IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Comparative Analysis in Content Based Image Retrieval System Using Color and Texture

K. Nirmala[*1], A. Subramani[2]

[*1]Department of Applied Science, K.S. Rangasamy College of Technology, Tiruchengode- 637 215, Tamil Nadu, India

[2]Department of Computer Applications, K.S.R College of Engineering, Tiruchengode- 637 215, Tamil Nadu

nirma.83@gmail.com

Abstract

Searching digital images from a large database can be solved by the Content Based Image Retrieval System based on the extraction of features such as color, texture and shape of the image. CBIR is the most efficient image retrieval method. Most of techniques for CBIR use numerical representations called feature vectors to allow content-based searching in large image collections. Analyzing images based on the colors is one of the most widely used techniques since the color does not depend on image size or orientation. Texture analysis is mainly used to identify the image given by the shape, size, brightness etc. The proposed method uses image features such as color and texture to retrieve the images from database. In addition, it uses auto color correlogram for identifying images based on the feature and Spatial texture for identifying images based on the texture. Moreover, this system has operated on a Corel database containing 1000 general-use color images. The study shows that the performance and accuracy of retrieving images in the large database.

Keywords: Color, Texture, Auto Color Correlogram (ACC), Spatial Texture.

Introduction

Content-based image retrieval (CBIR) systems mainly involves research on databases and image processing handling problems which varies from storage issues to user friendly interfaces [1].CBIR method has attracted users in image retrieval. CBIR process comprised of four steps such as data collection, building up the feature database, searching in the database, arranging the images in a particular order and deal with the results of the image retrieval [2].Several images are stored on the database. Image features such as color, texture, shape and structure are extracted from the images. Relevant images are retrieved from the database based on the similarity of the image features. Thus a CBIR system helps the user to retrieve the images which are relevant to the contents of the query image.

In CBIR systems, the images in the database are labeled by the feature vectors. These images can be extracted by means digital image processing techniques. The input query to a database is specified by an image in CBIR systems. The query's feature vector is computed and the closest items which are related to the query image in the database is computed by means of specific algorithms and the final output query is returned to the user.

Color is mainly associated with chromatic attributes of an image. Color is a part of the three dimensional coordinate system. The appropriate color space should be selected in order to identify the color based images. RGB (Red, Green, Blue), CMY(Cyan, Magenta, Yellow) and LHS (Luminance, Hue and Saturation) are the most popular color spaces[2]. The proposed method uses RGB color space for image identification.

Color based image retrieval can be extracted in many ways like Histogram, Color moments, Color Correlogram, color coherence vector etc [3-5]. The most commonly used color based image retrieval is color histogram which has limited discriminative power as it ignores the spatial organization of colors. Probability of finding color pairs at fixed pixel distances is known as correlation of colors which is used to provide an efficient spatial retrieval feature.

Several schemes for using spatial information about colors to improve upon the histogram method have been proposed recently [6]. Common approach is to divide images into sub regions and impose positional constraints on the image comparison (image partitioning). Another approach is to augment histograms

with local spatial properties (histogram refinement).The proposed method uses auto color correlogram for identifying the images based on the color.

Texture is the main property for the recognition of images. It is a repeated pattern of pixels over a spatial domain, of which the addition of noise to the patterns and their repetition frequencies result in textures that can appear to be random and unstructured. Its properties are the visual patterns in an image that have properties of homogeneity that do not result from the presence of only a single color or intensity. [7]. Texture is a natural property of surfaces and it provides visual patterns of the image. It contains vital information regarding the structural arrangement of the surface (example clouds, leaves bricks). It also gives the relationship between the surface and external environment [8].

