International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-12

https://doi.org/10.70917/ijcisim-2026-1850
Article

A Study of the Thematic Imagery and Cultural Implications of Taohawu New Year Paintings from an Iconographic Perspective

Yichao Yan [1,2] and Mohammad Khizal Bin Mohamed Saat [1,*]

1 School of the Arts, Universiti Sains Malaysia, Pulau Pinang, 11800, Malaysia
2 Henan University of Urban Construction, Pingdingshan, Henan, 467036, China
* Correspondence author: mkhizal@usm.my

Abstract: Taohuawu New Year paintings embody traditional spiritual culture and reflect the emotional needs of the people, occupying an important place in traditional Chinese art. This paper adopts an image-based approach to extract key image features from Taohuawu New Year paintings, including color histograms, SIFT-BoF features, and global contrast. Based on these results, a combined Corr-LDA model is employed to identify the probabilistic themes and cultural connotations of the New Year paintings. The results indicate that the average grayscale value in the color histogram of Taohuawu New Year paintings with the "warm/romantic" emotional theme is 0.04853, while those with the corresponding "festive/joyful" and "calm/relaxed" themes are 0.02932 and 0.00931, respectively. The thematic analysis of the New Year paintings yielded six probabilistic themes: folk life, agricultural activities, landscapes, supernatural beings, patterns, and opera. The most frequent keyword was "festival" under the folk life theme, with a frequency of 105. Taohuawu New Year paintings are a direct reflection of the aesthetic sensibilities of the common people and an intrinsic expression of national character and spiritual values.

Keywords: Iconography; Color histogram; Sift-BoF; Corr-LDA; Taohuawu New Year paintings

1. Introduction

According to German scholar Jan Assmann, cultural memory is an objective externalization fixed in the form of words, images and dances, which is presented in traditional symbolic coding and performance, with uniqueness and durability, and capable of constructing collective identity [1]. The core function of cultural memory is to construct collective multiple identities, including identity, emotional identity, cultural identity and value identity [2]. Cultural inheritance relies on cultural memory, which connects the past and the present and constructs collective multiple identities through cultural practices, oral traditions and material storage media [3].

Suzhou "Taohuayu" woodblock prints originated from the Song Dynasty engraving and printing process, from the earliest monochrome printing to the red, yellow, green, black and blue five-color engraving and printing, through the development of the Ming and Qing dynasties to form a rich subject matter, fine production, smooth lines, rich colors, delicate and elegant folk art style of the Jiangnan region [4]. The prosperity and development of woodblock prints in "Peach Blossom Garden" is attributed to the prosperous economy, colorful people's life, and blossoming poetic and pictorial arts in Gusu City during the Ming and Qing Dynasties. As a historical district of Gusu City, "Peach Blossom Dock" has been a gathering place for stores, workshops and folk artists of Su-style woodblock prints since ancient times [5-6]. After a development period of over 600 years, it has become the hallmark and symbol of the Su-style woodblock New Year paintings. In the subsequent innovative development of Taohuawu woodblock New Year paintings, whether it can still maintain the distinct characteristics of

"Taohuawu" and inherit the core spirit of "Taohuawu" is the key to the innovation of "Taohuawu" New Year paintings [7]. The distinctive regional characteristics of “Peach Blossom Dock” are mainly manifested in its subject matter, color, carving, modeling and artistic style, etc. [8]. For example, Taohuayu woodblock prints have the unique style of painting-type woodblock prints, such as full-bodied compositions, brilliant colors, and fine images, as well as the Jiangnan characteristics of delicate beauty, vividness and elegance, compared with those of Yangliuqing in Tianjin, Zhuxianzhen in Kaifeng, and Fengxiang in Shaanxi Province.

In terms of cultural connotation, Taohuayu New Year's Paintings, as an important carrier of historical memory and cultural identity in the Jiangnan region, is a visualized presentation of collective memory [9-11]. It transmits historical memory and shapes local identity and cultural identity through the symbols of images, textual memories, and ritual customs [12]. Cultural connotations are not static, but are constantly reconstructed and reproduced with the evolution of social history [13]. In the contemporary era of rapid cultural change, the Peach Blossom Wood New Year's Paintings are facing the double test of inheritance and innovation, and there is an urgent need to realize the reconstruction of memory and awaken people's emotional memories through digital dissemination and the application of innovative design [14-16]. The theory of cultural memory provides a new perspective for this process, which enhances the people's sense of identity and belonging to the local culture through the application of diversified memory fields and innovative memory media [17].

Image-based techniques were employed to process and analyze Taohuawu New Year paintings. The images were preprocessed using the Mean-shift clustering algorithm to segment the foreground and background. The color values of each component in the HSV color space were statistically calculated to extract color histograms. For texture features, the SIFT-BoF algorithm was used to detect keypoints and their dimensions across multiple scales, generating SIFT feature vectors. Finally, global contrast is determined through priority allocation based on regions of interest, completing the processing and extraction of the overall image features of the Taohuawu New Year paintings. Building on this, a combined Corr-LDA method is used to model image annotation based on these features, thereby extracting thematic imagery and expressions of cultural connotations.

