

Research on the application of virtual simulation technology in the protection and inheritance of intangible cultural heritage—taking chinese language and writing as an example

Li Chen^{1,✉}

¹ Department of Chinese Language and Literature, Pingdingshan Vocational and Technical College, Pingdingshan, Henan, 467000, China

ABSTRACT

Cultural heritage represents the historical and cultural achievements of a nation, playing a vital role in studying human civilization and preserving national languages and scripts. This study utilizes virtual simulation technology to design a virtual pavilion for Chinese language and writing, employing image and text feature extraction algorithms for feature fusion and 3D modeling. The effectiveness of Chinese character extraction is validated through feature point matching, while the virtual exhibition's impact is assessed via user experience scores. Results indicate that the proposed algorithm achieves accurate extraction with no misrecognition. User interest rankings highlight text images as the most influential factor, followed by visual imagery, pavilion experience, scene art, and language culture. Analysis of user feedback shows an average experience score exceeding 60 points, confirming the pavilion's effectiveness in preserving and promoting Chinese language and writing culture.

Keywords: virtual simulation technology, text feature extraction, three-dimensional modeling, user experience

1. Introduction

Intangible cultural heritage is closely related to people's lives and refers to practices, performances, expressions, knowledge and skills, as well as their associated tools, objects, artifacts and cultural spaces, which are recognized by groups, communities and individuals as their cultural heritage to be transmitted from generation to generation [12, 2, 5]. Intangible cultural heritage is a collection of ideas of each local folk culture and art creators, unique artistic achievements tested by time, and is

✉ Corresponding author.

E-mail address: 19937516797@163.com (L. Chen).

Received 19 August 2024; Revised 27 October 2024; Accepted 10 November 2024; Published Online 28 March 2025.

DOI: [10.61091/jcmcc125-21](https://doi.org/10.61091/jcmcc125-21)

© 2025 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

a valuable asset left by the development and practice of human society [14, 7]. In the past, based on the background of the times, the underdevelopment of science and technology has led to fewer channels for people to obtain forms of entertainment, and traditional arts are very popular with the general public, but now there are more or less intangible cultural heritage inheritance faults or even disappearance of the problem. With the addition of virtual reality interactive technology, the dissemination of intangible cultural heritage will be faster, and the dissemination effect will be more obvious compared to the creation of unilateral access to information forms, effectively solving the problem of difficult inheritance [15, 10, 6].

As an intangible cultural heritage in the process of human civilization development, Chinese language and writing are the treasures of Chinese culture and an important carrier of Chinese literature, and the richness and polysemy of the Chinese language provide a broad space and unlimited imagination for the creation of Chinese literature [16, 11, 8]. Chinese language and writing have rich cultural connotations, and one can gain a deeper understanding of the connotations and development of ancient Chinese culture by studying the forms, meanings, and development history of Chinese characters [17, 1]. Chinese characters also contain rich cultural symbols and meanings, the study of which can reveal the deeper connotations of ancient Chinese culture [4, 3]. Therefore, the inheritance of Chinese characters ensures the inheritance of the artistic charm and cultural value of literary works, which is of great significance to the Chinese nation.

In this paper, a virtual pavilion of Chinese language and writing is designed through virtual simulation technology to provide an immersive experience with the help of digital technology to promote the dynamic inheritance and protection of language and writing. The study adopts Unity 3D technology and image text fusion technology to construct a 3D model of the virtual pavilion, and analyzes the recognition effect of the proposed image text feature extraction algorithm. The cultural inheritance effect of the virtual exhibition hall is concluded from the users' interest in the virtual exhibition hall of Chinese language and their own experience feedback results.

2. Design of virtual simulation pavilion for chinese language and literature

2.1. *Virtual pavilion creation concept*

2.1.1. Design conceptualization. Firstly, at the level of user experience design, the first-person perspective is adopted to fix the user's visual focus, increase the user's sense of immersion and presence experience, and ensure that the user can maintain a positive mental exploration of the virtual pavilion.

Secondly, at the level of visual guidance, the two-dimensional first of all to ensure the quality of the mapping in the space so as to restore the realism of the scene as much as possible, but also need to ensure the consistency of the UI interface and the entire pavilion; modeling needs to be fully foreseen in the content of the interaction, the three-dimensional model to ensure the integrity of the space, and if necessary, to carry out artistic processing.

Finally, at the level of content narrative, design the visiting route in line with the user's cognitive logic, and at the same time, not too straightforward, and appropriately create "surprises" to enhance the user's enthusiasm for obtaining information, but also to ensure the correctness of the user's interactive logic.

In addition to the above three points, but also fully mobilize the lighting and sound to enrich

the design of the venue, increase the atmosphere, to provide users with a more favorable visiting experience.

2.1.2. System development process. The realization stage of the Chinese language and writing virtual display system integrates the plane visual content and 3D model into a scene, and then completes the design of the system interactive content through coding, so as to realize the complete function of the Oracle virtual display system, which is the process from the conceptual theory to the actual effect display.

First of all, according to the design architecture of the pavilion and the main story line to sort out the material induction partition; then complete the graphic display design content, covering the exhibition wall design, UI design, mapping processing, design elements, etc.; Next, according to the historical materials and historical scholars' research, the use of H-dimensional software to produce a restoration model of the Yinshang Palace in line with the study, and according to the characteristics of the building, the design architecture, and to improve the design of the relevant exhibits in the virtual pavilion, Next, we use H-dimensional software to create a restoration model of the Yin Shang Palace in line with this study, and according to the architectural characteristics and design structure, we improve the design of related exhibits, games, decorations and other models in the virtual exhibition hall, and integrate the three-dimensional space; then, we improve the UI interface in C4D, and preset all the interactable interfaces; we pack and export Unity3D, and according to the layout of the interface and the function modules, we improve the human-computer interaction, roaming, merging and collision and other function settings, and then we pack and release the apk file and export the VR glasses to conduct the function test.