Texture is difficult to examine by considering the value of a single pixel. It occurs mainly due to the variation in a neighborhood of pixels. There are several methods used for texture extraction. Feature extraction is used to measure local texture properties in an image, followed by boundary detection or segmentation to group the features into similar regions [9]. Once the texture features are extracted and feature vectors are formed, then the region in an image is grouped by texture segmentation which has similar texture properties. One aspect of texture is concerned with the spatial distribution among the gray levels of the pixels in a local area. The proposed method uses spatial texture for identifying texture properties.

The paper can be organized as follows: Section II describes the related works involved in content based image retrieval, Section III describes the methodology used to retrieve the images, and Section IV describes Experimental results obtained by using proposed methodology.

Methodology

Visual Image retrieval in large database is a complex task because different types of data can be associated with the images. CBIR method resolves this problem. The proposed methodology uses three methods to retrieve the images.

1. Color- based Image Retrieval
2. Texture- based Image Retrieval
3. Color and Texture based Image Retrieval
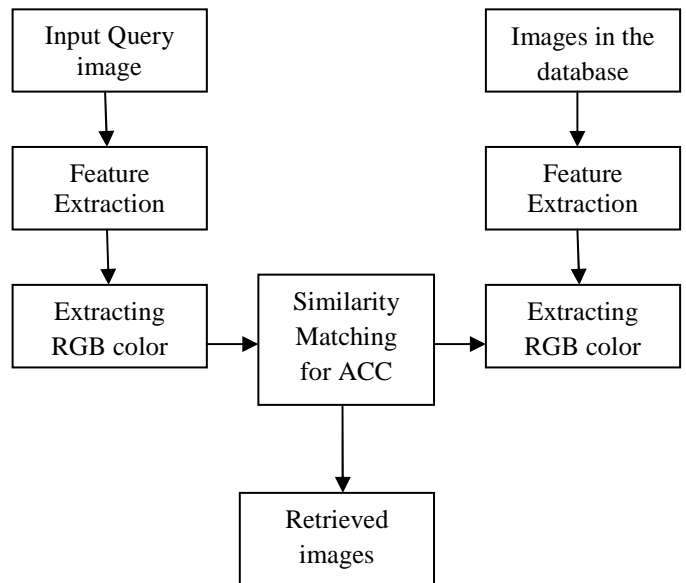
Color- based Image Retrieval

Color-based retrieval has been evolved from simple statistical measures such as average color to color histograms and spatial color descriptors. The proposed method uses Auto color correlogram for identifying color-based images. The method used for extracting relevant images is shown in the figure 1.

Auto Color Correlation Algorithm:

An auto color correlation defines how to compute the mean color of all pixels of color C_j at a distance k-th from a pixel of color C_i in the image. [10] Formally, the ACC of image $\{I(x,y), x = 1,2,\dots,M, y = 1,2,\dots,N\}$ is defined as

$$ACC(i,j,k) = MC_j\gamma_{C_iC_j}^{(k)}(I)$$

$$= \left\{ r_{mcj}\gamma_{C_iC_j}^{(k)}(I), g_{mcj}\gamma_{C_iC_j}^{(k)}(I), b_{mcj}\gamma_{C_iC_j}^{(k)}(I) \,\middle|\, c_i \neq c_j \right\} \qquad (1)$$

where the original image $I(x,y)$ is quantized to m colors $C_1, C_2, \dots, C_m$ and the distance between two pixels $k \in [min\{M,N\}]$ is fixed a priori. Consider MC_j be the RGB value of color m in an image I.

Fig 1. Steps involved in Color-based image retrieval

The mean colors are defined as follows:

$$r_{mcj}\gamma_{C_iC_j}^{(k)}(I) = \frac{\prod_{c_i,rc_j}^{(k)}(I)}{\prod_{c_i,c_j}^{(k)}(I)} |c_i \neq c_j \qquad (2)$$

$$g_{mcj}\gamma_{C_iC_j}^{(k)}(I) = \frac{\prod_{c_i,gc_j}^{(k)}(I)}{\prod_{c_i,c_j}^{(k)}(I)} |c_i \neq c_j \qquad (3)$$

