Object based Image Retrieval from Database using Combined Features

H. Kavitha[1]
[1]Assistant Professor Dept. of ISE
Siddaganga Institute of Technology,
Tumkur, India.

M. V. Sudhamani[2, Ph.D]
[2]Professor and Head Dept. of ISE
RNS Institute of Technology,
Bengaluru, India.

ABSTRACT

Content based image retrieval (CBIR) is a promising way to address image retrieval based on the visual features of an image like color, texture and shape. Every visual feature will address a specific property of the image, so the state of the art focuses on combination of multiple visual features for content based image retrieval. This paper proposes a content based image retrieval system based on the combination of local and global features. The local features used are Bi-directional Empirical Mode Decomposition (BEMD) technique for edge detection and Harris corner detector to detect the corner points of an image. The global feature used is HSV color feature. For the experimental purpose the COIL-100 database has been used. The result show significant improvement in the retrieval accuracy when compared to the existing systems.

General Terms

Content based image retrieval, object recognition.

Keywords

CBIR, Harris corner detector, BEMD edge detection technique, HSV color features.

1. INTRODUCTION

The rapid growth of Internet and the falling price of digital cameras and image scanners have contributed to the rapid growth of digital images. This criterion has paved the way for the efficient storage and image retrieval systems. In the early days in 1970s the image retrieval was text based. Due to the complexity in the naming convention and also the manual effort behind annotating the images, this method lead to the development of content based image retrieval systems in early 1980s [11]. CBIR is a technique for retrieving the similar images to that of a query image by comparing the features of the query image with the features of the images in the database. These features can be two types local or global [5]. The global features refer to the visual content of the entire image. So the global features fail in describing the important features of the components of that image. This makes the global features unsuitable for retrieving images based on parts of whole image. This drawback of the global features is overcome by the local features like interest point detectors. These interest points are preferred since they provided locally rich information about the image. The work in [2] and [4] are based on interest points for image retrieval. The local invariant points were traditionally used for stereo matching initially. The Harris [3], Hessian [8], scale invariant, affine invariant are some of the algorithms based on interest point detectors. Paper [7] and [8] have proposed scale and affine-invariant interest points, which can deal with significant affine transformations and illumination changes. A recent work has combined the local and global features of an image, for the object recognition [9]. In this work the interest points are detected based on the robust and rotation invariant Harris detector. The edge of an image is derived by the BEMD algorithm. Color features of the images are extracted in the form of histograms.

The remainder of the paper is organized as follows: In Section II the general methodology of our proposed system is dealt with. The feature extraction techniques adopted are discussed in Section III. Section VI deals with the similarity measures used. In Section V the experimental results are shown and we provide the conclusions and future works in Section VI.

