

# Content-Based Emotional Semantic Recognition and Retrieval of Male T-Shirt Images

Zhang Xiaomeng<sup>1, a</sup>, Zhang Haibo<sup>2, b, \*</sup> and Zhang Zekun<sup>1, c</sup>

<sup>1</sup>Information Center, Beijing Institute of Fashion Technology, Beijing, 100029, China

<sup>2</sup>Library, Beijing Institute of Fashion Technology, Beijing, 100029, China

<sup>a</sup>2540944916@qq.com, <sup>b</sup>hbdmzhh@126.com, <sup>c</sup>281323910@qq.com, \*corresponding author

**Keywords:** male T-shirt; emotional semantic; support vector machine; image recognition; image retrieval

**Abstract:** Emotional semantics identifying and retrieving of images is one of the hotspots in current research. It can provide certain basis for clothing design, clothing wearing and clothing matching. As a kind of male T-shirt, the recognition and retrieval of emotional semantics of images has also attracted much attention. On the basis of establishing the relationship between emotional semantic space and low-level features of T-shirt images based on the previous knowledge of clothing field, 40 T-shirt images were trained and a machine learning model was established by using support vector machine (SVM). The quantitative mapping relationship between emotional semantic space and low-level features of T-shirt images was established, which could be calculated automatically. The emotional semantics of the image is calculated to realize the emotional semantic recognition of the image. Similarity measure algorithm is also used to measure the emotional semantics of images, and image emotional semantic retrieval is realized. Through the experiment of software programming and development, a better recognition and retrieval effect have been achieved, thus verifying the feasibility and effectiveness of emotional semantic analysis of images combined with domain knowledge.

## 1. Introduction

As is known to all, information containing emotional semantics in images, for example, some images are “exciting”, some make people happy. As a type of clothing, male T-shirt images are also rich in emotional information, the recognition of these emotional information can provide certain basis for clothing design, clothing wear, clothing matching and so on. With the development of artificial intelligence technology, it has been promoted to one of the frontier hot research issues to let computers understand the emotional semantic content of images, so as to realize content-based emotional semantic recognition and retrieval of images. In this paper, based on the previous research on the factor space, the correlation between the factors and the low-level features of the emotional semantics of the male T-shirt image, the content-based emotional semantics recognition and retrieval of the male T-shirt image are realized by using support vector machine through programming.

## 2. Previous Studies

After subjective evaluation, 8 pairs of emotional words are determined for male T-shirt images, "simple - gorgeous", "calm - sportive", "formal - casual", "plain - fashionable", "classical-modern", "elegant - rugged", "gentle - strong", "implicit - flamboyant". Based on 8 pairs of sentiment description values, after the factor analysis process, the 3D emotion factor space of the male T-shirt sample image is built, the quantitative relationship between the three factors and the eight pairs of sentiment description values is shown in Equation (1), the formula for calculating the three factor values common to each image is shown in equation (2) [1]:

$$\begin{aligned}
E_1 &= 0.945f_1 - 0.056f_2 + 0.078f_3 \\
E_2 &= 0.203f_1 + 0.922f_2 + 0.152f_3 \\
E_3 &= 0.111f_1 + 0.962f_2 + 0.108f_3 \\
E_4 &= 0.914f_1 + 0.104f_2 - 0.189f_3 \\
E_5 &= 0.774f_1 + 0.494f_2 - 0.150f_3 \\
E_6 &= 0.159f_1 + 0.466f_2 + 0.823f_3 \\
E_7 &= -0.247f_1 - 0.047f_2 + 0.922f_3 \\
E_8 &= 0.829f_1 + 0.455f_2 + 0.006f_3
\end{aligned} \tag{1}$$

Among them,  $E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8$  in formula (1) represents the description value of 8 pairs of emotional words, and  $f_1, f_2, f_3$  represents the 3 factor values of the male T-shirt image.

$$\begin{aligned}
f_1 &= 0.413E_1 - 0.103E_2 - 0.156E_3 + 0.328E_4 + 0.188E_5 + 0.072E_6 + 0.033E_7 + 0.240E_8 \\
f_2 &= -0.259E_1 + 0.435E_2 + 0.487E_3 - 0.109E_4 + 0.130E_5 + 0.055E_6 - 0.158E_7 + 0.065E_8 \\
f_3 &= 0.193E_1 - 0.054E_2 - 0.106E_3 - 0.030E_4 - 0.101E_5 + 0.501E_6 + 0.620E_7 + 0.023E_8
\end{aligned} \tag{2}$$

Through the analysis of the relationship between the emotion factor and the low-level features of the male T-shirt image, the 11-dimensional (10-dimensional saturation-cold-warm fuzzy histogram plus 1D color contrast) feature, 257-dimensional (256-dimensional grayscale histogram plus) is considered. The 1D color contrast) feature, 4D (3D Tamura texture component parameter plus 1D average hue) feature can be used to characterize 3 factors [2]. The low-level feature extraction of emotional semantics is the key to image retrieval. Support vector machine technology and similarity measurement method are the theoretical and technical support for image retrieval.