$$b_{mcj}\gamma_{C_iC_j}^{(k)}(I) = \frac{\prod_{c_i,bc_j}^{(k)}(I)}{\prod_{c_i,c_j}^{(k)}(I)} |c_i \neq c_j \qquad (4)$$

where denominator $\prod_{c_i,xc_j}^{(k)}(I)$ is the total of pixels values of color C_j at distance k from any pixel of color C_i when x is RGB color space and denoted $C_j \neq 0$. N represents the number of accounting color C_j from color C_i at distance k is computed as follows:

$$N = \prod_{c_i,c_j}^{(k)}(I) \qquad (5)$$

$$= \begin{cases} P(x_1,y_1) \in C_i | P(x_2,y_2) \in C_j; \\ k = \text{MIN}\{|x_1 - x_2|, |y_1 - y_2|\}; \end{cases}$$

By reducing the size of color correlogram from $O(m^2d)$ to $O(3md)$, ACC can be able to find the local spatial correlation between color. To decrease the storage space required and increase the speed of retrieval, the size of ACC can be reduced from O (3md) to O (m). By using this algorithm, dominant RGB peaks values in any color bins are captured. The dominant elements are compared inorder to reduce the feature storage amount and speed of retrieval while processing similarity calculation of the two images.

```
For every K distance {
    For every X position
    For every Y position {
        Cᵢ ← current pixel
        While(C ⱼ ←
            Get neighbors pixel of  Cᵢ at
                distance K)
        {
        For every color Cₘ {
        If (Cₘ = Cᵢ and Cᵢ ≠ Cⱼ){
            countColor++
            colorR[Cₘ ] = colorR[Cₘ ] + colorRCⱼ
            colorG[Cₘ ] = colorG[Cₘ]+ colorGCⱼ
            colorB[Cₘ] = colorB[Cₘ] + colorBCⱼ
            }
        }
        }
    }
    meanColorR = sum ( colorR[Cₘ ])/countColor
    meanColorG = sum ( colorG[Cₘ ])/countColor
    meanColorB = sum ( colorB[Cₘ ])/countColor
}
```

The similarity of binary codes for auto color correlation can be measured using intersection technique. It measures the similarity of binary codes for the same color between the query and model images. Consider $Bc_m(I) = b_1^r, b_2^r, \ldots, b_m^r;\ b_1^g, b_2^g, \ldots, b_m^g;\ b_1^b, b_2^b, \ldots, b_m^b;$ represents the binary code of auto color correlation colors to color Cm in RGB space of query image I, then the intersection result of query image and model image concerning color Cm should be calculated.

The proposed method first computes the mean pixel value of the whole small block (4x4) and it compares each pixel to the block mean. If the pixel value is greater than or equal to the block mean, respective pixel position of the bitmap will have the value 1, otherwise it will be assigned as 0. When the RGB values in a color bin of ACC exceed a given threshold, then the bin is classified as effective, else it is classified a non-effective. Binary "1" is assigned to effective bin and, binary "0" is assigned to non effective bin. Thus by using feature vector of ACC in RGB color space, the accuracy of retrieval process can be improved.

Finally, system performs the similarity between the query image and images in the database and displays the retrieved images.

