# AIGC Multimodal Fusion Drives the Optimisation and Reshaping of Short Video Creation Process

M.D. Gao and W.Y. Leong*

*INTI International University, 71800 Negeri Sembilan, Malaysia*

In recent years, the development of artificial intelligence generated content (AIGC) technology has brought unprecedented changes to digital content creation. The continuous emergence of AIGC applications on short video platforms represented by TikTok enables content creators to create high-quality short video content at a lower cost and higher efficiency through technologies such as generative adversarial networks and natural language processing. With the widespread application of AIGC tools in the short video creation process, it is crucial to understand the impact of AIGC on content creation, creator experience, user interaction, and platform supervision. The integration of AIGC with multimodal fusion technologies is revolutionising the short video creation process. By seamlessly combining textual, visual, auditory, and contextual inputs, AIGC-driven multimodal fusion enhances creative workflows, enabling optimised content production with heightened efficiency and personalisation. This approach leverages deep learning models to interpret and synthesise diverse data streams, fostering dynamic storytelling and engaging user experiences. Key advancements include automated scene generation, real-time video editing, and adaptive narrative designs that resonate with audience preferences.

**Keywords:** AIGC; multimodal; short video; ChatGPT; process innovation

## I. INTRODUCTION

Since TikTok was officially launched in 2016, online streaming has officially entered the era of short videos (Savic, 2021). People have become accustomed to watching short videos and have begun to prefer short, fast and flat narrative content. Data shows that as of October 2018, TikTok has been downloaded by more than 800 million global users in more than 150 countries (Zeng *et al.*, 2021). The change in people's viewing habits has also led to a change in the creative thinking of creators. Compared with the previous long video creation, short videos have lower production requirements, but higher requirements for content creativity. Creators need to use a shorter duration to show more eye-catching content.

In December 2022, OpenAI launched the artificial intelligence chatbot program ChatGPT. ChatGPT can write articles similar to real people and give detailed and clear answers in many knowledge fields (Floridi & Chiriatti,

2020). It quickly gained attention. It had 1 million users 5 days after its launch and hundreds of millions of users two months after its launch. With the popularity of ChatGPT, the concept of AIGC has also aroused widespread discussion from all walks of life, and many industries have also tried to integrate AIGC into their daily work. In the field of short video creation, creators can use AIGC tools to create scripts, generate video content, produce background music, create virtual images, AI editing, etc. It can improve production efficiency, produce high-quality content at a lower cost, and achieve cost reduction and efficiency improvement.

The rise of OpenAI has proved the commercial value of the AIGC field, and many companies have also released their own AIGC tools. For example, Midjourney, an AIGC tool that generates pictures using text descriptions, SUNO, an AIGC tool that generates music, and Sora, an AIGC tool that will be launched soon for video generation (Ramesh *et al.*, 2022). These tools are all single-function AIGC platforms. Although they have strong content generation capabilities,

*Corresponding author's e-mail: waiyie@gmail.com

the content types are not rich. The future trend of AIGC must be multimodal fusion generation, using one tool to integrate multiple contents. Now there are also many companies making related attempts, such as Dreamnia, a multimodal AIGC tool that ByteDance has withdrawn, which can generate pictures, videos and music based on text descriptions (Zeng *et al.*, 2021).

This paper mainly studies the optimisation and reshaping of the short video creation process by multimodal fusion AIGC tools. It uses experimental testing, content analysis, A/B testing and other research methods to analyse how to use AIGC tools to assist short video creation, and uses the tools developed by the author to test the feasibility of AIGC replacing humans in short video creation and provides certain research references for subsequent researchers.

## A. Research Significance

With the development of short video platforms, short video content has gradually become an indispensable form of entertainment in users' daily lives (Liao, 2024). The rapid development of AIGC (AI generated content) technology, especially the progress in the field of multimodal content generation such as text, images, audio and video, has brought new opportunities and challenges to short video creation. Researching the application of multimodal fusion AIGC tools in the short video creation process is not only of theoretical significance, but also of practical application value.

### 1. Promoting innovation in short video content production

The introduction of AIGC tools can greatly optimise the short video creation process, allowing creators to improve the creative expression and diversity of content within limited time and resources (Cao *et al.*, 2023). Multimodal AIGC tools can help creators quickly generate scripts, images, music and videos, making the creation process more efficient and creative (Zhang *et al.*, 2024). Through this study, we can explore how AIGC tools can empower short video creation, help creators break through the limitations of traditional production models, and inspire more novel creative ideas.

### 2. Verify the feasibility and effect of AIGC technology in the field of short videos

Through experimental testing and content analysis of the application of AIGC in short video creation, this study aims to test the feasibility and effect of AIGC in video content generation, virtual image production, AI editing, etc., and provide data support for further promotion of AIGC technology in the industry. The study will evaluate the technical advantages and disadvantages of AIGC for short video creators and related platforms, help them use AIGC tools more reasonably, and improve content quality and user experience.

### 3. Promote the development of AIGC multimodal fusion

Most of the AIGC tools on the current market have a single function and are difficult to meet the diverse needs of creators. This study focuses on the application of multimodal fusion AIGC tools in short video creation, aiming to provide empirical analysis and practical experience for multimodal content generation. The research results will not only help promote the further development of AIGC tools but also provide reference for the design and optimisation of future multimodal AIGC platforms.

### 4. Promote cost reduction and efficiency improvement in the creative industry

The application of AIGC technology is expected to reduce the cost of short video content production and enable the creative industry to achieve a leap in content production efficiency. This study will analyse how AIGC tools can help creative content to be produced efficiently, verify its potential in reducing costs and increasing efficiency, and provide a reference for innovative content production models in the short video field.

