ISSN:1738-6764, eISSN:2093-7504

Study of Production Workflows in Interactive 3D Animation with AI Applications

Elisa Adriana [1], Balgum Song [2,*]

[1] Dongseo University, Visual Contents, Ph.D. Student; elisaadriana09@gmail.com
[2] Dongseo University, International College Animation, Professor; sbu1977@dongseo.ac.kr
* Correspondence

https://doi.org/10.5392/IJoC.2020.20.4.068

Manuscript Received 8 July 2024; Received 10 December 2024; Accepted 19 December 2024

Abstract: *The rapid advancement in technology has led to the creation of interactive media across various fields, including education, entertainment, advertising, film, gaming, and animation. However, interactive animations have not achieved the same level of popularity as interactive films and games, often due to their complex story structures, additional production steps, high costs, and the necessity for expertise in game engines to enable interactivity. This paper examines the use of artificial intelligence (AI) tools, particularly Convai within Unreal Engine, to establish a more efficient workflow and reduce production costs in interactive 3D animation. The study compares traditional manual production methods using Unreal Engine and ChatGPT with AI-enhanced workflows that incorporate Convai. The findings indicate that AI tools significantly reduce production time and simplify the creation of interactive features. However, Convai has limitations in flexibility and precision, particularly when it comes to customizing features and animations. While AI tools are beneficial for beginners and those with limited programming experience in Unreal Engine due to their user-friendly nature, manual workflows provide greater flexibility for complex interactions and customizations. The research concludes that AI has substantial potential to improve the production of interactive 3D animation, although further advancements are necessary to enhance support for character and animation customization.*

Keywords: Interactive; 3D Animation; Artificial Intelligence, Unreal Engine

1. Introduction

With the rapid development of technology, interactive media have flourished and branched into various domains, including film, game and animation. This evolution led to the creation of the interactive film 'Kinoautomat,' first introduced in 1967 [1]. In 1976, the first text-based story game, 'Colossal Cave Adventure,' was created, playing a significant role in the development of RPG and visual novel games [2]. Although the gaming industry has achieved significant success in establishing multiple genres based on interactive storytelling, the production of interactive animations has remained relatively limited. Netflix first ventured into producing interactive film with the release of 'Black Mirror: Bandersnatch,' which gained significant attention at the time. Following the relatively successful attempt with 'Black Mirror: Bandersnatch,' Netflix proceeded to produce several short interactive 3D animations, including 'Carmen Sandiego: To Steal or Not to Steal,' 'The Boss Baby: Get That Baby!' and 'Puss in Book: Trapped in an Epic Tale.' Unfortunately, these interactive 3D animations were not as successful, as evidenced by their relatively low ratings [3].

There are many factors that cause creators to hesitate to make interactive animations, such as complex story structures, extra steps in the production workflow, using game engines to enable interactivity, and high production costs. Therefore, this paper attempts to use AI tools such as Convai, supported in Unreal Engine, as a solution for creating a more efficient workflow and decreasing the production time in interactive 3D animation. Although it is expected that the application of AI will have a significant influence on the production workflow and time cost, there is a strong possibility of encountering several limitations in the settings and quality of the animation. This paper significantly contributes by exploring underdeveloped area in interactive 3D animation

by employing a free AI plugin to reduce production time and make it accessible to non-experts, while also introducing a new workflow to enhance efficiency and update production practices.

2. Research Background

2.1 Interactive Animation and Game

While interactive animation shares common ground with story-based game in allowing the audience to influence the story, the primary focus of these two fields differs. Although the audience is empowered to become more active, the main objective of interactive animation is to deliver a compelling storyline and convey messages to the audience. Thus, interactivity should support and help to achieve this primary objective rather than focusing on technology or the level of activity the audience will experience. On the other hand, games prioritize gameplay and autonomy, providing players with greater freedom and offering more active engagement during play [4].

Interactive animation receives some attention for its creativity experiences, interactive storytelling is more effectively integrated into the gaming medium due to gaming's active nature, contrasting with the passive observation of animation. The other reason is the duration and complexity to implement interactive storytelling. Animation is usually watched in one sitting, which means there is a limited time duration before the audience feel tired. It typically less than 2 hours for a full-length animation and less than 30 minutes for short animation. Games, on the other hand, are not bound by time limits, giving them more freedom to create complex and long story structure also the amount of interactivity provided within the game. The extensive differences between traditional animation and interactive animation can make it challenging for audiences to adapt. Considering this, less interactivity might be more suitable approach for interactive animation.

