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

Dr. Anna Saro Vijendran, S. Vinod Kumar,
PGCS, STC

***Abstract*---Content-based image retrieval (CBIR) is the application of computer vision techniques. It becomes an active and fast advancing research area. It differs from classical information retrieval in that image databases are essentially unstructured, since digitized images consist purely of arrays of pixel intensities, with no inherent meaning. This paper retrieves the image by separating images into its three color components R, G and B and for that Discrete Wavelet Transformation is applied. Those transformed sub-bands are given to Histogram of Oriented Gradient (HOG) for extracting its feature vectors with Relevant Feedback technique is used. The performance of this approach is calculated by accuracy and it proved that nearly maximum accuracy is produced within reduced iteration.**

Keywords –Content-Based Image Retrieval (CBIR), Relevant Feedback, Discrete Wavelet Transformation (DWT), Histogram of Oriented Gradient (HOG).

I. INTRODUCTION

Technical challenges are increasing day by day because of large collections of image storing, managing and retrieving of images in computer systems.

CBIR differs from classical information retrieval. In that approach, image databases are primarily unstructured, since digitized pictures consist strictly of arrays of pixel intensities, with no inherent. One of the key problem associated with any reasonably image process is that they got to extract helpful information from the large volume of available information (such as recognizing the presence of explicit shapes or textures) before concerning whether the image's contents is feasible. Image databases therefore disagree basically from text databases, wherever the staple (words keep as ASCII character strings) has already been logically structured by [3]. There's no equivalent level of one retrieval during text information.

Image Retrieval System can be divided mainly into two types first is text based approach in this method images are first annotated with text and then text-based image retrieval uses traditional database techniques to manage images. During text descriptions, images can be organized by topical or semantic hierarchies to facilitate easy navigation and browsing based on standard Boolean queries. However, since automatically generating descriptive texts for a wide spectrum of images is not feasible, nearly all text-based image retrieval systems require manual annotation of images. Evidently, annotating images manually is a bulky and expensive.

Normally, image features can be moreover local or global [13]. The global features describe the visual content of the entire image. The retrieval systems based on global features cannot represent all the characteristics of the image. Thus, the global features are not suitable for tasks like partial image matching or searching for images that contain the same object or same scene with different viewpoints. In order to evade using global features, the interest point's detectors were introduced to represent the local features of images in image retrieval systems. The interest points are the salient image patches that contain rich local information about an image.[15,16]

Many researches proposed many algorithms based on color, texture and shape based CBIR. In this process new algorithm of DWT sub bands with HOG is proposed. Individuality in this paper is the image which first subjected to separation of components and then the features are extracted for all the three components.

The paper is organized as follows: Section II describes the related works involved in content based image retrieval, Section III describes the methodology used to retrieve the images, and Section IV describes the Experimental results obtained by using proposed methodology.

II. RELATED WORKS

Pass, Zabih and Miller [5] proposed a technique which splits a global image histogram into coherent and scattered components. The measure of color coherence identifies the existence of connected colored regions. While this technique improves on color histogram indexing, it does not support querying by the spatial locations of the color regions.

Jacobs, Finkelstein and Salesin [6] devised an image match criteria and system which uses spatial information and visual features represented by dominant wavelet coefficients. Their system allows the user to sketch example images and provides for improved matching over image distance norms, though, their technique provides for little edibility in specifying approximate and relative spatial information.

Sanjay et al. [6] put forth an image mining technique using wavelet transform. The author proposed an image mining approach using wavelet transform. It uses common pattern identical, pattern identification and data mining models with the intention that a real life scene/image can be associated to a particular category, assisting in different prediction and forecasting mechanisms. It is a three-step procedure i.e. image gathering, learning and classification. Since wavelet transform uses time frequency association, it can be utilized for image mining as a substitute of Fourier transform. Wavelet transform is utilized to decompose an image into dissimilar frequency sub bands and a small frequency sub band is used for

Principal Component Analysis (PCA). Classification assists in recognizing the category to which an image relates with. They have constructed a prototype system for identification using DWT + PCA system. The conception of image mining as a consequence can be competently used for weather forecasting so that one can know the natural disasters that may occur in advance.

Patil et al.,[7] put forwards an overview of the technical achievements in the research area of relevance feedback (RF) in content-based image retrieval (CBIR).It also covers the current state of art of the research in relevance feedback in CBIR, various relevance feedback techniques and issues in relevance feedback.

Recent works have shown that efficient and robust shape-based cues can be obtained from histogram of oriented gradient (HOG) in images [8].