2.2. Virtual simulation display scene design

Scene immersion function: realizing scene immersion can satisfy the users to walk, watch and learn in the scene in the way of first-person perspective, through this function, the users can have an all-around viewing and learning of Chinese language and characters.

Display function: display includes the realization of immersive action display and scene display function, light and sound simulation of reality.

2.2.1. Virtual display scenario development. Virtual reality technology as a multidisciplinary fusion technology, the use of computer graphics, signaling and other technologies, to create and the real world as the same visual, auditory, tactile virtual environment. Users interact with the virtual environment, the user's relevant action data to the computer simulation of the virtual world, the virtual world perceived to make feedback, produce feedback signals to the user [9]. With a mathematical model to represent the virtual reality technology as:

Definition $F(x)$ is the selection of a transformation function to capture the user's action data in the real world:

$$Z_0 = F(O), \quad (1)$$

where o serves as the set of action states of the user in the real world and Z_0 represents the collected data.

To make the transmission of Z_0 convenient, Z_0 needs to be deformed and defined as the transformation function $G_0(x)$:

$$Z_l = G_0(Z_0). \quad (2)$$

The transformed user data is denoted by Z_l , and then the work machine parses Z_l and the parsed function is $G_l(x)$:

$$Z_2 = G_l(Z_l). \quad (3)$$

Here Z_2 represents the user data parsed by the working machine.

The working machine then reconstructs the virtual scene using the Z_2 data, defining the process as $P(x)$ function $W = P(Z_2)$

W represents the virtual scene after reconstruction by the working machine, resulting in a virtual scene representation as:

$$W = P(G_l(G_o(F(O)))). \quad (4)$$

Virtual reality technology as a mapping function, can be the input data scene, the real world behavior as the input side of the function, the virtual scene changes as the dependent variable as the output side of the function, the virtual scene will change because the user produces changes and make corresponding changes, the expression is:

$$V_i = (X_i, Y_i, Z_i)(i = 1, 2, 3, \dots, n), \quad (5)$$

where, X, Y represents the water plane coordinates of the ground and Z represents the vertical coordinate values.

2.2.2. Construction of scene models. Scene modeling process: three-dimensional modeling software settings \rightarrow import two-dimensional drawings \rightarrow create the main body of the building \rightarrow to the building model mapping \rightarrow adjust the way of mapping \rightarrow export the model and other processes, modeling mode process shown in Figure 1. The principles to be followed in model construction:

- 1) In order to maximize the appearance of the model, try to reduce the number of surface pieces of the model surface so that the model surface is fully simplified.
- 2) Ensure the rationality of the model mesh distribution.
- 3) When modeling, it is necessary to align points with points, lines with lines, and when welding the model, pay special attention to the precise correspondence of the welding points, so as to ensure the stability of the structure.
- 4) Reasonable use of mapping to ensure the appearance of the 3D model.

Environment modeling includes building modeling, ground modeling, and plant modeling. Building modeling adopts polygonal modeling method, and after establishing a simple model, texture mapping ensures the sense of realism. Plant modeling calls the vegetation resource package in the virtual reality engine and imports the prepared vegetation into the resource package of the virtual reality engine. In the virtual reality engine, the vegetation file is added to the ground through PlaceTress. When importing the plant model, design and plan the parameters of the size, position and number of plants.

2.2.3. Virtual scene roaming technology. Artificial roaming refers to the user through the manipulation of the keyboard, mouse and other input devices to realize the arbitrary roaming in the three-dimensional scene. Through artificial roaming, users can flexibly and accurately observe the scene in all directions, pick up virtual objects and obtain information. The process of artificial roaming is a process of continuously changing the viewpoint position or view direction and rendering the scene according to the interactive control commands. The interaction control between the user and the scene mainly consists of basic actions such as moving forward, backward, moving left, moving right, looking up, looking down, turning left, turning right, ascending, descending and so on.

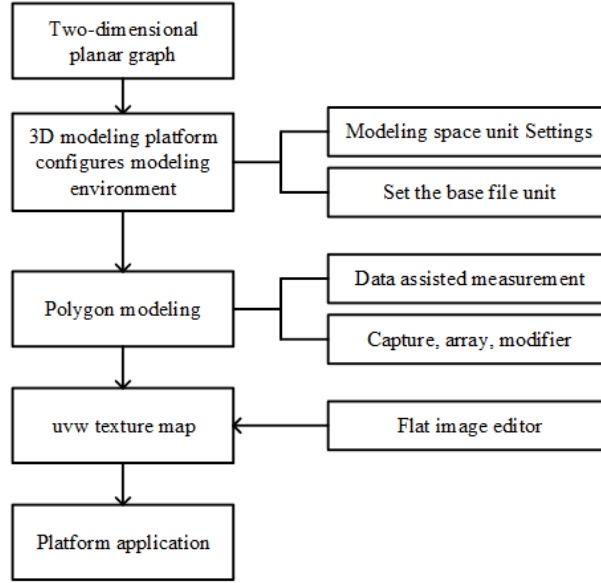


Fig. 1. Modeling method process

Forward and backward simply move the point of view a certain distance along the line of sight, the direction of the line of sight remains unchanged. Let the coordinates of the point of view before the transformation be $E(E_x, E_y, E_z)$, and the coordinates of the point of view after the transformation be $E'(E'_x, E'_y, E'_z)$, then

$$E' = E + vL W = L, \quad (6)$$

where v is the traveling speed and W is the new line-of-sight unit vector.

