

Research on Multimodal Interaction Paradigm of AI Painting

Zhao Yin^{1,a,*}

¹Guangzhou Huashang College, Guangzhou, Guangdong Province, China

^a1092553935@qq.com

*Corresponding author

Abstract: Based on a comprehensive analysis of the development status and problems of AI painting, this paper discusses how contemporary AI painting can grasp the meaning of traditional patterns and patterns under multi-modal interaction, and proposes the establishment of online AI painting designer circles and materials. It also visualizes the input information to create a portrait of the designer and shape the image of the designer's IP, creating an "AI painting multi-modal interactive inheritance paradigm" integrating design talents, resources, and projects, and integrates 3D virtual reality technology to build The multi-modal interactive space allows users to have a sense of belonging to collaborative design on the space platform. After registering an online account, users can use their own perspective to observe and select designers of their favorite styles, and enter the emotional elements that they ultimately want to obtain the finished product. Furthermore, visualized input data are utilized in this work to generate designer portraits and develop designer IP characters. An "AI Painting Multimodal Interactive Heritage Preservation Platform" is accordingly constructed, which provides a basis for the digital preservation of intangible cultural heritage as well as the digital inheritance of traditional culture.

Keywords: AI Painting, Multimodal Interaction, Paradigm, New teaching model

1. The Development Status Of AI Painting

Subsequent paragraphs, however, are indented. In 2015, the State Council identified artificial intelligence as a key development priority under the "Internet +" initiative, which was first incorporated into the Government Work Report in 2017^[1]. With advances in technology, artificial intelligence has been applied to creative production across music, film-television, literature, games, art and other domains, achieving rapid progress, especially in digital painting. In 2016, Google unveiled artworks created with the Deep Dream algorithm and achieved promising market returns^[2]. In February 2017, programmer Christopher Hesse built a deep-learning-based platform that transforms hand-drawn sketches into complete visual outputs, though the final rendering quality was limited by users' drawing skills^[3]. In April 2017, Google released AutoDraw, a machine-learning-assisted drawing tool that matches rough hand-drawn strokes against existing image databases to automatically generate finished artworks^[4]. In November 2017, Meitu launched Andy, the world's first AI-powered painting robot. Taking user selfies as inputs, it could generate illustrations in multiple styles with more than 30 visual effects^[5]. In October 2018, Edmond de Belamy — the first artwork produced via Generative Adversarial Network (GAN) by the French art collective Obvious — was successfully auctioned at Christie's New York^[6]. In 2019, Microsoft XiaoIce produced original paintings by learning artistic styles from 236 classical masters and held an exhibition at the Central Academy of Fine Arts^[7]. That same year, the SmartPaint system, jointly developed by Zhejiang University and Alibaba DAMO Academy, allowed users to rapidly produce high-quality paintings with artificial intelligence^[8]. In 2020, Microsoft upgraded XiaoIce to its eighth-generation version, which launched a serialized personal AI-painting portfolio and further enhanced its capabilities for intelligent artistic creation (see Figure 1).

In short, artificial intelligence can be deployed to learn diverse painting styles and generate stylistically consistent artworks. Meanwhile, as the dataset expands, the quality of AI painting will keep improving, thereby ushering in a new domain for pictorial art. AI painting represents a future development direction for painting. It implies that future computer systems will no longer rely solely on command lines and graphical interfaces for human-computer interaction; instead, they will engage with users intuitively through more natural, imperceptible multimodal interactions.

