

October 2013

AN OPTIMIZED FEATURE EXTRACTION TECHNIQUE FOR CONTENT BASED IMAGE RETRIEVAL

SAVITHA SIVAN

Electronics and Communication Engineering, Karunya University, Tamilnadu, India,
savita.sivan@gmail.com

THUSNAVIS BELLA MARY.I

Electronics and Communication Engineering, Karunya University, Tamilnadu, India, t.v.mary@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijipvs>



Part of the [Robotics Commons](#), [Signal Processing Commons](#), and the [Systems and Communications Commons](#)

Recommended Citation

SIVAN, SAVITHA and MARY.I, THUSNAVIS BELLA (2013) "AN OPTIMIZED FEATURE EXTRACTION TECHNIQUE FOR CONTENT BASED IMAGE RETRIEVAL," *International Journal of Image Processing and Vision Science*: Vol. 2 : Iss. 1 , Article 3.

DOI: 10.47893/IJIPVS.2013.1061

Available at: <https://www.interscience.in/ijipvs/vol2/iss1/3>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Image Processing and Vision Science by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

AN OPTIMIZED FEATURE EXTRACTION TECHNIQUE FOR CONTENT BASED IMAGE RETRIEVAL

SAVITHA SIVAN¹, THUSNAVIS BELLA MARY.²

¹M.Tech Student, ²Assistant Professor, Electronics and Communication Engineering, Karunya University, Tamilnadu, India

Abstract- Content-based image retrieval (CBIR) is an active research area with the development of multimedia technologies and has become a source of exact and fast retrieval. The aim of CBIR is to search and retrieve images from a large database and find out the best match for the given query. Accuracy and efficiency for high dimensional datasets with enormous number of samples is a challenging arena. In this paper, Content Based Image Retrieval using various features such as color, shape, texture is made and a comparison is made among them. The performance of the retrieval system is evaluated depending upon the features extracted from an image. The performance was evaluated using precision and recall rates. Haralick texture features were analyzed at 0°, 45°, 90°, 180° using gray level co-occurrence matrix. Color feature extraction was done using color moments. Structured features and multiple feature fusion are two main technologies to ensure the retrieval accuracy in the system. GIST is considered as one of the main structured features. It was experimentally observed that combination of these techniques yielded superior performance than individual features. The results for the most efficient combination of techniques have also been presented and optimized for each class of query.

Keywords-CBIR, Feature Extraction, GIST Descriptor, Co-occurrence, Color Moments

I. INTRODUCTION

Content-based image retrieval (CBIR) is defined as a task of searching similar images for query example from large-scale image databases through analyzing the actual contents of images rather than the affiliated metadata [11,12, 13]. Content based image retrieval (CBIR) has become an active research area which has many applications in crime prevention, medicine, image processing and pattern recognition. Since the companies are maintaining large image databases, the requirement there is to adopt a technique that can search and retrieve images which will be both time efficient and accurate. In order to meet these requirements, in general, perform the retrieval process in two steps. Feature extraction is the first step in which unique signatures, termed as feature vector, is identified for every image based on its pixel values. Visual features such as color, texture and shape are commonly used in this step. In the classification step, matching between the features extracted from the query image with the features of the database images are done and the images are grouped according to their similarity. The feature extraction is considered as a critical factor because the particular features made available for discrimination directly influence the efficacy of the classification task [7].

Among the low level features, the least specific feature is the color and shape is the most specific. Selection of various features for content based image retrieval is an important task because only the combination of some specific features would give the accurate results. Depending on the application, colors can be represented for using different color spaces such as Red-Green-Blue (RGB) and Hue-Saturation-

