International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 3290-3301

https://doi.org/10.70517/ijhsa464463a

Intelligent color matching method for smart buildings based on k-means clustering algorithm

Lushan Shi[1,*], Hui Xiong[1], Haoqi Qiu[1] and Rongbang Chen[1]
[1] Xiamen Institute of Technology, Xiamen, Fujian, 361021, China
Corresponding authors: (e-mail: sls76333@126.com).

Abstract Intelligent color matching methods are the future development direction of architectural design due to the extensive use value and development prospect. In this paper, we use web crawler technology to collect data from architectural images in various regions, and perform color recognition as well as data processing on the images based on Reinex theory. Adaptive K-means++ clustering method and intelligent color matching method are proposed. The study firstly screens 450 collected images of different types of buildings. The color characteristics of country house, gothic building, a European house and garden building images were used to do the grayscale histogram color evaluation, and the analysis was mainly carried out with country house, gothic building, a European house and garden building image types, in which the color range of country house building types is small, the color is dark, and there is little difference in the color of the images. Then selected Jiangsu, Zhejiang, Hangzhou three regions of the building image color as a practical case, the use of K-means + + clustering method of the three regions of the building color characteristics of the clustering, the calculation of the building color ratio, color dispersion indicators and constructed a color network model, based on the results of the calculation of the three regions of the color of the judgment of the similarity of the color, based on the similarity of the color of the completion of the building color intelligent Matching. The experimental results show that the building color similarity between Jiangsu region and Henan region is high, which verifies that the color intelligent matching method proposed in this paper has high efficiency and high matching accuracy.

Index Terms intelligent building, color matching, K-means clustering, image processing

I. Introduction

"Green, smart, and livable" smart buildings are essential to meet the sustainable development and smart operation of smart cities [1], [2]. From a structural point of view, smart cities have standardized ideas and complex components, aiming to accelerate the city to become smart. However, in the process of preconstruction of smart city, the first step needs to establish the foundation that supports the whole city - the building, when the existence of all the individual buildings through the hierarchical technical support, to achieve the advancement of intelligence, and then connect it to the other components in the city, that is, to form the foundation of the city's wisdom overall [3], [4]. From this point of view, smart buildings and smart cities belong to a kind of individual contained in the overall set of relations. It can be said that smart building is the most basic element of building a smart city, which is an intelligent building that uses the building as a basic platform to comprehensively manage various information contained in the building [5], [6]. As a new culture in the history of building development, smart building takes the realization of fine control and intelligent design and management of buildings, the integration and application of information as the core, synthesizes a variety of innovative modern technologies, and deeply integrates the building's own economic performance, cultural performance, management performance and use performance [7], [8]. Therefore, the development of smart buildings is an indispensable and important force to help build a smart city. With the rapid development of the economy, people's requirements for the construction of smart buildings have gradually increased [9], [10]. Color layout, as the main point of architectural design, has a greater impact on the human psychological effect. Therefore, it is necessary to examine how to combine the actual needs of smart building construction based on the k-means clustering algorithm, and design a new intelligent matching method for smart building colors to achieve a reasonable matching of color elements, to ensure a high degree of similarity of colors, to safeguard the aesthetics of the layout of the smart building, to improve the visual effect, and to enhance the efficiency of color matching [11]-[15].

The study proposes adaptive K-means++ clustering as well as color intelligent matching methods. Firstly, the architectural images were collected using crawler technology, and the collected data were integrated and processed, totaling 450 architectural images. Some representative images of different types of buildings were selected for gray histogram analysis, and the analysis was carried out with the architectural images of country houses, Gothic buildings, European residential buildings and garden buildings. In order to verify the effectiveness of the color intelligent matching proposed in this paper, the architectural colors of Jiangsu, Henan, and Hangzhou are selected as the research objects, and the color percentage and color dispersion of the architectural colors of each region are calculated, and the color similarity is judged based on the calculation results to complete the color intelligent matching.

II. Intelligent Building Color Intelligent Matching Design

II. A.Architectural Image Collection and Processing Based on Retinex Theory

II. A. 1) Architectural image collection

With the sudden increase in image data and coverage of buildings in the network, Internet photographs have been used as a source for building such systems in a fully automated manner. As far as authenticity is concerned, the collection of Internet photographs comes from a wide range of sources and has a wide range of time span and authenticity. At the same time, field photography is too extensive and costly, and the feasibility of collecting large-scale landmark building scene data is low.

Based on this, the data collection part of this paper selects the web crawler algorithm based on Scrapy framework to crawl the relevant image data with the landmark name as the keyword. Before mining the data on the Internet, it is necessary to establish a “landmark name keywords” catalog. Therefore, this paper first collects the text catalogs of buildings in different places through a hierarchical way.

II. A. 2) Building image analysis based on Retinex theory

Retinex theory is a landscape image segmentation tool that can split the image into two light components, namely incident light and reflected light. In general, the landscape image light feature extraction method is used to obtain a sub-block of landscape light information, add the incident component to it, go through the overlapping process, and at the same time, pan the sub-block, according to the results of the incident component calculation, remove the incident component of the image in the landscape, and ultimately obtain the image reflective component, so that the image can be enhanced [16]. This method can not only improve the efficiency of local color matching, but also shorten the time consuming image color recognition.

II. A. 3) Architectural image enhancement processing

In this paper, the Retinex theory is used as the research basis, assuming that the incident light and the reflected light compose the ecological landscape image, the former component is denoted as $W(x,y)$, the latter component is denoted as $Q(x,y)$, and the landscape image is denoted as $I(x,y)$. According to the relationship between the three before, the following formula is established:

$$I(x,y)=Q(x,y)\times W(x,y) \tag{1}$$

Equation (1): I image size is calculated as $W\times Q$. Assuming that the two image size of the same matrix as Count, After, the use of these two matrices to store the original landscape image information, after the calculation process to form the landscape image information.