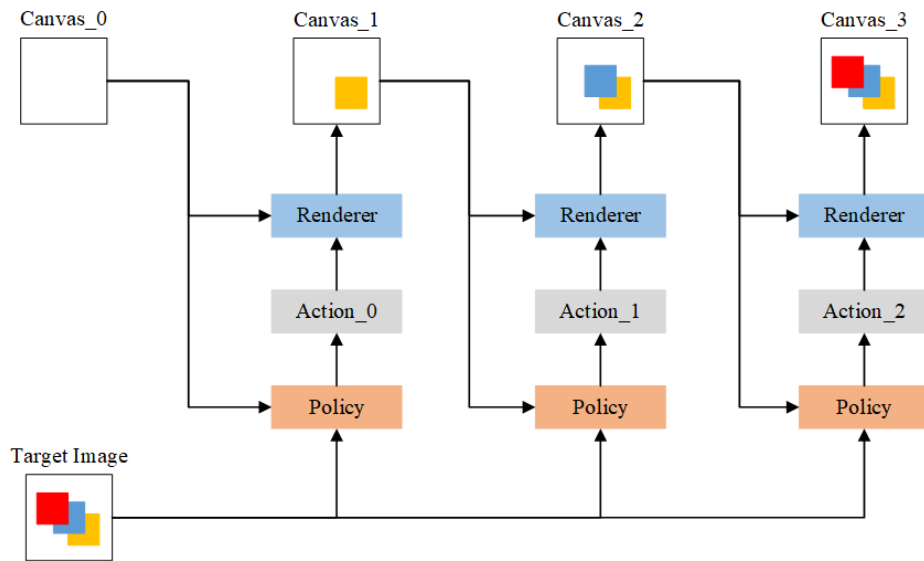


Figure 1. AI painting process

2. Multi-modal Interaction

The so-called "Modality" is a biological concept put forward by German physiologist Helmholtz, that is, the channel through which organisms receive information through perceptual organs and experience. For example, humans have vision, hearing, touch, taste and the olfactory modality, multi-modality is the fusion of multiple senses(see Figure 2).

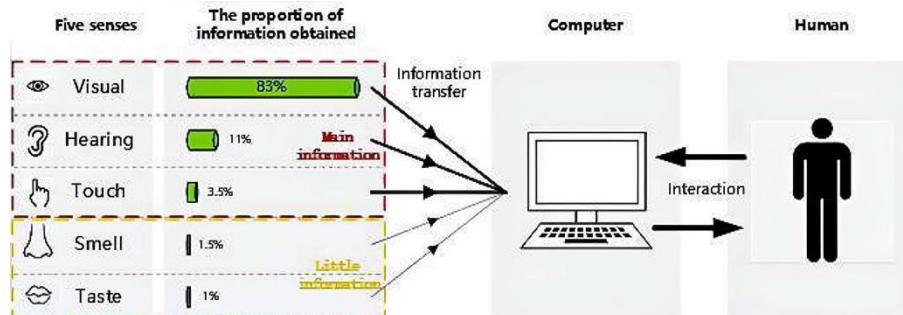


Figure 2. Multimodal interaction paradigm

According to research by scholars, among the ways that humans perceive information, the proportions of obtaining external information through sight, hearing, touch, smell and taste are 83%, 11%, 3.5%, 1.5% and 1%, respectively. Multi-modality refers to the fusion of multiple senses, and multi-modal interaction refers to the communication between people and the computer through multiple channels such as sound, body language, information carriers (text, pictures, audio, video), environment, etc., to fully simulate The way people interact with each other. This interaction method conforms to the morphological characteristics and user expectations of robot products, breaking the traditional PC-style keyboard input and the touch-sensitive interaction mode of smart phones. The multi-modal interaction method defines the exclusive interaction mode of the next generation of intelligent product robots. It laid the foundation for the research and development of related hardware, software, and applications.

3. The Construction form of Multi-modal Interactive Inheritance Paradigm

The "AI Painting Multimodal Interaction Inheritance Paradigm" leverages cloud computing technology to integrate design workflows into a Web-based resource pool, delivering professional design services for diverse user groups. Incorporating 3D virtual reality technology, it constructs a multimodal interactive space spanning visual, auditory and tactile dimensions, nurturing a sense of collaborative belonging among users. By building secondary platforms such as university youth designer teams and design material libraries, coupled with influencer marketing strategies, the paradigm capitalizes on the

brand value of the second classroom for AI painting talent cultivation. Furthermore, prioritizing the development of mini-programs and official WeChat accounts for user acquisition allows young designers to establish decentralized communities on the platform. This approach delivers services for relevant enterprises and end users that are better customized to practical demands and carry stronger emotional resonance.

3.1. Visual layer realization form

First, the system enables machines to fully learn diverse design styles by importing representative designer works and matching stylistic labels, thereby automatically generating exclusive IP images. Users entering the cloud design platform can select preferred designer styles and obtain personalized customized design solutions. Second, the AI platform learns the semantic connotation of diverse graphic elements and accumulates massive design resources, including flat, minimalist, hand-painted, classical, and new Chinese-style elements. These elements can be freely combined according to customer demands to produce high-quality customized designs. Taking the innovative brocade pattern as a typical case^[6], by inputting visual features of traditional red-flower-and-phoenix brocade and contemporary garment patterns, implicit folklore and cultural elements can be effectively extracted. On this basis, digital reorganization, partial pattern reconstruction, and style transformation are implemented while retaining the original color aesthetics and cultural connotation. Ultimately, traditional brocade aesthetics can be reconstructed and integrated into modern fashion contexts to form innovative contemporary patterns with traditional cultural charm. When inputting the visual modalities of the true red wearing flower and phoenix brocade pattern and contemporary clothing patterns, the representative folklore implied elements in the "new-like" brocade can be extracted, and then contemporary The design tool arranges and reorganizes these elements, or extracts part of the pattern based on the artistic conception for reorganization, or digitally transforms it while preserving the color and emotional meaning. In short, to create a top-tier modern pattern that can convey the charm of "new" brocade. The allegory of its elements should be integrated into the contemporary fashion context.



