



# An Innovative Technique for Content Based Image Retrieval Using Color and Texture Features

K.Valli Madhavi<sup>1</sup>, R.Tamilkodi<sup>2</sup>, R.Bala Dinakar<sup>3</sup>, K.JayaSudha<sup>4</sup>

Associate Professor, Department of MCA, GIET, Rajahmundry, India<sup>1</sup>

Associate Professor, Department of MCA, GIET, Rajahmundry, India<sup>2</sup>

Associate Professor, Department of MCA, GIET, Rajahmundry, India<sup>3</sup>

Associate Professor, Department of EXTC, SFIT, Mumbai, India<sup>4</sup>

**ABSTRACT** : The revolutionary internet and digital technologies have imposed a need to have a system to organize abundantly available digital images for easy categorization and retrieval. We determine relative importance of hue and intensity based on the saturation of an image with respect to the visual properties of the human eye. We effectively apply this method to the generation of a color histogram and use it for content based image retrieval applications. This paper presents the content based image retrieval using features like color and texture. The texture and color features are extracted through wavelet transformation and color histogram and the combination of these features is robust to scaling and translation of objects in an image. The experimental results show that this method is more efficient than the traditional CBIR method based on the single visual feature and other methods combining color and texture.

**Keywords:** CBIR, Texture, Wavelet Transform, HSV.

## I.INTRODUCTION

Multimedia data including images and videos have been dramatically increased in our life due to the popularity of digital devices and personal computers. Huge volumes of media data bring a lot of critical challenges of retrieving contents from large-scale multimedia databases effectively. A basic requirement of an image database is to perform content-based searches for images. The term Content Based Image Retrieval (CBIR) seems to have originated with the work of Kato [2] for the automatic retrieval of the images from a database, based on the color and shape present. Since then, the term has widely been used to describe the process of retrieving desired images from a large collection of database, on the basis of syntactical image features (color, texture and shape). The system first extracts and stores the features of the query image then it go through all images in the database and extract the features of each image. The results are the images that its features are most similar to the query image. Image Retrieval is a domain of increasing and crucial importance in the new information based society. Large and distributed collections of scientific artistic technical and commercial images are becoming a common ground thus requiring sophisticated and precise methods for users to perform similarity and semantic based queries. Reasons for its development are that in many large image databases, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming. These old methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete.

CBIR usually indexes images by low-level visual features such as color and texture. The visual features cannot completely characterize semantic content, but they are easier to integrate into mathematical formulations [3]. Extraction of good visual features which compactly represent a query image is one of the important tasks in CBIR. Color is one of the most widely used low-level visual features and is invariant to image size and orientation [1]. Each pixel of the image can be represented as a point in a 3D color space. Commonly used color space for image retrieval include RGB, HSV (or HSL, HSB), and opponent color space. However, one of the desirable characteristics of an appropriate color space for image retrieval is its uniformity [5]. Uniformity means that two color pairs that are equal in similarity distance in a color space are perceived as equal by viewers. In HSV (or HSL, or HSB) space is widely used in computer graphics and is a more intuitive way of describing color. The three color components are hue, saturation (lightness) and value (brightness). The hue is invariant to the changes in illumination and camera direction and hence more suited to object retrieval. RGB coordinates can be easily translated to the HSV (or HLS, or HSB) coordinates by a simple formula [6]. Since any pixel in the image can be described by three components in a certain color space (for instance, red, green, and blue components in RGB space, or hue, saturation, and value in HSV space), a histogram, i.e., the distribution of the number of pixels for each quantized bin, can be defined for each component. Clearly, the more bins a color histogram contains, the more discrimination power it has. When an image database contains a large number of images, histogram comparison will saturate the discrimination. To solve this problem, the joint histogram technique is introduced [7].



This paper is organized as follows, in this section 2, along with the analysis of color features; the quantization scheme for HSV color space and the proposed algorithm is presented. Section 3 presents the algorithm based on texture feature using wavelet in detail. Section 4 presents the performance evaluation as well as the similarity measurements. Section 5 presents the experimental results of the proposed algorithm and comparison with other algorithms. Therefore; we conclude remarks in section 6.

