# Content-Based Image Retrieval System By Multi-Dimensional Feature Vectors

*Shinya Asami and Hiroaki Kotera*
*Graduate School of Science and Technology, Chiba University*
*Chiba, Japan*

## Abstract

This paper introduces our approach to ***content-based image*** retrieval by means of key image in place of keywords. The color, spatial frequency, and shape features are extracted from the image sources and spanned in multi-dimensional feature vector space. The mutual color distances between segmented key color areas are compared to narrow the candidates for key image with desired color-tone. The spatial frequency features are represented by the limited DCT components, where low-to-middle frequency spectra are selected in circular zonal sectors. Here the mutual correlations were taken for only 7 DCT components to discriminate the differences in image textures. In addition, the compositional or shape features are simply extracted by down sampling and labeling process. The correlation between the mosaic bi-level patterns after down sampling was very useful to find the rough similarity in image structure. The paper presents experiments on the stamp or facial image retrieval.

## Introduction

Recent years, the opportunities to manipulate digital images have increased enormously by the rapid growth of imaging technologies, data telecommunication, and World Wide Web especially. So it has required system that desired images can be efficiently retrieved from image database.

In contrast to the text-based image retrieval so far, content-based image retrieval (CBIR) is the system that retrieve similar images from database by comparing features (e.g. color, shape, or structure, etc) automatically extracted from the images themselves. A typical current CBIR system allows users to formulate queries by submitting an example of key image to be retrieved. After that, as retrieval results, the system displays thumbnails of similar images on the screen. This paper presents a basic CBIR approach with multi-dimensional image feature extraction in simplified low-dimensional vector space. Fig.1 illustrates the overview of CBIR using multi-dimensional color image features.

## Image Feature Vectors

Multi-dimensional image features have been extracted to compare a key image with database images. We utilize the following three major image features.

[1] Mutual distance in CIELAB space between image ***Key Colors***
[2] 7-dimensional vector in low-to-middle Fourier ***Spectra***
[3] Down sampled rough image ***Structure*** and/or specified ***Shape*** for main object in attention.

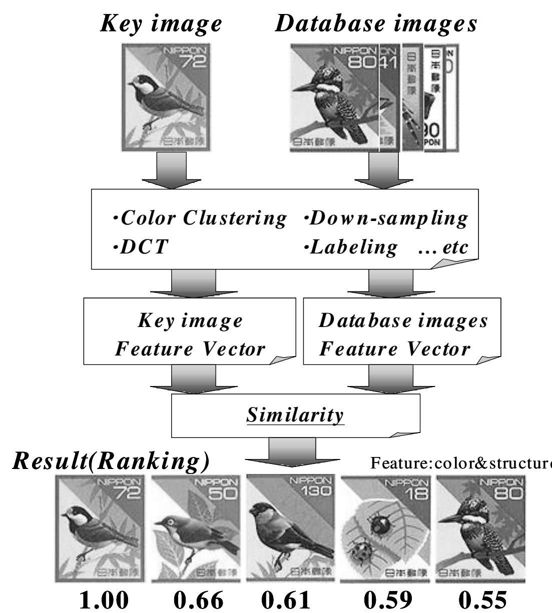


*Figure 1. Content-based image retrieval system*

## Color Clustering for Color Feature Extraction

We introduced well-known ***k-means*** clustering algorithm to extract the color feature from images in CIELAB space and modified it for quick clustering. The detailed process is given as follows:

[1] To start the clustering, the initial seed points are placed at the regular lattice points divided into M x M meshes in the target image. Here, totally 36 (M=6) initial seeds are uniformly distributed as the temporal class centers.

[2] All pixels are classified into the classes nearest to seed points and their locations are renewed by recalculating the gravity centers to be distributed in uniform.

[3] If the mutual color distance between the nearest two classes is less than a constant threshold, those two classes are merged, and a new class is generated. Then, return back to [2].

[4] If the color distance is larger than the constant threshold, the clustering is stopped.

[5] Finally, K most significant classes with higher population are selected and their gravity center color vectors are defined as the color feature for the target image.

Step [3] and [4] reduces the iteration number in conventio-nal ***k-means*** method and accelerate the clustering process.

***Color similarity*** between key and test images, is simply measured by the summation of Euclidian distances between the K pairs of nearest gravity center color vectors.

Figure 2 shows an example of classified color areas in case of K=3 and the selected three pairs of classes in stamp image.

