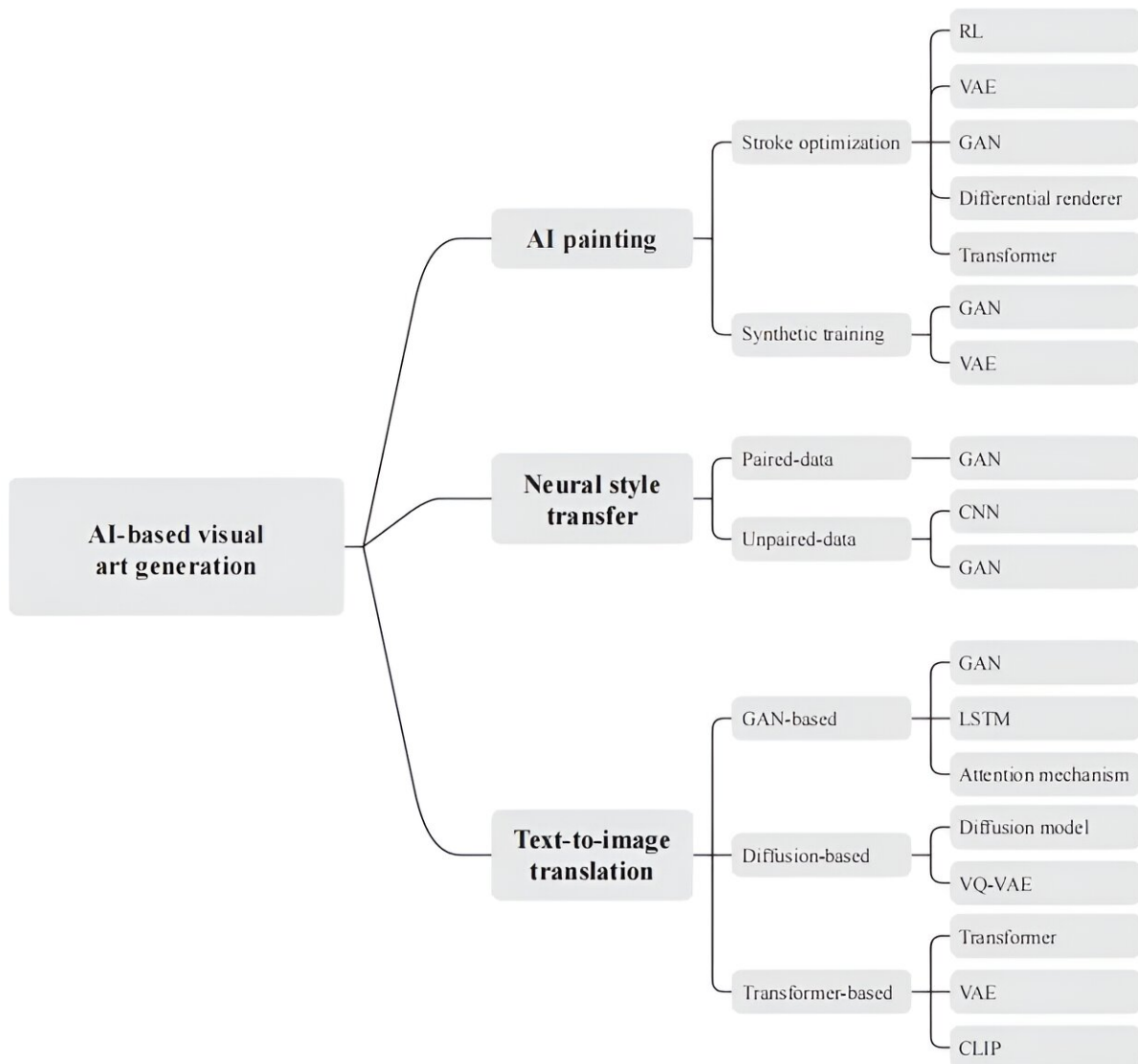


Cogeneration of innovative audio-visual content: A new challenge for computing art

March 22 2024



Summary of AI-based visual art generation. Credit: Beijing Zhongke Journal

Publishing Co.

