Fashion Color Verification System Based on Similarity Search Algorithm

Chengrui Xu¹, Yujing Tian¹*, Zeyu Zheng², Ge Li² and Wenxiang Shang²

¹ Donghua University, College of Fashion and Design, No.1882 West Yan’an Road, Changning District, Shanghai, P.R. China
² Hangzhou Zhiyi Technology Co., Ltd., Room 1201, Block F, Hangzhou Bay Information Port, Xiaoshan District, Hangzhou, P.R. China

* tinatian5186@163.com

Abstract. As the times progresses and the technology rapidly develops, it has become a trend to make use of artificial intelligence to accurately predict the trend of fashion color. In facing up with the demand of enterprises and design developers’ verification on the accuracy of color decision-making, a set of fashion color verification system based on similarity search algorithm is proposed on basis of the technical feasibility of ‘searching image by color’ and the prediction method of the fashion color in this paper. Through the reasonable construction of database and analysis framework, the known color is introduced into the system. Not only can the masonry layout of color image under the selected labels be gained, but also the visual report on the analysis of the color trend can be gained, in which the verification and correction of the color can be achieved. By comparing this set of system with the analysis process of the traditional artificial color trend, it’s concluded that the system can effectively improve the accuracy of predication and the efficiency of the verification.

1 Introduction

1.1 Research background

In the modern society, the update of the fashion is rapid, the timeliness is strong. Not only is the correct judgment and interpretation of the fashion trend associated with whether the products meet the social demand, and are marketable, but also associated with the economic benefits and survival of enterprises. To solve this problem, it’s necessary for brand manufacturers to predict the emerging trends in advance and predict the demand at the primary stage, or make the full verification before new products enter the market.

This paper begins with the attribute of the color. The basic flow for the verification of the fashion color is organized, the source of the fashion database is precise, comparing the advantages and disadvantages of key technologies such as ‘search image by color’, ‘search color by image’, ‘search color by color’ and ‘search image by image’. A scientific, fast and accurate model of the trend verification of fashion colors is established with computers’ deep learning, which can be used to judge the supporting relationship between the known color and fashion colors of products, and ensure that the correct decision is made when the new color is enter the market. It effective help enterprises to reduce inventory, and achieve commercial success as a result.
1.2. Current research
As the times develops, artificial intelligence allows the digitalization of fashion trend to come true. It is also the inevitable orientation of the industry development to predict the color trend by combining artificial intelligence with the insight of big data in consumption. Color is the most prominent feature in clothing image. The common method of describing color feature include color histogram, color moment and aggregative vector. Besides, there are the upgraded versions of these algorithms, for example, extraction of the blocked color moment feature and scheme of the retrieval based on blocked color histogram [1]. Lifang He et al. [2] proposed an effective EKH algorithm to search the optimal thresholds of color images at different levels in order to overcome the shortcomings of the multilevel threshold division algorithm. Worring et al. [3] carried out summarized research of the image retrieval technology based on the color, texture and features of shape; Deng et al. [4] conducted a study on image retrieval from color expression and description method. Hou et al. [5] describes the shape features of clothing with the invariant moments and descriptors and conducted the research of clothing image retrieval in combination with the color histogram.

But faced up with the huge space in market demand, a targeted ‘searching image by color’ system has not been developed that focuses on the fashion trend industry.

1.3. Opportunity moment
By combining the application of AI technologies with the methodology of fashion trends research, it not only can provide the basis for improving the accuracy of predicating color by means of constructing the database of fashion industry’s image; It can also optimize the existing model of fashion color verification and accelerate the speed of the prediction. It can also output the analysis results of fashion colors by visual methods, verify and analyze the color in the developing or developed process, and support or optimize the decision in designing color.

