

IJACT 23-9-34

## Enhancing Video Storyboarding with Artificial Intelligence: An Integrated Approach Using ChatGPT and Midjourney within AiSAC

Sukchang Lee<sup>1</sup>

<sup>1</sup>Prof., Dept. of Digital Contents, Konyang Univ., Korea  
E-mail [2stonespear@gmail.com](mailto:2stonespear@gmail.com)

### Abstract

*The increasing incorporation of AI in video storyboard creation has been observed recently. Traditionally, the production of storyboards requires significant time, cost, and specialized expertise. However, the integration of AI can amplify the efficiency of storyboard creation and enhance storytelling. In Korea, AiSAC stands at the forefront of AI-driven storyboard platforms, boasting the capability to generate realistic images built on open datasets foundations. Yet, a notable limitation is the difficulty in intricately conveying a director's vision within the storyboard. To address this challenge, we proposed the application of image generation features from ChatGPT and Midjourney to AiSAC. Through this research, we aimed to enhance the efficiency of storyboard production and refined the intricacy of expression, thereby facilitating advancements in the video production process.*

**Keywords:** Storyboarding, Video Storyboard, ChatGPT, Midjourney, AiSAC

### 1. Introduction

"Creating a storyboard is a complex, time-consuming, and labor-intensive process that demands skilled specialists[1]." Due to these challenges, the convergence of artificial intelligence (AI) with storyboarding has garnered significant attention, aiming to expedite the video production pipeline, enhance visual storytelling effects, and reduce labor. In other words, in pursuit of heightened efficiency in storyboard creation, there has been a rising trend in leveraging AI for video storyboard production over recent years. Both the industry and academia are, accordingly, showing keen interest in AI-assisted storyboarding.

The leading AI storyboard platform in Korea is 'AiSAC (AI Analysis System for Ad Creation)'. By utilizing AI and big data technology in the creation of storyboards, AiSAC significantly reduces the time spent on sketching and revising. Furthermore, as AiSAC is based on open datasets, its characters and backgrounds are realistic. However, despite these advantages, there are also noticeable shortcomings in the outcomes produced by AiSAC. Notably, AiSAC's representation of scene images tends to be rough, making it challenging to depict intricate details within the storyboard panels. Additionally, due to the limited number of open sources, there is a constraint that results in the storyboards being expressed rather simplistically.

To address the aforementioned challenges with AiSAC, we integrated the image generation capabilities of ChatGPT and Midjourney, proposing a solution by comprehensively applying the combined outputs to AiSAC. Prior to this integration, we explored the characteristics and challenges of storyboard production utilizing

---

Manuscript received: August 1, 2023 / revised: August 20, 2023 / accepted: August 30, 2023

Corresponding Author: [2stonespear@gmail.com](mailto:2stonespear@gmail.com)

Tel: +82-41-730-5332

Professor, Dept. of Digital Contents, Konyang Univ., Korea

AiSAC. Drawing on the functional features of ChatGPT and Midjourney, actual cases were derived to address AiSAC's limitations. We then compared the traditional outputs with those produced using the enhanced methods, highlighting key differences and advancements. Through this research, an elevated approach to video storyboard production using AI was anticipated.

## **2. The Importance of Video Storyboards and the Emergence of AiSAC**

Generally, the video production process undergoes stages such as development, pre-production, production, post-production, distribution, and screening. This journey is characterized by high costs and the need for significant time, making it a complex endeavor. Consequently, video-makers must meticulously plan during the pre-production stage to ensure the successful creation of the video.

