Noname manuscript No.
(will be inserted by the editor)

A method for extracting emotion using colors comprise the painting image

Dongwann Kang · Hyounoh Shim · Kyunghyun Yoon

Received: date / Accepted: date

Abstract Paintings can evoke emotions in viewers. In this paper, we propose a method for extracting emotions from paintings by using the colors that comprise the paintings. The proposed approach is based on a color image scale, which is one of the popular experimental scales focusing on the relation between colors and emotions. We first construct a color combination and emotional word dataset. To this end, we create a color spectrum from the input painting. We then search for the best matching color combination from the dataset, which is most similar to the color spectrum. The best matching color combination is mapped to the corresponding emotional word. Afterward, we extract the emotional word as the emotion evoked by the painting. To evaluate the proposed method, we compared the results of the proposed algorithm to those of a user study on the extraction of emotions from several paintings. Through several experiments, we show that the proposed method exhibits excellent performance with respect to predicting the emotions evoked by a painting. Finally, we propose an image exploration system based on the emotion extraction method mentioned above. In this system, users can explore painting images emotionally coherently.

Keywords color image scale · emotion of painting · color combinations · image exploration

D. Kang
Bournemouth University, Bournemouth, Dorset, UK

H. Shim
Affective Computing in Real Life (Acryl), Seoul, Korea

K. Yoon
Chung-Ang University, Seoul, Korea
E-mail: khyoon@cau.ac.kr

1 Introduction

When people view paintings, they may feel certain emotions. For example, they may feel comforted by a painting and feel sad by looking at another painting. Consequently, paintings evoke certain emotions. The evoked emotions may either depend on the person's background or be common universal emotions.

The components of a painting that evoke emotions are the subject and the expressional style, such as colors, composition, and texture. These affect the evoked emotions compositively. Therefore, it is difficult to predict an emotion evoked by a painting by using computational methods. Consequently, a method to identify the relationship between each component and the corresponding evoked emotion is preferentially required. Among the abovementioned components, not only is color the most intuitive and fundamental but its relationship with the evoked emotion also has attracted attention in the field of psychology. In this study, with respect to the relationship between colors and emotions, we aim to extract emotions from paintings.

Colors evoke certain emotions solely or through their relationship with other colors [28, 23]. Although this depends on a person's experience and cultural background, Kobayashi [11] surveyed emotional adjectives evoked by colors and showed that there are universal relationships between colors and emotions. In his study, the color image scale, which is a three-dimensional space for quantitatively measuring emotions, was proposed to find out the relationship between colors and emotions. Today, his studies [11, 12] are widely used in the field of design to determine colors according to emotions.

The main idea of this study is, based on the relationship between colors and emotions, to extract an emotion evoked by a painting by using the colors used in the painting. For this, on the basis of Kobayashi's studies [11, 12, 25], which suggested a number of color combinations and emotion pairs, we propose a method that generates a color spectrum from the input painting, finds the best matching pair in which the color combination is similar to the color spectrum, and determines the emotion of the best matching pair as the emotion of the input painting. Finally, we present a novel system that explores images by using the emotion extracted from the images. In our system, the users can explore design images with coherent emotion, so that they can find appropriate image more easily.

The remainder of this paper is organized as follows: In Section 2, we provide an overview of the related work on emotions and paintings. We then present a method that constructs a color combination and emotion pair dataset in Section 3 and propose an emotion extraction method based on the similarity between a color combination and the color spectrum generated from a painting in Section 4. We then present the results of the proposed method and evaluate the performance of this method in Section 5. In section 6, we then present our system that explores images based on its emotion. Finally, we conclude this paper with a summary of our ideas and discuss the limitations of the proposed method and briefly discuss the future work directions in Section 7.

2 Related Work

2.1 Paintings and Emotions

For a couple of decades, the inference of emotions evoked by an image has been one of the topics of interest in the fields of computer vision and image processing [8]. Zhao et al. [30] classified images according to the evoked emotions by using a machine learning-based classifier. They extracted the principle of art-based features from International Affective Picture System (IAPS), which is a standard emotion evoking photographic image set [13], by using image processing techniques, and generated a support vector regression-based classifier by training the features. They showed that the emotions evoked by a photographic image could be predicted by the features inherent in the images. Machajdik and Hanbury [16] achieved a better prediction performance by using the features extracted from IAPS images on the basis of photographic principles, such as the rule of third and the depth of field. Marchesotti et al. [17] developed various generic image features, applied learning on the collection of The Chinese University of Hong Kong [9], which is an image set that consists of images with positive/negative ratings, and assessed the aesthetic quality of images. The above mentioned studies showed that universal emotions could be extracted from images by using image processing and vision techniques. However, they focused on photographic images instead of paintings.

Previous studies on paintings have included classifications based on art movement or genre by using machine learning. Icoglu et al. [6], for example, proposed an algorithm for the classification of paintings on the basis of art movements by using the k-nearest neighbor, Bayesian, and support vector machine (SVM) classifiers. They employed several vision-based features, such as the correlation coefficient calculated using the gradient map, the maximum value of the intensity histogram, and the range of the intensity histogram. Zujovic et al. [31] proposed a method for classifying paintings based on genres. They extracted several features from images of paintings by using the HSV color system, Canny edges, and Gabor filters. These studies achieved reasonable classification accuracy but did not consider the evoked emotions.

Yanulevskaya et al. [29] proposed a method that recognizes the emotions of a painting. They constructed a descriptor by using the *Lab* color system and Scale-invariant feature transform, and generated a classifier through the SVM. In their study, 100 participants rated 500 abstract paintings on a seven-level positive/negative scale. These rated paintings were used for training the classifier. This study showed the long-known observation that the colors of a painting affected the emotions evoked by it.

Several works [5, 20, 10, 21] studied the emotion of multimedia such as music, image, and movie. Contrary to these studies, in this paper we focus on colors of painting images only.

