

# Content based Classification and Retrieval of Images

Dayanand Jamkhandikar  
Research Scholar., CS&E Dept.  
Niet, Nims University  
Jaipur, India

Anjali Anuveer  
CS&E Dept.  
Gndec, Bidar

Dr. Surendra Pal Singh  
Associate Prof. CS&E Dept.  
Niet, Nims University  
Jaipur, India

**Abstract**—CBIR is a challenging task. The SIFT and HARRIS corner detection approach takes image and transforms into large collection of local feature vectors. In addition, Sift is invariant to translation, rotation and scaling. K-means clustering is used to cluster the descriptors forming visual words (cluster center) and finally image is represented in the form of histogram where each histogram bin is called as visual word. The (dis)similarity is measured by comparing the histograms of query image with that of database images and displays the corresponding retrieved images. Construct classifier SVM, KNN for determining which category the image belongs. Our experiments shows high rate of retrieval effectiveness in terms of precision and recall.

**Keywords**— CBIR, SIFT, HARRIS, K-means clustering, SVM, KNN

## I. INTRODUCTION

Image Retrieval system is an effective and efficient tool for managing large image databases. Image retrieval is classified into two types: Text Based Image Retrieval and Content Based Image Retrieval. Text Based Image Retrieval is to retrieve based on text. Content-Based Image Retrieval (CBIR) is defined as a process to find similar image from the database when a query image is given. So, the user has to present a query image in order to retrieve images stored in the database according to the similarity of the query image.

Advances in data storage, data transmission, and image acquisition have enabled the creation of large image datasets. It has spurred great interest for systems that are able to efficiently retrieve images from these collections. This task has been addressed by the so-called *Content-Based Image Retrieval* (CBIR) systems [8]. Cbir uses three types of features to retrieve images:

(1)Color (2) Shape (3) Texture

Content based image retrieval system aims to retrieve the most similar images as related to the given query image by the user. Assuming “color” feature, given a query image we would like to obtain a list of images from the database which are most similar in color to the query image. Solving this problem, mentioned two concepts need to be developed, first the feature which represents the color information of

the image and the second one is similarity measure to compute the similarity between the feature values of two images. Research on image retrieval is in trend of high importance, as it has applications in various domains such as bio-medical, astronomy, remote sensing, and digital library.

Image classification is one of classical problems of concern in image processing. The objective of image categorization is to predict the categories of the input image by utilizing its elements. There are different methodologies for classification which are k-nearest neighbour (KNN), adaptive boost (adaboosted), artificial neural network (nn), support vector machine (SVM) etc. In our project we have used svm, knn classifier to classify images.

## II. RELATED WORK

We survey the techniques for image acquisition and methods related to object classification. The features used to retrieve images are based on color, shape, texture and in some papers it is used in combination. The feature extraction and matching algorithms are required to retrieve images. Some of the feature extraction algorithms are Canny edge, Fourier transform, Sift, Harris corner etc. And for matching Dynamic matching, histogram matching, Euclidean distance etc. For classification SVM, KNN, ADABOOST, NN etc. Nguyen *et al.* [2] utilize Support Vector Machine (SVM) for selecting the salient feature points on a shape and describe each point with well-designed local descriptors. Dynamic programming (DP) is then applied to match the local features in the given pair of shape for the evaluation of similarity. Then again, Li *et al.* [8] plan another local descriptor named ROMS, and after that develop a global representation with the extricated components by performing unsupervised adapting under the Bag-of-Words (BoW) system.

On the other hand, it is seen in their investigation that the discriminative power of the BoW representation worsens in examination to the immediate utilization of dynamic programming.

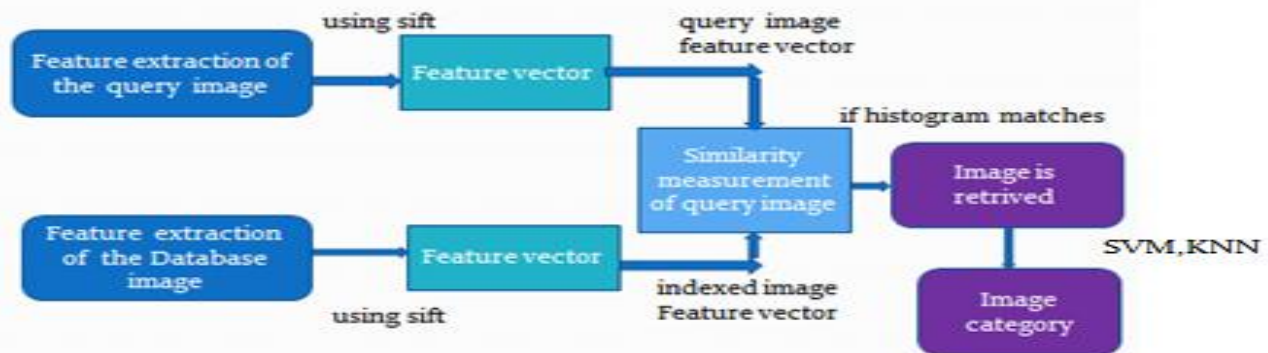


Fig 1: Block diagram of proposed model

As we can see, the presented learning based shape descriptors have shown their potential fit as a fiddle representation, yet there is still space for improvement with respect to time efficiency and descriptive capacity. CBIR solutions fail towards capturing some local components representing to the points of interest and subtleties of the scenes [10]. These facts can be gotten by mapping low-level components into middle-level and high-level semantics [8].

