International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 8, No. 4 (2015), pp. 297-306
http://dx.doi.org/10.14257/ijsip.2015.8.4.26

A New Content Based Image Retrieval System by HOG of Wavelet Sub Bands

Dr. Anna Saro Vijendran and S. Vinod Kumar

Director, Department of MCA, SNR Sons College, Coimbatore
Assistant Professor, PG Department of Computer Science, STC College of Arts & Science, Pollachi
saroviji.271@rediffmail.com, vinomsc38@rediffmail.com

Abstract

The term content-based image retrieval to describe his experiments into automatic retrieval of images from a database by color and shape feature. The term has since been widely used to describe the process of retrieving desired images from a large collection on the basis of features that can be automatically extracted from the images themselves. The features used for retrieval can be either primitive or semantic, but the extraction process must be predominantly automatic. Retrieval of images by manually-assigned keywords is definitely not CBIR as the term is generally understood – even if the keywords describe image content. And this paper deals, retrieve the images into three basic categories of color called RGB. After retrieving the image with their component the transformation can be applied. The HOG method is used to retrieve the feature of image vectors and others. In this paper, the HOG method is fully analyzed and proves its accuracy and efficiency of image retrieval with reduced number of steps.

Keywords: *Content Based Image Retrieval (CBIR), RGB, Discrete wavelet Transformation (DWT), Histogram of Oriented Gradient (HOG)*

1. Introduction

There are many resources on the internet which people can use to create process and store images. This has created the need for a means to manage and search these images. Image retrieval method is a technique for searching and retrieving images from a large database of digital images. A large collection of these images is referred to as image database. An image database is a system where image data are integrated and stored. Image data include the raw images and information extracted from images by automated or computer assisted image analysis. Development of methods which would increase retrieval accuracy and reduce retrieval time is the main challenges in CBIR.

Early techniques were not generally based on visual features but on the textual annotation of images. The images were first annotated by text and then searched using text based approach. However in many situations, text annotation scheme is inefficient. For the huge image data the vast amount of labor required in manual annotation. Also describing every visual feature within the images is very time consuming and difficult. So instead of manual annotations by text based keywords, images are indexed by their own visual features such as color, texture, shape *etc*.

In CBIR no additional information on images, such as text annotations, time or place of creation is available. The retrieval problem is solved only by analyzing content of the image based on the available characteristics of its pixels.

An alternative method of the content-based image retrieval is description based image retrieval (DBIR). In DBIR, retrieval is possible if all images of the collection have annotations describing their content. A general CBIR system makes use of different type

ISSN: 2005-4254 IJSIP

of queries such as query by example image, sketch or region and provides relevant images from a given database, based not exclusively on textual annotation or media metadata, but on a similarity function using low-level features.

The algorithms are act as a major role in image retrieval from databases. Many researchers proposed many algorithms and each algorithm has a unique feature and drawbacks too. In this paper, discrete wavelet transformation (DWT) algorithm is proposed. DWT algorithm retrieves the images from the database with HOG method. This method initially retrieves the images with components called RGB and then their features are extracted.

1.1. Architecture of CBIR System

Content-based retrieval uses the contents of images to represent and access the images from the large database.

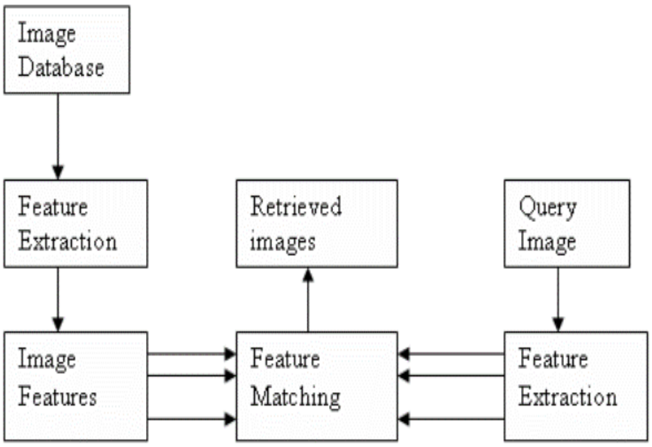


Figure 1.1. General Architecture of CBIR

The general architecture of CBIR system is for the given image database, features are extracted first from individual images. The features can be visual features like color, texture, shape, region or spatial features or some compressed domain features. The extracted features are described by feature vectors. These feature vectors are then stored to form image feature database. For a given query image, we similarly extract its features and form a feature vector. This feature vector is matched with the already stored vectors in image feature database.