In summary, this study places the application of multimodal AIGC tools in the actual needs and future development trends of short video creation, aiming to provide a technical application framework with practical value for the short video industry, while enriching academic research in the field of AIGC, and providing theoretical support and practical reference for subsequent researchers to explore the wider application of AIGC technology in short video content creation.

## II. CONCEPT AND DEVELOPMENT OF AIGC

AI-Generated Content (AIGC) refers to a form of digital content created automatically using artificial intelligence technologies (Brown *et al.*, 2020). AIGC encompasses multi-modal content generation, including text, images, audio, and video (Ramesh *et al.*, 2021). The core technologies behind AIGC include Generative Adversarial Networks (GANs) (Goodfellow *et al.*, 2014), Deep Learning (DL) (LeCun, Bengio & Hinton, 2015), Natural Language Processing (NLP), and Reinforcement Learning (RL) (Radford *et al.*, 2019). These technologies allow machines to generate content that aligns with human creative style and logic based on input parameters, text descriptions, or image prompts.

AIGC is marked by its "low-cost, high-efficiency, and diverse" characteristics. In traditional content creation, generating high-quality content requires significant human effort, time, and expertise. In contrast, AIGC relies on machine learning models to standardise and automate content creation processes, vastly improving the speed and diversity of generated content (Bommasani *et al.*, 2021). In Table 1, the biggest features of these contents can be summarised as high output efficiency and wide coverage (J Sun, 2024). For example, GANs learn from extensive human data to autonomously generate stylised images or videos (Zhang *et al.*, 2018). As technology advances, AIGC content increasingly approaches human-level quality and can even replace human content production in certain cases.

Table 1. The Features of AIGC

| Feature | The Details |
|---|---|
| High output efficiency | Artificial intelligence is highly reusable and can achieve large-scale, multi-threaded continuous output. With the continuous improvement of artificial intelligence algorithms and the training of large-scale data, its machine learning and creative capabilities are also constantly improving. Compared with humans, the output efficiency of artificial intelligence undoubtedly exceeds human capabilities. |
| Wide coverage | Artificial intelligence-generated content is not only efficient but also covers a wide range of fields. In addition to ordinary text and art content, artificial intelligence can also create works in the form of music, video, etc. In terms of writing, AI can produce various works, including novels, press releases, poems, etc. |

### A. Technical Principles of AIGC

The foundation of AIGC is based on AI-driven deep learning models. Below are several key AIGC technical principles:

**Generative Adversarial Networks (GANs):** GANs consist of two parts, the Generator and the Discriminator. The Generator creates content similar to real data, while the Discriminator evaluates its authenticity (Goodfellow *et al.*, 2014). Through adversarial training, the Generator continuously improves the quality of generated content to closely resemble human-created work. This technology is widely used in image generation and video synthesis, such as DeepFake and DALL-E (Ramesh *et al.*, 2021).

**Natural Language Processing (NLP):** In text generation, NLP is a core technology (Radford *et al.*, 2019). The GPT series models (e.g., GPT-3 and GPT-4) leverage massive text data training to produce high-quality, semantically coherent content. NLP models are widely applied in automated news generation and intelligent dialogue systems (Brown *et al.*, 2020).

**Reinforcement Learning (RL):** In interactive content generation, RL assists models in optimising generated content through rewards and feedback. RL is commonly applied in game development and interactive story generation (Yang *et al.*, 2021).

### B. Development Stages of AIGC

#### 1. Initial stage (1980s–early 2010s)

The concept of AIGC can be traced back to the 1980s during the early stages of AI development, when researchers began exploring machine-generated creativity (Schmidhuber, 2015). Early AI-generated content experiments focused on simple graphics and music synthesis, with output that was limited to basic mimicking.

Around 2010, with advances in computing power and deep learning algorithms, AIGC entered a new developmental phase. Deep learning enabled models to extract and mimic

human creative features from large datasets, laying the groundwork for further AIGC advances (LeCun *et al.*, 2015).

## 2. *Rapid development phase (2014–2020)*

In 2014, Goodfellow and colleagues introduced Generative Adversarial Networks (GANs), marking a significant milestone in AIGC development (Goodfellow *et al.*, 2014). The advent of GANs improved the quality and diversity of machine-generated content, leading to breakthroughs in visual content generation. GANs were applied in various domains, including facial synthesis, artistic style transfer, and 3D model generation, significantly advancing AIGC applications in both commercial and academic fields (Zhang *et al.*, 2018).

Simultaneously, the field of Natural Language Processing (NLP) also achieved notable progress. In 2018, OpenAI released the GPT-2 model, demonstrating AI's vast potential in text generation (Radford *et al.*, 2019). Based on an autoregressive language model, GPT-2 could generate logically consistent, semantically coherent long-form text. Since then, AIGC has seen wide application in automatic text generation, text summarisation, and news writing (Brown *et al.*, 2020).

## 3. *Rapid development phase (2014–2020)*

Since 2020, AIGC has entered an application diffusion phase, permeating various industries, especially in media, advertising, and entertainment. Below are a few notable technological advancements and application cases:

**Text Generation:** OpenAI released GPT-3 in 2020, which garnered attention for its remarkable text generation capabilities (Brown *et al.*, 2020). With 175 billion parameters, GPT-3 performs various tasks, from article writing and translation to code generation. Subsequently, GPT-4 further enhanced content accuracy and fluency, solidifying AIGC as a representative tool for text generation.

**Image Generation:** In 2021, OpenAI launched DALL-E, a model capable of generating images based on textual descriptions (Ramesh *et al.*, 2021). By combining NLP and GAN technologies, DALL-E generates images in various styles based on user-provided text prompts. This breakthrough overcame traditional image generation limitations and found widespread application in advertising,

artistic creation, and social media content production. In 2022, next-gen models like Midjourney emerged, further enhancing image generation quality through more powerful models and enriched datasets (Wei *et al.*, 2023).

