Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **4**, July **2018**

The Classification and Retrieval of the Image Affective Semantics Based on Integration of Multi Features and SVM

Hui Chen

College of Electronics and Information Science
Fujian Jiangxia University
Room 4#608,Jincheng Garden,xindian,Fuzhou, 350012, China
63292936@qq.com

Lin Xu

Innovative Information Industry Research Institute
Fujian Normal University
Fuqing Brunch of Fujian Normal University
Campus County 1#, Longjiang Road, Fuqing, Fuzhou, 350300, China
xulin@fjnu.edu.cn

Fuquan Zhang*

School of Computer Science & Technology
Beijing Institute of Technology
Fujian Provincial Key Laboratory of Information Processing and Intelligent Control
Minjiang University
Room A9#301, Jiyuan, Juyuan, Jinshan, Fuzhou, 350008, China
8528750@qq.com

Received November, 2017; revised March, 2018

Abstract. *Images contain rich emotions. The traditional image classification and retrieval methods mainly classify images according to the bottom visual features of the images. Most of them ignore the influence and effect of emotions. The image affective classification and retrieval method in this paper combines the bottom visual features with the high-level affective semantics. Aiming at the insufficiency of the retrieval based on the single bottom visual features, it combines the bottom image color feature with the shape feature, and establishes the image feature space. Aiming at the semantic gap between low-level visual features and high-level semantics of the image, it establishes the emotional semantic space through cognitive psychology experiments and factor analysis. The method in this paper uses SVM to fill the semantic blanks to achieve the affective semantics annotation of the image, maps feature space to the emotional space of the image to realize the classification and retrieval of the image affective semantics. The result of the experiment shows that this method improves the efficiency and performance of image retrieval, and provides an effective solution to the affective semantics research of images.*
Keywords: Image retrieval; Color feature; Shape feature; SVM; Emotional space

1. **Introduction.** With the rapid development of diversified information technology, various types of multimedia emerge constantly. Images are one of the most common types of media. Their number is growing at a geometric exponential way. With the approach

of "big data" era, most scholars pay close attention to the research on how to make accurate and fast image storage and retrieval[1][2]. At present, the classification and retrieval techniques of images are mainly divided into three categories, namely text-based image retrieval, content-based image retrieval and semantics-based image retrieval. Text-based image retrieval is mainly through manual annotation of images, using text retrieval technology for retrieval. Although the retrieval technology is simple, but the main disadvantage is that the manual annotation workload is too large and inaccurate. In early 1990s, content-based image retrieval (CBIR) has become the mainstream, CBIR uses the low-level visual features such as color, texture, shape, and so on to do the image retrieval[3]. The judgment of image similarity not only depends on the image visual features, but also depends on the understanding of the image content, namely the semantic features of the image expression. The semantics-based image retrieval is still in the exploratory stage. Colombo et al established a series of rules of mapping image visual content to image semantics[4], Nadia explored the matching relationship between low-level features and high-level affective semantics from the angle of multimedia data mining[5]. In October 2005, the first International Conference on affective computing and intelligent interaction was held in Beijing, China. The famous experts and scholars home and abroad in this field gathered together and studied the affective computing, artificial emotion and artificial psychology. In September 2011, Chinese HeFei University of Technology held the first forum of "affective computing and advanced intelligent machines", demonstrating the new developments in the frontiers such as the research and application of affective computing in advanced intelligent machines, which effectively promoted the research of the frontier of affective computing and intelligent interaction. In addition to its own original meaning, the image also has a lot of emotional information that can be mined. The semantic gap [6][7][8] between low-level visual features and high-level rich semantics is the difficult point of image classification and retrieval. To establish affective semantic space and annotate image affective semantics can effectively solve the problem of semantic gap. The effective interpretation and description of image emotions is the key to image emotion classification and retrieval. In many basic features of the image, color and shape features are closely related to human emotions. Some psychological researches show that colors have distinctive association and artistry, and have direct influence on people's psychology and emotions. Shape feature is not affected by the change of target and background. Line shapes and angles can also reflect the affective semantics of the image. In this paper, the color and shape features of the image are combined effectively, and SVM is used to annotating the affective semantics of the image to realize the image affective classification and retrieval.

