ISSN No. 0976-5697

Volume 4, No. 8, May-June 2013

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Usage of Color and Texture Features for Natural Image Indexing and Retrieval

N.Magesh (M.E.)
Assistant Professor, Dept of Computer Science and Engineering, Institute of Road and Transport Technology, Erode-638 316.
nallasamymagesh@gmail.com

Dr.P.Thangaraj (Ph.D.)
Professor and Head, Dept of Computer Science and Engineering,
Bannariamman Institute of Technology, Sathyamangalam – 638 401.

Abstract: The novel approach combines color and texture features for content based image retrieval (CBIR). This paper is used to retrieve the images from the huge collection of image databases. The image retrieval performed with basic concepts are not reaching to the user specifications and not attracted to the user, So a lot of research interest in recent years with new specifications such as feature indexing techniques are used in retrieval. The proposed system has focused on the computing the average and standard deviation value for color and entropy and tamura features for texture. By using above features, retrieval accuracy can be improved. The proposed method outperforms the other previously developed methods by providing the classification accuracy of more than 70% for the various types of images taken from coral database. Hence, this paper concentrates on color and texture features for image retrieval in different directions. The proposed method significantly improves efficiency with less computational complexity.

Keywords: Color, Texture, Tamura, Threshold, Retrieval, Image Database, Mean, Standard deviation, Median features.

I. INTRODUCTION

Nowadays, with increased digital images available on the Internet, efficient indexing and searching becomes essential for large image archives. Experiments into the automatic retrieval of images from a large collection database are based on the colors, shapes and texture features. Content Based Image Retrieval (CBIR) has drawn substantial research attention in the last decade. CBIR usually indexes images by low-level visual features such as color, texture and shape. The visual features cannot completely characterize semantic content, but they are easier to integrate into mathematical formulations. Extraction of good visual features which compactly represent a query image is one of the important tasks in CBIR [3].

A. Image Indexing:

Image indexing and its application to content-based image retrieval is a challenging research topic in the field of image processing. In content based image retrieval system, the user sets the query image by choosing a similar image from the "Sample Query Set" and a query can also be set by identifying the desired object from the segmented query image. The image retrieval system evaluates the similarity of each image in its data store to the query image in terms of textural and color characteristics and returns the retrieved images with ranking within a desired range of similarity [9] [15].

Feature extraction is very crucial step in image retrieval system to describe the image with minimum number of descriptors. The basic visual features of images include color and texture. we have attempted to capture the color information by its color. The underlying technique of the image retrieval system is based on the adaptation of an entropy based approach to color and texture analysis. An optimal set of three features color, texture entropy and the dominant color of the image are extracted so as to render the feature vector for each image maximally informative, and yet to obtain a low vector dimensionality for efficiency in computation. Furthermore, we consider the entropy based texture entropies as constituting the feature vector for indexing the images. The indexing scheme is effectively used to generate an intermediate result set for any query image and this result set contract the search space by including only similar images [1] [12].

B. Image Retrieval:

Image Retrieval means of searching the original images related to the keywords. In order to retrieve the original image, they're in need of using the tool called Image Retrieval Engine. Color is one of the most widely used low-level visual features and is invariant to image size and orientation. As conventional color features used in CBIR, Texture is also an important visual feature that refers to innate surface properties of an object and their relationship to the surrounding environment. Many objects in an image can be distinguished solely by their textures without any other information. There is no universal definition of texture. Texture may consist of some basic primitives, and may also describe the structural arrangement of a region and the relationship of the surrounding regions. When we retrieve to identify the image, extract the corresponding features from a known image and then retrieve the image database to identify the images which are similar to it, also we can give some of the characteristics based on an image feature, then retrieve out the required images based on the given suitable values. In the whole retrieval process, feature extraction is essential; it is closely related to all aspects of the feature, such as color, texture and shape [4]. This kind of image retrieving process can be obtained in two phases

a. New image insertion with the features in database.

b. Searching new images in available database.

In order to make this process very efficient, a collection of ten thousand images is taken into the working directory of MATLAB. Using those MATLAB programming, all image features in R, G and B color projection values are extracted and stored in the database. It provides efficient retrieval services for various image databases [7].

