

Article

# Research on Optimization Method of VR Task Scenario Resources Driven by User Cognitive Needs

Qianwen Fu, Jian Lv \*, Zeyu Zhao  and Di Yue

Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, Guiyang 550025, China; fuqw9312@gmail.com (Q.F.); lsssgxty@gmail.com (Z.Z.); yued2566@gmail.com (Y.D.)

\* Correspondence: jlv@gzu.edu.cn

Received: 24 December 2019; Accepted: 24 January 2020; Published: 26 January 2020



**Abstract:** Research was performed in order to improve the efficiency of a user's access to information and the interactive experience of task selection in a virtual reality (VR) system, reduce the level of a user's cognitive load, and improve the efficiency of designers in building a VR system. On the basis of user behavior cognition-system resource mapping, a task scenario resource optimization method for VR system based on quality function deployment-convolution neural network (QFD-CNN) was proposed. Firstly, under the guidance of user behavior cognition, the characteristics of multi-channel information resources in a VR system were analyzed, and the correlation matrix of the VR system scenario resource characteristics was constructed based on the design criteria of human-computer interaction, cognition, and low-load demand. Secondly, analytic hierarchy process (AHP)-QFD combined with evaluation matrix is used to output the priority ranking of VR system resource characteristics. Then, the VR system task scenario cognitive load experiment is carried out on users, and the CNN input set and output set data are collected through the experiment, in order to build a CNN system and predict the user cognitive load and satisfaction in the human-computer interaction in the VR system. Finally, combined with the task information interface of a VR system in a smart city, the application research of the system resource feature optimization method under multi-channel cognition is carried out. The results show that the test coefficient CR value of the AHP-QFD model based on cognitive load is less than 0.1, and the MSE of CNN prediction model network is 0.004247, which proves the effectiveness of this model. According to the requirements of the same design task in a VR system, by comparing the scheme formed by the traditional design process with the scheme optimized by the method in this paper, the results show that the user has a lower cognitive load and better task operation experience when interacting with the latter scheme, so the optimization method studied in this paper can provide a reference for the system construction of virtual reality.

**Keywords:** virtual reality system; multi-channel cognition; cognitive load; QFD; human-computer interaction; prediction optimization

## 1. Introduction

The research on the user experience and cognitive load of human-computer interaction in virtual reality system has attracted attention. In the virtual reality (VR) system task scenario, the mapping relationship analysis between the visual expression of multi-channel information resources and the user's hidden cognitive needs is an important part of studying the user's cognitive load and user experience. At the same time, it is very important for designers to predict a user's cognitive load and satisfaction during the system construction process [1–3]. In the field of human-computer interaction, in order to accurately grasp the level of user experience perception, many scholars have provided a valuable research basis for the analysis of user cognitive behavior and satisfaction.

In the cognitive theory of user behavior, scholars have studied the cognitive load of human–computer interaction from a cognitive mechanism and an information coding aspect. Scholar Cheng Shiwei [4] and others proposed a resource model based on distributed cognition. In the research of user cognition theory, Lu Lu [5] and other scholars proposed a multi-channel information cognition processing model. Li Yang [6] et al. designed badminton experimental scenes under VR conditions and added seven modal clues to study the influence of multi-channels on moving target selection performance and subjective feelings. Paquier Mathieu [7] et al. discussed the self-centered distance perception method of users under the alternation of visual and auditory peaks of virtual objects in distance dimension. Lei Xiao [8] et al. summarized the use of tactile clues to interact with other sensory stimuli to predict potential perceptual experiences in multi-sensory environments. Geitner Claudia [9] and others extended the research on multimodal warning performances. The above research shows that the user’s information perception ability in multi-channel is greater than that in single channel, so this paper divides the information input in VR system into three channels: visual, auditory and tactile.

The quality function deployment (QFD) method plays a bridge role in the research of user satisfaction between user requirements and design elements. QFD is a process that dynamically converts user requirements into design, parts and manufacturing. The analytic hierarchy process (AHP) is usually used to process the collected user requirements, and matrix tools are combined to integrate various data, and the house of quality is used to form visual engineering feature proportions. Kathiravan [10] and others have improved the performance of QFD in the process of user-oriented product design. Shi, Yanlin and Qingjin Peng [11] et al. improve the feedback capability of QFD by distinguishing the needs of different passengers on a high-speed rail. Geng Xiuli [12] and others proposed a customer demand-driven module selection method for product service systems. Li Fei [13] et al. proposed a method for calculating and transforming the importance of user requirements based on double-layer correlation. In cognitive psychology, user cognitive load assessment is mainly divided into subjective assessment and physiological measurement. Lu Kun [14] and others have carried out experimental measurement and mathematical modeling research on mental load for the prediction of user mental load in an aircraft cockpit display interface. Shengyuan Yan [15] et al. analyzed the cognitive psychology of users in the emergency operation procedures of nuclear power plants through NASA mission load index and eye movement experiments, and then optimized the layout of the operation interface. Emami [16] et al. optimized the operation interface through brain–computer interface (BCI) to reduce visual interference and thus reduce the cognitive load of users. At the level of predicting user satisfaction, the use of the neural network method has attracted attention. Yan Bo [17] et al. used product usage data to establish a user perception evaluation model and predicted user perception satisfaction through back propagation neural network (BP). Diego-Mas Jose A [18] proposed a user experience modeling method based on neural network prediction.

To sum up, previous studies have not established a predictive feedback mechanism between resource elements and cognitive behaviors in the field of VR system resource optimization, and lack hierarchical analysis of the correlation between design resource elements and user’s cognitive behaviors in VR systems. Considering this, in order to coordinate information capacity and user cognition in human–computer interaction, Based on the research results and theories of previous scholars combined with the existing problems, this paper proposes a cognitive load forecasting model based on the mapping of user cognitive behavior and system design resource elements under VR system multi-perception channels. Taking a smart city as an example, a model is established to sort out explicit design resources in order to obtain implicit user needs.

## 2. Research Framework

- (1) Building VR system cognitive resource space: extracting user behavior characteristics and corresponding design resource characteristics from the visual perception channel, auditory perception channel, and tactile perception channel, and then analyzing the mapping relationship between explicit coding and implicit cognition of information representation under multi-channel.

In the mapping relationship analysis, users receive feedback information from virtual reality software and hardware through physical channels, then generate cognitive behaviors through feedback information, and then make decisions on tasks in the system. On this basis, users' VR system cognitive resource space is built.

- (2) Establishing QFD design element feature transformation space: focusing at a user's cognitive low-load demand, AHP and QFD are used to analyze the relevant importance of VR system visual resources, auditory resources and tactile resources, and obtain the importance ranking. Designers can refer to the ranking of the importance of each design resource aiming at the user's cognitive load demand when making design decisions, thus assisting designers to carry out efficient design.
- (3) Neural network model predicts user's cognitive load: according to the characteristics of the convolution neural network (CNN)'s nonlinear expression of variable relations, the cognitive load of users in VR system task scenarios is predicted and analyzed, thus assisting designers in building a VR system efficiently and accurately. In the neural prediction results, the system configuration scheme with the highest cognitive load value and the system configuration scheme with the lowest cognitive load value are retrieved, which can provide scheme reference for designers.

The research framework is shown in Figure 1. The specific research content is two to three chapters.

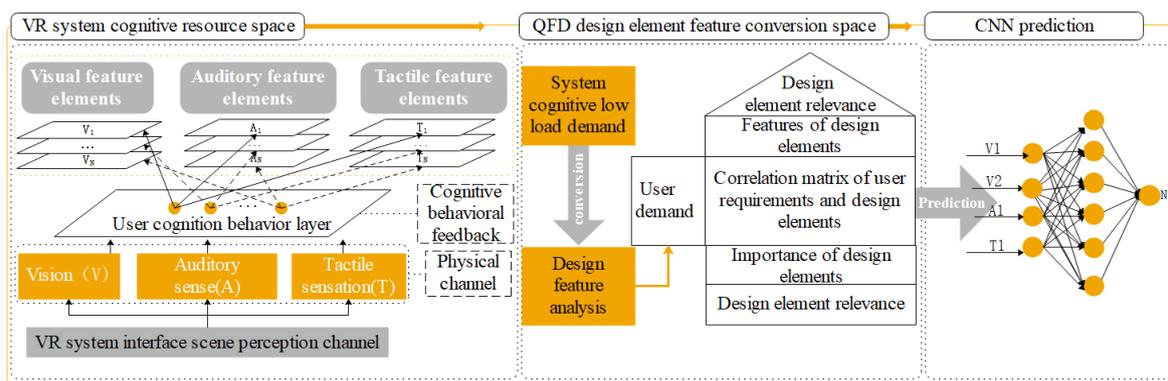


Figure 1. Theoretical framework.

