Zhang W, Zhang JW, Wong KK *et al*. Computational approaches for traditional Chinese painting: From the "Six Principles of Painting" perspective. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 39(2): 269-285 Mar. 2024. DOI: 10.1007/s11390-024-3408-x

Computational Approaches for Traditional Chinese Painting: From the "Six Principles of Painting" Perspective

Wei Zhang[1] (张　玮), Jian-Wei Zhang[1] (张建伟), Kam-Kwai Wong[2] (黄锦畦), Yi-Fang Wang[3] (王懿芳)
Ying-Chao-Jie Feng[1] (封颖超杰), Lu-Wei Wang[1] (王璐玮), and Wei Chen[1, 4, *] (陈　为), *Member, CCF*

[1] *State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058, China*

[2] *Department of Computer Science and Engineering, The Hong Kong University of Science and Technology Hong Kong 999077, China*

[3] *Kellogg School of Management, Northwestern University, Evanston 60208, U.S.A*

[4] *Laboratory of Art and Archaeology Image, Zhejiang University, Hangzhou 310058, China*

E-mail: zw_yixian@zju.edu.cn; zjw.cs@zju.edu.cn; kkwongar@connect.ust.hk; yifang.wang@kellogg.northwestern.edu
fycj@zju.edu.cn; ppwlwpp@zju.edu.cn; chenvis@zju.edu.cn

Received May 18, 2023; accepted March 12, 2024.

Abstract Traditional Chinese painting (TCP) is an invaluable cultural heritage resource and a unique visual art style. In recent years, there has been a growing emphasis on the digitalization of TCP for cultural preservation and revitalization. The resulting digital copies have enabled the advancement of computational methods for a structured and systematic understanding of TCP. To explore this topic, we conduct an in-depth analysis of 94 pieces of literature. We examine the current use of computer technologies on TCP from three perspectives, based on numerous conversations with specialists. First, in light of the "Six Principles of Painting" theory, we categorize the articles according to their research focus on artistic elements. Second, we create a four-stage framework to illustrate the purposes of TCP applications. Third, we summarize the popular computational techniques applied to TCP. This work also provides insights into potential applications and prospects, with professional opinion.

Keywords traditional Chinese painting (TCP), digital humanities, cultural heritage, computer vision, deep learning

1 Introduction

Originating from the Han Dynasty, traditional Chinese painting (TCP) has been a primary art form in China[1], characterized by artistic expressions depicted on paper and silk with brushes dipped in black ink and Chinese pigments. TCP has been used to express the author's artistic creativity and insinuate criticism of the society, philosophy, and politics of the time (Fig.1(a)). Existing studies on TCP mainly focus on its history, explanation, and appreciation of paintings, as well as various styles and techniques. However, these are typically conducted through a close-reading and case-study approach based on painting theories[4], which while insightful, is time-consuming and unscalable for studying the patterns and evolution trends of TCP that have emerged over the centuries.

Recent advances in computer technology have greatly improved the efficiency of TCP research. For example, convolutional neural networks (CNNs) detect and segment the elements in TCP[5–7], while generative adversarial networks (GANs) generate TCP and perform style transfer[8–12]. However, applying

Survey

Special Section of CGI 2023

This work was supported by the Fundamental Research Funds for the Central Universities of China under Grant No. 226-2022-00235, and the Key Program of the National Natural Science Foundation of China under Grant No. 62132017.

*Corresponding Author

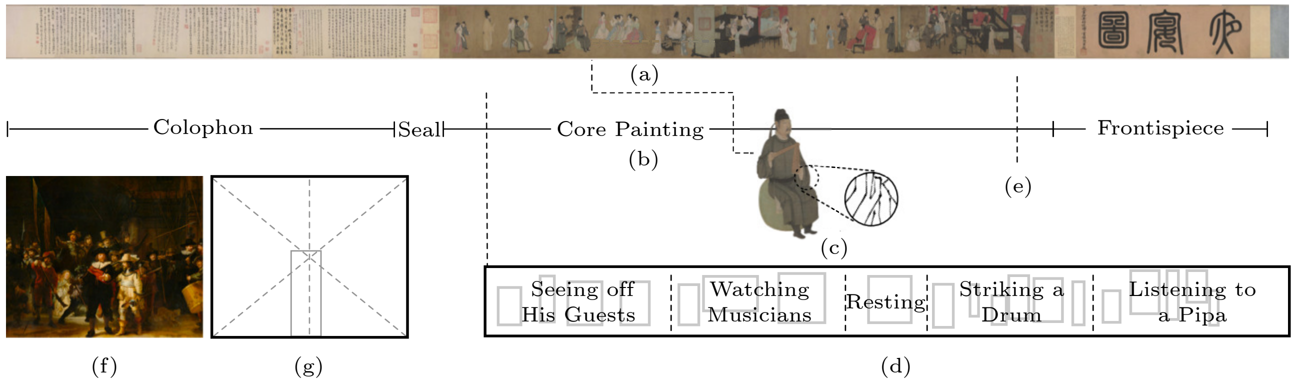


Fig.1. (a) "The Night Revels of Han Xizai" (Gu Hongzhong, 10th-century, Handscroll, 28.7 cm × 335.5 cm)[2] mainly depicts the scene of a night banquet hosted by Han Xizai, the prime minister of the Tang Dynasty. (b) Sections of a painting. (c) Iron wire. (d) Five distinct parts. (e) Ink and color on silk. (f) "The Night Watch" (Rembrandt van Rijn, 1 642 363 cm × 437 cm)[3] depicts the scene of a night patrol by the civic militia. (g) Focus perspective. Both paintings, (a) and (f), depict group portraits in terms of their subject matter, depicting dozens of characters. However, there are significant differences in terms of the materials, format, and techniques used in the paintings.

these deep learning techniques to TCP requires a comprehensive understanding of their characteristics. For instance, although AI-generated TCP may resemble the originals when viewed from afar, the brushstroke details can be far from natural. Moreover, the visual objects are often mislocated about the TCP composition style. Oil paintings have similar challenges and have been studied about their specific characteristics, such as stroke composition[13–18]. Nevertheless, such analyses and conclusions cannot be directly applied to TCP, since they differ greatly in terms of format, techniques, and material (see Table 1). They have different requirements for ink diffusion effects, point-of-view considerations, and painting techniques to depict textures.

Various papers on computer vision techniques have been conducted mainly from the perspectives of aesthetic judgment and stylization. DiVerdi[19] investigated the modular framework for digital paintings and loosely addressed TCP by calligraphy. Zhang *et al.*[20] surveyed systems of photographs and paintings from the perspective of aesthetic evaluation. Kyprianidis *et al.*[21] reviewed methods for transforming photos into aesthetically stylized renderings. Li *et al.*[22] conducted a review of the computer methods used during various stages of production and preservation of Chinese cultural heritage. TCP, however, is only viewed as a common format of image data in these papers, which overlooks the peculiarities of Chinese paintings' data attributes and analysis tasks.

To bridge the gap between general image data and TCP, we conduct a systematic review of the literature on key areas including data visualization (VIS), computer vision (CV), computer graphics (CG), and human-computer interaction (HCI). We draw inspiration from the TCP appreciation theory and adapt the "Six Principles of Painting" to modern concepts for categorizing the collected literature (Section 3). In collaboration with Chinese painting specialists, we propose a four-stage framework to review the purposes of using computational techniques in TCP (Section 4). We then classify the recent computer-aided techniques from the perspectives of task, feature, and rendering (Section 5). Lastly, we report the discussions with specialists about the current drawbacks and potential future applications of computer technol-

Table 1. Differences Between Traditional Chinese Painting and Oil Painting

Aspect	Traditional Chinese Painting	Oil Painting
Format	The point of views can come from different scenes (Fig.1(d)); contain several sections (Fig.1(b))	Depict the scene from a focal perspective (Fig.1(g))
Technique	Objects are depicted by different brushwork systems, e.g., wrinkling techniques for landscapes and calligraphic-line techniques for figures. (Fig.1(c))	Focus on realism, using light and dark tones to express the texture of objects (Fig.1(f))
Material	Painting black ink and Chinese pigments (similar to gouache paint) on paper and silk (Fig.1(e))	Applying the mixture of pigments and drying oils with diverse plasticity on wood and canvas (Fig.1(f))

ogy to TCP (Section 6). We believe this paper can offer an explanation, insights, and examples into every aspect of TCP, thus making way for a more comprehensive appreciation. An interactive browser of this paper is available online①.

2 Methodology

This section describes the paper collection methods, the expert collaboration process, and the procedures for coding articles.

2.1 Methods and Corpus

We construct a literature corpus based on keyword- and relation-search methods. We focus on TCP-related keywords (e.g., “Chinese landscape painting”, “Chinese ink wash painting”, and “Chinese brush painting”), resulting in an initial 88 papers. To expand our literature corpus, we further use the relation-search method. We identify eight influential papers from the initial corpus and exhaustively traverse their references and citations, which expands the corpus to 112 papers. We assess the corpus based on relevance, focusing on publications that probe methodologies and applications. Papers on theory[23] and evaluation[24] are excluded. Despite extensive research on brushes, we only include those relevant to TCP, discarding the calligraphy-related articles[25, 26]. Finally, the corpus comprises 94 papers.