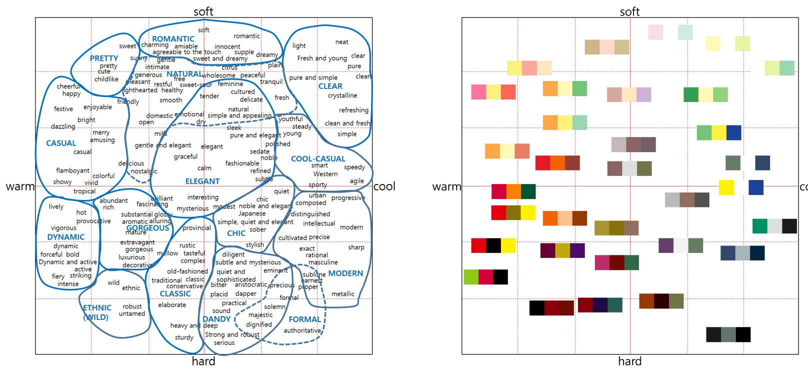
As discussed above, most studies on paintings have focused on the classification of the painting genre and the extraction of simple emotions, such as positive/negative emotions. Unlike these studies, we propose a method for extracting more complicated emotions from paintings.

2.2 Colors and Emotions

In the field of psychology, many studies on emotions evoked by a single color or a color combination were conducted for a long time. On the basis of these studies, it is known that a single color has its own meaning [28], evokes a measurable emotion [27, 18], and has a strong universal trend [1]. Furthermore, several studies [19, 11] have shown that color combinations evoke their own emotions. Kobayashi [11] developed the color image scale to match a color combination and the corresponding emotional word. He generated a three- dimensional space that consists of three psychological factors, such as soft/hard, warm/cool, and grayish/clear. He then surveyed the scale of the three factors of color combinations and emotional words from design students and general consumers. Finally, by using the three factors surveyed, he located the color combinations and emotional words on the color image scale. The color combinations have equivalent meaning with emotional words, which are located at the nearest position of the color combination.

On the basis of the above principle, Kobayashi [12] showed 900 pairs of three-color combinations and 180 emotional words on a two-dimensional color image scale, which does not contain a grayish/clear factor. In the design field, this practical study has been effectively utilized to choose colors on the basis of emotions. However, Kobayashi matched color combinations and emotional words on a two-dimensional color image scale without considering the grayish/clear factor; therefore, the range of possible emotions represented in the color image scale was limited. Consequently, the example pairs contained only positive emotional words. To overcome this problem, The Nippon Color & Design Research Institute(NCD)'s work [25], which is an extended study of [12], utilized a three-dimensional color image scale and showed that additional pairs of color combinations and the corresponding emotional words contained both positive and negative emotions.

The preliminary version of this work was presented in [22, 15], where we extracted emotions from paintings by using a two-dimensional color image scale and a dataset in Kobayashi's work [12] consisting of pairs of color combinations and emotional words. To expand the range of emotion, in this extension, we employ a three-dimensional color image scale, and extend the dimensions of the dataset to locate pairs in the dataset on the three- dimensional color image scale. The main difference is that more wide-range emotions including negative emotions can be extracted from paintings by using the proposed approach than by using the preliminary version. In addition, we present a novel image exploration system as an application of proposed emotion extraction method.



(a) Word image scale (b) Combination image scale

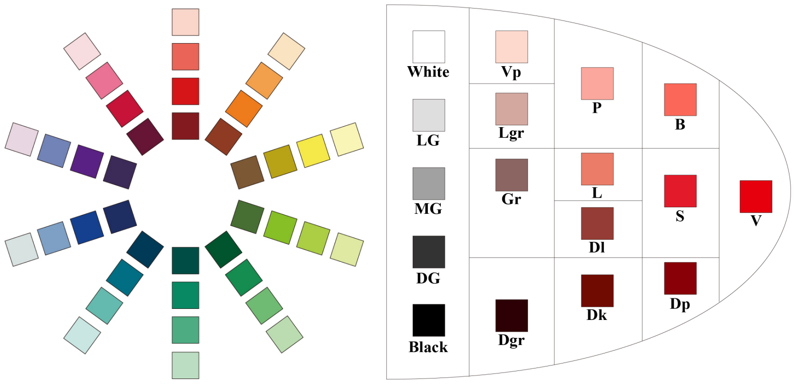
Fig. 1 Two-dimensional color image scale from [12]. On the basis of the coordinates that consist of two factors, namely warm/cool and soft/hard, three-color combinations and emotional words are paired.

3 Constructing Color Combination-Emotional Word Dataset

Kobayashi [12] created a dataset that contains 180 emotional words and 900 three-color combinations. They are located on the two-dimensional color image scale with the grayish/clear factor missing (Figure 1). On the basis of the two-dimensional coordinates, each color combination is paired with the closest emotional word. Consequently, 900 pairs of color combinations and emotional words are in the dataset. Because the grayish/clear factor is ignored, it is not guaranteed that the grayish/clear factors of the color combination and the emotional word in a pair are closest to each other, although they are paired. Moreover, all the emotional words employed in the dataset consist of positive emotions only, because positive and negative emotions are only distinguishable on a three-dimensional color image scale with the grayish/clear factor. Therefore, the dataset is not appropriate for extracting wide-range emotions.

To overcome this problem, we extend the dimensions of the dataset by estimating the missing grayish/clear factor. NCD's work [25], which extends Kobayashi's study [12], offers 50 pairs of three-color combinations and emotional words on a three-dimensional color image scale. The pairs include both positive and negative emotional words. From this dataset, we identify the three-dimensional coordinates of 50 emotional words. Consequently, 900 color combinations in the dataset presented in Kobayashi's study [12] can be paired with the 50 emotional words, if the three factors of the color combination are assigned.

Three factors of the color image scale are psychological values obtained by a user study. Therefore, it is not possible to calculate each factor by using a color combination. However, unlike the other factors, the grayish/clear factor is based on the color property. Consequently, we can estimate this factor from



(a) 10 hues (b) 12 tones of red hue

Fig. 2 Twelve tones of the red hue in the Hue & Tone 130 color system.

the colors. Kobayashi [12] used the Hue & Tone 130 color system (Figure 2), which consists of 10 hues and 12 tones based on the Munsell color system. He divided the 12 tones into 5 grayish tones, namely light grayish, grayish, light, dull, and strong; and 7 clear tones, namely very pale, pale, bright, dark grayish, dark, deep, and vivid. He defined that a grayish tone is a color mixed with gray, and a clear tone is a color mixed with black or white. As shown in Figure 2, grayish tones are located on the inside of the Hue & Tone 130 volume, and clear tones are located on the boundaries of the volume. On the basis of this observation and the *Lab* color system, which is similar to the Hue & Tone 130 color system, we propose Equation 1 that estimates the grayish/clear factor from a single color.

