Image (Artwork) Retrieval based on Similarity of Touch by Self-Organizing Map with Refractoriness

Kento KADOKURA and Yuko OSANA

School of Computer Science, Tokyo University of Technology
1404-1 Katakura, Hachioji, Tokyo, 192-0982, Japan Email: osana@stf.teu.ac.jp

Abstract—In this paper, we propose an image (artwork) retrieval based on touch by self-organizing map with refractoriness. Most of conventional similarity-based (content-based) image retrieval systems deal with only scenery image, landscape photographs and so on. Only a few retrieval systems for artwork (illustration) have been proposed. However, they can not realize the retrieval based on touch of artwork. The proposed system realizes the retrieval based on similarity of touch using the self-organizing map with refractoriness. In this system, the number of colors, color information, outline information and so on are used as the image feature. We carried out computer experiments in the proposed system which stores 500 images and confirmed that the proposed image retrieval system can retrieve artwork whose touch are similar.

1. Introduction

Recently, some similarity-based image retrieval systems using neural networks which make use of flexible and soft information processing ability of the neural networks have been proposed[1]-[5]. The main target of these systems are scenery images, and most of these image retrieval systems which use color information as feature can retrieve similar scenery images appropriately. However, most of the conventional similarity-based (content-based) image retrieval systems [1]-[6] deal with only scenery image, landscape photographs and so on. Although only a few retrieval systems for artwork (illustration) have been proposed, they can not realize the retrieval based on similarity of touch of artwork.

In the consideration of similarity of touch, hue in images does not need to be similar. If saturation and brightness are similar even if hue is different, the touch of two images has a high possibility of being similar. So, different suitable features which are used in the retrieval from the case of search of a scenery images have to be selected.

In this paper, we propose an image (artwork) retrieval based on similarity of touch by self-organizing map with refractoriness. In this system, the number of colors, color information, outline information and so on are used as the image features.

2. Image Features

2.1. The Number of Colors

In the proposed image retrieval system, the normalized number of colors is used as one of the image feature. The normalized number of colors of the image p, $x^{N1(p)}$ is given by

$$x^{N1(p)} = \frac{N_p^{N1} - N_{min}^{N1}}{N_{max}^{N1} - N_{min}^{N1}} \quad (1)$$

where N_p^{N1} is the number of colors in the image p, N_{max}^{N1} is the maximum number of colors in whole stored images and N_{min}^{N1} is the minimum number of colors in whole stored images.

In the proposed image retrieval system, the normalized average number of colors in each divided area is also used as the image feature. Here, the divided areas by the K-means algorithm[7] and the features in each divided area are considered. The normalized average number of colors in each area of the image p, $x^{N2(p)}$ is given as

$$x^{N2(p)} = \frac{N_p^{N2} - N_{min}^{N2}}{N_{max}^{N2} - N_{min}^{N2}} \quad (2)$$

where N_p^{N2} is the average number of colors in each divided area of the image p, N_{max}^{N2} is the maximum number of colors in each area in whole stored images and N_{min}^{N2} is the minimum number of colors in each area in whole stored images.

2.2. Color Information

2.2.1. *Saturation and Brightness*

In the consideration of similarity of touch, hue in images does not need to be similar. If saturation and brightness are similar even if hue is different, the touch of two images has a high possibility of being similar. So, the average saturation and brightness are also used as the image features. The saturation and brightness in the image p which has $N^{(p)}$ pixels, $x^{S(p)}$ and $x^{V(p)}$ are given by

$$x^{S(p)} = \frac{1}{N^{(p)}} \sum_{i=1}^{N^{(p)}} s_i^{(p)} \quad (3)$$

$$x^{V(p)} = \frac{1}{N^{(p)}} \sum_{i=1}^{N^{(p)}} v_i^{(p)} \quad (4)$$

where $s_i^{(p)}$ is the saturation of the pixel i in the image p, and $v_i^{(p)}$ is the brightness of the pixel i in the image p.

