Image Retrieval using Steerable Pyramid

Swapna Borde
Vidyavardhini's College of Engineering and Technology, Vasai (W),

Dr. Udhav Bhosle
Rajiv Gandhi Institute of Technology.
Mumbai,

ABSTRACT
The Content Based Image Retrieval System (CBIR) is a system, which retrieves the similar images from an image collection based on visual features such as color, texture and shape. It is an active and emerging research field in computer vision. In this paper, content based image retrieval (CBIR) is done using the image feature set extracted from Steerable Pyramid applied on the image at two levels (Level-1 and Level-2) of decomposition. The performance is evaluated using standard bench marks such as Precision and Recall. Our experiments are conducted on a database of 445 images with five different classes and successful matching results are obtained by using Steerable Pyramid Level-2.

Index

Content Based Image Retrieval (CBIR), Steerable Pyramid (SP).

1. INTRODUCTION
Content Based Image Retrieval (CBIR) is an important research area for manipulating large multimedia databases and digital libraries. It retrieves the similar images from an image collection according to the similarity between features extracted from the query image and each image in the database based on color, texture and shape. The features are automatically extracted from the images themselves. It is an alternative to the text based image retrieval systems. A human describes the images according to the image content, the caption, or the background information. However, the representation of an image with text requires significant effort and can be expensive, tedious, time consuming, subjective, incomplete, and inconsistent. To overcome the limitations of the text-based approach, Content-Based Image Retrieval (CBIR) techniques are used. High retrieval efficiency and less computational complexity are the desired characteristics of CBIR system. CBIR finds applications in advertising, medicine, crime detection, entertainment and digital libraries [6].

A key function in the CBIR system is feature extraction. A feature is a characteristic that can capture a certain visual property of an image either globally for the whole image, or locally for objects or regions. Different search engines use different features to retrieve images. One of the approaches to texture feature extraction is the filter bank approach that decomposes a texture image into subbands using a linear transform or filter bank.

Several previous works extract texture features based on wavelet packet signatures and wavelet frames. Although these methods allow for a multiresolution decomposition, they are limited in directional selectivity and not able to capture directional information.

The 2D Gabor transform, is a common tool for multiscale and multi-directional decomposition. Despite the high performance compared to the separable transforms, it produces over complete representations for images. Many other multiresolution and multi-directional image representation methods like the octave-band DFB, multiscale DFB and complex wavelets also have been used to form a feature vector. The focus of this paper is to extract the texture information for retrieval of images by using Steerable Pyramid. The Steerable pyramid provides a multi-scale, multi-directional decomposition of an image.

We consider a simple architecture of a typical Content Based Image Retrieval (CBIR) (Figure 1), where there are two major tasks. The first one is feature extraction (FE), where a set of features, called image signatures, is generated to accurately represent the content of each image in the database. A signature is much smaller in size than the original image, typically of the order of hundreds of elements (rather than millions). The second task is similarity measurement (SM), where a distance between the query image and each image in the database using their signatures is computed so that the top "closest" images can be retrieved [1], [2].

