# Image Retrieval Using Sequential-Clustering with Color-Spatial Information

Jungdong Jang* and Taihong Yi**

*Department of Information Communication, Korea Electric Power Corporation
E-mail: jdjang@kepco.co.kr
**Department of Electronics Engineering, Yeungnam University
E-Mail: taehong@ynucc.yeungnam.ac.kr

## Abstract

Recent advances in information technology have made an explosive increase in the number of images. In order to manage and retrieve these images efficiently, content-based image retrieval is appeared to be the core technology. In this paper we proposed to use Sequential Clustering method considering color and spatial information of the features in the image. We used color mean, variation and size of each divided region for feature vector, which have 18 features and that need smallest storage than conventional methods. Our experimental results showed that retrieval efficiency not only improved 8.8 percents better than conventional methods but also saved much computation and data storage. In addition, our method showed excellent Precision vs. Recall evaluation and better subjective judgment by visual perception.

**Content Areas** : vision, multimedia

**Tracking Number** : 182

“This paper has not already been accepted by and is not currently under review for a journal or another conference, nor will it be submitted for such during IJCAI’s review period.”

# Image Retrieval Using Sequential-Clustering with Color-Spatial Information

**Content Areas : vision, multimedia**
Tracking Number : 182


## Abstract

Recent advances in information technology have made an explosive increase in the number of images. In order to manage and retrieve these images efficiently, content-based image retrieval is appeared to be the core technology. In this paper we proposed to use Sequential Clustering method considering color and spatial information of the features in the image. We used color mean, variation and size of each divided region for feature vector, which has 18 features and that needs smallest storage than conventional methods. Our experimental results showed that retrieval efficiency not only improved 8.8 percents better than conventional methods but also saved much computation and data storage. In addition, our method showed excellent Precision vs. Recall evaluation and better subjective judgment by visual perception.


## 1 Introduction

Consider that everyday we view various forms of visual media, such as photographs, graphics, animations and videos. These media are provided increasingly in digital form. For example, the WorldWide Web (WWW) is one such source for viewing hundreds of gigabytes of digital visual information. While there is great accessibility to large stores of digital imagery, new systems need to be developed to manage and search better for the visual information. To cope with this problem, there have been two major approaches to study image retrieval from somewhat different angles.

The one is to annotate each image manually with text. This approach has however insuperable problems that creating annotations by hand can become hopeless because databases can be very large. Manual indexing is that the content of an image is a highly subjective concept, that is, different people may perceive it differently.

Another approach is to index images directly based on the visual image content. Content-based image retrieval was proposed in the early 90's and it has been under intensive research from then on. In general, visual feature extraction is fundamental bases in the content-based image retrieval paradigm. Visual features are usually divided into four classes depending on what kinds of properties they describe. The classes are color, texture, shape, and spatial relationship.

Automatic feature extraction would be desired in large-scale image retrieval system. Besides, the manual annotation of text is likely to be error prone with incompatible or inconsistent text description. In addition, some of the image databases are not confined to specific domains. In these situations, content-based retrieval approaches are preferred.

However, the use of color information has been used widely in content-based retrieval systems. It is clear that color plays a very important role in defining an image. But color alone is insufficient to characterize an image. In order to overcome these problems, we discuss a technique of considering color-spatial information with spatial knowledge to obtain an overall impression of the image.

In this paper, content-based image retrieval considering color-spatial information is proposed. Our goal is to improve the above-mentioned shortcomings through following process. In the first, Sequential-Clustering schemes, which are quite straightforward and fast method, are introduced to cluster the homogeneous property of any images. Then, clustered regions are used as the basis for obtaining relevant color-spatial information with mean-colors, variations and region sizes. Finally, the feature vector is used to retrieve relevant images from image database.

The rest of the paper is organized as follow. In the next section, we briefly review various color-spatial retrieval techniques. In section 3, concepts related with Sequential Clustering are introduced and discuss our color-spatial features. Section 4 describes similarity as well as performance evaluation. In section 5, to demonstrate efficiency and effectiveness of the proposed approach, experiments on a total of 418images are described. Finally, we discuss the conclusions and suggestions for future work in section 6.

