

An Integrated approach to CBIR using multiple features and HSV Histogram

N.Neelima^{#1}, E.Sreenivasa Reddy^{#2}, Kalpitha.N^{#3}

^{#1}Department of Electronics & Communications Engineering, Jain University, Bangalore, India

^{#2}Department of Computer Science Engineering, Acharya Nagarjuna University, Guntur, India.

^{#3}Department of Computer Science Engineering, Jain University, Bangalore, India.

¹ neelima.niz@gmail.com

² esreddy67@gmail.com

³ kalpitha.n@gmail.com

Abstract— An efficient search for semantically relevant images has always been thirst in computer vision and processing specially in large scale image retrieval. We propose an integrated approach for fast and effective image retrieval system using multiple features and hsv histogram. Relevance feedback allows user interaction to improve the performance. Features used in this work are improved lbp and modified fourier descriptors, plays vital role in effective retrieval. Experimental results on CALTECH-101 and MPEG CE shape 1 datasets proves that our framework provides better retrieval efficiency compared to the state of art methods.

Keyword- semantic retrieval, extended lbp, modified fourier descriptors, similarity distance.

I. INTRODUCTION

CBIR is well known as Content Based Image Retrieval is one of the hot thirst areas in computer vision and processing. Several applications involves such as medical diagnosis, document analysis, space applications and entertainment. QBIC (Query By Image Content) also has wide applications in semantic image retrieval for art collections, galleries, logo search etc. These systems focus on the original content and the semantic gap between the images. This content is described by shape, texture and color features. Extraction of features and measuring the relevance between the features gives the efficient retrieval. This can be used in medical image retrieval for treatment of the patients who suffers with similar type of problem. Rouhollah Rahmani et.al proposed [2] localized content-based image retrieval, where the user is only interested in a portion of the image, and the rest of the image is irrelevant. In this they used multiple-instance learning algorithm to identify the desired object and the feature weights. Then the images are ranked in the database using a similarity measure that is based upon only the relevant portions of the image. Relevance feedback is a query modification technique which attempts to capture the user needs through iterative feedback and query refinement [3]. Texture analysis and retrieval has gained wide attention in the field of medical, industrial, document analysis [4]. Santosh kumar et.al proposed a descriptor which includes both color and texture information. The individual R, G and B channel wise directional edge information between reference pixel and its surrounding neighbourhoods are extracted by computing its grey-level difference based on quinary value (-2, -1, 0, 1, 2) instead of binary and ternary value in 0°, 45°, 90°, and 135° directions of an image [4]. Xiang [6] proposed the retrieval system with color, texture and shape features. They used pseudo-zernike moments for extracting the shape and texture feature were extracted by steerable filters. But the proposed method is more superior as it uses generic fourier descriptor to extract the shape features. Ammar huneti [7] proposed another retrieval system using color and texture features. These features were extracted by using discrete wavelet transforms and trained by using self-organizing maps. But in some of the applications like medical images and textures color may not be the appropriate choice to retrieve the similar images. Youngeun An [8] proposed the system with color histogram quantization in HSV space. In this the input image is segmented and from each segment the local histogram is calculated by using maximum color occurrence. Kazuhiro Kobayashi [11] proposed the system in which features are extracted from both spatial and frequency domain. In this LBP is used as spatial domain and wavelet transform is used to extract the features in frequency domain.

II. PROPOSED FRAMEWORK

The proposed framework consists of three steps. Preparing the feature database, Extracting features from the image and comparing the relevance between the features. The content of the images is described by its features. Color feature is selected only when the images are rgb images. For the rest we considered shape and texture features to retrieve the relevant images. Fig.1 shows the proposed system frame work. If the input images are rgb images along with the shape and texture hsv histogram is also calculated otherwise shape and texture features are extracted and the feature vectors are stored individually. This is done for the entire images in the database and the feature vectors database is prepared. This process is done in offline. When the user enters the query image the same features are calculated and these features are compared with the features in database. According to the similarity distance all the images were ranked and displayed in the ascending order.

