Semantic Analysis and Modeling of the Color of Stage Costume – A Case Study of "The Painting Journey"

Wenli Huang[1, 2, 3], Bowen Liu[1, 2, 3], Fang Sun[1, 2, 3], Shuang Wang[1, 2, 3, *]
[1]State Key Laboratory of Media Convergence and Communication, Communication University of China, 100024, China
[2]Key Laboratory of Acoustic Visual Technology and Intelligent Control System (Communication University of China), 100024, China
[3]Beijing Key Laboratory of Modern Entertainment Technology (Communication University of China), 100024, China
Email: wangshuang@cuc.edu.cn
[*]Corresponding author: Shuang Wang

***Abstract*—The creation of Chinese dance dramas plays an important role in the dissemination of Chinese culture in the context of pluralistic unity. As one of the outstanding poetry dramas in the past two years, "The Journey of a Legendary Landscape Painting" (JLLP) has rich national cultural connotations and aesthetic value in its stage costume colors. This article first studied the data extraction method of stage costume colors based on clustering and association rule mining, and constructed the color dataset for JLLP. On this basis, the color semantic mining and extraction including subjective and objective factors were conducted. Finally, combined with color semantics, the character personality of JLLP was further analyzed. The research results show that the color extraction method designed in this article can accurately preserve the color characteristics of stage costumes. In addition, the costume colors of JLLP play a good supporting role in shaping the character personality.**

Keywords—Stage Costume, K-means, Association Rules, color semantic, characteristic analysis

I. Introduction

In recent years, with the improvement of people's living standards, more and more people start focusing on aesthetic enjoyment, in which the stage art, with its feature of multiple artistic elements, bears the responsibility of spreading Chinese traditional culture and the creation of aesthetics. However, there are few studies on stage art combined with computer network technology. Color is one of the most intuitive elements of stage art, so the digital storage of stage costume colors and the study of the relationship between stage costume colors and character personality of them are of certain significance for the digital preservation of the Chinese traditional culture and for the research of whom works in related design fields.

The popular drama JLLP debuted in August, 2021, bringing a visual feast to the audiences. It originated from the ancient painting, "A Panorama Mountains and Rivers" (PMR), the colors of which are of great research value. This paper took its costume colors as the object, focused on the construction of stage costume color dataset, extracted the subjective and objective attribute of the color, analyzing the relationship between them and the character personality.

II. Construction of the Color Dataset

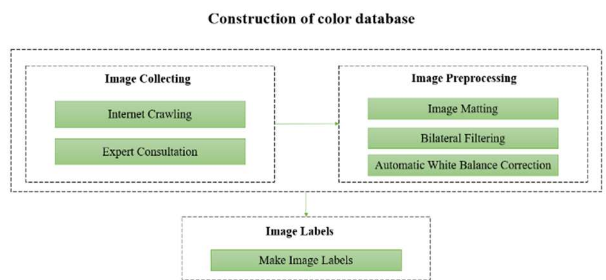


Fig. 1. The Flowchart of Color Dataset Construction

A. Image pre-processing

Datasets are the basis for digital image processing and analysis. As shown in Fig.1, we collected three types of images by Internet Crawling and expert consultation at first. With the help of professional image processing software, a total of 210 costume images were extracted from all images and stored in PNG format. Furthermore, we carried out bilateral filtering [1] and automatic white balance correction [2], reducing the influence of noise and uneven lighting conditions on the color.

In the JLLP, the scroll-exhibitor leads the audiences into the world of Ximeng through the chapters of "Scroll Unfolding, Seal Asking, Silk Singing, Stones Search, Penmanship Practice, Ink Milling, and Drawing". The costumes of the whole drama correspond to the following characters: the scroll-exhibitor, Ximeng, the seal-carver, the juan-weaver, the stone-polisher, the brush-maker, the ink-maker, the group-dancers, the teachers and students of the Hanlin Academy, and the female officials, etc. After completing the image preprocessing, we labeled the costume images according its scene, character and main color, which is helpful to analyze the color of each costume for different character. The data structure is shown in Fig. 2 Calculating the percentage of each costume and color, we obtained the proportion of all kinds of images. As is shown in the Fig. 3, white occupies a large proportion in all costumes although there exists 36% costume of the group-dancers.