Texture- based Image Retrieval

Image texture can be defined by the number and types of its primitives and the spatial organization or layout of its primitives. There are several approaches for extracting texture from an image. The proposed method uses Spatial texture for identifying texture based image retrieval. The method used for extracting relevant texture images is shown in the figure 2.
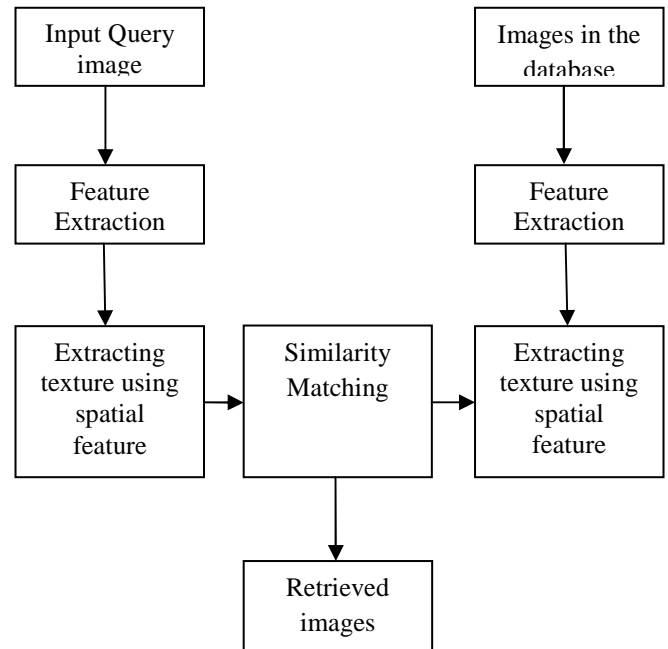
Fig 2. Steps involved in Texture-based image retrieval

Identifying Texture primitive

Let I be a $M \times N$ image. Divide the image into $m \times m$ pixel non-overlap blocks. The mean value μ and the standard deviation σ of gray in an image block is calculated for each block based on the following equation[11].

$$\mu = \frac{\sum_{\forall i,j} p(i,j)}{4} \qquad (6)$$

$$\sigma = \frac{\sum_{\forall i,j} \|p(i,j) - \mu\|}{4} \qquad (7)$$

where p(i,j) is the gray value of the pixel located in (i,j) for image I . By using the principle of BTC Block Truncation Code, make pixel values in each block equal to "1" if the gray value of the pixel is bigger than μ . Else make the pixel value as "0". A series of binary blocks are gained by this way. Binary blocks represent the texture feature for image blocks. These binary blocks is also states the shape distribution to some extent. Image blocks whose standard deviation is smaller than threshold β are regarded as even blocks and make their primitive value as "0". Else the primitive value is calculated using the method described above.

Spatial Feature Identification of the Texture Primitive

The type of texture primitive should be defined first. Then an image of $M \times N$ is corresponding to a matrix of $[M/m] \times [N/m]$ is denoted by P. $P(x,y)$ is the index of the texture primitive which is located in (x,y) in P . For certain kind of texture primitive in $P(x,y)$, keep its value and make other values equal to zero to extract the spatial information of the texture primitive. Then the spatial distribution map of texture primitive is constructed.

$$A_i = \{(x,y)| (x,y) \in P, P(x,y) = i, 0 \le i \le 2^{m*m} - 1\}$$

be the set of points with index i in P and $|A_i|$ be the number
of elements in A_i. Let $C_i = (x_i, y_i)$ be the centroid. x_i and y_i are computed as follows:

$$x_i = \frac{1}{|A_i|} \sum_{(x,y) \in A_i} x; \qquad (8)$$

$$y_i = \frac{1}{|A_i|} \sum_{(x,y) \in A_i} y; \qquad (9)$$

Consider r_i be the radius of the point whose index i is defined by,

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \qquad (10)$$

The sum of the distance between all points whose index is i and centroid can be measured as follows:

$$R_i = \sum r_i = \sum_{(x,y) \in A_i} \sqrt{(x - x_i)^2 + (y - y_i)^2} \qquad (11)$$

By using this method, texture primitives can be extracted.

Similarity matching :

The similarity distance between the two images is given by

$$D(B,Q) = \sqrt{(b_1 - q_1)^2 + (b_2 - q_2)^2 + \cdots (b_{2^{m*m}-1} - q_{2^{m*m}-1})^2} \qquad (12)$$

where, B is an image in the database and Q is a query image. $B = (b_1, b_2, \ldots b_{2^{m*m}-1})$ and $Q = (q_1, q_2, \ldots q_{2^{m*m}-1})$ are the spatial distribution feature of image B and Q respectively.