## 2 Color-Spatial information

Color histograms are commonly used in content-based retrieval systems and have proven to be very useful; however, when an image is transformed into a histogram, all the spatial information is discarded. Indexing using color histograms has significant limitations of this lack of spatial information [Pass *et al*., 1996; Stricker and Dimai, 1996].

To overcome this problem, a histogram refinement method using the color coherence vectors (CCV) was presented [Pass *et al*., 1996], which partitions pixels based on spatial coherence. Another method was presented to incorporate color and spatial information, called color correlograms, which express how the spatial correlation of pairs of color changes with distance [Huang *et al*., 1997].

On the other hand, [Stricker and Dimai, 1996] presented a method of dividing an image into “five fuzzy regions”. The first three-color moments of each region are used to form a feature vectors for an image.

[Hsu *et al*., 1995] attempts to capture the spatial arrangement of the different colors in the image. The image is partitioned into rectangular regions using maximum entropy, where each region is predominantly a single color. The similarity between two images is the degree of overlap between regions of the same color. [Ma and Manjunath, 1997] performs retrieval based on segmented image regions. Their segmentation is not fully automatic, as it requires some parameter tuning and hand pruning of regions.

Our approach to segmentation uses the Sequential-Clustering algorithm [Theodoridos and Koutroumbas, 1999] to cluster images and to compute mean, variation, and size of each region. Figure 1 presents schematic of our image retrieval technique.

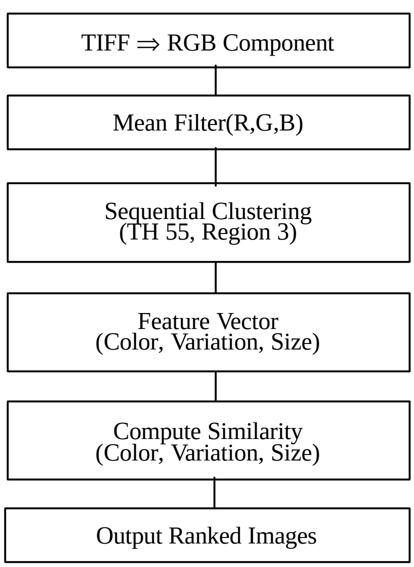



Figure 1. Schematic of our system

## 3 Sequential-Clustering and Feature Extraction

A content-based image retrieval that is solely based on the color is found to be insufficient. In this section, we try to find the clustered regions of the homogenous property in the image and use them as basis for similarity computation.

Let $d(f,C)$ denote the distance between a feature vector $f$ and a cluster $C$. When $C$ is represented by a single vector, $d(f,C)$ is given by

$$d(f,C) = d(f,\overline{C}) \tag{1}$$

where $\overline{C}$ is the representative of $C$. In our case the mean vector is used as a representative. The updating may take place in an iterative fashion. That is obtained as

$$\overline{C}_k^{new} = \frac{(nC_k^{new} - 1)\overline{C}_k^{old} + f}{nC_k^{new}} \tag{2}$$

where $nC_k^{new}$ is the updated component of $C_k$ after each addition of $f$ and $\overline{C}_k^{new}(\overline{C}_k^{old})$ is the representative of $C_k$ after(before) each iteration. (See Figure 2)

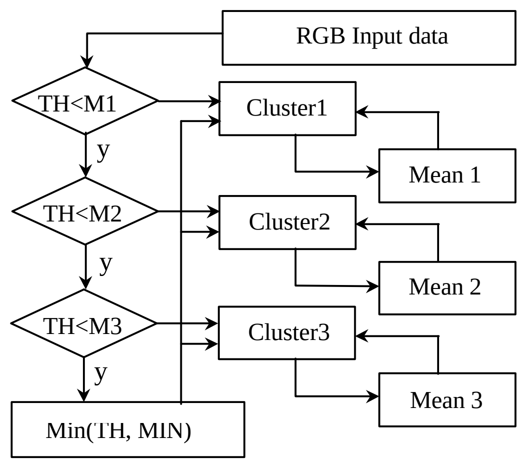



Figure 2. Sequential-Clustering Algorithm

Minimum distance and color-mean value are defined as

$$\overline{R}_k = \frac{\sum r_k}{n_k}, \overline{G}_k = \frac{\sum g_k}{n_k}, \overline{B}_k = \frac{\sum b_k}{n_k} \tag{3}$$

