IMAGE RETRIEVAL USING MULTI-SCALE COLOR CLUSTERING

Se-Hwan Kim, Woontack Woo and Yo-Sung Ho

Kwangju Institute of Science and Technology (K-JIST)
1 Oryong-dong, Puk-gu, Kwangju, Korea
E-mail: shkim@gogh.kjist.ac.kr, {wwoo, hoyo}@kjist.ac.kr

ABSTRACT

A fundamental issue in content-based image retrieval is how to select image features that can represent image contents appropriately. In this paper, a multi-scale color clustering algorithm based on human perceptual properties of color images is proposed for image retrieval. The multi-scale clustering algorithm is an unsupervised clustering method that utilizes the perceptual uniformity property in the (p,q) color space. The proposed color clustering algorithm produces a small set of representative color vectors for each image that capture color properties of the image, and a set of correlogram values that contain the spatial information of the image.

1. INTRODUCTION

Recent explosion of the internet usage and wide availability of multimedia services in the world wide web (WWW) have remarkably increased the need for image retrieval systems that can index a large amount of images effectively and retrieve them based on their visual contents efficiently [1]. When we design an image retrieval system based on the visual content, critical issues are what to define as image features and how to use them for image indexing and retrieval. They largely affect remaining aspects of the system design, and greatly determine capabilities of the image retrieval system.

The color histogram is popularly used in many image retrieval systems due to its good performance in characterizing the global color content. It is easy to compute with only few constraints when applied to images. However, it has several inherent problems in image indexing and retrieval. Most of all, it does not include any spatial information. For example, it cannot distinguish a single large yellow region from a large number of scattered yellow pixels. Furthermore, it is big in size; which makes it rather difficult to create an effective database indexing scheme.

Several approaches have been proposed to overcome the problems associated with the traditional histogram method. Pass and Zabih proposed the color coherence vectors (CCV) to incorporate spatial information into color histogram representation [2]. Huang, et al., proposed the color correlogram to consider the local color spatial correlation as well as the global distribution of this spatial correlation [3]. However, these methods did not consider how to obtain the segmented images.

In this paper, we propose a multi-scale color clustering algorithm based on human color perception for an image retrieval system. By presenting the multi-scale clustering in an appropriately chosen color space, we can preserve the perceptual uniformity and decide the number of classes for clustering depending on the image. Furthermore, we can prevent clusters from being oversegmented due to fine textures while preserving the spatial information by the proposed algorithm. Finally, a new similarity measure, which incorporates color and spatial information, is represented.

2. COLOR SPACE AND TRANFORMATION

Among several low-level features, color information has been extensively studied due to its invariance under image scaling and orientation. Color features used in image retrieval include global and local color histograms, their mean and higher order moments. However, there are common issues underlying all color-based retrieval methods; the selection of a proper color space, the use of a proper color quantization scheme, and the development of efficient feature representation.

Perceptual uniformity is one of the most important factors to be satisfied when choosing a proper color space. Perceptual uniformity means that two color pairs that are equal in distance in a color space are perceived as equal in distance by viewers [4]. This has two implicit meanings: the pixel distribution is uniform, and the color changing is smooth in terms of human perception.

0-7803-6725-1/01/$10.00 ©2001 IEEE

It is known that colors within MacAdam ellipses are visually indistinguishable [5]. Any color lying on the perimeter of a MacAdam ellipse is just noticeably different (JND), as compared to the center of that ellipse. Unfortunately, the size and shape of these ellipses vary considerably in the color models used for image processing, including the HSV and YIQ models. In other words, the same distances in different parts of the color space denote different amounts of perceived color shifts. Consequently, it is not appropriate to perform image clustering in the above color spaces. The so-called uniform chromaticity scale (UCS) model is the best linear transformation model that has been devised for perceptual uniformity. However, this model still does not provide equal distances throughout its color space.

