International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-13
© Cerebration Science Publishing

https://doi.org/10.70917/ijcisim-2026-0116
Article

Graphics Algorithm Based Virtual Reality Immersive Interactive Interface Design for International Chinese Language Teaching and Learning

Xiaoyan Wang *

College of Literature, Dalian University, Dalian 116622, Liaoning, China; wtqwxy7117@126.com

Abstract: This paper addresses the issue of cultural experience deficits in international online Chinese language instruction by proposing a layered optimization method for VR immersive interactive interfaces based on computer graphics algorithms and human visual characteristics. By analyzing human color sensitivity, spatial perception, and multi-channel characteristics, a layered optimization model centered on visual perception intensity is constructed, and a genetic algorithm is employed to solve the layout of interface elements. The experimental section focuses on Chinese language teaching scenarios, determining that the optimal combination is a feature space capacity of 128MB (accuracy rate of 93.18%) and a dictionary space of 512MB (accuracy rate of 78.89%). When the weight parameter is set to 0.35, the recognition accuracy rate reaches 97.23%. Quantitative analysis of 15 teaching elements (e.g., the 3D model of the Hall of Supreme Harmony, cultural explanation subtitles) shows that the core teaching elements (A1/A8) have a visual communication index (VCI) > 3700, and high-intensity visual perception elements ($\mathrm{VPI} \geq 4.5$) are concentrated in the core teaching area (accounting for 48%). Among the 1,520 element layouts, 32% of high-intensity areas carry 87% of high-VCI elements. The proposed method significantly outperforms existing methods in Chinese cultural image processing, with an average gradient of 6.65–8.05, 1.9–2.8 times higher than contrast methods, a PSNR improvement of 29.7–40.2%, a peak of 21.23, and an MSE of only 2.68%–4.80%, which is 1/7 of contrast methods.

Keywords: computer graphics algorithms; international Chinese language education; VR immersive interactive interface; hierarchical optimization model

1. Introduction

In recent years, the Chinese language has been undergoing rapid globalization. International Chinese language education currently faces significant challenges, including uneven distribution of educational institutions, teachers, and teaching resources, as well as a large and dispersed population of learners. To meet the global demand for Chinese language learning, relying solely on traditional classroom instruction is no longer feasible; a significant portion of this demand must be addressed through internet-based online teaching. However, online teaching faces challenges such as a lack of classroom atmosphere and language interaction environments, insufficient interactivity, and weak classroom management, which severely limit its development [1-2]. Utilizing virtual reality (VR) technology for international Chinese language education is a promising solution to these issues.

Virtual reality (VR) technology integrates multiple scientific and technological disciplines across various application scenarios, creating a virtual information environment in a multi-dimensional information space that provides users with an immersive, lifelike experience [3]. VR leverages human psychological and sensory characteristics to generate 3D virtual images via computer technology, resulting in highly realistic and immersive environments [4-5]. In international Chinese language education, the use of VR wearable devices such as data suits, data gloves, and head-mounted displays places users in highly realistic virtual environments, enabling teachers and students to engage in interactions and communications with the virtual environment that are nearly indistinguishable from reality [6-8]. Virtual reality content is diverse, including panoramic images, large-screen movies, panoramic tours, or videos. International Chinese learners can use virtual reality technology to explore realistic historical sites, natural landscapes, and urban attractions, giving learners a sense of being there. Additionally, the entire process can be replayed, not only providing learners with an excellent Chinese language learning environment but also expanding their Chinese language horizons and knowledge systems, with vast application prospects [9-11]. It is evident that the application of VR technology in international Chinese language education offers more opportunities for the deep integration of the internet and international Chinese language education.

How to effectively conduct research on international Chinese language teaching based on VR technology and promote the innovative development and transformation of international Chinese language teaching in the new era has become a new direction for academic research. Ma, C. indicates that under the current trend of economic globalization, international Chinese language education needs to shift from a theory-centric teaching model to incorporating VR technology into language classrooms. By constructing realistic environments, VR technology can provide effective support for oral language instruction, thereby driving the development of international Chinese language education [12]. Liu, W. revealed the potential benefits and effectiveness of applying virtual reality technology in teaching Chinese as a foreign language. By analyzing the impact of real learning scenarios and interactive learning methods on students' language learning outcomes, it can provide valuable insights for educators [13]. Barrett, A. et al. used the Technology Acceptance Model to investigate learners' attitudes toward VR Chinese learning tools, finding that learners have a high acceptance rate for VR technology, but optimizing VR learning environment design and enhancing learners' virtual reading and writing abilities are key to integrating VR technology into international Chinese language education [14]. Chen, C. and Yuan, Y. designed a virtual reality-based Chinese vocabulary learning theoretical framework, pointing out that by leveraging virtual reality technology to create an immersive environment with high interactive engagement and multimodal scaffolding learning methods, it can effectively promote students' vocabulary learning and memory, thereby enhancing learners' Chinese proficiency [15]. Xiaoning, X. et al. outlined the primary challenges facing international Chinese language education and explored effective pathways for integrating VR technology into Chinese language education, providing guidance for educators and practitioners to conduct sustainable teaching activities [16]. Xie, Y. et al. conducted an empirical study and found that interactive VR tools effectively stimulate students' interest in learning the Chinese language and culture by providing authentic learning scenarios, demonstrating significant effectiveness [17]. Based on this, we must recognize the current state of technological advancement in international Chinese language education in the new era, continuously optimize the immersive and interactive learning experiences enabled by virtual reality technology, and inject new vitality into the innovative development of international Chinese language education, opening up new prospects.