C. Image Retrieval Types:

The process of retrieving and displaying relevant images is based on user's queries from the web or image database. Generally Image retrieval methods are classified into two types. They are

a. Text-Based Image Retrieval
b. Content-Based Image Retrieval

a. Text-Based Image Retrieval (TBIR):

The image is annotated by using text descriptions like Creator, File name, Place, Date, Time, objects. The image retrieval is done by using one or more text descriptors. The high-level retrieval involves retrieval of an image based on the name of objects, emotions and actions that can be associated with the image. Nowadays many se4arch engines retrieve the images using text based image retrieval methods [21].

b. Content-Based Image Retrieval (CBIR)

The TBIR does not depend on the actual object present in the image. So the process is further required to analyze the content of the image. CBIR uses low level features such as color, texture, shape and object location. CBIR is an extraction of images based on image content [6] [9] . The various levels of CBIR are

Level 1: Retrieval by primitive features such as color, texture, shape and spatial location.
Level 2: Retrieval of objects of given type. Example: find the picture of the flower.
Level 3: Retrieval of abstract attributes that involves high level reasoning g. Example: 'Find a picture of a baby smiling'.

D. Advantage of CBIR over TBIR:

Content-Based image retrieval is one of the efficient retrieving processes where the images are retrieved on the content of the image such as color, texture and shape. Whereas in the text based image retrieval process, images are retrieved on the keyword declaration. The key word declaration does not always represent context sensitive information of an image. So the search process leads to irrelevant garbage resulting. CBIR search for a query images without any information. It also finds an image with color and easily identifying duplicate image from the image pool [9].

II. IMAGE FEATURES

It mainly deals with the color and the texture features [11] . A color image is a combination of some basic colors. In MATLAB breaks each individual pixel of a color image (termed 'true color') down into Red, Green and Blue values. Image color values are stored in matrix form. Using Image matrix, all R, G, B components in the image are extracted. Instead of using single color projection we can select different combinations of color methods applying we can achieve good performance. The application can execute an image search based on a query image, The images are stored in a folder containing image files which are extracted by the comparison of features in real time [8] [13].

This kind of features can be obtained in several ways listed below

a. In calculating the R, G, B mean values (i.e.,) rm, gm, bm while preparing the database itself.
b. Using of RG and BG average values
c. By calculating the Blue's standard deviation (i.e.,)
d. The texture entropy value is used in retrieving the images.
e. Tamura0 and Tamura1 features can be used.

A. The Image Retrieval using Color Features :

Color is one of the most important features in CBIR. It is most widely used for both human perception and computer vision . As we know image is a collection of pixels i.e. row and column, Currently RGB i.e. Red, Green, Blue color model is used in digital image because it is more convenient for displaying images in CRT. In content based Image retrieval, color descriptor has been one of the first choices because if one chooses a proper representation, it can be partially reliable even in the presence of changes in lighting, view angle, and scale. In image retrieval, the color values are the most commonly used color features. It denotes the probability of the intensities of the three color channels. It deals with mainly colors in the digital images. These are stored in a feature vector [11]. Relatively the images will be retrieved to maintain the similarities. The following statistical measures are used in obtaining the features [6] [21].

a. Mean:

This feature is related to the brightness of an image. The image is a collection of row and column i.e. pixel values. Each level consists of some pixel values, and then finds out all the pixel values of all the levels and calculate average i.e.mean of all the pixel values that represent the brightness of the image. If the mean is higher, then the image is bright and if mean value is lower, then the image is dark [21]. Formulas for calculating mean is

$$\mu_j = \frac{1}{N}\sum_{i=1}^{N} x_{ji} \qquad (1)$$

b. Standard Deviation and Variance:

It represents the contrast of each level of image. It is calculated by using mean values of the image and pixel values of each level and variance is the square of standard deviation..

$$\sigma_j = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_{ji} - \mu_j)^2} \qquad (2)$$

If the value of SD is high, then the image will be in high contrast and if SD is low, then it will be low contrast.