Figure 3. Extraction of Elements from the Patterns of True Red Flower and Phoenix Brocade

In terms of the extraction of the implied elements of the patterns, according to Huang Xiuzhong's "Shu Brocade Weaving Techniques", the "new" brocade pattern elements include: baoxiang flowers, grapes, hibiscus flowers, honeysuckle, mandarin ducks, ribbon birds, red birds, herons, cranes, etc(see Figure 3). These pattern elements are selected by people in their form, meaning or sound in the historical accumulation, and they are combined into a pattern with a certain symbolic significance by their imagination, that is, the combination pattern. According to Li Jigao's "History of Bashu Silk", the main embossed patterns in this brocade are honeysuckle pattern, treasure-like pattern, ribbon bird pattern, and heron pattern. Extracting the flower pattern in the "new pattern" brocade can output a new graceful and luxurious tangled branch pattern, elegant and gorgeous jewel pattern; extracting the bird pattern in the "new pattern" can output a ribbon bird pattern that symbolizes power and status, The heron pattern symbolizing a bright future and a broad career. In addition, the input of "auspicious", "Buddha", "rich and honor", "longevity" and other implied vocabulary can facilitate the re-creation of the "new one" in the future to randomly copy and combine new patterns that express the emotions of the aesthetic subject.

In terms of color and emotional element extraction, the development of digital printing and CAD technology has brought unprecedented development opportunities for textile printing and dyeing. The magnificent aesthetics of the Tang Dynasty urged the "new look" to be more colorful. However, the aesthetics of contemporary young people tend to be simple and elegant. If you want to meet the elegant

aesthetics of contemporary young people, you can try the yellow and royal blue in the brocade. , Violet is extracted to form a more simple and elegant new color combination. This method can also see the close connection between the pattern color arrangement and the visual beauty brought by the pattern. Therefore, when innovating pattern color matching, the CAD arrangement sequence of the pattern color should be preset to ensure that the design pattern color meets expectations. The effect of, enriches its emotional context in the reconstructed pattern. Then based on the predecessor's chromatographic analysis of the unearthed true red flower phoenix brocade, it has a wide variety of colors, as many as 24 hues, with eight-color silk threads such as bright red, pink white, dark green, light green, yellow, sapphire blue, and purple. The new craft weft yarn is woven by the flower method, and the color elements that are rich in folk customs in the brocade are extracted (see Figure 4), and then summarized and sorted according to the colored and achromatic systems, and it is found that the colored ones in the brocade are red. , Yellow, dark green, light green, sapphire blue, purple; white in achromatic series. Therefore, it is necessary to pay attention to the grasp and inheritance of its color expression in the process of recreating contemporary "new samples".

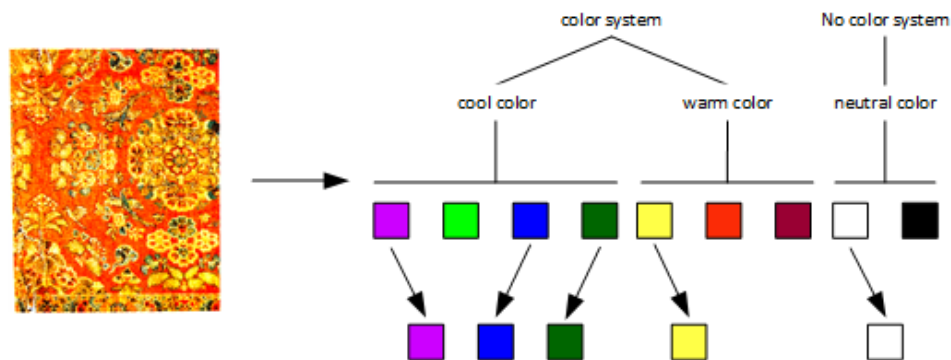


Figure 4. Extraction of color elements from true red flower and phoenix pattern brocade

