



## KNN AND STEERABLE PYRAMID BASED ENHANCED CONTENT BASED IMAGE RETRIEVAL MECHANISM

Bohar Singh <sup>(1)</sup>, Mrs.Mehak Aggarwal <sup>(2)</sup>

<sup>(1)</sup> Research Scholar, Department of Computer Science & Engineering, LLRIET, Moga

boharwalasra@gmail.com

<sup>(2)</sup> Associate Professor, Department of Computer Science & Engineering, LLRIET, Moga

mehakaggarwal@llriet.ac.in

### ABSTRACT

Recently, digital content has become a significant and inevitable asset of or any enterprise and the need for visual content management is on the rise as well. There has been an increase in attention towards the automated management and retrieval of digital images owing to the drastic development in the number and size of image databases. A significant and increasingly popular approach that aids in the retrieval of image data from a huge collection is called Content-based image retrieval (CBIR). Content-based image retrieval has attracted voluminous research in the last decade paving way for development of numerous techniques and systems besides creating interest on fields that support these systems. CBIR indexes the images based on the features obtained from visual content so as to facilitate speedy retrieval. Content based image retrieval from large resources has become an area of wide interest nowadays in many applications. In this thesis work, we present a steerable pyramid based image retrieval system that uses color, contours and texture as visual features to describe the content of an image region. To speed up retrieval and similarity computation, the database images are classified and the extracted regions are clustered according to their feature vectors using KNN algorithm. We have used steerable pyramid to extract texture features from query image and classified database images and store them in feature features. Therefore to answer a query our system does not need to search the entire database images; instead just a number of candidate images are required to be searched for image similarity. Our proposed system has the advantage of increasing the retrieval accuracy and decreasing the retrieval time.

### Keywords

Content based image retrieval (CBIR), color, shape, texture features, DWT, KNN, relevance feedback.

Language: English

Date of Publication: 16-08-2018

DOI: 10.24297/ijct.v17i2.7606

ISSN: 2277 - 3061

Volume: 17 Issue: 02

Journal: International Journal Of Computers & Technology

Website: <https://cirworld.com>



This work is licensed under a Creative Commons Attribution 4.0 International License.



## INTRODUCTION

Content based image retrieval is based on (automated) matching of the features of the query image with that of image database through some image-image similarity evaluation. Therefore, the images will be indexed according to their own visual content in the light of the underlying (chosen) features like color (distribution of color intensity across image, texture (presence of visual patterns that have properties of homogeneity and do not result from the presence of single color, or intensity), shape (boundaries, or the interiors of objects depicted in the image), or any other visual feature or a combination of a set of elementary visual features. Needless to say, the advantages and end users of such systems range from simple users searching a particular image on the web as well various type of professional bodies, like police force for picture recognition, journalists requesting pictures that match some query etc. From historical perspective, probably the first use of CBIR goes back to Kato in early nineties where he implemented what sounds to be the first automated image retrieval system using color and shape features. Since Kato's pioneer work, many prototypes of CBIR systems have been developed, and some of them did go to commercial market, e.g., IBM's QBIC system, which supports color, shape and texture feature, Virage developed by VirageInc and supports color, texture, color layout and shapes. However, it has been acknowledged the lack of maturity of current technology, which limited its large-scale deployment. This motivated the intensive research carried out in many aspects of CBIR including image indexing, feature selection and extraction, image similarity calculus and user's feedback/interactions.

A typical task solved by CBIR systems is that a user submits a query image or series of images and the system is required to retrieve images from the database as similar as possible. Another task is a support for browsing through large image databases, where the images are supposed to be grouped or organized in accordance with similar properties. Although the image retrieval has been an active research area for many years this difficult problem is still far from being solved. Any CBIR system involves at least four main steps: -

- Feature extraction and indexing of image database according to the chosen visual features, which form the perceptual feature space, e.g., color, shape, texture or any combination of the above.
- Feature extraction of query image (s).
- Matching the query image to the most similar images in the database according to some image-image similarity measure. This forms the search part of the CBIR
- User interface and feedback which governs the display of the outcomes, their ranking, the type of the user-interaction with possibility of refining the search through some automatic or manual preference (weighting) scheme, etc.

In a typical CBIR system (Figure 1), image low level features like color, texture, shape and spatial locations are represented in the form of a multidimensional feature vector. The feature vectors of images in the database form a feature database. The retrieval process is initiated when a user query the system using an example image or sketch of the object. The query image is converted into the internal representation of feature vector using the same feature extraction routine that was used for building the feature database. The similarity measure is employed to calculate the distance between the feature vectors of query image and those of the target images in the feature database. Finally, the retrieval is performed using an indexing scheme which facilitates the efficient searching of the image database.