B. Image Retrieval using Texture Features:

Texture refers to a structural collection of pixels of an image. Many features of image can be extracted by texture features. For texture feature extraction in 1973, Haraclick proposed Gray level co-occurrence matrix (GLCM). He proposed some parameters for texture extraction. Texture elements involve with Entropy, Energy, Inertia, and Homogeny values in calculating with all those fields can achieve the desired tasks [20].

a. Contrast (Moment of inertia): Image contrast can be a sharpness of image. Contrast is higher when image grooves are deep [3].
b. Energy: It can be measured by the gray distribution of the image. Coarness of image depends on high energy level .
c. Entropy: It is a quantity which is used to describe the amount of information of an image. If the entropy level is low, then image having more black area. An image that is perfect having entropy zero [4].
d. Correlation: It is used to calculate the degree of similarity of the elements of an image.

Like color, the texture is a powerful low-level feature for image search and retrieval applications. Much work has been done on texture analysis, classification and segmentation for the last four decades, So far, there is no unique definition of texture; however, an encapsulating scientific definition as given in can be stated as, "Texture is an attribute representing the spatial arrangement of the gray levels of the pixels in a region or image". The commonly known texture descriptors are Wavelet Transform, Gabor-filter , co-occurrence matrices and Tamura features [10],

The image segmentation is the process of dividing an image into multipart's. This is typically to identify objects or other relevant information in the digital images. There are different ways to perform image segmentations like thersholding methods, clustering methods, texture methods and Transform methods. It can also perform in MATLAB with an image processing tool which gives a segmentation algorithm, tools and comprehensive environment for data analysis, visualization and algorithm development. The techniques that are used to find the objects of interest are usually referred as segmentation techniques. The Goal of image segmentation is to change the representation of an image into something that is more meaningful and easier to analyze Image segmentation is typically used to locate the objects and boundaries in images. The segmentation algorithm is applied to the image and the maximum size of the segment is found.

III. EXPERIMENTAL SET UP

An RGB image, sometimes referred to as a true color image, is stored in MATLAB as m-by-n-by-3 data array that defines red, green, and blue color components for each individual pixel [11] .

A. Image Retrieval using R, G and B average value:

The color of each pixel is determined by the combination of the red, green, and blue intensities stored in each color plane at the pixel's location. Graphics file formats store RGB images as 24-bit images, where the red, green, and blue components are 8 bits each. This yields a potential of 16 million colors. The precision with which a real-life image can be replicated has led to the commonly used term true color image .Most color images are recorded in RGB space, which is perhaps the most well-known color space [2] [4]. So for a given query image the color spaces are shown in Fig. 1

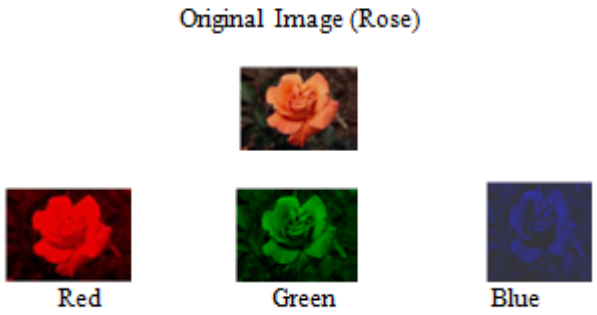


Figure 1. Color Image representation

The color is taken for representing the feature since the entropy of image patches is helpful for their separation from uniform, and smoothly varying object surfaces. The entropy of image patches is helpful for their separation from uniform, and smoothly varying object surfaces. The RGB values are given in Table 1. The image retrieval using the RGB average entropy value is shown in table 2. In this paper, as far as Precision and recall value is concerned, we have achieved a 80% of precision and recall value.

