2014 International Symposium on Nonlinear Theory and its Applications
NOLTA2014, Luzern, Switzerland, September 14-18, 2014

Similarity-based Image Retrieval considering Shape and Texture by Self-Organizing Map with Refractoriness

Ikko MIURA 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 a similarity-based image retrieval considering shape and texture by self-organizing map with refractoriness. Most of the conventional image retrieval systems are used color information mainly, and feature on shape and texture are not fully considered. We carried out a series of computer experiments and confirmed the effectiveness of the proposed system.

1. Introduction

Recently, some similarity-based image retrieval systems which make use of the flexible information processing ability of artificial neural networks have been proposed[1]-[10]. Most of these systems use color information as feature, the image retrieval for scenery images can be realized. However, in most of these systems, feature on shape and texture are not fully considered.

In this paper, we propose a similarity-based image retrieval considering shape and texture by self-organizing map with refractoriness.

2. Image Features

In the proposed system, color, color and size of artifacts, shape (distance from circumcircle)[12], SIFT (Scale-Invariant Feature Transform)[13], HOG (Histograms of Oriented Gradients)[14], spectrum, LBP (Local Binary Pattern)[15] and keywords are used as image features.

2.1. Color

In the proposed system, each image is divided into some regions by the K-means algorithm[11]. Then, normalized average x, y and z coordinates in the HSV color space on each region are trained in the self-organizing map, and are used as image feature. Here, two self-organizing maps for natural objects and artifacts are used. In the proposed system, the features on color are calculated per nine areas.

The feature vector on the color of the natural objects at the area s in the image p, $\boldsymbol{x}^{N(p,s)}$ is given by

$$x_i^{N(p,s)} = g\left(\sum_{l\in C_s^N} x_i^{N(p,s,l)}\right) \tag{1}$$

$$g(u) = \begin{cases} 1 & (u > 0) \\ 0 & (u = 0) \end{cases} \tag{2}$$

where C_s^N is the set of the regions which belong to the area s for the natural objects, $x_i^{N(p,s,l)}$ is the output of the neuron i in the self-organizing map which learns the color of the natural objects when the color information at the region l which belongs to the area s in the image p is given.

The feature vector on the color of the artifacts at the area s in the image p, $\boldsymbol{x}^{A(p,s)}$ is given by

$$x_i^{A(p,s)} = g\left(\sum_{l\in C_s^A} x_i^{A(p,s,l)}\right) \tag{3}$$

where C_s^A is the set of the regions which belong to the area s for the artifacts, $x_i^{A(p,s,l)}$ is the output of the neuron i in the self-organizing map which learns the color of the artifacts when the color information at the region l which belongs to the area s in the image p is given.

2.2. Color and Size of Artifacts

The feature vector on color of artifacts $\boldsymbol{x}^{A2(p)}$ is calculated by

$$x_i^{A2(p)} = \sum_{l\in C^A} x_i^{A(p,l)} r^{A(p,l)} \tag{4}$$

where C^A is the set of artifact areas, $x_i^{A(p,l)}$ is the output of the neuron i of the self-organizing map which learns the color of the artifacts when the color information at the region l which belongs to the area s in the image p is given, $r^{A(p,l)}$ is the rate in the whole artifacts areas of the area l of the image p.

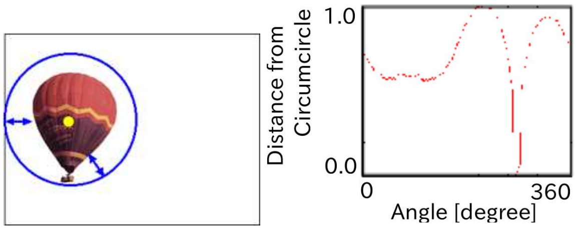
The feature on size of artifacts $x^{A3(p)}$ is calculated by

$$x^{A3(p)} = \frac{\sum_{l\in C^A} S^{(p,l)}}{\sum_{l} S^{(p,l)}} \tag{5}$$

where $S^{(p,l)}$ is the number of pixels of the region l of the image p.

2.3. Shape (Distance from Circumcircle)

As the method to describe the shape, we use the distance between the point on the circumcircle and the edge of the object [12]. In this method, first, the center of the object is found by the moments, and the circumcircle centering on the point is drawn. Then the distance from the point on the circumference to the edge of the object toward the center is



(a) Circumcircle (b) Distance from Circumcircle

Figure 1: Shape (Distance from Circumcircle).

calculated. Figure 1 shows an example of the circumcircle of the object and the corresponding shape information.

