.6 subjects completed the task after failing 2 attempts. Most of the other subjects needed 2 attempts or completed the task in discontinuous situations. These results indicate that natural gesture interaction faces some challenges and difficulties in the field of virtual reality. The results of the operation behavior time statistics are shown in Table 3, the average operation behavior time refers to the immediate time required for a single operation, the total operation elapsed time refers to the time from the end of the previous operation behavior to the completion of the current operation behavior, and the total elapsed time of all the operations adds up to the total time taken by the subjects to complete the whole cognitive task. It can be seen that the average time taken for most operations was less than 0.15 seconds, and the shortest total time taken was 0.9 seconds, indicating that the basic operations were responsive.

Table 3. Statistical results of operation behavior time

Operational behavior	Average operation behavior time /Total operation time
Grab→Move→Put down	0.12/2.9
Pull upwards→Grab→Put down	0.19/4.1
Grab→Rotate counterclockwise→Put down	0.15/4.5
Grab→Move→Put down	0.09/2.1
Grab→Put down	0.08/0.9
Grab→Rotate counterclockwise→Put down	0.13/3.8
Grab→Rotate counterclockwise→Put down	0.12/3.4

After the test assessment task was completed, each subject was interviewed and allowed to score multiple assessment dimensions, and the scores are shown in Table 4. In the subjects' feedback, three subjects reflected that the gesture interaction lacked a sense of control and feedback. They were more accustomed to the interaction with force feedback, and felt that the interaction in this experiment was not natural, and they were not satisfied with the operation experience. Four subjects mentioned that they occasionally lost control of the object when grasping the object. However, overall, the subjects were excited about this system test. The scores of all the dimensions were above 4.0, and the “interface clarity” (4.36) was the best performance, indicating that the system was highly recognized for its interface clarity.

Table 4. Evaluation Score Results

Dimension	Influencing factor	Evaluation outline	Mean score (5 points)
Scene	Interface	Interface clarity level	4.36
	Authenticity	The authenticity of the scene situation	4.12
Character	Guide	Provide clear and understandable instructions	4.28
	Delay	Gesture recognition response rate perception	4.32
	Smooth	Gesture recognition of limb coordination	4.40
Interaction	Fault tolerance rate	Rationality of the feedback mechanism for non-standard behaviors	4.16
	Binding force	User interaction collision authenticity	4.32
	Feedback	Feedback mechanism operates quickly and accurately	4.24
Task	Streamline	Streamlining of the task process	4.08
	Expectation	The task feedback results are in line with the expected situation	4.32
	Rationality	The operation is naturally in line with human ergonomic principles	4.28

3.2.2. Gesture Recognition

In the process of operational interaction to accomplish cognitive tasks, gesture recognition as the core technology of communication between users and virtual environment, its accuracy directly affects the experience of task completion and the practicality of the system. There are some gestures whose change curves are extremely similar, which are called similar gestures in this paper. The recognition rate of the following similar gestures is compared between the interactive system in this paper and the interactive system using four traditional gesture recognition algorithms, such as Hausdorff Distance, Fréchet, DTW, and Pearson's correlation coefficient, and the recognition results are shown in Table 5. The average recognition rate of this system reaches 97.03%, which is significantly higher than the systems using other algorithms. Especially in complex gestures such as five-finger rotation, the recognition rate of this system are over 97%, which is an outstanding advantage. It shows that the hierarchical feature fusion and similarity thresholding strategy effectively improves the differentiation ability of similar gestures, and verifies the advantages of the proposed gesture recognition method in terms of accuracy and robustness.

Table 5. Comparison results of recognition rates for ten similar gestures (%)