Take the upper left region of the original landscape image as the information acquisition area, extract a sub-block of graphic information from the rest of this, recorded as $W\times Q$, the sub-block to take the filtering process, to get the incident component, and then take the constant logarithm, to get the localized image, the following is the calculation formula:

$$I'(x,y)=Q(x,y)+W(x,y) \tag{2}$$

The local image of Eq. (2) is taken as a Gaussian function, and the smoothing filtering is done according to the calculation results to obtain the incident component, which is given below:

$$I_{low}(x,y)=I'(x,y)*E(x,y) \tag{3}$$

The result of the calculation in Equation (3) will be stored in the After matrix, and at the same time, the data information involved in the above calculation will be stored in the Count matrix in the matrix sub-block.

Take the original landscape image as the processing object, move to the right direction, set the moving distance as sub-block B, and the control step size shall not exceed m. Observe the relationship between the right side of the sub-block and the boundary of the landscape image at this time, if the right side exceeds the boundary position, then move to the left direction, and the moving object is still sub-block B, and the control step size shall not exceed n. If the right side fails to exceed the boundary of the landscape image, then continue to move to the right. Repeat the operation according to the above steps to obtain the corresponding incident component of the landscape image.

Regarding the acquisition of the enhanced landscape image, take the incident component of the image as the processing object, remove this part of the incident component in the logarithmic domain, and then the enhanced landscape image can be obtained, which is denoted as Q', and the calculation formula is as follows:

$$Q'(x,y) = I'(x,y) - \log I_{bon}(x,y) \tag{4}$$

The color spaces commonly used in image processing are RGB, CMYK, CIE-Lab and HSV color spaces, and different color spaces can be converted by formula.

II. B.Building color extraction based on k-means clustering algorithm

II. B. 1) Color space and conversion

(1) Color space

RGB color space is based on the principle of additive color mixing, the color space composed of the three primary colors of light, red, green and blue, for the division of the three colors of the range specified as 0-255, CMYK color space is based on the principle of subtractive color mixing, the color space of the color shown through the reflection of light, widely used in the industrial printing industry [17]. RGB and CMYK color space are both oriented to the computer and other hardware. The HSV color space is a model based on the characteristics of human observation of color, the space makes the hue, saturation, and brightness to describe the color characteristics, which can directly reflect the relationship between colors, and higher flexibility.

(2) Color space conversion

Generally images are stored in RGB form in the computer, image processing with the help of computer c language involves opencv computer vision library, in which the image format defaults to BGR mode, so first the image BGR mode is converted to RGB mode, then RGB is used as a medium and finally converted to HSV mode. In this, the formulas involved in RGB to HSV are shown below. Let $\max$ and $\min$ be the largest and smallest of r, g and b respectively, i.e. (h,s,v) in HSV color space:

$$h = \begin{cases} 0 & if \max = \min \\ 60 \times \dfrac{g-b}{\max - \min} + 0 & if \max = randg \geq b \\ 60 \times \dfrac{g-b}{\max - \min} + 360 & if \max = randg < b \\ 60 \times \dfrac{b-r}{\max - \min} + 120 & if \max = g \\ 60 \times \dfrac{r-g}{\max - \min} + 240 & if \max = b \end{cases} \tag{5}$$

$$s = \begin{cases} 0 & if \max = 0 \\ \dfrac{\max - \min}{\max} = 1 - \dfrac{\min}{\max} & Other \end{cases}$$

$$v = \max$$

II. B. 2) Creation of color histograms

There are about 224 colors in the real color space, but from the point of view of image segmentation, the interval in which the human eye can perceive the existence of different shades in a color image is narrow, so most of the colors in the image are redundant in terms of the segmentation results conforming to human visual perception. Moreover, the dimension of the global color histogram established by counting each color contained in a color image is too high, which obviously increases the computational volume and reduces the efficiency of the algorithm when calculating the similarity between pixel blocks. Therefore, in order to reduce the computational complexity and better distinguish similar colors, non-uniform quantization of the color space is an effective measure.

Quantization refers to grouping all similar colors within a certain range into one main color, and using this one main color to represent the original color. As a preliminary work for statistical analysis of image color values, quantization of the color space can reduce the number of colors to be processed and color redundancy while maintaining the main color of the image, so that the amount of data to be processed during clustering of the image is greatly reduced. In this paper, the HSI color space components are divided non-uniformly, where the hue H is divided into eight parts, and the saturation value S and luminance value I are both divided into three parts. Then all the colors in the image are mapped into 8 × 3 × 3 = 72 color Bin, and similar colors are mapped into the same division region. The specific division principle is shown in Equation (6):

$$H=\begin{cases}0 & if \quad h\in[316,20]\\1 & if \quad h\in[21,40]\\2 & if \quad h\in[41,75]\\3 & if \quad h\in[76,155]\\4 & if \quad h\in[156,190]\\5 & if \quad h\in[191,270]\\6 & if \quad h\in[271,295]\\7 & if \quad h\in[296,315]\end{cases} \quad S=\begin{cases}0 & if \quad s\in[0.0.2)\\1 & if \quad s\in[0.2,0.7)\\2 & if \quad s\in[0.7,1)\end{cases} \quad I=\begin{cases}0 & if \quad i\in[0.0.2)\\1 & if \quad i\in[0.2,0.7)\\2 & if \quad i\in[0.7,1)\end{cases} \tag{6}$$

Based on the above quantization level, different weights are set for H, S, and I respectively, so as to transform this three-dimensional feature vector into a one-dimensional feature vector G.

