Exploring Human-AI Collaboration Continuum in Augmented Reality Applications

Shokoufeh, SB, Bozorgmehrian, Virginia Tech, shbozorgmehrian@vt.edu

Joseph, JG, Gabbard, Virginia Tech, jgabbard@vt.edu

This position paper delves into the realm of human-AI collaboration within the context of AI, Generative (Gen) AI, and augmented reality (AR) user experiences. It investigates the evolving role of AI from problem-solving to problem-finding, particularly focusing on the definition of different scenarios in human-AI AR-based interactions. By exploring the intersection of AI, generative AI, AR, and human centered design, this paper aims to highlight a continuum in which we can define sectors of human-AI interactions in Augmented Reality, from development to interfaces and interaction considerations.

CCS CONCEPTS • Augmented Reality • Generative AI • Artificial Intelligence • Human-AI Collaboration

ACM Reference Format:

Shokoufeh, SB, Bozorgmehrian, Joseph, JG, Gabbard. 2024. Exploring Human-AI Collaboration Continuum in Augmented Reality Applications. In CHI 2023 Workshop on Generative AI and HCI), May 11, 2024, Hybrid. ACM, Honolulu, HI, USA, 4 page

1 INTRODUCTION

The integration of artificial intelligence (AI) has various impacts in augmented reality (AR) applications. AI in AR has transformed human-AI collaboration and co-creative systems in many areas [1]. As a result, AR applications now have far more potential to support and create new, immersive, and interactive experiences. Furthermore, with the introduction of Generative AI—such as those that make use of Large Language Models (LLMs) and Generative Adversarial Networks (GANs)--new, plausible media, such as 3D models, pictures, music, text, design, and motion can be integrated into AR more easily, both during design and in real-time [2]. Additionally, the use of generative AI has reinterpreted AI's function as a human support system rather than just a decision-maker, highlighting the significance of collaborative interactions between agents of generative AI and humans [3] as well as co-creative applications [4]. By studying the exemplar applications of AI in AR and scenarios, we propose a continuum of AI engagement in user experience in Augmented Reality (AR) applications categorizing the role of AI/GenAI from behind-the-scenes supporter to an explicit assistant agent, aiming to guide further research and development in human-AI collaboration in AR applications.

2 AI, GENERATIVE AI AND COLLABORATION IN THE AUGMENTED REALITY INTERFACES

The role of AI as a human assistant in Augmented Reality (AR) experiences has been a subject of extensive research. Recent research, for example, has examined how experiential augmented reality applications affect consumers' intentions on fashion purchases. The study used the stimulus-organism-response model to investigate how AR affects consumers'

GenAICHI: CHI 2024 Workshop on Generative AI and HCI

affective and behavioral responses, underscoring the changing role of artificial intelligence in influencing consumer experiences in AR environments [5]. Additionally, the development of an AI ecosystem for aerospace AR applications has highlighted how AI can be used to introduce dynamic approaches for AR-based instruction modules through predictive machine learning algorithms, highlighting how it is changing the way humans and AI interact in aerospace applications [6]. Furthermore, the combination of artificial intelligence with augmented reality has impacted robotics applications by facilitating intuitive control and feedback, clarifying the purpose of robot motion, and redefining human-robot interactions [1].

Because of their increased usability and flexibility, AI-based generative models are expected to be used more frequently in a variety of applications, such as digital art production and automatic code generation [7]. That makes them a perfect tool to be integrated into Augmented Reality platforms as well. Generated content by either text-based models for conversational chatbots, diffusion models for image generation, or other areas can perform levels of assistance and collaboration with humans as the developer or the user [8, 9, 10].

In the next section, we describe a proposed continuum that describes the potential roles that AI or generative AI can play in this field. By reviewing recent practices in the field, we look into the different types of Human-AI Collaboration in Augmented Reality (AR) field.

3. DIFFERENT TYPES OF HUMAN-AI COLLABORATION IN AR CATEGORIZATION—PROPOSED CONTINUUM OF AI ENGAGEMENT IN USER EXPERIENCE

In this section, we move further in defining the human-AI collaboration categorization in the context of Augmented Reality applications. We look deeper into the three components of AI/GenAI models, AR Applications, and the user. We aim to define the major areas in which AI/GenAI presents itself to the user in the context of AR hence specifying the configurations in which these three components are shaping the overall system.