Walter Benjamin came up with aura and authenticity in "The Work of Art in the Age of Mechanical Reproduction" in 1936 to describe the value of original artworks created by artists instead of mechanical copies. He wanted to defend artificiality and support traditional fine arts.

Not singly but in pairs, Mitchell's core concern was about the work of art in the age of biocybernetic reproduction in "What Do Pictures Want?: The Lives and Loves of Images" (2005) followed Benjamin's train of thought. He especially mentioned the first cloned sheep, Dolly, and regarded it as a living image. Mitchell is one of the most representative and influential theorists in the arts and human sciences.

He developed a series of concepts to explain the value of biocybernetic artworks. However, based on carbon, artists in this field are limited by the development of biotechnology, and their creation is sometimes repressed instead of released. Some disappointing cooperation between biologists and artists seems to popularize scientific knowledge by entertaining visual culture and show a lack of critical thinking. Therefore, researchers need to reflect on what artists can do with advanced technology.

Humans have faced new challenges in the era of the metaverse. Humans have not only mechanical copies of artworks or biocybernetic replicas but also avatars of humans themselves.

A [paper](#) in *Machine Intelligence Research* proposes the concept of artificial intelligent (AI) art, thus summarizing the main features of artworks produced by artificial intelligent technology, such as extended reality (XR, the combination of VR/AR/MR), cyber-physical system

(CPS), cloud computing, and blockchain.

The cooperation between AI technical staff and artists is more intimate than that between biologists and artists. AI technology discharges artists from laborious work, which was primarily accused by Marx and others, and encourages them to realize their full potential in art. As a result, AI is like a capable partner in the team who always understands the artist in time and does intense work to bring the artist's romantic conceiving into reality.

With bright prospects for development, AI technology plays essential roles in design, creation, and exhibition in art circles. The concept of AI art may be easily confused with computer art. It is important to note that AI art is more advanced than computer art and can cover more perceptual requirements, including optical and acoustical requirements. AI art usually presents a fusion of the human senses.

Art appreciators obtain visual, aural, and tactual feelings simultaneously. In other words, AI technology provides a rich audio-visual feast for modern art exhibitions.

The development plan for AI art is still in its infancy. There are some concerns in society about the general applicability of AI technology. Here is an interesting question that has always been mentioned. Do AI dream of electric sheep? In 1968, Philip K. Dick first put this question forward in his science fiction: "Do Androids dream of electric sheep?" The inspiration for the films "Blade Runner" and "Blade Runner 2049."

After discussing AI ethical issues, this title has become the core question that represents the fear that AI would replace humans. Such fears soon spread to the humanities. Some intellectuals believe there should be limits to AI technology. However, if one has sufficient knowledge of AI technology, one will find such fears laughable. People's fear is nothing

more than a rejection of the unknown.

Cutting-edge AI technology still needs to reach the emotional level of humans. The urgent need to work on developing and applying AI technology remains as strong as possible.

AI technology is currently used in the art field for technique classification, style migration, interactive design, manufacturing, cultural industry, and so on. AI art has produced AI-generated poetry, VR painting, digital media art, AI voiceovers, and smart electrical appliances. These examples show the solid creative power of AI art.

However, some artists are not happy with it. In 1972, the German artist Joseph Beuys gave a speech at Documenta in Kassel, presenting the idea that "everyone is an artist." His views have caused an uproar. In those days, it was nothing more than an imagination. After all, not everyone was skilled in creating art. With the development of AI art, this idea seems to be becoming a reality. AI is powerful enough to allow anyone to become an artist.

It should be noted that the creative ability of AI is not endless. It comes from humans who have talents in creating art. The development of AI art is, therefore, not incompatible with the training of artists. In contrast, the spread of AI art enables artists to do what they do best. In this way, AI art development and traditional art innovation can hold a win-win situation.

