Artistic Color Matching Technology Based on Silhouette Coefficient and Visual Perception

Huizhou Li[1]*, Wubin Zhu[2]
School of Fine Arts and Design, Hefei Normal University, Hefei, 230601, China[1]
Zhejiang Uniview Technologies Co., Ltd, Hangzhou, 310051, China[2]

***Abstract*—Now-a-days, traditional color matching methods cannot meet the current market demand. Meanwhile, there are many factors to consider in the design, which affect the design efficiency. Therefore, it is necessary to seek more efficient design methods. So, this study proposed an improved K-Means based on silhouette coefficients and designed an image main color adaptive extraction model. Subsequently, an evaluation method for artistic color matching schemes based on visual perception and similarity measurement was introduced. Finally, Pix2Pix based on visual aesthetics was designed to develop color matching schemes. These results confirmed that in the objective evaluation of the main color extraction results, the structural similarity of the main color images generated using the silhouette coefficient was superior to other methods. The maximum structural similarity of this method was 0.675, with an average of 0.663. Meanwhile, the peak signal-to-noise ratio of the main color image generated by this method reached a maximum of 21.49 dB, with an average of 21.05 dB. In the validation of Pix2Pix based on visual aesthetics, the average color palette similarity of Pix2Pix's design scheme based on visual aesthetics was 0.807. Meanwhile, the average comprehensive evaluation index of this method was 0.798, which was better than Pix2Pix without integrating visual aesthetics. In the experimental verification of computational efficiency, the average color matching time of the Pix2Pix network model based on visual aesthetics is only 13.75ms. The average time consumption of the K-Means clustering algorithm model is as high as 135.67ms. Overall, the designed image main color adaptive extraction model and color matching model have strong practical applicability. These methods provide effective auxiliary design solutions for the design and development of artistic products, which helps to improve design efficiency.**

***Keywords*—Silhouette coefficient; visual perception; K-Means; color matching; similarity measurement; Pix2Pix**

I. Introduction

Due to the rapid development of industrial technology, significant achievements are made in the application of color science in engineering. Color plays a very important role in interior design, cultural and creative industries, printing and packaging, dyeing and finishing chemistry, and textile and clothing [1-2]. Colors are usually presented in groups of multiple forms in the application field, namely color matching schemes. Human Visual Perception (VP) has an impact on color, so there are certain patterns in color matching schemes under sensory recognition. VP also has advantages and disadvantages in human perception. How to design excellent color matching schemes is a key issue faced by designers and related industry personnel [3]. In response to the challenges encountered in design, computer-aided design patterns based on image processing and computer vision increasingly emerge and become a research focus in intelligent design. Through image processing technology and artificial intelligence algorithms, color processing and digital representation can be achieved, while intelligent design, matching, and recommendation can be achieved based on the attribute elements of color. However, color matching is influenced by physical, physiological, and psychological factors, which increases the complexity of implementing intelligent color matching design algorithms [4-5]. Currently, the evaluation of color matching schemes mainly relies on subjective methods such as sensory engineering. Although these methods can reflect user preferences, the evaluation is often cumbersome and easily influenced by subjective factors [6].

At present, domestic and foreign researchers mainly introduce the application research of color matching algorithms in the field of color science from three aspects: image main color extraction, color matching method research, and color matching evaluation methods. The main color of an image is the representative color of the image. Generally speaking, the number of main colors in an image is much smaller than the total number of colors in the source image. Meanwhile, the main color of the image is the color description that best fits the visual perception of the source image. In the research of color matching algorithms, the extraction of image main colors plays a crucial role in stimulating designers' design inspiration and obtaining high-quality color matching schemes. Due to the different color information in each image in nature, the corresponding number of main colors should have certain differences. Therefore, how to perform adaptive main color extraction on images is an important research topic in this field. In the research of color matching methods, due to the various attributes related to the human visual perception system such as psychology, physics, and physiology, how to combine human visual perception to generate color matching schemes in the field of computer-aided design is an important issue in this field of research. In addition, in the field of intelligent design, the intelligent matching and design of colors is a subjective design requirement. However, subjective evaluation methods are often more complex, which significantly affects product design efficiency. Therefore, using a combination of subjective and objective evaluation methods to evaluate color matching schemes is of great significance. How to conduct a comprehensive and multidimensional analysis of color matching schemes is currently a research focus and difficulty. Therefore, an image main color adaptive extraction method based on Silhouette Coefficient (SC) was adopted in this study, and a visual aesthetics based Pix2Pix was designed to formulate color matching schemes. Meanwhile, a multi-dimensional color matching evaluation method combining VP and similarity measurement was proposed. The contribution of this study lies in proposing an intelligent computer-aided design method for artistic color matching, which fully integrates VP and intelligent algorithms, helping to design color matching schemes that are more in line with the human eye VP.

The research content mainly includes six sections. Firstly, a review is conducted on K-Means based on SC and color image processing techniques. Secondly, an improved K-Means based on SC is introduced, and an image main color adaptive extraction model is designed. Meanwhile, a method for evaluating artistic color matching schemes and a model for developing color matching schemes are designed in Section III. In Section IV, the proposed method is experimentally validated. Finally, Section VI is summarized and future prospects are proposed.

II. Related Works

SC can effectively improve clustering performance, which is widely used in the improvement of K-Means. Regarding the clustering problem of student data, A. Yudistira et al. proposed using K-Means to analyze student datasets. These results confirmed that a total of 3 clusters were obtained, including 59, 94, and 1 student, respectively. The elbow method was used to determine a good classification number of 3, with an SC of 0.489, indicating that the clustering results of this study were superior [7]. Xiang et al. introduced K-Means and SC to address the low quality of answering questions in Brainly and segmented users by tracking the records of answering mathematical topic questions. These results confirmed that K-Means performed better in silhouette score, with a score of 0.9081 [8]. Ben Marzouk et al. proposed a K-Means-based energy consumption structure analysis method to address the large and chaotic data in analyzing energy consumption structures in different regions. Meanwhile, this team analyzed energy consumption using the elbow method and SC. These results confirmed that this algorithm efficiently conducted data mining, greatly improving the convenience of energy consumption structure analysis [9]. Qu et al. proposed the Bert-CK mode to address the instability of traditional K-Means. This pattern combined Bert to extract semantic, syntactic features, and SC, improved the CK mean+algorithm, and solved the instability of K value and initial centroid selection. These results confirmed that the Bert-CK model outperformed the baseline model, improving the accuracy of user classification and topic features [10]. To address the instability and local optima of traditional K-Means, Agustino et al. proposed a relative mass algorithm based on data fields. This algorithm selected high-quality points as the initial clustering centroids and used SC to improve K-Means to improve the analysis performance. These results confirmed that the acceleration ratio of the improved algorithm increased to 1.91, and the computational efficiency increased by 33.03% [11].

Currently, color plays an important role in image processing. Tabatabaian et al. proposed an algorithm based on a combination of color and shape features to address the insufficient image retrieval efficiency. Cumulative histograms were used to calculate color features, 7 Hu invariant moments were used to calculate shape features, and weights were combined to measure similarity using Euclidean distance. These results confirmed that this algorithm effectively improved the accuracy of image retrieval [12]. In response to the poor performance and resource consumption of traditional photo enhancement methods on high-resolution images, Zeng et al. proposed a learning-based image adaptive 3D lookup table method. By learning 3D lookup tables and small convolutional neural networks, fast and flexible photo enhancement was achieved, with model parameters less than 600,000 and processing 4K images in less than two milliseconds. The PSNR, SSIM, and color difference metrics of this method were superior to state-of-the-art photo enhancement methods, demonstrating efficient and superior performance [13]. Berman et al. found that color distorting et al. happened in underwater images, and their study considered different water types' spectral profiles. By assessing only the blue-red and blue-green color channels' attenuation ratio, this challenge was simplified. These results confirmed that this dataset achieved strict quantitative evaluation of natural image restoration algorithms for the first time [14]. To address underwater images' color deviation and low contrast, Li et al. proposed a multi-color space embedding underwater image enhancement network called Ucolor guided by medium transmission. This network was combined with attention mechanisms to adaptively integrate and highlight the most discriminative features extracted from multiple color spaces. These results confirmed that the network outperformed other methods in both visual quality and quantitative metrics [15]. Zhang et al. found that the traditional color perception and recognition methods for Cantonese embroidery images had poor three-dimensional color restoration. They introduced a discrete mathematical model to design a new color perception and recognition method for Cantonese embroidery images. These results confirmed that the color pixel image curve of this method had 800 pixels for each color, and the color pixel image curve distribution was the most dense, with a high color restoration degree [16]. To deal with the encryption of RGB color images, MA Tahiri et al. used 3D fractional order modified Henon mapping and discrete fractional order Krawtchuk moments. Meanwhile, this team put forward a new method to optimize the proposed Henon map's parameters. These outcomes indicated this mixed algorithm's optimization efficiency compared to other metaheuristic methods [17].

