

QUERY-BY-EMOTION SKETCH FOR LOCAL EMOTION-BASED IMAGE RETRIEVAL

Young-Chang Lee

Department of Industrial Design, Yeungnam University, Gyeongsang-do, Korea

Kyoung-Mi Lee

Department of Computer Science, Duksung Women's University, Seoul, Korea

Keywords: Image retrieval, Emotion, Color sketch, Content-based retrieval, Semantic-based retrieval.

Abstract: This paper has proposed an image retrieval system by using an emotion sketch in order to retrieve images that locally hold different emotions. The proposed image retrieval system divides an image into the 17×17 non-overlapping sub-areas. In order to extract the emotion features in sub-areas, the proposed system has used the emotion colors that correspond to the 160 emotion words that suggested by the color imaging chart of H. Nagumo. By calculating the distribution of emotion colors corresponding to the emotion words from the sub-areas, the system takes the emotion word that holds the largest value among the histogram values of the emotion words of each sub-area. The image retrieval system with using the proposed emotion sketch query has demonstrated excellent retrieval precision and recall functions that are better than the global approach by evaluating the validity of the Corel database.

1 INTRODUCTION

The content-based image retrieval system retrieves the images of similar features by analyzing the images and extracting the features (Marsicoi *et al.*, 1997; Yoshitaka and Ichikawa, 1999). However, the degree of awareness in understanding the images may vary from person to person. Moreover, a difference exists between the content representation of images and the semantic requirements of users. Accordingly, many studies have been recently conducted in relation to the semantics-based image retrieval system that detects the semantic representation of images and retrieves the similar images to the semantic requirements of users (Wang and He, 2008). Such semantics-based image retrieval system needs the following technologies (Wang, 2001):

- a function that can automatically extract the semantic representation of image data,
- an interface that can accurately obtain the semantic query request of users, and
- a method that can quantitatively compare the semantic similarities between two images.

The emotional expression of an image refers to

the highest level of semantic expressions, and is expressed in emotional word-like adjectives. In 1998, Japanese researchers attempted to use Kansei information to build image retrieval systems with impression words (Yoshida *et al.*, 1998). Kansei is a Japanese word for emotional semantics. Colombo *et al.* in Italy proposed an innovative method to get high-level representation of art images, which could deduce emotional semantics that consist of action, relaxation, joy and uneasiness (Colombo *et al.*, 1999). K.-M. Lee *et al.* have proposed the emotion retrieval of textile images by using 160 emotional words that are presented by the color image chart (Lee *et al.*, 2008).

As for the studies of emotion-based image retrieval systems as in the above, most of emotion retrieval systems have used a global approach that extracts the emotion of entire images. However, this global approach is not appropriate to extracting the meaning of images that locally hold different emotions. Figure 1 shows part of the retrieved result by using the global emotion that's referred to as the emotion word 'sweet.' Although these images globally hold the 'sweet' features, the meaning of images can vary by how and where the 'sweet' features are distributed. On the other hand, since the

local approach extracts the features of each sub-area within the image, it is appropriate to more accurately retrieve the desired emotion image by user.



Figure 1: Retrieved results by a global approach: 'sweet'.

In order to resolve this problem, this paper proposes the local emotion extraction and retrieval method with using the sketch query method. The images are retrieved by using the 160 diversified emotion words and the emotional similarities between two images are calculated by considering the semantic relations between emotion words.

2 EMOTION COLORS

In order to implement the emotion-based image retrieval system, the emotions in images have to be represented first. This paper uses the method that substitutes colors into emotion words to extract the emotional meaning that's intrinsic to the images. These emotion colors are obtained from the color imaging chart of Nagumo (Nagumo 2000). The imaging chart sets the vertical axis and horizontal

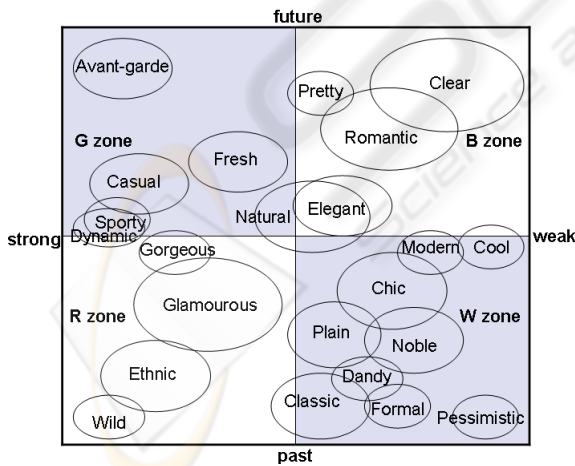


Figure 2: Imaging chart of Nagumo.

axis to time and energy respectively (Figure 2). The chart has intended to understand the emotion of images by analyzing them into these two elements and by obtaining their coordinates. The time axis

indicates future as it moves up (+) and represents past as it moves down (-) from the center. The axis 'energy' gets stronger as it moves to the left and gets weaker as it moves to the right.

