The Generative AI Revolution in Immersive Technologies

Arlindo R. G. Filho*
Rodrigo Z. Fanucchi*
arlindogalvao@ufg.br
rodrigo_fanucchi@ufg.br
Advanced Knowledge Center for Immersive Technologies/ AKCIT
Universidade Federal de Goiás
Goiânia, Goiás, Brasil

Ingrid Winkler
SENAI CIMATEC University Centre;
Advanced Knowledge Center for Immersive Technologies/ AKCIT
Salvador, Bahia, Brasil
ingrid.winkler@doc.senaicimatec.edu.br

Alexandre G. de Siqueira
Universidade da Flórida
Gainesville, Estados Unidos da América

Abstract

Generative Artificial Intelligence has revolutionized various aspects of the modern world, ranging from individual users leveraging tools to answer queries to companies developing advanced chatbots. However, it may still be premature to assert that all the potential applications of Generative AI have been fully explored. Among these unexplored possibilities, the integration with immersive technologies, such as virtual reality, augmented reality, and mixed reality, remains a domain with substantial room for development. This paper explores the opportunities related to Generative AI in the context of immersive technologies, as well as the challenges associated with its broader adoption in virtual worlds.

1 Introduction

Researchers define immersive technologies from various perspectives, one of which is the concept of the *reality-virtuality continuum*. According to this concept, immersive technologies are positioned at the intersections between the physical and virtual worlds, with the goal of enhancing the realism of experiences [6].

Since the popularization of the Internet in the 1990s, the concept of immersion in virtual worlds has significantly evolved. Several well-established virtual environments have emerged, such as software applications, mobile apps, social networks, videoconferencing platforms, 2D and 3D virtual worlds, and augmented reality applications, among others. Although these virtual environments are not perpetual and often disconnected, they have driven varying degrees of digital transformation, allowing the immersion of different actors and experiences in virtual worlds throughout different stages of development.

With the rise of Generative Artificial Intelligence (AI), this evolution has become even more apparent through the multimodality of models, which bridge previously isolated domains—such as the integration of video with text or text with images, among other possibilities. The conjunction of extended reality, blockchain, and Generative AI has redefined human interaction with both the physical and digital worlds [3], particularly through its ability to personalize user preferences and experiences based on individual needs.

Moreover, virtual reality has contributed to achieving several United Nations Sustainable Development Goals (SDGs), including SDG 4 (inclusive and equitable education for all), SDG 3 (Health and Well-being), SDG 13 (Climate Action), and SDG 12 (Sustainable Consumption and Production) [1]. By enabling more sustainable and intuitive immersive technologies, Generative AI thus plays a role in fostering a more sustainable society overall.

To further elucidate how the revolution of Generative AI can enhance immersive technologies, the following sections will explore the opportunities and challenges associated with this intersection.

2 Opportunities

Although there is not yet a consensus on a definitive framework, some key technologies are reasonably well-positioned for the construction and conception of a new generation of immersive technologies for virtual worlds, with Artificial Intelligence possibly being the most prominent among them.

Gartner's 2022 Hype Cycle examined the evolution and expansion of immersive experiences through the use of AI. The report highlights decentralized identity, digital humans, customer digital twins, non-fungible tokens (NFTs), superapps, and Web3 as key thematic areas.

More specifically, [4] analyze, across various scenarios, the impacts of Generative AI in the context of virtual worlds. In general terms, the study provides a broad view of the innovation process within virtual worlds, concluding that Generative AI holds significant opportunities related to the development of applications for human-machine and machine-machine interactions. However, issues such as scalability—primarily due to the still limited dissemination of extended reality devices—are highlighted by the authors as factors that need to be addressed over time.

In this sense, it can be considered—through an analysis similar to that in [4]—that scalability in immersive technologies will likely be achieved when the democratization of digital human creation occurs, supported by easy-to-use applications for non-programmers. Applications with intelligent composition tools would allow anyone to create digital humans with human-like expressions on devices accessible to the general public.

The challenge in developing immersive technologies that enable this goal lies in the combination of physical and digital existence. It is believed that immersion will involve a decentralized, scalable, and autonomous iteration, allowing users to connect to a unified existence through physically decentralized devices that cooperate in the digital environment. Ideally, it is expected that the new generation of immersive technologies will be interactive, intelligent, intuitive, digitally enhanced, and either artificially created or augmented in real-time by Artificial Intelligence.

The use of AI, particularly deep learning, has significantly boosted productivity in the development of immersive solutions for virtual

*Both authors contributed equally to this research.

worlds. Until a few years ago, AI tasks were primarily focused on automating data analysis, but a new generation of architectures has enabled the development of AI-driven content creation.