Table 1 – Red, Green and Blue average values

S.No	Image	Name	R-Mean	G-Mean	B-Mean	RGB Average
1		Drink/1.jpg	6.4432	6.5181	6.5572	6.5062
2		Car/5.jpg	6.9523	7.2181	6.6757	6.9487
3		Duck/0.jpg	7.3824	7.323	7.34	7.3485
4		Dino/0.jpg	4.5731	4.7665	5.5696	4.9697
5		Tiger/T1.jpg	127.47	115.23	115	119.21

Table 2 – The precision and Recall value for RGB average

Input Image	Number of Testing images	Relevant Images in the test set	Total Retrieved Documents	Retrieved Documents		Precision	Recall
				Relevant	Irrelevant		
	1000	10	9	3	6	0.33	0.3
	1000	10	13	4	9	0.3	0.39
	1000	10	18	5	13	0.27	0.5
	1000	10	32	7	25	0.22	0.7
	1000	10	44	8	36	0.18	0.8

B. Performance in RGB average

The image retrieval using the RGB_{Mean} value is shown in Table 2. In this work, as far as precision and recall value is concerned, it average recall value is 0.63 Which is greater than average precision value (0.36). It shows that this method tries to extract maximum relevant images.

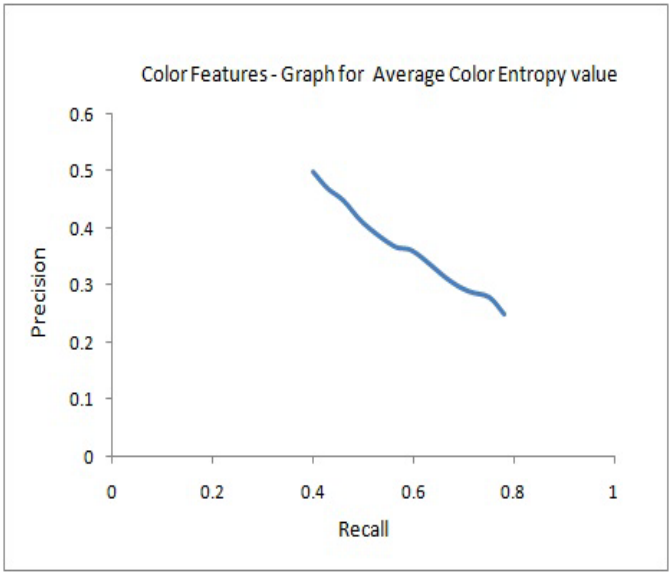


Figure: 2 Red, Green and Blue average values

C. Image Retrieval using Rg And By Average Value:

Assuming that the original data is in RGB color space, the red–green (rg) and blue–yellow (by) components [3] can be calculated as

$$rg = R - G, \quad (3)$$

$$by = ½ (R + G) – B \quad (4)$$

The above formula is based on the color theory concept. rg and by values are very sensitive to human stimulus. The Table 3 shows RedGreen and BlueYellow values for various images.The image retrieval is done in order to check the relevancy using rg and by values. The image retrieval using the rg and by average value is shown in Fig 3.

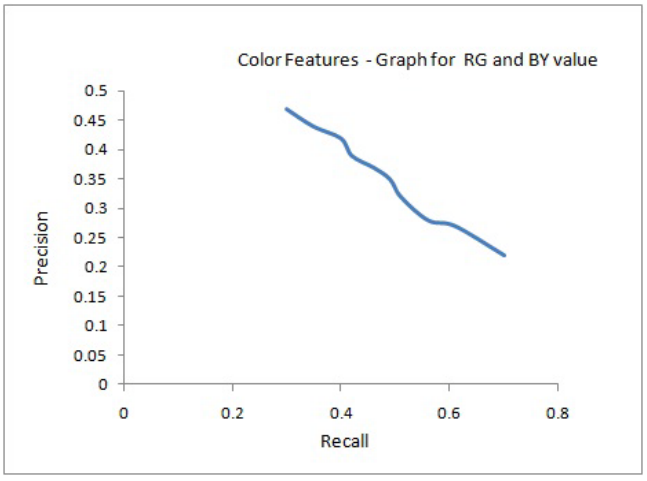


Figure: 3 RedGreen and BlueYellow values

The image retrieval using the RG and BY average value is shown in table 3. In this paper, as far as Precision and recall value is concerned, we have achieved a 56% of average precision value and 58 % as the average recall value.