2. GENERAL METHODOLOGY OF THE PROPOSED SYSTEM

The general methodology adopted is focused in this section and is shown in Fig 1. Firstly we extract the features like edge feature using BEMD concept, corners of the objects by Harris corner technique and the HSV color features are extracted using the histograms. These features are stored in the database. This process is considered the offline phase of our work. In the real time phase the user provides the query image for which the BEMD edge, Harris corners and HSV color features are extracted. The three individual features of the query image are compared with the three individual features of the database image features. Then the features are combined together with some weightage and the images are ranked after sorting these combined feature. The most relevant images are displayed as the resultant images.
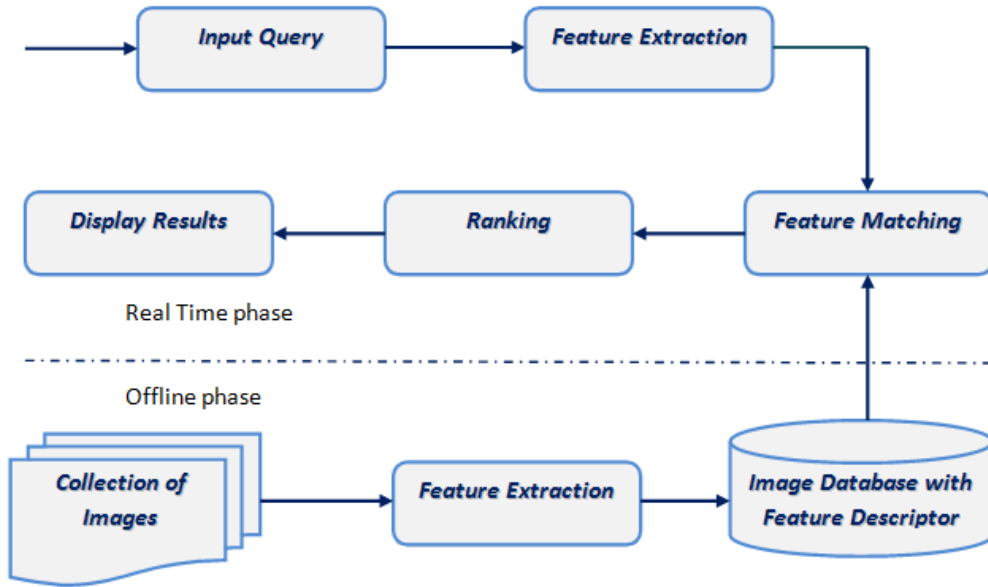
Fig1. General methodology

3.FEATURE EXTRACTION

This work uses three different features, namely Harris corner, HSV color and the BEMD edge features. In the following sections these techniques are discussed.

3.1. Color Features

One of the most commonly used visual feature for image retrieval is the color feature. There are many color models like RGB, HSV, HIS used to represent the color models. Since the HSV color space is more similar to human vision system, this colorimetric approach is been used. The global color histogram for extracting the color features has been adopted. The image in the general RGB (Red, Green, and Blue) color space is converted into the HSV (Hue, Saturation, and Value) color space. Further the HSV color space is quantized to form 11 bins each for the H, S, V components.

For similarity measure the Histogram Intersection Technique (HIT) has been adopted. When two histograms are compared using HIT, it returns the number between 0 and 1. 1 represents the most similar histograms and the similarity decreases as this value reaches 0. Let H_Q and H_B denote the histograms of the query image and an image in the database, respectively and S (H_Q, H_B) denote their similarity. Then S (H_Q, H_B) is given by the formula (1),

$$S (H_Q, H_B) = \frac{\sum_{x \in X, y \in Y, z \in Z} \min (H_Q(x,y,z), H_B(x,y,z))}{\min [\sum_{x \in X, y \in Y, z \in Z} H_Q(x,y,z), \sum_{x \in X, y \in Y, z \in Z} H_B(x,y,z)]} \quad (1)$$

here *X*, *Y* and *Z* are the arguments of the discretized color channels. This metric satisfies the associativity condition. Finally the distance $d_C(Q, B)$ between the query image and the database image according to the extracted color feature is given by the equation 2.

$$d_C (Q, B) = 1 - S (H_Q, H_B) \quad (2)$$

3.2. Harris Corner Detector

In the field of image processing interest point detection is one of the corner stone. Interest points are those points of an image where the intensity of an image will be high compared to other parts of an image [3]. Corner points are also a part of interest points. Harris and Stephens are the master minds behind the Harris corner detector. The Harris corner detector was found by Harris and Stephens, in 1988. The Harris corner detector concept was proposed as the interest points for the image retrieval by Schmid and Mohr. The core idea in the Harris corner detector is the usage of the auto correlation function to find the pixels were the signals change in the two directions. A matrix is constructed based on the auto correlation function as shown in equation 3, where σ_B denotes the derivation scale, σ_A denotes the integration scale, G the Gaussian and L stands for the Gaussian kernel smoothed image. This matrix is based on the signal's first derivatives. The auto correlation function's principal curvatures are the eigenvectors of the matrix. If both the eigenvector values are large and distinct, then there exists a corner. Fig 2. depicts the Harris corner detected for a frog image.

