Copyright ⓒ 2022 KSII

Research of 3D image processing of VR technology in medicine based on DNN

Gong ZhaoZhe[1*], Li XiaoDong[1], Shi XiaoYing[1], Liu Geng[1] and Chen Bin[1]
[1]Hangzhou DianziUniversity
[email :gongzhaozhe@hdu.edu.cn]
[*]Corresponding author: Gong ZhaoZhe

Received September 13, 2021; revised Janaury 1, 2022; revised February 23, 2022; accepted May 8, 2022; published May 31, 2022

Abstract

According to a survey published in an authoritative journal in January 2020, the globalincidence rate of mental illness is 8.3% for men and 10.6% for women, which indicates thatmental illness has become a globally recognized obstacle. Therefore, specific psychotherapy including mental illness will become an important research topic. It is very effective forpatients with special mental diseases, such as mental illness, to reduce their mental reaction byexposure therapy; the system uses the virtual reality system of medical images processed by learningalgorithm to reproduce the effect of virtual reality exposure method of the high scene of transparent ladder. Compared with the old invasive exposure scene, the results show that theimprovement of both conditions has obvious effect, and the effect of human treatment underthe two conditions is not good. There are obvious differences, which show that virtual reality model will gradually replace the on-the-spot feeling. Finally, with more and more researchers have put forward a variety of other virtual reality image processing models, the research of image processing has gradually become more and more interested.

Keywords: DNN, Virtual Reality, Medical image, Image processing, Optical imaging

http://doi.org/10.3837/tiis.2022.05.009 **ISSN : 1976-7277**

1. Introduction

Exposure psychotherapy is a kind of psychotherapy technology aiming at reducing or improving the symptoms or bad behaviors of patients [1]. Its development has a history of more than one hundred years. It has the characteristics of strong pertinence, easy operation, short course of treatment and quick effect. As there are many schools of thought, their kinds and names also present a situation of contention. Recently, in the treatment report of virtual reality computer image exposure, 200 participants with acrophobia were studied, including 34 in virtual reality condition group and 32 in immersive condition group. In immersive condition, the test was carried out from the height of open overpass or tower to the ground[2]. On the other hand, under the condition of virtual reality, immersive virtual reality computer three-dimensional image reproduction is carried out on site, the test content is the measurement of disease degree and the actual movement monitoring of overpass. After the test, the comparison results show that under both experimental results, the symptom improvement has obvious effect, and there is no obvious difference between the two treatment conditions. Under this condition, after six months of cycle maintenance treatment, the treatment effect of the disease, VR condition and immersion condition are the same.

At present, the research of VR image processing, from medical imaging technology to medical image processing, from big data analysis to artificial intelligence, more and more new technologies are applied to the medical field, which also improves the safety and effective treatment of patients in the process of medical treatment[3]- [6].

With the deepening dependence of human beings on deep learning technology, the big data computing technology of VR image algorithm [7] - [10]. Model has a far-reaching impact on the development and application of deep learning. In this paper, visual analysis technology is used to show the working process of the model. It is a better way for researchers to understand the working mechanism of deep learning design, and to develop an effective VR processing system using this model.

On the other hand, for the system that used virtual reality image before, the hardware level also tends to be mature, and the related originals have been fixed [11] - [14]. It is quite difficult for developers to make their own originals. If they purchase images for tens of seconds, hundreds of thousands of Yuan will be incurred. Therefore, it is very difficult to prepare all kinds of on-site hardware very expensive. However, for the new system, the computer image application software package of virtual reality system is free of charge, with few equipment and low specific requirements for patients.

In this way, in the process of the continuous development of virtual reality image theory and application, virtual reality technology has been widely concerned in medical treatment.

2. Related Work

2.1 About the effective range of the virtual reality visual range

Construction and visualization of medical knowledge maps under deep learning, and the vision of virtual reality, not only the projected image prompts, but more importantly, the experience of the experience, therefore, the system was developed based on the effective range of visual effects.