Table 3 – RedGreen and BlueYellow values

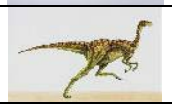
S.No	Image	Name	Rm	Gm	Bm	rg	by
1		Ship/4.jpg	196	186	184	10	7
2		Drink/0.jpg	186	174	178	12	2
3		Car/0.jpg	41	46	35	-5	8.5
4		Duck/6.jpg	165	164	175	1	-10.5
5		Dino/1.jpg	227	227	226	0	1

Table 4 – The precision and Recall value

Input Image	Testing images	Relevant Images	Total Retrieved Documents	Retrieved Documents		Precision	Recall
				Relevant	Irrelevant		
	1000	10	6	3	3	0.47	0.3
	1000	10	9	4	5	0.44	0.35
	1000	10	10	4	6	0.39	0.42
	1000	10	22	6	16	0.27	0.61
	1000	10	32	7	25	0.22	0.7

D. Image retrieval using Standard Deviation Blue:

The paper [10] shows that the standard deviation Blue gives the best accuracy for natural images. Table 5 gives the standard deviation Blue value for various images [4] [18].

Table 5 gives the SD_{Blue} value for various images. It is calculated for various images. The SD_{Blue} gives the verage precision value as 0.38 and the average recall value as 0.68. The precision and recall graph for SD_{Blue} is drawn as shown in the Figure 4. When compared to all the above color features, the SD_{Blue} gives the maximum relevancy with respect to color features.

E. *Performance in Standard Deviation Blue*

The image retrieval using the blue standard deviation value is shown in Table 6. In this paper, as far as Precision and recall value is concerned. It gives the maximum recall value.

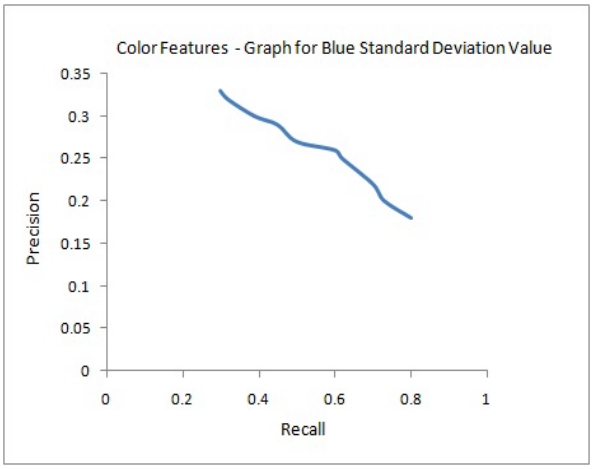


Figure: 4 Standard Deviation Blue values

Table 5 – Standard Deviation Blue values

S.No	Image	Name	Blue Standard Deviation
1		Light/3.jpg	67.163
2		Flag/1.jpg	56.252
3		Bus/2.jpg	70.717
4		Doll/3.jpg	40.704
5		Mask/2.jpg	65.524

Table 6 – The precision and Recall value

Input Image	Number of Testing images	Relevant Images in the test set	Total Retrieved Documents	Retrieved Documents		Precision	Recall
				Relevant	Irrelevant		
	1000	10	19	7	12	0.36	0.6
	1000	10	23	7	16	0.31	0.67
	1000	10	24	7	17	0.29	0.71
	1000	10	29	8	21	0.28	0.75
	1000	10	32	8	24	0.25	0.78

Table 7 – Texture Entropy values

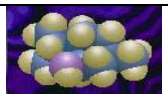
S.No	Image	Name	Entropy	S.No	Image	Name	Entropy
1		Drink/0.jpg	9.6237	6		Molecular/2.jpg	11.769
2		Duck/0.jpg	10.767	7		Mask/4.jpg	8.4071
3		Dog/1.jpg	12.235	8		Ship/0.jpg	12.002
4		Doll/3.jpg	7.3234	9		Plate/2.jpg	11.131
5		Dino/0.jpg	12.002	10		Bus/2.jpg	12.461

Table 8 – The precision and Recall value

Input Image	Testing images	Relevant Images	Total Retrieved Documents	Retrieved Documents		Precision	Recall
				Relevant	Irrelevant		
	1000	10	10	6	4	0.6	0.6
	1000	10	11	6	5	0.57	0.62
	1000	10	13	7	6	0.54	0.68
	1000	10	16	8	8	0.5	0.75
	1000	10	17	8	9	0.47	0.79
	1000	10	21	8	13	0.39	0.88

F. Image Retrieval using Texture Entropy Value:

For texturally uniform image, entropy is small. The entropy of a random variable indicates the degree of uncertainty associated with the variable. Higher the entropy more uncertain is the value of random variables.