2. Construction of system

2.1. Flow of fashion color verification
In general, there are 9 steps in completing the verification of the known color of the brand company:
(1) Establishing the database of image within the clothing industry
(2) Classification of database information
(3) Color identification
(4) Comparison of color in the database
(5) Database proportion analysis
(6) Determination of color threshold
(7) Accurate extraction of color-coding data
(8) Comparison of time series
(9) Presentation of report

2.2. Establishment and retrieval of image database within fashion industry

2.2.1. Source of the database. The correct and systematic database source can better demonstrate the objectivity and accuracy of fashion color prediction. Figure 1 shows it includes four levels: social background, data of the consumers, data of products, and professional predicative institute.
- Level-1 source: social background
  It includes politics, economy, culture, art and design. These social backgrounds will fundamentally affect the pattern of the times, consumption power, ideology, and aesthetic form, etc.
- Level-2 source: source of consumer data
  Modern consumers like to share themselves through social media such as Instagram, influencers, twitter, Weibo, Bilibili, etc. These hot social media are one of the main sources of
consumer data now, followed by VOGUE street style, HARPERS BAZAAR Street style comes from consumer insight data sources from magazines and web.

- **Level-3 source: source of product data**
  According to the structure of the fashion industry chain, product data mainly comes from Fabric exhibitions, Fashion brands, Garment exhibitions, Runway, Fashion markets, etc. and related data from product fabrics, finished products to sales end.

- **Level-4 source: professional institute of predating trend**
  Many fashion trend forecasting agency have formed their own unique prediction contents and trends under the development and accumulation of long-term predictions, which are expressed in the form of professional consulting reports, forecasting report, sketchbooks, and Runway analysis.

![Figure 1. Source framework of databases.](image1.png)

### 2.2.2. Retrieval of labels of image data.

Figure 2 shows the storage of fashion images requires time labels, color labels, image source labels, and product type labels for classification and labeling.

![Figure 2. Index label for image database based on specific market.](image2.png)

- **Time labels**
  At the beginning of constructing database, it’s necessary to collect images according to the time series, which makes the readable time one of the important labels for images. For brand companies with different perception on color, different time cycles should be selected to meet
different retrieval requirements by days / weeks / months / quarters / half-year / one-year / two-year / ten-year units.

- **Color label**
  In this system, the prevailing color system is divided into two broad categories. One is the current common color-coding system, including RGB value, CMYK value, HEX value, LAB value and so on; The other category includes color labels from different color brands, including Pantone, NCS, HKS, DIC, and more.

- **Label of product types**
  It can be divided into fashion type, product type, space type and others. In which the fashion type can also be specifically subdivided into fashion, fashion accessories, make-up, home textile, etc. The product type can be subdivided into automobile, furniture products, designer toys and games, etc. The space field can be subdivided into architectural, interior design, retail space and installation art, etc. In addition, the content of the cross-domain such as graphic design, packaging, and the current art trading can also be part of the data source for platform design, packing and current arts fairs.

- **Labels of image source**
  According to the image source, the influence factor to the final brand can be determined.

2.3. **Construction of model**

Figure 3. Operational logic of model retrieval.

Input 1: Select the known the color card picture or color model information into the system.

Input 2: Select a database (use labels to filter high-impact, time-series-matching data and product types);

Operation 1: Analysis of color.

Operation 2: Output the data of the complete color card and color coding (coding of the different brands’ color card and RGB value, CMYK value, HEX value, LAB value).

Operation 3: Based on the key algorithm of ‘searching image by color’, all images with the approximate color / similar color in the database are extracted, and the images are arranged from high to low according to the color threshold as Figure 4.

Figure 4. Database output masonry layouts for the approximate / similar color image.
Operation 4: The share of the number of images with the color threshold in the database during the cycle is shown by the pie chart of the color ring.

Operation 5: The number of images with the color threshold and the proportion of other color pictures in the database are shown by the pie chart of the color ring.

Operation 6: By comparing with the same color threshold near to the pie chart of hue ring, the proposed color threshold and the specific color card are output as Figure 5.

Operation 7: By comparing with the color threshold in a longer period, the fashion trend of the color can be predicted.

2.4. Key technologies
Since the model of the retrieval system is based on the relationship between the color and image, the existing research is mainly based on the binary relationship between the color and image. In addition to the traditional way of inputting descriptive keywords to achieve the retrieval of image and color, the existing retrieval system of the color and image almost cover four key technologies of ‘searching image by image’, ‘searching color by color’, ‘searching color by image’ and ‘searching image by color’.

