I J C T A, 10(9), 2017, pp. 511-517 ISSN: 0974-5572

Semantic Image Retrieval: An approach for Image Retrieval using Object Recognition

Sushila Aghav*

ABSTRACT

Content based Image Retrieval is well studied area for retrieving the similar images based on the contents of the image. Colour, Shape, Texture descriptors are used in CBIR for similar image retrieval. These approaches face the problem of semantic gap in image retrieval. Image understanding plays important role to cover the semantic gap. Image understanding by means of object recognition approach attempts to detect and identify the Object players in images and their relationships. Hypotheses here are to understand objects in images enables filling semantic gap in image retrieval. Bag of Word model is discussed for object identification & recognition. Each image represented using Object name, their counts, and described relationship in Feature vector. Based on the similarity of feature vectors Images are retrieved. Image retrieval accuracy is depends on the accurate object recognition.

Keywords: Image Retrieval, Object Recognition, BOW

1. INTRODUCTION

The repository of Digital Images is becoming huge day by day. Huge number of images and videos are uploaded in social media and needs different approach to handle to search the meaningful images from this huge repository. CBIR systems are helpful in retrieving similar images from the repository. Basic contents of the image like Color, Shape and Texture are useful to analyze the similarity of the images. These Features have been used in literature to find out similar images in successful way. Color feature is been used in enormous researches with the different feature extraction approach like Histogram. Similarly Texture Features are successfully used in textured images to find out similar images.

1.1. Current CBIR System Overview

Content based image retrieval is based on the similarity of images [1]. To check the similarity between images, various low level content descriptors are used. Commonly used descriptors are color descriptors,

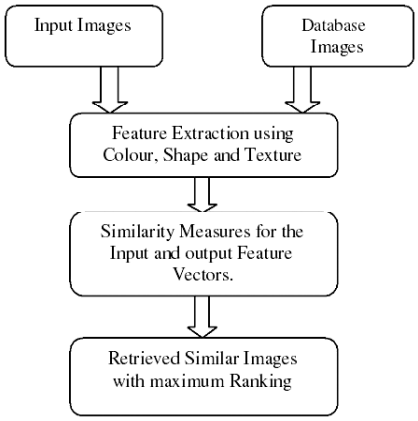

Figure 1: CBIR System

* Research Scholar, NMIMS, Assistant Professor MITCOE, Pune

shape descriptors and texture descriptors. In the following section brief overview of these descriptors are discussed.

1.2. Image Descriptors

Descriptors are tools to describe image. Basically in the CBIR applications, color, shape, textures are used ad Image descriptors.

1.2.1. Color Descriptors

Each digital image holds the color information as RGB color plane. These RGB plane treated as color descriptor. The image represented as RGB values and pixels of images with each value. These are color descriptors. These color descriptors are further minimized to reduce feature vector size. Color histograms, Color Moments, Color coherence vector, Colour Correlogram are few examples to represent Color Image descriptors. Histogram based Color descriptors are popular and easy to implement.

1.2.2. Texture Descriptors

Spatial distribution of some basic primitives creates some visual pattern called as texture descriptors [2]. These patterns are extracted using some methods and considered as Image descriptors. The various methods for modeling textures and extracting texture features are of four broad categories of problems: texture segmentation, texture classification, texture synthesis, and shape from Texture.

1.2.3. Shape Descriptors

Images have components like objects and regions. Image represented using various shapes inside it. Shape descriptors are the regions inside the image that follow certain properties. Shape descriptors are popular and is further used to detect objects inside the Image. Methods to extract shape descriptors are boundary based and regions based. More information regarding shape descriptor can be found in [3]

1.3. Semantic Gap in Retrieved Image

Current CBIR systems with the above discussed descriptors have limitation of semantic gap between the input and retrieved images. Though these CBIR are successful in retrieving the similar images, these systems may retrieve the images that are totally dissimilar. Two images have similar color descriptors may be different. Images that have similar shapes inside it may have different context. Texture images are useful only in good textured images.

1.4. Approaches to Fill Semantic Gap in image retrieval

To retrieve semantically similar images, one of the successful approach is image annotation. In image annotation, each image is explicitly annotated with it meaning. Thus all images that matches the input image contents meaning, gets retrieved. [1][4].

Problem with this approach is, it is tedious and impossible to annotate images in huge database. So the method that may annotate the image, automatically, by understanding the image with high level contents, is required.

Object recognition play important role by finding objects and their relationship [4] [5]. Images with similar objects, their association are semantically similar. The objective of the approach is to understand the image by recognizing the objects present in it. Then from collection of images retrieve those images with similar objects.