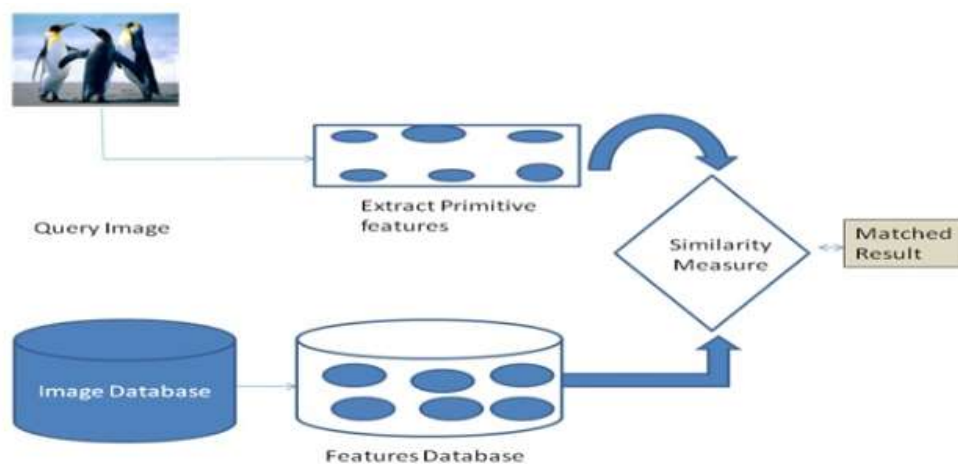


Figure 1. Architecture of CBIR



Recently, user's relevance feedback is also, incorporated to further improve the retrieval process in order to produce perceptually and semantically more meaningful retrieval results. In this chapter, we discuss these fundamental techniques for content-based image retrieval. CBIR is the application of computer vision to aid the image retrieval process of searching for digital images in large database based on the comparison of low level features of images. The search is carried out by using contents of the image themselves rather than relying on human-inputted metadata such as caption or keyword describing the image. Compared to text-based retrieval systems, CBIR is more feasible in large-scale databases and is usually used in environments which require fast retrieval and real-time operations. Software's which implements CBIR are known as content-based image retrieval systems (CBIRS). CBIR came to the interest of researchers as it offers the ability to index images based on content of the image itself [4]. In a standard CBIR based machine (Figure 1), image based features like color, texture, shape of an image and spatial locations are shown and represented in the form of a multidimensional feature vector. The characteristic vectors of images inside the database form a feature database. The retrieval process in CBIR system is started whenever a consumer question the system using query image or provides the sketch of the image [3]. The question photograph or the query image is converted into the internal illustration of feature vector using the equal characteristic extraction process that was used for constructing the feature database. The similarity measure or the degree is hired to calculate the distance among the feature vectors of query image and those of the target images inside the characteristic database of images. Finally, the retrieval is achieved by using an indexing scheme which facilitates the efficient searching of the image database. Recently, consumer's relevance remarks or we can say the feedback is likewise included to further enhance the retrieval method so one can produce perceptually and semantically greater meaningful retrieval effects using CBIR system [4]. CBIR retrieves images based on visual features such as colour, texture and shape [3]. In this method, colour, shape and texture of an image are classified automatically or semi-automatically with the aid of human classifier. Retrieval results are obtained by calculating the similarity between the query and images stored in the database using predefined distance measure. The results are then ranked according to the highest similarity score.

