Development Status and Future Challenges of Human-Computer Interaction Devices in the Metaverse

Yuhao Ma

Suzhou NO.3 High School, Suzhou, 215335, China
Corresponding Author: Yuhao Ma, Email: kma736020@gmail.com,

Abstract
With the development of technologies such as human-computer interaction, augmented reality, blockchain technology, and digital twins, as well as the growing demand for social interaction and digital assets, the Metaverse is developing rapidly. It has become an important development direction of the future digital world. Human-computer interaction is the core to realize virtual-real integration in the Metaverse, which solves the challenges of multi-user collaboration, communication, and socializing. To deeply explore the Metaverse as a digital space that integrates the virtual and real worlds, how its combine with human-computer interaction devices to create rich, immersive and personalized user experience, this article first introduces the definition of the Metaverse and human-computer interaction technology. And then, some relevant methods and devices of human-computer interaction are summarised to point out the key factors of current devices. The development of human-computer interaction devices, such as virtual reality and augmented reality devices, is discussed. The findings suggest that the major challenge is accurately capturing and understanding users' natural interaction behaviors and state perceptions in a virtual environment. Furthermore, the portability and costs of hardware should also be considered for expanding the development of human-computer interaction devices in Metaverse.

Introduction

In today's society, technological advancements are progressing rapidly. The concept of the Metaverse, as a virtual shared space (Mystakidis, S, 2022), has garnered widespread attention in recent years. The metaverse can provide users with an immersive virtual environment, which can be applied to online education, entertainment, and so on. It also promotes the emergence of new business models and the development of technology, especially in the fields of virtual reality, augmented reality, and human-computer interaction (HCI) (Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. 1994). HCI devices serve as the bridge between the real world and the virtual world, and have become a popular topic in recent years. HCI devices include head-mounted displays (HMDs), handheld controllers, motion capture devices, and haptic feedback devices. These devices play a crucial role in the construction of the Metaverse because they are the basic tools that we should use to enhance the sense of experience.

Citation:Yuhao Ma. (2025) Development Status and Future Challenges of Human-Computer Interaction Devices in the Metaverse .The Journal of Young Researchers 1(4): e20250523

Received on April 28, 2025; Accepted on May 8, 2025; Published on May 23, 2025

By summarizing and exploring this topic, this paper will first introduce the definition and development of Metaverse and HCI devices, then discuss the current issues faced by HCI devices and analyze the current status and challenges of interaction devices through case studies, and finally summarize the possible future development directions of HCI devices.

In addition, some companies have already developed HCI devices. Apple Inc. recently released a product called VisionPro, which is one of the most advanced technologies available today. It has many representative features, but it also has some shortcomings. Its impact on the market and future technological trends are worth studying. This article will also analyze the challenges faced by existing HCI devices through statistical analysis of the global HCI device shipments.

A review of relevant literature reveals that although recent studies have conducted much research on HCI, such as the applications in Augmented Reality (AR), Virtual Reality (VR), and so on. This paper differentiates itself from existing literature by focusing on the development and challenges of HCI in the Metaverse, while also analyzing real-world cases and envisioning future challenges and applications.

Literature review

Metaverse

Metaverse is a collective virtual shared space created by the convergence of virtually enhanced physical reality and physically persistent virtual spaces, including the sum of all virtual worlds, augmented reality, and the internet. (Mystakidis, S. 2022) We can understand it as a digitally immersive space built through virtual reality, augmented reality, and the internet, allowing users to interact and engage in activities through virtual identities. Its concept was born in 1992. At present, the recognized core attributes of the Metaverse include the following: boundless, persistent, immersive, decentralized, economic system, and social experience. It provides users with a virtual universe full of possibilities.

