Evaluation of Eye Movement Features and Visual Fatigue in Virtual Reality Games

Yuwei Ji
Department of Basic Education, Zhengzhou Urban Construction Vocational College, Zhengzhou-451263, China

***Abstract*—VR games make people happy physically and mentally, but also lead to eye health problems. At present, the existing VR systems lack fatigue detection technology, which makes it difficult to help users use their eyes reasonably. In order to improve the user experience of VR gamers, this paper proposes a visual fatigue detection algorithm based on eye movement features, which uses the relationship between the lateral and longitudinal displacements of the human head and the displacement of the center point of the human eye to locate the position of the human eye. Moreover, in this paper, the human eye position tracking model is input into the three-frame difference algorithm to detect eye movement features. In addition, for tiny motion interference such as eyebrows, the image opening operation of eroding first and then expanding is used to remove it. Through experiments, it is found that the eye movement feature detection method adopted in this paper can greatly improve the detection speed with less accuracy loss, meet the sensitivity requirements of eye movement feature capture, improve the real-time performance of the system, and effectively improve the real-time analysis of player status. Therefore, integrating this algorithm into the virtual game system can help players adjust their own state, which has a positive effect on improving the game experience and reducing eye damage.**

Keywords—Virtual reality; games; eye movement features; visual fatigue

I. INTRODUCTION

In three-dimensional game design, the application value of VR technology cannot be ignored. First of all, VR technology can provide an immersive gaming experience, which is one of its most significant advantages. Through technical means such as headsets and surround sound effects, VR technology can completely immerse players in a virtual environment, making them feel as if they are in the three-dimensional world of the game. This immersion greatly enhances the player's gaming experience, allowing them to get a more realistic feel for everything that's happening in the game. From a technical point of view, VR technology adjusts the game screen in real time by tracking the movement of the player's head, making the player's field of vision and observation angle in the game more natural. At the same time, surround sound technology can make players feel sounds from all directions, further enhancing the realism of the game.

Secondly, immersive experiences bring more possibilities to game design. In traditional game design, there is a certain sense of distance between players and the game world, and the application of VR technology can break this boundary and enable players to integrate more deeply into the game world. This provides a broader creative space for game designers, and can design richer and more complex game scenes and plots, thus enhancing the attractiveness and interest of the game [1].

However, long-term use of VR (more than 30 minutes) may have a certain negative impact on the visual health of the whole body and eyes. The specific manifestations include dizziness, nausea and other symptoms. At the same time, users are accompanied by dry eyes, and even symptoms such as diplopia, tearing, eye pain, eye soreness, and inability to concentrate, as well as related symptoms such as visual asthenopia and video terminal syndrome. The illusion caused by the virtual environment can produce uncomfortable symptoms, such as eye fatigue, dizziness, and other visual fatigue symptoms [2].

There are depth cues in virtual environment scenes, which stimulate eye movements. There is a strong correlation between eye movements and asthenoptic fatigue. Many studies have shown that eye movement behavior can reflect people's thinking movement, and rich information can be obtained from eye tracking movement. Its core purpose is to obtain the gaze point trajectory of human eyes during observation. Combined with knowledge in various fields, we can conduct in-depth analysis of users' visual behavior. Eye trackers are instruments that carry eye tracking technology to track and analyze eye movements, and eye tracking is the core function of eye trackers. From a mental and physical perspective, eye movements are a fundamental reflection of the human state. Meanwhile, eye movements are arguably the most frequent of all human movements, and eye movements are essential to the work of the human visual system. In addition, multiple observations of the eyes are not smooth movements, but multiple eye movement patterns are performed concurrently [3].

Visual fatigue and visual discomfort can be used alternately, but there are still differences between them. Visual fatigue refers to the decline of human visual system performance, which can be measured objectively, and visual discomfort is the subjective response of the visual system. Some researches on eye movement behavior in virtual reality focus on comparing the States before and after use, and some researches also compare the eye movement differences under watching different video content. However, up to now, there has been no research on the differences of visual fatigue and eye movement in different interactive environments of VR. Eye movement behavior is closely related to visual fatigue and the physiological and pathological state of the eye. The changes of eye movement speed, fixation time and blink frequency determine fatigue and mental load.

In response to the design requirements of high speed, small size and non-contact of detection equipment, the face image is preprocessed by image processing technology such as noise reduction; Then the YCbCr color space domain conversion algorithm is used to segment and locate the face and eye region; The blink frequency is counted by frame difference multiple moving object detection algorithm in the face region after positioning; By comparing the real-time detected blink frequency data with the given threshold, it can provide real-time fatigue data reference for VR game users, which is convenient for timely and effective control of game duration and timely and effective protection.

In order to improve the user experience of VR gamers, this paper proposes a visual fatigue detection algorithm based on eye movement features, which uses the relationship between the lateral and longitudinal displacements of the human head and the displacement of the center point of the human eye to locate the position of the human eye. Moreover, in this paper, the human eye position tracking model is input into the three-frame difference algorithm to detect eye movement features. In addition, for tiny motion interference such as eyebrows, the image opening operation of eroding first and then expanding is used to remove it.

II. Related work

1) Research on fatigue: A lot of research has been done on fatigue at home and abroad. In early foreign studies, fatigue was defined as the loss of energy resources after overwork. The resources here are reflected in two aspects. On the one hand, it represents the loss of internal motivation and action from the psychological aspect, and on the other hand, it represents the decrease of external work performance. The study in [4] hold that the decrease of efficiency caused by excessive physical or mental activities is a manifestation of fatigue. Researchers divided fatigue into psychological fatigue and physical fatigue, central fatigue and peripheral fatigue, cognitive fatigue and exercise fatigue, subjective fatigue and objective fatigue, overall fatigue and local fatigue from different dimensions, among which psychological fatigue and physical fatigue are the most widely studied. At present, the most widely studied is to divide fatigue into mental fatigue and physical fatigue. At the same time, it involves psychological state changes in multiple dimensions such as behavioral response, attention, emotion and motivation [5].

2) Visual fatigue: Visual fatigue is one of the types of fatigue. Studies have shown that the characteristics of visual display terminal (VDT) work are directly related to eye discomfort and psychological symptoms. Visual fatigue is defined as subjective symptom syndrome produced when working with eyes, while VDT visual fatigue is more due to eye discomfort and other comprehensive symptoms caused by eyes staring at video terminals for a long time, such as dry eyes, astringent eyes, tingling, eye fatigue, soreness, photophobia and tears, frequent eye movement behaviors, diplopia, blurred vision, heavy eyelids, etc. At the same time, it is also accompanied by headache, dizziness, loss of appetite, memory loss, neck, shoulder, waist, back, joint dysfunction, etc. [6]. Visual fatigue caused by electronic screen operation has become one of the key research fields of human factors engineering since 1970s. The most common fatigue symptom in video display task is eye fatigue, which includes both physiological fatigue and psychological fatigue. The former is manifested as the general symptoms of eye fatigue and fatigue caused by excessive eye use with the extension of working hours, including symptoms that occur after excessive eye function is stressed. It is manifested by low function of central nervous system, a large decrease in flicker fusion frequency, long-term tension of ciliary muscle, significantly low function of eyeball accommodation system, eye fatigue, eye tingling, and temporary decrease of eyesight. The latter, like mental fatigue, is manifested as cognitive fatigue, such as difficulty in maintaining the initial state and continuing to complete the current task, decreased task performance, decreased attention, etc. [7].