This study proposes and elaborates on a design methodology for immersive VR interaction interfaces based on computer graphics algorithms and human visual characteristics. By leveraging a deep understanding of the human visual system, particularly its spatial perception and intensity sensitivity, a hierarchical optimization model is constructed to guide the intelligent layout of interface elements, thereby enhancing learners' information acquisition efficiency and cultural immersion in VR environments. Specifically, the study first focuses on the specific needs of online Chinese language cultural education, analyzing the advantages of VR panoramic display technology in compensating for the lack of on-site cultural experiences. It then proposes a hierarchical optimization model for graphical human-computer interaction interfaces based on visual perception intensity. The core idea of this model is to divide the VR interface into regions with different levels of visual perception intensity and, based on the importance levels of interface elements, optimize their spatial layout through mathematical modeling. Furthermore, a genetic algorithm is proposed for solving the problem. The model's chromosome encoding rules (encoding element arrangement sequences as genes) and decoding process (gradually placing elements into appropriate perception intensity regions) are detailed, along with the steps for iteratively calculating optimal or suboptimal layout solutions.

2. VR Chinese Teaching Interface Design Method Based on Visual Perception Layered Optimization

2.1. Teaching Design and Practice for Online International Chinese Classes Based on VR Panoramic Display

Unlike general language courses, cultural courses should place greater emphasis on explaining and disseminating the elements and spirit of Chinese culture. Due to the limitations of online teaching, many international students have lost the opportunity to visit cultural sites in person for close-up observation and learning, which has impacted the effectiveness of cultural experiences. At this point, teachers need to adopt more diverse teaching methods in online instruction and utilize other more vivid and concrete teaching tools to assist students in their learning, thereby enhancing the cultural experience during class. Based on an analysis of the characteristics and advantages of VR panoramic displays, which align well with the needs of online cultural education, the author has applied VR panoramic displays in international Chinese language online cultural education and presents specific teaching designs and practices in this chapter.

In selecting teaching themes, analysis of survey questionnaires revealed that international students have a stronger interest in learning about Chinese daily culture, social life, and literary arts. Therefore, these three areas became the focus of the teaching content selected by the author. Chinese architectural art has a long history, not only demonstrating unique achievements in group layout, space, structure, and decorative art, but also embodying the ideological spirit of the Chinese nation, which plays a crucial role in helping international students understand Chinese culture. Taking ancient Chinese architectural culture as the entry point for the teaching content in this instructional design, the most representative ancient Chinese architectural site—the Forbidden City—was selected to design this online cultural teaching practice course.

This instructional design focuses on cultural knowledge instruction, integrating VR panoramic displays as a teaching method into the instructional process, enabling international students to better acquire cultural knowledge and develop related language skills during the online cultural experience.

2.2. Human Eye Visual Characteristics and Image Processing

This paper will explore the visual characteristics of the human eye and their significant impact on image processing. It will examine the imaging principles and visual perception capabilities of the human eye, and study the differences in visual perception between the two main forms of image representation in computers: bitmaps and vector graphics. Additionally, it will analyze the basic characteristics of human vision to understand how the human eye perceives and processes images.

2.2.1. Principles of Human Eye Imaging

The principle of human eye imaging involves the cornea, lens, and retina. The cornea and lens focus external light onto the retina inside the eyeball, forming an inverted, reduced real image. The photoreceptor cells on the retina convert light signals into electrical signals, which are then transmitted through the nerves and ultimately interpreted by the brain as visual images.

2.2.2. How Images Are Represented on Computers

Images typically exist in the form of bitmaps and vector graphics. Bitmaps are composed of pixels, each containing specific color information, representing complex images that may become distorted when enlarged. Vector graphics represent graphics and icons that can be enlarged infinitely without distortion.

2.2.3. Basic Characteristics of Human Vision

(1) Color sensitivity. The human eye has different perception capabilities for light of different wavelengths, resulting in varying degrees of color sensitivity, which affects color processing and correction in images. The human visual system's response to stimulus signals does not depend on the absolute brightness of the signal but rather on the degree of stimulation relative to the background brightness.

(2) Spatial perception. When viewing an image, the human eye can only resolve details in a specific region of the image and cannot simultaneously resolve other regions of the image.

(3) Multi-channel characteristics. The human visual system includes multiple channels that can decompose image signals by direction and frequency, enabling us to perceive details such as edges and textures in an image.

2.3. Interaction Interface Layered Optimization Model Based on Visual Perception Intensity

Based on the above in-depth understanding of the human visual system, particularly its perception of spatial distribution characteristics, the core innovation of this chapter lies in proposing a layered optimization design model for VR Chinese teaching interfaces.

2.3.1. Building a Layered Optimization Model

A layered optimization model for graphical human-machine interface based on visual perception intensity. The visual perception intensity of regions such as flat areas, non-flat areas, curved surfaces, and single-frame distributions in graphical human-machine interfaces is set as the optimization target. The model is established based on the visual perception intensity of each region in the graphical human-machine interface and the importance of visual perception elements at the edges. To achieve hierarchical optimization of different regions in the human-machine interface, the key parameters in each region are first defined.

(1) $D=\{d_1,d_2,\cdots,d_a\}$, where the current rank value of the i th visual perception element and the set of importance ranks of all visual perception elements in the graphical human-computer interaction interface are d_i and D, respectively, with i taking values of $1,2,\cdots,n$

(2) $X=\{x_1,x_2,\cdots,x_\alpha\}$ where x_j and X represent the perceived intensity level of the current j th visual perception region and the perceived intensity levels of the various regions in the graphical human-computer interaction interface, respectively, with j taking values from $1,2,\cdots,m$

(3) When arranging visual perception element i, the area it occupies in the j th intensity level region is denoted by q_{ij};

(4) $R=\{r_1,r_2,\cdots,r_a\}$ represents the total area occupied by the current i th visual perception region and the visual perception intensity indices of each visual perception element after being arranged on the graphical human-machine interface, denoted as r_i and R, respectively, where i takes values from $1,2,\cdots,n$. Then,

$$r_i=\sum_{j=1}^{\infty} d_i x_j q_{ij} \tag{1}$$

In this context, d_i and x_j represent the importance level and the perceived intensity level of the area occupied by the placement position of any visual perception element when it is arranged on a graphical human-machine interface, respectively. The area occupied by the placement position is denoted by q_{ij}.