2. MOTIVATION

Many applications based on computer vision systems require retrieving the similar images. Major applications of CBIR like medical image retrieval, biometric, face recognition, biodiversity image retrieval; digital libraries are using image retrieval in their domain.

Along with these applications, it is important to track the image database for some activity happening, or surveying the image repositories on Social Networking sites for some particular meaning / activity images, that may first require understanding images semantically. By means of Object recognition, it is possible to track such details. Applications of semantic image retrieval are Terrorism activities, some sensitive / violence images tracking and retrieval based on some related input images.

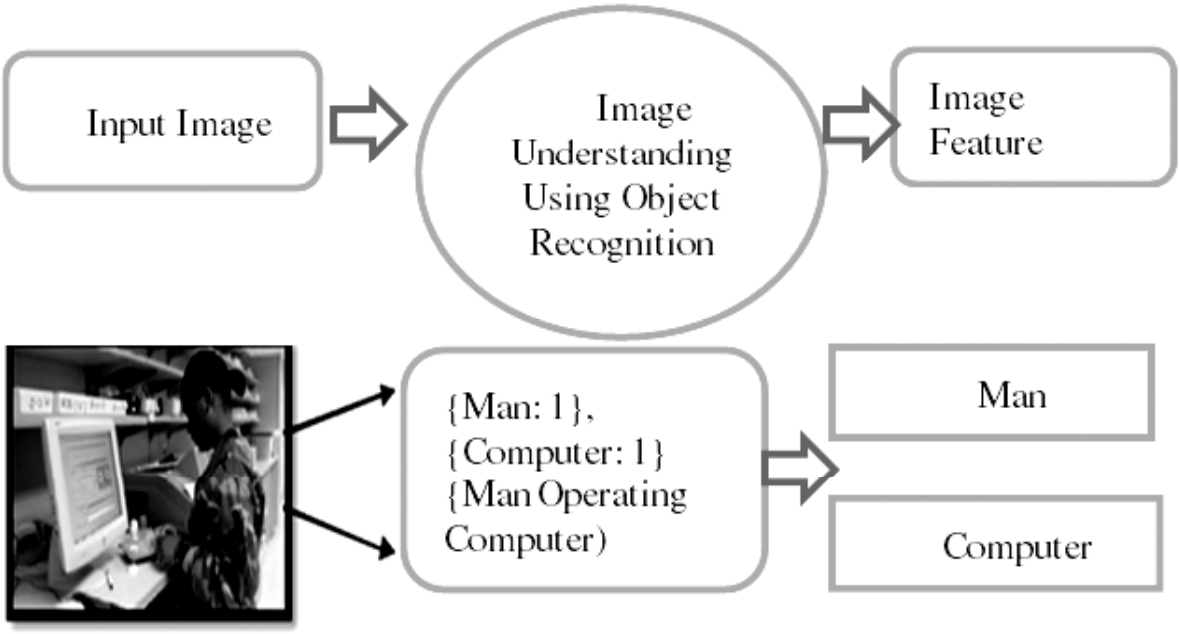

Figure 2: Extracting Meaning from Image

3. METHODOLOGY

To retrieve semantically similar images, Object recognition play important role by fin ding objects and their relationship. Images with similar objects, their association are semantically similar. The objective of the approach is to understand the image by recognizing the objects present in it. Then from collection of images retrieve those images with similar objects. As shown in figure.3 following are Steps:

Steps 1: Object Recognition u sing Bag of Word:

1. Image is scanned for the localization of the objects.
2. Detect the Interest points of the located region.
3. Extract the small areas around those interest points. Find the Local descriptors for these patches.
4. Cluster the Local descriptors. Identify mean of each cluster. Collection of the means is Visual words for the Object.
5. Train the classifiers using Visual Words of Training Objects.
6. Run the classification model for the object recognition using the above Visual words for Object recognition.
7. Annotate the Image with the Recognized Object and its relationships.

Step 2: Image Retrieval:

1. Input the Image.
2. Run the Step1 for it.
3. Search Image database and find images with similar objects as the query image with similarity measures.
4. Output the Images.

