IMAGE RETRIEVAL SYSTEM BASED ON COLOUR HISTOGRAM AND MEDIAN CUT SEGMENTATION

Jehad Q. Odeh[1] Fatimah Binti Ahmad[2] Mohamad Othman[2] Rozita Johari[2]

1. Multimedia University, Melacca, Malaysia
Faculty of information science and technology
jehad.alnihoud@mmu.edu.my

2. University Putra Malaysia, Serdang
Faculty of Computer Science and Information Technology
fatimah@fsktm.upm.edu.my, mothman@fsktm.upm.edu.my,
rozita@fsktm.upm.edu.my

Abstract

The need for an image retrieval system became a challenging research topic. While, general object recognition still difficult, it is easy and practical to capture and use some image features like color distribution to identify objects with some accuracy.

In our research we used the color histogram as the main feature to extract from both image query and image database. Then, by using Distance function (City block), the similar images will be retrieved. Color histogram calculation is a fully statistical approach, so we perform some testing on a set of images to calculate the mean, standard deviation, energy, and skew. Furthermore, we test two segmentation algorithms (histogram thresholding and median cut).

Our final version of the system will use the SOM (Self-Organizing Map) to cluster the database based on the values extracted from the colour histogram. Furthermore, we study the different colour space available and through out an extensive study and testing we decided to use the CIE model because it is closer to the human perception of the color and it is more suitable to be use in any image retrieval system. Moreover, we did consider the interfacing part with some details that because any image retrieval system should have an easy interface to allow the user to retrieve the relevant images based on his/her query. We applied a new technique to allow the user to compose the image query based on a set of templates available. The interface allows the user to create their own image query, as well as an example image, an approach known as QBE (query by example).

Keywords: colour Histogram, CBIR, thresholding, median cut, QBE, Euclidean, City block.

1. Introduction

Large image databases were created and used in many applications including multimedia encyclopedia, geographical information system and others. The need for an image retrieval system became a challenging research topic.

There are several problems with the traditional approach using textual information. First and the most important drawback of this approach is the difficulty to describe images based on content, that because the same image has different meanings for human. Second depicting the spatial relationship among the objects is not easy task. Last the keywords to describe images are subjective and we don't have a unique and standard description. Therefore, an efficient and automatic procedure is required for retrieving images from databases. So, the image retrieval based on content is more desirable and promising area of research.

General object recognition among image still difficult, the review for the content based image retrieval done by Arnold [1] shows that object segmentation for broad domains of general images is not likely to succeed, with a possible exception for sophisticated techniques in a narrow domains. While, object recognition still difficult using colour as the base for retrieval is promising and may give valuable results. The problem with the colour histograms that they do not provide any information related to shape, texture or location [2].

Colour histograms are a way to represent the distribution of colours in images where each histogram bin represents a colour in a suitable colour space (RGB, Lab, etc). A distance, usually represented by quadratic form, between a query image histogram and a data image histogram can be used to define the similarity match between the two distributions. Typically, 256 colours are adequate to capture the color distributions of most natural scene [3].

In our approach to content-based image retrieval, colour has been considered as one of the primitive features of the image. The contribution of this paper can be drawn easily through a multiple layer image retrieval system based on median cut segmentation and using the mean and standard deviation as the main features to be calculated and compared in the colour histogram. That because the mean will give an idea about the brightness of the image and the standard deviation give an idea about the contrast and these two features, all what we need to compare image similarity based on colour in any CBIR (content based image retrieval) system.

2. Related Research Work

Enser [4] reviewed methods for providing subject access to pictorial data, developing a four-category framework to classify different approaches. He discusses the strengths and limitations both of conventional methods based on linguistic cues for both indexing and searching. His conclusions are that, while there are serious limitations in current text-based techniques for subject access to image data, significant research advances will be needed before visually based methods are adequate to achieve this task.

Aigrian [5] discussed the main principles of automatic image similarity matching for database retrieval, emphasizing the difficulty of expressing this in terms of automatically generated features. They review a selection of current techniques for both still image retrieval and video data management. They conclude that the field is expanding rapidly, but that many major research challenges remain, including the difficulty of expressing semantic information in terms of primitive image features, and the need for significantly improved user interfaces.

Eakins [6] proposed a framework for image retrieval, classifying image queries into a series of levels, and discussing the extent to which advances in technology are likely to meet the users' needs at each level. He concludes that automatic CBIR techniques can already address many of users' requirements.