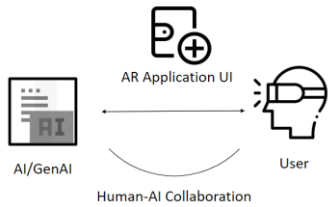

Figure 1: Three Components of the Human-AI Collaboration in Augmented Reality Context

3.1- In behind the Scenes: Ideation, Desing Inspiration, and Development Phases:

A well-developed Augmented Reality application—like any other product design and development—relies heavily on ideation, creativity, efficiency, and effectiveness of the design and development process. AI, including Generative AI, can play a pivotal role in augmenting these processes by collaborating on offering solutions and insights, inspiring design concepts and implementation processes. In this category, AI is quiet, it is just a reinforcing tool or supporter of development process for anything that happens behind what is presented to the user. AI helps designers and developers to quickly and efficiently explore a wide range of ideas by automating the brainstorming phase. Moreover, with advancements in text-based image generation models like DALL-E, Mid Journey, Google Gemini, they are also able to move from ideation to rapid visualization of corresponding interface designs based on their top selected ideas. Later in the process, the assistant tools can serve as a collaborator in code development process by speeding up the code generation process, optimizing the code, debugging, and improving overall efficiency.

For further description of this end side of the continuum, we look at the application of AI or Generative AI here as tool for any sort of product design as it is discussed by Hong, et al [11]. They explain the advantages and disadvantages of using Generative AI models as an assistant tool for inspiration, idea diversity, and high-fidelity design.

3.2- As a Method Supporting the Adaptive Intelligent AR (IAR) System:

Here, the AI's potential is integrated in the AR application design toward creating an Intelligent Augmented Reality (IAR) platform that can adapt themselves to the user's choice of action and environment in real-time. The designed AR application includes context-aware computation and is actively driving the application. It is important to mention that in our definition of this sector, the AI's engagement with the user's task is passive and trained intentionally by the designers and developers for the desired task and environment. In this category, AI is quiet. It does not show itself to the user explicitly but supports the AR system to be adaptive to some extent based on the user's interaction and the environment.

As we move toward the middle of this continuum, we notice intelligent interactions that Augmented reality applications begin to gain. For example, Augmented Reality interface design for bridge inspection project by Smith, et al specifically concentrates on developing a prototype Augmented Reality (AR) framework capable of addressing the features of Geo-located Annotations and Automated/Semi-Automated Defect Labeling. In this sense, by using computer vison algorithms, they are providing a framework that understands and adapts the interface based on the environment [12]. Another Example here can be a platform developed for assembly process in Augmented Reality mentioned by Sabin, et al [13]. Their suggestion involves the use of cutting-edge deep learning methods for object detection, along with a regression-based mapping approach to determine the 3D positions of assembly components. After the automatic identification of machine parts, a multimodal interface incorporating eye gaze and hand tracking was introduced to help with the manual assembly process. Specifically, we suggested employing an eye cursor to lead the user through the task, while also utilizing fingertip distances and object sizes to identify any errors made during the process. Many works have been done covering this area of transition toward intelligent Augmented Reality environments. Another example that touches this approach for a helping the visually impaired users is explained by Lee, et al, incorporating real-time computer vision (CV) to improve the visibility of tennis balls [14].

3.3- AI Presented in a Form of an Assistant Agent to the User in Augmented Reality Environment

Using Chat GPT [15] for answering the questions of a homework, asking Gemini [16] to explain a diagram quickly while searching on Google or asking Copilot [17] for code completion or suggestion are samples of having an agent on the side for helping one with a specific task that they are trying to complete. They, and many other assistant tools need to be presented to the user in a shape and form and need to have methods of interaction designed to get the human's prompt input. 2D displays are currently being used and they are presenting themselves in different graphical designs. In this category, AI is explicit, and the user is actively engaged with it. It presents itself to the user and Augmented Reality interfaces and related interaction techniques can be used as an environment that we provide for humans to interact with the Generative AI assistant agents. We are thinking of the methods and the graphics that AI is presenting to the user and interacts with them embedded on an Intelligent Augmented Reality (IAR) platform. It can be in the shape of a chatbot appearing on the screen, understanding the context and answering the user's prompts in text, voice, or graphical visualization format appearing on the screen, understanding the human voice input and providing human-like voice instruction, or having the shape of a human-like avatar with the sound and corresponding gestures.

