

Assignment 3

Submitted By: 2019PCS0003 - Alokedu Mazumder

Platform used: Google Colab

Library used: numpy, sklearn, openCV

Approach:

Cooking the Codebook

- 1) Divide the input image into a set of sub-images.
- 2) Apply sift on each sub-image. Each key point will have 128 dimensional feature vectors.
- 3) Encode these vectors to create a codebook by simply applying k-means clustering with a chosen k . Each image will produce a matrix \mathbf{V}_i ($i \leq n$ and n is the number of images used to create the codeword.) of size $128 \times m$ where m is the number of key points gathered from the image. The input to K-means is therefore, a big matrix \mathbf{V} created by horizontal concatenation of \mathbf{V}_i , for all i . The output of K-means is a matrix \mathbf{C} with size $128 \times k$.

Computing the histograms

For each image in the dataset, do the following:

- 1) Create a histogram vector h of size k and initialize it to zeros.
- 2) Apply dense SIFT as in step 2 in stage 1.
- 3) For each key point's vector find the index of its "best match" vector in the codebook matrix \mathbf{C} (can be the minimum in the Euclidian distance) .

Now, simply I have feeded the feature vectors and labels to a SVM classifier with RBF kernel.

Note: Number of clusters is recommended as $3 \times k$. More the number, the better the results.

Current result: 63% efficiency. Checked last time, but wasn't able to run due to hardware limitations as the model takes a very very huge amount of time to train in my laptop and mostly falls out due to memory overflow.