## RELATED WORK

Savvas A. Chatzichristofis et al. (2008) deals with a new low level feature that is extracted from the images and can be used for indexing and retrieval. This feature is called "Color and Edge Directivity Descriptor" and incorporates color and texture information in a histogram. CEDD size is limited to 54 bytes per image, rendering this descriptor suitable for use in large image databases. Chuen-Horng Lin et al. (2008) proposes three feature vectors for image retrieval. In addition, a feature selection technique is also brought forward to select optimal features to not only maximize the detection rate but also simplify the computation of image retrieval. Michal Perdoch et al. (2009) proposes a novel method for learning discretized local geometry representation based on minimization of average reprojection error in the space of ellipses. HerveJegou et al. (2010) addresses the problem of image search on a very large scale, where three constraints have to be considered jointly the accuracy of the search, its efficiency, and the memory usage of the representation. They first proposed a simple yet efficient way of aggregating local image descriptors into a vector of limited dimension, which can be viewed as a simplification of the Fisher kernel representation. SwapnaliniPattanaik et al. (2012) gives an overview idea of retrieving images from a large database. CBIR is used for automatic indexing and retrieval of images depending upon contents of images known as features. The features may be low level or High level. The low-level features include color, texture and shape. Yanzhi Chen et al. (2012) proposed a discriminative criterion for improving result quality. This criterion lends itself to the addition of extra query data, and they showed that multiple query images can be combined to produce enhanced results. Experiments compare the performance of the method to state-of-the-art in object retrieval, and show how performance is lifted by the inclusion of further query images. ReljaArandjelović et al. (2012) made the following three contributions: (i) a new method to compare SIFT descriptors (RootSIFT) which yields superior performance without increasing processing or storage requirements; (ii) a novel method for query expansion where a richer model for the query is learnt discriminatively in a form suited to immediate retrieval through efficient use of the inverted index; (iii) an improvement of the image augmentation method proposed by Turcot and Lowe where only the augmenting features which are spatially consistent with the augmented image are kept. Sumaira Muhammad et al. (2012) has given comparison of three different approaches of CBIR based on image features and similarity measures taken for finding the similarity between two images. Results have shown that selecting an important image feature and calculating that through a meaningful way is of great importance in image retrieval. GhanshyamRaghuwanshi et al. (2015) proposes a novel technique for texture image retrieval based on tetrolate transforms. Tetrolates provide fine texture information due to its different way of analysis. Jitendra Singh et al. (2016) proposes the content based image retrieval as one of most technique of data and multimedia technology. As image collections are growing at a rapid rate, and demand for efficient and effective tools for retrieval of query images from database is increased significantly. Zhijie Zhao et al. (2016) proposes a scheme which is based on three noticeable algorithms: color distribution entropy(CDE), color level co-occurrence(CLCM) and invariant moments. CDE takes the correlation of the



color spatial distribution in an image into consideration. CLCM matrix is the texture feature of the image, which is a new proposed descriptor that is grounded on co-occurrence matrix to seize the alteration of the texture.

## RESEARCH MOTIVATION

Image retrieval is an extension to traditional information retrieval. Approaches to image retrieval are somehow derived from conventional information retrieval and are designed to manage the more versatile and enormous amount of visual data which exist. Low-level visual features such as color, texture, shape and spatial relationships are directly related to perceptual aspects of image content. Since it is usually easy to extract and represent these features and fairly convenient to design similarity measures by using the statistical properties of these features, a variety of content-based image retrieval techniques have been proposed in the past few years. High-level concepts, however, are not extracted directly from visual contents, but they represent the relatively more important meanings of objects and scenes in the images that are perceived by human beings. These conceptual aspects are more closely related to users' preferences and subjectivity. Concepts may vary significantly in different circumstances. Subtle changes in the semantics may lead to dramatic conceptual differences. Needless to say, it is a very challenging task to extract and manage meaningful semantics and to make use of them to achieve more intelligent and user friendly retrieval.

## KNN

An instance based learning method called the K-Nearest Neighbor or K-NN algorithm has been used in many applications in areas such as data mining, statistical pattern recognition, image processing. Successful applications include recognition of handwriting, satellite image and EKG pattern. In data mining, we often need to compare samples to see how similar they are to each other. For samples whose features have continuous values, it is customary to consider samples to be similar to each other if the distances between them are small. Other than the most popular choice of Euclidean distance, there are of course many other ways to define distance. The k-means clustering algorithm attempts to split a given anonymous data set (a set containing no information as to class identity) into a fixed number ( $k$ ) of clusters. Initially  $k$  number of so called centroids are chosen. A centroid is a data point (imaginary or real) at the center of a cluster. In context each centroid is an existing data point in the given input data set, picked at random, such that all centroids are unique (that is, for all centroids  $c_i$  and  $c_j$ ,  $c_i \neq c_j$ ). These centroids are used to train a KNN classifier. The resulting classifier is used to classify (using  $k = 1$ ) the data and thereby produce an initial randomized set of clusters. Each centroid is thereafter set to the arithmetic mean of the cluster it defines. The process of classification and centroid adjustment is repeated until the values of the centroids stabilize. The final centroids will be used to produce the final classification/clustering of the input data, effectively turning the set of initially anonymous data points into a set of data points, each with a class identity.  $Cl$  is a categorical variable, and there is a scalar function,  $f$ , which assigns a class,  $y = f(x)$  to every such vectors. We do not know anything about  $f$  (otherwise there is no need for data mining) except that we assume that it is smooth in some sense. We suppose that a set of  $T$  such vectors are given together with their corresponding classes.