Intensity (HSI). HSI color space is easy to understand and calculate than RGB. Most of the work related to color feature extraction is based on color histogram, but it has various drawbacks such as CH is very sensitive to noise, its not possible for high dimensional data because of increased complexity. So color moments can be used for color extraction. Image contents are mostly concentrated on the low-order moments. They are first moment (mean), second moment (variance) and third moment (skewness) and hence they are only considered for color feature extraction. Texture is another important visual feature which is an innate property of virtually all surfaces. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment [2]. Features derived from GLCMs are useful in medical image retrieval. Gray level co-occurrence method use gray-level co-occurrence matrix to sample statistically the gray-levels of an image. This matrix can be created from the information about each pixel and its neighbour pixels. Visual features are automatically extracted from images which save a lot of human effort [5]. Haralick texture features are then extracted from this matrix which is having a reduced computational complexity and biggest discriminatory power [3]. Shape plays an important role in human recognition and perception. An important ability for an agent (both human and robot) is to build an instantaneous concept of the surrounding environment. Such holistic concepts activate top-down knowledge that guides the visual analysis in further scene interpretation tasks. For this purpose, Gist features were introduced [6] that represent the spatial envelope of a scene given as a 2D image. A variety of scene recognition approaches have been proposed that work well for outdoor scenes

(buildings, street, mountains etc) but frequently break down for indoor scenes(kitchen, living room etc). Latest advances in the field of computer vision have spanned the approach to understanding the semantics of natural scene images into two directions: global representation and an order less bag of features (BOF) model [11]. The former, proposed in [11], attempts to capture the “gist” of a scene without object segmentation and recognition. The Gist descriptor is a low-dimensional representation of the attributes of a scene, namely naturalness, openness, roughness, expansion and ruggedness. The result of the scene classification paradigm based on this holistic perspective on natural scene images compared to human performance in a rapid scene classification experiment [9], has provided evidence that the concept of representing the global attributes of a scene is in parallel to the human visual and cognitive system. The findings from further experiments [6] in psychology and cognitive science explain why the Gist descriptor performs remarkably well on scene recognition tasks, (especially on outdoor categories,) with applications extending to place recognition [8]. The Gist descriptor imposes strong constraints on spatial layout by dividing the image into an N-by-N grid, but fails to delineate the spatial structure in each block. Consequently, mismatch occurs due to averaging. It has been widely used to represent the image for classification, copy detection, and object recognition operation in individual blocks. On the other end of the spectrum is the BOF model. Inspired by the bag of words model in text categorization, this paradigm represents each image as an occurrence histogram of visual words that are local descriptors of regions or patches in the image.

II. FEATURE EXTRACTION

Feature extraction is principal components of all content-based image retrieval (CBIR) systems. The input data is transformed into a set of features. This process is called feature extraction. In order to effectively retrieve an image, appropriate features should be selected.

A. Texture Feature extraction

Gray level co-occurrence matrix (GLCM) is a statistical tool to extract the texture of an image. The number of texture features available from GLCM is large and some of the features are strongly correlated with each other [3], which show undesirable computational complexity. A feature selection procedure may be applied in order to select the optimal features to make the retrieval efficient [4]. The calculation of texture features involves two steps, the co-occurrence matrices (co-matrices) and the Haralick Texture Features. The co-matrix is computed from an image and the features are computed based on the co-matrices.

The generation of the co-occurrence matrices is based on second order statistics as described in [14] and [22]. This approach computes histogram matrices for different pixel pair orientations. Together with the pixel pairs along a specific angle (horizontal, diagonal, vertical, co-diagonal) and the distance (one to five pixels), a two-dimensional symmetric histogram of the gray levels is generated. The co-matrix indexes are represented by the gray levels of the pixel pair and it is incremented by one, an example can be found in [4]. For each specific angle/distance combination a separate matrix must be generated. That means one side of the square co-matrix is as long as the gray level range in the image. In this texture feature extraction is done at different pixel orientation such as 0° , 45° , 90° , 180° . For each direction, Haralick texture features are estimated from GLCMs [4]. Therefore, these features values correspond to texture feature vector. Haralick features refer to the fourteen textural features extracted from GLCMs. The textural features are Angular Second Moment (ASM), Contrast, Correlation, Variance (VAR), Inverse Difference Moment (IDM), Sum Average (SA), Sum Variance (SV), Sum Entropy (SE), Entropy (ENT), Difference Variance (DV), Difference Entropy (DE), Information Measure of Correlation I (IMC1), Information Measure of Correlation II (IMC2) and Maximal Correlation Coefficient (MCC). The mathematical expressions to estimate the Haralick features are given in [4].