3) Eye movement behavior: There are three main types of eye movement behaviors: Saccade, Fixation and SmoothPursuit. The saccades are rapid eye movements that align the fovea with the target. During the experiment, it is important to ensure that the subject does not have saccades while chasing the target smoothly. This eye movement is called catch-up saccades and is more common when catching up at high speeds. Fixation is to keep the foveal visual field on the target for a certain period of time to obtain sufficient visual image details. Smooth pursuit is a kind of fairly slow eye movement that minimizes the movement of retinal targets, and it keeps the eyes fixed on moving objects. The saccade eye movement differs from the smooth pursuit eye movement in that the initial acceleration and deceleration and peak velocity of the former are both higher [8]. A large number of related studies in physiology and psychology have confirmed that some behaviors of human eyes are related to the degree of visual fatigue. Therefore, when collecting eye movement data, it is collected by instruments, and the visual fatigue state of subjects is detected by analyzing the data of eye movement indexes. Through the collation of a large number of references, this paper mainly selects six eye movement parameters, namely, average eye movement duration, average fixation duration, number of eye movement behaviors, number of fixation points, average saccade amplitude and average pupil area, for analysis and research [9].

The average duration of eye movement behavior (unit: ms) refers to the duration of the average single eye movement behavior in the sample. The length of eye movement behavior can also reflect the current level of mental activity and drowsiness, which is closely related to fatigue. By watching different types of videos, study in [10] found that although there was no significant difference in the average duration of eye movement behavior on the whole, it showed an increasing trend with time. The study in [11] studied the continuous viewing of movies, and compared which display can induce visual fatigue more in two kinds of visual display terminals (linear polarization and circular polarization), and found that the average duration of eye movement behavior in the two display terminals increased significantly. The study in [12] explored the effects of different polarized light displays on human visual comfort, and found that the average duration of eye movement behavior increased with the extension of viewing time of subjects who watched the video content displayed by linearly polarized light and the video content displayed by circularly polarized light liquid crystal.

The average fixation duration (unit: ms) refers to the average of the time allocated by subjects at each fixation point in an experimental process, usually in milliseconds. The length of fixation time can reflect the difficulty of information capture and processing, and can indirectly reflect whether the subjects are tired [13]. Generally speaking, the longer the average duration of the fixation point, the deeper the processing degree of the fixation point, and it also reflects the more concentrated the attention of the current subjects. Therefore, the quality of the subjects' fixation ability also reflects the depth of information processing degree and the concentration and distraction of attention to a certain extent. The study in [14] conducted a comparative study on mental fatigue between the elderly and young people, and found that the average fixation duration of both the elderly and young people increase with the development of fatigue, and the average fixation duration of the elderly is significantly higher than that of the young people. The study in [15] divided the subjects into fatigue group and non-fatigue group, and found that the average fixation duration in the fatigue group is significantly lower than that in the fatigue group. Relevant studies have proved that the average fixation duration can indirectly reflect the degree of fatigue of the subjects.

Number of eye movement behaviors (unit: units/min) refers to the number of times the upper and lower eyelids are closed per unit time. Number of eye movement behaviors involves the interaction of efferent nerves between the brain mechanisms responsible for controlling eyelid muscles and other muscle groups, and is closely related to mental activity and visual fatigue. A large number of studies have proved that eye movement behavior is related to visual fatigue, which can be used as an index to evaluate visual fatigue. Generally speaking, the number of eye movement behaviors will increase with visual fatigue. The study in [16] found that the number of eye movement behaviors increases significantly with time by watching different types of videos. The study in [17] studied the continuous viewing of movies, and compared which display can induce visual fatigue more in two kinds of visual display terminals (linear polarization and circular polarization). The results found that the number of eye movement behaviors in the two display terminals increases significantly with time. The study in [18] evaluated the visual fatigue caused by long-term viewing of visual display terminal (VDT) and reading hard-copy, and found that the number of eye movement behaviors under VDT increases significantly from the 2nd hour. The study in [19] studied fatigue driving and found that the number of eye movement behaviors increases with the increase of driving fatigue.

To sum up, visual fatigue includes two meanings: on the one hand, it is eye fatigue caused by some reason, and on the other hand, it refers to psychological fatigue caused by boredom of something and cognitive load.

III. Visual Fatigue Detection based on Eye Movement Features

In terms of limitations, although a lot of research has been done on VR game fatigue, there are still some challenges and limitations. Firstly, the hardware limitation of VR technology is an important factor leading to fatigue. For example, the weight and volume of VR head display equipment are relatively large, and wearing it for a long time can easily lead to discomfort and fatigue. The resolution and scene rendering ability of the device also need to be improved to provide a clearer and more realistic visual experience. Secondly, the rendering amount of VR game images is much higher than that of general games, which requires higher computing power to support, which is a technical bottleneck in the actual development. In addition, the current research on how to effectively alleviate the fatigue of VR games lacks systematic solutions and guidelines.

To sum up, the research on VR game fatigue has made some progress, but it still needs more in-depth research and exploration in hardware technology, image rendering and mitigation strategies. Therefore, this paper attempts to summarize the anti-fatigue system suitable for VR games based on eye movement feature recognition, so as to reduce the impact of game fatigue on players' physical and mental health.

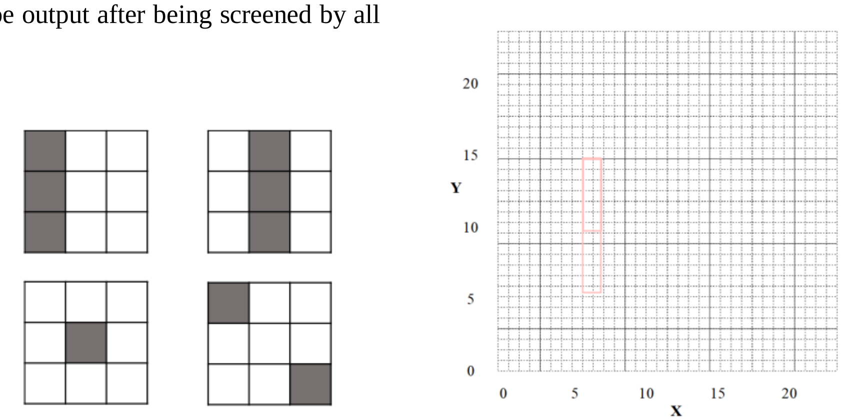
By analyzing the relationship between binocular center point displacement and face displacement, a dynamic video eye position tracking model is established, and then the eye movement behavior rate is detected by frame difference algorithm in the human eye area, and finally the game fatigue judgment is completed.

A. Eye Movement Behavior Detection Algorithm

In this project, the eye movement behavior detection algorithm based on eye features, the eye movement behavior detection algorithm based on background differential moving target detection algorithm and the eye movement behavior detection algorithm based on inter-frame differential moving target detection algorithm are analyzed and studied. The eye movement behavior detection algorithm is comprehensively evaluated from three aspects: effective detection rate, false detection rate and detection speed, and finally the eye movement behavior detection algorithm suitable for this project is selected.