Experimental Results

This experiment is evaluated using MATLAB for the famed coral dataset. Corel database [12] contains large number of images of various contents ranging from animals and outdoor sports to natural images. The main aim of this proposed method is to detect the maximum number of relevant images for the given user input query. The accuracy of the image can be calculated by the following formula:

$$Accuracy = \frac{N-X}{N} * 100$$

where N is number of relevant images in the database which are known to the user and X is the number of irrelevant images in the database which are known to the user.
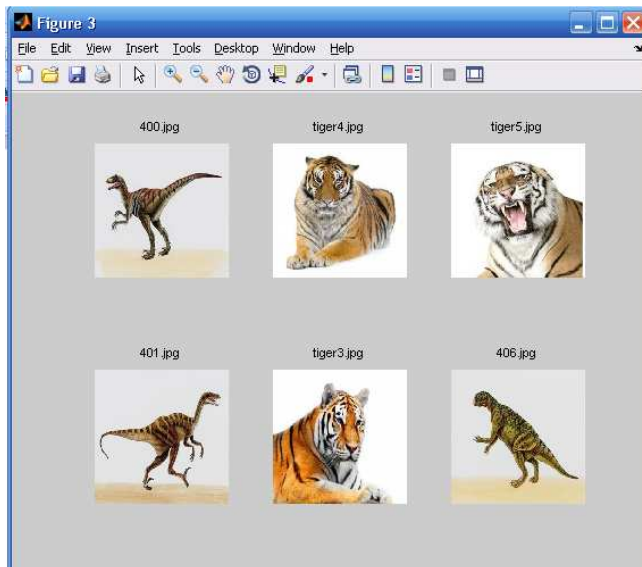
Color based Image Retrieval:

The user will give input query image in the system. The image is searched automatically in the database based on the color feature and detected images will be displayed to the user based on the auto color correlogram. The input query is shown in the figure 3. For example consider bus as an input query given by the user.
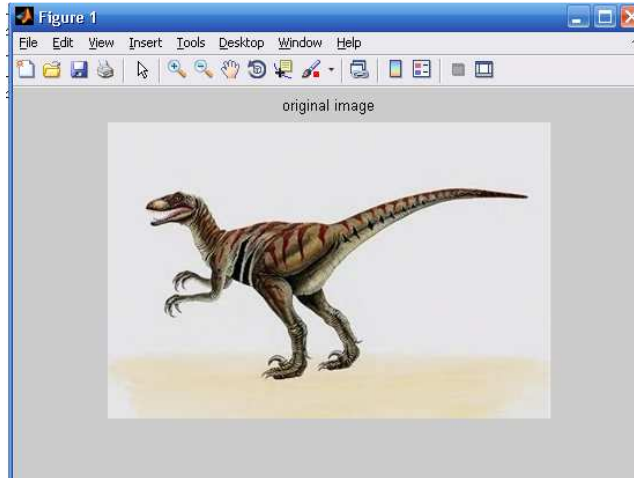
Fig 3. Input Query image

The similarity measure between the images depends on the techniques used for feature extraction. The similarity of binary codes for auto color correlation method use intersection technique to measure the same color between the query image and the database images.