## STEERABLE PYRAMID

The Steerable Pyramid is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image-processing and computer vision applications. We developed this representation in 1990, in order to overcome the limitations of orthogonal separable wavelet decompositions that were then becoming popular for image processing (specifically, those representations are heavily aliased, and do not represent oblique orientations well). Once the orthogonality constraint is dropped, it makes sense to completely reconsider the filter design problem (as opposed to just re-using orthogonal wavelet filters in a redundant representation, as is done in cycle-spinning or undecimated wavelet transforms!).

The basic functions of the steerable pyramid are  $K$ th-order directional derivative operators (for any choice of  $K$ ), that come in different sizes and  $K+1$  orientations. As directional derivatives, they span a rotation-invariant subspace, and they are designed and sampled such that the whole transform forms a tight frame. An example decomposition of an image of a white disk on a black background is shown to the right. This particular steerable pyramid contains 4 orientation sub bands, at 2 scales. The smallest sub band is the residual low pass information. The residual high pass sub band is not shown. The block diagram for the decomposition (both analysis and synthesis) is shown to the right. Initially, the image is separated into low and high pass sub bands, using filters  $L_0$  and  $H_0$ .

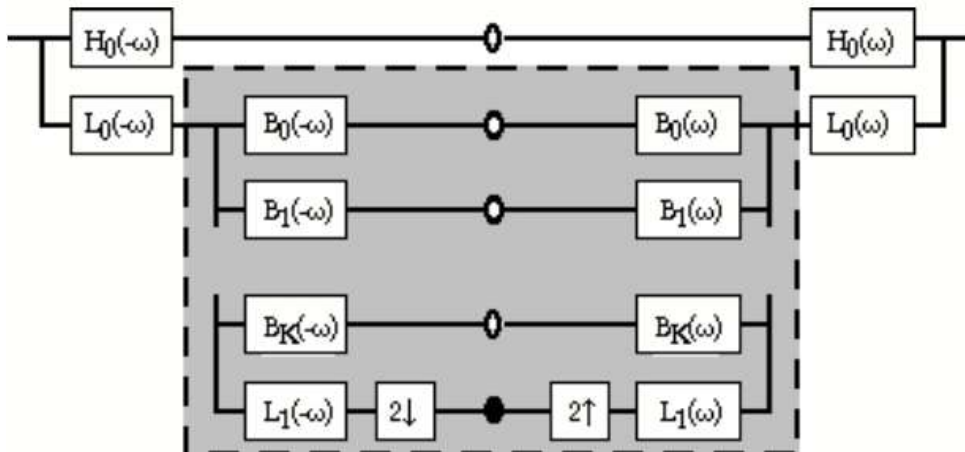


Figure 2. Decomposition using Steerable Pyramids

The low pass sub band is then divided into a set of oriented band pass sub bands and a lower-pass sub band. This lower pass sub band is subsampled by a factor of 2 in the X and Y directions. The recursive (pyramid) construction of a pyramid is achieved by inserting a copy of the shaded portion of the diagram at the location of the solid circle (i.e., the low pass branch). The right side of the diagram is the synthesis part. The synthesized image is reconstructed by up sampling the lower low-pass sub band by the factor of 2 and adding up with the set of band-pass sub bands and the high-pass sub band.

### RESEARCH METHODOLOGY

- D represents the number of images in the database and Q represents the set of query images.
- Extract the low level features like color, shape and texture of all database as well as query images.
- Match all the database images with the query images by using KNN. Then the result will comes as classified images and unclassified images. Classified images are stored in labeled folder and unclassified images are stored in unlabeled folder.
- Steerable pyramid mechanism is used to extract the various features like color, texture and contour from the query image and classified database images
- Multiple scales and rotation invariance is used by steerable pyramid to extract the features.
- Then Match the classified images with query images and display output.