In summary, numerous scholars have conducted extensive research on the application of SC-based improved K-Means and color image processing. On this basis, the improved K-Means based on SC is applied to the extraction of main colors in images for future development and evaluation of artistic color matching schemes. This paper hopes to provide effective color matching design for the field of art and design.

III. Artistic Color Matching Technology based on Silhouette Coefficient and Visual Perception

To deal with the problem that traditional algorithms cannot adaptively extract the main color of images, an improved K-Means based on SC is proposed, and a corresponding image main color adaptive extraction model is designed. Next, a color matching evaluation scheme combining VP and similarity measurement is introduced. Finally, a Pix2Pix based on visual aesthetics is designed to achieve intelligent color matching scheme formulation.

A. Adaptive Extraction and Evaluation of Image Main Colors based on Silhouette Coefficients

Adaptive extraction of image main colors is crucial in artistic color matching, which affects the normal operation of the entire color matching. In image processing, using K-Means for image main color extraction can efficiently separate the main colors of images in complex color spaces. The core of this algorithm is to minimize the sum of squared distances between each pixel in the class and its corresponding Cluster Center (CC) [18]. K-Means is a clustering algorithm based on Euclidean distance. The closer the Euclidean distance between two data points, the greater the similarity between these two. Given a sample set, this algorithm first randomly selects k initial CC and then measures their Euclidean distance from the remaining data. Subsequently, the CC with the closest Euclidean distance was determined and the target object was assigned to the corresponding cluster. Next, the average value of data within each cluster is calculated and used as the new CC. Finally, the running is iterated to the maximum to end [19-20]. The calculation of spatial Euclidean distance is represented by Eq. (1).

$$d(v,C_i)=\sqrt{\sum_{j=1}^{n}(v_l-C_{il})^2} \tag{1}$$

In Eq. (1), C_i refers to the i th CC. v refers to the target data. n refers to the data dimension. x_l refers to the l th attribute value of the target data. C_{il} refers to the l th attribute value of CC. The objective function used by K-Means for image main color extraction is represented by Eq. (2).

$$J=\sum_{n=1}^{M}\sum_{k=1}^{K}r_{nk}\left\|C(m)-\mu_k\right\|^2 \tag{2}$$

In Eq. (2), N refers to the total distance of all classification categories. M refers to the quantity of sample colors. K refers to the quantity of color classification. m is the pixel index of the image. k refers to the k th class color. r_{nk} refers to two components. μ_k refers to CC for the k th class color. Given the unique color information of each image, the dominant colors and their quantities also vary. Traditional K-Means cannot adaptively extract the main color, which has certain limitations. Therefore, the study proposes SC to improve K-Means [21]. Fig. 1 shows an improved K-Means based on SC.

In this improved algorithm, each cluster corresponds to an SC value. The clusters corresponding to the maximum SC value are the optimal clusters for the sample. The distance within the cluster is represented by Eq. (3).

$$a(i)=\frac{1}{|C_i-1|}\sum_{j\in C_i,i\neq j}d(i,j) \tag{3}$$

In Eq. (3), $a(i)$ refers to the intra cluster distance, which is the average Euclidean distance between the current pixel and all pixels in the cluster. $d(i,j)$ refers to the spatial Euclidean distance between the targets i and j. The distance between clusters is represented by Eq. (4).

$$b(i)=\min_{k\neq i}\frac{1}{|C_k|}\sum_{j\in C_k}d(i,j) \tag{4}$$

The SC value of the target sample i is calculated using Eq. (5).

$$s(i)=\begin{cases}1-\dfrac{a(i)}{b(i)}, & a(i)<b(i)\\ 0, & a(i)=b(i)\\ \dfrac{b(i)}{a(i)}-1, & a(i)>b(i)\end{cases} \tag{5}$$

In Eq. (5), $s(i)$ refers to the SC value, with a range of [-1, 1], representing the best and worst clustering effects, respectively. The overall SC value of the cluster is represented by Eq. (6).

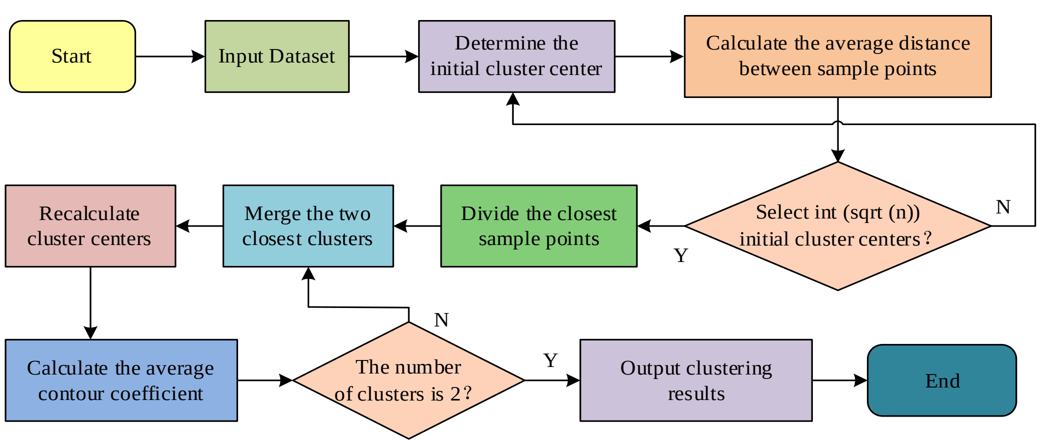


Fig. 1. The process of improving K-Means algorithm based on silhouette coefficients.

$$s = \frac{\sum_{i=1}^{N} s(i)}{N} \tag{6}$$

In Eq. (6), N represents the total target sample. At present, the evaluation methods for main color extraction mainly include questionnaire survey or manual visual inspection, which can effectively combine VP to evaluate the results of main color extraction, but often require a lot of user time. To address these shortcomings, a comprehensive evaluation method is proposed, which utilizes the main color image features and pairwise image similarity to evaluate the effectiveness of main color extraction. The main color image refers to the image obtained by replacing the color of each pixel in the source image with the target main color. To construct a main color image, the first step is to convert the image from the RGB color space to the Lab color space. Next, each pixel of the image was traversed and the extracted main color Lab values are input. Then, the Lab color difference between each pixel's color and all target main colors is calculated separately. The Lab value corresponding to the minimum color difference is used to replace the Lab value of the current pixel. Finally, the main color images are output. After obtaining the main color images corresponding to each sample, the Peak Signal-to-Noise Ratio (PSNR) and structural similarity between paired images are calculated using similarity measurement method. This method is mainly used to characterize the effectiveness of main color extraction [22]. The objective quantitative evaluation indicators are represented by Eq. (7).

$$\begin{cases} E_1 = PSNR(I_0, I_1) \\ E_2 = SSIM(I_0, I_1) \end{cases} \tag{7}$$

In Eq. (7), E refers to the quantitative evaluation indicator. I_0 refers to the source image. I_1 is the main color image. On this basis, qualitative analysis of the main color results extracted by various methods is conducted by combining subjective sensory engineering experiments. Table I shows the five-point scale of perceptual engineering.

TABLE I. FIVE-POINT SCALE OF PERCEPTUAL ENGINEERING

Number	Effect description	Rating range
I	Extremely poor	0-1
II	Poor	1-2
III	General	2-3
IV	Better	3-4
V	Perfect	4-5

Fig. 2 shows the adaptive extraction of image main colors based on SC.