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 a image region and the combined features are classified by using linear SVM. (Support Vector Machine) They showed that the grids of HOG descriptors significantly outer-performed existing feature sets for human detection. Keet al. [10] applied Principal Components Analysis (PCA) to reduce the dimensionality of the feature vectors and tested them in an image retrieval application.

III. METHODOLOGY

Image retrieval in a larger database is quite complex process. For this efficient features must be extracted from the images for its training.

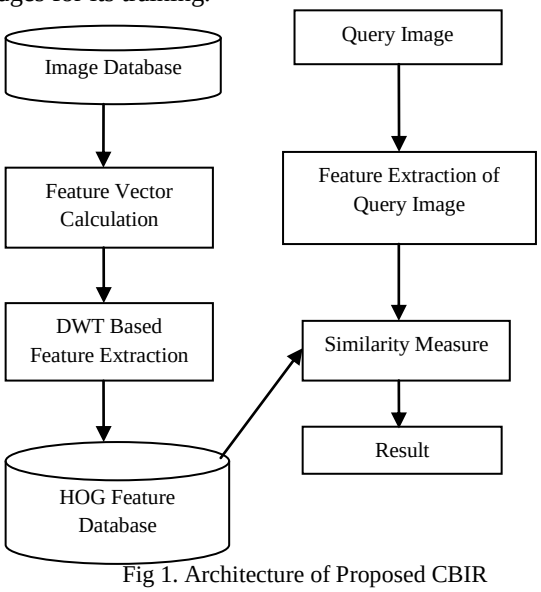


Fig 1. Architecture of Proposed CBIR

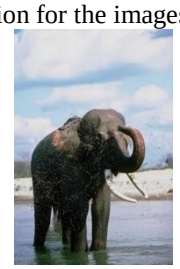
IV. *Feature Extraction:*

Each image has its own unique features which makes them different from other images.

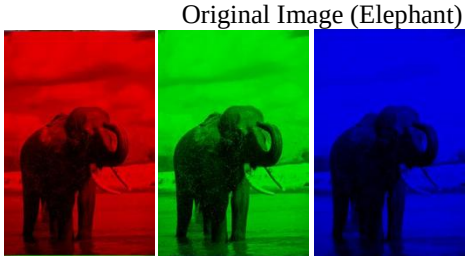
This proposed methodology utilizes two techniques.

1. DWT and
2. HOG

Based feature extraction for the images.



Original Image (Elephant)



Red Green Blue

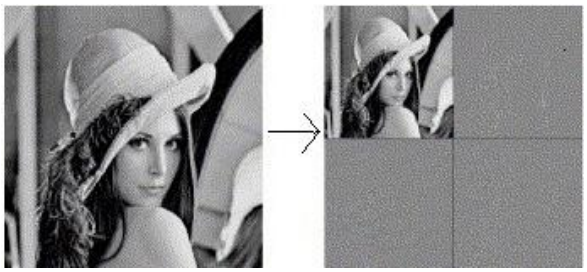
Fig2: Separation of R, G & B component

Each input image is separated into its three components Red, Green and Blue from the image. Sample of separation of R, G and B component is shown in above Figure 2. This image is subjected to DWT individually and features values are extracted.

V. DISCRETE WAVELET TRANSFORM:

Wavelet transform [1, 4, and 5] is a relatively recent signal processing tool that has been successfully used in a number of areas. Wavelet becomes more popular tool for image compression. Wavelets provide multi resolution capability, good energy compaction and adaptability to human visual system characteristics.

The discrete wavelet transform (DWT) is a linear transformation that operates on data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that separates data into different frequency components, and then studies each component with resolution matched to its scale.[16]



a) Input Image b) Single Level DWT

Fig 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.

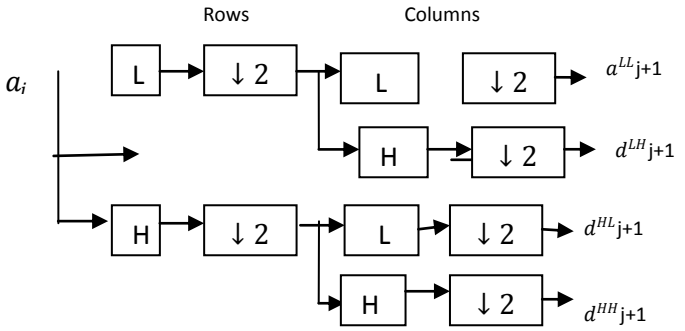


Fig 4: Wavelet decomposition for two-dimensional pictures

On scale j+1 there is only half from number of a and d elements on scale j. This causes that DWT can be done until only two a_j elements remain in the analyzed signal. These elements are called scaling function coefficients.