### 3. Emotional Semantic Recognition and Retrieval Framework for Male T-Shirt Images

Support Vector Machine (SVM) [3] the image low-level feature vector and the high-dimensional emotion factor spatial map are fitted, and the degree of fitting affects the experimental results of image recognition. The realization of emotional semantic recognition and retrieval of male T-shirt images is to automatically calculate the actual factor values according to equations (1) and (2), and then calculate the emotional descriptor values and emotional factor prediction values of the images after machine learning. The calculation flow shown in Figure (1) and Figure (2) shows the image recognition and retrieval process in this paper.



Figure 1. Emotional semantic recognition process for male T-shirt images

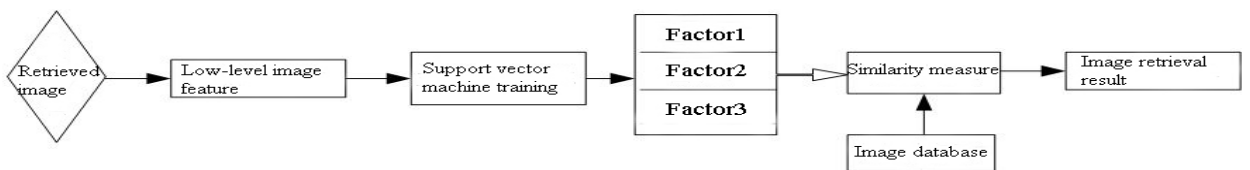


Figure 2. Emotional semantic retrieval process of male T-shirt images

Figure 1 and Figure 2 both extract the low-level features of 40 male T-shirt image samples, establish the mapping relationship between low-level features and emotional semantics through support vector machine, estimate new factor values to achieve image recognition, and subjectively with emotional words. Evaluation of search images is different. This paper finally realizes image

retrieval based on the similarity of sentiment factors.

#### 4. Support Vector Machine Technology

Relatively speaking, the excessive amount of image emotion semantic value calculation will affect the recognition efficiency of the image and affect the recognition effect of the image. SVM was first proposed by Corinna Cortes and Vapnik [3] in 1995. It is mainly used to solve the problems of less data, nonlinear fitting and high-dimensional pattern recognition. It can be added to machine learning for best solution. The application idea of SVM is to find an optimal hyperplane after mapping the function vector to the high-dimensional emotional factor space, achieve the optimal fitting in the low-dimensional space, predict the emotional factor value of the image, and realize the emotional semantic recognition of the image. SVM technology converts the original space into a nonlinear space and uses the penalty coefficient  $C$  to reduce the missing information of the fit [4-5]. In the process of machine learning, the SVM regression prediction model is built, and the measure of the fitting error in this learning or training process is called the penalty function. After the calculated image feature values are learned by different penalty coefficients  $C$ , the obtained training models will have different results. Therefore, it is necessary to go through multiple experiments to compare the effects of learning under different penalty coefficients, and then obtain the best effect under the number of times.

$$R(\omega, \xi, \xi^*) = \frac{1}{2} \omega \cdot \omega + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3)$$

In equation (3),  $\xi_i, \xi_i^*$  are relaxation factors. When there is an error in division,  $\xi_i, \xi_i^*$  are greater than 0, the error does not exist as 0, and the constant  $C > 0$  indicates the degree of punishment for the sample exceeding the error, which becomes a punishment coefficient. The linear fitting function is:

$$f(x) = \omega \cdot x + b = \sum_{i=1}^n (\alpha_i' - \alpha_i^*) x_i \cdot x + b \quad (4)$$

The best effect of image sentiment semantic recognition and retrieval needs to be added to the support vector machine technology through linear kernel function, polynomial kernel function, radial basis function and other functions. Nonlinear regression prediction is used to establish the model of machine learning, and then in high dimension. A space is transformed into a nonlinear fitting function. The linear fitting function is  $f(x) = \omega \cdot \Phi(x) + b(x, y_i)$ ,  $i = 0, 1, 2, \dots, n$ . The Radial Basis Function (RBF) used herein:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (5)$$