**Video Generation:** As computational resources improved, video generation technology matured. In 2022, Meta introduced Make-A-Video, a model capable of generating short videos from text (Wang *et al.*, 2022).

**Virtual Characters and Virtual Anchors:** AIGC technology has also seen wide application in developing virtual characters and virtual anchors. Platforms like TikTok and YouTube now host numerous virtual AI-generated avatars that engage with audiences, greatly enriching platform content.

At present, AIGC has penetrated into many industries, especially those that require content output, such as text, music and image generation in the media field, AI design and special effects generation in the game industry, etc (Figure 1).
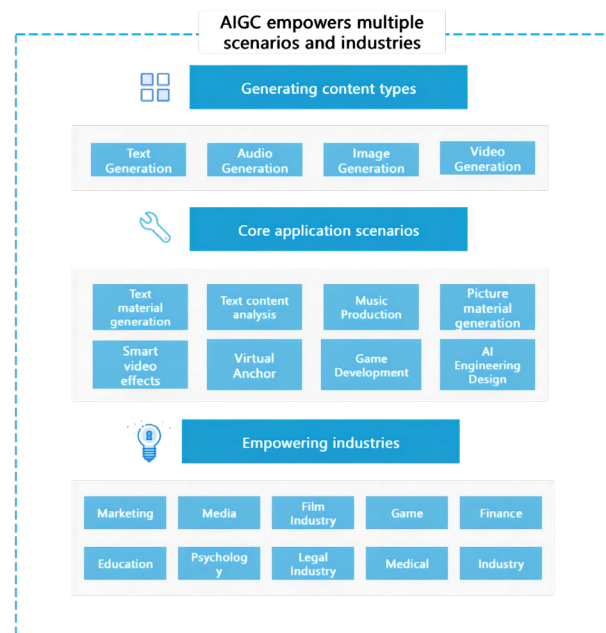


Figure 1. AIGC Empowers Multiple Scenarios and Industries.

## 4. *Future Trends of AIGC*

Currently, AIGC technology continues advancing toward multi-modal fusion, personalised content generation, and automated creation tools (Floridi & Chiriatti, 2020), as shown in Figure 2. The following outlines future trends for AIGC:

**Multi-modal Generation:** Future AIGC technologies will further achieve multi-modal integration, producing not only single-type content (e.g., text or image) but also content combining text, images, videos, and audio, enabling more interactive and immersive experiences (Yang *et al.*, 2021), as shown in Figure 3.

**Real-time Generation and Interaction:** Future AIGC content generation will become increasingly real-time, allowing users to interact with AIGC content in real-time, with applications in virtual meetings, online education, and interactive entertainment (West, 2023).

**Personalised Content Generation:** With growing user demand for personalised content, AIGC will further enable customised content to meet diverse user needs. Future AIGC content will be generated automatically based on user data and preferences, enhancing user experience and engagement (Bommasani *et al.*, 2021).

**Automated Content Production Processes:** AIGC's widespread application will further simplify content production workflows. AI will automate every phase, from ideation to editing and publishing, significantly reducing production costs and enabling individuals and small teams to create high-quality content easily.

## III. THE IMPACT OF AIGC ON SHORT VIDEO CREATION

The emergence of AIGC tools has greatly changed the creation of short videos. Due to the short and concise characteristics of short videos, users are more accustomed to watching a large number of short video contents in a period of time. This has higher requirements for the efficiency of short video production. The creators need to create as much content as possible in the shortest possible time and ensure a stable update frequency. This high-intensity creation is a huge challenge for short video creators. The emergence of AIGC tools can optimise the creation process and greatly alleviate the anxiety of creators about high-intensity creation. However, the content produced by AIGC also needs to withstand market tests in terms of dissemination and user acceptance. This article will analyse the impact of AIGC on the creation process, the impact of AIGC on the psychology and behaviour of creators, and the dissemination and user response of AIGC-generated content.
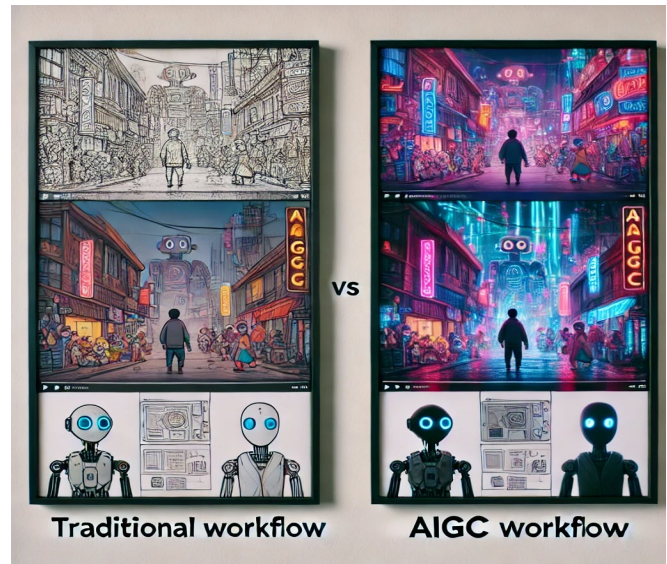


Figure 2. A side-by-side comparison of AIGC-generated vs. traditionally created video clips.
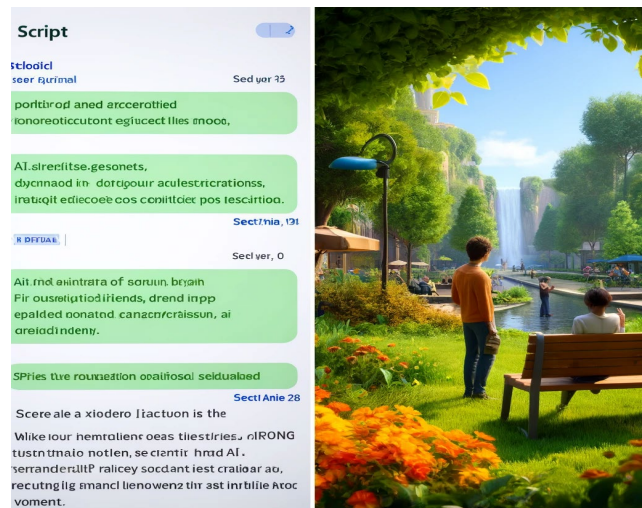


Figure 3. AIGC-generated scripts and visuals.