The entropy is used for a set of numbers to calculate degree of randomness in the set and the entropy is calculated over a neighborhood matrix of all the pixels. The entropy of a set is calculated as:

$$ENT = -\sum\sum P(i,j)\log[P(i,j)] \quad (5)$$

The image retrieval using the Texture Entropy values is shown in table 8. In this paper, as far as Precision and recall value is concerned, we have achieved 52% as average precision value.

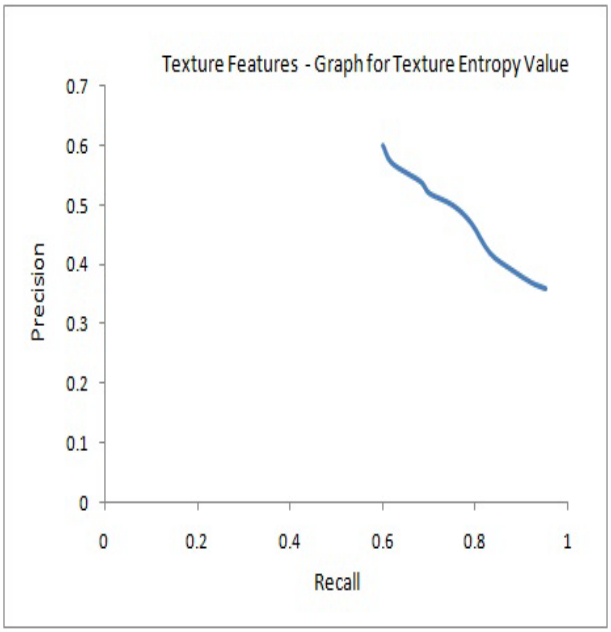


Figure: 5 Texture Entropy values

G. Image retrieval using Tamura0 And Tamura1:

Tamura feature is designed in accordance with psychological studies on the human perception of texture. The important Tamura features are Coarseness, Contrast and Directionality. The various Tamura Features are formulated and calculated from the input image. In the database creation phase, two Tamura features are extracted to describe each texture image in the database. The query can be expressed Tamura feature values. Tamura Image is a notion where we calculate a value for the three features at each pixel and treat these as a spatial joint coarseness-contrast-directionality (CND) distribution, in the same way as images can be viewed as spatial joint RGB distributions [14] [17]. Table 9 gives the Tamura0 and Tamura1 values. Fig 6 gives the precision and recall value when retrieving the images using tamura0 and tamura1 features.

Table 9 – Texture Tamura0 and Tamura1 values

S.No	Image	Name	Tamura0	Tamura1
1		Ship/5.jpg	5.7671	53.8863
2		Molecular/3.jpg	6.6478	65.8218
3		Car/0.jpg	5.4255	62.7219
4		Mask/6.jpg	6.8395	61.0807
5		Duck/3.jpg	7.1944	51.9997