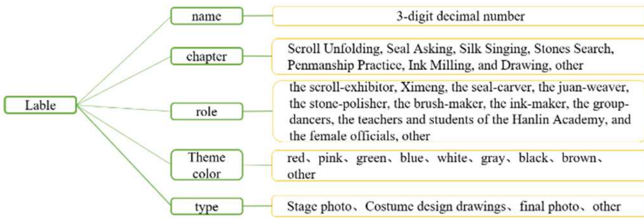


Fig. 2. Costume Data Tagging Structure

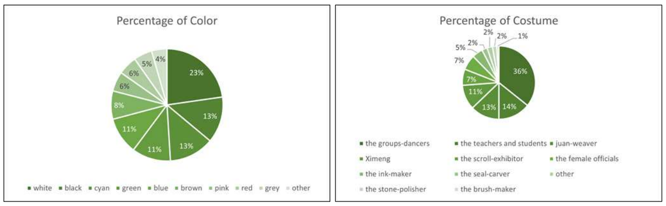


Fig. 3. Costume Percentage and Color Percentage

B. Theme color extraction

Color is the most intuitive impression the costumes bring to human. It can convey the emotion. Using computer network technology, we can realize the quantification of costumes theme color, such as theme color extraction based on clustering, color histogram, etc. Among them, K-means clustering algorithm (K-means) [3]is widely used in image theme color extraction as its high efficiency.

In this research, the costume theme color extraction was carried out in RGB color space. Since the transparent pixel points will be treated as white during clustering, they have been deleted first. In addition, there were shadow colors in some images. To weaken the influence of them during theme color extraction, each pixel whose brightness is less than 20 was deleted.

In the traditional k-means, the selection of K relies on manual selection, but usually the number of theme colors of each costume is different. That not only consumes a lot of manpower, but also requires rich personal experience, so we introduced the Elbow Method [4] to estimate the number of theme colors for each image, thus reducing the input to improve efficiency. The basic principle of it is to calculate the Sum of Square Error (SSE) of each pixel to the corresponding clustering centers. The smaller the SSE is, the more convergent the cluster centers are. In more detail, the steps are:

① Preset the value range of k, and traverse all k values for K-means;

② Calculate the SSE value of each k value;

③ With the increase of k value, the SEE value shows a decreasing trend until its decrease suddenly becomes smaller, at which point the k value is the most suitable. Finally, carry out the K-means again by this k value, to realize the K-means that best adapts the image to k.

We used a rectangular color palette to carry the theme colors of each costume, which contain not only color but also its ratio. Fig. 4 shows the theme color palettes of the group-dancer's costume and the Juan-weaver's costume, respectively.

Fig. 4. Group-dancer Costume Color Palette and Juan-weaver Costume Color Palette

C. Mining the rules for color matching

Since the number of images in each set of costumes varies greatly, all the images were firstly classified according to the roles corresponding to the costumes in the image data label table in Section 2 of this paper.

Association Rule Mining (ARM) is one of the methods in data mining. ARM is also widely used in the field of color, determining the association between colors based on the co-occurrence frequency of multiple colors. We adopted Eclat to mine the rules of color matching.

However, the traditional ARM acts on the structured database. To mine the association rules on the unstructured data such as image colors, firstly, the unstructured data should be transformed into the structured data; moreover, the mining of the association rules will be affected under the huge amount of the color data, so it is necessary to construct a commonly used color database first, and associate it with the original color dataset.

1) Community Partition

We used Leiden Algorithm [5], a classic algorithm in Community Partitioning, to divide the original color database into similar colors. As a result, we obtained the commonly used color database, whose colors are shown in Fig. 5. There exist 78 commonly used colors, which are highly condensed from the original colors and are representative.