The color imaging chart is divided into 4 image zones such as B (Budding), G (Growth), R (Ripen) and W (Withering) by the axes of time and energy. There are 23 standard groups ($g_1 \dots g_{23}$) in 4 image zones. Each standard group is indicated by the center, major-axis and minor-axis of oval. The center coordinates of each group were obtained by considering the intersection of two axes as the origin and by calculating the relative position and size of each group in the imaging chart.

3 IMAGE RETRIEVAL USING EMOTION SKETCH

This section introduces a method that automatically extracts the emotion features of images in the sketch method by using the color emotion-word in the Section 2 and this method calculates the similarity between images.

3.1 Extraction of Emotion Words

Color R	Color G	Color B	Romantic	Innocent	Limber	Tender
0	157	165	1	-1	-1	-1
67	164	113	-1	2	-1	-1
150	200	172	1	-1	3	4
247	197	200	1	2	-1	4

Figure 3: Relation between colors and emotion words.

The emotion-based image retrieval system needs to extract the emotion words that are included in the image. For each emotion word, Nagumo has suggested a color palette that is made up of 6~24 emotion colors. The emotion colors correspond to the 160 color emotion words, and each color can hold multiple color emotion words. In other words, one color can hold romantic and lively emotions at the same time, as is shown in the Figure 3. Therefore, the relations of 160 color emotion words that correspond to each color were obtained by being considered the emotion number, if holding an emotion word for each color, or by being set -1. The total emotion colors that are arranged in this way is 120 units.

In the sketch method, the image is divided into a total of 289 sub-areas (17×17) ($E_{1,1} \dots E_{17,17}$). In this paper, in order to obtain the emotion words of

each sub-area, we take a histogram ($H_1 \dots H_{160}$) of 160 emotion words by using all the pixels within this sub-area, and we determine the emotion word that has the largest number of emotion words as being the representative emotion word. In other words, as for the sketch value of sub-area, this paper takes the emotion word that holds the greatest histogram value for being the emotion word of a sub-area:

$$E_{xy} = \max_{1 \leq b \leq 160} H_{xy,b} \quad (1)$$

Since the color difference may occur in the input process, such as image scanning, we consider all the colors that have the distance within the allowable error range for the RGB value as identical emotion colors. Since the color change is too small to be visually distinguished by humans, this paper has set the error range as '10'.

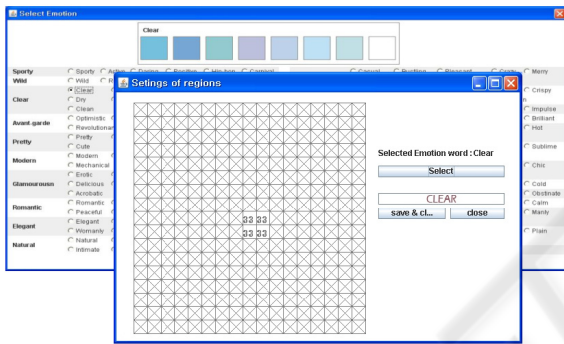


Figure 4: Screenshot of emotion sketch query : 'clear'.

Figure 4 shows an input screen of emotion sketch query. This paper has provided a means of selecting emotion words for the desired part of an image to be retrieved from the images that are divided into 289 sub-areas.

3.2 Similarity of Emotion Words

In order to retrieve the images of similar emotion from the sketch query, the distance d_E between two images A and B can be calculated using Eq. (1) as:

$$d_E(A, B) = \sqrt{\sum_{y=1}^{17} \sum_{x=1}^{17} (E_{xy}^A - E_{xy}^B)^2} \quad (2)$$

Here, E_{xy}^A and E_{xy}^B refers to the x^{th} , y^{th} emotion word from the emotion sketch of images A and B.

Since the sketch value E_{xy}^A and E_{xy}^B of Eq.(2) is one of 160 emotion words, the similarity of emotion words can be calculated by digitizing emotion words. It would be difficult to correctly reflect the actual relationships between emotion words with using the

simple differences between the numbers that are set to emotion words as in Eq.(2). Thus, this paper has used the relationship of standard groups that are obtained from the color imaging chart that's shown in the Figure 2. In other words, the system calculates the distance between standard groups G_{xy} that include the corresponding emotion words and the emotion words within the group are then compared with each other. The Eq.(2) may be converted as follows:

$$d_E(A, B) = \sum_{y=1}^{17} \sum_{x=1}^{17} \{d(G_{xy}^A, G_{xy}^B) + d(E_{xy}^A, E_{xy}^B)\}$$

Here $E_{xy} = e_k \in g_n = G_{xy}$. The distance between standard groups can be obtained by using the center (cx_n, cy_n) of these standard groups.