The representation formula of the nonlinear fitting function is [6]:

$$f(x) = \omega \cdot \Phi(x) + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x, x_i) + b \quad (6)$$

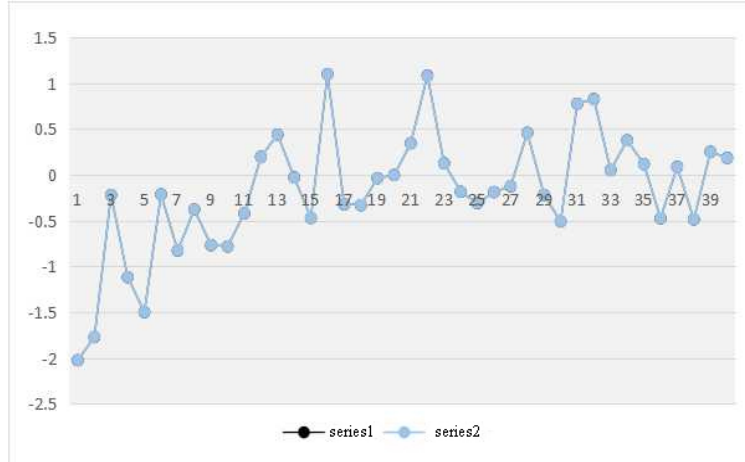
The basic principle of support vector machine regression training is to map the nonlinear vector to the high-dimensional factor space, and then perform the nonlinear regression training fitting in this space to realize the prediction of the image emotion factor.

#### 5. Emotional Semantic Recognition of Male T-Shirt Images

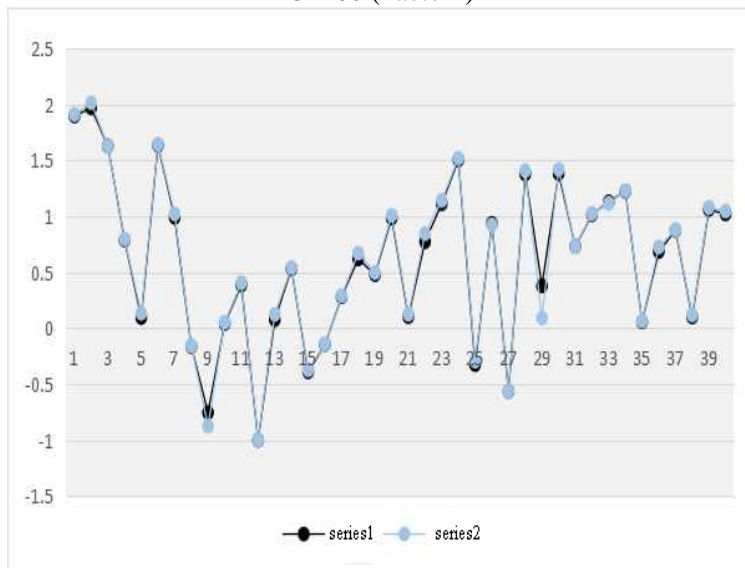
##### 5.1 Male T-shirt Sample Image SVM Regression Training Experiment

In the literature [2], the low-level features of emotional semantics of male T-shirt images are analyzed. Before realizing the image recognition of male T-shirt samples, 40 sample images need to

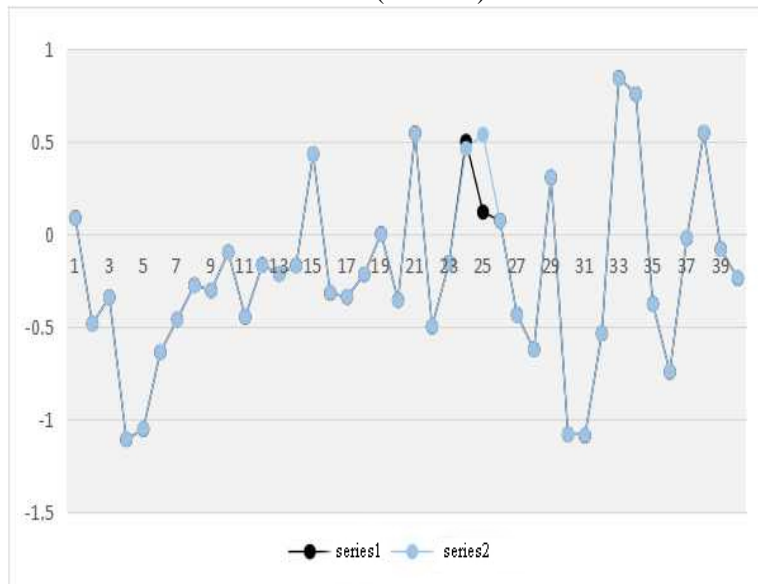
be subjected to SVM regression prediction training, and the experimental results are analyzed. The SVM regression prediction uses the developed LIBSVM [7] software. First, the low-level feature values of the 40 male T-shirt images are calculated as vectors, and input into the SVM for regression training analysis, using "C=100, 200, 300", respectively. The "400, 500, 1000" and other penalty coefficients were tested. The results showed that different SVM models were obtained for different times, and the new factor values of the male T-shirt images were also different.



