DESIGN OF A FEEDBACK MECHANISM FOR CONTENT BASED IMAGE RETRIEVAL USING WK-MEANS CLUSTERING

Dr. M. Nester Jeyakumar
Head, Dept. of Computer Science& MCA, Loyola College, Chennai-34, India
Dr. Jasmine Samraj
Associate Professor, PG & Research Dept. Of Computer Science, Quaid-E-Millath Govt. College for Women (A), Chennai-2, India.

Abstract: Content-Based Image Retrieval (CBIR) with relevance feedback is the most emerging research field now. The power of dynamic learning ability can be used by subsequent incidents concerning adaptive knowledge. The proposed methodology can be applied to medical as well as non-medical images for retrieval. The similarity between the images is based on the visual content descriptors such as color shape and texture. The system proposed has three phases of the process; first, the native or primitive features are calculated. Second the indexing process is done with the desired class or group. Finally the retrieval and adaptive learning process takes place. When compared to the state of art methods, the proposed method in the field of CBIR gives prominent relevance results. The system has experimented with different images of the available datasets.

Keywords: Content Based Image Retrieval (CBIR), Relevance Feedback, Clustering, Similarity Measure.

Introduction: Content-Based Image Retrieval (CBIR) is a technique which uses visual contents to search images from large scale image databases according to users' interest. It has been an active and fast advancing research area since the 1990s. During the past decade, remarkable progress has been made in both theoretical research and system development.

Visual Content-Based Image Retrieval: Researchers from various communities of computer vision, database management, human-computer interface, and information retrieval were attracted to visual information management field. To overcome the difficulties faced in the text-based retrieval the new approach CBIR was introduced. The visual contents of the image such as color, shape, texture, and spatial layout are extracted from the images to represent and index the image. The visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. These extracted feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with a query image [1]. The similarity distance between the query image feature vectors and those of the images in the database are calculated, and the retrieval is performed using an indexing scheme. The indexing scheme provides an efficient way to search in the image database. To support efficient indexing, the dimensionality reduction techniques are applied to feature vectors [2]. The following sections cover the detailed process of CBIR.

Image Visual Content Descriptors: The content of images may include both visual and semantic content. General visual content includes color, texture, shape, spatial relationship, etc. An efficient visual content descriptor should be invariant on the random variance introduced by the imaging process [3]. These visual content descriptors can be either global or local. A global descriptor takes into account the visual features of the entire image, whereas a local descriptor uses the visual features of particular regions or objects to describe the image content. A better method is to divide the image into homogenous regions according to some criterion using segmentation algorithms. In the following section, will see some widely used techniques for extracting color, texture, shape and spatial relationship from images.

Color: Color feature is extensively utilized for the efficient extraction of the visual content of images. Before selecting an appropriate color description, color space must be determined. Each pixel of the given

image can be represented as a point in a 3D color space. RGB, Munsell, CIE L*a*b*, CIE L*u*v*, HSV (or HSL, HSB), and opponent color space are some of the commonly used color spaces.

Color moments have been used successfully in many retrieval systems. Color distributions in an image can be efficiently and effectively represented by the first order (mean), the second order (variance) and the third order (skewness) color moments. If the color pattern is unique compared with the rest of the data set, then the color histogram can be effectively used to represent the color content of the image. In [4] a different way of incorporating spatial information into the color histogram, color coherence vectors (CCV), due to its additional spatial information, it has been shown that CCV provides better retrieval results than the color histogram.

The color correlogram [5] has been not only used to characterize the color distributions of pixels, but it is also used to represent the spatial correlation of pairs of colors that are present. In a three-dimensional histogram, the first and the second dimension are the colors of any pixel pair, and the third dimension is used to depict their spatial distance.

Invariant color representation has been introduced to content-based image retrieval recently. In [6], a set of color invariants for object retrieval was derived based on the Schafer model of object reflection. Color Edge [7], Color Texture [8], are the color feature descriptors with various representation scheme. In [9], specular reflection, shape, and illumination invariant represented based on blue ratio vector (r/b, g/b, l) is given. In [10], surface geometry invariant color features are used. These invariant color features, when applied to image retrieval, may yield illumination, scene geometry and view geometry independent representation of color contents of images.

Texture: The texture is another important property of an image that is used for retrieval. Investigations in the field of pattern recognition and computer vision have resulted in various texture representations. Texture representation can fall into these two categories: structural and statistical. Structural methods, includes morphological operator and adjacency graph. Statistical methods, includes the Tamura feature, Markov random field, Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), Wold decomposition, fractal model, and multi-resolution E-filtering techniques such as Gabor and wavelet transform.