In this paper, we use a nonlinear transformation in geodesic chromaticity color space which has been shown to provide almost equally perceived color shifts throughout the space. The nonlinear transformations are defined by [5]:

$$p = 3751a^2 - 10a^4 - 520c^2 + 13295c^3 + 32327ac$$
$$-25492a^2c - 41672ac^2 + 10a^3c - 5227a^{1/2} + 2952a^{1/4}$$
$$q = 404d - 185d^2 + 52d^3 + 69b(1-d^2) - 3b^2d + 30bd^3$$

where

$$a = \frac{10x}{2.4x + 34y + 1}, \quad b = \frac{10x}{4.2y - x + 1}$$
$$c = \frac{10y}{2.4x + 34y + 1}, \quad d = \frac{10y}{4.2y - x + 1}$$
$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z}$$

and X, Y and Z are the tristimulus CIE color coordinates, respectively, derived from RGB tristimulus values via the following transformation matrix:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.490 & 0.310 & 0.200 \\ 0.177 & 0.813 & 0.010 \\ 0.000 & 0.010 & 0.990 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

3. COLOR SEGMENTATION

The chromaticity diagram is defined as a two-dimensional representation of an image where each pixel produces a pair of values (p,q). A three-dimensional distribution is defined as a histogram associated with the chromaticity diagram. We can observe that most values are concentrated on some regions, producing very few high peaks in the (p,q) space. Figure 1 shows the flow diagram for the feature vector extraction procedure.

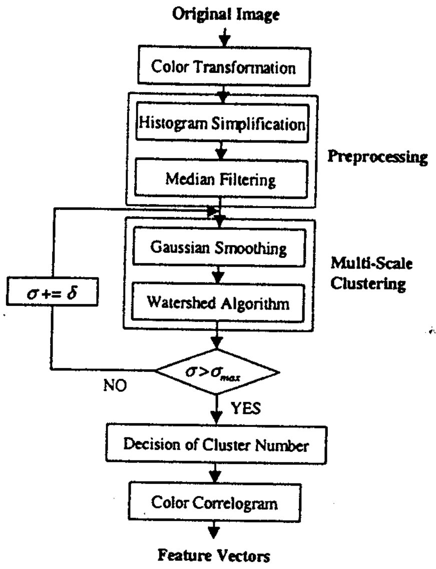


Figure 1. Feature Vector Extraction

3.1. Preprocessing

After the color transformation into the (p,q) color space, we may have undesired histogram patterns which produce false positive results. In order to eliminate those patterns, we perform histogram simplification and median filtering. Histogram simplification removes unnecessary information from the input for easier segmentation. That is, we can prevent clusters from being oversegmented due to fine textures while preserving the spatial information by accomplishing this procedure. Histogram simplification is accomplished by morphological filters, typically opening and closing by reconstruction.

Opening by reconstruction of erosion: $r^{(rec)}(\varepsilon_n(f), f)$

Closing by reconstruction of dilation: $\varphi^{(rec)}(\delta_n(f), f)$

For opening by reconstruction of erosion, simplification operation is performed by the erosion which eliminates all components that are smaller than the structuring element. Then, the reconstruction process restores the contour of components that have not been totally removed by the erosion. The histogram obtained by reconstruction is a good starting point for segmentation: it is much simpler than the original histogram and also peaks are defined precisely. The remaining isolated peaks in the histogram are eliminated using the median filter.

3.2. Multi-Scale Clustering

Consider a set of n color points, with each point having k pixels after preprocessing. A scale-space representation

[6] of these points can be realized by convolving them with a Gaussian kernel $\phi_\sigma(x)$ with a scale size σ,

$$\phi_\sigma(x) = \frac{1}{\sigma^r\sqrt{(2\pi)^r}} \exp\left(-\frac{1}{2}\sum_{j=1}^{r}\left(\frac{x_j}{\sigma}\right)^2\right)$$

After the Gaussian kernel is applied to the histogram, the histogram image is segmented into clusters by the watershed algorithm in the (p,q) space. In the efficient immersion simulation, we flood the surface from its local minimum. When water coming from two different minima merges, an imaginary dam is built to prevent any mixing of water. Watershed lines partition the space by associating a region called catchment basin to each local minimum [7].