2.2. Animation and Interactive Animation Production

In traditional animation production, the storyline typically follows a linear structure fully controlled by the creator, resulting in a single predetermined ending. In contrast, interactive storytelling facilitates a two-way communication between the audience and the creator. Interactive storytelling introduces a more complex structure where the audience can intervene in the storyline by making personal choices regarding the characters' decisions. These decision-making points create a more complex story structure with multiple possible endings, known as branching trees, as can be seen in fig 1 [5].

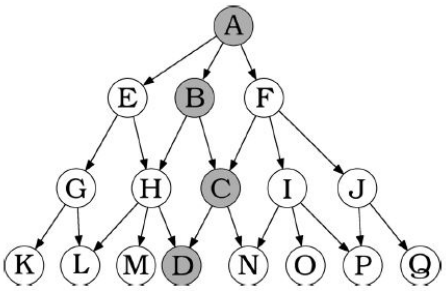

Figure 1. Branching tree in interactive story structure according to Chris Crawford.

Branching narrative trees expand exponentially and can become difficult to manage, as each choice can lead to multiple distinct storylines. Each decision typically results in at least two branches, with subsequent branches further multiplying, necessitating meticulous planning. Creators must consider the full spectrum of audience actions and possible story outcomes, ensuring each choice meaningfully impacts the narrative and avoids repetition. Considering the story structure, animators are required to produce more scenes and potentially create additional 3D assets, such as varied environments and interactive objects.

Furthermore, the audience's ability to influence the story necessitates several additional steps in animation production, including considerations for the platform and tools for interactivity (e.g., touch screens, text communication, motion tracking, or other devices) [6]. While animation typically uses 3D software like Maya, ZBrush, and Blender, interactive animation requires additional tools such as Unreal Engine and Unity for interactivity, which increases production costs. This approach creates a new storytelling experience, allowing

audiences to view narratives from different perspectives, with technological advancements continually enhancing and promising further developments [7]. For example, the shift in consumer behavior following COVID-19, where people prefer watching films at home on their devices instead of in cinemas, creates a promising environment for the growth of interactive animation.

2.3 AI application in Animation Production

Currently, several AI tools have been created to help with animation, offering significant potential to enhance animation production workflows and reduce costs. For instance, Cascadeur features automatic physics animation, auto-posing, in-between, and creating secondary actions, allowing animators to focus solely on key posing rather than starting from scratch. Additionally, AI enables the management of complex, branching narratives in interactive animation, providing real-time interaction and character behavior adjustments [8]. This advancement enhances the appeal of interactive animations by granting audience's greater narrative control.

The development of ChatGPT and integration with Unreal Engine enables real-time interaction with avatars via plugins. In story-based games like "Suck Up," players use voice recordings to engage AI villagers, who respond using developer-created Generative AI [9]. Interactive storytelling with AI is also evident in games such as "Storyteller", where players arrange 2D characters and environments to progress narratives, and in projects like "1001 Nights" [10, 11]. For 3D interactive game and animation, Unreal Engine behavior tree and blueprint and Unity ML-Agents are commonly used by many such as "Life is Strange" series and "Until Dawn".

Recent developments have integrated ChatGPT into game engines to improve interactive character communication, but this often requires programming skills in C# or C++. Convai, an AI plugin supported by Unreal Engine, presents a more accessible option, offering automated functions, task recognition, and lip-sync without requiring programming expertise. This paper will compare Unreal Engine's integration of ChatGPT through manual blueprints and behavior trees with the Convai plugin to assess their contributions and limitations in animation production. Given the rapid advancements in technology and AI tools that enhance the efficiency and innovation of interactive animation production, the future of interactive animation appears promising.

3. Materials and Methods

The method used in the paper is using Unreal Engine 5.2 to enable more interactivity within interactive 3D animation. Then compare the manual production step using blueprint with ChatGPT 3.5 Turbo and Convai production step to analyze the contribution and limitation of AI application within interactive 3D animation production workflow. Convai is an AI plugin designed to provide creators of virtual worlds with advanced conversational AI tools [12]. It allows users to create their own characters, complete with backstories, voices, lip-sync capabilities, knowledge, personalities, and the ability to perform actions. Similar to ChatGPT, it offers an AI assistant that can provide real-time feedback.

