

Sub-Block based Color Moments, Wavelet and Edge Histogram for Image Retrieval

Naveena AK¹, N K Narayanan²

¹Computer Science and Engineering Department, College of Engineering Trikaripur, Kasaragod, Kerala, India

naveenaak@gmail.com

²School of Information Science and Technology, Kannur University, Kerala India

nknarayanan@gmail.com

Received 05 June 2017; Accepted 30 June 2017

ABSTRACT

This paper proposes a novel image retrieval algorithm using local color feature of image sub-block and global texture and shape features. Image sub-blocks are identified by partitioning the image into blocks. Color Texture and shape are the low level image descriptor in Content Based Image Retrieval. These low level image descriptors are used for image representation and retrieval in CBIR. In this paper a Content Base Image Retrieval (CBIR) System using the image features extracted by color moments, wavelet and edge histogram is presented. Combining the color, texture and shape feature leads to a more accurate result for image retrieval. Moreover the color moments are taken by partitioning the images into blocks hence it also gives spatial color information. Here SVM classifier is used to classify the images into different class and the similarity measure is taken only with the images in the same class.

Keywords: sub-block CBIR; color moment; wavelet; edge histogram; SVM

INTRODUCTION

With the emergence of the multimedia content over the internet recent research attracts efforts in providing tools for effective retrieval of image database. There is a growing interest in developing effective methods for searching large image database [1]. Image database areas include a variety of fields such as medical, entertainment, crime detection, digital libraries etc. As the database grows larger, the traditional keyword based search and retrieval of similar images become inefficient as it does not consider the image feature and suffers from the following limitations. 1) It is difficult to express the visual content like color, texture, shape and object within the image precisely 2) For large dataset annotations requires more skilled labour and very large sophisticated keyword system 3) key word increase linguistic barrier to share image data globally.

To overcome these limitations image retrieval based on content is emerged as a promising alternative. The system which retrieves the desired images from a large collection of images automatically using the visual features is called a

Content Based Image Retrieval (CBIR) system. CBIR becomes an important area of research with increasing demand and use in digital images in various fields [2]. Two major functions of CBIR are the feature extraction and similarity matching. In the feature extraction phase the set of features which are fruitful to represent each image uniquely is extracted from the image [3]. Similarity matching measures the similarity between features of the query image and the images in the database. In CBIR the user queries are in the form of query image and the features are extracted from the query image and the stored as a features vector. These features are compared with the previously extracted and stored feature vectors of the database images. The comparison is made by computing the distance values and these values are used to rank the images according to their similarity with the query images. The most similar images are finally displayed to the user. The challenge in content based retrieval is to represent each image in a unique way to make accurate identification of the image. Therefore, the key to successful retrieval system depends on choosing the right features to accurately represent the image

and the size of feature vector. Here an initial classifier is used to identify the class label of the query image. Then distance measure between the query image and the images in that class is computed to retrieve the most similar images.

Many earlier image retrieval systems performed the retrieval based only on the global features of the images [4, 5,6]. Global features alone cannot represent the image and hence in such systems the performance was low [7]. Most of the early studies on CBIR used only a single feature among various features. But it is hard to attain satisfactory retrieval result by using a single feature. Recent active research in image retrieval use a combination of color and texture feature [8]. A novel approximate indexing using high dimensional image descriptor vector by employing value cardinalities of their dimension is proposed by Tiakas *et al.* in [9]. A multi sort algorithm is used in this paper to reorder the descriptors' dimension according to their value cardinalities, in order to increase the probability of two similar images to lie within a close constant range. A variety of relevant feedback scheme have been developed in recent years to improve the performance of content based image retrieval [10,11,12]. Another active research on image retrieval focused on region based technique. Region based technique either involve sub-divide the images into fixed blocks or segmenting the image into meaningful regions based on the pixel values [13,14]. Performance of the segmentation based method depends highly on the quality of the segmentation because the features are taken from pixels that belong to a particular segment. Small area of incorrect segmentation might leads to a representation very different from the real object. Also accurate segmentation is a challenging problem and cost of segmentation is higher. A new type of CBIR approach is presented in [15] which uses a block truncation coding (BTC) technique. This method was first introduced by Guoping Qiu in [15] and a number of variations of BTC is now recently used in the CBIR [16][17][18]. Image retrieval using the BTC extracts the image feature descriptor from the compressed data stream. Lot of computations is required to encode the image into compressed data. In this paper a block based color feature, global texture and shape feature are used for feature extraction. This retrieval method is relatively simple and avoids unnecessary computations. The rest of the paper is organized as follows. A description of the proposed method is provided in section II. Section III presents the experimental results on image retrieval and a comparison of the method with the existing

technique. Concluding remarks are given in section IV.

