Scientific and Social Research, 2024, Volume 6, Issue 9
https://ojs.bbwpublisher.com/index.php/SSR
Online ISSN: 2981-9946
Print ISSN: 2661-4332

Reproduction of a Digital Avatar using AI Smart Glasses

Yancheng Wu*

Fujian Changle No. 1 Middle School, Fuzhou 350001, Fujian Province, China

***Corresponding author:** Yancheng Wu, 18859118000@139.com

Abstract: In response to the rapid development of artificial intelligence, this article is committed to researching the creation of AI glasses that combine augmented reality technology and artificial intelligence big model reasoning ability. The study uses a dimensionality reduction model to record human experiences in multidimensional space, continuously feeds behavioral habits through machine learning, and constructs a virtual digital human that conforms to the characteristics of digital avatars.

Keywords: AI smart glasses; Big language models; Machine learning; Digital avatar

Online publication: September 27, 2024

1. Background

In recent years, significant technological breakthroughs have been made in the field of artificial intelligence (AI), particularly in areas such as deep learning, computer vision, speech recognition, and natural language processing. In addition to the rapid development of cloud computing, edge computing, and other technologies, the maturity of these technologies provides good technical conditions for the training of AI digital avatars. With the development of technology, people's demand for personalized services is increasing day by day, and they are increasingly eager to achieve more natural and smooth interaction with machines. They are also increasingly eager to cultivate a digital avatar that fits their personality traits, behavioral habits, and risk preferences, to help them complete higher-level and complex mental labor. Cultivating personalized digital individuals is imperative and will soon become the focus of various industries.

2. Research on the current situation at home and abroad

With the development of artificial intelligence, AI smart glasses, VR glasses, Eye tracker, near-infrared brain scanner, and other devices have been released and continuously iterated and updated both domestically and internationally.

AI smart glasses integrate technologies from multiple fields such as sensors, artificial intelligence, microcomputers, communication technology, and biomedical engineering to achieve a wide range of functions including human-computer interaction, intelligent perception, and augmented reality (AR). It not only provides users with a convenient way to obtain information but also greatly enriches their daily life and work experience.

VR glasses, also known as virtual reality glasses, use head-mounted display devices to close off a person's visual and auditory perception of the outside world, guiding users to create a feeling of being in a real environment [1].

An eye tracker, also known as an eye tracking analyzer, is a digital video analyzer device that uses machine vision technology to analyze the eye movement trajectory of humans or animals. It can record and analyze a person's eye movement trajectory in real time and accurately through precise gaze measurement technology, including key data such as fixation point, fixation time, and eye movement distance. These data are of great significance for revealing human cognitive processing processes, psychological activity states, and visual information extraction mechanisms.

Near-infrared brain scanner, more commonly known as near-infrared functional brain imaging (fNIRS) or high-density near-infrared brain imaging (HD-NIRS), is a non-invasive neuroimaging technique used to study brain functional activity. It analyzes brain functional activity, particularly changes in blood oxygen levels and blood flow, by detecting near-infrared light reflection and scattering signals within the brain [2].

The combination of these device functions makes it possible for people to explore and create digital humans that truly fit my characteristics.

3. Exploration of implementation theory

3.1. Fundamentals

Firstly, it is important to understand the fact that only two-dimensional nonlinearity can be quantified. All information needs to be reduced in dimension and transmitted based on two-dimensional codes (010011010011011). In high-dimensional applications, the output can be restored. However, once three-dimensional models are reduced in dimension, the information will inevitably be lost. What researchers need to consider and study is how to reversibly preserve multidimensional dynamic behavior in multiple two-dimensional systems. For example, a three-dimensional device requires at least three orthogonal projection planes (front, side, and top) to store information to ensure that the main information is not lost after dimensionality reduction of complex information and to achieve relatively complete restoration in applications. If higher accuracy is required, six or more projection surfaces are needed, including up, down, left, right, front, back, and even more. Therefore, researchers need to understand and demonstrate a mathematical formula: $R^n = nR^2$ or $R^n = (n + \infty)R^2$. If the total dimension increases by 1, such as adding a time dimension, a two-dimensional plane mapping system needs to be added on this basis.