Idris [7] provided an in-depth review of CBIR technology, explaining the principles behind techniques for colour, texture, shape and spatial indexing and retrieval in some detail. They identify a number of key unanswered research questions, including the development of more robust and compact image content features, more accurate modeling of human perceptions of image similarity, the identification of more efficient physical storage and indexing techniques, and the development of methods of recognizing objects within images.

Guojun[8] presented the problem associated with the original technique of using colour histogram for image retrieval. He stated that the original technique, in which the comparison - between the query and the images in the database- using bin-to-bin difference, yield to misleading result in a lot of circumstances. That because such a calculation ignored the relationships between the bins in the image. They used the perceptually weight histograms (PWH) to overcome the problem. In PWH, a pixel contributes weights to a number of perceptually similar bins instead of a single bin. The contributing weights are inversely proportional to the distance between the pixels and the bins.

Niblack [9] characterized the color composition of an image based on the colour histogram regardless to its scale and orientation.

Tat-seng [10] proposed a computer-aided segmentation technique to segment the image content based on modification of the colour pair segmentation technique.

3. Preprocessing

In the preprocessing stage of the system, we concentrate on resizing the image database, change the color space from RGB to CIE, and applying histogram equalization.

3.1 Color Space

Color image can be modeled as three monochrome image data, where each band of data corresponds to a different color. The actual information stored in the digital image data is the brightness information in each spectral band. Typical color images are represented as red, green, and blue (RGB images). Using the 8bit monochrome standard as a model, the corresponding color image would have 24bits/pixel (8 bits for each band). Most of the time the RGB color transformed into mathematical space that decouples the brightness information from the color information. Then, the image information will consist of a one-dimensional brightness (luminance) space, and two-dimensional color space. As an example for such a transform is the HSL color transform, which allow us to describe colors in terms of hue/saturation/lightness. There are various methods to transform RGB to HSL color space. Most of these methods are algorithmic in nature and are geometric approximations to mapping the RGB color cube into HSL color space.

One problem associated with the previous color spaces is that they are not perceptually uniform. This means that two different colors in one part of the color space will show the same degree of perceptual differences as two colors in another part of the color space, even though they are the same "distance" apart. Therefore, we cannot define a metric to show how close, or far apart, two colors are in terms of human perception. The following figure shows example to illustrate this problem. Color A may be orange and B may be green; colors C and D may be slightly different shades of same color, but as we can easily notice that $\triangle AB = \triangle CD$.

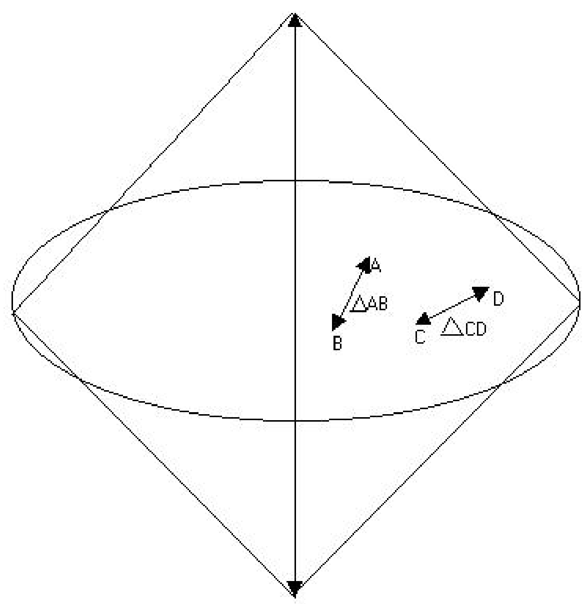


Figure 1: HSL color space

In computer imaging applications a perceptually uniform color space is important. CIE is one of the most common standards, which considered as a uniform color space. The basic concepts developed by the CIE involve chromaticity coordinates. For the RGB color space, chromaticity coordinates are defined as follows:

$$r = R/(R+G+B) \quad \ldots (1)$$
$$g = G/(R+G+B) \quad \ldots (2)$$
$$b = B/(R+G+B) \quad \ldots (3)$$

These equations basically normalize the individual components to the sum of the three values, which is a way to represent the brightness information. This decouples the brightness information from the coordinates, and the CIE uses chromaticity coordinates as basis of color transforms they define. These include the standard CIE XYZ color space and the perceptually uniform L*u*v, L*a*b color space. Since, the last two color spaces are perceptually uniform, they are more suitable to be used in our system. All images will be transformed to one of these color spaces.