Grey Level Co-occurrence Matrices (GLCM) is one of the earliest methods for texture feature extraction [11]. The features such as Contrast, Correlation, Entropy, Homogeneity and Energy can be calculated from the co-occurrence values derived from the image with distance and direction.

The Tamura features [12], includes Coarseness, Contrast, Regularity, and Roughness, Directionality, Line likeness that is designed in agreement with psychological studies on the human perception of texture. Tamura features have been used in some early well-known image retrieval systems, such as QBIC [13, 14] and Photo book [15].

Wold decomposition [16, 17] provides another approach to describing textures regarding perceptual properties. The three Wold components, harmonic, evanescent, and in deterministic, correspond to periodicity, directionality, and randomness of texture respectively. Periodic textures have a strong harmonic component, highly directional textures have a strong evanescent component, and less structured textures tend to have more substantial in deterministic component.

The SAR model is an instance of Markov random field (MRF) models, which have been very successful in texture modeling in the past decades. Compared with other MRF models, SAR uses fewer parameters. The SAR model is not rotation invariant. To describe textures of different granularities, the multi-resolution simultaneous autoregressive model (MRSAR) [18] has been proposed to enable multi-scale texture analysis. An image is represented by a multi-resolution Gaussian pyramid with low-pass filtering and sub-sampling applied at several successive levels.

Another method used extensively for extracting texture features is the Gabor filter [19, 20]. Gabor filter proves to be optimal regarding minimizing the joint uncertainty in space. And the frequency is often used as an orientation and scale tunable edge and line detector.

Shape: Extraction of shape features of objects or regions in images has been used in many content-based image retrieval systems [11-24]. Compared with color and texture features, shape features are usually described after images have been segmented into regions or objects. Since robust and accurate image segmentation is difficult to achieve, the use of shape features for image retrieval has been limited to specialized applications where objects or regions are readily available. Classical shape representation uses a set of moment invariants. This central moment can be normalized to be scale invariant. Based on these moments, a set of moment invariants to translation, rotation, and scale can be derived [25].

The contour of a 2D object can be represented as a closed sequence of successive boundary pixels. The turning function or turning angle measures the angle of the counterclockwise tangents as a function of the arc-length according to a reference point on the object's contour. One major problem with this representation is variant to the rotation of object and the choice of the reference point. The minimum distance needs to be calculated over all possible shifts and rotations.

Fourier descriptors describe the shape of an object with the Fourier transform of its boundary. Three types of contour representations, i.e., curvature, centroid distance, and complex coordinate function, can be defined. The curvature at a point along the contour is defined as the rate of change in tangential direction of the contour.

The centroid distance is defined as the distance function between boundary pixels and the centroid of the object. The complex coordinate is obtained by merely representing the coordinates of the boundary pixels as complex numbers.

Circularity is computed with size and perimeter of an object. This value ranges between 0 (corresponding to a perfect line segment) and 1 (corresponding to a perfect circle). The major axis orientation can be defined as the direction of the largest Eigen vector of the second-order covariance matrix of a region or an object. The eccentricity can be defined as the ratio of the smallest Eigen value to the largest Eigen value.

Spatial Information: Objects or regions having similar texture or color properties can be distinguished easily by applying spatial constraints. For instance, parts of the clear blue sky and regions of ocean may have similar color histograms, but their spatial locations in images will be different. Therefore, the spatial location of regions (or objects) or the spatial relationship between multiple regions (or objects) in an image is very useful for searching images. The most widely used representation of spatial relationship is the 2D strings proposed by Chang et al [26].

The 2D G-string [27] cuts all the objects along their minimum bounding box and extends the spatial relationships into two sets of spatial operators. One defines local spatial relationships.

The other defines the global spatial relationships, indicating that the projection of two objects is disjoint, adjoin or located at the same position. Besides, 2D C-string [28] is proposed to minimize the number of cutting objects. 2D-B string [29] represents an object by two symbols, standing for the beginning and ending boundary of the object. In addition to the 2D string, spatial quad tree [30], and the symbolic image is also used for spatial information representation.

