The Investigation of Generative AI on Role-Playing Games: Applications, Challenges, and Future Trends

Zhenpu Zhao

Civil Engineering, Ritsumeikan University, Kusatsu, Japan

zzplei@gmail.com

Abstract. In response to problems of homogenized experience caused by the solidification of development logic and interaction methods in desktop RPG games, this paper summarizes the specific application of current generative artificial intelligence in the development process and content of RPG games by reviewing relevant literature; and obtains solutions to optimize the creation paradigms of text and these two types of content in the game industry. In the creation of text content, artificial intelligence generated content (AIGC) technology demonstrates potential auxiliary creation capabilities on two levels: in the narrative dimension, the dynamic task generation system based on GPT solves the problem of mass production of text in open-world games. Compared with traditional manual scripts, the generated content ensures content diversity while reducing expansion costs (low marginal benefit); in the long process of plot interaction, by pre-training dialogue logic to ensure character consistency, task generation that conforms to human player experience habits, has storylines, and is logical is also achieved. Regarding the application of text-to-image generation (TTIG), this paper describes the industrial benefits brought by this technology. The text generation module provides convenience for the iteration of proposals and concept diagrams. TTIG tools automatically convert neutral sketches to stylized renderings, reducing human input in the prototype design stage by 80% and lowering resource costs, while SketchMetaFace's 2D to 3D deep model generation shortens the modeling cycle of medium-fidelity models. For the above application scenarios of generative AI, this paper also sorts out current obstacles and challenges faced and discusses targeted solutions.

Keywords: Generative AI, RPG, TTIG, AIGC.

1. Introduction

Today, after more than three decades of development in the Desktop-device products, most of the gold in the mine has been mined - whether it is those astonishing, ingenious designs or the visual shock brought by the development of hardware. Perhaps people have to accept the fact that before new terminals and their corresponding interaction methods become mainstream, game developers can only make minor adjustments to the masterpieces that have emerged over the past three decades, as they did before. In most cases, this has little effect. Players often find somethings are familiar when they encounter new products - just as one cannot publish the theory of relativity again, those great designs are inevitably borrowed or referenced. At the same time, in some large-scale 3A games, manufacturers try to integrate several proven mature gameplay elements into the game system to demonstrate the richness of the game content and progress compared to previous works. However, for the player experience, these attempts may be negative - it may lead to contradictions in the core gameplay, the game rhythm may become incoherent under the drive of multiple goals, and players may lose focus in this complex game system.

The above-mentioned issues can reflect a major predicament of desktop-device game products: the homogenization of experience caused by weakened innovation capabilities. In this case, is there a way to provide players with interesting and novel content experiences on mature platforms where interaction mechanisms have almost been fully explored and lack innovation space? Perhaps the development of generative artificial intelligence can offer a universal solution to this problem.

For instance, there is a process of an RPG game experience - players carry a secret that can change the world and embark on adventures in a vast map; the NPCs in this game do not merely repeat their lines as the game progresses, but behave like staff in LARP (Live Action Role Playing)

- although they generally do not think about the meaning of their existence, they respond to players while remaining true to their character designs. With prompts pre-tested and integrated into the game itself, NPCs act like real people who have thoroughly read the rule book, enriching the game's world view, enhancing players' immersion, and bringing greater replayability to the game. With the iteration of high-quality large language models (LLMs) such as GPT, creators can efficiently generate character background stories, cultural backgrounds, and behavioral logics using the text generation function [1].

Generative AI can also provide assistance in the art design process for development - for instance, GenAI can directly generate images, 3D models, and animations based on text descriptions, reducing reliance on traditional modeling techniques. Meanwhile, through its game character design tool Sketchar, communication barriers between designers and illustrators during the development process will be mitigated [2].

This paper will summarize the current research on the application and development trends of generative AI in games or interactive media and based on the actual situation of Desktop-device products, discuss the feasibility and specific utility of various application technologies and solutions as a carrier.

2. AIGC+RPG

2.1. Text Generation

In most RPG games, the game process is composed of the interaction system and the story plot is mainly presented to players through text content, while the application scenarios of generative artificial intelligence in text content mainly focus on endowing the game system with the function of natural communication. By breaking down natural communication from shallow to deep, the demands and actual application situations that generative artificial intelligence faces in RPG games can be summarized on two levels:

2.1.1 Narrative

In the main storyline of strongly narrative-driven RPG games, the advancement of the game plot highly relies on the narrative scripts between players and NPCs. In contrast, in other open-world-themed products that emphasize free exploration; to enhance the narrative immersion, the presentation of the story background also cannot do without a large amount of flavor text. With the popularity of open-world RPG games (such as The Elder Scrolls V: Skyrim), players' demand for dynamic and massive game content has soared, while the traditional way of manually writing tasks is facing bottlenecks such as high time cost and creative exhaustion [3].