### A. The Influence of AIGC on the Creation Process

#### 1. Efficiency and production models

AIGC technology has been shown to substantially enhance content creation efficiency, as the automation of tasks such as content generation allows creators to significantly reduce production time, making it easier for non-professional users to produce high-quality content. This automation facilitates streamlined workflows, enabling creators to focus on ideation and storytelling rather than time-consuming manual adjustments, Figure 4. Tools for automatic subtitle generation, background music composition, and filter recommendations further reduce workload, ensuring that

creators can quickly refine their projects with minimal effort. Additionally, AIGC promotes a shift toward more dynamic and varied content types, with interactive and engaging formats like virtual characters and AI-driven avatars that allow for new ways of connecting with audiences. As a result, AIGC has broadened production models from traditional, labour-intensive methods to faster, more interactive, and automated processes, as shown in Figure 5.

Such as, virtual idol HeZ, also known as Hezi, is a widely discussed character in the entertainment industry in recent years. It debuted in the music talent show "Tomorrow's Stars" as a virtual singer without a physical form, sparking controversy and attention. HeZ's voice is synthesised by computers, its image is designed in a two dimensions anime style, and it is accompanied by voice actors, enabling it to perform singing and dancing on stage (Yuhan Chu,2024).



Figure 4. Virtual Idol HeZ in "Tomorrow's Stars".

Interestingly, when discussing whether AIGC tools can truly improve creative efficiency, some creators feel that AIGC tools may lead to inefficiency. These inefficiencies manifested as either direct or indirect. Direct inefficiencies are those associated with the shortcomings of AIGC tools outputs. For example, incorrectly worded content or hallucinated information require human intervention to resolve, which may cancel out the efficiencies gained (Osadchaya *et al.*, 2024).
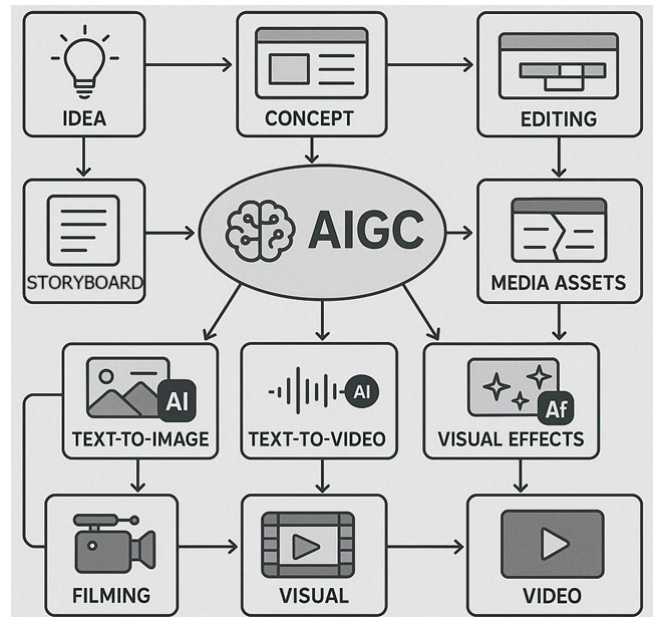


Figure 5. The workflow diagram illustrating the integration of AIGC in video production.

Indirect inefficiencies, on the other hand, are not immediately linked to AI outputs themselves but rather to the broader consequences of AIGC use. First, participants feared problems related to intellectual property and data security, as they could not be certain that protected or sensitive information fed into the AI would not be exposed to others. Indeed, some companies have already fallen victim to this. At Samsung, there have been three separate instances of employees accidentally leaking sensitive data to ChatGPT (Mauran, 2023).

### 2. Creative expression and content diversity

AIGC technology also empowers creators by expanding the possibilities for creative expression. With AI's assistance, creators can quickly generate content that aligns with different themes, styles, and visual effects, allowing for highly customised creative outputs that cater to diverse audience preferences. For instance, advanced AI filters and special effects can generate unique visual elements, which help create richer and more varied visual narratives (Zhang Y, 2024). However, the content produced by AIGC tools is knowledge of answers obtained from limited training data, which may hinder the ability of creators to generate truly novel ideas and simply perpetuate old content that existed before. Some researchers warn that the widespread use of AIGC tools may inadvertently contribute to content

homogenisation. Standardised AI-generated templates could lead to repetitive styles, reducing originality and the distinctive quality of content. Thus, while AIGC enhances creative potential, it also presents a challenge for creators to maintain the uniqueness of their work amidst increased automation (Liu & Chen, 2023).

## B. AIGC's Impact on Creators' Psychology and Behaviour

### 1. Creator experience and satisfaction

By simplifying and accelerating the content creation process, AIGC tools significantly enhance creator experience and satisfaction. Creators using AIGC tools report feeling more productive and less constrained by technical limitations, allowing them to focus on the artistic and strategic aspects of their work (Davis, 2022). This is particularly beneficial in the early stages of content production, where creators often struggle with ideation and initial design. However, despite the positive impact on workflow efficiency, some creators have expressed concerns over the authenticity and originality of AI-generated content. According to Wang and Hong (2024), a growing number of creators are wary of AIGC potentially diluting their personal style, as they fear over-reliance on AI could result in work that feels less unique or personally meaningful. This dichotomy suggests a need for balance between leveraging AIG's capabilities and retaining individual creative autonomy.