2. **Multi Features Integration.**

2.1. **Extraction of Color Features.**

To extract color features needs to choose the appropriate color space first. The three elements of color hue, saturation and brightness correspond to three components of HSV color space, and accord with color understanding of human visual perception. The non-uniform quantization of HSV color space can save space and reduce complexity. Visual perception is most sensitive to the change of hue H, so the quantization number of this value is set to 9, and the saturation S and brightness V are set to 3 and 1 respectively. A feature vector L is formed. $L = 9H + 3S + V$, and the numerical range is [0,1,..., 71]. Quantization process is shown in Figure 1.

Color features can be described by color histograms as:

$$H = \begin{cases} 0 & \text{if} \quad h \in [316, 20], \\ 1 & \text{if} \quad h \in [21, 40], \\ 2 & \text{if} \quad h \in [41, 75], \\ 3 & \text{if} \quad h \in [76, 155], \\ 4 & \text{if} \quad h \in [56, 190], \\ 5 & \text{if} \quad h \in [191, 270], \\ 6 & \text{if} \quad h \in [271, 295], \\ 7 & \text{if} \quad h \in [296, 315], \end{cases} \quad S = \begin{cases} 0 & \text{if} \quad s \in [0, 0.2], \\ 1 & \text{if} \quad s \in [0.2, 0.7], \\ 2 & \text{if} \quad s \in [0.7, 1], \end{cases} \quad V = \begin{cases} 0 & \text{if} \quad v \in [0, 0.2], \\ 1 & \text{if} \quad v \in [0.2, 0.7], \\ 2 & \text{if} \quad v \in [0.7, 1]. \end{cases}$$

Figure 1. Quantization Process

$$H = (h[c_1], h[c_2], \cdots, h[c_n]), \quad \sum_{k=1}^{n} h[c_k] = 1, \quad 0 \le h[c_k] \le 1 \tag{1}$$

The pixel number of the kth color is represented by h[Ck]. The Euclidean distance method is used to measure the similarity of color histogram between two images. Iq is used to query the image. Iq represents the target image. The formula is as follows:

$$D_h = \left(\sum_{i=1}^{n} (h_q[c_i] - h_t[c_i])^2 \right)^{1/2} \tag{2}$$

2.2. **Extraction of Shape Features.** In this paper, the quaternary tree block method is used to extract the shape feature of the image. Each sub block image feature reflects the local feature of the image. The quaternary tree block method takes the information of the spatial relationship of image into consideration, and improves the overall features of the image through analysis and description of the local feature of space.

In image segmentation, the quaternary tree block decomposition method is used from coarse to fine, that is, the image center is determined as the center of gravity, the horizontal axis and the vertical axis divide the image B_0 into four sub blocks first, and then each sub block is decomposed in the same way, as shown in Figure 2.

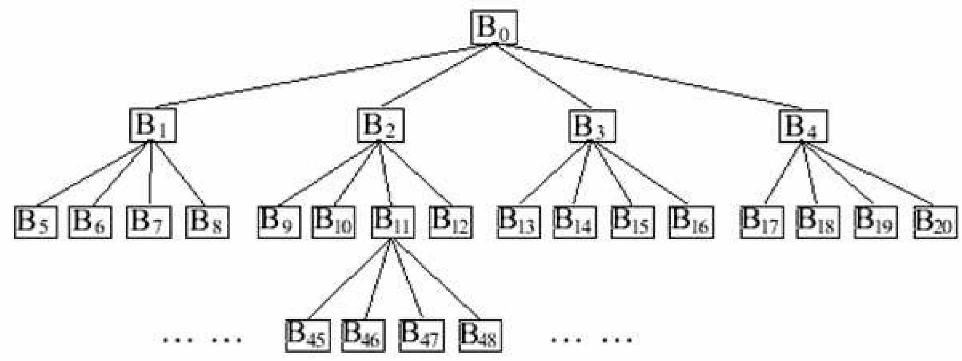

Figure 2. The Quaternary Tree Block Decomposition

Finally, two shape features of Hu invariant moments and information entropy are extracted from each sub block. It is defined as follows:

The central moment of $p+q$ order of the digital image $f\left(x,y\right)$ is defined as:

$$u_{pq}=\sum_{x=1}^{M}\sum_{y=1}^{N}\left(x-\bar{x}\right)^{p}\left(y-\bar{y}\right)^{q}f\left(x,y\right)\ \ ,p,q=0,1,2\cdots \tag{3}$$