II. PROPOSED BLOCK METHOD

Image features can be taken by segmenting the images. Segmentation can be done in different ways. A pixel wise segmentation is computationally costly and accurate segmentation is very challenging. In this paper a fixed sized block method is used for feature extraction. Since the features are extracted from each block it also gave the spatial information of the pixels. The images are partitioned into 5x5 blocks and the first three color moments are extracted from each plane of each block. The figure 1 shows the original image and the image partitioned in to 5x5 blocks in a single plane.



Figure 1: Image Partitioned into 5x5 block

A. FEATURE EXTRACTION

Feature extraction is an important unit in any classification problem. An image can be represented by its features and these features are used for classification. The overall accuracy of the image retrieval is based on the selected feature. Two types of generic features are: global features and local features. Global features include color, texture and shape features extracted from the whole image. It is the easier method, but global representation does not give information about

spatial location of the pixels. To overcome the limitation of the global representation the images are segmented into set of regions or the features are extracted from the sub images, which is called local image representation [19]. In some case the image is divided into several parts and the feature coefficients are extracted from these parts and the final image representation is obtained by concatenating the representation of the image parts. The present paper includes global and local feature extraction methods. A successful categorization of the images into semantically meaningful categories based on the low level features greatly improve the performance of the CBIR by filtering out the images from irrelevant class during matching.

1. COLOR MOMENT

Color is an important feature for human perception. Color is the most widely used visual feature in image retrieval. An important attribute of color is that it is not sensitive to rotation, translation and scale changes [20]. Moreover the three dimensional values make its description superior to one dimensional gray level values. Each pixel of the image can be represented as a point in 3D color space. Widely used color model in CBIR are RGB, CIE L*a*b* and HSV. RGB is the commonly used color space for image processing. Other color space representation can be obtained by transforming the RGB color model. The idea of color transformation is to develop a model of color space that is perceptually similar with human color vision. CIE L*a*b* space are considered to be device independent and perceptually uniform. This consists of Luminance or lightness component L* and chromatic components a* and b*. HSV color space is approximately perceptually uniform. L*a*b* color space is used to extract color feature in this paper.

L*a*b* model is a three dimensional model, it can be properly represented only in three dimensional space. The three coordinates represent the lightness of the color L*, its position between red and green a* and its position between yellow and blue b*.

Color moment is a measure that can be used to differentiate images based on the feature of color. The basis of color moment is based on the assumption that the distribution of color in an image can be interpreted as a probability distribution. Probability distribution is characterized by number of unique moments [4]. The color in an image also follows a certain probability distribution, thus the moments of that distribution can be used as features to identify the

image based on color. Three channel color moment in the L*a*b* color space is used for color feature extraction in this paper. Color moments perform better in L*a*b* color space [21]. In this paper first moment, second moment and third moment of each color channel is used. The first moment is the average and hence average color of the image is taken as a feature. The second and third central moments are the standard deviation and skewness of each color channel.

Let μ_i , σ_i and s_i represent the mean, standard deviation and skewness respectively.

Mathematically these three moments are described as

$$\mu_i = \frac{1}{N} \sum_{j=1}^N P_{ij}$$

(1)

Where P_{ij} is the pixel value of the j^{th} pixel in the i^{th} color channel.

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (P_{ij} - \mu_i)^2 \right)^{1/2}$$

(2)

and

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (P_{ij} - \mu_i)^3 \right)^{1/3}$$

(3)

2. WAVELET FEATURE

Texture is an innate property of virtually all surfaces used in identifying objects or regions of interest in an object. Texture is a primitive set of pixels in some regular or repeated relationship in an image. Wavelet transform provide a multi resolution approach to texture analysis and classification [1][5]. The wavelet transform split the mage into high and low frequency components. At the first level of decomposition there are 4 sub-bands LL1, LH1, HL1 and HH1 where L denote low frequency and H denote high frequency. For the successive level of decomposition, the LL sub-band of the previous level is used as the input. To perform the second level decomposition wavelet transform is to the LL1 band. This wavelet decomposition is known as the pyramid structure wavelet transforms (PWT). PWT recursively decomposes LL band. However for some textures the most important information often appears in the middle frequency channels. For this Tree wavelet transform (TWT) is used which decomposes other bands such as LH, HL and HH [1,22,23].