Left shift and right shift are just translations of the viewpoint along the x_e -axis direction of the viewpoint coordinate system, and the direction of the line of sight remains unchanged. The direction of the line of sight remains unchanged.

$$E' = E + S_x N W = L, \quad (7)$$

where S_x is the translation

Looking down and looking up refers to the process of changing the scene that the user sees with his head down and his head up. Looking down and looking up do not change the position of the viewpoint, but only the direction of the line of sight changes so that the direction of the line of sight is rotated by a certain angle around the x_e axis of the viewpoint coordinate system. Let the angle of rotation be θ , then:

$$E' = E \quad W = LR(\theta) \quad R(\theta) = TR_x(\alpha)R_y(\beta)R_z(\theta)R'_x(\alpha)R'_y(\beta)T', \quad (8)$$

where T is the transformation matrix of translating the origin of the viewpoint coordinate system to the origin of the world coordinate system, $R_x(\alpha), R_y(\beta)$ is the transformation matrix of rotating the x_e -axis of the viewpoint coordinate system around the x_w -axis by an angle of α to the $x_w O z_w$ -axis, and rotating the y_w -axis by an angle of β to the z_x -axis, and $R_x(\alpha), R_y(\beta), T'$ is the transformation matrix of rotating around the x_w -axis and the y_w -axis in the opposite direction and translating in the opposite direction, respectively.

Turning left and turning right refers to the process of scene change when the user turns left and turns right. Turning left and turning right does not change the position of the viewpoint, but only

the direction of the line of sight changes, so that the direction of the line of sight rotates around the y_e axis of the viewpoint coordinate system by a certain angle. Let the angle of rotation be θ , then:

$$E' = E \quad W = LR(\theta). \quad (9)$$

The calculation of $R(\theta)$ is similar to that of $R(\theta)$ in the overhead and elevation views, except that the rotation around the viewpoint coordinate system x_e is replaced by rotation around the viewpoint coordinate system y_e axis.

Ascending and descending only increase or decrease the height value of the viewpoint (y coordinate), while the direction of the line of sight remains unchanged, i.e.,

$$E' = E + \nu_y M \quad W = L, \quad (10)$$

where ν_y is the amount of change in the height of the viewpoint.

The above cases are the basic ways for users to interact with the scene, and in practical applications, it is often a combination of the above ways to show the user a broader view.

Auto roaming is the process of pre-setting the roaming path so that the viewpoint travels along the roaming path, thus achieving the purpose of roaming in the 3D scene. The roaming path can be set in the scene plan by clicking the control points with the mouse. Users use the mouse to pick up a series of control points on the floor plan, and specify the elevation and speed of each control point, and then through the 2D screen coordinates to the 3D scene coordinates of the conversion to get a 3D space control point sequence. A roaming path is a curve in three-dimensional space, which is determined by the control points according to a certain interpolation method.

A common method for collision detection in virtual scenes is to use a wraparound box pseudo that wraps around the object object. The basic idea of the enveloping box method is to use simple geometry instead of complex geometry. First of all, a rough detection of the object's enclosing box, when the enclosing box is intersected by the enclosing geometry may intersect when the enclosing box is not intersected by the enclosing geometry must not intersect, so that a large number of impossible to intersect the geometry and geometric parts, so as to quickly find the intersecting geometric parts.

Typical types of bounding box are: bounding sphere, bounding box AABB along the axes, directional bounding box OBB and so on. The choice of a wraparound box usually depends on its application area and the different constraints it implies. For collision detection algorithms, they can be analyzed with the help of a consumption function.

$$T = N_v C_v + N_p C_p + N_u C_u, \quad (11)$$

where T is the total cost of collision detection, N_v is the number of pairs of enclosing boxes involved in overlap detection, C_v is the cost of a pair of enclosing boxes to do the overlap test, N_p is the number of pairs of geometric elements involved in intersection test, C_p is the cost of a pair of geometric elements to do the intersection test, N_u is the number of nodes of enclosing box hierarchy that need to be modified after the object's motion, and C_u is the cost of modifying a node.

3. Virtual presentation of chinese language and writing

3.1. Design process

By summarizing and summarizing the contents of the previous chapters, a preliminary understanding of the overall processing flow of language and text image processing can be made, and the specific process schematic is shown in Figure 2:

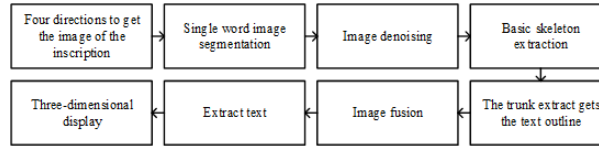


Fig. 2. Image processing flow of inscription

For each step specific descriptions are as follows:

1) Language and text image acquisition. The research samples in this paper are all acquired by using high-definition shooting, thus reducing the interference factors in the captured language text images and thus highlighting the text edge features in the images. In addition, according to the carving characteristics of the language text, the image is acquired by oblique lighting from the top, bottom, left and right directions of the same language text.

2) Image Preprocessing. This paper first analyzes and discusses the collected images, analyzes the characteristics and problems of the language text images, and then carries out the language text image denoising, single word image segmentation and basic text skeleton extraction.

3) Text outline extraction. On the basis of language and text image preprocessing, according to the theoretical methods of mathematical morphology and wavelet transform, the text outline in the language and text image is extracted using the backbone extraction algorithm based on Chinese character strokes.

4) Image Fusion. A simple image fusion operation is carried out on the language text images after font outline extraction, i.e., the language text images of the same target in the four oblique directions of the upper, lower, left and right lightening are processed and fused into the image, so as to obtain a more complete binary image of the font outlines. Finally, manual checking is carried out to delete a small number of unreasonable areas to get the language text results.

5) Three-dimensional display of language text. For the three-dimensional display of the text outline of the extracted language text image, a three-dimensional world museum scene model is first constructed, and then the acquired language text is embedded for three-dimensional display.