2. Literature Review

Enser [1995] reviews 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 search, and experimental systems using visual cues for one or both of these. His conclusions are that, while there are serious limitations in current text-based techniques.

Eakins [1996] proposes 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 users' needs at each level. His conclusion is that automatic CBIR techniques can already address many of users' requirements at level 1, and will be capable of making a significant contribution at level 2 if current research ideas can be successfully exploited. They are however most unlikely to make any impact at level 3 in the foreseeable future.

Idris and Panchanathan [1997a] provide an in-depth review of CBIR technology, explaining the principles behind techniques for color, texture, shape and spatial indexing and retrieval in some detail. They also discuss the issues involved in video segmentation, motion detection and retrieval techniques for compressed images.

Recent works have shown that efficient and robust shape-based cues can be obtained from histogram of oriented gradient (HOG) in images.

Dalal, *et al.*, [9] proposed a human detection algorithm using histograms of oriented gradients (HOG) which are similar with the features used in the SIFT descriptor. HOG features are calculated by taking orientation histograms of edge intensity in a local region. They are designed by imitating the visual information processing in the brain and have robustness for local changes of appearances, and position.

Dalal, *et al.*, extracted the HOG features from all locations of a dense grid on an image region and the combined features are classified by using linear SVM.

3. Methodology

Color feature is the most significant one in searching collections of color images of arbitrary subject matter. Color plays very important role in the human visual perception mechanism. All methods for representing color feature of an image can be classified into two groups: color histograms and statistical methods of color representation.

The most frequently used color spaces are as follows: RGB (red, green, and blue used in color monitors and cameras), CMY (cyan, magenta and yellow), CMYK (cyan, magenta, yellow, and black used in color printers), Lab (CIE L*a*b, lightness, a and b are two color dimensions, from green to red and from blue to yellow) HSI, HSV (hue, saturation, and value).

The Lab space relies on the international standard of color measurement developed by the International Commission on Illumination CIE (Commission International de Eclairage). The HSV space is similar to spaces HSI, HSL, and HSB. The HSV space is used more frequently because the RGB to HSV transformation is simpler from the computational standpoint compared to the RGB to Lab transformation.

The following architecture provides the steps for proposed CBIR methods and algorithm.

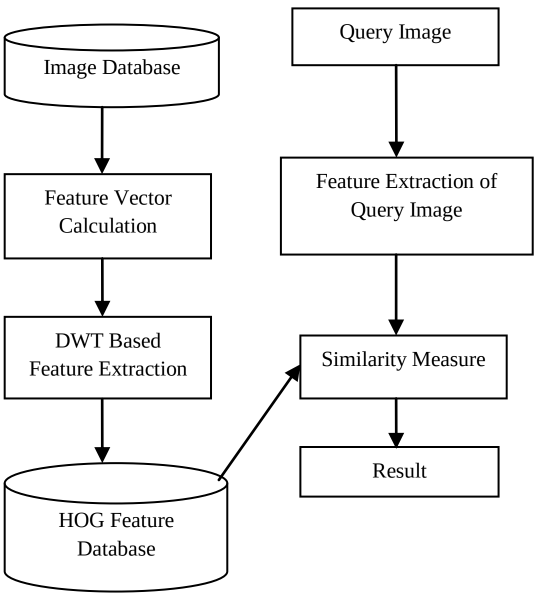


Figure 3.1. Architecture of Proposed CBIR

The each image has its own properties and components. These properties and components are extracted using the following two techniques, such as,

1. Discrete Wavelet Transformation (DWT).
2. Histogram Oriented Gradient (HOG).

Based feature extraction for the images. The following architecture shows the general architecture of future extraction and storage process.