## II.COLOR QUANTIZATION

A color histogram represents the distribution of colors in an image, through a set of bins, where each histogram bin corresponds to a color in the quantized color space. A color histogram for a given image is represented by a vector:

$$H = \{H[0], H[1], H[2], H[3], \dots \dots H[I], \dots H[N]\}$$

Where i is the color bin in the color histogram and H[i] represents the number of pixels color i in the image, and n is the total number of bins used in color histogram. Typically, each pixel in an image will be assigned to a bin of a color histogram in which the value of each bin gives the number of pixels that has the same corresponding color. The normalized histogram H' is given as

$$H' = \{H'[0], H'[1], H'[2], H'[3], \dots \dots H'[I], \dots H'[N]\}$$

Where  $H'[I] = \frac{H[I]}{p}$ , p is the total number of pixels of an image [13].The block diagram for the proposed method is shown in Figure 1.

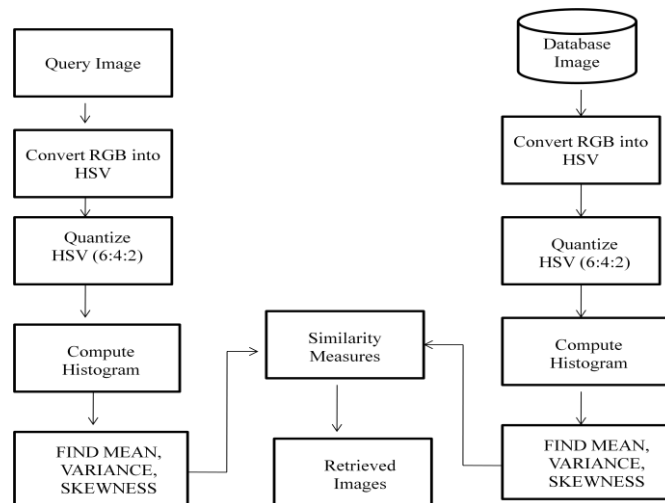


Fig.1 Block diagram of proposed color histogram without DWT

### A. Proposed algorithm for color histogram without DWT

Take the query image, convert from RGB color space to HSV color space. Quantize each pixel in HSV to bins then normalize it by dividing with total pixels, the query image is occupied. After calculating mean, variance and skewness for normalized histogram. Store the values in color feature vector.

The following steps are followed.

1. Read the input image
2. Convert RGB color space into HSV color space
3. Quantize each pixel in HSV space to 48 histogram bins.
4. The normalized histogram is obtained by dividing with the total number of pixels.
5. Calculate mean, variance and skewness for the Normalized histogram.
6. Store the values as color feature vector(C).
7. Repeat Steps 1 to 6 for the image present in the database.
8. Calculate the similarity measure of query image and the image present in the database
9. Repeat the steps from 7 to 8 for all the images in the database
10. Retrieve the images



### III.DISCRETE WAVELET TRANSFORMATION

Discrete Wavelet Transformation (DWT) [14] is used to transform an image from spatial domain into frequency domain. Wavelet transforms exact information from signal at different scales by passing the signal through low pass and high pass filters. Wavelets provide multi-resolution capability and good energy compaction. Wavelets are robust with respect to color intensity shifts and can capture both texture and shape information efficiently. It can be computed linearly with time and thus allowing for very fast algorithms. This paper uses Haar wavelets to compute feature signatures, because they are the fastest to compute and also have been found to perform well in practice [16]. Haar wavelets enable us to speed up the wavelet computation phase for thousands of sliding windows of varying sizes in an image. The proposed method finds effort to extract the primitive features of the query image and compare them to those of database images. The image features under consideration are texture by using the DWT concept. The block diagram for the proposed method is shown in Figure 2.