3.2. Extraction of chinese language image text

3.2.1. Preprocessing of chinese language image texts. Digital cameras shoot color images, and text extraction does not require color information, so the first color image grayscale, and then make a low-pass filter. Because of the text to try to retain the information, especially the text edges, so here can only do a slight low-pass filtering.

Because oblique light may cause uneven illumination of the image, so it is necessary to apply homomorphic filtering to the image. The process of homomorphic filtering is as follows: for an unevenly illuminated language text image $f(x, y)$, firstly, the incident light and reflected light are partially separated using the logarithmic transformation, as shown in Eq. (12):

$$f(x, y) = i(x, y) \cdot r(x, y). \quad (12)$$

Then there is $z(x, y) = \ln f(x, y) = \ln i(x, y) + \ln r(x, y)$, where $i(x, y)$ represents the incident

light and $r(x, y)$ represents the reflected light. The Fourier transform is then performed to convert the representation of the image from the spatial domain to the frequency domain, followed by the selection of a suitable filter for filtering and denoising the image, followed immediately by the Fourier inverse transform and exponential transform.

3.2.2. Chinese language image text stem extraction. The process of text outline extraction for language text images can be regarded as a process of information extraction, and for preprocessed language text images, it is the process of extracting text outlines from the background with noise to get new language text images. For the research samples in this paper, on the basis of the main extraction algorithm of existing research, combined with the analysis of the text structure in the language and text images in the previous paper, the original algorithm is improved, and an outline extraction method based on the stroke characteristics of Chinese characters is proposed.

The algorithm in this paper completes the extraction of text outline in three steps, the specific steps are as follows: the first step, the determination of the width of the horizontal, vertical, apostrophe, and downstroke of Chinese characters, which is to utilize the principle of wavelet multi-scale edge detection of the image to carry out the wavelet analysis of the text outline in the language image, and then to determine the width of the corresponding Chinese characters; the second step of the main extraction algorithm, which is to filter out the noise and details together, and to get the noise and details from the language image, and the main extraction algorithm is to filter out the noise and details together, and then to extract the text structure. The second step of the trunk extraction algorithm is to filter out the noise and details together to get the trunk of the text; the third step utilizes the trunk expansion algorithm, the method is to take the trunk of the text obtained in the second step as the starting point, and based on the widths of the Chinese characters determined in the first step in the horizontal, vertical, apostrophe, and downward widths of the Chinese characters, the details of the text that satisfy the conditions are included to get the whole outline of the text.

1) Wavelet analysis of Chinese character outline. In this paper, the smooth function chosen is the cubic B-spline function, and thus the obtained quadratic B-spline wavelet with tight branching set is characterized by odd symmetry and no time or space offset, which in turn can be used to detect the edges of the Chinese characters in the image, and the specific implementation process is shown as follows.

For the grayscale figure $f(x, y)$ of a language text image, in order to detect and obtain the edges of the four stroke structures of its Chinese characters: horizontal, vertical, apostrophe, and press, so the image is subjected to a discrete binary wavelet transform, which is specified as a J-layer decomposition on the rows and columns: The transform formula in the row direction is:

$$\left\{ \begin{array}{l} SH_j f(x, y) = \sum_k h_k SH_{j-1} f(x, y + 2^{j-1}k) \\ WH_j f(x, y) = \sum_k g_k SH_{j-1} f(x, y + 2^{j-1}k) \end{array} \right\} \quad (j = 1, 2, \dots, J; \quad y = 1, 2, \dots, n). \quad (13)$$

The transformation formula in the column direction is:

$$\left\{ \begin{array}{l} SL_j f(x, y) = \sum_k h_k SL_{j-1} f(x, y + 2^{j-1}k) \\ WL_j f(x, y) = \sum_k g_k SL_{j-1} f(x, y + 2^{j-1}k) \end{array} \right\} \quad (j = 1, 2, \dots, J; \quad y = 1, 2, \dots, m), \quad (14)$$

where, j is the number of layers for which the wavelet transform is performed.

When $j - 1$, we have

$$SH_{j-1}f(x, y) = SH_0f(x, y) = f(x, y), \quad (15)$$

$$SL_{j-1}f(x, y) = SL_0f(x, y) = f(x, y), \quad (16)$$

where $SH_jf(x, y)$ and $SL_jf(x, y)$ denote the corresponding smoothing filtering results of the language text image, while $WH_jf(x, y)$ and $WL_jf(x, y)$ denote the results of the wavelet transform of the language text image, respectively.

2) Trunk Extraction Algorithm. For the process of extracting the text trunk of the image, the essence is the process of using structural elements to perform corrosion operations on the language text image. In the actual experimental processing, the appropriate structure operator can be selected for the operation according to the width of the character backbone and the width of the image noise lines. It is set that in the language text image, 0 is used to denote the white background, and 1 is used to denote the black pixel point corresponding to the text and noise points in the image. Based on the analysis of the image noise of the study sample, the 3×3 all 1 template is chosen as the structure operator in this paper. Further observation of the language text image shows that the region formed by the connection of individual noise points in the image may be relatively large, in order to better obtain the textual backbone, so a secondary judgment is made on the 5×5 neighborhoods of pixels that satisfy the structural operator, and if there are fewer pixel points that satisfy the structural operator, they are regarded as noise points, which makes the extracted textual backbone free of noise regions.

3.2.3. Image fusion. In this paper, four binary language and text images of the same object are fused, and in the specific processing, the four images are first divided into two groups of two binary language and text images, respectively, up and down diagonal lightening language and text images as a group, and left and right diagonal lightening language and text images as a group; and then the corresponding image fusion is carried out for each group: finally, the binary language and text images of the two groups are fused again, and the overall effect is better. Finally, the two groups of binary language and text images are fused again to obtain a better overall effect of the language and text images.