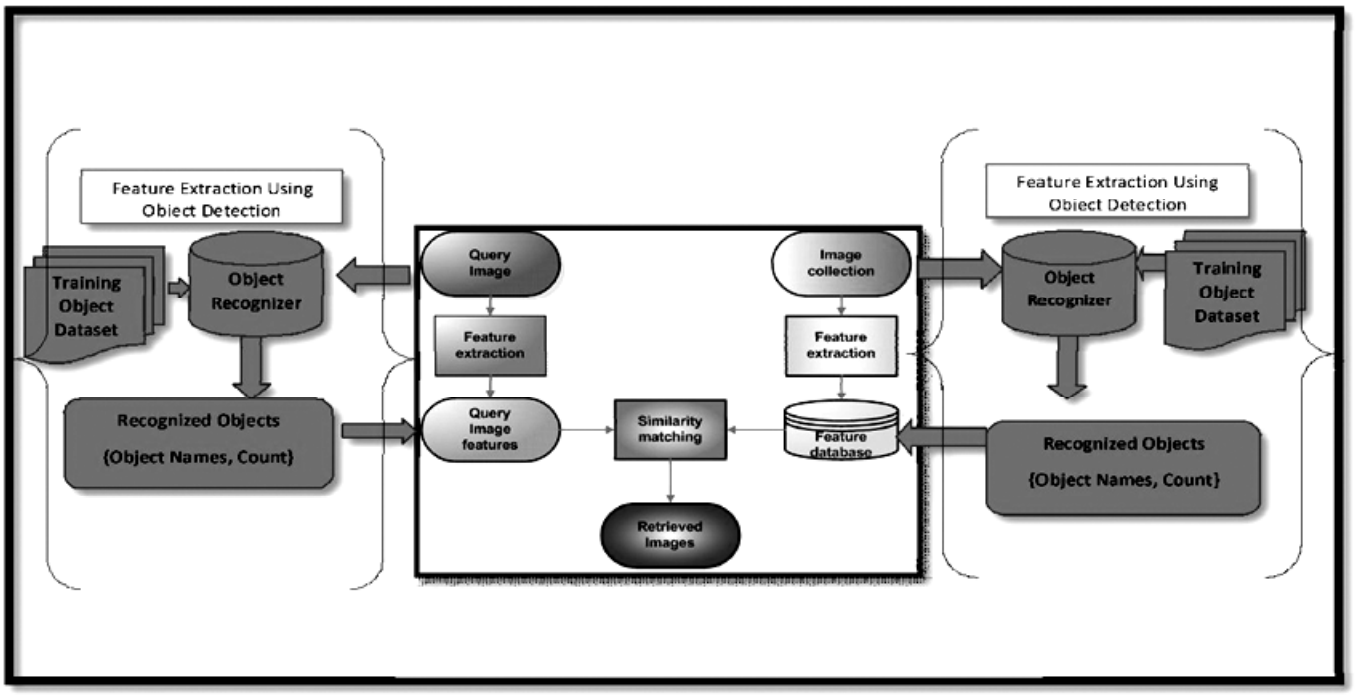

Figure 3: Semantic Image Retrieval Methodology

3.1. Object Recognition

Object recognition is two phase process, in first phase the presence of the objects are located called object localization, in second phase the located object is labelled with it identification called as Object Identification

Object recognition is basically classification process. In this process the Objects and their feature are learned in learning phase. In testing phase the target objects features are matched with the learned Object Features. For best match, the label of learned is assigned to target object.

Learning phase: In this Object detection extracts meaningful features of the objects that uniquely identify the objects. Mainly features are observer centric or object centric. Observer centric features consider low level of the object in each view of objects. Limitation of this approach is that this method requires learning each different view of the object.

Another way to represent the object is Object-cantered representation. In these approach primitives of a original object is considered as feature. Geometric primitives, collection of sub regions, multiview representation, bounding boxes are the basic primitives used in this approach. This approach faces the limitation of complexity for Model building.

3.2. Object Features

Global Feature: Global Feature are the characteristics of regions in images such as area (size), perimeter, Fourier descriptors, and moments. Global features can be obtained either for a region by considering all

points within a region, or only for those points on the boundary of a region. In each case, the intent is to find descriptors that are obtained by considering all points, their locations, intensity characteristics, and spatial relations.

Local Features: Local features are usually on the boundary of an object or represent a distinguishable small area of a region. Curvature and related properties are commonly used as local features. High curvature points are commonly called corners and play an important role in object recognition. Local features can contain a specific shape of a small boundary segment or a surface patch. Some commonly used local features are curvature, boundary segments, and corners

Bag of Features or Bag of word is another type of feature that may be used to identify the object. In this approach, interest points are detected using various interest point detection techniques as LoG, DoG, Harris Laplace, MSER etc. The patch surrounding the interest point is described using Local descriptors. The major property of the Local descriptor is Invariance to scale, illumination, rotation changes. SIFT, SURF descriptors are frequently used in the bag of feature approach to describe the local patches. Similar descriptors are clustered using the clustering algorithm and each cluster is represented by it mean descriptor. Collection of such mean Descriptors is known as Bag of Feature or Bag of Word. Descriptor performance affects with the Interest point detector. So selection of correct interest point detector is crucial in this case.

