2006 IEEE International Conference on
Systems, Man, and Cybernetics
October 8-11, 2006, Taipei, Taiwan

Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction

Wang Wei-ning(corresponding), Yu Ying-lin and Jiang Sheng-ming

Abstract—**Image emotional semantic research is a promising and challenging issue. This paper analyzes the emotional space and builds a novel scheme that can annotate the image emotion semantic automatically and realize emotional image retrieval. Based on psychological experiments evaluating evoked feelings for art paintings, we first identify an orthogonal three-dimension emotional factor space of images through 12 pairs of emotional words. Then, the following three novel image features are designed for each emotional factor to predict it. They are luminance-warm-cool fuzzy histogram, saturation-warm-cool fuzzy histogram integrated with color contrast and luminance contrast integrated with edge sharpness. The values of emotional factors can be predicted from the image features automatically by using support vector machine of regression. Finally, we design and implement an emotion-based image retrieval system, which enables the users to perform retrieval using emotional semantic words. Experimental results show the effectiveness of our model.**

I. Introduction

CONTENT-BASED Image Retrieval (CBIR) system supports image search based on perceptual features such as color, texture and shape. However, users prefer using keywords at semantic level to using low-level features (which is more efficient) to conduct search [1]. Accordingly, some semantic-based retrieval systems have been proposed following CBIR in recent years with lots of progress reported in the literature [2]. However, most of them overlook the effect of emotions. Colombo et al. [3] proposed an innovative method to get a high-level representation of images and videos by aggregating low-level information according to a set of rules and transforming it into emotional level phrases. Researchers also attempted to build image retrieval systems with impression words [4]-[6] and investigated emotional vocabularies and semantics [7]-[10]. Although semantic description for emotion has become remarkable in recent years, studies are still required in this field.

This paper analyzes the emotional space and proposes a novel scheme to annotate the image emotion semantic automatically and realize emotional image retrieval. Based on psychological experiments measuring evoked feelings by art paintings, we first identify an orthogonal three-dimension emotional factor space of image through 12 pairs of emotional words. Then, three novel image features are designed for each emotional factor to predict them. By using Support Vector Machine of Regression (SVR), the values of emotional factors can be predicted based on image features automatically. Finally, an emotion-based image retrieval system is designed and implemented, in which the users can perform retrieval using semantic words. Experimental results demonstrate the effectiveness of our model.

The rest of the paper is organized as follows. The psychological experiments and the data analysis are described in detail in Section II. Section III describes the three novel image features and the methods to evaluate the factor values using the image features. The image query scheme and the experimental results are presented in Section IV followed by the conclusion and future work in Section V.

II. Emotional Factor Space Construction

Adjective words are required to describe image emotions. From the statistical analysis, we first collected 12 emotional word pairs. Then in a psychological experiment using semantic differential(SD) technique[11][12], which is often used to evaluate human feelings evoked by inducements, 42 observers assessed images on the 12 emotional word pairs to build up an orthogonal three-dimension image emotional factor space.

A. Experiments Design

Emotional word pair selection: Emotional adjectives are widely used in emotion processing studies while different researches select different words. From the previous studies [4]-[6] [13] and some dictionaries, we collect 150 emotional words. After refining, these emotional words were organized into 50 pairs. We performed psychological experiments for these 50 word pairs, and selected 12 of them from 80 effective feedbacks as listed in column 1 of Table II, which are the most significant emotions invoked by the images.

Image database selection: To achieve better experimental results and algorithm generalization ability, a large number of image samples are needed. However, people's emotions may change with time and people will feel tired after a long time test.. To get the exact result, the estimate process should be finished within an acceptable time duration, thus the number of sample images is limited. As a trade-off, 100 sample images are selected, and an observer needs one or two hours to finish the test. The images are art paintings because they contain more colors and clues that can invoke sophisticate emotions.