Having said that, we can categorize this form of Assistant Agents design in the IAR platforms with three major considerations: 1) What is the modality of interaction that the system is designed to get the human input as the GenAI's input prompt; 2) What is the type of graphics that is presented on the screen that represents this agent; 3) What is the modality of interaction that the system is designed to provide the outputs as the answers for human inquiry. To further look into this approach in the state of the art, Zhao, et al, describe a project based on Image generation for in-situ fashion design in Augmented Reality. The user interface offers various functions such as adjusting landmarks, sketching, erasing, cleaning, generating images, and providing assistance. Users have the flexibility to embellish either a reference image or a real human model through the interface, with their drawing inputs overlaying on the canvas alongside the chosen image or live video stream [18].

Another example in this area is the study done by Jones, et al, in the area of using generative AI in Augmented Reality environment for action recommendation. In this study, the focus lies on harnessing the potential of Generative AI models, particularly large language models (LLMs), to offer recommendations aligned with individuals' overarching objectives. These recommendations are tailored to fit seamlessly within relevant contexts, such as the appropriate time, place, or alongside relevant tools. The research revealed that users appreciate the passive nature of these AI-generated suggestions, recognizing their capacity to inform, motivate, and inspire [19]. Moving along through this continuum, we can mention AI-Generated content exploration in Augmented Reality application published by Hu, et al. focusing on the automatic generation of media content, including images and text, through AI algorithms. Recognizing the increasing quality of AI-generated content (AIGC), the study aims to explore the possibilities of incorporating such content into AR displays [20].

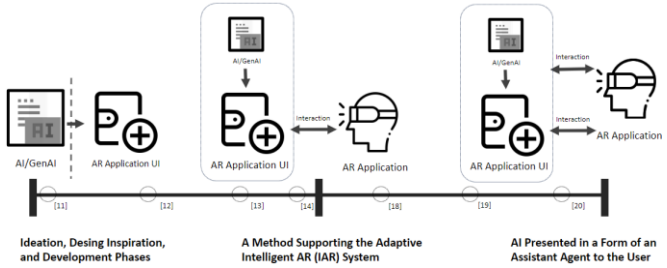

Figure 2: Continuum of AI/GenAI Engagement in User Experience in the Context of Augmented Reality

1. Hong et al. (2023): Focus on Generative AI for product design. 2. Smith et al. (2022): Wearable AR interface design for bridge inspection. 3. Raj et al. (2024): AR and deep learning for assembly process assistance. 4. Lee et al. (2023): Real-time CV and AR for low vision sports support. 5. Zhao & Ma (2018): Two-stage image generation for AR fashion design collaboration. 6. Jones et al.: Generative AI for situated action recommendations in AR. 7. Hu et al. (2023): Exploration of AI-generated content for AR display.

The proposed continuum can serve as a ground for further study in the area of AI engagement with the user within Augmented Reality (AR) applications realm. It underscores the critical roles that AI/GenAI can have across various stages of AR development and user interaction from behind-the-scenes supporter to an explicit, actively engaging agent.

4. DISCUSSION

In the proposed continuum, by categorizing AI engagement into three main areas—behind-the-scenes ideation, as a method supporting adaptive intelligent AR systems, and as assistant agents we aimed to define visual graphics of different roles that AI or GenAI can have when they are integrated into the process. It will generate a mind map for researchers and developers to define where they want to stand in designing the human-AI collaboration in the Augmented Reality environments regarding the experience they will have or design for the users.

3 5. CONCLUSION

With a growing focus on applications of AI and Generative AI in the field of applications development and user engagement, we believe it is crucial to know where we stand when we talk about human-AI collaboration and what type of user engagement we are dealing with. In this regard, this paper specifically looks into this topic in the field of Augmented Reality (AR) application and defines a continuum that explains the featured human-AI collaboration categorization. It delineates three key areas of AI engagement—behind-the-scenes ideation, as a method supporting adaptive intelligent AR systems, and as assistant agents. We have demonstrated the transformative impact of generative AI on AR applications and the importance of considering human-AI interactions in the design and development process through a review of relevant literature and examples from recent studies.