This method uses a four level DWT for feature extraction. Feature vector can be constructed using the mean and standard deviation of the energy sub-band at each level. Mean and standard deviation of each sub-band levels using pyramid type and tree type decomposition are computed

and a 60D features for pyramid and 64D feature for tree structure are constructed and used as feature vector.

3. SHAPE FEATURE

Shape of an object can be composed of one or more regions, and sometimes may have holes also. Here shape feature is extracted using the edge histogram descriptor. Edge can be described as image positions where local intensities change along a particular orientation. The stronger the intensity changes the higher the evidence for an edge at that position. Edges are categorized into five types: vertical, horizontal, 45° diagonal, 135° diagonal and non-directional edges [24].

Canny edge detector is used to detect edge in this paper. Canny edge detector detects edges by double thresholding. Canny operator applies two thresholds to the gradient, a high threshold for low edge sensitivity and low threshold for high edge sensitivity. It starts with low sensitivity results and grows it to include connected edge pixels from the high sensitivity result. This help in filling gaps in detected edges. The Figure 2 shows an example of edge extracted image.



Figure 2: (above) Original image (below)Edge extracted from the image

The canny edge detector first smooths the image to eliminate the noise. It then finds the image gradient to highlight regions with high spatial

derivatives. The algorithm then tracks along these regions and suppresses any pixel that is not at the maximum (nonmaximum suppression). The gradient array is now further reduced by hysteresis. Hysteresis is used to track along the remaining pixels that have not been suppressed. Hysteresis uses two thresholds and if the magnitude is below the first threshold, it is set to zero (made a nonedge). If the magnitude is above the high threshold, it is made an edge. And if the magnitude is between the 2 thresholds, then it is set to zero unless there is a path from this pixel to a pixel with a gradient above T2 [24]. Here the image is first converted into binary image and the edges are extracted from the binary image. Edge histogram is used as the shape feature in this paper.

B. SVM CLASSIFIER

Support vector machine has been proposed by Vapnik [26] which is a very effective method for general purpose pattern recognition. SVM has the inherent ability to solve a pattern classification problem in a manner close to the optimum for the problem of interest. The main idea of SVM is to construct a hyperplane as the decision surface in such a way that the margins of separation between positive and negative samples are maximized. The Figure3 shows support vector machine that uses an optimal hyper plane to separate two classes.

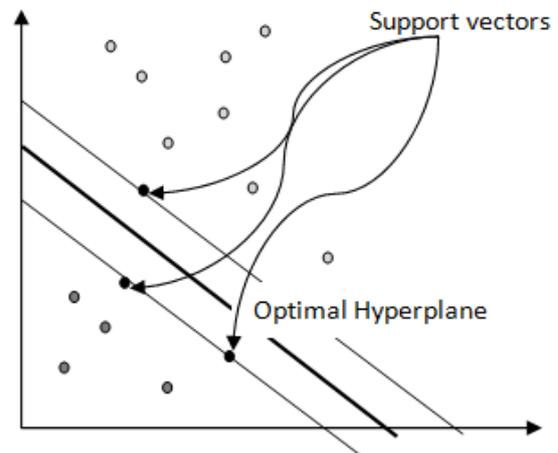


Figure 3: SVM with optimal Hyperplane

Image retrieval can be considered as an N class classification where the query image is compared with the number of images in the databases and the query image is classified into most appropriate class according to its distance score. Machine learning is of two types unsupervised clustering and the supervised classification. SVM belong to the later one and introduced as a learning machine for two group classification [25][26]. SVM is later extended to do preference learning and multiclass classification.

A classification task usually involves separating data into training and testing sets. Each instance in

the training set contains one class label and several features. The SVM predicts the class label of the test data given only the test features.

C. IMAGE RETRIEVAL

This image retrieval system uses an SVM classifier to find the class of the query image. Figure 4 shows the components on image retrieval system used in this paper. The three main unit in this CBIR is Query unit, Database unit and Retrieval unit.