1) Haar-Adaboost human eye coarse positioning: Haar-Adaboost algorithm combines Haar-like features with Adaboost cascade classifier. The core idea of the algorithm is to take the Haar-like sub-window as the input of the weak classifier, and use the window template to traverse each region of the image to calculate the features of the window. It then uses the trained Adaboost cascade classifier to screen the feature. If the feature passes each strong classifier screening in the cascade classifiers, the region is determined to be the human eye. The process of Haar-like sub-window traversing the image is shown in Fig. 1.

Fig. 1 shows a Haar-like linear feature template traversing from bottom to top and left to right in an image. Adaboost iterative algorithm is to train weak classifiers to form a strong classifier with better classification effect. Multiple strong classifiers are arranged from low to high according to their complexity, and the detection results of each level can only be passed to the next level classifier after being screened, and the detection results can only be output after being screened by all strong classifiers. The flow of human eye coarse positioning of Haar-Adaboost algorithm is shown in Fig. 2.



(a) Four feature windows of Haar-like. (b) Traversal process of Haar-like feature window for image.

Fig. 1. Haar-like sub-window and its process of traversing the image.

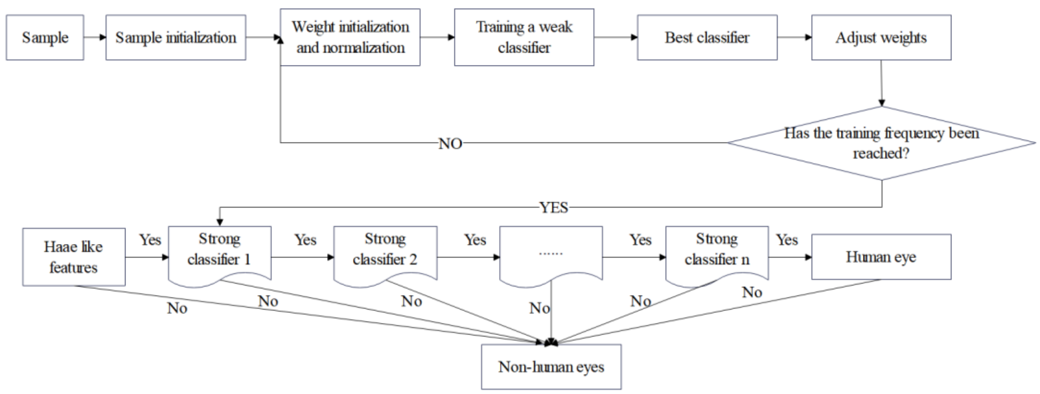


Fig. 2. Human eye coarse positioning process of Haar-Adaboost algorithm.

2) ERT human eye fine positioning and eye movement behavior detection: On the basis of coarse localization of human eye area, ERT algorithm is used to accurately segment human eye area and judge eye movement behavior. The ERT algorithm needs to establish a GBDT (Gradient Boosting Decision Tree), which is an object detection method based on the idea of face alignment. Human eye detection using face feature point matching is shown in Fig. 3 [20].

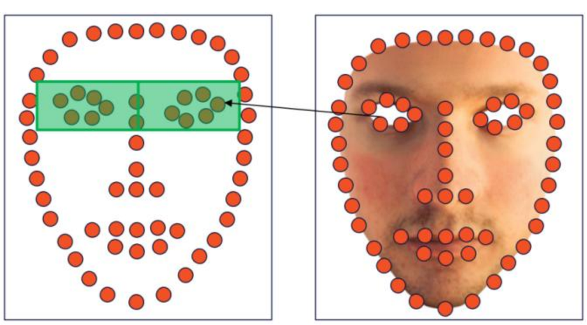


Fig. 3. Accurate positioning of human eye area based on face feature points.

The background difference algorithm can detect moving objects only by comparing the gray value of the current frame image with the standard background gray value. The algorithm is simple and easy to implement, and it is widely used in the field of video surveillance.

The video frame image at time t is $F(x,y,t)$, the standard background image is $G(x,y,t)$, the binarized image is $B(x,y,t)$, the dynamic judgment threshold is T, and the binarized gray difference image is $D(x,y,t)$, which can be expressed as Eq. (1).

$$D(x,y,t)=\begin{cases} 1, \left|F(x,y,t)-G(x,y,t)\right|>T \\ 0, Othes \end{cases} \tag{1}$$

After the image is processed by Eq. (1), the part with gray value of 1 is the moving target, and the portion with a gray value of 0 is the background of the image. The limitation of background difference algorithm is that it needs to set the standard background in advance, and it requires high stability of the background. Because of the dynamic changes of the background, it is difficult to selectively reconstruct the background. In addition, when the image environment is shaken, the illumination changes suddenly, and the like, it is also considered that a moving target appears after background difference.

Using inter-frame difference algorithm to detect moving objects in face area can achieve the purpose of detecting eye movement behavior. Similar to the principle of background difference algorithm, it assumes that the image function of the current video frame is $F(x,y)$, the image function of the first frame of the current image is $F_{k-\tau}(x,y)$, and the gray difference function $D_k(x,y)$ of the two images can be expressed by Eq. (2) [21]:

$$D_k(x,y) = \left|F(x,y) - F_{k-\tau}(x,y)\right| \tag{2}$$

The target region gray function in (2) is binarized, as shown in Eq. (3).

$$B_{Dk}(x,y) = \begin{cases} 1, D_k(x,y) > T \\ 0, Others \end{cases} \tag{3}$$

Three-frame difference method obtains the same part of two difference images by AND operation of two difference images, which avoids the image hole phenomenon of adjacent difference algorithm and enhances the robustness of the algorithm. When the three-frame difference method is used, Eq. (3) becomes Eq. (4).

$$\begin{cases} D_k(x,y) = \left|F_k(x,y) - F_{k-1}(x,y)\right| \\ D_{k-1}(x,y) = \left|F_{k-1}(x,y) - F_{k-2}(x,y)\right| \end{cases} \tag{4}$$

In Eq. (4), $D_k(x,y)$ is the expression of the difference between the k-th frame image and the $k-1$-th frame image, and $D_{k-1}(x,y)$ is the expression of the difference between the k-th frame image and the $k-2$-th frame image. By bringing Eq. (4) into Eq. (3), the binarization, Eq. (5) of the two difference images is obtained.

$$\begin{cases} B_{D_k}(x,y) = \begin{cases} 1, D_k(x,y) > T \\ 0, Others \end{cases} \\ B_{D_{k-1}}(x,y) = \begin{cases} 1, D_{k-1}(x,y) > T \\ 0, Others \end{cases} \end{cases} \tag{5}$$

In Eq. (5), T is the moving target judgment threshold. When $B_{D_k}(x,y)$ and $B_{D_{k-1}}(x,y)$ are equal to 1, it indicates that there are moving targets in two adjacent frames. Because the movement of the moving target is continuous and continuous, the calculation results of two consecutive differences will be 1. If the false detection is caused by noise such as holes, it will be eliminated by AND operation. The target object to be detected obtained by doing and operation in Eq. (5) is shown in Eq. (6).

$$A(x,y) = B_{D_k}(x,y) \& B_{D_{k-1}}(x,y) \tag{6}$$

In Eq. (6), $\&$ represents the AND operation, and $A(x,y)$ is the detection objective function. When the value of $A(x,y)$ is 1, it indicates that there is a moving target in the measured picture, otherwise it is stationary.