$$d(G_{xy}^A, G_{xy}^B) = \sqrt{(cx_n^A - cx_n^B)^2 + (cy_n^A - cy_n^B)^2}$$

It is not easy to digitize the distance E_{xy} between emotion words. This paper has considered all the emotion words within the same group as being individual emotions and it has defined them as follows:

$$d(E_{xy}^A, E_{xy}^B) = \begin{cases} 0 & E_{xy}^A = E_{xy}^B \\ t & E_{xy}^A \neq E_{xy}^B \end{cases}$$

Here, t can be smaller than the distance between the closest standard groups. As shown in Figure 2 and Table 1, this paper has set t to '0.3' since the distance between 'sporty' and 'dynamic' that are the closest groups within the color imaging chart is about 0.36.

4 EXPERIMENTAL RESULTS

The database of the proposed image retrieval system was designed by using JAVA and MySQL on Windows XP platform of Pentium @ 4 CPU 3.0GHz and 1GB RAM. As for the test images, we have used the Corel database that is most commonly used for the performance evaluation of the content-based image retrieval method. The Corel database is comprised of 9,908 images that belong to the semantic category such as butterfly, cosmos, sunrise/sunset, flower, character, mountain, national flag and boat.

Figure 5 shows the retrieval results of the center areas of the images that hold the emotion 'clear' by using the emotion sketch of Figure 4. Also by using the emotion sketch, it is possible to retrieve the images where two emotions are mixed. Figure 6

shows the image retrieval result that holds the ‘clear’ emotion in the center and the ‘street-fashion’ emotion in the circumference.



Figure 5: Retrieved results using Figure 2.



Figure 6: Retrieved results using two emotions ‘clear’ and ‘street-fashion’.

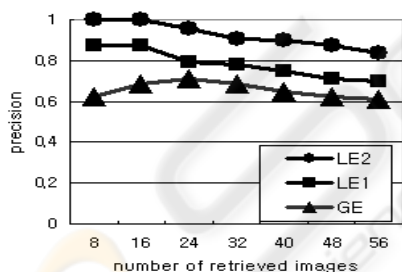


Figure 7: Retrieved results on the Corel database.

This paper has assessed the retrieval results according to the precision and the number of retrieved images. Precision is defined as the ratio of the properly retrieved images to all the retrieved images, and it is used to evaluate the ability of a system to retrieve only proper images. The number of retrieved images refers to the number of images that are retrieved from the Corel database.

Figure 7 shows the average that has queries of 20 emotion words from the Corel database. The experiment has been made for global emotion retrieval (GE), local emotion retrieval by using only one sub-area (LE1), and local emotion retrieval by

using two sub-areas (LE2). Figure 7 shows that the retrieval result of using the proposed emotion sketch gives better precision than the global approach.

5 CONCLUSIONS

This paper has proposed the image retrieval system of using emotion sketch. The proposed retrieval system as a content-based image retrieval system includes the following technologies:

- Automatically extracting emotions from images by using emotion colors and emotion color words.
- Allowing the user to easily select the emotions of each sub-area by providing convenient GUI as is shown in Figure 4.
- Enabling the comparison of the emotional similarities between images by suggesting a method that can quantitatively calculate the similarity of emotion words.

Also as compared to the results of existing emotion-based image retrieval studies, this paper has proposed a method that can find various emotions that are scattered on the image rather than finding one global emotion from one image.

REFERENCES

Colombo, C., Del Bimbo, A., and Pala, P., 1999. "Semantics in visual information retrieval," *IEEE Multimedia*, vol. 6, no. 3, pp. 38-53.

Lee, Kyoung-Mi, et. al, 2008. "Textile image retrieval integrating contents, emotion and metadata," *Journal of KSII*, vol. 9, no. 5, pp. 99-108.

Marsicoi, M. De, Cinque, L. and Levialdi, S., 1997. "Indexing pictorial documents by their content: A survey of current techniques," *Image vision computing*, vol. 15, pp. 119-141.

Nagumo, H., "Color image chart," Chohyung, 2000.

Wang, W. and He, Q., 2008. "A survey on emotional semantic image retrieval," In *International Conference on Image Processing*, pp. 117-120.

Wang, S., 2001. "A robust CBIR approach using local color histograms," *Technical report TR 01-13*, U. of Alberta Edmonton, Alberta, Canada.

Yoshida, K., Kato, T., and Yanaru, T., 1998. "Image retrieval system using impression words," *IEEE Transactions on Systems, Man, and Cybernetics*, vol.3, no.11-14, pp. 2780-2784.

Yoshitaka, A. and Ichikawa, T., 1999. "A survey on content-based retrieval for multimedia databases," *IEEE transactions on Knowledge Data Engineering*, vol. 11, pp. 81-93.