

CBIR Using Colour Layout Descriptor and Coiflets Wavelets

Muhammad Imran¹, Rathiah Hashim¹, Aun Irtaza², Noor Elaiza Abd Khalid³

¹FSKTM, University Tun Hussein onn Malaysia

²National university of Science and emerging Science, Islamabad

³FSKM University Teknologi MARA Malaysia

³National University of Computer Science and emerging Sciences Islamabad, Pakistan

Received: January 14 2014

Accepted: April 3 2014

ABSTRACT

Nowadays a lot of information in the form of digital content is easily accessible but finding the relevant image is a big problem. Content Based Image Retrieval (CBIR) comes in to solve the image retrieval dilemma. But a CBIR system faces certain problems such as a strong signature development, gap between the low level features and high level semantics. One of the major challenges of CBIR is to bridge the gap between the low level features and high level semantics. Previously, several researchers have been proposed to improve the performance of a CBIR system but they have answered image retrieval problem to an extent. This paper proposed a new CBIR signature development that uses MPEG-7 descriptor and wavelet packets. The results of the proposed method namely CLD-cw are compared with the five well reputed systems (i.e. SIMPLIcity, FEI, Histogram based, FIRM, and Variance Segment etc) from the industry. The results of the CLD-cw demonstrates high accuracy rate than the previous systems. The proposed CLD-cw has significant performance in term of accuracy. The CLD-cw can retrieve semantically similar images because it used the Color Layout Descriptor (CLD) of MPEG-7.

INTRODUCTION

In this era, lot of information is available in the form of digital content. Electronic storage and the computing power are increasing with every passing day. The end result of this rapid increase in content storage is making it easier for the users to access information in the form of images and videos. The image and video contents provides basis for the many educational and commercial applications. Currently, image and video databases are available on a large scale but to search relevant images from them is quite challenging. Two major systems are available to retrieve information from image databases. One is the Keyword Based Image Retrieval (KBIR) the other is the Contents Based Image Retrieval (CBIR). In KBIR, search is dependent on the keywords i.e. who adds the keyword and also the same image can have different meanings for different users. Due to these reasons performance of KBIR is not good [1]. CBIR retrieves similar images in terms of contents instead of keyword. Although CBIR is a better way to retrieve the relevant image from any database but still there is lot of challenges to be answered in this area of research. One of the biggest challenges is to retrieve an image with high accuracy from a database.

The large collections of digital images are being created in different areas such as government, commerce, hospitals and academia. In the past simple browsing or KBIR was used to search images from these collections [34]. To search images with CBIR, user has to provide the image or sketch as a query and the system will return the similar images based on the matching features of the images. CBIR is an application of computer vision technique which is also known as query by image content (QBIC) and content based visual information retrieval system (CBVIR).

Feature extraction and similarity measure are the key components of the CBIR system. Usually three types of features are extracted to measure the distance between the two images which are Color feature, Texture featured and Shape features.

Various CBIR techniques have been proposed in the literature. Few of them use local features while others use global features. The researchers also segment the image into regions based on color and texture. Many machine learning techniques are also applied on the CBIR system. The concept map of CBIR system can be seen in the Figure 1.