System Architecture:

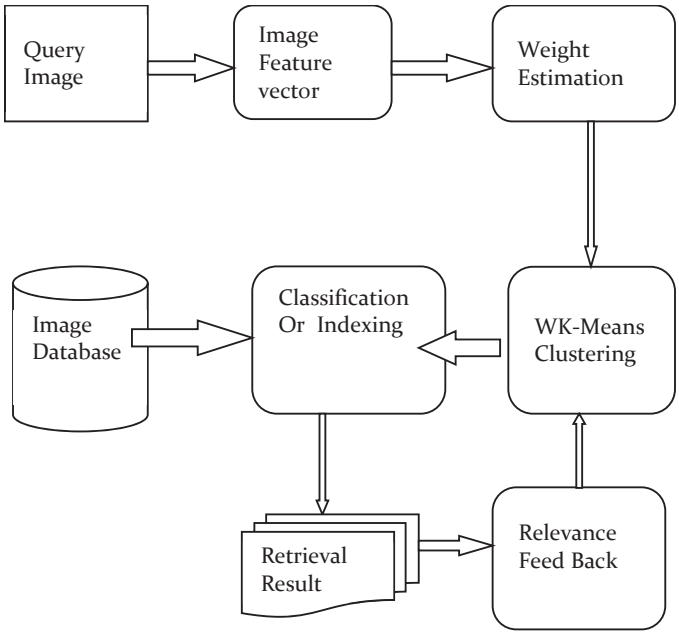


Figure 1: Proposed System Architecture

Data Set Description: The public WANG database is used for the experiment. The database comprises of 1000 images and which form 10 classes of 100 images each. Figure. 1 describes the proposed system architecture. All the 10 classes are used for relevance estimation by the given query image. It is assumed that the user is searching for images from the same class, and thus the remaining 99 images that belong to the same class are considered relevant and the images from all other classes are considered irrelevant. Sample images are shown below.

Query Image: The data set contains 1000 different images from public WANG Database. The query image is the essential and initial part of the CBIR system. The image which is taken for querying the database, is considered as the query image. Some examples of sample query images as shown in Figure. 2 a, b & c.

(a)

(b)

(c)

Figure 2: a, b & c: Sample Images

Image Feature Vector Specification: The given input image I is termed as the input vector to the clustering process. If the image consists of rows and columns then the generated feature vector to be in size where size l where l=m*n.

Weight Estimation &WK-Means clustering: K-means is one of the simplest unsupervised learning algorithms. The main idea is to classify a given data set into the number of K clusters. The goal is to define k centers, one for each cluster. These centers should be placed in a cunning way because different location causes a different result. So, it is better to separate and put them as far as possible away from each other. Each point that belongs to a given data set is associated with its nearest center in the next step. We need to re-calculate k new centroid. A new binding has to be done between the same data set points and the nearest center after finding the K new centroid. The resulting iteration, we may notice that the K centers change their location step by step until no more changes are done.

This algorithm primarily aims at minimizing an objective function known as squared error function with

$$W(S,C,w)=\sum_{k=1}^{K}\sum_{i\in I}\sum_{v=1}^{M}\mathrm{s}_{\mathrm{ik}}w_v^{\beta}(y_{iv}-c_{kv})^2$$

the weights given by [31]:

Subject to,

$$\begin{cases} S_{ik}\in\{0,1\} \\ \sum_{k=1}^{k}S_{ik}=1 \\ \sum_{v=1}^{M}W_v=1 \end{cases}$$

The weight w_v should be non-negative and add to the unity. ß is the user defined variable.

WK-Means updates the feature weights according to:

$$W_v=\frac{1}{\sum_{u\in V}\left[\frac{D_v}{D_u}\right]^{\frac{1}{\beta-1}}}$$

Where D_v is the sum of within-cluster variances of v weighted by clusters' cardinalities:

$$D_v=\sum_{k=1}^{K}\sum_{i\in I}(y_{iv}-c_{kv})^2 s_{ik}$$

The final updated centroid C is to be used for classifying the new query sample. Feature vector by identifying the minimum distance centroid is supposed to be that particular class.

Classification Or Indexing (Similarity Measure): The similarity between the cluster centroid and the input feature vector is based on L1 and L2 similarity measures.
The **L1 similarity** also known as L1 norm is calculated between the centroid and feature vector based on the formula given below.

$$\boldsymbol{Distance} = |X1 - X2| + \lceil Y1 - Y2 \rceil$$

The **L2 similarity** also known as Rular Distance is calculated between the centroid and feature vector based on the formula given below.

$$\boldsymbol{Distance} = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$

The minimum distanced cluster is activated as a classified class, and also it will be used for further retrieval process.