According to this paper, the fused images are obtained from different angles for the same object, in the process of image fusion, this paper adopts the control point correction method, i.e., through geometric correction, to eliminate the differences of the language and text images in the scale, translation, rotation, etc., and then the binary language and text images will be subjected to the phase-or operation at pixel level to obtain a synthesized binary language and text image.

3.3. Three-dimensional modeling and display of Chinese language image characters

3.3.1. VAML language. VAML as a new modeling language, its main role is used for modeling three-dimensional virtual scenes, which is closely related to multimedia communication, virtual reality and other technical fields. Virtual reality is the ultimate goal of combining with multimedia technology, the biggest feature of virtual reality is expressed in the fact that it allows people to be in a virtual realistic three-dimensional world, can interact naturally with this virtual world, and can change the point of view arbitrarily. Bring the interaction into the, so that visitors can interact with it, there is an immersive intuitive feeling on the exchange of documents, description of the standard, based on the realization of the virtual reality environment provides a viable solution [13].

3.3.2. Modeling process. First, we build various 3D objects in 3DStudioMAX, and then export these objects in .wrl format into a VRML world, edit their positions and relationships in the virtual reality world to form a unified whole world. Since the rectangular language text belongs to regular geometry, we use direct programming vrml code to generate the geometric model of the language text, and then add texture mapping for the texture node of VRML.

4. Chinese language and writing virtual simulation experience and feedback

4.1. Analysis of the effect of image text extraction algorithm

For any input text image, the algorithm extracts its feature points and traverses the comparison with the language and text database, and if the number of matching feature points is greater than the set subordinate judgment threshold, then it determines that the input text image matches the text in the standard library. Therefore, the setting of the judgment threshold determines the accuracy of the recognition judgment, with a high rate of omission if the threshold is too large and a high rate of misjudgment if the threshold is too small.

Through the comparison and analysis of repeated experimental data, the judgment threshold of the feature point of text image subordinate recognition is set reasonably. First of all, from the text data information standard library randomly selected sample set of 355 words, a total of 3267 text variant images for the training of feature point extraction, each word sample set to extract an image as a test sample shown in Figure 3. The average number of feature points of a single language text image in the figure is 82.

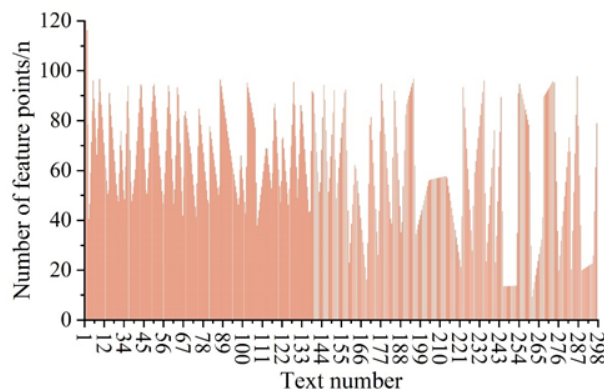


Fig. 3. Characteristic point curve

The language text images in the test sample set are removed in the standard library, the language text images in the sample set are compared with their variants of the language text images for feature point comparison respectively, assuming that the language texts in the sample set are all unaffiliated in the standard library, traversal feature point matching is carried out in the sample set with the standard library after the removal of the data, and the matching result is shown in Figure 4. From the figure, it can be seen that the highest number of feature points for language text matching each other is 11.

According to the above analysis results, the feature points with matching thresholds of 8, 10 and 11 are selected and tested respectively, and the test results are shown in Table 1. According to the analysis of the table, it is concluded that in the process of matching the feature points of the language

text image, the number of feature points that fail to match the original image is not greater than 11, and when the threshold value is greater than 11, it can accurately match the original image, and the leakage rate of the original image is almost 0, and the leakage rate of the variant is 99.5%, which ensures that the algorithm will only accurately match a pair of images of the language text variant and differentiate the variant under the same prefix, so in the highest accuracy and low leakage rate, the threshold value of 11 is the optimal value. Using the set feature point matching threshold 11, no misrecognition occurs under this set threshold.

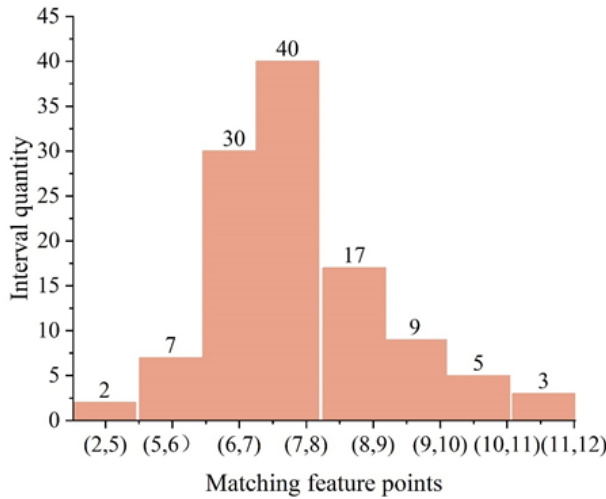


Fig. 4. Language matching feature points

Table 1. Match results of different thresholds

Threshold selection	Original missing rate	Error selection rate	False rate of variation	Leakage rate of variant
8	0%	25%	93.4%	81%
10	0.4%	4%	31%	98.2%
11	0.6%	0%	0%	99.5%

4.2. Analysis of users' interest in language and text modeling

The proportion of interest in the exhibits in the scene when 500 people visit the virtual scene is shown in Figure 5, which shows that users are interested in the text image itself, and the order of interest from high to low is as follows: text image (12)>visual image (9)>experience in the pavilion (8)>scene art (7)>language and culture (6)>image color (5)>fonts charisma (3)>culture and art (1).

4.3. Analysis of user experience results

4.3.1. Analysis of test degree data. This section conducts a quantitative study of user experience through the virtual pavilion of Chinese language and writing constructed in Chapter III. The questionnaire adopts a percentage system with five scoring intervals and sets corresponding scores, and the scoring intervals are shown in Table 2.