The information entropy of the sub block image is defined as follows:

$$H(p_0,p_1)=-p_0\log p_0-p_1\log p_1 \tag{4}$$

$$p_1=\sum\sum F(x_i,y_j)/Size(i,j)p_0=1-p_1 \tag{5}$$

p_1 and p_0 are derived from the geometric distribution of image pixels, $F\left(x,y\right)=1$ represents the actual pixels of the image, $F\left(x,y\right)=0$ represents the blank pixels. If $F=\{F_0,F_1\cdots,F_i,\cdots\}$stands for the shape feature sequence, I and J stand for the image to be retrieved and the query image respectively, and M is the total number of hierarchical sub blocks, the similarity between images is calculated by Euclidean distance:

$$d_2\left(I,J\right)=\sum_{k=0}^{M}w_k\left|F_k\left(I\right)-F_k(J)\right| \tag{6}$$

2.3. **Feature Integration Process.** Since there is no direct comparability between the different underlying features, this paper uses Gauss normalization method to normalize the features, as shown in the following formula:

$$t'(i)=\left[\frac{t\left(i\right)-\mu}{3\sigma}+1\right]/2 \tag{7}$$

In the formula,$t\left(i\right)$ represents the first feature of the image; $t'(i)$ represents the first normalized feature, σ is the standard variance, and μ is the mean value of features. The similarity measure of color feature and shape feature after normalization are set as $D_1\left(I,J\right)$ and $D_2\left(I,J\right)$, the similarity weight of color feature and shape feature are set as w_1 and w_2 Aand the total similarity of image is set as $D\left(I,J\right)$,shown as the following formula:

$$D(I,J)=w_1D_1(I,J)+w_2D_2(I,J) \tag{8}$$

The feature integration process is illustrated in figure 3:

3. **Emotional Semantic Annotation Based On SVM..**

3.1. **The Establishment Of Emotional Space.** The analysis thought of dimension space put forward by psychologists can be applied to the study of emotional semantics, and different emotions in space correspond to emotion vectors, so as to establish the emotion space[9-10]:

1. According to the image characteristics of perceptual database, the adjectives describing the psychology are collected to prepare for the subsequent quantitative experiments. As shown in Table 1, 10 pairs of antonyms of adjectives are set carefully:

2. To carry on an experiment of semantic quantification, select sample images to conduct a detailed questionnaire survey. Take “warm - cold” for example, in the questionnaire survey, the users have 5 levels of options: warm, a little warm, neutral, a little cold and cold. By collecting the experimental data, the user perception database is formed.

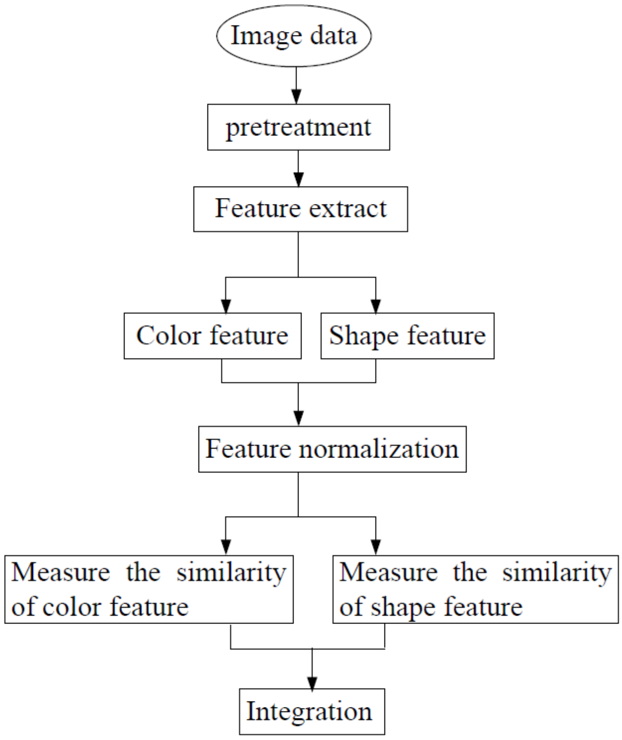

Figure 3. Feature Integration

Table 1. Antonyms of Adjectives

Bright - Dark	Soft - Rigid
Simple - Luxury	Lively - Serious
Comfortable - Uncomfortable	Tiny - Huge
Warm - Cold	Excited - Calm
Melancholy - Cheerful	Dry - Wet