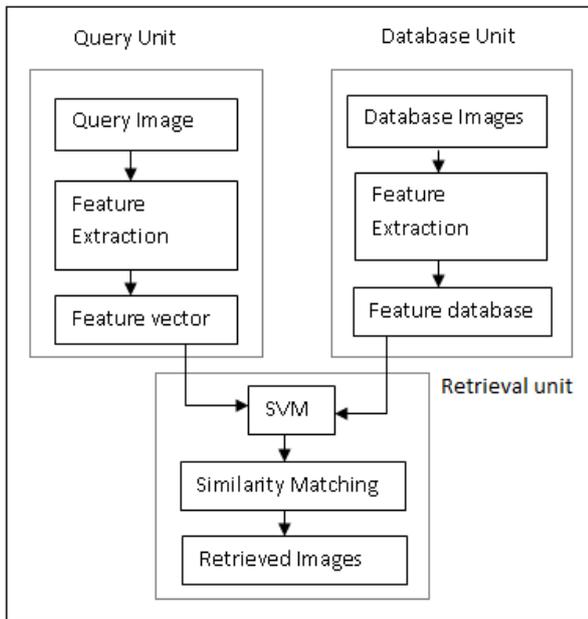


Figure 4: Components of CBIR

Query unit: Query unit randomly select an image from the train dataset and features are extracted and represented as a feature vector. All the features described in the section II A is extracted and stored as a features vector. Since the block method is used for color feature extraction, the query image is partitioned into 5x5 block and the three color moments are extracted from each channel of the L*a*b* color space. Before extracting the texture feature the query image is transformed to gray scale, the mean and standard deviation of the 4-level DWT form the texture feature. Edge detected using canny edge detector is used to extract the edge histogram descriptor. The three types of features combined to form the feature vector of the query image.

Database unit: Feature vectors of all the training images are extracted and stored as the feature database in database unit. The features of each image in the train image is extracted in the similar way as the query image, so feature vector of each image in the database and query image feature vector must be of the same size.

Retrieval unit: Retrieval unit takes the feature vector of the query image and feature database as the input and decide the most appropriate class for the query image using SVM classifier. Then based on the similarity score the most similar image to the query images are displayed to the user. The algorithm for this image retrieval system is given in Figure 5.

Algorithm 1

Input: Database L with labeled images.

1. Extract the visual features x_i ($i= 1, 2, 3$) of the images in the database L.
2. Each feature vector x_i is stored in the feature database DB.
3. Input query image y from the test image.
4. Extract all the features y_i of y.
5. Load the feature database DB.
6. Train the SVM classifier using the one against one method.
7. Classify the query image into most appropriate class in the image database L say C_i .
8. Compute the Euclidean distance between the query image and all the images in that class C_i .
9. Sort the images in the class C_i based on distance measure.
10. Display the most similar images from the class C_i .

Figure 5: Algorithm for image retrieval

III. RESULTS AND SYSTEM EVALUATION

The simulations are performed on Wang image database [28] which contains 1000 images of 10 different classes, each class with 100 images. An experimental used sample dataset along with the label is shown in Figure 6. A retrieved image is considered to be correct if it belongs to the same class as the query.

For the simulation experiment five images from each class are chosen randomly out of 1000 images to test the system. Hence a subset of 50 images is used as the test image.