Moreover, the average saturation and brightness in the pixels which has similar hue are also used as the image feature. The average saturation and brightness in the pixels which have hue belongs the group j (hue is divided into N^H

groups), $x^{S_j(p)}$ and $x^{V_j(p)}$ are given by

$$x^{S_j(p)} = \frac{1}{N^{H_j}} \sum_{i:h_i^{(p)} \in C^{H_j}} s_i^{(p)} \quad (5)$$

$$x^{V_j(p)} = \frac{1}{N^{H_j}} \sum_{i:h_i^{(p)} \in C^{H_j}} v_i^{(p)} \quad (6)$$

where N^{H_j} is the number of pixels whose hue belongs the hue group j, and $h_i^{(p)}$ is the hue of the pixel i in the image p. And C^{H_j} is the set of hue which belongs to the hue group j and it is given by

$$C^{H_j} = \left\{ h \,\middle|\, \frac{360(j-1)}{N^H} \leq h < \frac{360(j)}{N^H} \right\} \quad (j = 1, \ldots, N^H). \quad (7)$$

2.2.2. Contrast

The contrast is a difference of the brightest portion and the darkest portion in the image. Here, the contrast is calculated using the value instead of the brightness. The average and maximum contrast in the image p, $x^{C1(p)}$ and $x^{C2(p)}$ are given by

$$x^{C1(p)} = \frac{1}{N^{R(p)}} \sum_{m=1}^{N^{R(p)}} \frac{v_{max}^{(p,m)} - v_{min}^{(p,m)}}{v_{max}^{(p,m)} + v_{min}^{(p,m)}} \quad (8)$$

$$x^{C2(p)} = \max_{m} \left(\frac{v_{max}^{(p,m)} - v_{min}^{(p,m)}}{v_{max}^{(p,m)} + v_{min}^{(p,m)}} \right) \quad (9)$$

where $N^{R(p)}$ is the number of divided areas in the image p. In the proposed image retrieval system, each stored image is divided into some areas by K-means algorithm. And, $v_{max}^{(p,m)}$ is the maximum value of the area m in the image p, and $v_{min}^{(p,m)}$ is the minimum value of the area m in the image p.

2.3. Outline Information

In the proposed system, the existence or non-existence of outline ($x^{L1(p)}$) is used as the image feature. Moreover, the rate that the pixel which is judged as outline ($x^{L2(p)}$), the value of the outline ($x^{L3(p)}$), the rate of the transparent color ($x^{T(p)}$) are also used as the image features.

3. Image (Artwork) Retrieval System based on Similarity of Touch

3.1. Structure

The proposed image retrieval system is based on the self-organizing map with refractoriness[2], and it has the input layer and the map layer. The input layer is composed of three parts corresponding to (1) the number of colors, (2) color information and (3) outline information. And each stored image is trained to correspond to one of the neuron in the map layer. In the proposed system, the feature vector for the key image is given to the neurons in the input layer as the query, and the neurons corresponding to the images which have similar touch fire sequentially.

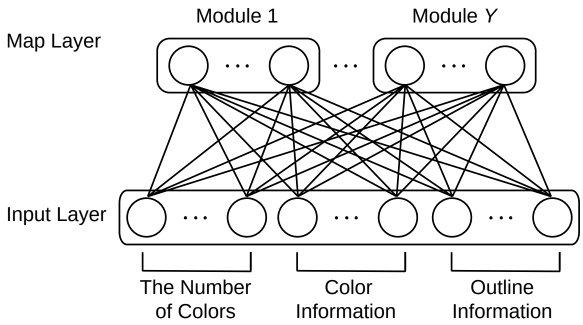


Figure 1: Structure of Proposed Image Retrieval System.

3.2. Learning Process

The learning process of the proposed image retrieval system is composed of two steps; (1) generation of feature vector and (2) learning of feature vectors of the self-organizing map with refractoriness.