3. Use factor analysis to analyze data and establish emotional space. Set Z_{mn} as user K's nth adjective evaluation of the image m. Calculate the normalized value X_{mn} and the average value Y_{mn} according to Formula (9) and obtain the matrix X.

$$y_{mn} = \frac{1}{K}\sum_{k=1}^{K} Z_{mnk}\ ,\ x_{mn} = \frac{y_{mn} - \overline{y_n}}{s_n} \tag{9}$$

The factor analysis is applied to matrix X to obtain the load matrix A and the common factor F, shown as formula (10):

$$X = FA + UD \tag{10}$$

If the dimension is reduced from N dimension to L dimension, the coordinate of the image m in the emotional space corresponds to the line m of the matrix F. fm=(fm1 ,fm2 ,K, ..., fmL), and the coordinate of the adjective n in the L dimensional space corresponds to the line n of the matrix A. an=(an1 ,an2 ,K, anL) [11].

3.2. **Support Vector Machine (SVM).** Support vector machine proposed by Vapnik and Cortes in 1995 is based on statistical theory. The main idea of support vector machine (SVM) is to establish a classification hyperplane as the decision surface, which maximizes the separation boundary between the positive and negative examples, and solves the problem of small sample, nonlinear and high-dimensional recognition [13][14]. It is a good classifier. The specific algorithm is as follows:

Given a sample training assemblage $X = \{x_i, y_i\}$,$x_i \in R^n$ is the n-dimensional input vector, and $y_i \in \{-1, 1\}, i = 1, 2, \ldots, m$ is the output classification identifier. Assuming that the training assemblage is linearly separable, all the plane assemblages which can separate the two different types of samples are $y_i (wx_i + b) \geq 1, i = 1, 2, \ldots, m$.

Of all the split planes, those that can classify the heterogeneous samples and maximize the classification interval are Optimal Separating Hyperplane. The sample point closest to the hyperplane is called support vector, and the distance to the hyperplane is $1/\|w\|$ .The hyperplanes that meet the condition $y_i (wx_i + b) \geq 1, i = 1, 2, \ldots, m$, and make $\frac{1}{2}\|w\|^2$ the minimum is the optimal hyperplane. Lagrange optimization method is used to turn the optimal classification plane problem into its dual problem:

$$\max Q(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1,j=1}^{n} a_i a_j y_i y_j (x_i \cdot x_j) \tag{11}$$

at the same time satisfy the constraint condition: $\sum_{i=1}^{n} y_i x_i = 0$ and $a_i \geq 0, i = 1, 2, \cdots, n$ The *Lagrange* multiplier corresponding to each sample is a_i. After solving the above problem, the optimal classification function is obtained:

$$f(x) = \operatorname{sgn}\left\{ \sum_{i=1}^{n} a_i y_i (x_i \cdot x) + b \right\} \tag{12}$$

For nonlinear samples, the nonlinear problem is transformed into a linear problem in the high dimensional space, and an optimal hyperplane is constructed in the high dimensional space. Support vector machine (SVM) solves the linear constrained two programming problem by using proper kernel function $K(x_i, x_j)$. The formula (11) is converted into:

$$Q(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1,j=1}^{n} a_i a_j y_i y_j K(x_i, x_j) \tag{13}$$

The corresponding SVM optimal classification function is:

$$f(x) = \operatorname{sgn}\left\{ \sum_{i=1}^{n} a_i y_i K(x_i, x) + b \right\} \tag{14}$$

The selection of kernel function in SVM method affects the generalization ability of the classifier. After studying and comparing, Gauss kernel function has better learning effect in the image retrieval.

3.3. **Affective Semantic Annotation.** When the new image is input, the affective semantic annotation of the image[15][16] is realized through SVM outputing the corresponding value of the image in the emotional space. The main steps of emotional annotation are shown in Figure 4:

4. Experiment on Emotional Classification and Retrieval of the Images.

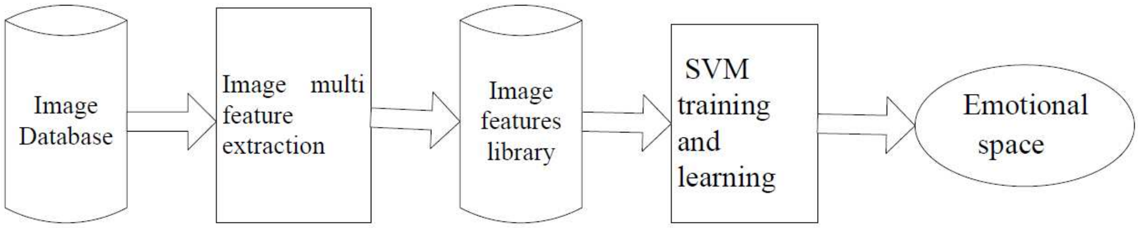

Figure 4. Emotional Annotation

4.1. **System Block Diagram.** The image library is composed of 300 color images, and the images are pretreated with format conversion and denoising. 50 images are selected from the image library as training samples, and the remaining 250 are used for testing. Through the questionnaire survey of the five-grade evaluation method, 200 respondents participate the emotional evaluation of the images, and the perceptual information database of the image is set up for training SVM to collect data. The system diagram of emotional classification and retrieval of images is shown as figure 5:

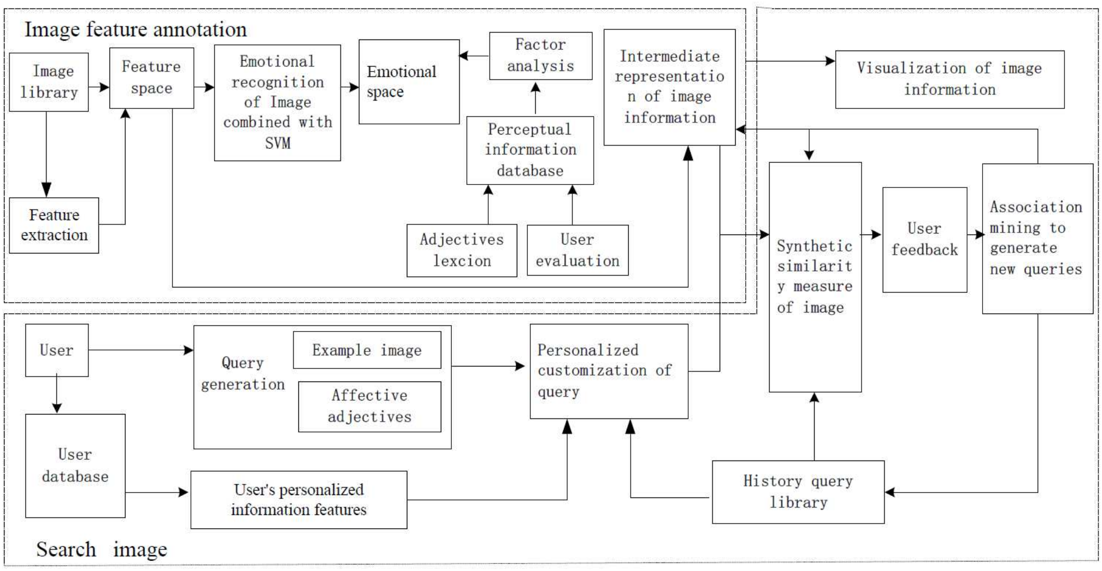

Figure 5. System Block Diagram

The retrieval algorithm flow of this paper is as follows: (1) Color histogram is used to extract the color feature of the image, and quaternary tree partition method is used to extract the shape feature of the image. (2) The multiple feature integration method is used to form an integrated feature vector, which is set as the input vector of SVM. (3) SVM is used to annotate the image affective semantics to realize the classification and retrieval of image affective semantics. (4) The system performance evaluation method is used to compare the retrieval results and user evaluation results, verifying the accuracy of the algorithm. (5) Compare the retrieval results of single feature method and the method proposed in this paper to verify the effectiveness of the algorithm.

4.2. **Analysis of Experimental Results.** Recall and precision are commonly used system performance evaluation methods. Taking the two adjectives "bright" and "melancholy" for example, the retrieval result shows the first 12 images closest to the adjectives, as shown in figures 6 and 7. It can be seen from Table 2 that the retrieval result of the

TABLE 2. Comparison of Retrieval Results

Affective Semantics	Precision ratio	Recall ratio
Bright	89.7	71.2
Melancholy	78.3	64.3

adjective “bright” has higher precision and recall, and the user satisfaction is higher. Due to the subjectivity and complexity of affective features, some dark images in the image library also give the impression of melancholy, but are not selected.