2. Iconographic Analysis of Taohuawu New Year Paintings

2.1. Preprocessing of Taohuawu New Year's Paintings

Let the input Taohawu New Year painting image be denoted by I, and the output image by I'. There are a total of s regions, denoted by $\{r_1, r_2, \ldots, r_s\}$ where r_i represents the i th region. This paper employs the Mean-Shift algorithm to perform foreground-background segmentation on Taohuawu New Year’s painting images. The Mean-Shift clustering algorithm is applied to the $L*a*b$ color channels and two spatial dimensions. Two Gaussian kernels are used in both the color domain and the spatial domain, respectively. The Mean-Shift algorithm consists of an iterative process: first, the mean shift value for the current point is calculated; then, the point is moved to its mean shift value; finally, this new point serves as the starting point for the next iteration, and the process continues until certain conditions are met.

2.2. Color Histogram

Given the unique production techniques of woodblock New Year prints—block carving, multi-color printing, and hand-coloring—as well as the rich color palette of these prints, color is undoubtedly a key factor in classifying their imagery. Furthermore, since previous research has shown that woodblock New Year prints from different regions generally differ in their use of color, color characteristics also serve as a crucial reference point when classifying these prints by region of origin.

There are many ways to represent color features, such as the cumulative histogram method, the histogram method, and the central moment method. This paper adopts the histogram method. Histograms can clearly illustrate the proportion of different colors in an entire image without concerning oneself with the spatial location of each color; however, this approach cannot describe the objects or subjects within the image. When selecting a color space, this paper chose the HSV model, which aligns with human visual perception. This model comprises three components: H (Hue), S (Saturation), and V (Value) [18]. Additionally, this paper made several modifications to the color histogram, such as setting the values of the white region to zero and using only the H and S components.

Since the color of the paper in New Year’s paintings can significantly interfere with image processing and classification, this paper has performed background-foreground segmentation on the paintings, and the paper color has been uniformly converted to white. In this section, we will directly perform feature extraction on the image I' mentioned in the previous section. However, since white often constitutes a large proportion of the resulting color histogram, which can affect classification accuracy, this paper sets the values in the white range of the New Year's painting color histogram to zero to eliminate the interference caused by white.

Furthermore, since the brightness of the New Year's painting images does not affect classification, and to improve the algorithm's processing efficiency, this paper only performs region segmentation on the H and S components of the image, i.e., quantizing the HS color space. The H and S components are quantized into 16 levels and 8 levels, respectively.

The quantization of the H component is performed according to the following formula:

$$H = (h \bmod 350 + 10)/22 \tag{1}$$

where h denotes the hue value before quantization, taking the range of $h \in [0,360)$; and H denotes the value after quantization. The value range is $H \in [0,15]$. Finally H takes the value of the integer part of the calculation result.

The quantization of the S component is according to the following formula:

$$S = (s - \varepsilon)/0.125 \tag{2}$$

where s denotes the saturation before quantization and takes the value in the range of $s \in [0,1)$, and S denotes the saturation value after quantization and takes the value in the range of $S \in [0,7]$. ε denotes the smallest floating point number supported by this computer. As with the H component, S takes the integer portion of the calculation result.

2.3. Sift-BoF Characterization

In addition to the color features, there are also texture features in Peach Blossom New Year paintings.Sift-BoF is an important tool for texture feature extraction in woodblock prints, and its feature extraction process is shown in Fig. 1.

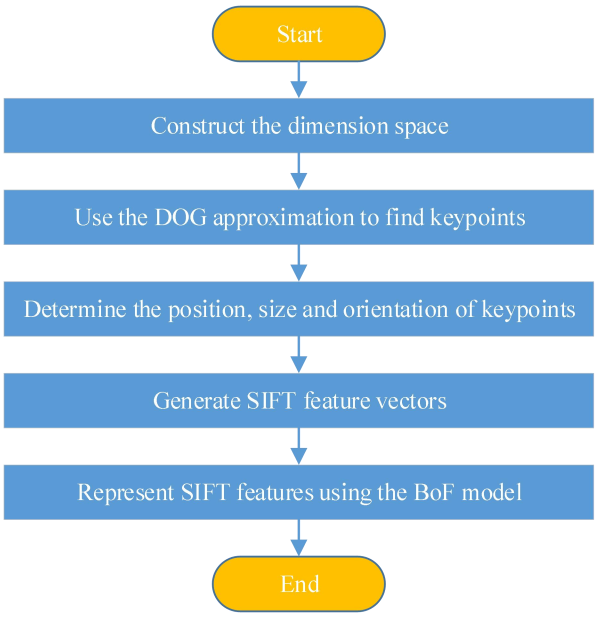


Figure 1. Sift-BoF Feature Extraction Process

(1) Constructing the scale space

The Sift-BoF algorithm in this paper realizes a multi-group scale space for yearbook images. The Gaussian convolution kernel is the only linear kernel to realize the scale transformation. The scale space of a two-dimensional image can be defined by the image and Gaussian kernel convolution as:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \tag{3}$$

where $G(x,y,\sigma)$ is the scale-variable Gaussian function and the expression of this function is:

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \tag{4}$$

where $I(x,y)$ is the original image, σ is the scale coordinate and (x,y) is the spatial coordinate. The size of σ determines the degree of smoothness of the image, the larger the value of σ indicates the overall characteristics of the image, while corresponding to a low resolution; the smaller the value of σ indicates the local details of the image characteristics, while corresponding to a high resolution.