2.2 Collaboration with Domain Experts

Over the past year, we have been working closely with two experts to strengthen our understanding of the specialized and unique domain of TCP. The experts include a professor with over 20 years of expertise in TCP, and a doctoral candidate in TCP theory with five years of experience. Our collaboration consists of the following stages. Firstly, we consult experts about the domain knowledge of TCP and the domain interpretation of some exemplar papers. Secondly, we iteratively refine the analysis framework for applying computer technology to TCP. Finally, we explore research challenges and opportunities by discussing the findings and proposing future research projects.

2.3 Coding and Classification

Through iterative discussions with experts, we analyze the corpus from three perspectives: research scope in TCP (Section 3), specifically-targeted problem (Section 4) and the use of computer-based methods (Section 5). To differentiate the TCP research from image data analysis properly, the TCP characteristics should be considered. We notice that the “Six Principles of Painting” has summarized the most important considerations for drawing and appreciating TCP. We use it as the coding scheme for categorizing the literature with different targeted problems, i.e., the concerned artistic elements. In addition, we evaluate the current analysis of TCP and conclude with a paradigm that outlines the components and purposes of TCP analysis that are supported by modern computer techniques. Specifically, the paradigm includes the digitalization, interpretation, creation, and exhibition of TCP. Lastly, we code the papers in the corpus according to the types of computational techniques they utilize rather than the concrete algorithms that are largely interchangeable.

During the paper analysis, three authors independently code 94 papers over four weeks②. The classification criteria are refined during the coding process. In cases where there are disputes regarding categorization, all authors are involved in debates to reach a consensus. For instance, we first restrict the classification of the papers using the “Resonance of the Spirit” in the Six Principles of Painting to the subjective evaluation of static Chinese painting images. After deliberation, we decide this idea should be expanded to include animations, as this is deemed a more expressive art form that could be evaluated by the principle of “Resonance of the Spirit”.

3 Six Principles of Painting

The Six Principles of Painting[27] were proposed in the sixth century to serve as the grading standards of TCP. Xie proposed and used these principles to rate previous paintings in the book *Gu Hua Pin Lu* (Notes on the Criticism of Old Paintings)[27]. They have remained influential today and shaped how TCP is drawn and appreciated. We select the translations collected in [28] to clarify these principles in the following subsections of this section.

①https://ca4tcp.com, Mar. 2024.
②https://ca4tcp.com/overview.html, Mar. 2024.

3.1 Resonance of the Spirit, Movement of Life

The first principle emphasizes the delivery of vitality in the objects and emotions. This principle is considered the most important, upon which the remaining principles were developed[27]. While the resonated spirits and lively movements are difficult to evaluate objectively, we adopt their semantic meaning and define them as techniques that aim to induce emotional arousal and engage audiences.

Papers within this category analyze the conveyed emotions and use this information to recreate captivating paintings on various devices. For example, Zheng *et al.*[29] utilized machine learning methods to extract relevant features and classify TCP with the conveyed emotions. To enhance the emotional expressions, other work[30–32] attempts to make the objects in the paintings "alive" (e.g., animations). Consequently, the paintings can be interpreted more vividly. In recent years, a significant amount of efforts[33–37] have utilized mixed reality technology to display TCP, giving viewers new perspectives on appreciating the paintings.

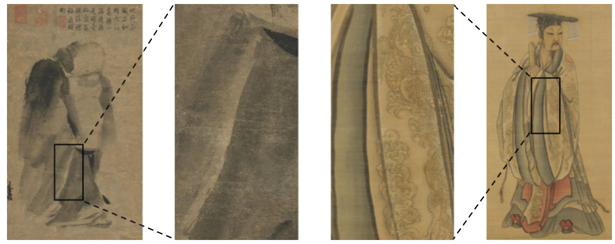
3.2 Bone Manner, Structural Use of Brush

The bone method corresponds to the use of the brush. Calligraphy and paintings highly influence one another, with each brush stroke having its own structure, texture, and meaning. We define this principle as rendering techniques that involve brush strokes. There are 36 articles discussing the bone method from the following aspects:

Stroke Extraction. Some articles focus on identifying and replicating the distinctive brushstrokes of TCP[38–48]. Frequent subjects in TCP, including mountains[49], rocks[50], and trees[51], are imitated.

Ink Simulation. TCP is created using brushes dipped in black ink or Chinese pigments, diffusing on rice papers or silk, making the simulation of ink effects a heavily studied area[52–62]. On the other hand, the physical model of the brush itself also deserves in-depth analysis[63–66].

Style Recognition. Xieyi and Gongbi are two representation styles of TCP[67]. Xieyi uses techniques that privilege the spontaneity of the line (Fig.2(a)). Gongbi uses highly detailed brushstrokes that precisely delimit details (Fig.2(b)). These stroke characteristics classify diverse painting styles[70–72].



(a) (b)

Fig.2. (a) Liang Kai's "Immortal in Splashed Ink"[68] exemplifies the Xieyi style, using wet strokes of monochromatic ink to create the immortal's cloth, while (b) Ma Lin's "King Yu of Xia"[69] exemplifies the more elaborate Gongbi style for clothing patterns.

3.3 Conformity with the Objects, Obtaining Their Likeness

Chinese painters develop a distinct way of depicting objects in the world. Compared with their Western counterparts, objects in Chinese paintings are more surreal. This principle is reinterpreted as extracting objects from TCP and depicting and sketching natural objects in the TCP artistic style. Six articles[73–78] study the TCP classification according to the extracted objects, and three articles[79–81] produce text descriptions (i.e., instance-level captions) for the objects in TCP. Additionally, 17 articles conduct the style transfer for natural objects with TCP art styles[8–12, 82–93].

3.4 According to the Species, Applying the Colors

The suitability of a type often appears to judge the correct use of colors. Inks must be applied in multiple layers to achieve the desired tone in the drawing process. Since inks and papers react to the environment and receive damage, ancient paintings have lost their original states and must be adequately conserved. This principle corresponds to color analysis and restoration of paintings. Papers that fall into this category concentrate on understanding the object classes[94–99] and performing further actions (e.g., color enhancement[100–102]). Additionally, understanding the color of TCP is crucial for repairing them. Some papers employ computer technologies to address the issues of deteriorating paper and fading color[6, 7, 103–106].

3.5 Plan and Design, Place and Position

Division and planning refer to the positioning and arrangement of objects respectively. Papers in this category concentrate on analyzing and enhancing the composition. A unique layout characteristic in Chinese painting is the concept of "void" (white space). It is believed that leaving some part of the paper blank could induce more imagination in readers' minds[107, 108]. The seals, preface, and postscript are also significant components of the composition of TCP[109, 110].

3.6 Transmit Models by Drawing

Prior to the invention of printers, the manual replication of paintings was required to facilitate their distribution and use as commodities. Copying the classics and antique masterpieces also helps amateurs improve their skills by closely observing the techniques. Papers in this area have concentrated on digitizing TCP and displaying them on various devices, as well as replicating the process of painting for the purposes of painting practice and education. The primary distinction between this principle and the others is that these methods, such as high relief[5], virtual reality reconstruction[111], and interactive devices[112, 113], focus on precise reproduction and minimal modification from the original copies. Understanding the painting process[114–117] of TCP is particularly useful for painting practice and education. It also provides a practical basis for painting generation.

4 Analytical Framework

In this section, we propose a framework for applying computational techniques in TCP based on the state-of-the-art and the expertise of domain experts. As shown in Fig.3, the framework involves four typical stages, from the "digitalization" and "interpretation" of existing paintings to the "creation" of new artworks. These three stages will also serve the purpose of "exhibition".

4.1 Digitalization

The digitalization stage involves transforming physical raw TCP into digital signals or codes, primarily as digital images. These images could form a large corpus of TCP, comprising hundreds and thousands of artworks that can hardly be accessed physically in one place. Therefore, this stage presents technical requirements for storage, retrieval, and restoration.

Storage requires storing a large number of TCP digital images in the database. Many image-related techniques are developed to store and show paintings smoothly with different levels of detail. However, there is little research targeting the storage of TCP digital images, which incorporates specific query and analytical requirements. For these large databases, information retrieval efficiently searches for interesting series of paintings from different dimensions. In addition to the metadata of TCP, such as authors and themes, content-based image retrieval utilizes similarities of paintings in terms of visual features[73, 93, 95]. Restoration of TCP deals with pigment fading and paper aging[103] in the digitalization stage. Chen *et al.*[7] simulated the aging and reverse-aging phenomena. Several studies apply image recovery techniques to restore the electronic forms of TCP in terms of stroke and brush[103] and colors[7, 104–106].

4.2 Interpretation

After obtaining the digital formats of the original TCP, the next stage is to interpret them by applying computational approaches. The aspects of interpretation vary from micro-level analysis (e.g., colors and objects[79, 80]), meso-level analysis (e.g., emotion extraction[80]), to macro-level analysis (e.g., the layout of the painting and white space analysis[107]).

In terms of techniques, the majority of the current work falls into two categories: local and global feature extraction. The former emphasizes the identification of relevant features, including handcrafted features (e.g., brushwork[87, 94] and color[88, 97]) and learned features with deep learning technologies (e.g., object detection[76, 84]). The latter focuses on classification's effectiveness, such as accuracy, in which most

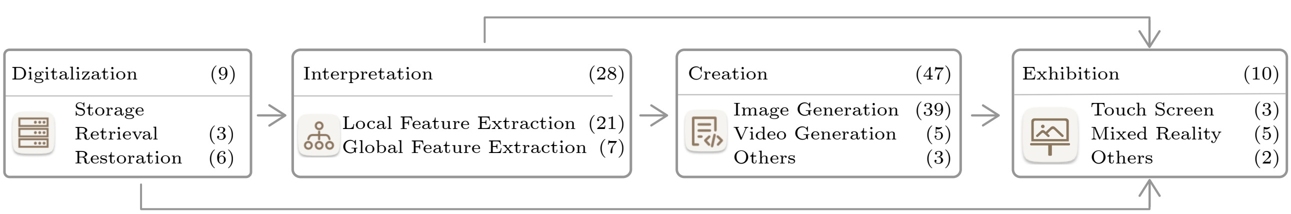


Fig.3. Analytical framework for applying computational techniques in TCP analysis. The numbers represent the paper count of each category and show the distribution of research focus.

work utilizes learned features[74, 84].

