International Journal of Computer Science and Informatics

Volume 1 | Issue 3 Article 3

January 2012

Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture

Swati V. Sakhare
Dept. of Electronics Engg., Bapurao Deshmukh College of Engg., Sevagram, Wardha (India), swatisakhare123@rediffmail.com

Recommended Citation
Sakhare, Swati V. (2012) "Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture," *International Journal of Computer Science and Informatics*: Vol. 1 : Iss. 3 , Article 3.
DOI: 10.47893/IJCSI.2012.1029
Available at: https://www.interscience.in/ijcsi/vol1/iss3/3

Available online at www.interscience.in

Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture

Swati V. Sakhare & Vrushali G. Nasre

Dept. of Electronics Engg., Bapurao Deshmukh College of Engg., Sevagram, Wardha (India)
E-mail : swatisakhare123@rediffmail.com

Abstract - Retrieval of images based on visual features such as color, texture and shape have proven to have its own set of limitations under different conditions. Various techniques have been implemented using these features like fuzzy color histogram, Tammura texture etc. In this paper we propose a novel method with highly accurate and retrieval efficient approach which will work on large image database with varied contents and background.

Keywords— *Fuzzy-Color Histogram, Tammura Texture, Boundary and shape representation.*

.

I. INTRODUCTION

CBIR or Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps:

Feature Extraction: The first step in the process is extracting image features to a distinguishable extent.

Matching: The second step involves matching these features to yield a result that is visually similar

COLOR

One of the important features that make possible the recognition of images by humans is color. Color is a property that depends on the reflection of light to the eye and the processing of that information by the brain. We use color to tell the difference between objects, places, and the time of day. Usually colors are defined in three dimensional color spaces. These could either be ***RGB*** (Red, Green, and Blue), ***HSV*** (Hue, Saturation, and Value) or ***HSB*** (Hue, Saturation, and Brightness). The last two are dependent on the human perception of hue, saturation, and brightness.

Most image formats such as ***JPEG***, ***BMP***, ***GIF***, use the RGB color space to store information. The RGB color space is defined as a unit cube with red, green, and blue axes. Thus, a vector with three co-ordinates represents the color in this space. When all three coordinates are set to zero the color perceived is black. When all three coordinates are set to 1 the color perceived is white. The other color spaces operate in a similar fashion but with a different perception.

The method of representing color information of images in CBIR systems is through color histograms. A color histogram is a type of bar graph, where each bar represents a particular color of the color space being used. In MatLab for example one can get a color histogram of an image in the RGB or HSV color space. Bars in a color histogram are referred to as bins and they represent the x-axis. The number of bins depends on the number of colors there are in an image. Y-axis denotes the number of pixels in each bin. In other words it gives the count of pixels in an image representing a particular color.

Quantization in terms of color histograms refers to the process of reducing the number of bins by taking colors that are very similar to each other and putting them in the same bin. Default, the maximum number of bins one can obtain using the histogram function in MatLab is 256. For the purpose of saving time when trying to compare color histograms, one can quantize the number of bins. Obviously quantization reduces the information regarding the content of images but it is a trade off between processing time and quality.

There are two types of color histograms, Global Color Histograms (***GCH***s) and Local Color Histograms (***LCH***s). A GCH represents one whole image with a single color histogram. An LCH divides an image into fixed blocks and takes the color histogram of each of those blocks. LCHs contain more information about an image but are computationally expensive when comparing images. “The GCH is the traditional method for color based image retrieval. However, it does not include information concerning the color distribution of the regions” of an image. Thus when comparing GCHs one may get inconsistent result in terms of similarity of images.

TEXTURE

Texture is an innate property of all surfaces that describes visual patterns, each having homogeneity. It contains important information about the structural arrangement of a surface, such as; clouds, leaves, bricks, fabric, etc. It also describes the relationship between the surfaces to the surrounding environment. In short, it is a feature that describes the distinctive physical composition of a surface.

Textures can be modeled as quasi-periodic patterns with spatial/frequency representation. The wavelet transform transforms the image into a multi-scale representation with both spatial and frequency characteristics. This allows for effective multi-scale image analysis with lower computational cost. According to this

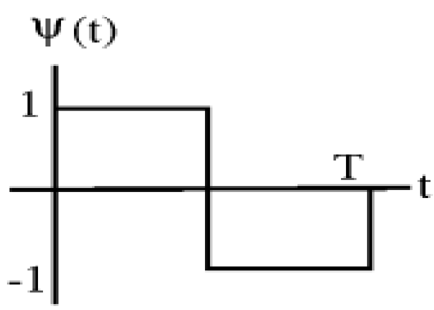


Figure 1 : Haar Wavelet Example

Transformation, a function, which can represent an image, a curve, signal etc., can be described in terms of a coarse

Level description in addition to others with details that range from broad to narrow scales.