First of all, the experimenters entered the “Chinese Language and Writing Virtual Pavilion” on the PC to browse freely, and then rated each question according to the questions on the questionnaire after they finished browsing. The arithmetic average and weighted average of each evaluation index

were calculated as the results of user experience evaluation. The calculation results are shown in Table 3. As can be seen from the table, the arithmetic mean scores of the five dimensions are in the following order: cultural experience (95.34) > emotional experience (86.23) > interactive experience (83.52) > functional experience (83.21) > aesthetic experience (78.56).

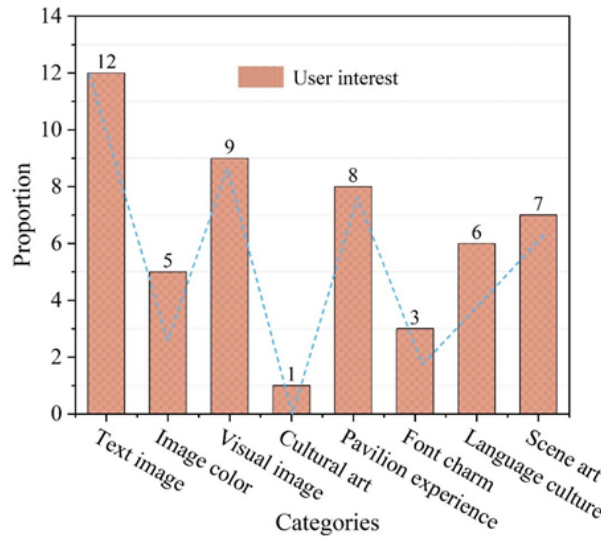


Fig. 5. The user’s ratio of interest to the exhibit

Table 2. Questionnaire scoring interval

Very bad	Bad	General	Good	Great
50-60	60-70	70-80	80-90	90-100

Table 3. User experience evaluation score

	Evaluation dimension	Evaluation index	Index average	Index weight	Index weight average score	Dimensional arithmetic Average score	Dimensional weighting Average score
Sensory level	Aesthetic experience	Color collocation	85	0.059	6.81	78.56	25.11
		Font design	83.53	0.0667	6.51		
		Icon symbol	74.2	0.0632	6.48		
		Page layout	65	0.0562	6.17		
		Relic pattern	86.37	0.0481	6.82		
Interaction level	Immersive experience	Ease of use	79.37	0.0653	6.94	83.21	16.77
		usefulness	91.87	0.0504	6.38		
		autonomy	77.87	0.0291	7.87		
	Interactive experience	System flow	89.03	0.0571	6.83	83.52	14.41
		interinteract	77.53	0.0573	6.24		
	Timely feedback	83	0.0543	6.33			
Reflective level	Emotional experience	novelty	90.2	0.0605	8.06	86.23	23.56
		kindred	83.03	0.0674	7.33		
		Pleasure sense	87.2	0.0536	6.44		
		satisfaction	87.2	0.0482	6.83		
	Cultural experience	Scene sense	93.03	0.066	7.84	95.34	15.76
		identity	96	0.0485	6.43		
		Sense of value	96.53	0.0491	6.48		

Combined with the weight values of each dimension in Table 3, the weighted average score is calculated as follows: aesthetic experience (25.11) > emotional experience (23.56) > functional experience (16.77) > cultural experience (15.76) > interactive experience (14.41), and the radar chart can clearly show the weights of each dimension as shown in Figure 6. In the dimension of aesthetic experience, the arithmetic mean score of user experience is the lowest, while its weighted mean score is the largest, indicating that the user’s experience on the dimension of aesthetic experience is poor, and the aesthetic experience has a greater impact on the user experience and needs to be improved.

The weighted average score of the whole user experience evaluation model of “Chinese Language and Character Virtual Pavilion” is 80.67, which indicates that the overall experience is good, but there is still some room for optimization.

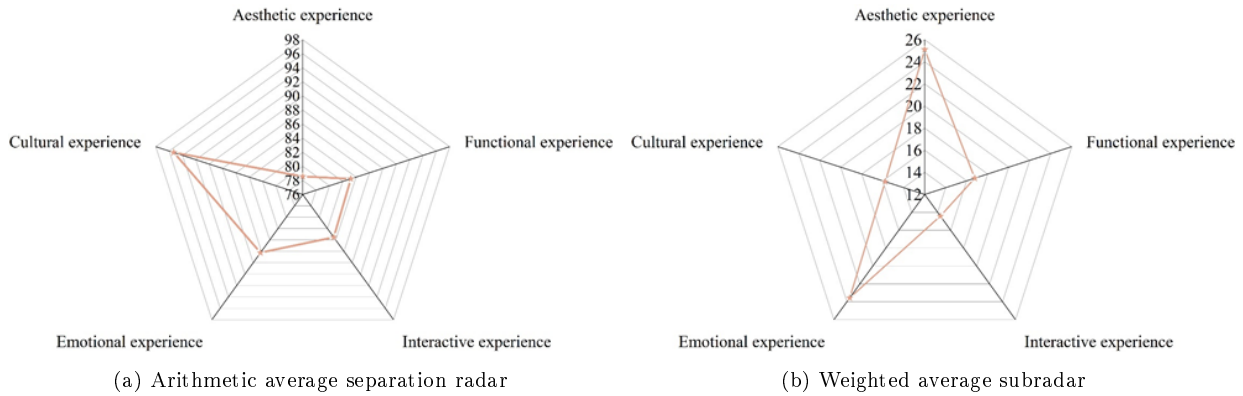


Fig. 6. User experience scoring radar

4.3.2. User experience model analysis.

1) Analysis of indicators. From the model, you can also get the weight of each evaluation index under each dimension. Taking the sensory level as an example, the weights of the five second-level evaluation indicators in this dimension were 0.18, 0.25, 0.14, 0.23 and 0.2, respectively, indicating that the importance of "color matching" and "cultural relics pattern" indicators in this dimension was slightly lower than that of "font design", "page layout" and "icon symbol", indicating that these two indicators are not rigid indicators to improve aesthetic experience. The weights of each indicator dimension are shown in Figure 7, Figure 8, and Figure 9.