The main feature of DWT is multi scale representation of function. By using the wavelets, given function can be analyzed at various levels of resolution. The DWT is also invertible and can be orthogonal. Wavelets seem to be effective for analysis of textures recorded with different resolution.

Similarity comparison technique is used for good performance of the CBIR system and the better similarity in the query image and the images which are retrieved. For this Euclidean Distance measurement is used.

Euclidean Distance (q, p) = $\sqrt{\sum_{i=1}^{n}(q\hat{}i - p\hat{}i)^2}$

Where, $p = (p1, p2, \dots, pn)$ and $q = (q1, q2, \dots, qn)$

By this process feature vectors are extracted for the images then they are given to HOG based feature extraction.

VI. *Histogram of Oriented Gradient (HOG):*

The HOG are the orient descriptor used for extracting features for the images. The technique counts occurrences of gradient orientation in localized portions of an image.

HOG features have been initiated by NavneedDalal and Bill Triggs [2] who have developed and tested several variants of HOG descriptors, with differing spatial organization, gradient computation and normalization methods.

Now the image is 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 using the most common method is to apply the 1D centered point discrete derivative mask in both the horizontal and vertical directions. Specifically, this method requires filtering the grayscale image with the following filter kernels:

$$D_x = [-1\ 0\ 1]$$

$$D_y = [-1\ 0\ 1]^T$$

Next, for the given an image say I, then obtain the x and y derivatives using a convolution operation:

$$I_x = I * D_x \text{And} I_y = I * D_y.$$

The magnitude of the gradient is $|G| = \sqrt{I_x^2 + I_y^2}$

The orientation of the gradient is given by:

$$\theta = \arctan\left(\frac{I_y}{I_x}\right)$$

Finally, combing all the Gradient orientation Histogram of the interest point's area together to form a feature vector of size 128-dimension.

These features are extracted for each DWT subbandsare combined and formed a feature vector. This feature vector is taken for the relevance feedback process.

VII. *Relevance Feedback*

In relevance feedback-based approaches, a CBIR system learns from feedback provided by the user. Relevance feedback on the image retrieved in response to the initial query. This feedback is used subsequently in improving the retrieval effectiveness. [11, 12]

VIII. 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. Many researchers think that Corel database meets all the requirements to evaluate an image retrieval system, because of its large size and heterogeneous content.

The accuracy of the image can be calculated by the following formula which is expressed in percentage (%):

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

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.

A. *DWT Based Image Retrieval:*

The query input image is given by the user. Based on the query image the DWT based retrieval system extracts the images.

The input query is shown in the figure 5. Here the bus is taken as the query image.

Fig 5. Input Query (Building)

Result using DCT

Fig 6. Output image produced by using DWT image retrieval

The figure 6 shows the result based on DWT transform. In this figure two images are not extracted correctly for the query image building. The accuracy of the output images is 80%.

B. *HOG BY DWT SUB BANDS BASED IMAGE RETRIEVAL:*

Here the image searches from the database based on the combined algorithm of DWT and HOG based feature extraction.

The same query image of fig 5 building is given also for this process.

Result of DWT with HOG at first iteration:

Fig 6. Output image before relevance feedback using texture-based image retrieval

Figure 6 illustrates the output images based on DWT with HOG features for the user's input query. The output figure shows that for building query image one image of people is extracted by it. This became error image. This image has to be replaced. This can be done by using relevance feedback process.

The accuracy of the output images obtained in the first iteration is 88%.

Iteration 2 (After Relevance Feedback):

Then Figure 7 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.

Fig 7. Output image 1 after relevance feedback used

C. *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%.

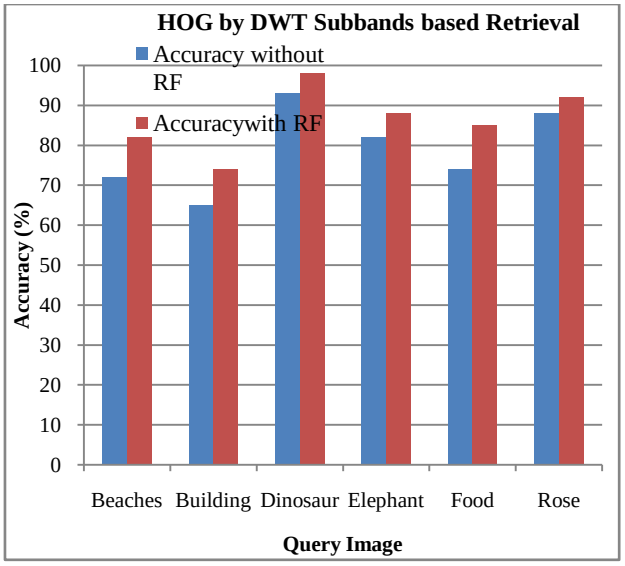


Fig 8. Accuracy Comparison of the HOG by DWT subbandsbased image retrieval before and after relevance feedback