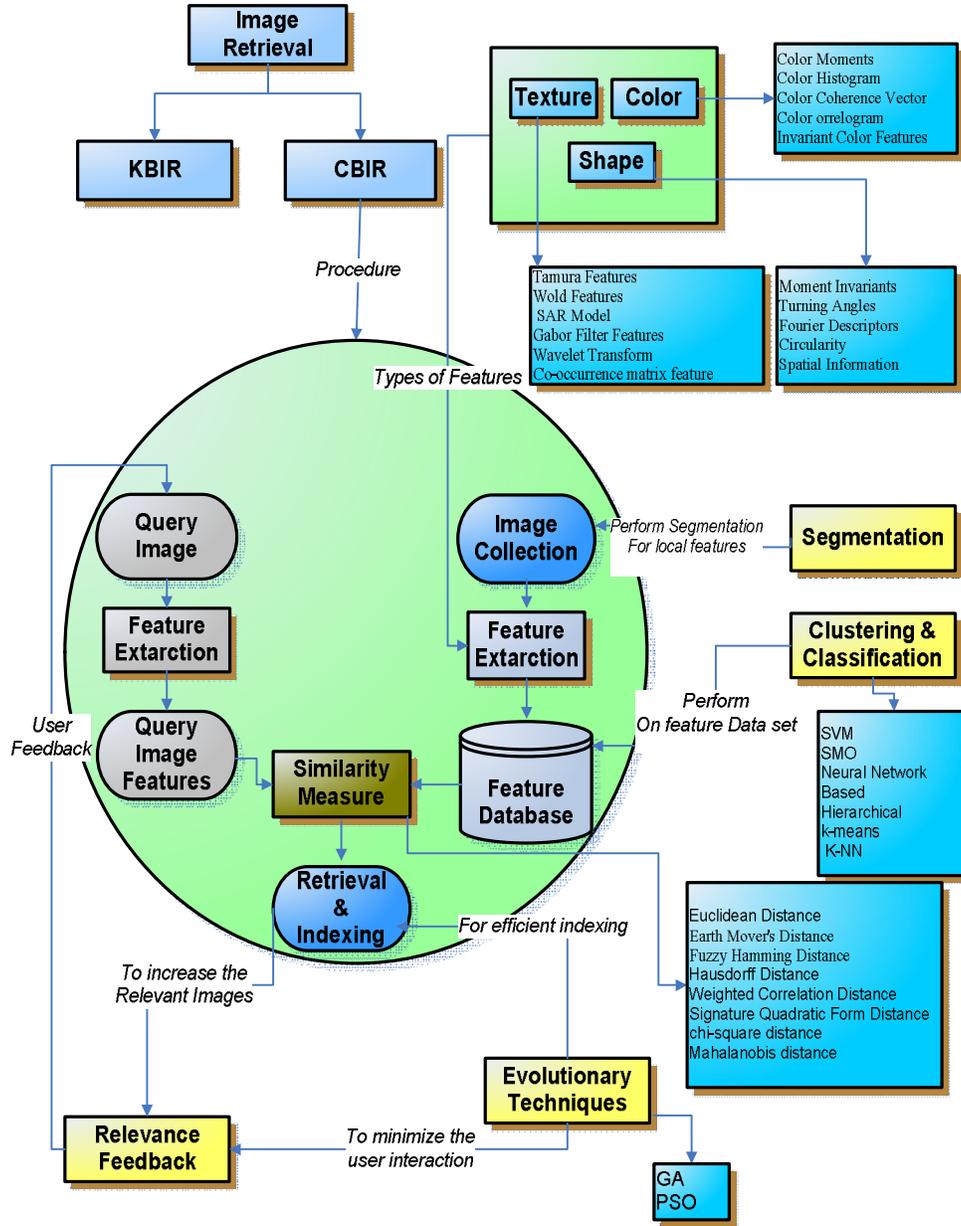


Figure 1: Concept Map of CBIR

1. Related Work

Similarity measure is very important for any CBIR system and is considered as the core of any CBIR system. Beecks et al [3] performed a comparison study of the similarity measures. The authors evaluated the performance of Hausdorff Distance, Perceptually Modified Hausdorff Distance, Weighted Correlation Distance, Earth Mover's Distance, and Signature Quadratic Form Distance on four different databases available in the literature. The results of their experiments show that the precision rate of each similarity measure depends on the feature space and the database. In case of only color features, Perceptually Modified Hausdorff Distance (PMHD) and the EarthMover's Distance (EMD) perform well. Hausdorff Distance achieved the lowest average precision. Authors also checked the computation time for each similarity measure technique. The lowest computation time was taken by Hausdorff Distance and Perceptually Modified Hausdorff Distance whilst the lowest efficiency was shown by the EMD. Finally, authors conclude that Quadratic Form Distance achieved the highest precision while HD and MHD takes the lowest computation time.

Singhai et al [4] perform a survey on the functionality of the CBIR. They conclude that most of the system used comprised of color and texture features while a few came up with the shape features and little bit are with layout features. Abubacker [5] used color, texture and shape features for the image. For the color feature they used the spatial based color moments. First they divide the image into 25 blocks then calculate the R G B values of each block. R, G, B values are converted to H, S, I. According to the author the three color moments; mean, variance and skewness are effective and efficient for the color distribution of images. The formula for mean, variance and skewness are given below:

$$Mean (\mu_i) = \frac{1}{N} \sum_{j=1}^N f_{ij} \dots\dots\dots (1)$$

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

$$Skewness (S_i) = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right)^{1/3} \dots\dots\dots (3)$$

Where, f_{ij} is the value of the i^{th} color component of the image block j and N is the number of blocks in the image. For the texture feature author used the Gabor filter. And applied the 2D Gabor function to obtain the set of Gabor filter with different scale and orientation. By using Gabor filter, he performs convolution on the image to obtain the Gabor transform. Invariant shape features are used to extract the shape features. Following are the steps taken by the author to extract the shape feature.

- 1: Based on the threshold value the image is converted to binary image.
- 2: Using canny algorithm the edges of the binary image are detected.
- 3: The centroid of the object is obtained by arranging the pixels in clockwise order and forming a closed polygon.
- 4: The centroid distance and complex coordinate function of the edges is found.
- 5: The farthest points are found and Fourier transform is applied on them.