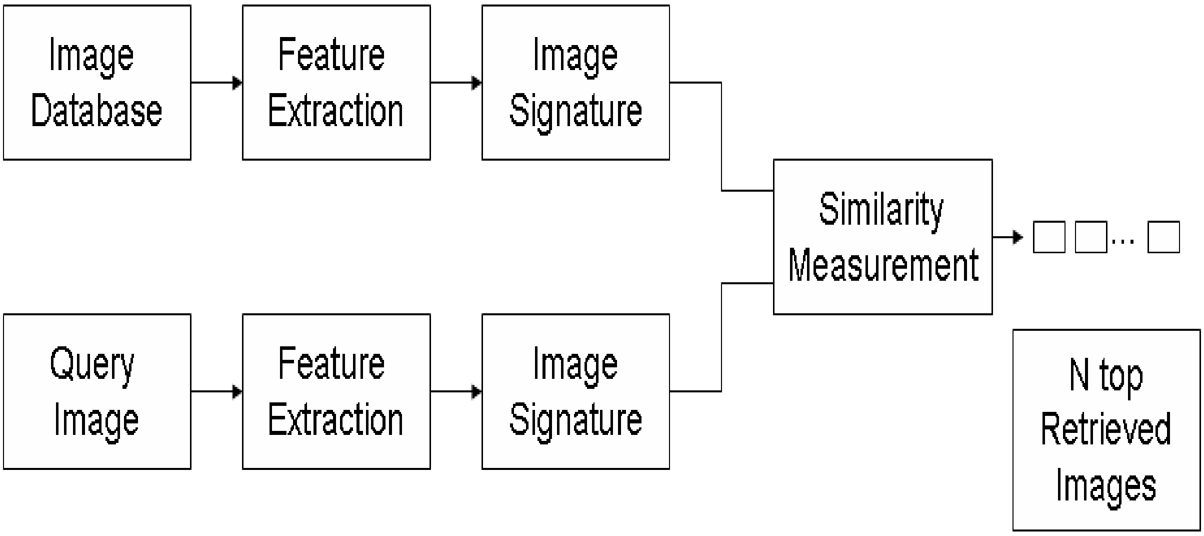

Figure 1: Content Based Image Retrieval Architecture

2. STEERABLE PYRAMID

The Steerable Pyramid generates a multi-scale, multi-directional representation of the image [3], [5]. The image is decomposed into low-pass sub-band and high-pass sub-band and decomposition is iterated in the low-pass sub-band. The Steerable Pyramid decomposition is similar to the two-dimensional discrete wavelet transform but with directional subbands.

Steps to extract features from a color image by using Steerable Pyramid are as follows:

1. This method first resizes the image to 128 ×128. Then divide the image into R, G, and B components.
2. Apply the low pass filter and high pass filter on each component (R, G, and B).
3. The low frequency content is the most important part of the signal. It is what gives the signal its identity. High frequency component contains less information as compared to low frequency component. Hence, the output of the low pass filter received from the first stage can be down sampled by a factor of 2. The process of filtering and down sampling can be repeated to get a multi-level decomposition.
4. Then directional subbands are obtained from the output of low pass filter of each stage. Figure 2 shows two stage (Level-2) Steerable Pyramid Transform.
5. Compute the features such as mean and standard deviation of directional subbands of query image as well as images in the database.
6. Then we use the Normalized Euclidean Distance [4] to compute the similarity measure between query image and images from the database.

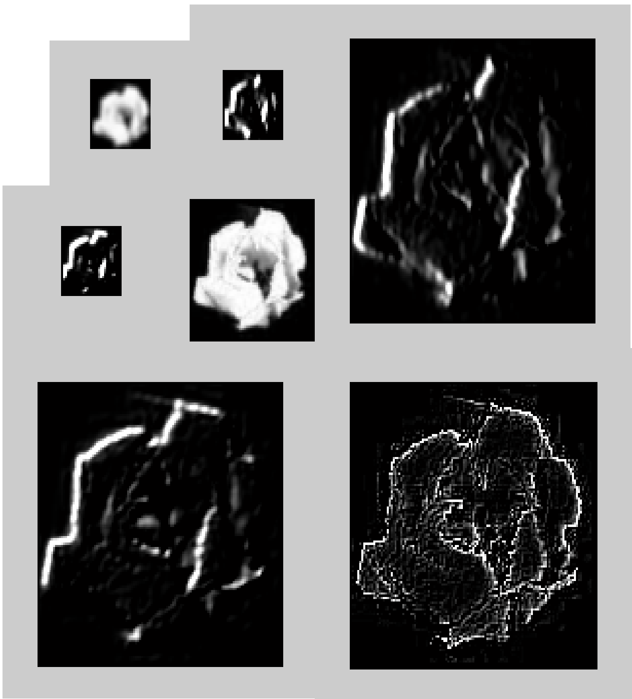

Figure 2: Two-stage steerable pyramid transform

In Figure 2, the output of the (single) highpass filter fills the LR quadrant. The two directional sub bands of the first stage appear in the UR and LL quadrants, while the lowpass output falls at the LR corner of the UL quadrant. The three output images of the second stage follow the same format in the UL quadrant

