

Emotion Recognition-Based AIGC Tools and Methods in Virtual Experience Applications: An Analysis

Yu Yao ¹
Kaiqiang Sun ^{2,*}
Long He ³
Hongchao Wei ⁴
Mingfang Zhang ⁵
Yujing Yang ³
Tong Sun ³

*Corresponding author

¹ Hongik University, Seoul, South Korea, yaoyu@g.hongik.ac.kr

² Hongik University, Seoul, South Korea, sunkaiqiang@g.hongik.ac.kr

³ Shenyang Qiyuan Technology Co., Ltd., Shenyang, China, Howard@originxr.cn

⁴ Dali University, Dali, China, Doctor, 912152@dali.edu.cn

⁵ Dalian University, Dalian, China, Lecturer, zhangmingfang@dlu.edu.cn

ABSTRACT

With the development of immersive technologies such as virtual reality, virtual experiences are becoming increasingly popular in fields such as education, healthcare and exhibitions. This places higher demands on the emotional adaptability and interactive expressiveness of content generation. The development of generative artificial intelligence (AIGC) technology has endowed virtual experience systems with the capability to generate efficient, multimodal content. This paper focuses on the 'emotion cue-driven AIGC generation mechanism', using 'anxiety–calm' as a representative emotion type. It then constructs an experimental framework covering both text and image tasks. The paper compares and analyses the generation performance of seven mainstream AIGC platforms. The study involves designing structured prompt templates and proposing a three-dimensional expert evaluation system comprising 'generative richness', 'emotional fit', and 'detail expressiveness', taking into account the contributions of language construction and visual rendering to virtual immersion. Experimental results show that Google Gemini and ChatGPT 4o demonstrate strong emotional expression and scene construction capabilities in text generation, while Midjourney excels in image generation tasks. It accurately responds to prompts and generates concrete, stylistically distinct visual content, making it suitable for emotion-driven virtual experience design. However, some platforms still exhibit significant gaps in detail representation and emotional consistency. This study proposes a fixed, prompt-driven mechanism based on emotion recognition which demonstrates cross-platform adaptability and virtual experience transfer value. It provides theoretical support and methodological pathways for the future application of AIGC in immersive education, virtual exhibitions and affective computing scenarios.

Keywords: AIGC, emotion prompting, virtual experience, multimodal generation, immersive content.

INTRODUCTION

Research Background

The rapid development of immersive technologies such as virtual reality (VR), augmented reality (AR) and mixed reality (MR) has seen virtual experiences gradually expand from professional applications into the mass market (Liberatore & Wagner, 2021). Leveraging core advantages such as immersion, interactivity and visualisation, these technologies are penetrating an increasing number of industries, including design, education, healthcare, entertainment and retail (Pavithra *et al.*, 2020). To further enhance immersion, technologies such as head-mounted displays (HMDs), motion tracking systems and virtual interaction interfaces are being employed to enable users to transcend physical space constraints and immerse themselves in highly realistic digital environments. This transformation expands the boundaries of content presentation and drives the development of user-centred experience design principles, empowering users to actively participate in virtual content experiences (Yang *et al.*, 2023). The emergence of artificial intelligence-generated content (AIGC) technology has further enhanced the richness and dynamic adaptability of virtual experiences (Wu *et al.*, 2023). AIGC technology relies on large-scale deep learning models to automatically generate multimodal content, such as text, images, audio and three-dimensional scenes. This provides an efficient mechanism for producing content for virtual experiences (Wan *et al.*, 2024). The rapid development of emotion recognition technology has enabled the personalisation and optimisation of virtual experience content by providing real-time perception and analysis of multimodal data such as facial expressions, voice tone and physiological signals. This technology can capture users' psychological changes in virtual environments and convert them into interactive inputs (Pal *et al.*, 2021). However, despite significant progress in emotion recognition and AIGC technology,

their deep integration remains in its exploratory phase. Traditional AIGC systems generate content based on prompts, but lack the ability to link user emotions. It is crucial to effectively map users' typical emotional states to AIGC input parameters and validate the regulatory effects of different emotional prompts on generated content in order to advance intelligent virtual experiences and emotional recognition-driven key interactions (Song, 2024). Therefore, integrating emotion recognition and AIGC can enrich cross-disciplinary research between affective computing and generative AI and promote the development of empathetic human-machine interaction modes. At the application level, this integration mechanism can be widely applied in immersive therapy, digital education, virtual exhibitions and intelligent interactive entertainment. This would enhance the personalisation and emotional resonance of virtual experiences and have significant application value and industrial prospects.

Research Objectives and Core Issues

This study focuses on the controllability and adaptability of generative artificial intelligence (AIGC) for creating emotion-driven content, with applications in immersive virtual experiences. The aim is to use emotion recognition-based prompt generation strategies to guide AIGC systems in generating text and image content with specific emotional styles, thereby enhancing users' emotional immersion and personalised experiences in virtual environments.

To this end, the study sets out three objectives:

- (1) Constructing emotion-recognising prompt language suitable for virtual experience tasks, and clarifying its semantic regulation and style guidance mechanisms in text and image generation tasks;
- (2) To develop a three-stage 'prompt generation - evaluation' process model to systematically assess the response capabilities of mainstream AIGC platforms to typical emotion-recognition prompts and the quality of emotional expression in generated content;
- (3) To explore the applicability and transferability of emotion-recognition-prompt-driven mechanisms in virtual experience scenarios, providing a theoretical basis and experimental paradigms for the future integration of intelligent narrative, interaction design and affective computing.

Through platform comparison experiments and expert-based evaluations, this study validated the effectiveness of structured prompts in controlling emotions in multimodal generation tasks and explored their feasibility and practical value as a central mechanism for content generation in virtual experience systems.

THEORETICAL BASIS AND TECHNICAL OVERVIEW

Evolution of AIGC Technology and Integration with Virtual Experiences

Artificial Intelligence Generated Content (AIGC) has emerged as a significant development in generative artificial intelligence in recent years. Relying on large-scale deep learning models, it can efficiently generate multimodal content, including text, images, audio, video and three-dimensional environments (Summaira *et al.*, 2021). Initially, AIGC focused on generating single-modal content, such as image-generation adversarial networks (GANs) and the WaveNet model for speech. The primary objective of this phase was to improve the quality and realism of the generated content, with applications primarily in image restoration, style transfer, and speech cloning. However, the content generation process was static and lacked interactivity, making it difficult to adjust the generated content dynamically based on user input (Masood *et al.*, 2023). However, the widespread adoption of Transformer architectures and the rapid development of large language models (LLMs) have led to AIGC entering a cross-modal generation phase, in which language serves as the core control parameter (Cao *et al.*, 2023). Representative models such as DALL-E, Stable Diffusion and Midjourney can generate images, audio or other modal content based on natural language prompts, enabling users to actively participate in the creative process (Kwon, 2024). Such systems have demonstrated strong content adaptability and creative flexibility in virtual exhibitions, game design and personalised education.

Currently, AIGC is evolving towards multimodal collaborative generation, with the ability to jointly model multiple inputs, such as text, images and speech (Cao *et al.*, 2025). For instance, models like GPT-4 and Gemini can integrate multimodal information within contextual frameworks to produce more immersive and interactive content. This stage emphasises the unity of 'semantic consistency', 'user context awareness' and 'generation control', enabling AIGC to transition from a tool-based role to an intelligent, collaborative entity for virtual content creation (Rane *et al.*, 2024). The introduction of AIGC in virtual experience applications has greatly improved content generation efficiency and personalisation levels (Cao *et al.*, 2023). Users can construct complex environments, virtual characters and interactive storylines from brief descriptions, transitioning from passive experience to active co-creation. However, although prompt control mechanisms have enhanced content adaptability to some extent, current AIGC systems generally lack the ability to perceive and respond to users' psychological states, particularly emotional ones (Puerta-Beldarrain *et al.*, 2025). Therefore, in virtual experience environments, AIGC can be used to dynamically construct highly personalised, immersive content such as contextual narratives, virtual character behaviour and scene atmosphere adjustment. For instance, in virtual exhibitions or educational spaces, changes in user emotions can prompt the system to provide suitable responses in terms of voice narration or image presentation, thereby enhancing user immersion and psychological resonance. Through AIGC technology, virtual experience systems can produce multimodal outputs that align with users' psychological states, thereby endowing the experience with the triple characteristics of 'interactivity, dynamism and emotional responsiveness' (Leong, 2023). AIGC is no longer merely a content generation tool, but has become a critical, real-time, content-driving mechanism within virtual experience scenarios.

Emotion Recognition Prompt Word Setting Basics

In the 'emotion-driven AIGC content generation' experimental framework constructed in this study, the design of emotion-recognition-representative prompt phrases is a key factor affecting both the style of the generated content and how

emotions are presented. The selection of emotion types is based on Russell's (1980) two-dimensional 'valence–arousal' emotion model. Valence measures the pleasantness or 'good–bad' direction of emotions. High-valence emotions refer to positive, pleasant emotional states, such as happiness, calmness, and satisfaction, while low valence corresponds to negative, unpleasant emotional experiences, such as anger, anxiety, and sadness. Arousal reflects the level of activation or physiological excitement experienced by an individual in a specific emotional state. High-arousal emotions, such as anger, excitement, and anxiety, manifest as a state of high physical and mental activation, while low-arousal emotions, such as calmness, fatigue, and depression, reflect a state of relatively low energy, relaxation, or inhibition. This model effectively illustrates the spatial relationship between different emotions and provides an important theoretical basis for designing the prompts and experimental variables in this study (Russell, 1980). It is widely used in emotion recognition and regulation research, effectively characterising the distribution of complex emotional states at psychological and behavioural levels. Therefore, 'anxiety' and 'calmness' were selected as representative emotions, leveraging their opposing attributes in the model space to enhance the distinguishability and detectability of prompts in the experimental task. In terms of constructing emotion prompts, this study drew inspiration from recent practices in prompt engineering (Reynolds & McDonell, 2021) by adopting a three-layer combination structure of 'emotion keywords + scene description + perceptual description'. This structure balances semantic clarity and stylistic refinement to facilitate the model's accurate understanding of emotional intent. For instance, 'a quiet lakeside cabin at sunset with a peaceful atmosphere and warm light' is used to generate text and image content depicting a tranquil atmosphere, whereas 'a chaotic city street at night with flashing red lights, crowds and tension' conveys a tense and anxious scene. This type of prompt design demonstrates good stability of expression and adaptability to different platforms in both text and image tasks. Considering the differences in the parsing mechanisms of natural language prompts across platforms, the study focuses particularly on the consistency of prompt performance in cross-modal tasks. Previous studies have shown that mainstream text generation models (e.g. GPT and Gemini) and image generation models (e.g. Midjourney and DALL·E) can handle multi-level semantic modification structures (Hao *et al.*, 2023). Therefore, the prompts in this study were constructed using neutral, descriptive language to avoid specific stylistic terms or platform-specific instructions, thereby minimising platform response bias. To further ensure the expressive stability and cross-platform adaptability of prompt phrases for emotion recognition in emotion control tasks, this study references empirical research in the field of prompt engineering (Gu *et al.*, 2023) and conducts multi-platform pre-experiments on candidate prompt phrases. The generated content was evaluated based on basic criteria such as semantic consistency, style stability and the effectiveness of expressing emotions. These criteria served as input templates for the main experimental phase.