### 3. Cognitive Behavior-Design Resource Mapping Model of VR System Combining Sensory Multi-Channels

#### 3.1. Channel Theory of Cognitive Resources

The theoretical basis of user cognition is mainly the theory of limited resources and graphic perception, which expresses the explicit resources and implicit cognition of VR system information representation. Due to the limited capacity of the user's cognitive resources, it is necessary to reduce a user's cognitive load through multiple channels during information identification [6,8,9], thus improving the cognitive efficiency of a user's experience and task operation scenarios. Therefore, this paper selects visual channels, auditory channels and tactile channels to study a user's cognitive behaviors and design resource characteristics. In system information reading and task operation, the computer perceives the user's behavior and converts it into encodable data. The operation process is a multi-channel perceived information input: firstly, the user receives information stimulation through multi-channel senses and stores it; then, short-term and long-term memory is called through operation perception to compile information and make decisions. Finally, the user executes corresponding actions according to the decision results to realize information output, and the user's cognitive information flow is shown in Figure 2. This paper analyzes user behavior through information multi-channel fusion to deconstruct VR system resource characteristics.

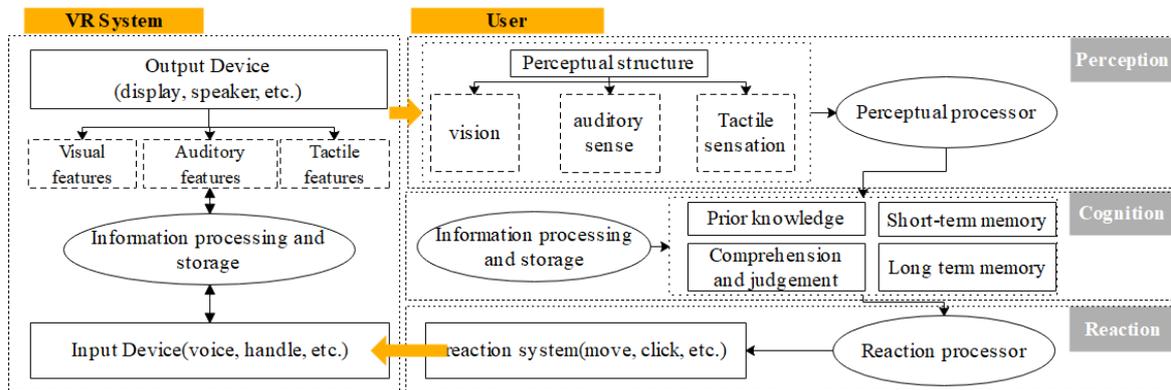


Figure 2. User cognitive information process.

3.2. Construction of Cognitive Behavior Design Feature Model

Based on the user’s cognitive psychology and VR resource characteristics in information transmission, the mapping relationship between the user’s cognitive behavior and resource features can be analyzed, and a cognitive behavior-design element feature model framework can be established, as shown in Figure 3. Firstly, a low-load cognitive channel domain, a cognitive behavior feedback domain and a design resource feature domain are established in order to obtain physical perception information, and then the importance of the user’s cognitive low-load requirements is transferred to the importance of design resource feature elements.

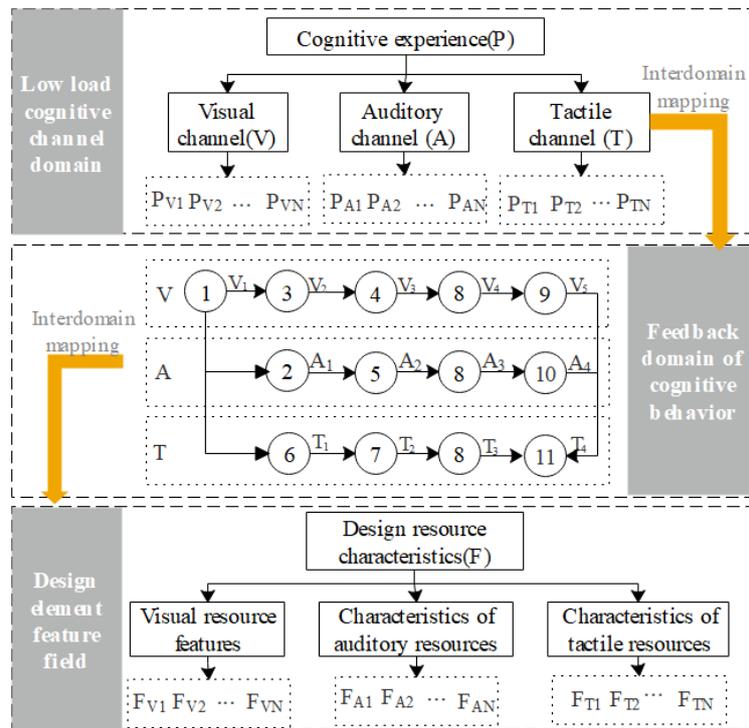


Figure 3. Cognitive behavior-design feature model.

1. Low-load cognitive channel domain: in VR task scenarios, explicit visual codes such as interface data pass through visual perception channels, background music and voice reminders pass through the auditory perception channel, and VR handle vibration feedback and task operation pass through the tactile sensing channel. The reception of explicit knowledge in the three channels affects each other, and there is a parallel, dependent and enabling relationship, which reduces

the cognitive resources in a single channel dimension, thus reducing the cognitive load of users. The low-load cognitive channel domain is shown in Figure 4, wherein P represents the user’s cognitive experience, V, A, and T respectively represent the visual channel, auditory channel, and tactile channel,  $\{P_{V1}, P_{V2} \dots P_{VN}\}$  represents the user’s cognitive experience under the visual channel,  $\{P_{A1}, P_{A2} \dots P_{AN}\}$  represents the user’s cognitive experience under the auditory channel, and  $\{P_{T1}, P_{T2} \dots P_{TN}\}$  represents the user’s cognitive experience under the tactile channel.

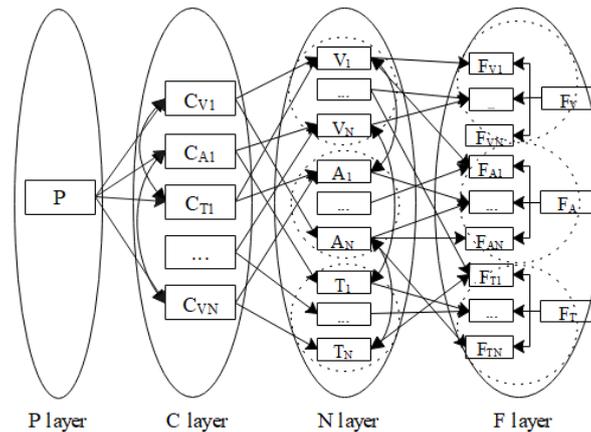


Figure 4. Mapping between domains.

2. Cognitive behavior feedback domain: the human–machine system based on task requirements uses the cognitive behavior-design resource feature network modeling method [19] to establish the user cognitive-behavior library. The user behavior selected by VR task is modeled and regulated in time sequence, and the operation and information channels are organically combined to intuitively reflect the interrelation between various behavior elements. The behavior elements are shown in Table 1 and the feedback field is shown in Figure 4. The numbers represent the sequence of the user’s actions during VR operation, and V, A, and T represent visual channel, auditory channel, and tactile channel respectively. Through the decomposition of user behavior, the corresponding behavior element requirements are obtained, such as easy discovery, easy understanding, convenient regulation, etc.

Table 1. Elements of user behavior feedback in cognitive channel.

Vision (V)	Auditory Sense (A)	Tactile Sensation (T)
V <sub>1</sub> Find	A <sub>1</sub> Find	T <sub>1</sub> Regulation
V <sub>2</sub> Browse	A <sub>2</sub> Understand	T <sub>2</sub> Operation
V <sub>3</sub> Search	A <sub>3</sub> Check feedback	T <sub>3</sub> Check feedback
V <sub>4</sub> Check feedback	A <sub>4</sub> Analysis	
V <sub>5</sub> Contrast		

3. Design resource feature domain: A VR system contains multi-dimensional information perception resource features. As shown in Figure 4, VR task scenario design resource features are deconstructed, where visual channel information is expressed as  $\{F_{V1}, F_{V2} \dots F_{VN}\}$ , and includes schema shape, color, etc. Auditory channel information is represented as  $\{F_{A1}, F_{A2} \dots F_{AN}\}$ , and includes background music, prompt tones, etc.; tactile channel information is expressed as  $\{F_{T1}, F_{T2} \dots F_{TN}\}$ , and includes the frequency and amplitude of the operating lever vibration.