Relevance Feedback: With the process of clustering the image, each iteration has to refer the previous step knowledge in terms of the target class derived by the user. Certainly, this is a supervised learning mechanism, but the effectiveness of the system has been tremendously improved when combining with the unsupervised clustering process.

Consider, we have a user query and partial knowledge of known relevant and non-relevant documents based on the previous step clustering process. The algorithm proposes using the modified query

$$\overrightarrow{q_m} = \alpha\overrightarrow{q_0} + \beta\frac{1}{\lceil D_r \rceil}\sum_{\overrightarrow{d_j} \epsilon D_r} \overrightarrow{d_j} - \gamma\frac{1}{\lceil D_{nr} \rceil}\sum_{\overrightarrow{d_j} \epsilon D_{nr}} \overrightarrow{d_j}$$

Where q_o is the original query image vector, D_r and D_{nr} are the set of known relevant and non-relevant image vectors respectively, and α, β, and γ are weights attached to each term. If we would like to higher β and γ starting from q_o, the new query moves you some distance towards the centroid of the relevant image and some distance away from the centroid of the non-relevant image. This new query can be used for retrieval in the further step of iterations and classification.

Experiment And Analysis: A potentially useful method for retrieval should yield good accuracy with relatively fewer feature vectors. In this connection are listed here the accuracy, sensitivity, specificity and F-Measure which are listed in Table 1. The conclusion drawn is as follows. First, the proposed retrieval method's classification accuracy is higher when compared with existing work [32]. Second, the proposed approach has better retrieval accuracy in comparison to the existing work [32] as per performance measures given in Table1. The proposed approach yields a specificity of 84.2 % and a sensitivity of 88.2 %.

Table 1: Performance Analysis

WK-Means				
Class	Accuracy	Sensitivity	Specificity	F-Measure
1	0.895	0.900	0.843	0.851
2	0.968	0.840	0.834	0.946
3	0.845	0.850	0.861	0.810
4	0.962	0.950	0.872	0.953
5	0.916	0.801	0.84s2	0.865
6	0.951	0.905	0.925	0.947

7	0.824	0.690	0.878	0.891
8	0.890	0.934	0.820	0.912
9	0.932	0.979	0.820	0.923
10	0.971	0.973	0.765	0.943
Overall	**0.915**	**0.882**	**0.842**	**0.904**

However, our proposed approach is still able to retrieve all cases and has least false positive. Thirdly, the proposed approach has a better robustness in terms of illumination and camera characteristics. Finally, the proposed relevance feedback mechanism is useful in many image retrieval methods to make the retrieval efficient. Figure. 2 shows the sample query image class bus from WANG database. The proposed work has a better discrimination power than typical retrieval methods such as the existing work [32]. The Figure. 3 shows the curve (Receiver Operating Characteristic (ROC)) analysis of the proposed work with the existing method [32] which specifies the sensitivity and specificity of the proposed work. The value of specificity and sensitivity are given in table 1. The Figure. 4 shows the Precision and Recall analysis of proposed method with work [32]. The Table 1 contains specificity, sensitivity, accuracy and F-Measure with different class set. Figure. 5 a, b, c & d shows the sample screen shots of the proposed relevance feedback based content based image retrieval system.

Figure 2: Sample Query Image Class Bus from WANG Database

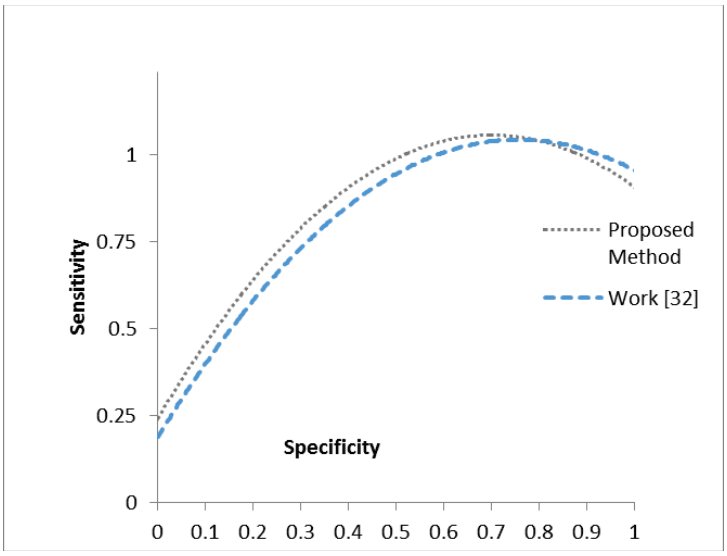


Figure 3: ROC Analysis of Proposed Method With Work [32]