In order to meet the requirements of typical application scenarios for virtual experiences, such as virtual museums, therapeutic gardens, stress simulation training or soothing interactive interfaces, the system prompts users to input appropriate emotion-oriented words in order to achieve emotion regulation. The AIGC platform actively generates style-matched content to construct a highly empathetic, immersive virtual space for users. Artificial intelligence-generated content (AIGC) is an important development area in generative artificial intelligence 2.3 Emotion recognition-driven prompt language mechanism for AIGC generation.

Previous studies have shown that in image and text generation tasks, the design of emotion type representative prompts not only determines the output style but also plays a crucial role in the accuracy of emotion recognition expression (He *et al.*, 2025). Based on this, this study proposes an emotion-prompt-driven generation process mechanism centred on 'emotion recognition prompts', comprising four stages: emotion type setting; emotion recognition prompt generation and input mapping; content generation and platform response; output evaluation and feedback.

Stage 1: Emotion type setting. The first stage involves setting the emotion type. For the experiment, this study selected 'anxiety' and 'calmness' as the core emotion recognition categories. Based on Russell's (1980) valence–arousal two-dimensional emotion model, the emotion space was divided into two quadrants: negative valence–high arousal and negative valence–low arousal. This configuration enhances the observability of platform response differences and controls variable influences (Russell, 1980).

Stage 2: Construction and Input Mapping of Emotion Recognition Representative Prompt Generators. Referring to the structured prompt generation strategy proposed by Reynolds and McDonell (2021), the prompts in this study adopt a 'three-part structure': emotion keyword + scene description + style modifier (Reynolds & McDonell, 2021). In text generation, for example, the prompt 'A peaceful evening at the lakeside with warm orange light and a gentle breeze' can guide the model to generate content with a calm atmosphere. In image generation, 'A crowded urban alley at night with flickering lights and a tense mood' can guide the visual expression of an anxious atmosphere. This structure has been shown to be semantically compatible and controllable in terms of emotion on platforms such as Midjourney and ChatGPT (Kreps *et al.*, 2023).

Stage 3: Content generation and platform response. After standardising the emotion-recognised prompt phrases, they are input into multiple mainstream AIGC platforms, including ChatGPT, Google Gemini, Midjourney and DALL·E. The response behaviour to the same prompts in text and image tasks is then observed. This verifies that, due to differences in model structure, training data and style control mechanisms, different platforms will exhibit different emotional expression capabilities for the same prompt phrases (Chen & Ibrahim, 2023). Building on this, the present study established a cross-modal experimental framework to validate prompt–content mapping comparisons.

Stage 4: Output evaluation and feedback. To systematically evaluate the quality of responses from AIGC platforms to prompts for emotion recognition and generation, this study designed a three-dimensional subjective evaluation system covering 'generative richness', 'emotional alignment', and 'detail expressiveness'. The output results were quantitatively analysed using an independent, anonymous expert scoring mechanism (Palinkas *et al.*, 2015). This feedback mechanism not only evaluates platform performance, but also serves as an important basis for future system tuning and user satisfaction models.

Research Progress and Problem Analysis on Emotion Recognition and AIGC Integration

Emotion recognition technology has been widely applied in fields such as intelligent interaction, educational tutoring, and psychological intervention, providing a potential support path for generative artificial intelligence (AIGC) in context understanding and output control. Research indicates that systems can preliminarily infer users' emotional states by identifying facial expressions, voice intonation, or physiological signals, and adjust response content accordingly (Picard, 1997; Zeng *et al.*, 2009).

In the AIGC context, some studies have attempted to embed emotion recognition results into the content generation process to enhance the emotional consistency of language, images, or multimodal content. For example, dialogue systems driven by emotion recognition tags can adjust tone style or response strategies based on the user's emotional state, thereby enhancing the naturalness and immersion of interactions to some extent (Calvo & D'Mello, 2010). However, such 'emotion-response' mechanisms still face challenges in practical applications, including high data collection costs, strong user perception intrusiveness, and weak model transferability.

Especially in AIGC applications deployed across multiple platforms, different models lack unified standards for responding to emotional inputs, and the role pathways of emotion recognition are often opaque, leading to limited system control capabilities and unstable output styles.

Additionally, in scenarios without user participation or offline generation, the feasibility and necessity of real-time emotion recognition are both limited. However, compared to emotion-driven mechanisms dominated by prompts, this approach offers a more general and flexible alternative. Prompts are predefined by the system or selected by users, combined with specific emotion descriptions (e.g., 'calm,' 'anxious') to form input content with semantic instruction functions, which can directly influence text or image generation processes. In actual generation tasks, this mechanism does not rely on real-time user emotion input, demonstrating high cross-modal adaptability and experimental controllability (Reynolds & McDonell, 2021). Although this study has not directly utilised real-time emotion recognition technology, the proposed emotion prompt-driven mechanism is essentially a 'transferable emotion mapping model' that can serve as a bridge for future integration of emotion recognition and generation systems.

In virtual experience environments, this mechanism demonstrates excellent application scalability, particularly in immersive interactive interfaces, where it can quickly adapt to scene atmospheres and generate content through predefined or user-selected prompts. Examples include soothing conversations in virtual therapeutic spaces and tension-building narratives in virtual performances.

AIGC PLATFORM AND GENERATION TOOLS

For AIGC applications based on emotion recognition, it is crucial to collect data on emotions and AIGC. Various intelligence on data collection and organization AI, as well as image-generating AI, was therefore collected, including advanced technologies from different countries and release times. These AI technologies were used to conduct different tests, and the results were analyzed to determine the most suitable technology for emotion recognition and virtual scene realization.

Text and Dialogue Generation AI Tools

This study selected seven popular AI tools: Replika (Luka Inc., 2022); Microsoft Copilot (Microsoft, 2023); Character.AI (Character.AI, 2022); Google Gemini (DeepMind, 2023); Poe (Quora, 2022); ERNIE Bot (Baidu, 2022); and ChatGPT 4o (OpenAI, 2023). Their basic information is shown in Table 1.

Table 1: Mainstream Text AI Tool Survey Form

AI Name	source country	Release time	Official website address	Official source literature/technical reports
Replika	USA	2017	https://replika.com/	(Luka Inc, 2022)
Microsoft Copilot	USA	2021	https://copilot.microsoft.com/	(Microsoft, 2023)

Character AI	USA	2022	https://character.ai/	(OpenAI, 2023)
Google Gemini	USA	2022	https://gemini.google.com/	(DeepMind, 2023)
Poe	USA	2022	https://poe.com/	(Quora, 2024)
ERNIE Bot	China	2022	https://yiyan.baidu.com/	(Baidu, 2022)
ChatGPT 4o	USA	2023	https://chat.openai.com/	(OpenAI, 2023)

Source: Drawn by the author.

Image Generation AI

This study also selected six mainstream image generation AI tools, including Imagine (Imagine AI, 2022), Midjourney (Midjourney Inc., 2022), Stable Diffusion (Stability AI, 2022), Openart (Openart AI, 2022), GPT 4o (OpenAI, 2023), and Blockade Labs (Blockade Labs, 2023). Their basic information is shown in Table 2.

Table 2: Image Generation AI Survey Form

AI Name	source country	Release time	Official website address	Official source literature/technical reports
Imagine	USA	2022.05	https://www.imagine.art/	(Imagine, 2022)
Midjourney	USA	2022.07	https://www.midjourney.com/	(Midjourney, 2022)
Stable Diffusion	Germany	2022.08	https://www.diffus.me/	(Stability AI, 2023)
Openart	USA	2022.11	https://openart.ai/	(Openart, 2022)
GPT 4o	USA	2023.01	https://openai.com/dall-e-3	(OpenAI, 2024)
Blockade Labs	UK	2023.04	https://www.blockadelabs.com/	(Blockade Labs, 2023)

Source: Drawn by the author.

EXPERIMENTAL DESIGN AND PLATFORM VALIDATION

Experimental Objectives and Research Framework

This study uses a dual-path experimental design to validate differences in performance between platforms in terms of their emotional response capabilities across two task categories: text and image generation. Accordingly, two experimental tasks were established:

(1) Emotion-Driven Conversational AI Generation Experiment

In this experiment, emotion recognition types are used as inputs for user emotions. Emotion prompt words are provided to seven mainstream conversational AI platforms (Microsoft Copilot, Replika, Character.AI, Google Gemini, Poe, ChatGPT 4o and ERNIE Bot) to compare their performance in understanding emotional intent, shaping language style and expressing emotions.

(2) Emotion-driven image-generating AI experiment

A fixed prompt combined with two contrasting emotion keywords was input into three major image-generating AI platforms (Midjourney, Stable Diffusion and DALL·E 3) to observe the effect of emotion prompts on the atmosphere, tone and level of detail of the images produced.

Expert Evaluation

Overview of sample information

To ensure a scientific assessment of AIGC-generated content in terms of emotional adaptability and performance quality, this study employed purposive sampling. Ten experts with backgrounds in artificial intelligence, virtual experiences, and user emotion research were invited to serve as evaluators (Palinkas *et al.*, 2015).