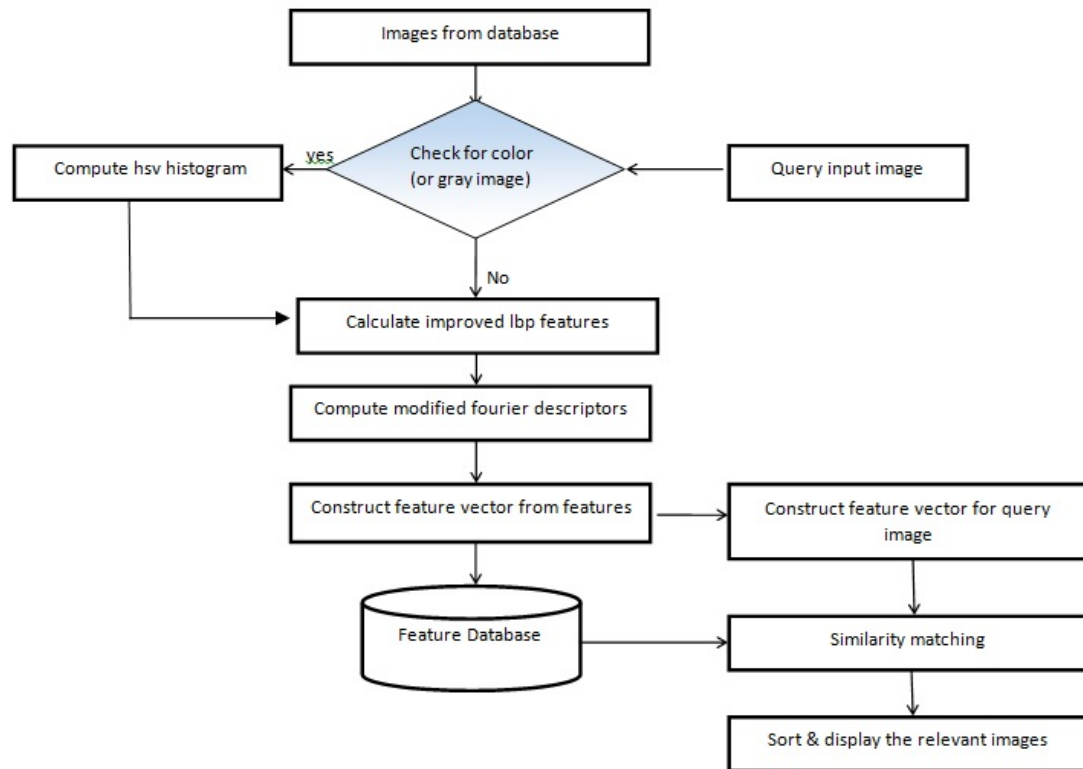


Fig.1. Block diagram of the Proposed Image retrieval system

A. Shape features

The shape of the objects in images gives important information in retrieval. Fourier descriptors are one of such shape feature which is fast and effective. Wide variety of shape descriptors exist, classified into contour based and region based descriptors. Contour based descriptors give the boundary information in detail but they fails in case of disjoint shapes. Region based descriptors use the pixel information with in the entire region. Several descriptors such as hu moments, zernike moments and legendre moments were discussed. Shape variations are represented by the sample points along the three most dominant principle axes in the feature space [9]. Zernike moments gives efficient image retrieval, but it has two disadvantages, the first one is it has redundant features at each repetition order. The second one is it cannot compute features in radial directions. Fourier descriptors are powerful features for the recognition of two-dimensional connected shapes as well as for broken shapes [10]. So we propose a Modified Fourier Descriptor (MFD), which provides multi resolution analysis. In general, FT (Fourier Transform) is used in shape analysis. But the features which are calculated from FT are not rotation invariant. In order to achieve the rotation invariance the polar form is used. For the given image $f(x,y)$ the modified polar transform is given by eq. (1).

$$MPF(\rho, \phi) = \sum_r \sum_t f(r, \theta_t) \exp [j2\pi(\frac{r}{R}\rho + \frac{2\pi t}{T}\phi)] \quad (1)$$

Where ρ and ϕ are radial and angular frequencies and

$$r = [(x - x_c)^2 + (y - y_c)^2]^{1/2}, \quad 0 \leq r < R, \quad \theta_t = t(\frac{2\pi}{T}) \text{ where } 0 \leq t < T, \text{ and } 0 \leq \rho < R. \quad (2)$$