If a visual perception element is to be placed in the core area of the graphical human-computer interaction interface, the value of r_i should be as large as possible. Based on this method, a hierarchical optimization model for graphical human-computer interaction interfaces based on visual perception intensity is established, as shown below:

$$Z=\max\left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} d_i x_j q_{ij}\right) \tag{2}$$

$\sum_{i=1}^{\infty} q_{ij}=q_j \cdot \sum_{j=1}^{\infty} q_{ij}=s_i, \sum_{j=1}^{\infty} q_j=\sum_{i=1}^{\infty} s_i$ are the constraints for this model.

The definition of the visual communication index is obtained from equation (2). If we want to indicate that high-importance visual perception elements are arranged in areas with high visual perception intensity in graphical human-computer interaction interfaces, then the value of z should be larger.

2.3.2. Model Solving Based on Genetic Algorithms

Use a genetic algorithm to solve the hierarchical optimization model of the graphical human-computer interaction interface based on visual perception intensity established in the previous section.

(1) Model encoding

When solving chromosomes using a genetic algorithm, use encoding rules based on visual perception elements to ensure that visual perception elements are appropriately arranged in different areas of the graphical human-computer interaction interface and that the arranged areas are connected, thereby achieving hierarchical optimization of the interface. If there are 8 visual perception elements, their serial numbers are treated as gene fragments for chromosome solution. If integers from 1 to 8 are used to encode the serial numbers of the perception elements, a feasible chromosome encoding is represented as follows:

$$\begin{aligned} p_1 &:(2 \quad 3 \quad 5 \quad 6 \quad 7 \quad 4 \quad 8 \quad 1) \\ p_2 &:(5 \quad 6 \quad 2 \quad 3 \quad 7 \quad 4 \quad 8 \quad 1) \end{aligned} \tag{3}$$

(2) Chromosome coding solution

Based on the above coding rules, the following steps are used to solve the graphical human-computer interaction interface layout:

1) The visual perception elements S_i and d_i information for the current implementation layout can be obtained based on the unassigned gene codes extracted from the gene fragments.

2) The information of each visual perception area for the unassigned visual perception elements can be obtained based on the visual perception area hierarchy from high to low, and the visual perception elements are arranged within the visual perception area according to the area matching rule.

3) Determine whether the current visual perception area can completely arrange the visual perception element. If so, jump to step 5); if not, jump to step 4).

4) Based on the area matching rule, arrange the visual perception elements in the next level visual perception area and the previously unarranged visual perception area together, and jump to step 3).

5) Calculate the r_i value after the visual perception element is completely arranged.

6) Determine whether the decoding of all chromosomes in the current encoding has been completely finished. If so, jump to step 7); if not, return to step 1).

7) The solution calculation for a single chromosome can be achieved by obtaining the current chromosome encoding Z value based on the calculation results and the distribution information of each visual perception element in each perception region.

3. Experimental Optimization and Analysis of VR Interactive Interfaces for International Chinese Language Teaching

Based on the hierarchical optimization model and genetic algorithm solution framework proposed in Chapter 2, which are based on visual perception intensity, this chapter will conduct systematic experiments and analyses to verify the practical effectiveness and performance advantages of this model in the design of VR interactive interfaces for international Chinese language teaching.

3.1. Experiment on Optimizing Interface Parameters for Chinese Language Teaching VR Systems

3.1.1. Establishing the Dataset and Setting Experimental Parameters

To rigorously validate the reliability and effectiveness of the visually communicative design method proposed in this paper, which is based on computer graphics algorithms, within an international Chinese language teaching VR environment, this study constructed a specialized graphic image dataset integrated with international Chinese cultural teaching content. The method's performance was comprehensively tested through experimental comparisons.

The core of the dataset construction is the actual application scenarios of the international Chinese language teaching VR system. We systematically collected five categories of image samples covering core cultural elements from the Forbidden City VR teaching module included in the system: architectural panoramas and environments, architectural component details, cultural symbols and elements, interactive interface elements, and dynamic effect frames. A total of 8,966 original image samples across these five categories were initially collected. To simulate the embedding effect of key information in the interface (such as explanatory subtitles, interactive prompts, and navigation markers), we specially designed and generated 100 synthetic images containing typical Chinese teaching interface elements (such as architectural diagrams with Chinese subtitles, detailed diagrams with overlaid cultural annotation markers, and scene diagrams with interactive controls). The 100 synthetic images were randomly inserted into the 8,966 base images, resulting in a specialized dataset containing 9,066 image samples. Using a random sampling method, 2,000 images were selected from the 9,066 images as training samples, with the remaining 7,066 images serving as test samples.

When describing the original features of the images, a scale division method based on image blocks was employed. The images were divided into image blocks with a side length of 2R pixels (R set to 16), and all feature interval pixel data were uniformly set to 2, resulting in a scale area of (2R x 2R) for each image block. Assuming that the feature distribution conforms to a Gaussian distribution, a fixed-scale spatial pyramid scheme was used for feature organization. All experimental programs were run on specified personal computers with consistent configurations to ensure the accuracy of the training process and the comparability of the results.

3.1.2. Evaluation of Experimental Parameters

During the evaluation of experimental parameters, it is necessary to focus on testing and analyzing the relevant parameters of the method described in this paper, such as the dictionary space capacity parameters of the feature space, in order to obtain the effects of parameter changes on other indicators, as shown in Figures 1 and 2. Figures 1 and 2 show the accuracy rates of the algorithm under different feature space capacities and dictionary space capacities, respectively.