A storyboard is fundamentally a visual document crafted through illustrations and is paramount in the production of visual media, including advertisements and animations. In videomaking, the storyboard's detailed depiction is essential for a meticulous video production process. "The storyboard serves a significant function in video direction, enabling directors and staff to exchange opinions. Additionally, producers utilize storyboards to estimate the budget and get a preliminary visual concept[2]." Moreover, before having the production stage, the storyboard provides an overview of the visual sequence, offering the production team a practical roadmap. "One of the vital stages is how well-structured the storyboard is. This is because the storyboard transforms the writer's creative vision into the visual scenes[3]." "To successfully manifest a director's visual language, it is imperative that the storyboard backs and clearly visualizes it[4]." In essence, a storyboard should not be perceived merely as a basic sketch based on the screenplay, but as a medium that initially expresses the video in a visual language. Therefore, not only should a storyboard be well-structured, but it also needs to be of considerable quality.

However, crafting a well-structured storyboard is a challenging endeavor. For the visualization of the storyboard to be meticulous, it demands detailed considerations like spatial arrangements aligned with the storyboard's composition, descriptions of characters and backgrounds, camera settings and movements, lighting, props, dialogues, and music. For these reasons, leveraging AI in the storyboard creation process can be seen as an effective approach, especially in light of the storyboard's visual functionality. "AI is a technology used to enable computers to do things that can be done with human intelligence[5]". Angana and Goswami (2020) believed that using AI could boost video production efficiency, given that storyboarding in the pre-production stage is a tedious and laborious task[6]. O'Meara and Murphy (2023) regarded the Text-to-Image AI Generator in storyboard creation as a critical bridge between video production and a director's vision[7].

Internationally, the use of AI in storyboarding has sparked robust discussions, and there are also instances of its application in domestic markets. AiSAC stands as a notable example, harnessing AI for storyboard creation. "Blending advertising content data with AI image recognition technology, AiSAC auto-detects characters, objects, locations, and landmarks from over 27,000 commercials. It compiles these into metadata, generates Ad copy, and facilitates storyboard creation, positioning itself as a comprehensive service platform[8]." Services like these, which provide unrestricted panel creation and detailed scene descriptions in both visuals and audio, have potential applications beyond broadcast content and can be valuable assets to the video industry.

The reasons, AiSAC is effective in movie storyboard production, include the following. Firstly, there is efficiency. AI can automate manual storyboard tasks, significantly speeding up the process. This includes converting written scripts into visual scenes or extrapolating visual elements based on narratives. Secondly, there is consistency. AI ensures that visual elements, themes, and characters maintain consistency throughout the storyboard. This is especially useful in projects or series where uniformity is crucial. Thirdly, there is data-driven insights. AI can suggest appropriate images for scenes based on vast amounts of data. Lastly, it is accessible even for non-experts. AI simplifies the storyboard creation process, allowing individuals without drawing skills to easily engage. In essence, AiSAC has transformed the traditionally manual process of storyboard creation into a more efficient one through the use of AI.

## **3. The Limitations of AiSAC's Representation**

Despite the numerous advantages of AiSAC previously mentioned, it suffers from a limitation termed as the 'lack of precision in representation.' As Turner (2022) highlighted, accurately rendering the tone of a movie into the storyboard using AI is essential[9]. A storyboard, as a crucial communication tool for staff, demands a meticulous and precise representation. However, AiSAC sometimes misses the nuanced details crucial for visualization. In cinema, this visualization is pivotal not only for conveying the director's intent but also for ensuring clear communication with both staff and actors. Specifically, when emotions of a character are intricately depicted within the storyboard panels, it becomes invaluable for actors, guiding them in interpreting and portraying those emotions during their performances.

To illustrate the lack of precise representation inherent to AiSAC, consider the following example. Imagine producing a romantic comedy set on a college campus. A storyboard panel image created through AiSAC might look as shown in Figure 1. The AI-generated image was based on the search keyword 'college', and since it was rooted in an actual photograph, the composition and background appeared grounded in reality.



**Figure 1. AI-derived imagery produced by AiSAC**

Despite the advantage of easily generating images, the visual representation was coarse, limiting its ability to depict intricate visual details. This limitation was believed to arise from the AI's reliance on online image data, likely designed this way to avoid copyright issues. Moreover, against such a backdrop, the characters consistently displayed neutral expressions. As a result of these characteristics, staff members cannot infer the scene's ambiance or a character's emotions from the illustrations alone, and they must rely on accompanying descriptions to understand the mood. Another constraint of AiSAC was its limited pool of AI-generated images, due to the lack of abundant open datasets materials online. For instance, when using AiSAC's fill function with the keyword 'college', only about 40 results were generated.

