

Automatic Segmentation and Indexing Image Colors

NIDHAL KHDHAIR EL ABBADI

Education College
University of Kufa, Najaf
Iraq

LAMIS AL SAADI

Education College
University of Babylon, Babylon
Iraq

Abstract:

The growth of digital image and video archives is increasing the need for tools that effectively filter and efficiently search through large amounts of visual data.

Towards this goal we suggested new method to automatically extract the color for the image to form a class of meta-data that is easily indexed. The algorithm of indexing the color based on the binary image to extract color regions from image. To achieve this goal the image segmented to many parts after converting color image to binary image based on Y component from the YCbCr color space. Color factors determined for each part by first convert color image for each part to HSV color space and then from the H component define the dominant color for that part. The proposed algorithm tested with many different images and gives very promised results.

Key words: color indexing, segmentation, HSV, YCbCr, image processing.

1. Introduction

The growth of using digital images and with increasing the size of images databases and image management tool such as

Google's image search and photo album tools such as Google's Picasa project, as well as image search applications in general social networking environment, the request emerged for new tools and techniques that effectively online search the huge database and retrieval of images [1].

Content based image retrieval (CBIR) has been one of the most important research areas in computer science for the last decade.

Application of World Wide Web (www) and the internet is increasing exponentially, and with it the amount of digital images data accessible to the users. A huge amount of image databases are added every minute and so is the need for effective and efficient image retrieval systems [4].

Content-based image retrieval has become a prominent research topic because of the proliferation of video and image data in digital form. The increased bandwidth availability to access the internet in the near future will allow the users to search for and browse through video and image databases located at remote sites.

One of the important techniques in this context is the color indexing which retrieve the images based on image color content. Some of the important objectives required for color indexing process such as: automated extraction of color, efficient indexing, and effective retrieval.

There are two methods for color indexing, the first one is global color distribution and the second one is the local or region color. The difference between these two methods is that the global distribution relays on comparing the color for the whole images, while in the regional we look for matching colors for the localized regions within images. Both methods are very useful in retrieve of images.

Many researchers' works in this field and many algorithms proposed; some of them are:

Sharma, et al. presented two methods for describing the contents of images. The first one characterizes images by global

descriptor attributes, while the second is based on color histogram approach. To compute feature vectors for Global descriptor, required time is much less as compared to color histogram. Hence cross correlation value & image descriptor attributes are calculated prior histogram implementation to make CBIR system more efficient. It is founded that further modifications are needed to produce better performance in searching images [6].

Singh and Hemachandran, proposed a content-based image retrieval method which combines color and texture features. To improve the discriminating power of color indexing techniques, they encode a minimal amount of spatial information in the color index. As its color features, an image is divided horizontally into three equal non-overlapping regions. From each region in the image, they extract the first three moments of the color distribution, from each color channel [7].

Vilvanathan and Rangaswamy, proposed an image retrieval technique based on color indexed image histograms. The RGB image is converted to an indexed image with 256 colors and the color map of a single image is used to decompose the entire dataset images. After that, the images are indexed so that each bin in a histogram corresponds to a specific color based on the index value [8].

2. Color Feature Extraction

Feature extraction means obtaining useful information that can describe the image with its content. We mean by image features the characteristic properties. For example, the image of a forest can be described by its green color.

Color is the most extensively used visual content for image retrieval. Its three-dimensional values make its discrimination potentiality superior to the single dimensional gray values of images. Before selecting an appropriate color description, color space must be determined first [2].

In image retrieval, the color is commonly used feature. Varieties of color spaces are available and different colors are defined on selected color space. To make image recognition possible by human, important feature is color.

Color has been the most effective feature and almost all systems employ colors. Although most of the images are in the red, green, blue (RGB) color space, this space is only rarely used for indexing and querying as it does not correspond well to the human color perception. It only seems reasonable to be used for images taken under exactly the same conditions each time such as trademark images [5].

Other spaces such as hue, saturation, value (HSV) or the CIE Lab and Luv spaces are much better with respect to human perception and are more frequently used. This means that differences in the color space are similar to the differences between colors that humans perceive. Much effort has also been spent on creating color spaces that are optimal with respect to lighting conditions or that are invariant to shades and other influences such as viewing position [3].

3. Proposed Methodology

One of the important factors used in retrieved image is the colors of different image regions. In this paper we suggested new algorithm for automatically segmented the images to different image partitions, and then detecting the partitions colors.

In this paper we suggested image segmentation which is differ from the region extraction, where the process of partition the image such that each point is assigned to one segment is called segmentation, while in the case that an image point may be assigned to many regions or to none called region extraction.

Segmentation is done to decompose an image into meaningful parts for further analysis, this step convert the color image to the binary image, at this case the image

background are changed as black color, while the white pixels represent the objects.

We suggested new method to segment the image based on converting the color image to 3 bands (red, green, and blue) and then transforming the blue band image to YCbCr color space. From YCbCr we use the Y component to determine the threshold to convert the image to binary, the threshold determined by the following relation:

$$\text{Threshold} = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N Y(i, j)$$

Where, Y (i, j) is the Y component value at the coordinate (i, j), and (M, N) is the image dimensions.

Now, each value in the blue band image will be compared with threshold, if its value greater or equal to threshold then it will be changed to the value one (white color), otherwise change to zero (black color), the resulted image from this process showed in Fig. 1. In this case the binary image mostly segmented the color image to the partitions according to the colors.



Fig. 1: color image and its binary image.