(2) Keypoint Detection

In order to efficiently detect the more stable feature points in the multi-scale space created above, we use the DoG poles (i.e., Difference of Gaussains) in the scale space as the basis for selection. In this paper, the DoG operator is used because DoG has been shown to be very suitable for finding keypoints of interest in images, and the DoG operator is computationally simple, performs efficiently, and detects relatively stable points of interest.

(3) Determination of key point location and size

In this paper, a 3D quadratic function is fitted in order to accurately determine the location and scale of keypoints. Since some of the feature points generated by the DoG operator are on the boundary, some do not have enough contrast in brightness. These are not useful features and we need to remove them. In this paper, we remove the lower contrast feature points by detecting the luminance, setting a certain threshold, and using approximate Harris Conner detection to remove the feature points located on the boundary.

When removing the lower contrast feature points, the derivative of the spatial scale function is made equal to 0 to get the exact location of the feature points:

$$D(x,y,\sigma)=D(x,y,\sigma)+\frac{\partial D^T}{\partial x}x+\frac{1}{2}x^T\frac{\partial^2 D}{\partial x^2}x \tag{5}$$

$$\hat{x}=-\frac{\partial^2 D^{-1}}{\partial x^2}\frac{\partial D}{\partial x} \tag{6}$$

Substituting Eq. (6) into Eq. (5) yields:

$$D(\hat{x})=D(x,y,\sigma)+\frac{1}{2}\frac{\partial D^T}{\partial x}\hat{x} \tag{7}$$

In this paper, the threshold is set to 0.05, if $|D(\hat{x})|\geq 0.05$, the feature point is considered to be of high contrast and can be retained; if $|D(\hat{x})|<0.05$, the feature point is considered to be of low contrast and needs to be removed.

When removing unstable boundary feature points, this paper realizes it by calculating the gradient of the Gaussian fuzzy image near the feature point.

(4) Determine the direction of the key point

Stable and reasonable feature points have been obtained through the above steps, and in order to be able to make these features rotationally invariant, it is necessary to assign an orientation to each feature point. In this paper, the orientation parameter of each key point is determined by using the gradient direction distribution of the pixels in the neighborhood of the key point:

$$\begin{aligned} m(x,y)&=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2} \\ \theta(x,y)&=\tan^{-1}\left(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right) \end{aligned} \tag{8}$$

The above equations are the modulus and direction of the gradient at the (x,y) pixel point, respectively, where L is the scale where the key point is located. In this paper, the gradient histogram is used to count the orientation of the feature points. Where the range of the gradient histogram is from 0 to 360 degrees, with one bar for every 10 degrees, totaling 36 bars. The farther away from the center point of the field will contribute less to the histogram. Therefore, in this paper, the Gaussian function is used to smooth the histogram first, and the value of the highest peak in the histogram is the direction of the key point.

(5) Generate Sift feature vector

Decompose the 16×16 window around the feature point into 16 4×4 sub-windows. In each 4×4 sub-window, the magnitude and direction of the gradient are computed and the average direction of the sub-window is counted using a histogram with 8 intervals.

In this paper, a Gaussian weighting function is used to generate a weighted value, which is then multiplied by the gradient size of each pixel point in the window of 16×16 to get the weighted gradient size, the further away from the feature point, the smaller the gradient size of the pixel point to be added to the histogram.

In this way, each 4×4 sub-window corresponds to a histogram of 8 intervals, and the value added to the histogram is the gradient size of the pixel weighted with a Gaussian, and the 16×16 window around the feature point contains 16 4×4 sub-windows, a total of $16\times8=128$ numbers, and then the 128 number of numbers composed of the vectors are unitized, and the unitized 128-dimensional vector is the SIFT feature descriptor.

2.4. Global Contrast Ratio

Detecting the saliency region of an image by computation is significant since we can prioritize the allocation of computational resources required for image analysis and synthesis through the saliency region. Saliency is widely used in vision domains including image segmentation, target recognition, adaptive compression, content-aware image editing, and image retrieval.

We improve on the histogram contrast (HC) based method by incorporating spatial information to detect saliency.HC-maps assign saliency values to pixels based on their color difference from other pixels and are used to produce full resolution saliency images. Specifically, the saliency value of a pixel in an image is defined in terms of its color contrast with other pixels in the image. The saliency value of a pixel I_p in an image I is defined as:

$$S(I_p)=\sum_{\forall I_i\in I} D(I_p,I_i) \tag{9}$$

where $D(I_p,I_i)$ is the color distance measure between pixel I_p and pixel I_i in HSV space. Since each pixel calculates the contrast with all the pixels, it makes the pixels with the same color value also have the same saliency.