## 4. Storyboard Creation Using ChatGPT and Midjourney

### 4.1 The Need for AI Imagery and the Process of AI Storyboard Creation

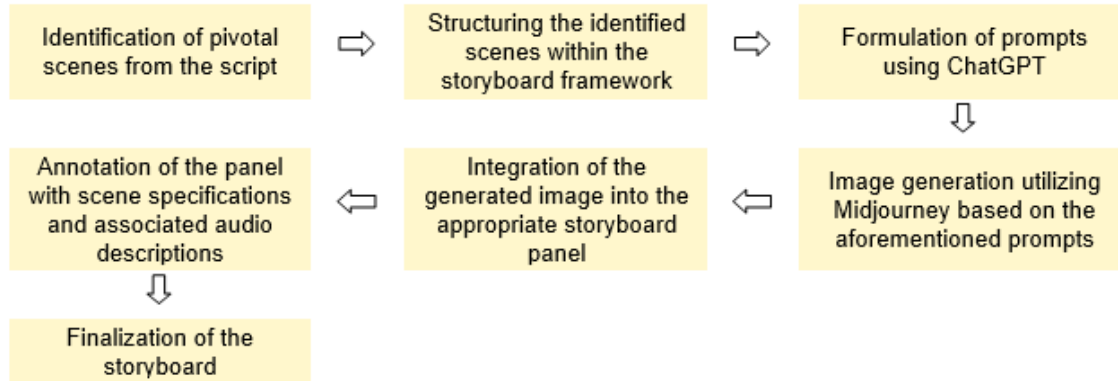
Due to the previously mentioned limitations of AiSAC, storyboards produced by AiSAC were constrained in their expressiveness. If there were a method to depict this with greater detail, it would lead to increased productivity and efficiency in storyboard creation. However, devising solutions to compensate for these constraints within AiSAC itself was challenging, primarily due to its inherent technical limitation in generating realistic images. This issue could be mitigated through the use of ChatGPT and Midjourney.

OpenAI's conversational AI, ChatGPT, can emulate human-like conversations to generate text. And the AI image generation platform, Midjourney, is designed to create visual images based solely on textual descriptions. Together, they can collaboratively generate AI images. If ChatGPT generates textual commands for Midjourney, then Midjourney can create AI images based on those textual descriptions. Although Midjourney alone is capable of producing AI images, the utilization of the 'Prompt' feature from ChatGPT aids in refining the image creation process, thereby enhancing productivity.

In this study, to address the limitations of AiSAC, the author utilized ChatGPT and Midjourney to produce a storyboard based on a short video script written by the author. The storyboard creation process was undertaken following the steps depicted in Table 1 below. The short video was about a youth drama set on a university campus, and the scenes included in the storyboard were from the video's opening segment. Initially, critical scenes from the script meant for the storyboard were identified, and these key scenes were mapped out on the storyboard. Subsequently, to generate the visualization of the scenes, prompts related to scene imagery were created using ChatGPT. These generated prompts were then inputted into Midjourney to produce the images. After creating the images, the desired final images were selected and stored, then incorporated into

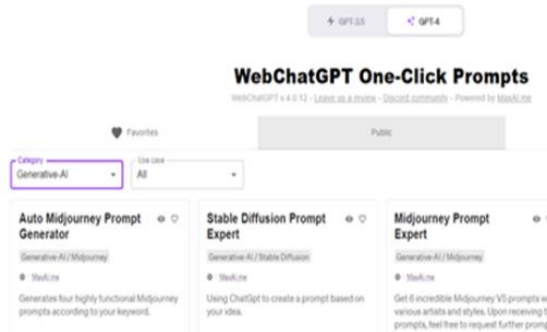
the storyboard panels in AiSAC. Following this, each storyboard panel was annotated with scene details and audio descriptions. After undergoing these steps, the storyboard was finalized and then exported as a file.

