Content-based Image Retrieval for Image Indexing

Md. Al-Amin Bhuiyan
Department of Computer Engineering
King Faisal University
Hofuf, Al-Ahsa 31982, Saudi Arabia

Abstract—Content-based image retrieval has attained a position of overwhelming dominance in computer vision with the advent of digital cameras and explosion of images in the Internet and Clouds. Finding the most relevant images in a short time is a challenging job with many big cloud sites competing in image search in terms of accuracy and recall. This paper addresses an image retrieval system employing color information indexing. The system is organized with the hue components of the HSV color model. To assess the precision of the image retrieval system, experiments have been carried out on a database consisting of 450 images drawn by Japanese traditional painters, namely Sharaku, Hokusai, Hiroshige, and the images obtained from the World Wide Web (WWW) multicolor natural scenes. In order to query the database, the user specifies an object on which the same color attributes are evaluated and all similar looking images are exposed as the outcomes of the query.

Keywords—color indexing; HSV color model; color histogram; Minkowski distance metric; fuzzy clustering; Color Quantization

I. INTRODUCTION

Content-based image retrieval is emerging tremendous interest in image processing, computer graphics, computer vision, pattern recognition, image management system, and so on [1-4]. In distinction to the traditional text-based approach, several benefits have been reported in literature for content-based access to images, such as automatic identification, classification, recognition, and retrieval of large digital libraries with photographic images, satellite images, medical images for remote searching and browsing over the ever increasing World Wide Web (WWW).

A fair amount of developments were carried on over the last couple of decades in image retrieval system due to the enormous interest of establishing multimedia information systems and database systems. The convergence of image processing/computer graphics and database technology yields the basis for the creation of such digital image archives. Moreover, with the proliferation of WWW, a remarkable amount of visual information is ready accessible publicly. As a result, it has become a promising demand for search strategies retrieving pictorial entities from large image documentations [5]. Over the past half century there has emerged an increasing interest in the cultural heritage of Japanese society. In tandem with this, and indeed a logical consequence of this, Japanese traditional painting pictures, known as Ukiyoie pictures [6], have been a growing concern with preserving, disseminating, displaying and effectively exploiting the rich cultural resources embodied in many museum and art gallery collections. Benefits of the technology are numerous and the most important points which are generally given are that the use of digital versions of surrogate representations of works of art can: assist security, provide a central database of information to provide easy retrieval of relevant material, assist in preservation of originals and provide a networkable resource of images which greatly increases availability and access to the system.

Image contents include color, shape and texture. Among these contents of images, color provides an efficient visual clue for image retrieval. Managing image data in this regard entails processing, storage, and retrieval of pictorial representations [5]. Due to its graphical illustration, the color histogram becomes the most frequently used technique for image indexing. It provides a convenient tool for computing the similarity between different images, since it proves to be robust to object translation, scaling, rotation, occlusion, deformation, and so on [7].

A substantial amount of research works have been reported in literature [8-16] on content-based image retrieval (CBIR). Swain and Ballard [17] proposed a color-based object recognition employing color histograms for matching between image regions and query objects. Kiedla and Kender [18] applied Gaussian kernels to smooth the histograms on finding skin in color images. Funt and Finlayson [19] developed a color indexing algorithm for object recognition to take into account the influence of lighting conditions. Ennesser and Medioni [20] proposed a local histogram method to localize objects in images. Chang and Wang [21] developed a texture segmentation algorithm employing color histogram. McKenna, Raja and Gong [22] employed an adaptive Gaussian mixtures to model the color allocations of objects. Androutitos, Plataniotis and Venetsanopoulos [23] established a cosine metric based distance measure for color image indexing and retrieval. Their query method is very flexible and provides single and multiple color queries. Liu and Ozawa [24] proposed the spatial neighborhood adjacency graph (SNAG) which could serve as a basis for detecting object by color contents from the candidate images. Sharma et al. [25] have represented images by global descriptor and developed a CBIR system that used color histogram processing. This system is not yet a commercial success. Because the distribution of RGB values changes proportionally with the illumination, which is suitable to some images but have low precision on others.

This paper addresses a color-histogram based method for indexing and retrieving color images. Different dominant and perceptually relevant colors have been extracted from each image in RGB model and are stored in the respective database files. Images are being identified and classified in HSV space depending on the color contents prevailing in these dominant...
colors, that is, whether a particular color component is significantly present in an image or not. Similarity between different images has been calculated on the basis of Minkowski distance metric. Experimental results demonstrate that the method is capable of indexing, classifying, and retrieval of images with distinct color properties.