4.3 Creation

After obtaining useful features and insights from the analysis of existing paintings, another large proportion of studies focuses on the creation of new artworks. Image generation and video generation are two mainstream creation outcomes.

Image Generation. The majority of studies focus on generating new paintings, considering the unique characteristics in style and the artistic elements of TCP based on the Six Principles of Painting. Style transfer takes existing inputs (e.g., photos and prepared sketches) and outputs generated TCP artworks[9, 11, 12]. Several studies also modify the content of the original input such as face replacement[11]. In addition to generating from existing materials, a few studies have experimented with creating artworks from scratch, including both content and style generation[8].

Video Generation. With the development of computer animation, a line of work has explored creating videos in a TCP style, aiming to express the Spirit Resonance in the Six Principles of Painting from a new perspective. They could be classified into two categories according to the input and animation entities. The first category is to input an existing video and apply video style transfer to transform the whole frame[30, 32]. The second category is to input a TCP and animate entities such as characters and animals on the painting[30, 46, 87]. In addition, 2.5D artworks[5, 92] and video scribing showing the construction of TCP for educational purposes[114] are explored.

4.4 Exhibition

As an essential type of artwork, exhibition is a typical stage for promoting TCP to the general audience. In addition to the traditional approach of arranging items one by one in the museum, studies have been exploring interactive approaches to engage audiences in the exhibition. Existing work could be classified into three categories according to the interactive platforms, namely, touch-screen-based, XR-based, and others.

Touch screens are commonly applied in today's museums. Hsieh *et al.*[113] presented an interactive tabletop for audiences to view detailed regions of TCP. Subramonyam *et al.*[112] developed an iPad application, "Rice Paper", for artists to highlight and annotate key information of the TCP for the general public. They also printed a tangible booklet based on this application to guide audiences. CalliPaint[89] is a system that allows audiences or artists to conveniently create TCP-based digital artworks.

With the development of immersive devices, XR (Mixed Reality) has become a new creation platform for curators and artists. Several studies[34, 35, 111] reconstruct TCP in the VR (Virtual Reality) environment with 3D or 2.5D characters and objects. They are intended to provide an immersive experience that the static TCP cannot fulfill. Jin *et al.*[33] evaluated the engagement of audiences when showing TCP on the touch screen and the VR platform.

In addition to the touch screen and XR, other interactive installations are also studied, such as using sensors to capture audiences' walking in the 3D space to generate Chinese Shanshui Paintings[83] and applying a real-time projector-camera system for audiences to interact with TCP[37]. Sound is also important supplementary information for the exhibition, as it can help users better understand the content of the painting. Ma *et al.*[36] embedded the audio explanation into the local area of the Chinese painting, thus enabling the user to move the focus to get an explanation of the corresponding area while enjoying the painting.

5 Computational Techniques

In this section, we will discuss computational techniques applied to TCP. Although TCP and natural images are similar in their modality to pictures, they differ in terms of techniques used for interpretation and creation. We organize and elaborate on the techniques from three perspectives: tasks for which computational techniques are used (Subsection 5.1), extracted features which the models use for these tasks (Subsection 5.2), and rendering techniques in which the model generates new paintings (Subsection 5.3). These categories of each aspect are listed in Fig.4.

5.1 Tasks

Previous work on TCP mainly focuses on tasks that resemble those in computer vision. Nevertheless, considering that TCP has distinct characteristics (as presented in Section 3) compared with natural images and videos, handling these tasks requires more

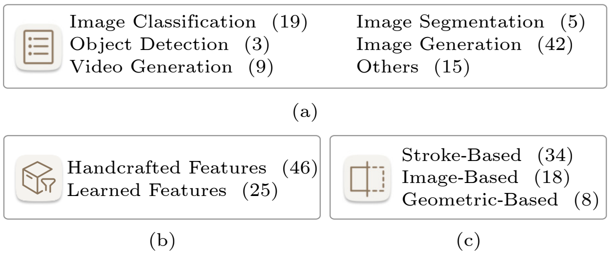


Fig.4. Categories of computational techniques applied to TCP with paper counts. (a) Task. (b) Features. (c) Rendering.

TCP-specific designs and contributions.

Image Classification. TCP can be classified according to multiple attributes (e.g., artists, painting techniques, and painting subjects). Annotating TCP with attributes can improve the retrieval experience and help understand the painting. Distinguishing TCP typically requires expert knowledge, which is time-consuming and expensive. Therefore, it is necessary to train automatic models for accurate TCP classification.

Many studies[39–42, 44, 97] classify TCP according to the artists in that the painting styles of different artists tend to be distinct. Specifically, for describing artists' painting styles, Li and Wang[40] adopted a mixture of multiresolution hidden Markov models, and Liu and Jiang[97] adopted various algorithms, such as Bayes, FLD, and SVM classifiers. Sun *et al.*[42] proposed artistic descriptors with Monte Carlo Convex Hull for feature selection and used SVM for classification. Previous methods typically utilize traditional image processing techniques for classification. In contrast, [44], [42], and [39] utilize MLPs or CNNs for distinguishing artists' styles.

TCP has two mainstream painting techniques, Gongbi (a meticulous style, focusing on details) and Xieyi (an ideographic style, expressing artists' feelings). Jiang and Wang *et al.*[70] applied discrete cosine transformation and CNNs for classifying Gongbi and Xieyi paintings, achieving promising performance. Some other work focuses on classifying landscape paintings, bird-and-flower paintings, and human/figure paintings. Meng *et al.*[96] applied a modified VGG network for painting subject classification, achieving 93.8% accuracy. Since TCP is closely related to calligraphy, [110] distinguishes TCP from calligraphy according to the Chinese characters' structures and the differences in image composition. Li and Zhang[75] proposed an LSTM-based model to classify TCP into five categories: ancient trees, people, flowers-and-birds, Jiangnan water-bound town, and ink paintings. However, these categories overlap with each other in the TCP concept, which inevitably limits the model's generalization ability.

Image Segmentation. Image segmentation was studied in TCP with traditional morphological methods, yet recent neural network-based methods have not been explored. There are two reasons: 1) it is hard to collect large-scale training datasets of TCP, which requires domain knowledge for annotation; 2) the object boundaries of TCP (specifically a key category, Xieyi painting) are hard to determine, as shown in Fig.2(a). Despite these difficulties, Hu and Wang[100] tried to extract the foreground objects from a human-designed saliency map, which has a smaller dependency on the scale of data. Some work[5–7] decomposes the painting into multiple layers to obtain foreground objects or stroke segmentation. On the other hand, the frontispiece and colophon are vital components of TCP (as shown in Fig.1(e)), thus Bao *et al.*[109] proposed a rule-based method to extract these scripts.

Object Detection. Directly adopting natural image-tailored deep learning models for detecting objects (e.g., figure, plant, flower) in TCP tends to have poor performance[76]. [84] utilizes modified YOLOv3 and RetinaNet, and [76] proposes a modified RPN by assembling low-level visual information and high-level semantic information. Apart from the categories that also appear in natural images, a traditional Chinese painting may contain many seals that identify the owners and collectors in a long history, and automatically detecting these seals can greatly help understand the artwork[98].

Image Generation. TCP has its styles (e.g., ink wash painting, white space) compared with other painting types, such as oil painting. Early work tries to transfer a natural image into ink wash paintings by adjusting colors and textures[6, 48, 88, 90] based on tuned hyper-parameters. These early methods are learning-free, thus typically requiring tuning hyper-parameters for each image. Recent studies[8–12] have made efforts to create TCP with generative adversarial networks that transfer noises or natural images into paintings by adversarial training. Some other studies[85] perform style transfer methods with CNNs by separately learning semantic information and styles from two source images and generating a blending image. These methods are machine learning models, requiring several training samples for learning millions of parameters.

Apart from regarding the painting as a whole to generate, another group of work considers that Chinese paintings employ brush strokes and ink to depict objects on paper or silk. Specifically, some studies[50, 56, 62, 64, 66, 116] model either the brush or various stroke shapes, pursuing better texture simulation of real brush strokes. With the specifically modeled brushes, users can draw Chinese paintings stroke by stroke on the screen, instead of drawing on paper with a real brush[71, 83, 89]. Considering the characteristics of rice paper and silk, a large number of early work[54, 57, 58, 60, 91, 93, 101] models the ink diffusion on the rice paper and silk, seeking to improve the realism of paintings. In addition, previous methods focus on creating digital Chinese paintings. Yao and Shao[117] built a painting robot to handle the brushes and drew real paintings by simulating human actions.