The frame difference method has a good effect in detecting eye movement behavior, but it can't judge the non-eye movement area. The three-frame difference method will still be affected by small moving objects such as hair and clothes pleats. Therefore, it is necessary to improve the three-frame difference method to improve the detection accuracy.

B. *Three-Frame Difference Eye Movement Behavior Detection Algorithm*

In this section, an eye movement behavior detection algorithm based on the combination of binocular position tracking in face region and three-frame difference method is proposed. The improved eye movement behavior detection algorithm is shown in Fig. 4.

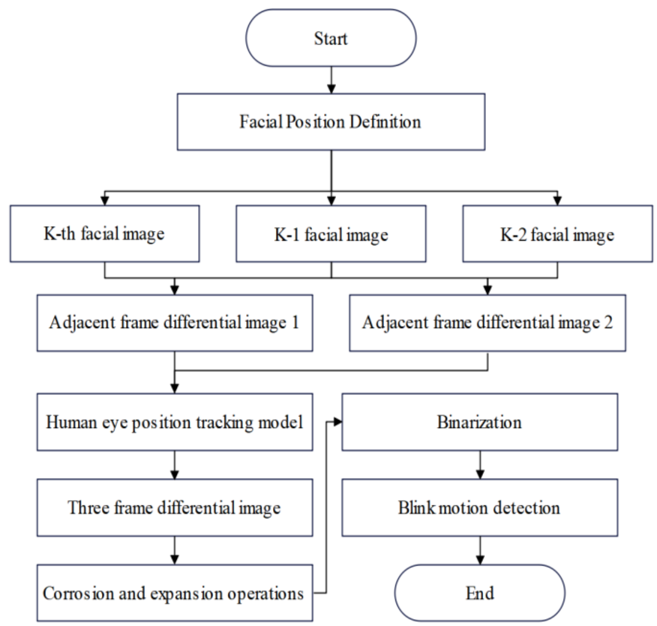


Fig. 4. Flowchart of improved algorithm.

As shown in Fig. 4, the optimization of the three-frame difference method in this project mainly includes the following points.

1) Definition of face area: In the face detection algorithm, the coordinate vertices of the face area are successfully found and input into the three-frame difference algorithm, and only the areas within the coordinates are differentiated. The gray value of the non-face areas outside the coordinates is 0, so that the interference of moving targets outside the face area on eye movement behavior detection is eliminated, as shown in Fig. 5.

2) Human eye position tracking: Before obtaining the three-frame difference map, the human eye position tracking model is introduced to estimate the position of both eyes, and the difference operation is only carried out in the human eye area, and the rest areas are set to 0, so as to eliminate the interference of other moving targets in the face area on eye movement behavior detection.

3) Corrosion expansion: The image obtained by further processing the corrosion expansion operation is used to remove sharp noise around the eyes and eliminate motion noise such as eyebrows that may affect the detection results.

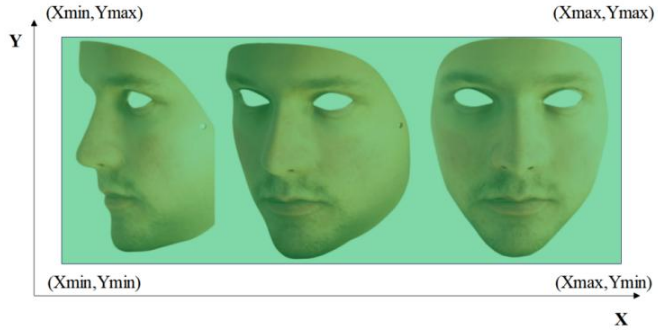


Fig. 5. Face area coordinate vertices.

All the gray values of pixels in areas other than the face are set to 0, and the coordinate level values (X_{min}, y_{min}), (X_{min}, y_{max}), (X_{max}, y_{min}), and (X_{max}, y_{max}) selected by the face frame are found. When the adjacent difference operation is performed, only the difference operation is needed on the targets falling within the coordinate level value range, as shown in Eq. (7).

$$\bar{F}_k(x,y) = \begin{cases} F_k(x,y) \& M_k(X_i, Y_i), \text{Human face area} \\ 0, \text{Others} \end{cases} \quad (7)$$

In Eq. (7), $\bar{F}_k(x,y)$ represents the first image used as the adjacent difference operation, $M_k(X_i, Y_i)$ represents the detected face region function, $F_k(x,y)$ represents the first frame image, $\&$ represents an AND operation, which eliminates other regions other than the face by the AND operation of the image to be detected with the face position, and X_i and Y_i represents the face region coordinate level value:

$$\begin{cases} X_i = X_{min}, X_{max} \\ Y_i = Y_{min}, Y_{max} \end{cases} \quad (8)$$

After introducing the face region coordinate level values, the following form is obtained.

$$\begin{cases} D_k(x,y) = \left|\bar{F}_k(x,y) - \bar{F}_{k-1}(x,y)\right| \\ D_{k-1}(x,y) = \left|\bar{F}_{k-1}(x,y) - \bar{F}_{k-2}(x,y)\right| \end{cases} \quad (9)$$

After the operation of Eq. (8) and Eq. (9), the difference operation of two adjacent frames is only performed in a specific area containing the face, which can avoid bringing moving targets other than the face into the three-frame difference algorithm, and eliminate the interference of moving targets other than the face to eye movement behavior detection.

A human eye position estimation algorithm under head movement is proposed. The algorithm ideas are divided into the following points.

1) A head surveillance video of a player playing a game is selected, and the video is decomposed into 100 frames of images.

2) The improved YCbCr algorithm is used to extract the player's face image, and the lateral displacement Δx_k^h and Δy_k^h of the player's head in two consecutive frames is calculated by combining the adjacent frame difference method, and a total of 99 groups of displacement data are extracted, as shown in Eq. (10).

$$\begin{cases} \Delta x_k^h = \left|x_k^h - x_{k-1}^h\right| \\ \Delta y_k^h = \left|y_k^h - y_{k-1}^h\right| \end{cases} \quad (10)$$

3) The minimum area of the human eye position in each frame image is manually framed, the center position of the box is found, the displacement of the center position of the box is used to represent the displacement of both eyes, and the sum of the transverse and longitudinal displacement differences of the center position of the box is calculated by using the method of step 2, as shown in Eq. (11).

$$\begin{cases} \Delta x_k^e = \left|x_k^e - x_{k-1}^e\right| \\ \Delta y_k^h = \left|y_k^e - y_{k-1}^e\right| \end{cases} \quad (11)$$

4) Two groups of data of transverse displacement of human head and transverse displacement of binocular center point, longitudinal displacement of human head and longitudinal displacement of binocular center point are extracted respectively, and the functional relationship fitting is carried out to find the relationship expression of the two groups of data. The calculation of the horizontal and longitudinal displacement of the face and the horizontal and longitudinal displacement of the center of the human eye in two consecutive images is shown in Fig. 6.

