International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-16

https://doi.org/10.70917/ijcisim-2025-0202
Article

The Design of Virtual Exhibition and Interactive Teaching Mode of Cave Art Supported by Augmented Reality Technology

Junlei Li *

School of Education, Longdong University, Qingyang, Gansu, 745000, China; ljl19852000@163.com

Abstract: The increasing development and maturity of augmented reality technology has injected new vitality into the model updating of digital display and educational inheritance of grotto art. In this paper, ORB algorithm is selected as the benchmark algorithm for image recognition of digital display of art works. Aiming at the low matching efficiency of the original ORM-SLAM algorithm and the redundant volume of the algorithm, the adaptive histogram equalization algorithm is used to preprocess the target image to form an improved ORM-SLAM algorithm that can efficiently track the matching of key points. Combined with the grotto artwork model matching method, the grotto artwork model based on augmented reality technology is constructed. The Hololens device is used as the realization carrier to design the realization scheme of virtual exhibition of cave art works under space mapping. Accordingly, the practical strategy of grotto art works in interactive teaching mode is proposed, under the guidance of this strategy, the mean value of the performance rating of the experimental subjects' grotto art works is 85.57, which is rated as excellent. The experimental results show that the interactive teaching mode under the virtual exhibition program in this paper can comprehensively display the contents and connotations of the cave art works, and effectively promote the enhancement of students' artistic literacy.

Keywords: ORM-SLAM improvement algorithm; cave art virtual exhibition; interactive teaching; spatial mapping; augmented reality

1. Introduction

Chinese cave art originated in the Qin Dynasty, developed in the Wei and Jin Dynasties, flourished in the Sui and Tang Dynasties, and continued into the Song and Yuan Dynasties [1-2]. China's cave art has unique artistic characteristics, including sculpture, painting and murals and other forms of art. First of all, the sculpture of cave art is exquisite, and the exquisite sculptures of Buddha statues are so lifelike and vivid that they are breathtaking [3-4]. Secondly, paintings and murals were also widely used in cave art, and the colorful paintings and murals displayed rich stories and religious imagery with their unique colors and lines [5-7]. In addition, grotto art is compatible with the characteristics of Chinese and ethnic art, forming a unique and charming Chinese character [8]. However, due to natural disasters, economic construction, tourism development and other factors, many precious and rare grottoes have been in an endangered situation, so how to effectively protect, research and develop these cultural relics resources has become very urgent and important [9-11].

With the continuous development of science and technology, augmented reality (AR) technology, as a kind of interactive way to integrate the virtual world and the real world, has been widely used in the field of cave art [12-13]. Through AR technology, people can add virtual elements to the real scene to provide viewers with a new experience of cave art [14-15]. Using augmented reality technology, the communication of grotto art can be carried out in a more vivid and rich form, which enhances the audience's sense of participation and experience [16-18]. AR technology can also provide feasible solutions for the protection and inheritance of fine art. Through AR technology, cultural heritage institutions can digitize precious grottoes and present them in AR form [19]. In this way, people can still appreciate the details and artistic beauty of the actual grotto artifacts without touching them. For those caves that are vulnerable to damage or cannot be displayed, AR technology provides an alternative that allows them to be preserved and passed on in a virtual environment [20-21]. And in the teaching of grotto art, it is of great significance to improve the teaching effect through the help of AR technology, which can provide students with an immersive experience of grotto art and enable them to learn in interactive teaching, thus improving the interest and quality of learning [22-25].

In this paper, we first sort out the framework of augmented reality system one by one and the realization process of image recognition matching technology to lay the theoretical foundation of research. Then it focuses on the mathematical principle and operation steps of ORB algorithm, and proposes ORB-SLAM improvement algorithm. The model of grotto art works is established and the matching process of the matching model algorithm is analyzed to construct the model of grotto art works based on augmented reality technology. Then, based on the spatial mapping, we describe in detail the spatial scanning and data labeling processing, surface mesh planarization processing and mapping, and form the virtual exhibition implementation scheme of the cave art works. After that, the performance of the proposed ORB-SLAM improvement algorithm is evaluated in terms of alignment effect and localization accuracy performance. The feasibility of the virtual exhibition scheme of grotto art is also analyzed through user satisfaction research. Finally, with the support of the virtual exhibition program in this paper, the practical path of applying the interactive classroom teaching mode to the grotto art works is clarified and its practical application effect is examined.

2. AR Technology Basis for the Exhibition of Grotto Art Works

2.1. Augmented reality technology and image recognition matching technology

Augmented reality technology is through the camera to collect image data, after a series of processing to extract the target object sign after matching, and maintain real-time access to the location of the camera and other related information, and then will be pre-computer-generated about the target object of the model, video and text and other virtual information and the real environment, so as to achieve augmented reality, which is conducive to the user a more vivid, imaginative and profound understanding of the The target object. The augmented reality system framework is mainly composed of six modules, including real environment acquisition, image detection and recognition, 3D tracking and registr,ation, virtual-reality fusion, intelligent terminal display and human-computer interaction, and the augmented reality system framework is shown in Fig. 1.

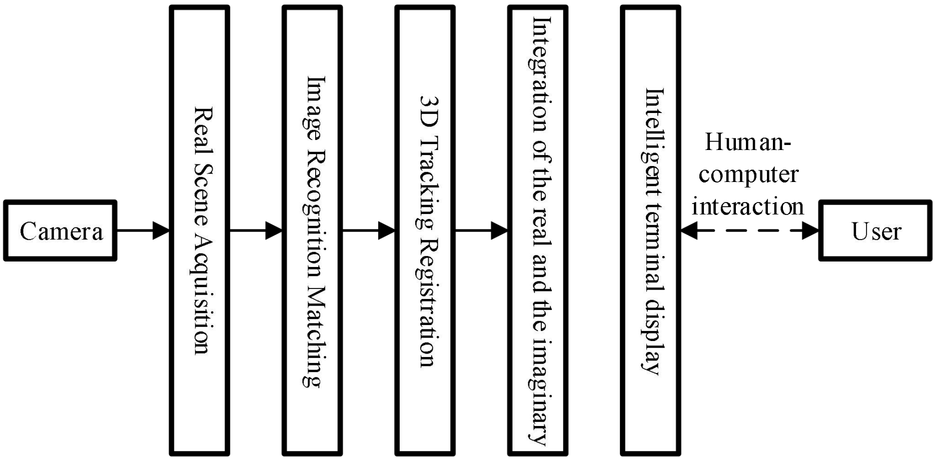


Figure 1. Augmented reality system framework.

The main technologies of the realization process of augmented reality include reality fusion technology, virtual modeling technology, sensor technology and 3D rendering, etc., of which the most critical technologies are 3D tracking and registration technology, image recognition and matching technology, reality fusion technology, intelligent display technology and human-computer real-time interaction technology. Virtual-reality fusion technology is related to the user's immersion and experience effect; virtual modeling technology is related to the success of the establishment of the system model; real-time human-computer interaction technology is related to the smooth communication between the user and the system, whether the use of the feeling of perfection; and the most critical technology among these key technologies is the image detection and recognition technology and three-dimensional tracking and registration technology.