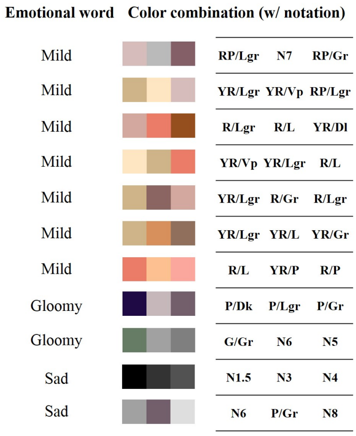
$$G(a) = max(1 - \sqrt{2max(|\mu - L_c| - \alpha, 0)^2 + \beta((a_c)^2 + (b_c)^2)}, 0) \quad (1)$$

Here, L_c, a_c, and b_c denote the components of *Lab* color c. We utilize the distance on the ab plane as the saturation value. Following Equation 1, if the *ab* distance of a color decreases, that is, if a color is desaturated, then the color is regarded as grayish. Further, if the luminance value of the color is close to μ, which is the grayish luminance value, the color is also regarded as grayish. In our experiments, we found that 0.4, 0.19, and 1.15 are suitable values of μ, α, and β, respectively.

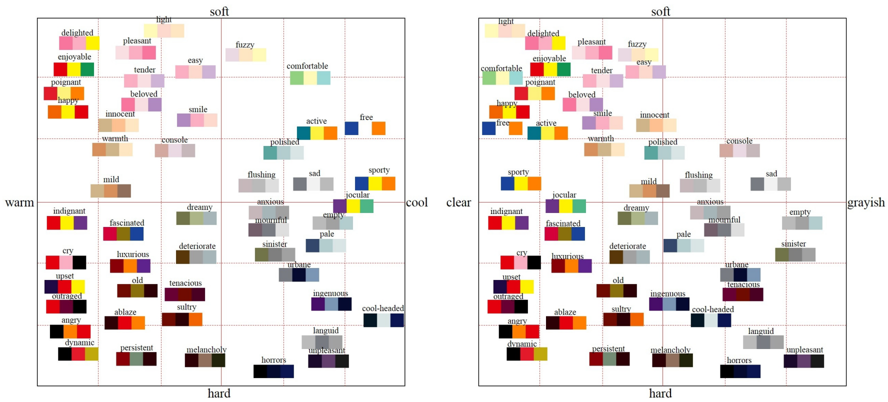
To obtain the emotions of color combinations in Kobayashi's work [12], the grayish/clear factor of the color combinations is required instead of that of single colors. We propose Equation 2 that estimates the grayish factor of a color combination, which consists of three colors, namely a, b, and c.

$$G(c_1, c_2, c_3) = (avg(G(c_1), G(c_2), G(c_3)) - 0.5) \times 6 \quad (2)$$

Here, $avg()$ denotes an average function. The range of each dimension of the color image scale is $[-3 : 3]$; therefore, we also scale the value of the grayish/clear factor.



(a) Pairs in the merged dataset



(b) Merged dataset on the color image scale

Fig. 3 Twelve tones of the red hue in the Hue & Tone 130 color system.

For the color combinations in Kobayashi's work [12], we obtain the warm/cool and soft/hard factors from the corresponding emotional words in the two-dimensional color image scale, and calculate the grayish/clear factor by using Equation 2. By using these three factors, we locate each color combination on the three-dimensional color image scale. Then, we pair each of the 900 color combinations with its nearest combination among the 50 emotional words. Finally, by merging the data of the two datasets, that is, the 900 pairs in Kobayashi's work [12] and the 50 pairs in NCD's work [25], we obtain a 950-pair dataset. Figure 3 shows our dataset on the three-dimensional color image scale. As shown in the figure, both positive and negative emotional words are included in our dataset.

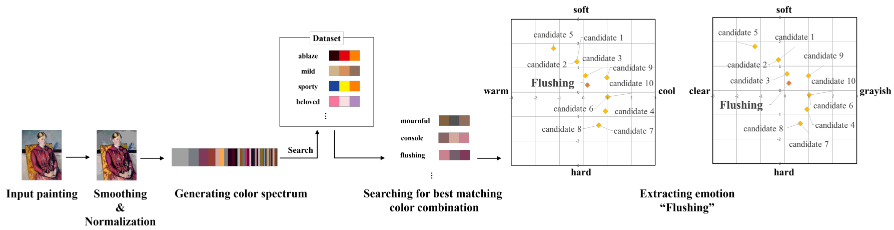
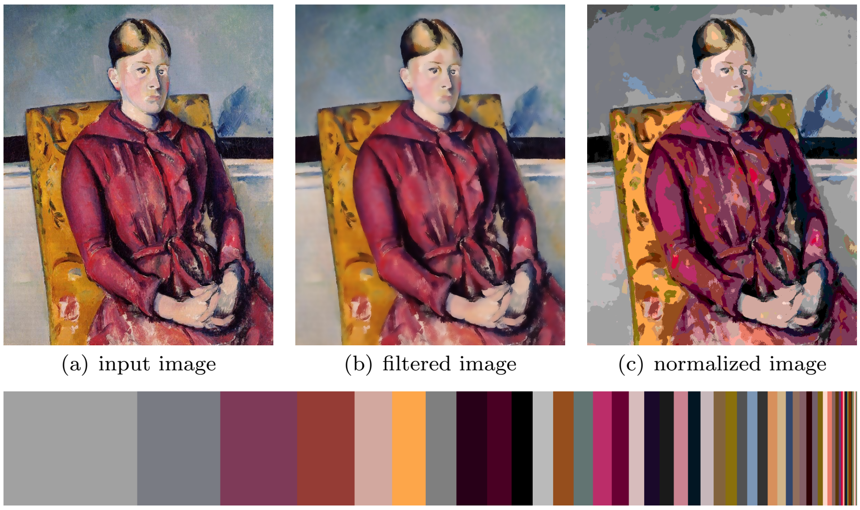


Fig. 4 Process overview of the proposed system.