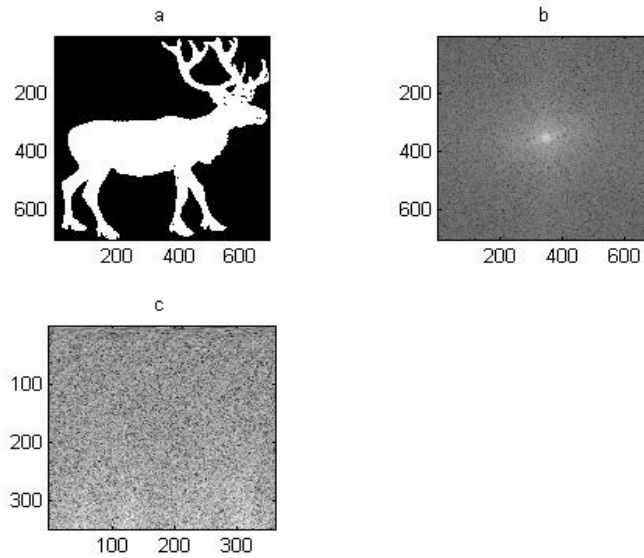


Fig. 2. (a) input image (b) FFT of the input image (c) polar FFT of the input image

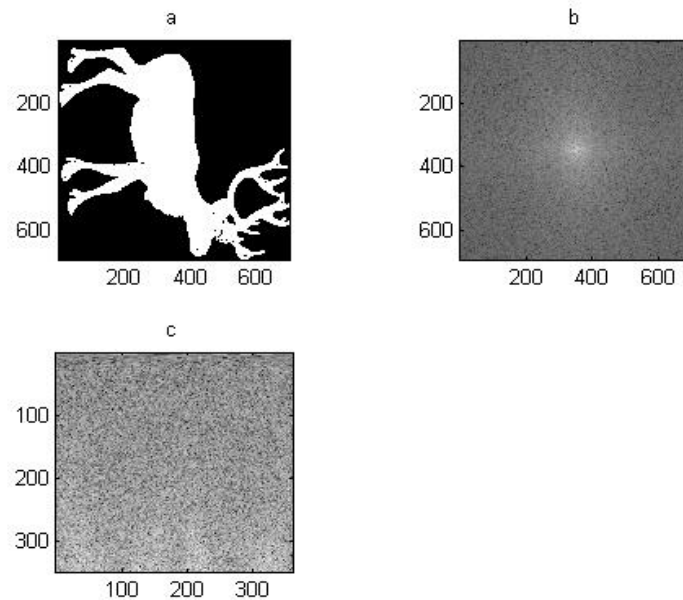


Fig.3. (a) rotated input image (b) FFT of the rotated image (c) polar FFT of the rotated image

The radial and angular resolutions are denoted as R & T, given by eq. (2). To achieve the translation invariance the first coefficient can be discarded. For rotation and scale invariance the following normalization is performed. To achieve scale invariance all the coefficients are divided by the absolute value of non-zero coefficient.

The MFD (Modified Fourier Descriptor) is given by eq. (3)

$$MFD = \left\{ \frac{|MPF(0,0)|}{area}, \frac{|MPF(0,1)|}{|MPF(0,0)|}, \dots, \frac{|MPF(0,m)|}{|MPF(0,0)|}, \frac{|MPF(m,0)|}{|MPF(0,0)|}, \dots, \frac{|MPF(m,n)|}{|MPF(0,0)|} \right\} \quad (3)$$

Where area is the area of the boundary with in the shape resides, m is the maximum no of radial frequencies and n is the no of angular frequencies selected. For efficient implementation few MFD features can be used. In our framework we used 36 features. The similarity comparison between the features is done by using euclidean distance. Fig. 2(a) shows the query image and Fig.2. (b) & (c) represents the corresponding FFT and the polar FFT forms. Fig.3.(a) shows the rotated version of the query image and Fig.3. (b) & (c) represents the corresponding FFT and polar FFT forms. It is clearly observed that the FFT is different for the two images, but there is no difference in the polar forms. So the rotational invariance is achieved by using the proposed modified fourier descriptor.

B. Texture Feature extraction

The texture is a powerful low-level feature for image search and retrieval applications [1]. In the field of computer vision, Local binary patterns are one of the best texture based visual descriptor. LBP is used to analyse the texture of images in the dataset and retrieve the same more appropriately. The computation of an LBP can be described as; the labels for each image pixels are formed by thresholding the surrounding 3 * 3 pixels with the centre value in order to generate the binary number to each pixel of an image. Finally the histogram plotted for all 256 pixels is obtained for texture description and analysis. The efficiency of the original LBP operator is enhanced by two main extensions namely discriminative capability and its robustness. The local structures of an image are defined by LBP.