**Table 1. Procedure for improved AI-assisted storyboard production method**



#### 4.2 Storyboard Creation Using ChatGPT and Midjourney

To generate AI images, it was imperative to first create a command to input into Midjourney. As demonstrated in Figure 2, within ChatGPT's WebChatGPT One-Click Prompts, 'Generative-AI' was selected and then a prompt was formulated using keywords relevant to the desired image.



**Figure 2. ChatGPT's extraction of commands**

Upon examining the prompt, it was clear that it not only specified the character's appearance and movements but also included settings such as camera, lens, and screen ratio. Given this context, it became possible for directors to accurately manifest their intended scenes into precise images. ChatGPT facilitated this by crafting the optimal command for Midjourney to generate such images, and based on this, Midjourney produced images that closely align with the director's intended vision.

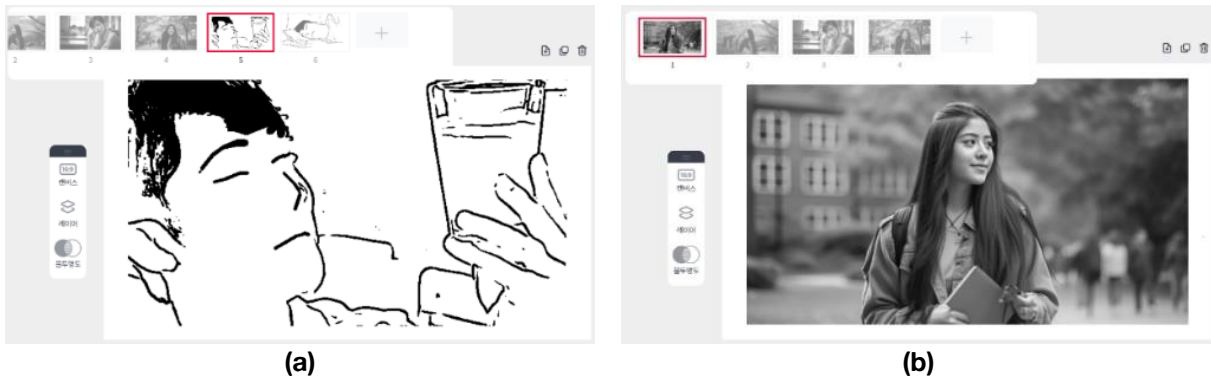
To generate AI images, an initial step involved copying the prompt command created in ChatGPT and entering it into the Midjourney chat window. As shown in Figure 3, this resulted in four images being generated by Midjourney. We could select the image that most closely resembled the intended vision from these generated images. The image selection and modification process were as follows: choosing U1 enlarged the first image, while selecting V1 recreated the first image with a similar feel. If one wishes to generate a different image beyond the provided four, simply click on the refresh button located on the right.