Akgül et al [6] completed a survey of CBIR in the medical imaging. Authors discussed the current state of the art techniques of CBIR in medical imaging. They came up with the new challenges and opportunities of CBIR in medical diagnoses process. Authors tried to focus the attention of the researcher on operation issues in medical CBIR and proposed certain strategies to tackle them. Huang et al [7] proposed the new technique of the CBIR using color moment and Gabor texture feature. To obtain the color moment they convert the RGB image to HSV image, then by getting the equalized histogram of the three HSV components calculated the three moments for each color space. Modified form of the Euclidian distance was used to measure the similarity between the query image and the database image. The equation is given below;

$$D(q, s) = \frac{1}{L} \sum_{i=0}^{L-1} \left(1 - \frac{|q_i - s_i|}{\max(q_i, s_i)} \right) \dots\dots\dots (4)$$

The global distance is computed as the weighted sum of similarities as:

$$D(q, s) = \frac{\omega_c \cdot D_c(q, s) + \omega_t \cdot D_t(q, s)}{\omega_c + \omega_t} \dots\dots\dots (5)$$

Through experiments Huang et al. [7] showed that the overall result of the proposed technique was better than other techniques. Zhao et al [8] combined relevance feedback with SVM and explored the different combinations of feature set. They extract three different kinds of texture feature and combined with the color feature in different combinations. Three combinations were contracted having one color moment and two texture features. The effect of region based filtering tested by the Pujari [9]. Pujari used the wavelet based texture features. For the color images the texture features of each color space R G B ware extracted separately and integrated region matching was used as similarity measure. The experiments performed using 0%, 3%, 6% and 8% filtering. There is no measureable difference between performance of the 6% and 8% filtering. Oliveira et al [10] used breast density for the image retrieval to help the radiologists in their diagnosis. Particle swarm optimization (PSO) with relevance feedback is used by the Broilo [11] to enhance the performance of CBIR. Broilo [11] formulates the image retrieval processes as an optimization problem and applied PSO on CBIR. This paper used color and texture features as a combination for signature development.

2. The CLD-cw Approach

To get semantically better results this paper proposed a new signature development technique by combining the color and texture features. To extract the color features, color layout descriptor of MPEG-7 is used while wavelets packets are used for the texture feature extraction. Finally both feature set are combined to represent the image in term of feature vector. Detail of feature extraction is provided in following sub sections.

2.1. MPEG-7

The MPEG-7 Visual Standards specifies content-based descriptors that allow users to measure similarity in images or video based on visual criteria, and can be used to efficiently identify, filter or browse images or videos based on visual content [45]. MPEG-7 has different descriptor for color, texture, shape and motion. In this paper, color layout descriptor is (CLD) used to extract color features of the images.

2.2. Color Layout Descriptor (CLD).

CLD is the best descriptor to describe the spatial distribution of color in an arbitrary shaped region [12]. There are four stages to extract the CLD [13]. In the first stage input image is partitioned into 64 (8x8) blocks. In second stage a single representative color is selected for each block. As a result a tiny image representation of size 8x8 is obtained. In the third stage each of the three color components are transformed by 8x8 DCT. Three sets of 64 DCT coefficients are obtained. To calculate the DCT in a 2D array, following formulas are used.

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N} \quad \begin{matrix} 0 \leq p \leq M-1 \\ 0 \leq q \leq N-1 \end{matrix} \dots\dots\dots (6)$$

$$\alpha_p = \begin{cases} 1/\sqrt{M}, p = 0 \\ \sqrt{2/M}, 1 \leq p \leq M-1 \end{cases} \quad \alpha_q = \begin{cases} 1/\sqrt{N}, q = 0 \\ \sqrt{2/N}, 1 \leq q \leq N-1 \end{cases}$$

The values B_{pq} are called the DCT coefficients of A. In the final stage zigzag scanning is performed for each set of coefficients.

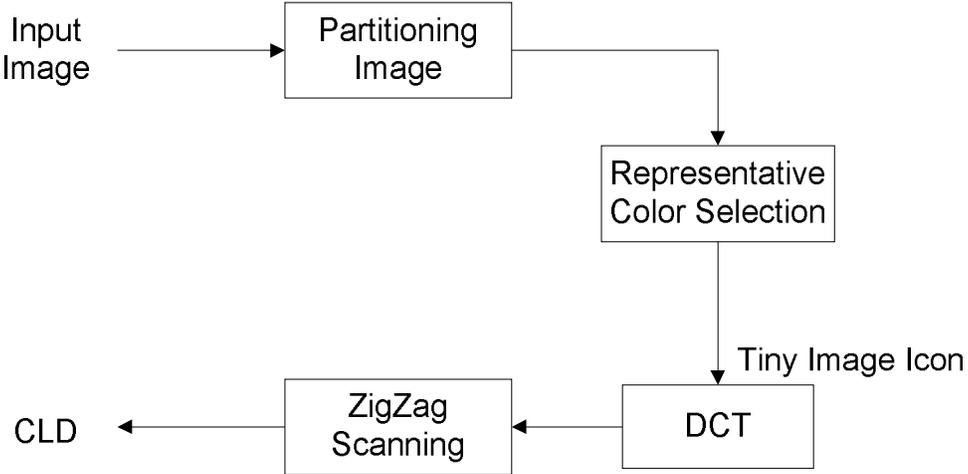


Figure 2: Extraction process of the CLD Scanning

2.2.1. Color Feature Vector

The steps of extracting the color layout descriptor are given in Figure 2. After zigzag scanning which is the last step. Three matrices for each block of Y, Cb and Cr color space are obtained. Three feature vectors for an image are acquired by taking the sum of the each matrix. Finally, the resulting feature vector is obtained by horizontally concatenating the all three feature vectors.