The development history of the Metaverse mainly has five stages (Gilbert et al.,2013). The first stage is a text-based environment (1970s-1980s). It relied entirely on text input and output, and users interacted through the command line interface. The second one is the graphic environment (1980s-1990s). It introduced a 2D graphical interface and improved the user experience. The third stage is the 3D virtual world (1990s). It started to use 3D graphics, but the interaction and immersion are limited. The fourth stage is advanced 3D virtual space (2000s-2010s), which could provide richer interaction and higher immersion, using more advanced 3D graphics technology. And the fifth stage is integrated into a Metaverse (2020s-present). It integrates multiple virtual worlds and develops towards a unified meta-universe.

Human-computer interaction devices

HCI devices play a vital role in the development of the Metaverse as they build a bridge between humans and the Metaverse (Lyu, Z. 2023). Human-computer interaction is the channel to realize the information exchange between the user and the computer, and the user interface is the bridge to transmit and exchange information in this channel. In an ideal state, human-computer interaction will no longer rely on machine language. In the absence of intermediate devices such as keyboard, mouse, and touch screen, human-computer interaction can be freely communicated anytime and anywhere, thus achieving the final integration of people's physical world and virtual world. However, due to the limitation of technical limitations, this ideal cannot be reached at present. The user's behavior and state (input) are transformed into a representation that the computer can understand and operate, and the computer's behavior and state (output) are transformed into an expression that people can understand and operate through the interface. In the following sections, this paper will delve into the current development of various HCI devices, such as VR, AR, and brain-computer interface technologies, and explore the technology of current HCI devices, including their interactive methods.

HCI devices refer to any device that enables users to interact with computer systems. These devices utilize advanced interaction technology such as VR (Card, Moran & Newell, 1983), which fully immerses users in a computer-generated three-dimensional digital environment, providing a highly realistic and immersive experience. The development of VR has been nearly 30 years in history, but because it usually requires the use of expensive specific input devices, while the lack of new interaction paradigms in actual use, VR has not entered the mainstream application in the historical development of display devices. Nowadays, with the change of key people, events, ideas and paradigms in the field of human-computer interaction and the continuous improvement of user models under the guidance of psychology, VR is gradually moving from science fiction to reality, and its importance is not limited to the field of game entertainment, it will be from military, aerospace, computer-aided design, surgical simulation, scientific calculation visualization, remote control and education And other fields to change the world. In the future, there may be such a wearable product that will observe your daily actions, learn your way of dealing with people, gradually understand the behavior patterns between people, and quickly provide you with the help you need, like a symbiote living with you.

AR (Caudell & Mizell, 1992), on the other hand, enhances the user's perception of their surroundings by overlaying digital information onto the real world. It is a technology that skillfully integrates virtual information with the real world. AR extensively uses multimedia, 3D modeling, real-time tracking and registration, intelligent interaction, sensing, and other technical methods to simulate computer-generated text, image, 3D model, music, video, and other virtual information. Applied to the real world, the two kinds of information complement each other, thus achieving "augmented" the real world.

Additionally, brain-computer interfaces (BCI) can directly decode electrical signals from the brain and convert them into commands to control external devices (Nicolas-Alonso, L. F.,& Gomez-Gil, J. (2012). Brain computer interfaces, a review. sensors, 12(2), 1211-1279.), thereby establishing a direct communication link between the brain and machines. In the future, brain-computer interfaces may replace the current expression production relying on face modeling and speech recognition, and convey emotional elements through brain activity, making the user's experience in the meta-universe more real and immersive. At the same time, the brain-computer interface may also provide users with visual, auditory, touch and other aspects of experience through neurofeedback, and truly realize the free switch between the objective real world and the virtual world composed of algorithms and numbers, which can not only improve the authenticity of user experience, but also make the operation more convenient and efficient.

Human-Computer Interactive Methods

HCI devices employ various interactive methods to enable seamless communication between humans and computers (Zhang Fengjun, Dai Guo-zhong,& Peng Xiao-lan, 2016), including touch, gesture, posture, controller-based interactions, voice interaction, visual interaction, force and tactile interaction, multi-modal interaction, brain-computer Interaction, and so on. By analyzing the current interaction methods of HCI devices, it can provide a better reference for the discussion of the development status of HCI equipment in the following paper.