The formula is shown in (7):

$$G=w_1\cdot H+w_2\cdot S+w_1\cdot I \tag{7}$$

w_1, w_2 and w_3 are the quantization levels of components H, S and I respectively. The experimental results prove that among the above three components, the dependence of G on component S is greater than or equal to three times the dependence on component I, and the dependence on component H is greater than two times the dependence on component S, so the conclusion shown in (8) can be drawn:

$$w_2\geq 3\cdot w_3, w_1\succ 2\cdot w_2 \tag{8}$$

In this paper, if w_3 is set to 1, then w_2 is 3, w_1 is 9, and G is calculated as shown in (9):

$$G=9\cdot H+3\cdot S+I \tag{9}$$

In this way, the one-dimensional feature vector G composed of three-dimensional feature vectors is less noisy and regionally distinct, and Bin corresponding to color values that contain more blocks of pixels in the space represents a more color-concentrated region of the image.

II. B. 3) Adaptive K-means++ Color Feature Extraction

The adaptive K-means++ algorithm to extract the main color features of each wire region of the wire harness consists of two parts: the first one is to determine the optimal number of clusters; the second one is to use the optimal number of clusters to automatically obtain the stable clustering centers of each wire region of the wire harness, and the clustering centers with the largest weights are taken as the main color features of each wire region [18].

(1) Determination of optimal clustering number

Before applying the K-means++ clustering algorithm, it is necessary to preset the number of clusters K. If the number of clusters of the research problem is known in advance, the value of K can be set directly. The value of K is a core variable in K-means++, and whether it is chosen appropriately affects the final clustering effect. This topic determines the optimal number of clusters by combining the K-means++ algorithm and quantitative indicators.

The contour coefficient indicator tests the clustering effect by combining intra-cluster cohesion and inter-cluster separation. The process of implementing this metric is:

Step1: If any data point x_i in the data set is classified into cluster A, compute the average distance between point x_i and all other data sample points in the same cluster A as $a(i)$, $a(i)$ is used to measure intra-cluster cohesion.

Step2: Calculate the average distance between point x_i and all other data points in the cluster except cluster A, denoted as $b_j(i)(j=1,2,\cdots,K)$. Select the cluster $b(i)$ with the closest distance to measure the inter-cluster separation.

Step3: Calculate the contour coefficient denoted as $s(i)$ for point x_i, denoted as:

$$s(i)=\frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \tag{10}$$

Step4: Calculate the profile coefficients of all data points in the dataset and find the average profile coefficient, denoted as $\overline{s}(i)$, then there exists a $\overline{s}(i)=\frac{1}{N}\sum_{i=1}^{N}s(i)$ relationship, and the range of values of $\overline{s}(i)$ is $[-1,1]$. In this range, the closer the value of $\overline{s}(i)$ is to 1, it means that the smaller the value of $a(i)$ is, the higher the similarity between the data points within the cluster, and the larger the value of $b(i)$ is, the higher the difference between the data points between the clusters, and the better the clustering effect is.

(2) Adaptive K-means++

After the image preprocessing operation, the edge contour of the target region in the wire harness image can be obtained. The filtered image eliminates the noise interference on the surface of the wire harness image, and the corresponding wire region in the filtered image can be obtained according to the contour position of each wire, and at the same time, in order to shorten the time spent on the clustering algorithm and improve the operational efficiency, a part of the rectangular region of each wire in the image is selected for clustering analysis as shown in Figure 1.