2.3. Wavelet Packets

To analyze the signals in a rich way the wavelet packet method is used which is a generalization of wavelet decomposition and can be described by the collection of functions $\{WJ(x)|J \in Z^+\}$ obtained from [14]

$$2^{\frac{p-1}{2}} W_{2n}(2^{p-1} x-l) = \sum_m h_{m-2l} 2^{\frac{p}{2}} W_n(2^p x-m) \dots\dots\dots (7)$$

$$2^{\frac{p-1}{2}} W_{2n+1}(2^{p-1} x-l) = \sum_m g_{m-2l} 2^{\frac{p}{2}} W_n(2^p x-m) \dots\dots\dots (8)$$

Where ‘p’ is a scale index and ‘l’ is the translation index. $W_0 x = \Phi(x)$ is the scaling function. $W_1 x = \psi(x)$ is the basic wavelet function. h_k and g_k are the quadratic mirror filters. Wavelet packets are well localized in both time and frequency and thus provide an attractive alternative to pure frequency (Fourier) analysis. For a given orthogonal

wavelet function, a library of bases is obtained called wavelet packet bases. Each of these bases offers a particular way of coding signals, reconstructing exact features and preserving global energy. The inverse relationship between wavelet packets of different scales can be shown through: [14]

$$2^{\frac{p}{2}}W_n(2^p x - k) = \sum_l h_{k-2l} 2^{\frac{p-1}{2}} W_{2n}(2^{p-1} x - l) + \sum_l g_{k-2l} 2^{\frac{p-1}{2}} W_{2n+1}(2^{p-1} x - l) \dots \dots \dots (9)$$

Equation (9) can be used to calculate the wavelet packets. Coefficients of coarser scale can be calculated using eq. 7 and eq. 8

$$S_{2n,l}^{p-1} = \sum_m h_m - 2l S_{n,m}^p \dots \dots \dots (10)$$

$$S_{2n+1,l}^{p-1} = \sum_m g_m - 2l S_{n,m}^p \dots \dots \dots (11)$$

The main difference between normal wavelet decomposition and wavelet packets decomposition is that despite of just splitting the approximation components, wavelet packets decomposes the detail components as well. So by using wavelet packets, rich analysis becomes possible.

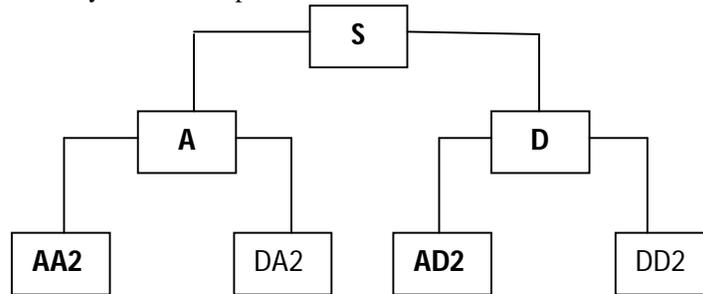


Figure 3: Wavelet packets decomposition at level 2

Wavelet packets procedure results in a large number of decompositions and its explicit enumerations are unmanageable. So it is necessary to find the optimal decomposition with respect to some reasonable criterion. One convenient criterion can be the selection of tree nodes on the basis of best entropy values. In this paper Shannon entropy measure is used to calculate the entropy. This can be calculated as:

$$E(S_i) = -\sum_i s_i^2 \log(s_i^2) \dots \dots \dots (12)$$

Using the Shannon entropy the optimal or the best tree can be calculated using the following scheme. A node N will be split into two nodes N₁ and N₂ if and only if the sum of the entropy of N₁ and N₂ is less than the entropy of N. This is a local criterion based only on the information available at the node N. It will create a form of a tree which is of much smaller size than the actual tree.

For the signature generation Shannon entropy based wavelet packets is used up to the 3rd level. Following formula is used to generate the Coiflets based features.

$$f_r = \sqrt{\frac{\sum c_{ij}^2}{i*j}} \dots \dots \dots (13)$$

Where f_r is the computed Wavelet signature (texture feature representation), C_{ij} is the representation of the intensity value of all elements of sub image and $i \times j$ is the size of the sub [14].