After the watershed algorithm is applied, the number of clusters is governed by the scale size σ in the (p,q) color space. The larger the scale size, the lower the resolution and consequently the fewer the number c of clusters. In the c vs. σ plot, long plateau-like segments denote the range over which the number of clusters remain the same or survive for relatively long periods of time. Based on the lifetime measure, a histogram or density function can be derived to reflect the frequencies with different numbers of clusters. This histogram is then used to identify the prominent numbers of clusters. Non-prominent numbers of clusters are separated by arranging the lifetimes in the descending order and finding the largest relative drop of lifetime.

4. SPATIAL INFORMATION

We include spatial information by adopting the color correlogram algorithm [3], where the color correlogram describes how the spatial correlation of pairs of colors changes with distance.

If a distance $d \in \{1,2,\ldots,m\}$ be fixed a priori, the correlogram is defined for $i, j \in \{1,2,\ldots,m\}$, $k \in \{1,2,\ldots,d\}$ by

$$\gamma_{c_i,c_j}^{(k)}(I) \triangleq \Pr_{p_1 \in I_{c_i}, p_2 \in I}\left[p_2 \in I_{c_j} \,\middle|\, |p_1 - p_2| = k\right]$$

Given any pixel of color c_i in the image I, the color correlogram gives the probability that a pixel at distance k from the given pixel has color c_j. In our paper, we use, for simplicity, the autocorrelation that captures spatial correlation between identical colors.

$$\alpha_c^{(k)}(I) = \gamma_{c,c}^{(k)}(I)$$

5. DATABASE INDEXING

The multi-scale clustering produces a small set of clusters for each input image. This cluster set collectively represents color properties of the original image. Another effect of the color clustering is that the image is segmented into a set of contiguous regions. These regions certainly capture the information of the object contained in the image.

To exploit both color cluster and spatial information of each image, we construct the following indexes: for each color cluster, the average p, q values and the pixel count normalized by the image size and for spatial information, the autocorrelation values with d different distances. Therefore, each image will be indexed by $c\times(3+d)$ feature vectors, that is, representative vectors in the database.

6. SIMILARITY MEASURE

During the query process, we calculate the distance to each representative vector for each color of multiple-colored query image. We obtain the minimum value of these distances.

$$D_C(q_m) = \min(\delta(q_m, i_1), \ldots, \delta(q_m, i_n), \ldots, \delta(q_m, i_c))$$

where q_m is the m^{th} color of query image and i_n is the n^{th} indexed representative color of each database image. $\delta(q_m, i_n)$ represents the Euclidean distance between them. If the index for $D_C(q_m)$ is found, the corresponding $D_S(q_m)$ is calculated for the same index.

$$D_S(q_m) = \sum_{k\in[d]} \frac{\left|\alpha_c^{(k)}(q_m) - \alpha_c^{(k)}(i_n)\right|}{1 + \alpha_c^{(k)}(q_m) + \alpha_c^{(k)}(i_n)}$$

$$D(q_m) = \exp(-\alpha D_C(q_m))\exp(-\beta D_S(q_m))D_H(q_m)$$

where $D_H(q_m)$ is the number of minimum pixels between q_m and i_n, and α, β are scale factors. The similarity measure, $D(Q,I)$, between the query image Q and one of the database images I is given as follows. The maximum $D(Q,I)$ for all the database images is chosen as the best-matched image.

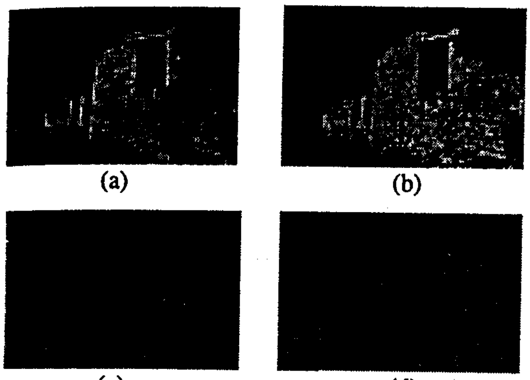
$$D(Q,I) = \sum_m D(q_m)$$

7. EXPERIMENTAL RESULTS

A statistical analysis of our image database has revealed that the average number of colors extracted was 7.62, and the maximum and the minimum number of extracted colors were 15 and 1, respectively. Quality of the reconstructed images is analyzed to evaluate performance of the three approaches for one sample of database images.