2.4. SIFT (Scale-Invariant Feature Transform)

SIFT is an algorithm to detect and describe local features in images. In the proposed system, the feature vector on SIFT is generated based on the idea of the Bag-of-Features[16].

2.5. HOG (Histograms of Oriented Gradients)

HOG is feature descriptors which is used for object detection. The feature vector on HOG is generated based on the idea of the Bag-of-Features as similar as the feature vector on SIFT.

2.6. Spectrum

An image spectrum is used as one of image features. Here, the calculated spectrum is divided into $N^k \times N^l$ areas, and binarized average spectrum in each area is used.

2.7. LBP (Local Binary Pattern)

LBP is the texture features which uses patterns that shows the magnitude relation of local brightness in an image. In this system, normalized histogram of LBP is used as one of image features.

2.8. Keywords

Keywords such as sky, cloud, tree also can be used as the query.

3. Similarity-based Image Retrieval using Self-Organizing Map with Refractoriness

3.1. Structure

The proposed system is based on the self-organizing map with refractoriness[2] and it has two layers; (1) Input Layer and (2) Map Layer. The neurons in the Input Layer receives the feature vector of key image(s) and the neurons in the Map Layer whose connection weights are similar to the input feature vector fires. In the proposed system, each neuron in the Map Layer corresponds to one of the stored images.

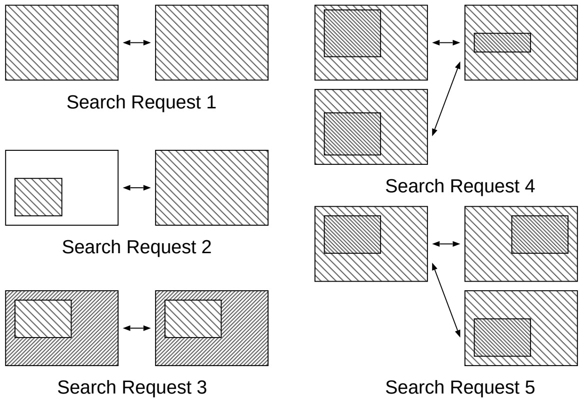


Figure 2: Search Requests 1~5.

3.2. Learning Process

In the learning process of the proposed system, image features of the images to be stored are trained in the self-organizing map with refractoriness.

Step 1 : Extraction of Artifacts

In **Step 1**, the original image is divided into some regions by the K-means algorithm, and whether or not artifacts are included is judged for each divided area.

Step 2 : Generation of Feature Vectors

In **Step 2**, the image features are extracted from the images to be stored, and the feature vectors are generated.

Step 3 : Learning of Self-Organizing Map with Refractoriness

In **Step 3**, the feature vectors generated in **Step 2** are trained in the self-organizing map with refractoriness.

3.3. Image Retrieval Process

In the proposed system, the following five search requests are considered (Fig.2) as similar as the conventional system[10], and the feature vector is generated for the key image based on the search request which is selected by a user.

(1) Retrieval of images which are similar to key image

(2) Retrieval of images which are similar to a part of key image

(3) Retrieval of images which have similar feature of same positions

(4) Retrieval of images which have similar features to common features in plural key images

(5) Retrieval of images which include similar artifacts in different position

(1) Generation of Feature Vector

The feature vector is generated for the key image(s) based on the search request which is selected by a user.

If Search Request 3 is selected, the feature vector on color information of natural objects is given by

$$x_i^{N(p,s)} = \begin{cases} -1 & \text{(if all regions in area } s \text{ are outside the selected part by users)} \\ g\left(\sum_{l \in C_s^N} x_i^{N(p,s,l)}\right) & \text{(otherwise)} \end{cases} \quad (6)$$

where C_s^N is the set of the regions which belong to the area s for the natural objects, $x_i^{N(p,s,l)}$ is the output of the neuron i in the self-organizing map which learns the color of the natural objects when the color information at the region l which belongs to the area s in the image p is given. In the proposed system, if all regions in area s are outside the selected part by users are set to -1. In the same way, the feature vector for the color information of artifacts is generated. And, spectrum and LBP are not used.