B. GIST Feature Extraction

The dominant spatial structure of a scene can be represented by a set of perceptual dimensions like naturalness, openness, roughness, expansion and ruggedness and it is known as GIST feature [6]. GIST are efficient to compute and quite compact to store. It describes the spatial structure of the image. In order to preserve dominant structural details, firstly the gray scale image is pre-processed by a whitening filter to and then normalized with respect to local contrast. The pre-processed image is then passed through a cascade of Gabor filters with fixed parameters.

The parameters for image pre-processing (image whitening and local contrast normalization) are kept the same as that of the GIST, and so are the parameters for Gabor filter. Then the image is filtered and segmented into grid cells where orientation histograms are extracted. There are S scales with O orientations at each scale. Each of these $S \times O$ images (orientation maps), representing the original image at one orientation in each scale, is then divided into an N-by-N grid. Within each block on the grid, the average intensity is calculated to represent the feature in that block. The images are filtered by Gabor filter at 4 scales, with 8 orientation channels at each scale. The response of Gabor filters on each grid cells are concatenated into GIST feature vector of $S \times O \times N \times N$ dimensions [11].

C. Color Feature Extraction

Color is the most extensively used visual content for image retrieval. Its three dimensional values make its discrimination potentiality superior to the single dimensional gray values of images. Before selecting an appropriate color description, color space must be determined first. RGB color system is commonly used color space and is ideally suited for hardware implementation. Unfortunately, the RGB is not well suited for describing color in terms that are practical for human interpretation. Contrary, the HSI (hue, saturation, intensity) model is an ideal tool for developing image processing algorithms based on color descriptions that are natural and intuitive to humans. Color moment is a compact representation of the color feature to characterize a color image. Most of the color distribution information is captured by the three low-order moments. The first order (mean), the second order (variance) and the third order (skewness) color moments have been proved to be efficient and effective in representing color distributions of images. In this paper, RGB image is converted to HSI color space. Then the three color moments are calculated for each hue, saturation and intensity which constitutes the feature values for color feature extraction. The mathematical expressions to estimate these color moments are given in [13].

III. DATABASE USED

A standard database is used for testing, the Corel database [7].The dataset used in the performance evaluation consists of 1000 Corel images .The image set comprises of 100 images in each of 10 categories. Here, five classes were taken and each class consists of 10 images. The 5 classes are seashore, buildings, roses and food. The images are of size 256x384 or 384x256. The images considered here are of size 256x384. These classes were used for relevance estimation.

IV. METHODOLOGY

The image database contains the raw images. The database images are the reference images which are used for similarity comparison. From the image database, three descriptors are retrieved by the feature extraction block. The three selected descriptors are Color moments, Haralick texture descriptors and GIST descriptor. The descriptors are fused to form the hybrid features and are termed as the feature vector which forms the feature database. When a user presents a query, the features of the query image is extracted and the search and retrieval block searches the feature database to find a group that best matches the query feature vector. The most common metric used to measure the similarity between two images in CBIR is the Euclidean distance defined as [6]. The user interface components display these images as query result to the user. In order to find the optimum

combination of techniques to be used for each class of query, five parameters [18] was evaluated and a comparative analysis was done as discussed in Section VI. This enables the retrieval system to adapt itself according to query image given by the user and use the relevant techniques for the image retrieval process to produce the best results [7].

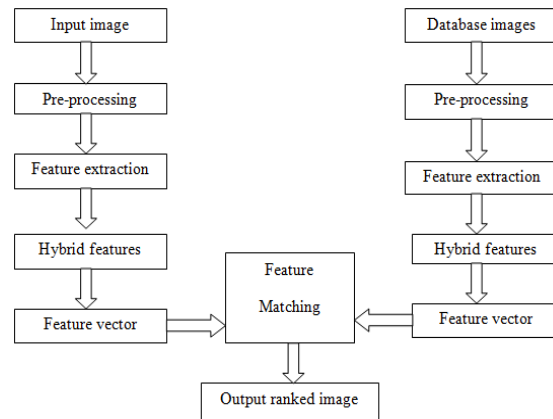


Figure1. Framework of hybrid image retrieval

V. PARAMETERS

A. Time

It is the time taken in seconds for the retrieval task to complete, at the end of which the system returns the images which are matched with the features of the query images, according to the technique used.