2.4.1. Searching color by image. The color tool web page Ginifab supports for ‘searching color by image’, users upload pictures. By specifying the different location of the picture, the color at the specified location is identified and converted into the code of RGB, CMYK, HEX color, which is shown in Figure 6.

2.4.2. Searching color by color. On the color tool website Qtcclor, users can input a color code of different forms or brands, in which they not only can get a color card picture, but also can get this
color code of other forms and it’s easy to enquiry. On this basis, there is a color scheme based on different color rules for professional reference and use, which is shown in Figure 7.

2.4.3. Searching image by color. Image trading site Shutterstock is a typical ‘searching image by color’ site with a searching function called Spectrum, users can search in combination with keywords and color. By controlling the color, select the image under the same color, which is shown in Figure 8.

2.4.4. Searching image by image. As Figure 9 shows that Google’s image retrieval tool can also complete the search by inputting images in addition to inputting the keywords. By inputting the existing images, the system can automatically associate the key words, and complete a series of operations, such as retrieval for the source of the existing images, and the similar picture, etc.

3. Experimental procedure
Use the retrieval system developed in this article to verify the popular colors of a certain clothing color with poor sales feedback.

3.1. Color input
The color card picture or color coding of the known color (HEX: # 1B2071) is introduced into the retrieval system, and the color card picture and color value are obtained. Not only will the color card image be shown in figure 10, but the different coding of the color can be provided: RGB: 27/32/113; CMYK: 100/0/43; HEX: #1B2071.

3.2. Filtering of database label
The labels are used to filter data and product types with high impact factors and time series matching. Since the brand company usually takes half a year as a cycle to launch products, 6 months can be
taken as a verification cycle. The picture source and the product type may be freely selected according to the decision-maker’s demand and the judgment.

3.3. Masonry layout and color analysis
According to the integration and induction of the database, the relevant pictures within the range of color threshold are obtained, which constitute a masonry layout for showing the color’s threshold. At the same time, it shows the proportion (HEX: # 1B2071) of the number of images in the database, as well as the proportion relationship of number of all images in color image. It shows that the known color (HEX: # 1B2071) has the lower proportion, however, there are more approximate / similar type color images under the same hue.

3.4. Matching results and auxiliary reference
Through the comparison and verification of the database, it can be found that the known color in the current quarter is not popular, the system provides color decision-making suggestions for brand company through two dimensions.

(1) Comparison of color range—refine the color performance of the color threshold, and to calibrate the known color of the brand, so as to generate an accurate threshold of the color. As shown in the figure 11-12, a color correction proposal is made by analyzing the approximate / similar color data of the brand’s known color of the hue ring in the database, and the specific color codes are output as color thresholds for reference, and verify them in the form of masonry layout of images.

(2) Comparison of time series—by expanding the range of the time cycle for the database and comparing with the color threshold of the last development quarter, the trend of the known brand color can be predicted. The output color threshold is compared with the previous period (the first 12-6 months) with 6-month time as a cycle, it is shown in figure 13. The threshold of fashion color output by the system shows that there is an obvious upward trend compared with the quantity in the last quarter in the image database based on the specific market, which means that the color is valuable in development.
3.5. Experimental results
The retrieval is carried out by inputting the company’s known colors (HEX: # 1B2071) into the system. The verification method of fashion color based on similarity algorithm is adopted to verify that the color deviates from the fashion trend at the market and that the poor sales data is caused by improper color decision. After verification, the fashion color threshold of its approximate color / similar color is obtained. Through the comparison based on the time series, it is verified that the fashion trend for the threshold of the output color is upward.

4. Comparison
Under the correct fashion industry database, comparing the popular color verification system based on the approximate search algorithm with the artificial color verification, it can be found that the former one is superior to manual in efficiency and accuracy, which is shown as Table 1.

