Content-based Image Retrieval of Color Image Considering Color and Spatial Information

Jungdong Jang, Taihong Yi

Department of Electronics Engineering, Yeungnam University
E-Mail: jdjang@kepco.co.kr

Abstract

Recent in information technology has made an explosive increase in the number of images. In order to manage and retrieve these images efficiently, content-based image retrieval of color image is appeared to be the core technology. Color histograms are widely used for content-based retrieval, considering only color have many disadvantages. In this paper we consider color and spatial information of the features in the image and we proposed to use Sequential Clustering method for homogeneous property regions. Also we proposed to use color mean, variation and size of each divided regions for feature vector, which have 18 features and its storage space is smaller than conventional methods. 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 subjective fine judgement of human perception.

Keywords: content-based retrieval, spatial retrieval, image clustering

1 Introduction

Consider that everyday we view various forms of visual media, such as television broadcasts, photographs, graphics, animations and videos. These media are provided and stored increasingly in digital form. For example, the World Wide 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 better manage, index and search for the visual information. To cope with this problem, there have been two major approaches that study image retrieval from somewhat different angles.

The traditional approach is to annotate each image manually with text whish describes the content of the image; some image attributes such as number and date can also be included in the annotation. 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. Besides, in some cases it is very difficult to describe the relevant content of an image by just a couple of words.

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. Content-based image retrieval means image annotation and retrieval based on their visual and semantic content instead of text labels and attributes. In general, visual feature extraction is fundamental bases in the content-based image retrieval paradigm. Feature extraction means the process of determining the relevant content of images. 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, even though image processing and pattern recognition have not advanced enough to make this reality. 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 for image retrieval has been used widely in many content-based retrieval systems with some success. It is clear that color plays a very important role in defining an image. Because it seems to be similar two images those color composition are similar, it is also equally clear that color alone is not sufficient to characterize an image. Histogram-based color retrieval techniques super from a lack of important spatial knowledge. 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, a content-based image retrieval of color image considering color-spatial information is proposed. Our goal is to improve the above mentioned shortcomings through next process. In the first, a Sequential-Clustering scheme, which is quite straightforward and fast method and depending on the distance metric used, is invited to cluster the homogeneous property of any images. Then, clustered regions are used as the basis for obtaining relevant color-spatial information through mean-colors, variations, region-sizes. Finally, the feature vector obtained in above process is used to retrieve relevant images from an 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, some Sequential Clustering related concepts in computational distance are introduced and we discuss our color-spatial features. Section 4 describes similarity function as well as retrieval evaluation. In section 5, we describe various experiments on a total of 418 images to demonstrate efficiency and effectiveness of the proposed approach. Finally, we discuss the conclusions and suggestions for future work in section 6.

2 Color-Spatial information

Color is a very important cue in extracting information from images. 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 spatial information is discarded. Indexing using color histograms have significant limitations of this lack of spatial information [1][2].

To overcome this problem, a histogram refinement method using the color coherence vectors (CCV) was presented [1], which partitions pixels based on spatial coherence. They claimed that two non-relevant images may have similar color histograms; however, their CCVs are different. Furthermore, another method was presented to incorporate color and spatial information, called color correlograms, which expressed how the spatial correlation of pairs of colors changes with distance [3].

On the other hand, Stricker and Dimai[2] presented a method of dividing an image into "five fuzzy regions", and claimed the regions are natural for an encoding of minimal spatial information. The first three-color moments of each region are used to form a feature vectors for an image.

Hsu et al. [4] 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 is 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 [5] perform retrieval based on segmented image regions. Their segmentation is not fully automatic, as it requires some parameter tuning and hand pruning of regions.

Classical object recognition techniques usually rely on clean segmentation of the object from the rest of the image or are designed for fixed geometric objects such as machine parts. Neither constraint holds in the case of natural images: the shape, sizes, and color of objects like tigers and airplanes are quite variable, and segmentation is imperfect. Clearly, classical object recognition does not apply. More recent techniques can identify specific objects drawn from a finite (on the order of 100) collection, but no present technique is effective at the general image analysis task, which requires both image segmentation and image classification.