The rest of the paper is organized as follows. Section II describes color model. Section III illustrates histogram and image retrieval. Section IV and Section V describes color quantization and image query, respectively. Section VI highlights experimental results and finally, Section VII draws the overall conclusions of this paper.

II. COLOR MODEL

Numerous color models have been justified for color specification, such as CIE (R,G,B), (X,Y,Z), (L*,u*,v*), and so on. The main drawback of the CIE (R,G,B) model is that it is not perceptually uniform and the proximity of colors in RGB space does not indicate color similarity. The (X,Y,Z) color space is not uniform, that is, the Euclidean distance between two colors is not proportional to the color difference perceived by humans. Although the (L*,u*,v*) space is uniform, but nevertheless, is not intimately related to the way in which humans perceive color [21].

A color is represented in HSV color space by the three features: hue, saturation and value. Hue is the characteristic of visual perception that corresponds to color sensation related to the dominant color, saturation indicates the comparative purity of color content and value specifies the brightness of a color. The HSV color model organizes similar colors under similar hue alignments. The transformation from RGB color space to HSV space is given by the equations [23,26-28]:

\[ H = \cos^{-1} \left( \frac{1}{2} \left( \begin{array}{c} \frac{1}{2}((R-G)+(R-B)) \\ \sqrt{(R-G)^2 + (R-B)(G-B)} \end{array} \right) \right), \]

where \([0, 2\pi]\).

\[ S = \max(R,G,B) - \min(R,G,B); \]

\[ V = \frac{\max(R,G,B)}{255}, \]

where \(R,G,B\) are the red, green, and blue component values which exist in the range \([0,255]\).

This research employs HSV color model for classification of pictures drawn by the painters namely Sharaku, Hokusai, Hiroshige and natural pictures. The RGB model has been used to identify the Ukiyoe pictures because these are being distinguished according to their red, green and blue color components in the face parts.

III. COLOR HISTOGRAM AND IMAGE RETRIEVAL

Color histogram [29,30] represents the distribution of colors in an image. A color histogram is a stable object illustration which is unaffected by occlusion and changes in viewing conditions, and that a color histogram has the advantage of being insensitive to scaling, rotation, and small deformation of objects and being immune to noise [31].

The basic idea to image retrieval by color content is to extract the characteristic colors from target images which are matched with those of the query. Different images from the database are then searched to check whether a specific color feature value is prominently existing in an image or not. If a number of images contain the substantial amount of that query color, then these are marked according to the priority basis. The architecture of the proposed system is shown in Fig. 1.

Since the hue component of the HSV color model performs better with human chromatic perception [17], the hue component has been chosen to designate the colors of images. Pixels in the image are characterized in the RGB space, so it appears natural to define the color attributes as the red, green, and blue value at each pixel. Let a color image \(Q(x, y)\) consists of three color channels \(Q(x, y) = (Q_R(x, y), Q_G(x, y), Q_B(x, y))\), or \(Q = (Q_R, Q_G, Q_B)\), at \((x, y)\) of size \(M \times N\). A hue histogram \(H(i)\) of a color image is achieved by counting the number of pixels which have got a hue value \(H(Q_R, Q_G, Q_B) = i\) in the image:

\[ H(i) = \frac{\#(H(Q_R, Q_G, Q_B) = i)}{M \times N} \]

(4)

where \# denotes the number of pixels with a hue value \(H(Q_R, C_G, Q_B) = i\) and \(M \times N\) is the total number of image locations.

A few sample color images and the hue histograms computed from the respective images are illustrated in Fig. 2.

Images are being classified depending on the prominent colors. Color segmentation has been employed to extract the regions of dominant and perceptually relevant colors. Natural pictures are being separated from those of the painting pictures on the basis of the ratio \(r_a\) of the area containing ten dominant colors \(a_{dc}\) to that of the total area containing all colors \(a_{dc}\). The reason behind this choice appears from the fact that the painting pictures contain only a few number of colors (the painters use only a limited number of colors during painting) in comparison to that of the infinite number of colors in nature. So the dominant colors contribute more to the images in comparison to those of the natural pictures. Painting pictures are being identified and classified according to the name of the painters, such as Sharaku, Hiroshige and Hokusai depending on the dominant color components because it has been found from the experimental results that the pictures are being fashioned with different colors according to the color choice of the painters. So ten prominent colors are being extracted from the hue histogram in RGB space for each image and the representative vectors are identified as, \(Q_i = (Q_{i,R}, Q_{i,G}, Q_{i,B})\), \((i = 1,2,...,N)\). The set
$Q_1 \cap Q_2 \cdots \cap Q_N$ of colors belong to the images that resembles the most widely used colors of a given painter.