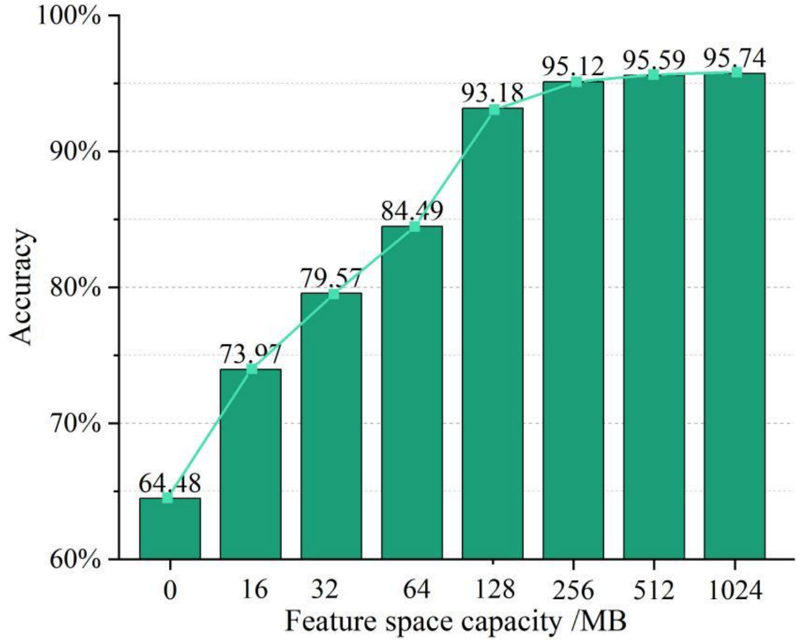

Figure 1. The influence of different feature space capacities on accuracy.

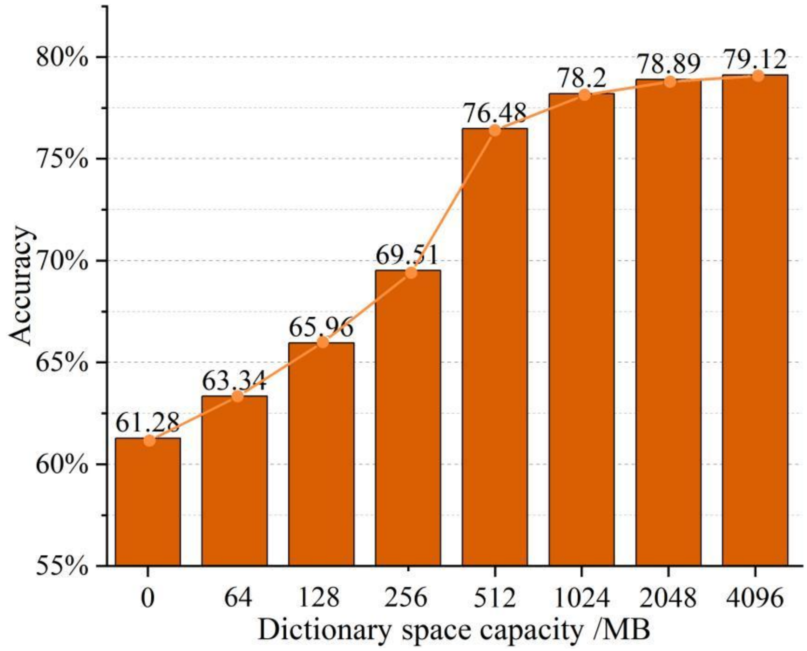

Figure 2. The influence of different dictionary space capacities on accuracy.

As can be seen from the figure above, the algorithm's accuracy rate shows an upward trend as the number of feature space distributions increases. When the feature space distribution reaches 128MB, the algorithm accuracy is 93.18%. However, when the space capacity is doubled to 256MB, the accuracy increases by less than 2% to 95.12%. When the feature space capacity is 1024, the algorithm accuracy is 95.74%. Although the number of parameters increases significantly, the accuracy does not improve significantly. Therefore, the optimal feature space capacity is considered to be 128MB. Additionally, the algorithm's accuracy rate continues to increase as the dictionary space capacity increases. When the dictionary space capacity reaches 512MB, the growth rate of the algorithm's accuracy gradually slows down, with an accuracy of 78.89%. When the dictionary space capacity reaches 4096, the accuracy is only 79.12%. It is concluded that the optimal dictionary space capacity is 512MB.

It can be seen that whether it is the feature space capacity or the number of dictionary space distributions, when there is a certain range of changes, the image recognition accuracy will exhibit synchronous changes. When both parameters are set to 0, the image recognition accuracy reaches 64.48% and 61.28%, respectively, indicating that the image recognition accuracy reaches its minimum value. To ensure that the feature space maintains optimal performance, it is necessary to retain as much information as possible to significantly increase the saturation of the feature space. At this point, if the dictionary space capacity is uniformly set to 512 MB, the corresponding feature space distribution should be set to 128 MB. As the image recognition accuracy continues to increase and reaches its maximum value, to effectively balance image recognition accuracy and recognition efficiency, it is necessary to analyze and determine the optimal communication of weight parameters, thereby obtaining the spatial clustering image shown in Figure 3.

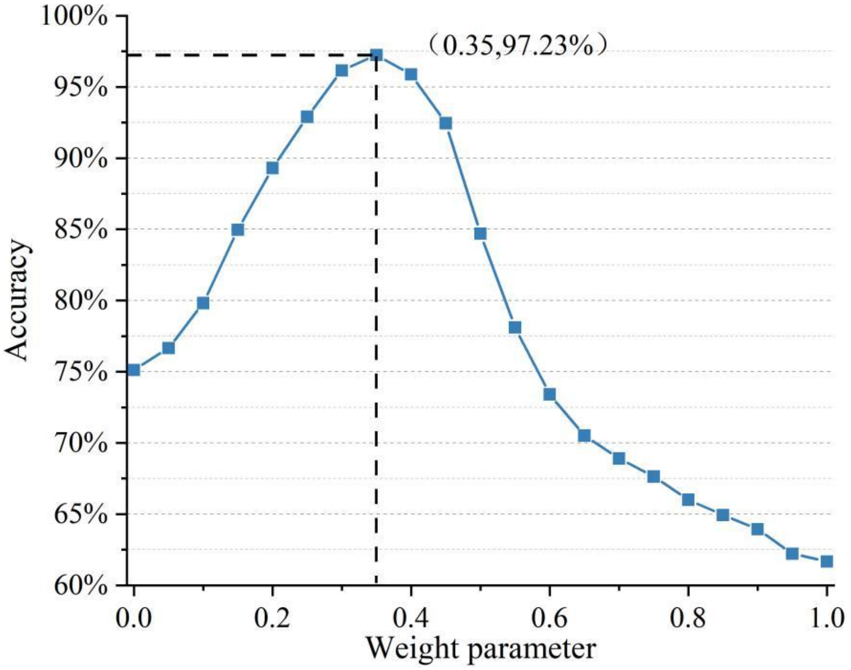

Figure 3. The result of the variation of accuracy rate with the weight parameters.