B. Error rate

It is a common measure for object retrieval. It is the ratio of number of non relevant images retrieved to the total number of images retrieved. The error rate should be low for good performance.

$$ER = \frac{\text{Number of non relevant images retrieved}}{\text{Total number of images retrieved}} \quad (1)$$

C. Normalised Average Rank

It is the measure of normalized average rank at which *i*th relevant image is retrieved. This measure is 0 for perfect performance and approaches 1 as performance worsens. For random retrieval, result would be 0.5. It is expressed as:

$$R = \frac{1}{(N \times Nr)} \left[\sum Ri - Nr(Nr-1)/2 \right] \quad (2)$$

R_i is the rank at which *i*th relevant image is retrieved and varies from 1 to *Nr*. *N* is the collection size. *Nr* is the number of relevant images for a given query.

D. Retrieval Efficiency

It is defined as the ratio of number of relevant images retrieved to the total number of images retrieved, if the no. of retrieved is greater than no. of relevant, else it is defined as the ratio of number of

relevant images retrieved to the total number of relevant images retrieved.

E. F-measure

It is the harmonic mean that combines the precision and recall .It is the measure of test’s accuracy. It is expressed as:

$$F = \frac{2 \times P \times R}{(P + R)} \tag{3}$$

where P is the precision and R is the recall.

VI. RESULT ANALYSIS

The goal here is to find the optimum combination of techniques for each class of query. Different applications require different needs. Apart from image retrieval, the retrieval system performance depends upon certain other factors such as time, efficiency, accuracy, error rate etc. So depending upon the user satisfaction, certain combination of features for certain classes yields better results.

A. Co-occurrence

TABLE I.RESULTS FOR CO-OCCURRENCE TECHNIQUE

Image class	Retrieval efficiency (%)	ER (%)	F-measure	R (%)	Time (sec)
Class 1	50	61	29.17	0.31	4.48
Class 2	61.67	47	42.5	0.098	5.12
Class 3	76.67	30	57.83	0.02	7
Class 4	48.33	50	40	0.084	8
Class 5	39	59	29	0.238	8

It was observed from the evaluation that better results were obtained for the combination of four Haralick texture features such as energy, entropy, contrast and homogeneity.

The parameters vary for each class. It is assumed that F-measure above 50% or more is termed as ‘good’ performance and below as ‘bad’ performance. From the table it was observed that for certain classes, F-measure is low, which affects the system accuracy.

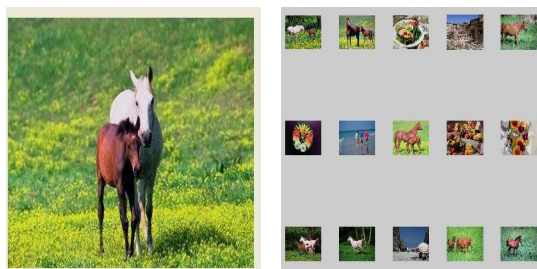


Figure 2.Illustration of query and retrieval using Co-occurrence method

B. Color Moments

TABLE II.RESULTS FOR COLOR MOMENTS

Image class	Retrieval efficiency (%)	ER (%)	F-measure	R (%)	Time (sec)
Class 1	63.33	59	43.17	0.08	3.41
Class 2	70	36	51.83	0.084	4.03
Class 3	49	41	44.33	0.05	6
Class 4	37.83	61	31.83	0.196	7.14
Class 5	30.5	48	39.5	0.122	7

From the table it was observed that even though the retrieval efficiency is good, for most of the classes, F-measure is low, which affects the system accuracy. For example , Class 1 has high retrieval efficiency but the error rate is very high which makes the system to underperform.