The success of image recognition is directly related to the realization of the augmented reality effect of the whole system, which can be said to be an indispensable part of the whole system, the image detection and recognition process is shown in Fig. 2.

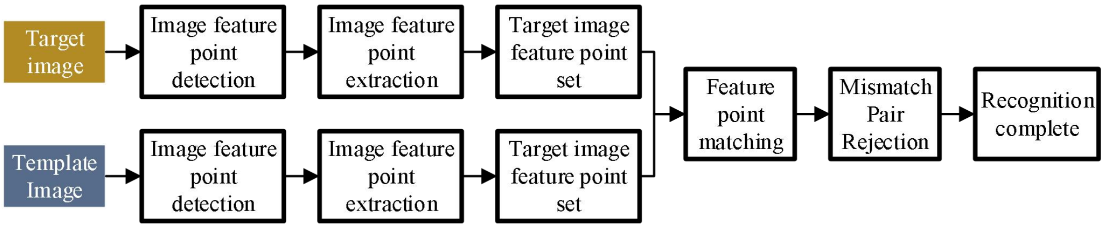


Figure 2. Image recognition and matching process.

One of the key processes of augmented reality system to realize information enhancement is the recognition and matching of the target image, and the good or bad recognition effect of the target image directly affects whether the later image matching link is successful or not. Image matching methods can be divided into image gray scale matching methods and image feature matching methods in general. Image grayscale-based matching methods mainly use the covariance function, the absolute value of the difference and other similarity measures to determine the similarity between the two images, the more classic method is the normalized grayscale matching method. Image feature-based matching method is to compare two images by identifying more obvious features such as color, position and special texture of the image to achieve the correct matching between the two images. Due to the matching rate of image feature-based matching method is more rapid, high accuracy and better robustness to lighting, shape changes and image occlusion, so compared to the matching method based on image features in real life wide range of applications, relatively speaking, based on image features algorithms are also more common, the current image-based algorithms in the more classic, more research when it comes to image feature-based algorithms are SIFT, SURF and ORB algorithms. In this paper, ORB algorithm is selected as the image feature algorithm used in the research, and the algorithm is introduced and improved in the following.

2.2. Principles of the ORB algorithm

ORB algorithm is the direction of the improved FAST detection algorithm and binary to describe the composition of the BRIEF, the algorithm running speed performance is very good. ORB algorithm operation process mainly includes corner detection, feature point extraction, and the generation of feature point descriptors and image matching and other processes.ORB algorithm implementation process is shown in Figure 3.

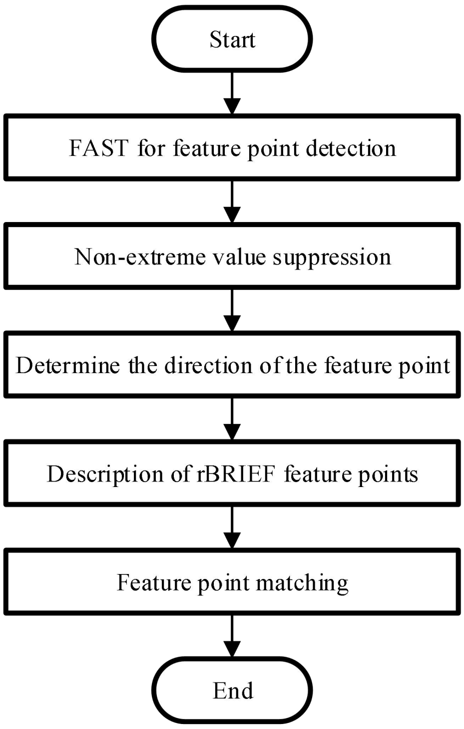


Figure 3. ORB Algorithm Process.

FAST determines whether a point is a desired corner point or not by using whether the change of pixels in the comparison area is abrupt or not; pixels that are too dark or too bright in the image may be corner points. This detection method only needs to compare the pixel size, so the detection is very fast. In ORB algorithm FAST feature point detection process is as follows:

Step 1: Select any point P in the target image, set the brightness of point P as Ip and set the threshold T.

Step 2: Take a point P as a fixed point and make a circle with a radius of R (usually taken as 3) and select 16 pixel points on the drawn circular region.

Step 3: In order to improve the computing speed of the algorithm, generally when determining whether a point P is a corner point, the point is first compared with the pixels of the 1st, 5th, 9th, and 13th points on the detected circular region, and a viewpoint is possible when and only when at least three of these four point pixels are greater than Ip+T or less than Ip-T. If it is not satisfied, then the point P should be excluded directly.

Step 4: If step three is satisfied, then point P and the remaining 12 points continue to be compared, if there are N (generally N takes 12) points in the circular region of 16 points when the brightness is within the range of (Ip-T, Ip+T), then it can be concluded that point P is a feature point in the target image.

Step 5: Cycle the previous four steps, so that each point in the image carries out the above steps to get all the eligible feature points in the target image.

After the initial detection of the corner points, it is very easy for the corner points to gather, so after the completion of the initial detection of the non-maximum value suppression to solve the problem of excessive aggregation of the corner points, which can make the response in the region to leave only the point of the maximum value of the FAST algorithm feature point speed there are also some shortcomings in the detection of the number of feature points there is uncertainty about the subsequent arithmetic process to increase the difficulty of the process. ORB algorithm in order to solve this problem, the introduction of the number of corner points N, in the calculation of the Harris response value of the corner points, according to a certain order of arrangement, select the maximum of the first N response value of the corner points to do the final need for the corner point set. Since the FAST algorithm is not very stable when detecting rotational changes in the image, the ORB algorithm is used to improve the rotational stability, so the gray scale center of mass method is chosen to solve the problem. The center of mass in the circular neighborhood of a feature point is calculated by moments, and the vector formed by the detected feature point and the center of mass is used as the direction of the feature point.

The grayscale center of mass of an image is centered by the grayscale value to be used as a weight, usually the moments of the image block B are expressed as equation (1):

$$m_{pq} = \sum_{x,y\in r} x^p y^q I(x,y) \tag{1}$$

In the above equation, r is the radius, x, y then represent the coordinate values, and $I(x, y)$ is the gray scale expression of the image. The center of mass of the image block B is obtained by moments as equation (2):

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \tag{2}$$

Assuming that the coordinates of the feature point are O, the angle of the direction vector is the direction of the feature point. The direction of the angular point can be calculated as equation (3):

$$\theta = \arctan\left(\frac{m_{01}}{m_{00}} / \frac{m_{10}}{m_{00}} \right) = \arctan\left(m_{01} / m_{10} \right) \tag{3}$$

With the introduction of the gray scale center of mass, the corners detected by FAST are then stabilized for rotational transformations, and at the same time the overall performance of the ORB algorithm is improved. The improved FAST by gray scale center of mass gives it a new name in ORB algorithm, called Oriented FAST.