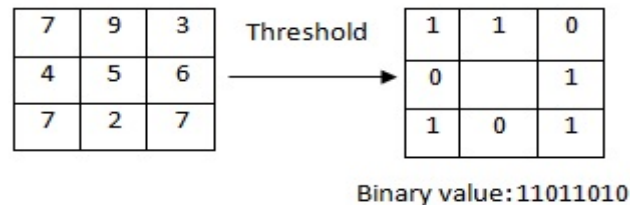


Fig. 4. The basic operation of LBP operator

Fig.4 describes the general LBP operator on a 3*3 neighbourhood. For a given pixel a 3*3 neighbourhood is considered and each pixel intensity value is subtracted from the centre value and the positive values are encoded as '1' and the negative values are encoded as '0'. All these binary values are considered clockwise and the resulting pattern is known as LBP codes. Fig.5 shows the next modification, ILBP is improved by operating on neighbourhoods of different radius from the circle. The notation in Fig.5 represents (P,R) where P denotes the no. of neighbourhood points on a circle of radius R. The ILBP (Improved LBP) operator makes the comparison among all the pixels with their mean intensity values. This is given in Fig.3. This above mention calculation was performed only on 3x3 matrixes. To improve the performance and making rotational invariant the Modified LBP operator is proposed in this paper. In Fig. 6, ILBP is represented as the modified version of original LBP operator. It performs the comparison among the neighbourhood pixels and central pixel along with the calculation of grey-value differences. Fig. 7 represents the modified LBP (MLBP) operator. In this the feature extracted consists of several layers of LBP codes along with the grey-value differences between the neighbouring and central pixels.

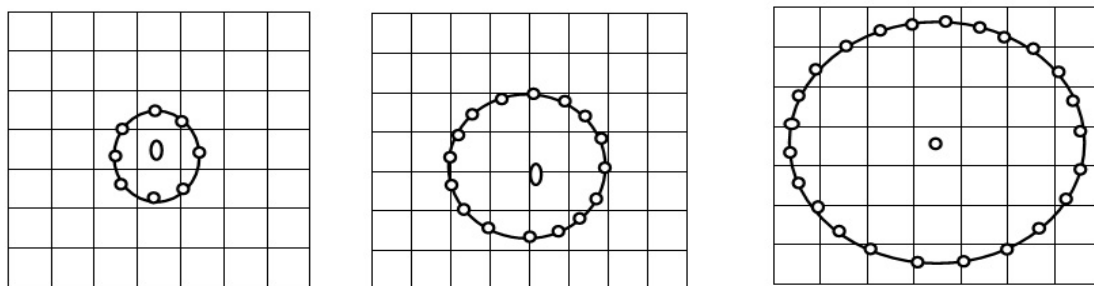


Fig. 5. Examples of ELBP operator with different radius and sampling points (8,1), (16,2) and (24,3)

Firstly all the grey-scale difference values are encoded to its binary values representation. Later all the binary values formed in a layer expresses as modified LBP. Hence, compared to first LBP layer the information encoded in additional layers tends to be more discriminative. This distinguishes the robustness of the operator. MLBP highlights the dimensionality of the feature to a greater extent.