The system detects the relevant images of the bus in the database based on the auto color correlation algorithm and displays the detected image as an output to the user. The retrieved images are shown in the figure 4.The accuracy of the retrieved images obtained by the auto color correlation algorithm is 67%. Thus the images are retrieved in color based image retrieval method and the output is displayed to the user.
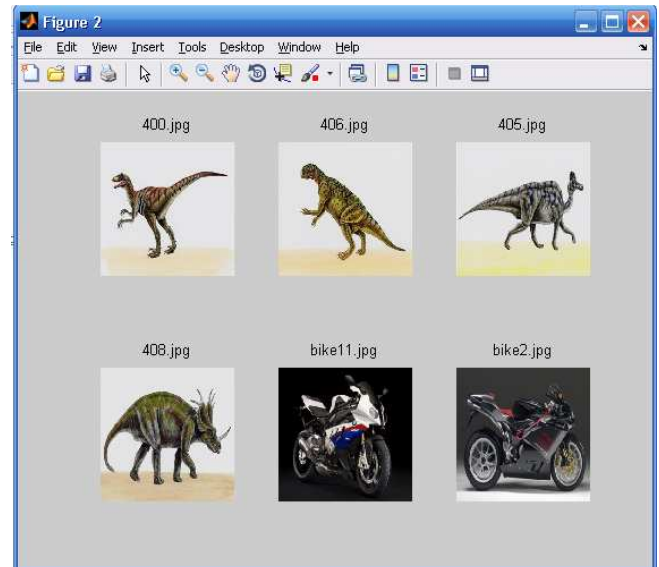

Fig 4. Retrieved images based on the color

Texture based Image Retrieval:

Based on the texture given by user's input query image The system search images in the database. The images are detected in the database based on the texture

and the output images are displayed to the user. The system finds the images based on the spatial texture primitive method.

Fig 5. Retrieved images based on the texture

Figure 5 reveals the output of the retrieved images for the given input query as shown in the figure 3. The accuracy of the retrieved images using the proposed method is 67%

Performance Evaluation

The dataset comprises of six different images. The corresponding accuracy of the query images and time taken to display these images has been studied for color-based, Texture-based and color and texture-based image retrieval. The accuracy and the time values obtained of the proposed three approaches is shown in the table 1.

Table 1: Accuracy And Time Comparison Of The Proposed Three Approaches

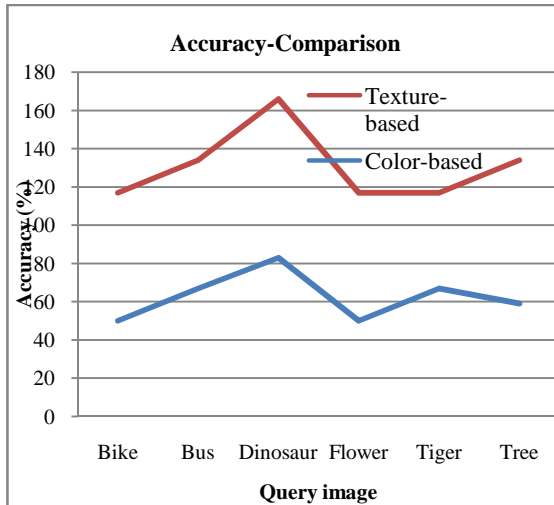| Query Image	Color- based		Texture-based	
	Acc (%)	Time (sec)	Acc (%)	Time (sec)
Bike	50	17	67	68
Bus	67	25	67	147
Dinosaur	83	32	83	131
Flower	50	22	67	126
Tiger	67	22	50	133
Tree	59	33	75	144
Average	61	25.1	68.7	124.8

Accuracy-Comparison

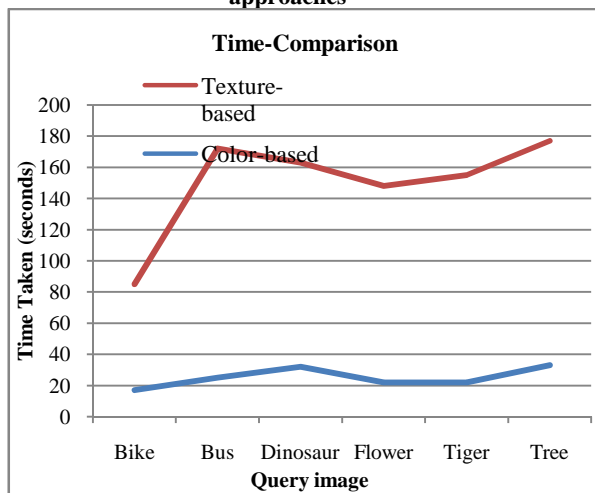Fig 7. Accuracy Comparison of the proposed two approaches

Time-Comparison

Fig 8. Time Comparison of the proposed two approaches

Figure 7 shows the accuracy comparison of the proposed approaches. Color & Texture based approach gives maximum accuracy for all the six datasets. Texture based approach gives maximum results than Color based approach.