(a) input image (b) filtered image (c) normalized image

(d) color spectrum generated from the input image

Fig. 5 Generating color spectrum from painting image.

4 Extracting Emotion from Painting

By utilizing the merged dataset discussed in Section 3, we extract emotions from the colors used in a painting. Figure 4 presents a process overview of the proposed system. First, we generate a color spectrum from an image of a painting to identify the colors used in the painting. Then, from the dataset, we search for the best matching color combination, which is the most similar to the color spectrum. Finally, we extract the emotional word corresponding to the color combination as the emotion of the painting.

4.1 Paintings and Emotions

To search for the best matching color combination, we need to identify the colors used in a painting. For this, we generate a color spectrum that consists of the colors used in the painting, as shown in the Figure 5. In the spectrum, the size of the area of each color is in proportional ratio to the amount of the

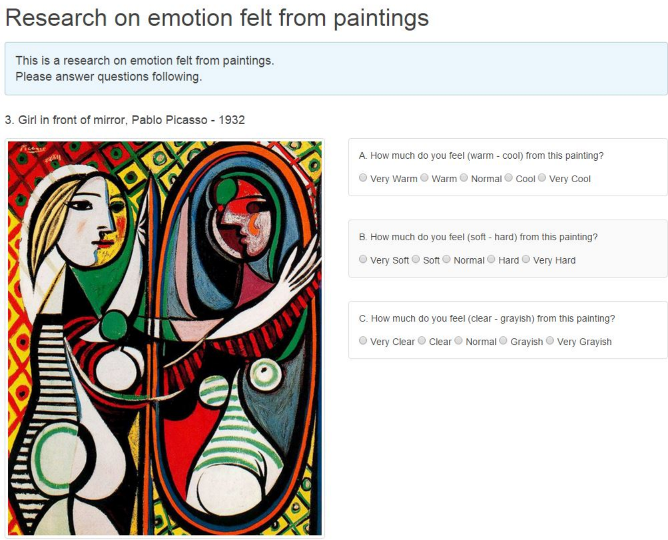
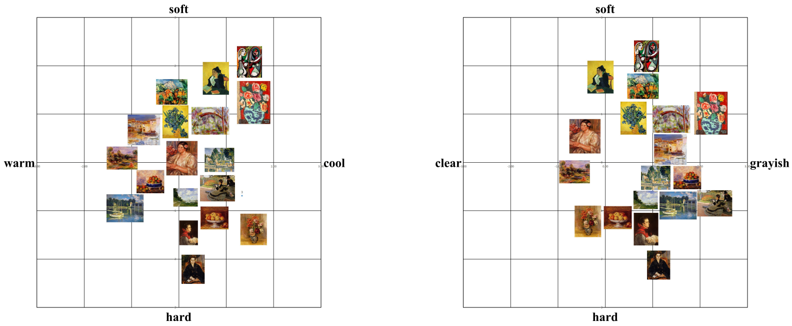


Fig. 6 Example question for gathering the three factors of the color image scale of a given image.

color in the painting. To remove noises that affect not the overall impression of the painting but the spectrum, we first smooth the colors of the painting (Figure 5(b)). For color smoothing, we utilize a bilateral filter [26]. As mentioned in Section 3, the color combinations in Kobayashi's works [11, 12, 25] are composed of colors from the Hue & Tone 130 color system. To represent the color spectrum by using the Hue & Tone 130 color system, we normalize the colors of the painting. For each color used in the painting, we find the most similar color in the Hue & Tone 130 color system and replace the original color with it. Here, we use the L_2 distance in the Lab color system to calculate the similarity. After normalizing all pixels of the input painting, we find that the normalized image is composed of a maximum of 130 colors (Figure 5(c)). For each normalized color, we then calculate the ratio used in the painting. Figure 5(d) shows an example of our color spectrum generated from the painting. The area of each color in the spectrum is proportional to the ratio.

4.2 Searching for Best Matching Color Combination

To search for the best matching color combination from the dataset and use its emotion as that of the painting, we compare each color combination and the color spectrum. For this, we need to calculate the similarity between each color combination and the color spectrum. A naïve approach to calculate the similarity is to compare the top three colors of the color spectrum and the color combination and then, calculate the color distances between them. However, in many cases, the top three colors in the color spectrum are not appropriate



(a) warm-cool/soft-hard scale (b) clear-grayish/soft-hard scale

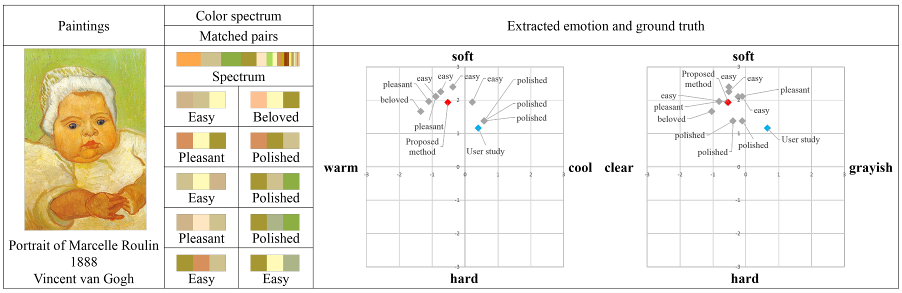
Fig. 7 Three-dimensional coordinates of the ground truth of paintings generated by the user study.

for representing the colors of the painting. For example, the main colors of Figure 5(a) are blue for the hat, orange and brown for the face. However, top three colors of Figure 5(d) are blackish, so the top three colors do not reflect the main colors. To address this problem, we use all the colors in the color spectrum to calculate the similarity. Equation 3 shows the similarity function that we use.