$$R_k = \left|r - \overline{R}_k\right|, G_k = \left|g - \overline{G}_k\right|, B_k = \left|b - \overline{B}_k\right| \tag{4}$$

$$D = \sqrt{{R_k}^2 + {G_k}^2 + {B_k}^2} \tag{5}$$

$$Min(TH,D) = \begin{cases} C_k, & TH \le D \\ C_{k+1}, & TH > D \end{cases} \quad (6)$$

where $k$ is the number of clustered regions, $n_k$ is the number of pixels, $\overline{R}_k, \overline{G}_k, \overline{B}_k (r_k, g_k, b_k)$ are mean value(present value of input pixel), $D$ is Euclidean distance, and $TH$ is a threshold value for clustering images.

The original and the clustered images by Sequential-Clustering method are shown in Figure 3.

After clustering of the image, we compute feature vectors as

$$RD_k = \frac{\sum r_k}{255 \cdot n_k}, GR_k = \frac{\sum g_k}{255 \cdot n_k}, BL_k = \frac{\sum b_k}{255 \cdot n_k} \quad (7)$$

$$XVR_k = \frac{\sqrt{\sum (x_i - \overline{X})}}{N_k}, YVR_k = \frac{\sqrt{\sum (x_i - \overline{Y})}}{N_k} \quad (8)$$

$$SZ_K = N_K \quad (9)$$

where $N$ is the number of pixels, $RD$,$GR$, and $BL$ are mean value of the primary colors of each regions. $\overline{X}$ and $\overline{Y}$ are center of the $X$, $Y$ coordinate for regions, $XVR$, $YVR$, and $SZ$ are the variation of $X$, $Y$ and the size of regions, respectively.

(a) (b) (c) (d) (e)

Figure 3. Sample images by Sequential-Clustering
(a) Original (b) Composite Clusters (c) Cluster1
(d) Cluster2 (e) Cluster3

## 4 Similarity Measure and Performance Evaluation

The distances between each region in a query and target images are given by

$$dC(Q,T) = \sum_{k=1}^{3} \sqrt{Rc_k^2 + Gc_k^2 + Bc_k^2} \quad (10)$$

$$dV(Q,T) = \sum_{k=1}^{3} \sqrt{Xv_k^2 + Yv_k^2} \quad (11)$$

$$dS(Q,T) = \sum_{k=1}^{3} \left| \frac{SZ_{Q,k} - SZ_{T,k}}{SZ_{Q,k} - SZ_{T,k}} \right| \quad (12)$$

$$S_D = \frac{w_1 \cdot dC + w_2 \cdot dV + w_3 \cdot dS}{w_1 + w_2 + w_3} \quad (13)$$

where $Rc_k, Gc_k, Bc_k, Xv_k, Yv_k$, and $SZ$ are the distance for color-mean, variation, and size for each color channel. $dC$, $dV$, $dS$, and $S_D$ are the distance of color-mean, variation, size, and overall distance of each region. $w_1, w_2$, and $w_3$ are feature weights and $Q$ is a query image and $T$ is a target image.

To evaluate performance, we use *Precision vs. Recall* [Smith, 1998] and *Efficiency* metrics. Formally, these metrics are defined as follows:

$$Recall = \frac{Retrieved\ and\ relevant}{All\ relevant\ images} = \frac{n}{N} \quad (14)$$

$$Precision = \frac{Retrieved\ and\ relevant}{Retrieved\ images} = \frac{n}{R} \quad (15)$$

$$Efficiency = \begin{cases} \frac{n}{N}, & if \quad N \le R \\ \frac{n}{R}, & otherwise. \end{cases} \quad (16)$$

We note that if $N \le R$, then *Efficiency* becomes the *Recall*, while if $N > R$, this becomes the *Precision*.