After initially obtaining the feature points of the target image, the feature point descriptor of the ORB algorithm is improved from the BRIEF descriptor by adding a rotation factor.The BRIEF descriptor manifests itself as a binary, consisting of a combination of 0s and 1s arranged in a certain way. The encoding after encoding through a certain sequence is used to represent the size comparison relationship between any two points (a, b) pixels in the domain range of the key point. In order to reduce the interference of noise, it is necessary to first perform a Gaussian filtering operation on the image, and then take a $S \times S$ domain window centered on the key point and perform a random selection of point pairs. If $a > b$, a binary code 1 is generated, and if $a < b$ or $a = b$, a binary code 0 is generated.After comparing all the points, a length of n is generated. Binary string, in the usual case n is taken as 128,256 or 512, in opencv the value of n is generally taken as 256 by default, with equation (4):

$$T\left(p:\ x,\ y\right) = \begin{cases} 1 : p(a) < p(b) \\ 0 : p(a) \geq p(b) \end{cases} \tag{4}$$

The coordinate system established in the original BRIEF algorithm lacks flexibility, and is composed of the horizontal and vertical planes plus the origin (feature point), which leads to the fact that the coordinate system does not change when the image undergoes rotational changes, so that the feature points and the feature point descriptors also change, which makes the algorithm not rotationally invariant. The improved BRIEF algorithm establishes the coordinate system by connecting the feature points with the form center of the region where the points are taken, which makes the improved algorithm have better stability when the image undergoes rotational changes, and the feature points detected by the improved FAST algorithm contain more information such as the angle of direction θ , etc., and the improved BRIEF algorithm performs the rotational discrimination of the point pairs first, and then performs the binary Encoding. The n point pairs form a $2 \times n$ matrix as equation (5):

$$T = \begin{pmatrix} x_1, x_2, \ldots\ldots, x_n \\ y_1, y_2, \ldots\ldots, y_n \end{pmatrix} \tag{5}$$

The new pair of points after rotation of the rotation matrix R_θ formed through θ is equation (6):

$$T_\theta = R_\theta T \tag{6}$$

A new binary descriptor is formed by comparing the sizes of pairs of points in the new set of pairs, so that pixels in T_θ can be used directly when obtaining the descriptors of feature points. The improved descriptor has the rotational invariance but also makes the differentiation performance very poor, in order to solve this problem, the learning method and greedy search are introduced to obtain the rBRIEF. After obtaining the feature point descriptor, the success of the matching is determined by judging the size of the Hamming distance between the feature point pairs.

2.3. ORB-SLAM Improved Algorithm

The original ORB-SLAM algorithm has low matching efficiency and large algorithm size. Aiming at these problems, this paper designs an improved ORB-SLAM algorithm aiming at high efficiency and small size, and completes the target feature recognition module within the system based on this algorithm. The structure of this improved algorithm is designed according to ORB-SLAM², which implements a total of three functions: tracking camera position and target features, implementing target feature mapping, and optimizing feature keypoints. The Tracking Camera Position and Target Feature module runs in the front-end of the system to provide observation information for the remaining two function modules, and the Target Feature Mapping and Optimizing Feature Keypoints modules are activated in the background when a target feature that serves as a keypoint is detected in order to eliminate the local drift error caused by ARCore rendering of the 3D model. The target feature mapping module contains information such as target keypoints, real-time video information, co-views, and spanning trees, etc. In order to achieve the goal of efficient computation, the module eliminates the misrecognized keypoint information among them by the 2-NN cross-matching method based on the Hamming distance.

Since the target image features in the real environment will be affected by the ambient light, resulting in an insufficient number of keypoints available for detection. Therefore the target images in the database are preprocessed using the adaptive histogram equalization algorithm. The processed target image will have enough keypoints for the system to detect and track, if there are enough describable points in the image, the algorithm does not need to extract the keypoints for the continuous input target video stream, it only needs to detect the keypoints of the stable image frames in the video stream, and the keypoints of the non-neighboring frames and the target image can be matched by the target feature descriptors processed in advance. After this processing, this improved algorithm has better detection accuracy and higher computational efficiency compared to the ORB algorithm.

Before describing the complete model of the improved algorithm, it is necessary to determine the speed of the cell phone camera at the video input during the system operation, and the specific speed is projected as in equation (7). Where ϕ_t denotes the motion state of the camera during the video stream input, and ρ_t denotes the relative motion change matrix of the camera between the initial frame image and the final frame image.

$$v_t = f\left[\phi_t\right] \odot \rho_t \tag{7}$$

Assuming that the camera's velocity increment between any two frames in the video stream is equal, let the initial frame camera motion position be Z_1, the final frame camera motion position be Z_t, and the displacement relation between the two be Z_t^1, and the camera displacement relation reacted by the camera's velocity is shown in Eq. (8):

$$\exp\left(v_t\right) = Z_1 \otimes Z_t = Z_t^1 \tag{8}$$

Let the camera internal reference matrix be N, the projection error between the 2D coordinates (x, y) of the keypoints to be tracked and matched in the video stream and the 3D coordinates r of the corresponding keypoints in the reality be $r_{x,y}$, and the robust kernel function be $\mathbb{Z}Hub$, whose function serves to reduce the error when the matching accuracy of the keypoints is insufficient, the The optimization formula for key point matching error is shown in equation (9):

$$W = \min_t N\left(\sum_{y=0}^{n-1}\sum_{x\in n}\mathbb{Z}Hub\left(\left\|r_{x,y}\right\|_{\Sigma_{xy}^{-1}}\right)\right) \tag{9}$$

The set of optimized keypoints is P_L and the set of image frames is V_L, and let P_k be the set of keypoints for which the matching is completed, and F_{xy} be the visible points other than the keypoints, the local keypoint mapping formula is as in equation (10).

$$E_{xy} = \left\|P_k + Z_t^1 \odot \left(P_L - \arg\min\sum_{P_k\in P_L}\sum_{V_L}\rho\left(F_{xy}\right)\right)\right\| \tag{10}$$

For all the target image keypoints of the target image feature database in the system, their spatial coordinates can be determined by the mapping formula to get the correspondence of the image feature keypoints in the real environment.

2.4. Cave art work model matching method

Cave artwork model matching is a very important task in establishing a virtual exhibition environment for cave art, and the accuracy of the matching result represents whether the system can display the corresponding model to visitors. The core idea of the matching model algorithm lies in how to determine the correspondence between the cultural relic model database and the target image feature database in the system. The matching process of the grotto art works is shown in Fig. 4. The same number is given to the model database of the grotto art works and the target image database, respectively. When the ORB-SLAM improved algorithm retrieves the similar target feature keypoints in the video stream acquired by the camera, it sends commands to the backend through the cell phone chip to locate the corresponding target image in the database, retrieves the matching model through the same number, and renders it to the target location.