### 2. Creative freedom and technological dependency

The highly automated nature of AIGC raises questions about its impact on creative freedom and dependency on technology. Scholars suggest that the templated approach of AIGC tools might inadvertently limit creative exploration by encouraging standardised design elements (Park & Lee, 2022). This reliance on pre-set models and templates may lead to a dependency on technology for creative output, possibly stifling creators' willingness or ability to experiment with new ideas. Furthermore, the ease of AI-driven generation could shift the role of the creator from an originator of ideas to a curator or editor of AI-generated content (Leong, 2024a). This shift has implications for the creative process itself, as creators may increasingly become

intermediaries, overseeing and refining AI output rather than engaging directly with content creation (Leong, 2025a). While this might streamline workflow, it could also restrict the hands-on, iterative process that is often crucial to artistic innovation, Figure 6.
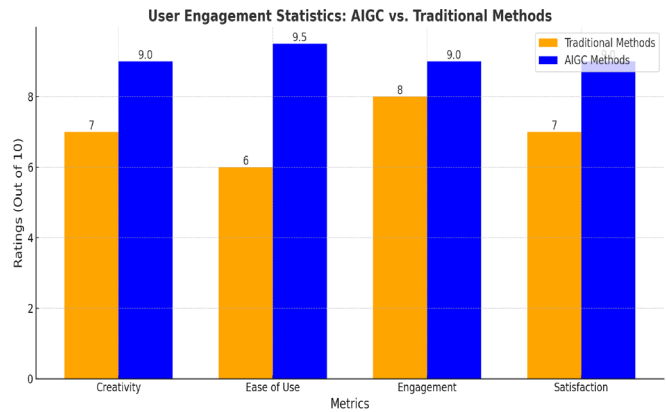
Figure 6. User engagement statistics for AIGC vs. traditional methods.

## C. Content Dissemination and User Response to AIGC-Generated Content

### 1. Content dissemination effectiveness

AIGC-generated content demonstrates notable efficacy in terms of dissemination, often achieving higher levels of engagement than traditional content. Mohamed *et al.* (2022) found that AIGC-generated videos, due to their interactive and innovative qualities, generally garner more views, likes, and shares. These high engagement levels are especially evident in AIGC-generated virtual avatars and AI-hosted segments, which attract younger audiences and foster a sense of connection and loyalty (Mingda & Wai Yie, 2024). The ability to instantly generate and tailor content for specific audience segments enhances its potential for viral reach and immediate impact. These findings underscore the role of AIGC in amplifying content distribution and engagement, as it meets the audience's demand for interactive, real-time experiences.

Of course, the dissemination effect of AIGC content also depends to a large extent on the quality of the content. The requirement for content quality to some extent represents the user's demand for realism, that is, whether AIGC content can be close to the look and feel of artificial content (Leong, 2025e). For example, the latest Christmas advertisement

launched by Coca-Cola Company, the 1-minute video is all generated by AI and produced with manual polishing, Figure 7. The visual presentation of the entire advertisement has a strong sense of AI, which is far from artificial videos, and the effect is not satisfactory.



Figure 7. Coca-Cola Christmas ad generated by AI.

### 2. User interaction and acceptance

Audience reception of AIGC-generated content is a focal point in current research, as user interaction and perception largely determine the success and longevity of AIGC applications in content creation. Some researchers report that while AIGC-generated content enjoys high interaction rates, partly due to its novelty, there are concerns about authenticity. Some viewers perceive AI-generated virtual hosts and characters as lacking emotional depth, which may affect their trust in the content and its creator. Such scepticism could have implications for brand image, as the perceived artificiality of AIGC content may undermine viewer trust and affect their overall experience. Balancing the benefits of AI's efficiency with genuine, relatable content is crucial for maintaining a loyal and engaged audience, as consumer preferences shift towards authenticity in digital interactions.

In studies on user acceptance, human-made artistic media content is perceived as more authentic and engaged with more, Figure 8. Among them, perceived authenticity does have a significant effect on the engagement for both human-made and AI-generated content (Bosch, 2024). The changes in authenticity and engagement sometimes come from whether people label AIGC-generated content, which also indirectly proves that at this stage, the difference in

perception between some AIGC content and artificial content is not particularly obvious, Table 2.

Table 2. Comparative Key Performance Indicators (KPIs)

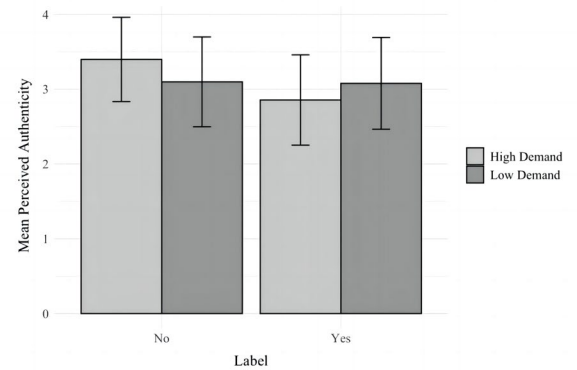| KPI | Traditional Methods | AIGC Methods |
|---|---|---|
| Time Efficiency (Hours/Video) | 8 | 4 |
| Cost Efficiency (USD/Video) | 1000 | 600 |
| Creativity (Rating/10) | 7 | 9 |
| Ease of Use (Rating/10) | 6 | 9.5 |
| Engagement (Rating/10) | 8 | 9 |
| Satisfaction (Rating/10) | 7 | 9 |



Figure 8. Estimated Marginal Means of AIGC Perceived Authenticity (Bosch, 2024).