Image evaluation: The evaluation sheets are designed using the semantic differential technique [11][12] and posted in web site and can be finished online. Each observer was

This work was supported in part by the National Natural Science Foundation of China (NO.60372068).

The authors are with the School of Electronic and Information Engineering, South China University of Technology, China (e-mail: {wnwang, ecylyu ,shmjiang}@scut.edu.cn).

asked to examine every image in the database, evaluate the emotion evoked bye this image, and put a score to each impression word pair. The value of the score is -2,-1, 0, 1 or 2. The Final evaluation value of each word pair for an image is the average score of all observers.

Moreover, we used the following methods to reduce the interference and confusion in subjective evaluation. 1) A standard explanation was provided for each adjective based on the dictionary. 2) Since the previous image may affect people's feeling of the later one, the images were displayed in a random order to lower the impact. Each observer will see different image series. 3) The observer had a chance to break the evaluation every time finishing a section of 10 images to lower the impact of fatigue.

After the psychophysical experiment, we got the evaluation values of 12 emotional word pairs for the 100 images in database.

Statistic analysis: 42 observers finished all the evaluation sheets. Table I shows their background distribution.

Table I. Statistics of observers

	Number of observers
Gender	male (16) ; female (26)
Age	bellow 20 (13) ; 21~25 (15) 26~30 (12) ; above 30 (2)
Major	art related (12) ; non-art related (30)
Education	undergraduate (30) ; post graduate (12)

The emotions are affected by the gender, education and major background. To evaluate the difference, we use Eq. (1) to compute the correlation coefficient.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{Y})^2}} \quad (1)$$

In general, the correlation coefficient is high if $0.7 \leq |r| < 1$, middle if $0.4 \leq |r| < 0.7$, low if $0.2 \leq |r| < 0.4$ and very low if $|r| < 0.2$.

Table II. Statistics of semantic words correlations

Emotional Semantic Words Pair	Gender	Major	Retest
Exhilarated-depressive (E1)	0.86	0.79	0.87
Warm-cool (E2)	0.89	0.87	0.89
Happy-sad (E3)	0.85	0.82	0.87
Light-heavy (E4)	0.86	0.81	0.88
Hard-soft (E5)	0.78	0.80	0.78
Brilliant–gloomy (E6)	0.90	0.89	0.91
Lively-tedious (E7)	0.82	0.78	0.80
Magnificent-modest (E8)	0.86	0.87	0.82
Vibrant-desolate (E9)	0.87	0.87	0.91
showy-elegant (E10)	0.88	0.91	0.87
Clear-fuzzy (E11)	0.82	0.75	0.81
Fanciful-realistic (E12)	0.86	0.81	0.84
Average of correlations	0.82	0.79	0.83

Note: All the emotional words originally used in the psychophysical experiment are Chinese words.

Table II shows the statistics of emotional semantic word correlations, where Column 2 and Column 3 indicate a little difference in gender and major. The results suggest that we can analyze the common emotional factors although there are some individual differences. To evaluate the test reliability, we invited some of the observers to redo the evaluation after one month, and compared the results with their previous ones. Column 4 in Table II indicates that the stability is very high.

B. Factor Analysis

The factor analysis is used to study the underlying factors of the 12 emotional word pairs and to clarify the interrelations between them. The extraction method applied here was the principal component analysis jointly using an orthogonal rotation technique [12]. The factor analysis requires that the original variables have strong correlations. We performed the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test, and the results showed that it was suitable to perform the factor analysis.

Three factors labeled F1,F2 and F3 were extracted from the experimental data, accounting for 88.6% of the total variance. Based on the factor matrix (Rotated Component Matrix) show in Table III, the emotional work pairs of light-heavy, happy-sad, exhilarated-depressive, lively-tedious, vibrant-desolate, brilliant–gloomy, warm-cool have a strong correlation with the first factor(F1), and showy-elegant, magnificent-modest, hard-soft with the second one(F2) while clear-fuzzy and fanciful-realistic with the third one(F3).