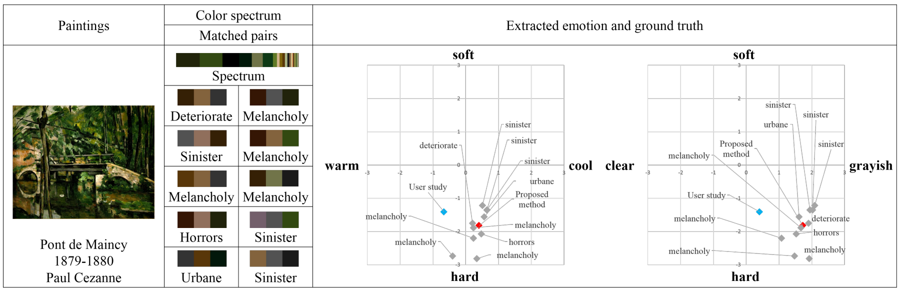
$$sim(c, s) = \sum_{i}^{3} \sum_{j}^{k} exp(D(c_i, s_j)^2/(-2\sigma^2)) \times r_j \tag{3}$$

where c denotes a color combination that consists of three colors $c_i = 0$ ... $c_i = 2$, s represents a color spectrum from the input painting that consists of k colors $s_j = 0$ ... $s_j = k$ with the ratio of colors r_j, and $D()$ denotes the color distance function. For calculating the color distance, the L_2 distance in the *Lab* color space is used. When colors that are similar to a color combination are used at a high ratio in the painting, this equation has a relatively high value. By using a Gaussian weight function in the color distance term, we avoid a color that is not similar to the color combination but has a high ratio from affecting the similarity. In this study, σ is set to 0.2.

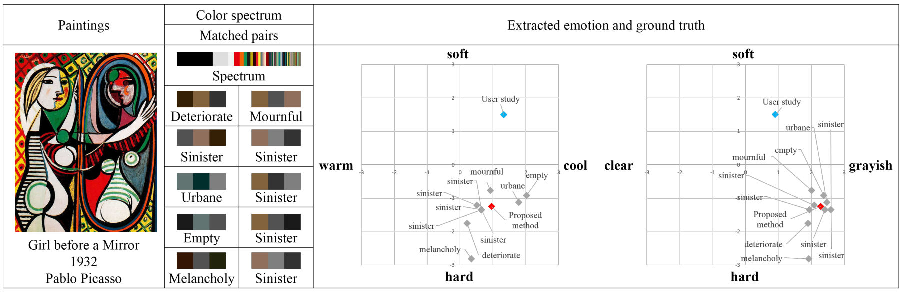
For the color spectrum of the painting, we calculate the similarities to all the color combinations in the dataset and extract the top N emotional words of the best matching color combinations, which are most similar to the color spectrum. This is based on the assumption that a viewer can experience complex emotions, not just a single emotion. Then, we calculate the average of the coordinates of N emotional words and find the nearest emotional word to the average coordinates as the emotion of the painting.



(a)



(b)



(c)

Fig. 8 Results of extraction of top three color combination and emotion pairs from the images of paintings.

5 Experimental Results

5.1 Evaluation of Grayish/Clear Factor Estimation

To evaluate the performance of the grayish/clear factor estimation proposed in Section 3, we compute the error of the factor between the ground truth and the estimated value. As mentioned in Section 3, our dataset contains 50 pairs of color combinations and emotional words, which have the ground truth coordinates on the three-dimensional color image scale; we utilize these pairs for validating the proposed method. For each of the 50 color combinations, we

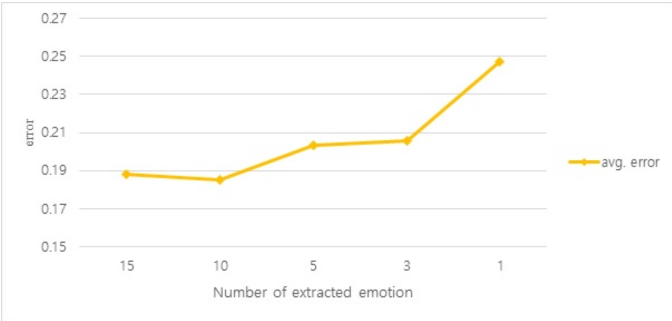


Fig. 9 Performance according to parameter N.

calculate the distance on the color image scale between color combination's estimated grayish/clear factor and the ground truth value.

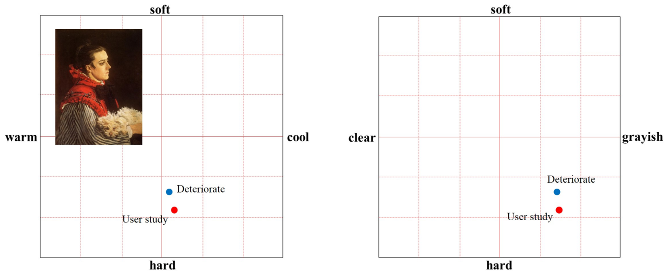
In our experiment, the average distance was 0.478. The range of each of the three factors including the grayish/clear factor was $[-3 : 3]$; thus, the average error rate was about 8% of the whole range. Figure 6 shows a couple of the best and the worst cases of estimation. As shown in this figure, our estimation method works well for both positive and negative emotions indiscriminately.

5.2 User Study for Collecting Ground Truth Emotions of Paintings

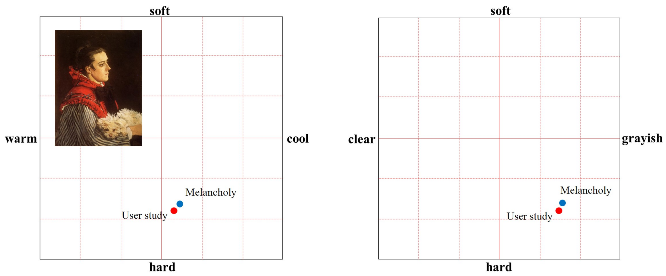
To evaluate the performance of the proposed emotion extraction method, the ground truth emotions of the painting are required. To obtain them, we conducted a user study that asks the user for the three factors of the color image scale for each painting image by using Amazon mTurk [7], which is a crowdsourcing web marketplace for surveys. For 100 painting images, the ground truth values of the three factors were generated by aggregating the workers' scores for each image. Figure 6 shows an example question for scoring on the color image scale of a given painting. For each image, we asked more than 100 workers to select the degree of the three factors considering only the color and the tone. To guarantee the reliability of the response, we set the minimum value of the total approved HITs and the HIT approval rate as 5,000 and 98%, respectively. Figure 7 shows several examples of the gathered ground truth coordinates of the paintings on the color image scale.