If Search Request 4 is selected, the feature vector is generated from the plural key images. In the proposed system, first, the feature vector for each key image $\boldsymbol{x}^{(p)} (p = 1, \cdots, N^{key})$ is generated. Here, N^{key} is the number of key images. Then, the feature vector for the plural key images $\boldsymbol{x}$ is generated from $\boldsymbol{x}^{(p)}$ as follows:

$$x_i = \begin{cases} x_i^{(p)} & \left(\sum_{p=1}^{N^{key}-1} \sum_{q=p+1}^{N^{key}} |x_i^{(p)} - x_i^{(q)}| = 0\right) \\ -1 & \text{(otherwise)} \end{cases} \quad (7)$$

As shown in Eq.(7), the proposed system uses only the common features in all key images.

(2) Image Retrieval

The image retrieval process of the proposed system has four steps.

Step 1 : Input of Feature Vector

Image features of the key image(s) are given to the Input Layer.

Step 2 : Calculation of Internal States of Neurons in Map Layer

When the image feature of the key image(s) is given to the Input Layer, the internal state of the neuron i of the module y in the Map Layer at the time $t+1$, $u_i^y(t+1)$ is calculated by

$$u_i^y(t+1) = 1 - \frac{D_r(\boldsymbol{w}_i^y, \boldsymbol{x})}{\sqrt{F'}} - \alpha \sum_{d=0}^{t} k_r^d x_i^{MAP(y)}(t-d) \quad (8)$$

where F' is the number of image features which is used in the retrieval process, α is the scaling factor, k_r is the damping factor, and $x_i^{MAP(y)}(t)$ is the output of the neuron i of the module y in the Map Layer at the time t. $D_r(\boldsymbol{w}_i^y, \boldsymbol{x})$ is the distance between the weight vector of the neuron i of the module y in the Map Layer $\boldsymbol{w}_i^y$ and the input $\boldsymbol{x}$, and is given by

$$D_r(\boldsymbol{w}_i^y, \boldsymbol{x}) = \sqrt{\sum_{f=1}^{F} \mu(f) \sum_{j: x_j \in C_f} (\phi(w_{ij}^y, x_j))^2} \quad (9)$$

where $f (= 1, \cdots, F)$ is an image feature (1 : color (natural objects), 2 : color (artifacts), 3 : color ratio of artifacts, 4 : size of artifacts, 5 : shape (distance from circumcircle), 6 : SIFT, 7 : HOG, 8 : spectrum (natural objects), 9 : spectrum (artifacts) 10 : LBP, 11 : keywords). C_f is the set of the inputs corresponds to the feature f, w_{ij}^y is the connection weight from the neuron j in the Input Layer to the neuron i of the module y in the Map Layer, and x_j is the jth element in the input feature vector. $\mu(f)$ is the weighting coefficient, and it is given by

$$\mu(f) = \begin{cases} \dfrac{1}{N'^{(f)}} & (N'^{(f)} > 0) \\ 0 & (N'^{(f)} = 0) \end{cases} \quad (10)$$

where $N'^{(f)}$ is the number of neurons which correspond to the feature f and is not equal -1.

$\phi(w_{ij}^y, x_j)$ is given by

$$\phi(w_{ij}^y, x_j) = \begin{cases} k_w & \left(x_j \in C_{11} \text{ and } \left(\left(w_{ij}^y = 0 \text{ and } x_j = 1 \text{ and } \sum_{k: x_k \in C_{11}^{g_j}} w_{ik}^y > 0\right) \text{ or } \left(w_{ij}^y = 1 \text{ and } x_j = 0 \text{ and } \sum_{k: x_k \in C_{11}^{g_j}} x_k > 0\right)\right)\right) \\ 0 & (x_j = -1) \\ w_{ij}^y - x_j & \text{(otherwise)} \end{cases} \quad (11)$$

where C_{11} is the neuron set corresponding to keywords, $C_{11}^{g_j}$ is the neuron set corresponding to the keyword which belongs to the same group to j, and k_w $(0 < k_w < 1)$ is the constant.

In each module, the neuron whose internal state calculated by Eq.(8) is maximum is selected as the winner neuron $c^{(y)}$.

Step 3 : Calculation of Outputs of Neurons in Map Layer

The output of the neuron i of the module y in the Map Layer at the time t, $x_i^{MAP(y)}(t)$ is calculated by

$$x_i^{MAP(y)}(t) = \begin{cases} 1, & (i = c^{(y)}, u_i^y(t) > \theta_{s1} \text{ and } D_{max}^{y(i)} < \theta_{s2}) \\ 0, & \text{(otherwise)} \end{cases} \quad (12)$$

where $c^{(y)}$ is the winner neuron in the module y, and θ_{s1} and θ_{s2} are the thresholds. $D_{max}^{y(i)}$ is the maximum distance for the image feature in the neuron i of the module y, and is given by

$$D_{max}^{y(i)} = \max_f \left(\mu(f) \sum_{j=1}^{N} (\phi(w_{ij}^y, x_j))^2\right) \quad (13)$$

In the proposed system, each stored image corresponds to a neuron in the Map Layer. So, the images corresponding to the fired neurons in the Map Layer are output.