Through the factor matrix, the values of emotional word pairs can be computed from the values of emotional factors (i.e. F1, F2 and F3).

Table III. Factor matrix of image emotional word pairs

Emotional Semantic words pair	Factors		
	F1	F2	F3
Exhilarated-depressive (E1)	**0.97**	-0.09	0.12
Warm-cool (E2)	**0.96**	0.17	0.13
Happy-sad (E3)	**0.94**	-0.11	0.24
Light-heavy(E4)	**0.90**	0.22	0.15
Hard-soft(E5)	**0.85**	0.36	0.07
Brilliant –gloomy(E6)	**0.84**	0.36	0.07
Lively-tedious(E7)	**0.67**	0.55	-0.01
Magnificent-modest (E8)	0.18	**0.94**	-0.08
Vibrant-desolate(E9)	-0.04	**0.93**	0.07
showy-elegant(E10)	0.35	**0.89**	-0.10
Clear-fuzzy(E11)	0.18	0.21	**0.91**
Fanciful-realistic (E12)	-0.18	0.36	**-0.81**

Furthermore, we can also get the Component Score Coefficient Matrix using the regression method to compute the emotional factors from the values of emotional word pairs.

Through factor analysis, we transformed the space of the 12-dimension emotional word pairs into a 3-dimension emotional factors space, which discovers the underlying factors of emotions and describe emotion semantics more simply.

III. FEATURE EXTRACTION

Feature extraction is a key issue in image retrieval. However, few researchers extract image features in an emotional perspective. Even in the image emotion analysis, researches also use the common features [4][5][6]. Therefore, the feature vectors are always very large due to the implicit relationship between the features and the emotional semantics. Early studies suggest that different semantic categories need different specific features [14] [15]. After a deep analysis of the relationship between emotional factors and the characters of image, three novel image features are designed for the three emotional factors respectively. By using Support Vector Machine of Regression (SVR), the values of emotional factors can be estimated automatically, and the results show great correlations with the psychophysical experiment data, which demonstrates the effectiveness of our approach.

A. Feature One for the First Emotional Factor

Being one of the main visual cues, color has been frequently used in image analysis and retrieval. L*C*H* space is selected to study the relationship between human emotions and image color characters because its definition and measurement are suited for vision perception psychology. L is luminance, which denotes the color lightness; C is chroma, which denotes the color saturation; H is hue, which denotes different colors. L*C*h* space is very similar to Munsell color space, and consistent with color psychology model. Each dimension can be perceived independently, and the color changing is smooth in terms of human perception. Thus, we converted the image from RGB to L*C*h* space,

The first emotional factor has a strong relationship with emotional word pairs including exhilarated-depressive, brilliant-gloomy, happy-sad, light-heavy, lively-tedious, vibrant-desolate, warm-cool. Based on the theory of color psychology [13][17], the high light colors invoke brilliant, light and lively emotions. We selected the top 30% images with a large value of the first factor and the bottom 30% images with a low value of the first factor, and then perform the statistical analysis. The images with a high value of the first factor have the high lightness, while the images with a low value of the first factor have the low lightness. T-test is performed and the result shows that the difference is statistically significant.

However, some images with the similar lightness have very different values on the first factor. Further investigation indicates that these images are different in warm and cool color distribution. Warm and cool colors have an apparently different effect on the feelings. The warm color images always invoke happy, lively and exhilarated feelings, while the cool color images always invoke sad, heavy and depressive feelings.

Therefore, the features of the first factor should have the ability to describe lightness and warm-cool of images.

1) Lightness description

Human recognize images in semantic words, so we describe the lightness by semantic words, such as very dark, dark, middle, light, very light. These semantic words are mapped by fuzzy membership function because of the fuzziness of human perception and natural language.