Max view of human eyes is 210 ° in parallel direction and 125 ° in vertical direction, but the

effective field of vision for text and color is 30 ° in horizontal direction and 40 ° in vertical direction. With the expansion of the field of vision and the influence of image on vision, immersion is also increasing. However, with the saturation of the beyond degree and effect, the immersion will not increase even if the image range is enlarged. Therefore, this range is called the induced field of vision, which is 100 ° in the horizontal direction and 85 ° in the vertical direction. It is suggested that the visual field of this study should be more than 100 ° for horizontal field of vision and 85 ° for vertical field of vision.

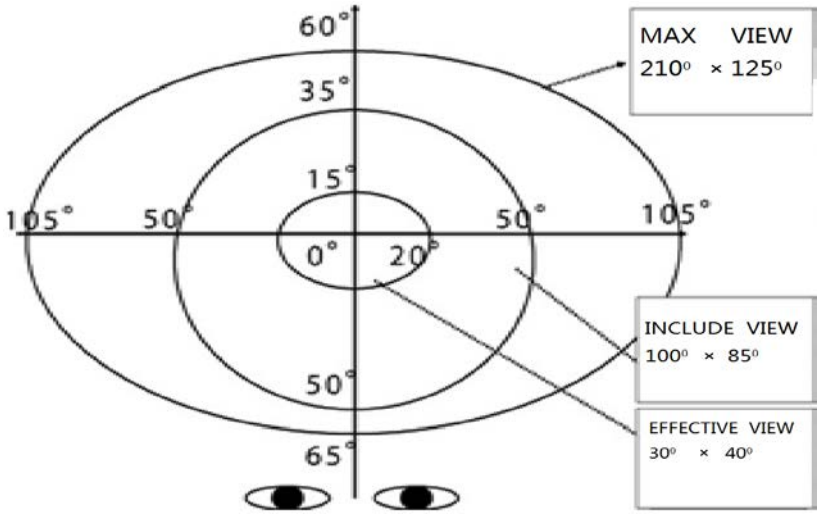

Fig. 1. Field of view

2.2 Proposal virtual realization system

Based on the above field of view, this research develops the view system equipment revealed in **Fig. 2** According to the principle of low cost and simplicity of equipment, a projector is used. After a concept simple projector projects from frontone, observer will produce intense, so, in order to enhance the stereoscopic feeling, the system is a hemispherical system.

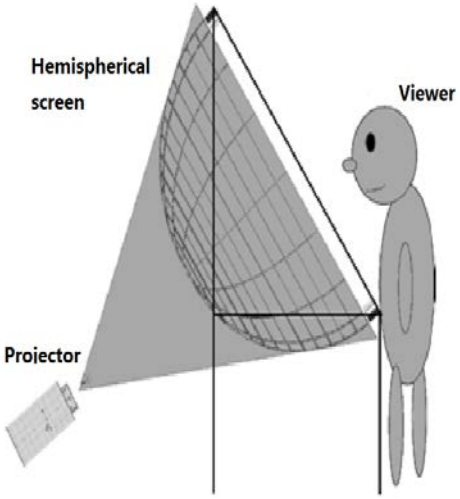

Fig. 2. System concept

Fig. 3. System actual

In **Fig. 3**, the real system diagram after the real installation according to **Fig. 2** is composed of one PC computer, one projector, one semi transparent screen and other auxiliary equipment.

With such simple equipment, the virtual reality image display can be achieved, which not only simplifies the hardware, but also facilitates the transportation and installation.

Forthe system uses a curvedsurface screen, the projection tilt must be corrected. We correct the system in two ways, the first is to calibrate the computer view, and the second is projection correction coordinate to achieve projection oblique correction.

3. Projection oblique correction processing

Generally speaking, projection of the raw data is the plane screen image, and we choose the hemispherical screen projection to realize the stereoscopic projection of virtual reality. Because the projection of hemispherical screen is inclined, we need to correct the original data image and projection. Therefore, in order to have a more effective effect on the virtual reality projection image, two correction processes should be carried out for the original data image and projection, namely, the computer image joint processing and the optical image joint processing.

3.1 Computer image joint processing

In order to display the scene image in a wider area, the wide-angle camera is used when collecting the image. Therefore, as shown in **Fig. 4**, the light passing through the centre of the lens will bend more than the light far away from the centre, which is called radial distortion. In addition, as shown in **Fig. 5**, the photography position section and the lens section are not parallel, which is called circular distortion, so it is necessary to process and correct the image by computer.