Table 1. Comparison of fashion color verification system based on similarity search algorithm and artificial color verification.

| Prediction steps               | Artificial color verification                        | Fashion color verification system based on similarity search algorithm |
|-------------------------------|-----------------------------------------------------|-----------------------------------------------------------------------|
| Establishing the database     | It has a low efficiency, there is a risk of omission, it needs to be downloaded manually. | With the artificial intelligence, ten thousand images can be retrieved a day. |
| Classification of database information | The picture needs to be renamed and labelled. The efficiency is low. It needs to be roughly grouped by Hue. With the expansion of database, the grouping definition will gradually harder. | Image’s naming and labelling can be completed quickly and in batches with high efficiency. No need to repeat grouping manually, images can be automatically grouped by brand, time, label. |
| Color identification          | With the help of third-party platforms.              | No third-party platform as required.                                  |
| Comparison of color in the database | Compare the known colors with all the images in the database. On the basis of the first grouping, pick out all the pictures that contain this color. The efficiency is lower, the accuracy of human recognition is lower. | With the algorithm of ‘searching image by color’, the approximate / similar color can be extracted from database. |
| Database proportion           | Sum and calculate the number of images with the help of third-party platforms or devices. It is difficult to | Automatically calculate the proportion of the known color (including the approximate color / similar color) in the |
| Analysis | calculate the area ratio of color in the picture. | image database, and it is easier to calculate the proportion of the color in the image area. |
| --- | --- | --- |
| Determination of color threshold | The range of color threshold is filtered through manual correction with the help of the third-party platform. It needs to repeat, efficiency is lower, accuracy is lower. | Artificial intelligence collecting and sorting, extract the range of the color threshold with most value for reference, without the help of third-party platform. |
| Accurate extraction of color-coding data | Manual correction, there is the possibility of errors and it needs the help of third-party platform. | Artificial intelligence collecting and sorting, extract the range of the color threshold with most value for reference, without the help of third-party platform. |
| Comparison of time series | It needs to expand the workload of data collection and analysis steps. That is, for some groups it is more difficult to achieve reuse with the time. | It does not need consume over labors. |
| Presentation of report | It takes a lot of time in typography, but it is more flexible and can customize the report style for different audience. | After the format of report is set, it can be generated quickly as the data is updated and iterated over time. |

5. Conclusion
The fashion color verification system based on the similarity search algorithm proposed in this paper is based on the theoretical model of ‘searching image by color’. After constructing the database for the performance of pictures in the different fields at the specific markets, it controls time periodic variable, verify the rationality of brand company’s color decision and the tool to correct the deviation. The established model and the calculated results can verify the fashion of the known colors quickly and accurately. Not only can it adjust the market environment for brand color in the next quarter, but also predicate color trend in the next quarter, which can provide the accurate judgment in production, promotion, retail for different enterprises, and reduce inventory and increase profits.

Acknowledgement
Project Support by the Fundamental Research Funds for the Central Universities, Shanghai Style Fashion Design &Value Creation Collaborative Innovation Center & the Fundamental Research Funds for the Central Universities & Shanghai Summit Discipline in Design (DB18212) , Shanghai Education Commission & Shanghai Education Development Foundation Chenguang Plan (13CG74) .

Reference
[1] Bao Q 2017 D. Clothing Image Classification and Retrieval Based on Deep Learning Zhejiang University P 13
[2] He L and Huang S 2020 J. An efficient krill herd algorithm for color image multilevel thresholding segmentation problem Applied Soft Computing 89 1060-63
[3] Zhang D, Wong A, Indrawan M, et al. 2000 Content-Based Image Retrieval Using Gabor Texture Features Proceedings of Pacific-Rim Conference on Multimedia vol 1 (Sydney NSW Australia: University of Sydney) p 392
[4] Deng Y, Manjunath B, Kenney C, et al. 2001 J. An efficient color representation for image retrieval IEEE Transactions on Image Processing 10 140-147
[5] Hou A, Zhao L and Shi D 2010 Garment Image Retrieval based on Multi-features Proc. Int. Conf. on Computer, Mechatronics Control and Electronic Engineering vol 6 (Changchun) p 517