In fuzzy logic applications, choosing membership functions to reflect the data distribution is the first and an essential step; the choice may be achieved by using unsupervised learning algorithms [16]. The algorithm is presented as follows [7]:

Input data sequence $x_1, x_2, ..., x_n$, where x_i denotes feature L* (or C* in part B. 2) of the i-th region and n is the number of color regions.

1) Initialize 5 membership functions.

Let $c_0 = \min\{x_1, x_2, ..., x_n\}$, and $c_6 = \max\{x_1, x_2, ..., x_n\}$, and compute $c_1, c_2, ..., c_5$ as follows: $c_j = c_0 + j \cdot (c_6 + c_0)/6$.

The initialized membership functions are shown as Fig. 1, in which $c_1, c_2, ..., c_5$ denotes class centroids of the initial fuzzy partition.

2) Set U=0. For each x_i, compute u_{ij} using the following rules, where u_{ij} ($1 \le i \le n$ and $1 \le j \le 5$) is the membership value that the i-th pattern belongs to the j-th semantic word.

Rule 1 : if $x_i \le c_1, u_{i,1} = 1$ and $u_{i,k\neq 1} = 0$;

Rule 2 : if $x_i > c_5, u_{i,k\neq 5} = 0$ and $u_{i,5} = 1$;

Rule 3 : if $c_j < x_i \le c_{j+1}$, $u_{i,j} = c_{j+1} - x_i / c_{j+1} - c_j$,

$u_{i,j+1} = 1 - u_{i,j}$ and $u_{i,k\neq j,j+1} = 1 - u_{i,j}$;

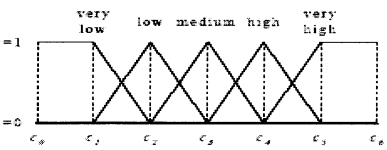

Fig. 1. Initialized membership functions

3) Update the class centroids $c_1, c_2, ..., c_5$ using

$$c_j = \sum_{i=1}^{n} u_{ij} x_i \Big/ \sum_{i=1}^{n} u_{ij} \tag{2}$$

4) Repeat step 2 and step 3 until $c_1, c_2, ..., c_5$ are unchanged.

In the above algorithm, each semantic word is a fuzzy set and represented as a triangular membership function. In step 1, five distributed triangular membership functions are chosen as an initial fuzzy partition, as shown in Fig. 3. In step 2, each element in U, i.e. u_{ij}, is computed according to the three rules. In step 3, the class centroids, i.e., $c_1, c_2, ..., c_5$, are updated to reflect new data distribution. The algorithm will terminate if the class centroids are unchanged, otherwise, step 2 and step 3 are repeated. A representative word set is thus generated, and the conversion from low-level image feature into features with semantic meaning is accomplished.

Fig. 2 shows the membership functions of the lightness. For example, if a pixel's lightness is 68, its membership for "very dark", "dark", "middle", "light", "very light", is [0, 0, 0.25, 0.75, 0].

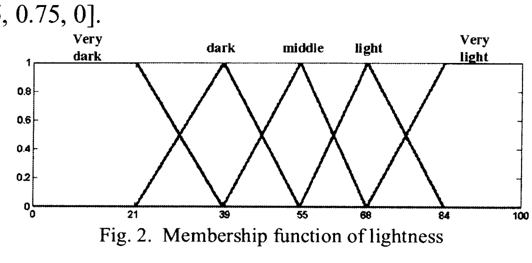

Fig. 2. Membership function of lightness

2) warm-cool description

Colors invoke the feelings of warm or cool. For example, the red, orange or yellow invoke the warm feelings, while the blue and green invoke the cool feelings. Refer to the early study performed by Ou and Sato [13], we defined the membership function of warm-cool as bellow:

$$\mu_{warm}(x)=\begin{cases}\cos(h-50^\circ) & 0^\circ \le h < 140^\circ \ or \ \ 320^\circ \le h \le 360^\circ \\ 0 & others\end{cases}$$

$$\mu_{cold}(x)=\begin{cases}\cos(h-230^\circ) & 140^\circ \le h < 320^\circ \\ 0 & others\end{cases} \tag{3}$$

Where $\mu_{warm}(x)$ is the membership function of warm color while $\mu_{cold}(x)$ is the one of cool color, and h is the hue of color x.