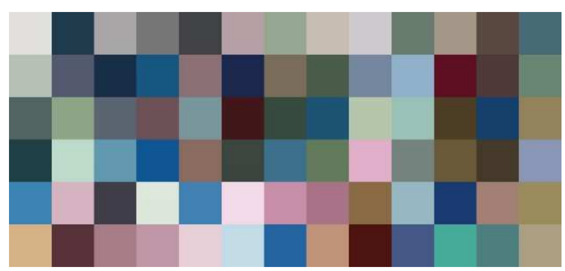


Fig. 5. Common Color Database

2) Eclat-based color association rule mining

After the establishment of common color database. We select the color with the smallest chromatic aberration in the common color database to replace the original color, establishing the Color Database shown in Table 1 and achieving the transformation of non-structural data to structural data.

Table 1 Color Database

index	Common color numbers corresponding to single color							
1	36	1	41	56	21	7	44	76
2	46	22	23	9	62	17	16	

Eclat [6] is a depth-first search algorithm, which adopts vertical data representation, using the important idea of inverted rows. Firstly, convert the horizontal data in Color Database into vertical data, which is shown in Table 2. "1-item" indicates the common color numbered i, and the number j in TID indicates the i^{th} common color appears in the j^{th} image.

Table 2 1-item set

1-item	TID
1	1, 5, 7, 9……
2	18, 21, 29, 42, 49, 80, 83, 84, 85……

Secondly, by directly calculating the number of items in the TID corresponding to each color, the support of the item was calculated according to the formula (1), where $C(AB)$ represents the number of co-occurrences of A and B, and N represents the total amount of data. If its value was less than the preset threshold 0.2, delete it. After screening all 1-item sets, 2-item sets are constructed from them. Repeating the above steps, we obtained the maximum frequent itemset.

$$\sup(A \Rightarrow B) = \frac{C(AB)}{N} \tag{1}$$

Thirdly, arrange and combine the color numbers in the maximum frequent item set, we calculate their confidence by the formula (2). If the value was less than 0.65, delete it.

$$con(A \Rightarrow B) = P(B \mid A) = \frac{P(AB)}{P(A)} \tag{2}$$

Table 3 The Examples of Matching color

NO.	Rule	Role	Color Matching
1	28 => 13	Group-dancer	
2	7 => 1	juan-weaver	
3	9 => 6	student	
4	8 => 1	Ximeng	
5	61 => 12	the stone-polisher	

Finally, we got a total of 241 color matching rules, including 143 binary rules, 99 ternary rules and 9 quadratic rules. Table 3 shows some of the color matching rules.

III. OBJECTIVE ATTRIBUTE ANALYSIS OF COLOR

A. Costume color objective attribute mining

The human eye's perception for color is affected by many factors, which we divide into subjective and objective factors. Objective factors refer to the intrinsic attributes of color, which can quantify the characteristics of color. Subjective factors, the additional attributes given by human, which usually refer to the subjective feelings of human beings towards color.[7]

As the three most important and stable elements of color, hue, lightness and chroma conform to human logical psychology and visual characteristics. Therefore, we chose the three elements of color to excavate the objective attributes of stage costume colors. Under the CIELAB color space, we extracted the L^*, a^*, b^* values of the colors, where L^*represents brightness, a^* represents red-green contrast, and b^* represents yellow-blue contrast. Then the a^*, b^* values are used to quantify the hue h and chroma C, whose calculation formulas are shown in (3) and (4).

$$h = \tan^{-1}\left(\frac{b^*}{a^*}\right) \tag{3}$$

$$C = \sqrt{a^{*2} + b^{*2}} \tag{4}$$

In addition, for color matching pictures, they contain much richer attributes than monochromatic colors. Therefore, we added color richness, space density, hue contrast, and brightness contrast to quantify the features of matching color pictures.[22] The calculation formulas are shown in (5), (6) (7) and (8), where $p(i)$ represent the percentage of pixel with gray value i in image, and n_i represents the number of pixels in the i_{th} segmented region under the watershed algorithm, and $L^*(x, y)$ represents the lightness of the lightness of the eight neighborhoods centered on the pixel(i, j).