Unlike the usage of Sine functions to represent signals in Fourier transforms, in wavelet transform, we use functions known as wavelets. Wavelets are finite in time, yet the average value of a wavelet is zero. A wavelet is a waveform that is bounded in both frequency and duration. While the Fourier transform converts a signal into a continuous series of Sine waves, each with constant frequency and amplitude and of infinite duration, most real-world signals (such as music or images) have a finite duration and abrupt changes in frequency. This account for the efficiency of wavelet transforms.

Examples of wavelets are Coiflet, Morlet, Mexican Hat, Haar and Daubechies. Of these, Haar is the simplest and most widely used, while Daubechies have fractal structures and are vital for current wavelet applications. These two are outlined below:

Haar Wavelet

The Haar wavelet family is defined as:

Ψ ((t) = 1 if 0 ≤ t and lt; T/2

-1 if 0 ≤ T/2 and lt; T

Daubechies Wavelet:

The Daubechies wavelet family is defined as

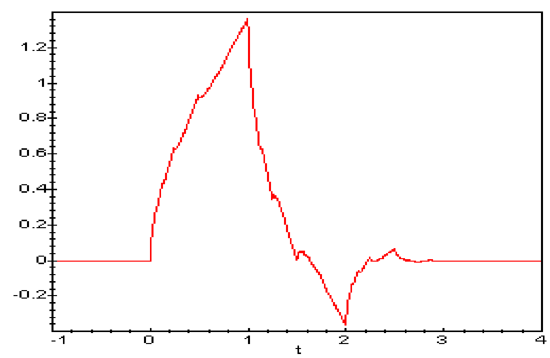


Figure 2: Daubechies Wavelet Example

II. PROBLEM STATEMENT

The problem involves entering an image as a query into a software application that is designed to employ CBIR techniques in extracting visual properties, and matching them. This is done to retrieve images in the database that are visually similar to the query image.

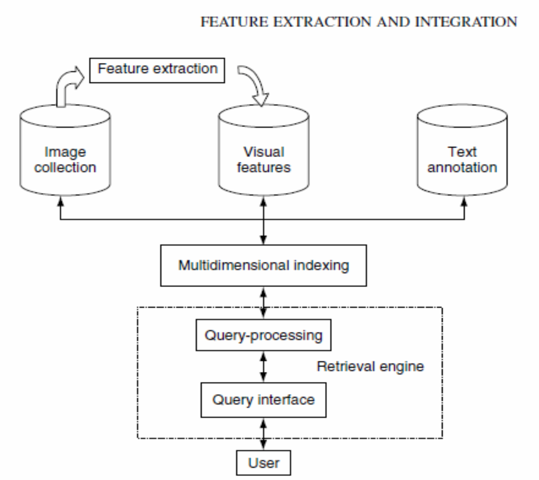
III. PROPOSED WORK

This section summarizes the culmination of the literature survey carried out in order to put forth the necessity for carrying out the proposed research work [2][3]. It is proposed:

- Design and Develop a system for Features Extraction of images for Content Based Image Retrieval (CBIR)
- Extract Color based and Texture based features of image database
- Optimize the effectiveness and accuracy of the CBIR system

To implement and evaluate proposed CBIR system for Image Retrieval of real-time image databases.

IV. ARCHITECTURAL DESIGN



An image retrieval system architecture.