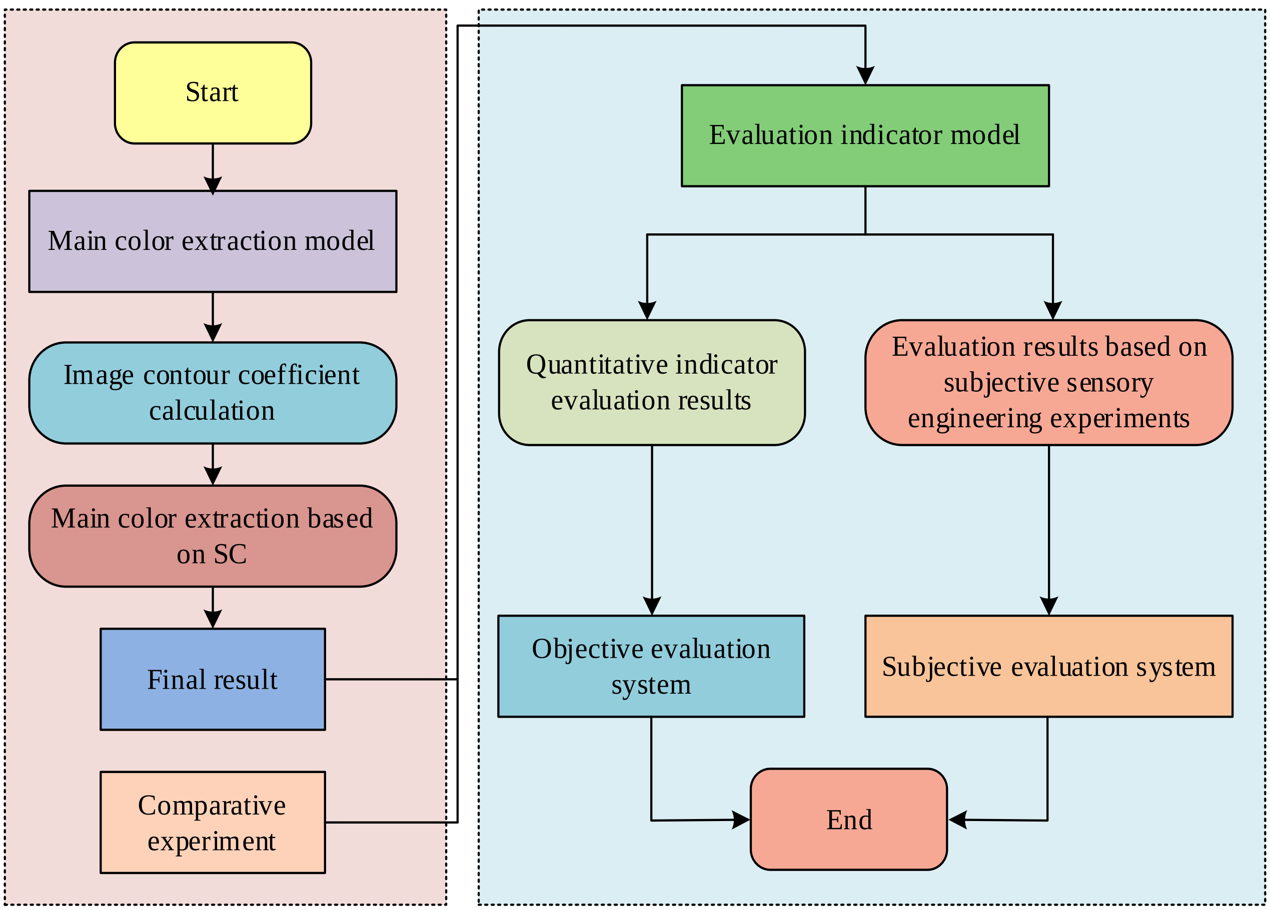


Fig. 2. Adaptive extraction process of image main color based on silhouette coefficients.

B. Color Matching Evaluation Based on Similarity Measurement and Visual Perception

Traditional methods for measuring the similarity of color matching schemes fail to take into account the different orientations of each color block in the palette, resulting in inaccurate color difference calculation for paired palettes [23]. To scientifically and effectively describe the similarity between color palettes composed of multiple color combinations, a minimum color difference model combining position information is proposed. This method mainly adopts the middle color palette strategy to avoid changes in the calculation results caused by differences in position information. The specific calculation steps include six steps. Firstly, it is assumed that the number of main colors in both the source and target palettes is u. The source palette is labeled as P1, and all colors in P1 are traversed and their Lab values are calculated. The target color palette is marked as P2, and all colors in P2 are traversed and their Lab values are calculated. In the second step, the color difference is calculated for the first color of the source palette P1 using all the colors in the target palette P2, and the results are recorded with the corresponding color in P2. The color block corresponding to the minimum value of the calculation result is set as the first color of the middle palette P3. In the third step, the first two stages are repeated, using all colors of P2 to calculate the color difference for the remaining u-1 colors of P1 one by one. Based on the first step, the remaining colors of the middle palette P3 are determined, and the complete middle palette P3 is ultimately obtained. In the fourth step, the average color difference between the source palette P1 and the middle palette P3 is calculated. Fig. 3 shows the specific calculation.

In the fifth step, for the u colors in the source palette P1, the color difference is calculated for each color in the target palette P2, and then a new intermediate auxiliary palette is obtained according to the processing method in the second step, denoted as P4. In the sixth step, the average color difference between P2 and P4 is calculated using the method in the third step. Fig. 4 shows the specific calculation.

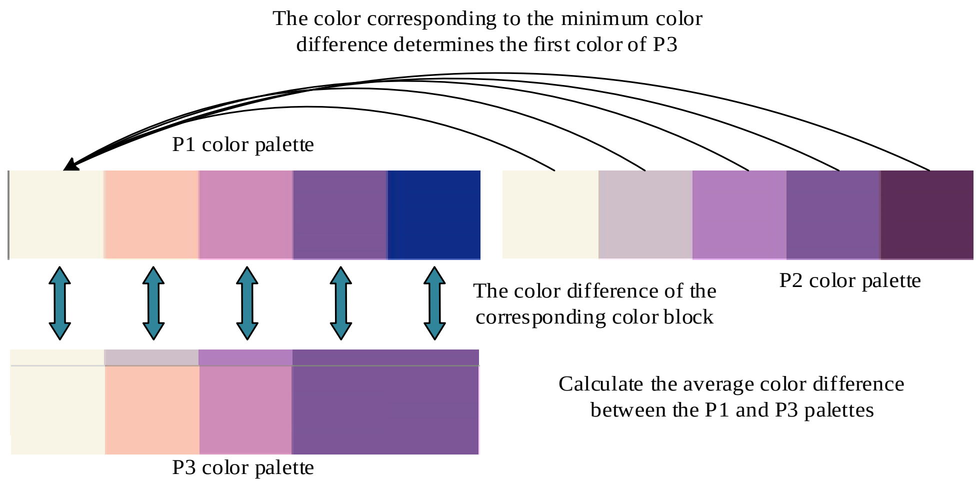


Fig. 3. Calculation method for middle auxiliary color palette P3.

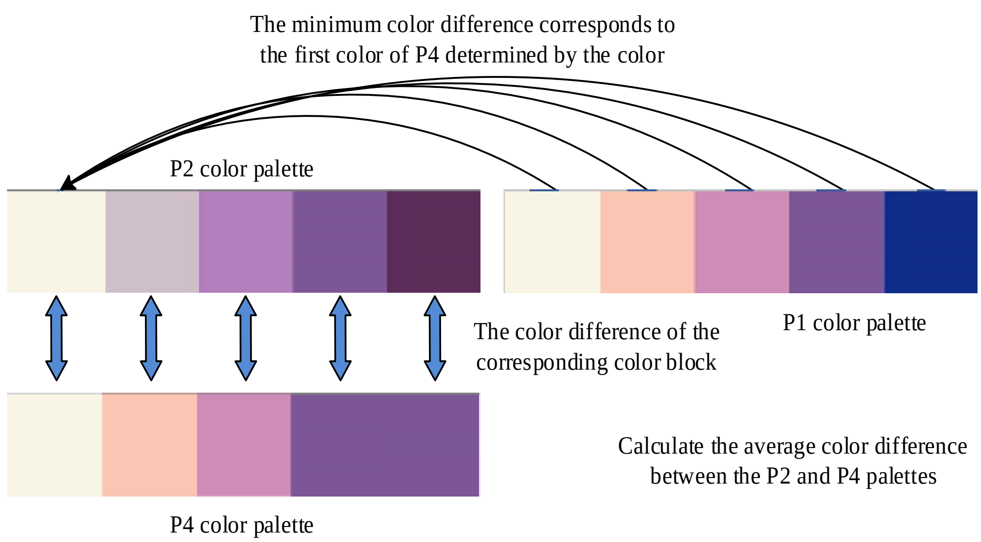


Fig. 4. Calculation method for intermediate auxiliary color palette P4.

Color is a comprehensive effect of the complex physiological and psychological reactions produced by the human visual system under the influence of light. The evaluation of different color schemes should take into account both physical properties and physiological and psychological reactions [24-25]. Therefore, eye tracking technology is used in this study to compare the color matching scheme generated by this intelligent color recommendation system with the original scheme. Based on computational aesthetics and experimental aesthetics, an aesthetic dimension evaluation is conducted on the color matching scheme. Firstly, eye tracking stimulation materials are constructed using the original sample images and the re-colored images. Subsequently, an eye tracking experiment is conducted, recording the basic information of the subjects and each participant's eye movement behavior measure for each group of image. Then the mean first fixed time, mean number of fixed points, and mean fixed time for each sample are calculated. Subsequently, the eye movement behavior indicators of each sample are normalized and these three indicators are weighted and fused. The longer the first fixed time, the lower the attractiveness. Therefore, it is necessary to use the normalized theoretical maximum value minus the normalized first fixed time. Finally, the VP data of the heavily colored image are calculated in proportion to its corresponding source image, which is the VP measure.