The expert panel primarily consisted of young researchers aged between 25 and 33, with an average age of approximately 29. Most were recent PhD graduates, postdoctoral researchers or young faculty members/engineers at the forefront of artificial intelligence and human-computer interaction. They had strong interdisciplinary research capabilities and experience in emerging technology applications. In terms of gender composition, 60% of the panel were male and 40% were female.

Their professional backgrounds spanned artificial intelligence algorithms, natural language processing, virtual reality design, human-computer interaction, emotion psychology and digital art design. This balanced perspectives from system development, content generation and user experience analysis. In terms of institutional composition, the experts were drawn from comprehensive universities (30%), design or engineering-focused institutions (30%), and corporate research departments and applied research institutions (40%), thus balancing theoretical depth with practical experience. Expert composition details are provided in Table 3.

Prior to the scoring process, all experts thoroughly understood the research objectives, evaluation dimensions and scoring procedures. They also signed informed consent forms to ensure the autonomy and ethical compliance of their participation. The scoring process was conducted independently and anonymously, without any mutual discussion or interference from platform preferences, to ensure the objectivity of the results and the scientific validity of the evaluation. Additionally, the experts had no economic or emotional ties to the AIGC platforms or the generated content being evaluated. The scoring activity did not involve collecting personal information or posing emotional risks, thus aligning with the humanistic and ethical requirements of virtual experience-related research. Future studies will incorporate ordinary users, combining multimodal behavioural data and interaction feedback to expand the representativeness of the sample and refine the evaluation framework for the adaptability of users to emotion-driven AIGC generation systems.

Table 3: Expert Composition Information Table (anonimised)

Specialist No.	Area of specialization	Title/position	Type of affiliation	Research and Practice Direction
Expert A	Artificial Intelligence and Affective Computing	Professor / PhD	comprehensive university	AIGC generation mechanism, multimodal sentiment recognition, and large language model tuning
Expert B	Virtual Reality and User Experience Design	Senior Researcher / UX Director	Research Institute / Enterprise	Modeling of virtual interactive environments, XR content development, assessment of user emotional adaptation
Expert C	Visual Communication and Digital Art	Associate Professor / Design Consultant	art school	AI Visual Generative Expression, Mood Atmosphere Construction, Digital Narrative Design

Expert D	Human-computer interaction and interface emotional design	Senior Engineer / Product Manager	technology company	Intelligent Interaction System, Semantic Driven UI Design, Virtual Character Personification
Expert E	Computer Graphics and Image Generation	Associate Researcher / Technical Leader	research organization	Image generation control mechanisms, style migration, generation model tuning
Expert F	Language Modeling and Natural Language Processing	Professor / Director of NLP Laboratory	engineering college	Conversational AI, text sentiment modeling, prompt engineering
Expert G	User psychology and digital behavior analysis	Associate Professor / Doctor of Psychology	comprehensive university	User emotional response assessment, behavioral data modeling, psychological mechanisms of immersive experiences
Expert H	Multimodal Perception and Emotion Recognition System	Postdoctoral / Systems Engineer	research organization	EEG emotion recognition, heart rate/electrical skin feature fusion, and emotion monitoring hardware development
Expert I	Digital Media Technology and AI Authoring Systems	Senior Lecturer / Head of Program Development	Colleges and universities of applied sciences	AI graphic creation platform integration, artistic expression and algorithmic tuning
Expert J	Cognitive Science and Intelligent Educational Technology	Associate Professor / Program Director	teacher-training college	Emotionally aware educational system, immersive learning environment, intelligent feedback interaction

Source: Drawn by the author.

Expert dimension settings and scoring criteria

To ensure that the expert evaluation dimension settings are both theoretically sound and practically valid, this study's evaluation dimension system design drew upon subjective assessment practices from multiple text-image generation tasks. This framework is applicable not only to subjective evaluations of text and image generation quality, but also provides an analytical framework for regulating emotional consistency and immersion in virtual experience scenarios. This enhances the system's emotional adaptability and the quality of content in immersive interactions.

'Emotional consistency' is derived from the 'emotional consistency' dimension in the emotion-controlled image generation model EmoDiff by Guo *et al.* (2023); 'detail representation' originates from the 'detail realism' and 'naturalness' metrics in image and text generation quality evaluation (Zhang *et al.*, 2018). 'Generation richness' is inspired by Greene and Botros *et al.* (2015), who researched the complexity of visual cognitive structures, emphasising the information hierarchy and structural diversity of generated content (Greene *et al.*, 2015). A three-dimensional expert evaluation system centred on 'Generative Richness (Scene Richness)', 'Emotional Alignment', and 'Descriptive Fineness' has therefore been constructed.

(1) Scene Richness:

This measures the diversity, spatial integrity, and visual layering of elements in the generated image.

For example, it considers whether architectural elements, characters, technological elements, and environmental details are rich and complete.

(2) Emotional Alignment:

This measures the image's ability to convey the target emotion (e.g. anxiety or calmness) through atmosphere, tone, lighting and other aspects.

For example, it considers whether the image creates a tense and stimulating atmosphere or a soothing and peaceful one.

(3) Descriptive Fineness:

This measures the level of detail, visual coherence and artistic expression in the image.

For example, it considers whether the details are clear and the elements are harmonious and natural.

For detailed definitions of the dimensions, application guidelines and scoring mechanisms, please refer to Table 4.

Table 4: Evaluation dimensions of emotion-driven generation

evaluation dimension	Definitions in brief	Theoretical sources	source (of a document)
Scene Richness	Measuring the diversity, hierarchy and spatial integrity of elements in generated content	Structural complexity enhances perceptual visualisation efficiency High structural stimuli elicit stronger encoded responses	(Greene <i>et al.</i> , 2015)
Emotional Alignment	Determine whether the generated content accurately expresses the emotional goal of the cue	"Emotional Consistency" in Emotional Consistency Generation Mechanisms is widely used in subjective assessments.	(Guo <i>et al.</i> , 2023; Li <i>et al.</i> , 2023)
Descriptive Fineness	Evaluate the level of sophistication and natural coherence of the generated content at the micro level	Detail completeness is one of the key dimensions in users' subjective judgement of the quality of image and speech production	(Zhang <i>et al.</i> , 2018; Li <i>et al.</i> , 2022)

Source: Drawn by the author.

To ensure consistency and comparability in expert ratings, this study uses a five-point Likert-type scale, widely adopted for subjective evaluations in natural language generation (NLG), image generation (IG) and emotional interaction systems. The scoring criteria are based on the 'Mean Opinion Score' system proposed by the International Telecommunication Union (ITU-T) and the evaluation practices of Guo *et al.* (2023), Huang *et al.* (2022) and Zhang *et al.* (2023) in the field of AIGC emotion generation. This clearly defines the performance range and subjective perception characteristics corresponding to each level, enhancing the horizontal comparability of multi-platform and multi-modal generation results. The scoring criteria are presented in Table 5.

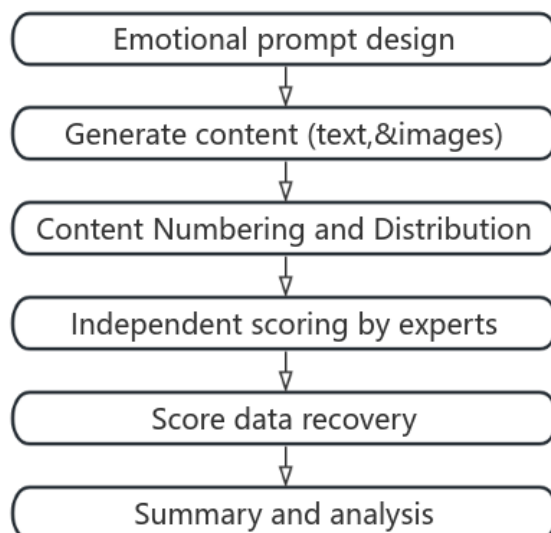
Table 5: Scoring criteria (1-5 points)

value of a score	Description of evaluation criteria	Literature base
5 points	The performance is extremely good and fully meets the requirements of the evaluation dimensions, with rich details, accurate expression of emotions, and excellent overall effect of the picture.	(ITU-T, 2003; Guo <i>et al.</i> , 2023)
4 points	Good performance, mostly meeting the evaluation dimensions, with good detail and emotion, with slight room for improvement	(Gatt & Kraemer, 2018; Zhang <i>et al.</i> , 2024)
3 points	Average performance, partially meets the requirements of the evaluation, with significant deficiencies in detail or emotional expression	(Zhang <i>et al.</i> , 2023)
2 points	Weak performances, most of which do not meet the requirements of the evaluation, and images that lack emotional expression or are poorly detailed	(Guo <i>et al.</i> , 2023)
1 point	Extremely poor performance, completely unable to meet the requirements of the evaluation dimension, and the mood does not match the picture seriously	(Li <i>et al.</i> , 2023)

Source: Drawn by the author.

Expert scoring process

To ensure the independence of the scoring process and the objectivity of the results, this study used an anonymous, decentralised expert scoring system. All scoring tasks were distributed separately via a unified platform. Experts completed evaluations of the generated content item by item, independently and without knowing the identities of other participants or the scoring results. No discussion or consultation sessions were held. The scoring materials were de-identified, retaining only task-related content numbers, in order to avoid cognitive bias (as shown in Figure 1).



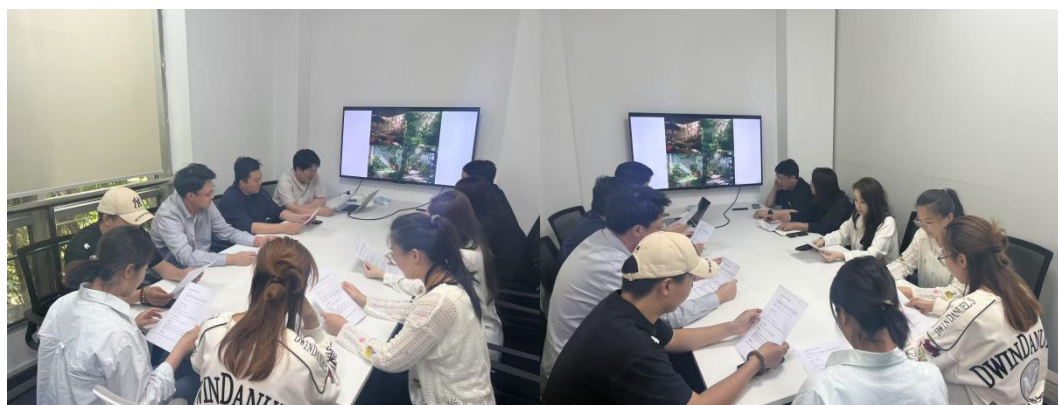
Source: Drawn by the author.

