IMAGE SEGMENTATION USING MULTI-SCALE COLOR CLUSTERING FOR IMAGE RETRIEVAL SYSTEMS

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

Kwangju Institute of Science and Technology (K-JIST)
1 Oryong-dong, Puk-gu, Kwangju, 500-712, Korea
Tel: +82-62-970-3157, Fax: +82-62-970-2204
E-mail: shkim@gogh.kjist.ac.kr, {wwoo, hoyo}@kjist.ac.kr

ABSTRACT

For content-based image retrieval systems, we need to select image features that can represent image contents appropriately. Since semantic description, rather than low-level feature description, is more meaningful for human perception, we can employ image segmentation for feature extraction. In this paper, a multi-scale color clustering algorithm based on human perceptual properties for color is proposed for image segmentation. The multi-scale clustering algorithm is an unsupervised clustering method that utilizes the perceptual uniformity property in the (p,q) color space. The proposed 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 to apply to the image retrieval system.

1. INTRODUCTION

Content-based image retrieval (CBIR) has been an intensive research area in recent years because of outburst of accessibility and increase in storage and communication volumes [1]. For content-based image retrieval systems, we need to index a large amount of images effectively and retrieve them based on their visual contents efficiently. 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.

In our daily life, we are accustomed to exploiting high-level concepts, such as objects, people, building, etc. Even though a human observer can easily understand these concepts, the computer system has a significant challenge in performing content-based image indexing and retrieval in an automatic fashion [2]. Object-based analysis, which performs semantically meaningful object segmentation on images, is an essential step to reduce the gap between low-level and semantic-level features.

The color histogram intersection method was developed by Swain and Ballard [3] and popularly used in many image retrieval systems due to its good performance in characterizing the global color content. However, it does not include spatial information. Various approaches have been proposed to overcome problems associated with the traditional histogram method. Pass and Zabih proposed the color coherence vectors (CCV) to incorporate the spatial information into the color histogram representation [4]. Huang, et al. proposed the color correlogram to consider the local color spatial correlation as well as the global distribution of the spatial correlation [5]. However, these methods do not consider how to obtain segmented images. Most systems just use a simple quantization method in a certain color space, such as RGB, HSV, HVC, YUV and L*a*b*. However, as most images do not possess a uniform color distribution, a typical histogram usually contains many empty or nearly empty bins. Even though the perceptual characteristic was considered in Blobworld system [6], they just used the pre-defined color ranges to segment the color image.

In this paper, we propose a multi-scale color clustering algorithm based on human color perception for image segmentation. 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 presented for image retrieval systems.

2. COLOR SEGMENTATOIN

2.1. Color Space Transformation

The characteristics of color have been extensively studied because of its invariance under image scaling and orientation even though there are several low-level features. However, there are common issues underlying all color-based retrieval methods. An appropriate color space and color quantization scheme should be used and an efficient feature representation method has to be developed.

Perceptual uniformity is one of the most important factors to be addressed when choosing a proper color space for image retrieval. Perceptual uniformity means that two color pairs that are equal in distance in a color space are perceived as equal in distance by viewers [7]. This has two implicit meanings: the pixel distribution is uniform, and the color changing is smooth in terms of human perception.

It is known that colors within MacAdam ellipses are visually indistinguishable [8]. Any color lying on the perimeter of a MacAdam ellipse is just noticeably different (JND), as compared to the center of that ellipse. However, the size and shape of these ellipses vary considerably depending on the color models. That is, the same distances in different parts of the color space denote different amounts of perceived color shifts. Even in the HSV and YIQ models, we cannot guarantee the perceptual uniformity. The so-called uniform chromaticity scale (UCS) model is the best linear transformation model that was devised for perceptual uniformity. However, it is still not proper to perform image clustering.

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 [8]:

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

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:

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

2.2. Preprocessing

We can define the chromaticity diagram as a two-dimensional representation of an image where each pixel produces a pair of values (p,q). A histogram associated with the chromaticity diagram can be represented in a three dimensional space. In this case, most values are gathered on some regions, thus we obtain very few high peaks in the (p,q) space. Figure 1 shows the flow diagram for the feature vector extraction procedure.

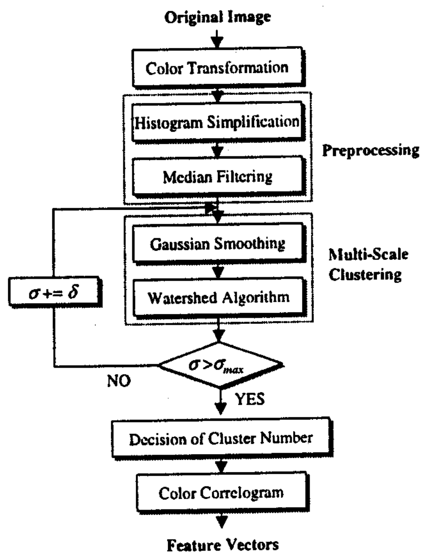


Figure 1. Feature Vector Extraction

After the color transformation, we may need preprocessing steps because unproper histogram patterns, which produce false positive results, appear. To remove these patterns, we perform histogram simplification and median filtering. Histogram simplification eliminates 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.

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.

2.3. Multi-Scale Clustering

After preprocessing steps, we have a set of n color points, with each point having k pixels in the (p,q) color space. A scale-space representation [9] of these points can be realized by convolving them with a Gaussian kernel $\phi_\sigma(\mathrm{x})$ with a scale size σ,

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

The histogram image is segmented into clusters by the watershed algorithm after the Gaussian kernel is applied to the histogram in the (p,q) space. The immersion simulation, an efficient watershed algorithm, consists in flooding the surface from its local minimum. When the water coming from two different minima would merge, an imaginary dam is built to prevent any mixing of water. As can be seen in Figure 2, watershed lines partition the space by associating a region called catchment basin to each local minimum [10].

