Human Interaction and Emerging Technologies (IHIET 2024), Vol. 157, 2024, 460–467
https://doi.org/10.54941/ahfe1005505

Interpolation and Depth Extraction: A Case Study of Shan Shui Artwork Generated by AI

Man Lai-Man Tin

School of Arts and Social Sciences, Hong Kong Metropolitan University, Ho Man Tin, Kowloon, Hong Kong

ABSTRACT

Traditional Shan Shui artworks (Chinese landscape paintings) have been static representations of the beauty and tranquillity of landscapes, and they have a long history and significance in Chinese art. The advancement of artificial intelligence (AI) technologies brings new possibilities to artwork creation and innovation to tradition. This study proposes using AI technologies, specifically artificial neural networks and computer vision, to learn from traditional paintings, generate new landscapes and extract depths in Shan Shui paintings. The research aims to go beyond using AI technology solely to create new artwork. Instead, it explores the ability of AI to generate dynamic landscapes with perspectives, allowing more immersive and engaging experiences, and through analysis of the depths embedded in the AI-generated Shan Shui paintings, trying to gain insights into understanding interpolation of spatial and dimensional aspects in the work and address the limitation of 2-dimension in art. This research signifies the convergence of art and technology, explores novel ways of creating and viewing traditional Shan Shui paintings, and explores the possibilities of understanding the landscapes generated through the lens of AI and computer vision technology.

Keywords: Artificial intelligence, Creative AI, Shan Shui painting, Interpolation, Depth extraction, Computer vision, Machine learning art generation

INTRODUCTION

Beginning in around the 4th century, Shan Shui is a traditional form of Chinese painting that depicts natural landscapes and scenery. The term 'Shan Shui' means 'mountain and water,' and such paintings always consist of key elements such as mountains, waterfalls, plants, bridges, houses, flowing water, and sometimes feature refined or official scholars. The Shan Shui paintings reflect the artist's unique observations, travel experiences, artistic representations, imagination, and philosophical representation of the landscape. Gradually, Shan Shui has evolved into a distinct art form characterised by its portrayal of depth and perspective of nature. The composition and geographic elements of Shan Shui paintings represent changes in perspective, signifying the spatial and dimensional travel experience of the artist in nature.

With the advancement of technology, there have been many attempts to use new tools and technologies, including artificial intelligence, to create Shan Shui paintings, which paved a good foundation and exploration of how technology can influence traditional Shan Shui paintings. For example, Alice Xue in her research 'End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks' (2021) introduced the use of Sketch-And-Paint GAN (SAPGAN) to generate Chinese landscape paintings from end to end, without conditional input, demonstrates the capability of AI and deep learning technology to create traditional Shan Shui paintings. However, there has been limited research on applying AI technologies to address and understand the perspective, spatial, and depth aspects of generated Shan Shui paintings. This research proposes the use of artificial intelligence technologies to learn from and generate Shan Shui paintings, aiming to explore the possibilities of creating dynamic landscapes, which seeks to address the limitation of traditional 2-dimensional static artwork and to gain a better understanding of the transformation of perspectives. In the second part of this research, the depth within Shan Shui paintings is further investigated through AI and computer vision technologies, which allows an analysis of the AI-generated landscape correspondence to the composition and understanding of the traditional Shan Shui landscapes.

TEST OF SHAN SHUI PAINTING GENERATION

To train AI, particularly machine learning, in the learning and generation of Shan Shui paintings, tests were conducted using Feng Shui (Chinese Geomancy, which means wind and water) landscape paintings with Shan Shui elements sourced from the internet as datasets. The datasets were used to generate new sets of paintings. 30 Feng Shui landscape paintings were employed in this attempt.

The Feng Shui landscape painting represents a variation of the traditional Shan Shui painting. It embodies some key principles of Shan Shui painting and incorporates the rules of Feng Shui. The structure of the painting and the geographical layout of Shan Shui and Feng Shui elements within the paintings have clear and standardised formats. Thus there is relatively less variation among different Feng Shui landscapes compared to traditional Chinese landscape paintings. This makes Feng Shui landscape paintings a suitable data set for the initial AI training and test to generate pictures with Shan Shui elements.

In this test, the key elements of landscape paintings, such as mountains, waterfalls, and flowing water were extracted and formed new combinations. The results demonstrate that the fundamental layout of Feng Shui landscape paintings is manifested in the generated pictures, indicating the ability of artificial intelligence to learn and generate landscape works, with the landscapes and Shan Shui elements visible in the results (see Figure 1).

Figure 1: Samples of Feng Shui landscape paintings generated by AI.

The next phase of this research aims to investigate the potential for creating dynamic, multi-perspective landscapes within Shan Shui paintings, in order to capture the spatial variations experienced across different artworks. The depth embedded in perspectives will be analyzed using AI technology in a subsequent stage.

Traditional Shan Shui paintings, due to the static nature, capture only a static moment of landscapes. To overcome this limitation, Shan Shui artists often employ multiple vanishing points to introduce multi-perspective views of landscape within the artwork. This study aims to explore the possibilities of generating dynamic work by transforming 2-dimensional static images and merging multi-perspective views from different Shan Shui paintings.