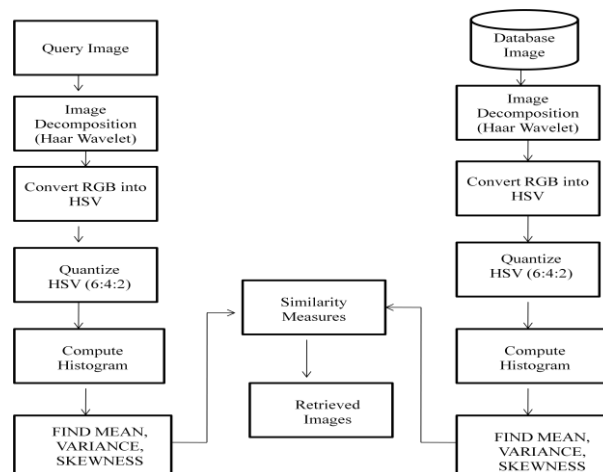


Fig.2 Block diagram of proposed method with DWT

#### A. Proposed Algorithm with Wavelet based color histogram

The wavelet transform provides a multi-resolution approach to texture analysis and classification. Studies of human visual system support a multi-scale texture analysis approach, since researchers have found that the visual cortex can be modelled as a set of independent channels, each tuned to a particular orientation and spatial frequency band. That is why wavelet transforms are found to be useful for texture feature extraction. The following steps are followed.

1. Extract red, green and blue components from an image.
2. Decompose red, green and blue component using haar Wavelet transformation at 3<sup>rd</sup> level to get approximate coefficient and vertical, horizontal and diagonal coefficients.
3. Combine approximate coefficient of red, green and blue component
4. Similarly combine the vertical, horizontal and diagonal coefficients of red, green and blue component
5. Covert the approximate, vertical, horizontal and diagonal coefficients into HSV plane
6. Quantize each pixel in HSV space to 48 histogram bins.
7. The normalized histogram is obtained by dividing with the total number of pixels.
8. Calculate mean, variance and skewness for the Normalized histogram.
9. Store the values as Texture feature vector (T).
10. Repeat Steps 1 to 6 for the image present in the database.
11. Calculate the similarity measure of query image and the image present in the database
12. Repeat the steps from 7 to 8 for all the images in the database
13. Retrieve the images

### V.PEFORMANCE EVALUATION

The performance of retrieval of the system can be measured in terms of its recall and precision. Recall measures the ability of the system to retrieve all the models that are relevant, while precision measures the ability of the system to



retrieve only the models that are relevant. It has been reported that the histogram gives the best performance through recall and precision value [17, 18]. They are defined as

$$\text{Precision} = \frac{\text{Number of relevant image retrieved}}{\text{Total number of images retrieved}}$$

$$\text{Recall} = \frac{\text{Number of relevant image retrieved}}{\text{Total number of relevant images}}$$

**A. Similarity Measurement**

Similarity measurement is a key to CBIR algorithms. These algorithms search image database to find images similar to a given query so that they should be able to evaluate the amount of similarities between images. To measure the similarity the Direct Euclidian Distance between the image in the database and the query image is given by

$$ED = \sum (V_{pi} - V_{qi}). (V_{pi} - V_{qi})$$

Where  $V_{pi}$  and  $V_{qi}$  be the feature vectors of image p and query image q respectively with respect to size 'n'.