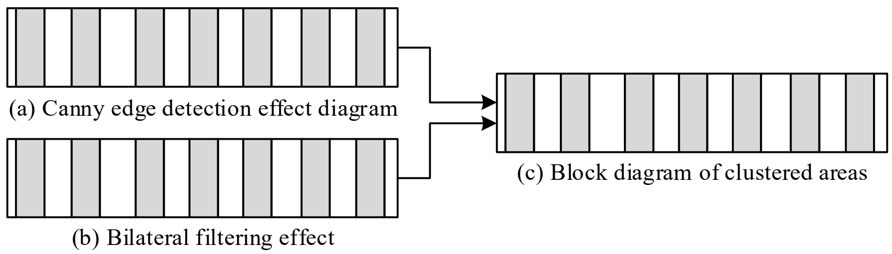


Figure 1: Cluster area effect schematic

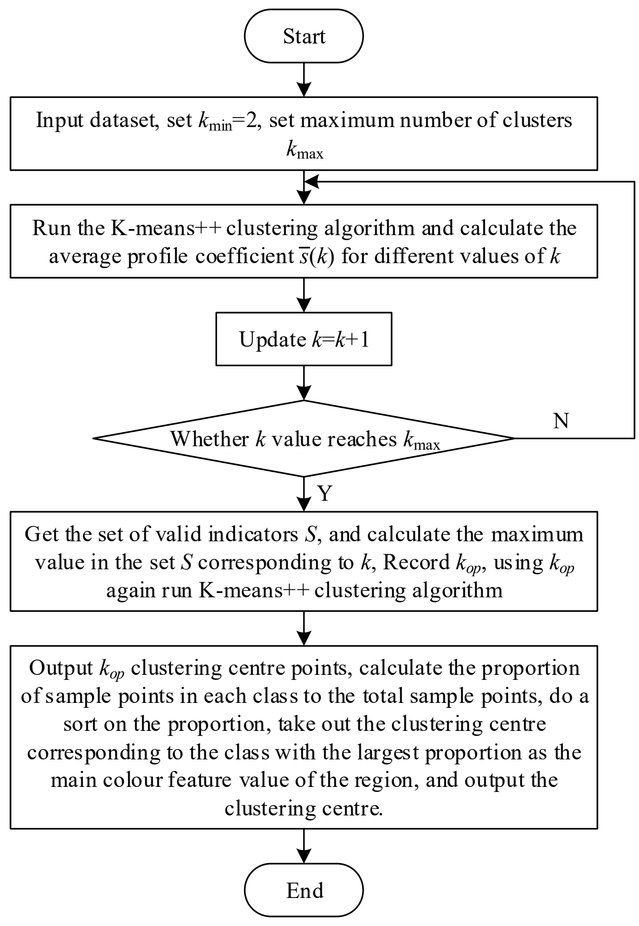


Figure 2: Adaptive k-means + + clustering algorithm flowchart

The clustering objective of K-means++ is to classify N data points in dataset X into a dataset $C=\{c_k \mid k=1,2,\cdots,K\}$ consisting of K classes, where the center of clustering of class c_k is denoted as μ_k, and the Euclidean distance between any data point x_i classified into class c_k and the center of clustering μ_k is denoted as:

$$dist(x_i,\mu_k)=\|x_i-\mu_k\|^2, x_i \in c_k \tag{11}$$

The sum of Euclidean distances from all data points x_i classified into class c_k to the clustering center μ_k of that class is:

$$J(c_k) = \sum_{x_i \in c_k} dist(x_i, \mu_k) = \sum_{x_i \in c_i} \|x_i - \mu_k\|^2 \tag{12}$$

Thus the sum of squared errors (SSE) of the clustering algorithm is obtained using equation (13), which is the sum of the distances of each data point contained in class K to its corresponding cluster center, denoted as:

$$SSE = \sum_{k=1}^{K} J(c_k) = \sum_{k=1}^{K} \sum_{x_i \in c_k} \|x_i - c_k\|^2 \tag{13}$$

The clustering objective is to minimize the value of SSE.

The specific flow of the adaptive K-means++ clustering algorithm is shown in Figure 2:

II. C.Building Color Intelligent Matching Design

All the colors in architectural images can be used as vector sets and exist in the color space. Any vector exists in the landscape image with corresponding color, calculating the feature vector similarity, judging the color similarity based on the result, and realizing the intelligent matching of landscape layout color on this basis.

Settings $W'(r',g',b')$ and $W''(r'',g'',b'')$ characterize the different colors present in the color space, respectively; η represents the chromatic aberration factor, on the basis of which the different color differences can be elaborated, calculated as:

$$\eta(W',W'') = \sqrt{(r'-r'')^2 + (g'-g'')^2 + (b'-b'')^2} \tag{14}$$

The greater the color similarity between W' and W'' elaborations, the smaller the difference in color vectors and the smaller the color factor values. There is an inverse correlation between the remaining color similarities. W' and W'' describes that the colors are the same when the color difference factor value is 0 and the color similarity is maximum, which represents that the colors are exactly the same.

Based on the inverse correlation between color similarity and color factor, the similarity of landscape layout colors can be analyzed based on the metric of color difference factor.

Gray-scale image matching is limited by compatibility, polar lines, parallax range and uniqueness. In order to accurately and reasonably use the color information of the landscape image, it is necessary to add the color similarity to limit the color matching, i.e., the best matching point should be controlled in the parallax search range and the reference point with the smallest color difference between the points, the reference point and the best matching point with the smallest color difference factor.

The color intelligent matching method introduces color similarity constraints based on the product between the squared pixel gray difference (Δk) and the color difference factor, i.e.:

$$\sigma\Delta k = \min(\sigma\Delta_1, \sigma\Delta_2, \sigma\Delta_3, \cdots\cdots, \sigma\Delta_N) \tag{15}$$

When the color difference factor of the point to be matched is in the minimum state, it belongs to the best matching point and satisfies Equation (15).

When color matching, comparing the color difference between the point to be matched and the datum point, the provided constraints cannot satisfy the relevant standards. The intelligent matching method for landscape layout color is based on the sum of the color difference factors between the pixel neighborhoods of the points to be matched and the datum point of $m \times n$. The color similarity range is further limited, i.e:

$$\alpha = \sum_{n+1}^{m} \sum_{n=1}^{n} \sigma_{MW} \tag{16}$$

where σ_{MW} represents the color difference factor between the reference point and the corresponding pixel point in the pixel neighborhood of the point to be matched.

The landscape layout color matching formula is obtained based on Eq. (16), i.e:

$$\alpha\Delta k = \min(\alpha\Delta_1, \alpha\Delta_2, \alpha\Delta_3, \cdots\cdots, \alpha\Delta_N) \tag{17}$$

III. Intelligent building color analysis and intelligent color matching empirical analysis

III. A. Experimental sample selection

The main research content of this paper is to color analyze the color matching colors of selected regions and countries' architectural images, mine their color information and their coloring laws, etc., and color collect their color matching methods, and use the color matching proposed in this paper for the automatic recommendation of the color matching scheme. Among them, based on the division of Chinese architecture and western architecture, the research object with representative architectural image color characteristics is selected, and the number of collected samples is 150 for each region, with a total of 450, and the images are obtained from the relevant websites of network data.