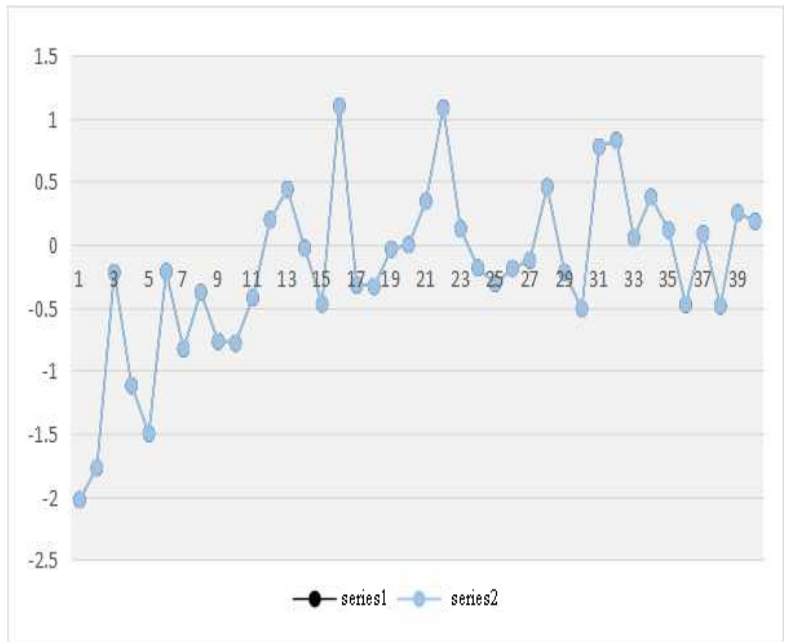
C=100 (Factor 1)



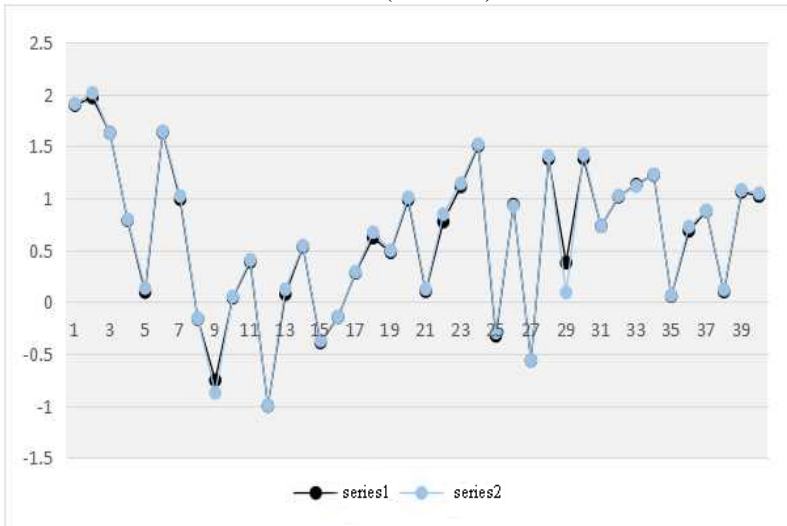
C=100 (Factor 2)