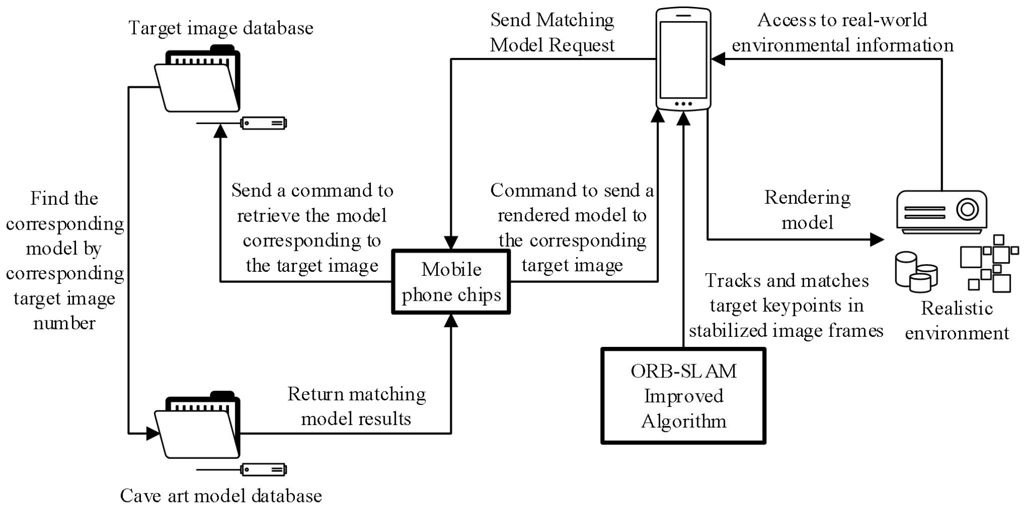


Figure 4. The matching process of grotto art works.

Let the artwork model matrix in the grotto artwork model database be U_s, and the feature data matrix in the target image feature database be Y_s, and notate the number of each data within each set of matrices $A, A \in A_s$, then the model relationship matrix within the artwork model database is shown in equation (11):

$$U = U_s \odot A_s \tag{11}$$

The feature data relationship matrix in the target image feature database is shown in equation (12):

$$Y = Y_s \odot A_s \tag{12}$$

According to the information labeling of A_s, the artwork model corresponding to the target image features can be found, and the correspondence between them is shown in equation (13):

$$U\left(A_s\right) \overset{A_s}{\rightleftarrows} Y\left(A_s\right) \tag{13}$$

After the projection of Eqs. (11)~(13), the artwork model will find the corresponding target image features, combined with the generation plane detected by the PDM tool, so as to determine the generation position of the artwork model in the exhibition environment, and then complete the establishment of the virtual exhibition environment of the entire grotto artwork.

3. Realization of virtual exhibition of grotto art based on spatial mapping

3.1. Spatial scanning and labeling of spatial data

Hololens is a version of the more favorable experience of mixed reality devices, which, with the support of the underlying SLAM technology, is manifested in the product form as Spatial Mapping, which is capable of scanning spatial scenes in real time and generating scene surface grids. In the functional development of the virtual system in this paper, the device is required to identify the real spatial scene, save the virtual object position through the spatial anchor point to ensure the positioning, and the ability of spatial comprehension is required to judge the ground, wall, and ceiling in order to realize the spatial placement of the virtual work.

(1) Spatial Scanning

Spatial scanning and understanding are mainly constructed with the help of Spatial mapping in HoloToolkit, which is a set of tool collections encapsulated by Microsoft based on the underlying APIs built into Unity to help quickly develop HoloLens applications using Unity integration. The spatial mapping component is mainly composed of three scripts, namely: spatial mapping observer script, spatial mapping management script, spatial mapping data source script. Their functions are as follows: the spatial mapping observer script inherits the spatial mapping data source script, regularly scans the surrounding environment, and generates or updates surface data. The spatial mapping management script is responsible for controlling whether the spatial mapping observer script works or not, and providing an interface to get the environment mesh list. For example, setting the mesh material, setting the shadow, setting the display and other operations. The spatial mapping data source script is responsible for storing and managing the mesh obtained from scanning, such as adding, updating, removing, extracting and other operations.

(2) Spatial surface mesh labeling

The spatial position Y coordinate of the spatial surface grid is compared with the spatial position Y coordinate of Hololens, and with the spatial position of Hololens as the coordinate origin, different spatial surfaces are judged to be the distance from the origin of Hololens, and then divided into different grid surface types and labeled, and the main types of grid surfaces divided into are the ceiling, the floor, and the wall, similar to walls, similar to floors, and horizontal platforms that are between ceilings and floors. During scanning, if the scanning ray collides with a known surface mesh, the "PlayspaceRaycast" function returns a result in the form of a "RaycastResult" that displays labeled information about the surface mesh.

3.2. Surface Mesh Planarization and Mapping

(1) Scene network planarization

According to the label information returned from the acquisition, confirm that the scanned mesh is a wall, floor, or ceiling, the surface triangle mesh acquired by scanning is not completely smooth, so it is necessary to do planarization of the mesh, the SpatialProcessing prefabrication of HoloToolkit components contains components to process the spatial mapping data. One of the SurfaceMeshesToPlanes.cs scripts will find and generate planes based on the spatial mapping data. Flat planes are used in the system to represent walls, floors and ceilings. This prefab also includes the RemoveSurfaceVertices.cs script, which removes vertices from the space-mapped mesh, as well as the ability to create holes in the mesh or remove redundant triangles that are no longer needed. By creating a flat precast as a master, the flat precast is used to overlay the scene mesh surface at the same locations.

(2) Generate maps for flat surfaces

After the scanned scene mesh is planarized, in order to make the scene more visual and distinguish different types of planes, it is necessary to give materials to different planes, for example, a material for the wall, a material for the ground, so we need to determine the specific types of planes, and here will be based on the planar normals and Y coordinates to determine whether it is a wall, a floor or a desktop.

4. Functional test of virtual exhibition of grotto art

The main content of this chapter is the performance test of the proposed ORB-SLAM improvement algorithm, and the application effect analysis of the virtual exhibition program of cave art. The performance test of the algorithm is in the form of comparing similar algorithms, and the application effect of the virtual exhibition program is reflected by the user satisfaction survey.

4.1. Performance test of ORB-SLAM improved algorithm

4.1.1. Assessment of alignment effects

In order to visualize the superiority of the proposed (A5) ORB-SLAM improvement algorithm, (A4) ORB algorithm, (A3) SIFT algorithm, (A2) SURF algorithm, and (A1) RANSAC algorithm are selected as the control, and the five algorithms are evaluated for the proportion of correctly estimating the keypoints (PCK), the root-mean-square-error (RMSE), and alignment time (Time) metrics on the two different experimental images of the grotto art. The performance of PCK, RMSE and Time of the five different algorithms on two different cave art experimental images, namely, (B1) illumination change and (B2) scale change, is shown in Table 1.On the B1 image, the PCK, RMSE and Time of the (A5) ORB-SLAM improved algorithm are 49%, 11.14 and 99ms in turn, which is the best among the five algorithms. On the B2 image, the PCK, RMSE and Time metrics of (A5)ORB-SLAM improved algorithm are 56%, 10.92% and 85ms in turn, which are still the best among the 5 algorithms. The percentage of correctly estimated keypoints in the two images of (A5)ORB-SLAM improved algorithm is always kept at 49.00% and above, and the alignment time of (A5)ORB-SLAM improved algorithm is always kept at 100ms and below, which shows that the algorithm in this paper can effectively balance the accuracy and efficiency of the alignment of the images of the cave works of art, and outputs more accurate alignment results.

Table 1. The registration performance of different algorithms.