Figure 2 displays the original and its quantized images by different quantization schemes. Figure 2(b) shows the reconstructed image derived by quantizating the original color space from 256 levels to 2 levels in each axis. Figure

2(c) and Figure 2(d) demonstrate the reconstructed images by performing color space quantization with the agglomerative hierarchical clustering approach [4] and with the proposed clustering approach, respectively. From Figure 2, we can observe that objects with similar colors are well clustered in the proposed clustering approach even though quantization levels of each approach are 8, 8 and 6, respectively.



(a) Original Image (b) Uniform Quantization
(c) Agglomerative Hierarchical Clustering
(d) Proposed Clustering
Figure 2. Perceptual Comparison

We have performed a variety of queries using a set of 1000 images. Figure 3 shows an example of the retrieval results by the proposed color-based image retrieval system. The image at left top corner shows the original image. As we proceed to the right and bottom direction, images differ from the original image. Figure 4 shows the performance for several images. The retrieval effectiveness was compared by measuring retrieval recall and precision scores. Recall signifies the proportion of relevant images in the database that are retrieved in response to a query. Precision is the proportion of the retrieved images that are relevant to the query.

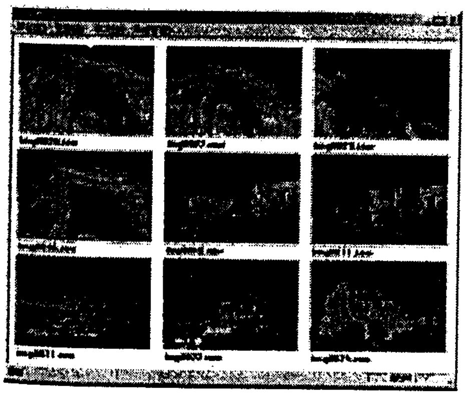


Figure 3. Query Results of the Proposed Approach

8. CONCLUSIONS

In this paper, we propose a new segmentation procedure, multi-scale color clustering algorithm based on human color perception. By using the proposed segmentation algorithm, we can prevent clusters from being oversegmented due to fine textures while preserving the spatial information. The spatial information of the clustered image is also maintained by applying color correlogram method. According to the experimental results, we can see that the proposed approach shows improved results over the histogram matching (HM) even though the averages of feature vectors are 45.72 and 64, respectively.

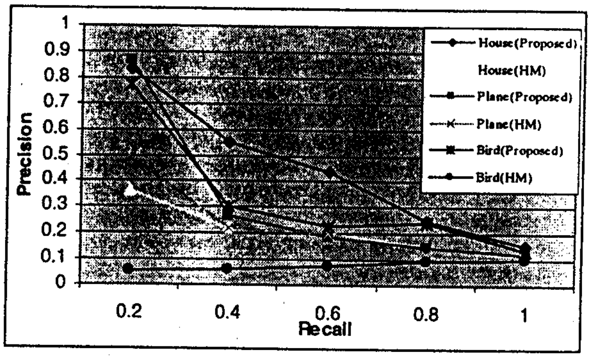


Figure 4. Precision vs. Recall

ACKNOWLEDGMENT

This work was supported in part by KOSEF through UFON, and in part by MOE through BK21.

REFERENCES

[1] Y. Gong, G. Proietti, and C. Faloutsos, "Image Indexing and Retrieval based on Human Perceptual Color Clustering," *Proc. of CVPR'98*, pp. 578-583, 1998.

[2] G. Pass and R. Zabih, "Histogram Refinement for Content-based Image Retrieval," *Proc. of the Third IEEE Workshop on Applications of Computer Vision*, 1996.

[3] J. Huang, S. Kumar, M. Mitra and W. Zhu and R. Zabih, "Image Indexing using Color Correlograms," *Proc. of CVPR'97*, pp. 762-768, 1997.

[4] J. Wang, W. J. Yang and R. Acharya, "Color Clustering Techniques for Color-Content-based Image Retrieval from Image Databases," *Proc. of Multimedia Computing and Systems '97*, pp. 442-449, 1997.

[5] D. MacAdam, *Color Measurement*, Springer-Verlag, 1981.

[6] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, 1994.

[7] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Trans. PAMI*, vol. 13, no. 6, pp. 583-598, June 1991.