### 3.3. Mapping Relationship between Domains of Cognitive Behavior-Design Feature Model

The mapping relationship between domains is required to transfer the importance of a user’s cognitive load to the resource characteristics of each information channel. In the model, P represents

the user’s cognition of low-load targets, C represents the tacit knowledge characteristics of VR cognitive criteria under target constraints, N represents the user’s behavioral needs under each information perception channel, and  $N = [N_\alpha \ N_\beta \ N_\gamma \ \dots]$ , where  $\alpha$ ,  $\beta$ , and  $\gamma$  represent a class of cognitive channels. The user feedback behavior under the action of each cognitive channel is expressed by X, assuming that the user behavior set X is expressed as:

$$X = [x_1 \ x_2 \ x_3 \ \dots \ x_m]^T$$

$$N_\mu = X_\mu$$

For the virtual reality research object, under the cognitive visual, auditory, and tactile channels, the user’s behavioral needs under each information perception channel are expressed as follows:

$$N = [N_\alpha \ N_\beta \ N_\gamma \ \dots] = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mn} \end{bmatrix}. \tag{1}$$

If F is the explicit knowledge feature of the design resource feature under each perception channel, then the general model of ontology knowledge of a VR system selection task scenario is formally characterized as follows based on Backus–Naur form (BNF):

$$\{C\}:: = [ \text{immersion}][ \text{Fluency}][ \text{Sense of pleasure} [ \dots ] ],$$

$$\{N\}:: = [ \text{Interactive nature}][ \text{Timely feedback}][ \text{Visual beauty}][ \dots ] ,$$

$$\{F\}:: = [ \text{Graphic shape}][ \text{Main tonal}][ \text{Prompt tone}][ \text{Earthquake frequency}][ \dots ] .$$

Among them, the mapping relationship is an abstract expression of the relationship between cognitive channels, user behaviors and design resource features. The specific implementation method is QFD transfer and the allocation of cognitive low-load user demand value elements. Mapping relationships have one-to-one, one-to-many and many-to-many relationships and interrelated influence relationships within the hierarchy, as shown in Figure 4. The P layer is the cognitive load layer, which transfers the user requirements of the P layer to the C layer, and the C layer is the virtual reality sensory criterion layer. The criterion layer will have intra-group association influence, and the importance of the criterion layer is transferred to the user behavior requirement layer of the N layer. The user behavior is fed back into the virtual reality resources, and the virtual reality resource layer is the F layer.

#### 4. Scenario User Cognitive Load Forecasting in VR System

##### 4.1. Design Resource Feature Priority Calculation Model with Cognitive Low Load

This paper uses the analytic hierarchy process (AHP) to analyze a user’s cognitive load and design resource characteristics in VR task selection scenarios. The AHP is a hierarchical weight decision analysis method, which integrates expert experience and theoretical data, and can realize the effective and unified combination of qualitative and quantitative aspects, and more objectively transfer the importance of user cognitive load in virtual reality. Based on the research of virtual reality cognitive theory, the specific implementation steps are as follows:

**Step 1:** A correlation model for calculating the importance of a user’s cognition of low-load demand is established. The model consists of four levels: target P, criterion level  $C_i$  ( $i=1, 2, \dots, n$ ), cognitive behavior requirement level  $N_{i1}, \dots, N_{in}$  and design feature level  $F_i$ .

**Step 2:** Taking the target layer P as the judgment criterion, the criterion layer correlation matrix is constructed, the criterion layer elements  $C1, C2 \dots Cn$  are compared with  $C1$  in turn, and the correlation comparison matrix A of user cognitive low-load demand  $N_{11}$  based on VR situation is established.

$$A = \begin{matrix} & C_1 & C_2 & \cdots & C_n \\ \begin{matrix} C_{11} \\ C_2 \\ \vdots \\ C_n \end{matrix} & \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1m_1} \\ e_{21} & e_{22} & \cdots & e_{2m_1} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n_11} & e_{n_12} & \cdots & e_{n_1m_1} \end{bmatrix} \end{matrix} . \tag{2}$$

The matrix A is weighted to judge the correlation between each element in the criterion layer for P, and the correlation degree value is e, which is assigned by a 0–9 scale method. The importance judgment index is shown in Table 2 The square root method is used to calculate the maximum eigenvalue  $\lambda_{max}$  and eigenvector W of the judgment matrix. First, we calculate the product of each row in matrix A,  $M_i = \prod_{j=1}^n b_{ij}$ ,  $i = 1, 2, \dots, n$ . Then we find the orientation quantity  $\tilde{w} = [\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_n]^T$ ,  $i = 1, 2, \dots, n$ . It is found that  $W = [\omega_1, \omega_2, \dots, \omega_n]^T$  is the eigenvector of the judgment matrix A, and the eigenvector is the importance degree of the criterion layer  $C_1, C_2 \cdots C_n$  with P as the judgment standard. Secondly, the maximum eigenvalue is calculated,  $\lambda_{max} = \sum_{i=1}^n \left[ \frac{(B \cdot W)_i}{n\omega_i} \right]$ , where  $(B \cdot W)_i$  is the i-th component of the product of the judgment matrix A and the eigenvector W. Finally, the eigenvectors are accumulated into the calculation of the lower correlation degree.

Table 2. Importance judgment index.

Scale	Meaning
1	The two elements are of equal importance compared to each other.
3	Compared with the two elements, the former is slightly more important than the latter.
5	Compared with the two elements, the former is obviously more important than the latter.
7	Compared with the two elements, the former is much more important than the latter.
9	Compared with the two elements, the former is more serious and important than the latter.
2, 4, 6, 8	The intermediate value of the above-mentioned adjacent judgment

**Step 3:** Based on the criterion layer C, the elements  $\{N_{11}, N_{12}, \dots, N_{in}\}$  in the cognitive behavior demand layer are compared with  $N_{11}$  in turn, and the matrix B of  $N_{11}$  correlation comparison based on cognitive criteria is constructed.

$$B = \begin{matrix} & N_{11} & N_{12} & \cdots & N_{1m_1} \\ \begin{matrix} N_{11} \\ N_{21} \\ \vdots \\ N_{1n_1} \end{matrix} & \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1m_1} \\ e_{21} & e_{22} & \cdots & e_{2m_1} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n_11} & e_{n_12} & \cdots & e_{n_1m_1} \end{bmatrix} \end{matrix} . \tag{3}$$

**Step 4:** Method obtains the maximum eigenvalue of each judgment matrix and its corresponding eigenvector, and classifies the eigenvectors of cognitive behavior demand layer into the matrix  $W_{ij}$  formula in turn, so that matrix  $W_{ij}$  represents the correlation information between user cognitive behavior demand  $N_{in}$  and  $N_{jn}$ . By comparing the link relations of cognitive behavior requirements in turn, the weightless relation matrix  $W_p$  is obtained to gain the priority weight value.

$$W_{ij} = \begin{bmatrix} w_{j1}^{i1} & w_{j1}^{i2} & \cdots & w_{j1}^{in_i} \\ w_{j2}^{i1} & w_{j2}^{i2} & \cdots & w_{j2}^{in_i} \\ \vdots & \vdots & \ddots & \vdots \\ w_{jn_1}^{i1} & w_{jn_2}^{i2} & \cdots & w_{jn_i}^{in_i} \end{bmatrix} , \tag{4}$$

$$W_p = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ & & \ddots & \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}. \tag{5}$$

**Step 5:** Taking the cognitive behavior demand layer N as the judgment criterion, the correlation degree of design features  $\{F_1, F_2, \dots, F_n\}$  and  $F_1$  is compared in pairs in turn, and a judgment matrix of  $F_1$  correlation degree comparison based on the cognitive criterion is constructed to express the link relationship of the design features  $\{F_1, F_2, \dots, F_n\}$ .

$$W_R^F = \begin{matrix} & & F_1 & F_2 & \cdots & F_a \\ \begin{matrix} N_{11} \\ N_{12} \\ \vdots \\ N_{nn} \end{matrix} & \begin{bmatrix} e_{11}e_{12} & \cdots & X_{1a} \\ e_{21} & e_{22} & \cdots & e_{2a} \\ & & \ddots & \\ e_{n1} & e_{n2} & \cdots & e_{na} \end{bmatrix} \end{matrix}. \tag{6}$$

According to the above analysis, the design feature importance  $w_b$  ( $b = 1, 2, \dots, a$ ) can be obtained through the criterion importance  $r_{in_i}^w$  and the correlation degree  $X_{ib}$  between the user’s cognitive behavior requirements and technical features, and the design feature layer feature vector and priority ranking can be calculated according to the above steps.

$$w_b = \sum_{i=0}^{n_n} r_{in_i}^w X_{ib}. \tag{7}$$

The above four steps consider the correlation between the user’s cognitive behavior requirements and design resource features in a VR task selection scenario system, which requires low cognitive load. In this paper, the importance of a user’s cognitive behavior needs is transformed into the importance of specific design resource features. According to the quantified priority ranking, under the scenario demand of low cognitive load, the resource features with higher importance are considered first, and the unimportant design resource features and conflicting design resource features are considered second. In this process, through objective and accurate analysis of a user’s experience and cognitive needs and clear design direction, innovative methods are used to realize the VR system task scenarios and build a specific design direction, thus ensuring the effectiveness of the design scheme.

#### 4.2. Forecast Model Task Flow

In the prediction model, the convolution neural network is used to predict the cognitive load value of users. CNN is a kind of artificial neural network with a high efficiency recognition ability. It adopts the method of local linking and sharing weights. It obtains representations from the original data by alternately using each pooling layer of convolution layer, automatically extracting local features of the data, and establishing feature vectors. The application of the CNN method is to enter the convolution layer first, and extract the spatial information between features through the convolution + pooling method. The convolution layer convolves the overall data and extracts the spatial information through the convolution kernel. The pooling layer reduces the parameter dimension of the model and improves the training efficiency of the model. CNN’s training algorithm is divided into two stages: the forward propagation stage, which takes a sample from the sample set and inputs it into the network, and then calculate the corresponding actual output, and the backward propagation stage, which calculates the difference between the real result and the expected result, and then adjusts the weight matrix according to the method of minimizing the error.