Our approach to segmentation uses the Sequential-Clustering algorithm [6] to cluster images to homogeneous property regions and to compute mean, variation, and size of each region. Figure 1 presents schematic of our image retrieval technique.

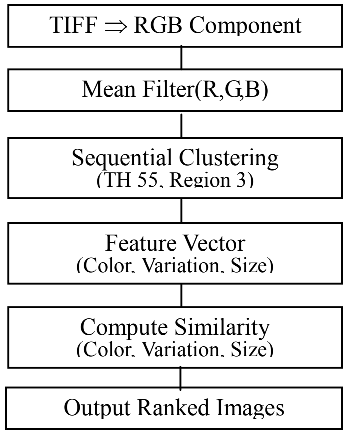


Figure 1. Schematic of our image retrieval technique

3 Sequential-Clustering and Feature Extraction

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

3.1 Sequential Clustering

A Sequential Clustering is quite straightforward and fast method. Let $d(f, C)$ denote the distance between a feature vector *f* and a cluster *C*. This may be defined by taking into account either all vectors of *C* or a representative vector of it. The user-defined parameters required by the algorithmic scheme are the threshold of dissimilarity and the maximum allowable number of clusters. When *C* is represented by a single vector, $d(f, C)$ becomes

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

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

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

where nC_k^{new} is the component of C_k after the assignment of f to it and $\overline{C}_k^{new}(\overline{C}_k^{old})$ is the representative of C_k after(before) the assignment of *f* to it. (see Figure 2)

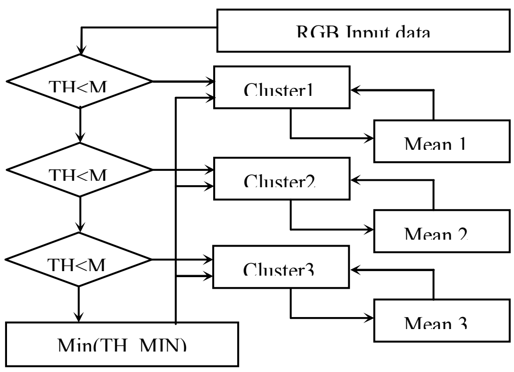


Figure 2. Sequential-Clustering Algorithm

We used two phase of the algorithm scheme. The first phase involves the determination of the clusters, via the assignment of some of the vector f to them. During the second phase, the unsigned vectors are presented for a second time to the algorithm and are assigned to the appropriate cluster. In this paper, each image pixel has a three-dimensional color descriptor in the RGB color space. We define minimum distance and color-mean as a representative below.

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

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

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

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

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

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

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

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

3.2 Feature Extraction

After clustering of the image, we compute the color-mean, variation, and relative size of clustered 3 regions as follows.

$$RD_k = \frac{\sum_{i=0}^{N_k - 1} r_k}{255 \cdot n_k}, \quad GR_k = \frac{\sum_{i=0}^{N_k - 1} g_k}{255 \cdot n_k}, \quad BL_k = \frac{\sum_{i=0}^{N_k - 1} b_k}{255 \cdot n_k} \tag{7}$$

$$XVR_k = \frac{\sqrt{\sum_{i=0}^{N_k-1}(x_i - \overline{X})}}{N_k}, \quad YVR_k = \frac{\sqrt{\sum_{i=0}^{N_k-1}(x_i - \overline{Y})}}{N_k} \tag{8}$$

$$SZ_k = N_k \tag{9}$$

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

4 Similarity Measure and Performance Evaluation

4.1 Similarity Measure

At query time, to determine the similarity of two images we measure the distance between of the given image and indices stored in the database. We use the *L1* and L2 distance measures when comparing three feature vectors. The distances between each region in a query and target images are given by

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

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

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

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

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

4.2 Performance Evaluation

To evaluate performance, we use *Precision vs. Recall* [7]and *Efficiency* metrics. Formally, these metrics are defined as follows:

$$Recall = \frac{Retrieved\ and\ relevant}{All\ relevant\ \ images} = \frac{n}{N} \tag{14}$$

$$Precision = \frac{Retrieved\ and\ relevant}{Retrieved\ images} = \frac{n}{R} \tag{15}$$

$$Efficiency = \begin{cases} \dfrac{n}{N}, & if \quad N \le R \\ \dfrac{n}{R}, & otherwise \end{cases} \tag{16}$$

where, the *relevant images* are predefined by the ground truth database. We note that if $N \le R$, then *Efficiency* becomes the *Recall* value, while if $N > R$, and then *Efficiency* becomes the *Precision* value. Because *Efficiency* is simpler than *Precision vs. Recall*, we use *Efficiency* for overall performance evaluation and use *Precision vs. Recall* to evaluate performance about each query image.