FIGURE 6. Example of “Bright” Images Retrieval

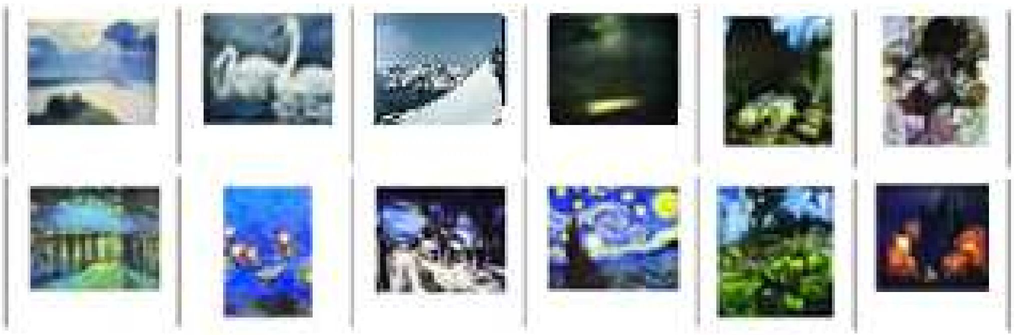

FIGURE 7. Example of “Melancholy” Image Retrieval

The retrieval result of single color feature and single shape feature method is compared with the retrieval result of the method of this paper. The average recall and average precision of the 30 retrieval experiments are as shown in Figure 8 and Figure 9. The retrieval method of this paper integrates color, shape and semantic features, which is better than the single color feature and single shape feature retrieval method. The result indicates the effectiveness of the method of this paper.

5. **Conlusion.** The retrieval combining human affective features and image low-level features has become a hot research topic in the field of image retrieval and artificial intelligence. This paper effectively combines the color features and the shape features, using SVM to fill the semantic blanks to achieve the affective semantics annotation of the image, and maps feature space to the emotional space of the image, realizes the classification and retrieval of the affective semantics of the image. Meanwhile the performance evaluation method is used to verify the effectiveness of the algorithm. The experiment result shows

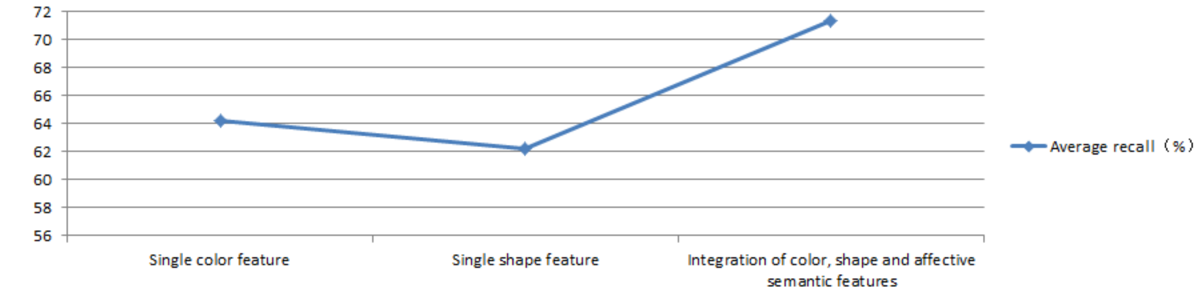

Figure 8. Comparison of Average Recall

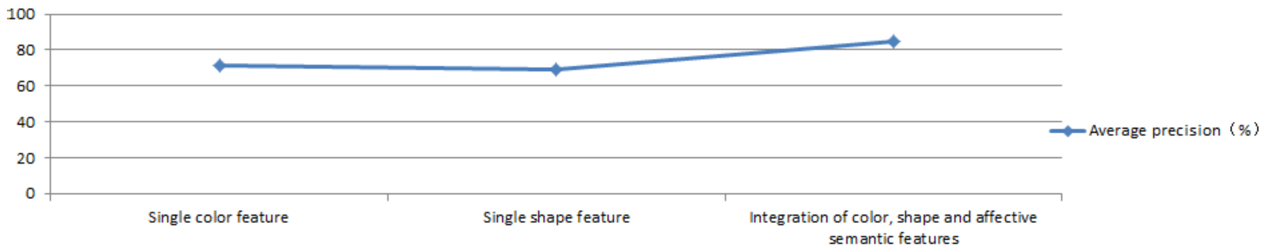

Figure 9. Comparison of Average Precision

that this method improves the retrieval efficiency and performance, and is a useful exploration of the affective semantics research of the image. The research on image classification and retrieval based on affective semantics, including the subjects of computer vision, artificial intelligence, cognitive science, aesthetics and psychology, is a brand-new field of interdisciplinary research. The research of this field has broad application prospects and economic value, and is of great practical significance.