**V.I.EXPERIMENTAL RESULTS**

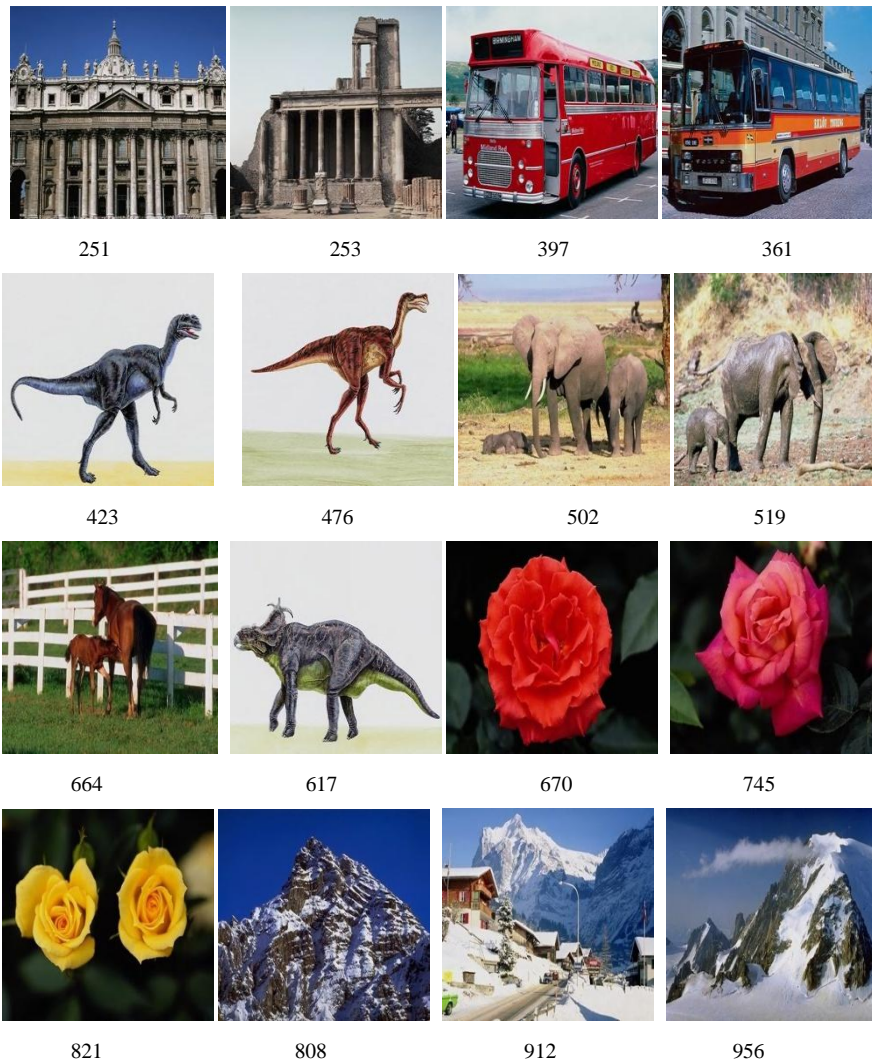




Fig3 Sample of WANG Image database



Fig 4 Query Image: Building

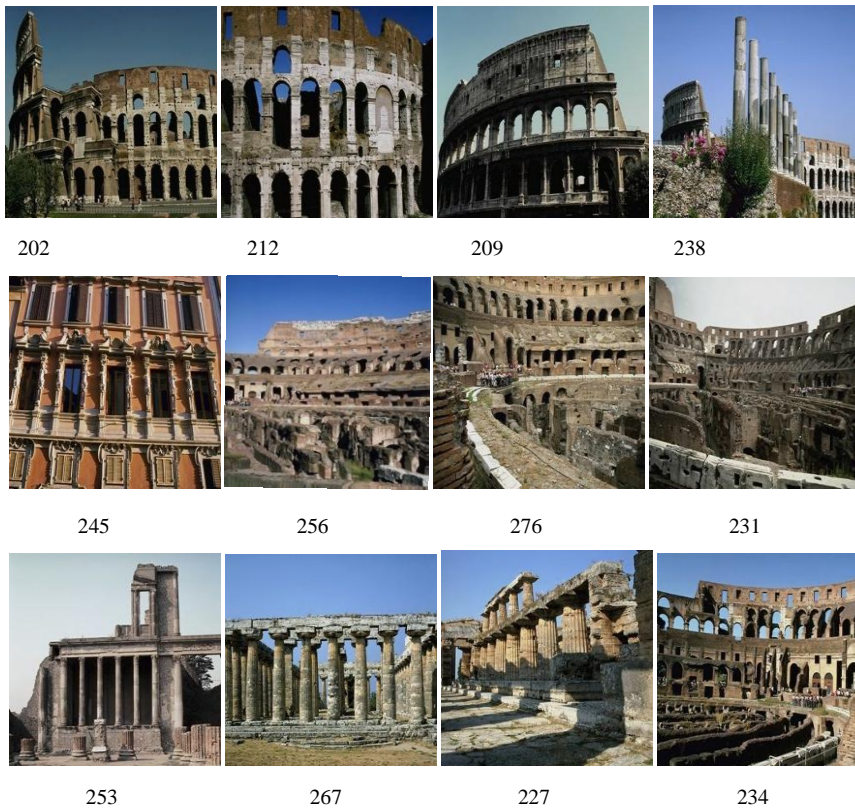


Fig5 Sample of WANG Image database



Fig 6 Query Image: Mountain

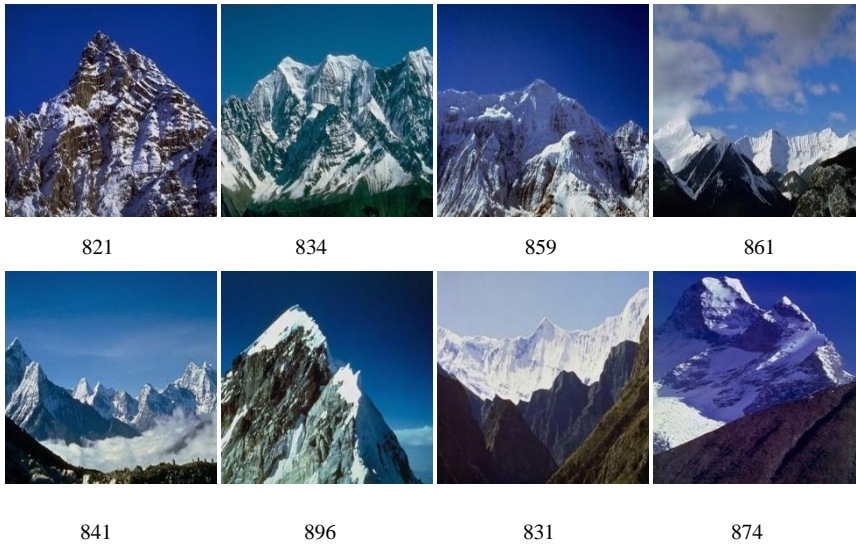


Table 1 Performance comparison of HSV & DWT Methods

Query Image	No.of Images Retrieved	HSV		DWT	
		No.of relevant	Precision (%)	No.of relevant	Precision(%)
Buses	20	17	0.850	19	0.950
Building	24	15	0.625	17	0.708
Flowers	27	22	0.815	24	0.889
Elephants	17	14	0.824	15	0.882
Mountains	34	29	0.853	31	0.912
Dinosaurs	19	14	0.737	17	0.895
Food	18	11	0.611	14	0.778
Beaches	26	18	0.692	19	0.731
African	30	19	0.633	24	0.800
Horses	28	20	0.714	22	0.786
Average			0.735	0.833	

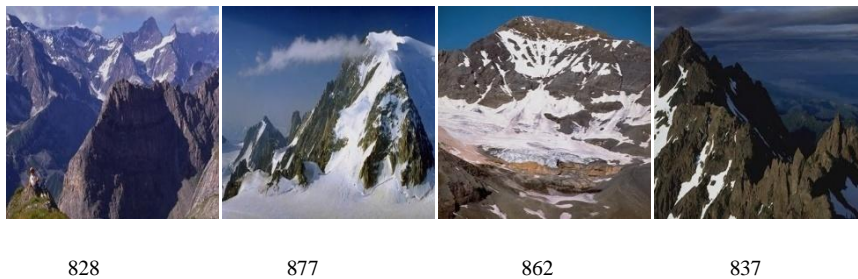


Fig 7 Query Image: Mountain



## VII.CONCLUSION

Automatic analysis and retrieval of images from a database is a challenging task. The difficulty arises from a number of issues, such as the complex mix of manmade and natural objects in an image. Most of the early studies on CBIR have used only a single feature among various color and texture features. However, it is hard to attain satisfactory retrieval results by using a single feature because, in general, an image contains various visual characteristics. This method provides a simple color based search in an image database for an input query image, using color and texture to give the images which are similar to the input image as the output. This experimental results show that the proposed method with DWT outperforms well when compared to normal histogram.

