Proceedings of CONF-MLA 2024 Workshop: Securing the Future: Empowering Cyber Defense with Machine Learning and Deep Learning
DOI: 10.54254/2755-2721/110/2024MELB0098

The Application and Practice of Artificial Intelligence in the Entertainment Field

Yuang Tian

Faculty of Animation, Arts & Design, Sheridan College, Oakville, ON L6H 2L1 Canada

Tiayuang@sheridancollege.ca

Abstract. Artificial intelligence (AI) technology has witnessed unprecedented advancements and a gradual penetration into civilian applications. This paper aims to thoroughly investigate the application of AI in the entertainment industry, with a particular focus on the principles and cross-disciplinary implementations of 3D real-life scanning, AI for non-player characters (NPCs), and AI video generation. By synthesizing how these technologies streamline content creation processes, lower technical barriers, and inspire novel approaches to game design, we observe that AI is not only reshaping the ecosystem of the entertainment sector but also facilitating the entry of newcomers into game development. However, alongside the benefits, this study identifies several challenges and limitations associated with current AI technologies, such as accuracy, cost-effectiveness, and ethical concerns, which require attention and resolution in future research and practice. Through a detailed examination and synthesis of these phenomena, this research provides a reference for practitioners and suggests directions for subsequent studies.

Keywords: Artificial intelligence, 3D scanning, game design, 3D modeling, NPC.

1. Introduction

Nowadays, artificial intelligence is increasingly involved in various fields, whether it is art modelling or video production. These new technologies can not only improve developer efficiency and lower the threshold for entertainment design but also bring players a more authentic and diverse experience. In addition, players' demands for games are becoming increasingly high, and designers need to use constantly developing technologies to meet the diverse needs of players. Through real-time scanning or NPC binding with GPT, they can create a dazzling gaming experience for players. Designers can convert real-life buildings, items, or furniture into 3D models and import them into game engines for secondary editing [1]. They can use these scattered elements to piece together a brand new map, or they can replicate existing objects in reality through 3D printing. NPCs that have imported GPT can also make different responses based on the player's behaviour, sometimes even exceeding the developer's expectations, providing players with a different experience than before. AI video generation has also brought new blood to the self-media community. Short videos generated by AI are flourishing on video websites, and operators almost do not need knowledge related to video production to create a video with realistic effects in a short period.

DOI: 10.54254/2755-2721/110/2024MELB0098

This article will explore the use of artificial intelligence in various fields such as game design, video production, and cultural heritage preservation, and introduce the principles of these technologies and their relationships with other industries.

2. 3D real scene scanning and its application in game design neighborhood

In the past two years, a large number of games have used 3D reality scanning in the production process. Designers scan vehicles, cultural relics, or buildings in reality, convert them into 3D models, and then modify textures and improve model loopholes to achieve a level that can be used as game engine materials. 3D real scene scanning is a technology that generates 3D models through multiple photos. In this process, artificial intelligence processes data and optimizes point cloud images to enhance the quality of the model [1]. 3D reality scanning is commonly used in exhibitions to protect 3D printing and restore cultural heritage [2]. The most commonly used measurement method is based on motion structure (SfM) to create 3D models. Users need to take photos of the modelled object from multiple angles, and then professional software will be responsible for creating the 3D model and texture mapping. Operators usually install cameras on airplanes, drones, or satellites, and then use the cameras to take photos of objects for 3D measurement [2]. This technology can be applied to games, cultural relic preservation, house sales, and VR exhibitions.

When users import many images into the generation software, artificial intelligence will calculate and scan the angles and shooting positions of these images, and generate sparse point clouds by identifying the features of these photos. After further processing the sparse point cloud, more detailed dense point clouds can be generated, making the model more refined. Finally, artificial intelligence will construct a 3D mesh (usually triangles or parallelograms, which are the small flat shapes that make up the model we see) from the point cloud, and then map the photo's texture onto the model to complete it [3].

The scans described above are based on large relics, but students and professors start with small artifacts at Compton University. In 2021-2022, Compton's University conducted an "Experimenting Digital Antiquity Project" activity. Although students used replicas, the study aimed to preserve relevant information before cultural relics were damaged through scanning. Students must take 20 to 80 photos of cultural relics, import them into Agisoft Metashape, use software such as Blender to correct the details in the 3D model, and finally 3D print them [4].

Similar to the above case, in the game Star Wars Battlefront, many of the items players use are made by 3D real-life scanning. The designers scanned a large number of props used in the movie and converted them into 3D game materials. The production team stated that this method not only has a low cost but also saves a lot of time [5]. At EA SPORTS In the PGA Tour, designers use LiDAR (3D laser) drones to record the terrain of each course, process impurities through programming software and then import them into the Frostbite editor (a terrain compiler). The designer did not simply replicate the terrain of the stadium and use it but rather used it as material for secondary creation [6].