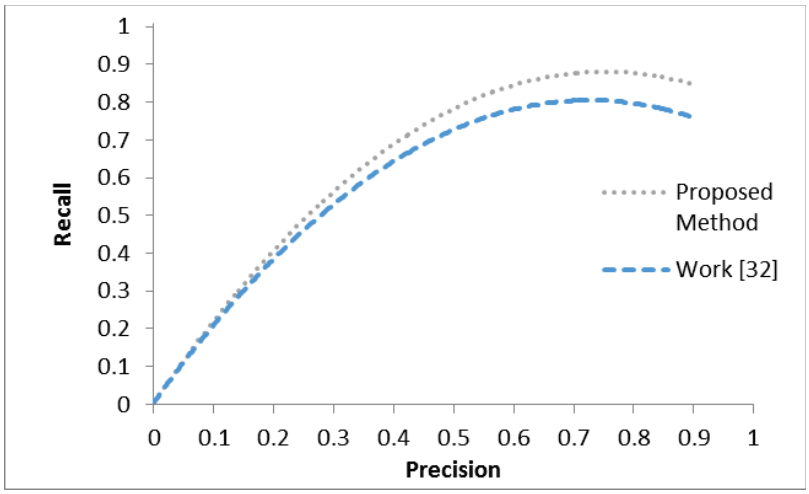
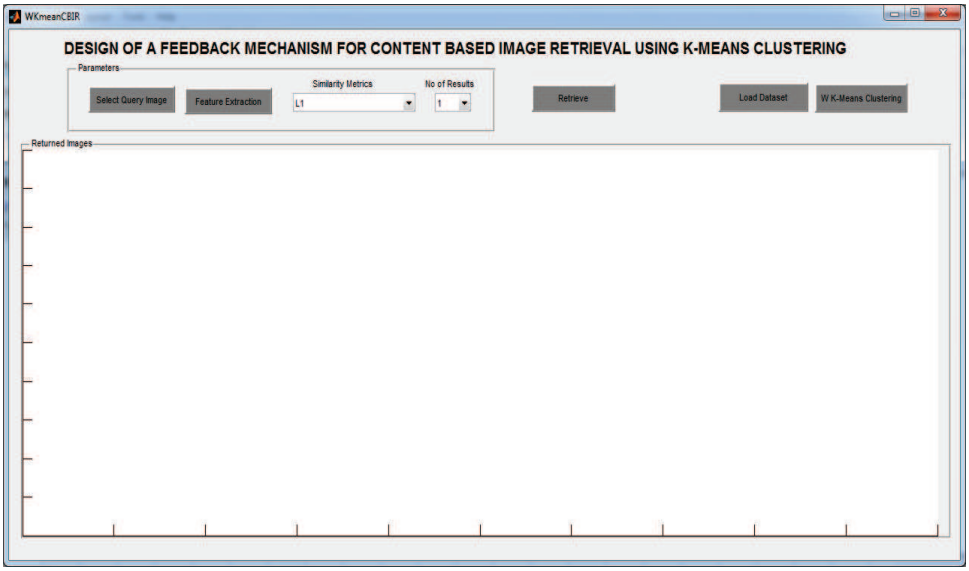
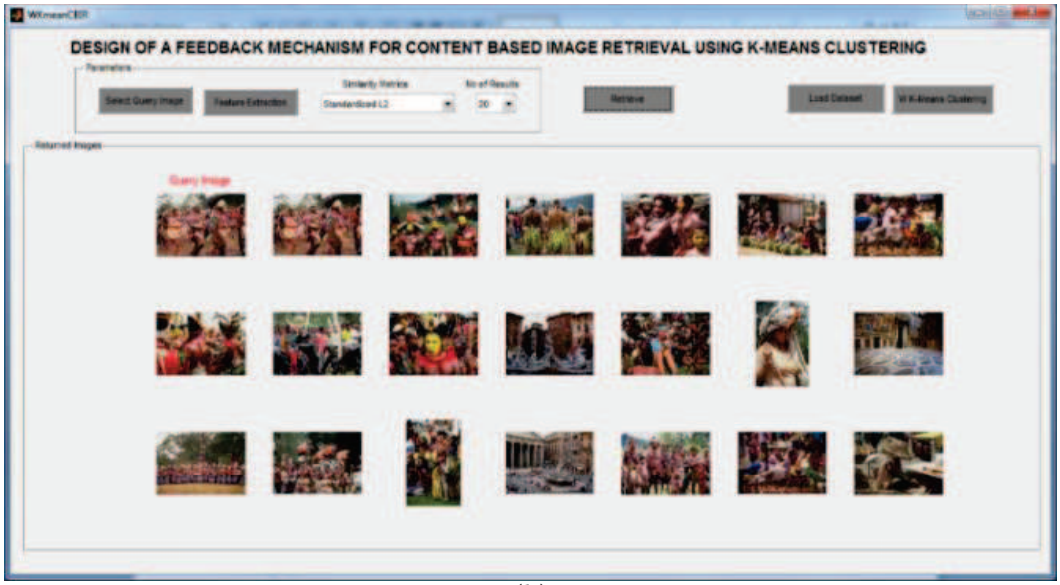