REFERENCES

[1] Bassyouni, Z., & Elhajj, I. H. (2021). Augmented reality meets artificial intelligence in robotics: a systematic review. Frontiers in Robotics and AI, 8. https://doi.org/10.3389/frobt.2021.724798

[2] Antol, S., Agrawal, A., Lü, J., Mitchell, M., Batra, D., Zitnick, C., ... & Parikh, D. (2015). Vqa: visual question answering. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 279). https://doi.org/10.1109/iccv.2015.279

[3] Seeber, I., Bittner, E., Briggs, R. O., De Vreede, T., De Vreede, G. J., Elkins, A., ... & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. Information & Management, 57(2), 103174.

[4] Zhou, Y., Koyama, Y., Goto, M., & Igarashi, T. (2020). Generative melody composition with human-in-the-loop Bayesian optimization. arXiv preprint arXiv:2010.03190.

[5] Watson, A., Alexander, B., & Salavati, L. (2018). The impact of experiential augmented reality applications on fashion purchase intention. International Journal of Retail & Distribution Management, 48(5), 433-451. https://doi.org/10.1108/ijrdm-06-2017-0117

[6] Venkatesh, V. (2021). Development of an AI ecosystem for AR applications within aerospace. Retrieved from https://doi.org/10.32920/ryerson.14641182

[7] Hu, Y., Yuan, M., Xian, K., Elvitigala, D. S., & Quigley, A. (2023). Exploring the Design Space of Employing AI-Generated Content for Augmented Reality Display. arXiv preprint arXiv:2303.16593.

[8] Dale, R. (2021). GPT-3: What's it good for? Natural Language Engineering, 27(1), 113–118. https://doi.org/10.1017/S1351324920000347

[9] Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 30(4), 681–694. https://doi.org/10.1007/s11023-020-09541-0

[10] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10684–10695).

[11] Hong, M. K., Hakimi, S., Chen, Y. Y., Toyoda, H., Wu, C., & Klenk, M. (2023). Generative AI for Product Design: Getting the Right Design and the Design Right. arXiv preprint arXiv:2306.01217.

[12] Smith, A., Duff, C., Sarlo, R., & Gabbard, J. L. (2022). Wearable Augmented Reality Interface Design for Bridge Inspection. In 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (pp. 497-501). https://doi.org/10.1109/VRW55335.2022.00111.

[13] Raj, S., Murthy, L. R. D., Shanmugam, T. A., et al. (2024). Augmented reality and deep learning based system for assisting assembly process. Journal of Multimodal User Interfaces, 18, 119–133. https://doi-org.ezproxy.lib.vt.edu/10.1007/s12193-023-00428-3

[14] Lee, J., Sarda, D. P., Lee, E., Lee, A., Wang, J., Rodriguez, A., & Froehlich, J. E. (2023, October). Towards Real-time Computer Vision and Augmented Reality to Support Low Vision Sports: A Demonstration of ARTennis. In Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (pp. 1-3).

[15] OpenAI. (n.d.). ChatGPT. Retrieved from https://openai.com/chatgpt

[16] Google. (n.d.). Gemini. Retrieved from https://gemini.google.com/app

[17] Microsoft. (n.d.). Copilot. Retrieved from https://copilot.microsoft.com/

[18] Zhao, Z., & Ma, X. (2018, December). A compensation method of two-stage image generation for human-ai collaborated in-situ fashion design in augmented reality environment. In 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR) (pp. 76-83). IEEE.

[19] Jones, B., Xu, Y., Hood, M. A., & Kader, M. S., Eghbalzadeh, H. Using Generative AI to Produce Situated Action Recommendations in Augmented Reality for High-Level Goals.

[20] Hu, Y., Yuan, M., Xian, K., Elvitigala, D. S., & Quigley, A. (2023). Exploring the Design Space of Employing AI-Generated Content for Augmented Reality Display. arXiv preprint arXiv:2303.16593.