The membership functions are shown in Fig. 3. The solid curve represents the membership function of warm, while the dashed curve represents the membership function of cool.

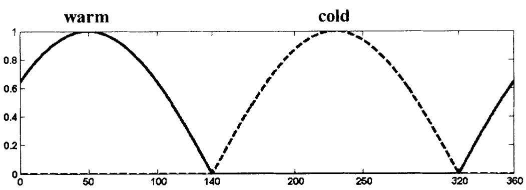

Fig. 3. Membership function of warm and cool colors

Combining the membership functions of lightness, warm and cool, we had 10 combinations including "very dark -warm", "very dark-cool", "dark-warm", "dark-cool", etc. For each pixel, we could compute a 10-dimension vector to describe its lightness and warm-cool. Computed all the pixels in one image, we could get a feature vector of a 10-dimension histogram, which had an ability to describe the distribution of lightness and warm-cool of the image. This feature has a good relation with first emotional factor.

B. Feature Two for the Second Emotional Factor

The second emotional factor has a strong relationship with emotional word pairs such as magnificent-modest, showy-elegant and hard-soft. Based on the theory of color psychology, the high saturation colors invoke showy, magnificent emotions, while the low saturation colors invoke modest and elegant emotions. At the same time, hard-soft feeling has a close relationship with color contrast [17].

We selected the top 30% images with high value of the second factor and the bottom 30% images with low value of second factor to perform the statistic analysis. The images with high values of the second factor are warm, high saturation as well as high contrast, while the images with low values are cool and low saturation as well as low contrast. The T-test is performed and the results verify the correlations.

Therefore, feature that can express image saturation, contrast and warm-cool are needed for the second factor. We designed an integrated feature of saturation-warm-cool fuzzy histogram and color contrast as followings:

1) Saturation description

We define three words for saturation including "low saturation"(LS), "middle saturation"(MS) and "high saturation"(HS). We have tried more words for saturation, e.g., 5 words, but the results are not improved because human beings are not so sensitive with the saturation.

Eq. (4) and Fig. 4 is the definition of membership functions of saturation.

$$\mu_{LS}(C)=\begin{cases}1 & C<10 \\ (27-C)/17 & 20\le C\le 27 \\ 0 & C>27\end{cases}$$

$$\mu_{MS}(C)=\begin{cases}(C-10)/17 & 10\le C<27 \\ (51-C)/24 & 27\le C<51 \\ 0 & others\end{cases} \tag{4}$$

$$\mu_{HS}(C)=\begin{cases}0 & C<27 \\ (C-27)/24 & 27\le C\le 51 \\ 1 & C>51\end{cases}$$

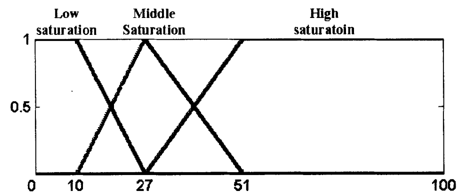

Fig. 4. Membership function of saturation

2) warm-cool description

Refer to the description for the first feature.

3) Contrast description

We define the contrast of an image as below:

$$Lab\_con=\left[\frac{1}{N-1}\sum_{i=1}^{N}(a_i^*-\overline{a^*})^2+(b_i^*-\overline{b^*})^2\right]^{1/2} \tag{5}$$

a_i^*, b_i^* represent the values of a*, b* in L*a*b* color space, $\overline{a^*}$, $\overline{b^*}$ are the average values of a* and b* of the whole image, N is the number of pixels in the image. Lab_con describes contrast of an image.