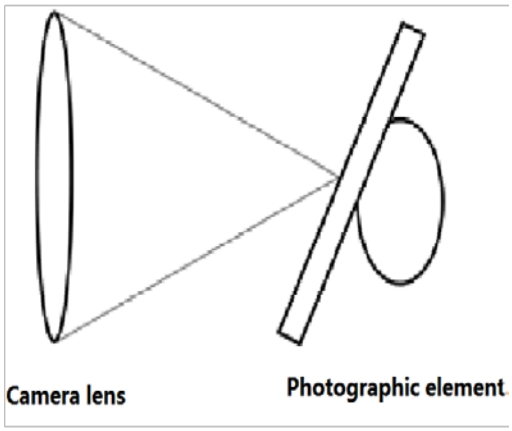

Fig. 4. Radial distortion

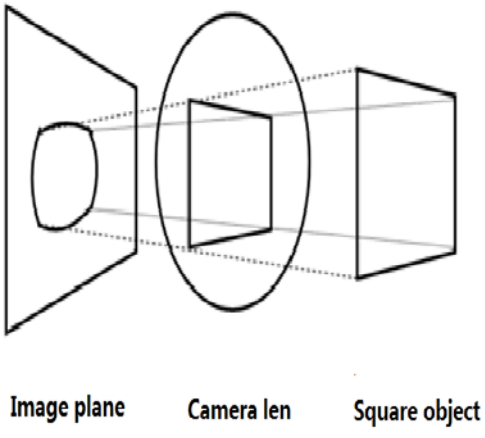

Fig. 5. Circumferential distortion

3.1 .1 Deep learning algorithm

The principle of deep learning algorithm is proposed in the application of image measurement and machine vision. The information obtained in learning and training plays a very important role in the interpretation of image, text, sound and other data. The ultimate goal of deep learning is to enable machines to analyze and train like humans, and recognize sound, text, image and other data. Deep learning is a complex learning algorithm. Compared with previous related technologies, it has achieved better results in speech and image recognition.

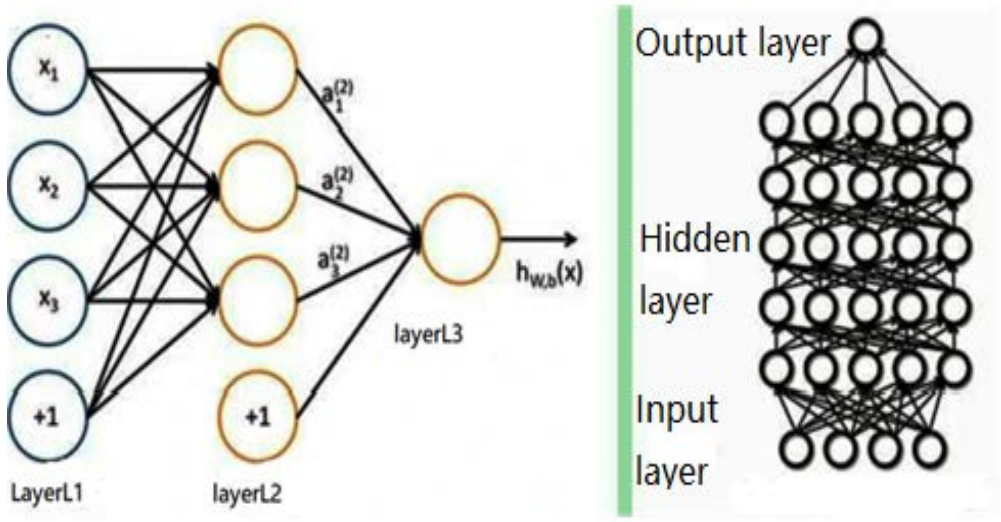

Fig. 6. Deep learning model with multiple hidden layers. [15]

In this study, there is a simple relationship between the view obtained by the camera and the object in three-dimensional space, as shown below: [image] = m [object], matrix M can be regarded as the geometric model of the camera view, and the parameters in M are camera parameters. The parameters are generally obtained through experiment and calculation. The process of obtaining the parameters is called camera calibration.

Camera calibration principle

The template surface is on the surfacewith the world coordinate system z = 0.