4. PROBLEM STATEMENT

Semantic Image Retrieval problem is the computer vision and machine learning problem. This problem is defined and stated using mathematical modelling as follows with Object Recognitions.

Mathematical Model for the Object Recognitions can be explained as follows

Hypothesis: Object in question is Present in the Image.

To prove this hypothesis Following is the Mathematical Model.

Let S be the system set consisting of solution for the problem

S = {S, I, X, O, DD, NDD, success, failure}

S = start state

E = end state {Image with Recognized Objects, Counts}

I = set of inputs = {Input Image}

X1 = Object Class Descriptors(Bag of words for each Object). {Bus, Dinosaur, Flower, ….}.... Wang Dataset.

Y1= Recognized Objects

X1• Y1

DD =deterministic data, All Objects in the Images are of deterministic

NDD = non deterministic data: Wrong Recognition error, power failure

success = All Objects are correctly recognized

failure = Object not detected Properly

5. TEST RESULT

Object Recognition phase is tested using BOW model. The Retrieved Images are ranked based on the Object Present in it. 100 Images of Wang Dataset is tested over for Object Recognition and Retrieval.

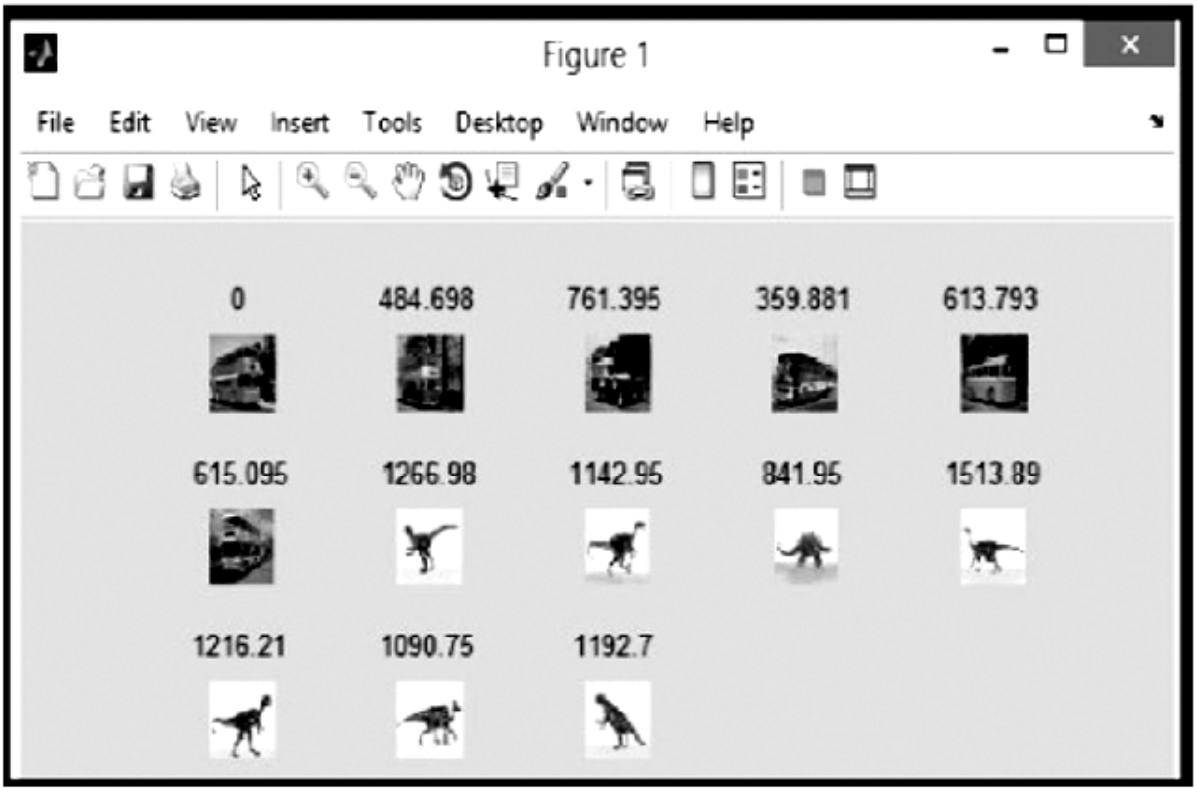

Figure 4: Test Result using Wang Dataset

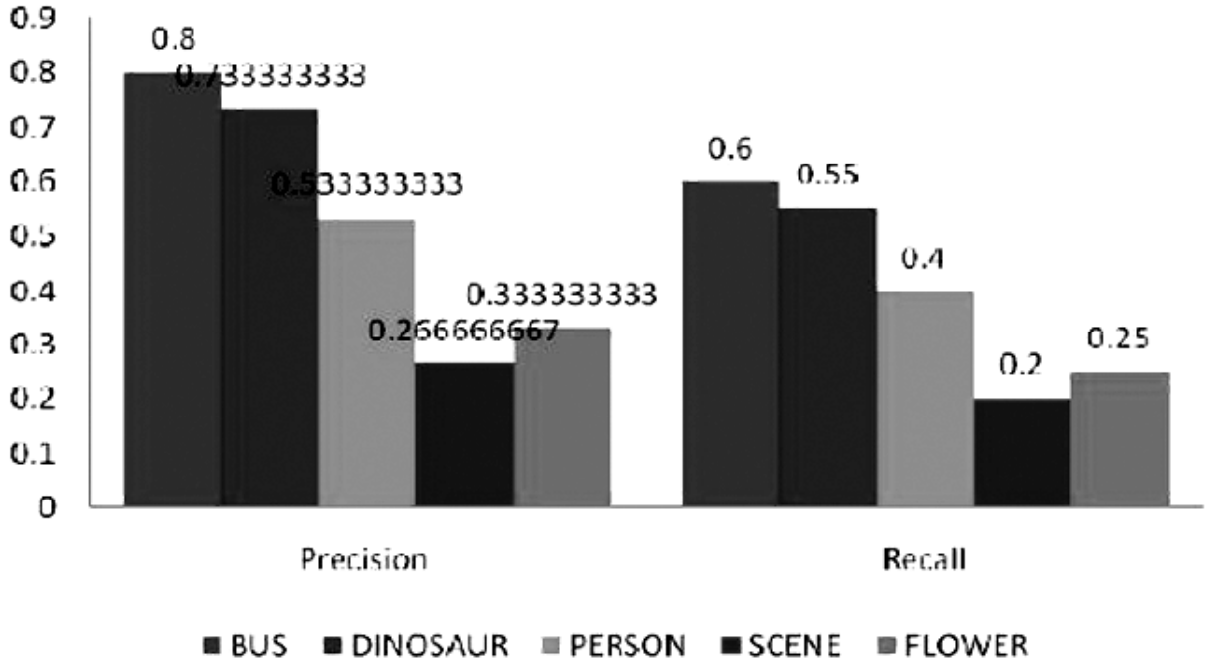

Figure 5: Test Result using Wang Dataset

Fig 4 showing that the top ranked images are very similar with the input image with BUS as object in it. Precision recall chart for the test are given in Fig 5.