3.2. Training steps

Step 1: Slice collection of causal information. Through the wearable terminal device with a built-in large model, it constantly collects complex causal relationships between oneself and the outside world (parental characteristics, words, and deeds, personal image, personal audio and video, school atmosphere, teacher teaching, family atmosphere, family relationships, online activities, information reception, classmate interactions, roommate interactions, club activities, social activities...). These factors have complex causal relationships with personal growth, and the causal relationships are complex and intertwined. There may be multidimensional interactive relationships, and researchers need to slice and record this information at different periods [3]. For example, recording the teaching situation of the chemistry teacher at a certain time in the morning, recording the chemistry issues discussed during personal interactions with classmates at noon, and recording the chemistry knowledge one has read and received in the afternoon will enhance understanding and cognition of a certain field of chemistry [4].

Step 2: The dynamic fusion of trajectory and causal information. Arrange complex relationship curves in high-dimensional space by time slicing, and project them onto a series of plane trajectories for event intersection and combination. Based on the information preservation in step one, it further helps to reconstruct the correlation relationship through reasoning and infer the influence relationship between factors. In this way, over time, it can fully and dynamically record one's thinking characteristics, behavioral habits, hobbies, and risk preferences.

Table 1. Related factor results

Factor 1							1, 7
Factor 2		2, 2					
Factor 3				3, 4			
Factor 4		4, 2	4, 3			4, 6	
Factor 5	5, 1		5, 3				
Factor 6				6, 4		6, 6	
	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6	Time 7

Step 3: Reasoning brings about the development of cognition. Continuously feeding dimensionality reduction data into the large model base to form continuous memory, utilizing the inference ability of the large model for sensitivity analysis, and stimulating the emergence ability of deep cognition in the trained individuals.

Step 4: Establish the L1 layer of the large model. Build the above training results into a stable L1 layer model that can flexibly switch CHAT bases, and can switch open source bases according to the characteristics of the large model base and personal preferences.

Step 5: Continuously repeat the above steps. Through a period of continuous training, The digital personal world was perfectly replicated, forming a digital avatar that fits their personality.

4. Exploration of technical paths

4.1. Implantation training pathway

4.1.1. Primary: Touch me mode

Configuration: Traditional Bluetooth smart glasses + recording + OCR semantic recognition text conversion function + online or offline large model personalized fine-tuning function + BeiDou positioning + memory storage tool based on registered users (**Figure 1**).