C. Intelligent Artistic Color Matching based on Visual Aesthetics

Due to the susceptibility of the human visual system to more visually attractive objects, eye movement behavior measurement is considered an important technical means in the field of quantitative visual aesthetics [26]. In quantifying visual aesthetics using eye movement behavior measurement, the three most critical eye movement behavior indicators include first fixed time, mean number of fixed points, and mean fixed time. These indicators reveal the characteristics of test images or videos from three aspects: visual comfort, attractiveness, and impact. The calculation of mean fixed time in eye tracking experiments is represented by Eq. (8).

$$T = \frac{\sum_{q=1}^{Q} T(AOI)}{Q} \tag{8}$$

In Eq. (8), Q refers to the number of test images. $T(AOI)$ refers to the testing time for dividing regions of interest in eye movement experiments. In visual search, more emphasis is usually placed on targets that have visual aesthetic appeal. The corresponding eye movement behavior indicators include first fixed time, mean number of fixed points, and mean fixed time. Eye movement behavior indicators are utilized to construct a visual aesthetic data stream, and three eye movement behavior indicators are subjected to normalization preprocessing. The mean number of fixed points and mean fixed time are positively correlated with the level of visual aesthetic preference [27-28]. The processing method for mean fixed time is represented by Eq. (9).

$$h' = \frac{h - h_{min}}{h_{max} - h_{min}} \tag{9}$$

In Eq. (9), h refers to the mean fixed time data that currently requires normalization preprocessing. h and h_{max} refer to the mean fixed time's maximum and minimum values, respectively. The first fixed time measurement index is negatively correlated with the visual aesthetic preference in interactive tasks, represented by Eq. (10).

$$g' = 1 - \frac{g - g_{min}}{g_{max} - g_{min}} \tag{10}$$

In Eq. (10), g refers to the first fixed time data that currently requires normalization preprocessing. g_{max} and g_{min} represent the first fixed time's maximum and minimum values, respectively. The visual aesthetic parameter is represented by Eq. (11).

$$W = 10 \cdot (\alpha h' + \beta q' + \gamma g') \tag{11}$$

In Eq. (11), W refers to the visual aesthetic parameter that integrates three different eye movement behavior data, with a range of values between [0, 10]. α, β, and γ represent the weights of mean fixed time, mean number of fixed points, and first fixed time, respectively. The image translation model Pix2Pix is a representation of conditional generative adversarial networks in image translation tasks. This model consists of U-Net and Markov discriminator as conditional generative adversarial networks for generator and discriminator, respectively [29]. This generator's input is a real sample image, and the output is a generated image. The discriminator needs to determine the authenticity of the output image of the generator, so its input is a paired image composed of the generated and real images. Fig. 5 shows the structure of Pix2Pix.

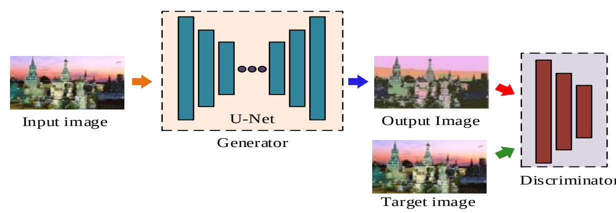


Fig. 5. The network structure of Pix2Pix.

Pix2Pix introduces $L_1 Loss$ to judge the global image based on the conditional generative adversarial network, represented by Eq. (12).

$$G^* = \arg\min_{G}\max_{D} L_{CGAN}(G,D) + \lambda L_{L_1}(G) \tag{12}$$

In Eq. (12), G refers to the generator. D refers to the discriminator. $L_{L_1}(G)$ refers to $L_1 Loss$. λ refers to the weight of $L_1 Loss$. $L_1 Loss$ is represented by Eq. (13).

$$L_{L_1}(G) = E_{x,y,z}\left[\left\| y - G(x,z) \right\|_1\right] \tag{13}$$

In Eq. (13), x refers to the true sample. y refers to conditional probability. z refers to random noise. A visual aesthetic evaluation model for color palettes is constructed using visual aesthetic parameters and image main color palettes.

This evaluation model mainly scores the input color palette and optimizes the loss function of the Pix2Pix backbone network using the score values [30]. The network model used is SE Inception V3, which compresses image features through global average pooling and scores the probability distribution of image aesthetic quality. This model mainly uses the real visual aesthetic parameters of the palette as the corresponding palette labels to train the rating model. Fig. 6 is a visual aesthetic evaluation model for color palettes based on SE-Concept V3.

In updating the loss function of the backbone network Pix2Pix, the aesthetic loss function is effectively optimized by introducing a palette aesthetic score, represented by Eq. (14).

$$S(G) = 10 - Score \tag{14}$$

In Eq. (14), $Score$ refers to the rating of the visual aesthetic rating model for the color palette, with 10 points being the maximum rating. Fig. 7 is an intelligent color matching algorithm based on visual aesthetics.

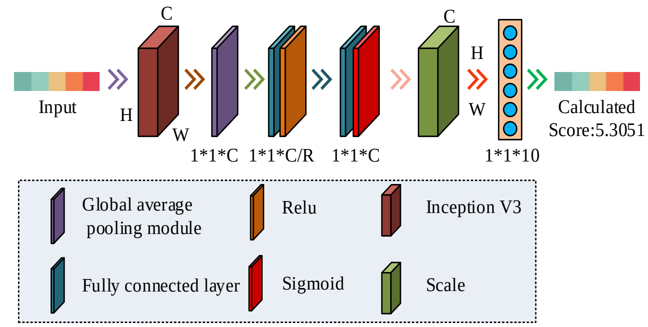


Fig. 6. A visual aesthetic evaluation model for color palettes based on SE-Concept V3.

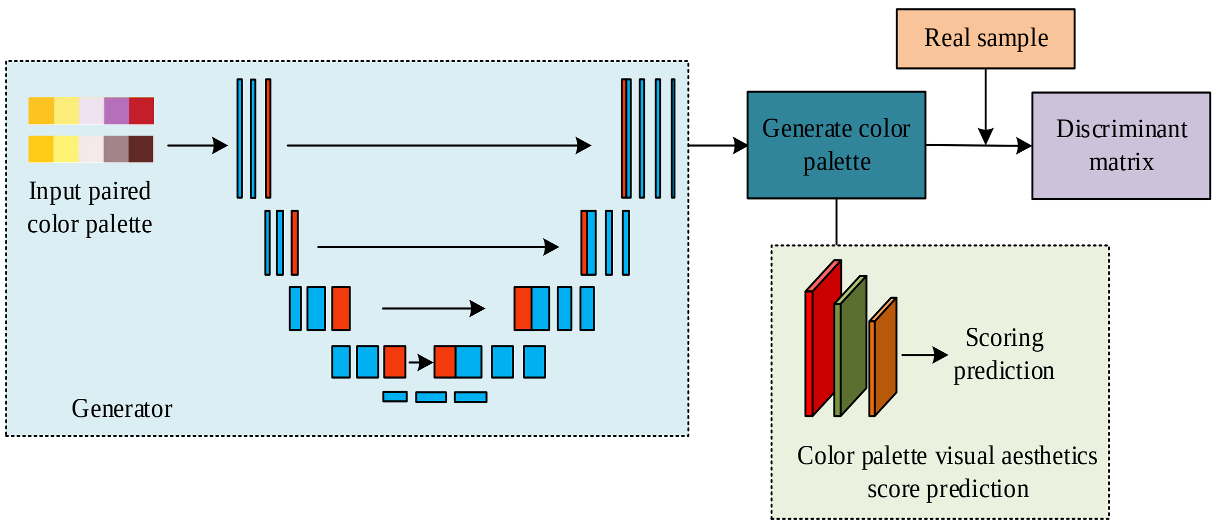


Fig. 7. A network model of intelligent color matching algorithm based on visual aesthetics.

IV. Experimental Analysis of Artistic Color Matching Based on Silhouette Coefficients and Visual Perception

Firstly, the study validated the effectiveness of the SC-based adaptive extraction method for image main colors and evaluated the main color extraction method from both subjective and objective perspectives. Subsequently, the effectiveness of the aesthetic evaluation model for color palettes was verified. Finally, experimental analysis was conducted on Pix2Pix based on visual aesthetics.

A. *Adaptive Extraction and Analysis of Image Main Colors based on Silhouette Coefficients*

To validate the SC based image main color adaptive extraction method's effectiveness, traditional K-Means, Median cut, and Octree algorithms were selected for comparison in main color extraction. To ensure the scientific and stable nature of the experiment, 58 sample images were collected from three types of samples: complex natural scenes, animal and human images. Table II shows the relevant parameters.