## IV. RESEARCH STATUS AND SHORTCOMINGS

### A. Research Status

As a cutting-edge content generation technology, AIGC has garnered extensive attention in recent years, particularly in its applications for short video content creation. The existing literature on AIGC primarily focuses on the following aspects:

### 1. Development of AIGC technology and algorithm optimisation

Numerous studies concentrate on optimising core algorithms of AIGC technology, including advancements in Generative Adversarial Networks (GAN), autoregressive models, and autoencoders to enhance the quality and diversity of generated content. For example, Goodfellow *et al.* (2014) introduced GAN, which significantly advanced

AIGC's capabilities in image and video generation. With the release of OpenAI's GPT series, particularly GPT-3 and GPT-4, AIGC's potential in text generation has greatly expanded, providing theoretical support and technical foundations for widespread applications in content generation. Additionally, some studies explore the optimisation of AIGC models in terms of generation speed, stability, and multi-modal integration, enabling the production of diverse content closely aligned with human creative styles (Brown *et al.*, 2020).

### 2. Application of AIGC tools and user experience

In terms of application, several studies approach AIGC tools from the perspective of both users and creators, examining specific uses and user experiences in content creation (Leong, 2025d). For instance, some studies gather feedback from content creators and industry practitioners through interviews and surveys to analyse how AIGC tools are utilised. Short video creators, for example, report that AIGC tools facilitate rapid script generation and visual effects, significantly reducing production costs and improving efficiency (Chen *et al.*, 2022). Furthermore, features such as automatic subtitle generation and smart filter recommendations decrease manual workload, allowing creators to focus more on creative design.

### 3. User acceptance and engagement

Several studies investigate user acceptance and engagement with AIGC-generated content through questionnaires and experimental design. For instance, designed experiments to evaluate user interaction and feedback on AIGC-generated content, finding that users generally have a positive attitude toward AIGC short videos, especially in terms of entertainment value and novelty. These studies often use metrics such as interaction rate, satisfaction scores, and viewing duration to quantify user acceptance of AIGC content. Research also indicates that AIGC-generated virtual characters and hosts are particularly appealing to younger audiences, enhancing user retention (Mingda & Wai Yie, 2024).

### 4. Policy, regulation and ethical considerations for AIGC content

With the increasing application of AIGC, issues concerning content originality, copyright, and ideological implications have attracted researchers' attention (Leong, 2025b). Questions such as whether AIGC-generated content qualifies as original work and how its copyright should be assigned—whether to the technology provider or user—are focal points in academic and industry discussions. Some researchers explore the ideological impact of AIGC content, especially regarding the potential influence of misinformation and content homogenisation on user values (Mohamed *et al.*, 2024). Additionally, studies have proposed regulatory frameworks and ethical guidelines for AIGC, suggesting enhanced content screening to ensure compliance with ethical standards and reduce the spread of misleading information (Park & Lee, 2022).

### 5. Application of AIGC in the short video industry

In recent years, AIGC applications in short video creation have increasingly garnered academic attention. Researchers have examined how AIGC can drive innovation in short video production by facilitating AI-generated scripts, audio, effects, and other tools that make the production process more efficient and cost-effective. Some researchers believe that AIGC enables short video creators to implement a "rapid creation-rapid release" model, significantly improving creative efficiency and promoting diversity and interactivity in content. Particularly on user-generated content (UGC) platforms, AIGC tools provide ordinary users with professional-level creative capabilities, enriching the content ecosystem and enhancing user engagement.

With advancements in multi-modal content generation, AIGC research has expanded to explore integrated applications of multi-modal tools for generating text, image, audio, and video content. Some studies discuss how multi-modal AIGC tools can enhance creative effects in short videos, such as generating videos from text, automatically creating background music, or adding emotional filters to videos. The emergence of multi-modal AIGC tools enables AI support throughout the short video creation process, offering creators greater flexibility and productivity. This trend suggests that multi-modal integration may become a

crucial direction for AIGC's future development, though research on this topic remains in its early stages.

## B. Insufficient Research

Although existing studies have provided valuable insights into the application of AIGC in the field of short videos, there are still some shortcomings that need further in-depth exploration.

### 1. Lack of empirical research on the specific impact on the creative process

Although some studies have involved the impact of AIGC on content creators, most of them remain at the qualitative analysis level and lack systematic experimental data to verify how AIGC tools specifically affect the creation process. For example, whether AIGC tools significantly improve the efficiency of short video creation, how to change the structured process of creation, and whether they really reduce the cost of creation have not been fully verified. Future research can use empirical methods such as experimental methods and content analysis to examine in detail the specific impact of AIGC on each link of the short video creation process.

### 2. Limited research on multimodal AIGC tools

Currently, most of the research on AIGC still focuses on the application of single-function tools, such as text generation or image generation, ignoring the potential of multimodal AIGC tools. However, the integrated application of multimodal tools can achieve comprehensive generation from scripts, audio to video, thus bringing a new model for short video content production. Therefore, more research is needed to focus on the integration of multimodal AIGC tools and their practical application in short video creation, in order to explore the application potential of this technology in content creation and cross-platform publishing.

In subsequent research on multimodal AIGC tools, we can consider establishing relevant mathematical models and obtaining the work efficiency coefficient of performance improvement through calculation of experimental data to obtain more comprehensive research results, Figure 8.

Such as, for quantify this efficiency gain, let the total time for traditional content creation be $T_{traditional}$, with n

different types of content (e.g., text, images, audio, and video), where the time required for each content type in the traditional process is represented by $T_i$. The total time for traditional content creation is then given by

$$T_{traditional} = \sum_{i=1}^{n} T_i \qquad (1)$$

When using AIGC multi-modal tools, the content generation process is automated and parallelised, reducing the total time to

$$T_{AIGC} = \alpha \cdot T_{traditional} = \alpha \cdot \sum_{i=1}^{n} T_i \qquad (2)$$

where α is the efficiency factor of the AIGC tool, with $0 < \alpha \le 1$. The smaller the value of α, the greater the time-saving effect of the AIGC tool, with values approaching zero indicating extremely high efficiency.