One significant interactive method of HCI devices is gesture and posture interaction, which utilizes video recognition to detect specific gestures and postures. For gestures, Cerney et al. (Cerney, M.,& Vance, J. (2005). Gesture recognition in virtual environments: A review and framework for future development. Human Computer Interaction Technical Report ISU-HCI-2005, 1, 28.) gave a broad definition: Gesture refers to the physical movement of a person's hand, arm, face or body to convey information or intention. Gesture recognition not only includes human motion tracking but also includes the interpretation of motion as semantic commands. For example, Microsoft's Kinect revolutionized gesture-based interaction by using an RGB camera and a depth sensor to track human movements with high precision.

With the progress of electronic information technology, the computing power and graphic display capability of handheld mobile devices represented by smartphones have been continuously enhanced, and different kinds of sensors such as cameras, GPS, accelerometers, gyroscopes, and electronic compass have also been integrated, improving the device's ability to perceive the environment and user behavior (Hinckley, K., Pierce, J., Sinclair, M.,& Horvitz, E. (2000, November). Sensing techniques for mobile interaction. In Proceedings of the 13th annual ACM symposium on User interface software and technology (pp. 91-100). In addition to the interaction of the devices themselves, attention has also been paid to the three-dimensional interaction capabilities of smart handheld mobile devices for use in VR/AR. There are several different prototype systems based on handheld device interaction, each of which has its own characteristics, but a common idea is to get rid of traditional desktop interaction and realize interaction between users and the environment, the core of which is the design problem of three-dimensional interaction based on handheld mobile devices. There are two types of interactive metaphors based on handheld mobile devices: the information lens metaphor and the direct pointing controller metaphor. Handheld devices facilitate interaction through portable devices, enhancing mobility. It gets rid of the traditional desktop interaction and realizes the interaction between users and the environment.

Voice interaction technology enables users to communicate with systems via speech recognition (Klevans, R. L.,& Rodman, R. D. (1997). Voice recognition. Artech House, Inc.), offering a hands-free experience, such as the voice assistants like Siri and Google Assistant, being good examples. Voice interaction takes an unused input channel, allowing efficient and precise input of large amounts of text, which is a completely natural and familiar way. In virtual reality user interfaces, voice input is particularly suitable for non-graphical command interaction and system control, that is, users issue voice commands to request the system to perform specific functions, change interaction modes, or system states. In addition, speech input also provides a complete means for symbolic input in HCI devices. The key to voice interaction is the speech recognition engine. Some existing speech recognition software, including Microsoft Speech API, IBM ViaVoice, Nuance, and the domestic iFlytek, have reached very good performance, and now, with the gradual opening of speech recognition technology and the open source speech technology threshold gradually lowered. In recent years, the rise of Google Glass, wearables, smart homes, and in-car devices has pushed voice recognition technology to the forefront of applications. Since Apple's iPhone 4s built-in Siri, almost all mobile phones have started to build voice assistant applications.

Visual interaction technologies, such as eye-tracking and head-tracking, display content by detecting the user's eye or head movements, making the experience more natural (Duchowski et al.,2008). Visual interaction refers to the principle of obtaining information from human vision, using a computer in conjunction with a camera, and detecting and identifying user operations using non-traditional input devices, thereby realizing natural human-computer interaction in two or three dimensions (Zhang, Y. (2020). Computer-assisted human-computer interaction in visual communication. Computer-Aided Design and Applications, 18 (1), 109-119.). This technology is particularly important in AR and VR devices.