To achieve this, interpolation is employed to create interpolated frames between the generated images, creating transitional visual effects between the landscapes. The results demonstrate the ability of artificial intelligence to learn and generate dynamic landscapes (see Figure 2).

Figure 2: Samples of interpolated frames created between the generated pictures.

The learning, creation, and generation of paintings featuring Shan Shui elements, along with the presentation of multi-perspectives, are primarily accomplished using AI. The next phase is to learn and generate more complex

Shan Shui paintings of various styles and compositions, based on traditional Shan Shui paintings sourced from the internet.

AI-GENERATED DYNAMIC SHAN SHUI PAINTING

In traditional Shan Shui paintings, instead of directly depicting the landscape, the artists incorporate their own artistic system and expression into the paintings, making them a unique art form of landscape expression. For example, Christin Bolewski, in her paper "'Shan-Shui-Hua' – Traditional Chinese Landscape Painting Reinterpreted as Moving Digital Visualisation" (2008) discusses the characteristics of traditional Chinese landscape painting, unlike Western use of single-point perspective, traditional Chinese landscape painting employs multi-perspectives, creating 'spatial illusions' in the artwork.

Building upon the findings of the previous tests, this section aims to integrate machine learning with interpolation technology to create dynamic landscapes within Shan Shui paintings. The generated Shan Shui paintings incorporate some of the original compositions and traditional Shan Shui painting elements such as mountains, trees, bridges, and flowing water, as well as the traditional artistic systems of multiple vanishing points, indicating the capacity of AI to learn and generate dynamic Shan Shui paintings while incorporating the distinctive expression of multi-perspectives (see Figure 3 and Figure 4).

Figure 3: Samples of the generated traditional Shan Shui paintings with multi-perspectives.

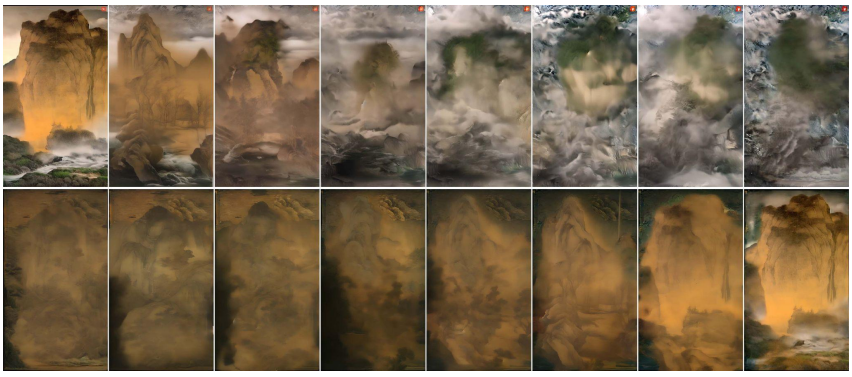


Figure 4: Samples of interpolated frames created between the generated Shan Shui paintings.

The dynamic landscapes signify the changes in natural scenery across different landscapes and the travel experiences in nature with multi-spot views from different Shan Shui paintings. The interpolated frames connect the generated Shan Shui, forming an ongoing loop of landscapes. The work extends the spatial and temporal aspects of traditional Shan Shui painting, in which the travel experiences of different sceneries, times, and perspectives are expanded from the traditional static form of artwork. The variation of Shan Shui composition and dimension can guide the viewer's observation through landscapes in more immersive and engaging experiences, allowing for a deeper integration with nature, which echoes the spirit of traditional Shan Shui artistic imagination and philosophical representation of landscapes. For example, 'Daoists, in particular, had a great affinity for experiencing man in the natural environment and seeking a kind of harmonious unification between man and nature,' as highlighted in Sophia Suk-mun Law's research 'Being in Traditional Chinese Landscape Painting' (2011).

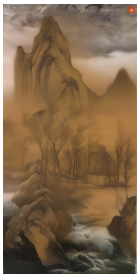
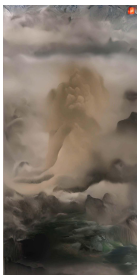
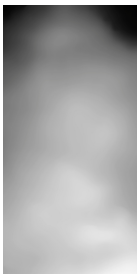
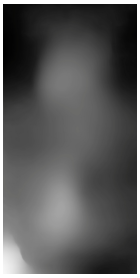
DEPTH EXTRACTION THROUGH AI AND COMPUTER VISION

Traditional Chinese Shan Shui painting features a unique representation of landscape composition and perspective. In addition to key landscape elements such as mountains, waterfalls, and plants, the depiction of spatial and depth aspects is also crucial for understanding how AI-generated works resonate with viewers. To better understand the depth embedded in AI-generated landscapes, AI and computer vision were used to extract the depth from dynamic Shan Shui paintings, attempting to analyse the interpolated spatial and dimensional visual information, which are particularly essential to examine the depth created in the AI landscape and how things can make sense to human in the age of machine intelligence.