$$E(x, \sigma_A, \sigma_B) = \sigma_B^2 \, G(x, \sigma_A) * \begin{pmatrix} L_x^2(x,\sigma_B) & L_x L_y(x,\sigma_B) \\ L_x L_y(x,\sigma_B) & L_x^2(x,\sigma_B) \end{pmatrix} \quad (3)$$
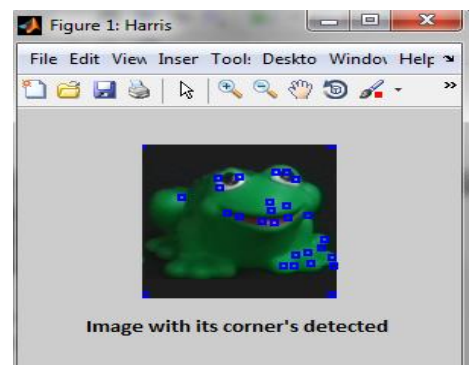

Fig. 2: **Harris corners detected for a frog image**

3.3. Bi-directional Empirical Mode Decomposition

Edge is an important feature in an image and carries important information about the objects present in the image. Extraction of edges is known as edge detection. Edge detection aims to localize the boundaries of objects in an image and significantly reduces the amount of data to be processed. In this work the edge detection is based on the BEMD technique [6]. During the decomposition of the image by the BEMD technique, the first Intrinsic Mode Frequency (IMF) is obtained and this gives the better characteristics of an edge. A clear edge of an image is obtained by processing the first IMF with a threshold. This process is known as shifting and is explained below:

Lets us assume X (t) be the original signal and
R (t) = X (t), k = 0 and i = 0

1. Compute the R (t)'s local maxima and minima.
2. The upper envelop Emax (t), lower envelop Emin (t) is computed by interpolating the maxima and minima respectively.
3. An approximation of the local average is computed and acts as the mean envelop.

$$M \text{ (t) }_1 = \frac{E_{max} \text{ (t)} + E_{min} \text{ (t)}}{2} \quad (4)$$

4. Increment i and the intermediate mode function is defined as:

$$P_i \text{ (t)} = T \text{ (t)} - M \text{ (t)} \quad (5)$$

5. The steps 1 to 4 are repeated until P_i (t) turns to be an IMF, after that the IMF is recorded.

$$C_I = P_i \text{ (t)} \quad (6)$$

6. Assign R (t) = R (t) - C_k (t). The shifting process is terminated based on the evaluation of the stopping condition, if not then k is incremented by 1, i assigned 0 and return back to step 1.

At the end of the shifting process, X (t) the original signal will take the form below:

$$X \text{ (t)} = \sum_{m=0}^{N}(C_n(t) + R_n) \quad (7)$$

The process discussed above is suitable for the single dimension signal. In case of two dimensions the stopping criteria for the shifting process is the standard division.

$$SD_k = \sum_{m=0}^{N} \frac{|P_{i-1} \text{ (m)} - P_i \text{ (m)} |^2}{P_{i-1}^2(m)} \quad (8)$$

Fig. 3: Ten Category of images used for the experiment

3.4. Extraction of the Edge Feature with Median Filtering

The edge feature extraction process goes as given below:
1. Convert the input image into gray scale image.
2. On the gray scale image apply the Histogram Equalization technique.
3. The empirical mode decomposition process detects the edge.
4. The Median filter of size 3 X 3 is used to smoothen the histogram equalized image.
5. Only the edge values of step 4 are replaced by the BEMD edge values.
6. Thus generated image is converted into the 64 bins feature vector and is put aside in the database for further usage.