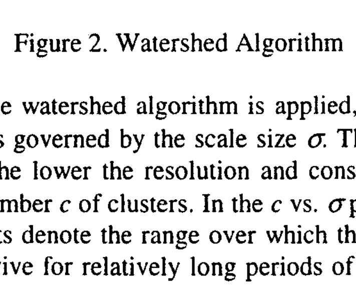


Figure 2. Watershed Algorithm

After the watershed algorithm is applied, the number of clusters is governed by the scale size σ. The larger the scale size, the lower the resolution and consequently the fewer the number c of clusters. In the c vs. σ plot, plateau-like segments denote the range over which the number of clusters survive for relatively long periods of time. Based on the lifetime measure, a histogram 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.

3. IMAGE RETRIEVAL

3.1. Database Indexing

We include spatial information by adopting the color correlogram algorithm [5], where the color correlogram describes how the spatial correlation of pairs of colors changes with distance.

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.

3.2. Similarity Measure

For each color of multiple-colored query image, we calculate the distance to each representative vector and 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.

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

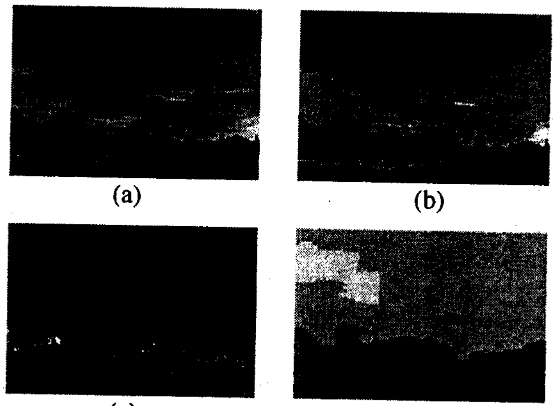
$$D(q_m) = \exp(-\alpha D_C(q_m))\exp(-\beta D_S(q_m))D_H(q_m)$$

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.

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

4. EXPERIMENTAL RESULTS

According to the statistical analysis, the average number of colors extracted was 7.62, and the maximum and the minimum number of extracted colors were 15 and 1, respectively. The original and its quantized images using different quantization schemes are shown in Figure 3. The reconstructed image derived by quantizating the original color space into 2 levels in each R, G and B axis is shown Figure 3(b). Figure 3(c) and Figure 3(d) demonstrate the reconstructed images derived by color space quantization with the agglomerative hierarchical clustering approach [7] and with the proposed clustering approach, respectively. From Figure 3, we can observe that objects with similar colors are well clustered by the proposed clustering approach even though the number of quantization levels of the proposed algorithm is smaller than those of other methods.



(a) Original Image (b) Uniform Quantization
(c) Agglomerative Hierarchical Clustering
(d) Proposed Clustering

Figure 3. Perceptual Comparison

Performance for a few queries from a set of 1000 images is compared in Figure 4, where we measure the effectiveness of image retrieval by recall vs. precision scores. While the horizontal axis (recall) represents the proportion of relevant images in the database that are retrieved in response to a query, the vertical axis (precision) is the proportion of the retrieved images that are relevant to the query. Experimental results demonstrate that the proposed scheme outperforms over the histogram matching (HM) even though the averages of feature vectors are 45.72 and 64, respectively.

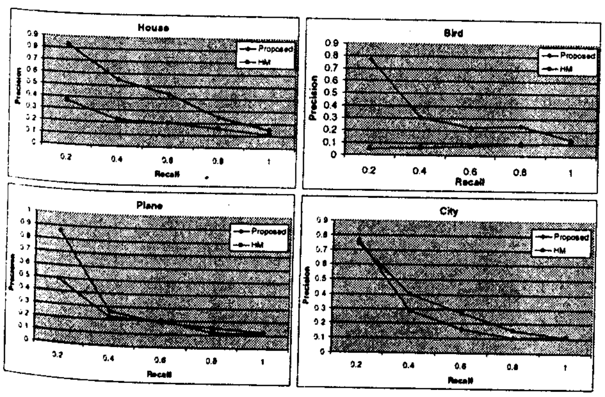


Figure 4. Performance Comparison

5. CONCLUSIONS

In this paper, we propose a new image segmentation scheme using multi-scale color clustering based on human color perception. While preserving the spatial information, we can prevent clusters from being oversegmented due to fine textures. The proposed scheme provides better segmentation results compared to uniform quantization and agglomerative hierarchical clustering methods. For the image retrieval systems, we apply color correlogram method to the segmented image and in such a way we can maintain the spatial information. The proposed approach provides improved performance over the histogram matching for image retrieval.

ACKNOWLEDGMENT

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

REFERENCES

[1] D. Dori, "Cognitive Image Retrieval," *Proc. of Pattern Recognition*, vol. 1, pp. 42-45, 2000.

[2] Y. Xu, P. Duygulu, E. Saber, A. Tekalp and F. Yarman-Vural, "Object-based Image Retrieval based on Multi-level Segmentation," *Proc. of ICASSP*, vol. 4, pp. 2019 –2022, 2000.

[3] M. Swain and D. Ballard, "*Indexing via Color Histograms*," *Proc. of Computer Vision*, pp. 390-393, 1990.

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

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

[6] C. Carson, S. Belongie, H. Greenspan and J. Malik, "Region-based image querying," *Proc. of CVPR*, pp. 42-49, 1997.

[7] 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*, pp. 442-449, 1997.

[8] D. MacAdam, *Color Measurement*, Springer, 1981.

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

[10] 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.