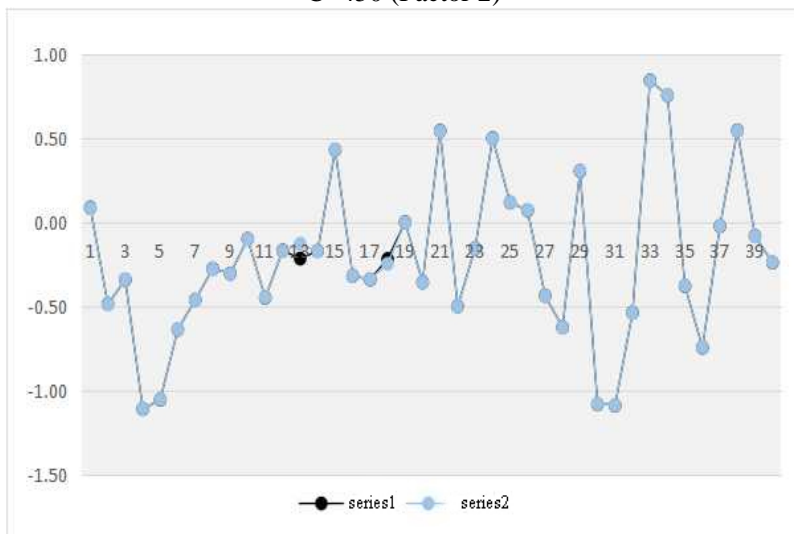
C=100 (Factor 3)



C=450 (Factor 1)



C=450 (Factor 2)



C=450 (Factor 3)

Figure.3 Prediction results of different penalty coefficient C training

In Fig.3, the black broken line indicates series 1, which is the actual emotional factor value quantified by subjective emotion description of 40 male T-shirt images, and the blue broken line

represents series 2, which is the number of training factors in the penalty factor "C=100" and "C= In the case of 450", the predicted value obtained after SVM training. As can be seen from the figure, the regression accuracy of training 450 times is higher. On the whole, the regression training of C=450, the two polylines of the three factors are basically coincident. It can be seen that the rationality and effectiveness of the selected low-level characterization methods also indicate the rationality of the results analysis.

### 5.2 Emotional Semantic Recognition of Male T-Shirt Images

According to FIG. 3, after the SVM model is established, the model can be used to perform emotional semantic recognition on the male T-shirt image. In Fig.4, click "Import Picture", after the image is imported, the system can automatically estimate the three factor values of the figure. Then click "image emotion recognition" to automatically calculate the eight pairs of sentiment description values of the image, and finally put the three factor values of the image into the library. Finally achieve the purpose of image recognition and save data.



Figure.4 Male T-shirt sample image emotional semantic retrieval system page

### 6. Emotional Semantic Retrieval of Male T-shirt Images

After the above analysis, it is concluded that the training effect of "C=450" is better. Some experts and scholars study image retrieval. It is common to use image sentiment description words to retrieve similar images. This paper is based on the new factor value of image recognition estimation. Retrieve images with similar emotional semantics. It can be known from equations (1) and (2) that the three factor vectors are independent of each other, and the similarity of image sentiment semantics can be determined by the degree of distance difference between vectors, if the interval between vectors is larger, The worse the emotional semantic similarity of the image. After reading the literature, collecting data, and defining the emotional semantic similarity metrics of male T-shirt images according to the definition of various distance functions, see equation (8) [8-10].

$$Sim = \frac{a \cdot b}{|a| \cdot |b|} \cdot \frac{\min(|a|, |b|)}{\max(|a|, |b|)} \quad (7)$$

In equation (8), the similarity measure ranges from [-1, 1], where is the coordinate of the image sample in the sentiment factor space, representing the angle cosine distance value between the image vectors, representing the two coordinates respectively. Vector mode, where the formula also represents the difference between the two vector modes.

As shown in Figure 5, click "Import Picture" to import the retrieved male T-shirt image, and the system will automatically display the emotional factor value of the image (the new factor value after recognition). When the image is retrieved, we will sort the similarity of the output, then the

ranking is set to "0.6" in the manner of emotional semantic similarity, and the image similar to the retrieved image is output at 0.6 or more according to the similarity.