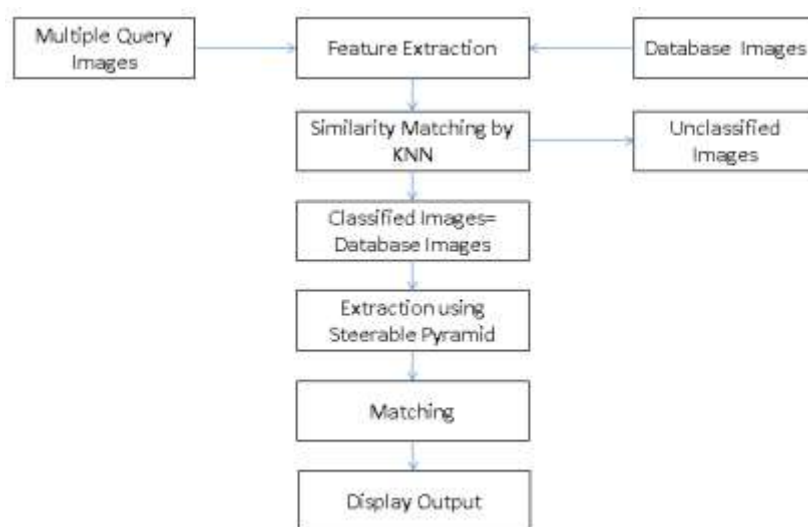


Figure 3. Flowchart of proposed work

### EXPERIMENTAL SETUP



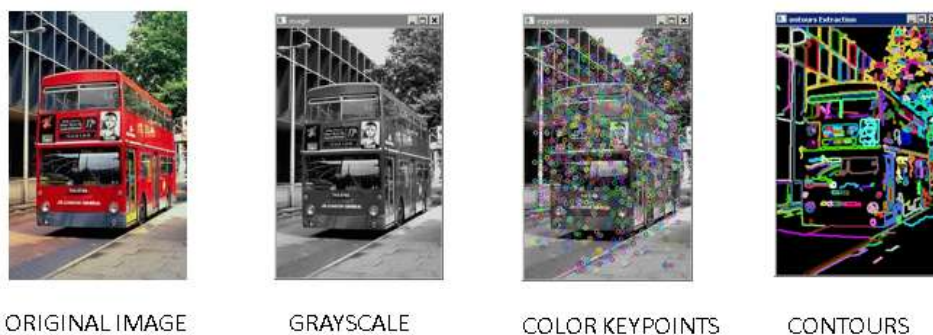
The user query is used to express the user's information needed to retrieve images in collection of database that conform to human perception. Querying by visual example is a paradigm, particularly suited to express perceptual aspects of low/intermediate features of visual content. The initial process is preprocessing the input image, for better performance. The input image has different contents and different format. Sometimes the same image is given twice with some small change in that image, but the comparison of these two images using pixel by pixel may not give the positive result. So the proposed methods with few enhancement techniques that enhance the image and then compare the image using only the modified images not the original one. WANG Database contains 1000 images which can be classified into different domains namely Buses, Dinosaurs, Flowers, Building, Elephants, Mountains, Food, African people, Beaches and Horses with JPEG format which used in a general purpose for experimentation. These images are stored with size 256x256 and each image is represented with RGB color space.



**Figure 4. WANG dataset images**

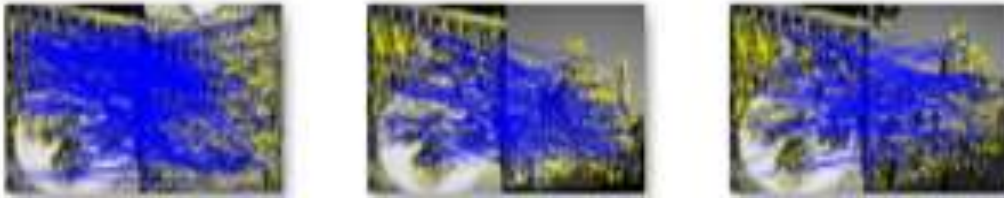
The open source computer vision library, OpenCV, began as a research project at Intel in 1998. It has C++, C, Python, Java and MATLAB interfaces and supports Windows, Linux, Android and Mac OS. OpenCV leans mostly towards real-time vision applications and takes advantage of MMX and SSE instructions when available. A full-featured CUDA and OpenCL interfaces are being actively developed right now. There are over 500 algorithms and about 10 times as many functions that compose or support those algorithms. OpenCV is written natively in C++ and has a templated interface that works seamlessly with STL containers. Officially launched in 1999, the OpenCV project was initially an Intel Research initiative to advance CPU-intensive applications, part of a series of projects including real-time ray tracing and 3D display walls. The main contributors to the project included a number of optimization experts in Intel Russia, as well as Intel's Performance Library Team

A retrieved image is considered to be correct if and only if it is in the same category as the query. The experiments are carried out in a personal computer with Intel Core i5 processor with 8GB RAM. The program is developed using OpenCV libraries and Visual Studio IDE. The features of the query image and database images are extracted using OpenCV libraries. Figure 5 shows the actual image along with some another features like grayscale version, extraction of color key points and contours of the sample image.