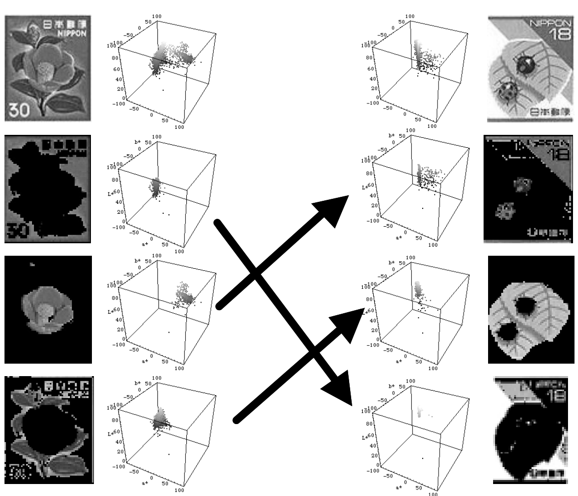

*Figure 2. Similarity measurement between key color areas*

## Spectral Feature Extraction

Fourier spectra of images carry the spatial frequency characteristics, which reflect the texture features. In order to extract image texture features, here we utilized DCT (Discrete Cosine Transform), because DCT is easy to calculate and gives real values without imaginary part. The detailed process to measure the similarity in texture features is described as follows.

In the color image, the spatial frequency information is mainly carried by luminance component. Here RGB image is transformed luminance image by the following equation.

$$Y=0.299R+0.587G+0.114B \tag{1}$$

The luminance image $Y=g(x, y)$ is transformed into spatial frequency components by forward DCT as

$$G(u,v)=\frac{2C_U C_V}{\sqrt{N_1 \times N_2}}\sum_{x=0}^{N_1-1}\sum_{y=0}^{N_2-1} g(x,y) cos\frac{(2x+1)u\pi}{2N_1} cos\frac{(2y+1)v\pi}{2N_2} \tag{2}$$

where, $N_1$ and $N_2$ denote $x$ and $y$ sizes of the image

Here, 2-D DCT coefficients are transformed into 1-D spectral vector in fan-shaped radial frequency domain.

First, the effective DCT spectra are extracted in the limited range of radial frequency, $f_L \le f \le f_H$.

Second, the circular doughnut region is divided into $P$ sectors by tangential angle $\theta = \tan^{-1}(u/v)$ (see Fig.3). Finally, the DCT coefficients within the each sector are averaged. Thus 2-D DCT spectra is simply transformed to 1-D vector $\boldsymbol{G}_P$ as follows.

$$\begin{gathered}
\boldsymbol{G}_P = \left[\ G_p\ \right];\ \ p = 1,2,\cdots,P \\
G_p = \sum_u \sum_v G(u,v) \\
for\ \ f_L \le f \le f_H\ ,\ ;\ \ f = \sqrt{u^2+v^2} \\
\frac{2K}{\pi} tan^{-1}(v/u) \le p \le 1 + \frac{2K}{\pi} tan^{-1}(v/u)
\end{gathered} \tag{3}$$

The similarity of spatial frequency distribution between two images **g** and **h** is calculated by the following cross correlation coefficient.

$$C = \frac{\sum_{i=1}^{P}(G_i - \overline{G})(H_i - \overline{H})}{\sqrt{\sum_{i=1}^{P}(G_i - \overline{G})^2}\sqrt{\sum_{i=1}^{P}(H_i - \overline{H})^2}} \tag{4}$$

where, $P$ denotes the dimension of feature vector G or H.

Figure 3 illustrates the transform of 2-D DCT spectra to 1-D $P$=7 dimensional vector.