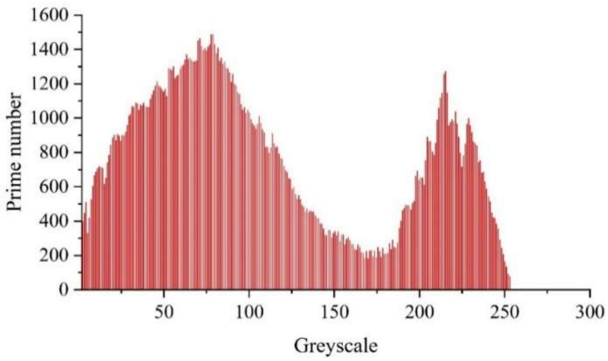
III. B. Analysis of regional architectural color composition

III. B. 1) RGB color feature map analysis

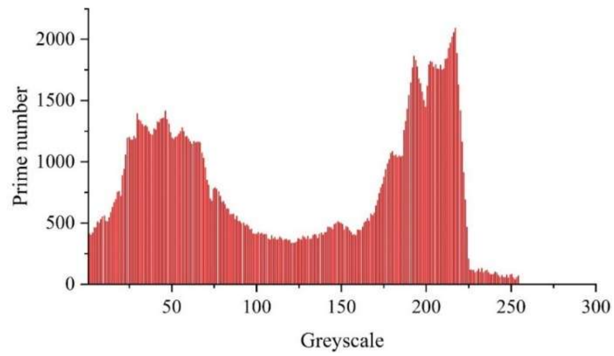
In this section we have done the histogram analysis of various types of collected architectural images and extracted the grayscale histogram in the model and the global color histogram in the model respectively. The properties present in the histograms of architectural images are described below respectively.

A: Gray scale histogram characteristics of country house

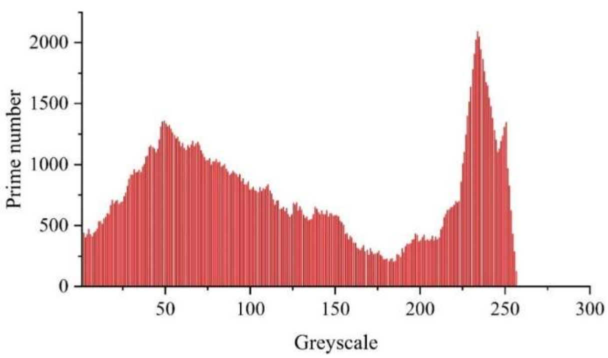
In the model, the grayscale histogram of each architectural image is extracted separately, and now the grayscale histogram of a representative image of a small country house is given, as shown in Fig. 3.



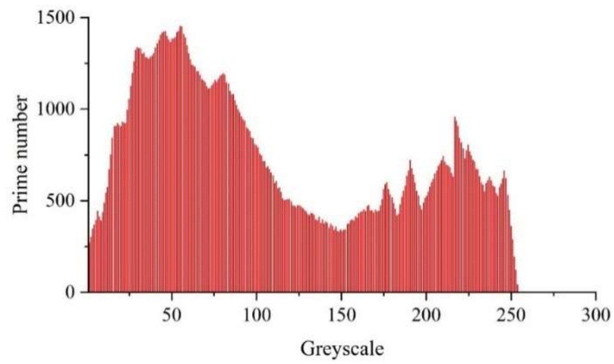
(a) Image 1



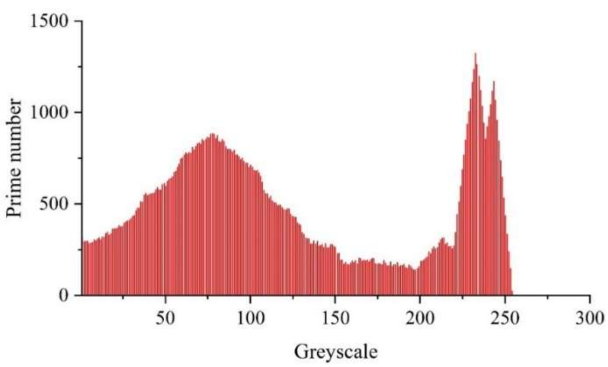
(b)Image 2



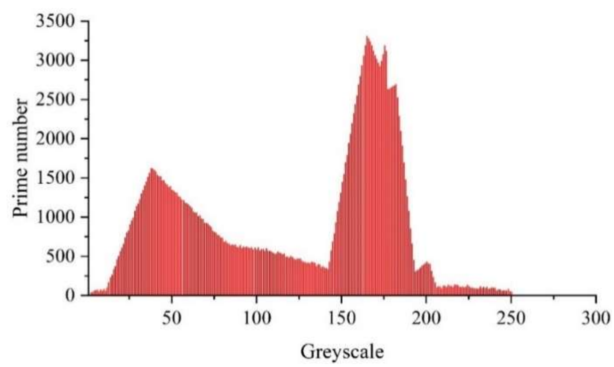
(c)Image 3



(d)Image 4



(e)Image 5



(f)Image 6

Figure 3: Color gray histogram

As can be seen from Fig. 3, in this class of images, the histograms of the grayscale of each image show a clear bimodal distribution, and the detailed data are shown in Table 1.

Combining Fig. 3 and Table 1 for the histogram curve with two gray level maxima, the first gray level maximum appears on the left side, which accurately describes the image with excellent resolution. The good brightness of the image is mapped to the second maximum on the right side of the curve. At the same time, the area ratio of the curve corresponding to these two maxima should be kept essentially constant.

Figure 3(b) is darker, so the second maximum is slightly to the left; in addition, this image has been shot twice and noise has been added, so the histogram distribution is presented with the first peak occupying a significantly smaller total area. Figure 3(f) is darker overall, and its second peak is obviously to the left, indicating that the proportion of houses and backgrounds in this image of the country house building image is quite equal, so the two peaks occupy almost the same area. For this type of architectural image, the visual effect of the image is best when the proportion of the first peak in the total area is roughly 0.67-0.73.