Basic principle:

$$s\begin{bmatrix}u\\v\\1\end{bmatrix} = K[r_1 r_2 r_3 t]\begin{bmatrix}X\\Y\\0\\1\end{bmatrix} = K[r_1 r_2 t]\underbrace{\downarrow}_{H}\begin{bmatrix}X\\Y\\1\end{bmatrix} \quad (1)$$

K is the internal matrix parameter of the camera, [xy1] t is the homogeneous coordinate of the point on the image processing template plane, [u v 1] t is the homogeneous coordinate of the corresponding point projected on the image processing template plane, [R1 R2 R3] and T are the rotation matrix and translation vector of the world coordinate system relative to the camera coordinate system respectively.

$$H=[h_1 h_2 h_3] = \lambda K[r_1 r_2 t] \quad r_1 = \frac{1}{\lambda}K^{-1}h_1,\ r_2 = \frac{1}{\lambda}K^{-1}h_2 \quad (2)$$

Due to the properties of rotation matrix, such as r1tr2 = 0 , the following two internal parameter matrix basic constraints can be obtained for each image

$$h_1^T K^{-T} K^{-1} h_2 = 0 \quad (3)$$

$$h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \quad (4)$$

Because the camera has five internal parameters, when the number of photographed images is greater than or equal to 3, K can solve the linear unique solution.

Algorithm description

(1) The convolution based neural network system is convolution neural network (CNN).

(2) Self coding neural networks based on multilayer neurons include auto encoder and sparse coding, which are widely concerned in recent years.

(3) The neural network is pre trained in the way of multi-layer self coding neural network, and then the depth confidence network (DBN) of neural network weight is further optimized by combining the identification information.

Because the camera has five internal parameters, when the number of photographed images is greater than or equal to 3, K can solve the linear unique solution.

Through multi-layer training, the "low-level" feature representation of the original data is gradually transformed into "high-level" feature representation, and then the "simple model" is used to complete complex classification and other training tasks.

In this study, the algorithm principle of deep learning convolution neural network (CNN) is adopted:

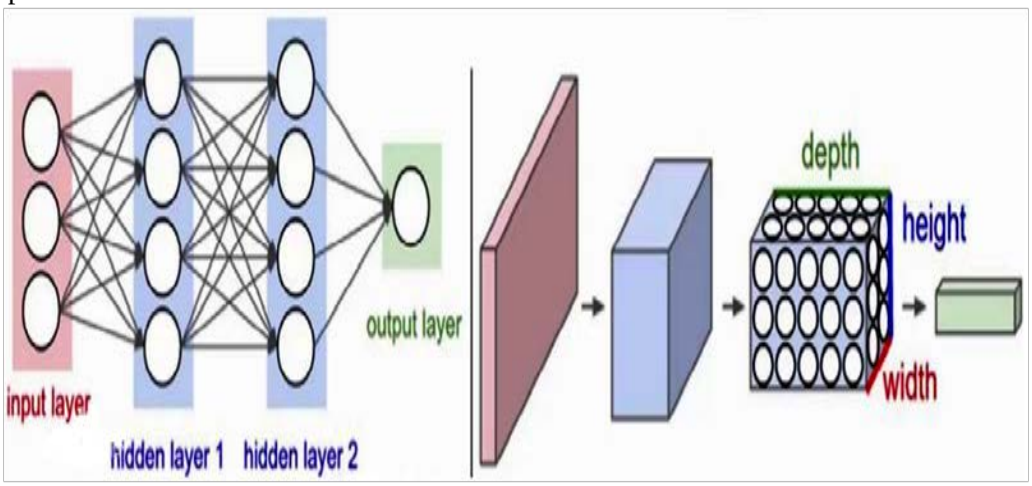

Fig. 7. Deep learning CNN model with multiple hidden layers [15]

Convolution neural network is inspired by the structure of vision system. The first convolution neural network computing model is proposed in Fukushima (D's neural cognitive machine). For the local connection between neurons and hierarchical organization image transformation, neurons with the same parameters are applied to different positions of the original neural network to obtain the translation invariant neural network structure. Then, the convolution neural network with error gradient is designed and trained, and the pattern recognition task has achieved excellent performance. The pattern recognition system based on convolution neural network is one of the best implementation systems so far, especially in image processing and recognition.