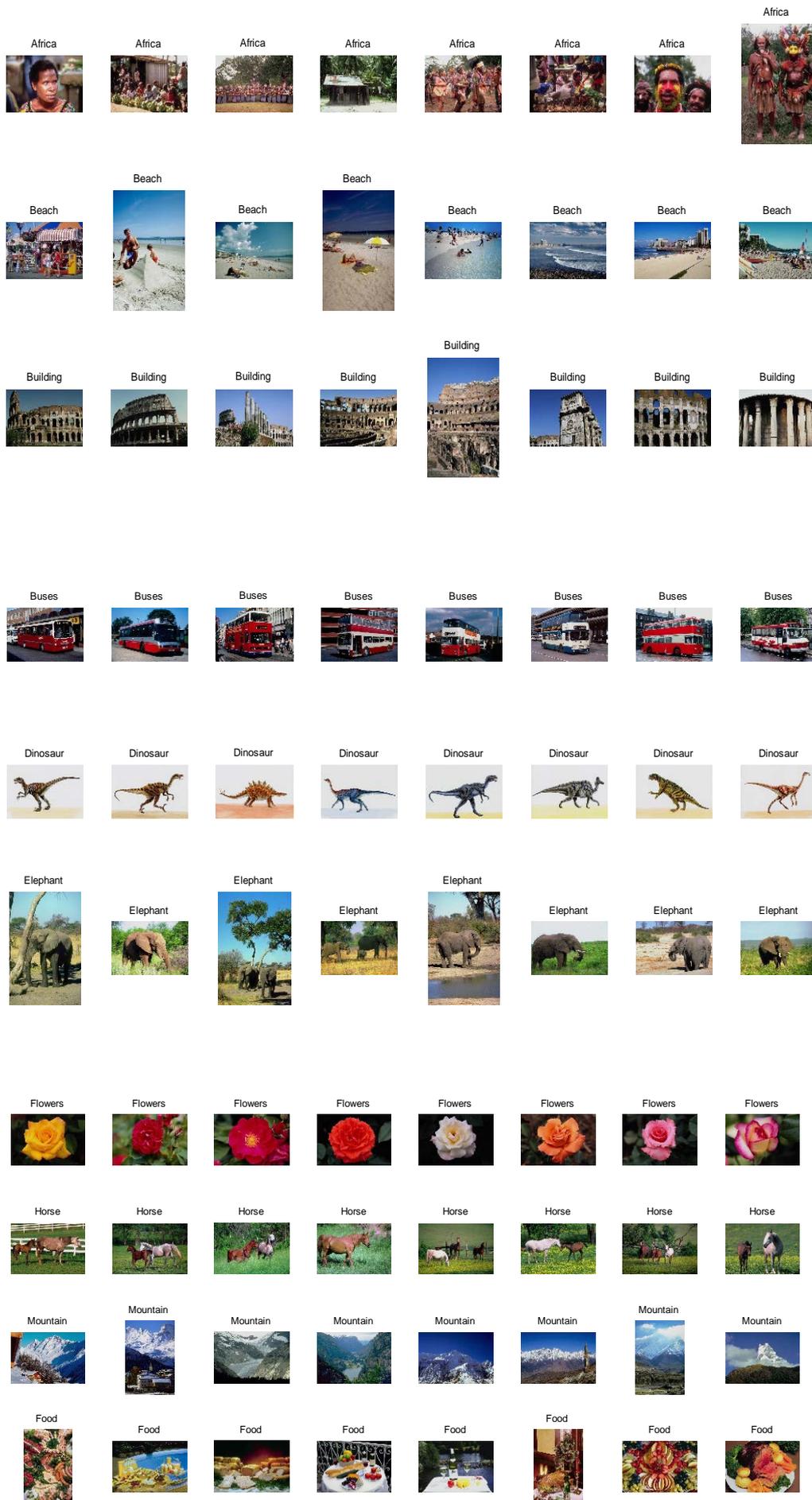


Figure 6: Sample Image database

Precision and recall can be used as a measure for analyzing the image retrieval accuracy which can be defined as follows [2].

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

$$Recall = \frac{\text{No. of relevant images retrieved}}{\text{Total no. of relevant images in database}}$$

A sample output is shown in Figure 7. The precision and recall of each class are separately calculated. From Table 1. shows that the precision based on single feature either color, texture or shape are less than the combined color, texture and shape feature. Table 2 shows the comparison of average precision of different categories like Africa, Beaches, Building, Buses, Dinosaur, Elephant, Flowers, Horses, Mountain and food using the methods of Jalab[7], CTDCIRS[29] and Vimina[30].