Force and tactile interaction technologies allow users to feel physical sensations when interacting with virtual objects. Haptic feedback technology enables interaction with users by simulating physical touch and resistance

(Minamizawa et al.,2008). It makes the experience more realistic through vibrations, damping sensations, and other methods. Compared with the traditional visual interaction and auditory interaction, haptic interaction can make users have a more real sense of immersion, and plays an irreplaceable role in the interaction process. In the traditional human-computer interface, force/haptic interaction is regarded as a special input and output method in the interface interaction. The user is provided with a haptic experience. Force/haptic interaction in virtual reality focuses on natural interaction, which is an important direction of HCI in the future.

Multi-modal interaction in VR integrates various input methods, offering a more immersive and dynamic experience. Multi-modal interaction refers to the collaboration of two or more input channels (such as voice, video, touch, and gesture, etc.) in a system, because it makes full use of different human sensory channels, so that the interaction is more natural and effective. In the multi-channel user interface, users can use natural interaction methods such as voice, gesture, eyes, expression, lip movement, etc., to cooperate with the computer system. Human and machine are regarded as active participants in information exchange, and there are serial/parallel, complementary/independent and other ways between the input channels. Human-computer interaction is closer to the form of human-human interaction, which greatly improves the interaction naturalness and efficiency will be the mainstream form of human-computer interaction in virtual reality in the future.

Recently, with the focus of Metaverse research turning to content communication and social interaction, how to break the bottleneck of the current audio-visual media interaction has become an urgent problem to be solved, and the use of brain-computer interface for sensory simulation is one of the solutions. Brain-Computer Interaction technology reads brainwaves or other neural signals, allowing users to control devices with their thoughts, showing a lot of potential in fields like healthcare and entertainment (Lebedev et al.,2006). The advantages of brain-computer interaction in the Metaverse are mainly embodied in three aspects: extreme immersion, natural interaction, and cognitive enhancement. It allows users to control virtual characters or environments with their minds by directly connecting the brain, freeing them from the limitations of traditional input devices for fast, intuitive interaction. Besides, brain-computer interfaces can simulate or enhance multi-sensory experiences such as sight, hearing, and touch, significantly enhancing the sense of reality and engagement in virtual worlds. Furthermore, brain-computer interfaces can assist learning and cognition, improve memory, concentration, and other abilities, and broaden the potential of metacognition in education, healthcare, and work scenarios. This combination will open up a new era of virtual experiences for users that have never been possible before.

These diverse interaction methods significantly enhance the usability and accessibility of HCI devices in the Metaverse. Multimodal interaction combines gestures, voice, and haptic feedback to give users a richer and more flexible experience. These different interaction methods are driving the development of HCI devices, making them easier to use and more natural, and laying the groundwork for future technologies like the Metaverse.

The Development of HCI Devices

Input Devices

Input devices are fundamental components of HCI, enabling users to convey commands and information to digital systems effectively. Over the years, these devices have seen significant improvements, greatly enhancing the user experience. Voice recognition technology (Klevans, R. L.,& Rodman, R. D., 1997) has also transformed the way users interact. Virtual assistants like Apple Inc.'s Siri allow users to interact with devices through simple voice commands, improving convenience and enhancing accessibility for individuals with disabilities. In addition, gesture control technology utilizes motion sensors and cameras to provide a more immersive experience, allowing users to interact with digital environments through physical movements. This has played a significant role in gaming and virtual reality applications, enabling users to manipulate their environments with natural gestures. With continuous advancements in technology, the accuracy and low latency of input devices will further improve, and more forms of devices will emerge, providing users with a better experience.

In the design of human-computer interaction in the Metaverse, a crucial aspect is selecting a suitable combination of input devices to ensure seamless communication between the user and the system. There are several types of input devices available for different tasks. For example, discrete input devices generate specific events based on user actions, such as keyboards or pinch gloves. Continuous input devices, like 3D tracking systems (Weng, X., Wang, J., Held, D.,& Kitani, K. 2020), follow the user's continuous movements and capture data on orientation and acceleration, including devices like Kinect and Leap Motion. Voice and physiological signal input devices, such as smart speakers like Google Home or Amazon Echo, enable voice interaction, while brainwave input devices capture EEG signals to control systems directly, such as NeuroSky MindWave and Emotiv Insight. Each input method plays a significant role in enhancing the interaction experience in the virtual world.