In the binary image the suggested algorithm traces the exterior boundaries of each object, and then all the components (partitions) in the binary image are labeled by using image label algorithm, which scan all image pixels, assigning preliminary labels to nonzero pixels and recording label equivalences in a union-find table. Then, resolve the equivalence classes using the union-find algorithm (The Union-Find algorithm is used for maintaining a number of non-

overlapping sets from a finite universe of elements). Finally, re-label the pixels based on the resolved equivalence classes.

The final step is to recognize each labeled partition from the other similar partitions by recover its original colors from the origin image. This is achieved by transform each segmented color partition to the HSV color space; we use the HSV color space for color feature generation. The hue (H) represents the dominant spectral component (color in its pure form), as in red, blue, or yellow. Adding white to the pure color changes the color: the less white, the more saturated the color is, this corresponds to the saturation (S). The value (V) corresponds to the brightness of color. The hue (color) is invariant to the illumination and camera direction, and thus suitable for object color recognition.

The H values of a pixel can be extracted from its RGB representation according to the following formula:

$$H = \arctan \frac{\sqrt{3}(G - B)}{(R - G) + (R - B)}$$

Later determine the color factor parameter (E), color factor use to recognize the color for each labeled partition. This can be achieved by using the following relation:

$$E = \frac{\sum_{k=1}^n (P_k)^2}{\sum_{k=1}^n P_k}$$

Where (P_k) is the value of pixel k, and n is the number of pixels in the partition.

This relation applied for many pure colors to determine the color factors, the results are summarized in the table 1.

Table 1: Values of color factor

Color Name	Color Factor
Yellow	0.1
Brown	0.2
Green	0.3
Blue	0.6
Purple	0.8
Red	0.9

5. Results

The proposed algorithm tested with many images, and the results were as follow:



Figure1: images with different colors and objects



Figure 2: images with colors mainly blue and green



Figure 3: images with colors mainly yellow and green



Figure 4: image with different colors

Table 2: results of the tested images

Fig. No.	Color factor	color
1	0.9	red
	0.3	green
2	0.6	blue
	0.3	green
3	0.1	yellow
	0.3	green
4	0.6	blue
	0.3	green
	0.8	purple

6. Conclusions

In this paper we suggested and evaluate new method for automated segmentation and extraction of local color partitions in image, efficient color indexing for each partition.

We use the color space YCbCr for segmentation which gives very accurate results. Also we use the hue (color) which is invariant to the illumination and camera direction, and thus suitable for object color recognition.

The algorithm tested with many images and gives promised results. We think this algorithm need more work to define color factors for more colors which can represents in the comuter.

REFERENCES

- Chakravarti, Rishav and Meng Xiannong. 2009. "A Study of Color Histogram Based Image Retrieval." *Sixth International Conference on Information Technology: New Generations*. 1323- 1328. [1]
- Long, F., Zhang, H., Dagan, H., and Feng, D. 2003. "Fundamentals of content based image retrieval." In *Multimedia Information Retrieval and Management Book*, 1-26. Berlin, Heidelberg: Springer-Verlag. DOI: 10.1007/978-3-662-05300-3_1. [2]

- Müller, Henning, Michoux, Nicolas, Bandon, David, and Geissbuhler, Antoine. 2004. "A review of content-based image retrieval systems in medical applications—clinical benefits and future directions." *International Journal of Medical Informatics* 73: 1-23. [3]
- Reddy P. V. N. and Satya Prasad, K. 2011. "Color and Texture Features for Content Based Image Retrieval." *Int. J. Comp. Tech. Appl.* 2(4): 1016-1020. [4]
- Rehman, Mehwish, Muhammad, Iqbal, Muhammad, Sharif and Mudassar, Raza. 2012. "Content Based Image Retrieval: Survey." *World Applied Sciences Journal* 19(3): 404-412. [5]
- Sharma, Neetu, Paresh, Rawat, and Jaikaran, Singh. 2011. "Efficient CBIR Using Color Histogram Processing." *Signal & Image Processing: An International Journal (SIPIJ)* 2(1): 94-112. [6]
- Singh S. Mangijao, and Hemachandran, K. 2012. "Content-Based Image Retrieval using Color Moment and Gabor Texture Feature." *IJCSI International Journal of Computer Science Issues* 9(5): 299- 309. [7]
- Vilvanathan, K. and Rangaswamy, R. 2013. "Bi-level classification of color indexed image histograms for content based image retrieval." *Journal of Computer Science* 9: 343-349. [8]



Nidhal El-Abbadi. University of Kufa, Education College, Najaf, IRAQ. He received BSc in chemical engineering, MSc, and PhD in computer science, worked in industry and many universities, he is general secretary of colleges of computing and informatics society in Iraq, Member of Editorial board of Journal of Computing and Applications, reviewer for a number of international journals, has many published papers and three published books (Programming with Pascal, C++ from beginning to OOP, Data structures in simple language). his research interests are in image processing, biomedical, and steganography, He's Associate Professor in Computer Science at the University of Kufa – Najaf, IRAQ. E-mail: nidhal.elabbadi@uokufa.edu.iq, nkhdhair@yahoo.com, Tel: +964-7706821230

Lamis Al Saadi. University of Babylon, Education College, Babylon, IRAQ. She received her BSc. in Computer Science from Babylon University; currently she is PhD student in computer science department, College of Science, Babylon University. She is working as a lecturer in College of Education, Babylon University. She has a number of research papers in the computer science field.