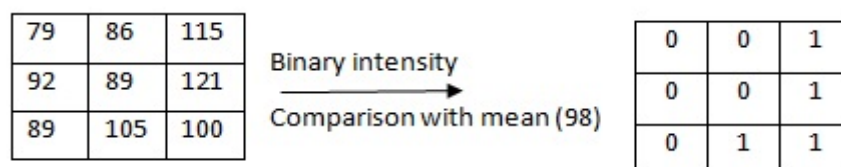


Fig. 6. ILBP operator example

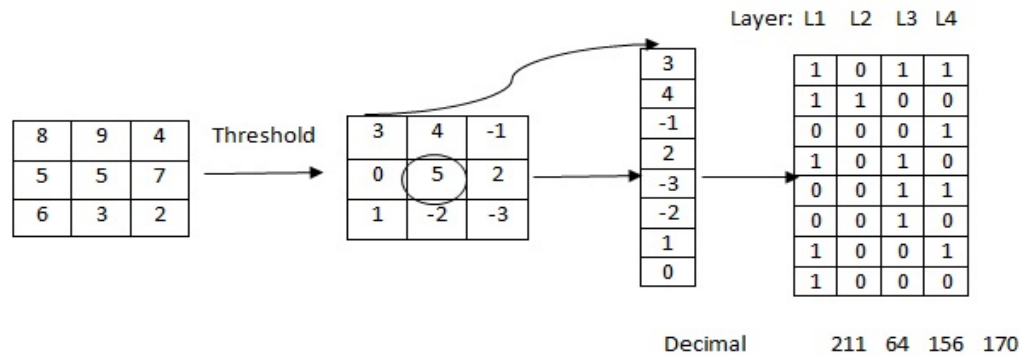


Fig. 7. The example of MLBP (Modified LBP) operator

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (4)$$

For a given pixel(x,y) in the given input image the LBP is calculated by eq.(4). where gp is the gray level intensity of the particular pixel and gc is the gray level value with in the 3*3 neighbourhood of the centre pixel. The rotational invariance can be achieved by eq. (5).

$$LBP_{P,R}^i = \min\{ROR(LBP_{P,R}, i)\} \quad i = 0, 1, \dots, P-1. \quad (5)$$

Where ROR performs bit wise circular right shift in the specified neighbourhood. The features from the rotation invariant lbp are stored in the feature vector.

C. HSV histogram

Hsv histogram is the feature which is used for the color image database. The rgb images from the database are converted into hsv space by using the following expressions.

$$H = \cos^{-1} \frac{1/2[(R-G)+(R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \quad (6)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \quad (7)$$

$$V = \frac{1}{3} (R + G + B) \quad (8)$$

The H,S,V values are normalized into [0,1] scale by using eq.(9) and (10) and quantized. The number of quantization levels for H,S and V are 8, 4 and 4 respectively.

$$H = \left(\left(\frac{H}{255} \right) * 360 \right) \bmod 360 \quad (9)$$

$$V = V/255 \text{ and } S = S/255. \quad (10)$$

The quantized histogram values of the query image are compared with database images by using chi-square distance. All the three features are stored individually and the similarity matching distance is calculated. The similarity matching for the shape and texture features is done with euclidean distance. The final distance with which the images are sorted is given by $D_{final} = D_{mfd} + D_{lbp} + D_{hsv}$. With this distance the images are indexed in ascending order and displayed.

III. SIMILARITY MATCHING

The similarity matching between the features of query image and the database images is calculated by using the following distance metrics.

Chi-square distance:

The chi-square distance between the feature vectors of query image (Qi) and database images (Dbi) is given by eq. (11).

$$\text{Chi-square } (Q_i, D_{b_i}) = \frac{1}{2} \sum_{j=1}^{n_i} \frac{(x_j - y_j)^2}{(x_j + y_j)} \quad (11)$$

Where $Q_i = [Q_1, Q_2, Q_3, \dots, Q_i]$ $D_{b_i} = [D_{b_1}, D_{b_2}, \dots, D_{b_i}]$

Euclidean distance:

The euclidean distance between the feature vectors of query image (Qi) and database images (Dbi) is given by eq. (12)

$$\text{Euclidean } (Q_i, D_{b_i}) = \sqrt{\sum_{j=1}^n (Q_i - D_{b_j})^2} \quad (12)$$

Where $Q_i = [Q_1, Q_2, Q_3, \dots, Q_i]$ $D_{b_i} = [D_{b_1}, D_{b_2}, \dots, D_{b_i}]$

IV. RESULTS

The proposed framework is tested with two different databases CALTECH-101 and MPEG CE-shape 1. MPEG-7 CE shape 1 dataset consists of 70 different categories of images each with 20 similar images. Fig. 8. shows the top ten retrieval results for different classes of images. The experiments were conducted by taking 20, 30 and 50 different classes randomly from the dataset. The precision is calculated by using the following equation.

$$\text{Precision} = \frac{\text{No. of similar images retrieved}}{\text{Total no. of images retrieved}}$$