Figure	Algorithm	PCK(%)	RMSE	Time(ms)
B1	A1	46	13.29	111
	A2	41	18.07	124
	A3	41	17.68	146
	A4	40	13.62	246
	A5	49	11.14	99
B2	A1	48	16.55	196
	A2	40	13.12	174
	A3	33	18.28	176
	A4	42	11.7	145
	A5	56	10.92	85

Three different grotto art experimental images, namely, (B3) lighting variation, (B4) scale variation and (B5) rotation variation, are selected to further evaluate the number of feature point matches (MP), the number of mismatches (MM), the matching correctness (MA) and the alignment time performance of five different algorithms as shown in Table 2. (A1)The RANSAC algorithm is good at extracting the number of feature points (>1000), but the number of mis-matches is high (>100) and the matching time is long (>100ms). (A2) The SURF algorithm similarly excels in extracting the number of feature points (>1000), while the number of mis-matches performs erratically (37~128), and the correct matching rate is low (<75.00). (A3) The SIFT algorithm extracts an insufficient number of features (<250), while the other metrics are moderate within the five algorithms, with a mediocre performance. (A4) The ORB algorithm is able to generate enough feature pairs, but the number of mis-matches is >75 and the alignment time fluctuates widely. In contrast, (A5) ORB-SLAM improved algorithm extracts a moderate number of feature points (500~600) and has the highest matching correctness (≥90.00%), and the shortest alignment time (≤95ms), which provides good real-time performance and the best overall performance.

Table 2. The feature extraction results of different algorithms.

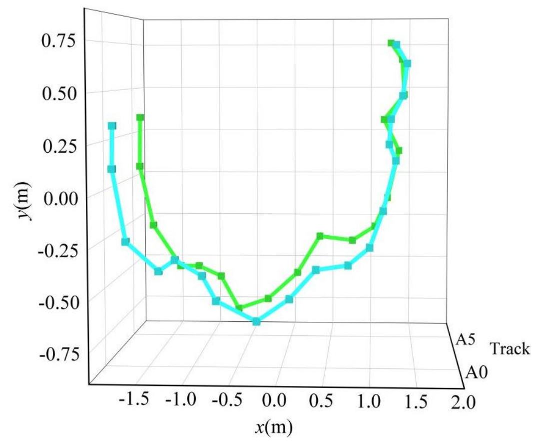
Figure	Algorithm	MP	MM	MA(%)	Time(ms)
B3	A1	1323	125	85.12	298
	A2	1244	37	58.08	88
	A3	240	117	68.36	95
	A4	1046	79	76.72	135
	A5	586	15	91.67	81
B4	A1	1180	116	85.18	268
	A2	1103	125	73.74	189
	A3	124	69	86.65	88
	A4	1020	86	82.72	161
	A5	568	25	96.41	95
B5	A1	1381	120	72.78	132

	A2	1061	128	67.23	97
	A3	148	98	76.66	89
	A4	980	88	74.21	95
	A5	516	25	94.14	78

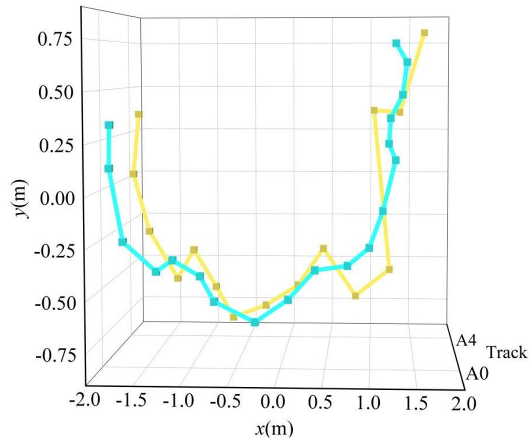
4.1.2. Positioning Performance and Error Analysis

In order to test the localization accuracy of this paper's (A5)ORB-SLAM improved algorithm in the process of practical application of virtual exhibition of grotto art, this subsection takes the B3, B4 and B5 images in the classic grotto art dataset as the experimental samples, and draws the performance of measured trajectories with the (A4)ORB algorithm and the (A3)SIFT algorithm. The positioning accuracy of the algorithm is quantitatively evaluated by comparing the measured trajectory of the algorithm with the real trajectory error. Among them, Fig. 5(a)-(c) shows the measured trajectory performance of the three algorithms on the B3 image, Fig. 6(a)-(c) shows the measured trajectory performance of the three algorithms on the B4 image, and Fig. 7(a)-(c) shows the measured trajectory performance of the three algorithms on the B5 image, and A0 is the true trajectory performance.

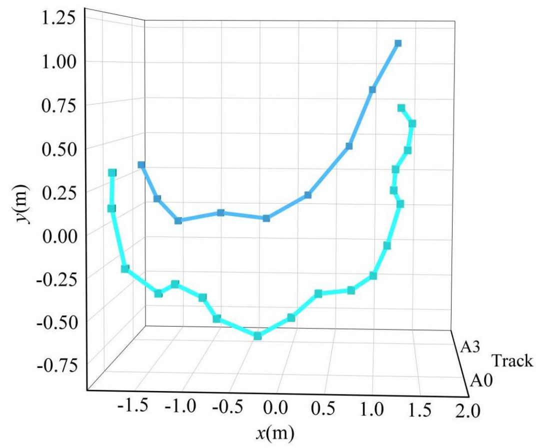
In Fig. 5, the maximum distance between the measured trajectory and the real trajectory of the (A5) ORB-SLAM improved algorithm is <0.01m, the maximum distance between the measured trajectory and the real trajectory of the (A4) ORB algorithm and the (A3) SIFT algorithm are both >0.2m, and the (A3) SIFT algorithm has the largest overall deviation. The (A5)ORB-SLAM improved algorithm measures the trajectory with the smallest deviation in the light change scenario.



(a) A5 Algorithm



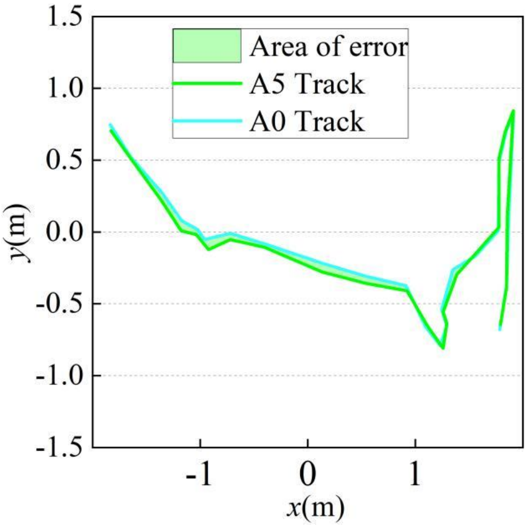
(b) A4 Algorithm



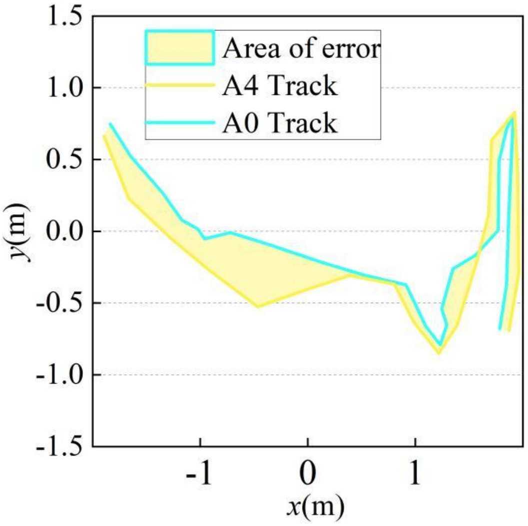
(c) A3 Algorithm

Figure 5. Sequence trajectory of B3.