Acknowledgment. This material is based upon work supported by the Research Program Foundation of Minjiang University under Grants No. MYK17021 and supported by the Major Project of Sichuan Province Key Laboratory of Digital Media Art under Grants No. 17DMAKL01 and supported by Fujian Province Guiding Project under Grants No. 2018H0028 and also supported by Fujian Education Department Class A Science and Technology Project under Grants No. JA14321.

REFERENCES

[1] E. G. Karakasis, A. Amanatiadis, A. Gasteratos, et al. , lmage moment invariants as local features for content based image retrieval using the Bag-of-Visual-Words model, *Pattern Recognition Letters (PRL)*, vol. 55(C), pp. 22-77, 2015.
[2] G. H. Liu,J. Y. Yang,Z. Y. Li, Content-based image retrieval using computational visual attention model, *Pattern Recognition(PR)*, vol. 48, no. 8, pp. 2554-2566, 2015.
[3] Anu Bala,Tajinder Kaur, Local texton XOR patterns:a new feature descriptor for content-based image retrieval, *Engineering Science and Technology, an lnternational Journal (JESTECH)* , vol.19, no. 1, pp. 101-112, 2016.
[4] C. Colombo, A.D.Bimbo, P.Pala, Semantics in Visual Information Retrieval, *IEEE Multimedia*, vol. 6, no. 3, pp. 38-53, 1999.
[5] N. Bianchi-Berthouze, Mining multimedia subjective feedback, *Journal of Intelligent Information System*, vol. 19, no. 1, pp. 43-59, 2002.
[6] X, Y. Wang, L. L. Liang, W. V. , et al., A new SVM-based relevance feedback image retrieval using probabilistic feature and weighted kernel function, *Journal of Visual Communication and lmage Representation*, vol. 38, pp. 256-275, 2016.

[7] M. K. Kundu, M. Chowdhury, S. R. Bulo , A graph-based relevance feedback mechanism in content-based image retrieval, *Knowledge-Baced Systems(KBS)*, vol. 73, pp. 254-264,2014.

[8] J. C. Pereira, N. Vasconcelos, Cross-modal domain adaptation for text-based regularization of image semantics in image retrieval systerns, *Computer Vision and lmage Understanding(CVIU)*, vol. 124, pp. 123-135,2014.

[9] J. Chen, Analysis Technology of Image Affective Semantics, *Beijing: Publishing House of Electronics Industry*, pp. 13-15, 2011.

[10] J. J. Chen, Analysis Technology of Image Affective Semantics, *Beijing: Publishing House of Electronics Industry*, pp. 13-15, 2011.

[11] Z. R. Liu,X. L. Yu, Research on Linguistic Computing Model Ior Image Emotion Semantic, *IEEE SystemsMan&Cybernetics Society*, vol. 19, pp. 3490-3493, 2011.

[12] H. Chen, Perceptual Image Retrieval Based on Color and Affective Semantic Features, *Journal of Putian University*,(2013,20), no. 5, pp. 70-73.

[13] Q. Zhang, T. Yang , S. Yang, Robust twin support vector machine for pattern classification, *Pattern recognition*, vol. 46, no. 1, pp. 305-316,2013.

[14] Z. Qi, T. Yang Y, S. Yang, Structural twin support vector machine for classification, *Knowledge-based systems*, vol. 43, no. 2, pp. 74-81,2013.

[15] S. Wang, F. Chen, Z. Wang, et al., Research on Image affective Semantic Annotation and Retrieval Algorithm Based on Support Vector Machines, *Pattern Recognition and Artificial Intelligence*, vol. 17 , no. 1, pp. 27-32, 2004.

[16] Y. L. Zhu, J. Y. Yang, J. Xu, et al., Multi-Feature Image Retrieval Involving Emotion and Semantics, *Journal of Xi'an Engineering University*, vol. 25, no. 4, pp. 527-532, 2011.