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

IX. CONCLUSION

This paper proposed a method based onHOG by DWT Subbands to extract the features of the Image. The RGB image based features has some drawbacks in extracting the features, So they are separated into its components and from that features are extracted. This HOG descriptor is used to extract the features around the interest points. DWT is efficient in extracting subbands of images and extracts information from the images at different scales.Theextracted features of the database is compared with the query image and based on that image is retrieved.The performance of this is evaluated using accuracy in calculation. By the obtained result, this method can be concluded as a proficient method.

REFERENCES

[1] M[1] Corel 1000 and Corel 10000 image database. [Online]. Available: http://wang.ist.psu.edu/docs/related.shtml
[2] Dalal, N. Triggs, B.: Histograms of oriented gradients for human Detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005
[3] David G. Lowe. Distinctive image features from scale-invariant key Points. International Journal of Computer Vision, 60(2):91– 110, 2004
[4] Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2005)
[5] A. Halawani, A. Teynor, L. Setia, G. Brunner, and H.Burkhardt, "Fundamentals and Applications of Image Retrieval: An Overview", Presented at DatenbankSpektrum, 2006, pp.14-23..
[6] M. K. Mandal, T. Aboulnasr, and S. Panchanathan,, "Image Indexing Using Moments and Wavelets", IEEE Transactions on Consumer Electronics, Vol. 42, No. 3, August 1996V.Gudivada and V. Raghavan. Content-based image retrieval systems. IEEE Computer, 28(9):18 – 22, September 1995
[7] Pushpa B. Patil, Manesh B. Kokare(2011), "Relevance Feedback In Content Based Image Retrieval: A Review", Journal Of Applied Computer Science & Mathematics, Vol.5, No. 10 .[8]Rocchio JJ (1971), Relevance Feedback in Information Retrieval in the SMART Retrieval system.313-323. Prentice Hall, 1971.
[9] Santini, S And Jain, R C (1997) "The Graphical Specification Of Similarity Queries" Journal Of Visual Languages And Computing 7, 403-421
[10] SitaramBhagathy, kapil, KapilChhabra, "A Wavelet Based Image Retrieval System", Vision Research Laboratory Department of Electrical and Computer Engineering University of California, Santa Barbara, Project Report.
[11] Sanjay T. Gandhe, K. T. Talele and Avinash G. Keskar,(2007) "Image Mining Using Wavelet Transform," Knowledge-Based Intelligent Information and Engineering Systems, Springer link book chapter, pp. 797-803.
[12] Ke, Y., Sukthankar, R.: PCA-SIFT: A more distinctive representation for local image descriptors. In: Proc. of Computer Vision and Pattern Recognition, Washington, pp. 66–75 (2004).
[13] Samuel Rota Bulo, Massimo Rabbi and Marcello Pelillo, "Content-Based Image Retrieval with Relevance Feedback using Random Walks", Pattern Recognition, June 6, 2011.
[14] T.Tuytelaars and L.VanGool, "Content-based Image Retrieval based on Local AffinelyInvariant Regions".3rd Int. Conf. on Visual Information Systems, Visual99, Amsterdam, The Netherlands, 2-4 June 1999,pp.493-500
[15] C. Teh and R. T. Chin., "On image analysis by the method of moments", IEEE Trans. on Pattern Analysis and Machine Intelligence, 10(4):pp. 496-513, 1988

About The authors

Dr. Anna Saro Vijendran

Dr. Anna Saro Vijendran received her Ph.D. Degree in Computer Science from Mother Teresa Women's' University, TamilNadu, India, in 2009. She has 20 years of experience in teaching. She is currently working as the Director, MCAin SNR Sons College,Coimbatore, TamilNadu, India.

She has presented and published many papers in International and National conferences. She has authored and co-authored more than 30 refereed papers. Her professional interestsare Image Processing, Image fusion, Data mining and Artificial Neural Networks.

Mr.S.Vinod kumar

Mr. S.Vinod kumar is working as a Assistant Professor in PG computer Science at Sree Saraswathi Thyagaraja College and completed 8 years of Teaching profession. He has completed his M.Sc. M.Phil and has guided more than 5 M.Phil Scholars.

He has presented more than 10 papers in the national and international conferences. He is doing his PhD in the Image processing domain under the guidance of Dr. Anna Saro