Combining with the saturation, warm-cool and the contrast, we got a 7-dimension vector, which had the ability to describe saturation, warm-cool and contrast, and deduced the second emotional factor.

C. Feature Three for the Third Emotional Factor

The third emotional factor has a strong relationship with emotional word pairs of clear-fuzzy and fanciful-realistic. Based on the theory of visual perception, high lightness contrast and sharp edges invoke clear and realistic emotions, while the low lightness contrast and fuzzy edges invoke fuzzy and fanciful motions [17].

The clear images have clear edges with high grads values. Human always focus on the sharp part of the image that has a significant impact on the emotions. However, the edges are only a small part of images. Therefore, we choose the top average of 0.5% grads values as the measurement of image sharpness. After performing the statistical analysis, the T-test results verified the correlations.

Combining with the light contrast and the sharpness, we got a two-dimension vector, which could deduce the third factor.

D. Image Emotion Evaluation Using Machine Learning

Through methods described in section 3, we got the image features that are related to respective emotional factors. The learning machine can build up the mapping between the image features and emotional factors. SVR [18][19] is chosen because it can achieve excellent performance and generalization ability on limited samples. For each factor, we construct a SVR machine respectively to train. Then the three factors' values of each image can be evaluated by the trained SVRs. The value of 12 emotional word pairs will be computed through the factor matrix discussed in Section II.

To sum up, there are two steps in machine evaluation, 1) three emotional factors evaluation from image features through trained SVRs; 2) twelve emotional word pairs evaluation from three factors by using factor matrix. In this way, automatic emotional evaluation of one image could be achieved. And the value of evaluation can be used in image retrieval based on emotional semantic words.

IV. Experiment Results and Analysis

A. Estimation of Image emotional factors

We separate the images into a training set and a test set. 75% of images were in training set while the 25% were in the test set. Radial Basis Kernel is chosen for the SVR, whose performances are shown in Table IV. The correlations between the three emotional factors of SVR outputs and those analyzed in Section II by human evaluation results are all above 85%. The results show the effectiveness of our image features and the image emotional semantic extracting model.

Table IV. Results of SVR training

EF	SVR	DV	SVR Parameters			correlations		
			ε	σ	C	RS	TS	WS
F1	SVR1	10	0.10	2.80	10	0.91	0.79	**0.87**
F2	SVR2	7	0.10	3.00	50	0.93	0.82	**0.88**
F3	SVR3	2	0.10	0.30	10	0.89	0.81	**0.86**

Notes: EF is the emotional factor. DV is the dimension of the feature vector. RS is the training set. TS is the test set, WS is the whole set

B. Image Retrieval based on Emotional Semantic Words

We can deduce the 12 emotional word pairs from the emotional factors by the factor matrix. Comparing the evaluated values with the original values from the psychophysical experiment, we got the accuracy as shown in Table V.

As indicated by the second column in Table V, differences between the machine-evaluated emotional words and the observer-evaluated ones are small for all the images. The average of mean square error (AMSE) of all the 12 pairs is 0.34.

In addition, the system computed mean value deviated from the observer emotional evaluation value is also small shown by columns 3 and 4 of the table. Given the emotional word pair j of image i, E_{ij} denotes the average value of the evaluation values of all observers and σ_{ij} denotes the variance. For each emotional word pair j and image i, comparing the computed value with the evaluated E_{ij}, we can get the ratio of the computed value in the accepted range. The average for the 12 subjective is 68% for the range $[E_{ij} - 0.3\sigma_{ij}, E_{ij} + 0.3\sigma_{ij}]$ while 86% for the range $[E_{ij} - 0.5\sigma_{ij}, E_{ij} + 0.5\sigma_{ij}]$. The results indicate that our model can match the emotions of human perception well.