Among them, the model algorithm of convolution neural network (CNN) takes pictures ofa certain calibration plate in different directions and levels (above three levels), without knowing the movement mode of the calibration plate. The internal parameters (matrix A in reference) and distortion coefficients of the camera are obtained directly. The accuracy of the calibration method is higher than that of the single calibration method, and there is no need for

a positioning instrument with very high accuracy.

3.2 Optical image joint processing

The vision of the image projected by the projector is affected by the observation point and the position of the projection. In this experiment, as shown in **Fig. 8**, the distance from the center of the hemispherical screen to the display projector is set to 180cm, and the position of the observation point is the center of the outer diameter of the hemispherical screen. For this observation point, **T** is the projection position of the projector, and**L** is the position of the ideal observation point.

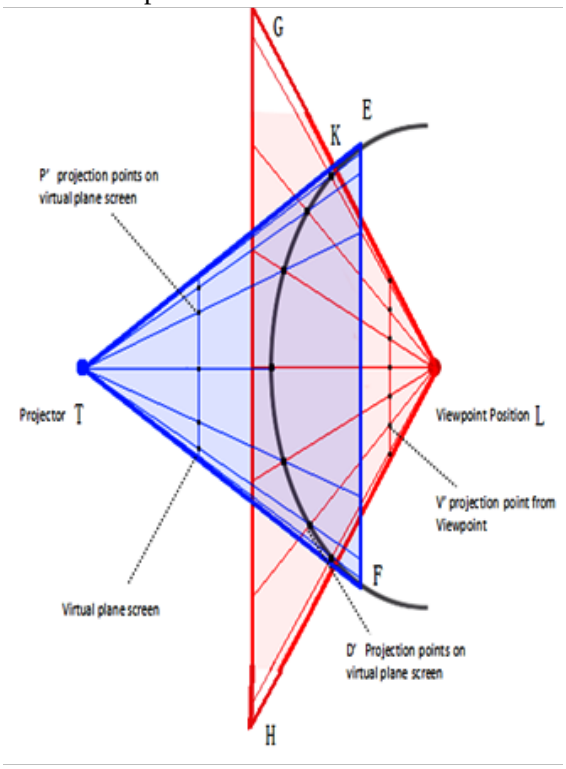

Fig. 8. Before geometric optics correction

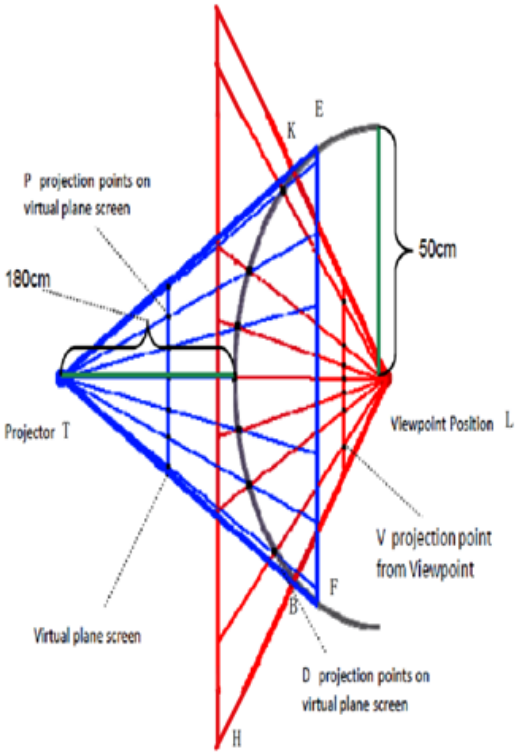

Fig. 9. After geometric optics correction

Triangle **TFE** area and the triangle **HLG** area are the areas set as the field of vision by the projection area of each part of the projector. The curved gray arc ekbf is the hemispherical .

First, from the initial point of view, the image projected by the projector is projected onto the plane screen, while the simulated image is equidistant between the projection points **P** of the plane. Intersection of the projector point and the extension line of the projection point **P** with the hemispherical screen is the projection point **D**. The projection point **D** is connected with the ideal viewpoint, that is, the projection point **V** seen from the ideal viewpoint. As a result, the equidistant projection point (**P**) projected by the projector is the bending point (**V**) bent outwards from the ideal viewpoint.