	Hausdorff distance	Fréchet	DTW	Pearson correlation coefficient	The proposed
Three-finger grip	92.53	92.62	95.11	94.32	97.42
Three-finger twist	94.12	93.18	96.23	95.44	96.75
Five fingers grasping	93.01	93.45	95.78	94.27	97.84
Five-finger rotation	92.75	92.81	96.42	94.82	97.16
Draw a wavy line	92.34	93.04	96.17	95.26	96.83
Draw an arc	91.67	91.42	95.48	95.18	96.72
Draw a triangle	93.03	92.89	94.96	96.23	97.04
Draw a circle	92.84	93.06	95.22	95.75	96.49
One and two fingers open	93.29	93.17	96.18	94.92	97.25
One or three fingers	92.88	92.65	95.75	95.61	96.84

open Average	92.85	92.83	95.73	95.18	97.03

3.3. Analysis of application effectiveness

In order to further test the actual effectiveness of the optimization path of red culture ideological education driven by artificial intelligence, this study conducted a semester-long comparative experiment of teaching application with University A as the practice field. The experiment was conducted in two parallel classes of a certain major in the university. Under the premise that the teaching hours, assessment methods, and the question types and difficulties of the final exams were identical, the experimental group adopted the teaching strategies designed in this paper throughout the whole course, while the control group followed the traditional multimedia lecture mode. After the implementation of the teaching optimization strategy for one semester, the teacher team will understand the teaching effect through the analysis of students' performance. The final exam was a university-wide standardized exam with questions from the Civics Exam Database created by the Civics Teaching and Research Center of University A. The students in the experimental group and the control group were tested on the Civics Exam Database. The students' grades of the experimental group and the control group were compared and analyzed, and the results of the analysis are shown in Figure 6. It was found that after the implementation of the teaching optimization strategy, the regular grades of the students in the experimental group were significantly improved, with an increase of 7.59 points over the control group class, and the final grades were also slightly improved. The average increase in the total grade is more obvious, and the total grade of the experimental group class reaches 84.26±3.05, which is far more than the 77.12±4.11 of the control group class.

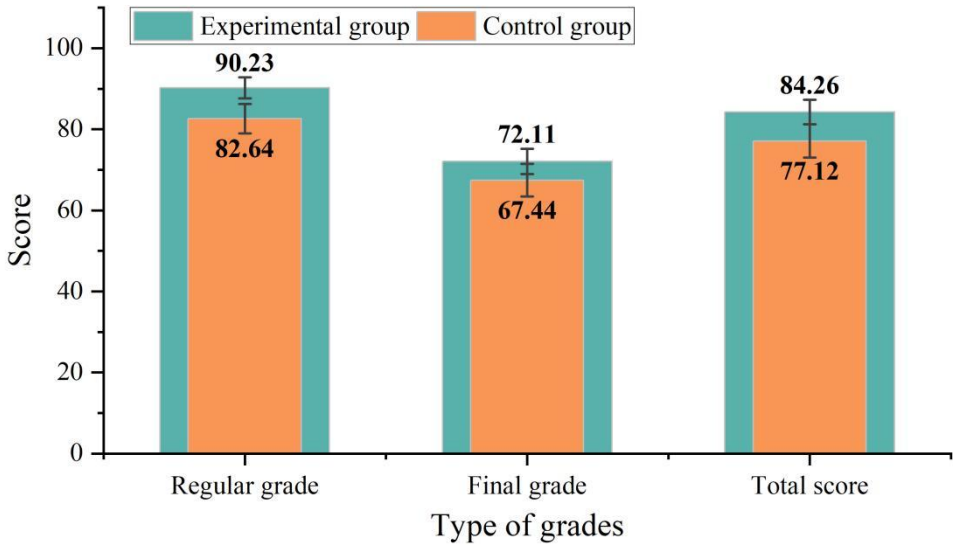


Figure 6. Analysis results of students' course grades

4. Conclusion

This paper constructs an artificial intelligence-driven interactive model for learning red culture in college civic education, and carries out experimental validation around key technical aspects.

In the 3D high-precision scene building virtual reality modeling experiments, the proposed method tests the highest scene resolution among the three methods, and when the number of scene images is 200, the resolution of the proposed method still reaches 1,323.53 ppi, which is far superior to the control method. When the number of images is 200, the proposed method saves 62.35ms in construction time compared to method A and 12.11ms compared to method B. The average frame rate stability reaches 71.81FPS, the average action recognition accuracy is improved to 92.25%, the average delayed response time decreases from 140.48ms to 90.48ms, and the semantic region matching rate increases from 80.77% to 94.78%, and the semantic region matching rate increases from 80.77% to 94.78%, and the average delayed response time decreases from 140.48ms to 90.48ms. The semantic region matching rate increased from 80.77% to 94.78%, the average resource utilization rate was