Among them, the proposed Scale Invariant Features Transform (SIFT) approach depends on the decision of a few parameters which straightforwardly affect its adequacy when smeared to retrieve images. Sift gives feature vectors of image which is extremely useful in matching stage, of query image with that of the database images. Our proposed work uses combination of SIFT and Harris for feature extraction and given descriptors to K-means algorithm which forms Bag-of-Words (BoW) paradigm in natural image matching. Like the system taken in Spatial Pyramid Matching (SPM) [3], a component division based procedure is applied to encoded element pooling and additionally shape vocabulary (or code book) learning. It expects to help better recognize the visual primitives in the corresponding feature space and meanwhile incorporate local and global characteristics of image.

### III. IMPLEMENTATION

It is more suitable to explain the whole system with the help of a block diagram. The complete block diagram of the whole system is shown in Fig. 1.

We have divided our work into three modules which are:

- 1) Feature extraction-SIFT HARRIS corner detection.
- 2) Quantization-K-Means clustering.
- 3) Classification-SVM AND KNN.

#### A. Feature extraction-SIFT,HARRIS

SIFT takes image and transforms into feature vectors. Firstly Sift generates keypoints and then form descriptors for keypoints with each keypoint having 128 feature vectors. Keypoints were then taken as maxima / minima of the Difference of Gaussian (DoG) that occur at multiple scales [13] [15]. Specially, a DoG image  $D(x, y, \sigma)$  given by:

$$D(x, y, \sigma) = L(x, y, k_i \sigma) - L(x, y, k_j \sigma) \quad (1)$$

Where  $L(x, y, k\sigma)$  is the convolution of the original image  $I(x, y)$  with the Gaussian blur  $G(x, y, k\sigma)$  at scale  $k\sigma$  i.e the scale space of an image is defined as function,  $L(x, y, \sigma)$  which was derived from the convolution of a variable-scale Gaussian [13] [15],  $G(x, y, \sigma)$  with an input image,  $I(x, y)$ .

$$L(x, y, k\sigma) = G(x, y, k\sigma) * I(x, y) \quad (2)$$

To have magnitude for keypoint

$$m(x, y) = (\sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}) \quad (3)$$

To have direction for keypoint

$$\theta(x, y) = \tan^{-1} \{ (L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y)) \} \quad (4)$$

So keypoint transforms into feature vector with both the magnitude and direction. And each keypoint is said to have 128 feature vectors called descriptors.

Harris corner: To detect corners in image and corners are called interest points. It uses sift descriptor.

#### B. Quantization:- Kmeans -Clustering

Create the visual vocabulary or bag of features, by taking all the feature descriptors from both Harris and Sift, use k-means clustering to cluster our descriptors. Each cluster center is called as visual word and histogram of visual words (bag of words) is final representation of image.

Typically, a large number of clusters (e.g., 500, 1000, and 2000) have been used in the literature. In our experiment we have used  $k=500$ . each category has 200 training images and 50 test image. We have conducted our experiments using Matlab 7.14 (R2012a) version tool. Matlab is a good tool to implement image processing related projects, so have used that tool.

C. Classification:-SVM, KNN

Constructing the SVM classifier and KNN (K-nearest neighbor) to determine the category to which the image belongs. And we have done the analysis to determine which classifier is more efficient.

IV. RESULTS AND ANALYSIS

We have used 800 Caltech database images <http://www.robots.ox.ac.uk/~vgg/data3.html> , <http://www.vision.caltech.edu/html-files/archive.html> with 4 categories aeroplane, car, face, motorbike, and



Fig 2: Query image

The first step for project execution is to select the query image from database. In fig2, the given query image is mortorbike image.