Video Generation. We divide the work on TCP video generation into three classes according to their targets: 1) displaying the painting process, 2) animating objects, and 3) natural video style transfer. For the first target, a few studies[114, 115] focus on the creation of Chinese painting, proposing to display the painting process of brush strokes for TCP. In this way, brush trajectory can be animated for both education and appreciation purposes. For the second target, some other studies[31, 32, 46, 87] present methods to animate figures, flowers, and water for a vivid representation of elements in Chinese paintings. Zhao and Ma[35] built a visualization system to build 2.5-dimensional stories about Chinese poetry, displayed by 360-degree videos, which is expected to provide an immersive appreciation of poetry in Chinese painting styles. For the third target, Liang *et al.*[30] displayed a deep learning based multi-frame fusion framework to stylize natural videos with ink wash styles. In the process of transferring, object coherence between adjacent frames is specifically considered for semantic consistency.

Others. Apart from the discussed tasks in CV and HCI above, there are various tasks involving TCP, such as color recovery[106], poet generation from TCP[79, 80], Chinese painting retrieval[73, 78, 95], white space understanding[107], and digital image enhancement[7, 103–105].

5.2 Feature Extraction

There are abundant features in TCP that distinguish them from many other painting genres. To help users analyze and learn from TCP, many researchers extracted features for downstream tasks, such as painting classification and creation. We summarize the features of TCP into two categories, handcrafted features, and learned features, according to the methodology of feature extraction.

5.2.1 Handcrafted Feature

The handcrafted features are extracted by rule-based methods and reflect the specific aspects of TCP.

Brushwork is an important feature in depicting the bone method of the paintings. Typically, the TCP is created with brushes dipped in ink, and the ink permeates through the rice paper, creating the unique shape of the brush strokes. To automatically generate the TCP, a wide range of studies[52, 63, 87, 90, 92, 94, 101, 117] focus on simulating the diffusion effect of color ink. Wang and Wang[93] proposed a physically-based model with a texture synthesis method to simulate the color ink diffusion. Chu and Tai[54] introduced a fluid flow model to calculate the percolation in the paper. In addition, the brushwork is related to the visual complexity of the paintings. Dense thin strokes can increase the complexity while sparse thick strokes lower the complexity. Fan *et al.*[108] measured stroke thickness based on the calculation of color change. Combining the analysis of stroke structures with the ink dispersion densities and placement densities, Lai *et al.*[87] generated animations for water flow in the TCP according to the stroke pattern groups of the flow field.

Color is another significant factor and implies the types of the TCP style[7, 77, 80, 88, 93, 97]. Liu *et al.*[97] extracted the color information of the paintings by calculating the mean and variance values of the image pixels and used them to support painting classification tasks. Color can also be used in painting retrieval[95], painting style modeling[80], and painting enhancement[7]. Over time, ancient Chinese paintings have faded and aged, requiring human restoration. Pei *et al.*[104, 105] designed color enhancement schemes to improve the image contrast, making the paintings more vivid and bright.

Objects, such as the scenery in the paintings, are the basic elements of the painting composition and contain semantic information. Zhang *et al.*[78] extracted objects by labeling pixels according to their connectivity in a pre-processed image and using them for image retrieval. Feng *et al.*[80] extracted the objects in the TCP and used them to describe the painting content and create the painting poetry. Zhao and Ma[35] built a TCP-style image repository for basic objects, supporting users to create immersive videos for poetry appreciation.

Scripts are written in the empty space of the paintings and serve as complementary expressions of the creators' artistic ideas. Bao *et al.*[109] automatically identified and extracted the scripts from the paintings according to their colors and regions. Several studies also focus on other features of the TCP, such as the white space[108], composition[41], and seal images[98].

5.2.2 Learned Feature

With the fast development of deep learning technology, many studies have introduced deep learning models to learn the features of TCP. Based on the labeled data of TCP, supervised learning methods, and YOLOv3 are applied in object detection[76, 80, 84] and image classification[74, 84, 96]. As the stylistic features of the TCP are unique from other paintings, it is valuable to learn the stylistic features to transfer neural images into TCP. A range of studies[8, 10, 12, 30] focus on capturing the stylistic features of TCP with adversarial training, a classical learning strategy in unsupervised learning. In contrast, Li and Zhang[75] introduced weakly-supervised learning for semantic classification in the scenario with a limited number of training images.

5.3 Rendering

TCP rendering techniques are adopted in the process of image and animation generation. According to the focus on used techniques in rendering, we classify the TCP rendering methods into three classes: stroke-based, image-based, and geometry-based.

5.3.1 Stroke-Based Rendering

Users can create TCP through simulated paint brushes to be rendered on a digital canvas. Strassmann and Hairy[62] proposed a realistic model of painting including Brush (a series of bristles with ink supply and positions), Stroke (a set of parameters like position and pressure), Dip (a procedure to assign states to each bristle of the brush), and Paper (the carrier of ink as it comes off the brush). With the four elements, one can build an interactive or automatic painting software on the computer.

Typically, there are two types of methods to generate brush strokes. The first method models the stroke boundaries with Bézier or B-spline curves and then fills the closed curves with designed textures. The other method directly models the two-dimensional brush, such as the work of Strassmann and Hairy[62]. However, the brush bristles are visually fixed in shape, and thus users cannot apply such an e-brush with their realistic painting skills. Some studies[65, 66] develop "soft" brushes in which the shape of the bristle bundle varies in response to the forces given by users. Furthermore, Xu *et al.*[64] modeled brushes with writing primitives (a bundle of hair bristles), instead of each single brush bristle, to improve the simulating realism. To further simplify the model complexity, Bai *et al.*[63] proposed a geometry model to simulate the entire brushes, instead of a large amount of bristles. A dynamic model was also introduced to simulate the brush deformation under internal and external forces. Previous methods focus on modeling general brush strokes, while some methods propose tailored algorithms to model specific object shapes and textures such as rocks[50], trees[49], bamboos[117], and water[32]. Instead of small brush strokes, Fu *et al.*[5] decomposed the painting into image layers with each layer representing a class of specific strokes. With these stroke layers, they can create a new high relief, an art form between 3D sculpture and 2D painting.

For animation generation, Xu *et al.*[46] built a brush stroke library obtained from painting experts and animated the paintings by decomposing them into brush strokes and changing these strokes. Zhang *et al.*[32] created running water animations with a novel proposed painting structure generation method, which is used to estimate water flow line positions. Previous methods create brush trajectories relying on manual inputs. Yang and Xu[115] automated this by modeling the brush footprint from paintings. In the automation process of extracting brush trajectory, some methods[114] reconstruct the drawing process by estimating and animating brush stroke order.

Apart from the brush strokes, ink diffusion in paper fibers structure is also a critical characteristic, which has been studied in literature[51, 54, 57, 59–61, 91, 93, 101]. Specifically, Kunii *et al.*[61] proposed a multidimensional diffusion model to simulate the ink density distribution as in real paper. Some studies[53, 58, 60, 93] further simulate the ink of brush strokes on various types of paper based on physical-based models. Considering the potential blending of multiple strokes, Yeh *et al.*[65] and Yu *et al.*[59] built the ink diffusion simulation model with multi-layered structures of brush and paper, respectively. Chu and Tai[55] created paintings with intricate ink diffusion effects using a lattice Boltzmann equation and accelerated the algorithm in real-time with both CPU and GPU.

5.3.2 *Image-Based Rendering*

Previous methods mainly focus on modeling the brush strokes and ink diffusion for interactive painting creation. From another technical route, one can directly synthesize Chinese painting from existing images. For instance, Yu *et al.*[48] proposed a framework for image-based painting synthesis. Specifically, the authors built a brush stroke texture primitive collection and mapped those texture primitives to a constructed mask image. Apart from blending strokes, some studies propose style transfer methods to transform natural pictures into paintings, involving handcrafted feature-based flow[88, 90, 91] or deep neural networks. Typically, style transfer models that utilize deep neural networks are data-driven[8–12, 85], characterized by training the model with tailored training program and large-scale datasets. For instance, ChipGAN[10] consists of a generator and a discriminator, where the generator is trained to transfer photos into paintings while the discriminator is trained to discriminate the generated paintings and real paintings. Meanwhile, ChipGAN requires thousands of images for training the model due to the large-scale trainable parameters.

For animating Chinese paintings, Liu *et al.*[31] proposed a sample point processing method to preserve the style of brush strokes and determine control bones, and a skeleton-based deformation method for animation generation. Liang *et al.*[30] leveraged deep neural networks for transferring natural videos into ink wash painting-style videos. To enhance temporal consistency between video frames, the authors introduced multi-frame fusion and implemented instance-aware style transfer, which helps generate paintings with proper white space.

5.3.3 *Geometry-Based Rendering*

Stroke-based and image-based rendering are both from the view of computer vision. Instead, geometry-based methods adopt the view of computer graphics. Chan *et al.*[38] decomposed the brush stroke like features into layers of procedural shaders and then mixed different layers to construct desired effects in 3D models. Some studies[47, 49, 52] synthesize objects with ink painting styles by building polygonal models or extracting silhouettes, and then mapping specific textures on the models. Amati and Brostow[92] developed a webcam-based system to capture the process of user drawing plants and build corresponding 2.5-dimensional digital models. Different from previous methods, Shi[86] built landscape paintings from a 3-dimensional city model by generating mountains from buildings and assigning ink painting styles.

6 Challenges and Opportunities

6.1 Lack of Large-Scale and High-Quality Datasets

The lack of large and high-quality open-access datasets is a crucial reason hindering the further development of traditional Chinese painting research. According to our paper, some studies have announced that they have produced a few Chinese painting datasets[8, 12, 73, 74, 95]. For example, Liong *et al.*[74] constructed an unlabeled dataset containing more than 1 000 Chinese paintings, and Dong *et al.*[73] collected a labeled dataset. However, these datasets are limited in size and have not been made open-source. Building a large and high-quality Chinese painting dataset faces several challenges.