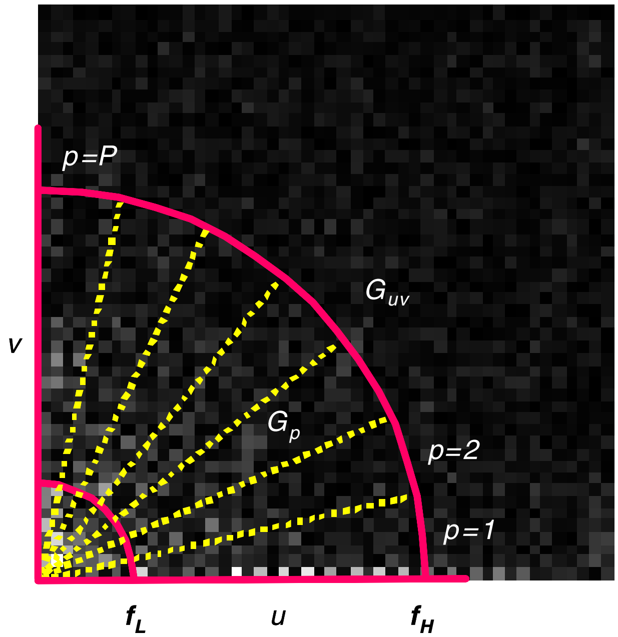


*Figure 3. P=7 DCT spectra in low-to-middle frequency domain*

## Structural Feature Extraction

The structural sketch is often useful to retrieve the images intuitively. The simplest way to catch the rough image structure is to use down sampling. The color image with M*N pixels is resized to 5*5 small mosaic bi-level pattern and its 1 or 0 values are used as structural feature.

The structural similarity procedure is measured as follows

[1] A color image is transformed into the luminance image by Equation (1).
[2] The Luma image is transformed into bi-level image.
[3] Down-sampling is applied for the bi-level image.
[4] The down-sampled image is transformed the 5*5 bi-level image, again.

The similarity between the mosaic patterns $G_i$ and $H_i$ for image **g** and **h** is given by the cross correlation coefficient in Equation (3). Figure 4 illustrates the structural similarity measurement.

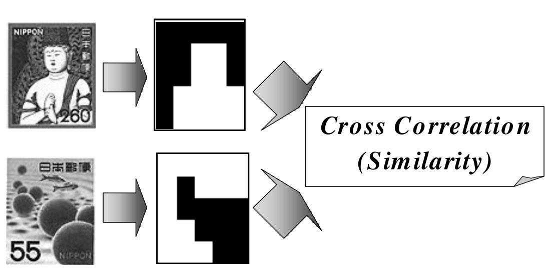


*Figure 4. Similarity measurement by down-sampling*

## Experimental Results

We conducted the image retrieval experiments on the test image databases in Fig. 5 and Fig. 7. Figure 6 shows a retrieved example for “Peking Opera Mask” in Fig. 5. Both similarities of ***Color*** and ***Spectrum*** are combined by half and half. In this sample, the DCT spectra are extracted in the range of $f_L$**=10** ~ $f_H$**=20**. Figure 8 shows another result for “Stamp” image in Fig. 7, where the similarities of ***Color*** and ***Structure*** are also mixed by half and half. These examples resulted in the better retrieval in the combination of two different features than in their single use.