Figure 8 shows the time comparison of the proposed approaches. Color based approach takes less time than all the proposed three approaches. Color & Texture based approach takes maximum time than all the three approaches.

It is clear from the experimental results that the color and texture based retrieval method achieves higher accuracy than both the proposed color based retrieval and texture based retrieval methods.

The experimental results proved that the proposed method achieves higher accuracy and better performance in retrieving the images.

Conclusion

Detecting relevant images from a huge database is the major problem nowadays. Content based Image Retrieval methods are used for this purpose. The study retrieves the images based on color feature and texture feature. The auto color correlation algorithm and spatial texture primitive algorithms are used in the proposed method to detect the color feature and texture feature respectively. It is evident from the experimental results by the texture properties achieves maximum accuracy for the Corel Dataset and can be applicable for other dataset also. In future, new algorithms can be integrated with the proposed method to detect the images accurately. The other image features like shape or other can be combined thus that the detection of the most relevant images will become effective.

References

[1] M. Worring, S. Santini, A. Gupta, and R. Jain. "Content-Based Image Retrieval at the End of the Years". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no.12 pp.1349–1380,December 2000.

[2] Nidhi Singhai, K. Shandilya "A Survey On: Content Based Image Retrieval Systems", International Journal of Computer Applications vol 4, no.2, pp.975 – 8887 , July 2010.

[3] C.H. Lin, R.T. Chen and Y.K. Chan, "A smart content-based image retrieval system based on color and texture feature", Image and Vision Computing vol.27, pp.658–665, 2009.

[4] G. Pass and R. Zabih, "Refinement Histogram for Content-Based Image Retrieval", IEEE Workshop on Application of Computer Vision, pp. 96-102. 1996.

[5] Kalyan Roy and Joydeep Mukherjee, "Image Similarity Measure using Color Histogram, Color Coherence Vector, and Sobel Method", International Journal of Science and Research (IJSR), India Online ISSN: 2319-7064, Volume 2 Issue 1, January 2013.

[6] Jing Huang, S Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, Ramin Zabih, "Image Indexing Using Color Correlograms", In IEEE Conference on Computer Vision and Pattern Recognition, pp. 762-768, 1997.

[7] Mr. Milind et al, "Efficient Content Based Image Retrieval Using Color and Texture", International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013 121 ISSN 2229-5518

[8] B. Ramamurthy and K.R. Chandran, "Content Based Medical Image Retrieval with Texture

Content Using Gray Level Co-occurrence Matrix and K-Means Clustering Algorithms", Journal of Computer Science vol. 8 ,no.7, pp. 1070-1076, ISSN 1549-3636, 2012 .

[9] P.P. Buch, M.V. Vaghasia, S.M. Machchchar, "Comparative analysis of content based image retrieval using both color and   texture", Engineering(NUiCONE),   Nirma   University International Conference, 2011, pp. 1-4.

[10] Anucha Tungkasthan, Sarayut Intarasema, Wichian   Premchaiswadi,   "Spatial   Color Indexing using ACC Algorithm", Seventh International   Conference   on   ICT   and Knowledge Engineering, 2009.

[11] Zhao Shan, Hai-Xia Zhai, "Image retrieval based on spatial feature of texture Primitive", Application research of computers, vol. 26, no.1,pp.378-379, 2009.

[12] Corel 1000 and Corel 10000 image database. [Online].   Available: http://wang.ist.psu.edu/docs/related.shtml.