Table 10 – The precision and Recall value

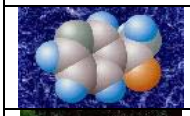
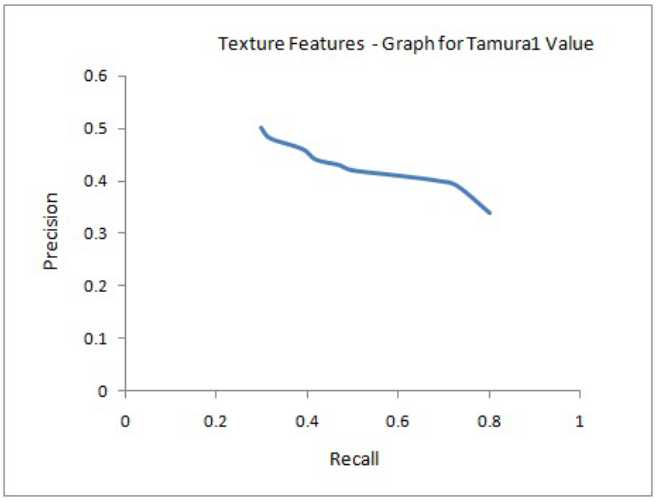
Input Image	Number of Testing images	Relevant Images in the test set	Total Retrieve Documents	Retrieved Documents		Precision	Recall
				Relevant	Irrelevant		
	1000	10	5	3	2	0.6	0.3
	1000	10	7	4	3	0.57	0.35
	1000	10	8	4	4	0.5	0.4
	1000	10	15	6	9	0.4	0.6
	1000	10	18	7	11	0.4	0.68

Table 11 – The precision and Recall value

Input Image	Number of Testing images	Relevant Images in the test set	Total Retrieved Documents	Retrieved Documents		Precision	Recall
				Relevant	Irrelevant		
	1000	10	6	3	3	0.5	0.3
	1000	10	9	4	5	0.46	0.39
	1000	10	12	5	7	0.43	0.47
	1000	10	15	6	9	0.41	0.6
	1000	10	18	7	11	0.4	0.69



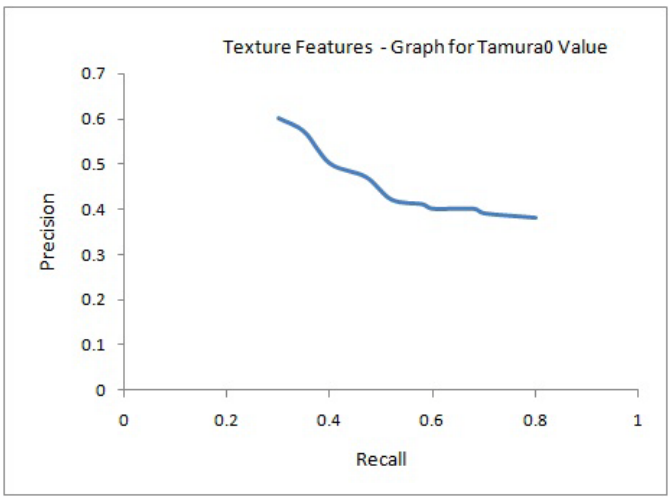


Figure: 6 Texture Precision Recall for Tamura0 and Tamura1 values

IV. CONCLUSION

The coral images are taken for image retrieval. The image retrieval is done by using color and texture features. When considering color feature as a primary component in the blue standard deviation gives better accuracy as compared with other features. When considering texture as a primary component the entropy value gives the best accuracy. Generally texture features give the more accuracy then color features. The overall performance can be improved by considering Blue standard deviation and Entropy value for natural images taken from the coral database. So the combined (Color - SD and Texture - Entropy) approach reaches the average precision value as 55% and recall value as 52%. The precision can be further improved by considering more images in the image database. Generally texture features give the more accuracy then color features.

V. FURTHER ENHANCEMENT

The approach can extend to include the additional color features like green standard deviation and red standard deviation. The other texture features like energy, moment of inertia, homogeneity. The segmentation algorithm can be used to include all the objects in a photograph. If the color normalization can be applied to various objects, some irrelevant objects can be omitted. If the shape features are added for recognizing the object, some general images like sun, ball, circle are recognized in an efficient manner. This can be extended to include indoor and outdoor images and Google earth map data. The approach can be extended to include human faces by considering the texture and shape features. The image mining can be done by using clustering algorithm and decision tree algorithm using the above features.

VI. REFERENCE

[1]. Alaa M Riad, Hamdy Elminir and Sameh Abd-Elghany," Web Image Retrieval Search Engine based on Semantically Shared Annotation", International Journal of Computer Science Issues, Vol. 9, Issue 2, No 3, 2012.