$$ColorRichness = -\sum_{i=0}^{255} p(i)\log_2 p(i) \tag{5}$$

$$SpaceDensity = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(n_i - \bar{n})^2} \tag{6}$$

$$ColorToneContrast = \left[\frac{1}{N}\sum_{1}^{N}\left((a_i^* - a^*)^2 + (b_i^* - b^*)^2\right)\right]^{1/2} \tag{7}$$

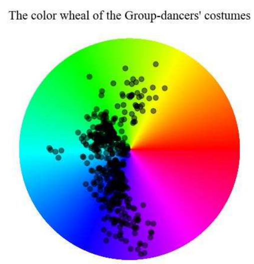
$$LightnessContrast = \frac{\sum_{(i,j)}\sum_{(x,y)\in\Omega}\left|L^*(i,j) - L^*(x,y)\right|^2}{N} \tag{8}$$

B. Character analysis based on color objective factors

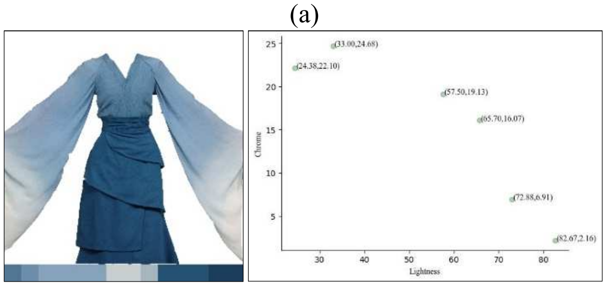
The design of the JLLP is based on the ancient famous painting PMR, in which the green costumes reproduce the rolling hills in the painting. In this paper, we mainly analyzed the objective physical characteristics of the Group-dancers' color.

As shown in Fig. 6(a), the costumes of the Group-dancers are mainly dark green, whose chrome are relatively low. It makes the color soft and subtle, which is similar with the impression the Group-dancers give the audiences. Low-chrome colors don't catch the audiences' eyes, but in the dance drama performances, they are less likely to cause the audiences' visual fatigue.

To analyze the distribution of chrome and brightness of the costumes further, we selected the costume shown in Fig. 6(b) to analyze. The ColorToneContrast of it is 7.843018, which belongs to weak contrast, easily giving person a sense of elegance and softness. The scatter plot shown in Fig. 6(c) reflects the distribution of chrome and lightness, where the chrome distributed in 0~25 belongs to low chrome contrast, containing the natural and untraditional color image, where the lightness is distributed in 20~85. It's LightnessContrast is 1.185044. The low lightness contrast makes the mountain shown by the Group-dancers appear softer and quieter.



(a)



(b) (c)

Fig. 6. Objective Factor. Figure 6(a) shows the distribution of hue and chrome in all costumes of the Group-dancers, Figure 6(b) shows one of the costumes of the Group-dances, Figure 6(c) shows the distribution of lightness and chrome of the image shown in Figure6(b).

IV. SUBJECTIVE ATTRIBUTE ANALYSIS OF COLOR

A. Costume color subjective attribute mining

Colors stimulate human associations, for example, red lets people associate with warmth, and such association is the subjective attribute that human beings give to colors. To analyze the subjective attributes of the costumes color in JLLP, we firstly excavated the semantic words of its color emotion.

Semantic Keyword Extraction Algorithm for Chinese Text Based on Semantics [8] takes Chinese text as the processing object, mapping it into a semantic network to establish a semantic similarity network of words based on the semantic similarity between the words. In the specific implementation process, we collected the texts about JLLP, performed the preprocessing include word segmentation, lexical annotation, lexical filtering and other preprocessing work, building the word semantic similarity network. Besides, in the word statistical features module, we improved the adjective weight, ultimately filtered to get eight semantic keywords. Matching them with antonyms, we got "elegant – boorish", "doughty - gentle", "hard - cozy", "rich – poor" four groups of character-image adjective pairs.