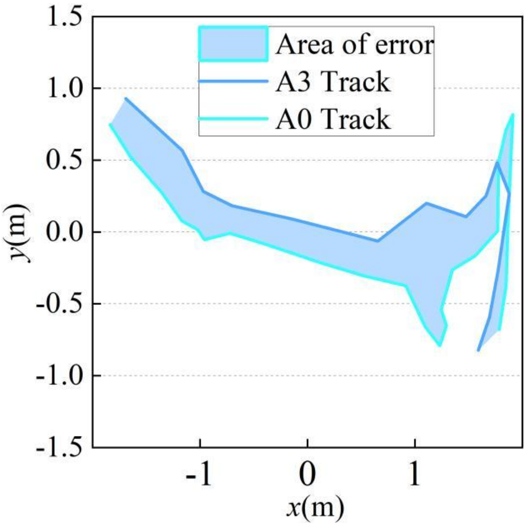
Observing the deviation area between the measured trajectory and the real trajectory of the three algorithms in the scale change scenarios in Fig. 6(a)-(c), the (A5) ORB-SLAM improved algorithm tends to be close to 0, while the deviation area between the (A4) ORB algorithm and the (A3) SIFT algorithm and the real trajectory are both >0.5m², and the (A3) SIFT algorithm has the largest deviation area and the lowest localization accuracy.



(a) A5 Algorithm



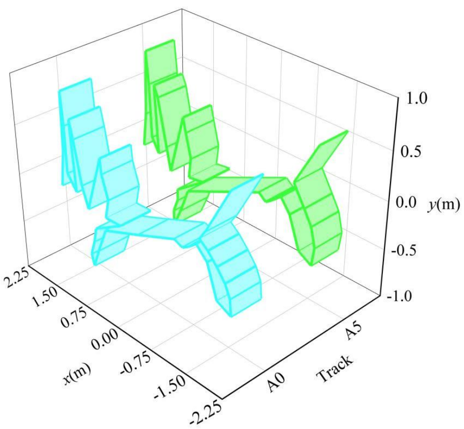
(b) A4 Algorithm



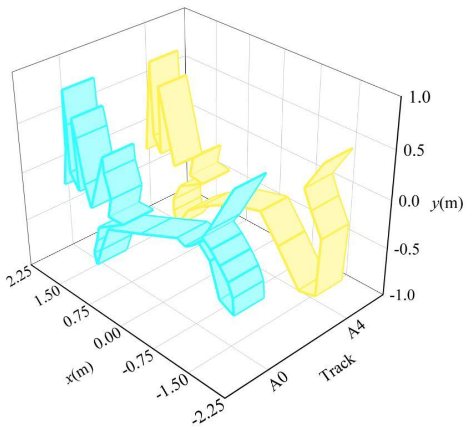
(c) A3 Algorithm

Figure 6. Sequence trajectory of B4.

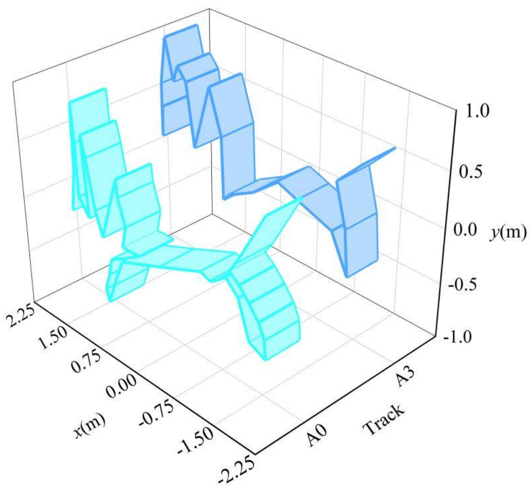
The performance of the three algorithms' measured trajectory deviation from the real trajectory in the rotational change scenario in Fig. 7 is consistent with the performance of the previous two scenarios, with the deviation distance of the (A5) ORB-SLAM improved algorithm consistently remaining <0.1m, converging to 0, and the (A3) SIFT algorithm's maximum deviation distance being the largest of the three algorithms (>0.3m).



(a) A5 Algorithm



(b) A4 Algorithm



(c) A3 Algorithm

Figure 7. Sequence trajectory of B5.

After the above analysis, the three algorithms' positioning accuracy performances are ranked from best to worst: (A5) ORB-SLAM improved algorithm, (A4) ORB algorithm, and (A3) SIFT algorithm. Among them, the ORB-SLAM improved algorithm can effectively cope with the challenges of rich scene changes, maintain stable and good positioning accuracy performance, and meet the demand for alignment accuracy in practical applications.

4.2. Assessment and analysis of the application of virtual exhibitions of cave art

In order to evaluate the actual experience of using the virtual exhibition of cave art, and to provide reference opinions for subsequent improvement and optimization. A trial run of the virtual exhibition of this paper on grotto art was carried out locally, and students from H Middle School were invited to the site to experience it. From (D1) immersion, (D2) fun, (D3) interactivity, (D4) information cognition, (D5) multidimensionality, and (D6) emotionality, a total of six dimensions, the content of the user satisfaction questionnaire was designed as follows:

(1) (D1) Immersion: (D11) I think the scene construction of grotto art has a sense of immersion, and (D12) I think the display of grotto art has a sense of realism.

(2) (D2) Interestingness: (D21) I think the exhibition process of the grotto art has a sense of freshness, (D22) I think the added related mini games are interesting.

(3) (D3) Interactivity: (D31) I think the dynamic modeling of the grotto art is necessary, (D32) I have a strong sense of participation in the exhibition process.

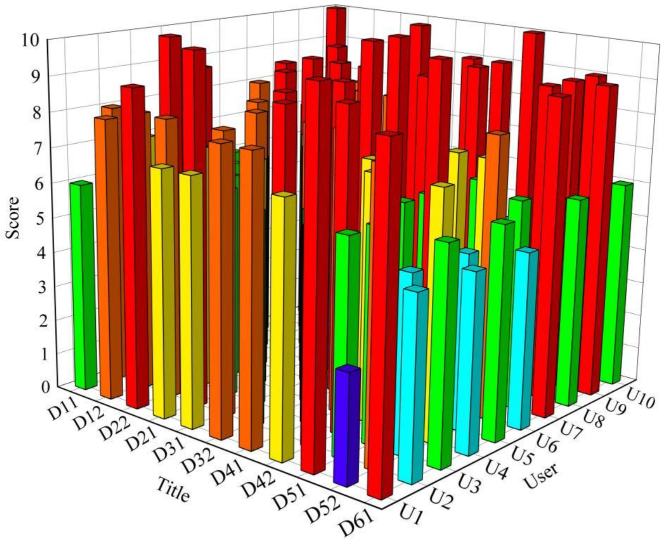
(4) (D4) Information Cognition: (D41) I think the audio announcement made my cognition clear, (D42) I think the textual prompts that appeared were complete.