On the whole, the pattern elements on the true red brocade with flowers and phoenixes are the crystallization of the wisdom of the Tang Dynasty brocade designers and craftsmen. It not only inherited the heroic and majestic Tang style, but also showed the plump and gorgeous Tang Yun. It indirectly explained the changing aesthetic tastes of the people of the Tang Dynasty, opened up a new generation of fashion in aesthetic style, and centered on myths in the past. The concept of decoration was transformed into an aesthetic trend that faces natural life and depicts real scenes. This craftsmanship is inseparable from the prevalence of "Buddhism and Confucianism" at that time. Each pattern carries Buddhist symbolism; for instance, the honeysuckle and Baotian flower motifs exhibit a strong Buddhist orientation. Contemporary aesthetic scholars have limited understanding of the historical evolution and unique cultural connotations of Tang Dynasty patterns. If pattern design innovations merely rely on the direct reproduction of ancient patterns, the resultant products will be stripped of their intrinsic cultural connotations. Accordingly, modern innovative design should be consistent with traditional cultural values and undergo reinterpretation to adapt ancient heritage to contemporary contexts.

3.2. Realization form of auditory layer

Human beings have the ability to create beautiful symbols and give them meaning. Recreation refers to both the process of constructing form and the product of this process. Therefore, in a sense, the reconstruction of the "new form" is tantamount to using symbolic interaction to rebuild a clear order for the transmission of information. And because the meaning of some patterns in the "new pattern" needs to be presented through concrete media, the changes in the media will prompt people to use computer information technology to break people's traditional patterns of inheritance, and draw viewers closer to the aesthetic subject. In addition to building the visual design layer and user models, adopting speech recognition and voice interaction for information delivery, and setting up professional influencer sales and after-sales services can effectively cut the time and design costs involved in order intake, product production and promotion. This also constitutes an optimal approach to inheriting the "New Chinese Style" and traditional Chinese aesthetic concepts. For example, the voice recognition function is used to match specialists in planning, design, promotion, sales and after-sales service. With order follow-up, customers can choose platform designers to conduct "one-to-one" communication services according to their needs. Designers should recommend popular elements for customers to choose according to customer needs and the style preference of cloud design platform evaluation, and use voice or video to update the design drawings Modifications to grasp the needs of consumers in real time and improve

customer satisfaction.

3.3. Realization form of tactile layer

In addition to voice and video communication, remote technology that uses modern technology to reproduce "new" weaving techniques and achieve "time and space integration" long-distance transmission is also one of the multi-modal interactive transmission methods. As an art form pursuing remote presence, telepresence technology is an artistic means of long-distance dissemination, which can enable the aesthetic subject to obtain a remote presence touch experience, and it can also make people produce a three-dimensional look and feel. For designers, this kind of interaction that breaks through the limitations of time and space has important teaching significance for the reconstruction of traditional patterns. In 1992, Paul Selmon launched a landmark work called "Dream of Long-distance Communication", which uses a bed as a medium to interact with two people in different spaces remotely and in real time via the Internet. In the same way, this method can be applied to the sharing of the "new pattern" weaving process, allowing the aesthetic subject and object to achieve "time and space integration", so that the inheritance of the "new pattern" weaving technique can break through the inherent pattern of precepts and deeds. It enables remote communication and enriches its dissemination forms to satisfy contemporary people's demand for AI painting.

For example: using PS to artistically process the pictures, get the clothing pattern that the individual expects again, and use the three-color fading method and the 3D pleating production process to produce a strong three-dimensional touch. The background color of the brocade is red with high color saturation. The gorgeous and festive colors reflect the prosperous, peaceful and peaceful social atmosphere of the Tang Dynasty. The design of contemporary clothing patterns needs to take the designer's original intention as the emotional keynote, and use it as the emotional main line of the design throughout the design activities. Obviously, it is not suitable to use overly complicated patterns for design, so designers can digitally deform and apply the stylized patterns on the clothing on the basis of retaining the original color charm, which not only retains the meaning of the original patterns but also Expresses personal aesthetic inclinations(see Figure 5).