From the perspective of desirable products, the main scientific and technological opportunities and challenges that are likely to remain unresolved over the next decade can be identified [2] [7] [8], and their advancements may significantly transform immersive technologies:

- Presence and User Experience: Immersive technologies are designed to enhance the user experience [8]. What sets immersive systems apart from other technologies is the realism of the experience. Understanding how the system and its characteristics influence presence is complex, given the various parameters, individual differences, and the diffuse nature of the experience. Digital humans should interact, learn, and express themselves in ways that resemble human behavior. A review of the main methods proposed for developing user presence immersion, which analyzed 1,214 articles published over the past 30 years, identified 29 key factors that evoke presence in virtual environments, categorized into four groups: engagement, personal characteristics, interaction fidelity, and display fidelity. This same study, for instance, highlights that one of the major limitations lies in the use of subjective metrics to assess the user's sense of presence [8].
- Character Creation and Narrative: The creation and evolution of characters in immersive environments is not a recent development. However, the task still faces challenges in the scalability of character construction. The current process, based on coding in computational frameworks, is still slow and relatively costly for potential mass adoption. Immersion is partial and dependent on input generated from an external event. Creating more scalable immersive content remains an open challenge. In virtual worlds, character development is a central part of immersive experiences. Natural interaction will likely involve techniques beyond point-and-click, such as dialogue generation, speech synthesis, and automated animation. Computing and engineering often draw inspiration from biological processes, and in this case, the search for imitating human behavior through a digital human twin emerges as a competitive hypothesis. Some unresolved challenges include the creation of intelligent agents capable of emotionally connecting with users and the dynamic adaptation of narratives in real time.
- Multimodality of Models: The ability of Large Language Models (LLM) to interact with various formats—voice, text, photo, video—enables the development of new forms of immersion, including fostering greater social inclusion.
- Automated Content Generation: A key point is the availability of content creation tools. Scientific and technological advancements have the potential to reduce production costs and allow access to non-specialists [5], which used a generative neural network model to accelerate the construction of scanpaths, a sequence of two-dimensional values corresponding to the coordinates of each point of human gaze on an image. For realistic experiences or the depiction of real places, tools that quickly capture geometric and material information from real scenes will be essential.

3 Challenges

The development, training, and validation of large Generative AI models rely on specific computational architectures, namely Graphics Processing Units (GPUs). This is due to the high capacity of these architectures for performing mathematical processes at high speeds. However, as a side effect, there is an extremely high energy consumption, which directly impacts the sustainability of virtual worlds. Therefore, sustainability in the relationship between Generative AI and immersive technologies is of utmost importance for the feasibility of their widespread use.

Additionally, another important point related to the integration of Generative AI and immersive technologies is the need to create security barriers—whether during training or in applications—to ensure that no discriminatory bias is present in the developed models [9]. This concern has been a significant issue guiding the development of UNESCO's guidelines [9].

As for Generative AI models, there are still problems to be solved to enable broader use. Specifically considering LLMs, these models are based on a probabilistic approach to defining the sequence of tokens in relation to a given context or question. Therefore, incorrect responses and hallucinations, as they are called, remain significant challenges in real-time applications, even with the use of techniques like Retrieval Augmented Generation (RAG) or Knowledge Graphs.

4 Final Considerations

This paper explored how Generative AI is related to immersive technologies in the current context. Furthermore, it provided a detailed account of the opportunities associated with this interaction, as well as the challenges that need to be addressed. Ultimately, it is evident that the Generative AI revolution is not an end in itself. There is vast potential still to be explored, and improving the immersion experience, with human-machine interactions that are closer to reality, is one of the most crucial points in our current digital reality.

Acknowledgments

This work has been fully supported by Advanced Knowledge Center of Immersive Technologies (AKCIT) with financial resources from the PPI IoT/Manufatura 4.0 / PPI HardwareBR of the MCTI grant number 057/2023, signed with EMBRAPII. The authors also thank the SENAI CIMATEC University Centre, the Bahia State Research Support Foundation (FAPESB) and the National Council for Scientific and Technological Development; IW is a DT-1D productivity grant fellow (308783/2020-4).

References

[1] Iolanda L. Chamusca, Yiyu Cai, Pedro M. C. Silva, Cristiano V. Ferreira, Thiago B. Murari, Antonio L. Apolinario, and Ingrid Winkler. 2024. Evaluating Design Guidelines for Intuitive, Therefore Sustainable, Virtual Reality Authoring Tools. *Sustainability* 16, 5 (2024). https://doi.org/10.3390/su16051744

[2] Yogesh K Dwivedi, Laurie Hughes, Elvira Ismagilova, Gert Aarts, Crispin Coombs, Tom Crick, Yanqing Duan, Rohita Dwivedi, John Edwards, Aled Eirug, et al. 2021. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International journal of information management* 57 (2021), 101994.

[3] Farhang M Hamzah. 2024. The Future of the Digital Social Economy: Navigating the Confluence of Blockchain, Metaverse, and Artificial General Intelligence. *Journal of Science and Technology* 29, 1 (2024), 31–36.
[4] Jussi S Jauhiainen. 2024. The Metaverse: Innovations and generative AI. *International Journal of Innovation Studies* 8, 3 (2024), 262–272.
[5] Daniel Martin, Ana Serrano, Alexander W Bergman, Gordon Wetzstein, and Belen Masia. 2022. Scangan360: A generative model of realistic scanpaths for 360 images. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (2022), 2003–2013.
[6] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.
[7] Sang-Min Park and Young-Gab Kim. 2022. A metaverse: Taxonomy, components, applications, and open challenges. *IEEE access* 10 (2022), 4209–4251.
[8] Vinicius Souza, Anderson Maciel, Luciana Nedel, and Regis Kopper. 2021. Measuring presence in virtual environments: A survey. *ACM Computing Surveys (CSUR)* 54, 8 (2021), 1–37.
[9] Bernd Carsten Stahl, Doris Schroeder, and Rowena Rodrigues. 2022. The ethics of artificial intelligence: An introduction. In *Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges*. Springer, 1–7.