**Figure 5. Feature Extraction**

Figure 6 shows the matching of images after steerable pyramid feature extraction. The labelled images are matched with the query images and results are shown in figure 6.



**Figure 6. Matching**

## PERFORMANCE EVALUATION

The performance of a retrieval system is evaluated based on several criteria. Some of the commonly used performance measures are average precision, average recall, average retrieval rate. All these parameters are computed using precision and recall values computed for each query image.

**Table 1. Processing Time Comparison**

S.No.	Dataset Name	Existing Processing Time	Proposed Processing Time
1	Human	3228	1662
2	Beech	4049	2411
3	Building	6173.0	2146
4	Bus	6459	2047
5	Horse	3422	2310
6	Mountain	5596	1550
7	Food	3955	2093

Table 1 shows the processing time of the existing work and the proposed work for the same set of images. The processing time is calculated in milliseconds. Processing time has been reduced in the proposed work that shows the effectiveness of the proposed work. The precision of the retrieval is defined as the fraction of the retrieved images that are indeed relevant for the query. The recall is the fraction of relevant images that is returned by the query.

**Table 2. Precision Comparison**

Category	Existing Work Precision	Proposed Work Precision
Human	0.3	0.7
Beech	0.1	0.7
Building	0.1	0.9
Bus	0.7	1.0
Horse	0.4	0.8
Mountain	0.1	0.5
Food	0.7	1.0

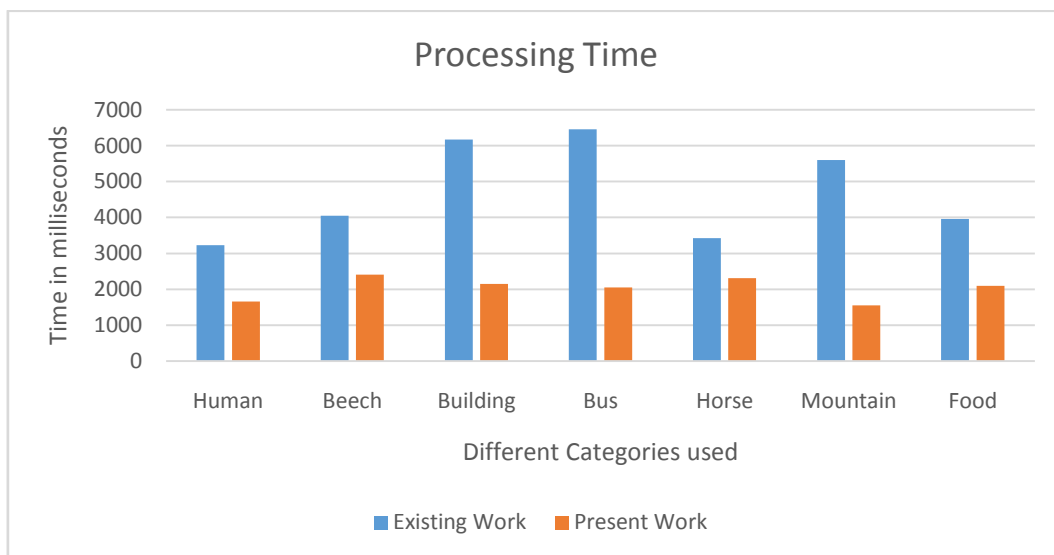


In Table 1, we have specified the precision comparison results for the different number of experiments for the existing work and proposed work. Different categories of images are taken in each experiment and the number of multiple query images are also different for some of the experiments. First of all KNN algorithm has been used for classification. KNN uses the query images for classifying the database images. More relevant images are grouped into one container and the remaining are grouped into another container. After applying the KNN, steerable pyramid has been applied. It is clear from the table 2, the precision of the proposed work has been increased for all the different categories of images.