Convai can be integrated and installed for free in Unreal Engine and Unity, with all plugins available on the website. While in the manual production step, the experiment will only use ChatGPT 3.5 Turbo for enabling communication between characters and NPCs. Interaction with objects and character animations are made using blueprints and behavior tree. The production testing was conducted by an individual with a background in 3D animation and limited programming experience. Things that would be considered in the experiment to be compare is the complexity and efficiency in production steps, the ability to communicate and object interaction, features and limitation.

4. Results

4.1. Production Workflow Using Unreal Engine and ChatGPT

The usual workflow in creating interactive animation within Unreal Engine is by using blueprint and behavior tree as can be seen in Fig 2. The user needs to create blueprint to connect all the features desired in AI character such as connecting ChatGPT, text to speech, make interface and buttons, animation blueprint for actions, trigger for interactive objects, and other setting needed. All the animation used for the character has to be imported and connected through animation blueprint. It is possible to connect ChatGPT to Unreal Engine, as there is a specific plugin available in Unreal Engine. Then make a behavior tree to create a trigger so that a certain actions or choices can influence the story to progress in a specific direction.

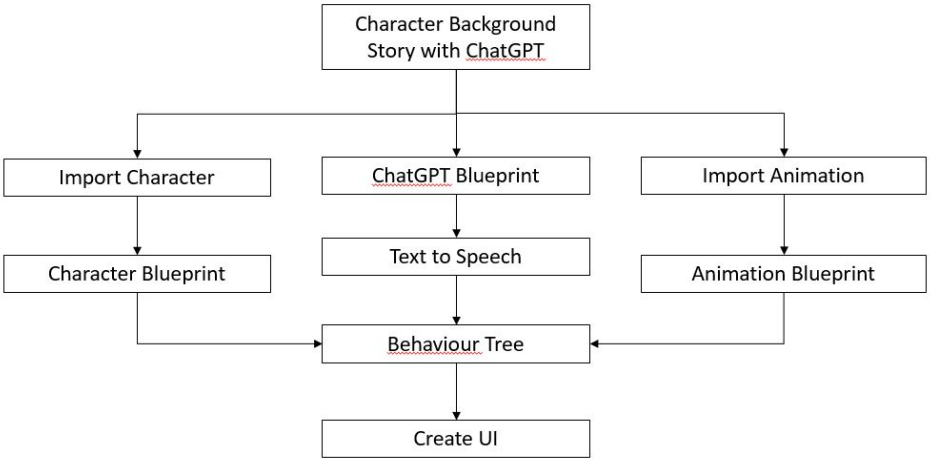

Figure 2. Manual production step in Unreal Engine and ChatGPT

4.2. Production Workflow Using Convai

Considering the many steps the user needs to take when creating interactive animations manually in Unreal Engine, Convai offers a faster and simpler workflow. First, the user needs to create an account on the Convai website to obtain the API key and character ID, which are crucial for connecting to the game engine. Then, the user can create their own character and configure settings such as backstory and knowledge. The user needs to install the plugin from the Unreal Engine marketplace and activate it in the project file beforehand. To utilize features such as voice extraction, lip-sync, and other additional capabilities, the user must download the plugins provided by Convai on the website. Finally, the user sets the interactive objects through Convai info details available within Unreal Engine, as can be seen in Fig 3.

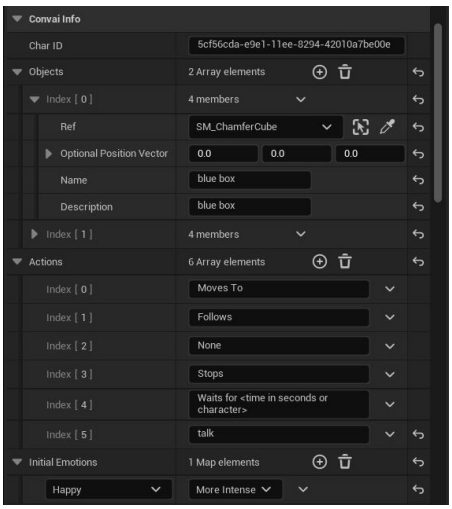

Figure 3. Convai info details to automatically set interactive objects and actions.

The user can set the interactive objects by clicking the eyedrop tool and adding the name of the objects in the details so that the object can be commanded whenever its name is mentioned. Convai provides a few characters and animations on the website that can be automatically added. However, other software must be used if the user wants to add their own custom characters and animations, as can be seen in Fig 4. Omniverse is compatible with Convai for this purpose. Although the software can be downloaded for free, the device used for Omniverse should fulfill certain requirements such as has minimum 16GB RAM and RTX GPU.