Output Devices

With continuous technological advancements, the performance of output devices has seen significant improvements. Modern monitors now offer higher resolutions and better color accuracy, along with higher refresh rates and low blue light technology, which greatly reduces eye strain for users. The development of projectors has led to clearer images and more portable designs, making them suitable for a variety of applications. Holographic imaging technology (Benton, S. A.,& Bove Jr, V. M. (2008). Holographic imaging. John Wiley & Sons.) is gradually maturing, providing three-dimensional visuals that significantly enhance user immersion and interaction. Moreover, multimodal output devices are becoming increasingly important; these devices integrate sound, vibration, and haptic feedback, allowing users to interact with systems through multiple senses, thus enhancing the immersive and realistic aspects of the user experience.

In virtual reality systems, output devices play a crucial role in providing sensory feedback to users. These devices mainly target visual, auditory, and tactile senses. Visual display devices, for example, vary based on factors like field of view, resolution, and refresh rate (Yin, K., Hsiang, E. L., Zou, J., Li, Y., Yang, Z., Yang, Q.,...& Wu, S. T. 2022). Common visual displays include HMDs and panoramic monitors. Auditory output devices provide spatial audio, enabling users to locate sound sources and immerse themselves in a 3D auditory environment. Additionally, tactile/haptic feedback devices simulate physical interactions, enhancing the sense of presence in virtual environments. The careful selection of output devices is essential to creating an effective immersive experience.

In the design of HCI devices in the Metaverse, the selection of output/input devices needs to consider the needs of specific interaction technologies, the mutual constraints between input devices and output devices, and the complementarity between multi-channel interactions, etc. In practice, cost is often the biggest factor, and it is also necessary to consider whether the design of interaction technology is limited by a given device, and whether it is necessary to purchase advanced interaction technology, and whether it is necessary to make new interactive devices for interactive technology.

Key Factors of HCI Devices

The current key factors influencing the HCI experience include resolution, refresh rate, clarity, latency, field of view (FOV), and interaction quality (Baker, B. J. (2024). Virtual reality. In Encyclopedia of Sport Management (pp. 1021-1023). Edward Elgar Publishing.).

Model	Resolution	Refresh rate	Release year	Field of view	Weight	Price
Pico 4 Pro	2160× 2160	90Hz	2023	105°	597g	$525.38
Oculus Quest 3	2064× 2208	120Hz	2023	110°	515g	$299
PS VR2	2000× 2040	120Hz	2023	110°	560g	$549.99
Apple Inc. Vision Pro	4K	90Hz	2023	120°	650g	$3,499

Resolution is critical for determining image sharpness, measured by the pixels per degree, with insufficient resolution potentially causing pixelation and reduced immersion. Refresh rate, ideally at 90- 120Hz or above, affects motion smoothness and reduces issues like dizziness or motion blur. Clarity, a more subjective parameter, is influenced by resolution, display technology, optics, and image processing, impacting the visual quality of the virtual environment. Latency refers to the time taken for pixels to change color, affecting motion-to-photon delay, with lower latency enhancing realism and reducing motion sickness. The FOV represents the maximum visible angle of the virtual image, with wider FOVs improving immersion and ensuring a natural viewing experience. Lastly, interaction quality involves the responsiveness and fluidity of user actions within the virtual environment, requiring a well-integrated combination of hardware, software, transmission, and interface design. These factors collectively define the realism, comfort, and interactivity of HCI systems.

Discussion

Overall Development Situations of HCI Devices

Current Status of VR Glasses

Here are some VR glasses produced by companies. This paper summarises their respective parameters, functions and prices, as shown in Table 1.