3. The possibility of AI as a game NPC

Nowadays, in some games like Minecraft, it is not uncommon to import large language models such as GPT into NPCs (non-player characters) or even player accounts. Therefore, is there a possibility that future games will extensively use artificial intelligence to replace monotonous procedural NPCs and create a more realistic and diverse gaming environment for players?

Usually, NPC AI within games can be divided into two types: behavioural logic and path logic. Behavioural logic is used to tell NPCs what to do, and AI can detect various items in the game and then use behavioural logic to determine where they are going. The most classic path logic is Navigation Mesh, which generates a grid of walkable areas in the game map. AI will search for paths through the grid nodes to assist NPCs in path planning. NPCs will find the best position to reach their targets in these spaces [7]. This is the current action principle of traditional game AI, but it is different from the AI that will be discussed next, as it focuses more on language and dialogue interaction with players.

DOI: 10.54254/2755-2721/110/2024MELB0098

Traditional conversational NPCs are mainly responsible for telling stories and assigning tasks through interactions with players, which are usually predetermined by the program. They do not do anything beyond the program's requirements and have no memory of the player's behaviour. However, with the emergence of big language models, designers seem to have found a new path. In 2023, a project called AI Town was conducted at Stanford University to explore the possibility of AIGC-driving intelligent agents. They created resumes of 25 citizens and imported them into GPT, generating 25 NPCs with different names, habits, memories, and interpersonal relationships, and then allowed them to move freely in Pixel Town to observe their behaviour. Due to its unique architecture, proxy NPCs generated by GPT can make choices based on changing environments and have the ability to remember, reflect, and retrieve. Pixel Town is a simple sandbox with all the facilities needed by NPCs on the map. Players can also join the game as outsiders or existing NPCs. At the same time, players can control the environment on the map, such as turning off stoves or damaging NPCs' water pipes, and observe NPC feedback. NPCs also make decisions about daily affairs based on their environment, short-term memory, and long-term memory.

This study provides a new possibility for future game development, which is to use large language models to replace traditional NPC dialogue programs. If NPCs with realistic behaviour can be created, it will help enhance players' gaming experience and present interactive novels in a completely new way. By combining the use of action AI and AI agents, game designers can create NPCs with more diverse personalities and behaviours, providing players with more unexpected fun during the game process.

4. The development and application of AI-generated videos

The development of AI-generated videos has made rapid progress in recent months. Six months ago, AI was unable to understand how humans use utensils and eat spaghetti. However, with the emergence of video generation AI such as Sora and Kling, generative videos have officially entered a new stage. This technology is not yet mature, but in the near future, users will not need to spend a lot of time and money shooting videos, setting up scenes, and creating special effects. Users only need to provide pictures and keywords to generate the desired videos (Figure 1).

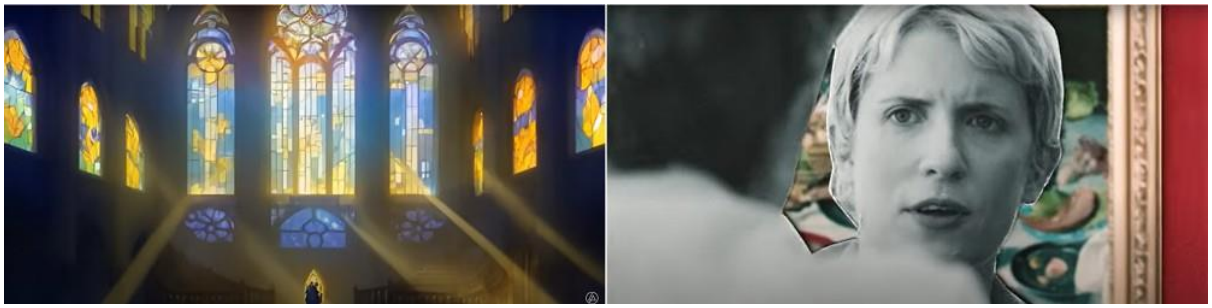

Figure 1. Use cases of text-to-video generation applications for the creative industry: music video (left) and film (right) [8].

The model for generating text-to-video mainly involves three constructs: language interpreters, temporary handlers, and vision processors. Language interpreters are responsible for processing input text and converting text content into visual objects, vision processors are responsible for generating images and scenes, and temporary handlers are responsible for processing video time and ensuring that frame rates are coherent [8].

Although the quality of videos generated by AI is becoming increasingly accurate and realistic, the drawbacks are also obvious. Videos do not understand the space outside the lens, which means that AI does not have the memory of the endless space outside the lens. As long as the lens is moved, it will be affected, indirectly leading to the inability to use AI to generate long videos. These are all minor issues that will be resolved with the updates and integration of the larger model (Figure 2).