METHODOLOGY

2.4. Color Feature Extraction algorithm

- Image Partitioning
 - Divide the image into 8x8 blocks
- Representative Color Selection
 - A single representative color is selected from each block
 - The selection results in a tiny image icon of size 8x8
 - The color space conversion between RGB and YCbCr is applied.
- DCT Transformation
 - The luminance (Y) and the blue and red chrominance (Cb and Cr) are transformed by 8x8 DCT
- Zigzag Scanning
 - A zigzag scanning is performed with these three sets of DCT coefficients
 - As a Result we obtain three matrixes for each block of Y, Cb and Cr color space.
 - Take sum of each matrix
 - Horizontally concatenate the three feature vector to obtain a final feature vector for an image

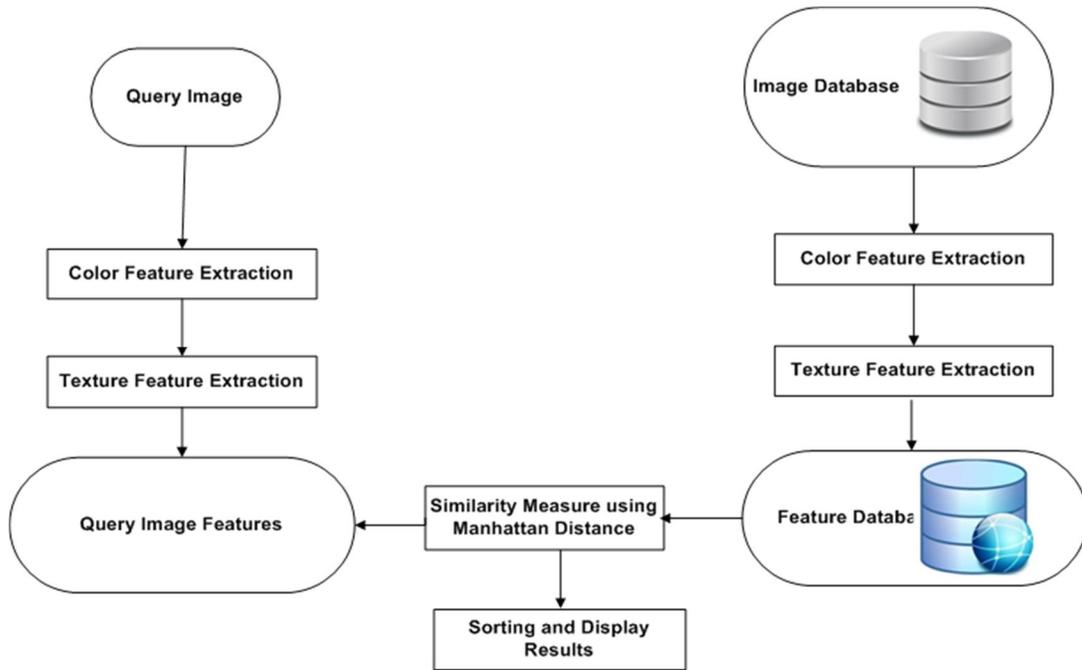
2.5. Texture Feature Extraction Algorithm

Let I be the image of size $w \times w$
 Divide the image I into four bands I_1, I_2, I_3, I_4
 based on Coiflets wavelet of size $w/2 \times w/2$
 Compute Signatures f_r for I_2, I_3, I_4
 Now take the image I_1 and divide it into 4 bands
 Namely $I_{11}, I_{12}, I_{13}, I_{14}$ of size $w/4 \times w/4$
 Compute signatures f_r for I_{12}, I_{13}, I_{14}
 Again take the I_{11} and divide it into 4 bands
 Namely $I_{111}, I_{112}, I_{113}, I_{114}$ of size $w/8 \times w/8$.
 Now we obtain 10 signatures then stop the process

$$f_r = \sqrt{\frac{\sum c_{ij}^2}{i * j}}$$

Where f_r is the computed Wavelet signature (texture feature representation), C_{ij} is the representation of the intensity value of all elements of sub image and $i \times j$ is the size of the sub

By using color feature extraction algorithm, color feature vector is generated while texture feature algorithm is used to extract the texture features. Features of the entire image database are extracted and saved to disc. When user input the query image, the system measures the similarity between query image and the database image by using Manhattan distance. After similarity measure the system sort the result and display to user. The flow chart of CLD-cw approach is shown below.



3. RESULTS AND ANALYSIS

Results of the CLD-cw approach are compared with other standard benchmark systems namely FEI, SIMPLiCity, FIRM [15-17], Variance Segment Method [18] and Histogram based taken from FEI. Famous coral dataset is used to assess the CLD-cw. The database contains 10 classes and each class has 100 images. CLD-CW algorithm is implemented under Matlab R2010a.

3.1. Metrics used for the Evolution

To measure the performance of the CBIR system different metrics are available. Precision is one of the metric which has been used in the several previous works such as Hiremath et al[19], Banerjee et al [15] and Wang et al [16]. Precision can be calculated as;

$$Precision = \frac{Number\ of\ True\ Positive}{Number\ of\ True\ Positive + False\ Positive}$$

Recall is also one of the metrics used for the evolution of the CBIR system.

$$Recall = \frac{Number\ of\ True\ Positive}{Total\ number\ of\ True\ positive}$$

3.2. Performance in Terms of precision

As illustrated above, precision is one of the metric used to check the performance of the CBIR system. Table-1 shows the performance of the proposed algorithm with different P @ n evaluation. The precision is calculated using the top most 50, 30 and 10 ranked results. In the experiments, 10 images from each category are selected randomly and used as query image. The average precision of 10 query image is then calculated for each category and presented in table 1.