5 Experimental Results

We implemented the proposed approach on Pentium III (650 MHz) computer with the Windows 98 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 weight between feature distance is 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 confirmation. Also, in order to compare results and performance, the Swain's method [8], CCV, and Correlogram were fully implemented. Table 1 and Table 2 are summarized for image database and method between other approaches.

Table 1. Classification of image database

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

Items	**Relevant**	**Query**
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

Table 2. Comparison of Methods to evaluate

Method	**Space**	**Vector**	**Distance**	**Clustering**	**Storage**(byte)
Swain	Opponent RGB	2048	Histogram Intersection	No	3,456,000
CCV	RGB	128	L1	No	226,560
Correlogram	RGB	256	L1	No	547,584
Proposed	RGB	18	L2	3	52,672

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, we see that the proposed approach generally performs better than other methods.

Table 3. Performance comparison by *Efficiency*

Method	**5th Rank**	**10th Rank**	**15th Rank**	**20th Rank**
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 5 gives Precision vs. Recall to evaluate performance for query images by "Starfish"(y0096), "Landscape"(y0185), "Landscape"(y0273), and "flower"(y0373) image. From the figure, like *Efficiency* results, our approach generally performs better than other methods.

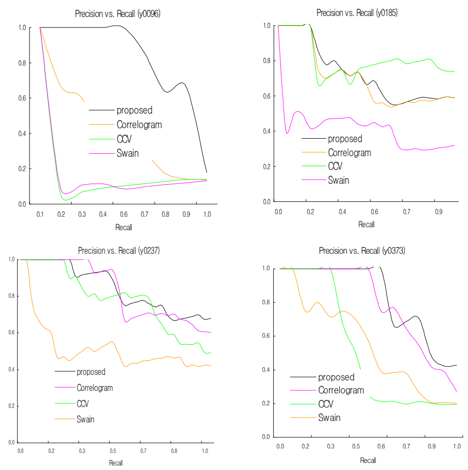
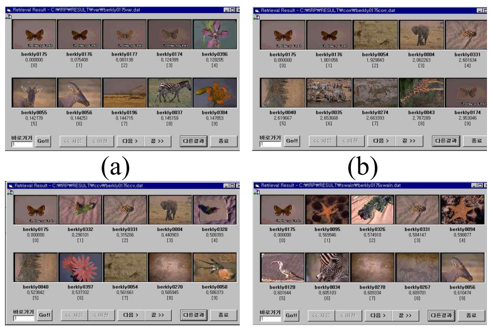


Figure 5. Precision vs. Recall



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

Figure 6. Retrieval results for "Butterfly" image
(a) Proposed (b) Correlogram (c) CCV (d) Swain

Figure 6 gives retrieval results to evaluate subjective performance by human eye judgement for query images by "Butterfly". From the figure, the correct images have been ranked in from the 1st to the 4th orders in our system. But 2nd, 10th, and 20th in Correlogram, 22nd, 41st, and 60th in CCV, and 11th, 71st, and 221st in Swain.

6 Conclusions

In this paper, we have studied content-based image retrieval considering color-spatial information, which is a very important thing of content-based retrieval area. We have proposed a feature extraction approach considering color-spatial information using Sequential-Clustering technique that was used to pattern recognition and have extracted features as color-mean, variation, and size in the three clustered regions. Our proposed method uses only eighteen features, which is much smaller than Histogram Intersection, CCV, and Correlogram method 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 subjective fine judgement of human perception.

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

References

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

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

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

[4]. 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.

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

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

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

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