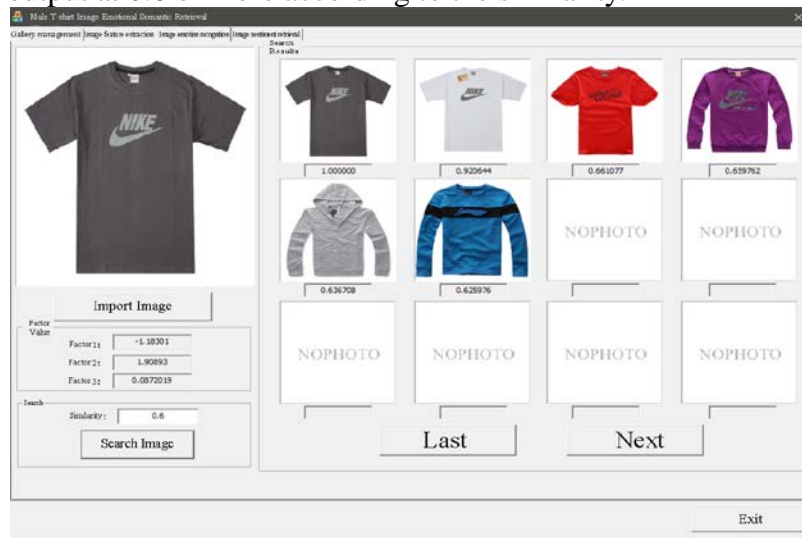


Figure.5 Male T-shirt image emotional semantic retrieval page

The similarity measure algorithm is a method applied to image retrieval. Different similarity measures, the number of images retrieved will be different. In order to test the effect of the emotional semantic retrieval of the male T-shirt image, the images were retrieved with the "0.6", "0.7", and "0.8" similarity values according to the steps of the search flow chart shown in FIG. During the search process, the first image displayed first on the right side of the system page is the "retrieved image", and then similar images are sequentially retrieved in descending order of similarity. The search result of the image is analyzed, and in the statistics of Table 1, the unreviewed image is excluded from the case of "only the retrieved image is retrieved".

Table.1 Emotional Semantic Retrieval Statistics of Male T-shirt Images

C	Similarity	Statistics(40)	Search Rate
100	0.6	31	77.5%
	0.7	30	75%
	0.8	26	65%
450	0.6	30	75%
	0.7	29	72.5%
	0.8	24	60%

As shown in Table 1, the results of the retrieval of male T-shirt images are different under different penalties and different similarities. It can be seen from the table that the similarity degree "0.6" has the highest retrieval rate; the "0.8" similarity image is less, but it exceeds 1/2 of the sample image, indicating that the method of realizing emotional semantic retrieval of male T-shirt images is effective, reasonable.

## 7. Conclusion

The establishment of the emotional factor space of male T-shirt images and the correct extraction of low-level features, using support vector machine technology and similarity measurement algorithm, through software programming, the emotional semantic recognition and retrieval of male T-shirt images. After verification, the recognition and retrieval effect is better, which provides some research experience for the research of images, especially the emotional semantic recognition and retrieval of clothing images. In short, under the premise of combining domain knowledge, the semantic analysis of the domain image, and then using the machine learning algorithm for training and testing, content-based image emotional semantic recognition and retrieval will get an ideal effect. Based on the research in this paper, the sample size of the image of the male T-shirt can also be increased to further improve the recognition and retrieval effect.

## Acknowledgments

The work described in this article was supported by grants from the Science and Technology General project of Beijing Municipal Commission of Education (No.AJ2016-11).

## References

- [1] H. B. Zhang, C. Y. Zhang. Study on the Emotion Factor Space of Men's T-shirt[C]. 2016 International Conference on Information Science and Control Engineering, Beijing, July, 2016.
- [2] X. M. Zhang, H. B. Zhang. Emotional Semantic Analysis of Male T-shirt Image Based on Color and Texture Features. 6th International Conference on Economics, Management, Education and Computer, Nov.17-18, 2018.
- [3] V. Vapnik. The Nature of Statistical Learning Theory [M]. New York: Springer, 1995.
- [4] A. J. Smola, B Scholkopf. A Tutorial on Support Vector Regression [J]. Statistics and Computing Archive, 2004, 14(3): 199-222.
- [5] N. Cristianini, J. S. Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods [M]. Beijing: Publishing House of Electronics Industry, 2004.
- [6] H. B. Zhang, T. J. Huang, L. Liu, ect. Fabric Image Emotional Semantic Recognition Based on SVM [J]. Tianjin Polytechnic University, 2013, 12, 32(6):23-27.
- [7] C. C. Chang, C. J. Lin. LIBSVM: a Library for Support Vector Machines [J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3):A2701-A2727.
- [8] W. N. Wang. Research on Image Feature Extraction, Retrieval and Classification Based on Emotional Semantics [D]. Guangzhou: South China University of Technology, 2005.
- [9] M. Ling, H. Zhang. Bilinear Image Similarity Matching Algorithm Based on Deep Feature Analysis [J]. Journal of Computer Applications, 2016, 36(10):2822-2825+2831.
- [10]H. B. Zhang. Emotional Semantic Recognition and Retrieval of Clothing Image Content-based [M]. Beijing: China Textile & Apparel Press, 2014, 4:69-99.