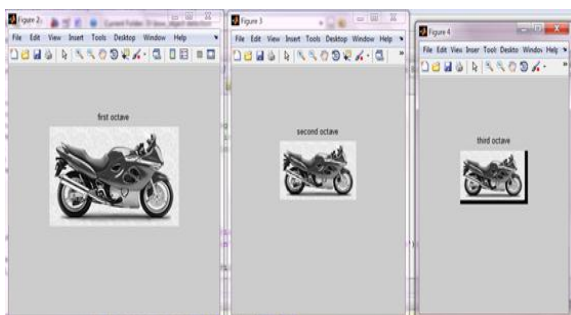


Fig 3: Octave generation

RGB image has to be first converted to gray image for faster processing. In fig3, the 3 octaves are formed for the query image using the sift algorithm. And then, the keypoints have to be formed for the octaves being generated. We have partioned our database into training and testing images. It is found that we can take any one of the 4 category of image from internet or any album and be put in test folder. Then it's to run the project, retrieve the

similar set of images and will classify the image to which category it belongs.

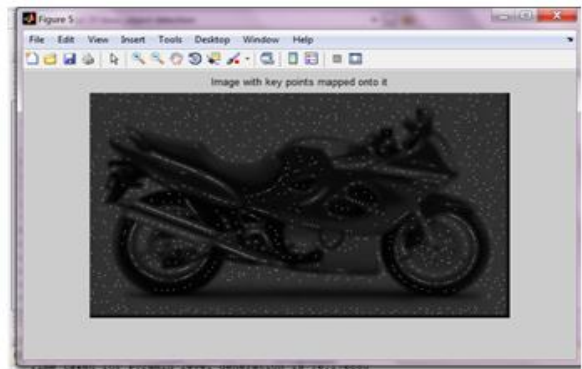


Fig 4: Keypoints for image

Keypoints are formed for the query image which are features for image.

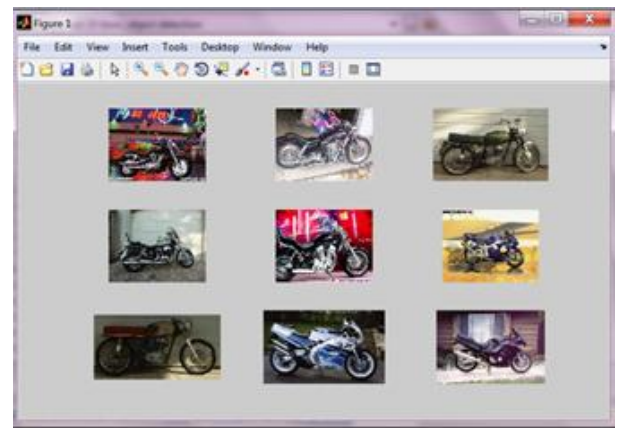


Fig 5: Images retrieved

In fig5, 9 motorbike images are retrieved for the query image of fig2 and the category to which the image belongs is also being displayed and shown in fig6 which is shown in command window that image belongs to category motorbike using K-NN classifier.

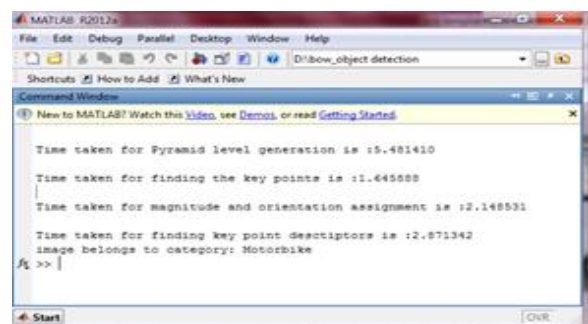


Fig 6: K-NN classifier for displaying category for retrieved images

The time taken for pyramid level generation, keypoints, keypoint magnitude and orientation, keypoint descriptors are shown in command window along with which category the image belongs are shown in fig6.

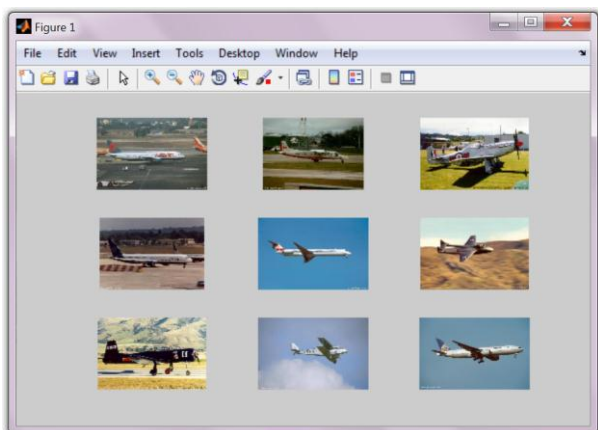


Fig 7: Images retrieved for category aeroplane

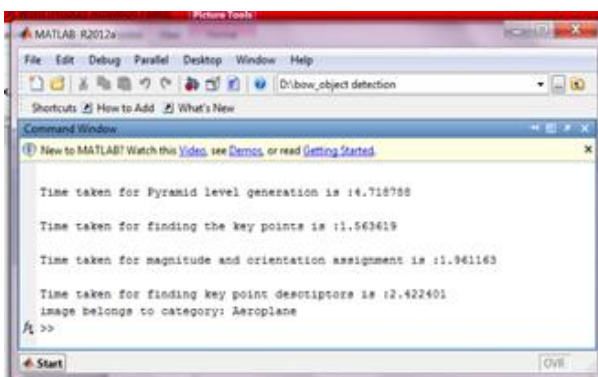


Fig 8: SVM classifier used for displaying category of image

In fig7, 9 aeroplane images are retrieved for query image aeroplane and the category to which the image belongs is also being displayed and shown in fig9 which is shown in command window that image belongs to category aeroplane using SVM classifier. We have done the analysis to determine the efficiency for our algorithms and performance evaluation is done with parameters precision and recall.