3.2.1. Generation of Feature Vector

The feature vectors of stored images are generated. The feature vector is composed of three parts; (1) the number of colors, (2) color information and (3) outline information. The feature vectors are generated by the method described in **2**. The feature vector of the image p, $\boldsymbol{x}^{(p)}$ is given by

$$\boldsymbol{x}^{(p)} = \Big(x^{N1(p)}, x^{N2(p)}, x^{S(p)}, x^{V(p)}, x^{S_1(p)}, \cdots, x^{S_{NH}(p)}, x^{V_1(p)}, \cdots, x^{V_{NH}(p)}, x^{C1(p)}, x^{C2(p)}, x^{L1(p)}, x^{L2(p)}, x^{L3(p)}, x^{T(p)} \Big)^T \quad (10)$$

where $x^{N1(p)}$ and $x^{N2(p)}$ are the features on the number of colors, $x^{S(p)}$ and $x^{S_1(p)}$, $\cdots$, $x^{S_{NH}(p)}$ are the features on saturation, $x^{V(p)}$ and $x^{V_1(p)}$, $\cdots$ $x^{V_{NH}(p)}$ are the features on value, $x^{C1(p)}$ and $x^{C2(p)}$ are the features on contrast, and $x^{L1(p)}, x^{L2(p)}, x^{L3(p)}$ and $x^{T(p)}$ are the features on outline.

3.2.2. Learning of Feature Vectors in Self-Organizing Map with Refractoriness

In this step, the generated feature vectors of stored images are memorized in the self-organizing map with refractoriness. In the proposed system, each feature vector is memorized to correspond to one of the neurons in the map layer.

3.3. Retrieval Process

In the retrieval process, first, the feature vector of the key (query) image is generated. And then, it is given to the input layer of the self-organizing map with refractoriness. The retrieval procedure is as follows:

Step 1 : Input of Feature Vector of Key Image
The feature vector of the key image $\boldsymbol{x}$ is generated and it is given to the input layer of the proposed system.

Step 2 : Calculation of Internal States in Map Layer

The internal state of the neuron i of the module y in the map layer at the time t, $u_i^y(t)$ is calculated by

$$u_i^y(t) = 1 - \frac{D_r(\boldsymbol{w}_i^y, \boldsymbol{x})}{\sqrt{F'}} - \alpha \sum_{d=0}^{t} k_r^d x_i^{MAP(y)}(t-d) \quad (11)$$

where F' is the number of image features which are used for the retrieval, α is the scaling factor of the refractoriness, k_r is the damping factor of the refractoriness and it takes between 0 and 1. $x_i^{MAP(y)}(t)$ is the output of the neuron i of the module y at the time t. And $D_r(\boldsymbol{w}_i^y, \boldsymbol{x})$ is the distance between the weight vector of the neuron i of the module y in the map layer and the input feature vector $\boldsymbol{x}$, and it is given by

$$D_r(\boldsymbol{w}_i^y, \boldsymbol{x}) = \sqrt{\sum_{f=1}^{F} \mu(f) \sum_{j: x_j \in C_f} (\phi(w_{ij}^y, x_j))^2} \quad (12)$$

where f $(= 1, \cdots, F)$ denotes the image feature (1 : the number of colors, 2 : color information, 3: outline information). C_f is the input set for the feature f, w_{ij}^y is the connection weight between the neuron j in the input layer and the neuron i of the module y in the map layer, and x_j is the jth element of the input feature vector. And, $\mu(f)$ is the weighting coefficient for the image feature and it is given by

$$\mu(f) = \begin{cases} \dfrac{1}{N'^{(f)}} & (0 < N'^{(f)}) \\ 0 & (N'^{(f)} = 0) \end{cases} \quad (13)$$

where $N'^{(f)}$ shows the number of neurons which are used in the distance calculation in Eq.(12). And $\phi(w_{ij}^y, x_j)$ is given by

$$\phi(w_{ij}^y, x_j) = \begin{cases} 0, & (x_j = -1 \text{ or } w_{ij}^y = -1) \\ w_{ij}^y - x_j & (\text{otherwise}) \end{cases} \quad (14)$$

In the proposed image retrieval system, the feature which does not exist in the image is set to -1.

In each module, the neuron whose internal states is maximum in the map layer is selected as the winner neuron $c^{(y)}$.