Table 1. Current status of VR glasses

It can be seen from Table 1 that VR glasses have made significant advancements in hardware performance, particularly in resolution, refresh rates, and field of view, which greatly enhance user immersion (Anthes et al.,2016), but the price is generally at a high level such as the Vision Pro costs a whopping $3,499.

Currently, only the Vision Pro has achieved a single-eye resolution of 4K, while other mainstream VR products generally offer a resolution of 2K per eye, but their combined binocular resolution can reach 4 K. The Vision Pro has a refresh rate of only 90Hz, whereas the Quest 3 achieves 120Hz. The field of view (FOV) for these devices typically ranges from 96 to 120. Most of them support gesture recognition, facial recognition, and eye-tracking. Overall, their weight exceeds 500g, making them relatively heavy. Interaction methods primarily include gesture recognition, facial recognition, and eye-tracking, offering diverse interaction modes and enhanced flexibility. However, many of these technologies are still in the experimental stage, with their stability and cost-effectiveness yet to be improved.

From the perspective of applications, such as the field of education, VR is used to provide immersive learning experiences, allowing students to interactively explore complex subjects. In gaming, VR glasses enable users to feel fully immersed, offering a stronger visual experience and more enjoyment. In enterprise applications, VR is gradually becoming an important tool for virtual meetings, facilitating collaboration in remote environments. Future developments may focus on improving comfort through lighter designs, increasing practicality, and incorporating AI.

Current Status of AR Glasses

In recent years, AR glasses have made significant technological breakthroughs, particularly in display technology and optical systems, making AR more practical in real-world environments (Billinghurst et al.,2015). The current status of AR glasses is shown in Table 2. AR glasses generally have a lower refresh rate, but they are lighter compared to VR headsets.

Table 2. Current status of AR glasses

Brand/model	Resolution	Refresh rate	Release year	Price	Weight
Microsoft HoloLens 2	2048×1080	60Hz	2019	$3,500	566g
Google Glass Enterprise 2	640×360	30Hz	2019	$999	46g
Vuzix Blade	480×480	60Hz	2020	$799	85g

AR glasses have also seen rapid advancements, especially in display technologies and application areas. However, they still face challenges in achieving the same level of visual fidelity and comfort as VR glasses. Despite this, AR glasses are gaining traction in industries like healthcare and education due to their ability to overlay digital information onto the physical world, making them highly useful for professionals and learners alike.

In terms of practical applications, AR glasses demonstrate their unique value across multiple industries. In healthcare, AR glasses are used to assist in surgical procedures, allowing surgeons to view 3D anatomical models and physiological data of patients directly during operations, thereby enhancing precision and safety. In education, students can learn complex concepts by observing and interacting with 3D models, which greatly increases their engagement and understanding. As technology continues to advance, AR glasses will play an important role in training, innovation, and improving productivity.

Analysis of the Global VR Device Shipments

From the current situation of the comic market, Human-computer interaction devices have played a key role in the development of the Metaverse. As these virtual environments expand, the way users interact with them has significantly evolved. Over time, HCI devices have made notable progress both in hardware and software. For example, advancements in VR and AR headsets have improved the way people experience digital worlds. Gesture recognition and voice input systems have also become more accurate, allowing for smoother interactions. Furthermore, brain-computer interfaces have shown great potential for the future, promising more natural and direct forms of communication with virtual systems.