**Figure 3. Image generation via Midjourney**

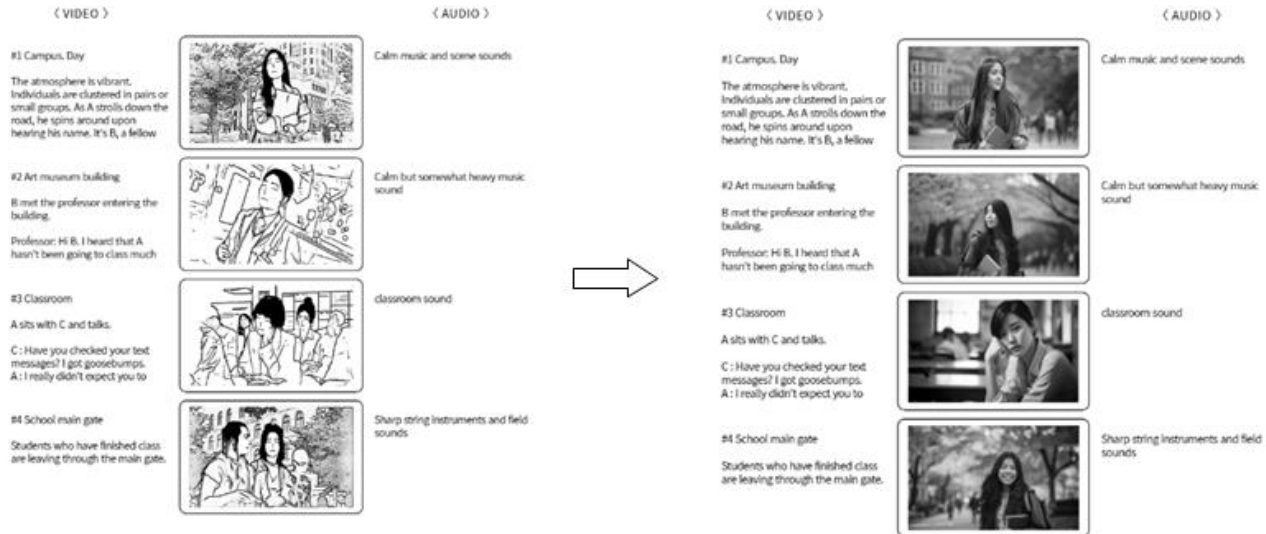
Although the images generated through Midjourney were drawn randomly, they could evolve towards the user's desired image by demanding subtle distinctions. Midjourney offered the capability to minutely modify the selected image. Even scenes that appeared nearly identical could be altered based on command variations. For instance, a small change in the direction where a background character looks and subtle modifications to a character's face were induced by adjusting the command.

The image results produced through this process exhibited a significant difference in quality and intricate representation compared to the default images within AiSAC panels. As revealed in Figure 4(a), the AI image lacked specific detail. Notably, the character's expression was generic and emotionless, necessitating additional scene descriptions for interpretation. In contrast, Figure 4(b), produced using ChatGPT and Midjourney and applied to AiSAC, showcased realistic image quality. Detailed facial expressions and backgrounds allowed for intuitive comprehension of the scene.



**Figure 4. Fine-tuned image data output from Midjourney**

To integrate images produced by ChatGPT and Midjourney into a storyboard, they should first be uploaded as original images to AiSAC and then imported onto the panel. By sequentially constructing such panels, an entire storyboard could be made, as show in Figure 5. Importantly, scenes centered around images, accompanied by audio descriptions, significantly aided in conveying the storyline. AiSAC offered the capability to numerously generate storyboard panels and had a feature allowing the aggregation of these panels into pdf or png files. This versatility suggested a high potential for industrial application. Within the panel images, users could utilize the drawing function to provide supplemental scene explanations or annotations regarding camera positioning and movement.



**Figure 5. Outcomes derived from the suggested methodology**

### 4.3 Comparison of Traditional and Improved Methods

To address the limitations inherent in the AiSAC system, the author conducted research on the process of deriving results from AiSAC, ChatGPT, and Midjourney, comparing their outcomes, and the storyboard creation. The following insights were gained. Firstly, the traditional AiSAC system often produced monotonous AI images, failing to clearly convey the user's intent in scene direction. However, the novel approach utilizing ChatGPT and Midjourney, while requiring an additional step of manually integrating images into AiSAC and thus slightly increasing processing time, allowed for a more meticulous and realistic representation of scenes. Secondly, the conventional method had a limitation in terms of the open datasets images available for AI creation, often leading to a shortage of images for storyboard panels. In contrast, the new approach allowed for the infinite generation of AI images, thereby eliminating restrictions on the range of image choices for storyboards. Lastly, due to limited visualization in the original method, scenes required detailed explanations using alternative means. This could increase the content explaining the scenes in the storyboard, potentially complicating the staff's understanding of the video. On the other hand, the updated technique could precisely visualize character emotions and backgrounds, incorporate camera and lens settings, and reflect shooting conditions. Consequently, scenes were visualized more comprehensively in the storyboard, allowing staff members to intuitively grasp the narrative.