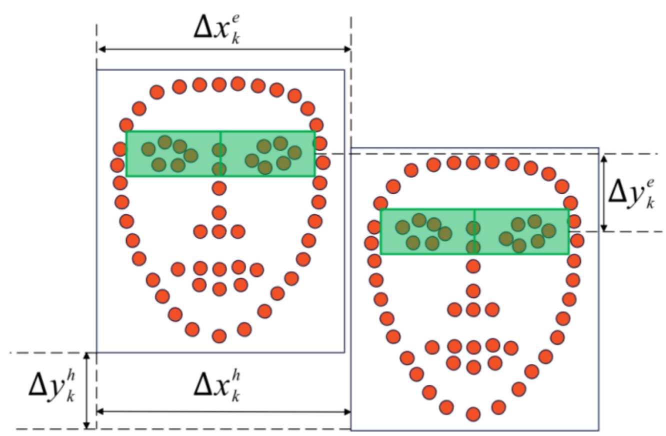
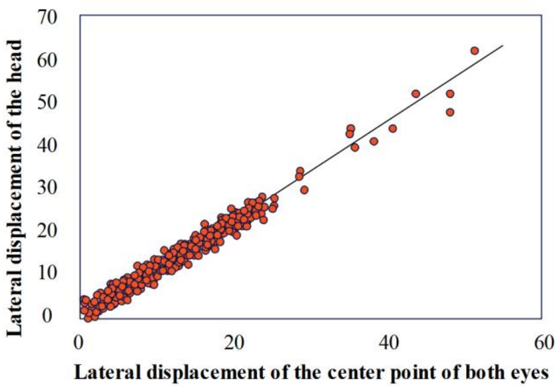
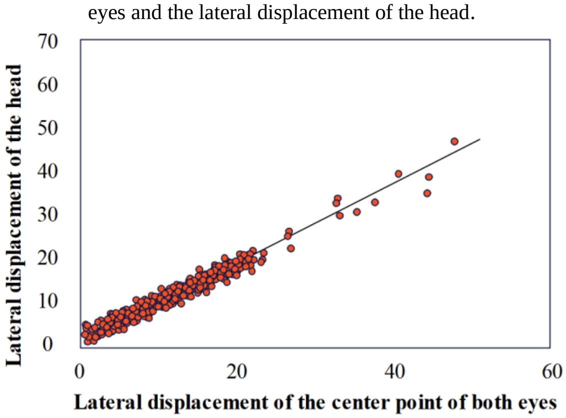


Fig. 6. Head displacement and binocular center point displacement.

The 99 groups data of Δx_k^h, Δx_k^e, Δy_k^h and Δy_k^e are counted and data relationship fitting is performed, and the results are shown in Fig. 7. Fig. 7 (a) is Relationship between the lateral displacement of the center point of both eyes and the lateral displacement of the head, Fig. 7 (b) is Relationship between longitudinal displacement of center point of binocular eyes and longitudinal displacement of head.



(a) Relationship between the lateral displacement of the center point of both eyes and the lateral displacement of the head.



(b) Relationship between longitudinal displacement of center point of binocular eyes and longitudinal displacement of head

Fig. 7. Relationship between human eye center displacement and head displacement.

As can be seen from Fig. 6, the distribution of the two sets of data shows a strong linear relationship, and the relationship between Δx_k^h and Δx_k^e and Δy_k^h and Δy_k^e is shown in Eq. (12) after data fitting.

$$\begin{cases} \Delta x^e = 1.22\Delta x^h - 1.33 \\ \Delta y^e = 0.94\Delta x^h + 0.15 \end{cases} \tag{12}$$

In Eq. (12), Δx^e represents the lateral displacement of the eye to be measured, Δy^e represents the longitudinal displacement of the eye to be measured, Δx^h represents the lateral displacement of the head of two adjacent images calculated by the adjacent frame difference method, Δy^h represents the longitudinal displacement of the head of two adjacent images calculated by the adjacent frame difference method. Through Eq. (12), the detection range is further compressed to the human eye area, the motion noise in the skin color area is filtered, and the detection accuracy is improved.

IV. System Construction and Experiment

A. System Construction

Based on the MCIA architecture designed in this paper, the sharing framework of eye tracking data is shown in Fig. 8.

In this paper, the eye tracking and gesture data of users in the scene are captured. After that, this paper builds a collaborative scene to transmit and visualize the data of both users in the scene, thus realizing the sharing of users' eye movements and gestures in VR scenes and providing data support for visual fatigue analysis in VR games.

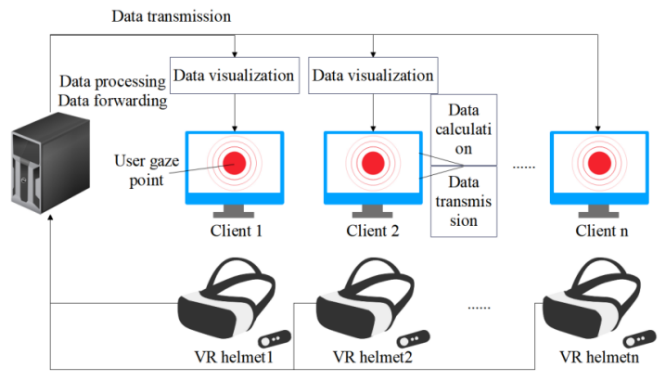


Fig. 8. Eye tracking data sharing framework diagram.

In order to verify whether the algorithm model designed in this paper is suitable for visual fatigue analysis in VR games, and to verify the influence of the best eye-hand visualization mode on the communication degree of both parties, this paper designs a VR game visual fatigue analysis system as shown in Fig. 9. The system is mainly divided into two modules: eye-hand data processing module and virtual scene interaction module. The eye-hand data processing module mainly provides data support for the virtual scene interaction module. The virtual scene interaction module will visually display user data and give users visual feedback. After receiving the visual feedback, the user's behavior is corrected again, so that the ability of collaborative interaction between the two sides is trained through the real-time interaction and real-time feedback of the system.

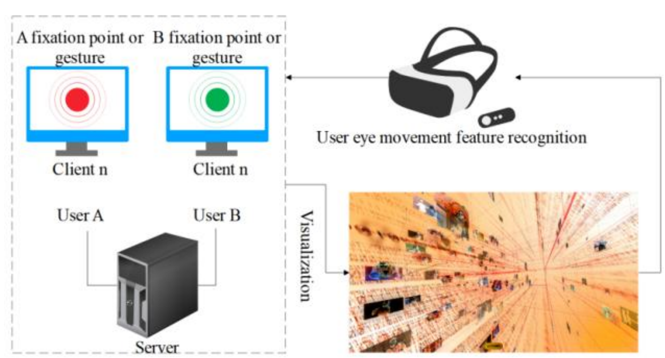


Fig. 9. System structure.

The data set in this article is a self-built data set, which is tested by 40 volunteers. These 40 volunteers are all VR game enthusiasts, so they meet the experimental verification needs of visual fatigue in this article. This paper conducts experiments through popular games, and selects two popular plot games to conduct experiments.