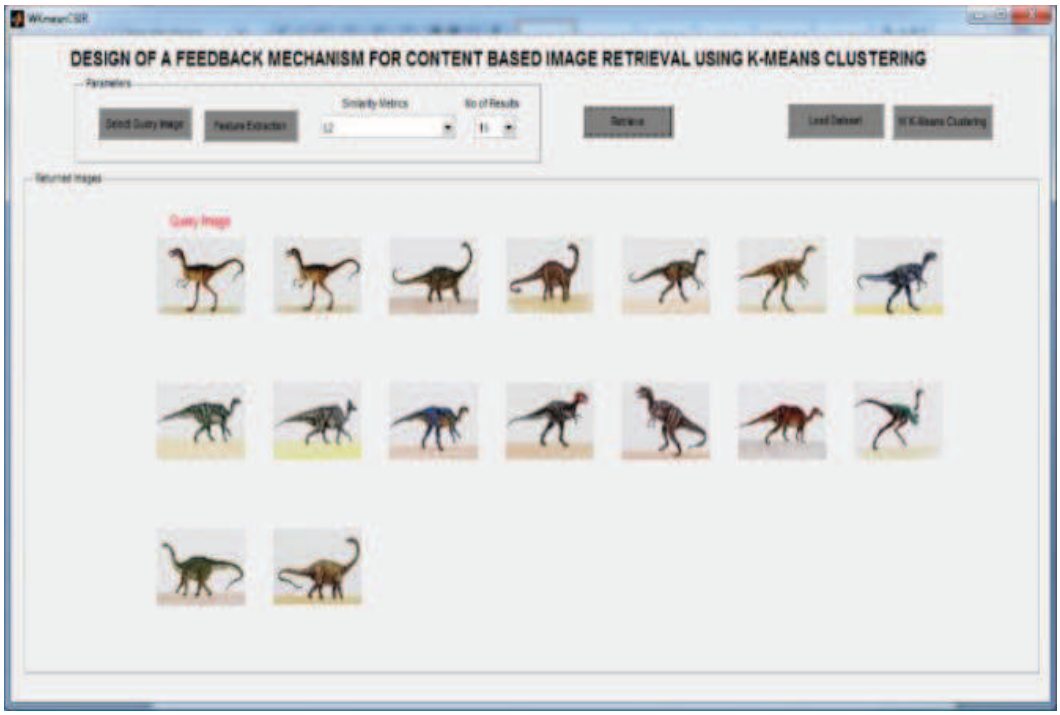
Figure 4: Precision Recall Analysis of Proposed Method With Work [32]