5.3 Evaluation of Emotion Extraction

In this experiment, we constructed a dataset containing 950 color combination and emotional word pairs by estimating the grayish/clear factor of the color combination described in Section 3. By using the method proposed in Section 4, we then generated the color spectra from 100 sample painting images and found pairs that were similar to the spectra. Finally, we extracted



(a) 50-pair dataset



(b) Extended and merged 950-pair dataset

Fig. 10 Comparison between the emotions extracted using two datasets.

the emotions of the paintings by averaging the coordinates of the pairs. By comparing with the ground truth emotions of the sample paintings, defined as the coordinates on the color image scale, we evaluated the performance of the proposed method that extracts the emotions of a painting from the viewpoint of accuracy. In this experiment, we evaluated the performance by computing the error as the distance between the coordinates of the ground truth and the averaged emotion. For 100 painting images, the average error and its standard deviation were found to be 1.10 and 0.0084, respectively. As the range of each factor of the color image scale was $[-3 : 3]$, the average error was 18%. Figure 8 shows a couple of examples of the paintings, generated color spectra, matched pairs, extracted emotions, and ground truth emotions. As shown in the figure, compared with the ground truth emotions, the extracted emotional words and the corresponding averaged emotions seem to be reasonable. In some cases (the bottom of Figure 8), the extracted emotions were far from the ground truth. However, in our analysis, the emotions evoked by most of these paintings seemed to be affected more by the subject than the colors of the painting.

Figure 9 shows the performance according to parameter N, which is the number of the selected best matching pairs. In our experiment, when we set the value of N as 10, the best performance was achieved. Hence, in all the other experiments conducted as part of this study, the value of N was set to 10.

Table 1 Examples of the estimated grayish/clear factor of color combinations and its ground truth value.

emotional word	color combination			ground truth	estimated	error
urbane	PB/Gr	PB/Dgr	PB/L	1.576	1.603	0.027
sporty	PB/V	Y/V	YR/V	-2.084	-2.059	0.025
pale	PB/Dl	PB/P	PB/Vp	0.349	0.374	0.025
upset	R/V	RP/Dp	N1.5	-0.807	-0.788	0.019
sad	PB/Gr	N9	N7	1.579	0.779	0.800
cool-headed	B/Dgr	PB/Vp	PB/Dk	-0.867	0.105	0.972

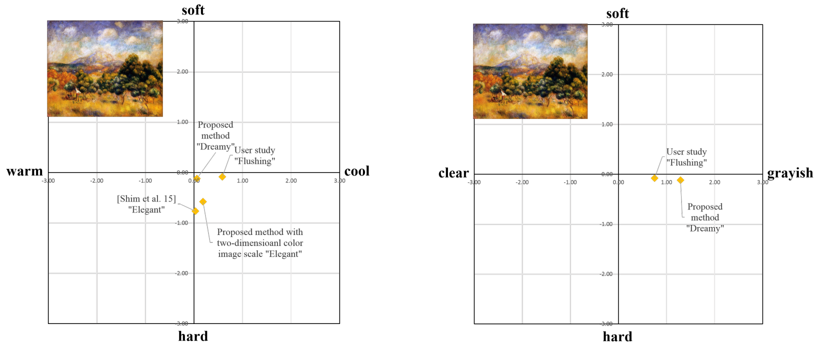
Table 2 Performance of two datasets.

Dataset	Avg. error	Std. dev.
50 pairs	1.22	0.0121
950 pairs	1.10	0.0084

When we extended the dimensions of the dataset, we utilized the 50 pairs presented in NCD's work [25], because they had three-dimensional coordinates. Consequently, without extending the dimensions of the 900-pair dataset defined on a two-dimensional image scale, we can use the set of 50 pairs as the entire dataset for the proposed method. In this experiment, we compared the performance of two datasets, namely a set of 50 pairs, which were defined on a three-dimensional color image scale and a set of 950 pairs; the dimensions of 900 pairs of the latter set were extended, and then, the pairs were merged with the other 50 pairs. For this experiment, the 100 paintings used above were employed to evaluate the performance. Table 1 shows the evaluation results. The performance of the merged dataset was superior to that of the other. Although the 50-pair data set contained the ground truth emotional words that were surveyed in NCD's work [25], it was not sufficient to find a color combination similar to the color spectrum of the input painting, as shown in Figure 10.

As mentioned in Section 2, the preliminary version of this study employed a two-dimensional color image scale and a 900-pair dataset. In this experiment, we compared the performance of method discussed in the preliminary study with that of the method proposed in this work. For the comparison, we considered three methods: one was the preliminary version, another was method proposed in this paper, and the other was the proposed method using a two-dimensional color image scale. Table 2 shows the results of the performance comparison. As shown in this table, the performance of our method is superior to that of the others, because only our method employed a color image scale with all the required dimensions. Figure 11 shows that the method using a two-dimensional color image scale could not extract the emotions of a painting that evoked negative emotions.

The IAPS dataset is widely used in many studies focusing on the emotions evoked by photographic images. Although the IAPS dataset does not contain images of paintings, the proposed method can be adopted by using the colors



(a) (b)

Fig. 11 Comparison between the emotions extracted using two- and three- dimensional color image scales.

Images	Extracted emotion by using our method	Semantic meanings labeled on image
	Tender	Sweetheart
	Horrors	Immediate

Fig. 12 Comparison between ground truth and extracted emotion for the IAPS dataset.

of the photographic images. In this experiment, we extracted the emotions of IAPS images by using the proposed method. Because the ground truth emotion of the IAPS dataset is measured using Russell's emotion model, it is not easy to directly compare emotions in two different models. Therefore, we compared the semantic meanings of the ground truth and the extracted emotion. Figure 12 shows a couple of examples. As shown in this figure, the semantic difference between the ground truth and the extracted emotion does not seem to be significant. Therefore, this experiment shows that the proposed method can be applied to general images with reasonable performance.