3.2 Histogram Equalization

The following figure shows two images and the related colour histogram plots. It is clear that applying histogram equalization enhance the brightness of the image.

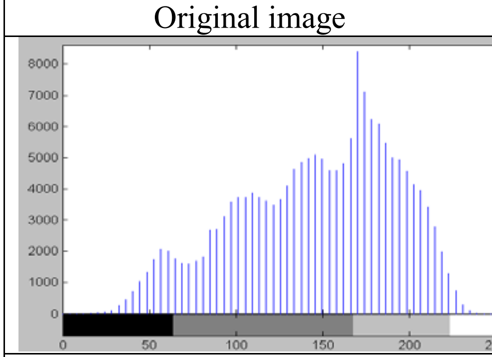
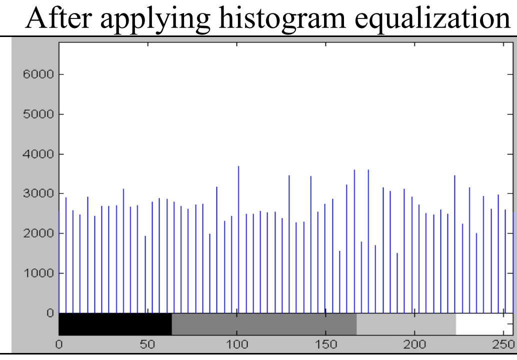
	
Original image	After applying histogram equalization
	
Colour histogram	Colour histogram (after applying histogram equalization)

Figure 2: Histogram Equalization

4. Histogram Feature

The histogram of an image defines as a plot of the gray level values versus the number of pixels. Figure 1, shows the different features we may extract from the colour histogram.

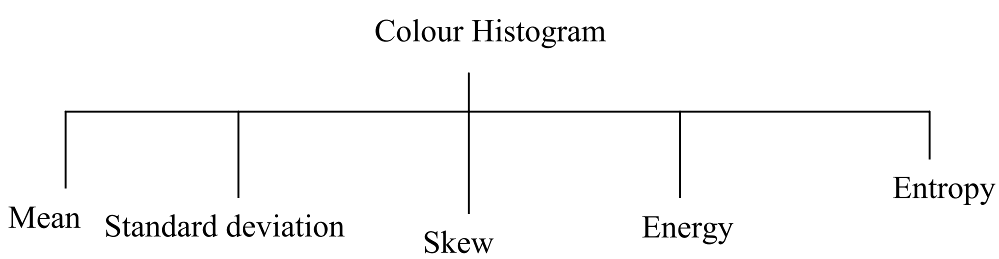


Figure 3: histogram features

The colour histogram provides information about the nature of the image, or an object within the image. We may obtain the following cases through simple analysis of the histogram.
Narrow histogram –implies a low contrast image. Histogram skewed toward the ends implies a bright image. Histogram with two major peaks, called ***bimodal***, implies an object that has a contrast with the background. Histogram features are statistically based features. It used as a modal of the probability distribution of the gray level.

Since colour histogram is a joint probability distributions, the first order histogram probability *P*(g) can be defined as:

$$P\,(g)= \frac{N(g)}{M} \quad \text{……………………………………………………………………(4)}$$

Where M is the number of pixels in the image and $N(g)$ is the number of pixels at gray level g. As any other joint probability distributions, the following conditions should be true.
$\forall g$, $P(g) \leq 1$ and $\sum_{g} P(g) = 1$.
Now it is the time to consider the different features of the color histogram with some details:

Mean: The value of the Mean shows the general brightness of the image. As a general rule bright images has high mean, while dark image has low mean. The mean define as:

$$\text{M}= \sum_{g=0}^{l-1} g\,p(g) \quad \text{…………..………….….……………………..…..(5)}$$

Where l is the gray level.
If we combine formula 1 and 2, we will obtain the following definition for the mean.

$$\text{M}= \sum_{g=0}^{l-1} g\,p(g) = \frac{1}{M}\left(\sum_{g=0}^{l-1} g\,N(g)\right) \quad \text{…………………..………..(6)}$$

Standard Deviation: Through the value of the standard deviation, we can obtain clear idea about the image contrast. As a general rule high standard deviation means high image contrast, while small standard deviation means low image contrast. The standard deviation define as:

$$\sigma_g = \sqrt{\sum_{g=0}^{l-1} (g-\bar{g})^2 P(g)} \quad \text{…………..……………………………………..(7)}$$