Figure 1: Expert Scoring Flowchart

The complete process is as follows:

- (1) Designing the emotion prompt: Researchers design emotion prompt texts based on the experimental objectives, such as 'anxiety' or 'calm' scenarios.
- (2) Content generation (text/images): The prompts are input into various AIGC platforms to obtain the corresponding text or image generation results.
- (3) Content numbering and distribution: Assign a unique ID to each piece of generated content and remove any platform information or prompts to prevent expert bias. Then, distribute the content to all experts.
- (4) Independent expert scoring: Each expert reads and views the content independently and scores it according to the scoring criteria, without any group discussion or communication.
- (5) Score data collection: Collect all expert scoring forms and organise them into structured data tables.
- (6) Summary and analysis: Conduct analyses such as calculating the average score and variance, and use the results for comparison and model optimisation.

The expert review site is shown in Figure 2.



Source: Photographed on site by the author

Figure 2: Expert review site

Comparative Analysis of Text and Dialogue Generation AI Tools

This experiment aims to verify the regulatory ability of emotional prompt words on the performance of AIGC content generation. Focusing on 'happy' and 'quiet and comfortable' as examples of high-value, low-arousal composite emotions (Russell, 1980), a unified prompt information input was constructed for seven mainstream conversational AI platforms to observe differences in prompt understanding and emotion-adaptive generation. This also facilitates cross-platform comparative analysis under semantically consistent input conditions. While this study focuses on positive emotions, it recognises the broad influence of different types of emotional prompts on style control. Future research could introduce negative emotions such as

'anxiety' and 'sadness' to expand the scope of emotional dimensions, enhance the breadth of comparisons within the experimental system and improve the comprehensiveness of adaptability assessments.






(1) Experimental task setting:


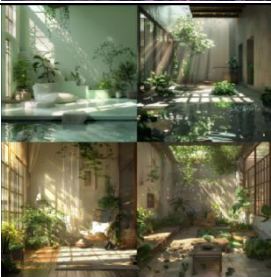
User emotion recognition requirements: AI is required to generate a 'pleasant, quiet and comfortable virtual environment'.

Prompting strategy: Based on the emotional needs of the user, provide each platform with prompts that are tailored to its language understanding mechanisms, ensuring that different AI systems respond comparably to the same emotional goals.

(2) Experimental platform and prompt configuration (See Table 6)

Table 6: Experimental platform and prompt configuration

AI Name	Cue word content	Picture effect
Microsoft Copilot	An environment without loud sounds, warm temperatures and soft light, some green elements, natural landscapes, plants and birdsong	
Replika	Relaxing, serene, calm, peaceful, tranquil	
Character AI	Happy, quiet, cozy, peaceful, safe, relaxed	
Google Gemini	Starry Sky Tent: Set up a tent under the starry sky far away from the hustle and bustle of the city, lie on a comfortable bed, feel tranquility	
Poe	Tranquility, serene environment, quiet atmosphere, comfort, soft materials, comfortable lighting, comfortable temperature	

ChatGPT 4o	Tranquil lakeside setting, warm sunset hues, orange and pink sky gradient, gentle water reflections, winding cobblestone path	
ERNIE Bot	An environment free of noisy sounds, using soft colors, soothing layouts, and natural elements like green plants, running water, birdsong	

Sources: Microsoft Copilot, Replika, Character AI, Google Gemini, Poe, ChatGPT-4o, ERNIE Bot, AI Testing Platform.

In terms of content generation, there are significant differences between platforms regarding the completeness of language construction, the subtlety of atmosphere creation and the depth of emotional expression. Google Gemini and ChatGPT 4o demonstrate high semantic generation capabilities in their responses to prompts, producing clearly evocative descriptive texts based on prompts such as 'sunset by the lake' and 'a cottage in the woods under the stars'. The generated content demonstrates emotional relevance and attention to detail, and exhibits excellent immersion and visualisation capabilities in terms of narrative pacing and scene composition. This makes it highly adaptable for virtual experiences. When applied directly to voice guides, emotional guidance or scene scripting in immersive systems, such content can effectively support users' emotional immersion and environmental perception, demonstrating high transferability.

In contrast, the text output by Replika and Character AI is more formulaic. It often uses simple emotional words such as 'calm' and 'peaceful' without providing specific contextual support or spatial imagination, which makes it difficult to construct virtual scenes that feel realistic and participatory. The content largely remains at the level of abstract emotional labels and therefore fails to support users' perception and immersion in emotional atmospheres during virtual experiences effectively. While Microsoft Copilot, Poe and ERNIE Bot demonstrate some scene-building capabilities, their text semantics are relatively loose, limiting their applicability in emotion-driven virtual spaces (see Table 7).

Table 7: Comparison of content generation features across platforms

Platform name	linguistic integrity	Ability to create an atmosphere	Richness of emotional expression	Typical content performance
Google Gemini	Specific and coherent	Outstanding quiet and cozy atmosphere	Subtle, figurative	Full scene descriptions of forest cabins, starry sky tents and more
ChatGPT 4o	Specific and coherent	Good immersion	Figurative and graphic	Detailed descriptions of lakeside sunsets, winding paths and more
Microsoft Copilot	basically complete	Medium atmosphere	Average emotional performance	Brief description, lack of detail
Poe	basically complete	Medium atmosphere	Average emotional performance	Generalized language, not enough detail
ERNIE Bot	The language is natural but generalized.	Limited atmosphere creation	Average emotional performance	Generic description, average emotional fit
Replika	Fragmentation, vocabulary stacking	Insufficient atmosphere creation	stereotyped	Simple words like relaxing, calm, peaceful, etc. are piled on top of each other.
Character AI	Fragmentation, vocabulary stacking	Insufficient atmosphere creation	stereotyped	Simple words like relaxing, calm, peaceful, etc. are piled on top of each other.

Source: Drawn by the author.

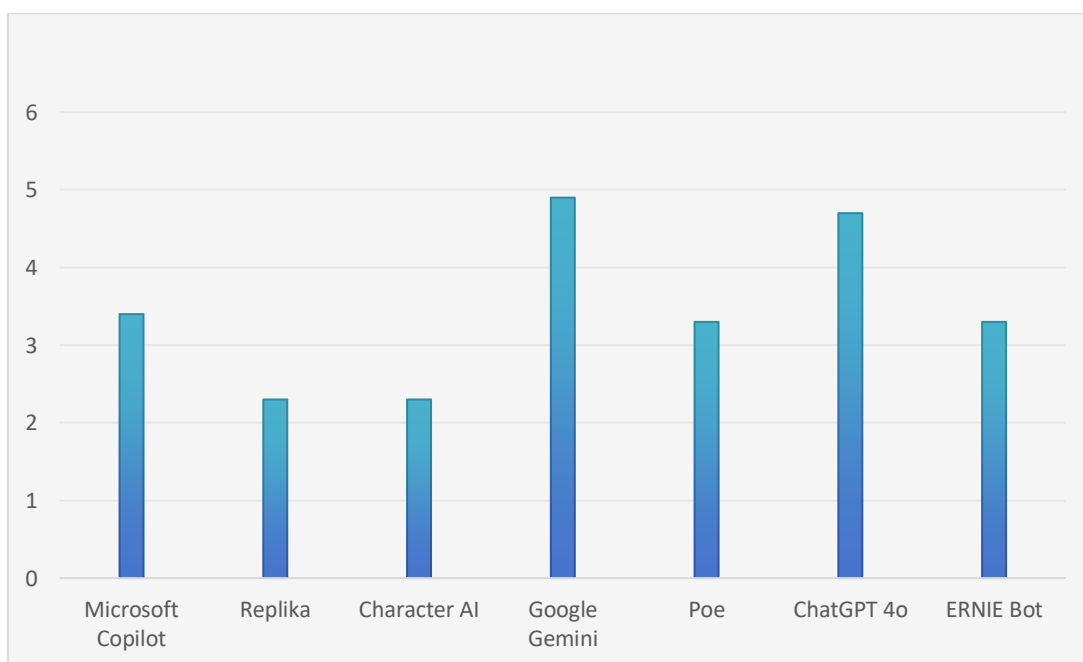
Experimental Scoring Results

Table 8 and Figure 3 shows the performance of the seven conversational AI platforms in terms of language completeness, atmosphere creation and emotional expression richness under the prompt ‘pleasant and quiet virtual environment’.

Table 8: Evaluation results of emotional adaptability generation by conversational AI tools

Platform name	Generation richness (1-5)	Emotional fit (1-5)	Description of finesse (1-5)	Average score (1-5)
Microsoft Copilot	2.9	4.1	3.1	3.4
Replika	1.9	2.6	2.3	2.3
Character AI	2.2	2.5	2.3	2.3
Google Gemini	4.9	4.9	4.9	4.9
Poe	2.9	4	3	3.3
ChatGPT 4o	4.9	4.8	4.5	4.7
ERNIE Bot	3.6	3.4	2.9	3.3

Source: Compiled by the author based on experimental evaluation.



Source: Screenshot of experimental results from the platform.

Figure 3: Generate comparison chart

The scoring structure indicates that Google Gemini and ChatGPT 4o received the highest ratings in all three dimensions: richness, emotional accuracy and descriptive detail. This demonstrates their excellent emotional understanding and language expression capabilities, enabling them to generate concrete, emotionally charged virtual scene descriptions, which indicates that both platforms have strong practicality and application potential when addressing users' emotional needs. Microsoft Copilot, Poe and ERNIE Bot performed moderately; while they could generate content with a certain emotional atmosphere based on prompts, their language expression was relatively general and lacked detail and visual imagery. There is room for improvement in their overall performance. Replika and Character AI scored the lowest; their generated content primarily consisted of single emotional words or simple sentence structures and lacked complete contextual descriptions and concrete expressions. These two platforms are therefore not suitable as primary tools for emotional interaction or virtual experience scenarios, as they are weak in their responsiveness to emotional prompts in the current task.