Table 1: Performance at different n

	The Precision at n (calculated using n top most results)		
	n=50	n=30	n=10
Africa	.70	0.606	.66
Beach	.50	0.679	.66
Buildings	.49	0.306	.35
Buses	.36	0.455	.51
Dinosaurs	1	1	1
Elephant	.28	0.541	.64
Flower	.64	0.720	.66
Horses	.44	0.668	.76
Mountains	.33	0.375	.5
Food	.14	0.268	.32
Avg	.488	.561	.606

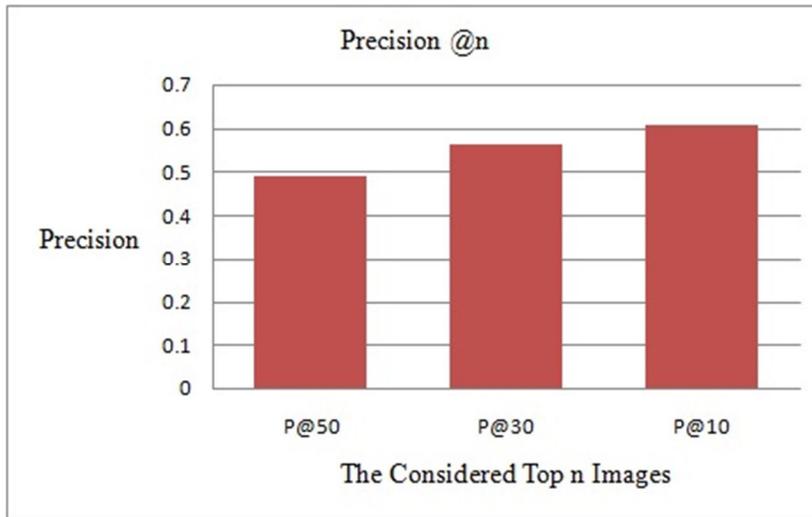


Figure 5: Performance at different n

From Table-1 and the graph shown in Figure-5, it is clear that precision is affected when number of top most n images is changed. However, performance of CLD-cw than SIMPLIcity, FIRM and FEI for n=30.

3.3. Comparison with Previous Methods

Comparison of the proposed method with FEI, SIMPLIcity, Simple Hist, FIRM and Variance Segment Method is shown in table 2.

Table 2: Comparison of proposed method with previous methods

Class	FEI	SIMPLIcity	Simple Hist	FIRM	Variance Segment	Proposed Method
Africa	.45	.48	.30	.47	0.13	0.60
Beach	.35	.32	.30	.35	0.26	0.67
Buildings	.35	.35	.25	.35	0.11	0.30
Buses	.60	.36	.26	.60	0.17	0.45
Dinosaurs	.95	.95	.90	.95	0.96	1
Elephant	.60	.38	.36	.25	0.34	0.54
Flower	.65	.42	.40	.65	0.49	0.72
Horses	.70	.72	.38	.65	0.20	0.66
Mountains	.40	.35	.25	.30	0.25	0.37
Food	.40	.38	.20	.48	0.15	0.26
Avg	.545	.471	.36	.505	0.174	.561

The FIRM and SIMPLIcity are cited in several earlier works. Both methods are based on segmentation therefore both have better performance. From Table-2, it is clear that FEI has better results than other methods. But the proposed methods have best performance then all five methods listed in the above table.

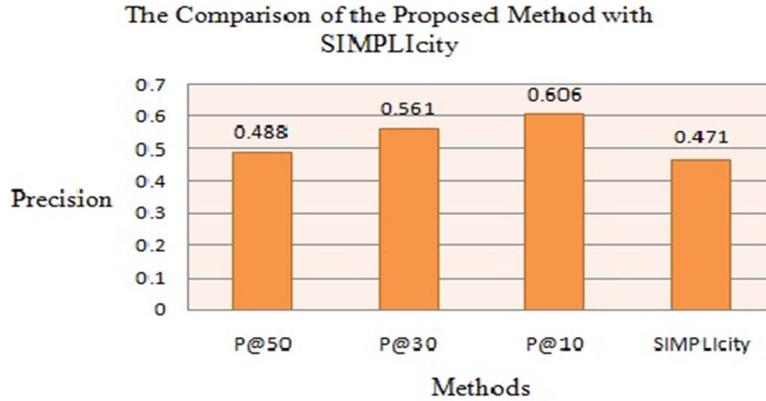


Figure 6: Comparison with SIMPLIcity

In the graph shown in Figure-6, the proposed method is compared with SIMPLIcity for $n = 50, 30,$ and 10 . From the above graph, it is observed that the proposed method has better average precision than SIMPLIcity for all different top retrievals. The SIMPLIcity and FIRM used sophisticated set of features. But CLD-cw with simple texture and MPEG-7 descriptor outperforms these methods.