4. THE SIMILARITY MEASURES USED

The corners of an image are extracted and this data is stored in the feature vector FH. Since the corners of each image differ from each other images the feature vector FH is of varying size. The edge features are extracted by applying the BEMD technique and the feature dimension is reduced to 64 bins. These 64 bins are stored in feature vector FE. For comparing the feature vectors, let's consider I as the image present in the database and Q is the query image. Then I (n) and Q (n) represent the average pixel values for the bins, with n ranging from 1 to 64. The difference between the values of the input image
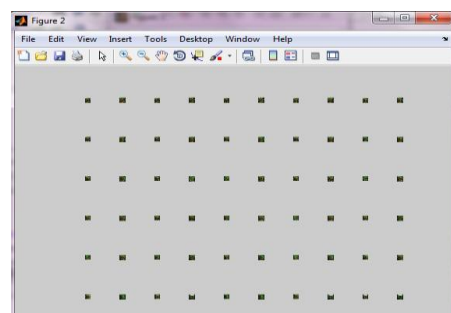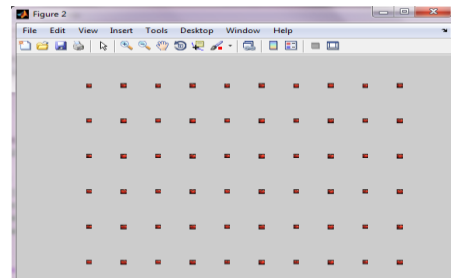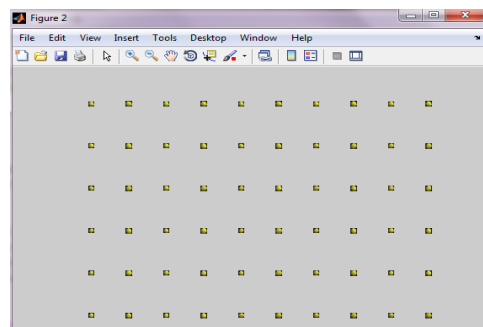

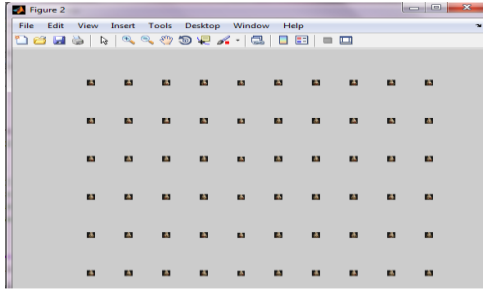**Fig. 4: First image is the query image, followed by retrieved images for a tomato and a mug**

Fig. 5: **First image is the query image, followed by retrieved images for a balloon and a vase**

and query image bins is calculated using the formula given below.

$$difference (n) = abs (I(n) - Q(n)) (9)$$
$$where\ n= 1,2, ….,64$$

This value is recorded in the feature vector FE. The HSV color features are extracted and quantized to form 11 bins for each component. The color features are stored in the FC feature vector. The final distance between the query and the input image is computed by the weighted sum of the three feature vectors FH, FC and FE, with the similarity measure given below.

$$SM = α.FE + β.FH + γ.FC$$

where the sum of α, β and γ will be 1.

The experiments were carried out with different values of α, β and γ. Good results were obtained with α = 0.5, β = 0.1 and γ = 0.4, since the edge feature is predominate in the case of scaling and rotation than the corners and also global color feature.

5. EXPERIMENTAL RESULTS

The Columbia Object Image Library (COIL-100) dataset was used for the experimental purpose. COIL-100 is a database of 7,200 color images of 100 objects. This corresponds to 72 different orientation of each object. The sample database of COIL-100 is shown in Fig. 2. In [6] the Harris edge detector and the BEMD techniques were used and the retrieval results were computed for each individual feature and the combination of these features. We have compared the current work's results with the results of [6]. Since the edges are more prominent than that of corners and color feature of an image. The edges are given more weight compared to that of color and corners features. The combination of Harris Corner (HC), BEMD and the color features of the query image and the above three features of the database images were compared and the resultant images were displayed. The precision and recall values of the retrieved images for the existing system [6] and the proposed work are tabulated in Table 1. The Fig. 4 shows the input query images and the resultant images of the tomato and a mug. Similarly the Fig.5 is depicting the balloon and vase query images and their corresponding retrieved images. The results prove that the combination of HC, color and BEMD technique are substantially better than the combination of HC and BEMD techniques. Fig. 6 depicts the precision and recall of the proposed and existing systems in the graphical form.