Comparison and Analysis of Image Generation AI Effects

Test settings

To further verify the regulatory role of emotional keywords on image generation style and atmosphere, this study used a single prompt word and introduced two types of emotional keyword — 'anxious' and 'calm' — to analyse the performance of different image generation AI platforms in response. The fixed prompt word was a futuristic city centre at night with a clearly visible starry sky, throughout which high-tech facilities were ubiquitous, including fashionable glass buildings, floating billboards and flying cars. People wear smart clothing and interact with various robots and forms of artificial intelligence on the streets. In the centre of the city, there is a huge LED screen displaying the latest advances in artificial intelligence technology. This scene

integrates high technology with everyday life, presenting a vision of the future in which artificial intelligence is deeply integrated into human life.

(2) Emotional keywords: anxiety, calm

Combine the above prompt with the emotional keywords and input them into three AI platforms (Midjourney, Stable Diffusion and DALL·E 3) to generate comparison samples. Summarise their atmospheric characteristics and stylistic expressions, as shown in Table 9.

Table 9: Comparison of mood performance based on emotion keywords across different AI platforms

AI platform	Emotional Keywords	Picture effect
Midjourney	Anxious	
	Calm.	
Stable Diffusion	Anxious	
	Calm.	



Source: Midjourney / Stable Diffusion / GPT-4o AI testing platforms.

Midjourney offers superior style control and visual expression. It can produce distinct stylistic differences and high-quality visual effects in response to various emotional prompts. It is well-suited to virtual experience design tasks that demand rigorous attention to visual detail and emotional atmosphere. In contrast, Stable Diffusion and GPT 4o are better suited to basic visual generation applications, although there is room for improvement in terms of emotional adaptability and the depth of visual expression. When given an anxious prompt, Midjourney can generate visually rich content with abundant details, strong lighting and high tension, successfully creating a tense, stimulating and information-dense futuristic urban atmosphere. In contrast, while Stable Diffusion can convey a mechanical aesthetic and high-contrast effects, it lacks detail and spatial depth, resulting in limited visual impact. GPT 4o is capable of dense information representation, but its visuals are flat and lack depth and subtlety, resulting in an overall conservative style.

When given calm emotion prompts, Midjourney also demonstrates strong visual expressiveness. It generates warm-toned, softly lit scenes with open spaces that evoke a soothing atmosphere, creating a quiet, relaxed futuristic urban setting. Stable Diffusion has improved in terms of soft colour tones and simple composition, but still falls short in terms of detail rendering and spatial depth. GPT 4o exhibits a certain degree of soft atmosphere but has a relatively monotonous composition lacking depth and richness with limited emotional alignment. Specific features are detailed in Table 10.

Table 10: Comparison of content features generated by image-generating AI platforms

Platform name	Emotional cues	Picture integrity	Ability to create an atmosphere	Visual presentation features
Midjourney	apprehensive	full and rich	Tense, exciting, information-dense	Cool colors, high contrast, strong light and shadow, rich in detail, outstanding sense of space
Stable Diffusion	apprehensive	basically complete	Compact and mechanical	Metallic gray, blue tones, high contrast, medium detail, average spatiality
GPT 4o	apprehensive	basically complete	Information-intensive but weakly layered	Cold colors, neon effect, the picture is more flat, lack of detail and space performance
Midjourney	undisturbed	full and rich	Soothing, quiet, open space	Warm colors, soft light, low contrast, rich detail, and spatial layering
Stable Diffusion	undisturbed	basically complete	Nature, simplicity, fusion of technology and nature	Soft blue-green tones, low saturation, clean layout, medium detail, average sense of space
GPT 4o	undisturbed	basically complete	Soft and smooth, but limited performance	A combination of warm and neutral tones, the image is smooth but lacks depth and richness of detail

Source: Drawn by the author.

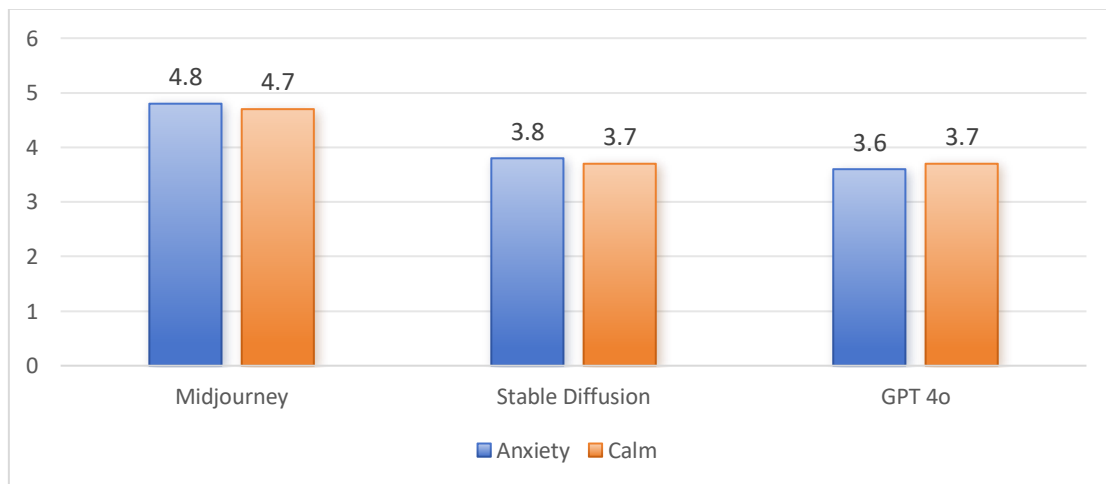
Analysis and summary of experimental scoring results

The expert scoring results for generative richness, emotional accuracy and detail expressiveness under the ‘anxious’ and ‘calm’ emotional prompt conditions for the three image generation AI platforms (see Table 11) were averaged and compared horizontally to produce Figure 3.

Table 11: Emotional Adaptability Scores for Image-Generating AI Platforms

Platform name	Emotional cues	Generation richness (1-5)	Emotional fit (1-5)	Expressive detail (1-5)	Average score (1-5)
Midjourney	apprehensive	4.9	4.8	4.9	4.8
Stable Diffusion	apprehensive	4	4	3.5	3.8
GPT 4o	apprehensive	3.5	4.2	3.2	3.6
Midjourney	undisturbed	4.9	4.9	4.3	4.7
Stable Diffusion	undisturbed	4	3.7	3.5	3.7
GPT 4o	undisturbed	3.5	4.6	3	3.7

Source: Compiled by the author based on experimental evaluation.



Source: Screenshot of experimental results from the platform.

Figure 4. Image generation AI platform score comparison

The evaluation results indicate that Midjourney performed exceptionally well in response to prompts representing both anxiety and calmness. In terms of generative richness, emotional alignment and expression of detail, Midjourney achieved scores close to or at the maximum. This effectively distinguishes emotional characteristics and produces high-quality visual works that are consistent in their emotional expression. This demonstrates outstanding emotional adaptability and control of style.

Stable Diffusion and GPT 4o performed relatively moderately. Both platforms demonstrated acceptable emotional alignment and were capable of generating images with distinct stylistic differences based on prompts. However, they lagged behind Midjourney in terms of generative richness and detail expression. Notably, while GPT 4o exhibited decent emotional recognition capabilities, it suffered from flat visual details and spatial representation, lacking visual impact and immersion.

When comparing anxiety and calmness prompts, the three platforms exhibited consistent performance trends. Under anxiety prompts, they tend to generate images with cool colour tones and tense atmospheres. Under calmness prompts, they produce images with warm colour tones and soothing rhythms. This indicates that emotional prompts strongly influence the style of images generated for these tasks.

DISCUSSION

Based on expert ratings of conversational AI platforms, Google Gemini and ChatGPT 4o performed the best overall with average scores above 4.8, significantly outperforming other platforms. Both platforms excel in the logical coherence and richness of the language they generate and effectively transform 'pleasant-calm' emotional cues into vivid, immersive scenes such as 'a forest cabin', 'a sunset by the lake' and 'a gentle breeze at dusk'. This type of concrete, narrative-rich language output is highly suitable for creating virtual experiences with emotional guidance functions, particularly in scenarios requiring emotional stability support, such as virtual healing, meditation training or educational tours (Brown *et al.*, 2020).

However, Replika and Character AI scored an average of only 2.3 points; their content consists primarily of generic adjectives such as 'relaxing', 'calm', and 'peaceful', lacking contextual construction and detailed descriptions. Consequently, their emotional expression lacks depth and layers, failing to meet the dual demands of narrative completeness and psychological resonance in virtual experience environments. Microsoft Copilot, Poe and ERNIE Bot scored around 3.4 and 3.3 respectively. While they could generally respond to emotional prompts, the generated language was abstract and neutral in style. This lacks the expressive tension needed to actively guide users into an immersive scenario, thereby limiting performance in complex virtual experience tasks.

Expert evaluations of image generation tasks show that Midjourney achieved an average score of 4.8 for both 'anxiety–calm' emotional prompts, demonstrating exceptional style control and expression of visual detail. Under the 'anxiety' prompt, it generates futuristic urban landscapes with cool colour tones, high contrast and compressed spatial composition. Under the 'calm' prompt, it creates natural environments with soft lighting, open spaces and warm tones. The visual elements are rich and highly aligned with the target emotional settings. These visuals have strong sensory guidance capabilities and are ideal for use in immersive virtual experiences to convey specific psychological states or scene atmospheres, such as anxiety-inducing training, emergency response simulations, or the creation of soothing scenes.

In comparison, Stable Diffusion and DALL·E 3 scored between 3.6 and 3.8. While Stable Diffusion can respond to emotional prompts in terms of tone and structural layout, it lacks visual depth and detail. DALL·E 3 possesses some structural logic capabilities but suffers from severe image flattening issues and lacks the spatial depth and emotional penetration required for immersive scenes. These limitations in visual expression restrict their potential applications in high-interaction, high-emotional dependency virtual experience systems.