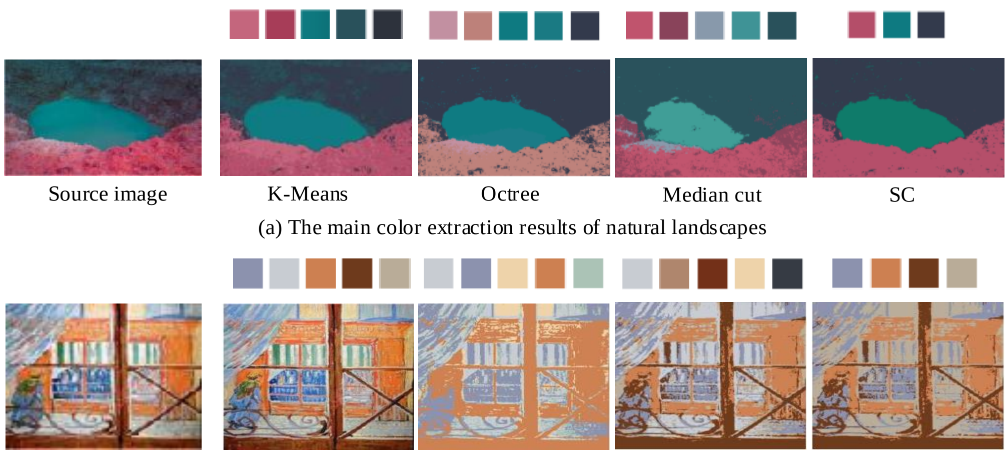
TABLE II. Relevant Parameters

Equipment and environment	Model number
Operating system	Windows10
Processor	ntel(R)Xeon(R)E5-2678V3
Graphics calculation card	NVIDIA GeForce GTX 1080 Ti
Develop compilation tools	Visual Studio Code
Server	Ultrascopic 7048 GR-TR
Algorithm framework dependencies	Numpy、OpenCV
Memory	Samsung 32G 2RX4 2400T
Algorithm development environment	Python 3.7

Firstly, the study evaluated the main color extraction methods from a subjective visual perspective, selecting natural scenes, animal and human images for main color extraction. Meanwhile, corresponding main color images were constructed based on the extracted results. Fig. 8 shows the main color extraction and image generation results of some natural scenes. In Fig. 8 (a), among the main color extraction results of natural landscapes, the image based on SC had the best main color extraction result. This method only characterized the color scheme of the entire lake with three main colors. In Fig. 8 (b), the image main color extraction method based on SC only characterized the color scheme of the entire scene with four main colors. Meanwhile, the main color extracted by SC matched better with the source image, and Octree had the lowest matching degree. In summary, compared to other methods, SC was more effective in extracting representative colors.

Fig. 9 shows the main color extraction and image generation results of some animals and characters. In Fig. 9 (a), in the main color extraction of animal scenes, the image main color extraction method based on SC extracted the three most representative colors of the source image. In Fig. 9 (b), the image main color extraction method based on SC only characterized the color scheme of the entire character scene with four main colors. Compared with other methods, the proposed method extracted colors that matched better with the source image. Octree had the worst extraction performance.

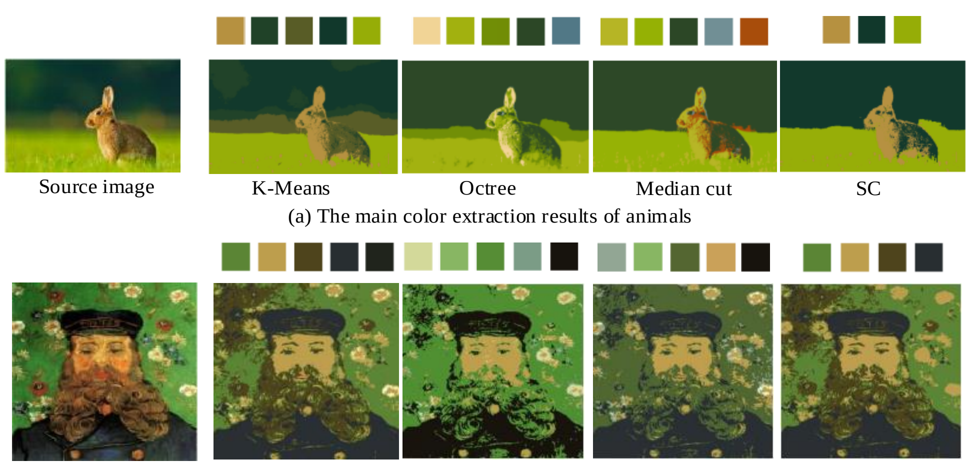
The study continued to objectively evaluate the main color extraction results of various methods. The evaluation indicators include the structural similarity and PSNR between the main color and source images. Fig. 10 shows the structural similarity and PSNR measured by various methods for 58 images. In Fig. 10 (a), the structural similarity of the main color image generated by SC was better than other methods. The maximum structural similarity of this method was 0.675, with an average of 0.663. The main color images generated by Octree had the lowest structural similarity, with an average of only 0.617. In Fig. 10 (b), the PSNR of the main color image generated by SC was also optimal. The PSNR of this method reached a maximum of 21.49 dB, with an average of 21.05 dB. The PSNR of K-Means ranks second, with an average of 20.54 dB. Therefore, the main color extraction method based on SC was optimal, and the generated main color image had a higher similarity with the source image.



(a) The main color extraction results of natural landscapes

(b) The main color extraction results of still life

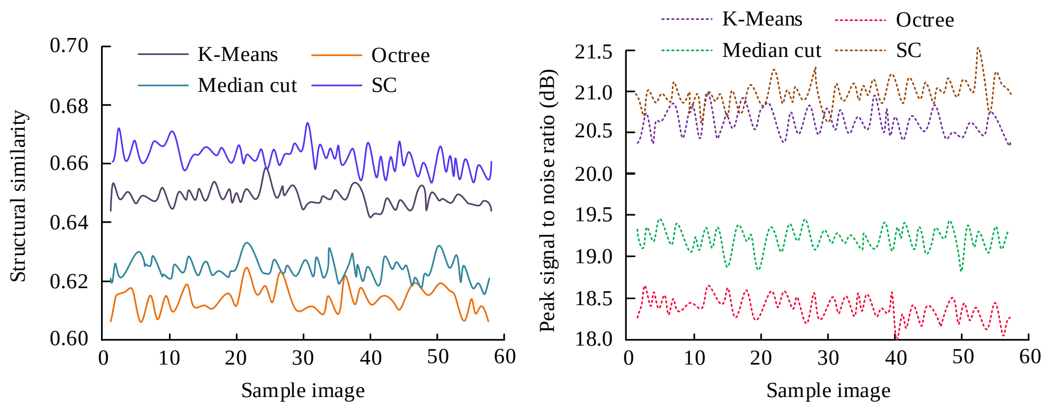
Fig. 8. Partial natural scene main color extraction and image generation results.



(a) The main color extraction results of animals

(b) The main color extraction results of characters

Fig. 9. Main color extraction and image generation results of some animals and figures.

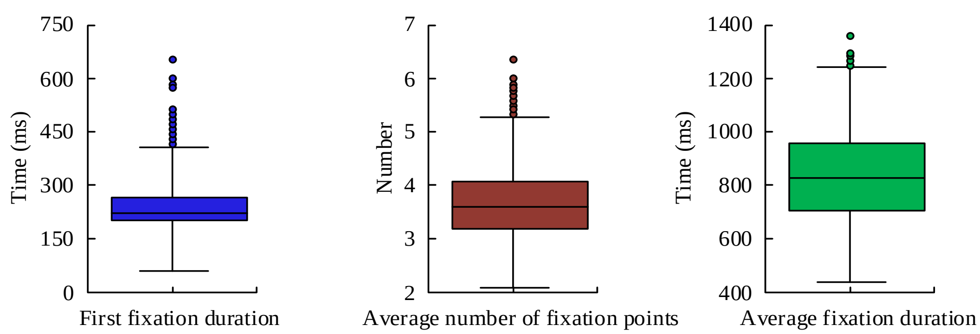


(a) The structural similarity of various methods (b) Peak signal to noise ratio of each method

Fig. 10. Structural similarity and peak signal-to-noise ratio of different methods.

B. Analysis of Intelligent Art Color Matching Evaluation Based on Visual Aesthetics

To verify the effectiveness of the aesthetic evaluation model for color palettes, a dataset of 2800 color palettes was collected in Adobe Color CC. The study conducted eye tracking experiments on each dataset to obtain its visual aesthetic parameters. These samples were randomly divided into 700 groups of eye tracking stimuli as experimental data for eye tracking. The study invited 50 undergraduate students to participate in eye tracking experiments. All participants had good vision, no color blindness, weak color, and other physical effects that affected their vision. The playback time of each group of stimulating materials on the screen was 5,000 milliseconds. After the experiment, the average eye movement data of each designated area of interest was obtained based on the eye tracking data of the subjects. Fig. 11 is a box plot of the first fixed time, mean fixed time, and mean number of fixed points. In Fig. 11 (a), the mean first fixed time of the experimenter was 182.58 ms. In Fig. 11 (b), the mean number of fixeds of the subjects was 3.5. In Fig. 11 (c), the mean fixed time of the subjects was 826.36 ms. There were outliers in the upper limits of the first fixed time, mean fixed time, and mean number of fixed points. This is due to the influence of psychological and physical factors on eye movement behavior, resulting in fluctuation.