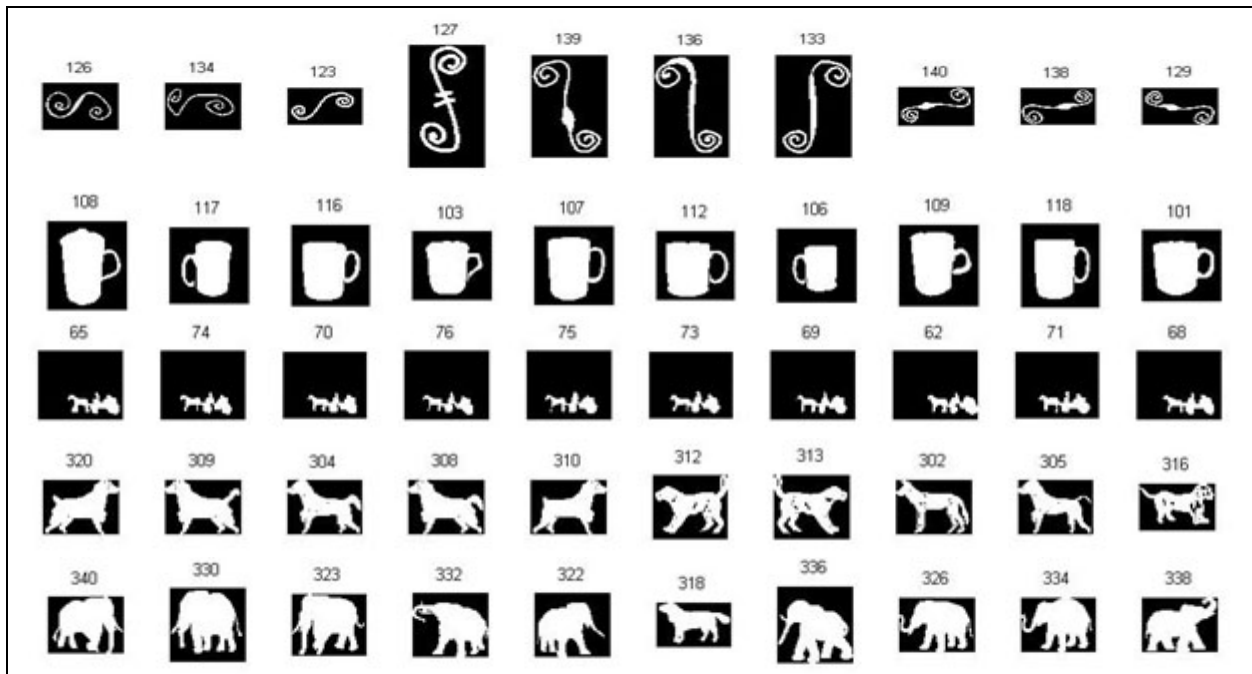


Fig.8. Top ten retrieval results for different categories of query images from MPEG-7 CE shape 1 dataset

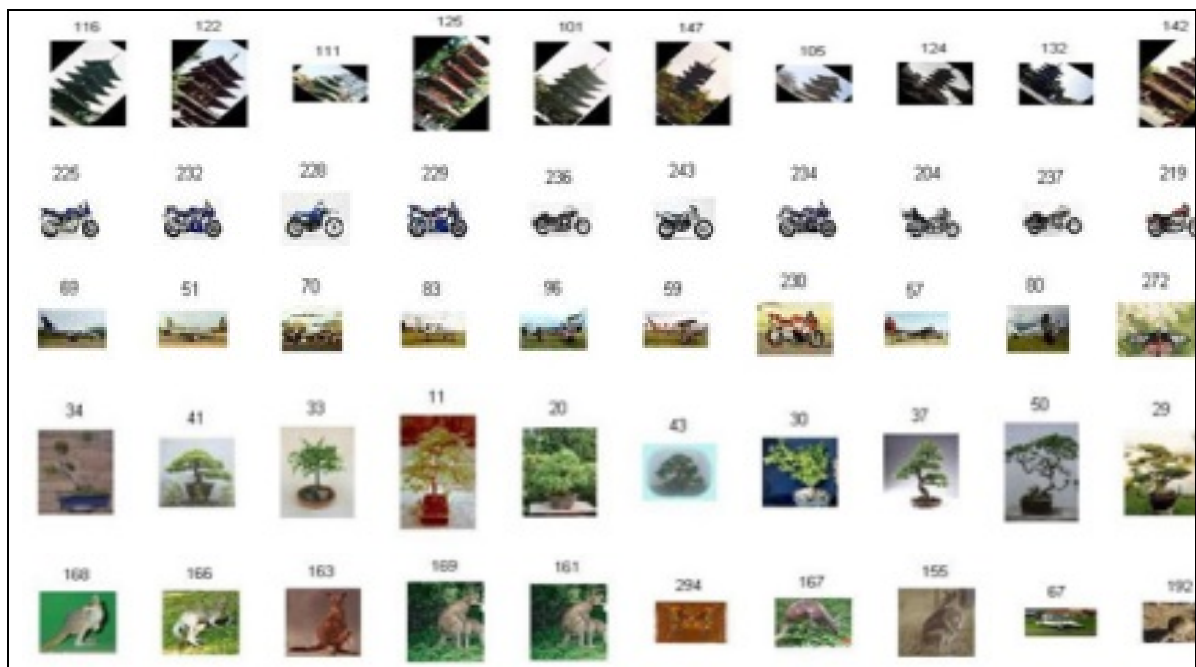


Fig.9. Top ten retrieval results for different categories of query images from CALTECH-101 dataset