B. Results

In this project, the head images of 40 volunteers in VR game state are taken, and the video is decomposed into 10,000 frames. The number of eye movement behaviors is counted by using the ERT algorithm based on Haar-Adaboost feature, the unimproved three-frame difference method and the improved three-frame difference method based on binocular position tracking in face area. The results are shown in Table I and Fig. 10 and Fig. 11.

Since this project judges whether there is visual fatigue based on the eye movement behavior rate of VR games, it is particularly important to obtain the frequency data of eye movement behavior under normal circumstances. Through the VR game test of the respondents in this paper, the frequency of eye movement behavior is artificially counted to explore the relationship between the frequency of eye movement behavior and fatigue. The results are shown in Table II.

TABLE I. COMPARISON OF STATISTICAL EFFECTS OF DIFFERENT ALGORITHMS ON EYE MOVEMENT BEHAVIOR

Algorithms	**Game 1**			**Game 2**		
	Detection rate (%)	**False detection rate (%)**		**Detection rate (%)**	**False detection rate (%)**	
Human eye feature ERT algorithm	94.45	4.26	Human eye feature ERT algorithm	94.45	4.26	Human eye feature ERT algorithm
Traditional three frame difference method	85.93	20.69	Traditional three frame difference method	85.93	20.69	Traditional three frame difference method
Improve the three frame difference method	91.18	8.61	Improve the three frame difference method	91.18	8.61	Improve the three frame difference method

TABLE II. CORRESPONDENCE BETWEEN EYE MOVEMENT BEHAVIOR RATE AND FATIGUE STATE (+ INDICATES NORMAL STATE, − INDICATES FATIGUE STATE)

Serial Number	**Number of Times**	**Frequency (Times/Min)**	**State**	**Serial Number**	**Number of Times**	**Frequency (Times/Min)**	**State**
1	53	18	+	21	69	24	-
2	43	15	+	22	30	11	-
3	41	14	+	23	16	6	-
4	48	17	+	24	57	27	-
5	43	15	+	25	22	22	-
6	48	17	+	26	27	27	-
7	34	12	+	27	25	7	-
8	50	17	+	28	62	21	-
9	39	14	+	29	28	10	-
10	49	17	+	30	52	18	-
11	33	12	+	31	21	8	-
12	40	14	+	32	60	31	-
13	42	15	+	33	58	20	-
14	53	18	+	34	27	10	-
15	30	11	+	35	53	18	-
16	37	13	+	36	46	16	-
17	41	14	+	37	35	12	-
18	57	20	+	38	28	10	-
19	29	10	+	39	61	21	-
20	31	11	+	40	19	7	-

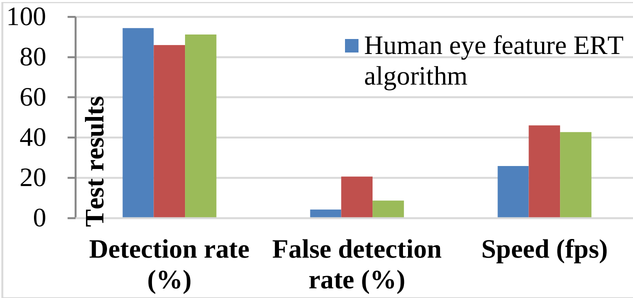


Fig. 10. Comparison of detection effects of different algorithms on eye movement behavior (Game 1).

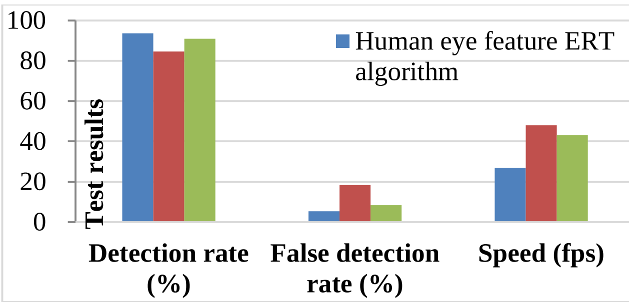


Fig. 11. Comparison of detection effects of different algorithms on eye movement behavior (Game 2).

In order to verify the influence of players' body movements on the detection results of eye movement features in VR games, this paper takes racing games as an example to analyze, and designs the tilt experiment of VR racing games passing through speed bumps. When the vehicle speed is 10 km/h, the tires on one side of the car body pass through the speed bump, and the tester's head also shakes violently. The experimental results obtained by processing are shown in Fig. 12. Fig. 12 (a) shows the moving directions of the right and left eyes in the X axis, Fig. 12 (b) shows the moving directions of the right and left eyes in the Y axis, and Fig. 12 (c) shows the moving directions of the right and left eyes in the Z axis.

In the Y direction, the black line and the red line obviously cross, indicating that the movement states of the left eye and the right eye are different at this time. Furthermore, the movements of the right and left eyes are also different in the X and Z directions. The root cause of these differences is that the right and left eyes move in opposite directions on the same axis (namely, X, Y, and Z axes). For example, the difference between the right and left eye on the Y axis indicates that the left eye moves down and the right eye moves up, or the left eye moves up and the right eye moves down. Therefore, by detecting the motion states of the right eye and the left eye, it is possible to determine whether there is a lateral tilting motion of the human body.

In order to further verify the accuracy of this model in VR game fatigue detection, this model is compared with study [4] (visual tracking), study [10] (EEG), study [16] (brain computer interface), and the accuracy of the above methods in VR game fatigue detection is verified through comparative tests, A total of six groups of tests were conducted, and the comparison results in Table III are obtained.

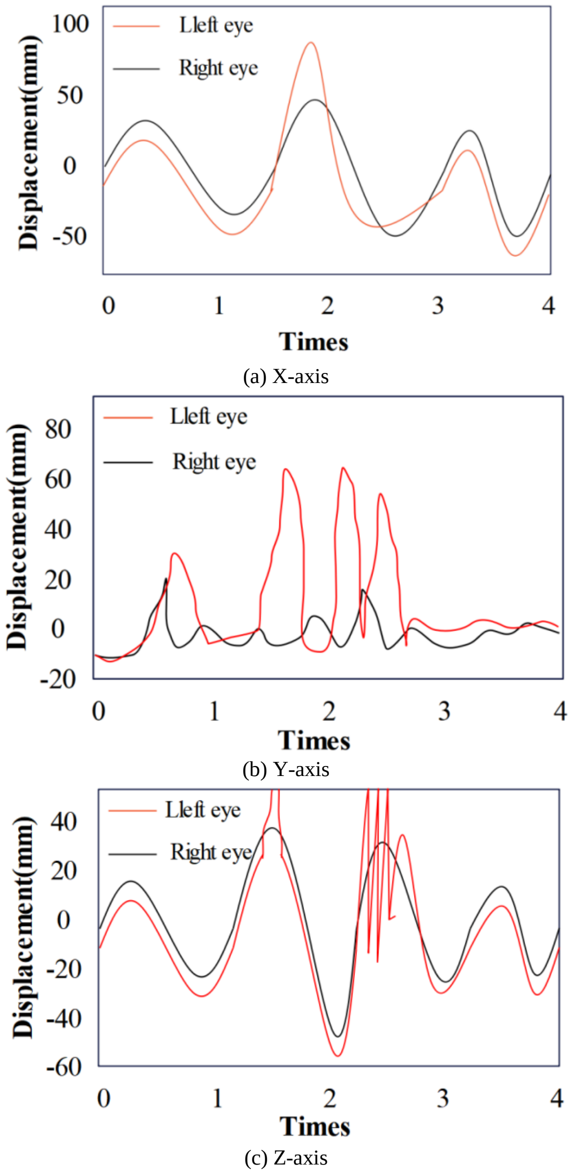


Fig. 12. Experimental results of roll motion in VR game.