(a) Box plot of first fixation time (b) Box plot of average number of fixation points (c) Box plot of average fixation duration

Fig. 11. Box plots of three eye movement fixed indicators.

The study continued to use a color palette visual aesthetic evaluation model to evaluate the visual aesthetics of sample images. To reduce the interference of outliers on the results, the upper and lower limits of these data were used as the maximum and minimum values for normalization processing. Fig. 12 shows the true values of visual aesthetic parameters and the visual aesthetic evaluation model scores of the color palette for 20 samples. The data predicted by the color palette visual aesthetic evaluation model were relatively close to the actual values. The predicted data of sample 8's model only differed by 0.05 from the actual value, indicating that the designed color palette visual aesthetic evaluation model had good predictive performance.

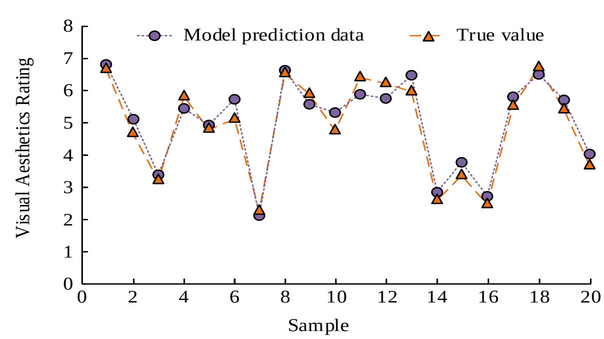
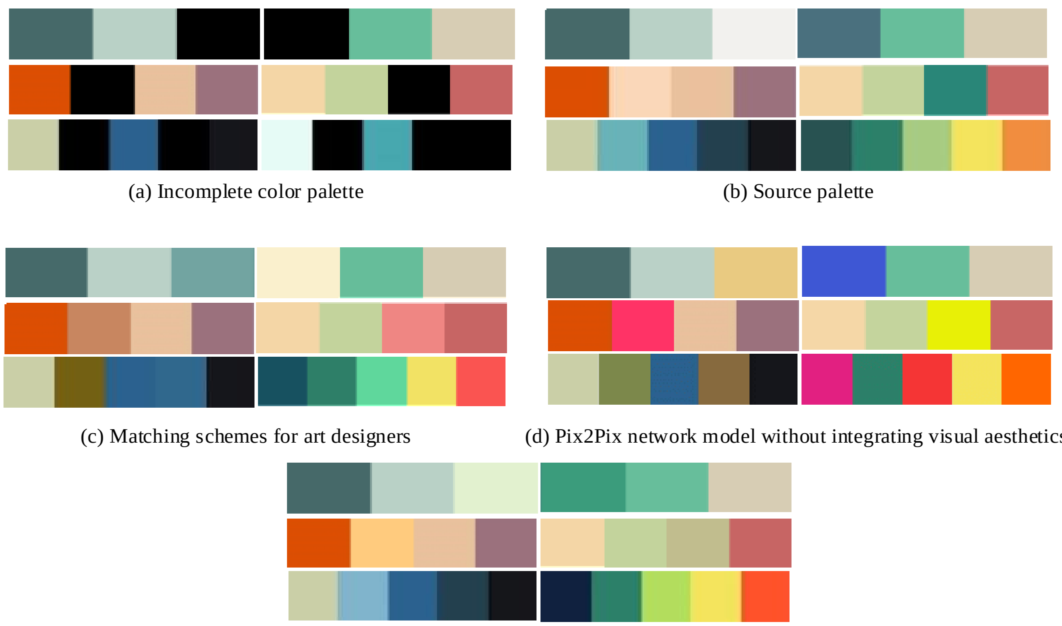


Fig. 12. Color palette visual aesthetic evaluation model rating results.

To verify the effectiveness of Pix2Pix based on visual aesthetics, this study selected art and design designers with matching schemes and Pix2Pix without integrating visual aesthetics for experimental comparison. 1680 sample images were extracted in the experiment, and the main color palette was obtained by extracting the main color from the sample set. Due to Pix2Pix requiring input into paired color palettes, the main color palette was fragmented to construct a paired color palette. To ensure consistent size of the paired color palettes, black replaced the missing color blocks, which were the colors that needed to be matched. The palette samples were divided into training and testing sets in an 8:2 ratio, with a total of 1344 training samples and 336 testing samples. In addition, the Adam optimizer was used for training, with an initial learning rate of 0.0002, an initial momentum term of 0.5, a network epoch count of 200, Batch_Size of 8, an initial parameter training weight of 100 for $L_1 Loss$, and an initial weight of 0.1 for $S(G)$. Fig. 13 shows the results of some color matching tests. Fig. 13 (a) means the incomplete palette, and Fig. 13 (b) refers to the source palette. In Fig. 13 (c), the complete color palette matched by the designer exhibited high harmony in color attributes and superior visual effects, but its design took a long time. In Fig. 13 (d), the color palette of Pix2Pix without incorporating visual aesthetics was vivid, but the visual effect was relatively cluttered. The reason is that Pix2Pix, which does not integrate visual aesthetics, mainly focuses on the sample distribution of all training images, but fails to effectively combine human eye VP, resulting in poor harmony in the generated color matching scheme. In Fig. 13 (e), the designed model had a visual effect similar to that of the source palette and a professional designer in terms of color matching. For the real-time performance of color matching, due to the sample size of color matching images being 32×1, the computational complexity of the designed model basically met the real-time requirements.



(a) Incomplete color palette (b) Source palette

(c) Matching schemes for art designers (d) Pix2Pix network model without integrating visual aesthetics

(e) Pix2Pix network model based on visual aesthetics

Fig. 13. Partial color matching test results.

Further research is conducted to verify the computational efficiency of the Pix2Pix network model based on visual aesthetics. Compare with K-means clustering algorithm, Gaussian Mixture Model (GMM), and Generative Adversarial Networks (GANs). The experiment measures CPU time, which includes the time required for palette generation and pixel mapping stages. The efficiency of color matching generation for sample images by each model is shown in Fig. 14. As shown in Fig. 14, the average color matching time of the Pix2Pix network model based on visual aesthetics proposed in the study is only 13.75ms. The average time consumption of K-means clustering algorithm model, GMM model, and GANs model is as high as 135.67ms, 90.36ms, and 54.69ms, respectively, which is significantly higher than the algorithm proposed in the study. The color matching scheme of the Pix2Pix network model based on visual aesthetics can complete tasks at a faster speed and has efficient computational performance.

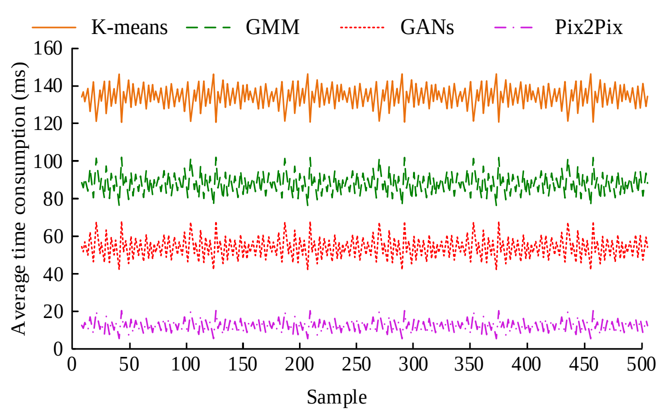
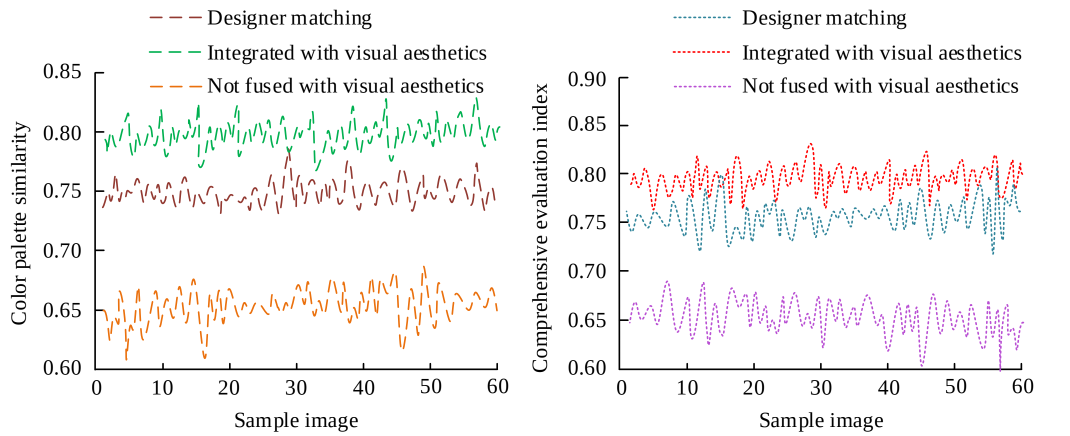


Fig. 14. The color matching time of each model for sample images.