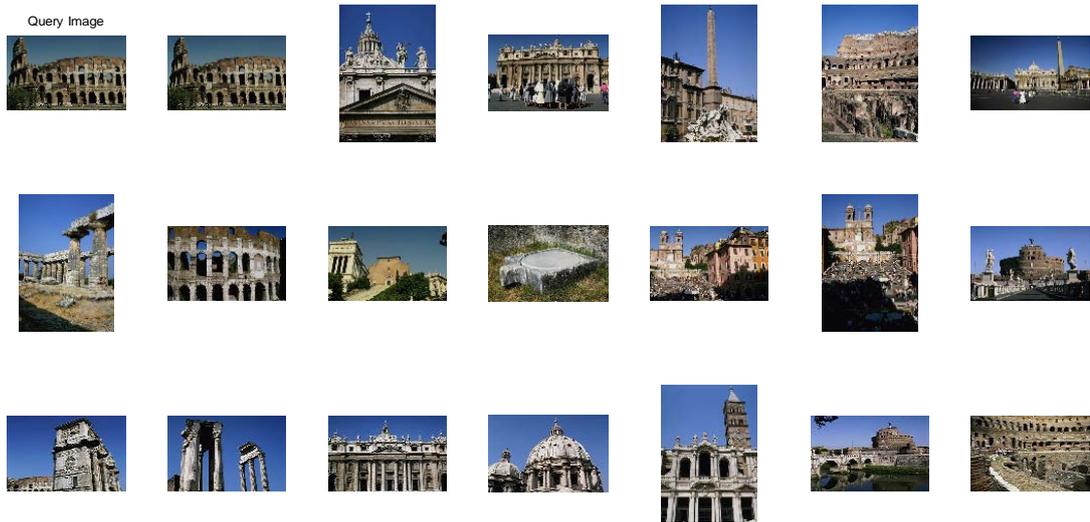


Figure 7: A Sample output of Image Retrieval

Table 1: % Average precision of color, texture, shape feature and the color+texture + shape feature

Class	Average precision			
	Color	Texture	Shape	Color+Texture+Shape
Africa	35.09	28.33	34.85	50.07
Beach	56.1	38.78	46.67	57.45
Building	34.69	32.56	44.0	60.0
Bus	63.79	48.28	83.02	87.5
Dinosaur	96.08	98.04	90.91	98.04
Elephant	66.04	62.22	60.42	65.52
Flower	88.24	73.68	82.93	93.02
Horse	88.89	70.37	60.0	90.2
Mountain	52.27	36.36	32.5	55.56
Food	41.86	38.0	53.66	53.45
Average	62.31	52.66	58.9	71.08

Table 2: Comparison of Average precision of Different Methods

Class	Average precision			
	Jalab [7]	CTDCIRS [29]	Vimina [30]	Proposed Color+Texture+Shape
Africa	32.3	56.2	71.56	50.07
Beach	61.2	53.6	46.56	57.45
Building	39.2	61	56.20	60.0
Bus	39.5	89.3	87.40	87.5
Dinosaur	99.6	98.4	99.95	98.04
Elephant	55.7	57.8	58.45	65.52
Flower	89.3	89.9	95.15	93.02
Horse	65.2	78	92.65	90.2
Mountain	56.8	51.2	35.65	55.56
Food	44.1	69.4	67.35	53.45
Average	58.29	70.48	71.05	71.08

IV. CONCLUSIONS

An efficient image retrieval system based on combining three different low level visual feature and SVM classifier is presented. The image retrieval is performed after classifying the image in to most appropriate class. Simulation results shows that this initial classification gives the correct class of the query image. This method was evaluated on the Corel 1000 image database and shows a superior performance than some of the existing image retrieval system.

REFERENCES:

1. Kokare, Manesh, Prabir K. Biswas, and Biswanath N. Chatterji. "Texture image retrieval using new rotated complex wavelet filters." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 35.6 pp.1168-1178, (2005). Conference on System Sciences (HICSS '00), January 2000.
2. Datta, Ritendra, et al. "Image retrieval: Ideas, influences, and trends of the new age." *ACM Computing Surveys (CSUR)* 40.2 (2008).
3. ElAlami, M. Esmel. "A new matching strategy for content based image retrieval system." *Applied Soft Computing* 14 pp407-418, (2014).
4. Stricker, Markus A., and Markus Orengo. "Similarity of color images." *IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology*. International Society for Optics and Photonics, 1995.
5. Manjunath, Bangalore S., and Wei-Ying Ma. "Texture features for browsing and retrieval of image data." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 18.8 pp 837-842, (1996).
6. Jhanwar, N., et al. "Content based image retrieval using motif cooccurrence matrix." *Image and Vision Computing* 22.14 pp 1211-1220, (2004).
7. Jalab, Hamid A. "Image retrieval system based on color layout descriptor and Gabor filters." *Open Systems (ICOS), 2011 IEEE Conference on*. IEEE, 2011.
8. Chun, Young Deok, Nam Chul Kim, and Ick Hoon Jang. "Content-based image retrieval using multiresolution color and texture features." *IEEE Transactions on Multimedia* 10.6 (2008): 1073-1084.
9. Tiakas, Eleftherios, et al. "MSIDX: multi-sort indexing for efficient content-based image search and retrieval." *IEEE Transactions on Multimedia* 15.6 (2013): 1415-1430.
10. Wang, Yu-Chen, et al. "Biased Discriminant Analysis With Feature Line Embedding for Relevance Feedback-Based Image Retrieval.