## REFERENCES

- [1] Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain, "Content-based multimedia information retrieval: State of the art and challenges," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 2, no. 1, pp. 1–19, 2006.
- [2] T. Kato, "Database architecture for content-based image retrieval", In Proceedings of the SPIE - The International Society for Optical Engineering, vol.1662, pp.112-113, 1992.
- [3] J. Vogel, B. Schiele, Performance evaluation and optimization for content-based image retrieval, *Pattern Recognition* 39 (5) (2006) 897–909.
- [4] Young Deok Chun, Nam Chul Kim, Ick Hoon Jang, Content-based image retrieval using multiresolution color and texture features, *IEEE Transactions on Multimedia* 10 (6) (2008) 1073–1084.
- [5] E.Mathias,"comparing the influence of color spaces and metrics in content based image retrieval," proceedings of international symposium on computer graphics,image processing and vision,pp.371-378,1998.
- [6] A.Blaser,databasetechniques for pictorial applications,lecturer notes in computer science,Vol.81,Springer Verlag GmbH,1979.
- [7] G.pass and R.Zabith,"Comparing images using joint hstograms,"*Multimedia systems*, Vol7, PP.234-240, 1999.
- [8] V.N.Gudivada and J.V.Raghavan, Special issue on content based image retrieval systems,*IEEE Computer Magazine* 28(9),1995.
- [9] A.Pentland and R.Picared, Special issue on digital libraries, *IEEE Trans. Patt.Recog. and mach. Intell.*, 1996.
- [10] A.D.Narasimhalu, Special issue on content based retrieval, *Multimedia systems*, 1995.
- [11] Special issue on Visual information management (R.Jain,Guest Ed.),*Comm.ACM*,Dec.1997.
- [12] B.Schatz and H.Chen,Buiding large scale digital libraries,*Computer*,1996.
- [13] J.R.Smith and S.F.Chang,"Tools an dtechniques for color image retrieval".in:IST/SPIE-Storage and retrieval for image and video databases IV,san Jose,CA,2670,426-437,1996.
- [14] R. C. Gonzalez, R. E. Woods and S. L. Eddins. *Digital Image Processing Using MALAB*, By Pearson Education, 2008.
- [15] A. Natsev, R. Rastogi and K. Shim, "WALRUS: A Similarity Retrieval Algorithm for Image Databases", In Proceeding. ACM SIGMOD Int. Conf. Management of Data, pp-395–406, 1999.
- [16] C. E. Jacobas, A. Finkelstein and D. H. Salesin, "Fast Multiresolution image querying", In Proc. Of SiGGRaPH 95, Annual Conference Series, pp-277-286, 1995.
- [17] G.Raghupati,R.S.Anandand M.L.Dewal,"color and texture features for content based image retrieval using motif", second international conference on multimedia and content based image retrieval, july21-23,2010.
- [18] T.Gevers,color in image database,intelligent sensory informationsystems,university of Amsterdam, the Netherlands 1998.

## BIOGRAPHY

K.Valli Madhavi is Head &Assoc. prof. in the Dept of Computer Applications, GIET,Rajahmundry. She has done MCA, M.Phil, and M.Tech. Her research area includes Data Mining & Image processing.

R.Tamilkodi is Assoc. prof. in the Dept of Computer Applications, GIET,Rajahmundry. She has done MCA,M.Tech.. Her research area includes Image processing & Pattern Recognition.

R.Bala Dinakar is Asst. prof. in the Dept of Computer Applications, GIET,Rajahmundry. He has done M.Sc.,M.Tech..

K.Jaya Sudha is Assoc. prof. in the Dept of EXTC, SFIT. She has done B.E, M.E. Her area of interest includes Signal processing and Web Mining.