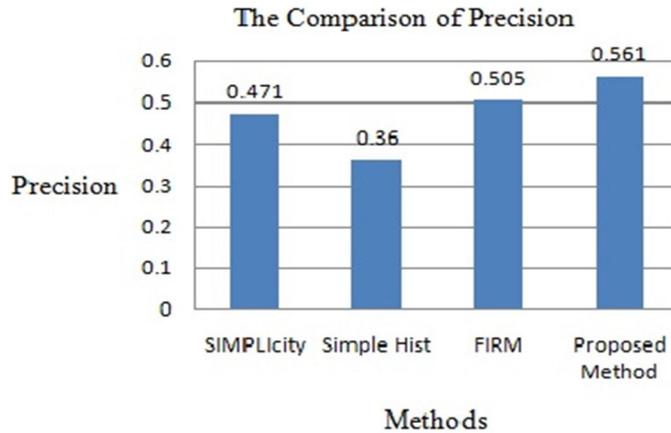


Figure 7: The Comparison of CLD-cw with Earlier Methods

The better performance of the proposed system can be observed from Figure 7 which the average precision comparison of CLD-cw with SIMPLICITY, Simple Hist and FIRM.

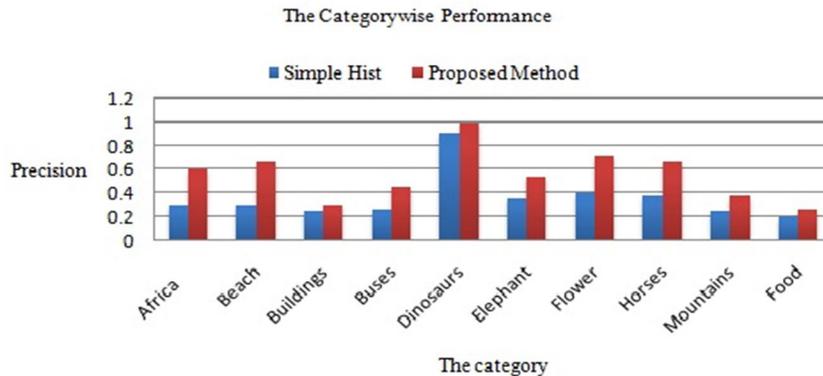


Figure 8: Category-wise Comparison of Proposed Method ($n @ 30$) with Simple Hist

In the above graph, we can see that the proposed method has better result than simple Hist method on all 10 categories. For the class dinosaurs the proposed method has 100% retrieval rate as the precision reaches to 1.

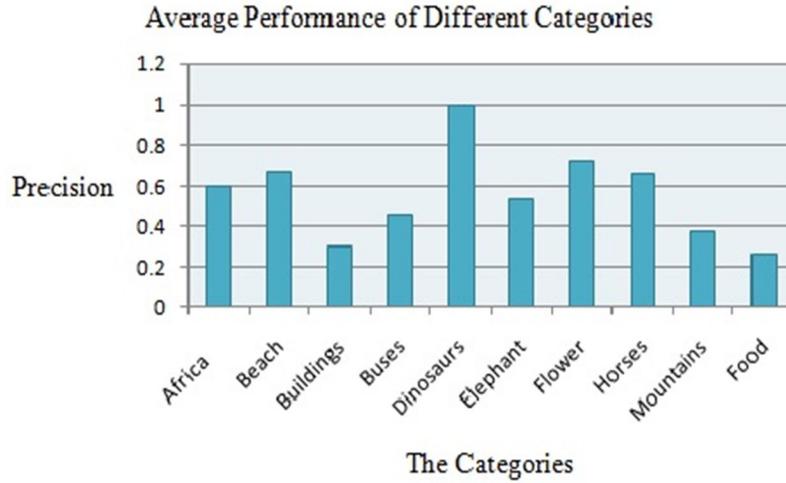


Figure 9: Average Performance of different categories

Figure 9 illustrated the average retrieval rate of all 10 classes achieved by the proposed system. The precision of dinosaur’s class is highest as it reaches to 1, while the food class has lowest precision, except food, building and buses the proposed system has better retrieval rate for all other classes.

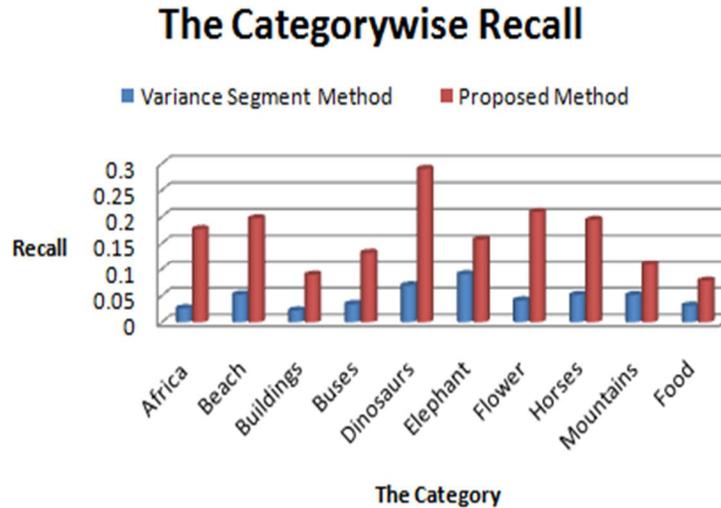


Figure 10: Average Recall of category wise

Figure 10, shows the category wise recall rate of the proposed method and variance segment method. It is obvious that proposed method has much better recall rate then the variance segment method.