In the convolution layer, a different convolution check input sets are used for convolution operation, and the corresponding feature data will be obtained by activating the function. The general mathematical expression of convolution is as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (8)$$

where:  $l$  is the Layer  $l$  network;  $b_j^l$  is biased;  $k_{ij}^l$  is the weight matrix;  $x_j^l$  is the Layer  $l$  output;  $x_i^{l-1}$  is the Layer  $l$  input;  $M_j$  is the  $j$ th convolution region of  $l-1$  layer characteristic graph;  $f(\cdot)$  is an active function. In CNN, ReLU is usually selected as the activation function, and its mathematical expression is:

$$f(x) = \max(0, x) \quad (9)$$

After passing through the convolution layer, the number of features will increase. If multiple convolution operations are carried out, the feature dimension will explode. In order to solve this problem, the common method is to add a pooling layer after the convolution layer. Its function is to reduce the amount of data processing while maximizing the retention of effective information. Common pooling methods include: meaning pooling, maxing pooling, stochastic pooling, The general mathematical expression for pooling is:

$$x_{i+1} = f(\beta \text{down}(x_i) + b) \quad (10)$$

where:  $x_i$  is the input,  $x_{i+1}$  is the output,  $\beta$  is multiplicative bias,  $b$  is additive bias,  $\text{down}()$  is pooling function, and  $f(\cdot)$  is activation function.

The data of the input set will obtain the advanced features of the input set after convolution pooling operation. The full connection layer weights these advanced features and then obtains the output through the activation function. The general mathematical expression of the full connection layer is:

$$y^k = f(\omega^k x^{k-1} + b^k) \quad (11)$$

where:  $x^{k-1}$  is the input of the full connection layer,  $y^k$  is the output of the full connection layer,  $\omega^k$  is the weight coefficient,  $b^k$  is the additive bias,  $k$  is the serial number of the network layer, and  $f(\cdot)$  is the activation function. In the full connection layer, Softmax activation function is often used for multi-classification prediction tasks.

The logical task process of the prediction model is a process of predicting and evaluating an interactive selection through the neural network according to the relationship between VR system resource characteristics and cognitive load. The specific process is shown in the Figure 5.

**Step 1:** Takes the VR system information interface scheme feature target as input, and based on it, the design resource features of interface scenario and multimodal perception channel are selected.

**Step 2:** Based on AHP-QFD, the priority ranking of design resource features is taken as a reference-aided design for design schemes.

**Step 3:** Involves building the virtual reality task selection scenario system.

**Step 4:** Use the input of the CNN neural network to detect whether the built design scheme meets the user's need to recognize low-load and design requirement constraints, and return to Step 2 if it does not.

**Step 5:** If the design constraints are not violated, the scheme is saved and implemented.

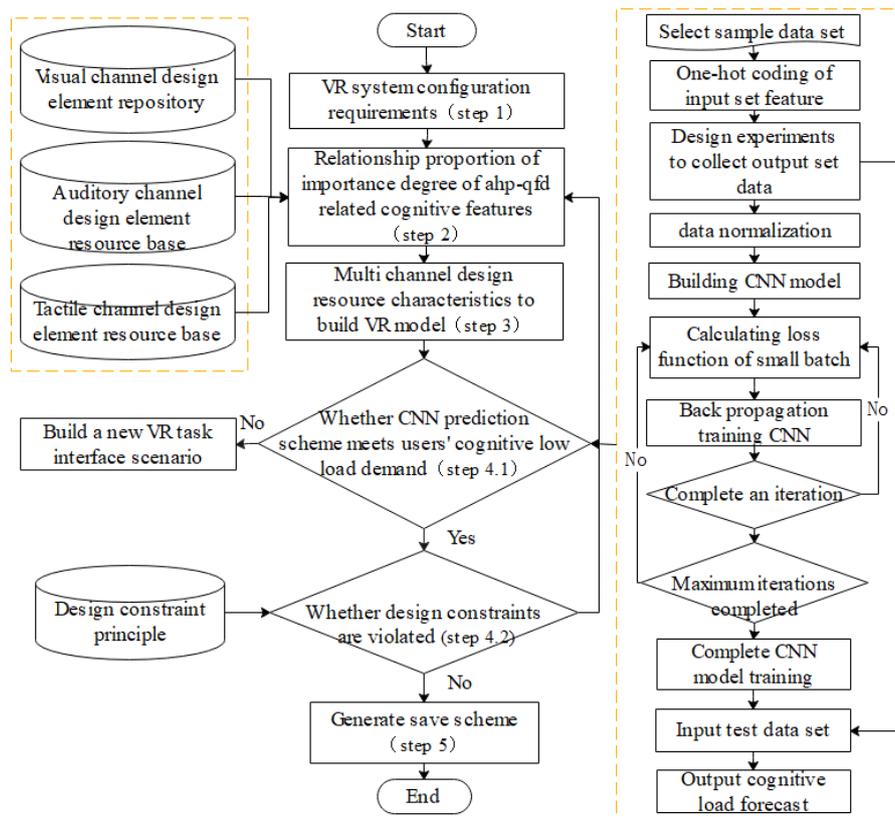


Figure 5. Prediction model task flow.

## 5. Application Case

In the information interface scenario of a VR system in a smart city, the main task requirement of users is to understand the general city layout and various index information. The user behavior is distributed as; reading, listening, searching, and tactile perception behavior modules. Because the information in the smart city system interface is dense, and the users who carry out the interactive experience are mostly users who are unfamiliar with VR system operation, they are prone to confusion and obstructions in their experience. Therefore, reducing the cognitive load of users and then improving the operation efficiency of users to help them complete tasks in a VR system is an urgent problem to be solved in VR interface task scenarios.

The ontology knowledge of virtual reality system is as follows: the VR system in this study is modeled by Rhinoceros, and the virtual reality development environment is as follows: the experimental platform is built with AMD 1800X CPU; GPU is NVIDIA® GeForce® GTX 1070; 16 GB of RAM; the operating system is Windows 10; the system development platform is UNREAL ENGINE 4.21. 1. Corresponding virtual reality equipment: HTC VIVE/HTC VIVE PRO; VIVE head-mounted equipment; VIVE control handle; VIVE locator. The construction of cognitive system is based on user cognitive theory and user cognitive behavior analysis.

### 5.1. Acquisition of User Cognitive Behavior Requirements in VR Task Selection System

In this paper, the principle of analytic hierarchy process is selected to carry out stratification, analytic hierarchy process (AHP) takes a complex multi-objective decision-making problem as a system and decomposes the target into several levels of multiple indexes. This method obtains the priority weight of each element in each level to a certain element in the previous level, and finally calculates the single ranking and the total ranking of the levels after weighting. The cognitive and design features of a VR system are layered, and the analytic hierarchy process (AHP) representation method is used to

divide the VR system into four levels, namely, target layer (P), criterion layer (C), cognitive behavior requirement layer (N), and design features layer (F).

Firstly, the VR system cognitive low-load is set as the target layer and secondly, the criterion layer is set. This paper collects the key words of VR system usage cognitive criteria through literature inquiry, system construction expert interview, VR system user interview and other channels, and 16 subjects were invited to determine the most suitable words in expressing the virtual reality situation. According to the number of votes, 33 criteria images were preliminarily screened, as shown in Table 3, and then the structure relation of virtual reality scene criterion words was further excavated. The subjects were invited to carry out the semantic grouping experiment. After evaluation and scoring, the subjects used artificial classification to classify the words they thought had similar meanings into a group (the number of words in each group could be different). After counting the same number of groups, Matlab obtained the matrix through the following operations, as shown in Table 4, and then imported the obtained data into SPSS data statistical software for clustering analysis to obtain the tree diagram shown in Figure 6. In this figure, the criterion words are divided into four groups in Table 5, with the group representative being the one closest to the center point, which are “immersion”, “visualization”, “fluency” and “pleasure”. Then, through interviews with system building experts and users, we use behavioral needs to set the cognitive behavioral needs layer. The user cognitive behavior requirements for the VR system interaction criteria are “natural interaction operation”, “real scene space”, “data visualization”, “matching of functional scene elements”, “clear information level”, “timely feedback”, “visual aesthetics”, “easy mastery” and “learning”. Finally, the design features of visual channel, auditory channel, and tactile channel in the virtual reality system are deconstructed respectively. Table 6 show a list of relationships establish according to design objectives and criteria.

Table 3. Criterion word set.

Criterion Word	Vote	Criterion Word	Vote	Criterion Word	Vote	Criterion Word	Vote
1 Technological	13	10 Lively	9	19 Enjoyable	15	28 Clear	11
2 Comfort	7	11 Charismatic	11	20 Casual	9	29 Rational	15
3 Visualization	14	12 Cheerful	8	21 Safe	16	30 Neat	8
4 Fluency	14	13 Gorgeous	8	22 Advanced sense	8	31 Natural	13
5 Easy to use	6	14 Smooth	13	23 Beautiful	9	32 Immersion	18
6 Relaxed	12	15 Substitution	10	24 Dreamy	17	33 Simple	9
7 Intuitive	15	16 Real	11	25 Endurable	9		
8 Sequential	11	17 Novel	12	26 Dynamic	13		
9 Ingenious	12	18 Pleasure	12	27 Trustworthy	12		

Table 4. Criterion word clustering matrix.