## 5 Experimental Results

We implemented the proposed approach on Pentium III (650Mhz) computer with the Windows98 OS and designed using Microsoft Visual C++ 6.0 compiler. The database consists of 418 TIFF images of size 192× 128 and experimented 53 query images by the example image. For the Sequential-Clustering, the number of clustering region is

3, threshold is 55, and weights $w_1, w_2, w_3$ are 4, 2.5, 2, respectively.

In order to evaluate proposed system, we used two methods. One is objective evaluation by equations (14), (15), and (16). The other is subjective evaluation by visual inspection. Also, in order to compare results and performance, Histogram intersection [Swain and Ballard, 1991], CCV, and Correlogram were fully implemented. Image database is summarized as Table 1. Table 2 compares the methods between other approaches.

Table 1. Classification of image database

| Items | Relevant Images | Query Images |
|---|---|---|
| Bird | 34 | 2 |
| Butterfly | 50 | 9 |
| Landscape | 68 | 17 |
| Native life | 38 | 4 |
| Tree | 48 | 3 |
| Flower | 56 | 6 |
| Animal | 83 | 5 |
| Sea, Fall | 16 | 4 |
| Fish | 12 | 1 |
| The others | 12 | 1 |
| Total | 418 | 53 |

Table 2. Methods used in evaluation

| Rank / Method | Feature | Cluster | Storage |
|---|---|---|---|
| **Swain** | 2048 | No | 3,4MByte |
| **CCV** | 128 | " | 226KByte |
| **Correlogram** | 256 | " | 547KByte |
| **Propose** | 18 | 3 | 52KByte |

For output lists of various Rank sizes and *Efficiency* results by equation (16), we show the corresponding comparison in Table 3 for overall 53 queries. From the table 3, we see that the proposed approach generally performs better than other methods.

Table 3. Performance comparison by *Efficiency*

| Rank / Method | 5th | 10th | 15th | 20th |
|---|---|---|---|---|
| **Swain** | 81.7 | 77.4 | 80.6 | 91.5 |
| **CCV** | 63.6 | 58.4 | 62.0 | 70.0 |
| **Correlogram** | 70.9 | 65.6 | 68.8 | 77.6 |
| **Proposed** | 74.6 | 68.7 | 73.1 | 82.7 |

Figure 4 gives Precision vs. Recall to evaluate performance for query images of "Starfish" (y0096), and "flower"(y0373) image. From the figure 4, our approach generally performs better than other methods.

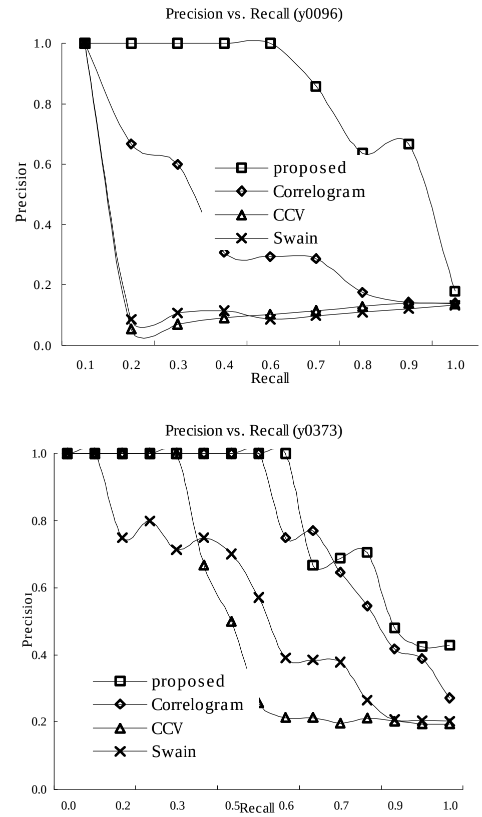
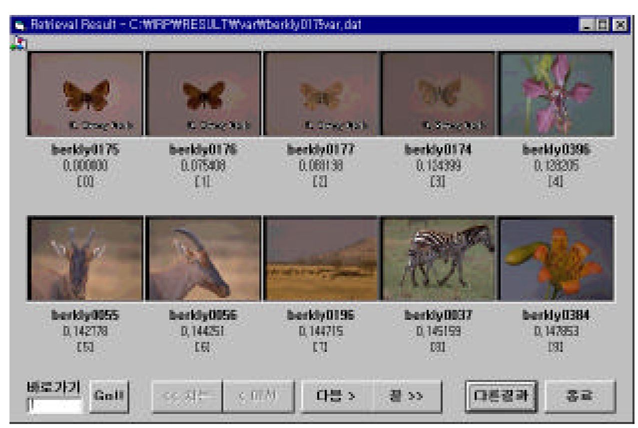
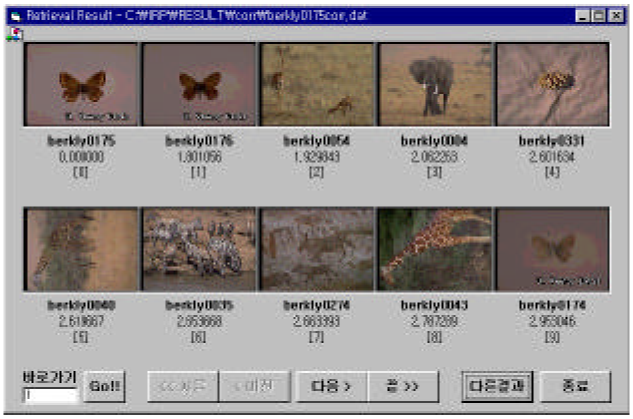