2) Importance ranking of evaluation indicators. The model also shows the importance weight of each evaluation indicator in the whole model as shown in Figure 10, with the indicators on the horizontal axis. In the figure, the weight value of “color matching” is 0.0799, which is one of the 18 evaluation indicators with the highest total weight, indicating that it has a great impact on improving the overall user experience.

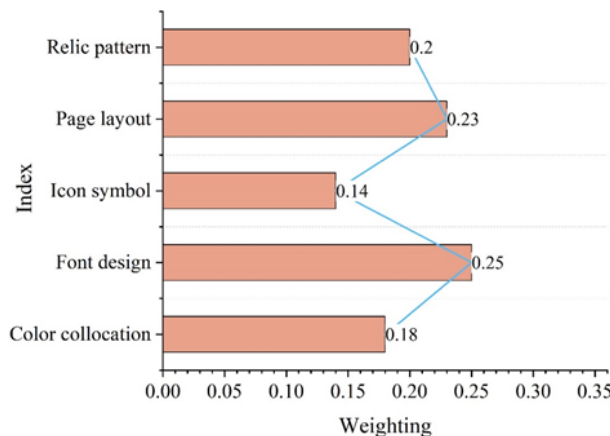


Fig. 7. Sensory level index weight

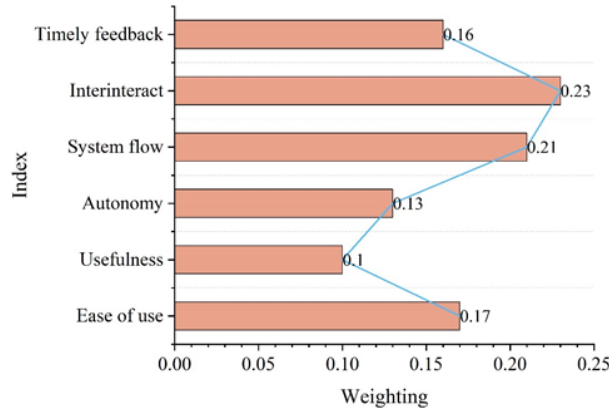


Fig. 8. Interactive level index weight

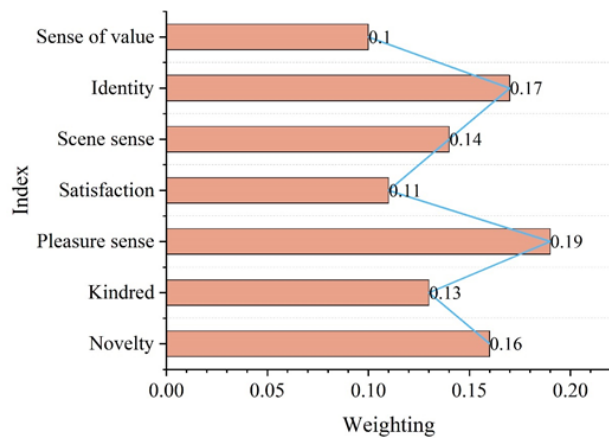


Fig. 9. Reflection level index weight

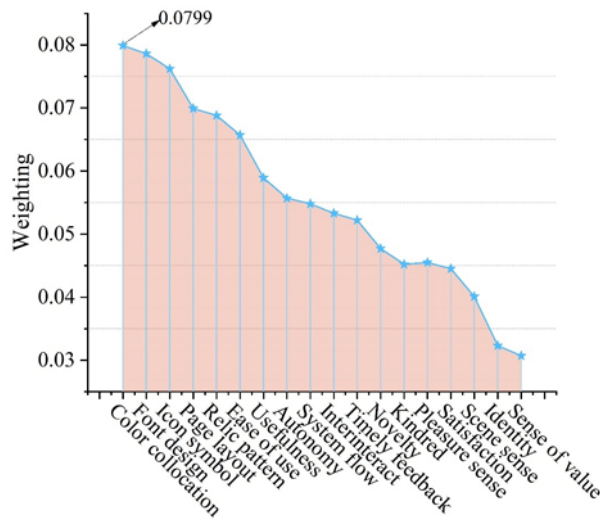


Fig. 10. Each index weight column chart

4.4. Feedback on user experience satisfaction results

The User Interaction Satisfaction Scale was adopted and appropriately modified and partially deleted, and the modified scale was divided into 5 parts with a total of 20 questions, and the score of each question ranged from 1-7, with an average score of 4. By calculating the ratings of each part of the questionnaire, the average of the ratings of each part of the scale of the 8 subject users after the completion of the statistics is shown in Figure 11. The ratings of each part of the questionnaire

of the subject users were higher than 5 and significantly higher than the average rating of 4. The highest rating was given to the system capability, which indicated that the users were satisfied with the cultural dissemination of the virtual pavilion of Chinese language and scripts, the use of the virtual technology, the pattern design and the color, and other aspects of the virtual pavilion.