Figure 3.Illustration of query and retrieval using Color Moments

C. GIST Descriptor

TABLE III.RESULTS FOR GIST DESCRIPTOR

Image class	Retrieval efficiency (%)	ER (%)	F-measure	R (%)	Time (sec)
Class 1	42.83	56	34.33	0.22	2.83
Class 2	56.17	48	41	0.134	2.52
Class 3	76.67	30	57.83	0.02	3.15
Class 4	70	36	51.83	0.068	2.27
Class 5	44.5	54	36	0.182	3.30

From the table it was observed that the time taken for Gist Descriptor is less compared to other features. Normalized average rank is low for every classes which indicates that the relevant images are retrieved in a faster rate.

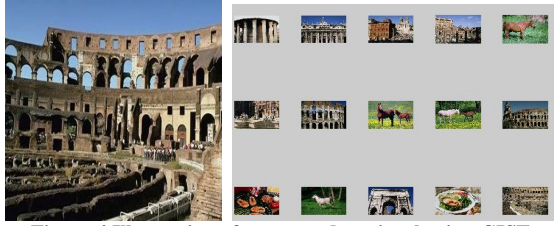


Figure 4. Illustration of query and retrieval using GIST Descriptor

D. Hybrid Technique

TABLE IV. RESULTS FOR HYBRID TECHNIQUE

Image class	Retrieval efficiency (%)	ER (%)	F-measure	R (%)	Time (sec)
Class 1	56.67	47	44.17	0.092	21.08
Class 2	83.33	32	61.67	0.024	22.77
Class 3	76.67	30	57.83	0.02	25.15
Class 4	61.67	47	42.33	0.068	28.27
Class 5	64.53	44	41	0.12	28.3

It is observed that when all the three features such as Color moments, Co-occurrence and Gist descriptor are combined, the retrieval efficiency of the system increases. But the main problem is the time taken will be more in order to meet the user satisfaction .So ,we go for optimization technique.



Figure 5. Illustration of query and retrieval using Hybrid technique

E. Optimization Technique

It has been observed that the combined approach gives excellent accuracy but long retrieval times. Therefore, optimization of the image retrieval process is necessary which can be attained by selecting the most efficient combinations to give the best possible retrieval performance.

TABLE V. RESULTS OBTAINED FOR OPTIMIZATION TECHNIQUE

CLASS 1	
Parameters	Features
Retrieval Efficiency (%)	Color moments, Co-occurrence
Error Rate (%)	Hybrid
F-measure	Color moments, Gistdescriptor

Normalized Average Rank	Color moments
Time(sec)	Color moments, Co-occurrence
CLASS 2	
Parameters	Features
Retrieval Efficiency (%)	Hybrid
Error Rate (%)	Color moments, Co-occurrence
F-measure	Gist descriptor
Normalized Average Rank	Color moments
Time(sec)	Color moments, Gistdescriptor
CLASS 3	
Parameters	Features
Retrieval Efficiency (%)	Co-occurrence
Error Rate (%)	Co-occurrence, Gistdescriptor
F-measure	Co-occurrence, Gistdescriptor
Normalized Average Rank	Co-occurrence, Gistdescriptor
Time(sec)	Color moments, Co-occurrence
CLASS 4	
Parameters	Features
Retrieval Efficiency (%)	Hybrid
Error Rate (%)	Gist descriptor
F-measure	Co-occurrence, Gistdescriptor
Normalized Average Rank	Color moments, Gistdescriptor
Time(sec)	Color moments, Co-occurrence
CLASS 5	
Parameters	Features
Retrieval Efficiency (%)	Hybrid
Error Rate (%)	Color moments, Gistdescriptor
F-measure	Hybrid
Normalized Average Rank	Color moments, Gistdescriptor
Time(sec)	Color moments, Co-occurrence

As discussed in the table, depending upon the user needs, the parameters can be combined and system performance can be improved. This enables a detailed performance analysis of the system.

VII. CONCLUSION

In this work, retrieval system evaluation is done based on different features for coral database. A

comparative analysis of the various features extraction techniques and their drawbacks when used individually were shown successfully using two performance measures like precision and recall. The performance of the retrieval system was observed to be better when the image retrieval was done based on hybrid features. An effective feature selection was done by optimizing the techniques for each class of images, resulting in an 'adaptive' retrieval system which results in a balanced performance in terms of image retrieval time, accuracy and redundancy factor. Such an image retrieval system can effectively recognize the class of the image query given by the user and can produce the best results according to it. In future work, other feature-based search can be performed for other database. The retrieval performance can further be improved by incorporating keyword annotations along with image retrieval so that an optimized set of results can be presented to the user.