Figure 4. Precision vs. Recall

Figure 5 gives retrieval results to evaluate subjective performance by visual judgment for query image of "Butterfly". From the figure 5, the correct images have been ranked ranging from the 1st to the 4th ranks in our system, but, in the other hand, 1st, 2nd, 10th, and 20th in Correlogram, 1st, 22nd, 41st, and 60th in CCV, and 1st, 11th, 71st, and 221st in Swain.



(a)




(b)


(c)


(d)

Figure 5. Retrieved results for "Butterfly" image
(a) Proposed (b) Correlogram (c) CCV (d) Swain

## 6 Conclusions

We have proposed a feature extraction approach considering color-spatial information using Sequential-Clustering technique.

Features such as color-mean, variation, and size have been extracted in the three clustered regions. Our proposed method uses only eighteen features, which is much smaller than Swain, CCV, and Correlogram methods that are used to compare with our retrieval performance. However, our experimental results showed that retrieval efficiency not only improved 8.8 percents better than conventional methods but also saved much computation and data storage. In addition, our method showed excellent Precision vs. Recall evaluation and better subjective judgment of visual perception.

In the future, we will study a geometric approach to represent the features extracted from the clustered image regions, and an influence on retrieval performance in different color space.

## References


[Huang *et al*., 1997] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image Indexing Using Color Correlograms," *IEEE Computer Vision and Pattern Recognition Conferenc*e, San Juan, Puerto Rico, June, 1997.

[Hsu *et al*., 1995] W. Hsu, T. S. Chua, and H. K. Pung, " An Integrated Color-Spatial Approach to Content-based Image Retrieval," *Proceedings of ACM Multimedia*, pp. 305-313, Sanfrancisco, November, 1995.

[Ma and Manjunath, 1997] W. Ma and B. Manjunath, "NeTra: A toolbox for navigating large image databases," *Proc. IEEE Int. Conf. on Image Proc.*, pages 568–571, 1997.

[Pass *et al.*, 1996] Greg Pass, Ramin Zabih, and Justin Miller, "Comparing Images Using Color Coherence Vectors," *ACM Conference on Multimedi*a. Boston, Massachusetts, November 1996.

[Smith, 1998] J. R. Smith, "Image Retrieval Evaluation," *Proceedings of IEEE Workshop on Content-Based Access of Image and Video Librarie*s, pp. 112-113, Santa Barbara, California, 21 June 1998.

[Stricker and Dimai, 1996] M. A. Stricker and A. Dimai, "Color indexing with weak spatial constraints," *Storage and Retrieval for Still Image and Video Databases*, volume 2670, pages 29-40, 1996.

[Swain and Ballard, 1991] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Visio*n, 7(1). pp. 11-31, 1991.

[Theodoridos and Koutroumbas, 1999] Sergios Theodoridos and Kostantinos Koutroumbas, *Pattern Recognition*, Athens: Academic Press, 1999.