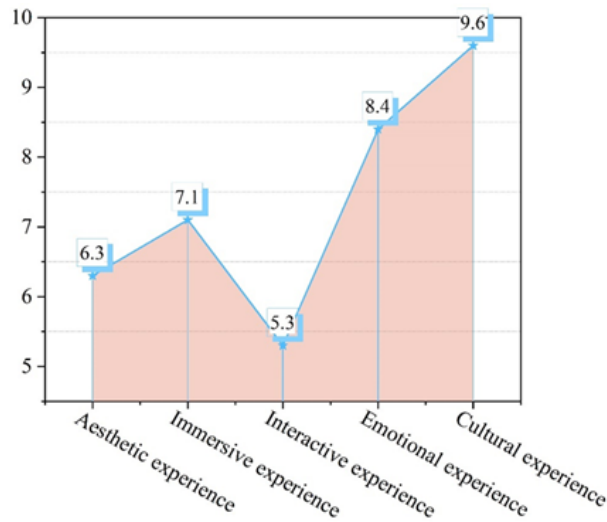


Fig. 11. The trial family qis score

5. Conclusion

In order to inherit and protect the intangible cultural heritage of Chinese language and writing, the research puts forward the concept of creating a virtual exhibition hall of Chinese language and writing, and designs the visiting route, design structure and development process of the virtual exhibition hall of Chinese language and writing based on user analysis. In the mid-term practice process, the image text extraction algorithm and 3D design software are used to complete the design of the virtual exhibition of Chinese language and writing, and the Unity 3D engine is used to improve the interaction module and the integration of the content of the exhibition space.

From the analysis of the effect of image and text feature extraction algorithm, in the process of matching the feature points of language and text images, the algorithm can accurately match the original image, the leakage rate of the original image is almost 0, and the leakage rate of the variants is 99.5%, and there is no case of misrecognition.

According to the feedback of the user experience results, the average of the total experience scores of the eight subjects exceeded 60 points, and the experience was effective. It shows that the users are satisfied with the effect of spreading characters and culture in the virtual exhibition hall of Chinese language and writing, and at the same time, it also plays a role in the inheritance and protection of the intangible cultural heritage of Chinese language and writing.

References

- [1] D. Ben-Amos. Between intangible cultural heritage and folklore. *Folklor/Edebiyat*, 29(114):347–386, 2023.
- [2] O. Bialostocka. Inhabiting a language: linguistic interactions as a living repository for intangible cultural heritage. *International Journal of Intangible Heritage*, 12:18–27, 2017.

- [3] W. Cao. Chinese characters and cultural inheritance: exploring the important role of chinese characters in the inheritance of chinese culture. *Cultura: International Journal of Philosophy of Culture and Axiology*, 21(1), 2024.
- [4] Z. Chen. Chinese language and literature studies in culture the embodiment of value in inheritance. *Education and Educational Research*, 4(1):114–116, 2023. <https://doi.org/10.54097/jeer.v4i1.10225>.
- [5] M. Cornu. Defining the perimeter of the intangible cultural heritage: focus on language. In *Intangible Cultural Heritage Under National and International Law*, pages 54–67. Edward Elgar Publishing, 2020. <https://doi.org/10.4337/9781839100031.00012>.
- [6] J. Deng. A brief analysis of the path of intangible cultural heritage inheritance and innovative development under digital technology. *Journal of Innovation and Development*, 3(3):29–32, 2023.
- [7] J. Eichler. Intangible cultural heritage, inequalities and participation: who decides on heritage? *The International Journal of Human Rights*, 25(5):793–814, 2021. <https://doi.org/10.1080/13642987.2020.1822821>.
- [8] H. Fan et al. Exploring the value of chinese language and literature research in cultural inheritance. *Frontiers in Educational Research*, 4(16):60–62, 2021. <https://doi.org/10.25236/FER.2021.041612>.
- [9] K. W. Griffin, G. J. Botvin, C. Williams, and S. M. Sousa. Using virtual reality technology to prevent substance misuse and violence among university students: a pilot and feasibility study. *Health Informatics Journal*, 30(4):14604582241290712, 2024. <https://doi.org/10.1177/14604582241290712>.
- [10] Q. Guo. Research on the ways to protect and inherit intangible cultural heritage in the information age. In *Journal of Physics: Conference Series*, volume 1575 of number 1, page 012169. IOP Publishing, 2020. <https://doi.org/10.1088/1742-6596/1575/1/012169>.
- [11] G. M. Mu and K. Dooley. Coming into an inheritance: family support and chinese heritage language learning. *International Journal of Bilingual Education and Bilingualism*, 18(4):501–515, 2015. <https://doi.org/10.1080/13670050.2014.928258>.
- [12] M. S. Nasirabadi. The role of intangible cultural heritage in cultural globalization. *Guidelines for Authors*, 2588:32, 2017.
- [13] G. Obourdin, S. De Maeyer, and P. Van den Bossche. Unlocking the power of immersive learning: the fairi instructional design proposition for adaptive immersive virtual reality. *Computers & Education: X Reality*, 5:100084, 2024. <https://doi.org/10.1016/j.cexr.2024.100084>.
- [14] S. Oeter. Intangible cultural heritage in need of protection: the case of regional and minority languages. *Cultural Heritage and International Law: Objects, Means and Ends of International Protection*:59–83, 2018. https://doi.org/10.1007/978-3-319-78789-3_4.
- [15] T. Petronela. The importance of the intangible cultural heritage in the economy. *Procedia Economics and Finance*, 39:731–736, 2016. [https://doi.org/10.1016/S2212-5671\(16\)30271-4](https://doi.org/10.1016/S2212-5671(16)30271-4).
- [16] L. Sagart, G. Jacques, Y. Lai, R. J. Ryder, V. Thouzeau, S. J. Greenhill, and J.-M. List. Dated language phylogenies shed light on the ancestry of sino-tibetan. *Proceedings of the National Academy of Sciences*, 116(21):10317–10322, 2019. <https://doi.org/10.1073/pnas.1817972116>.
- [17] R. Shen. Interactions between min and other sinitic languages: genetic inheritance and areal patterns. In *The Palgrave Handbook of Chinese Language Studies*, pages 1–16. Springer, 2021. https://doi.org/10.1007/978-981-13-6844-8_33-2.