To quantify color subjective images for further research. We designed subjective evaluation experiment and scored each group of character-image adjective pairs with the five-point Likert scale. Based on the frequency of the color occurrence, we selected 25 monochrome images and 25 color matching images as the experiment object to design questionnaire. A total of 32 subjects' data were collected, and the ratio of male to female was 1:1, 65% of which the subjects had some knowledge about color. Deleting two data that did not pass the Ishihara color blindness test and the four invalid data, we calculated the Cronbach Alpha of each adjective pairs to verify the reliability of the experimental data. The Cronbach Alpha is commonly used to measure the consistency of the data, and is usually considered reliable when the value is larger than 0.7.[9] In our experiment, the mean of Cronbach Alpha for four adjective pairs are 0.9165.

Due to the bias in the understanding of color emotions among different individuals, there may be values that are obviously inconsistent with the statistics in the evaluation results, so the outliers that meet the condition of $\left|X_{ij}^k - \bar{X}_j^k\right| > 3\sigma^{2k}_j$ in the experimental data were removed, where $\bar{X}_j^k$ and σ^{2k}_j represent the mean and variance of the image k, and X_{ij}^k represent the score of the i_{th} subject on the j_{th} pair word in the k_{th} image.

B. Correlation analysis on subjective and objective factors

As we mentioned in the previous chapter, the human eye's perception for color is affected by both subjective factors and objective factors. To explore whether there exists relationship between them, we used the Random Forest regression model to calculate the correlation between them, which usually performs well in nonlinear correlation problems.

Before fitting, we divided the data into training set and test set according 7:3. Fig. 7 shows the correlation coefficient of subjective and objective factors of color under each adjective pair. The average correlation coefficients in monochrome and matching color are both higher than 0.7, which indicates that there exists correlation between subjective and objective factors of color. What's more, it is easy to find that the adjective pair, rich-poor performed best in the fitting while the hard-cozy performed more inferior. That may because the former shows a more obvious subjective emotion in the costumes color of JLLP while the latter has the weaker emotional tendency in JLLP.

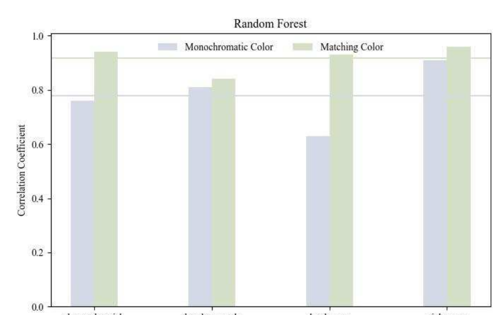


Fig. 7. Random Forest

Furthermore, the average training error in the test set is 0.21, explaining that the model can predict the character personality to some degree.

C. Character analysis based on color subjective factors

To further study the character expressed by color, counting the subjects' average scores for each image and each group of adjective pairs, we focused on analyzing the characters of the Ximeng, the group-dancers, the stone-polisher, and the students of Hanlin Academy.

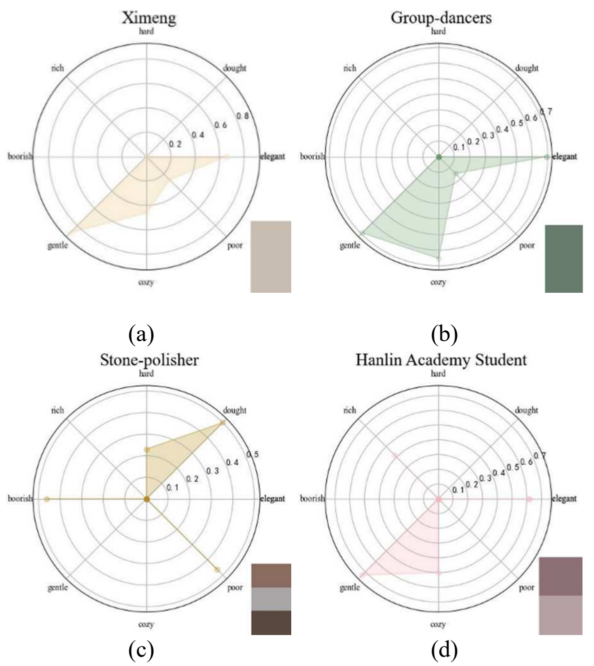


Fig. 8 Character Analysis

In the JLLP, Ximeng is the painter of AMR. His costume is mainly in umber. According to the statistical results of the subjects' evaluation shown in Fig. 8(a), the color gives the impression of elegant and poverty to the subjects during the experiment, which is consistent with the feeling Ximeng himself brings to the audience.