- " *IEEE Transactions on Multimedia* 17.12 (2015): 2245-2258.
11. Grigorova, Anelia, et al. "Content-based image retrieval by feature adaptation and relevance feedback." *IEEE transactions on multimedia* 9.6 (2007): 1183-1192.
 12. Tao, Dacheng, et al. "Direct kernel biased discriminant analysis: a new content-based image retrieval relevance feedback algorithm." *IEEE Transactions on Multimedia* 8.4 (2006): 716-727.
 13. Wang, James Ze, Jia Li, and Gio Wiederhold. "SIMPLicity: Semantics-sensitive integrated matching for picture libraries." *IEEE Transactions on pattern analysis and machine intelligence* 23.9 pp 947-963, (2001).
 14. Chen, Yixin, James Ze Wang, and Robert Krovetz. "CLUE: cluster-based retrieval of images by unsupervised learning." *IEEE transactions on Image Processing* 14.8 pp 1187-1201, (2005).
 15. Qiu, Guoping. "Color image indexing using BTC." *IEEE Transactions on Image Processing* 12.1 pp 93-101, (2003).
 16. Guo, Jing-Ming, Heri Prasetyo, and Jen-Ho Chen. "Content-based image retrieval using error diffusion block truncation coding features." *IEEE Transactions on Circuits and Systems for Video Technology* 25.3 pp 466-481, (2015).
 17. Guo, Jing-Ming, Heri Prasetyo, and Nai-Jian Wang. "Effective Image Retrieval System Using Dot-Diffused Block Truncation Coding Features." *IEEE Transactions on Multimedia* 17.9 pp1576-1590, (2015).
 18. Guo, Jing-Ming, and Heri Prasetyo. "Content-based image retrieval using features extracted from halftoning-based block truncation coding." *IEEE Transactions on image processing* 24.3 pp1010-1024 (2015).
 19. Wang, Xiang-Yang, et al. "An image retrieval scheme with relevance feedback using feature reconstruction and SVM reclassification." *Neurocomputing* 127 pp 214-230, (2014).
 20. Yue, Jun, et al. "Content-based image retrieval using color and texture fused features." *Mathematical and Computer Modelling* 54.3 Pp 1121-1127,(2011).
 21. Duan, Lixin, et al. "Improving web image search by bag-based reranking." *Image Processing, IEEE Transactions on* 20.11 pp 3280-3290, (2011).
 22. Kokare, Manesh, B. N. Chatterji, and P. K. Biswas. "Wavelet transform based texture features for content based image retrieval." *Proc. 9th Nat. Conf. on Communications (NCC 2003), Chennai, India.* 2000.
 23. Kokare, Manesh, P. K. Biswas, and B. N. Chatterji. "Rotation-invariant texture image retrieval using rotated complex wavelet filters." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 36.6 pp 1273-1282. (2006).
 24. Canny, John. "A computational approach to edge detection." *IEEE Transactions on pattern analysis and machine intelligence* 6 pp 679-698, (1986).
 25. Jain, Sonali. "A Machine Learning Approach: SVM for Image Classification in CBIR." *International Journal of Application or Innovation in Engineering & Management (IJAIEM)* 2.4 (2013).
 26. Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." *Machine learning* 20.3 pp273-297, (1995).
 27. Hsu, Chih-Wei, and Chih-Jen Lin. "A comparison of methods for multiclass support vector machines." *Neural Networks, IEEE Transactions on* 13.2 pp 415-425, (2002).
 28. <http://wang.ist.psu.edu/docs/related>
 29. Rao, M. Babu, B. Prabhakara Rao, and A. Govardhan. "CTDCIRS: content based image retrieval system based on dominant color and texture features." *International Journal of Computer Applications* 18.6 pp 40-46, (2011).
 30. Vimina, E. R., and K. Poulouse Jacob. "CBIR using local and global properties of image sub-blocks." *Int. J. Adv. Sci. Technol* 48 pp11-22. (2012).