As can be seen from Figure 3, when the dictionary space capacity and the number of feature space distribution are 512MB and 128MB, respectively, the accuracy of the algorithm will show the shape of "several" with the increase of the weight parameters. When the weight parameter is 0.35, the image recognition accuracy rises to 97.23%.

3.2. Analysis of VR Teaching Interface Element Layout Based on a Hierarchical Optimization Model

Through the experiment in Section 3.1, we determined the optimal configuration of feature space capacity (128 MB), dictionary space capacity (512 MB), and weight parameters (0.35) in the context of a Chinese language teaching VR system. Based on this optimized parameter system, this section will focus on applying a layered optimization model to analyze and design the layout of visual perception elements in the actual teaching interface.

3.2.1. Visual Communication Index Analysis of Visual Perception Elements

To validate the effectiveness of the method proposed in this paper for optimizing the design of virtual reality immersive interactive interfaces in international Chinese language education, the method was applied to the VR immersive interactive interface of an actual international Chinese language education system.

Fifteen visual perception elements were selected: A1: Forbidden City panoramic dynamic display; A2: Hall of Supreme Harmony 3D model; A3: Chinese cultural explanation subtitles; A4: Interactive disassembly animation of bracket structure; A5: Glazed tile color gradient display; A6: Student real-time interactive question-and-answer box; A7: Navigation map (palace location markers); A8: Idiom and anecdote pop-up window; A9: Traditional pattern zoom observation area; A10 : Voice control button; A11: Background music switch icon; A12: Learning progress indicator bar; A13: Share function button; A14: Help prompt icon; and A15: User avatar settings. The statistical results of the visual communication index are shown in Figure 4.

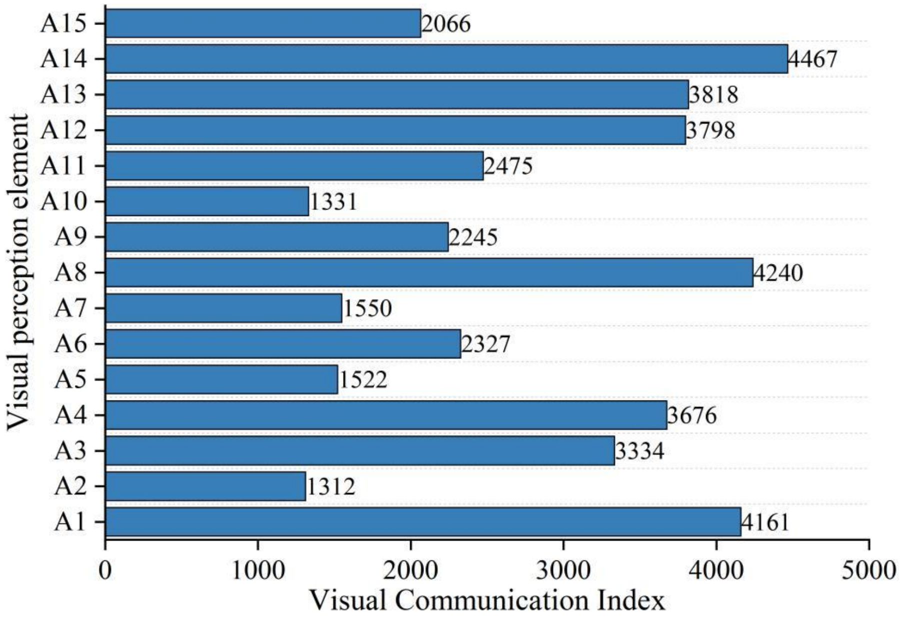

Figure 4. The visual communication index of visual perception elements.

Figure 4 shows the visual communication index (VCI) for 15 visual perception elements, with values ranging from 1312 (A10 voice control button) to 4467 (A14 help prompt icon). Among these, core instructional functional elements (such as A8 Idiom and Anecdote Pop-up Window, A1 Forbidden City Panorama Dynamic Display, and A13 Share Function Button) all have VCIs exceeding 3,700, indicating that the model prioritizes the visual communication efficiency of critical instructional content. Interactive auxiliary elements (such as A14 Help Tip Icon and A12 Learning Progress Indicator Bar) have the highest VCI values (4467 and 3798), highlighting their importance in guiding user operations within the VR environment. Secondary display elements (e.g., A5 glazed tile color gradient display, A7 navigation map) have lower VCI values (1522–1550), indicating that the model reasonably allocates interface resource priorities.

As shown in Figure 4, the proposed method can effectively calculate the visual communication index of different visual perception elements. The visual communication index reflects the importance of visual perception elements. Through reasonable color coordination, font selection, layout design, etc., the interface becomes more aesthetically pleasing and comfortable, thereby attracting users' attention and increasing their willingness to use it.

3.2.2. Visual Perception Intensity of Visual Perception Elements

Further analysis of different visual perception elements in the VR immersive interactive interface of the international Chinese language teaching system was conducted to statistically determine the importance of different visual perception elements. The statistical results of the visual perception intensity of some visual perception elements in the VR immersive interactive interface are shown in Figure 5. The visual perception intensity was divided into five levels from 1 to 5.