The above method has a time complexity of $O(N^2)$, which is relatively expensive to compute, and the speed of the algorithm can be greatly improved if the total number of image pixel colors is reduced. However, the true color space contains 256^3 possible colors, which is much larger than the total number of pixels in the image. According to the characteristics of the small number of colors used in the Chinese New Year paintings and the distinctive features of their origins, we also convert the image from RGB space to HSV space and quantize it into $16\times8=128$ colors when calculating the saliency features, just like calculating the weighted augmented principal color histogram. In this way, the number of colors is greatly reduced.

Although we can use the color histogram after color quantization to compute color contrast efficiently, the quantization itself may produce flaws. This is because some similar colors may be quantized to different color values. To reduce the noise that this randomness brings to the calculation of the saliency values, we use a smoothing operation to improve the saliency value of each color. The significance value of each color is replaced with a weighted average of the significance values of similar colors. This process is essentially a smoothing process in the color space. We choose $m=n/4$ nearest neighbor colors to improve the significance value of color c, see Eq:

$$S'(c)=\frac{1}{(m-1)T}\sum_{i=1}^{m}(T-D(c,c_i))S(c_i) \tag{10}$$

We used a linearly varying smoothing weight $(T-D(c,c_i))$ to assign larger weights to colors that are closer to c in the color feature space. Similar histogram intervals will be very close after smoothing, indicating that similar colors are very likely to be assigned similar saliency values, thus reducing quantization flaws. The following figure shows the saliency characteristics of a yearbook image:

By looking at the saliency map, we find that regions with large saliency values have a large contrast with the surrounding objects, and people tend to pay more attention to regions of the image that have a large contrast with the surrounding objects. In addition to contrast, spatial relationship is also a factor that draws human visual attention, that is, the distribution of saliency that we mentioned above.

3. Image-based Extraction of New Year's Paintings Theme Culture

3.1. LDA Subject Modeling

The LDA model is a three-level hierarchical Bayesian model, which consists of three layers: document, topic, and word. The LDA model is also a generative model, and its theoretical basis is the probabilistic graphical model. A generative model is a model that can randomly generate observable data. Using the LDA model, a document can be randomly generated. For example, suppose there are three topics: "science fiction", "sports", and "film and television". In a document describing the process of film production, it may contain both "science fiction" and "film and television" topics. There are a series of words related to the "science fiction" topic, and each word has a probability attached to it, representing the likelihood of each word appearing under the "science fiction" topic. Similarly, in the "film and television" topic, there is also a series of words related to "film and television", and each word also has a probability attached to it. If we want to generate a document about movie and television, the process is: (1) Choose a topic randomly, the probability of choosing two topics, science fiction and movie and television, will be high. (2) Choose words with high relevance to one of the topics chosen in step (1). (3) Repeat the steps (1) and (2) N times. In this way, a document consisting of N words can be generated. The process of generating a document is as follows:

First extract a topic scale (the topic scale varies from document to document) $\theta \sim Dir(\alpha)$;

Then for each vocabulary $w_n, n \in \{1,2,3,\cdots,N\}$:

(a) Extract a subject assignment $z_n \mid \theta \sim Mult(\theta)$;

(b) Extracting a text word $w_n \mid z_n \sim Mult(\pi_{z_n})$;

The joint probability distribution formula for the LDA model is shown in (11):

$$p(\theta,Z,W \mid \alpha,\pi) = p(\theta \mid \alpha)\sum_{n=1}^{N} p(Z_n \mid \theta)\, p(W_n \mid Z_n,\pi) \tag{11}$$

The LDA model generation process has only one stage, the text vocabulary extraction stage.

Text Vocabulary Extraction: The model extracts a potential topic proportion θ with a Dirichlet distribution with parameter α, and extracts a topic label z according to this proportion, and then extracts a text vocabulary according to the corresponding topic label, and finally repeats N times to generate a document.

3.2. Corr-LDA modeling for image imagery

Corr-LDA is a classical model for image annotation work, which extends LDA into a multimodal model that flexibly relates images and their corresponding texts. The model not only considers the conditional relationship between an image and its corresponding text, but also pays attention to the joint action between them.The Corr-LDA probabilistic graphical model representation is shown in Fig. 2.

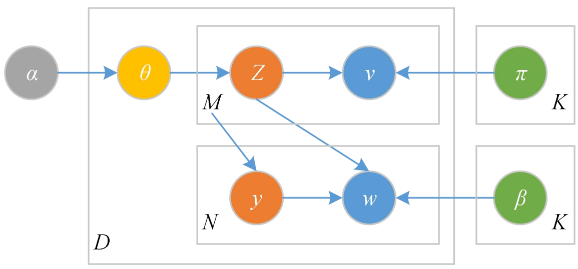


Figure 2. Corr-LDA Probabilistic Graph Model

The specific generation process is as follows:

In the first step, a subject scale $\theta \sim Dir(\alpha)$ is extracted;

In the second step, for each image word $v_m, m \in \{1,2,3,\cdots,M\}$:

(a) Extract the subject assignment $z_m \mid \theta \sim Mult(\theta)$;