The similarity between different painters are calculated on the basis of Minkowski-form vector distance metric from their hue histograms. The generalized Minkowski-form distance metric ($L_M$ norm) is given by:

$$d_M(h_q, h_t) = \left( \sum_{i=1}^{N} |h_q^i - h_t^i|^M \right)^{\frac{1}{M}}$$

where $N$ is the dimension of the vectors $h_q$ and $h_t$, and $h_q^i$ is the $i$-th component of $h_q$. This research uses $M = 2$, (which is often used for $L_2$ metric).

Let $h_q$ and $h_t$ be the query and target histograms, respectively, then application of the histogram intersection operator introduced in [17] provides a simple way to match two different images $I_q$ and $I_t$ through their color histograms as [32]:

$$H(I_q, I_t) = \sum_{i=1}^{N} \min d_M(h_q^i, h_t^i).$$

![Fig. 1. Architecture of the image retrieval system](image)

Ukiyoe actors are distinguished from those of the natural and painting pictures on the basis of face colors. Human skin colors cluster in a small region in a color space. Although the color representation of a face obtained by a camera is influenced by many factors such as lighting conditions, facial expressions, etc. and the skin colors cluster and differ from person to person in different races [33,34], Ukiyoe actors are nevertheless drawn by some distinct colors by different painters. The presence of some colors in a specific zone provides information whether the images are of actors’ faces or not. Fig. 3(a) illustrates a face image, and Fig. 3(b) illustrates the skin color distributions in the RGB color model.
IV. COLOR QUANTIZATION

In order to reduce the computational cost in segmentation, an input color image is quantized so that the number of colors contained in the image is reduced while the primary chromatic information about the image still remains the same. In the quantization method [21], the number of quantized colors are first determined, say \( k \), by a histogram thresholding technique. Then fuzzy \( c \)-means classification algorithm is performed to classify each pixel in the image to one of the \( k \) colors [34]. The number of clusters are decided depending on the threshold values.

In fuzzy clustering, each color has a degree of belonging to clusters, rather than belonging completely to just one cluster. Thus, colors on the edge of a cluster, may be in the cluster to a lesser degree than the colors in the center of the cluster. For each color \( C \) we have a coefficient providing the degree of being in the \( j \)-th cluster \( u_j(C) \). Usually, the sum of those coefficients for any given \( C \) is defined to be 1:

\[
\forall C \left( \sum_{j=1}^{k} u_j(C) = 1 \right),
\]

where \( k \) is the number of clusters.
In fuzzy c-means, the centroid of a cluster is the mean of all colors, weighted by their degree of belonging to the cluster. Therefore, the center of the cluster, \( r_j \) will be:

\[
r_j = \frac{\sum C u_j(C)^m C}{\sum C u_j(C)^m}
\]

(8)

where \( r_j \) is the center of \( j \).

The degree of belonging is related to the inverse of the distance to the cluster center:

\[
u_j(C) = \frac{1}{d(r_j, C)},
\]

(9)

Then the coefficients are normalized and fuzzyfied with a real parameter \( m \geq 1 \) so that their sum is 1. Therefore,

\[
u_j(C) = \frac{1}{\sum \left( \frac{d(r_j, C)}{d(r_i, C)} \right)^{2(m-1)}}.
\]

(10)

This investigation uses \( m \) equal to 2, which is equivalent to normalizing the coefficient linearly to make their sum equals 1.

VI. IMAGE QUERY

The query process is to effectively find and retrieve those images from the database that are most similar to the user’s query image. In this case, \( r \)-nearest neighbor query is employed, which retrieves the \( r \) images that are most similar to the query image (which are typically sorted by lowest dissimilarity to the query image). Given a number \( N \) of \( I \) images and a feature dissimilarity function \( f_d \), find the images \( I_T \) \( \in N \) such that \( f_d(I_q : I_T) \leq T_{f_d} \), where \( I_q \) is the query image and \( T_{f_d} \) is the threshold for feature dissimilarity. In this case a query returns any number of images depending on the bounds defined by the threshold of feature dissimilarity \( T_{f_d} \).

VI. EXPERIMENTAL RESULTS

The effectiveness of the proposed method has been justified over some experimental results. The database furnished for this experiment contains a total of 450 images: 80 drawn by Sharaku, 80 by Hokusai, 80 by Hiroshige and 210 natural pictures (sea, flowers, sunrise, sunset, scenery, animals, architectures, towns, etc.) downloaded from the Internet. The snapshot of the CBIR software is shown in Fig. 4. When a user selects the CBIR software, up to 1.0\( \sigma \), where \( \sigma \) is the variance, for the five dominant colors, pictures drawn by Sharaku will get the \( r_a \) value within the range \([19.29,28.67]\), Hokusai within the range \([12.45,20.77]\), and Hiroshige within \([29.90,39.72]\), respectively. Relating Sharaku, there is an outstanding difference from other painters – almost all of his drawn pictures are of Ukiyoe actors.