To scientifically analyze and compare the experimental results, the color matching evaluation method was used to evaluate the color matching effects of each method. 60 samples were selected for color matching effect analysis, with evaluation indicators including similarity of color palettes and comprehensive evaluation index. Fig. 15 shows the evaluation indicators for three methods. In Fig. 15 (a), the average color palette similarity of the Pix2Pix design scheme based on visual aesthetics was 0.807, which was 0.043 higher than the scheme designed by the designer. In Fig. 15 (b), the average comprehensive evaluation index of the proposed method was 0.798, which was 0.158 higher than the Pix2Pix without integrating visual aesthetics. Pix2Pix based on visual aesthetics had a more significant color matching effect.



(a) Color palette similarity across different methods (b) Comprehensive evaluation index of different methods

Fig. 15. Color palette similarity and comprehensive evaluation index of three methods.

The study then combined perceptual engineering experiments to conduct psychological and physical analysis on the effect of color matching. Twenty sample sets were selected for the experiment, and 10 participants were invited to subjectively evaluate the test results. The average score was taken as the final result. Fig. 16 shows the average scores of all participants on the four color matching schemes. The subjective score of the color palette generated by the proposed method reached a maximum of 4.25, with an average of 3.81, which is 027 and 0.39 higher than the Pix2Pix without integrating visual aesthetics, respectively. Meanwhile, the matching effect of this method was closer to the matching scheme designed by professional designers and the color matching effect of the source palette. The proposed method also performed well in subjective evaluation.

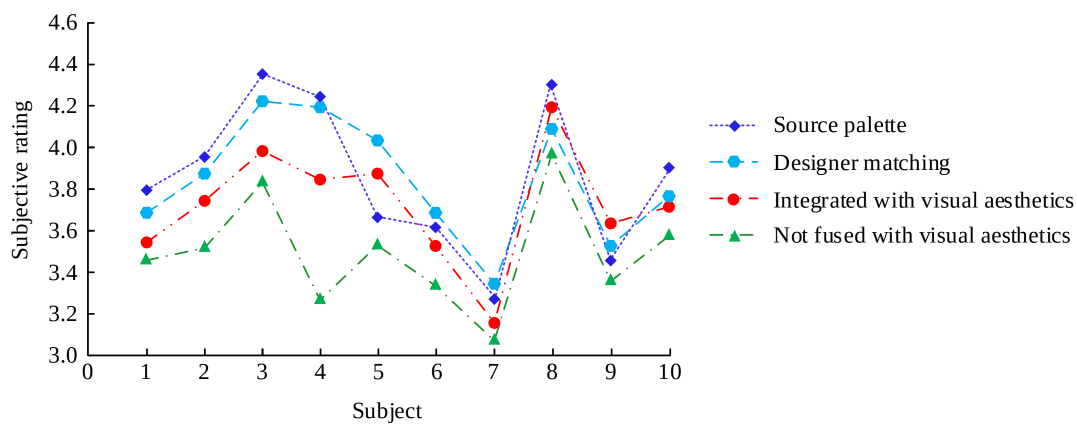


Fig. 16. The average rating results of four color matching schemes.

V. Discussion

Currently, in the field of color science, issues related to color matching include image main color extraction, color matching methods, and the selection of color matching evaluation methods. The purpose of the study is to achieve effective adaptive color extraction, combine human visual perception to generate the best color matching scheme, and conduct a more comprehensive and multidimensional analysis of the color matching scheme. For this purpose, research has successively proposed a SC based adaptive extraction method for image main colors, a Pix2Pix network model based on visual aesthetics, and a multi-dimensional color matching evaluation method combining visual perception and similarity measurement. The results showed that in the main color extraction results of natural landscapes, the image main color extraction based on SC was the best, and this method only represented the color system of the entire lake with three main colors. Meanwhile, this method only characterizes the color scheme of animal and human image scenes with four main colors. At the same time, it can be seen that the main color extracted by SC matches better with the source image. Related studies have shown that the image main color extraction method based on SC can represent the color scheme of the entire image with the most concise colors, which is concise and efficient [31]. In the effectiveness verification of the Pix2Pix network model based on visual aesthetics, the complete color palette matched by the designer showed high harmony in color attributes and superior visual effects, but the design process was time-consuming. The Pix2Pix network model, which does not integrate visual aesthetics, is paired with a palette of bright colors, but the visual effect is relatively messy. The model designed for research has a visual effect similar to that of the source color palette and professional designers in terms of color matching. The experimental results of K. Qiu et al. show that integrating visual aesthetics can effectively improve the color matching effect of network models, and has lower computational complexity [32]. In addition, in the experimental analysis of color matching effect evaluation, the subjective score of the color palette generated by the proposed method reached the highest of 4.25, with an average of 3.81. Compared with the Pix2Pix network model without integrating visual aesthetics, it improved by 027 and 0.39, respectively. Meanwhile, the matching effect of this method is closer to the matching scheme designed by professional designers and the color matching effect of the source palette. The method proposed by the research institute also performs well in subjective evaluation. And the data predicted by the color palette visual aesthetic evaluation model is relatively close to the actual values. Among them, the predicted data of sample 8's model only differs by 0.05 from the true value, indicating that the color palette visual aesthetic evaluation model designed in the study has good predictive performance.

VI. Conclusion

Currently, there are challenges in color matching, such as cumbersome design processes and severe homogenization. To achieve intelligent color matching scheme design and development, a computer-aided design method based on image processing and VP was introduced. For image main color extraction, a SC-based image main color adaptive extraction model was introduced. Meanwhile, a color matching evaluation method combining VP and similarity measurement was introduced. In addition, a color matching model based on visual aesthetics was designed. These results confirmed that in the main color extraction of natural landscapes, animals, and people, the image main color extraction based on SC had the best results. Meanwhile, the main color extracted by this method was more closely matched with the source image, which more effectively extracted representative colors. In the validation of the color palette aesthetic evaluation model, the data predicted by the color palette visual aesthetic evaluation model were relatively close to the actual values. The model prediction data of sample 8 only differ by 0.05 from the true value. In the validation of Pix2Pix based on visual aesthetics, the designed model achieved visual effects similar to those of the source color palette and professional designers in color matching. Moreover, its computational complexity basically met the real-time requirements. In addition, the subjective score of the color palette generated by the proposed method reached a maximum of 4.25, with a mean of 3.81, which was 027 and 0.39 higher than the Pix2Pix without integrating visual aesthetics, respectively. The average similarity of color palettes in the design scheme of the Pix2Pix network model based on visual aesthetics is as high as 0.807, which is 0.043 higher than the scheme designed by the designer. And the average comprehensive evaluation index of the method proposed by the research institute is as high as 0.798, which is 0.158 higher than the Pix2Pix network model without integrating visual aesthetics. Meanwhile, the matching effect of this method was closer to the matching scheme designed by professional designers and the color matching effect of the source palette. Overall, the proposed color matching computer-aided design method achieved significant results and had significant practical application value. However, the designed SC-based image main color adaptive extraction model has a high time complexity. Subsequent research can further improve algorithm efficiency while ensuring adaptive conditions are met.

Fundings

The research is supported by: The Philosophy and Social Science Research Project in Universities of Anhui Province in 2023, (NO. 2023AH040158); Horizontal scientific research project of Hefei Normal University in 2023, (NO. HXXM2023018); The Project of Supporting Outstanding Young Talents in Universities of Anhui Province in 2019, (NO. gxyq2019065).

References

[1] S. Lee, M. Cherel, S. Gougeon, E. Jeong, J. M. Lim, and S. G. Park, "Identifying patterns behind the changes in skin pores using 3-dimensional measurements and K-means clustering," SKIN RES TECHNOL, vol. 28, no. 1, pp. 3-9, Aug. 2022, DOI: 10.1111/srt.13082.

[2] L. Y. Hsu and H. -T. Hu, "QDCT-Based Blind Color Image Watermarking With Aid of GWO and DnCNN for Performance Improvement," Access, vol. 9, no. 1, pp. 155138-155152, May. 2021, DOI: 10.1109/ACCESS.2021.3127917.

[3] S. Wu, X. Hu, W. Zheng, C. He, G. Zhang, H. Zhang, and X. Wang, "Effects of reservoir water level fluctuations and rainfall on a landslide by two-way ANOVA and K-means clustering," B ENG GEOL ENVIRON, vol. 80, no. 7, pp. 5405-5421, May. 2021, DOI: 10.1007/s10064-021-02273-8.

[4] J. Chen, and M. K. Ng, "Color image inpainting via robust pure quaternion matrix completion: Error bound and weighted loss," SIIMS, vol. 15, no. 3, pp. 1469-1498, Apr. 2022, DOI: 10.1137/22M1476897.