(b) Extract image words $v_m \left| z_m \sim Mult\left(\pi_{z_m}\right)\right.$;

In the third step, for each text word $w_n, n \in \{1,2,3,\cdots,N\}$:

(a) Extract index variable $y_n \sim Unif\left(1,2,3,\cdots,M\right)$;

(b) Extraction of text words $w_n \sim Mult\left(\beta_{z_{y_n}}\right)$;

The joint probability distribution of the Corr-LDA model is shown in Equation (12):

$$P\left(v,w,\theta,Z,y\middle|\alpha,\pi,\beta\right)=P\left(\theta\middle|\alpha\right)\left(\sum_{m=1}^{M}A\right)\left(\sum_{n=1}^{N}B\right) \tag{12}$$

Among them:

$$A=P\left(z_m\middle|\theta\right)P\left(v_m\middle|z_m,\pi\right) \tag{13}$$

$$B=P\left(y_n\middle|M\right)P\left(w_n\middle|y_n,Z,\beta\right) \tag{14}$$

The Corr-LDA model generation process is divided into two stages: the stage of extracting image vocabulary and the stage of extracting text vocabulary.

Image vocabulary extraction: the model extracts a potential theme proportion θ with a Dirichlet distribution parameterized by α , and extracts a theme label z according to this proportion, then extracts an image vocabulary according to the corresponding theme label, and finally repeats the process M times to generate an image.

Text vocabulary extraction: according to the theme proportion of the image generated in the previous stage, a theme is selected to generate a labeled word, and repeat N times to generate a labeled word for the image.

4. Analysis of Image Features and Theme Extraction of Peach Blossom New Year Painting

4.1. Color Histogram Extraction Analysis

For the image of the new year painting of Peach Blossom Dock, one can make histograms of multiple components of the color space separately, or make histograms of only one determinant component in the color image, or transform the image to other color models and then make statistical histograms. A piece of Peach Blossom New Year painting work is randomly selected and its color is counted by using the imaging method, and Fig. 3 shows the obtained global histogram. It can be found that in the HSV color space, the color values of this Peach Blossom Dock New Year Painting are mainly concentrated in the range of gray values of (0,60), although there is also a sample density distribution near the gray value = 160.

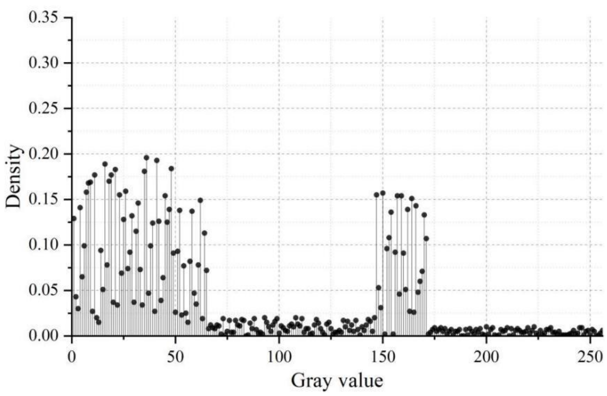


Figure 3. Global color histogram

In statistical histograms, the similarity of colors is usually used to measure the image differences, and when the image features cannot cover all the values taken, there will be zero values to analyze the similarity of the two colors, which will affect the calculation of the similarity metrics, and the resultant accuracy will be biased. The horizontal coordinate in the cumulative histogram is the color value and the vertical coordinate is the color cumulative frequency. Figure 4 shows the obtained cumulative color histogram. Different from the global histogram, the cumulative histogram can more intuitively see the color extension of the new year paintings in Taohuayu, and the final cumulative value reaches 9.783. The advantages of the cumulative histogram are that the common zero value is eliminated, and the neighboring colors are correlated in the frequency of the vertical coordinate, which is able to judge the similarity of the colors, and it also overcomes the defects of the general histogram that the quantization of the uneven matching is not effective, and the disadvantage is that the amount of calculation and the amount of storage is increased. The disadvantage is that it increases the amount of calculation and storage. Since the H, S and V channels are independently distributed, but their joint distribution is not unique, if the joint distribution method is not considered, it will lead to an increase in the number of dissimilar images in the result set.