controlled at 56.69%, and the rendering anomaly rate decreased from 2.95% to 0.54%. In the evaluation of the application of interactive system in virtual simulation teaching, the average time consumed for most operations was less than 0.15 seconds, and the shortest total time consumed was 0.9 seconds. The scores of all the evaluation dimensions were above 4.0, with the best performance in "interface clarity" (4.36). The average gesture recognition rate is 97.03%, which is significantly higher than that of the system using other gesture recognition algorithms. In the teaching practice, the experimental group's students' regular grades improved significantly, 7.59 points higher than that of the control group, and the final grades also improved slightly. The average improvement of the total grade is more obvious, and the total grade of the experimental group class reaches 84.26±3.05, which is much more than the 77.12±4.11 of the control group class.

The results of simulation experiments and system application show that the proposed method shows good application value and potential for popularization in terms of scene construction efficiency, user interaction experience and achievement of educational goals.

References

1. Zhong, G. (2025, May). The Role of Red Song Heritage and Red Culture in Primary School Music Education: a Literature Review. In Proceedings of the 2nd International Conference on Educational Development and Social Sciences (EDSS 2025) (Vol. 924, p. 159). Springer Nature.
2. Hu, R. (2025). Neutrosophic Cultural Resonance for Teaching Effectiveness: A Plithogenic Probabilities-Inspired Framework Model for University Fine Arts Programs Using Red Culture Resources. Neutrosophic Sets and Systems, 93, 660-670.
3. Hou, J. (2023). Educational psychology and red culture identity education for college students with borderline personality disorder: art design collaboration. CNS Spectrums, 28(S2), S60-S61.
4. Ma, Y. (2023). A Probe into the Path of Integrating Red Culture into Ideological and Political Education in Colleges and Universities. Adult and Higher Education, 5(14), 36-41.
5. Wu, W., & Dong, X. (2022). Research of Integrating Red Culture into the Party Spirit Education of College Student Party Members in the New Era. Frontiers in Educational Research, 5(15).
6. Xie, B. (2024). Digital protection and dissemination of red cultural resources based on the Internet of Things and big data. Journal of Computational Methods in Sciences and Engineering, 14727978251367169.
7. Huang, M., & Zeng, X. (2024). Digital Protection and Innovative Development Path of Red Culture Resources Based on Distributed Machine Learning Supported by Intelligent Information. Journal of Combinatorial Mathematics and Combinatorial Computing, 120, 381-391.
8. Wang, L. (2024). The Path of Integrating Red Culture into Party Construction in Colleges and Universities in the Digital Age. Critical Humanistic Social Theory, 1(2), 32-38.
9. Wang, S., & Zhang, S. (2025). Philosophical Reflections on the Role of Artificial Intelligence in Shaping Cultural Values in Ideological Education. Cultura: International Journal of Philosophy of Culture and Axiology, 22(3).
10. Fu, C., & Ou, M. (2024). Research on Digital Empowerment of Integrating Red Culture Resources into College Ideological and Political Courses. Advances in Humanities and Modern Education Research, 1(1), 115-121.
11. Sun, Y. (2022). Research on the Application of the Red Culture Immersion Teaching in Ideological and Political Education in Colleges and Universities. Academic Journal of Humanities & Social Sciences, 5(15), 28-32.
12. Zhao, Q. (2024). The Logic and Approaches of Enhancing University Students' Ideological and Political Quality through Red Culture. Journal of Social Science Humanities and Literature, 7(4), 76-82.
13. Zhang, W. (2018, December). Research of Red Culture Propaganda Platform and Ideological and Political Education in Colleges and Universities. In 2018 8th International Conference on Management, Education and Information (MEICI 2018) (pp. 1213-1217). Atlantis Press.

14. Tan, L., & Liang, R. (2023, July). Research on the Design of Red Cultural and Creative Products Based on Digitalization. In International Conference on Human-Computer Interaction (pp. 162-171). Cham: Springer Nature Switzerland.
15. Miao, F., & Zhang, N. (2024). Digitization, Preservation, and Dance Narrative Exploration of Red Music Cultural Heritage in Hebei Province: A Multidisciplinary Approach. International Journal of Social Sciences and Public Administration, 2(2), 211-221.