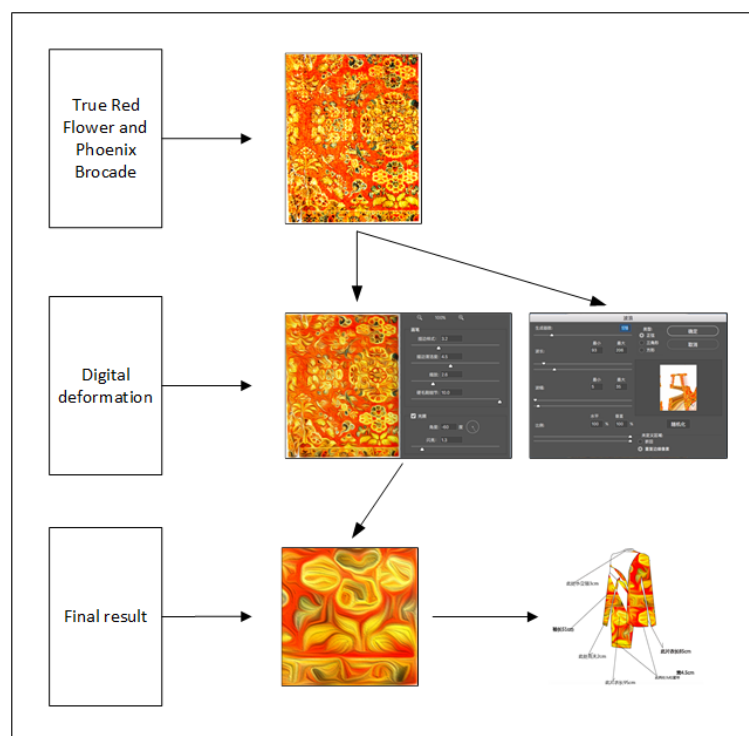


Figure 5. Digital transformation and application of pattern

4. Conclusions

To sum up, in the contemporary era of continuous integration and development of art and technology and information fragmentation, the design circle needs more integration and sharing of core information.

While integrating innovative design materials, the platform attaches greater importance to fostering a designer community. This not only effectively accelerates the development of new technologies, forms and models and the optimization and upgrading of design workflows, but also marks the focal point of future design development. The established multimodal interaction paradigm, which combines the advantages of manually designing networks and neural network search (NAS), can both describe some general design principles of simple networks and generalize them in various settings, while exploiting semi-automatic process to find simple pattern models that are easy to understand, build, and generalize. As a platform driver, it can not only allow young designers to share their design styles on the platform, but also take the "new pattern" brocade as an example, people can first draw their patterns with AI on the basis of pattern elements, and store them in "New Pattern Brocade Pattern Library", and then in the process of reconstruction, you can quickly select patterns with personal aesthetic tendencies for design reorganization, and then classify Shu Brocade patterns into "Shu Brocade Pattern Library" according to the age, and finally realize the AI painting material. Cloud integration can also indirectly guide the market-oriented trend of the reproduction of traditional patterns and their living inheritance under the contemporary aesthetic trend, trying to reduce the abuse of patterns in the above-mentioned film and television works that deviate from the times and the emergence of intangible culture in contemporary young people.

Acknowledgements

This study was funded by 2025 Project for Scientific Research Talents of Guangzhou Huashang College Research Project, Project Name is Digitalization of Intangible Cultural Heritage Exhibition Halls (Project No.:2025HSPY450), and the 2025 School-level Quality Engineering Project, Guangzhou Huashang College, Generative AI Empowering Classroom Teaching: Innovation and Practical Research on the Dual-teacher Collaborative Teaching Model (Project Number: HS2025ZLGC78).

References

- [1] *The State Council of the People's Republic of China. Guiding opinions on actively promoting the "Internet +" action*[EB/OL]. (2015-07-04)[2026-09-16]. http://www.gov.cn/zhengce/content/2015-07/04/content_10002.htm.
- [2] Mordvintsev A, Olah C, Tyka M. *Inceptionism: going deeper into neural networks*[EB/OL]. (2015-06-17)[2026-09-16]. <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>.
- [3] Hesse C. *Edges2cats: image-to-image translation demo*[EB/OL]. (2017-02-20)[2026-09-16]. <https://github.com/affinelayer/pix2pix-tensorflow>.
- [4] Google Creative Lab. *AutoDraw*[EB/OL]. (2017-04)[2026-09-16]. <https://experiments.withgoogle.com/autodraw>.
- [5] Meitu Company. *World's first AI robot Andy released*[EB/OL]. (2017-11)[2026-09-16].
- [6] Christie's. *An artwork created by an algorithm*[EB/OL]. (2018-08-20)[2026-09-16]. <https://press.christies.com/release-christies-announces-auction-of-an-artwork-created-by-an-algorithm/>.
- [7] People's Network. *AI portrait sold for \$430,000; XiaoIce holds solo exhibition at Central Academy of Fine Arts*[EB/OL]. (2019-07-25)[2026-09-16]. <https://ai.people.com.cn/n1/2019/0725/c422228-31255085.html>.
- [8] SUN L Y, CHEN P, XIANG W, et al. *SmartPaint: a human-computer collaborative painting system based on generative adversarial networks*[J]. *Frontiers of Information Technology & Electronic Engineering*, 2020, 21(10).