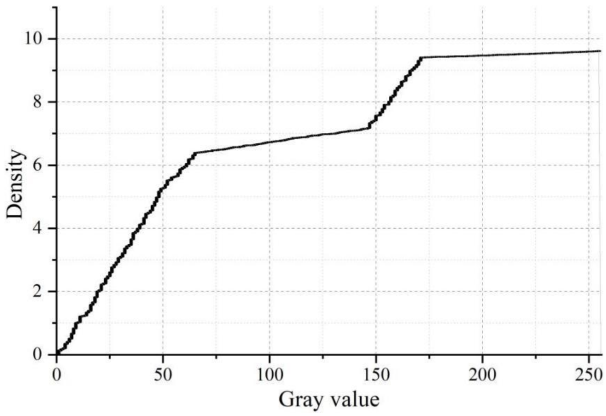
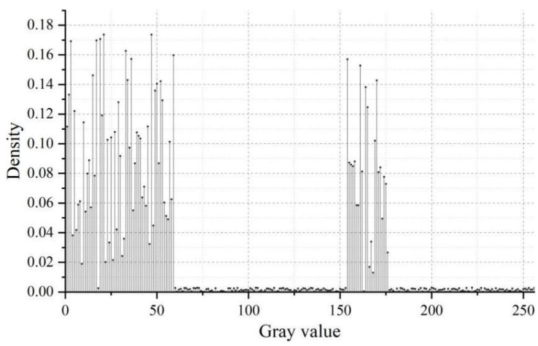
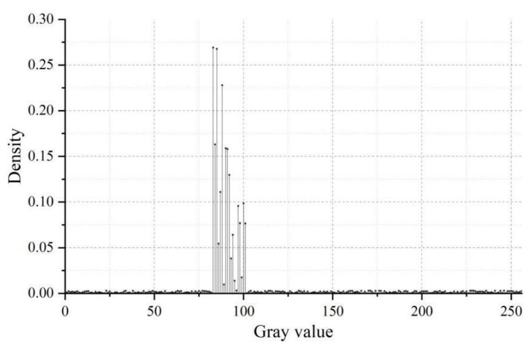


Figure 4. Cumulative color histogram

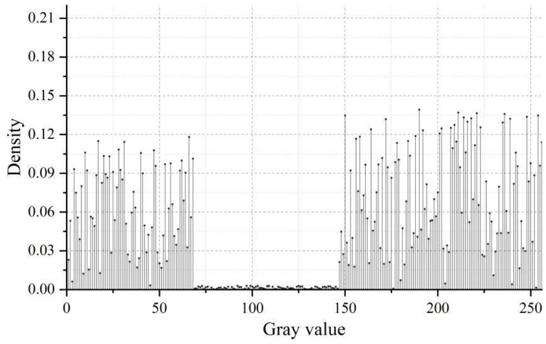
Since the task of this article is to extract the themes and cultural connotations of the Taohuawu New Year paintings, a comparison of the histogram distributions under different emotional themes of the Taohuawu New Year paintings is added here. The frequency of pixels on each color quantization level is counted to obtain the color histograms of the Taohuawu New Year paintings under the three emotional themes of "joyful/pleased", "calm/released", and "warm/romantic". The comparison of color statistics for different themes is shown in Figure 5. The "warm/romantic" emotional theme has the widest color distribution, with an average gray value of 0.04853. While the average gray values of the color histograms under the "joyful/pleased" and "calm/released" emotional themes are 0.02932 and 0.00931 respectively. Based on the color histograms, different Taohuawu emotional themes can be preliminarily distinguished.



(a) Joyful / Cheerful



(b) Calm/Relax

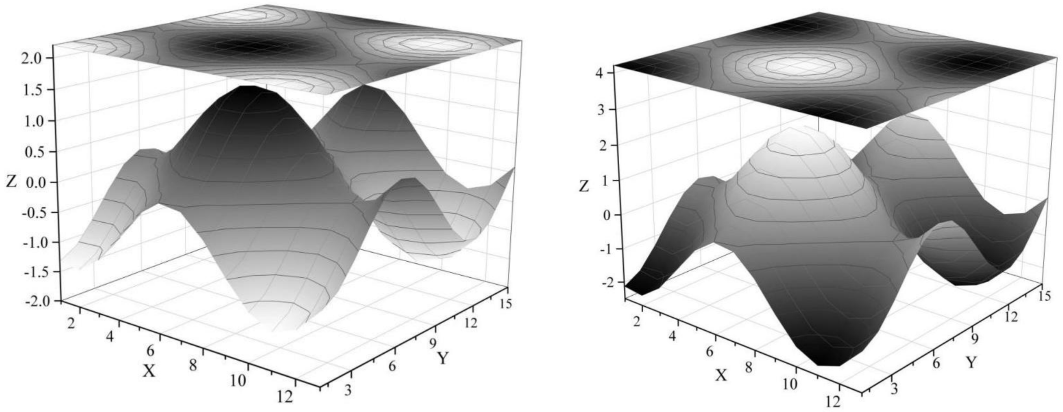


(c) Warmth / Romance

Figure 5. Color statistics comparison of different themes

4.2. Texture Sift-BoF feature extraction analysis

The Sift-BoF algorithm in this paper realizes the multi-group scale space of the New Year's picture image, and the appropriate 2D Sift-BoF filter function is selected, and the real and imaginary parts of the spatial-domain Sift-BoF filter are given in the schematic diagram in Fig. 6. The convolution output of the image is also in complex form, and we take the measure of this complex as the final Sift-BoF feature. The Sift-BoF filter applied to the image of the new year painting of Taohuayu contains wavelet kernels of different spatial frequency domains and orientations to be convolved with the image. So that the different scales and directions of the filters cover the entire frequency domain, with the help of such a complete set of filters to the image convolution, can effectively quantize the bandwidth within the local frequency information amplitude and phase.