By introducing text generation technology based on language models (GPT), creators can enhance the diversity and narrative quality of RPG games at a relatively low cost. For the periodic performance content in a fixed game process, the GPT model can generate more natural, coherent and creative task descriptions, reducing formulaic problems. For example, Quest-GPT-2 can batch produce more differentiated and contextually adapted content for homogeneous text requirements by anchoring specific details such as locations and character designs [4].

At present, a relatively reliable application method for AI-generated text is the pre-implantation of text - a hybrid content production process combining the generation model with manual screening: the model can quickly generate a large number of embryonic quest descriptions, and human developers can select high-quality or highly creative outputs from them for modification, replacing the completely manual way of writing task texts [3].

2.1.2 Task

In RPGs, quests are challenging tasks that players need to complete. Main quests are the cornerstone of the game's storyline, while side quests serve as supplementary elements to enrich the story background and world exploration. As players' demands for more game content increase,

traditional human quest designers face challenges in terms of time and creativity. When generating text to construct a complex, complete, and logical process, language models, unlike unitized narrative scripts, need to ensure the coherence of the generated content on a larger scale. At the same time, human creators also need to design a more universal task framework to ensure that the generated content can be reasonably utilized.

Systematically explored the application potential of GPT-2/GPT-3 in RPG task generation, and the following achievements were obtained [3]:

For the first time, the fine-tuned Quest-GPT-2 was used to generate complete task descriptions containing multiple elements (task issuers, locations, character relationships, and rewards). A large-scale user study (349 participants evaluated 500 generated descriptions) verified the potential of language models to generate content - 22% of the generated content reached a human-acceptable level, breaking through the formulaic problems caused by traditional rule/graph structure methods.

In the application scenarios of generation tasks, based on the intervention time point of AI and the degree of interference by human creators, three application modes can be classified [4]. Firstly, as a creative assistant, it collaborates with professional authors in the creation process (the model generates the initial draft, and humans add complex details). Secondly, automatically expands a complete task outline and description based on the starting sentence. Thirdly, semi-autonomous generation mode, which involves offline batch generation followed by light manual screening before being put into use in games, provides a practical solution for game development and has significant efficiency advantages over traditional fully automated or fully manual creation methods.

2.2. Image Generation

In an era where the performance of general hardware is advancing rapidly, players are becoming increasingly demanding about the visual effects of products. From exquisite realistic scenes to fantastical cosmic creatures, from original drafts with design elements that fit the character design to detailed models with fine materials and structures - all of these require the support of a strong art team. Or, in a large-scale RPG with a game process lasting over a hundred hours, will people get tired of the monotonous enemies they, see? And will they expect the game's art style to be closer to their own aesthetic preferences?

The application of text-to-image generation systems (TTIG) in the gaming industry has become commonplace. From the initial screening of art concepts during project initiation to the partial replacement of some strong alternative art assets in the later stages, creators have gradually become accustomed to generating content more efficiently with the help of generative AI. [Project NOX], a case study of a product developed with AI-assisted processes, demonstrates the efficiency gains brought by generative AI to project development through a comparison with traditional development processes - AI reduced the proposal preparation time by 50%, cut resource costs by 80%, and continuously improved iteration efficiency (for example, the second proposal for NOX only took 1.5 months to complete three new concept images) [5]. The following are two specific applications of TTIG in the creative process [1]:

Automated prototyping: Tools like EmoG can transform neutral sketches into expressive character visuals, streamlining the early prototyping process (saving time and manpower).

Quick 2D to 3D conversion: SketchMetaFace enables the rapid generation of detailed 3D depth models from 2D sketches, accelerating the mid-fidelity modeling stage.

Based on a survey of visual creators in the gaming industry [6], it was found that they maintain a curious and open attitude towards the integration of TTIG into their workflow and are optimistic about the quality of the content produced by the continuously evolving TTIG. This also reflects, from the side, the industry adaptability foundation that game practitioners have shown towards TTIG due to their long-term adaptation to technological change traditions (such as tool iterations and player feedback mechanisms).

3. Challenges and Future Prospects

3.1. Technology Application

Limited by the current performance of language models or due to the fact that the system of collaborative creation between human creators and generative AI is not yet well-established, the application of generative AI in RPG games is still in the experimental stage at present. However, with the rapid iteration of language models, game developers will eventually have to face the scenario of working with AI, and players will be willing to accept the content produced by AI once its quality stabilizes. Optimism about the future should be based on an honest and profound understanding of the current problems - currently, AIGC still has the following technical issues:

Firstly, consistency. When generating text content, the GPT model has difficulty distinguishing multiple entities (such as characters and locations) and cannot integrate task elements reasonably, leading to logical errors (such as "gluing unrelated information") [4]. Also, when generating image content, AI tools (such as DALL-E) require manual fine-tuning to match the preset artistic style and have difficulty maintaining consistency among multiple characters or across different scenes automatically [1]. For future prospect, detailed design anchor points need to be added and modelized generation curves need to be confirmed.