In addition, the details of the country house image are well represented in the gray scale histogram. Taking the histogram of the country house building image in Fig. 3(a) as an example, its histogram is distributed in the whole gray scale region, and the first peak is located in the left dark region, which corresponds to the lush trees in the background in the traditional country house; and the second peak is located in the right bright region, which corresponds to the long blue water and the houses in the country house building. The distribution of the two peaks on the histogram shows that the image is appropriately light and dark. This type of country house is suitable for human habitation and can make people feel relaxed and happy.

Table 1: Gray histogram distribution of small rural area

Image number	1st peak	Valley	2st peak	Left peak area
(a)	68	169	222	0.745
(b)	42	110	211	0.475
(c)	53	168	239	0.673
(d)	53	146	231	0.698
(e)	64	174	241	0.720
(f)	51	138	178	0.537

B: Histogram Characteristics of Other Types of Building Images

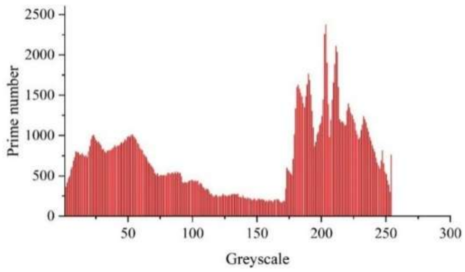
Examples are given to illustrate the uniqueness of different types of architectural images, which are analyzed by histogram plotting in order to analyze and display more intuitively, as shown in Figure 4.

After analyzing the examples in this section to illustrate several types of architectural images, for each type of image, although its gray-scale histogram does not have an obvious bimodal distribution characteristics, but also shows different light and dark characteristics. Examples are given to illustrate the uniqueness of each:

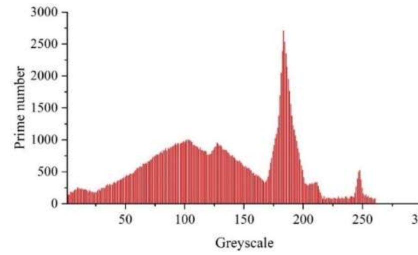
Figure 4(a) shows multiple sharp peaks in the bright area with a large percentage in the gray level histogram. The dark part of the change is more gentle, but the dark area also occupies a certain proportion. The reason for this is that Gothic churches have helicoidal lines, a soaring and majestic appearance, with pointed arches, large windows, slender bunched columns, and floral window panes painted with biblical stories as the main elements. These tall buildings are usually set against an upper background of blue sky and white clouds and a lower background of lawn and trees, with gray, white, and fawn as the dominant colors. The large area of sky and light-colored buildings in the image make the histogram have large light areas, while the dark green bushes and black openwork windows correspond to the grayscale distribution characteristic of its dark areas. In short, the histograms of such towering buildings will have large bright regions.

For the European residential building in Fig. 4(b), its houses are mainly in reddish-brown and orange, with a few white and green accents, and the background color is soft and mostly in warm tones, with a moderate degree of brightness, giving people a sunny and warm feeling. In the gray scale histogram of European residential buildings, the gray scale level is mainly concentrated in the middle region, and the higher spikes in the original figure correspond to the higher color brightness in the region. Therefore, the gray level of European residential buildings is located in the middle of the histogram, and its light and dark changes are not obvious.

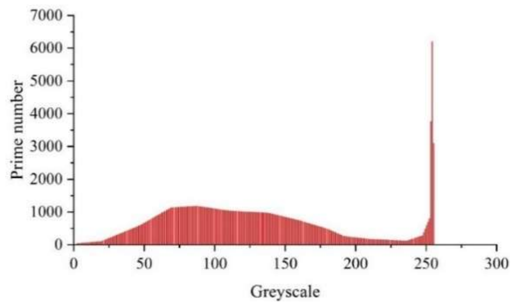
Figure 4(c) shows the grayscale histogram of the Summer Palace image. After investigation, the garden buildings are characterized by the buildings themselves using more big red, dark green, and dark blue colors, and a large green tree as the main background, with fewer light-colored elements, which makes the histograms of the gardens mainly distributed in the dark area, and the bright area is very small. The histogram of the garden does not have the obvious bimodal distribution characteristics of the country house, which also shows that the garden architecture has its own distinctive features.



(a)Gray histogram of Notre Dame



(b)Gray histogram of European house

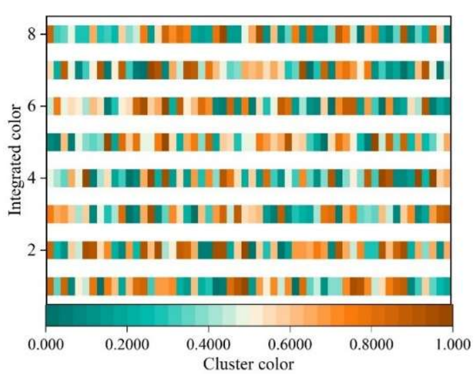


(c)Gray histogram of Summer Palace

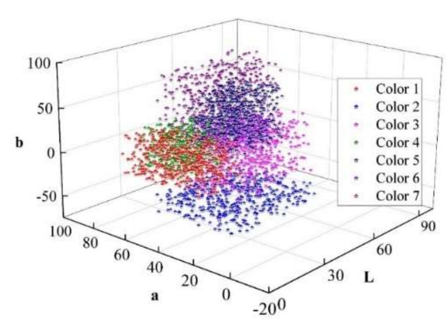
Figure 4: Different building type histogram analysis