(a) Sift-BoF.real component

(b) Sift-BoF.imaginary component

Figure 6. Spatial domain Sift-BoF filter

This Sift-BoF filter is used to extract the Sift-BoF features of the new year paintings of Peach Blossom Garden, and the Sift-BoF features extracted for 80 new year paintings of Peach Blossom Garden are shown in Fig. 7. The extracted Sift-BoF features are basically above 75%, with a maximum of 92.15%, and the extracted features in this paper can fully characterize the iconographic features of the new New Year paintings in Peach Blossom Garden.

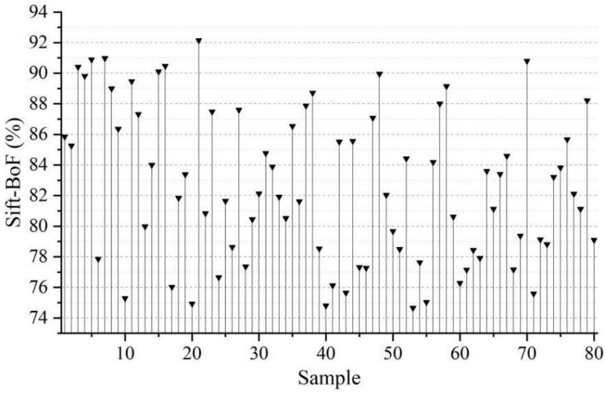


Figure 7. The characteristics of the New Year paintings of the Taohuawu area

4.3. Cultural Extraction Analysis of New Year's Paintings Theme

With the information of the new year paintings of Taohuawu obtained by color feature and texture feature extraction, the Corr-LDA model is used to extract the theme of the new year paintings of Taohuawu and discuss the cultural connotation behind it. The extracted keywords of the theme of the new year painting of Taohuawu are shown in Table 1. The first 8 keys under theme 1 are festival, birthday, marriage, marketplace, amusement, textile, children's play, and ladies in order, and the keyword frequency decreases in order, with the highest being festival (105) and the lowest being ladies (27). Similarly, from the table, we can get the keyword characteristics and frequency of the rest of the themes, and according to these keywords and frequency, we can name the themes as follows: Theme 1=folk life, Theme 2=farming, Theme 3=scenery, Theme 4=spirits, Theme 5=patterns, and Theme 6=plays and writings.

Table 1. Key Words for the New Year Painting Theme of Taohuawu

Topic 1	**Frequency**	**Topic 2**	**Frequency**	**Topic 3**	**Frequency**
festival	105	sowing	82	garden	95
longevity	63	plowing	81	landscape	77
marriage	63	mulberry picking	76	bridge	72
market	49	silkworm raising	75	street	56
recreation	45	irrigation	61	lake	51
weaving	45	harvest	57	pagoda	42
children's play	36	threshing	49	water town	38
ladies	27	grain drying	40	courtyard	35
Topic 4	Frequency	Topic 5	Frequency	Topic 6	Frequency
fox spirit	104	flowers	75	battle	103
snake spirit	98	birds & beasts	69	love	96
ghost	91	cloud pattern	68	loyalty	79
monster	89	water pattern	62	filial piety	69
dragon god	85	entwined branches	59	reunion	60
phoenix	84	ruyi	54	revenge	59
qilin	60	bat	41	court case	42
demon	44	longevity character	35	scholar and beauty	29

The most prominent aesthetic quality of Taohuayu New Year's paintings is that they are full of happiness and symbols, and contain the praise of life, which is a direct reflection of the Chinese people's character of being happy with life and loving peace. New Year's paintings often express people's hope for more children, more happiness, and a peaceful and stable life, and at this time, New Year's paintings perfectly serve as the carrier of these beautiful visions, carrying the praises of life and the love of life that people have been carrying for thousands of years. Peach Blossom Garden New Year Paintings also use the harmonic meanings of the things they depict or the hidden symbols behind them to express their beautiful meanings. Some of the paintings also use auspicious and beautiful words, and by naming or inscribing the works, they achieve the purpose of good meaning.

The Chinese people believe that balance is a kind of beauty, which is in line with the middle way advocated by Confucianism, so they have a strong attachment to even numbers and symmetry, which is externalized in poetry, architecture and folklore, forming the concept of stability. This is reflected in the

Taohuayu New Year's Paintings, whose compositions emphasize fullness and symmetry, balance and completeness.

The Chinese people have a long history of spiritually honoring the unity of heaven and man. The concept of the unity of heaven and man has guided the way of thinking and behavior of the Chinese people. The Chinese people believe that even as a separate individual, man is one with everything in the universe, and this idea of the unity of man and heaven is reflected in the ideas of Taoism, Confucianism and Buddhism. This idea is reflected in Taoism, Confucianism and Buddhism, and is reflected in the Peach Blossom Garden New Year's Paintings, where there are a lot of subjects such as statues of gods and fairies. In addition to drawing images of the gods and fairies, folklore activities have also been created to pray for good weather conditions and no illnesses in the coming year.