A comparison of task types shows that emotional prompts demonstrate good style guidance effects in both text and image dual tasks. However, the implementation paths show significant differences: text generation tasks rely more on language understanding and semantic construction capabilities, making them suitable for narrative guidance, character dialogue and emotional prompt descriptions in virtual experiences. In contrast, image generation tasks focus on visual style and scene rendering control, making them suitable for constructing immersive scene atmospheres and emotional spaces. These two types of generation task can complement each other in virtual experience systems, respectively fulfilling the functions of 'language context guidance' and 'visual emotion construction'.

Additionally, this study reveals significant differences among platforms in terms of emotional adaptability and content accuracy. This indicates that AIGC tools should not be selected as a one-size-fits-all solution for virtual experience systems. Instead, strategic combinations should be made based on specific task requirements and experience objectives. For instance, in a virtual museum, ChatGPT could provide contextual guidance and story narration, while Midjourney could generate atmospheric images and spatial renderings. This would create an emotion-driven experience through language-image synergy. Emotional prompts demonstrate significant adaptability in both text and image tasks as a cross-modal content generation control mechanism. However, uneven implementation capabilities across platforms remind us that, when constructing virtual experience systems, we must pay attention to differences in platform functionality and match task characteristics accordingly to achieve stable, personalised emotional immersion output (Watson *et al.*, 2023).

RESEARCH SIGNIFICANCE

Theoretical Contributions

This study focuses on the cross-disciplinary topic of 'emotion-prompt-driven generative artificial intelligence (AIGC) content construction' and proposes an empirically validated emotion-guided AIGC mechanism model consisting of 'prompt setting – content generation – performance evaluation'.

This model systematically integrates four stages: pre-intervention with emotion signals; structured prompt construction; multi-platform content generation control; and subjective effect evaluation. The model aims to address the current lack of theoretical integration and process modelling in the fusion of emotion computing and generation mechanisms. From a theoretical perspective, this study is the first to explicitly incorporate user emotions into the pre-generation stage, emphasising their active role in establishing generation intentions and regulating style.

This differs from traditional AIGC research, which treats emotion only as a variable for evaluating the acceptability of the generation result. Instead, the study constructs a 'emotion–prompt–generation–feedback' closed-loop mechanism that elucidates how emotional cues are converted into control signals recognisable by the generation platform through prompts. This influences the structure, style and expression logic of the content production process.

Additionally, based on cross-modal experimental practices, the study has established an expert evaluation system centred on core metrics such as 'generation richness', 'emotional alignment', and 'detail expressiveness'. This system considers the multidimensional expression characteristics of text and image tasks, providing a reusable subjective evaluation framework that supports the evaluation of performance, the comparison of platforms, and the standardisation of AIGC systems in emotion-driven application scenarios. The theoretical contribution of this study expands the research path for integrating

affective computing and generative AI, and lays a systematic foundation for modelling, evaluating and optimising emotion-regulated generation mechanisms.

Practical Value

At the application level, this study uses text and image AIGC platforms as experimental objects to systematically analyse the differences in generation performance caused by different emotion prompts on each platform. The established process model and evaluation system are highly practical and their specific application value is reflected in three key areas: firstly, when designing emotion-adaptive AI interaction systems, the study's summarisation of emotion prompt structures and construction of an evaluation mechanism can inform the development of AI text generation and voice interaction systems for specific emotion-driven scenarios (e.g. child companionship, elderly care, and emotional guidance). This is particularly relevant when designing functional modules such as 'tone control', 'style matching' and 'empathy response' (Calvo & D'Mello, 2010).

Second, in digital creation and visual content generation scenarios, the study revealed differences in the response mechanisms of AIGC platforms to 'anxiety-calm' emotion prompts, which can support creative platforms in refining the control of generated image style and expression accuracy. Secondly, in scenarios involving digital creation and visual content generation, the study revealed differences in the response mechanisms of AIGC platforms to 'anxiety-calm' emotion prompts. This information can support creative platforms in refining the control of generated image style and expression accuracy.

This mechanism could be embedded in emotion-sensitive content generation domains, such as brand customisation, digital art and narrative-based text-image creation. This would enhance compatibility between content and user states. Thirdly, in virtual spaces and immersive interaction systems, maintaining emotional consistency between generated content and the current state or task scenario of users has become essential for enhancing the realism and contextual coherence of the experience. The emotion-driven model and three-dimensional evaluation system constructed in this paper can serve as adjustment parameters and adaptation standards for system content generation modules, supporting various application scenarios, such as virtual tours, situational games, and metaverse interactions.

Additionally, the proposed cross-platform testing method and scoring mechanism can be employed for version comparison, algorithm evaluation, and the structured collection of user experience feedback during the development stage of AIGC systems. This demonstrates good replicability and platform universality.

This study provides a systematic modelling approach for AIGC generation mechanisms in theory and demonstrates their broad practical application, offering feasible solutions and technical frameworks for integrating emotion perception and intelligent content generation in areas such as educational technology, the content industry, digital health, and cultural creativity.

CONCLUSION

The study focused on the emotion-driven generation mechanism of AIGC and designed and implemented a dual-task comparison experiment covering conversational and image generation AIs. It also systematically verified the adaptability and regulatory effects of emotion prompts on different platforms and generation tasks. Through a combination of content feature analysis of text and image generation results and expert evaluation, the study revealed the expressive value and platform differences of emotion prompts in multimodal content generation. This provides empirical support for emotion control strategies in cross-modal AIGC systems.

The results indicate that emotional prompts can effectively regulate both the language style of text and the visual atmosphere of images. They also demonstrate cross-modal applicability and style consistency. Google Gemini and ChatGPT 4o performed excellently in text generation tasks, producing concrete and coherent language with an emotional tone. Midjourney demonstrated exceptional control over style and the ability to express detail in image generation tasks. It accurately responded to 'anxiety' and 'calm' prompts, generating high-quality visual works that vividly convey emotion.

Additionally, the experiments revealed significant differences in emotional adaptability across platforms. Replika, Character AI, Stable Diffusion and DALL·E 3 all exhibited weaknesses in terms of detail richness, emotional consistency and the creation of immersion. This highlights the need for task-specific selection and optimisation based on platform characteristics in practical applications (George *et al.*, 2023).

The main contributions of this study are twofold: First, at the theoretical level, we propose a chain-driven model of 'emotion recognition–prompt mapping–generation control–subjective feedback', which systematically characterises the mediating role of emotional factors in the AIGC generation process. This provides a clear, logically coherent theoretical framework for integrating Affective Computing and Generative AI. Second, at the practical level, we construct a replicable, cross-platform testing and evaluation mechanism that can be applied to testing emotion adaptability in the development stage of AIGC systems. In scenarios such as digital exhibitions, immersive education, psychological counselling and emotion therapy, AIGC combined with emotion prompt mechanisms can provide solutions for atmosphere control and adaptation to the user's state in virtual spaces. This drives the evolution of virtual experience systems from static content presentation to dynamic emotional empathy.

RESEARCH LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

This study constructed and validated a theoretical model of emotion cue-driven AIGC generation mechanisms through a dual-task experiment. However, the research design and application have several limitations that affect the broad applicability and in-depth promotion of the results.

The experiment focused solely on the 'anxiety–calm' pair of opposing emotions, resulting in limited emotional type coverage. The adaptability and transferability of the model across a broader emotional spectrum has not been validated. The prompt design relied on static, predefined prompts and lacked real-time interaction and dynamic adjustment mechanisms. This hindered the simulation of emotional evolution in real-world contexts and impacted the ecological effectiveness of the generation mechanism.

Furthermore, the integration of AIGC into virtual experience scenarios is underdeveloped, with the experimental settings mainly focusing on offline text and image generation. This means that AIGC output content has not been embedded into immersive and interactive virtual platforms, making it difficult to evaluate its suitability and user response in actual scenarios such as VR education, virtual exhibitions, and emotional healing. Concurrently, evaluation remains primarily based on expert scoring, lacking user behaviour data and emotional feedback, and failing to reflect the impact of generated content on virtual experience quality.

Research focuses on content-level differences without delving into the underlying logic of generation mechanisms or pathways for prompt understanding across platforms, thereby limiting the exploration of explainability analysis and prompt engineering optimisation for AIGC systems.

Future research could expand into the following areas: Expanding the range of emotions covered and constructing a multimodal generation system that supports real-time control; Testing the impact of generated content on the user experience in immersive environments to improve contextual adaptability; Introducing user behaviour and feedback data to enhance the evaluation system; Strengthening analysis of internal platform mechanisms to improve the controllability and empathy of AIGC models in virtual experiences.