Further analysis yields the efficiency improvement ratio E, which represents the proportion of time saved by using AIGC tools compared to traditional processes. This is calculated as

$$E = \frac{T_{traditional} - T_{AIGC}}{T_{traditional}} = 1 - \alpha \qquad (3)$$

If $\alpha = 0.5$, the AIGC tool reduces creation time by 50%, or $E = 0.5$, indicating a 50% increase in content creation efficiency. Such a mathematical model can more intuitively reflect the improvement in the efficiency of short video creation by multimodal AIGC tools.
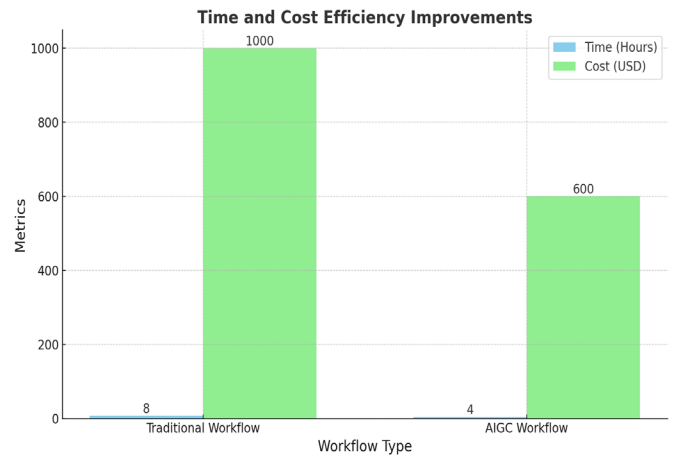


Figure 9. Time and cost efficiency improvements between traditional and AIGC workflows.

*3. Lack of research on the long-term effects of AIGC content*

Most existing studies focus on the short-term dissemination effects of AIGC-generated content, such as the audience's initial interaction rate, clicks, and sharing volume, but rarely explore the long-term impact of AIGC content on user behaviour and user loyalty. For example, will AIGC-generated content gradually lose its appeal over time? Will such content cause trust issues among viewers? Answers to these questions will help to fully understand the depth and breadth of the impact of AIGC content on short video platforms.

*4. Insufficient comprehensive analysis of user experience and satisfaction*

Although some studies have focused on users' acceptance of AIGC content, the analysis of users' deep experience and emotional response is still relatively weak. Especially in the field of short videos, how users emotionally resonate with AIGC-generated virtual images, automatically generated plots, and other content, and whether AIGC content can bring a higher sense of immersion, are all key issues worth exploring in future research. Further in-depth analysis of users' emotional feedback on AIGC content will provide data support for short video platforms to optimise user experience and improve content quality.

In summary, although existing research has laid a certain foundation for the application of AIGC in the field of short video creation, there is still room for in-depth exploration in terms of empirical data, multimodal tools, long-term effects, user emotional experience (Leong, 2025c). Future research can be further expanded in these aspects to more comprehensively reveal the potential and challenges of AIGC in short video content creation and provide richer references for industry practice and academic research.

## V. CONCLUSION

The rapid advancement of AI-generated content (AIGC) technology has brought significant changes to the content creation landscape, particularly within the short video industry. Through automated processes, multi-modal integration, and algorithmic improvements, AIGC enables creators to streamline production, diversify content formats, and enhance creative expression, addressing both practical and artistic challenges in short video creation. The current research indicates that AIGC can substantially increase production efficiency, making high-quality content creation accessible to both professional and non-professional creators. Tools for scriptwriting, automatic editing, and virtual character generation demonstrate the potential of AIGC to simplify the creative process while expanding the types of content available to audiences.

Future research on the application of AIGC in short videos can be approached from the perspective of multimodal fusion, integrating text, image, audio, and video generation capabilities into a single platform. This fusion not only enriches storytelling options but also enables creators to produce immersive and interactive content that resonates with different audiences. Multimodal AIGC tools are uniquely positioned to meet the changing needs of digital platforms because they enable a more flexible and dynamic creative process. Specifically, multimodal integration enables creators to leverage different types of content simultaneously, enhancing narrative depth and audience engagement. For example, tools that can generate text-based scripts, visual effects, soundtracks, and video clips can significantly reduce production time and increase creative output, which is particularly beneficial for time-sensitive platforms such as TikTok and Instagram.

In conclusion, while AIGC offers transformative capabilities for content creation, its successful integration into the short video industry will require a balanced approach. Future research should aim to develop more robust multi-modal tools, explore the long-term effects of AIGC on user behaviour and engagement. As AIGC continues to evolve, it has the potential to redefine content creation, enabling creators to push the boundaries of storytelling while meeting the diverse needs of digital audiences.

## VI. REFERENCES

Bommasani, R, Hudson, DA, Adeli, E, Altman, R, Arora, S, von Arx, S, ... & Liang, P 2021, 'On the opportunities and risks of foundation models', arXiv preprint arXiv:2108.07258.

Bosch, MJ 2024, 'Generation Z's Perceived Authenticity and Engagement with AI-Generated vs. Human-Made Artistic Media Content: The Roles of Content Labeling and Attitude toward Regulation', Bachelor thesis, University of Twente.

Brown, T, Mann, B, Ryder, N, Subbiah, M, Kaplan, J, Dhariwal, P, ... & Amodei, D 2020, 'Language models are few-shot learners', arXiv preprint arXiv:2005.14165.

Cao, Y, Li, S, Liu, Y, Yan, Z, Dai, Y, Yu, PS & Sun, L 2023, 'A comprehensive survey of AI-generated content (AIGC): A history of generative AI from Gan to ChatGPT', arXiv preprint arXiv:2303.04226.