Several attempts, such as deep learning and transformer models, which have been available on software platforms, were used to examine the depth in the artwork by estimating depth information from images and videos generated by AI. The results were compared to assess the ability of AI and computer vision to estimate depth accurately and determine whether the estimated depths are aligned (see Table 1).

The results indicate the ability of various AI and computer vision models to analyse spatial information and extract depth from AI-generated landscapes. The depths extracted in set 1 are the most pronounced compared to the other sets. The silhouette of the mountain can be clearly seen, and the levels of both near and far distances in the landscape can generally be differentiated. The deep learning transformer model also distinguishes the depth of the scenery, which can make sense to viewers. However, the depths extracted using the Depth Perception Transformer (DPT) in sets 2 to 5 are unclear. This is likely due to the unclear landscapes created in the interpolated frames, particularly noticeable in the original AI-generated landscapes in sets 3 to 5. Similar outcomes were observed in the use of the deep learning transformer model. As the landscapes become more difficult to interpret, the depth in the Shan Shui cannot be distinguished. For example, in set 5, the depth is essentially just a flat plane without any obvious depth of field or distance. Despite this, neural networks can analyse the dynamic landscape and capture the basic silhouette of the mountain as well as the distance and depth of field.

Table 1. Comparison results of depth estimation through AI and computer vision technologies.

Set	Original AI-generated landscapes	Depth extracted by Depth Perception Transformer from still landscapes	Depth extracted by deep learning transformer model from still landscapes	Depth extracted by neural networks from dynamic landscape
1				
2				
3				
4				
5				

With reference to the research by Chi-Min Hsieh and Hsiao-Ching Chou, 'Computational Three Distances: Exploring the Aesthetics of the Southern Song Dynasty and Its Adapted Simulation and Rendering' (2023), the difficulty in interpreting the landscape may also reflect the intersection between the landscape and fog or clouds. For example, more areas of the landscapes are covered by fog or clouds in sets 2, 3, and 4 compared with set 1, in which 'the view of distant mountains from the near bank of a wide body of water' is generated. Additionally, the fragmentary scenery in set 5 makes it difficult to identify the space and depth of the landscape. These align with the traditional theory of distances in Shan Shui painting, namely 'hidden distance,' 'board distance' and 'obscure distance' respectively.

The accuracy of the proposed methods still has room for improvement. Depth can be extracted by utilising AI and computer vision technologies, indicating a certain degree of comprehensibility in AI-generated content. This showcases the potential for AI technology to assist viewers in interpreting the content and therefore how the 'hidden' or 'potential' artistic spirits could be realised through computer program algorithms, as highlighted in 'Rendering and presentation of 3D digital ink landscape painting' Xunxiangby et al. (2023).

CONCLUSION

This research demonstrates the convergence of art and technology, particularly, traditional Shan Shui paintings and artificial intelligence to learn and create landscape paintings. By employing artificial neural networks and computer vision technologies, the study not only tried to create new artwork, but also explored the ability of AI to generate dynamic landscapes with multi-perspectives, and extraction of depth in the landscapes to enhance the viewer's engagement and examine the comprehensibility of AI-generated content.

The interpolation employed in this research demonstrated the ability of image processing and computer vision to create interpolated frames between the generated landscape paintings. This allowed viewers to engage with the artwork with an experience expanded from traditional spatial and temporal dimensions, in particular, the dynamic perspectives, scenes, and spaces in the landscapes.

Depth extraction using AI and computer vision was another focus of this research. Various attempts were made to analyse the depths embedded in the generated Shan Shui paintings. Although the depths were not always clearly distinguished in some attempts, neural networks generally captured the field of depth when extracted from the dynamic landscape. This indicates the possibility of comprehending AI-generated content through AI and computer vision technologies. This research provides a foundation for future studies on the integration of AI in creative processes and the elements that affect the comprehension of traditional artwork, such as traditional Chinese ink paintings.

REFERENCES

Bolewski, C. (2008). 'Shan-Shui-Hua' – Traditional Chinese Landscape Painting Reinterpreted as Moving Digital Visualisation. Electronic Workshops in Computing. https://doi.org/10.14236/ewic/eva2008.4

Hsieh, C.-M., & Chou, H.-C. (2023). Computational three distances: Exploring the aesthetics of the Southern Song Dynasty and its adapted simulation and rendering. SIGGRAPH Asia 2023 Art Papers. https://doi.org/10.1145/3610591.3616426

Law, S. S. (2011). Being in Traditional Chinese Landscape Painting. Journal of Intercultural Studies, 32(4), 369–382. https://doi.org/10.1080/07256868.2011.584615

Li, X., Huang, Y., Jiang, Z., Liu, Y., & Hou, Y. (2023). Rendering and presentation of 3d digital ink landscape painting. Computer Animation and Virtual Worlds, 35(1). https://doi.org/10.1002/cav.2215

Xue, A. (2021). End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). https://doi.org/10.1109/wacv48630.2021.00391