Data Availability. As most Chinese paintings are held in museums and private collections all over the world, there is a problem of copyright ownership. It is particularly difficult to collect online resources on Chinese paintings.

Data Quality. Many famous Chinese paintings are large in size, rich in details, and difficult to preserve, leading to the high cost of digitizing Chinese paintings and a high technical barrier for generating high-definition pictures.

Data Diversity. Chinese paintings contain relatively independent items, such as colophons and seals that can be used for analyzing historical events and collection paths. However, only a few articles discuss the extraction of colophons and seals[98, 109] without further exploration.

Data Annotation. Due to the domain professionalism, annotating Chinese painting data requires high-level expertise, especially for systematic annotations based on the Six Principles of Painting, which can be extremely expensive.

6.2 Insufficient Consideration of TCP Uniqueness

Analyzing TCP research tendencies (Section 3), most articles focus on Bone Manner, Structural Use of the Brush (32/94), (and) Conformity with the Objects, Obtaining their Likeness (26/94). This is primarily because stroke and object recognition, segmentation, and classification in paintings align with established computer vision tasks and models. Experts point out that "a significant aspect of the beauty of TCP lies in the beauty of brushwork" emphasizing the importance of a detailed analysis of brushwork. However, researchers have paid little attention to the unique stroke system in traditional Chinese painting (3/94)[49–51]. For example, there are 18 unique drawing methods in TCP techniques for depicting portraits (Fig.1(c)), and different wrinkling techniques for depicting mountains and rocks (Table 1). A detailed stroke system analysis illuminates the painter's style and mentoring relationships between painters, warranting attention in future computer fields.

Moreover, the Place and Position aspect is not given as much attention in the current work (4/94). According to experts, the composition of Chinese paintings is crucial. Painters often use white space to convey mood, and use inscriptions and seals to balance the picture's composition. Therefore, a computer-based systematic examination of the composition of TCP can aid specialists in understanding the compositional traits of paintings across time. Simultaneously, refining the compositional principles of TCP, as exemplified by the "beginning, transition, turning, and synthesis" in vertical landscape paintings, can provide improved guidance for artists in their creative process and enhance viewers' appreciation and understanding of the painting.

Regarding the study of Movement of Life (9/94), it is important to note that in recent years, the fusion of AR and VR technology into TCP has risen, allowing audiences to experience Chinese painting from a new perspective. The development of new technologies has increased the opportunities for the study and presentation of TCP, but more work of this type is required, and it may be reinforced in the future.

6.3 Disregard for Data-Linking in TCP Analysis

Cultural heritage has various types of data, including paintings, ancient books, sculptures, architecture, and more. As a form of cultural heritage, TCP has garnered attention in recent years. However, most current research has focused solely on TCP data, and rarely combines other datasets for cross-analysis. From a historical research perspective, TCP collections represent only a snapshot of a certain period. Snapshots of different artifacts may describe the same social landscape. To gain a more comprehensive historical understanding, different collections of cultural relics should be viewed together. Therefore, it is worthwhile to pay attention to how to integrate different data related to paintings to restore a more accurate historical picture.

On the other hand, the use of multi-modal data to construct deep learning models is a growing trend. Multi-modal data enables better feature representation construction in the latent space, which improves the fusion of textual, visual, and other forms of information like videos and knowledge graphs. This ultimately strengthens the model's performance and enhances its generalization capacity.

6.4 Insufficient Exploration of Machine Learning Methods and Large Models on TCP

TCP image data has unique characteristics compared with natural images, such as cross-domain, few annotated training samples, imbalanced classes, and variable sizes of objects. Advanced ML methods have taken profound discussions on related topics such as transfer learning, domain adaptation, domain generalization, few-shot learning, and learning with long-tailed data distribution. Therefore, applying these advanced methods to TCP can promote Chinese painting analysis from a computational view. Meanwhile, these methods can effectively reduce the demand for data annotations and alleviate the burden of collecting large-scale and high-quality annotated datasets.

In addition, large language models and large vision models, and Stable-Diffusion are becoming the new foundations of advanced research. Current models are not specifically adapted to TCP data, tending to generate images that ignore the Six Principles of Paintings, as well as textual descriptions that often lack detail and do not capture the essence of the painting. Unimodal or multimodal large models will inevitably be adopted in traditional Chinese painting research.

6.5 Inadequate Applications for Artwork Creation and Promotion

Although a line of work has explored the generation of TCP-styled paintings and videos, the quality of these AI-generated artworks could be doubtful. Involving artists in the creation process with a semi-automated creation style would be a promising direction in the future. In addition, advanced display and interaction techniques (e.g., immersive techniques) should be applied to promote TCP to the general public. New storytelling approaches should also be constructed to enhance the appreciation and understanding of TCP.

7 Conclusions

In this paper, we proposed a classification method (the Six Principles of Painting; analytic framework; computational techniques) to investigate the literature on computer applications in TCP. Our approach enables researchers to quickly locate relevant literature from different perspectives, facilitating in-depth research. The research findings indicate several challenges exist in the application of computer technology in the field of TCP. Firstly, existing studies mainly focus on the representation of objects and brushstrokes, while research on emotions and composition is relatively limited, lacking a comprehensive reflection of the unique characteristics of TCP. Secondly, there is a lack of a large-scale, high-quality dataset. Lastly, the utilization of large language models and large vision models in the context of TCP remains insufficient.

In future work, we will further explore these potential research questions. Additionally, with the continuous evolution of computer technology, we will regularly update the literature database to provide insightful perspectives to inspire further research.

Acknowledgements The authors would like to thank the anonymous reviewers for their constructive suggestions. They would also like to thank Prof. Huang-Huang Deng, Dr. Pei-Yi Jiang, Dr. Hang-Chao Gong, and Dr. Ru-Yue Wang from Zhejiang University, and Yi-Tong Zhang from Hangzhou City University for their valuable feedback and contributions to this project.

Conflict of Interest The authors declare that they have no conflict of interest.

References

[1] Cheng M, Tang W H, Choy E. Essential Terms of Chinese Painting. City University of HK Press, 2018.

[2] McCausland S. The night banquet: A Chinese scroll through time by De-Nin D. Lee (review). *Journal of Song-Yuan Studies*, 2012, 42: 486–492. DOI: 10.1353/sys.2013.0008.

[3] de Winter J C F, Dodou D, Tabone W. How do people distribute their attention while observing *The Night Watch*? *Perception*, 2022, 51(11): 763–788. DOI: 10.1177/03010066221122697.

[4] Bradley A J, El-Assady M, Coles K, Alexander E, Chen M, Collins C, Jänicke S, Wrisley D J. Visualization and the digital humanities. *IEEE Computer Graphics and Applications*, 2018, 38(6): 26–38. DOI: 10.1109/MCG.2018.2878900.

[5] Fu Y F, Yu H C, Yeh C K, Zhang J J, Lee T Y. High relief from brush painting. *IEEE Trans. Visualization and Computer Graphics*, 2019, 25(9): 2763–2776. DOI: 10.1109/TVCG.2018.2860004.

[6] Zhang J W, Zhang Y, Zhang S P, Yan L X, Chen J Y. Multispectral image matting of ancient Chinese paintings. In *Eurographics 2011*, Day A M, Mantiuk R, Reinhard E, Scopigno R (eds.), Eurographics Association, 2011, pp.9–15.

[7] Chen L H, Tsai M F, Hsu C H, Chen Y S. Simulating aging and reverse-aging phenomena of traditional Chinese paintings. In *Proc. the 26th Annual Conference of the Japanese Society for Artificial Intelligence*, Jun. 2012. DOI: 10.11517/pjsai.JSAI2012.0_4M1IOS3c5.

[8] Xue A. End-to-end Chinese landscape painting creation using generative adversarial networks. In *Proc. the 2021 IEEE Winter Conference on Applications of Computer Vision*, Jan. 2021, pp.3863–3871. DOI: 10.1109/WACV48630.2021.00391.

[9] Zhang F Q, Gao H M, Lai Y P. Detail-preserving CycleGAN-AdaIN framework for image-to-ink painting translation. *IEEE Access*, 2020, 8: 132002–132011. DOI: 10.1109/ACCESS.2020.3009470.

[10] He B, Gao F, Ma D Q, Shi B X, Duan L Y. ChipGAN: A generative adversarial network for Chinese ink wash painting style transfer. In *Proc. the 26th ACM International Conference on Multimedia*, Oct. 2018, pp.1172–1180. DOI: 10.1145/3240508.3240655.

[11] Li J Y, Wang Q, Li S J, Zhong Q, Zhou Q. Immersive traditional Chinese portrait painting: Research on style transfer and face replacement. In *Proc. the 4th Chinese Conference on Pattern Recognition and Computer Vision*, Oct. 29–Nov. 1, 2021, pp.192–203. DOI: 10.1007/978-3-030-88007-1_16.

[12] Wang R, Huang H B, Zheng A H, He R. Attentional wavelet network for traditional Chinese painting transfer. In *Proc. the 25th International Conference on Pattern Recognition*, Jan. 2021, pp.3077–3083. DOI: 10.1109/ICPR48806.2021.9413063.

[13] Zheng N Y, Jiang Y F, Huang D J. StrokeNet: A neural painting environment. In *Proc. the 7th International*

Conference on Learning Representations, May 2019.

[14] Litwinowicz P. Processing images and video for an impressionist effect. In *Proc. the 24th Annual Conference on Computer Graphics and Interactive Techniques*, Aug. 1997, pp.407–414. DOI: 10.1145/258734.258893.