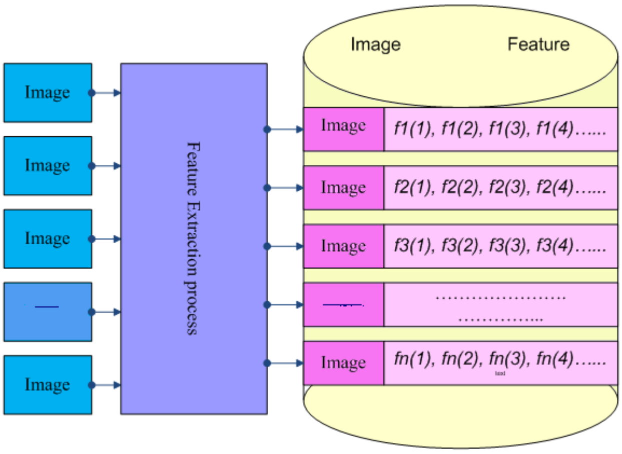


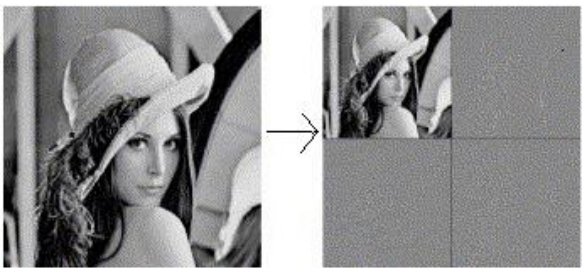
Figure 3.2. Feature Extraction and Storage Process for an Image Collection

3.1. Discrete Wavelet Transformation (DWT)

DWT is widely used for multi-scale image analysis. It decomposes an image into four sub-bands: an approximated image and horizontal (DH), vertical (Dv), and diagonal (Dd) detailed images. The detailed images measure variations along the columns (horizontal edges), rows (vertical edges), and diagonals (diagonal edges) respectively.

More than one decomposition level may be utilized for face recognition task to give reduced but meaningful information describing face image. The approximated image is decomposed again to wavelet sub-bands. Two or three decomposition levels may be used. The final resultant approximated image is used as a feature vector. It has three levels, one, level two, and level three of decomposition respectively.

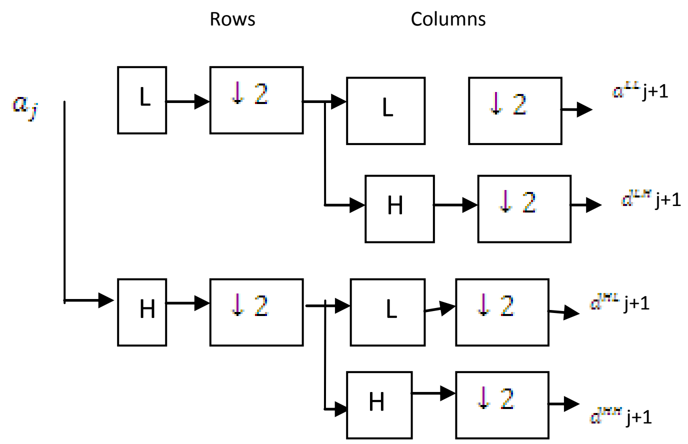
A color space is a model for representing colors in terms of intensity values. RGB color space is fundamental color space in imaging.



a) Input Image b) Single Level DWT

Figure 3. Single Level Decomposition Using DWT

The Figure 3 shows the single level transformation using DWT for input image Lena. Where H and L denotes high and low-pass filters respectively.



The most common method for comparing two images in content based image retrieval is by using an image distance measure, which compares the similarity of two images in various dimensions such as color, texture, shape, *etc.*

The Euclidean and City block distance measures are most common measures which are used to compare the similarity of any two images.

3.2. Histogram Oriented Gradient (HOG)

Histogram of Oriented Gradients (HOG) feature descriptor is used in this work to extract the features. It is very effective to represent objects and is widely used in human and face detection. The first step is detecting all interest points of the image using the Harris detector. This operator is based on the auto-correlation matrix that describes the local structure of the image. Then compute the Gradient Orientation Histogram around the 16 x 16 pixel region of each interest points. First, the region is divided in 4 X 4 sub-region, For each sub-region the 8- bin gradient orientation *h(k), k = 0 to 7* are calculated which forms a feature vector of size 128 dimension (4 x 4 x 8). The gradient oriented histogram is computed as follows:

$$h(k) = \sum_{d_{ij} \in d_k} m_{ij} \quad (4)$$

$$m_{ij} = \sqrt{dx_{ij}^2 + dy_{ij}^2} \quad (5)$$

$$d_{ij} = \arctan\frac{dy_{ij}}{dx_{ij}} - D \quad (6)$$

$$dx_{ij} = I_{ij} - I_{i+1,j} \quad (7)$$

$$dy_{ij} = I_{ij} - I_{i,j+1} \quad (8)$$

$$D = \arctan\frac{\sum_{ij} dy_{ij}}{\sum_{ij} dx_{ij}} \quad (9)$$

Where *Iij* is the *i, j* pixel value of each sub-region, *mij* is the Gradient magnitude of the pixel *i, j, dij* is the gradient direction at pixel *i, j, h (k)* is the *kth* dimension *h (k)* of the gradient histogram represents the total intensity of the pixel gradient whose direction lies in the *kth* direction bin *dk., k = 0 to7*. The direction bins are defined by the relative angle to the dominant gradient direction *D* of the image region. Finally, combing all the Gradient orientation Histogram of the interest point's area together to form a feature vector of size 128-dimension.

4. Experimental Results

The Evaluation is performed to find the relevant images for the given input query with reduced number of iterations. The experiment is done by using CORAL image Database with the software MATLAB.

Corel database contains large amount of images of various contents ranging from animals and outdoor sports to natural images. These images are pre-classified into different categories of size 100 by domain professionals. The accuracy of the image can be calculated by the following formula which is expressed in %:

$$Accuracy = \frac{N - X}{N} * 100$$

Where N is number of relevant images in the database which are known to the user and X is the number of irrelevant images in the database which are known to the user.

4.1. DWT based Image Retrieval

The image is retrieved from the databases by using query. The user is given the query to the database. The query image extracts all the features of the image, such as, color features and vector.

The following image is nothing but a query image given by the user.

Figure 4.1. Query Image

After implementing the Discrete Wavelet Transformation, it produces the following result,

Figure 4.2. Output Image Produced by using DWT Image Retrieval

In the above result, the two images not extracted fully. Its accuracy percentage is 80% only.

4.2. HOG by DWT Sub band based Image

Retrieval:-

Here, the HOG method is implemented with DWT algorithm and the images are searches from the database and produce the following result.

Figure 4.3. Output Image before Relevance Feedback using Texture-based Image Retrieval

The above result is obtained by HOG and DWT features. The user image query is implemented and searched from the databases based on the HOG and DWT. But the above result is error because it can only produce 88% of accuracy.

Then Figure 4.4. Shows the output images based on combined DWT and HOG after relevance feedback. The accuracy of images obtained in the second iteration is 98%. In this iteration itself the maximum output is produced, if maximum accuracy is not produced means next iteration can be done. This process can be carried till maximum accuracy is obtained.

Figure 4.4. Output Image 1 after Relevance Feedback Used

4.3. Performance Evaluation

The Corel image dataset consists of mainly six different images. The corresponding accuracy of the query images to display these images before and after relevance feedback has been observed for combined DWT and HOG image retrieval. The accuracy before and after relevance feedback for this image retrieval is shown in the Table 1.

Table 1. Accuracy and Time Comparison of Hog by DWT Sub Bandsbased Image Retrieval Before and After Relevance Feedback

Query Image	Accuracy (%) without RF	Accuracy (%) with RF	Number of Iterations
Beaches	72	82	3
Building	65	74	2
Dinosaur	93	98	2
Elephant	82	88	2
Food	74	85	3
Rose	88	92	2
Average	79	86	2

Table 1 illustrates that the average number of iteration for the six datasets is 2 and the maximum accuracy after relevance feedback is 86%.