As the imaginary scene in the dance drama, the Group-Dancer whose costumes are rich in colors is to show the continuity and grandeur of peaks. When the dancers' gentleness collides with the dark green of the peeks, a sense of elegance also emerges. The color shown in the Fig. 8(b) is one of the colors of group-dancer costumes, living the impression of elegance and coziness to the subjects.

The color of the stone-polisher's costume is mainly ochre. As one of the working people in the JLLP, the image of hard work and diligence gradually emerges in the gaps between their polishing stones, which are in line with the results of the subjects' evaluation shown in Fig. 8(c) of the given color.

In the chapter of "Penmanship Practice", a group of students appreciating flowers and practicing with the instruction from teacher, which reflects the students' lively and lovely characteristics. It is consistent with the results of the evaluation shown in Fig. 8(d). In addition, the image of the students shown in the JLLP also reflects their richness from the side, which is true with the feeling given by the pink color.

In a word, the result of experiment shows that the colors we extracted can well reflect the character personality in JLLP.

V. Conclusion

In this paper, we took the color of the stage costumes of the poetic dance drama JLLP as the research object, annotated the collected images with scenes, roles and colors, constructing its image dataset; we also extracted the theme color of the costumes by k-means clustering algorithm, regulated the theme color by Leiden community delineation algorithm, and carried out the color mining by Eclat algorithm, to construct the stage costumes colors database represented by JLLP. On this basis, this paper extracted three-dimensional and sixteen dimensional objective physical factors of monochrome color, and matching color, and mined the four groups of character personality adjective pairs from texts. The paper briefly analyzed the character based on the color objective and subjective factors, demonstrating that there exists correlation between them.

Acknowledgment

This paper was supported in part by the National Natural Science Foundation of China (NSFC) - General Program (Grant No. 62276240). This paper was Wang Shuang's result of the second phase of the Palace Museum's open project, supported by Longfor - Forbidden City Cultural Foundation and The Forbidden City Cultural Heritage Conservation Foundation. This paper was the result of project "Color Image Calculation Method Based on Stage Costumes", supported by Key Laboratory of Acoustic Visual Technology and Intelligent Control System of ministry of culture and tourism (Communication University of China).

References

[1] Y. Yin. Research on Color Transfer Algorithm Based on the Miao's Costume Color System. Beijing: Beijing University of Posts and Telecommunications, 2020.

[2] K. Wang. Color Correction and Recognition Method of Test Paper Based on Perfect Reflection and Gray World. Chongqing: Chongqing Jiaotong University, 2018.

[3] S. Chen. The color extraction and matching evaluation of Cloud Shoulder based on K-means. Hangzhou: Zhejiang Sci-tech University, 2021.

[4] W. Long, X. Zhang, and L. Zhang. "Business Process Clustering Method Based on k-means and Elbow Method," J. Jianghan Univ. (Nat. Sci. Ed.), vol. 48, no. 1, pp. 81-90, 2020.

[5] Liu Saisai. Research of Community Detection Algorithms. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020.

[6] F. Zhang et al., "Analysis and Improvement of Eclat Algorithm," Computer Engineering, vol. 36, no. 23, pp. 28-30, 2010.

[7] S. Wang, J. Liu, J. Jiang, Y. Jiang, and J. Lan. "Attribute analysis and modeling of color harmony based on multi-color feature extraction in real-life scenes," Frontiers in Psychology, vol. 13, 2022.

[8] L. Wang and X. Huai, "Semantic-based Keyword Extraction Algorithm for Chinese Text," Computer Engineering, vol. 38, no. 1, pp. 1-4, 2022.

[9] S. Wang, J. Liu, S. Liu, B. Yin, J. Jiang and J. Lan, "Research on Attribute Analysis and Modeling of Color Harmony Based on Objective and Subjective Factors," *IEEE Access,* vol. 10, pp. 75074-75096, 2022.