3. EXPERIMENTAL RESULTS

For evaluating the performance of the algorithms, we used an experimental database which consists of 445 images of size 128x128x3 with five different classes. The Classes and distribution of the images is shown in Table 1. Figure 3 shows sample views for the each of the classes in the database. To test the proposed system, one query image is selected from each category of images, so in all 5 queries are fired on the database.

Class	A	B	C	D	E
No.of Images	132	80	109	70	54

Table 1: Image database: Class-wise Distribution

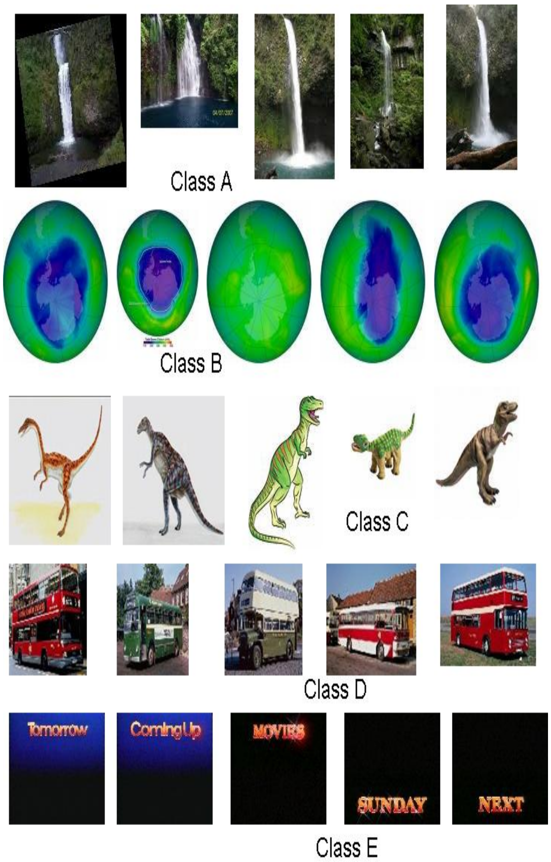

Figure 3: Sample images in the database (Five Classes A, B C, D, and E)

Features are extracted by applying Steerable Pyramid on query image as well as images in the database. To compare the similarity of two images, we use Normalized Euclidean Distance. The query image was selected from the database and it would be the first image in the result list. Other images in the result list were retrieved based on the similarity to the query images. The performance is evaluated using standard bench marks such as Precision, Recall. Precision is defined as the ratio of number of relevant images retrieved to total number of retrieved images. And Recall is defined as the ratio of number of relevant images retrieved to total number of relevant images in the database. To find the computational efficiency we measure retrieval time. The retrieval time primarily depends on the size of the image database, the software and the hardware versions that system runs on, size of the feature vector and the similarity measure used.

Figure 4 is the query image. Figure 5 (a) gives the results of first 30 retrieved images for a query image (shown in Figure 4) for the steerable pyramid Level-1. Figure 5 (b) gives the results of first 30 retrieved images for a query image (shown in Figure 4) for the Steerable pyramid Level-2.

Figure 4: Query Image (Class 'D')

Fig. 5 (a) First thirty retrieved images using Steerable Pyramid Level-1 (Total relevant Images retrieved:---26)

Fig. 5 (b) First thirty retrieved images using Steerable Pyramid Level-2
(Total relevant Images retrieved:---28)

(All result images are sequentially ordered from left to right with respect to their similarity with query image. Here total numbers of relevant retrieved images are higher in Steerable Pyramid-Level-2 based image retrieval technique as compared to Steerable Pyramid Level-1.)