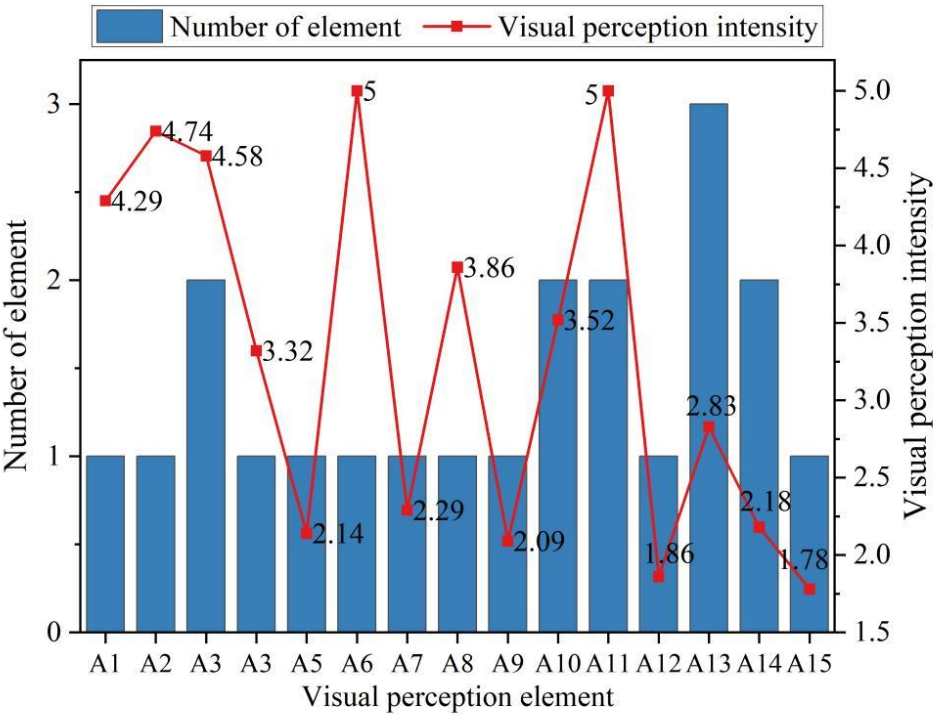

Figure 5. The visual perception intensity of visual perception elements.

Analysis of the statistical results in Figure 5 shows that high VPI elements (VPI ≥ 4.5) include A6 Student Real-Time Interactive Q&A Box (5.00), A11 Learning Progress Indicator Bar (5.00), A2 Taihe Hall 3D Model (4.74), and A3 Chinese Cultural Explanation Subtitles (4.58). These elements are concentrated in core interaction and teaching content, and should occupy the central area of the user's field of view. Medium-high VPI elements (VPI 3.5–4.0): A1 Forbidden City Panorama Dynamic Display (4.29), A8 Idiom and Folklore Pop-up Window (3.86), and A4 Dou Gong Structure Dissection Animation (3.72) are primarily used for showcasing cultural details. Low VPI elements (VPI ≤ 2.5): A15 User Avatar Settings (1.78), A12 Background Music Switch Icon (1.86), etc., which are auxiliary functions suitable for placement in the peripheral areas of the interface.

The proposed method effectively measures the visual perception intensity of different visual perception elements. When optimizing the design of an interactive interface, visual perception elements with higher visual perception intensity should be placed in areas of the interface with higher visual communication indices.

3.2.3. Layout of Visual Perception Elements in Interactive Interfaces

The virtual reality immersive interactive interface for international Chinese language teaching is divided into six major modules: the central teaching area, the panoramic display area, the interactive operation area, the cultural annotation layer, the environmental immersion layer, and the system function area. The particle swarm algorithm is used to solve the optimization model of the teaching VR immersive interactive interface, and the number of units of visual perception elements of different intensities in the six different modules is counted. The layout results of the visual perception elements of the human-computer interaction interface are shown in Table 1.

Table 1. Layout of visual perception elements in the interactive interface.

Module	Strength grade					Number of elements
	1	2	3	4	5	
Central Teaching Area	42	35	78	105	38	298

Panoramic display area	18	22	45	92	12	189
Interactive operation area	76	64	85	48	53	326
Cultural annotation layer	15	28	63	57	29	192
Environmental immersion layer	102	89	46	21	8	266
System Function area	118	72	35	18	6	249
Total	371	310	352	341	146	1520

It can be seen that the core teaching elements of this international Chinese language teaching virtual reality immersive interactive interface are highly concentrated. The central teaching area has a 4-5 level intensity unit ratio of 48% (143/298), with high VCI elements concentrated in this area to ensure that core teaching content occupies the user's field of vision. The cultural annotation layer has a 3-5 level intensity unit ratio of 78% (149/192), with high interactivity elements prioritized to support the immersive learning process. The interactive operation zone adopts a “dumbbell-shaped distribution,” with the highest number of units at intensity levels 1 and 5 (76+53=129), while the medium-intensity zone accommodates transitional functions, aligning with multi-channel characteristics. The system functional zone has 76% of units at intensity levels 1-2 (190/249), constraining low-priority elements to the edges to reduce visual interference. The environmental immersion layer has a 72% (191/266) proportion of units at intensity levels 1-2, placing non-interactive elements in low-perception zones to avoid GPU resource waste. The panoramic display zone has a high density of units at intensity level 4 (92/189) while limiting the number of units in the core zone (only 12 at intensity level 5) to prevent overload.

Global high-intensity units (levels 4-5) account for 32% (487/1520) but carry 87% of high VCI elements, significantly enhancing the visual communication index. Even at a scale of over 800 units, this layout maintains sparse high-intensity regions (level 5 intensity units account for only 9.6%) and modular aggregation in low-intensity regions, aligning with the human visual system's perception limits for information density and spatial distribution. This validates the scalability of the hierarchical optimization model in large-scale VR systems.

3.3. Comparison and Verification of Image Enhancement Effects in Chinese Language Teaching VR Scenarios

The previous section provided a detailed analysis of the Visual Communication Index (VCI) and Perceptual Intensity (VPI) of core teaching elements and interactive elements under an optimized layout, and presented the specific results of the layered layout. To further validate the effectiveness and reliability of the image enhancement method for the virtual reality immersive interactive interface in international Chinese language teaching proposed in this study, simulations were conducted. The image enhancement method based on MR-VAE, the image enhancement method based on the U-Net generative adversarial network, and the research method based on computer graphics algorithms proposed in this paper were compared to objectively and scientifically evaluate the image enhancement effects.