**Table 3. Recall comparison**

Dataset Name	Existing work Recall	Proposed Work Recall
Human	0.1	0.2
Beech	0.0	0.2
Building	0.0	0.2
Bus	0.2	0.4
Horse	0.1	0.2
Mountain	0.0	0.1
Food	0.2	0.3

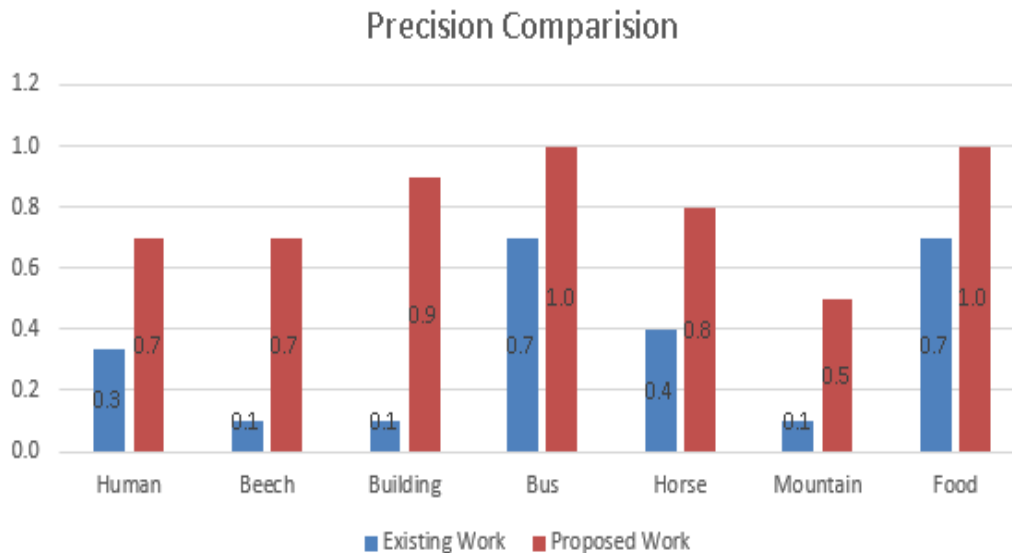
In the table 3, we have specified the recall values of existing work and the proposed work. It is evident from the values, that the recall has been improved in the proposed work.



**Figure 7. Processing Time comparison**

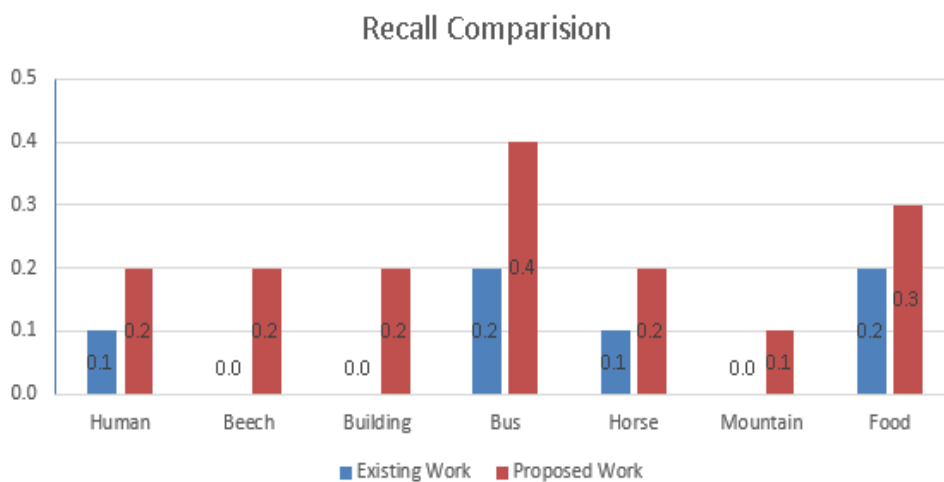
Figure 7. shows the significant improvement in the execution time for the proposed work. The execution time has been reduced because we have reduced the database images after applying the KNN classification mechanism. There is an improvement of approximately 30-40% in overall execution time.





**Figure 8. Precision Comparison**

In the figure 8 and figure 9, we have illustrated the precision and recall graphs of the proposed work. Both the graphs are depicting a significant improvement that shows the effectiveness of our proposed work. Recall is showing an upward trend in the proposed work that shows the improvement over the existing work.



**Figure 9. Recall Comparison**

## CONCLUSION

CBIR is a vast research area and has many open questions and challenges. Designing a CBIR system involves choosing particular feature representation techniques, optimal dimensionality and reliable similarity functions in order to achieve best results. The ultimate aim is to reduce the gap between semantic information in the image and the extracted low-level features. The developed CBIR system can be extended to include stronger features and additional learning capabilities. This will provide higher accuracy values thus facilitating the investigation of results. A larger database can also be used to increase confidence in the results obtained. Furthermore, experiments can be run on a different data set for more rigorous proof of concept. We have reviewed the main components of a content based image retrieval system by applying KNN with steerable pyramid and using multiple query images, including image feature representation, indexing, query processing, and query-image matching and user's interaction, while highlighting the current state of the art and the key - challenges. It has been acknowledged that it remains much room for potential improvement in the development of content based image retrieval system due to semantic gap between image similarity outcome and user's perception. After evaluating the results, we have reached up to the solution that we have been able to improve the CBIR mechanism using the proposed mechanism in this work. We will further test and benchmark this integrated image retrieval framework over various large image databases, along with tuning the relevance feedback to achieve optimal performance with highly reduced dimensionality.