[2]. Allan Hanbury " A Survey Methods for Image Annotation", Science Direct, Pattern Recognition and Image Processing Group, Institute of Computer-Aided Automation, Favoritenstra , A-1040 Vienna, Austria , 2008.

[3]. Awang Iskandar D. N. F., Thom A and Tahaghoghi S. M. M, " Content-based Image Retrieval using Image Regions as Query Examples", Proceedings of the nineteenth conference on Australasian database, Vol. 75, Australia, 2008.

[4]. Babu Rao M, Prabhakara Rao B, Govardhan A, "Content Based Image Retrieval using Dominant Color, Texture and Shape", International Journal of Engineering Science and Technology, ISSN: 0975-5462 Vol. 3 No. 4, 2011.

[5]. Dengsheng Zhang, Monirul Islam Md, Guojun Lu and Jin Hou," Semantic Image Retrieval Using Region Based Inverted File", IEEE, Digital Image Computing, Techniques and Applications, 2009.

[6]. Eakins J and Graham M , "Content-based Image Retrieval", Technical Report, University of Northumbria at Newcastle, 1999.

[7]. Feng H, Shi R and Chua T. S, "A Bootstrapping Framework for Annotating and Retrieving WWW Images," Proceedings of ACM International Conference on Multimedia: p.960-967, 2004.

[8]. Gonzalez R. C and Woods R. E, "Digital Image Processing ", 2nd edition, Pearson Education, 2000.

[9]. Huan Wang, Liang-Tien, Chia · Song Liu, "Image retrieval ++ - web image retrieval with an enhanced multi-modality ontology ", Springer Science, Business Media, LLC, 2008.

[10]. Kavitha Ch., Prabhakara Rao B, Govardhan A, "Image Retrieval Based on Color and Texture Features of the Image Sub-blocks ", International Journal of Computer Applications (0975 – 8887), Vol 15– No.7, 2011.

[11]. Compact Composite Descriptors for Content Based Image Retrieval.Basics, Concepts, Tools.ISBN-10: 363937391X ..

[12]. Kristına Lidayova, "Semantic Categorization and Retrieval of Natural Scene Images ", Proceedings of CESCG 2012, the 16th Central European Seminar on Computer Graphics, 2012.

[13]. Madhu G, Govardhan A and Rajinikanth T.V, "Intelligent Semantic Web Search Engines: A Brief Survey", International journal of Web & Semantic Technology Vol.2, No.1, DOI: 10.5121/ijwest.2011.2103 34, 2011.

[14]. Minakshi Banerjee, Malay K. Kundu and Pradipta Majia, " Content-Based Image Retrieval Using Visually Significant Point Features", Journal on Fuzzy Sets and Systems, Vol. 160, pp. 3323–3341, 2009.

[15]. Pal S. K and Pal N. R, "A Review on Image Segmentation Techniques", Pattern Recognition, Vol. 26, No. 9, pp. 1277-1294, 1993.

[16]. Ramamurthy and Chandran K.R, "Content based Image Retrieval for Medical Images using Canny Edge Detection Algorithm," International Journal of Computer Applications, Vol. 17, No. 6, pp. 32-37, 2011.

[17]. Rui Y, Huang T. S and Chang S. F, " Image Retrieval: Current Techniques, Promising Directions, and Open Issues," Journal of Visual Communication and Image Representation, 10(4): p. 39-62. 1999.

[18]. Venkata Ramana Chary R, Rajya Lakshmi D and Sunitha K.V.N "Feature Extraction Methods for Color Image Similarity" Advanced Computing: An International Journal (ACIJ), Vol.3, No.2, 2012.

[19]. Xin-Jing Wang, Lei Zhang, Xirong Li, and Wei-ying Ma, " Annotating Images by Mining Image Search Results ", IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol 30, NO.11, 2008.

[20]. Yanai K. "Web Image Mining toward Generic Image Recognition ", Proceedings of International conference of World Wide Web, 2003.

[21]. Ying Liu, Dengsheng Zhang, Guojun Lu and Wei-Ying Ma," A Survey of Content-Based Image Retrieval with High-Level Semantics", Science Direct, Pattern Recognition - 40, 262– 282, 2007.