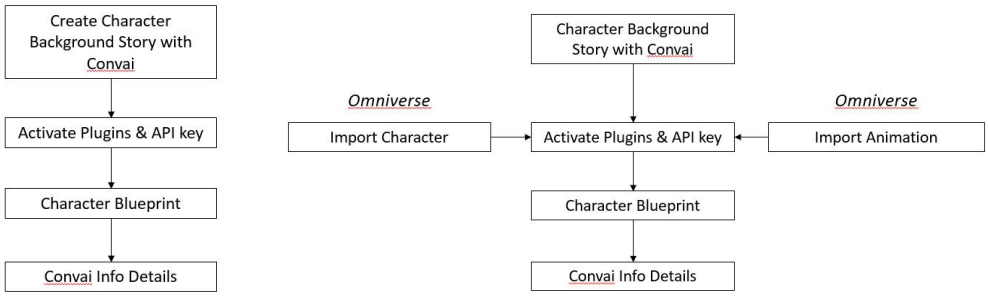

Figure 4. Interactive 3D animation production step using Convai (a) Production workflow using the default setting in Convai (b) Production workflow using Convai and Omniverse to add customize character and animation.

4.3. Comparison Result

In this paper, we present a simple interactive system featuring two characters: a player-controlled character and an interactive character capable of text-based communication, executing basic commands such as walking and running, and performing simple interactions like following and approaching objects. The player character is represented by the Mannequin character in Unreal Engine, while the interactive character is Echo character available in Unreal Engine. The interactivity was developed using two different methods: the first employed ChatGPT and Blueprints, taking 45 minutes to complete the interactive animation setup, including character communication and basic interactions such as walking and object following. The second method utilized the ConvAI plugin, requiring only 10 minutes to achieve comparable results. The comparison results in production time, communication, object interactivity, features, pros, and cons of manual production workflows with ChatGPT and Convai show that the use of AI tools significantly reduces time costs, as depicted in Table 1.

Table 1. Unreal Engine ChatGPT and Convai comparison table

	Unreal Engine ChatGPT (Manual)	**Convai**
Production Time	45 minutes	10 minutes
Communication	Consistent	Inconsistent
Object Interaction	Blueprint, behavior tree and triggers	Automatic drag and drop feature
Features	-Text to speech -OpenAI plugin	-Personality -Speaking style -Lip-sync
Pros	-More flexible control -Can be implemented with any character	-Automatic actions and animation -User friendly -Fewer step in production workflow -Free plugins
Cons	-More production steps -Requires extensive knowledge of blueprint and programming -Create triggers for every interaction. -No lip-sync feature -Requires payment for ChatGPT API	-Difficult to change the already set settings -Works best with human based character -Short interaction range -Lip-sync is rough with only basic shape (open and closing mouth) -Limited animations.

This study highlights key differences in production time, production steps, and flexibility in creating interactivity between two methods. Utilizing ChatGPT and manual Blueprints requires extensive programming knowledge, and significant time investment, as it involves constructing the system from the scratch. This process includes developing Blueprints to connect the character to ChatGPT, setting parameters such as chat history storage and API key integration, designing the user interface (UI) for text-based interactions, configuring animations, and creating a behavior tree for interactivity and commands with triggers. The usage of API key for connection to Unreal Engine also requires payment to ChatGPT Open AI.

Conversely, the ConvAI plugin offers pre-built functionalities, allowing users to set parameters easily and use it for free within Unreal Engine. Basic animations such as idle, walking and running and text communication features, including lip-sync, are pre-configured in ConvAI. It also provides settings to detail character personality traits that affect speech and responses. Simple commands, like following a character, are implemented without additional setup. ConvAI facilitates object interactivity through a simple drag and drop eyedropper tool, enabling the identification of objects for text-command interactions. However, ConvAI's pre-set functions, such as text box customization, animation changes, non-human character usage, and communication range adjustments, lack flexibility. In contrast, manual Blueprints offer complete control over function customization and visual design on the screen.