Criterion Word	1	2	3	4	...	30	31	32	33
1	0	0	0	0	...	1	5	13	0
2	0	0	0	0	...	0	3	0	1
3	0	0	0	0	...	3	2	1	2
4	0	0	0	0	...	9	0	0	13
...	...	...	...	...	...	...	...	...	...
30	1	0	3	9	...	0	1	0	5
31	5	3	2	0	...	1	0	4	1
32	13	0	1	0	...	0	4	0	1
33	0	1	2	13	...	5	1	1	0

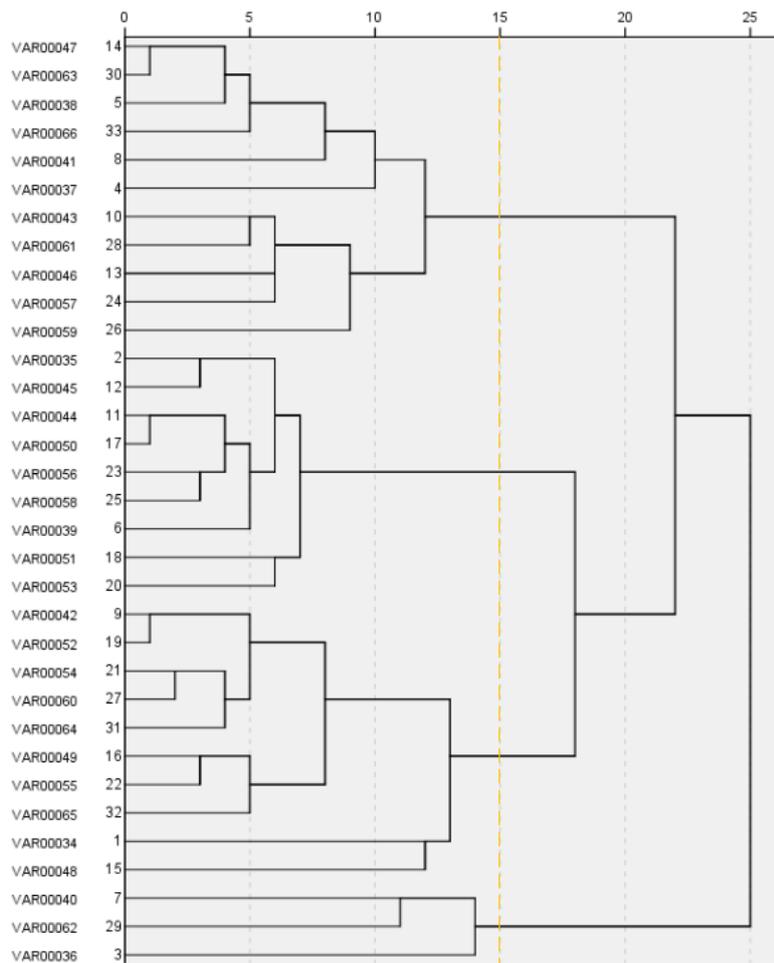


Figure 6. Cluster tree diagram.

Table 5. Criterion word group.

Group 1	Group 2	Group 3	Group 4
3 Visualization	1 Technological	9 Ingenious	2 Comfort
7 Intuitive	15 Substitution	19 Enjoyable	12 Cheerful
29 Rational	16 Real	21 Safe	11 Charismatic
	22 Advanced sense	27 Trustworthy	17 Novel
	32 Immersion	31 Natural	23 Beautiful
			25 Endurable
			6 Relaxed
			18 Pleasure
			20 Casual
			10 Lively
			28 Clear
			13 Gorgeous
			24 Dreamy
			26 Dynamic
			14 Smooth
			30 Neat
			5 Easy to use
			33 Simple
			8 Sequential
			4 Fluency

Table 6. Hierarchical list of the user’s cognition of low-load demand.

Target Layer (P)	Criteria Layer (C)	Cognitive Behavioral Demand Layer (N)	Design Feature Layer (F)
Cognitive low-load (P)	Immersion (C <sub>1</sub> )	Natural interaction operation (N <sub>1</sub> )	Interface Layout (F <sub>1</sub> )
		Real scene space (N <sub>2</sub> )	Graphic area chamfering feature (F <sub>2</sub> )
	Visualization (C <sub>2</sub> )	Data Visualization (N <sub>3</sub> )	Main Tone (F <sub>3</sub> )
		Functional scenario element matching (N <sub>4</sub> )	Color Contrast between Mission Area and Overall (F <sub>4</sub> )
	Fluency (C <sub>3</sub> )	Clear information level. (N <sub>5</sub> )	Interface transparency (F <sub>5</sub> )
		Timely feedback (N <sub>6</sub> )	Browse Order (F <sub>6</sub> )
	Pleasure (C <sub>4</sub> )	Visual aesthetics (N <sub>7</sub> )	Prompt Tone (F <sub>7</sub> )
		Easy to master and learn (N <sub>8</sub> )	Background music (F <sub>8</sub> )
			Vibrating tactile sensation (F <sub>9</sub> )

5.2. Recognition of Design Feature Priority Analysis

Taking the target layer’s cognitive low-load P as the judgment index, the correlation between the criterion layer {C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, C<sub>4</sub>} and P is analyzed to establish a judgment matrix, i.e., the correlation between immersion, visualization, fluency, pleasure, and cognitive low-load, and the eigenvector of the P-C judgment matrix is calculated, as shown in Table 7.

Table 7. P-C judgement matrix.

P	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	Eigenvector W	
C <sub>1</sub>	1	3	1/5	5	0.1178	λ <sub>max</sub> = 4.12 CR = 0.0449
C <sub>2</sub>	1/3	1	1/7	3	0.2634	
C <sub>3</sub>	5	7	1	7	0.5638	
C <sub>4</sub>	1/5	1/3	7	1	0.0550	

In the same way, the whole elements in the cognitive behavior demand layer are set in the criterion layer, and each element in the criterion layer corresponds layer by layer to establish a judgment matrix; they are the judgment matrices of cognitive behavior demand layer in immersion, visualization, fluency, and pleasure, namely the N-C<sub>1</sub> judgment matrix, the N-C<sub>2</sub> judgment matrix and the N-C<sub>3</sub> judgment matrix. The N-C<sub>1</sub> judgment matrix and its eigenvectors are shown in Table 8. The CR values of N-C<sub>1</sub>, N-C<sub>2</sub> and N-C<sub>3</sub> judgment matrices are 0.03565, 0.09576, 0.09336 and 0.9886 respectively, and the CR values are all less than 0.1, which verifies the validity of the matrix. According to Table 9, in the process of building the VR system information interface, users have higher requirements for natural interactive operation, clear information level, matching of situational functional elements, and easy mastery and learning in their cognitive behavior requirements.

Table 8. N-C1 judgment matrix.

C <sub>1</sub>	N <sub>1</sub>	N <sub>2</sub>	N <sub>3</sub>	N <sub>4</sub>	N <sub>5</sub>	N <sub>6</sub>	N <sub>7</sub>	N <sub>8</sub>	Eigenvector W	
N <sub>1</sub>	1	1/2	7	3	5	5	7	2	2.7903	λ <sub>max</sub> = 8.3731 CR = 0.0356
N <sub>2</sub>	2	1	7	4	5	6	8	3	3.7643	
N <sub>3</sub>	1/7	1/7	1	1/5	1/3	1/2	2	1/6	0.3503	
N <sub>4</sub>	1/3	1/4	5	1	2	3	5	1/3	1.1952	
N <sub>5</sub>	1/5	1/5	3	1/2	1	2	3	1/4	0.7401	
N <sub>6</sub>	1/5	1/6	2	1/3	1/2	1	3	1/5	0.5345	
N <sub>7</sub>	1/7	1/8	1/2	1/5	1/3	1/3	1	1/7	0.2701	
N <sub>8</sub>	1/2	1/3	6	3	4	5	7	1	2.1276	

Table 9. Sequence list of cognitive behavioral needs priority.

	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	Priority
	<b>0.1178</b>	<b>0.2634</b>	<b>0.5638</b>	<b>0.0550</b>	
N <sub>1</sub>	2.7903	0.2801	3.4451	1.2391	0.2198
N <sub>2</sub>	3.7643	0.5829	0.5623	3.1847	0.0972
N <sub>3</sub>	0.3503	2.9743	0.3606	0.2554	0.0938
N <sub>4</sub>	1.1953	3.3032	1.2491	1.2228	0.1612
N <sub>5</sub>	0.7401	1.7155	2.2247	0.8602	0.1679
N <sub>6</sub>	0.5346	0.4247	1.3712	0.3737	0.0885
N <sub>7</sub>	0.2701	0.8178	0.2739	0.9554	0.0412
N <sub>8</sub>	2.1277	1.0466	1.3712	2.4771	0.1301

The whole element set in the design feature layer and each element in the cognitive behavior demand layer correspond layer by layer to establish a judgment matrix; they are, respectively, the design feature layer and the judgment matrix with natural interactive operation, real scene space, visual

expression of data, clear information level, real-time feedback, beautiful vision and easy learning and mastering. The judgment matrices established are the F-N<sub>1</sub> judgment matrix, F-N<sub>2</sub> judgment matrix, F-N<sub>3</sub> judgment matrix, F-N<sub>4</sub> judgment matrix, F-N<sub>5</sub> judgment matrix, F-N<sub>6</sub> judgment matrix, F-N<sub>7</sub> judgment matrix and F-N<sub>8</sub> judgment matrix, of which F-N<sub>1</sub> judgment matrix and its eigenvectors are shown in Table 10, and the CR values of other seven groups of judgment matrices are 0.061452, 0.06794, 0.03961, 0.046382, 0.039876, 0.08085, 0.063423 and 0.082875 respectively. CR values less than 0.1 verify the validity of the matrix.