Table V. Statistics of accuracy for query results of 12 semantic words

Emotion	Our method			Hayashi's method		
	AMSE	0.3σ (%)	0.5σ (%)	AMSE	0.3σ (%)	0.5 σ(%)
E1	0.34	71.8	85.9	0.40	45.3	73.4
E2	0.38	56.5	83.5	0.42	48.4	78.1
E3	0.30	71.8	83.5	0.41	51.6	75.0
E4	0.34	72.9	85.9	0.43	48.4	71.9
E5	0.33	63.5	89.4	0.36	64.1	90.6
E6	0.39	61.2	82.4	0.46	42.2	62.5
E7	0.30	72.9	84.7	0.36	53.1	79.7
E8	0.31	76.5	90.6	0.36	51.6	79.7
E9	0.37	55.3	84.7	0.45	43.8	64.1
E10	0.34	70.6	82.4	0.40	43.8	73.4
E11	0.39	64.7	87.1	0.47	43.8	68.8
E12	0.31	76.5	91.8	0.36	67.2	87.5
Ave	0.34	67.8	86.0	0.41	48.7	74.3

Notes: AMSE is the average of mean square error

Hayashi et al.[4] developed methods for emotional computing. They used a 65-dimension feature vector, and neutral network to evaluate the emotional subjective directly. Using Hayashi's methods on the database, we got the results shown in Table V. The average of MSE is 0.41 for the range $[E_{ij} - 0.3\sigma_{ij}, E_{ij} + 0.3\sigma_{ij}]$ and the average is 49%. While for the range $[E_{ij} - 0.5\sigma_{ij}, E_{ij} + 0.5\sigma_{ij}]$, the average is 74%. Hayashi's features were extracted for the general color distribution and frequency distribution, and were not emotional specific. Moreover, they did not perform emotional space analysis, e.g. factor analysis as the original emotional space is used without dimension reduction. Because our features are more specific to the emotions, we achieved better results by less feature vectors.

Based on the evaluation results, the emotional word query system is implemented. The users can input one of the 24 emotional words, and the system will output the images with decent order of the emotional values. The preceding images have stronger emotions. Figs 5-6 are some examples of the query results. Fig. 5 shows the former 18 query results of emotional word "happy", and Fig. 6 shows the former 18 query results of emotional word "sad".

V. Conclusion and Future work

In this paper, we analyze the emotional space and build a novel scheme that can annotate the image emotion semantic automatically for emotional image retrieval. We build an image retrieval system based on our proposal, and the experimental results demonstrate the effectiveness of the proposed model.

In the future, we will focus on the following methods to improve our work: 1) considering the influence of image local features and ROI (Region of Interest) on emotions; 2) introducing mechanisms to adjust the system by collecting and analyzing feedbacks; 3) Developing mechanisms to recognize and adapt individual differences to customize the algorithms for different users.

Fig. 5. The former 18 query results of happy images

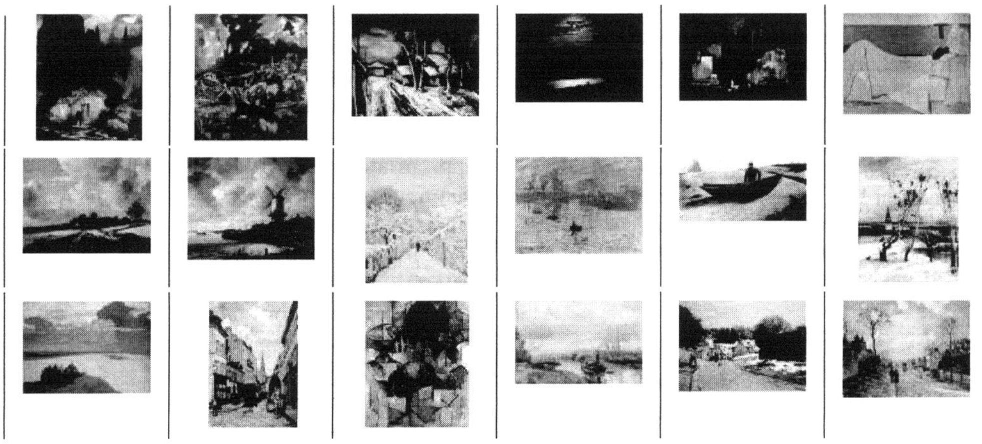

Fig. 6. The former 18 query results of sad images

REFERENCES

[1] John B., Kanav K., et al. "Indexing natural images for retrieval based on kansei factors", In Proc. Human vision and electronic image conference, California, 2004. SPIE vol. 5292, pp.363-375.