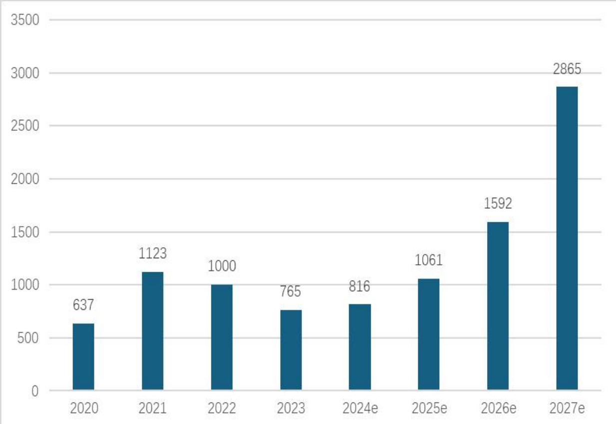

Figure 1. Global VR Device Shipments and Forecasts

Figure 1 summarizes the current sales of VR devices and predicts future shipment volumes. The horizontal axis represents the years, while the vertical axis indicates device sales (in tens of thousands of units). From Figure 1, it can be observed that VR device sales are increasing year by year, primarily due to the rapid advancements in VR technology in recent years and more affordable pricing.

However, despite these improvements, there are still challenges that need to be addressed. One major issue is user experience. While current devices offer more realistic interactions, they often come with limitations such as discomfort during long periods of use or a lack of precision in detecting certain gestures. Another challenge is compatibility. As the Metaverse continues to grow, users may encounter difficulties using devices across different platforms or systems. This lack of standardization makes it harder for users to fully immerse themselves in various virtual environments without experiencing technical issues. Lastly, cost remains a barrier. Many advanced HCI devices are expensive, preventing broader access and slowing the adoption of Metaverse technologies.

In conclusion, while HCI devices have made important strides in enhancing the virtual experience, there is still room for improvement in areas like user comfort, device compatibility, and affordability. Addressing these challenges is essential for the future of the Metaverse, ensuring that it becomes a fully accessible and seamless space for everyone.

Challenges

The above summarizes the current state of research on HCI devices in the Metaverse. Advancements in graphics and display technologies, coupled with the evolution of helmet-based devices and the integration of the Metaverse with the Internet and artificial intelligence, have transformed the Metaverse into a novel personal computing platform. This shift has spurred interest in developing more efficient and user-friendly HCI technologies to meet the diverse demands of virtual reality applications.

Despite these advancements, HCI in the Metaverse still faces significant challenges. One major challenge is accurately capturing and understanding users' natural interaction behaviors and state perceptions in virtual environments. While sensors can collect diverse data—such as appearance, depth, brain signals, physiological and biochemical parameters, and speech—these measurements are often limited to explicit representations of human behaviors or states. Developing methods to accurately and in real-time extract, characterize, and depict these behaviors remains a core issue for interaction perception in the Metaverse.

Another challenge is understanding users' interactive behaviors and intentions within virtual reality. While machine learning, pattern recognition, and similar technologies have enabled partial and simplified interpretations of user intent, significant uncertainties persist in modeling complex action intentions and state transitions. Addressing these uncertainties requires a deeper understanding of human cognitive processing and the abstraction of computational models to accurately interpret interactive intentions, a critical step toward achieving effective virtual reality interaction cognition.

Furthermore, the most significant challenges in the development of human-computer interaction devices for the Metaverse are the current hardware and technological limitations. Devices such as VR and AR headsets have undoubtedly made progress, but they remain far from perfect. The bulkiness of these devices makes them uncomfortable for extended use, and the high costs associated with advanced models are a major barrier to widespread consumer adoption. Moreover, the limited battery life of these devices further restricts their usability, especially when users aim for longer, uninterrupted experiences in the Metaverse. Another challenge is that many of these devices rely heavily on external computing power, which reduces their portability and the potential for a seamless, fully immersive environment. Furthermore, while brain-computer interaction technology is often hailed as the future of HCI, it is still in its early stages. Current brain-computer interaction devices struggle with accuracy and consistency in translating neural signals into actionable commands. This raises concerns about the reliability and precision needed for real-time interactions within the Metaverse. Alongside this, there are growing concerns about the safety and ethical implications of devices that interact directly with the brain. In addition, the challenge of ensuring that these devices are accessible and affordable to a broad population remains, particularly as the digital divide becomes more pronounced. These limitations in both hardware and broader technological infrastructure pose considerable barriers to achieving the vision of a fully integrated, immersive Metaverse. Therefore, in the future, I expect human-computer interaction to make breakthrough progress in the following aspects:

Basic research on human-computer interaction in the Metaverse. Focusing on the study of human perception mechanisms and cognition-based natural human-computer interaction theories and methods, including the study of human multi-sensory (sight, hearing, touch, force, smell, taste) channels' perception mechanism of VR, AR, and other generated environments.

Development of new sensing devices, providing basic theoretical support for the development and research of HCI devices as well as human-computer interaction methods. Support research on human physiological response mechanism, cognitive process and emotional mechanism in Metaverse environment, study multi-channel feedback presentation and adaptive adjustment methods, and build physiological computation theory.

Conclusion

Thanks to the rising popularity of the Metaverse concept, related HCI devices have garnered widespread attention and research. This paper first reviews the definitions and concepts related to the Metaverse and HCI devices, outlining the methods and current state of research, including some basic devices used in this field. Then, based on the key AR and VR devices employed in the Metaverse, this paper discusses the current state of the industry. Through analysis, the application of HCI equipment in the Metaverse still faces significant hardware and technical challenges. In my view, the key issue to address lies in the development of interaction methods. The ultimate goal of HCI devices should be to directly translate thoughts from the brain into the Metaverse, eliminating the need for complex interactive devices and enabling seamless access to the Metaverse anytime, anywhere.

While current HCI technology has made some progress, it remains in its early stages and is still far from achieving this ultimate vision. The current focus is on improving the accuracy and responsiveness of input devices, such as eye-tracking, motion sensing, and gesture recognition. These technologies are critical for creating more natural and intuitive user interfaces. However, as the Metaverse strives for a fully immersive and user-friendly experience, it is clear that the next frontier lies in the seamless integration of brain-computer interfaces to enable thought-based interaction.

Despite these challenges, the Metaverse represents a transformative concept for the future of society and civilization, offering vast potential for development and innovation. As both hardware and software continue to evolve, the promise of a fully realized Metaverse — where users can interact through their thoughts and seamlessly navigate virtual worlds — becomes more achievable. The intersection of neuroscience, AI, and immersive technologies holds the key to shaping the next generation of HCI devices in the Metaverse.

Conflict of Interests: the author has claimed that no conflict of interest exists.

References

1. Baker, B. J. (2024). Virtual reality. In Encyclopedia of Sport Management (pp. 1021-1023). Edward Elgar Publishing
2. Mystakidis, S. (2022). Metaverse. Encyclopedia, 2 (1), 486-497.
3. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. (1994). Human-computer interaction. Addison-Wesley Longman Ltd.
4. Lyu, Z. (2023). State-of-the-art human-computer-interaction in Metaverse. International Journal of Human –Computer Interaction, 1-19.
5. Zhang Fengjun, Dai Guo-zhong, & Peng Xiao-lan. (2016). A Survey of human-computer interaction in Virtual reality. Science China: Information Science, 046 (012), P.1711-1736.
6. Weng, X., Wang, J., Held, D. & Kitani, K. (2020, October). 3d multi-object tracking: A baseline and new evaluation metrics. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 10359-10366). IEEE.

7. Yin, K., Hsiang, E. L., Zou, J., Li, Y., Yang, Z., Yang, Q. & Wu, S. T. (2022). Advanced liquid crystal devices for augmented reality and virtual reality displays: principles and applications. Light: Science & Applications, 11 (1), 161.
8. Benton, S. A. & Bove Jr, V. M. (2008). Holographic imaging. John Wiley & Sons.
9. Klevans, R. L. & Rodman, R. D. (1997). Voice recognition. Artech House, Inc.