Report

The aim of the project was to build a search engine and object classifier using the Bag of Words Model (BoW) of SIFT/SURF features and a SVM classifier using GIST features respectively.

Data Set:

The data set used Caltech dataset which had 11 categories like AK47, compact discs, etc. The number of images in each category were around 100.

Method 1:

The Bag of Words Model is implemented in three steps using OpenCV API. The first step comprises of building a vocablary of visual words which are extracted using k-means clustering of SIFT/SURF features extracted from the image. Thirty images were taken from each category. The vocabulary size was taken as 7500 which gave better results compared to 5000 visual words. This step is time consuming for it is computationally expensive. The second step comprises of building document vectors of all images as well as building an inverted index. The final step is the querying process in which an input image is given and the top 8 images are displayed using cosine distance. The query process was very fast and the images are displayed almost instantaneously.

The results of the querying were satisfying. Objects like AK47 and American flag are being displayed with maximum correct answers.

Method 2:

The classification method was done using GIST features (which are glocal features unlike the local features like SIFT, SURF, etc which are extracted for the entire image, so there will only be one descriptor for the entire image) and SVM classifier in MATLAB.

The training data set when tested, the classification results were almost perfect. The testing data set results however were not that satisfactory.