[2] A. W. M Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval: the end of the early years", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, December 2000.

[3] C. Colombo, A. Del Bimbo, and P. Pala, "Semantics in Visual Information Retrieval", IEEE Multimedia, vol. 6, No.3, pp. 38-53, 1999

[4] Hayashi. T., Hagiwara, M., "Image query by Impression words-The IQI System", IEEE Trans. Consumer Electronics, vol.44, No.2, pp.347-352, May 1998.

[5] Yoshida, K., Kato, T., Yanaru, T., "Image Retrieval System Using Impression Words", in Proc. IEEE Systems, Man and Cybernetics, Tokyo, 1998, vol.3, No.11-14, pp. 2780-2784.

[6] Nadia Bianchi-Berthouze. "K-DIME: an affective image filtering system", IEEE Trans. Multimedia, vol.10, no.3, pp.103-106, July-Sept, 2003.

[7] Wei-Ning Wang, Ying-Lin Yu, "Image Emotional Semantic Query Based on Color Semantic Description", in Proc. Intl. Conf. Machine Learning and Cybernetics, Guangzhou, 2005, vol.7, pp. 4571-4576.

[8] Wang Wei-ning, Yu Ying-lin, Zhang Jian-chao, "Image emotional classification: static vs. dynamic", in Proc. IEEE Systems, Man and Cybernetics, Amsterdam, 2004, vol. 7, pp. 6407-6411.

[9] Alan Hanjalic, Li-Qun Xu, "Affective Video Content Representation and Modeling", IEEE Trans. Multimedia, vol. 7, no. 1, pp. 143-154, February 2005.

[10] Sunkyoung Baek, Miyoung Cho, PanKoo Kim, " Matching Colors with KANSEI Vocabulary Using Similarity Measure Based on WordNet", in Proc. Of Intl. Conf. Computational Science and Its Applications, Singapore, 2005, pp.37-45.

[11] Osgood CE, Suci GJ, Tannenbaum PH. The measurement of meaning. University of Illinois Press, 1957.

[12] Spearman C. "General intelligence: Objectively determined and measured," American Journal of Psychology, Vol.15, pp. 201-293, 1904.

[13] Li-Chen ou, M. Ronnier Luo, Andree Woodcock, Angela Wright. "A study of Colour Emotion and Colour Preference. Part I: Colour Emotions for Single Colours", Color research and Application, vol. 29, no. 3, pp. 232-240, June, 2004.

[14] Aleksandra Mojsilovic, Jose Gomes and Bernice Rogowitz, "Semantic-Friendly Indexing and Querying of Images Based on the Extraction of the Objective Semantic Cues", Intl. J. of Computer Vision, vol. 56, no. 1/3, pp. 79-107, January/February, 2004.

[15] Jia Li, James Z. Wang, "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, no.9, pp.1075-1088, 2003.

[16] Hsin-Chih Lin, Chih-Yi Chiu, and Shi-Nine Yang, "Texture Analysis and Description in Linguistic Terms", in Proc. 2nd IEEE Pacific-Rim Conference on Multimedia, Beijing, 2002, 205-209.

[17] J. Itern, Art of Color, Otto Maier Verlag, Germany,1961.

[18] V.Vapnik. The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.

[19] Alex J. Smola and Bernhard Schölkopf, "A tutorial on support vector regression", Statistics and Computing, Vol. 14, no. 3, pp.199-222, August 2004.