To be on target, a paper by Prof. Gao Feng from Peking University focuses on AI-generated video and AI-generated audio. Audio-visual ability is often thought of as a composite human sensory ability. Their combination has rapidly improved the production efficiency of industries such as movies, short videos, and games. The summary of AI visual and auditory technology and the presentation of existing results

can help practitioners in the industry determine art industry trends in the future.

Audio-visual art generation can be divided into visual art generation and auditory art generation. Section 2 of the paper provides a comprehensive overview of the datasets and methods in the two fields.

Visual art generation part: first, researchers introduce ten classic image datasets; then, based on three tasks of AI painting, style transfer, and text-to-image translation, researchers summarize the classic models in the field of visual content generation; finally, researchers show typical systems and products for it.

Auditory art generation part: they use the form of sound expression as an indicator, specifically listing eight classic music datasets in the field of auditory art generation; then, regarding the model structure as standard, the music generation methods are divided into two categories, general model, and composite model. Researchers outline nine classical frameworks for music generation and identify related models and products.

There are two types of evaluation methods for algorithm performance: objective evaluation and subjective evaluation. Objective evaluation applies several metrics based on mathematical theory, which is quantitative, efficient, and widely used, but it is not suitable for content that requires subjective feelings. Subjective evaluation usually requires the design of experiments, and observers evaluate the results of the algorithm, which is time-consuming, laborious, and difficult to quantify.

Nevertheless, subjective evaluation is consistent with subjective feelings. In the field of art generation, subjective evaluation plays an important role in evaluating the creativity of the model. In Section 3, researchers provide an overview of measuring the quality of generated results from

objective and subjective perspectives.

Section 4 introduces the proposed materials and mechanism.

Cogeneration of audio-visual content is a multimodal task and requires approaches to fuse information from different sources, including image, video, audio, text, etc. By weighing the strengths and limitations of various audio-visual art generation algorithms, researchers develop and propose a joint generation mechanism for generating digital audio-visual artworks using multiple types of algorithms.

The system is divided into a visual art generation module and an auditory art generation module. The former is responsible for generating dynamic video content of a specified style, and the latter generates the corresponding video soundtrack through the text features associated with the video. In Section 4.1, researchers introduce two datasets constructed for audio-visual joint tasks. In Sections 4.2 and 4.3 they demonstrate the visual art generation module and the auditory art generation module, respectively.

This paper has summarized the results of the technical development of audio-visual art generation. The technology of audio-visual art generation has a wide range of applications. It can be used at home to make entertainment more diverse. It can also be used in public places. For example, it can increase the attractiveness of commercial promotion and art exhibits.

A study proposed a new museum archiving system that applies AI technology to the service of art institutions, including museums. Studies such as this show widespread interest in AI-based computing art, which can facilitate people's daily lives and empower the development of cultural industries.

Furthermore, visual art generation and auditory art generation will

revolutionize the way art is produced and increase its productivity. However, this inevitably poses some challenging issues. Traditional artists have shown great anxiety about the development of computing art. They fear that computers will soon replace their jobs. This concern is not unwarranted. AI is increasingly replacing manual labor.

There are two aspects involved in Section 5. On the one hand, researchers need to clarify whether computing art qualifies to replace artificial art. On the other hand, researchers need to know whether computing art instead of artificial art is more beneficial to the well-being of society. In summary, the main challenges of AI-based computing art can be summarized as the artificial and intelligent aspects of computing art.

This paper has provided a comprehensive survey on audio-visual content generation. Researchers hope that this review will help people better understand the research field of audio-visual art and the development tendency of AI-based [generation](#).

More information: Mengting Liu et al, Cogeneration of Innovative Audio-visual Content: A New Challenge for Computing Art, *Machine Intelligence Research* (2024). [DOI: 10.1007/s11633-023-1453-5](https://doi.org/10.1007/s11633-023-1453-5)

Provided by Beijing Zhongke Journal Publishing Co.

Citation: Cogeneration of innovative audio-visual content: A new challenge for computing art (2024, March 22) retrieved 9 May 2024 from <https://techxplore.com/news/2024-03-cogeneration-audio-visual-content-art.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.