CALTECH-101 dataset has 101 different categories of images each with more than 50 similar images. The database is prepared by selecting 30 different classes of images each with 40 similar images. The relevant images retrieved for various categories of images are shown in Fig. 9. Fig.10 shows the average efficiency of the proposed framework for different classes of CALTECH dataset with different number of images. The maximum retrieval efficiency for CALTECH-101 dataset is 84% and for the MPEG-7 dataset is 92%.

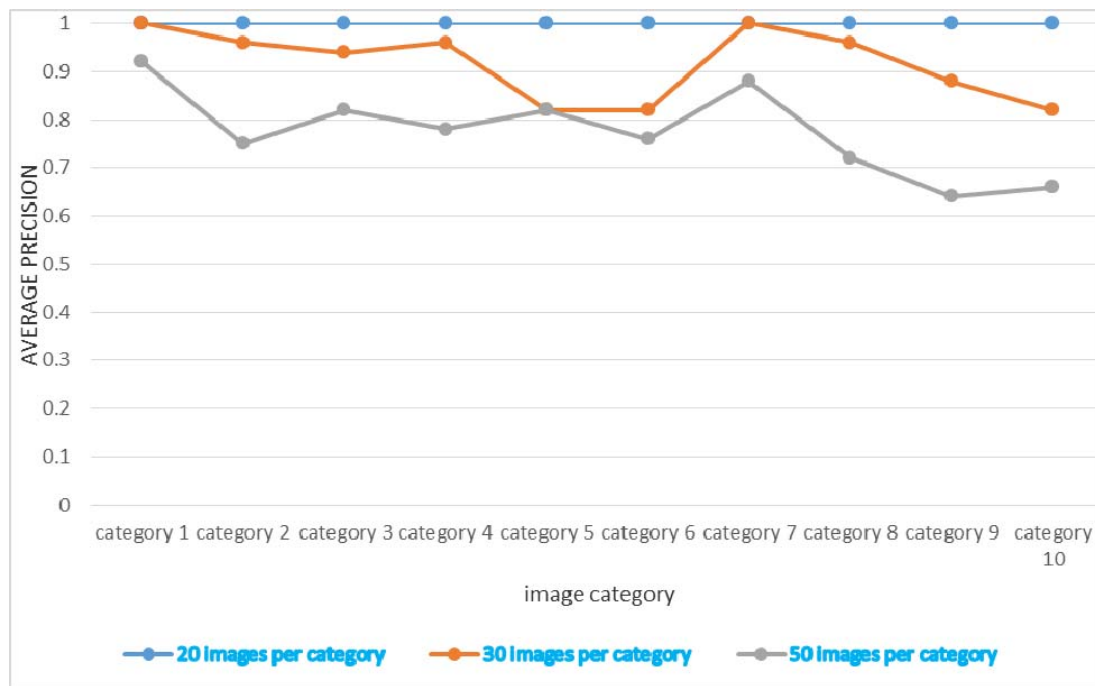


Fig.10 Average Retrieval efficiency of the proposed method for CALTECH data set for different classes

TABLE I RETRIEVAL EFFICIENCY FOR MPEG- 7 CE SHAPE I DATA SET

Number of categories used for evaluation	Average retrieval efficiency
20 categories	92%
30 categories	88%
40 categories	81%
50 categories	78%

Table I shows the average retrieval efficiency for MPEG-7 CE shape I dataset. In this different categories of images were considered and the proposed algorithm was tested. The maximum retrieval performance is 92%, which is optimum compared to the existing methods.

V. CONCLUSION

In this paper, an integrated and effective image retrieval system is proposed with hsv histogram, modified fourier descriptors and improved local binary patterns. The modified fourier descriptors are used in polar form in order to achieve the rotation and translation invariance. The shape descriptors are extracted from the modified fourier descriptors. The texture features are extracted from improved LBP. The hsv histogram is used for color feature matching. The proposed approach is an integrated approach as it can be used for both rgb images and gray-scale images. The retrieval efficiency is optimum for both the test databases considered. The efficiency is 88% and 72% which is greater than the existing methods.