4. CONCLUSION

A new CBIR system has been implemented to answer the shortcomings in the previous CBIR methods. Nowadays a lot of information in the form of digital content is easily accessible on the internet but finding the relevant image is a big problem using current CBIR systems. The proposed system used the Color Layout Descriptor (CLD) of MPEG-7 and wavelets packets to extract the feature set. CLD was selected for the color features and

texture features are extracted using wavelet packets. The performance of the system has been evaluated with the standard SIMPLIcity and other similar datasets. The system is also tested on the different top n image retrieval. During simulations it is observed that the proposed system outperforms, if the n is equal to 30 or less than 30. From the results, it can be clearly seen that the performance of the proposed CBIR system is better than the previous systems using sophisticated region, shape and texture matching techniques.

Acknowledgment:

The authors declare that they have no conflicts of interest in this research.

5. REFERENCES

- [1]. H. Zhi-Chun, et al., "Content-based image retrieval using color moment and Gabor texture feature," in International Conference on Machine Learning and Cybernetics (ICMLC) 2010, pp. 719-724.
- [2]. R. C. Veltkamp and M. Tanase, "Content-based image retrieval systems: A survey," Department of Computing Science utrecht University2000.
- [3]. C. Beecks, et al., "A Comparative Study of Similarity Measures For Content-Based Multimedia Retrieval," presented at the IEEE International Conference on Multimedia and Expo (ICME), 2010.
- [4]. N. Singhai and P. S. K. Shandilya, "A Survey On: Content Based Image Retrieval Systems " International Journal of Computer Applications vol. 4, July 2010.
- [5]. K. A. S. Abubacker and L. K. Indumathi, "Attribute Associated Image Retrieval and Similarity Reranking," in Proceedings of the International Conference on Communication and Computational Intelligence, 2010.
- [6]. C. B. Akgül, et al., "Content-Based Image Retrieval in Radiology: Current Status and Future Directions," Journal of Digital Imaging, 2010.
- [7]. Z.-C. Huang, et al., "Content Based Image Retrieval Using Color Moment and Gabor Texture Feature," in Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, 2010.
- [8]. L. Zhao and J. Tang, "Content-Based Image Retrieval Using Optimal Feature Combination and Relevance Feedback " presented at the International Conference on Computer Application and System Modeling, 2010.
- [9]. J. D.Pujari and A. S.Nayak, "Effect of Region Filtering on The Performance of Segmentation Based CBIR System," presented at the International Conference on Signal and Image Processing, 2010.
- [10]. J. E. E. d. Oliveiraa, et al., "MammoSys: A content-based image retrieval system using breast density patterns," computer methods and programs in biomedicine, p. 10, 2010.
- [11]. M. Broilo and F. G. B. D. Natale, "A Stochastic Approach to Image Retrieval Using Relevance Feedback and Particle Swarm Optimization," IEEE TRANSACTIONS ON MULTIMEDIA, vol. 12, p. 11, 2010.
- [12]. T. Sikora, "The MPEG-7 Visual Standard for Content Description An Overview," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, vol. 11, pp. 696-702, 2001.
- [13]. C. V. Royo, "Image-Based Query by Example Using MPEG-7 Visual Descriptors," Master, UNIVERSITAT POLITÈCNICA DE CATALUNYA, 2010.
- [14]. N. G. Rao, et al., "Texture Based Image Indexing and Retrieval," IJCSNS International Journal of Computer Science and Network Security vol. 9, pp. 206-210, 2009.
- [15]. M. Banerjee, et al., "Content-based image retrieval using visually significant point features," Elsevier, Fuzzy Sets and Systems, pp. 3323 – 3341, 2009.
- [16]. J. Z. Wang, et al., "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 23, 2011.
- [17]. Y. Chen and J. Z. Wang, "A region-based fuzzy feature matching approach to content-based image retrieval," IEEE Transactions on Image Processing Pattern Analysis and Machine Intelligence, vol. 24, pp. 1252-1267, 2002.
- [18]. H.H.Pavan Kumar Bhuravarjula, V.N.S Vijaya Kumar , "A NOVEL CONTENT BASED IMAGE RETRIEVAL USING VARIANCE COLOR MOMENT," International Journal of Computer and Electronics Research, Vol 1, Issue 3, October 2012.
- [19]. H. P. S and J. D. Jagadeesh Pujari, "Content Based Image Retrieval using Color Boosted Salient Points and Shape features of an image " International Journal of Image Processing, vol. 2, pp. 10-17, 2008.