[15] Huang Z W, Zhou S C, Heng W. Learning to paint with model-based deep reinforcement learning. In *Proc. the 2019 IEEE/CVF International Conference on Computer Vision*, Oct. 27–Nov. 2, 2019, pp.8709–8718. DOI: 10.1109/ICCV.2019.00880.

[16] Zou Z X, Shi T Y, Qiu S, Yuan Y, Shi Z W. Stylized neural painting. In *Proc. the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2021, pp.15689–15698. DOI: 10.1109/CVPR46437.2021.01543.

[17] Kotovenko D, Wright M, Heimbrecht A, Ommer B. Rethinking style transfer: From pixels to parameterized brushstrokes. In *Proc. the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2021, pp.12196–12205. DOI: 10.1109/CVPR46437.2021.01202.

[18] Liu S H, Lin T W, He D L, Li F, Deng R F, Li X, Ding E, Wang H. Paint transformer: Feed forward neural painting with stroke prediction. In *Proc. the 2021 IEEE/CVF International Conference on Computer Vision*, Oct. 2021, pp.6598–6607. DOI: 10.1109/ICCV48922.2021.00653.

[19] DiVerdi S. A modular framework for digital painting. *IEEE Trans. Visualization and Computer Graphics*, 2015, 21(7): 783–793. DOI: 10.1109/TVCG.2015.2403352.

[20] Zhang J J, Miao Y W, Yu J H. A comprehensive survey on computational aesthetic evaluation of visual art images: Metrics and challenges. *IEEE Access*, 2021, 9: 77164–77187. DOI: 10.1109/ACCESS.2021.3083075.

[21] Kyprianidis J E, Collomosse J, Wang T H, Isenberg T. State of the "art": A taxonomy of artistic stylization techniques for images and video. *IEEE Trans. Visualization and Computer Graphics*, 2013, 19(5): 866–885. DOI: 10.1109/TVCG.2012.160.

[22] Li M, Wang Y, Xu Y Q. Computing for Chinese cultural heritage. *Visual Informatics*, 2022, 6(1): 1–13. DOI: 10.1016/j.visinf.2021.12.006.

[23] Wu X M, Li G F, Liang Y. Modeling Chinese painting images based on ontology. In *Proc. the 2013 International Conference on Information Technology and Applications*, Nov. 2013, pp.113–116. DOI: 10.1109/ITA.2013.32.

[24] Bo Y H, Yu J H, Zhang K. Computational aesthetics and applications. *Visual Computing for Industry, Biomedicine, and Art*, 2018, 1(1): 6. DOI: 10.1186/s42492-018-0006-1.

[25] Mi X F, Xu J, Tang M, Dong J X. The droplet virtual brush for Chinese calligraphic character modeling. In *Proc. the 6th IEEE Workshop on Applications of Computer Vision*, Dec. 2002, pp.330–334. DOI: 10.1109/ACV.2002.1182203.

[26] Bai B D, Wong K W, Zhang Y N. An efficient physically-based model for Chinese brush. In *Proc. the 1st International Workshop on Frontiers in Algorithmics*, Aug. 2007, pp.261–270. DOI: 10.1007/978-3-540-73814-5_25.

[27] Lin Y T. The Chinese Theory of Art. G. P. Putnam's Sons, 1967.

[28] van Briessen F. The Way of the Brush: Painting Techniques of China and Japan. C. E. Tuttle Publishing, 1962.

[29] Zheng W, Li H Y, Xu H S, Sun M J. Chinese painting emotion classification based on convolution neural network and SVM. *Journal of Nanjing Normal University (Natural Science Edition)*, 2017, 40(3): 74–79, 86. DOI: 10.3969/j.issn.1001-4616.2017.03.011. (in Chinese)

[30] Liang H, Yang S, Wang W J, Liu J Y. Instance-aware coherent video style transfer for Chinese ink wash painting. In *Proc. the 30th International Joint Conference on Artificial Intelligence*, Aug. 2021, pp.823–829. DOI: 10.24963/ijcai.2021/114.

[31] Liu D S M, Cheng C I, Liu M L. Animating characters in Chinese painting using two-dimensional skeleton-based deformation. *Multimedia Tools and Applications*, 2020, 79(27): 20343–20371. DOI: 10.1007/s11042-020-08842-5.

[32] Zhang S H, Chen T, Zhang Y F, Hu S M, Martin R. Video-based running water animation in Chinese painting style. *Science in China Series F: Information Sciences*, 2009, 52(2): 162–171. DOI: 10.1007/s11432-009-0035-7.

[33] Jin S, Fan M, Kadir A. Immersive *Spring Morning in the Han Palace*: Learning traditional Chinese art via virtual reality and multi-touch tabletop. *International Journal of Human–Computer Interaction*, 2022, 38(3): 213–226. DOI: 10.1080/10447318.2021.1930389.

[34] Jin S, Fan M, Wang Y C, Liu Q. Reconstructing traditional Chinese paintings with immersive virtual reality. In *Proc. the 2020 CHI Conference on Human Factors in Computing Systems*, Apr. 2020, pp.1–8. DOI: 10.1145/3334480.3382934.

[35] Zhao Z J, Ma X J. ShadowPlay2.5D: A 360-degree video authoring tool for immersive appreciation of classical Chinese poetry. *Journal on Computing and Cultural Heritage*, 2020, 13(1): Article No. 5. DOI: 10.1145/3352590.

[36] Ma W, Wang Y Z, Xu Y Q, Li Q, Ma X, Gao W. Annotating traditional Chinese paintings for immersive virtual exhibition. *Journal on Computing and Cultural Heritage*, 2012, 5(2): Article No. 6. DOI: 10.1145/2307723.2307725.

[37] Jin M, Zhang H, Yang X B, Xiao S J. A real-time Pro-Cam system for interaction with Chinese ink-and-wash cartoons. In *Proc. the 2007 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2007, pp.1–2. DOI: 10.1109/CVPR.2007.383480.

[38] Chan C, Akleman E, Chen J. Two methods for creating Chinese painting. In *Proc. the 10th Pacific Conference on Computer Graphics and Applications*, Oct. 2002, pp.403–412. DOI: 10.1109/PCCGA.2002.1167884.

[39] Jiang W, Wang X Y, Ren J C, Li S, Sun M J, Wang Z, Jin J S. MTFFNet: A multi-task feature fusion framework for Chinese painting classification. *Cognitive Computation*, 2021, 13(5): 1287–1296. DOI: 10.1007/s12559-

021-09896-9.

[40] Li J, Wang J Z. Studying digital imagery of ancient paintings by mixtures of stochastic models. *IEEE Trans. Image Processing*, 2004, 13(3): 340–353. DOI: 10.1109/TIP.2003.821349.

[41] Sun M J, Zhang D, Wang Z, Ren J C, Jin J S. Monte Carlo Convex Hull model for classification of traditional Chinese paintings. *Neurocomputing*, 2016, 171: 788–797. DOI: 10.1016/j.neucom.2015.08.013.

[42] Sun M J, Zhang D, Ren J C, Wang Z, Jin J S. Brushstroke based sparse hybrid convolutional neural networks for author classification of Chinese ink-wash paintings. In *Proc. the 2015 IEEE International Conference on Image Processing*, Sept. 2015, pp.626–630. DOI: 10.1109/ICIP.2015.7350874.

[43] Sheng J C. Automatic categorization of traditional Chinese paintings based on wavelet transform. *Computer Science*, 2014, 41(2): 317–319. DOI: 10.3969/j.issn.1002-137X.2014.02.069.

[44] Sheng J C, Jiang J M. Recognition of Chinese artists via windowed and entropy balanced fusion in classification of their authored ink and wash paintings (IWPs). *Pattern Recognition*, 2014, 47(2): 612–622. DOI: 10.1016/j.patcog.2013.08.017.

[45] Sheng J C, Jiang J M. Style-based classification of Chinese ink and wash paintings. *Optical Engineering*, 2013, 52(9): 093101. DOI: 10.1117/1.OE.52.9.093101.

[46] Xu S H, Xu Y Q, Kang S B, Salesin D H, Pan Y H, Shum H Y. Animating Chinese paintings through stroke-based decomposition. *ACM Trans. Graphics*, 2006, 25(2): 239–267. DOI: 10.1145/1138450.1138454.

[47] Yeh J W, Ouhyoung M. Non-photorealistic rendering in Chinese painting of animals. *Journal of System Simulation*, 2002, 14(6): 1220–1224.

[48] Yu J H, Luo G M, Peng Q S. Image-based synthesis of Chinese landscape painting. *Journal of Computer Science and Technology*, 2003, 18(1): 22–28. DOI: 10.1007/BF02946647.

[49] Way D L, Lin Y R, Shih Z C. The synthesis of trees in Chinese landscape painting using silhouette and texture strokes. *Journal of WSCG*, 2002, 10: 499–506.

[50] Way D L, Shih Z C. The synthesis of rock textures in Chinese landscape painting. *Computer Graphics Forum*, 2001, 20(3): 123–131. DOI: 10.1111/1467-8659.00505.

[51] Zhang Q, Sato Y, Takahashi J Y, Muraoka K, Chiba N. Simple cellular automaton-based simulation of ink behaviour and its application to *Suibokuga*-like 3D rendering of trees. *The Journal of Visualization and Computer Animation*, 1999, 10(1): 27–37. DOI: 10.1002/(SICI)1099-1778(199901/03)10:1<27::AID-VIS194>3.0.CO;2-C.