*Figure 5. Test image sets (Peking Opera Mask)*

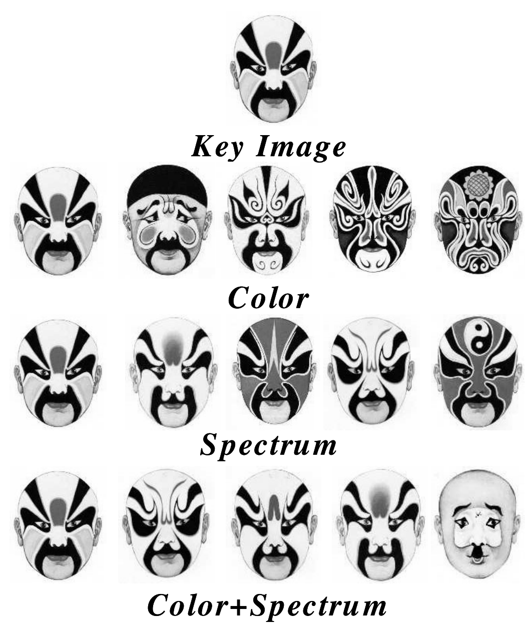


*Figure 6. Retrieval result (Peking Opera Mask)*

*Figure 7. Test image sets (Stamp)*

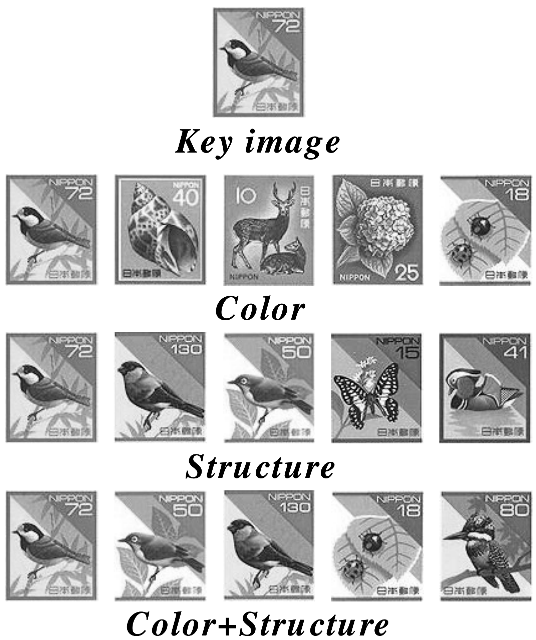


*Figure 8. Retrieval result (Stamp)*

## Shape Feature Extraction

A mosaic pattern will reflect the rough structural sketch of image but not the geometrical shape feature exactly. In order to extract the shape feature of a main object in attention, we tried to apply labeling process to it. For example, typical shape feature "Circularity" is measured as follows.

[1] A specified class containing a main object is selected by the color clustering procedure as stated above.

[2] A main object is separated by the labeling process. Here we defined the connected pixels occupying the maximum area as a main key object.

The "Circularity" is calculated by

$$e = \frac{4\pi S}{L^2} \quad \text{S: area size, L: around length} \qquad (5)$$

The difference in circularity between key and database images is defined as "circular" similarity as shown in Fig. 9.

Figure 10 shows an experimental result retrieved for "Fruit" images. The top five won the higher similarity scores are listed. The most left image in the list is used as a key. The numerical value in each image shows the total similarities put together the ***Color***, ***Shape***, and ***Circularity*** features.

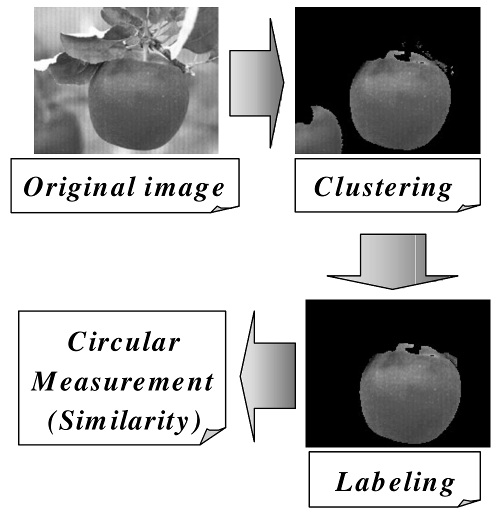


*Figure 9. Circularity measurement to extract the shape feature*

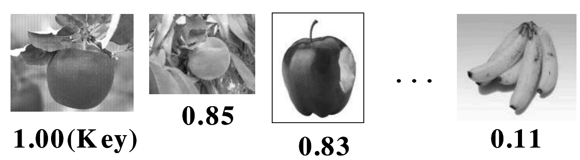


*Figure 10. Retrieval result (Fruit)*

## Conclusion

A content-based image retrieval system using multi-dimensional feature vectors is proposed. In all the test images, color feature was indispensable and useful to narrow the candidates at the first stage, and the combination with spatial or structural feature is expected to retrieve the reliable target intuitively. However, as shown in the experimental samples, which feature of ***Spectrum*** or ***Shape*** should be combined with ***Color*** will depend on the image contents. The structural feature by mosaic pattern worked well for all kinds of image but spectral feature by DCT wasn't always available when used in single. Shape feature may be applied for retrieving the specified objects with simple geometrical structure.

Future works are going to be continued to find the optimum dimension number for each feature vector, the better selection or combination of multiple feature vectors, and to test a variety of huge image databases.

## References


1. Qasim Iqbal and J. K. Aggarwal: "Perceptual Grouping for Image Retrieval and Classification" Third IEEE Computer Society Workshop (2001)
2. Dengsheng Zhang, Aylwin Wong, Maria Indrawan, Guojun Lu "Content-based Image Retrieval Using Gabor Texture Features"
3. S. Asami and H. Kotera, "Hierarchical Image Retrieval System Using Multi-dimension Characteristic Vectors", Proc. Color Forum Japan '02, (2002)


## Biography

**Shinya Asami** graduated from Department of Image Science, Faculty of Engineering, Chiba University, Japan in 2002. He is a student at Graduate school of Science and Technology, Chiba University. His work has primarily focused on image retrieval, including image segmentation, image recognition, and color image processing.