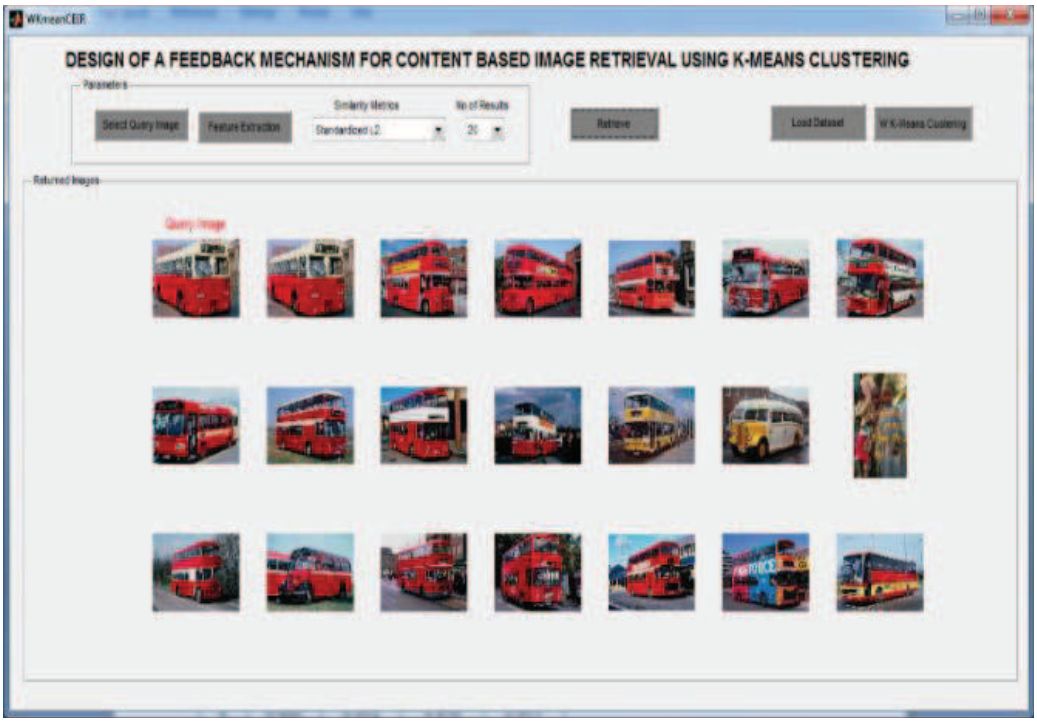
(a)



(b)



(c)



(d)

Figure 5 a, b, c & d: Sample screen shots of the proposed system

Conclusion: The method provided herein is a relevance feedback based CBIR which is a novel approach. Though CBIR is an emerging research, we cannot expect the entire upheaval of existing techniques with CBIR. But instead, can propose a complement that provides better results. The clustering information while grouping leads to adaptive knowledge for training the retrieval system in further effectiveness. The main aim of this work was to improve relevancy of a CBIR application by allowing the system to retrieve

images with earlier learned knowledge. The proposed CBIR method presented delivers the equivalent results. This mechanism has improved the results in terms of performance measures that are stated.

References:

1. G. Pass, and R. Zabith, "Histogram refinement for content-based image retrieval," IEEE Workshop on Applications of Computer Vision, pp. 96-102, 1996.
2. J. Huang, et al., "Image indexing using color correlogram," IEEE Int. Conf. on Computer Vision and Pattern Recognition, pp. 762-768, Puerto Rico, June 1997.
3. T. Gevers, and A.W.M.Smeulders, "Pictoseek: Combining color and shape invariant features for image retrieval," IEEE Trans. on image processing, Vol.9, No.1, pp102-119, 2000.
4. T. Gevers, and H. Stokman, "Classifying color edges in video into shadow geometry, highlight, or material transitions," IEEE Transactions on Multimedia, vol. 5, no. 2, pp. 237-243, Sep. 2003.
5. H. Guan, and S. Wada, "Flexible color texture retrieval method using multi resolution mosaic for image classification," Proceedings of the 6th International Conference on Signal Processing, vol. 1, pp. 612-615, Feb. 2002.
6. G. D. Finlayson, "Color in perspective," IEEE Trans on Pattern Analysis and Machine Intelligence, Vol.8, No. 10, pp.1034-1038, Oct. 1996.
7. T. Gevers, and A. W. M. Smeulders, "Content-based image retrieval by viewpoint-invariant image indexing," Image and Vision Computing, Vol.17, No.7, pp.475-488, 1999.
8. R. Haralick, K. Shanmugam, and I. Dinstein, (1973) "Textural Features for Image Classification", IEEE Trans. on Systems, Man and Cybernetics, SMC-3(6):610-621.
9. H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," IEEE Trans. On Systems, Man, and Cybernetics, vol. Smc-8, No. 6, June 1978.
10. M. Flickner, H. Sawhney, W. Niblack, et al., "Query by image and video content: The QBIC system." IEEE Computer, Vol.28, No.9, pp. 23-32, Sept. 1995.
11. W. Niblack et al., "Querying images by content, using color, texture, and shape," SPIE Conference on Storage and Retrieval for Image and Video Database, Vol. 1908, pp.173-187, April 1993.
12. A.Pentland, R.W. Picard and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases," Proc. Storage and Retrieval for Image and Video Databases II, Vol. 2185, San Jose, CA, USA February, 1994.
13. J. M. Francos. "Orthogonal decompositions of 2D random fields and their applications in 2D spectral estimation," N. K. Bose and C. R. Rao, editors, Signal Processing and its Application, pp.20-227. North Holland, 1993.
14. F. Liu, and R. W. Picard, "Periodicity, directionality, and randomness: Wold features for image modeling and retrieval," IEEE Trans. on Pattern Analysis and Machine Learning, Vol. 18, No. 7, July 1996.
15. J. Mao, and A. K. Jain, "Texture classification and segmentation using multiresolution simultaneous autoregressive models," Pattern Recognition, Vol. 25, No. 2, pp. 173-188, 1992.
16. J. G. Daugman, "Complete discrete 2D Gabor transforms by neural networks for image analysis and compression," IEEE Trans. ASSP, vol. 36, pp. 1169-1179, July 1998.
17. A. K. Jain, and F. Farroknia, "Unsupervised texture segmentation using Gabor filters," Pattern Recognition, Vo.24, No.12, pp. 1167-1186, 1991.
18. J. E. Gary, and R. Mehrotra, "Shape similarity-based retrieval in image database systems," Proc. Of SPIE, Image Storage and Retrieval Systems, Vol. 1662, pp. 2-8, 1992.
19. W. I. Grosky, and R. Mehrotra, "Index based object recognition in pictorial data management," CVGIP, Vol. 52, No. 3, pp. 416-436, 1990.
20. H. V. Jagadish, "A retrieval technique for similar shapes," Proc. of Int. Conf. on Management of Data, SIGMOID'91, Denver, CO, pp. 208-217, May 1991.
21. D. Tegolo, "Shape analysis for image retrieval," Proc. of SPIE, Storage and Retrieval for Image and Video Databases -II, no. 2185, San Jose, CA, pp. 59-69, February 1994.
22. M. K. Hu, "Visual pattern recognition by moment invariants," in J. K. Aggarwal, R. O. Duda, and A. Rosenfeld, Computer Methods in Image Analysis, IEEE computer Society, Los Angeles, CA, 1977.
23. S. K. Chang, Q. Y. Shi, and C. Y. Yan, "Iconic indexing by 2-D strings," IEEE Trans. on Pattern Anal. Machine Intell., Vol.9, No.3, pp. 413-428, May 1987.

24. S. K. Chang, E. Jungert, and Y. Li, "Representation and retrieval of symbolic pictures using generalized 2D string", Technical Report, University of Pittsburgh, 1988.
25. S. Y. Lee, and F. H. Hsu, "2D C-string: a new spatial knowledge representation for image database systems," Pattern Recognition, Vol. 23, pp 1077-1087, 1990.
26. S. Y. Lee, M.C. Yang, and J. W. Chen, "2D B-string: a spatial knowledge representation for image database system," Proc. ICSC'92 Second Int. computer Sci. Conf., pp.609-615, 1992.
27. H. Samet, "The quadtree and related hierarchical data structures," ACM Computing Surveys, Vol.16, No.2, pp.187-260, 1984.
28. Soo Beom Park, Jae Won Lee, Sang Kyoon Kim , " Content-based image classification using a neural network", Pattern Recognition Letters 25 (2004) 287-300
29. J. Z. Huang, Michael k. Ng, H. Rong, Z. Li, "Automated Variable Weighting in K-Means Type Clustering", IEEE Trans. Patrn. Anlys. And Mach. Inteln.,Vol. 27, No. 5., 2005
30. Jayant Mishra, Anubhav and Kapil Chaturvedi, An Unsupervised Cluster-based Image Retrieval Algorithm using Relevance Feedback, International Journal of Managing Information Technology (IJMIT) Vol.3, No.2, May 2011
31. P. Jeyanthi, V. Jawahar Senthil Kumar, Image Classification by K-means Clustering, Advances in Computational Sciences and Technology ISSN 0973-6107 Volume 3 Number 1 (2010)
32. Xiang Sean Zhou, Thomas S. Huang ,Relevance feedback in image retrieval: A comprehensive review, Multimedia Systems :536-544 (2003), Digital Object Identifier(DOI) 10.1007/ s00530-002-0070-3

***