REFERENCES

- Baidu. (2020). ERNIE bot technical white paper. <https://yiyan.baidu.com/>
- Blockade Labs. (2023). Skybox AI overview. <https://www.blockadelabs.com/>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901. https://papers.nips.cc/paper_files/paper/2020/hash/1457c0d6bfc4967418bfb8ac142f64a-Abstract.html
- Calvo, R.A., & D' Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18 - 37. <https://doi.org/10.1109/T-AFFC.2010.1>
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P.S., & Sun, L. (2023). A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT [Preprint]. arXiv. <https://arxiv.org/abs/2303.04226>
- Character.AI. (2024). Terms of service. <https://beta.character.ai/legal/tos>
- Chen, X., & Ibrahim, Z. (2023). A comprehensive study of emotional responses in AI-enhanced interactive installation art. *Sustainability*, 15(22), 15830. <https://doi.org/10.3390/su152215830>
- DeepMind. (2023). Gemini: A family of highly capable multimodal models [Technical report]. <https://gemini.google.com/>
- Gatt, A., & Krahmer, E. (2018). Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, 61, 65-170. <https://doi.org/10.1613/jair.5477>
- George, A.S., George, A.H., Baskar, T., & Pandey, D. (2023). The allure of artificial intimacy: Examining the appeal and ethics of using generative AI for simulated relationships. *Partners Universal International Innovation Journal*, 1(6), 132-147.
- Greene, M.R., Botros, A.P., Beck, D.M., & Fei-Fei, L. (2015). What you see is what you expect: Rapid scene understanding benefits from prior experience. *Attention, Perception, & Psychophysics*, 77(4), 1239-1251. <https://doi.org/10.3758/s13414-015-0855-0>
- Gu, J., Han, Z., Chen, S., Beirami, A., He, B., Zhang, G., & Torr, P. (2023). A systematic survey of prompt engineering on vision - language foundation models [Preprint]. arXiv. <https://arxiv.org/abs/2307.12980>
- Guo, J., Ma, Z., Li, W., Zhang, R., Zhu, Y., & Qian, C. (2023). EmoDiff: Emotion-guided diffusion for image generation [Preprint]. arXiv. <https://arxiv.org/abs/2307.11204>
- Guo, Y., Du, C., Chen, X., & Yu, K. (2023). EmoDiff: Intensity controllable emotional text-to-speech with soft-label guidance [Preprint]. arXiv. <https://arxiv.org/abs/2211.09496>
- Hao, Y., Chi, Z., Dong, L., & Wei, F. (2023). Optimizing prompts for text-to-image generation. *Advances in Neural Information Processing Systems*, 36, 66923-66939. <https://doi.org/10.48550/arXiv.2305.18264>
- He, Y., Dang, S., Ling, L., Qian, Z., Zhao, N., & Cao, N. (2025). EmotiCrafter: Text-to-emotional-image generation based on valence-arousal model [Preprint]. arXiv. <https://arxiv.org/abs/2501.05710>
- Imagine. (2022). Imagine AI official website. <https://www.imagine.art/>

- International Telecommunication Union. (2003). Recommendation P.800: Methods for subjective determination of transmission quality. <https://www.itu.int/rec/T-REC-P.800-200308-I>
- Kreps, S., McCain, R., & Brundage, M. (2023). All the news that's fit to fabricate: AI-generated text as a tool of disinformation. *Science Advances*, 9(4), eabq6494. <https://doi.org/10.1126/sciadv.abq6494>
- Kwon, D.H. (2024). Analysis of prompt elements and use cases in image-generating AI: Focusing on Midjourney, Stable Diffusion, Firefly, DALL • E. *Journal of Digital Contents Society*, 25(2), 341-354.
- Leong, W.Y. (2023). Dynamic storytelling and personalised film narratives using AIGC (Unpublished master's thesis). *INTI International University, Malaysia*.
- Li, C., Wang, J., Zhang, Y., Zhu, K., Hou, W., Lian, J., & Xie, X. (2023). Large language models understand and can be enhanced by emotional stimuli [Preprint]. arXiv. <https://arxiv.org/abs/2307.11760>
- Li, X., Liu, T., & Chen, Y. (2022). Text generation evaluation beyond BLEU: A survey. *Computational Linguistics*, 48(3), 567-598. https://doi.org/10.1162/coli_a_00436
- Liberatore, M.J., & Wagner, W.P. (2021). Virtual, mixed, and augmented reality: A systematic review for immersive systems research. *Virtual Reality*, 25(3), 773 - 799. <https://doi.org/10.1007/s10055-020-00453-2>
- Luka Inc. (2022). Replika white paper. <https://replika.com/>
- Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., & Malik, H. (2023). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence*, 53(4), 3974-4026. <https://doi.org/10.1007/s10489-022-04047-1>
- Microsoft. (2023). Microsoft Copilot technical overview. <https://copilot.microsoft.com/>
- Midjourney. (2022). Midjourney official website. <https://www.midjourney.com/>
- OpenAI. (2023). GPT-4 technical report (No. arXiv:2303.08774) [Technical report]. <https://openai.com/research/gpt-4>
- OpenAI. (2024). Introducing GPT-4o: Multimodal flagship model. <https://openai.com/index/gpt-4o>
- Openart. (2022). Openart AI official website. <https://openart.ai/>
- Pal, S., Mukhopadhyay, S., & Suryadevara, N. (2021). Development and progress in sensors and technologies for human emotion recognition. *Sensors*, 21(16), 5554. <https://doi.org/10.3390/s21165554>
- Palinkas, L.A., Horwitz, S.M., Green, C.A., Wisdom, J.P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533-544. <https://doi.org/10.1007/s10488-013-0528-y>
- Pavithra, A., Kowsalya, J., Keerthi Priya, S., Jayasree, G., & Nandhini, T. K. (2020). An emerging immersive technology: A survey. *International Journal of Innovative Research in Technology*, 6(8), 119-130.
- Picard, R.W. (1997). Affective computing. *MIT Press*.
- Puerta-Beldarrain, M., Gómez-Carmona, O., Sánchez-Corcuera, R., Casado-Mansilla, D., López-de-Ipiña, D., & Chen, L. (2025). A multifaceted vision of the human-AI collaboration: A comprehensive review. *IEEE Access*, 13, 29375 - 29405. <https://doi.org/10.1109/ACCESS.2025.3536095>
- Quora. (2024). Poe: Ask questions, get instant answers. <https://poe.com/>
- Rane, N., Choudhary, S., & Rane, J. (2024). Gemini versus ChatGPT: Applications, performance, architecture, capabilities, and implementation. *Journal of Applied Artificial Intelligence*, 5(1), 69-93.
- Reynolds, L., & McDonnell, K. (2021). Prompt programming for large language models: Beyond the few-shot paradigm [Preprint]. arXiv. <https://arxiv.org/abs/2102.07350>
- Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178. <https://doi.org/10.1037/h0077714>
- Song, X. (2024). Leveraging AIGC and human-computer interaction design to enhance efficiency and quality in e-commerce content generation. Unpublished manuscript.
- Stability AI. (2023). Stable Diffusion 3: Research paper. Stability AI.
- Summaira, J., Li, X., Shoib, A. M., Li, S., & Abdul, J. (2021). Recent advances and trends in multimodal deep learning: A review [Preprint]. arXiv. <https://arxiv.org/abs/2105.11087>
- Wan, J., Zhou, L., & Wang, Y. (2024). The concept and connotation of "AIGC+": A retrospect and prospect of AIGC. In *Proceedings of the Wuhan International Conference on E-Business* (pp. 118 - 129). Cham, Switzerland: Springer Nature.
- Watson, E., Viana, T., & Zhang, S. (2023). Augmented behavioral annotation tools, with application to multimodal datasets and models: A systematic review. *AI*, 4(1), 128 - 171. <https://doi.org/10.3390/ai4010008>
- Wu, J., Gan, W., Chen, Z., Wan, S., & Lin, H. (2023). AI-generated content (AIGC): A survey [Preprint]. arXiv. <https://arxiv.org/abs/2304.06632>
- Yang, Y., Zhong, L., Li, S., & Yu, A. (2023). Research on the perceived quality of virtual reality headsets in human - computer interaction. *Sensors*, 23(15), 6824. <https://doi.org/10.3390/s23156824>
- Zeng, Z., Pantic, M., Roisman, G.I., & Huang, T.S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 39-58. <https://doi.org/10.1109/TPAMI.2008.52>
- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 586-595. <https://doi.org/10.1109/CVPR.2018.00068>

- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 586 – 595). <https://doi.org/10.1109/CVPR.2018.00068>
- Zhang, S., Deng, Y., Li, Z., Li, X., & Sun, M. (2023). Aligning human judgment and evaluation metrics in NLG. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023)* (pp. 1234-1245). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.acl-long.89>
- Zhang, Y., Xu, M., & Wang, L. (2018). Subjective evaluation methods for GAN-generated images: A review. *IEEE Transactions on Multimedia*, 20(9), 2456-2472. <https://doi.org/10.1109/TMM.2018.2825955>
- Zhang, Y., Yang, X., Xu, X., Gao, Z., Huang, Y., Mu, S., ... & Yu, G. (2024). Affective computing in the era of large language models: A survey from the NLP perspective [Preprint]. arXiv. <https://arxiv.org/abs/2408.04638>

APPENDIX A: Background questionnaire on the participation of experts (anonimised)

I. Basic information (multiple choice/fill-in-the-blank questions, remain anonymous)

1. your age range:
 - 25-27 years
 - 28-30 years
 - 31-33 years
 - 34 and over
2. Your gender:
 - male
 - daughter
3. Your current occupation/title:
 - Professor/Associate Professor
 - postdoctoral
 - Senior Engineer/Technical Leader
 - Head of UX/Product Design
 - Lecturer/Mid-level in R&D
 - Other (please specify): _____
4. The nature of your organisation is:
 - Comprehensive university
 - Engineering or art colleges
 - Corporate Research Institute / R&D Team
 - Other research institutes
5. Your research interests (multiple choice):
 - AIGC algorithm/large language model
 - Virtual reality/immersive interaction
 - Artificial Intelligence and Emotion Recognition
 - User Experience and Emotional Psychology
 - Digital Art/Image Generation
 - Natural Language Processing/NLP
 - Other (please specify): _____

II. Have you been involved in any of the following types of projects or research (multiple choice):

- AIGC-related scientific research or product evaluation
- Emotion-based user research or interaction design
- Virtual Reality/Digital Interactive Scenario Development
- Multimodal data acquisition and behavioural analysis
- Evaluation Scale Design and Subjective Scoring Tasks
- No experience of participation, but relevant theoretical knowledge

III. Do you have practical experience in any of the following areas? (Multiple choice)

- Use or test generative models such as ChatGPT, Midjourney, Stable Diffusion, etc.
- Participate or lead projects related to emotion recognition or AI generation
- Prompts written or optimised for AIGC-generated content
- Systematic evaluations, reviews or user studies of AI-generated content have been conducted
- Other relevant experience (please briefly describe): _____

APPENDIX B:

serial number	1	2	3	4	5	6	7	8	9	10
Time of submission of answer sheets	5/22/20 25 12:45	5/22/20 25 12:23	5/22/20 25 12:44	5/22/20 25 12:35	5/22/20 25 12:46	5/22/20 25 12:27	5/22/20 25 12:33	5/22/20 25 12:48	5/22/20 25 12:19	5/22/20 25 12:40
Time used	21 seconds	49 seconds	76 seconds	57 seconds	58 seconds	28 seconds	41 seconds	41 seconds	43 seconds	64 seconds