**Table 10.** F-N<sub>1</sub> judgment matrix.

N <sub>1</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	F <sub>7</sub>	F <sub>7</sub>	F <sub>8</sub>	F <sub>9</sub>	Eigenvector W
F <sub>1</sub>	1	1	5	3	6	3	3	8	1/3	2.3469
F <sub>2</sub>	1/7		1/3	1/3	1/2	1/5	1/7	2	1/3	0.3762
F <sub>3</sub>	1/5	3	1	1/4	3	1/3	1/5	3	1/5	0.6399
F <sub>4</sub>	1/3	3	4	1	5	3	1/3	3	1/7	1.2696
F <sub>5</sub>	1/6	2	1/3	1/5	1	1/5	1/6	3	1/5	0.4241
F <sub>6</sub>	1/3	5	3	1/3	5	1	1/3	5	1/3	1.1856
F <sub>7</sub>	1/3	7	5	3	6	3	1	8	3	2.9133
F <sub>8</sub>	1/8	1/2	1/3	1/3	1/3	1/5	1/8	1	1/7	0.2724
F <sub>9</sub>	3	9	5	3	5	3	1/3	7	1	2.8925

$\lambda_{max}=10.0487$   
CR=0.0897

Based on the above research and Table 11, a QFD model for VR task selection clean interface design is established to build the importance of design features, and the correlation between design features is qualitatively analyzed, as shown in Figure 7.

**Table 11.** The overall design features priority sequence table.

	N <sub>1</sub>	N <sub>2</sub>	N <sub>3</sub>	N <sub>4</sub>	N <sub>5</sub>	N <sub>6</sub>	N <sub>7</sub>	N <sub>8</sub>	Overall Priority
	0.2198	0.0972	0.0938	0.1612	0.1679	0.0885	0.0412	0.1301	
F <sub>1</sub>	2.3469	1.5832	3.1781	4.1471	4.1471	1.2599	3.1781	1.9172	0.2102
F <sub>2</sub>	0.3762	0.2394	0.9954	0.6474	0.6551	1.0801	0.9799	0.5748	0.0456
F <sub>3</sub>	0.6399	2.6889	1.5711	0.4447	1.0021	1.1665	4.50576	0.5722	0.0856
F <sub>4</sub>	1.2696	0.4201	2.1131	1.5299	2.2821	1.48909	1.46716	1.1071	0.1098
F <sub>5</sub>	0.4241	3.3303	0.7253	0.2305	0.3141	0.88508	2.19359	0.2405	0.0583
F <sub>6</sub>	1.1856	0.7401	4.0166	3.2386	3.2488	1.53746	0.71299	1.4025	0.1565
F <sub>7</sub>	2.9133	1.2258	0.4538	2.2822	1.5217	3.50277	0.41341	4.4573	0.1753
F <sub>8</sub>	0.2724	0.3619	0.2106	0.3293	0.2335	1.69537	0.26973	0.3502	0.0296
F <sub>9</sub>	2.8925	2.1661	0.3418	0.9755	0.4447	3.05190	0.28638	2.3301	0.1287

As can be seen from Table 11, in the design of the VR system information interface task scenario, we need to first consider the layout of the information interface. once the task selection area and the data information reading area are rationally arranged, we must then consider the setting of the prompt tone, as the perceptual setting of bimodal information fusion under audio-visual consistency will reduce the cognitive load of users, and improve the correct rate of user operation. Then there is the design of the visual browsing sequence, where the focus is on the frequency of text and graphics. Next, setting the tactile vibration of the handle gives the user behavior feedback, and the comparison between the task area and the overall tone will affect the correct rate of the user’s reading information and task selection. Then, the sensory experience of the interface color and the setting of the interface transparency will be considered. A VR interface with transparency will increase the spatial authenticity of the scene through which one passes, and the setting of background music will affect the user’s pleasure. Therefore, designers can refer to the importance ranking provided in Table 12 for scheme design when building the system.

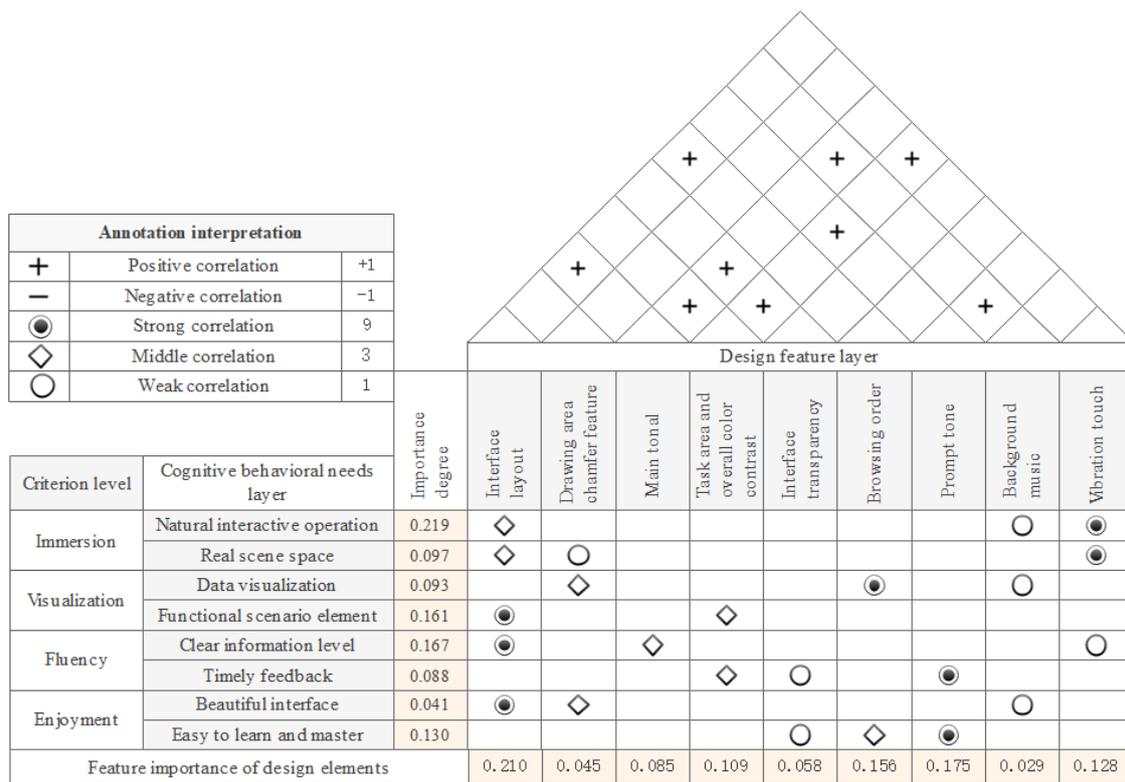


Figure 7. System quality function deployment (QFD).

Table 12. Priority ranking of technical features.

1	2	3	4	5	6	7	8	9
F <sub>1</sub>	F <sub>7</sub>	F <sub>6</sub>	F <sub>9</sub>	F <sub>4</sub>	F <sub>3</sub>	F <sub>5</sub>	F <sub>2</sub>	F <sub>8</sub>

5.3. Forecast Model Input Set Data Collection

The input set data uses the design element analysis method to deconstruct and reanalyze the design element features of the virtual reality interface samples. Under the same interface size and font format/size, the design items are shown in Table 13 as the layout of the operation area, visual browsing sequence, color, transparency, prompt tone, handle vibration, etc. On this basis, the design elements of the virtual reality interface are extracted and the distribution of each design category is determined according to their corresponding elements. Visual channel resource features are processed by artificial intelligence (AI), and auditory channel and tactile resource features are edited by UE4, as shown in Figure 8.