## REFERENCES

- [1] Amirani, "Framework for image retrieval using machine learning and statistical similarity matching techniques," IET Image Process, pp. 1-11, 2013.
- [2] P. MANIPOONCHELVI and K. MUNEESSWARAN, "Multi region based image retrieval system," Indian Academy of Sciences, pp. 333-344, 2014.
- [3] H. J'egou, M. Douze, C. Schmid and P. P'erez, "Aggregating local descriptors into a compact image representation," IEEE, pp. 3304-3311, 2010.
- [4] Y. Chen, J. Z. Wang and R. Krovetz, "CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning," IEEE, pp. 1187-1201, 2005.
- [5] R. Fergus, L. Fei-Fei, P. Perona and A. Zisserman, "Learning Object Categories from Google's Image Search," IEEE, 2005.
- [6] Y. Chen, X. Li, A. Dick and A. v. d. Hengel, "Boosting Object Retrieval with Group Queries," IEEE, pp. 765-768, 2012.
- [7] R. Arandjelovi'c and A. Zisserman, "Three things everyone should know to improve object retrieval," IEEE, pp. 2911-2918, 2012.
- [8] M. Perd'och, Chum and J. Matas, IEEE, pp. 9-16, 2009.
- [9] S. A. Chatzichristofis and Y. S. Boutalis, "CEDD: Color and Edge Directivity Descriptor: A Compact Descriptor for Image Indexing and Retrieval," Springer-Verlag Berlin Heidelberg, pp. 313-322, 2008.
- [10] Y. Chen and J. Z. Wang, "A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval," IEEE, pp. 1252-1267, 2002.
- [11] C.-H. Lin, R.-T. Chen and Y.-K. Chan, "A smart content-based image retrieval system based on color and texture feature," ELSEVIER, p. 658-665, 2009.
- [12] Z. Wang, J. Li and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE, pp. 947-963, 2001.
- [13] S. Gandhani, R. Bhujade and A. Sinhal, "AN IMPROVED AND EFFICIENT IMPLEMENTATION OF CBIR SYSTEM BASED ON COMBINED FEATURES," IET, pp. 353-359.
- [14] S. M. H. Khan, A. Hussain and I. F. T. Alshaikh, "Comparative study on Content Based Image Retrieval (CBIR)," IEEE, pp. 61-66, 2013.
- [15] Sreedevi S. and Shinto Sebastian, "Fast Image Retrieval with Feature Levels," IEEE, 2013.
- [16] S. Ezekiel, Mark G. Alford, David Ferris and Eric Jones,, "Multi-Scale Decomposition Tool for Content Based Image Retrieval," IEEE, 2013.
- [17] K. Juneja, A. Verma , S. Goel and S. Goel , "A Survey on Recent Image Indexing and Retrieval Techniques for Low-level Feature Extraction in CBIR systems," IEEE, pp. 67-72, 2015.
- [18] K. BELATTAR and S. MOSTEFAL, "CBIR using Relevance Feedback: Comparative Analysis and Major Challenges," IEEE, pp. 317-325, 2013.
- [19] D. Jeyabharathi and A. Suruliandi, "Performance Analysis of Feature Extraction and Classification Techniques in CBIR," IEEE, pp. 1211-1214, 2013.
- [20] H. Xie, Y. Ji and Y. Lu, "An Analogy-Relevance Feedback CBIR Method Using Multiple Features," IEEE, pp. 83-86, 2013.
- [21] B. Kaur and S. Jindal, "An implementation of Feature Extraction over medical images on OPEN CV Environment".
- [22] S. Kumar, S. Jain and T. Zaveri, "ARALLEL APPROACH TO EXPEDITE MORPHOLOGICAL FEATURE EXTRACTION OF REMOTE SENSING IMAGES FOR CBIR SYSTEM," IEEE, pp. 2471-2474, 2014.
- [23] K. BELATTAR and S. MOSTEFAL, "CBIR with RF: which Technique for which Image," IEEE, 2013.
- [24] G. Raghuwanshi and V. Tyagi, "Texture image retrieval using adaptive tetrolet transforms," ELSEVIER, pp. 1-8, 2015.