TABLE III. COMPARISON OF VR GAME FATIGUE DETECTION ACCURACY

	Visual tracking	EEG	Brain computer interface	This study
1	69.18	77.10	80.82	91.09
2	70.46	75.07	79.23	91.18
3	72.39	76.74	78.95	88.95
4	68.84	76.10	82.61	88.37
5	72.64	82.08	79.78	90.55
6	70.80	75.16	80.67	91.83

C. Analysis and Discussion

For humans, the position of the human eye on the head is relatively fixed, and the area other than the head needs to be removed. Therefore, the area of the human eye can be located only by estimating the position of the human eye on the head. Based on the fact that the human eye is fixed at the position of the face, when the player's head remains stationary, it is enough to detect moving targets in more than two-thirds of the skin color area of the face. However, in VR ordered scenes, the player's head is often in a constant state. In a state of shaking, finding the approximate position of the human eye when the head is shaking is the difficulty of algorithm improvement. Through observation, it can be found that when the player's head shakes, the movement trajectory of the center point of both eyes and the movement trajectory of the head show a certain regularity, and the relationship between the movement trajectory of the center point of both eyes and the displacement of the head can be located for the human eye only by finding the relationship between the movement trajectory of the center point of both eyes and the movement trajectory of the head.

The eye movement behavior detection algorithm based on human eye features needs to locate the human eye first, and then judge whether there is eye movement behavior according to the coordinate changes of human eye feature points. In the process of cascade classifier training and human eye feature matching, a large number of image samples need to be trained, and to traverse the whole detected image, the algorithm complexity is high, and it is extremely difficult to implement in FPGA development board. After observing the player's face in the game state, it is found that in the normal game state, the player's face is basically stationary, and only the eyes and mouth will appear tiny movements. Therefore, the detection of eye movement behavior can be completed only by detecting whether there is moving target in the face range, and no longer need to waste a lot of computing power for data sample training and feature point acquisition. At present, most algorithms have realized the location of face area. Therefore, the background difference algorithm is considered to detect the moving target of human face and then judge whether there is eye movement behavior.

It can be seen from Table I and Fig. 10 and Fig. 11 that the ERT algorithm based on Haar-Adaboost human eye features has the highest detection rate and the lowest false detection rate for human eye movement behavior, but the processing speed of the algorithm is the slowest among the three algorithms, and the algorithm is complex, which is difficult to implement in FPGA development board. The traditional three-frame difference algorithm has the fastest detection speed, but its detection is the lowest among the three algorithms, and the false detection rate reaches about 20%. The detection rate of eye movement behavior of the improved three-frame difference algorithm is about 91%, the average detection rate is nearly 7% higher than that of the traditional three-frame difference method, and the average false detection rate is 56% lower than that of the traditional three-frame difference method. Because the algorithm introduces the human eye position estimation model, the processing speed is lower than that of the traditional three-frame difference method, but the processing frame rate still reaches 42 fps, which meets the performance requirements of visual fatigue detection of VR games.

It is pointed out that under normal circumstances, the frequency of players' eye movement behavior is between 12-17, and with the increase of game duration and fatigue, the highest rate of eye movement behavior reaches 40 times per minute. As can be seen from Table II, under normal conditions, the frequency of normal people's eye movement behavior is mainly distributed between 11 and 19. Considering that VR gamers are more concentrated in the game situation, and the rate of eye movement behavior is relatively low, this project will set the threshold of vision through attention to 11-19. Combined with the actual experimental data, it is found that setting the threshold of fatigue state to 11-19 times per minute has the greatest correlation with the fatigue state of VR gamers. When the eye movement behavior fatigue is 5 to 10 times per minute or more than 18 times per minute, it is determined that the VR game player has abnormal eye movement behavior phenomenon and visual fatigue.

It can be seen from Fig. 12 that the experimental results in this paper are consistent with the actual situation, which further verifies that the proposed stereo vision measurement method can be used for tracking measurement of eye movement and verifies the robustness of the algorithm under real conditions.

In Table III, visual tracking has the lowest accuracy, and the highest accuracy is only 72.64%. The recognition results of EEG and brain computer interface are similar, and the highest recognition accuracy can reach 82%. The accuracy of this research model in VR game fatigue detection can reach more than 88%, and the highest can reach 91.83%.

Through comparative analysis, it is verified that this research model has excellent performance in VR game fatigue detection

Taken together, the eye movement feature method proposed in this paper has a good effect in visual fatigue monitoring of VR games and can effectively improve the real-time analysis of player status. Therefore, integrating this algorithm into the virtual game system can help players adjust their own state, which has a positive effect on improving the game experience and reducing eye damage.

There are significant differences in the fatigue state of different individuals while playing games.

This difference is mainly reflected in individuals' reactions to the side effects of VR games. For example, some people may experience severe dizziness, nausea, eye fatigue, and overall fatigue after playing high-intensity VR games, while others may have mild or almost no symptoms. This difference is influenced by various factors, including individual sensory adaptation, age, gaming experience, and duration of continuous use of VR devices.

Specifically, an individual's sensory adaptation plays a crucial role in understanding the severity of VR vertigo. Some people are more likely to experience subjective visual vertical line changes after exposure to VR, especially under high intensity, which may be related to their milder VR dizziness symptoms. On the contrary, those who suffer from the most severe VR vertigo are unlikely to change the way they perceive vertical lines.

In addition, the study also found that women are more likely to experience screen sickness when using VR than men, which may be consistent with statistical data showing that women are also more prone to motion and screen sickness in other environments. This gender difference is particularly important in the widespread application of VR technology, as it may affect the acceptance and user experience of VR technology among users of different genders.

In addition to gender and sensory adaptation, an individual's gaming experience can also affect their fatigue state. Novice players may feel more exhausted and uncomfortable due to unfamiliarity with the VR environment, while experienced players may be better able to adapt and reduce fatigue. Therefore, for different individuals, the fatigue state while playing games is a complex and variable problem, influenced by multiple factors.

The system has a good statistical effect on capturing eye movement behavior and eye movement behavior rate. However, although there is a great correlation between eye movement behavior rate and fatigue degree, it is difficult to accurately judge the mental state of VR game players only through this single index. In addition, due to factors such as living habits, there are differences in the frequency of eye movement behavior among individuals. Therefore, we should further eliminate individual differences through experiments, and make a comprehensive evaluation of fatigue state in combination with factors such as head posture and mouth characteristics.