![Fig. 3. RGB color distribution of a typical Ukiyoe actor’s face](image)

![Fig. 4. Snapshot of the CBIR interface](image)

<table>
<thead>
<tr>
<th>Name of painters</th>
<th>( C_1 )</th>
<th>( C_2 )</th>
<th>( C_3 )</th>
<th>( C_4 )</th>
<th>( C_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharaku</td>
<td>61</td>
<td>91</td>
<td>121</td>
<td>181</td>
<td>27,39</td>
</tr>
<tr>
<td>Hokusai</td>
<td>27,41</td>
<td>241</td>
<td>91</td>
<td>181</td>
<td>61</td>
</tr>
<tr>
<td>Hiroshige</td>
<td>91</td>
<td>181</td>
<td>209,215</td>
<td>241</td>
<td>45,56</td>
</tr>
</tbody>
</table>

The \( r_a \) versus the number of occurrences graph is shown in Fig. 5, which reveals that if the threshold value is taken up to 1.0\( \sigma \), where \( \sigma \) is the variance, for the five dominant colors, pictures drawn by Sharaku will get the \( r_a \) value within the range \([19.29,28.67]\), Hokusai within the range \([12.45,20.77]\), and Hiroshige within \([29.90,39.72]\), respectively. Relating Sharaku, there is an outstanding difference from other painters – almost all of his drawn pictures are of Ukiyoe actors.
The painters used only a few number of colors during drawing pictures. The number of colors used by different painters has been justified for different threshold values of $r_a$ and is shown in Fig. 6, which reveals that the natural pictures have got innumerable number of colors whereas those of the painting pictures are limited. Finally, the Ukiyo e actors are distinguished from the normal pictures according to the RGB color distribution in the face color. Fig. 7 shows the skin color distribution of 80 Ukiyo e actors faces, where two distinct color zones are found. For the larger cluster, mean values for red, green and blue are $m_R =103.65, m_G =104.13, m_B =104.14$, and the variances are $\sigma_R =39.07, \sigma_G =37.84, \sigma_B =40.15$, respectively.

![Graphs showing distribution of colors](image-url)

Fig. 5. $r_a$ versus number of occurrences

![Graph showing $r_a$ versus number of colors](image-url)

Fig. 6. $r_a$ versus number of colors
The image retrieval system has been accessed by two commonly used evaluation measures: (i) precision and (ii) recall, as defined by:

\[
\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of images retrieved}} \quad (11)
\]

\[
\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of images in the collection}} \quad (12)
\]

The precision versus number of images and recall versus number of images are shown graphically in Fig. 8 and Fig. 9, respectively, for different classes of images. Finally, the proposed method has been compared with the existing
influential methods where similar measures have been established. The comparison has been made in terms of precision and is shown graphically in Fig. 10. The graph reveals that the proposed method performs better for less number of images and for higher number of images the performance is more or less the same like existing methods.

VII. CONCLUSION

This paper describes the design and implementation of the content-based image retrieval system. So a quantitative analysis of color distribution has been presented for searching, indexing and retrieving color images. In the query process, the goal of the query is to retrieve images of interest. Prior to the trials, 450 images have been inspected, among them 320 are designated as painting pictures according to the ratio of the dominant colors to those of the total areas from the hue histograms. When the query is issued, the corresponding color index file is analyzed to select a set of candidate images containing regions with the similar colors of the query. The major limitation of the proposed method is that similarity measure for image retrieval has been established on the basis of Minkowski distance metric with $L_2$ norm. Other types of similarity measures like Mahalanobis metric or hausdorff distance can be considered in future for expressing similarity between colors. Our future plan is to develop a multimedia information system that will be able to perform the storage, browsing, indexing, and retrieval of multimedia data based on their text, sound and video contents.

ACKNOWLEDGMENT

The author would like to express his profound gratitude to Late Professor Susumi Matsudaia of Sonoda Womens' University, Japan for letting him use his enthusiastically collected Ukiyoe pictures. He is thankful to Professor Hiromitsu Hama of Osaka City University for his pragmatic suggestions regarding to this research work. The author is also grateful to all of the WWW users whose collected images enhanced his research work as a tool for experimental verification.

REFERENCES