[5] W. Zheng, L. Yan, C. Gou and F. Y. Wang, "An ACP-Based Parallel Approach for Color Image Encryption Using Redundant Blocks," T-CYB, vol. 52, no. 12, pp. 13181-13196, Dec. 2022, DOI: 10.1109/TCYB.2021.3105568.

[6] K. M. Hosny, S. T. Kamal and M. M. Darwish, "A novel color image encryption based on fractional shifted Gegenbauer moments and 2D logistic-sine map," VISUAL COMPUT, vol. 39, no. 3, pp. 1027-1044, Jan. 2023, DOI: 10.1007/s00371-021-02382-1.

[7] A. Yudistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering, "JAIR, vol. 1, no. 1, pp. 20-28, Mar. 2023, DOI: 10.58602/jaiti.v1i1.22.

[8] Y. Xiang, J. Hong and Z. Yang, "Slope-Based Shape Cluster Method for Smart Metering Load Profiles," IEEE T SMART GRID, vol. 11, no. 2, pp. 1809-1811, Mar. 2020, DOI: 10.1109/TSG.2020.2965801.

[9] M. Ben-Marzouk, S. Pelissier, G. Clerc, A. Sari and P. Venet, "Generation of a Real-Life Battery Usage Pattern for Electrical Vehicle Application and Aging Comparison With the WLTC Profile," TVT, vol. 70, no. 6, pp. 5618-5627, Jun. 2021, DOI: 10.1109/TVT.2021.3077671.

[10] K. Qu, B. Zou and J. Zhou, "Rapid environmental assessment in the South China Sea: Improved inversion of sound speed profile using remote sensing data," AOS, vol. 41, no. 7, pp. 78-83, Jul. 2022, DOI: 10.1007/s13131-022-2032-2.

[11] D. P. Agustino, I. G. Harsemadi and I. G. B. ABudaya, "Edutech Digital Start-Up Customer Profiling Based on RFM Data Model Using K-Means Clustering," ISJ, vol. 4, no. 3, pp. 724-736, Sept. 2022, DOI: 10.51519/journalisi.v4i3.322.

[12] F. Tabatabaian, E. Beyabanaki, P. Alirezaei and Epakchi S, "Visual and digital tooth shade selection methods, related effective factors and conditions, and their accuracy and precision: A literature review," J ESTHET RESTOR DENT, vol. 33, no. 8, pp. 1084-1104, Sept. 2021. DOI: 10.1111/jerd.12816.

[13] H. Zeng, J. Cai, L. Li, Z. Cao and L. Zhang, "Learning Image-Adaptive 3D Lookup Tables for High Performance Photo Enhancement in Real-Time," TPAMI, vol. 44, no. 4, pp. 2058-2073, 1 Apr. 2022, DOI: 10.1109/TPAMI.2020.3026740.

[14] D. Berman, D. Levy, S. Avidan and T. Treibitz, "Underwater Single Image Color Restoration Using Haze-Lines and a New Quantitative Dataset," TPAMI, vol. 43, no. 8, pp. 2822-2837, Aug. 2021, DOI: 10.1109/TPAMI.2020.2977624.

[15] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo and W. Ren, "Underwater Image Enhancement via Medium Transmission-Guided Multi-Color Space Embedding," T-IP, vol. 30, no. 1, pp. 4985-5000, Mar. 2021, DOI: 10.1109/TIP.2021.3076367.

[16] Y. Zhang and Q. Xiong, "Color perception and recognition method for Guangdong embroidery image based on discrete mathematical model," JIFS, vol. 40, no. 3, pp. 3887-3897, Mar. 2021, DOI: 10.3233/JIFS-191484.

[17] M. A. Tahiri, H. Karmouni, A. Bencherqui, A. Daoui, M. Sayyouri, H. Qjidaa and K. M. Hosny, "New color image encryption using hybrid optimization algorithm and Krawtchouk fractional transformations," VISUAL COMPUT, vol. 39, no. 12, pp. 6395-6420, Dec. 2023, DOI: 10.1007/s00371-022-02736-3.

[18] L. Zhang,M. Li, Y. Sun, X. Liu, B. Xu and L. Zhang, "Color matching design simulation platform based on collaborative collective intelligence," CCF TPAMI vol. 4, no. 1, pp. 61-75, Jan. 2022, DOI: 10.1007/s42486-022-00088-4.

[19] M. Yu, Z. Tang, X. Zhang, B. Zhong and X. Zhang, "Perceptual Hashing With Complementary Color Wavelet Transform and Compressed Sensing for Reduced-Reference Image Quality Assessment," IEEE T CIRC SYST VID, vol. 32, no. 11, pp. 7559-7574, Nov. 2022, DOI: 10.1109/TCSVT.2022.3190273.

[20] O. Tasar, S. L. Happy, Y. Tarabalka and P. Alliez, "ColorMapGAN: Unsupervised Domain Adaptation for Semantic Segmentation Using Color Mapping Generative Adversarial Networks," IEEE T GEOSCI REMOTE, vol. 58, no. 10, pp. 7178-7193, Oct. 2020, DOI: 10.1109/TGRS.2020.2980417.

[21] M. Abu-Faraj, Z. Alqadi, B. Al-Ahmad, K. Aldebei and B. Ali, "A Novel Approach to Extract Color Image Features using Image Thinning," AMIS, vol. 16, no. 5, pp. 665-672, Sept. 2022, DOI: 10.18576/amis/160501.

[22] S. K. Roy, S. Manna, T. Song and L. Bruzzone, "Attention-Based Adaptive Spectral–Spatial Kernel ResNet for Hyperspectral Image Classification," IEEE T GEOSCI REMOTE, vol. 59, no. 9, pp. 7831-7843, Sept. 2021, DOI: 10.1109/TGRS.2020.3043267.

[23] I. Konya, I. Shishido, Y. M. Ito and R. Yano, "Combination of minimum wiping pressure and number of wipings that can remove pseudo-skin dirt: A digital image color analysis," SKIN RES TECHNOL, vol. 26, no. 5, pp. 639-647, Mar. 2020, DOI: 10.1111/srt.12844.

[24] A. U. Rehman, A. Firdous, S. Iqbal, Z. Abbas, M. M. A. Shahid, H. Wang and F. Ullah, "A Color Image Encryption Algorithm Based on One Time Key, Chaos Theory, and Concept of Rotor Machine," Access, vol. 8, no. 1, pp. 172275-172295, Nov. 2020, DOI: 10.1109/ACCESS.2020.3024994.

[25] Q. Liu and L. Liu, "Color Image Encryption Algorithm Based on DNA Coding and Double Chaos System," Access, vol. 8, no. 1, pp. 83596-83610, Jun. 2020, DOI: 10.1109/ACCESS.2020.2991420.

[26] Z. Zhang, J. Tang, F. Zhang, H. Ni, J. Chen and Z. Huang, "Color Image Encryption Using 2D Sine-Cosine Coupling Map," Access, vol. 10, no. 1, pp. 67669-67685, May. 2022, DOI: 10.1109/ACCESS.2022.3185229.

[27] K. Bhosle and V. Musande, "Evaluation of Deep Learning CNN Model for Recognition of Devanagari Digit," Artif. Intell. Appl., vol. 1, no. 2, pp. 114-118, Feb, 2023, DOI: 10.47852/bonviewAIA3202441.

[28] X. Zeng, S. Tong, Y. Lu, L. Xu and Z. Huang, "Adaptive Medical Image Deep Color Perception Algorithm," Access, vol. 8, no. 1, pp. 56559-56571, Jul. 2020, DOI: 10.1109/ACCESS.2020.2982187.

[29] Q. Zhao, "Research on the application of local binary patterns based on color distance in image classification," MULTIMED TOOLS APPL, vol. 80, no. 18, pp. 27279-27298, May. 2021, DOI: 10.1007/s11042-021-10996-9.

[30] M. G. A. Malik, Z. Bashir, N. Iqbal and M. A. Imtiaz, "Color Image Encryption Algorithm Based on Hyper-Chaos and DNA Computing," Access, vol. 8, no. 1, pp. 88093-88107, Aug. 2020, DOI: 10.1109/ACCESS.2020.2990170.

[31] G. Sun, H. Fu, J. Ren, A. Zhang, J. Zabalza, X. Jia and H. Zhao, "SpaSSA: Superpixelwise Adaptive SSA for Unsupervised Spatial – Spectral Feature Extraction in Hyperspectral Image," T-CYB, vol. 52, no. 7, pp. 6158-6169, July. 2022, doi: 10.1109/TCYB.2021.3104100.

[32] K. Qiu, K. Bittkau, A. Lambertz, W. Duan, Z. Liang, H. Shen, and K. Ding, "The Impact of Reflectance Variation in Silicon Heterojunction Solar Cells and Modules on the Perception of Color Differences," IEEE J PHOTOVOLT, vol. 11, no. 2, pp. 306-311, Mar. 2021, doi: 10.1109/JPHOTOV.2020.3048240.