III. B. 2) Analysis of the Color Composition of Regional Buildings

This section first takes the color of architectural images in Jiangsu region as an example, and on the basis of adaptive clustering of each sample image, the clustered color is obtained after secondary clustering. 41 samples of Jiangsu architectural images of adaptive color clustering comprehensive color map is shown in Figure 5. In order to facilitate lateral comparison of the differences in color distribution in different regions, the K value of the secondary clustering in the experiment is set to 56, which gives the 30-bin color values in order. In addition, based on the comprehensive color map, the distribution of the clustered colors in the CIELaB space is plotted (color space distribution map colors to distinguish different categories of colors). Further, in order to demonstrate the extracted color two-by-two co-occurrence relationship, the clustered color full relationship matrix is shown in Table 2, i.e., the two-by-two co-occurrence frequency when the threshold value is zero. In order to facilitate the observation of high-frequency co-occurrence relationship, the experimental co-occurrence rate threshold ∂ is uniformly set to 0.2.



(a)Cluster 1



(b)Cluster 2

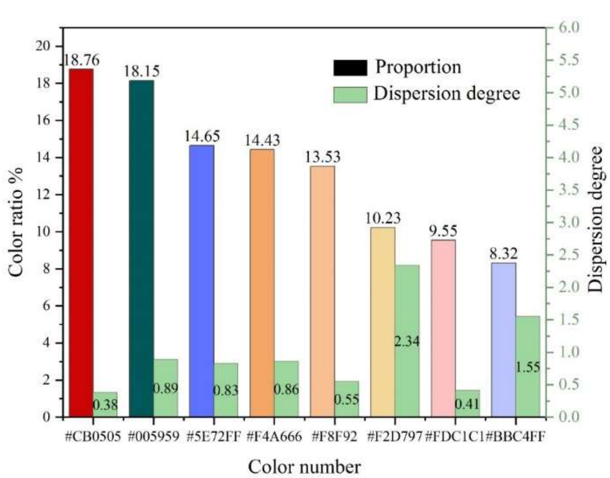
Figure 5: Color map of jiangsu area

Table 2: Matrix of architectural clustering in Jiangsu province

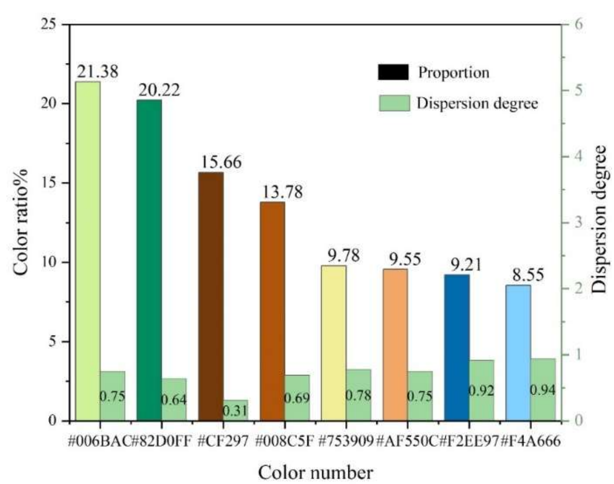
Color number	Color number 1	Color number 2	Color number 3	Color number 4	Color number 5	Color number 6	Color number 7	Color number 8
Color number 1	-	**0.35**	**0.41**	0.05	0.18	0.00	0.00	0.00
Color number 2	0.35	-	**0.48**	0.26	0.18	0.18	0.00	0.05
Color number 3	0.41	0.48	-	0.12	**0.99**	0.05	0.11	**0.41**
Color number 4	0.05	0.26	0.12	-	0.00	**0.48**	0.00	0.05
Color number 5	0.18	0.18	0.99	0.00	-	0.11	0.11	**0.48**
Color number 6	0.00	0.18	0.05	0.48	0.11	-	0.05	0.25
Color number 7	0.00	0.00	0.11	0.00	0.11	0.05	-	0.05
Color number 8	0.00	0.05	0.41	0.05	0.48	0.25	0.05	-

Note: The values in bold in the table are the pairwise color co-occurrence rates that are greater than the co-occurrence rate threshold and are shown as a line in the color network relationship model diagram.

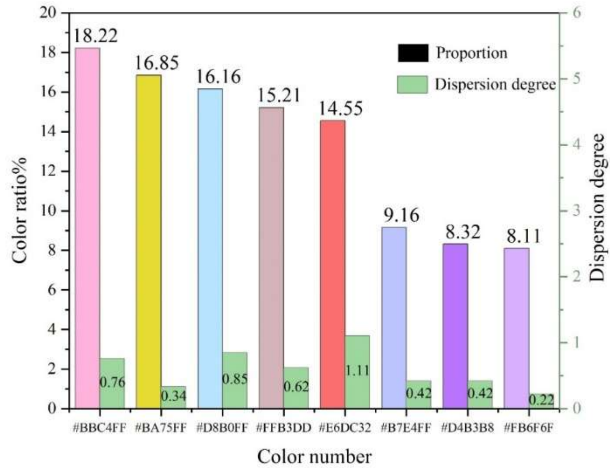
According to the above basic feature information, this paper constructs the color network relationship map. Figure 6 shows the dispersion of architectural image colors in Jiangsu, Henan and Hangzhou. In terms of hue performance, the overall performance of the three colors is black, gray, yellow, red, etc., and there is a correlation between the hues. However, in the proportion of colors used, there are some differences in their color sequences, indicating the differences in the coloring of the main, auxiliary and accent colors of their garments. From the perspective of architectural color composition, in order to allow designers to observe the main, auxiliary, and accent colors, the experiment delineates the 1~3 color as the main color, the 4~7 color as the auxiliary color, and the 8 color as the accent color, with the three main colors accounting for 47.21%, 53.69%, and 47.53%, respectively. The auxiliary colors were 42.58%, 36.22% and 43.88% respectively. The proportion of accent colors is 10.21%, 10.09% and 8.59% respectively. As can be seen from the data, Jiangsu and Hangzhou are closer in terms of color ratio, and Henan has a slightly higher proportion of primary colors. The value of the folded line in Figure 6 is the dispersion of the color, except for the No. 6 cluster color in Jiangsu region, which is more than 1.5, the dispersion of the color in the rest of the regions is more balanced. Further, by summing up the dispersion of each cluster color, we can get the overall dispersion of the colors used in architectural images in the three regions in order of 7.81, 5.78, 4.74, and the data show that Hangzhou region is the most stable in architectural colors.