Second, if the projection point **V**seen from the ideal viewpoint is corrected, it will become an equidistant arrangement. In this way, it is necessary to track the case that the projection point **V** seen from the ideal observation point position is equidistant. Therefore, it is necessary to correct the image **P** projected by the projector.

Fig. 9 shows the corrected model diagram. Set the arrangement of projection points (**V** ') seen from the ideal observation point to equal distance, and extend the ideal observation point

L and projection point (**V'**). The intersection between the extension line and the hemispherical screen is the projection point **D'**, and the intersection between the projection point **D'** and the projector position **T** and the image projection plane is **P'**. The position of the projection point is determined by tracking from the viewpoint. From this point of view, the distance between the projection points of the simulated projection plane is not equal. Suchimage transformation is geometric optical processing.

4. Experimental Classification Results and Analysis

As shown in **Fig. 10** and **Fig. 11**, the correction verification using a square lattice is shown. Because the projection screen is a hemispherical surface, the periphery of the hemispherical surface is not easy to correct. However, the necessary induced visual field seen in the experiment needs great correction to achieve the goal of the experiment.

About the experimental results: as shown in **Fig. 10** and **Fig. 11**, the correction verification using square lattice is shown. Since the projection screen is a hemispherical surface, the periphery of the hemispherical surface is not easy to correct. However, the necessary induced visual field seen in the experiment needs to be greatly corrected to achieve the purpose of the experiment.

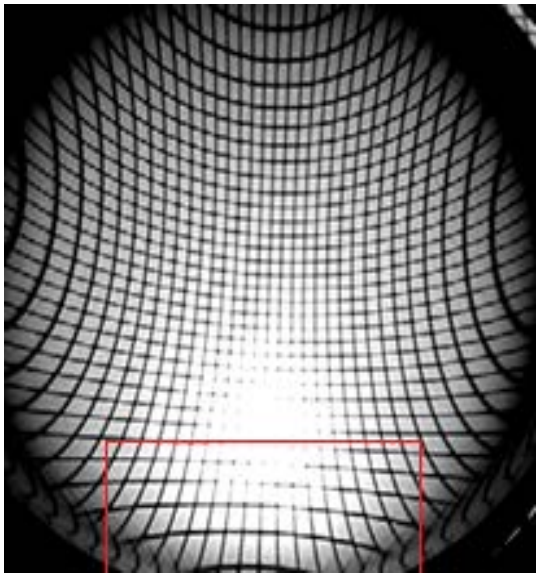

Fig. 10. Image before projection correction

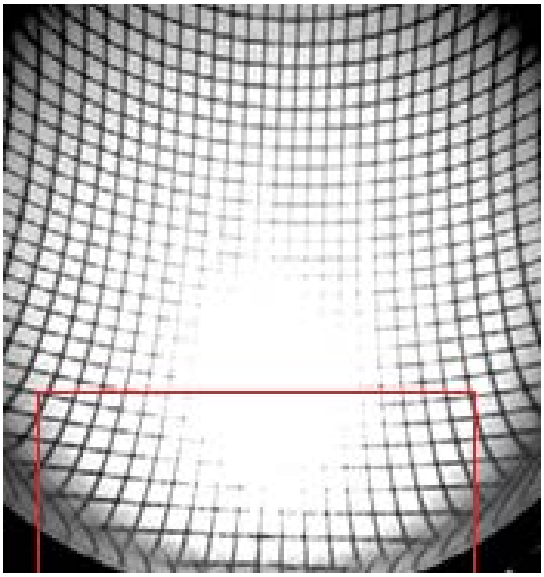

Fig. 11. Image after projection correction

Next, as shown in the figure below, is the corrected actual projection. The actual projection is the hemispherical projection animation image after correcting the photographic image from the elevated cable car to the ground. For the observer, the immersion and presence of this animated projection scene have been greatly improved, which shows that virtual reality has great expectations for the effectiveness of exposure therapy (such as acrophobia). The following are the pictures after actual projection:

Fig. 12. Image before projection correction

Fig. 13. Image after projection correction

The equipment for this experiment is shown in **Table 1**:

Table 1. Test equipment

Real time system correction		**Deep learning training correction**	
OS	**Windows7**	**OS**	**FreeRTOS**
CPU	**Gen Intel(R) Core™ i5-1155G7@2.50GHz**	**CPU**	**RISC-V Dual Core 64bit-Kendryte K210, mobilenet/TinyYOLOv3**

Two methods are used to correct the contrast image, as shown in **Fig. 14** and **Fig. 15**:

Fig. 14. Real time system correction

Fig. 15. Deep learning training correction

The evaluation of this system is realized by questionnaire survey. The application of the experiment is the projection of animation videos from the windows of high-rise buildings. The objects of the questionnaire survey are students. Through the investigation of 500 college students, we get an immersive evaluation. Considering the hemispherical projection characteristics, the necessary systematic analysis is carried out to speculate the immersion of all members.

The following two contents are analyzed and inferred:

a. Analysis of the projection field: effective view field and hemisphere

b. Analysis of discomfort

Analysis of these two points is carried out through two experiments, fixed at the ideal view point and freely changingview point at the same time.

Table 2. Analysis of ideal viewpoint

Immersion	**Effective field of view**	**hemisphere**	**No special difference**
	28%	55%	17%
Maladjustment	**Effective field of view**	**hemisphere**	**No special difference**
	14%	32%	54%

Table 2 and Table 3 are the questionnaire survey results of the experiment, **Table 2** is Analysis of ideal viewpoint, and **Table 3** is Analysis of free viewpoint.

Table 3. Analysis of free viewpoint

Immersion	**Effective field of view**	**hemisphere**	**No special difference**
	14%	77%	9%
Maladjustment	**Effective field of view**	**hemisphere**	**No special difference**
	18%	45%	37%

The analysis of the questionnaire survey results, the mapping in the global projection of the hemisphere has a more obvious effect than the observation of the students who are limited to the effective field of view, especially in the free view. Here, because the set induction range starts from the ideal viewpoint, there will be no such feeling at the free viewpoint, which is the reason why the edge is considered in consideration.

In addition, more than half of the observers from the ideal viewpoint have an unsuitable answer, which is the feeling of the same size as the free viewpoint. This is because the image at the ideal viewpoint is corrected, and the deviation around the perimeter of the hemisphere cannot be suppressed.

5. Conclusion

This study collects and analyzes the research on psychotherapy exposure therapy in related disciplines in advance, discusses the feasibility of establishing models in different ways of the simulation system, and this paper its impact on the subject's behaviour with consideration of relevant errors. Secondly, further consider the effects of joint computer image processing and optical joint imaging, and use comprehensive optical image theory to analyze the effects of special structures.

Collect data on the interaction structure of several social networks, design relevant rules, simulate and analyze the influence of the interaction structure of the subject on the behaviour of the subject, and analyze the reasons behind the phenomenon based on the above conclusions.Discuss the special points of specific diseases, and analyze the effectiveness of patients of the same type by changing the transformation of virtual reality images.

This method also has limitations. For example, the stability of this method is relatively poor. How to improve the stability of graphics and images and optimize the training method in real time without reducing the effect of the model is the direction of follow-up work. In addition, from the perspective of real scene, reconstructing 3D model and improving system resolution are the focus of future work.

Acknowledgement

The authors wish to thank Li XiaoDong, Shi XiaoYing, Liu Geng, Chen Bin. This work was supported in part by a grant from LY19F020042, LGF19F020015.