3.3.1. Evaluation Indicators

Objective evaluation mainly uses mathematical theorems to calculate corresponding values to objectively reflect some characteristics of the image. The main indicators used are as follows:

(1) Average gradient: This value mainly shows the obvious gray difference parameters at the image boundary, which is a parameter that characterizes the clarity of the image. The larger the value, the more complex the image layers and the higher the clarity. The calculation formula is as follows:

$$V=\frac{F\left(x,y\right)}{\left(M-1\right)\left(N-1\right)} \tag{4}$$

In the formula, F(x,y) represents the gray value of the image at point (x,y).

(2) Peak signal-to-noise ratio, the calculation formula for this value is as follows:

$$z=10\times\alpha\frac{L_{MAX}^{2}}{MSE} \tag{5}$$

In the formula, L_{MAX}^{2} represents the maximum gray value of pixels in the image.

(3) MSE: Compare the mean square error after processing the three methods. The calculation formula

is as follows:

$$Q=\frac{\sum_{i=1}^{r}\sum_{j=1}^{e}\left(r_i-e_j\right)^2}{M\times N} \tag{6}$$

In the formula, r_i,e_j represent image pixel values.

3.3.2. Image Performance Comparison Analysis

Ten images related to international Chinese language teaching were processed, and the results of comparing the average gradient, peak signal-to-noise ratio, and mean square error of the images enhanced by three image enhancement methods are shown in Figures 6 and 7, respectively.

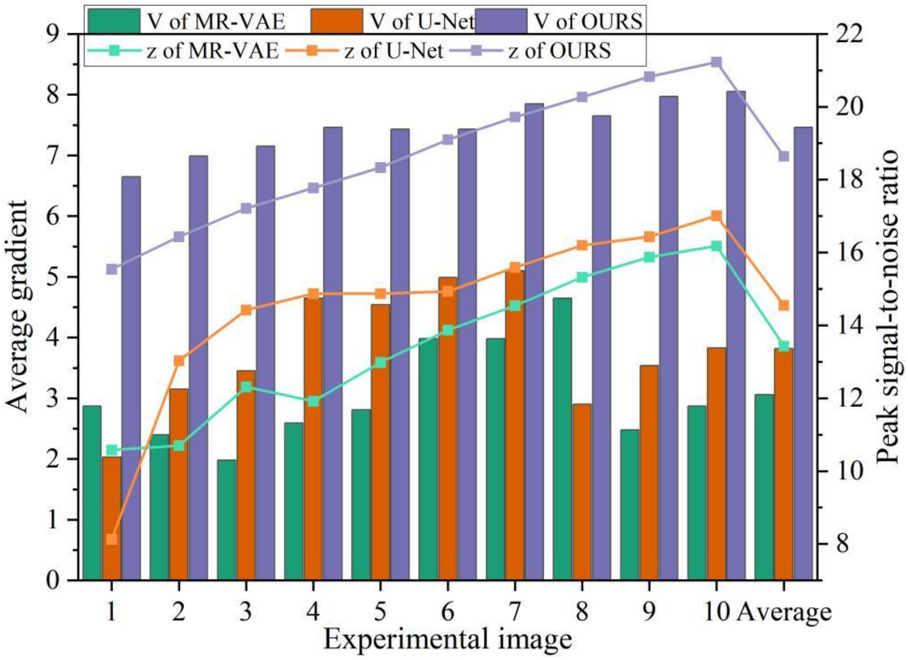

Figure 6. Results of the average gradient and peak signal-to-noise ratio of the image.

First, in terms of average gradient reflecting image clarity, the method proposed in this paper significantly outperforms the comparison methods. Among all 10 experimental images, the average gradient value of the proposed method (6.65–8.05) far exceeds that of the comparison methods, with an average improvement of 2.8 times compared to the MR-VAE method (1.98–4.65) and 1.9 times compared to the U-Net method (2.03–5.10). The gradient values of the proposed method consistently remain above 6.65, with a fluctuation range of only 1.4 (6.65 to 8.05), while the fluctuation ranges for MR-VAE and U-Net are 2.67 and 3.07, respectively, indicating that the proposed method has stronger adaptability to complex scenes.

In terms of peak signal-to-noise ratio (PSNR), which reflects image quality, the proposed method achieves an average PSNR of 15.54 to 21.23, representing an average improvement of 40.2% compared to MR-VAE (10.58 to 16.18) and an average improvement of 29.7% compared to U-Net (8.13 to 17.01). As image complexity increases, the PSNR of the proposed method steadily increases from 15.54 to 21.23, representing a 36.7% increase, while the increase for the comparison methods is less than 25%, demonstrating its sustained optimization capability in handling high-detail scenes.

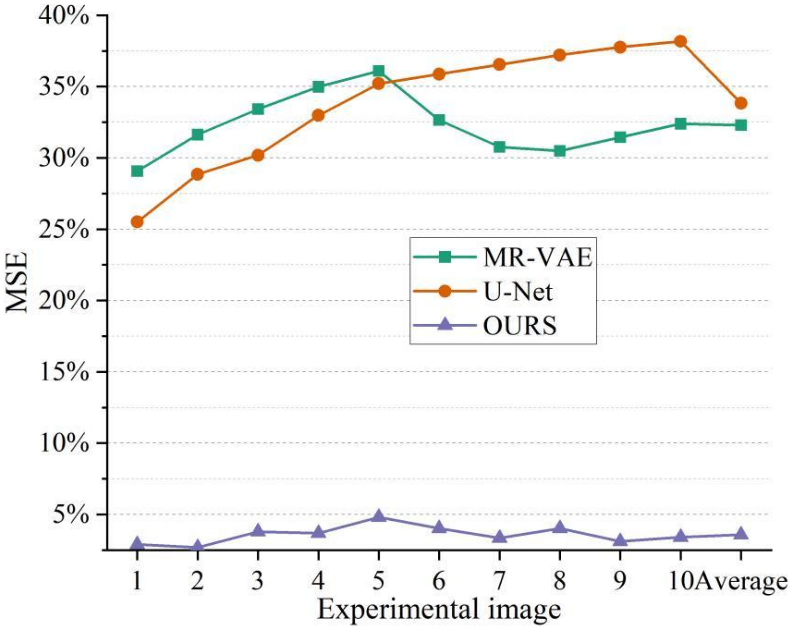

Figure 7. Comparison results of mean square error.