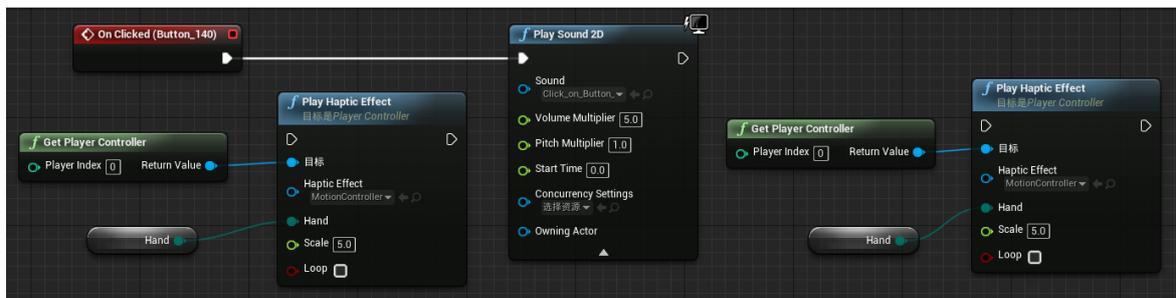
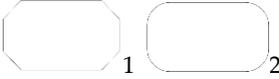
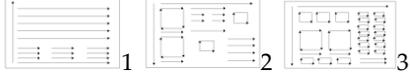


Figure 8. UE4 Sensory channel interactive experience coding blueprint.

**Table 13.** Virtual reality (VR) information interface scenario design element table.

Project	Characteristic Graph Interpretation
Interface Layout	
Graphic area chamfering	
Main Tone	
Color Contrast between Mission Area and Overall	
Visual browsing order	
Interface transparency	Existence -1 Does not exist -2
Prompt Tone	Existence -1 Does not exist -2
Vibrating tactile sensation	Existence -1 Does not exist -2

When sorting out the input set, due to the fact the design feature module is an explicit knowledge feature and belongs to the feature classification value, the one-hot coding method is adopted to extend the values of discrete features to the Euclidean space, with the number 0 representing irrelevant options and the number 1 representing relevant options. Taking the experimental sample 1 as an example, the design feature is decomposed into 21,112,111 corresponding configuration one-hot coding vectors (01000000, 10, 1000, 100, 01, 010, 10, 10, 10, 10). Thirty-nine scene design elements of VR system information interface are processed with input set information according to the one-hot coding mode.

#### 5.4. Data Acquisition of Forecast Model Output Set

The output set data are the VR system task situational cognitive load value and the task response time. For experimental preparation, 39 VR task information interface samples were selected and processed, imported into the UE4 system for task scenario construction, and set horizontal variable parameters. In the formal experiment, 16 people aged between 20 and 26 years old, nine boys and seven girls, were recruited for the cognitive load test. All the subjects had normal or corrected vision, with no defects in visual, auditory, and tactile perception, and were right-handed. Ten subjects had experience in using a VR system and six subjects had no previous experience in using VR. There were two experimental tasks; one was to read the interface data information, and the other was to click the “enter the system” selection area. The experimental scene is shown in Figure 9. We recorded the time (seconds) that the user clicked the task button and measured and recorded the cognitive load value using the NASA-TLX(National Aeronautics and Space Administration-Task Load Index) scale.

The cognitive load was the average cognitive load value and the average reflection time value of the 16-person experiment with the reflection time length, as shown in Table 14.



Figure 9. Experimental scene diagram.

Table 14. Experimental data set.

	One-Hot Encoding	Cognitive Load	Reaction Duration
	0100000010100010001010101010	59.33	1.262
	0000100001010010001010010101	60.8	1.349
	0000100010000110010010101001	63.53	1.362
	0100000001100001010100010101	71.66	1.352
	1000000010010011001010101010	46.6	1.252
...	...	...	...
	0010000010000100101001100101	109.6	1.534

### 5.5. Construction of CNN Prediction Model

Based on the feature analysis of input set and output set data, the CNN model structure oriented to the VR system information interface scenario mainly included the following levels with the following functions: the function of the convolution layer was feature scanning and extraction; the function of the pooling layer was feature filtering; the function of Flatten was to realize data flattening and dimension reduction. The input data of the neural network for predicting the cognitive load of the VR system users was 28 rows and one column. Firstly, six convolution layers were constructed, the first layer taking the form of a one-dimensional convolution, with the number of convolution kernels being 2048, the size set to seven, and the output information dimension being (222,048). The input value of the second layer convolution was the output of the first layer, the number of convolution kernels was 1024, the size was 5, and the output dimension was (181,024). The number of convolution kernels in the third layer was 512, the size was 5, and the output dimension was (14,512). The number of convolution kernels in the fourth layer was 256, the size was 5, and the output dimension was (10,256). The number of convolution kernels in the fifth layer was 128, the size was 3, and the output dimension was (8128). The number of convolution kernels in the sixth layer was 64, the size was 3, and the output dimension

was (664). A pooling layer with a step size of two was set up, and the data reading information of  $1 \times 2$  matrix in  $1 \times n$  was averaged, which reduced general parameters and improved learning efficiency. The output dimension is (364), followed by 192 neuron tiling layers in the eighth layer, 128 neuron full connection layers in the ninth layer, 20 neuron full connection layers in the tenth layer, and one neuron output layer in the eleventh layer.

### 5.6. Validation of Model Results

Samples were selected as test sets for performance testing, and the data in the output layer were normalized and then detected by the function mean square error (MSE). The function expression is:

$$MSE = \frac{1}{p} \sum_{k=1}^p (y_k - y_k^*)^2. \quad (12)$$

If the MSE value is less than 0.01, the CNN model of VR task selection scenario can be proved to be reliable. The user cognitive load test data and the output layer value of the established CNN model were detected, and the measured value was 0.00424 calculated by mean square deviation MSE. Given the MSE value was less than 0.01, test performance of the CNN model was proven to be good. The fitting situation is shown in Figure 10, which shows that the output cognitive load value is basically consistent with the cognitive load value data of the test, and it can be concluded that the established model can complete the correct mapping of the user cognitive load and the design features under the multi-channel behavior analysis.

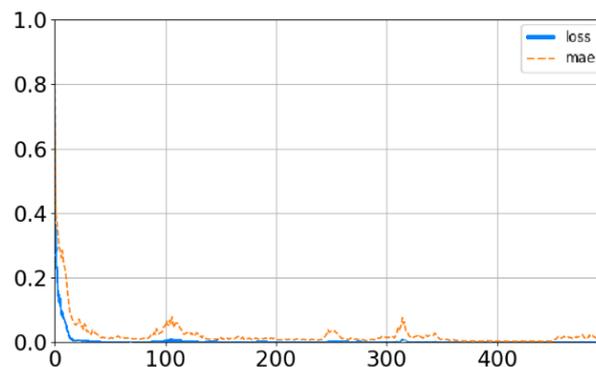
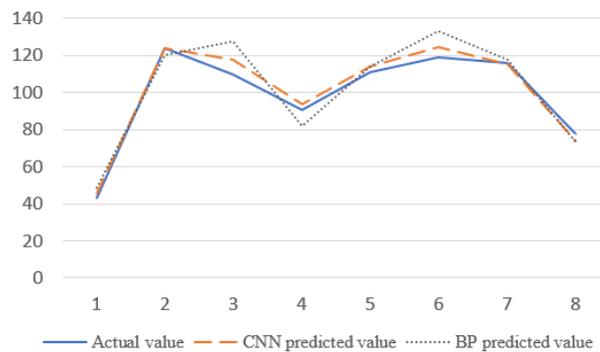


Figure 10. Python results.

There were  $8 \times 2 \times 4 \times 3 \times 3 \times 2 \times 2 \times 2 = 3072$  kinds of resources in the VR system information interface. When the input design feature resource code was 000100001001001001010101010, the cognitive load prediction value was the smallest, 42.74226. The design features are mainly as follows: the task selection area in the interface layout is mainly distributed in the lower right part of the interface, it is convenient for users to select tasks, the shape chamfering in the interface is mainly round chamfering, give users a soft feeling. Cold and light tones are adopted in the overall tone. Moreover, the task selection area has a lightness contrast with the overall tone. This makes it easier for the user to identify the target task. At the same time, when the system is built, it is set to have a transparent interface. So that the surrounding environment scene can be seen through the interface to increase the immersion experience of the user, interface graphics and characters should be properly matched to increase visual expression of the interface. In multi-channel information setting, information of auditory channel and visual channel should be added, such as setting prompt tone, background music and handle vibration, so that users can have natural interaction in VR system. When the input code was: 0000010010001010101010110011001, the value of cognitive load forecast was the largest; 125.55457. The design feature ratio of the two types of data provides reference for designers. Compared with the BP prediction, as shown in Figure 11, the error between the predicted value and the actual value of the

CNN prediction model in this paper is small, the accuracy is slightly higher than that of BP neural network, and the comprehensive performance is better.



**Figure 11.** Comparison of prediction results.

### 5.7. Comparative Analysis of Design Scheme Results

When building VR system interface scenarios, it is very important to reduce the user's cognitive load, considering that VR users with dense information and less experience are prone to depression when performing task operations in virtual reality scenarios. The traditional design process mainly relies on the designers' subjective experience to make decisions and judgments. The design method proposed in this paper enables designers to refer to the priority information of design resources under multi-mode to assist in the design of decision-making when building a system. Secondly, the completed design scheme can be input into the neural network for prediction before it is put into practical application, thus obtaining the user's cognitive load value, which can reduce the time required for users to receive feedback the cognitive experience of the scheme, thus improving the design efficiency and reducing the design cost.