6 An emotion-based image exploration system

Although our method proposed in section 4 aims to extract an emotion from painting image, this method can be utilized for extracting the emotion of

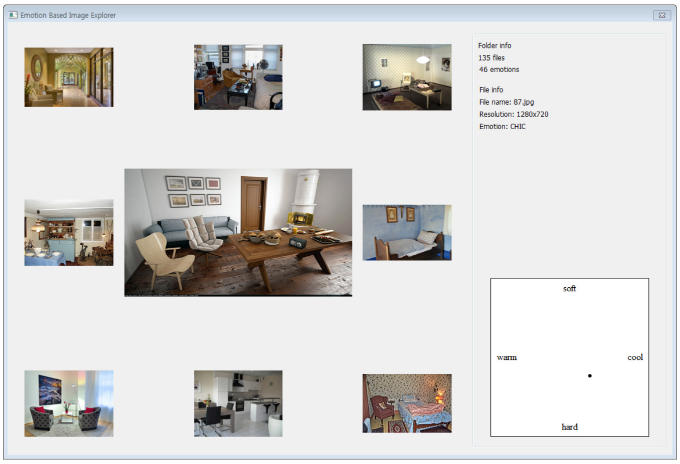


Fig. 13 User interface for our image exploration system.

general color images as well. Recently, many studies [24, 14, 2, 3] proposed image exploration and retrieval systems. In these studies, visual descriptor and semantic descriptor were usually used to find the desired image. Contrary to these studies, we present our exploration system that finds the desired image based on the emotion estimated as described in Section 3. For our purpose, we propose a two-dimensional parameter-based emotion exploration method. Our system consists of two parts: one is an image exploration window, and the other is an information panel that shows a selected image and its emotion on the color image scale in graph form.

In our method, we display images using the two parameters employed for establishing the axes of the color image scale, cool-warm and soft-hard, as shown in Figure 13. In this method, a two-dimensional graph is provided to adjust the parameters. On the left panel, the center image indicates the nearest image from the coordinates that correspond to the current parameters. Around the center image, more warmer-cooler and softer-harder images are displayed. The coordinates of these images are near to the coordinates of the center image. To find these images, we generate a Voronoi diagram [4] using the points on the image coordinates and obtain the neighboring cells of the center image. Initially, the value of each parameter is set to zero, which is the origin of each axis. When the user drags a point on the graph, the parameter value changes. Then, the center image on the left panel also changes. Consequently, images near to the current parameter are displayed around the center image.

7 Conclusions

In this paper, we proposed an algorithm for extracting emotions from paintings on the basis of a color image scale. In the proposed method, we first constructed a color combination and emotional word pair dataset from two different color image scale datasets. For this, we estimated a grayish/clear factor from the colors and merged two datasets on a three-dimensional color image scale. Next, we normalized the input painting with the Hue & Tone 130 color system and generated a color spectrum that describes the colors used in the painting and their ratio. We then searched for the best matching color combinations from the dataset, which were most similar to the color spectrum. Finally, we calculated the average of their coordinates and extracted the nearest corresponding emotional word as the emotion of the painting. Through various experiments, we showed that the proposed method exhibits reasonable performance with respect to predicting the emotions evoked by a painting. Based on this method, we finally propose a painting image exploration system, where users can explore painting images emotionally coherently

In this study, we considered only the colors of a painting among the various factors that affect the emotions evoked by the painting. Although the proposed method works well without considering the other components, an emotion is normally evoked by not only colors but also texture, composition, subject, etc. Consideration of these components along with colors can improve the performance of extracting more precise emotions. Therefore, painting components, such as texture and composition, must be studied. In the future, we will study the extraction of emotions from a painting by using these components.

Further, on the basis of the assumption that a view can evoke a complex emotion, we extracted the top N emotions and obtained their average as the emotion evoked by the painting. Nevertheless, the issue is not resolved because people can simultaneously feel many different emotions. Therefore, the problem of how to present these emotions still remains. To solve this problem, a more sophisticated user study for measuring these emotions is required. This will be also addressed in a future study.

Acknowledgements This research was supported by the Chung-Ang University Research Scholarship Grants in 2014 and Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2014.

References

1. Adams, F.M., Osgood, C.E.: A cross-cultural study of the affective meanings of color. Journal of Cross-Cultural Psychology **4**(2), 135–156 (1973). DOI 10.1177/002202217300400201
2. Ahmad, J., Sajjad, M., Mehmood, I., Rho, S., Baik, S.W.: Saliency-weighted graphs for efficient visual content description and their applications in real-time image retrieval systems. Journal of Real-Time Image Processing pp. 1–17 (2015). DOI 10.1007/s11554-015-0536-0
3. Ahmad, J., Sajjad, M., Rho, S., Baik, S.W.: Multi-scale local structure patterns histogram for describing visual contents in social image retrieval systems. Multimedia Tools Appl. **75**(20), 12,669–12,692 (2016). DOI 10.1007/s11042-016-3436-9. URL http://dx.doi.org/10.1007/s11042-016-3436-9
4. Aurenhammer, F.: Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Comput. Surv. **23**(3), 345–405 (1991). DOI 10.1145/116873.116880. URL http://doi.acm.org/10.1145/116873.116880