(5) (D5) Multi-dimensionality: (D51) I think the venue factor does not affect the viewing experience

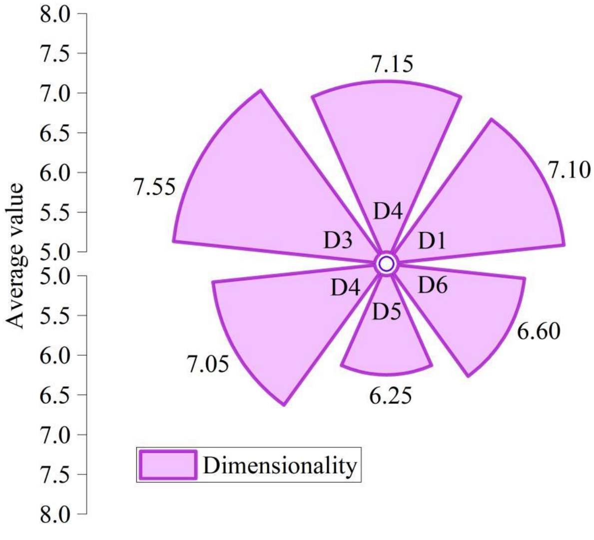
of the Cave Art, and (D52) I think the presentation angles of the Cave Art exhibits are comprehensive.

(6) (D6) Emotional: (D61) Emotions such as shock and awe appeared in my experience of the Cave Art exhibit.

The questionnaire contains a total of 11 questions, each of which has a full score of 10 points, and is divided into the following five grades: excellent (9-10), good (7-8), fair (5-6), passing (3-4), and failing (0-2). Ten students (U1-U10) were randomly invited to participate in the satisfaction survey of the virtual exhibition and interactive teaching of grotto art, and the satisfaction test was summarized in Fig. 8(a), and the average value of each dimension was calculated in Fig. 8(b).



(a) Satisfaction rating situation



(b) The average score of dimension satisfaction

Figure 8. Satisfaction test situation.

Overall the 10 students' rating values for this paper's virtual exhibition on cave art were mostly distributed in 6-10 points, i.e., the students thought that the overall performance of this paper's virtual exhibition on cave art was better. The average scores of (D1) immersion, (D2) fun, (D3) interactivity, (D4) information cognition, (D5) multidimensionality and (D6) emotionality are 7.10, 7.15, 7.55, 7.05, 6.25 and 6.60>6.00, which are all in the range of "average" or above. Students expressed some recognition of the effect of the virtual exhibition of grotto art in this paper, but the "multi-dimensional" and "emotional" dimensions should be optimized and improved.

5. Exploration of an Interactive Teaching Model for Virtual Cave Art

This chapter combines the form of the constructed virtual exhibition of cave art with a student-centered approach to propose a practical strategy for applying interactive teaching mode in the art classroom. It also sets up an application experiment to test the feasibility and reliability of the proposed interactive teaching model through controlled variables.

5.1. Practical Strategies for Applying Interactive Teaching Models in Art Classrooms

The importance of interest, as a source of motivation for students to learn, is self-evident. If art teachers can skillfully use interactive means to attract students' attention at the beginning of teaching, awakening their curiosity and desire to explore art knowledge, it will undoubtedly provide a solid foundation for the success of the whole class. Specifically, the interactive introduction strategy can be flexible, focusing on stimulating students' interest in participation, so that students from passive acceptance of knowledge to active exploration of knowledge.

First, according to the content of different grotto works of art carefully designed interactive topics. These topics should be close to students' life experiences or points of interest, such as discussing the story behind a piece of cave art, or exploring the connection between color and emotion, as a bridge to naturally transition to the content of the lesson, so that students can generate a strong desire for knowledge in a relaxed and pleasant atmosphere.

Second, gamification elements can be incorporated or performance activities can be organized. As an effective way of interactive introduction, games can not only quickly focus students' attention, but also help to cultivate students' observation, imagination and teamwork skills. For example, in simple and interesting art games such as “I Draw You Guess” and “Color Solitaire”, teachers can guide students to act out stories related to artworks in groups, so as to stimulate students' interest in learning and cultivate their artistic perception. Taking the teaching of “Ticket Design” as an example, teachers can design a lively and inspiring classroom introduction to stimulate students' interest in the design of tickets for grotto art exhibitions. Before the class, the classroom can be set up to simulate the construction of a miniature “Cave Art Exhibition Entrance”, and display a series of carefully designed ticket samples, each ticket is painted with the characteristics of different cave art works. As the bell rings, students enter the special “Cave Art Exhibit” and express their curiosity and excitement. Then, the teacher can guide the students to start a group discussion on the topic of “If you were the designer of the Cave Art Exhibition, how would you design an attractive ticket”, encouraging the students to share their own travel experiences or imaginative journeys, and exploring what elements make the ticket unique and attractive, such as color, pattern, text messages etc. In this way, interactions that are close to students' life experiences and interests can naturally lead them into learning about the design of tickets for cave art exhibitions.

5.2. Effectiveness of the implementation of the interactive teaching model

Classes H1 and H2 of secondary school H, each with 45 students, were selected as the experimental group (G1), while class H2, the control group (G2), was taught in the traditional mode of learning cave art. After one semester of teaching, a test on the theme of “Cave Art” was assigned, and each work was graded by the teacher. Each piece of work was graded out of 100 points and was categorized into three grades: excellent (85-100), moderate (60-84), and poor (0-59). The performance of the work of the two classes is summarized in Figure 9.

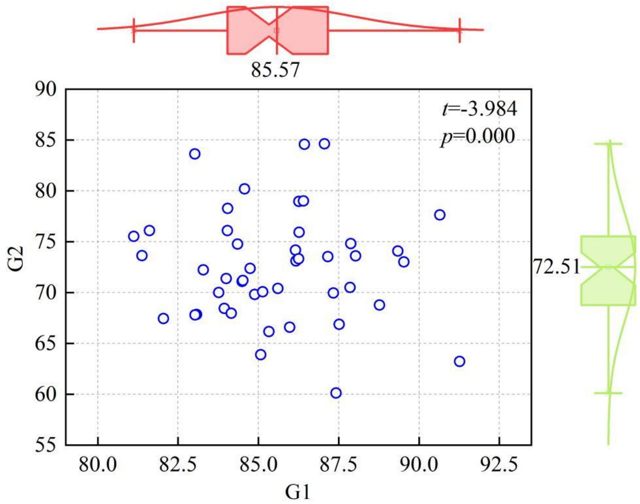


Figure 9. A comparison of students' works.

It can be seen that (G1) the experimental group of students' work ratings are concentrated in the [75,95] range, with a mean value of 85.57 and a rating of excellent. (G2) The control group students' work ratings are concentrated in the [55,95] range, with a large difference, the mean value is 72.51, and the rating is moderate. And the two groups of students' work scores performance $p=0.000<0.05$, with statistically significant differences, verifying that this paper's virtual exhibition and interactive teaching mode on the students' grotto art connotation of learning, absorption and application of the positive effect of promotion.