DOI: 10.54254/2755-2721/110/2024MELB0098

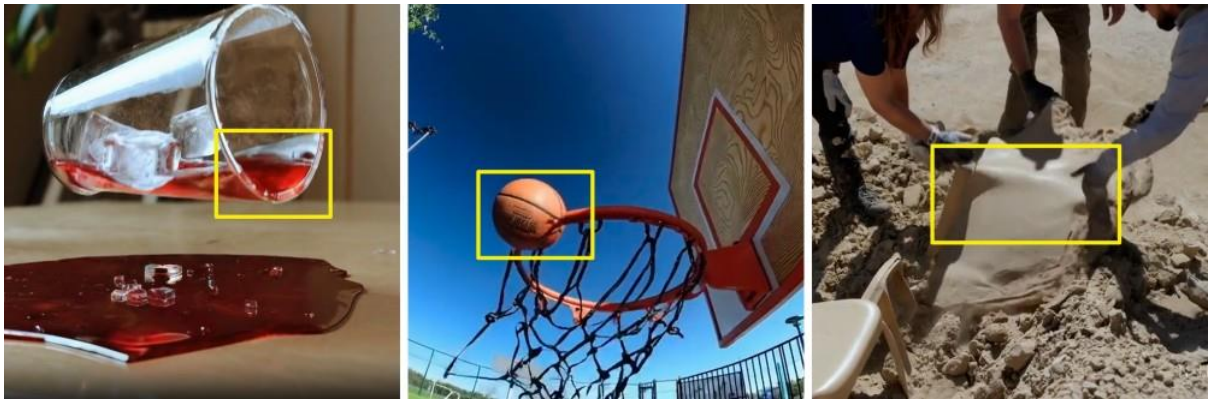

Figure 2. This image demonstrates Sora's lack of understanding at the physical interaction level. (Left) Water flows upwards, (middle) ball passes through the solid ring, (right) plastic chair made of clay [8]

5. Conclusion

This paper systematically explores the principles and cross-disciplinary applications of 3D real-life scanning, AI for NPCs, and AI video generation. The analysis demonstrates that these emerging technologies effectively streamline content creation processes, lower technical barriers, and inspire novel approaches to game design. With these advancements, artificial intelligence not only reshapes the ecosystem of the entertainment industry but also provides more opportunities for newcomers, thereby promoting diversity and inclusivity in game development. The integration of AI reduces the gap between experienced developers and aspiring creators, enabling a wider range of individuals to contribute original ideas and innovations. However, it is important to acknowledge that current technologies still face several challenges and limitations. Issues such as insufficient accuracy, cost control concerns, and ethical considerations need to be rigorously addressed in future research and practice. For instance, the reliance on AI for content creation raises questions about originality and ownership, which must be navigated carefully to ensure a fair and sustainable creative environment. Additionally, while AI tools enhance productivity, they must also be developed and deployed responsibly to avoid potential biases and adverse social impacts.

Looking ahead, as an essential production tool, the widespread application of AI will drive advancements in multiple fields, enabling users to create high-quality work without needing specialized skills. The potential for democratizing content creation means that more individuals, regardless of their backgrounds, will be able to engage in creative processes previously reserved for skilled professionals. In summary, the development of AI technology has significantly reduced the entry barrier for becoming a designer, and it is anticipated that in the future, an increasing number of individuals will be able to actively participate in the creative and development processes, ultimately fostering innovation and growth across the industry.

References

[1] Arts, E. (2023). Procedural Terrain in EA SPORTS PGA Tour. *Electronic Arts Inc*. https://www.ea.com/frostbite/news/procedural-terrain-in-ea-sports-pga-tour

[2] Kasapakis, V., Gavalas, D., & Dzardanova, E. (2024b). 3D Modelling Through Photogrammetry in Cultural Heritage. *In Springer eBooks* (pp. 23–26).

[3] Verykokou, S., & Ioannidis, C. (2023). An Overview on Image-Based and Scanner-Based 3D Modeling Technologies. *Sensors*, 23(2), 596.

[4] Chapinal-Heras, D., Carlos, D., Natalia, G., Sergio, E., Lucía, P. S., De Corselas Manuel, P. L., & Elías, R. Z. M. (2024). Photogrammetry, 3D modelling and printing: The creation of a

DOI: 10.54254/2755-2721/110/2024MELB0098

collection of archaeological and epigraphical materials at the university. *Digital Applications in Archaeology and Cultural Heritage*, e00341. https://doi.org/10.1016/j.daach.2024.e00341

[5] Arts, E. (2019). Photogrammetry and Star Wars Battlefront - Frostbite. *Electronic Arts Inc.* https://www.ea.com/frostbite/news/photogrammetry-and-star-wars-battlefront

[6] Arts, E. (2023). Procedural Terrain in EA SPORTS PGA Tour. *Electronic Arts Inc.* https://www.ea.com/frostbite/news/procedural-terrain-in-ea-sports-pga-tour

[7] Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., Bernstein, M. S. (2023). Generative Agents: Interactive Simulacra of Human Behavior. arXiv.org. https://arxiv.org/abs/2304.03442

[8] Cho, J., Puspitasari, F. D., Zheng, S., Zheng, J., Lee, L., Kim, T., Hong, C. S., Zhang, C. (2024). Sora as an AGI World Model? A Complete Survey on Text-to-Video Generation. *arXiv.org. https://arxiv.org/abs/2403.05131*