References

[1] Huiyu LI, Linwei FAN, "A flexible technique to select objects via convolutionalneural network in VR space," *Science China(Information Sciences)*, 63(01), 57-76, 2020. Article(CrossRef Link)
[2] EJie CHEN, Shixiong KAI, "Cooperative transportation control of multiplemobile manipulators through distributed optimization," *Science China(Information Sciences)*, 61(12), 5-21, 2018. Article(CrossRef Link)
[3] P.VIJAYA, Satish CHANDER, "LionRank: lion algorithm-based metasearch engines for re-ranking of webpages," *Science China(Information Sciences)*, 61(12), 65-80, 2018. Article(CrossRef Link)
[4] SpalloneR., Lamberti F., Guglielminotti Trivel M., Ronco F., Tamantini S., "3D RECONSTRUCTION AND PRESENTATION OF CULTURAL HERITAGE: AR AND VR EXPERIENCES AT THE MUSEO D'ARTE ORIENTALE DI TORINO," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVI-M-1-2021, pp. 697-704, 2021. Article(CrossRef Link)
[5] Ye Ding, LiJun Yang, MingHui Hong, "Enhancement of pulsed laser ablation assisted with continuous wave laser irradiation," *Science China(Physics, Mechanics & Astronomy)*, 62(03), 70-78, 2019. Article(CrossRef Link)
[6] Kaimaris D., Roustanis T., Karolos I. A., Klimantakis K., Patias P., "ALIGNING POSITIONING AND NAVIGATION OF SMART MOBILE DEVICES TO AR/VR ENVIRONMENTS," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-M-1-2021, 347-353, 2021. Article(CrossRef Link)
[7] Markus Antonietti, "Polymer coated nanowire network promises self-healing and superstretchable conductors," *Science China(Chemistry)*, 62(02), 151-152, 2019. Article(CrossRef Link)
[8] Chuan ZHANG, Lulu GE, Yuchen ZHUANG, Ziyuan SHEN, Zhiwei ZHONG, Zaichen ZHANG, Xiaohu YOU, "DNA computing for combinational logic," *Science China(Information Sciences)*, 62(06), 5-20, 2019. Article(CrossRef Link)
[9] Tingting Wu, Yuemeng Yang, Yu Cao, Yan Huang, Li-Ping Xu, Xueji Zhang, Shutao Wang, "Enhanced lateral flow assay with double conjugates for the detection of exosomes," *Science China(Chemistry)*, 61(11), 1423-1429, 2018. Article(CrossRef Link)
[10] Yinjuan Huang, Qichun Zhang, "Manipulating asymmetric photon transport through electrical control: a new strategy to construct optical diodes or isolators," *Science China(Chemistry)*, 61(11), 1351-1352, 2018. Article(CrossRef Link)
[11] Sparkes Matthew, "AI can tweak VR videos to stop cybersickness," *New Scientist*, 249(3326), 2021. Article(CrossRef Link)
[12] Jung Sungchul, Li Richard, Mckee Ryan, Whitton Mary C, Lindeman Robert W, "Floor-vibration VR: Mitigating Cybersickness Using Whole-body Tactile Stimuli in Highly Realistic Vehicle Driving Experiences," *IEEE transactions on visualization and computer graphics*, vol. 27, no. 5, pp. 2669-2680, 2021. Article(CrossRef Link)
[13] Matthew Sparkes, "AI can tweak VR videos to stop cybersickness," *New Scientist*, vol. 249, no. 3326, p. 16, 2021. Article (CrossRef Link)
[14] Song Hayeon, Kim Taenyun, Kim Jieun, Ahn Dohyun, Kang Youngcheol, "Effectiveness of VR crane training with head-mounted display: Double mediation of presence and perceived usefulness," *Automation in Construction*, vol. 122, 2021. Article(CrossRef Link)
[15] https://baike.baidu.com/item/deeplearn/3729729?fr=aladdin#ref_[5]_10105430

GONG ZhaoZhe received the Ph.D. degree in Computer Science from Chiba University from Japan in 2013. He is now a lecturer at School of Computer Science and Technology, Hangzhou Dianzi University, China. His research interests include Image Processing, Software Quality, Machine Learning andNatural Language Processing.

Xiaodong Li received the Ph.D. degree in Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, China. His current research interests include pattern recognition, machine learning, image processing and data mining.

XIAOYING SHI received the Ph.D. degree in control theory and control engineering from the Zhejiang University of Technology, China, in 2014. She is currently a Lecturer with the School of Computer Science and Technology, Hangzhou Dianzi University. Her research interests include visual analysis, machine learning, and urban transportation.

Geng LIU received the Ph.D. degree in Computer Science from University of Insubria (Italy) in 2014. He is now a lecturer at School of Computer Science and Technology, Hangzhou Dianzi University, China. His research interests include Software Metrics, Software Quality, Machine Learning andNatural Language Processing.

Chen Bin received his Ph.D degree in Tokyo University of Agriculture and Technology, Japan, in 2013. He is a lecture at the School of Computer Science, Hangzhou Dianzi University, China. His research interests include software engineering, machine Learning, and handwriting character recognition.