Meanwhile, data dependency and generalization also face difficulties. Model performance is limited by the theme of the training data (e.g., focusing on medieval fantasy RPGs), raising doubts about its generation capabilities for other settings (such as science fiction or modern). In addition, homogeneous datasets (e.g., using only World of Warcraft quests) can reduce model generalization [3]. For future prospect, structurally separate the theme from the content to ensure their independence during training and generation. Based on this plan, the execution of generative content will be separated into two steps: theme selection and constraints & structural plot generation; As a result, the specific story script will be composed of themes and elements that are in line with the world background.

3.2. Social Ethics

Unlike machines or other inanimate objects, humans are quite obsessed with the mental reflection of things beyond natural phenomena or objective facts - meaning; or in some perspectives, this is the only thing that humans can perceive and care about.

Just as many advancements in the biological field have been concealed or restricted, whether AIGC can be accepted or developed largely depends on the contemplation of the following ethical propositions within the industry and the formation of a social consensus:

Cultural biases and stereotypes - AI training data may contain cultural hegemony or stereotypical content (such as a lack of cultural nuances), leading to the generation of characters that perpetuate biases, which requires additional review by designers [1]

Balancing User Intentions and Public Value - AI needs to identify the conflict between the user's potential intentions and social ethics through dynamic grounding. For instance, when Alice suggests "blowing up a car", Jarvis not only considers the user's direct demand but also introduces eco-friendly solutions to reconcile personal creative freedom with public safety values [7].

4. Conclusion

In general, this paper reviews the current research on generative AI in games and summarizes the industry status and future development trends in the fields of AI text and image generation. AI text generation can be applied to simple dialogue scripts with NPCs or used in tasks closely related to the game process to provide better creativity and randomness. However, at the current development level, high-quality LLMs still have deficiencies in distinguishing various gaming elements and generating corresponding logical outputs. Therefore, in practical applications, creators need to design better constraints and select appropriate prompt words, while ensuring that the model

references objects from a coherent context. On the other hand, for the field of AI image generation, although it is still imperfect and has flaws, its value as a tool to assist creators in content creation and provide concept and inspiration references has been widely verified. Moreover, as game industry practitioners have long been in a creative environment of technological revolution, they can better accept and adapt to new production methods. Admittedly, at the current technical level, AI-generated image content still requires manual adjustment to achieve the desired visual effects. However, with the iteration of TTIG functions, the quality of the content it can produce will improve, and it will also be able to present a wider range of styles. It can be foreseen that this productivity tool will be more widely and multi-levelly applied in the game production process. Based on the actual application situation and potential possibilities of generative AI in the creation of RPG game products, the author believes that AIGC has the potential to open up a path of unlimited content innovation under limited hardware conditions, and if it is fully applied in the game industry, game creation work will also show a trend similar to other IT departments, that is, transforming from a labor-intensive industry to a smart-enhanced new industry.

References

[1] Wu Z, Chen Z, Zhu D, Mousas C, Kao D. A Systematic Review of Generative AI on Game Character Creation: Applications, Challenges, and Future Trends. IEEE Transactions on Games. 2025 Apr 28.

[2] Long LI, Xinyi CH, Ruoyu WE, Toby Jia-Jun LI, Ray LC. Sketchar: supporting character design and illustration prototyping using generative AI. Proceedings of the ACM on Human-Computer Interaction. 2024;8(CHI PLAY):337.

[3] Van Stegeren J, Myśliwiec J. Fine-tuning GPT-2 on annotated RPG quests for NPC dialogue generation. InProceedings of the 16th International Conference on the Foundations of Digital Games 2021 Aug 3 (pp. 1-8).

[4] Värtinen S, Hämäläinen P, Guckelsberger C. Generating role-playing game quests with GPT language models. IEEE transactions on games. 2022 Dec 12;16(1):127-39.

[5] Lee J, Eom SY, Lee J. Empowering game designers with generative AI. IADIS International Journal on Computer Science & Information Systems. 2023 Jul 1;18(2):213-30.

[6] Vimpari V, Kultima A, Hämäläinen P, Guckelsberger C. "An adapt-or-die type of situation": perception, adoption, and use of text-to-image-generation AI by game industry professionals. Proceedings of the ACM on Human-Computer Interaction. 2023 Sep 29;7(CHI PLAY):131-64.

[7] Vaithilingam P, Arawjo I, Glassman EL. Imagining a future of designing with ai: Dynamic grounding, constructive negotiation, and sustainable motivation. InProceedings of the 2024 ACM Designing Interactive Systems Conference 2024 Jul 1 (pp. 289-300).