[52] Xu T C, Yang L J, Wu E H. Stroke-based real-time ink wash painting style rendering for geometric models. In *Proc. the SIGGRAPH Asia 2012 Technical Briefs*, Nov. 2012, Article No. 19. DOI: 10.1145/2407746.2407765.

[53] Way D L, Lin W J, Shih Z C. Computer-generated Chinese color ink paintings. *Journal of the Chinese Institute of Engineers*, 2006, 29(6): 1041–1050. DOI: 10.1080/02533839.2006.9671203.

[54] Chu N S H, Tai C L. MoXi: Real-time ink dispersion in absorbent paper. *ACM Trans. Graphics*, 2005, 24(3): 504–511. DOI: 10.1145/1073204.1073221.

[55] Chu N S H, Tai C L. Real-time painting with an expressive virtual Chinese brush. *IEEE Computer Graphics and Applications*, 2004, 24(5): 76–85. DOI: 10.1109/MCG.2004.37.

[56] Mi X F, Tang M, Dong J X. Droplet: A virtual brush model to simulate Chinese calligraphy and painting. *Journal of Computer Science and Technology*, 2004, 19(3): 393–404. DOI: 10.1007/BF02944909.

[57] Guo Q, Kunii T L. "Nijimi" rendering algorithm for creating quality black ink paintings. In *Proc. the 2003 Computer Graphics International*, Jul. 2003, pp.152–159. DOI: 10.1109/CGI.2003.1214460.

[58] Way D L, Huang S W, Shih Z C. Physical-based model of ink diffusion in Chinese paintings. *Journal of WSCG*, 2003, 11(1/2/3).

[59] Yu Y J, Lee Y B, Cho H G, Lee D H. A model based technique for realistic oriental painting. In *Proc. the 10th Pacific Conference on Computer Graphics and Applications*, Oct. 2002, pp.452–453. DOI: 10.1109/PCCGA.2002.1167898.

[60] Lee J. Diffusion rendering of black ink paintings using new paper and ink models. *Computers & Graphics*, 2001, 25(2): 295–308. DOI: 10.1016/S0097-8493(00)00132-1.

[61] Kunii T L, Nosovskij G V, Hayashi T. A diffusion model for computer animation of diffuse ink painting. In *Proc. the 1995 Computer Animation*, Apr. 1995, pp.98–102. DOI: 10.1109/CA.1995.393542.

[62] Strassmann S. Hairy brushes. *ACM SIGGRAPH Computer Graphics*, 1986, 20(4): 225–232. DOI: 10.1145/15886.15911.

[63] Bai B D, Zhang Y N, Wong K W, Li Y. Chinese hairy brush: A physically-based model for calligraphy. *Chinese Journal of Electronics*, 2009, 18(2): 302–306.

[64] Xu S H, Lau C M F, Xu C F, Pan Y H. Virtual hairy brush for digital painting and calligraphy. *Science in China Series F: Information Sciences*, 2005, 48(3): 285–303. DOI: 10.1360/03yf0389.

[65] Yeh J S, Lien T Y, Ouhyoung M. On the effects of haptic display in brush and ink simulation for Chinese painting and calligraphy. In *Proc. the 10th Pacific Conference on Computer Graphics and Applications*, Oct. 2002, pp.439–441. DOI: 10.1109/PCCGA.2002.1167892.

[66] Lee J. Simulating oriental black-ink painting. *IEEE Computer Graphics and Applications*, 1999, 19(3): 74–81. DOI: 10.1109/38.761553.

[67] March B. Some Technical Terms of Chinese Painting. Waverly Press, 1935.

[68] Xie K. The form of no-form: Reconstructing Huineng in two paintings by Liang Kai [Ph. D. Thesis]. University of Oregon, 2017.

[69] Fong W C. Beyond Representation: Chinese Painting and Calligraphy, 8th-14th Century. Yale University Press, 1992.

[70] Jiang W, Wang Z, Jin J S, Han Y H, Sun M J. DCT-CNN-based classification method for the Gongbi and

Xieyi techniques of Chinese ink-wash paintings. *Neurocomputing*, 2019, 330: 280–286. DOI: 10.1016/j.neucom.2018.11.003.

[71] Yang L J, Xu T C, Du J X, Wu E H. Easy drawing: Generation of artistic Chinese flower painting by stroke-based stylization. *IEEE Access*, 2019, 7: 35449–35456. DOI: 10.1109/ACCESS.2019.2904669.

[72] Jiang S Q, Huang T J. Categorizing traditional Chinese painting images. In *Proc. the 5th Pacific-Rim Conference on Multimedia*, Nov. 30–Dec. 3, 2004, pp.1–8. DOI: 10.1007/978-3-540-30541-5_1.

[73] Dong Z H, Wan J, Li C Y, Jiang H, Qian Y G, Pan W X. Feature fusion based cross-modal retrieval for traditional Chinese painting. In *Proc. the 2020 International Conference on Culture-oriented Science & Technology*, Oct. 2020, pp.383–387. DOI: 10.1109/ICCST50977.2020.00080.

[74] Liong S T, Huang Y C, Li S M, Huang Z K, Ma J Y, Gan Y S. Automatic traditional Chinese painting classification: A benchmarking analysis. *Computational Intelligence*, 2020, 36(3): 1183–1199. DOI: 10.1111/coin.12328.

[75] Li D X, Zhang Y. Multi-instance learning algorithm based on LSTM for Chinese painting image classification. *IEEE Access*, 2020, 8: 179336–179345. DOI: 10.1109/ACCESS.2020.3027982.

[76] Gu Q Q, King R. Deep learning does not generalize well to recognizing cats and dogs in Chinese paintings. In *Proc. the 22nd International Conference on Discovery Science*, Oct. 2019, pp.166–175. DOI: 10.1007/978-3-030-33778-0_14.

[77] Lu G M, Gao Z, Qin D N, Zhao X, Liu M J. Content-based identifying and classifying traditional Chinese painting images. In *Proc. the 2008 Congress on Image and Signal Processing*, May 2008, pp.570–574. DOI: 10.1109/CISP.2008.477.

[78] Zhang D, Pham B, Li Y F. Modelling traditional Chinese paintings for content-based image classification and retrieval. In *Proc. the 10th International Multimedia Modelling Conference*, Jan. 2004, pp.258–264. DOI: 10.1109/MULMM.2004.1264994.

[79] Chen J Z, Huang K Y, Feng Y C J, Zhang W, Tan S W, Chen W. Automatic poetry generation based on ancient Chinese paintings. *Journal of Computer-Aided Design & Computer Graphics*, 2021, 33(7): 1038–1044. DOI: 10.3724/SP.J.1089.2021.18633. (in Chinese)

[80] Feng Y C J, Chen J Z, Huang K Y, Wong J K, Ye H, Zhang W, Zhu R C, Luo X N, Chen W. iPoet: Interactive painting poetry creation with visual multimodal analysis. *Journal of Visualization*, 2022, 25(3): 671–685. DOI: 10.1007/s12650-021-00780-0.

[81] Yuan S Z, Dai A J, Yan Z L, Liu R X, Chen M, Chen B Y, Qiu Z J, He X D. Learning to generate poetic Chinese landscape painting with calligraphy. arXiv: 2305.04719, 2023. https://arxiv.org/abs/2305.04719, Mar. 2024.

[82] Luo P X, Zhang J C, Zhou J. High-resolution and arbitrary-sized Chinese landscape painting creation based on generative adversarial networks. In *Proc. the 31st International Joint Conference on Artificial Intelligence*, Jul. 2022, pp.5015–5018.

[83] Zhou A L, Zhang K. Shanshui journey: AI reproducing the experience of Chinese "literati" ink paintings. *Leonardo*, 2023: 1–22. DOI: 10.1162/leon_a_02474.

[84] Meng Q Y, Li K Y, Zhou M Q, Zhang H H. The elements extraction on traditional Chinese paintings based on object detection. In *Proc. the 2nd Artificial Intelligence and Cloud Computing Conference*, Dec. 2019, pp.111–116. DOI: 10.1145/3375959.3375980.

[85] Wu B, Dong Q S. Research on the synthetic method of ink painting based on convolutional neural network. In *Proc. the 10th International Conference on Digital Image Processing*, Aug. 2018, pp.647–659. DOI: 10.1117/12.2503115.

[86] Shi W L. A generative approach to Chinese Shanshui painting. *IEEE Computer Graphics and Applications*, 2017, 37(1): 15–19. DOI: 10.1109/MCG.2017.13.

[87] Lai Y C, Chen B A, Chen K W, Si W L, Yao C Y, Zhang E. Data-driven NPR illustrations of natural flows in Chinese painting. *IEEE Trans. Visualization and Computer Graphics*, 2017, 23(12): 2535–2549. DOI: 10.1109/TVCG.2016.2622269.

[88] Guo F, Peng H, Tang J. A novel method of converting photograph into Chinese ink painting. *IEEJ Trans. Electrical and Electronic Engineering*, 2015, 10(3): 320–329. DOI: 10.1002/tee.22088.

[89] Li J J, Ngai G, Chan S C F, Hua K A, Leong H V, Chan A. From writing to painting: A Kinect-based cross-modal Chinese painting generation system. In *Proc. the 22nd ACM International Conference on Multimedia*, Nov. 2014, pp.57–66. DOI: 10.1145/2647868.2654911.