REFERENCES

- [1] Manimala Singh, K.Hemachandran, "Content Based Image Retrieval using Color and Texture", Signal & Image Processing : An International Journal, Vol.3, No.1, February 2012 .
- [2] Rouhollah Rahmani, Sally A. Goldman, Hui Zhang, Sharath R. Cholleti, and Jason E. Fritts, "Localized Content-Based Image Retrieval," IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 30, No. 11, pp. 1902-1912, Nov. 2008.
- [3] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang., "ImageRetrieval: Ideas, Influences, and Trends of the New Age," ACM Computing Surveys, Vol. 40, No. 2 article 5, pp.5:1-5:60, Apr.2008.
- [4] Santosh Kumar Vipparthi, Shyam Krishna Nagar, "Color Directional Local Quinary Patterns for Content Based Indexing and Retrieval",
- [5] Human-centric Computing and Information Sciences, Springer 2014.
- [6] Fazal Malik, Baharum Baharudin, "Analysis of Distance Metrics in Content-Based Image Retrieval using Statistical Quantized Histogram Texture features in the DCT domain", Journal of King-Saud University-Computer and Information Sciences, Elsevier pub, vol.25, no.2, pp.207-218, 2013.
- [7] Xiang Yang Wang, Yong Jian Yu, Hong Ying Yang, "An Effective Image Retrieval Scheme using color, texture and shape features", Journal - Computer Standards & Interfaces, Elsevier, vol.33, no.1, pp.59-68, 2011.
- [8] Ammar Huneiti, Maisa Daoud, "Content Based Image retrieval using SOM and DWT", Journal of Software Engineering and Applications, vol.8, pp. 51-61, 2015.

- [9] Youngeun An, Muhammad Riaz, Jongan Park, "CBIR based on adaptive segmentation of HSV color space", Proceedings of IEEE International Conference on Computer Modelling and Simulation, IEEE computer society, ACM pub, pp.248-251, 2010.
- [10] Yan Zhang, Bogdan J. Matuszewski, Aymeric Histace, Frédéric Precioso, "Statistical Model of Shape Moments with Active Contour Evolution for Shape Detection and Segmentation", Journal of Mathematical Imaging and Vision, vol.47, issue 1, pp: 35-47, 2013.
- [11] Christoph Dalitz, Christian Brandt, Steffen Goebbels, David Kolanus, "Fourier descriptors for broken shapes", EURASIP Journal on Advances in Signal Processing, 2013.
- [12] Kazuhiro Kobayashi Qiu Chen, "Content-Based Image Retrieval using Features in Spatial and Frequency Domains", 4th International Conference on Soft Computing, Intelligent Systems, and Information Technology, ICSIIT 2015, Communications in Computer and Information science, Springer pub, pp 269-277, 2015.
- [13] Youngeun An, Muhammad Riaz, Jongan Park, "CBIR based on adaptive segmentation of HSV color space", proceedings of IEEE International Conference on Computer Modelling and Simulation, IEEE computer society, ACM pub, pp.248-251, 2010.

AUTHOR PROFILE

Author1 N.Neelima received her B.E degree in Electronics & Communications Engineering from Andhra University and M.Tech degree from Bharath University, India. Currently she is pursuing Ph.D in the field of image processing in Acharya Nagarjuna University, Andhra Pradesh. She is working as Asst.Prof. in the department of Electronics & Communications Engineering, Jain University, Bangalore, India.

Author2 Dr.E.Sreenivasa Reddy received his M.S degree from Birla Institute of Technology & Science, Pilani, India and Ph.D degree from Acharya Nagarjuna University, India. He is working as Professor in the department of Computer Science & Engineering in Acharya Nagarjuna University, Guntur, India. He is the chairman for board of studies in Computer Science & Engineering department in Acharya Nagarjuna University. He is also the member of several professional bodies.

Author 3 Kalpitha.N received her B.Tech degree from Visveshwaraya Technological University, M.Tech from Jain University, Bangalore. Currently she is pursuing her Ph.D from Visveshwaraya Technological University in the field of Image Processing. She is working as Asst.Prof. in the department of Computer Science & Engineering, Jain University, Bangalore, India.