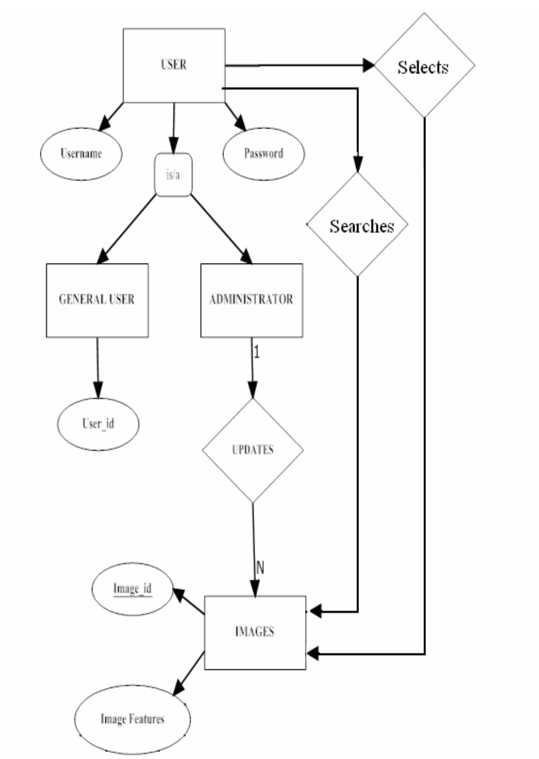


Figure : ER

V. ALGORITHM

1. **Collection of Image Database**

- We consider a database containing 100 images with any one of the formats .bmp, .jpg or .tiff.
- The images will be from RGB color model.

2. **Feature Extraction**

- Feature Extraction is carried out by using colours, using textures or by using shapes. For color feature extraction, color histograms such as Local Color Histogram (LCH), Global Color Histogram (GCH) and Fuzzy Color Histogram (FCH) are used. For extracting textures Statistical, Structural, Spectral approaches are used. In addition to this, Tammura Texture and Wavelet Transform are used. Boundary-based and Region-based color representation are used with Fourier Descriptor (Fourier transformed boundary) and Moment Invariants (Region-based moments).
- The images are registered with their corresponding features such as color, texture, shape.
- These extracted features will be forwarded to Feature Vector Module

3. **Similarity Measures**

The Direct Euclidian Distance between an image P and query image Q can be given as the equation below $ED=\sum (V_{pi} - V_{qi}).(V_{pi} - V_{qi})$.

Where, V_{pi} and V_{qi} be the feature vectors of image P and Query image Q respectively with size ‘n’

4. **Comparison of results with other techniques**

- When the user passes a query image, the composite feature vector of both query image and the image which is stored in database will go through Similarity Comparison

5. **Finally the image will be retrieved**.

VI. METHODOLOGIES

The solution proposed is to extract the primitive features of a query image and compare them to those of database images. The image features under consideration were color, texture and shape. Thus, using matching and comparison algorithms, the color, texture and shape features of one image are compared and matched to the corresponding features of another image. This comparison is performed using color, texture and shape distance metrics. Towards the end, these metrics are applied one after another, so as to retrieve database images that are similar to the query. The similarity between features is to be calculated using algorithms used by well known CBIR systems such as IBM's QBIC.

Following methodologies will be used in the proposed work

- Collection of Image Database
- Feature Extraction
- Similarity Measures
- Comparison of results with other techniques

VII. IMPLICATIONS

- The application performs a simple search in an image database for an input query image, using color, texture and shape to give the images which are similar to the input image as the output.
- The number of search results may vary depending on the number of similar images in the database.
- CBIR is still a developing science.

The development of powerful processing power and faster and cheaper memories contribute heavily to CBIR development. This development promises an immense range of future application using CBIR.

Possible outcomes

- Input → Query Image
- Output → The set of images from the image database with matching features with Query image