V. Conclusion

In order to improve users' immersion and experience in VR game environment, this paper proposes a visual skin fatigue recognition algorithm based on eye movement tracking, which uses the relationship between the lateral displacement and longitudinal displacement of the human head and the displacement of the center point of the human eye to locate the position of the human eye, and inputs the human eye position tracking model into the three-frame difference algorithm to detect eye movement behavior. In addition, for tiny motion interference such as eyebrows, this paper adopts the image open operation of eroding first and then expanding to remove it. The eye movement behavior detection method adopted in this paper greatly improves the detection speed, meets the sensitivity requirements of eye movement behavior capture, and improves the real-time performance of the system with less accuracy loss. Moreover, the correlation between eye movement behavior frequency and fatigue is based on relevant reference and actual experiments, and the data are reliable.

The system has a good statistical effect on capturing eye movement behavior and eye movement behavior rate. However, although there is a great correlation between eye movement behavior rate and fatigue degree, it is difficult to accurately judge the mental state of VR game players only through this single index. In addition, due to factors such as living habits, there are differences in the frequency of eye movement behavior among individuals. Therefore, we should further eliminate individual differences through experiments, and make a comprehensive evaluation of fatigue state in combination with factors such as head posture and mouth characteristics.

References

[1] Merzon, L., Pettersson, K., Aronen, E. T., Huhdanpää, H., Seesjärvi, E., Henriksson, L., MacInnes, W. J., Mannerkoski, M., Macaluso, E., & Salmi, J., “Eye movement behavior in a real-world virtual reality task reveals ADHD in children,” Scientific reports, vol. 12, no. 1, pp. 20308, 2022.

[2] Nam, Y., Hong, U., Chung, H., & Noh, S. R., “Eye movement patterns reflecting cybersickness: evidence from different experience modes of a virtual reality game,” Cyberpsychology, Behavior, and Social Networking, vol. 25, no. 2, pp. 135-139, 2022.

[3] Pastel, S., Marlok, J., Bandow, N., & Witte, K., “Application of eye-tracking systems integrated into immersive virtual reality and possible transfer to the sports sector-A systematic review,” Multimedia Tools and Applications, vol. 82, no. 3, pp. 4181-4208, 2023.

[4] Souchet, A. D., Philippe, S., Lourdeaux, D., & Leroy, L., “Measuring visual fatigue and cognitive load via eye tracking while learning with virtual reality head-mounted displays: A review,” International Journal of Human–Computer Interaction, vol. 38, no. 9, pp. 801-824, 2022.

[5] Rappa, N. A., Ledger, S., Teo, T., Wai Wong, K., Power, B., & Hilliard, B., “The use of eye tracking technology to explore learning and performance within virtual reality and mixed reality settings: a scoping review,” Interactive Learning Environments, vol. 30, no. 7, pp. 1338-1350, 2022.

[6] Kim, J., Jang, H., Kim, D., & Lee, J., “Exploration of the virtual reality teleportation methods using hand-tracking, eye-tracking, and eeg,” International Journal of Human–Computer Interaction, vol. 39, no. 20, pp. 4112-4125, 2023.

[7] Stoeve, M., Wirth, M., Farlock, R., Antunovic, A., Müller, V., & Eskofier, B. M., “Eye tracking-based stress classification of athletes in virtual reality,” Proceedings of the ACM on Computer Graphics and Interactive Techniques, vol. 5, no. 2, pp. 1-17, 2022.

[8] Alcañiz, M., Chicchi-Giglioli, I. A., Carrasco-Ribelles, L. A., Marín-Morales, J., Minissi, M. E., Teruel-García, G., Sirera, M., & Abad, L., “Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis,” Autism Research, vol. 15, no. 1, pp. 131-145, 2022.

[9] Mitre-Ortiz, A., Muñoz-Arteaga, J., & Cardona-Reyes, H., “Developing a model to evaluate and improve user experience with hand motions in virtual reality environments,” Universal Access in the Information Society, vol. 22, no. 3, pp. 825-839, 2023.

[10] Baceviciute, S., Lucas, G., Terkildsen, T., & Makransky, G., “Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study,” Journal of Computer Assisted Learning, vol. 38, no. 1, pp. 120-136, 2022.

[11] Dong, J., Ota, K., & Dong, M., “Why vr games sickness? an empirical study of capturing and analyzing vr games head movement dataset,” IEEE MultiMedia, vol. 29, no. 2, pp. 74-82, 2022.

[12] Banstola, S., Hanna, K., & O’Connor, A., “Changes to visual parameters following virtual reality gameplay,” The British and Irish Orthoptic Journal, vol. 18, no. 1, pp. 57, 2022.

[13] Ugwitz, P., Kvarda, O., Juříková, Z., Šašinka, Č., & Tamm, S., “Eye-tracking in interactive virtual environments: implementation and evaluation,” Applied Sciences, vol. 12, no. 3, pp. 1027, 2022.

[14] Lin, Y., Gu, Y., Xu, Y., Hou, S., Ding, R., & Ni, S., “Autistic spectrum traits detection and early screening: A machine learning based eye movement study,” Journal of Child and Adolescent Psychiatric Nursing, vol. 35, no. 1, pp. 83-92, 2022.

[15] Huygelier, H., Schraepen, B., Lafosse, C., Vaes, N., Schillebeeckx, F., Michiels, K., Note, E., Vanden Abeele, V., van Ee, R., & Gillebert, C. R., “An immersive virtual reality game to train spatial attention orientation after stroke: A feasibility study,” Applied Neuropsychology: Adult, vol. 29, no. 5, pp. 915-935, 2022.

[16] Hadjiaros, M., Neokleous, K., Shimi, A., Avraamides, M. N., & Pattichis, C. S., “Virtual reality cognitive gaming based on brain computer interfacing: A narrative review,” IEEE Access, vol. 11, pp. 18399-18416, 2023.

[17] Lu, A. S., Pelarski, V., Alon, D., Baran, A., McGarrity, E., Swaminathan, N., & Sousa, C. V., “The effect of narrative element incorporation on physical activity and game experience in active and sedentary virtual reality games,” Virtual Reality, vol. 27, no. 3, pp. 1607-1622, 2023.

[18] Zhou, Y., Feng, T., Shuai, S., Li, X., Sun, L., & Duh, H. B. L., “EDVAM: a 3D eye-tracking dataset for visual attention modeling in a virtual museum,” Frontiers of Information Technology & Electronic Engineering, vol. 23, no. 1, pp. 101-112, 2023.

[19] Jiménez-Rodríguez, C., Yélamos-Capel, L., Salvestrini, P., Pérez-Fernández, C., Sánchez-Santed, F., & Nieto-Escámez, F., "Rehabilitation of visual functions in adult amblyopic patients with a virtual reality videogame: a case series," Virtual Reality, vol. 27, no. 1, pp. 385-396, 2023.

[20] Luong, T., & Holz, C., "Characterizing physiological responses to fear, frustration, and insight in virtual reality," IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 11, pp. 3917-3927, 2022.

[21] Kristjánsson, T., Draschkow, D., Pálsson, Á., Haraldsson, D., Jónsson, P. Ö., & Kristjánsson, Á., "Moving foraging into three dimensions: Feature- versus conjunction-based foraging in virtual reality," Quarterly Journal of Experimental Psychology, vol. 75, no. 2, pp. 313-327, 2022.