[90] Dong L X, Lu S F, Jin X G. Real-time image-based Chinese ink painting rendering. *Multimedia tools and Applications*, 2014, 69(3): 605–620. DOI: 10.1007/s11042-012-1126-9.

[91] Liang L Y, Jin L W. Image-based rendering for ink painting. In *Proc. the 2013 IEEE International Conference on Systems, Man, and Cybernetics*, Oct. 2013, pp.3950–3954. DOI: 10.1109/SMC.2013.674.

[92] Amati C, Brostow G J. Modeling 2.5 D plants from ink paintings. In *Proc. the 7th Sketch-Based Interfaces and Modeling Symposium*, Jun. 2010, pp.41–48.

[93] Wang C M, Wang R J. Image-based color ink diffusion rendering. *IEEE Trans. Visualization and Computer Graphics*, 2007, 13(2): 235–246. DOI: 10.1109/TVCG.2007.41.

[94] Zhan Y, Gao Y, Xie L Y. Aesthetic feature analysis and classification of Chinese traditional painting. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, 45(12): 2514–2522. DOI: 10.13700/j.bh.1001-5965.2019.0375. (in Chinese)

[95] Hung C C. A study on a content-based image retrieval technique for Chinese paintings. *The Electronic Library*, 2018, 36(1): 172–188. DOI: 10.1108/EL-10-2016-0219.

[96] Meng Q Y, Zhang H H, Zhou M Q, Zhao S F, Zhou P B. The classification of traditional Chinese painting based on CNN. In *Proc. the 4th International Conference on*

Cloud Computing and Security, Jun. 2018, pp.232–241. DOI: 10.1007/978-3-030-00009-7_22.

[97] Liu C, Jiang H. Classification of traditional Chinese paintings based on supervised learning methods. In *Proc. the 2014 IEEE International Conference on Signal Processing, Communications and Computing*, Aug. 2014, pp.641–644. DOI: 10.1109/ICSPCC.2014.6986272.

[98] Bao H, Xu D, Feng S H. An effective method to detect seal images from traditional Chinese paintings. In *Proc. the 2009 International Conference on Wireless Communications & Signal Processing*, Nov. 2009. DOI: 10.1109/WCSP.2009.5371700.

[99] Guan X H, Pan G, Wu Z H. Automatic categorization of traditional Chinese painting images with statistical Gabor feature and color feature. In *Proc. the 5th International Conference on Computational Science*, May 2005, pp.743–750. DOI: 10.1007/11428831_92.

[100] Hu Z K, Wang T M. Object extraction in Chinese painting base on visual saliency. In *Proc. the 8th International Symposium on Computational Intelligence and Design*, Dec. 2015, pp.493–496. DOI: 10.1109/ISCID.2015.175.

[101] Xu S H, Tan H S, Jiao X T, Lau F C M, Pan Y H. A generic pigment model for digital painting. *Computer Graphics Forum*, 2007, 26(3): 609–618. DOI: 10.1111/j.1467-8659.2007.01084.x.

[102] Jiang S Q, Huang Q M, Ye Q X, Gao W. An effective method to detect and categorize digitized traditional Chinese paintings. *Pattern Recognition Letters*, 2006, 27(7): 734–746. DOI: 10.1016/j.patrec.2005.10.017.

[103] Guo F, Tang J, Peng H. Image recovery for ancient Chinese paintings. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2013, 6(5): 165–178. DOI: 10.14257/ijsip.2013.6.5.15.

[104] Pei S, Chiu Y. Background adjustment and saturation enhancement in ancient Chinese paintings. *IEEE Trans. Image Processing*, 2006, 15(10): 3230–3234. DOI: 10.1109/TIP.2006.877478.

[105] Pei S C, Zeng Y C, Chang C H. Virtual restoration of ancient Chinese paintings using color contrast enhancement and lacuna texture synthesis. *IEEE Trans. Image Processing*, 2004, 13(3): 416–429. DOI: 10.1109/TIP.2003.821347.

[106] Ding H Y, Ding H D. Research on computer color recovery system for traditional Chinese painting. In *Proc. the 2012 International Conference on Systems and Informatics*, May 2012, pp.1985–1988. DOI: 10.1109/ICSAI.2012.6223439.

[107] Fan Z B, Zhang K, Zheng X S. Evaluation and analysis of white space in Wu Guanzhong's Chinese paintings. *Leonardo*, 2019, 52(2): 111–116. DOI: 10.1162/LEON_a_01409.

[108] Fan Z B, Li Y N, Yu J H, Zhang K. Visual complexity of Chinese ink paintings. In *Proc. the ACM Symposium on Applied Perception*, Sept. 2017, Article No. 9. DOI: 10.1145/3119881.3119883.

[109] Bao H, Liang Y, Liu H Z, Xu D. A novel algorithm for extraction of the scripts part in traditional Chinese painting images. In *Proc. the 2nd International Conference on Software Technology and Engineering*, Oct. 2010, pp.V2-26—V2-30. DOI: 10.1109/ICSTE.2010.5608756.

[110] Liang Y, Bao H, Liu H Z. A simple method for classification of traditional Chinese painting and calligraphy images. In *Proc. the 2010 International Conference on Educational and Information Technology*, Sept. 2010, pp.V3-340—V3-344. DOI: 10.1109/ICEIT.2010.5608356.

[111] Yuan C, Yun Z. Tunable, a VR reconstruction of "Listening to a Guqin" from emperor Zhao Ji. In *Proc. the SIGGRAPH ASIA 2016 VR Showcase*, Nov. 2016, Article No. 6. DOI: 10.1145/2996376.2996379.

[112] Subramonyam H, Shen Y C, Jones S L. SIGCHI: Enabling context for traditional Chinese paintings with "rice paper". In *Proc. the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, Apr. 2015, pp.49–54. DOI: 10.1145/2702613.2726958.

[113] Hsieh C K, Hung Y P, Ben-Ezra M, Hsieh H F. Viewing Chinese art on an interactive tabletop. *IEEE Computer Graphics and Applications*, 2013, 33(3): 16–21. DOI: 10.1109/MCG.2013.50.

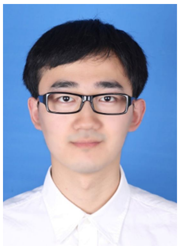
[114] Tang F, Dong W M, Meng Y P, Mei X, Huang F Y, Zhang X P, Deussen O. Animated construction of Chinese brush paintings. *IEEE Trans. Visualization and Computer Graphics*, 2018, 24(12): 3019–3031. DOI: 10.1109/TVCG.2017.2774292.

[115] Yang L J, Xu T C. Animating Chinese ink painting through generating reproducible brush strokes. *Science China Information Sciences*, 2013, 56(1): 1–13. DOI: 10.1007/s11432-012-4740-2.

[116] Xie N, Hachiya H, Sugiyama M. Artist agent: A reinforcement learning approach to automatic stroke generation in oriental ink painting. *IEICE Trans. Information and Systems*, 2013, E96.D(5): 1134–1144. DOI: 10.1587/transinf.E96.D.1134.

[117] Yao F H, Shao G F. Painting brush control techniques in Chinese painting robot. In *Proc. the 2005 IEEE International Workshop on Robot and Human Interactive Communication*, Aug. 2005, pp.462–467. DOI: 10.1109/ROMAN.2005.1513822.

Wei Zhang received her M.A. degree in visual communication design from Central China Normal University, Wuhan, in 2012. She is currently a Ph.D. candidate in the State Key Laboratory of CAD&CG at Zhejiang University, Hangzhou. Her current research interests include digital humanities visualization and visual analytics. More information can be found at: https://zwstart.github.io.



Jian-Wei Zhang received his B.S. degree in mathematics and applied mathematics at Zhejiang University, Hangzhou, in 2018. He received his Ph.D. degree in computer science and technology at Zhejiang University, Hangzhou, in 2023. He is currently employed at Tencent. His main research interests are AI generated content (AIGC), large language model (LLM), and representation learning.

Kam-Kwai Wong received his B.E. degree in computer science and engineering from The Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2016. He is currently a Ph.D. candidate in the Department of Computer Science and Engineering at The Hong Kong University of Science and Technology (HKUST), Hong Kong. His main research interests are data visualization, visual analytics, and data mining.

Yi-Fang Wang is currently a postdoctoral fellow in The Center for Science of Science & Innovation in Kellogg School of Management at Northwestern University. Before that, she obtained her Ph.D. degree in computer science and engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2022. Her research interests lies broadly in visual analytics of computational social science and digital humanities. Her current research focuses includes visual analytics of Science, social mobility and inequality, and data-driven art about social science and humanities.

Ying-Chao-Jie Feng received his B.E. degree in software engineering from Zhejiang University of Technology, Hangzhou, in 2020. He is currently a Ph.D. candidate in the State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou. His research interests include data visualization, human-computer interaction, natural language processing, and digital humanities.

Lu-Wei Wang obtained her Bachelor's degree in computer science from Zhejiang University, Hangzhou, in 2023. She is currently a Master student in University of Illinois Urbana-Champaign in the United States. She is interested in computer visualization and computer vision.

Wei Chen received his Bachelor's and Ph.D. degrees from Zhejiang University, Hangzhou, in 1996 and 2002, respectively. He is currently a professor in the State Key Lab of CAD&CG, Zhejiang University, Hangzhou. His research interests include visualization and visual analysis, and he has published more than 70 IEEE/ACM Transactions and IEEE VIS papers. He actively served as guest or associate editors of the ACM Transactions on Intelligent System and Technology, the IEEE Computer Graphics and Applications, and Journal of Visualization.