Energy: Energy shows how the gray level is distributed. The maximum value of energy is 1 and it become smaller as the pixel value distributed among the gray level. Energy define as:

$$\text{Energy} = \sum_{g=0}^{l-1} \left(p(g)\right)^2 \quad \text{……………………………………………..(8)}$$

Skew: Measures the asymmetry about the mean in the Gray level distribution. Skew define as:

$$\text{Skew} = \frac{1}{\sigma_g}\sum_{g=0}^{l-1}\left(g-\bar{g}\right)^3 P(g) \quad \text{.......................................(9)}$$

Entropy: It has been used in compression to lower the number of bits needed to store the image. Entropy define as:

$$\text{Entropy} = -\sum_{g=0}^{l-1} P(g)\log_2\left[P(g)\right] \quad \text{..(10)}$$

5. Median Cut Segmentation

This algorithm is based on an algorithm originally developed to map 24-bit color images to 8-bit color. It works by finding the maximum spread along the red, green or blue axes, and then dividing the color space with the median value along that axis. This division of the color space continues until the number of desired colors is reached. At this point, all the color vectors in a given subdivision of the color space are used to find an average color for that subdivision. After all the average colors are found, the algorithm goes back and maps each of the original color vectors to the closest one.

In our approach we used Median cut to segment all images in the database. Median cut segmentation enable us to reduce the image size dramatically at the same time the quality of the image still acceptable and no differences can be noticed through the human eye. Median cut allows us to control the number of colors in the image, which yield to reduce the image size if we reduce the amount of colors in that image. After testing the algorithm on some images we decide to use 128 colors as the maximum number of colors retained by the image. There are no significant changes in the appearance of the image and the human eyes as mentioned before do not recognize these changes. The following example shows three images the original image and two segmented images with 128 colors and 5 colors.

		
Original image with no segmentation Size: 26.0 KB	Same image after applying median cut with 128 colors Size: 20.7 KB	Same image after applying median-cut with 5 colors Size: 20.1 KB

Figure 4: Median-cut Segmentation Testing

6. Similarity Measurement and Ranking Technique

To measure the similarity between two images, distance function should be used. The difference can be measured by distance measure in the n-dimensional space (the bigger the distance between two vectors, the greater the difference). Given two vectors A and B, where $A = [a_1 a_2 a_3 .... a_n]$ and $B = [b_1 b_2 b_3 .... b_n]$. Then the distance between A and B can be calculated based on city block as follows:

$$D = \sum_{i=1}^{n} |a_i - b_i| \quad \text{..}(11)$$

City block is faster compare with other distance measure function like Euclidean distance function at the same time the result will be the same. And since we are looking to increase the speed of retrieval, city block gained our selection.

As a summary for our similarity and ranking approach we will measure the distance between two vectors (one for image query and the second for the image database). Each feature vector holds the mean and the standard deviation for the three bands (Red, Green, and Blue). Then a threshold value will be used to decide the similarity factor. This value can be flexible and based on user selection to allow increase or decrease the number of retrieved images, so the user can control the level of similarity. Since we are using distance function we will consider the image with closer distance as the most relevant, so the retrieved images will be arranged from left to right, up to down in display window based on the distance measurement.

7. System Design and Interface Consideration

7.1 System Design

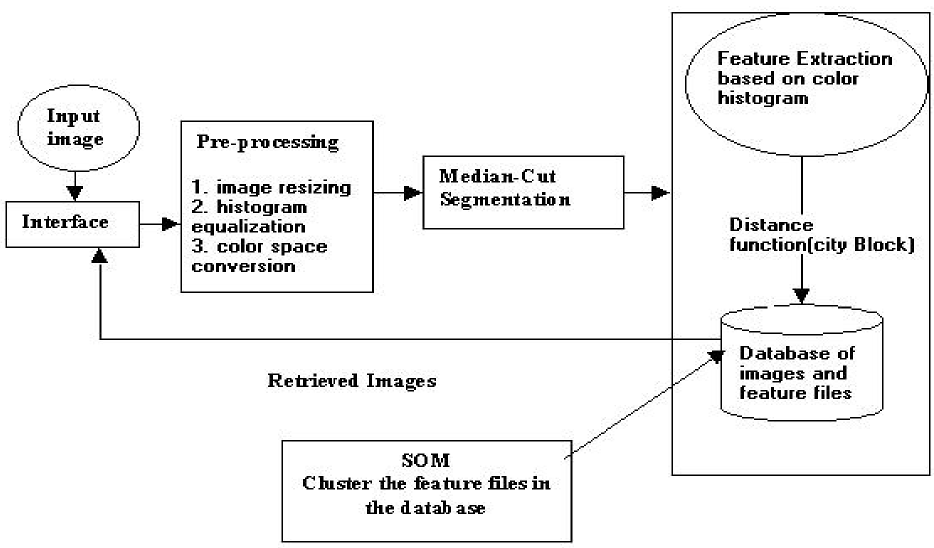


Figure 5: System Block Diagram

Figure 5 shows our approach for image retrieval by content through the system design. In the previous figure the whole system including the database and the self-organizing map has been shown.