Below table1 gives the average value of precision and recall taking all categories car, aeroplane, faces, motorbike.

Table 1. Average value of precision, recall

Input images	Values for precision	Values for recall
1	0.011	1
2	0.011	1
3	0.011	1
4	0.01	0.88
5	0.01	0.88
6	0.01	0.88
7	0.0087	0.77
8	0.0087	0.77
9	0.0087	0.77
10	0.75	0.6

In fig.9 the average precision-recall graph shows that we have obtained good precision-recall rate.

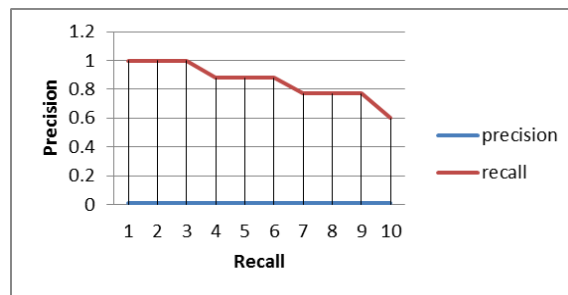


Fig 9: Average precision-recall graph

As per the analysis done for classifiers, we found from table2 that SVM classifier is more efficient than KNN classifier.

Table 2: Classification accuracy

Algorithm	Accuracy (%)
Class segment sets [11]	69.7
Inner Distance Shape Context(DP) [4]	73.6
Lim [5]	80.4
Our descriptor (KNN)	76.42
Our descriptor (SVM)	84

## V. CONCLUSION

Image Retrieval is gaining momentum among researchers working in image processing and computer vision areas because of the wide number of applications. In this paper, an algorithm for feature extraction and classification has been proposed for object recognition and retrieving similar images. The advantage of using Sift algorithm is that it provides 128 feature vectors for every single keypoint in image and is very useful during matching. In addition, Sift is invariant to translation, rotation and scaling, as well as faster in terms of processing time. We used K-means clustering for clustering descriptors generating bag of words (BOW). By experiments, we find that our algorithm is efficient with minimum delay and with high rate of retrieval accuracy measured in terms of precision and recall.

## VI. REFERENCES

- [1] Xing gang Wang and Xiang Bai, "Shape Vocabulary: A Robust and Efficient Shape Representation for Shape Matching" *IEEE transactions on image processing*, vol. 23, no. 9, September 2014 3935.
- [2] F. Porikli and H. Nguyen, "Support vector shape: A classifier based shape representation," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, vol. 35, no. 4, pp. 970-982, 2013.
- [3] J. Ponce, C. Schmid and S. Lazebnik, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2. IEEE, 2006, pp. 2169-2178.

- [4] H. Ling and D. W. Jacobs, "Shape classification using the innerdistance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.
- [5] K.-L. Lim and H. K. Galoogahi, "Shape classification using local and global features," in *Proc. 4th Pacific-Rim Symp. Image Video Technol.*, 2010, pp. 115–120.
- [6] J.-P. Vandeborre, M. Daoudi, H. Tabia, and O. Colot, "A new 3D-matching method of nonrigid and partially similar models using curve analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 852–858, Apr. 2011.
- [7] G. Lavoué, "Combination of bag-of-words descriptors for robust partial shape retrieval," *Vis. Comput.*, vol. 28, no. 9, pp. 931–942, 2012.
- [8] G. H. Freeman, N. Alajlan, M. S. Kamel, and, "Geometry-based image retrieval in binary image databases," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1003–1013, Jun. 2008.
- [9] L. J. Latecki, B. Feng, X. Wang, X. Bai and W. Liu, and, "Bag of contour fragments for robust shape classification," *Pattern Recognit.*, vol. 47, no. 6, pp. 2116–2125, 2014.
- [10] B. Sankur, C. B. Akgul, F. Schmitt and Y. Yemez, "3D model retrieval using probability density-based shape descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 6, pp. 1117–1133, Jun. 2009.
- [11] H. K. Galoogahi and K.-L. Lim and, "Shape classification using local and global features," in *Proc. 4th Pacific-Rim Symp. Image Video Technol.*, 2010, pp. 115–120.