5. Conclusion

In this paper, color histogram statistics, Sift-BoF feature algorithm, and global contrast calculation are applied to image science processing of the new year paintings in Peach Blossom Garden to obtain key features, and then Corr-LDA probabilistic theme model is applied to further extract the imagery theme and cultural expression behind the new year paintings from the obtained features. The study finds that:

(1) The color histogram can indirectly reflect the mapping of the color gray values of the new year paintings of Taohuawu under different themes of emotional expression, and the cultural expression of the new year paintings can be preliminarily classified by comparing the global histogram distributions and cumulative histogram distributions of different works.

(2) The Sift-BoF filter designed in this paper performs the extraction task of texture features of the new year paintings in Peach Blossom Garden well, and the Sift-BoF feature extraction rate reaches up to 92.15%, and the majority of them are more than 75%.

(3) The Corr-LDA combined model extracts six major themes of imagery in the new year paintings of Peach Blossom Lake, which are folk life, agricultural affairs, landscape, elves and monsters, patterns, and plays and texts, and by discussing and analyzing the keywords under these themes, we can excavate the cultural connotation and spiritual expression of the new year paintings of Peach Blossom Lake.

References

1. Assmann, J. (2015). Memory and culture. Memory: A history, 101, 325.
2. Ißler, R. A. (2019). Cultural memory. In The Bonn Handbook of Globality: Volume 2 (pp. 807-819). Cham: Springer International Publishing.
3. Shershova, T., & Chaika, V. (2024). Cultural memory and memory culture. Humanities Studios: Pedagogy, Psychology, Philosophy, 1(12), 188-195.
4. Zhu, J., & Kim, B. K. (2023). Female Characters in Chinese New Year Pictures of Ming and Qing Dynasties: Focusing on Taohuawu New Year Pictures. Journal of Humanities and Social Sciences Studies, 5(4), 72-86.
5. Fang, Y., Jiang, Z., Zhang, R., & Gu, L. (2023, July). Construction and Design Application of Taohuawu New Year Woodblock Prints Color Database. In International Conference on Human-Computer Interaction (pp. 38-51). Cham: Springer Nature Switzerland.
6. Dai, F., Jiang, Z., Fang, Y., & Guan, X. (2022, August). Research on User's Subjective Preference of Taohuawu New Year Painting Based. In Man-Machine-Environment System Engineering: Proceedings of the 22nd International Conference on MMESE (Vol. 941, p. 44). Springer Nature.
7. Peng, X. (2021, January). Study on the application and research of visual symbols of traditional New Year pictures in modern design. In The 6th International Conference on Arts, Design and Contemporary Education (ICADCE 2020) (pp. 580-583). Atlantis Press.
8. Lei, H. (2014, November). The digital protection and inheritance of Chinese woodblock new year prints. In 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (pp. 277-281). IEEE.
9. Zhu, Z., & Cui, M. (2020). Value Evaluation of Comprehensive Development of Woodblock New Year Prints Based on FAHP. In E3S Web of Conferences (Vol. 179, p. 02008). E3S Web of Conferences.

10. Huang, H. (2025). Depictions of Wopao zhuan in Chinese Woodblock Prints. CHINOPERL: Journal of Chinese Oral and Performing Literature, 44(1), 25-47.
11. Du, J., & Wang, Y. (2019, March). Analysis of the Color of Qin Qiong and Jing De Door-God Picture of Taohuawu and Fengxiang. In Advances in Graphic Communication, Printing and Packaging: Proceedings of 2018 9th China Academic Conference on Printing and Packaging (pp. 144-149). Singapore: Springer Singapore.
12. Zheng, Q., Zhang, S., Liang, J., Chen, Y., & Ye, W. (2023). The impact of cultural memory and cultural identity in the brand value of agricultural heritage: A moderated mediation model. Behavioral Sciences, 13(2), 79.
13. Assmann, J. (2011). Communicative and cultural memory. In Cultural memories: The geographical point of view (pp. 15-27). Dordrecht: Springer Netherlands.
14. Al Humam, A. A. (2025). Collective memory and cultural identity review. Scientific Journal of King Faisal University: Humanities and Management Sciences, 26(1), 3.
15. Zhang, S. N., Ruan, W. Q., & Yang, T. T. (2021). National identity construction in cultural and creative tourism: The double mediators of implicit cultural memory and explicit cultural learning. Sage Open, 11(3), 21582440211040789.
16. Gusevskaya, N., & Plotnikova, E. (2020, December). Historical memory and national identity. In 2020 6th International Conference on Social Science and Higher Education (ICSSHE 2020) (pp. 1026-1030). Atlantis Press.
17. Bouchat, P., & Rimé, B. (2018). National identity and collective memory: A social psychological perspective. In Bridging disciplinary perspectives of country image reputation, brand, and identity (pp. 89-103). Routledge.
18. Fahmi, M., Yudhana, A., & Sharkawy, A. N. (2025). Classification for waste image in convolutional neural network using Morph-HSV color model. Scientific Journal of Engineering Research, 1(1), 18-25.

About the Author
Yichao Yan is pursuing a Doctor of Fine Arts degree at Universiti Sains Malaysia. Her research interests also include the intangible cultural heritage and preservation of traditional Chinese art.