The following graphs (Figure 6 to Figure 10) shows precision and recall plotted against number of retrieved images. The graphs are plotted by randomly selecting a one query image from each Class, so in all 5 queries are fired on the database. The precision and recall of all queries are obtained for the feature sets of both the levels (Level-1 & Level-2) of Steerable pyramid.

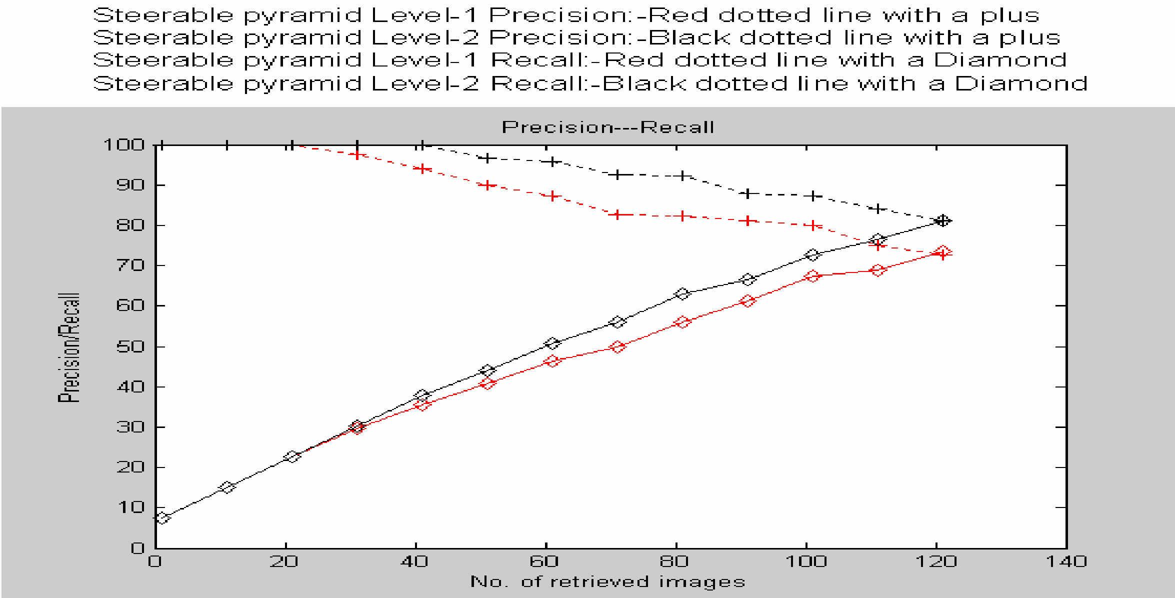

Figure 6: Precision /Recall for Query Image in Class 'A'