Chu, Y 2024, 'A Communication Research on Reshaping the Content Ecosystem of New Media Platforms Based on AIGC Technology', Frontiers in Computing and Intelligent Systems, vol. 9, no. 3, pp. 24-27.

Floridi, L & Chiriatti, M 2020, 'GPT-3: Its nature, scope, limits, and consequences', Minds and Machines, vol. 30, pp. 681-694.

Goodfellow, I, Pouget-Abadie, J, Mirza, M, Xu, B, Warde-Farley, D, Ozair, S, ... & Bengio, Y 2014, 'Generative adversarial nets', Advances in Neural Information Processing Systems, vol. 27, pp. 2672-2680.

LeCun, Y, Bengio, Y & Hinton, G 2015, 'Deep learning. Nature, vol. 521, no. 7553, pp. 436-444.

Leong, WY 2024a, 'Integrating AIGC for Automated Post-Production, 2024 RIVF International Conference on Computing and Communication Technologies (RIVF), 21 - 23 December 2024.

Leong, WY 2025a, 'Machine Learning in Evolving Art Styles: A Study of Algorithmic Creativity', Engineering Proceedings', 2025, vol. 92, no. 1, p. 45.

Leong, WY 2025b, 'Dynamic Storytelling and Personalized Film Narratives Using AIGC', ASM Science Journal, vol. 20, no. 1, 2025.

Leong, WY 2025c, 'AI-Driven Optical Illusions: Innovations in Perceptual Art and Design', Innovation on Design and Culture, vol. 4, no. 1, pp. 1-14. doi: 10.35745/idc2025v04.01.0001.

Leong, WY 2025d, 'Ethical and Philosophical Perspectives on Artificial Intelligence-Generated Art', International Journal of Social Sciences and Artistic Innovations, vol. 5, no. 2, 2025.

Leong, WY 2025e, 'Digital Reconstruction of Historical Artworks Using Artificial Intelligence', 2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI), Indonesia, 21 January 2025.

Liao, M 2024, 'Analysis of the causes, psychological mechanisms, and coping strategies of short video addiction in China', Frontiers in Psychology, vol. 15, p. 1391204.

Mauran, C 2023, 'Whoops, Samsung workers accidentally leaked trade secrets via ChatGPT', Mashable, <https://mashable.com/article/samsungchatgpt-leak-details>.

Mingda, G & Wai Yie, L 2024, 'Research on the Application of AIGC in the Film Industry', Journal of Innovation and Technology, vol. 2024, no. 22, pp. 1-12.

Mohamed, E, Osman, ME & Mohamed, BA 2024, 'The impact of artificial intelligence on social media content', Journal of Social Sciences, vol. 20, no. 1, pp. 12-16.

Osadchaya, E, Marder, B, Yule, JA, Yau, A, Lavertu, L, Stylos, N, ... & AlRabiah, S 2024, 'To ChatGPT, or not to ChatGPT: Navigating the paradoxes of generative AI in the advertising industry', Business Horizons.

Radford, A, Wu, J, Child, R, Luan, D, Amodei, D & Sutskever, I 2019, 'Language models are unsupervised multitask learners', OpenAI Blog, vol. 1, no. 8, p. 9.

Ramesh, A, Dhariwal, P, Nichol, A, Chu, C & Chen, M 2022, 'Hierarchical text-conditional image generation with clip latents', arXiv preprint arXiv:2204.06125, vol. 1, no. 2, p. 3.

Ramesh, A, Pavlov, M, Goh, G, Gray, S, Voss, C, Radford, A, ... & Sutskever, I 2021, 'Zero-shot text-to-image generation', arXiv preprint arXiv:2102.12092.

Savic, M 2021, 'Research perspectives on TikTok & its legacy apps from musically to TikTok: Social construction of 2020's Most downloaded short-video app', International Journal of Communication, vol. 15, p. 22.

Schmidhuber, J 2015, 'Deep learning in neural networks: An overview', Neural Networks, vol. 61, pp. 85-117.

Sun, J 2024, 'AIGC Fusion Exploration: The Intersecting Path of Digital Humanities and Artificial Intelligence', J. Electrical Systems, vol. 20, no. 2, pp. 327-335.

Wang, X, Hong, Y & He, X 2024, 'Exploring artificial intelligence generated content (AIGC) applications in the metaverse: Challenges, solutions, and future directions', IET Blockchain, vol. 4, no. 4, pp. 365-378.

Wang, Y, Zhang, L & Zhang, L 2022, 'Emerging applications of deepfakes in social media: A survey', IEEE Transactions on Knowledge and Data Engineering.

Wei, Q, Li, J & Zhang, Y 2023, 'Public emotional dynamics toward AIGC content generation across social media platform', arXiv preprint arXiv:2312.03779.

Yang, G, Zhang, M & Wang, X 2021, 'The future of generative AI in content creation: Challenges and opportunities', IEEE Internet of Things Journal, vol. 8, no. 23, pp. 2347-2355.

Zeng, J, Abidin, C & Schäfer, MS 2021, 'Research perspectives on TikTok & its legacy apps research perspectives on TikTok and its legacy apps—introduction', International Journal of Communication, vol. 15, p. 12.

Zhang, H, Goodfellow, I, Metaxas, D & Odena, A 2018, 'Self-attention generative adversarial networks', arXiv preprint arXiv:1805.08318.

Zhang, Y 2024, 'The Impact of AI-Generated Content on China's New Media Visual Creative Practitioners', in SHS Web of Conferences, EDP Sciences, vol. 207, p. 03002.

Zhang, Z, Li, X, Sun, W, Jia, J, Min, X, Zhang, , ... & Zhai, G 2024, 'Benchmarking AIGC Video Quality Assessment: A Dataset and Unified Model', arXiv preprint arXiv:2407.21408.