Figure 5. Interactive 3D animation final result using Convai

5. Discussion

The experimental comparison indicates that AI implementation in production significantly enhances workflow efficiency, reducing time costs by 35 minutes. The Convai website offers many automated features that enable users to create interactivity easily, such as voice, lip-sync, personality traits, emotion, text and speech UI, voices, and animation. These features are especially useful for beginners and those who are not professionals in using blueprints and programming in game engines. However, using customized characters and animations presents limitations, as it often requires additional plugins and software like Omniverse and Character Reallusion, which have specific device system requirements. Although updates have improved lip movement details, lip-sync accuracy remains low, demanding time-consuming manual adjustments for each syllable.

Conversely, manual work with blueprints in Unreal Engine offers greater flexibility, allowing users to create and organize interactivity according to their preferences. Users can create any conditions and triggers that affect the interaction and design their UI according to their needs. However, extensive knowledge of programming and operating Unreal Engine is required, as using blueprints is very practical. Everything needs to be set up individually within the blueprints using nodes, including all transition animations, such as from walking to running or idle to jumping. Therefore, it requires expertise and is time-consuming, as there are many additional steps in the production workflow.

In terms of conversational consistency, ChatGPT appears to perform better than Convai, providing uniform personal answers during repeated tests. For example, when asked about favorite food, it gives the same answer every time whereas Convai's responses can vary with characters sometimes answering politely and other times not. ChatGPT has a feature to control the randomness of completions called temperature, the lower the temperature value, the more accurate and repetitive the responses will be. However, ChatGPT lacks lip-sync capabilities and has limited text-to-speech voice options, with generated voices often sounding awkward.

Despite this, it is possible to integrate voices from other AI tools like Eleven Labs. While ChatGPT does not offer detailed settings for personality traits and emotions as Convai does, similar results can be achieved by providing more comprehensive descriptions in the ChatGPT assistant node.

6. Conclusion

Through experimental comparison, this research shows that AI tools significantly aid in creating simple interactivity and reducing production time. Despite some limitations, AI tools have great potential in interactive 3D animation production workflow where interactivity is less complex. Convai allows animators to focus on animation without needing programming expertise. However, for more complex interactivity and extensive condition-setting, a manual production workflow is recommended due to limitations and frequent errors in Convai's preset settings when users attempt to modify nodes. Further development of AI tools for interactivity could enhance technical support for character, lip-sync, and animation customization, particularly in stylized 3D cartoon animation.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] C. Hales, “Cinematic interaction: From kinoautomat to cause and effect,” Digital Creativity, vol. 16, no. 1, pp. 54-64, Aug. 2005, doi: https://doi.org/10.1080/14626260500147777.
[2] J. Lebowitz and C. Klug, Interactive storytelling for video games: A player-centered approach to creating memorable characters and stories, Taylor & Francis, 2011.
[3] Carmen Sandiego: To Steal or Not to Steal, 2020. [Online] Available: https://www.imdb.com/title/tt11767524/
[4] T. Abba, “As we might watch: What might arise from reconsidering the concept of interactive film?,” Journal of Media Practice, vol. 9, no. 1, pp. 19-27, 2008, doi: https://doi.org/10.1386/jmpr.9.1.19_1.
[5] C. Crawford, The Art of Computer Game Design, 1984.
[6] B. Tomlinson, “From linear to interactive animation: How autonomous characters change the process and product of animating,” Computers in Entertainment, vol. 3, no. 1, p. 5, Jan. 2005, doi: https://doi.org/10.1145/1057270.1057282.
[7] N. Simon, “Applications of Theoretical Models for Interactive Narrative Animation,” Ph.D. dissertation, RMIT Univ., Melbourne, Australia, 2011.
[8] R. Zhang, “Using Artificial Intelligence Assistant Technology to Develop Animation Games on IoT,” Computer Science and Information Systems, vol. 20, pp. 765-792, Jan. 2023, doi: https://doi.org/10.2298/CSIS220719021Z.
[9] Suck Up! (n.d). [Online] Available: https://www.playsuckup.com
[10] Storyteller on Steam. (n.d). [Online] Available: https://store.steampowered.com/app/1624540/Storyteller/
[11] Y. Sun, Z. Li, K. Fang, C. H. Lee, and A. Asadipour, “Language as Reality: A Co-Creative Storytelling Game Experience in 1001 Nights using Generative AI,” 19th AAAI Conference on Artificial Intelligence and Interactive Digital EntertainmentAugust, 2023, doi: https://doi.org/10.48550/arXiv.2308.12915.
[12] Convai-Conversational AI for Virtual Worlds. (n.d). [Online] Available: https://convai.com/