7.2 Interface

To construct a visual query the interface should be flexible and allow refinement for the search result. The user may use the query result as a new query. Furthermore, the same interface to construct the image query will show the query result, so the image in a thumbnail will be presented to the user and zooming capabilities should be provided to allow the user to view the image with the original resolution. The user feedback is necessary to refine the searching and to help the system in modifying the retrieving accuracy, this part will be considered to let the user give an informative feedback, which can be employed by the system. In general there are two scenarios for any image retrieval interface:

1. The user has a sample image to use. So, the user can use the whole image or part of it in searching.
2. The user has a vague idea a bout the image in his mind.

Based on the above two scenarios our interface will provide query by example as well as query by template images. Moreover, the user can use the template images and modify, alter, or add some parts using a sketchpad provided within the interface.

8. System Evaluation

Using precision and recall to measure the accuracy of retrieval system is still the most prominent technique. We will use recall and precision to evaluate our system as well.

$$Precision = \frac{M}{R} \quad \text{and} \quad Recall = \frac{M}{D} \qquad \ldots\ldots\ldots(12)$$

Where M is the number of retrieved images matching the query, D is the number of images matching the query in the database obtained through subjective testing, and R is the total number of retrieved images.

9. Conclusion and Expectations

There are many approaches for retrieving and classifying images in a huge image database, most of the recently research concerning in retrieving by content rather than keywords. We proposed a new approach to retrieve images from the database. Our approach uses colour histogram combined with the median cut segmentation. Since there are many factors affecting the accuracy and speed of any CBIR system, we did consider the pre-processing stage to enhance the overall performance of our system. These include histogram equalization technique, images resizing, and colour spaced conversion. To overcome the speed problem, the system will allow for off-line calculation to measure the feature vector and save it in the same image database. Using template images in which the related information saved in advance in the database will give the user ability to search based on cues provided through the interface. After the complete implementation of the system we expect to introduce new system with high level of accuracy, better speed and user friendly interface.

References

[1] Arnold W. M. Semeulders, Simone Sanati, Amarnath Gupta, and Ramesh Jain (2000). " Content-based image retrieval at the end of the early years". *IEEE* 016-8826.

[2] Apostol Natsev, Rajeev Rastogi, and Kyuseok Shim (1999). " WARLUS: A similarity retrieval algorithm for image database". *ACM* 1-58113-084-8.

[3] James Hafner, Harpreet S. Sawhney, Will Equitz, Myron Flickner, and Wayne Niblack (1995). "Efficient color histogram indexing for quadratic form distance measure". *IEEE* 0162-8828.

[4] Enser P G B (1995). "Pictorial information retrieval". Journal of Documentation, 51(2): 126-170.

[5] Aigrain, P et al (1996). "Content-based representation and retrieval of visual media - a state-of-the-art review". Journal of Multimedia Tools and Applications3 (3): 179-202.

[6] Eakins, J P (1996). "Automatic image content retrieval - are we getting anywhere?". Proceedings of Third International Conference on Electronic Library and Visual Information Research (ELVIRA3), De Montfort University, Milton Keynes, pp 123-135.

[7] Idris F and Panchanathan, S (1997a)."Review of image and video indexing techniques". Journal of Visual Communication and Image Representation 8 (2):146-166.

[8] Guojun Lu and Jason Philips (1998). "Using perceptually weighted histograms for colour-based image retrieval". In proceeding of ICSP, 1998.

[9] Niblack W. *et al* (1998). "Updates to the QBIC system" in Storage and Retrieval for Image and Video Databases VI, Proceeding SPIE 3312, 150-161.

[10] Tat-Seng Chua, Swee-Kiew Lim and Hung-Keng Pung (1994). "Content-based Retrieval of segmented Images". ACM 0-89791-686, Multimedia 94-10/94.