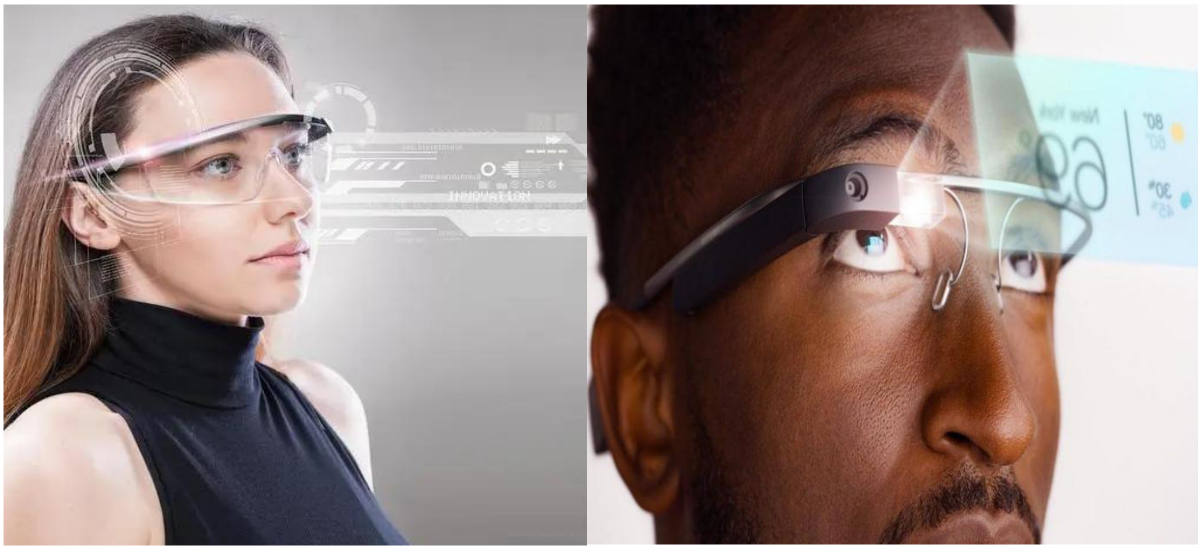

Figure 1. Application examples

Usage: Used to comprehensively receive multidimensional information from people's daily learning, work, life, footprints, and social circle activities.

4.1.2. Advanced: Look at me mode

Configuration: Touch me mode + built-in or external eye tracker (**Figure 2**).

Figure 2. Me mode examples

Usage: To emphasize the reception of the text, images, videos, etc. that people are particularly interested in within the intersection of the two eyes' line of sight triangle, and to enhance my personalized preferences.

4.1.3. Higher: Copy me mode

Configuration: Look at me mode + wearable near-infrared brain computer (**Figure 3**).

Figure 3. Look at me mode

Usage: By monitoring the blood flow intensity in different thinking activities through 52 brain regions, or monitoring the differences in the excitement status of skin pores, high-precision recognition of people's emotional states in a certain activity can be achieved, further enhancing the judgment ability of the large model on personal preferences.

4.2. Tips and deviation optimization

Prompted and Deviation Optimization utilizes prompt engineering to systematically optimize the deviation states generated during the training process of digital humans.

One is precise role positioning. Assign specific roles to digital individuals, such as consultants, analysts, planners, architects, rigorous scientists, certified public accountants, and so so, to meet the training needs in specific scenarios.

The second is to constrain the output boundary. Use appropriate constraints to help digital avatars provide accurate and comprehensive answers in legal, institutional, knowledge, and other Q&A scenarios, such as prompting "Please do not fabricate."

The third is to stimulate creative thinking. In a long-term constrained environment, positive incentives are needed to guide digital avatars to break rigid thinking patterns, enrich their knowledge levels, and extend their work scope with a closed-loop thinking approach.

5. Application prospects

By developing a digital person that fits behavioral characteristics, one can imagine that in many fields, it can easily help one work, study, and even entertain under the authorization of "oneself." For example, professionals such as teachers, doctors, auditors, and so on can use digital avatars to replace themselves and be active in the work environment, providing services such as classroom, after-school, guidance, follow-up, data analysis, report writing, and so on, greatly improving work quality and efficiency. Students can arrange digital avatars according to their plans to help "oneself" accurately subscribe to the knowledge they are interested in in the ocean of knowledge. Players can combine VR applications to allow digital avatars to roam freely in the virtual world, with countless application scenarios. It is believed that in the near future, digital human training will become a reality, and digital humans will also become the most capable assistants for people in various industries.

Disclosure statement

The author declares no conflict of interest.

References

[1] Ni H, Zhuo R, 2024, Smart Glasses, Still Waiting for a Real Explosion. Global Times New Media, published August 24, 2024.

[2] Pu Qin Intellectual Property, 2022, The Opening of the Metaverse — VR Glasses. https://www.sohu.com/a/514919927_120064369

[3] Yin ZQ, Yao JP, Yang H, et al., 2005, Digital Video Eye Movement Analyzer, patent, CN200510132753.9.

[4] Gong JY, Zhong Y, Fei DC, et al., 2024, Progress in the Application of Near Infrared Brain Functional Imaging in the Diagnosis and Rehabilitation of Mild Cognitive Impairment. Chinese Journal of Physical Medicine and Rehabilitation, 46(06): 556–560.