6 CONCLUSION

Semantic Image retrieval is discussed in the paper is new approach for image retrieval. This approach is useful in many of the real time applications in medical databases, social network image databases etc. It is

robust to intensity, rotation, and scale in images and is able to retrieve similar meaning images. The test result are convincing with the the regular shaped objects. For Objects with the variations in colour and shape are more confusing and system need to be trained further with this variation to get good results for them.

REFERENCES

[1] HHerve Jegou, Matthijs Douze, and Cordelia Schmid. Improving bag-of-features for large scale image search. International Journal of Computer Vision, 87(3):316–336, 2010.

[2] Yang Cao, Changhu Wang, Zhiwei Li, Liqing Zhang, and Lei Zhang. Spatial-bag-of-features.In CVPR, pages 3352– 3359, 2010.

[3] Alper Yilmaz,Omar Javed, and Mubarak Shah, Object Tracking: A Survey, ACM Computing Surveys, Vol. 38, No. 4, Article 13, Publication date: December 2006.

[4] Yang, Y. And Ramanan, D. 2011. Articulated pose estimation with flexible mixtures-of-parts. In IEEE Conf. on Computer Vision and Pattern Recog. (CVPR).

[5] Hui Hui Wang, Dzulkifli Mohamad, and N. A. Ismail, Semantic Gap in CBIR: Automatic Objects Spatial Relationships Semantic Extraction and Representation. International Journal Of Image Processing (IJIP), Volume (4): Issue (3).

[6] Jonathon S. Hare1, Patrick A.S. Sinclair1, Paul H. Lewis et el, Bridging the Semantic Gap in Multimedia Information Retrieval.