5. Han, B.J., Rho, S., Jun, S., Hwang, E.: Music emotion classification and context-based music recommendation. Multimedia Tools Appl. **47**(3), 433–460 (2010). DOI 10.1007/s11042-009-0332-6. URL http://dx.doi.org/10.1007/s11042-009-0332-6
6. Icoglu, O., Gunsel, B., Sariel, S.: Classification and indexing of paintings based on art movements. In: 12th European Signal Processing Conference, pp. 749–752 (2004)
7. Ipeirotis, P.G.: Analyzing the amazon mechanical turk marketplace. XRDS **17**(2), 16–21 (2010). DOI 10.1145/1869086.1869094. URL http://doi.acm.org/10.1145/1869086.1869094
8. Joshi, D., Datta, R., Fedorovskaya, E., Luong, Q.T., Wang, J., Li, J., Luo, J.: Aesthetics and emotions in images. Signal Processing Magazine, IEEE **28**(5), 94–115 (2011). DOI 10.1109/MSP.2011.941851
9. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1, CVPR '06, pp. 419–426. IEEE Computer Society, Washington, DC, USA (2006). DOI 10.1109/CVPR.2006.303. URL http://dx.doi.org/10.1109/CVPR.2006.303
10. Kim, R., Kwak, Y.J., Mo, H., Kim, M., Rho, S., Man, K.L., Chong, W.K.: Trustworthy movie recommender system with correct assessment and emotion evaluation. In: Proceedings of the International MultiConference of Engineers and Computer Scientists, vol. 2 (2015)
11. Kobayashi, S.: The aim and method of the color image scale. Color Research & Application **6**(2), 93–107 (1981). DOI 10.1002/col.5080060210. URL http://dx.doi.org/10.1002/col.5080060210
12. Kobayashi, S.: Color Image Scale. Kosdansha International (1991)
13. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (iaps): Technical manual and affective ratings. NIMH Center for the Study of Emotion and Attention pp. 39–58 (1997)
14. Lee, M.H., Rho, S., Choi, E.I.: Ontology based user query interpretation for semantic multimedia contents retrieval. Multimedia Tools and Applications **73**(2), 901–915 (2014). DOI 10.1007/s11042-013-1383-2
15. Lee, T., Lim, H., Kim, D.W., Hwang, S., Yoon, K.: System for matching paintings with music based on emotions. In: SIGGRAPH ASIA 2016 Technical Briefs, SA '16, pp. 31:1–31:4. ACM, New York, NY, USA (2016). DOI 10.1145/3005358.3005366. URL http://doi.acm.org/10.1145/3005358.3005366
16. Machajdik, J., Hanbury, A.: Affective image classification using features inspired by psychology and art theory. In: Proceedings of the 18th ACM International Conference on Multimedia, MM '10, pp. 83–92. ACM, New York, NY, USA (2010). DOI 10.1145/1873951.1873965. URL http://doi.acm.org/10.1145/1873951.1873965
17. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: Proceedings of the 2011 International Conference on Computer Vision, ICCV '11, pp. 1784–1791. IEEE Computer Society, Washington, DC, USA (2011). DOI 10.1109/ICCV.2011.6126444. URL http://dx.doi.org/10.1109/ICCV.2011.6126444
18. Ou, L.C., Luo, M.R., Woodcock, A., Wright, A.: A study of colour emotion and colour preference. part i: Colour emotions for single colours. Color Research & Application **29**(3), 232–240 (2004). DOI 10.1002/col.20010. URL http://dx.doi.org/10.1002/col.20010
19. Ou, L.C., Luo, M.R., Woodcock, A., Wright, A.: A study of colour emotion and colour preference. part ii: Colour emotions for two-colour combinations. Color Research & Application **29**(4), 292–298 (2004). DOI 10.1002/col.20024. URL http://dx.doi.org/10.1002/col.20024
20. Rho, S., Yeo, S.S.: Bridging the semantic gap in multimedia emotion/mood recognition for ubiquitous computing environment. J. Supercomput. **65**(1), 274–286 (2013). DOI 10.1007/s11227-010-0447-6. URL http://dx.doi.org/10.1007/s11227-010-0447-6
21. Seo, S., Kang, D.: Study on predicting sentiment from images using categorical and sentimental keyword-based image retrieval. J. Supercomput. **72**(9), 3478–3488 (2016). DOI 10.1007/s11227-015-1510-0. URL http://dx.doi.org/10.1007/s11227-015-1510-0
22. Shim, H., Kang, D., Yoon, K.: Extracting emotional adjectives from paintings using color combinations. In: Proceeding of Computer Graphics International 2015 (2015)

23. Sivik, L.: Research on the meanings of color combinations. In: Proceedings of the Congress of the Association Internationale de la Coleur (AIC), pp. 130–132 (1989)
24. Tak, Y.S., Rho, S., Hwang, E., Lee, H.: Tertiary hash tree-based index structure for high dimensional multimedia data. Multimedia Tools and Applications **61**(1), 51–68 (2012). DOI 10.1007/s11042-010-0687-8
25. The Nippon Color & Design Research Institute: A color scheme to convey the mind. Kodansha (2008)
26. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Proceedings of the Sixth International Conference on Computer Vision, ICCV '98, pp. 839–. IEEE Computer Society, Washington, DC, USA (1998). URL http://dl.acm.org/citation.cfm?id=938978.939190
27. Valdez, P., Mehrabian, A.: Effects of color on emotions. Journal of experimental psychology: General **123**(4), 394 (1994)
28. Wright, B., Rainwater, L.: The meanings of color. The Journal of General Psychology **67**(1), 89–99 (1962). DOI 10.1080/00221309.1962.9711531. URL http://dx.doi.org/10.1080/00221309.1962.9711531. PMID: 14008415
29. Yanulevskaya, V., Uijlings, J., Bruni, E., Sartori, A., Zamboni, E., Bacci, F., Melcher, D., Sebe, N.: In the eye of the beholder: Employing statistical analysis and eye tracking for analyzing abstract paintings. In: Proceedings of the 20th ACM International Conference on Multimedia, MM '12, pp. 349–358. ACM, New York, NY, USA (2012). DOI 10.1145/2393347.2393399. URL http://doi.acm.org/10.1145/2393347.2393399
30. Zhao, S., Gao, Y., Jiang, X., Yao, H., Chua, T.S., Sun, X.: Exploring principles-of-art features for image emotion recognition. In: Proceedings of the 22Nd ACM International Conference on Multimedia, MM '14, pp. 47–56. ACM, New York, NY, USA (2014). DOI 10.1145/2647868.2654930. URL http://doi.acm.org/10.1145/2647868.2654930
31. Zujovic, J., Gandy, L., Friedman, S., Pardo, B., Pappas, T.: Classifying paintings by artistic genre: An analysis of features & classifiers. In: IEEE International Workshop on Multimedia Signal Processing (MMSP), 2009, pp. 1–5. Rio de Janeiro, Brazil (2009)