In terms of mean squared error (MSE) for reconstruction accuracy, the proposed method exhibits extremely low and stable error. The MSE of the proposed method (2.68–4.80) is only 1/8 that of MR-VAE (29.07–36.09) and less than 1/7 that of U-Net (25.51–38.17). In images with significant noise interference (e.g., Figure 5), the MSE of the proposed method is only 4.80%, while MR-VAE and U-Net reach 36.09% and 35.20%, respectively, with an improvement in error control capability of over sevenfold.

4. Conclusion

This paper proposes a layered optimization method for VR immersive interactive interfaces based on computer graphics algorithms and human visual characteristics, significantly enhancing the visual communication efficiency and cultural immersion in international Chinese language education. By constructing a layered model centered on visual perception intensity, combining genetic algorithms to solve interface layout problems, and validating the system in the cultural education scenario of the Forbidden City, the model achieves a recognition accuracy rate of 97.23% after parameter optimization (feature space: 128 MB; dictionary space: 512 MB; weight: 0.35).

After parameter optimization, the model achieved a recognition accuracy rate of 97.23% (feature space: 128MB; dictionary space: 512MB; weight: 0.35).

In the VR interface layout with 1,520 elements, 32% of high-intensity regions (levels 4-5) accommodate 87% of high-VCI elements, achieving efficient aggregation of core teaching resources. High-intensity units account for 48% of the central teaching area and 78% of the cultural annotation layer. Approximately 76% of the system functional area units are constrained to the peripheral low-intensity zone, reducing visual interference while lowering GPU load.

The proposed graphics algorithm demonstrates significantly superior performance when processing Chinese educational images, with an average gradient value of 7.46 across 10 images—more than twice that of the MR-VAE method (3.06) and 1.9 times that of the U-Net method. The PSNR peak reaches 21.23, representing an average improvement of 29.7%–40.2% over comparison methods. The mean squared error (MSE) is only 2.68%–4.80%, with an average of 3.68%, less than one-seventh of the comparison methods.

Funding
This research was supported by the General Project of Liaoning Provincial Social Science Planning Office Fund: "Language Feature Analysis and Recognition of Online Emotions in Social Media" (L22BYY006).

References

1. Ho, W. Y. J. (2018). Mobility and language learning: A case study on the use of an online platform to learn Chinese as a foreign language. London Review of Education, 16(2), 239-249.
2. Zhang, Q. (2022). Interactive course design and development for cognitively inspired distance international Chinese education. Computational Intelligence and Neuroscience, 2022(1), 5040920.
3. Lampropoulos, G. (2025). Combining artificial intelligence with augmented reality and virtual reality in education: Current trends and future perspectives. Multimodal Technologies and Interaction, 9(2), 11.
4. Xie, Y., Chen, Y., & Ryder, L. H. (2021). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. Computer Assisted Language Learning, 34(3), 225-245.
5. Xia, W., Wang, G., & Zhao, Q. (2024, December). Research on the Application of Virtual Reality in Chinese Education and Future Development. In Proceedings of the 2024 2nd International Conference on Information Education and Artificial Intelligence (pp. 665-673).
6. Zhang, H., Cui, Y., Shan, H., Qu, Z., Zhang, W., Tu, L., & Wang, Y. (2020, June). Hotspots and trends of virtual reality, augmented reality and mixed reality in education field. In 2020 6th international conference of the Immersive Learning Research Network (iLRN) (pp. 215-219). IEEE.
7. Wei, Z., & Yuan, M. (2023). Research on the current situation and future development trend of immersive virtual reality in the field of education. Sustainability, 15(9), 7531.
8. Cheng, H. J. (2021). Integrating augmented reality and virtual reality technology into Chinese education: An observation from nine cases. In Contextual Language Learning: Real Language Learning on the Continuum from Virtuality to Reality (pp. 147-170). Singapore: Springer Singapore.
9. Zhao, J. H., & Yang, Q. F. (2023). Promoting international high-school students' C hinese language learning achievements and perceptions: A mind mapping-based spherical video-based virtual reality learning system in C hinese language courses. Journal of computer assisted learning, 39(3), 1002-1016.
10. Li, Q., Liu, Q., Jiang, M., & Wang, R. (2024). Virtual reality as a bridge: enhancing intercultural learning and communication skills among international students. International Communication of Chinese Culture, 11(3), 385-411.
11. DeWitt, D., Chan, S. F., & Loban, R. (2022). Virtual reality for developing intercultural communication competence in Mandarin as a Foreign language. Educational technology research and development, 70(2), 615-638.
12. Ma, C. (2018). Teaching application of computer virtual reality technology in international education of Chinese language. Kuram ve Uygulamada Egitim Bilimleri, 18(6), 3837-3843.
13. Liu, W. (2023). A systematic review of the use of virtual reality in teaching Chinese as a foreign language. Journal of China Computer-Assisted Language Learning, (0).
14. Barrett, A., Pack, A., Guo, Y., & Wang, N. (2023). Technology acceptance model and multi-user virtual reality learning environments for Chinese language education. Interactive Learning Environments, 31(3), 1665-1682.
15. Chen, C., & Yuan, Y. (2025). Effectiveness of virtual reality on Chinese as a second language vocabulary learning: Perceptions from international students. Computer Assisted Language Learning, 38(3), 397-425.
16. Xiaoning, X., Xiaoyan, L., & Nazarudin, M. N. (2024). Sustaining Chinese Education with Online VR Technology: A Systematic Review. Jurnal Pendidikan Malaysia, 49.
17. Xie, Y., Ryder, L., & Chen, Y. (2019). Using interactive virtual reality tools in an advanced Chinese language class: A case study. TechTrends, 63, 251-259.