The neural network predicts the situational cognitive load of VR system task selection, and extracts a corresponding better scheme. The design process is a process of cyclical improvement, as the better schemes are further refined and improved. The results of the traditional design method and the optimization design method proposed in this paper are compared to verify the effectiveness of the proposed method (see Table 15 for the comparison chart). For the construction of virtual reality task scenarios with the same design requirements, the resource characteristics corresponding to the traditional designed interface scenario system were 0000001010000110001010010101, the cognitive load predicted by CNN was 85.637 and the task response time was 1.429. The traditional design scheme was then brought into the model of this paper for optimization design, and by doing so it was improved. Firstly, the design feature library of key factors was retrieved through QFD importance ranking for redesign. The corresponding resource design feature of the improved interface scenario system was 0000001010000110010010101010, the cognitive load value was 62.06667, and the task response time was 1.12 s; these values were lower than the design scheme before improvement. Therefore, the effectiveness of the proposed method is verified.

**Table 15.** Comparison table of optimization schemes.

Category	Traditional Design Scheme	Scheme after Model Optimization
Element code	0010000010100010001010100101	000000101000011001001010101010
Virtual reality scene map		
Visual elements	<p>The interface layout is neat and balanced, and the operation area is in the upper right corner of the interface.</p> <p>Graphic area chamfer is characterized by ramp chamfer</p> <p>The main color is cryogenic</p> <p>The contrast between the mission area and the overall tone is lightness contrast.</p> <p>The visual browsing order is from left to right and from top to bottom, with balanced pauses and illustrations.</p> <p>No interface transparency</p>	<p>The interface layout is neat and balanced, and the operation area is in the middle and lower part of the interface.</p> <p>Graphic area chamfer is characterized by ramp chamfer</p> <p>The main color is cryogenic</p> <p>The contrast between the mission area and the overall tone is lightness contrast.</p> <p>The visual browsing order is from left to right and from top to bottom, with balanced pauses and illustrations.</p> <p>Interface transparency</p>
Blueprint Setting of Hearing Elements and Tactile Elements		
Cognitive load	<p>There is a warning tone, no background music, no tactile vibration.</p> <p>Mental needs 5.267</p> <p>Physical demand 6.267</p> <p>Time Demand 6</p> <p>Task Performance 6.267</p> <p>Degree of effort 5.667</p> <p>Frustration 5.4</p> <p>Total Load Value 85.637</p> <p>Task Selection Time 1.429</p>	<p>There are warning tones, background music and tactile vibrations.</p> <p>6.571</p> <p>3.714</p> <p>3.786</p> <p>3.929</p> <p>4.857</p> <p>3.428</p> <p>62.06667</p> <p>1.12</p>

## 6. Conclusions

In view of the large delay and lag of a user’s cognitive load feedback that occurs in the current virtual reality system, there is a belief that the design and setting of the questionnaire may be affected by the subjective prejudice of designers, which can lead to a long VR system construction process, a high cost, and low user satisfaction in the design scheme. This paper introduces VR multi-perception channel mapping to design resource features to establish a user cognitive load assessment and prediction model based on QFD-CNN, which leads to accurate modeling of user perception and timely feedback of interface scene cognitive load data features; this paper implements this through a smart city virtual reality system.

1. The application of cognitive psychology in a VR system is expanded: In this paper, visual, auditory, and tactile perceptual information is integrated into the task scenario research of the VR interface. Guided by cognitive psychology theory, the mapping relationship between the explicit coding of the visual representation of information and the implicit cognition of users under the VR system task selection operation is analyzed, and the user cognitive behavior demand model of virtual reality system is established.
2. The design cycle is shortened and the accuracy of the design scheme is increased: AHP-QFD is used to analyze the relevant importance of the design resource elements in the VR space, and key influencing factors are retrieved to assist designers in system construction. According to the user’s cognitive behavior stratification and its corresponding VR system resource characteristics, the cognitive load of users in VR system interface selection is learned through the nonlinear

expression of variable relationship characteristics of a neural network, which helps achieve the user experience cognitive low-load demand of prediction design, thus reducing the time cost and increasing the accuracy of the designer's scheme.

In future research, the influence of designing resource features on user goal finding and task learning in a VR task context is deeply discussed. VR resource features can be dynamically optimized according to user feedback, and the optimal interval value of receiving resource feature information from each perception channel under the condition of low cognitive load can be determined.

**Author Contributions:** Q.F. conceives research topics and writes manuscripts; J.L. manages the project; Z.Z. and D.Y. set up experiments to sort out the data set. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was supported by the National Natural Science Foundation (51865004) and the Guizhou Provincial Science and Technology Department ([2017] 1046, [2017] 2016, [2018] 1049, [2019] 2010).

**Acknowledgments:** The authors would like to convey our heartfelt gratefulness to the reviewers and the editor for the valuable suggestions and important comments which greatly helped us to improve the presentation of this manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Chiou, R.; Stelter, M.; Rich, A.N. Beyond colour perception: Auditory–visual synaesthesia induces experiences of geometric objects in specific locations. *Cortex* **2013**, *49*, 1750–1763. [[CrossRef](#)] [[PubMed](#)]
2. Zhuoluo, M.A.; Liu, Y.; Zhao, L. Effect of haptic feedback on a virtual lab about friction. *Virtual Real. Intell. Hardw.* **2019**, *1*, 428–434. [[CrossRef](#)]
3. Massiceti, D.; Hicks, S.L.; Rheede, J.J. Stereosonic vision: Exploring visual-to-auditory sensory substitution mappings in an immersive virtual reality navigation paradigm. *PLoS ONE* **2018**, *13*, e199389. [[CrossRef](#)] [[PubMed](#)]
4. Cheng, S.; Sun, S. Human computer interaction resource model based on distributed cognition. *J. Comput. Integr. Manuf. Syst.* **2008**, *14*, 1683–1703.
5. Lu, L.; Tian, F.; Dai, G. Multi channel cognitive and interactive model integrating touch, hearing and vision. *J. Comput. Aided Des. Graph.* **2014**, *26*, 654–661.
6. Li, Y.; Wu, D.; Huang, J.; Tian, F.; Wang, H.; Dai, G.Z. Influence of multi-modality on moving target selection in virtual reality. *Virtual Real. Intell. Hardw.* **2019**, *1*, 303–315.
7. Paquier, M.; Côté, N.D.; Devillers, F.; Koehl, V. Interaction between auditory and visual perceptions on distance estimations in a virtual environment. *Appl. Acoust.* **2016**, *105*, 186–199. [[CrossRef](#)]
8. Lei, X.; Zhang, Ti.; Chen, K.; Zhang, J.; Tian, Y.; Fang, F.; Chen, L.H. Psychophysics of wearable haptic/tactile perception in a multisensory context. *Virtual Real. Intell. Hardw.* **2019**, *1*, 185–200.
9. Geitner, C.; Biondi, F.; Skrypchuk, L.; Jennings, P.; Birrell, S. The comparison of auditory, tactile, and multimodal warnings for the effective communication of unexpected events during an automated driving scenario. *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *65*, 23–33. [[CrossRef](#)]
10. Kathiravan, N.; Devadasan, S.R.; Michael, T.B.; Goyal, S.K. Total quality function deployment in a rubber processing company: A sample application study. *Prod. Plan. Control* **2008**, *19*, 53–66. [[CrossRef](#)]
11. Shi, Y.; Peng, Q. A VR-based user interface for the upper limb rehabilitation. *Procedia CIRP* **2018**, *78*, 115–120. [[CrossRef](#)]
12. Geng, X.; Xu, S.; Ye, C. Module selection method of customer demand driven product service system. *Comput. Integr. Manuf. Syst.* **2016**, *22*, 55–61.
13. Lu, K.; Wei, Z.; Zhuang, D. Physiological model of brain load discrimination and prediction for aircraft cockpit display interface. *J. Beijing Univ. Aeronaut. Astronaut.* **2016**, *42*, 685–693.
14. Akyeampong, J.; Udoka, S.; Caruso, G.; Bordegoni, M. Evaluation of hydraulic excavator Human–Machine Interface concepts using NASA TLX. *Int. J. Ind. Ergon.* **2014**, *44*, 374–382. [[CrossRef](#)]
15. Yan, S.; Tran, C.; Chi, C.; Chen, Y.; Tan, K.; Habiyaemye, J.L. Effect of user interface layout on the operators' mental workload in emergency operating procedures in nuclear power plants. *Nucl. Eng. Des.* **2017**, *322*, 266–276. [[CrossRef](#)]

16. Emami, Z.; Chau, T. The effects of visual distractors on cognitive load in a motor imagery brain-computer interface. *Behav. Brain Res.* **2020**, *378*, 112240.
17. Yan, B.; Zhang, L.; Chu, X. Modeling of user perception evaluation based on convolution neural network. *J. Shanghai Jiaotong Univ.* **2019**, *53*, 844–851.
18. Diego-Mas, J.; Alcaide-Marzal, J. Single users' affective responses models for product form design. *Int. J. Ind. Ergon.* **2016**, *53*, 102–114. [[CrossRef](#)]
19. Liu, W.P.; Nie, J.F.; Jin, Y.; Bai, Y.J. Study on the evaluation method of mental load of armored vehicle crew based on task network model. *J. Mil. Eng.* **2015**, *36*, 1805–1810.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).