## 5. Conclusion

With the meteoric rise of AI technology, its applications are now seen in diverse fields, and its influence continues to grow. In tandem with this global trend, the video industry has been actively utilizing AI technology, leading to rapid advancements in the storyboard creation environment. This change signals video-makers to consider novel approaches in storyboard creation using AI. We explored the characteristics and limitations of an AI storyboard creation platform called AiSAC, which was utilized in the video industry. While AiSAC significantly contributed to the efficiency of the storyboard creation process through AI utilization, it had a limitation in meticulously capturing characters' emotions and scenes. To address this, we employed the features of ChatGPT and Midjourney to produce AI images. These images were then integrated into AiSAC, proposing an enhanced method of AI-driven storyboard creation. We aimed to address the limitations of AiSAC by creating more realistic images in Midjourney, grounded on the intricate prompts generated from ChatGPT, tailored to the director's vision. These images were then utilized in AiSAC's storyboard creation features, resulting in a superior storyboard product. Through such an enhanced method, video directors could depict their envisioned movie scenes more intricately on the storyboard. This level of

detail in the storyboard facilitated smoother communication among the crew, ultimately increasing the efficiency of the videomaking process. In this regard, we focused on practical solutions to overcome AiSAC's limitations, primarily by manually integrating AI-generated images into AiSAC. Future research might consider the adoption of more advanced AI technologies to facilitate the automatic generation of entire storyboards. Exploring and implementing the implications of these emerging AI technologies on storyboard creation can enrich interdisciplinary understanding and applications.

## Reference

- [1] J. Hart, *The Art of the Storyboard: A Filmmaker's Introduction*, Taylor & Francis, pp. 27-28, 2013.
- [2] P. Ye and T. Baldwin, "Towards Automatic Animated Storyboarding," in Proc. 23th AAAI Conference on Artificial Intelligence 2008, Chicago, Illinois, p. 578, July 13-17 2008.
- [3] T. Bartindale, A. Sheikh, N. Taylor, P. Wright, and P. Olivier, "StoryCrate: Tabletop Storyboarding for Live Film Production," in Pro. the SIGCHI Conference on Human Factors in Computing Systems, pp. 169-172, May 2012. DOI: 10.1145/2207676.2207700
- [4] H. Lee, "The 21st Century's Film Storyboard, Report on the Use of Pre-visualization," Design Research, p. 35, 2020. DOI: 10.46248/kidrs.2020.4.33
- [5] H. Jung, "A Study on the Current State of Artificial Intelligence Based Coding Technologies and the Direction of Future Coding Education," The International Journal of Advanced Culture Technology, Vol. 8, No.3, p. 187, 2023. DOI: 10.17703/IJACT.2020.8.3.186
- [6] D. Angana and R. Goswami, "The Film Industry Leaps into Artificial Intelligence: Scope and Challenges by the Filmmakers," in Proc. Rising Threats in Expert Applications and Solutions. Advances in Intelligent Systems and Computing, vol 1187. Springer, Singapore, pp. 665-667, 2021. DOI: 10.1007/978-981-15-6014-9\_80
- [7] J. O'Meara and C. Murphy, "Aberrant AI creations: Co-creating Surrealist Body Horror Using the DALL-E Mini Text-to-image Generator," *Convergence*, Vol. 29, No. 4, pp. 1085-1086, 2023. DOI: 10.1177/13548565231185865
- [8] BrandBrief, Easily create advertisements with AI/big data... Kobaco launches 'Isaac', <https://www.brandbrief.co.kr/>
- [9] Kenyon, When AI Meet Screenwriting: Can AI Generate Beat Sheets and Storyboards?, <https://www.digital.kenyon.edu/>