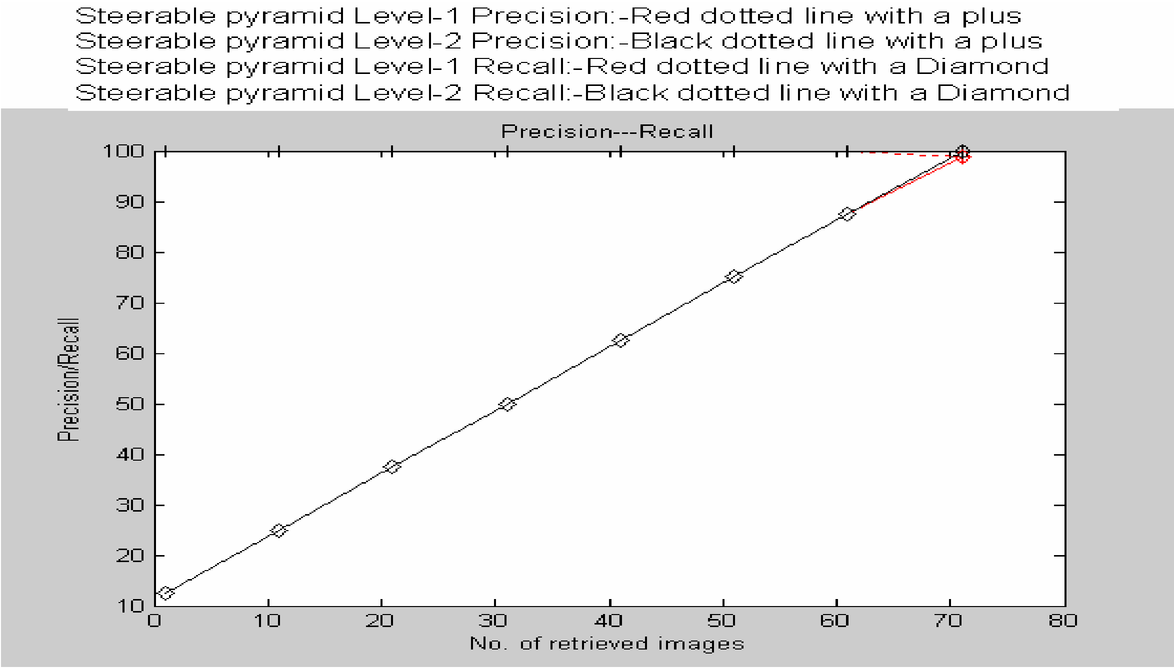

Figure 7: Precision /Recall for Query Image in Class'B'

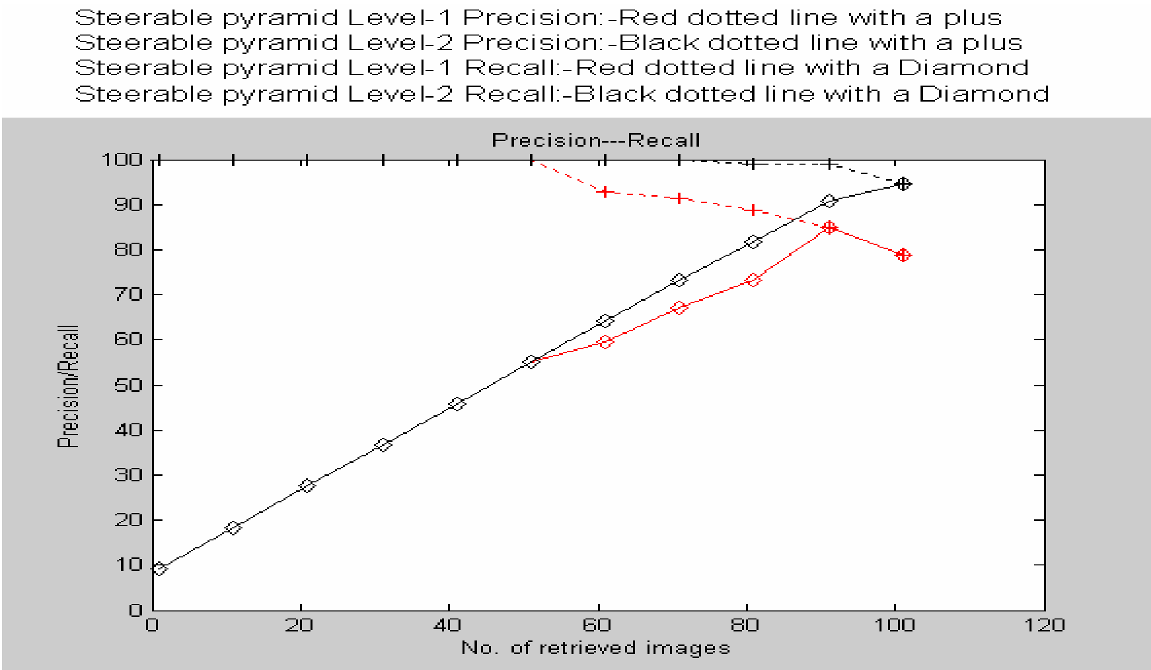

Figure 8: Precision /Recall for Query Image in Class'C'