Step 3 : Output of Neurons in Map Layer

The output of the neuron i of the module y in the map layer $x_i^{MAP(y)}(t)$ is determined based on the internal states of neurons in the map layer calculated in **Step 2** and it is calculated by

$$x_i^{MAP(y)}(t) = \begin{cases} 1 & (i = c^{(y)} \text{ and } u_i^y(t) > \theta_1 \\ & \text{ and } D_{max}^{y(i)} < \theta_2) \\ 0 & (\text{otherwise}) \end{cases} \quad (15)$$

where θ_1 and θ_2 are the thresholds. And $D_{max}^{y(i)}$ is the maximum distance per feature of the neuron i of the module y in the map layer and it is given by

$$D_{max}^{y(i)} = \max_f \left(\mu(f) \sum_{j \in C_f} (\phi(w_{ij}^y, x_j))^2 \right). \quad (16)$$

The image corresponding to the fired neuron in the map layer is output as the retrieval result.

Step 4 : Repeat

Steps 2 and **3** are repeated.

4. Computer Experiment Results

In this experiment, we used the proposed system which memorizes 500 images which can be divided into 25 groups based on the similarity of touch.

Figure 2 shows a part of the retrieval results of the proposed system. In this figure, top four images which are retrieved are shown. As shown in Fig.2, the proposed system can realize the image retrieval based on similarity of touch.

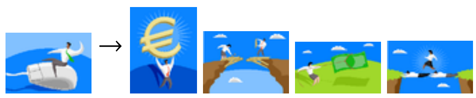
(a) Group 1

(b) Group 2

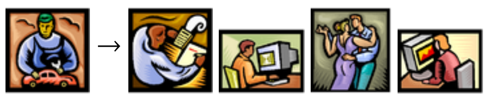
(c) Group 3

(d) Group 4

(e) Group 5



(f) Group 6



(g) Group 7

(h) Group 8

(i) Group 9

(j) Group 10

(k) Group 11

(l) Group 12

(m) Group 13

(n) Group 14

(o) Group 15

(p) Group 16

(q) Group 17

(r) Group 18

(s) Group 19

(t) Group 20

(u) Group 21

(v) Group 22

(w) Group 23

(x) Group 24

(y) Group 25

Figure 2: Retrieval Results of Proposed System.

5. Conclusion

In this paper, we have proposed the image (artwork) retrieval based on similarity of touch using self-organizing map with refractoriness. The proposed system realize the image retrieval based on similarity of touch by the self-organizing map with refractoriness. In this system, the number of colors, color information, outline information and so on are used as image features. We carried out some computer experiments in order to demonstrate the effectiveness of the proposed system and confirmed that the proposed system can retrieve artworks whose touch are similar.

References

[1] J. T. Laaksonen, J. M. Koskela, S. P. Laakso and E. Oja : "PicSOM –content-based image retrieval with self-organizing maps," Pattern Recognition Letters, Vol.21, No.13, 14, pp.1199–1207, 2000.

[2] H. Mogami, M. Otake, N. Kouno and Y. Osana : "Self-organizing map with refractoriness and its application to image retrieval," Proceedings of International Joint Conference on Neural Networks, Vancouver, 2006.

[3] Y. Kuramochi, R. Fukazawa and Y. Osana : "Similarity-based image retrieval considering artifacts from a part of images by self-organizing map with refractoriness," Proceedings of IEEE International Conference on System, Man and Cybernetics, Seoul, 2012.

[4] Y. Morimoto and Y. Osana : "Similarity-based image retrieval considering artifacts from plural key images by self-organizing map with refractoriness," Proceedings of International Symposium on Nonlinear Theory and its Applications, Palma de Mallorca, 2012.

[5] I. Kitano and Y. Osana : "Similarity-based image retrieval considering position fluctuation by self-organizing map with refractoriness," Proceedings of IASTED Artificial Intelligence and Applications, Innsbruck, 2013.

[6] M. Lux and S. A. Chatzichristofis : "Lire: Lucene image retrieval: an extensible Java CBIR library," Proceedings of the 16th ACM international conference on Multimedia, pp.1085–1088, 2008.

[7] J. B. MacQueen : "Some methods for classification and analysis of multivariate observations," Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol.1, No.297, University of California Press, 1967.

[8] T. Kohonen : Self-Organizing Maps, Springer, 1994.