1. Your age range.	3	4	2	1	1	4	1	2	1	3
2. Your gender.	1	2	1	2	1	1	1	1	2	2
3. Your current occupation/title.	1	4	1	3	3	1	1	2	5	1
4. The nature of your organisation is.	1	3	2	3	4	2	1	4	3	1
5. Your research interests (optional): (AIGC algorithms/large language modelling)	1	0	0	0	1	1	0	0	1	0
5 (virtual reality/immersive interaction)	0	1	1	0	0	0	0	0	0	0
5 (Artificial Intelligence and Emotion Recognition)	1	0	0	0	0	1	0	1	1	0
5 (User Experience and Emotional Psychology)	0	1	1	0	0	0	1	0	0	1
5 (Digital Arts/Image Generation)	0	0	1	1	1	0	0	0	1	0
5 (Natural Language Processing/NLP)	0	0	0	1	1	1	0	1	0	0
6. Have you participated in any of the following types of projects or research (multiple choice): (AIGC related scientific research or product evaluation)	0	0	0	0	0	1	0	0	1	0
6 (Emotion-based user research or interaction design)	0	1	0	0	0	0	1	0	1	1
6 (Virtual Reality/Digital)	0	0	1	1	0	0	0	0	1	1

1 Interactive Scenario Development)										
6 (Multimodal data acquisition and behavioural analysis)	1	0	0	1	1	1	1	1	0	0
6 (Evaluation Scale Design and Subjective Scoring Tasks)	0	0	1	1	1	1	1	1	0	0
6 (no participatory experience, but relevant theoretical knowledge)	0	0	0	0	0	0	0	0	0	0
7 Do you have hands-on experience with any of the following? (Multiple choice) (using or testing generative models such as ChatGPT, Midjourney, StableDiffusion, etc.)	1	1	1	1	1	1	1	1	1	1
7 (Participated in or led projects related to emotion recognition or AI generation)	1	1	1	0	0	1	1	1	1	1
7 (Prompts written or optimised for use in AIGC-generated content)	0	1	1	1	1	1	1	1	1	1
7 (systematic evaluation, review or user study of AI-generated content)	0	1	1	1	0	1	1	1	1	1

APPENDIX C: AIGC Emotionally Generative Content Expert Evaluation Form

Respected experts:

Greetings! Thank you for your participation in the evaluation of this study. In order to systematically evaluate the adaptive performance of the AIGC platform in generating content guided by emotional cues, we would like to ask you to independently and objectively rate the text or image content generated by the platform by combining the following dimensions:

Rating scale: 1 = very poor; 2 = poor; 3 = fair; 4 = good; 5 = very good

Your rating will be used as important data for the evaluation of the platform's capabilities. All information will be used for academic research only and the evaluation process will be anonymous and independent.

Basic information (filled in anonymously)

Research Directions:

Current title/status:

Any experience with AIGC or emotion recognition related programs? Yes No

I. Evaluation of the content of the text

Please read the following text generated by the AI platform and rate it based on professional judgment:

Tip Mood: Pleasure

Content Number: QY-MC-00001

Platform Source: Microsoft Copilot

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

Cue Mood: Pleasure

Content number: QY-RP-00001

Platform Source: Replika

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

Tip Mood: Pleasure

Content number: QY-CA-00001

Platform Source: Character AI

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

Cue Mood: Pleasure

Content number: QY-GG-00001

Platform Source: Google Gemini

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

Cue Mood: Pleasure

Content number: QY-PE-00001

Platform Source: Poe

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

Cue Mood: Pleasure

Content number: QY-CG-00001

Platform Source: ChatGPT 4o

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

Cue Mood: Pleasure

Content number: QY-EB-00001

Platform Source: ERNIE Bot

Generative richness: is the content structure complete, is the language expression full, and does it reflect some semantic depth?

1 2 3 4 5

Emotional fit: Does the language style, rhythm, and wording match the emotional characteristics of the cue? Does it have a clear emotional orientation?

1 2 3 4 5

Detailed expression: Is it equipped with delicate descriptions, portrayal of emotional fluctuations, and is the logic smooth and natural?

1 2 3 4 5

II. Evaluation of image content

Please observe the content of the following images and professionally assess their mood-generating performance:

Cue Mood: Anxiety

Image No.: QY-MJ-00001

Platform Source: Midjourney

Generating richness: are the image elements rich, the composition complete, and does it have a good spatial hierarchy?

1 2 3 4 5

Mood Fit: Does the color palette, lighting, and compositional style fit the mood cue? Does it create the target atmosphere?

1 2 3 4 5

Detailed expression: Is the texture of the picture, edge treatment, and details of the main subject clear, and does it have a certain artistic expression?

1 2 3 4 5

Cue Mood: Calm

Image No.: QY-MJ-00001

Platform Source: Midjourney

Generating richness: are the image elements rich, the composition complete, and does it have a good spatial hierarchy?

1 2 3 4 5

Mood Fit: Does the color palette, lighting, and compositional style fit the mood cue? Does it create the target atmosphere?

1 2 3 4 5

Detailed expression: Is the texture of the picture, edge treatment, and details of the main subject clear, and does it have a certain artistic expression?

1 2 3 4 5

Cue Mood: Anxiety

Image No.: QY-SD-00001

Platform Source: Stable Diffusion

Generating richness: are the image elements rich, the composition complete, and does it have a good spatial hierarchy?

1 2 3 4 5

Mood Fit: Does the color palette, lighting, and compositional style fit the mood cue? Does it create the target atmosphere?

1 2 3 4 5

Detailed expression: Is the texture of the picture, edge treatment, and details of the main subject clear, and does it have a certain artistic expression?

1 2 3 4 5

Cue Mood: Calm

Image No.: QY-SD-00001

Platform Source: Stable Diffusion

Generating richness: are the image elements rich, the composition complete, and does it have a good spatial hierarchy?

1 2 3 4 5

Mood Fit: Does the color palette, lighting, and compositional style fit the mood cue? Does it create the target atmosphere?
1 2 3 4 5

Detailed expression: Is the texture of the picture, edge treatment, and details of the main subject clear, and does it have a certain artistic expression?
1 2 3 4 5

Cue Mood: Anxiety
 Image number: QY-DE-00001
 Platform Source: GPT 4o

Generating richness: are the image elements rich, the composition complete, and does it have a good spatial hierarchy?
1 2 3 4 5

Mood Fit: Does the color palette, lighting, and compositional style fit the mood cue? Does it create the target atmosphere?
1 2 3 4 5

Detailed expression: Is the texture of the picture, edge treatment, and details of the main subject clear, and does it have a certain artistic expression?
1 2 3 4 5

Cue Mood: Calm
 Image No.: QY-DE-00001
 Platform Source: GPT 4o

Generating richness: are the image elements rich, the composition complete, and does it have a good spatial hierarchy?
1 2 3 4 5

Mood Fit: Does the color palette, lighting, and compositional style fit the mood cue? Does it create the target atmosphere?
1 2 3 4 5

Detailed expression: Is the texture of the picture, edge treatment, and details of the main subject clear, and does it have a certain artistic expression?
1 2 3 4 5

APPENDIX D:

serial number	1	2	3	4	5	6	7	8	9	10	average value
1	Virtual Reality and User Experience Design	Artificial Intelligence and Affective Computing	Human-computer interaction and interface emotional design	Visual Communication and Digital Media Arts	Computer Graphics and Image Generation	language model	User psychology, digital behavior analysis	multimodal perception	Cognitive Science and Intelligent Educational Technology	Digital Media Technology	
2	Senior User Researcher	Professor/BoD	product manager	Associate Professor/Design Consultant	Associate Researcher/Technical Leader	lecture on	Associate Professor, Ph.D. in Psychology	doctoral	Associate Professor/Course Leader	Lecturer, PhD	
3	be	be	be	be	be	be	be	be	be	be	
4	2	1	2	3	3	3	3	3	5	4	2.9
5	4	3	4	4	5	4	4	4	5	4	4.1
6	2	2	3	2	3	3	3	4	4	5	3.1
7	1	1	1	2	2	2	2	3	3	2	1.9
8	2	2	2	2	2	3	3	3	3	4	2.6
9	4	1	2	1	2	2	3	2	3	3	2.3
10	1	1	2	1	2	2	4	3	3	3	2.2
11	2	1	2	2	3	3	3	3	3	3	2.5
12	2	2	2	3	4	2	2	2	2	2	2.3
13	4	5	5	5	5	5	5	5	5	5	4.9
14	5	5	5	5	5	5	4	5	5	5	4.9
15	5	5	4	5	5	5	5	5	5	5	4.9

16	2	2	3	4	3	3	3	3	4	2	2.9
17	4	4	3	5	4	4	5	3	4	4	4
18	2	1	3	3	3	3	3	4	4	4	3
19	5	5	5	5	5	5	5	4	5	5	4.9
20	5	4	5	5	5	5	5	5	5	4	4.8
21	4	4	5	4	5	5	5	5	4	4	4.5
22	3	3	3	3	4	4	5	4	3	4	3.6
23	3	2	2	3	3	4	4	4	4	5	3.4
24	2	2	3	3	3	3	2	3	4	4	2.9
25	5	5	5	5	5	5	5	5	5	4	4.9
26	5	5	3	5	5	5	5	5	5	5	4.8
27	5	5	5	5	5	5	4	5	5	5	4.9
28	4	4	5	3	3	5	4	4	4	4	4
29	4	4	4	4	5	4	4	4	3	4	4
30	4	4	4	2	4	4	3	3	4	3	3.5
31	3	3	3	3	3	4	4	4	4	4	3.5
32	5	3	4	4	4	4	5	5	4	4	4.2
33	2	3	2	3	5	3	3	4	3	4	3.2
34	5	5	5	5	5	5	4	5	5	5	4.9
35	5	5	5	4	5	5	5	5	5	5	4.9
36	5	4	4	4	5	4	4	3	5	5	4.3
37	5	3	4	4	4	4	4	5	4	3	4
38	5	3	4	4	4	2	4	3	4	4	3.7
39	5	2	3	5	2	4	4	4	3	3	3.5
40	3	3	5	2	4	4	3	3	4	4	3.5