Step 4 : Repeat

Steps 2 and **3** are repeated. If the Search Request 5 is selected and there is no similar image corresponding to the key image or the user can not get enough similar images, go to **Step 5**.

Step 5 : Retrieval Images which include Similar Artifacts in Different Position

The cases when the artifacts are shifted to the left or right, and up and down position b ($\in \{\pm 0.5, \pm 1.0, \pm 1.5, \pm 2.0\}$) are considered. The image retrieval which include similar artifacts in different position is realized by the method proposed in ref.[9].

4. Computer Experiment Results

Table 1 shows the precision, the recall and F-measure of the proposed system and the conventional system[10] which stores 550 images. From these results, we confirmed that images can be searched with the proposed system more correctly than the conventional system.

5. Conclusions

In this paper, we have proposed the similarity-based image retrieval considering shape and texture by self-organizing map with refractoriness. Most of the conventional image retrieval systems are used color information mainly, and feature on shape and texture are not fully considered. We carried out a series of computer experiments and confirmed that images can be searched with the proposed system more correctly than the conventional system.

Table 1: Precision, Recall and F-measure.

(a) Search Request 1

	Precision	Recall	F-measure
Conventional	0.970264	0.819721	0.888662
Proposed	0.992346	0.823518	0.900084

(b) Search Request 2

	Precision	Recall	F-measure
Conventional	0.875231	0.801817	0.836917
Proposed	0.907935	0.839824	0.872552

(c) Search Request 3

	Precision	Recall	F-measure
Conventional	0.865801	0.769191	0.814642
Proposed	0.883741	0.774816	0.825702

(d) Search Request 4

	Precision	Recall	F-measure
Conventional	0.989465	0.935322	0.961632
Proposed	0.994253	0.955291	0.974383

(e) Search Request 5

	Precision	Recall	F-measure
Conventional	0.952857	0.823193	0.883292
Proposed	0.967825	0.835218	0.896645

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 IEEE International Joint Conference on Neural Networks, Vancouver, 2006.

[3] K. Nagashima, M. Nakada and Y. Osana: "Similarity-based image retrieval by self-organizing map with refractoriness," Proceedings of IEEE International Joint Conference on Neural Networks, Orlando, 2007.

[4] H. Tanimizu and Y. Osana: "Similarity-based image retrieval from plural key images by self-organizing map with refractoriness," Proceedings of IEEE and INNS International Joint Conference on Neural Networks, Hong Kong, 2008.

[5] S. Hayashida and Y. Osana : "Similarity-based image retrieval considering artifacts by self-organizing map with refractoriness – Image segmentation by K-means algorithm –," Proceedings of IEEE International Conference on System, Man and Cybernetics, Anchorage, 2011.

[6] M. Kobayashi and Y. Osana : "Similarity-based image retrieval considering color of artifacts by self-organizing map with refractoriness," Proceedings of International Symposium on Nonlinear Theory and its Applications, Kobe, 2011.

[7] 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.

[8] 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.

[9] 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.

[10] T. Yamane, R. Ohno and Y. Osana : "Similarity-based image retrieval considering size and color of artifacts by self-organizing map with refractoriness," Proceedings of IEEE International Conference on System, Man and Cybernetics, Manchester, 2013.

[11] 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.

[12] K. Kushima, H. Akama, S. Kon'ya, H. Kimoto and M. Yamamuro, "ExSight: An Object Based High Performance Image Retrieval System," Information Processing Society of Japan, Vol.40, No.2, pp. 732–741, 1999.

[13] D. G. Lowe : "Object recognition from local scale-invariant features," Proceeding of IEEE International Conference on Computer Vision, pp.1150–1157, 1999.

[14] N. Dalal, and B. Triggs : "Histograms of oriented gradient for human deteciton," Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, pp.886–893, 2005.

[15] T. Ojala, M. Pietikäinen and D. Harwood : "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition, Vol.29, pp.51–59, 1996.

[16] G. Csurka, C.R.Dance, L.Fan and C.Bray : "Visual categorization with bags of Keypoints," Proceedings of European Conference on Computer Vision, pp.1–22,2004.