Table 1: Precision and Recall for the existing and the proposed system

Category of Image	Existing System		Proposed System	
	Precision	Recall	Precision	Recall
Category1	99.17	82.64	99.17	82.64
Category2	69.50	57.92	70.17	58.47

Category3	98.83	82.36	99.50	82.92
Category4	43.00	35.83	45.17	37.64
Category5	97.17	80.97	98.83	82.36
Category6	100.00	83.33	100.00	83.33
Category7	100.00	83.33	100.00	83.33
Category8	67.33	56.11	69.67	58.06
Category9	47.33	39.44	49.83	41.53
Category10	100.00	83.33	100.00	83.33
Average	**82.23**	**68.53**	**83.23**	**69.36**

6. CONCLUSIONS

In this work the experiments have been carried out with the combination of local and global features. The experiments were applied on the ten categories of images each with seventy two different orientations from COIL-100 image database. The average precision and recall value of the current system and the previous system are
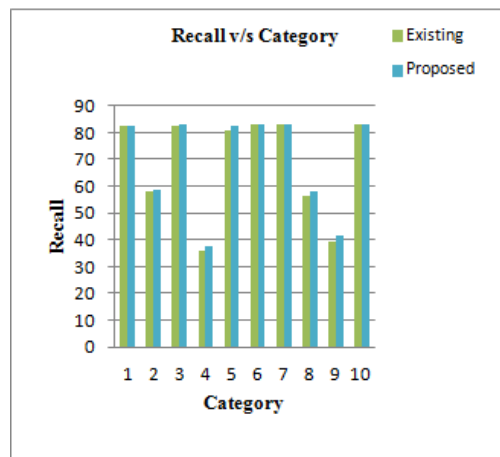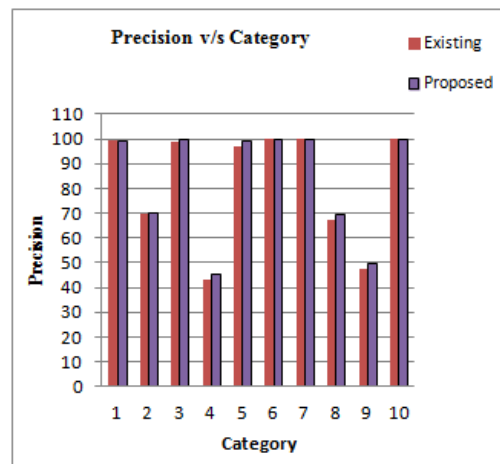

Fig. 6 Graphs showing the precision and recall of the existing and proposed systems respectively

tabulated in table 1. The BEMD and Harris corner features were taken into consideration with the weightage of .5 for each in the previous work. In the current work the above mentioned two features along with the HSV color features have been considered. The results show significant improvement over the existing systems. The future work includes database population and the consideration of the other features of the images to enhance the retrieval process, which is the focus of current research work.

7. REFERENCES

[1] Veltkamp, R. C. and Tanase, M. 2004. Content-based image retrieval systems: A survey, Utrecht University Tech. Rep., UU-CS-2000-34.

[2] Schmid, C. and Mohr, R. 1997. Local gray value invariants for image retrieval. In Proceedings of the IEEE Transaction on Pattern Analysis and Machine Intelligence, 19 (5), 530-534.

[3] Harris, and Stephens, M. J. 1988. A combined corner and edge detector. In Proceedings of the Alvey Vision Conference, 147- 152.