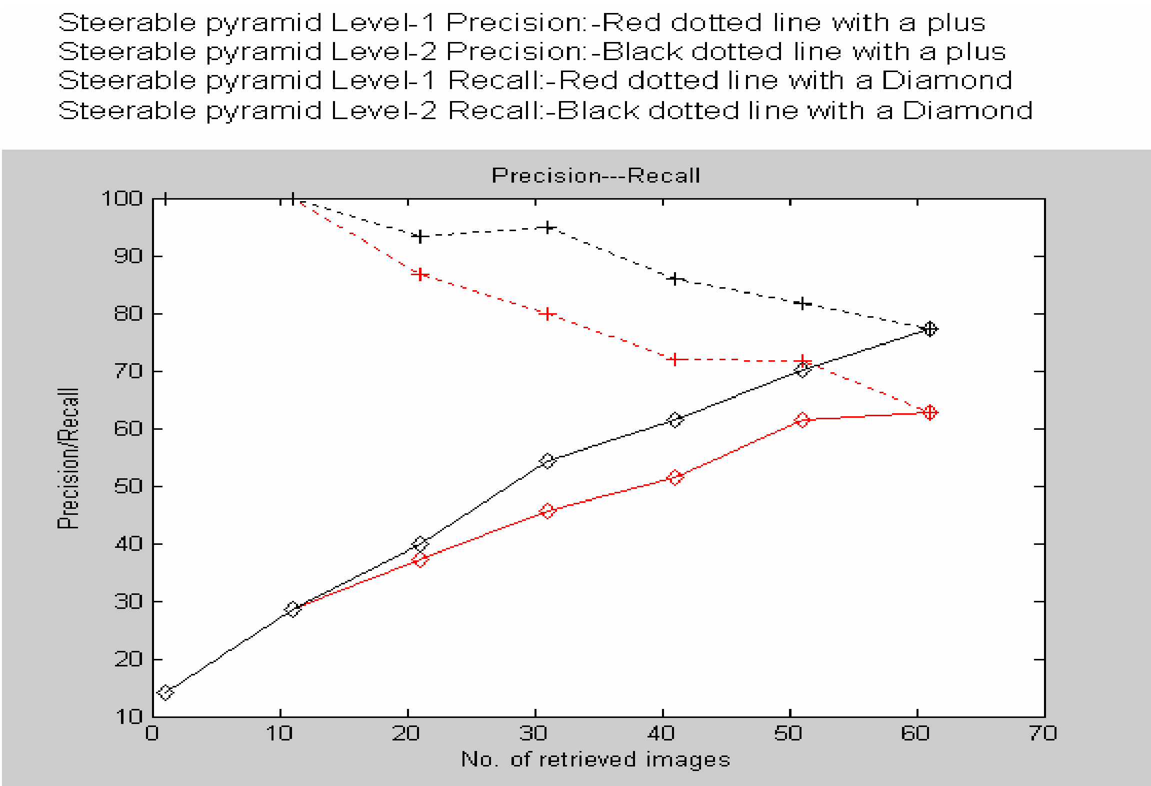

Figure 9: Precision /Recall for Query Image in Class 'D'