REFERENCES

- [1] R. Datta, D. Joshi, J. Li, and J. Z. WANG, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, 40(2): 1-60, 2008.
- [2] Y. Rui, T. S. Huang, and S.-F. Chang, "Imageretrieval: Current techniques, promising directions, and open issues," *Journal of Visual Communication and Image Representation*, 10(1): 39-62, 1999.
- [3] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349-1380, 2000
- [4] M. Robert Haralick, K. Shanmugam, Itshak Dinstein, 1973. "Textural Features For Image Classification," *IEEE Transactions on Systems, Man And Cybernetics*, 6, pp. 610-621.
- [5] J. Li, J.Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 9, 1075-1088, 2003.
- [6] Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. *IJCV* 42 (2001) 145-175.
- [7] Aman Chadha, Sushmit Mallik , Ravdeep Johar, "Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval", *International Journal of Computer Applications* (0975 – 8887) Volume 52– No.20, August 2012
- [8] M. R. Greene and A. Oliva, "Recognition of Natural Scenes from Global Properties: Seeing the Forest without Representing the Trees," *Cognitive Psychology*, 58(2):137-179, 2009.
- [9] M. R. Greene and A. Oliva, "The Briefest of Glances: The Time Course of Natural Scene Understanding," *Psychological Science*, 20:464–472, 2009.
- [10] C. Siagian and L. Itti, "Rapid Biologically-inspired Scene Classification Using Features Shared with Visual Attention," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):300–312, 2007.
- [11] Robust Scene Classification By Gist With Angular Radial Partitioning, Wei Liu, Serkan Kiranyaz and Moncef Gabbouj, *Proceedings of the 5th International Symposium on Communications, Control and Signal Processing, ISCCSP 2012, Rome, Italy, 2-4 May 2012*
- [12]. Z.Y.He,X.G.You,Y.Yuan, "Texture image retrieval based on non-tensor product wavelet filter banks", *Signal Processing*,Vol.89,pp.1501-1510,2009.
- [13] T.V. Saikrishna, Dr.A.Yesubabu, Dr.A.Anandarao, T.Sudha Rani, "A Novel Image Retrieval Method Using Segmentation and Color Moments", *Advanced Computing: An International Journal (ACIJ)*, Vol.3, No.1, 2012.
- [14] J. Sivic and A. Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos," *International Conference on Computer Vision*, 2:1470-1477, 2003.
- [15] Monika Jain,Dr.S.K.Singh, "A Survey On:Content Bsed Image Rterieval Systems Using Clustering Techniques For Large Datasets,International Journal of Managing Information Technology (IJMIT)Vol.3,No.4,November 2011.
- [16] Wey-Shiuan Hwang , Weng, J.J. ; Ming Fang ; Jianzhong Qian , "A fast image retrieval algorithm with automatically extracted discriminant features", *Content-Based Access of Image and Video Libraries*, 1999, pp. 8-12.
- [17] David.M.W.Powers, "Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation, *Journal of Machine Learning Technologies*,ISSN: 2229-3981 & ISSN: 2229-399X, Volume 2, Issue 1, 2011, pp.37-63.
- [18] Henning Muller, Wolfgang, David McG. Squire, Stéphane Marchand-Maillet, Thierry Pun, "Performance evaluation in content-based image retrieval: overview and proposals",*Pattern Recognition Letters*,Volume 22, Issue 5, April 2001, pp. 593–601.
- [19] Muller, S. , Rigoll, G., " Improved stochastic modeling of shapes for content-based image retrieval", *Content-Based Access of Image and Video Libraries*, 1999.
- [20] Sagar Soman, Mitali Ghorpade, Vrushali Sonone,Satish Chavan, "Contant Based Image Retrieval Using Advanced Color and Texture Features,International Journal of Computer Applications,2011.