[4] Tuytelaars, T. and Van Gool, L. 1999. Content-based Image Retrieval based on Local Affinely Invariant Regions. In Proceedings of the 3rd International Conference on Visual Information Systems, Visuals99, 493-500.

[5] Kavitha, H. and Sudhamani, M. V. 2012. Content Based Image Retrieval - A survey. In Proceedings of the National Conference on Trends in Advanced Computing and Information Technology, 49-56.

[6] Kavitha, H. and Sudhamani, M. V. 2013. Image Retrieval based on object recognition using the Harris corner and edge detection Technique. In Proceedings of the International Conference on Communication, VLSI & Signal Processing, 181-184.

[7] David Lowe, G. 2004. Distinctive image features from scale-invariant keypoints. In Proceedings of the International Journal of Computer Vision, 60, 91-110.

[8] Mikolajczyk, K. and Schmid, C. 2004. Scale and affine invariant interest point detectors. In Proceedings of the International Journal of Computer Vision, 1, 63-86.

[9] Muralidharan, R. and Chandrasekar, C. 2012. Combining local and global feature for object recognition using SVM-KNN. In Proceedings of the Pattern Recognition, Informatics and Medical Engineering (PRIME), 1-7.

[10] Dubuisson, S. 2010. The computation of the Bhattacharyya Distance between Histograms without Histograms. In Proceedings of the IPTA, 373–378.

[11] Smeulders, A. W. Worring, M. Santini, S. Gupta, A. and Jain, R. 2000. Content - based image retrieval at the end of the early years. In Proceedings of the IEEE Trans. Pattern Anal. Mach. Intell. 22, 12, 1349–1380.

[12] Samia Omar, G. Mohamed Ismail, A. Sahar Ghanem, M. 2009. WAY-LOOK4: A CBIR system based on class signature of the images color and texture features. In Proceedings of the IEEE/ACS International Conferences on Artificial Intelligence, 464-471.

[13] Aradhana Katare. Suman Mitra, K. Asim Banerjee. 2007. Content Based Image Retrieval System for Multi Object Images Using Combined Features. In Proceedings of the International Conference on Computing: Theory and Applications (ICCTA'07), 595-599.

[14] Huijsmans, D. P. and Sebe, N. 2005. How to complete performance graphs in content-based image retrieval: Add generality and normalize scope. In Proceedings of the IEEE Trans. Pattern Anal. Mach. Intell. 27, 2, 245–251.

[15] Shrinivasacharya, P. Kavitha, H. and Sudhamani, M. V. 2011. Content Based Image Retrieval by Combining Median Filtering and BEMD Technique. In Proceedings of the International Conference on Data Engineering and Communication Systems (ICDECS 2011), 2, 231-236.

[16] Kavitha, H. and Sudhamani, M. V. 2012. Edge Detection of an Image based on Bi-Level Histogram Equalization and Ant Colony Optimization Technique. In Proceedings of the International Conference on Computer Technologyand Science, 128-135.

[17] Kavitha, H. and Sudhamani, M. V. 2012. A Survey of clustering algorithms for CBIR systems. In Proceedings of the National Conference on Computer Networks and Soft Computing, 53-56.

[18] Flickner, M. Sawhney, H. Niblack, W. Ashley, J. Huang, Q. Dom, B. Gorkani, M. Hafner, J. Lee, D. Petkovic, D. Steele, D. and Yanker, P. 1995. Query by image and video content: The QBIC system. In Proceedings of the IEEE Comput. 28, 23–32.

[19] David G. Lowe. 1999. Object recognition from local scale-invariant features. In Proceedings of the International Conference on Computer Vision, 1150-1157.

[20] Meirav Adoram. and Michael S. Lew. 1999. IRUS: Image Retrieval Using Shape. In Proceedings of the Multimedia Computing and Systems IEEE International Conference on Digital Object Identifier, 2, 597-602.