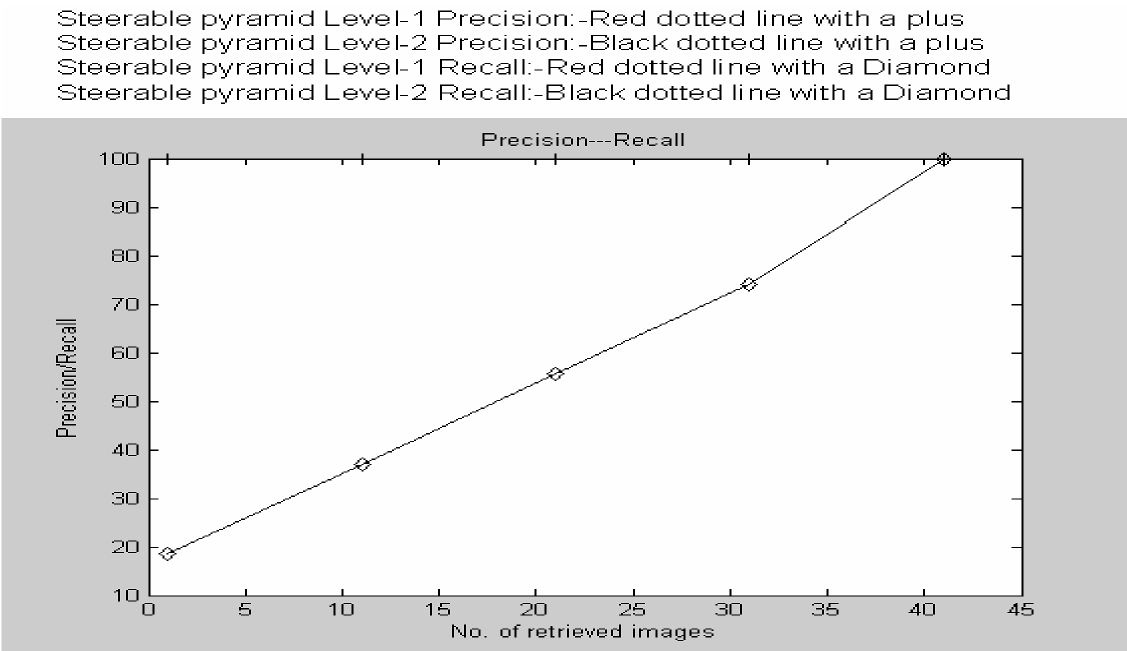

Figure 10: Precision /Recall for Query Image in Class 'E

From Figure 6 to Figure 10 it is clear that Recall increases as the number of retrieved items increases and the Precision is constant (i ,e 100% for query image in Class 'B' and Class 'E') and it decreases for query image in Class 'A' , Class 'C' and Class 'D' as the number of retrieved item increases. From Figure 6 to Figure 10, it is also clear that the precision/Recall values of Steeralbe Pyramid Level-2 are higher than Steerable Pyramid Level-1.

4. CONCLUSION

Efficient and effective retrieval techniques are needed to extract the similar images from the database. In this paper, the Steerable Pyramid (Level-1 and Level-2) based feature descriptor is used to retrieve the relevant images from the database. The Steerable Pyramid generates a multi-scale, multi-directional representation of the image. For evaluating the performance of the Steerable Pyramid Based Image Retrieval, we used an experimental database which consists of 445 images of size 128x128x3 with five different classes. The experimental results with the help of graphs clearly indicate that the Steerable Pyramid Level-2 is more efficient as compared to steerable pyramid Level-1.

5. REFERENCES

[1] Minh N. Do, Member, IEEE, and Martin Vetterli, Fellow, IEEE," Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance," IEEE Transactions On Image Processing, VOL.11, NO.2, February 2002.

[2] Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng," Fundamentals of Content-Based Image Retrieval,"

[3] Michele Saad," Content Based Image Retrieval Literature Survey ",EE 381K: Multi Dimensional Digital Signal Processing , March 18, 2008

[4] Lei Zhu, Chun Tang, Aibing Rao and Aidong Zhang,"Using Thesaurus To Model Keyblock-Based Image Retrieval ," Technical Report, Department of Computer Science and Engineering , State University of New York At Buffalo,Buffalo,NY 14260,USA.

[5] Truong T. Nguyen and Soontorn Oraintara ," Texture Image Retrieval Using Complex Directional Filter Bank ",Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019–0016.

[6] Ch.Srinivasa rao *, S. Srinivas kumar #, B.N.Chatterji ," Content Based Image Retrieval using Contourlet Transform ",Research scholar, ECE Dept., JNTUCE, Kakinada, A.P, India. , Professor of ECE, JNTUCE, Kakinada, A.P, India. Former Professor, E&ECE Dept., IIT, Kharagpur, W.B, India.