VIII. RESULTS

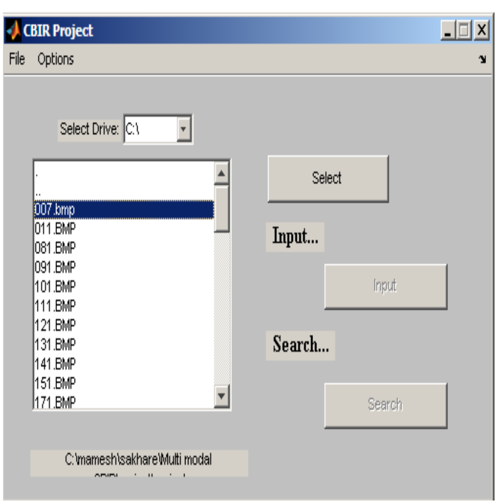


Fig. 8.1 : CBIR Front End

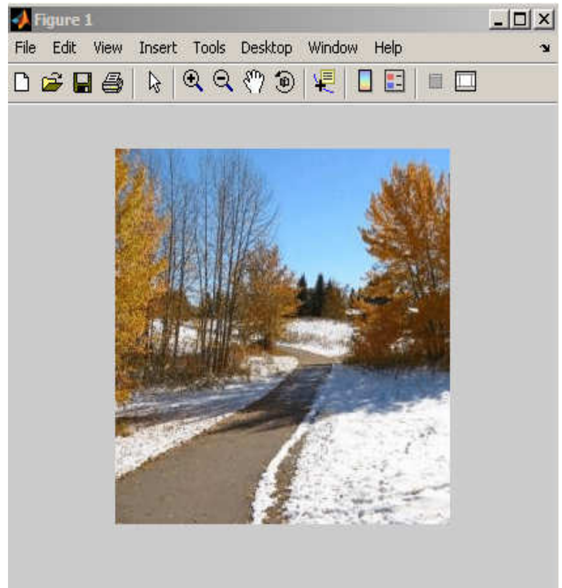


Fig. 8.2 : Select image for search

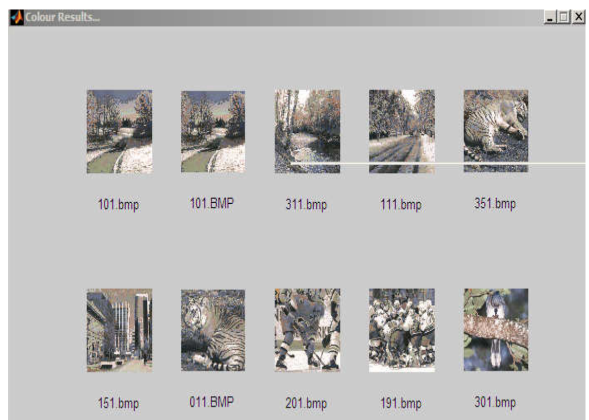


Fig. 8.3 : color results search

Fig. 8.4 Texture results search

Table 1: Performance Comparison

SR. NO	CBIR TECHNIQUE	Performance
1	COLOR FEATURES	35.25%
2	TEXTURE FEATURES	28.27%
3	COLOR & TEXTURE FEATURES	76.35%

IX. CONCLUSION

The dramatic rise in the sizes of images databases has stirred the development of effective and efficient retrieval systems. The application performs a simple color-based search in an image database for an input query image, using color, texture and shape to give the images which are similar to the input image as the output. The number of search results may vary depending on the number of similar images in the database.

CBIR is still developing science. As image compression, digital image processing, and image feature extraction techniques become more developed, CBIR maintains a steady pace of development in the research field.

REFERENCES

[1] Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng, "FUNDAMENTALS OF CONTENT-BASED IMAGE RETRIEVAL"

[2] Arti Khaparde, B.L.Deekshatulu, M.Madhavilath, Zakira Farheen, Sandhya Kumari V, "Content Based Image Retrieval Using Independent Component Analysis", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.4, April 2008.

[3] Tomasz Andrysiak, Michał Chora´, "Image Retrieval Based On Hierarchical Gabor Filters, Int. J. Appl. Math. Compute. Sci., 2005, Vol. 15, No. 4, 471–480

[4] S.A.Pine, R C Patil, "Content Based Image Retrieval Using Wavelet Transform" Telecommunication' Thadomal Shahani Engineering College, Bandra (East), Mumbai-51.

[5] X. Fu, Y. Li, R. Harrison, S. Belkasim, "Content-based Image Retrieval Using Gabor-Zernike Features", The 18th International Conference on Pattern Recognition (ICPR'06)

[6] Guoping Qiu, "Color Image Indexing Using BTC,"IEEE Transactions on Image Processing, VOL.12, NO.1, pp.93-101, January 2003

[7] H.B.Kekre, Sudeep D. Thepade, "Boosting Block Truncation Coding with Kekre's LUV Color Space for Image Retrieval", International Journal of Electrical, Computer, and Systems Engineering 2; 3 © www.waset.org summer 2008

[8] Kai-Krung Ma, and Sarah A. Rajala, 1991, Sub band Coding of Digital Images Using Absolute Moment Block Truncation, IEEE Transactions on Acoustics, Speech, and Speech, and Signal Processing, 1991. ICASSP-91., International Conference on, 1991 Vol. 4, pp. 2645-2648

[9] Dr.H.B.Kekre, Ms. Swapna Borde," Content Based Image Retrieval," National Conference on Applications of Digital Signal Processing, NCDSP-2007, January 19-20, 2007

[10] Lei Zhu, Chun Tang, Aibing Rao and Aidong Zhang,"Using Thesaurus To Model Keyblock-Based Image Retrieval ," Technical Report, Department of Computer Science and Engineering, State University of New York At Buffalo,Buffalo,NY 14260,USA.

[11] K Bhandari, N Dugar, N Jain, N Shetty, "A novel Hiogh Performance Multi-Modal approach for Content Based Image Retrieval", ICWET 2010

❑❑❑