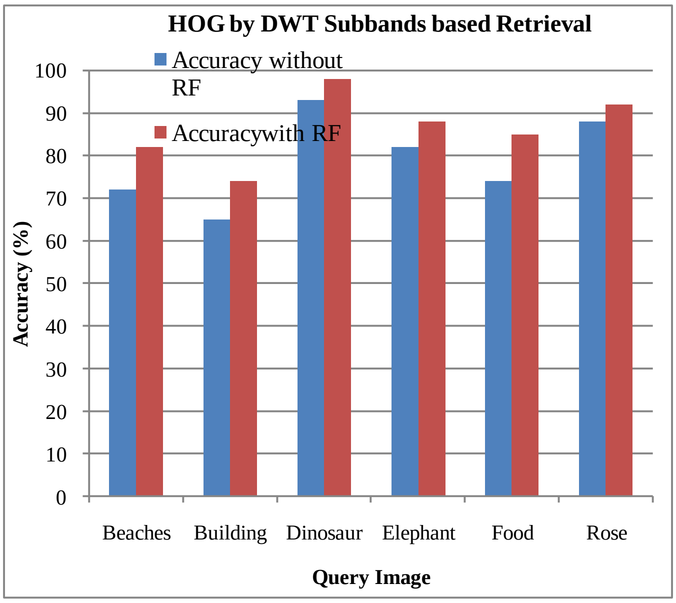


Figure 4.5. Accuracy Comparison of the HOG by DWT Sub Bands based Image Retrieval Before and After Relevance Feedback

The experimental result proved that the proposed method produces the nearly maximum accuracy for image retrieval.

5. Conclusion

The search for the relevant information in the large space of image databases has become more challenging. More précised retrieval techniques are needed to access the large image achieves being generated, for finding relatively similar images. This paper proposed a method of HOG and DWT for extracting the features of the images from the databases using user image query. The image features called R, G and B is extracted from the databases. Then extracted features of the database is compared with the query image and based on that image is retrieved. The accuracy of the image is measured using performance of the image.

References

[1] N. S. T. Sai and R. Patil, "Average Row and Column Vector Wavelet Transform for CBIR", Second international conference on Advance in Computer Vision and Information Technology ACVIT**, (2009)**, Aurangabad, India.

[2] N. S. T. Sai and R. Patil, "Image Retrieval using DCT Coefficients of Pixel Distribution and Average Value of row and Column Vector", IEEE International Conference on Recent Trends in Information, Telecommunication and Computing ITC, **(2009)**, Kochi, Kerala, India.

[3] N. S. T. Sai and R. Patil, "Moments of Pixel Distribution of CBIR", International Conference and Workshops on Emerging Trends in Technology ICWET, **(2010)**, Mumbai, India.

[4] N. S. T. Sai and R. Patil, "New Feature Vector for Image Retrieval: Sum of the Value of Histogram Bins", IEEE Conference on Advance in Computing, Control & Telecommunication Technologies ACT, **(2009)**, Trivandrum, India.

[5] N. S. T. Sai and R. Patil, "Image Retrieval using Equalized Histogram Image Bins Moment", International Joint Journal Conference in Engineering IJJCE, **(2010)**, Trivandrum, India.

[6] W. W. J. Z. Wang, G. Wiederhold, O. Firschein and S. X. Wei, "Wavelet Based Image Indexing Techniques with Partial Sketch Retrieval Capability", J. Digital Libraries, **(1997)**.

[7] M. Lamard, G. Cazuguel, G. Quellec, L. Bekri, C. Roux and B. Cochener, "Content Based Image Retrieval based on Wavelet Transform Coefficients Distribution", Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite International, Lyon, France, **(2007)** August 23-26.

[8] P. S. Hiremath, S. Shivashankar and J. Pujari, "Wavelet Based Features for Color Texture Classification with Application to CBIR", International Journal of Computer Science and Network Security (IJCSNS), vol. 6, no. 9A, **(2006)** September.

[9] C. S. Rao, S. S. Kumar and B. N. Chatterji, "Content Based Image Retrieval using Contourlet Transform", GVIP Journal, vol. 7, Issue 3, **(2007)** November.

[10] D. Lowe, "Object Recognition from Local Scale-Invariant Features", in Proc. of the 7th Int. Conf. on Computer Vision, vol. 2, **(1999)** September 20-25, pp. 1150.

[11] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points", International Conference on Computer Vision, **(2001)**, pp. 525-531.

[12] K. Mikolajczk and C. Schmid, "A performance evaluation of local descriptors", IEEE Conference on Computer Vision and Pattern Recognition.

[13] J. R. Smith and S. Chang, "Single color extraction and image query", in Proc. of IEEE Int'l Conf. on Image Processing, **(1995)**, pp. 528-531.