6. Conclusion

This paper constructs a model of grotto art works based on ORB-SLAM improved algorithm, proposes a virtual exhibition program based on spatial mapping, and designs a practical way of interactive teaching mode for grotto art works.

Among them, the ORB-SLAM improved algorithm shows image matching performance and robustness far beyond similar algorithms in terms of matching accuracy, feature point extraction efficiency, and localization trajectory error. In a variety of viewpoints, rotation, scale and lighting changes and other complex scenarios, the proportion of correct estimation of the key points is maintained at 49.00% and above, while the alignment time is always stable at 100ms and below, and the error between the image measurement trajectory and the real trajectory tends to be close to 0m. Cave art virtual exhibition program in the trial run, there is a sense of immersion, fun, interactivity, and information cognition of a total of four dimensions to obtain the user's satisfaction level “Good” evaluation (>7.00), to meet the user's experience needs.

Grottoes artwork modeling technology under the interactive teaching mode, not only to assist students' artwork up to the “excellent” rating (average value of 85.57), and demonstrated compared to the traditional teaching mode obvious advantage of the enhancement of the effect (students' work scores performance $p = 0.000 < 0.05$). The combination of virtual exhibition technology and interactive teaching mode can create a rich learning environment for students to learn cave art works, enhance students' understanding and utilization of the connotation of cave art works, and promote the inheritance and innovation of cave art.

Funding

This research was supported by the 2022 Longdong University Youth Science and Technology Innovation Fund Project "Gansu Modern and Contemporary Art Culture Research" (XYSK2204).

References

1. Runze, Y. (2023). A study on the spatial distribution and historical evolution of grotto heritage: a case study of Gansu Province, China. Heritage Science, 11(1), 165.
2. Li, G. S., Wang, W. F., Qu, J. J., & Song, M. H. (2013). Study on temperature and humidity environment of grotto 72 at the Mogao Grottoes in Dunhuang, China. International journal of climatology, 33(8).
3. Xia, J., Liu, H., Qiao, H., Zhu, L., & Liu, Y. (2024). A study for the distribution characteristics of surface temperature and the protection of grotto temples in China. Environmental Monitoring and Assessment, 196(12), 1248.
4. Yeshi, W., & Champadaeng, S. (2025). A Study of Art History: A Survey of Yungang Grottoes by History Students. Journal of Practical Studies in Education, 6(5), 1-7.
5. Zhigang, Y. (2023). An Analyses of Sculpture Patterns of Mount Xumi Grottoes in China. Pakistan Social Sciences Review, 7(3), 921-934.
6. Ling, J., & Gao, F. (2017). Cultural Analysis of Yungang Grottoes under Multicultural Background. In Asia International Symposium on Language, Literature and Translation (p. 498).
7. Wu, M. (2022). A Comparative Study of the Aesthetic Characteristics of Buddhist Cave Image Art in the Tang and Song Dynasties. Arts Studies and Criticism, 3 (1): 71, 73.
8. Chen, W. (2025). Aesthetic Consciousness Reflected in Prehistoric Art. In Study of Prehistoric Aesthetic Consciousness of the Chinese Nation: “Civilization” Before Civilization (pp. 143-187). Singapore: Springer Nature Singapore.
9. Liu, R., Fauzi, T. A., Suhaimi, S. N., & Rozlan, R. M. (2025). The Aesthetic of Ruins in Dunhuang Murals. Idealogy Journal, 10(1).
10. Zhang, J., Kang, K., Liu, D., Yuan, Y., & Yanli, E. (2013). Vis4heritage: visual analytics approach on grotto wall painting degradations. IEEE Transactions on Visualization and Computer Graphics, 19(12), 1982-1991.
11. Wu, L., Zhang, L., Shi, J., Zhang, Y., & Wan, J. (2022). Damage detection of grotto murals based on lightweight neural network. Computers and Electrical Engineering, 102, 108237.
12. Li, X., Yusof, M. J. M., Hiong, S. W., & Perumal, V. (2021, April). The use of digital advancement technology in protecting the Longmen Grottoes Art. In Journal of Physics: Conference Series (Vol. 1875, No. 1, p. 012011). IOP Publishing.
13. Tribhuvan, A. P., & Binnaser, A. A. (2024). Exploring the digital heritage of Ajanta Caves through VR technology. Digital Applications in Archaeology and Cultural Heritage, 35, e00381.
14. Rivero, O., Dólera, A., García-Bustos, M., Eguilleor-Carmona, X., Mateo-Pellitero, A. M., & Ruiz-López, J. F. (2024). Seeing is believing: An Augmented Reality application for Palaeolithic rock art. Journal of Cultural Heritage, 69, 67-77.
15. Blanco-Pons, S., Carrión-Ruiz, B., & Lerma, J. L. (2019). Augmented reality application assessment for disseminating rock art. Multimedia Tools and Applications, 78(8), 10265-10286.
16. Duval, M., Smith, B., Gauchon, C., Mayer, L., & Malgat, C. (2020). “I have visited the Chauvet Cave”: the heritage experience of a rock art replica. International Journal of Heritage Studies, 26(2), 142-162.
17. Li, J., & Kieanwatana, K. (2024). The Impact of Augmented Reality Tourism on Tourist Experience and Behavior—A Case Study of the Mogao Caves Cultural Heritage. International Journal of Sociologies and Anthropologies Science Reviews, 4(5), 103-122.
18. Boboc, R. G., Băutu, E., Gîrbacia, F., Popovici, N., & Popovici, D. M. (2022). Augmented reality in cultural heritage: an overview of the last decade of applications. Applied Sciences, 12(19), 9859.
19. Avci, E. (2019). Enhancing the cultural tourism experience through augmented reality. Co-Editors, 215.
20. He, Z., Wu, L., & Li, X. R. (2018). When art meets tech: The role of augmented reality in enhancing museum experiences and purchase intentions. Tourism Management, 68, 127-139.
21. Ieiri, Y., Tsai, H. Y., Mizobuchi, K., Tengfei, S., & Hishiyama, R. (2023). Design of traditional cultural experiences using augmented reality based on environmental presence. IEEE Transactions on Human-Machine Systems, 53(2), 390-400.
22. Kuo, Y. T., Garcia Bravo, E., Whittinghill, D. M., & Kuo, Y. C. (2024). Walking into a modern painting: The impacts of using virtual reality on student learning performance and experiences in art appreciation. International Journal of Human–Computer Interaction, 40(23), 8180-8201.
23. Bäck, R., Plecher, D. A., Wenrich, R., Dorner, B., & Klinker, G. (2019, March). Mixed reality in art education. In 2019 IEEE conference on virtual reality and 3D user interfaces (VR) (pp. 1583-1587). IEEE.
24. Nanu, A., Titieni, A., Nedelcu, M., Nedelcu, F., & Sarbu, C. (2013). Use of virtual reality for artistic education. In ICERI2013 Proceedings (pp. 47-56). IATED.
25. Erkut, C., & Dahl, S. (2019). Incorporating Virtual Reality with Experiential Somaesthetics in an Embodied Interaction Course. The Journal of Somaesthetics, 4(2).