(a) Jiangsu



(b)Henan



(c)Hangzhou

Figure 6: Regional architectural image clustering ratio and dispersion

IV. Conclusion

This paper proposes an intelligent color matching design method on image enhancement technology, and uses K-means++ combined with contour coefficient index method, i.e., adaptive K-means++ to automatically extract features of different types of architectural images, analyze the grey histograms of these types of architectural images as well as classify the architectural images of different regions for color matching by using the intelligent color matching method. It was found that when the color analysis of country house and other types of architectural images were analyzed, the gray histogram of the architectural image of country house had the best visual effect when the proportion of the first peak in the total area was roughly 0.67-0.73. Finally, using image technology and clustering methods to analyze the color matching of Jiangsu, Hangzhou and Henan, the results show that there are differences in the use of color in the three regions, and the main colors of buildings in Henan and Hangzhou account for a higher percentage.

References

[1] Jin, Y. , & Wang, J. . (2023). Integrated application strategy of large-scale intelligent building based on renewable energy technology. Intelligent Buildings International.

[2] Du, Y. , Zhou, Z. , & Zhao, J. . (2022). Multi-regional building energy efficiency intelligent regulation strategy based on multi-objective optimization and model predictive control. Journal of Cleaner Production, 349, 131264-.

[3] Li, Y. W. , & Cao, K. . (2020). Establishment and application of intelligent city building information model based on bp neural network model. Computer Communications, 153.

[4] Yigitcanlar, T. , Butler, L. , Windle, E. , Desouza, K. C. , Mehmood, R. , & Corchado, J. M. . (2020). Can building "artificially intelligent cities" safeguard humanity from natural disasters, pandemics, and other catastrophes? an urban scholar's perspective. Sensors (Basel, Switzerland), 20(10).

[5] Chen, W. , Yang, Q. , Zhao, S. , Xing, J. , & Zhou, Q. . (2020). A graphical programming language and its supporting tool for insect intelligent building. Scientific Programming, 2020(1), 1-18.

[6] Nathalie, Labonnote, Karin, & Hoyland. (2017). Smart home technologies that support independent living: challenges and opportunities for the building industry -a systematic mapping study. Intelligent Buildings International.

[7] Prusak, D. , Karpiel, G. , & Kuakowski, K. . (2021). The architecture of a real-time control system for heating energy management in the intelligent building. Energies, 14.

[8] Boodi, A. , Beddiar, K. , Benamour, M. , Amirat, Y. , & Benbouzid, M. . (2018). Intelligent systems for building energy and occupant comfort optimization: a state of the art review and recommendations. Energies, 11(10).

[9] Papantoniou, S. , Mangili, S. , & Mangialenti, I. . (2017). Using intelligent building energy management system for the integration of several systems to one overall monitoring and management system. Energy Procedia, 111, 639-647.

[10] Yang, M. , She, D. , & Guo, Y. . (2023). Multiagent game of intelligent building detection and its harm rumor analysis. Mathematical Problems in Engineering.

[11] Zhao, H. . (2022). Design and implementation of an improved k-means clustering algorithm. Mobile information systems, 2022(Pt.29), 6041484.1-6041484.10.

[12] Hussain, M. A. , Akbari, A. S. , & Ghaffari, A. . (2017). Colour Constancy Using K-Means Clustering Algorithm. 9th international conference on developments in esystems engineering. IEEE.

[13] Khan, Z. , & Yang, J. . (2024). Nonparametric k-means clustering-based adaptive unsupervised colour image segmentation. Pattern analysis and applications: PAA, 27(1), 17.1-17.20.

[14] Nongmeikapam, K. , Kumar, W. K. , & Singh, A. D. . (2018). Fast and automatically adjustable grbf kernel based fuzzy c-means for cluster-wise coloured feature extraction and segmentation of mr images. Iet Image Processing, 12(4), 513-524.

[15] Zhang, X. . (2022). Research on colour matching in art design based on neural network mathematics models. Mathematical Problems in Engineering.

[16] Zhichao Sun,Huachao Zhu,Xin Xiao,Yuliang Gu & Yongchao Xu. (2024). Nighttime image semantic segmentation with retinex theory. Image and Vision Computing105149-105149.

[17] FURUKAWA Shota,NAGASAKI Rei & UEDA Yoshiaki. (2022). Brightness and Saturation Enhancement by Using Color Vector on Equi-Hue Plane in RGB Color Space. Procedia Computer Science966-975.

[18] Kai Tian,Jiuhao Li,Jiefeng Zeng,Asenso Evans & Lina Zhang. (2019). Segmentation of tomato leaf images based on adaptive clustering number of K-means algorithm. Computers and Electronics in Agriculture104962-104962.