COMPUTER VISION AND IMAGE PROCESSING PROJECT #3

1. IMAGE BASED SEARCH ENGINE:

A dataset consisting of pictures from various landmarks at UB Campus like Lake la Salle, Capen Library, UB stampede, Davis Hall and the Centre for the arts were collected. The query images from these landmarks were also taken and stored separately.

The images in the dataset has to be indexed which is a process of quantifying using an image descriptor to extract features from each image and storing the resulting features for later use, such as performing search.

Goal: Given a query image from one of the categories, search engine has to return the category's corresponding images in the top 5 results.

Description:
The image descriptor used is a 3D color histogram in the RGB color space with 8 bins per red, green, and blue channel. The next step is to apply image descriptor to each image in the dataset. This simply means that we are going to loop over the images, extract a 3D RGB histogram from each image, store the features in a dictionary and write the dictionary to file. The key for the dictionary will be the name of the image file assuming that all filenames are unique. The value for the dictionary will be the computed histogram for the image.

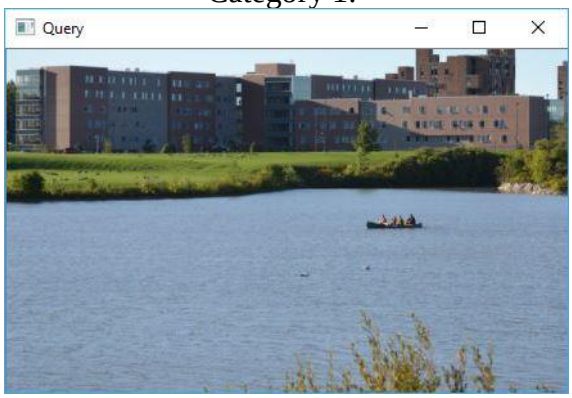
Next is searching. To do the search, we loop over the image filenames and corresponding features in the index then use the chi-squared distance to compare the color histograms. The computed distance is then stored in the results dictionary, indicating how similar the two images are to each other. The results are sorted in terms of relevancy (the smaller the chi-squared distance, the relevant) and returned.

RESULTS:

For each query image, the top 5 results are shown below:

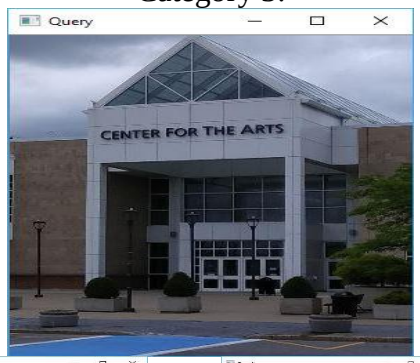
Shri Sai Sadhana Natarajan (50247664)

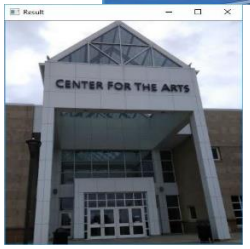
Category 1:

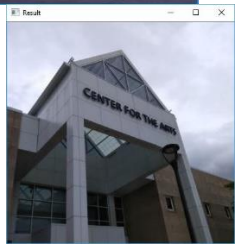


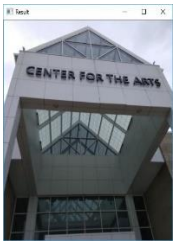
Category 2 :

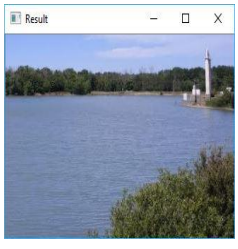
Category 3:



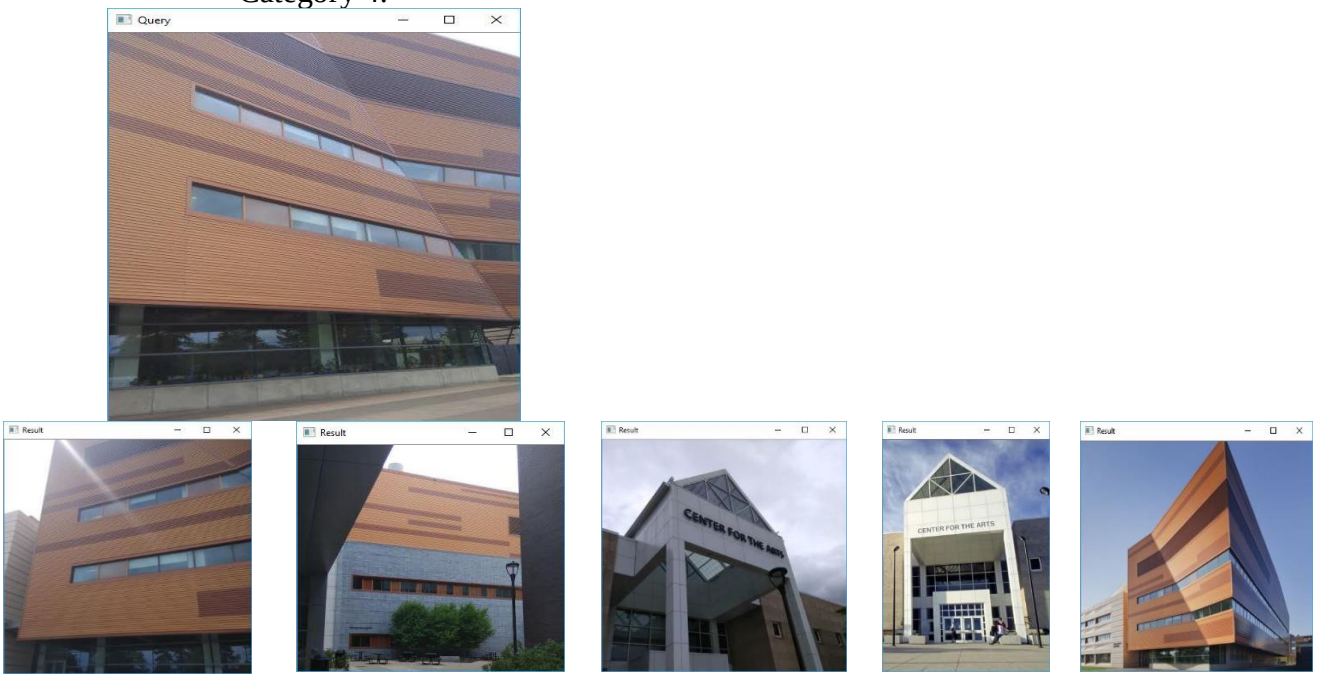




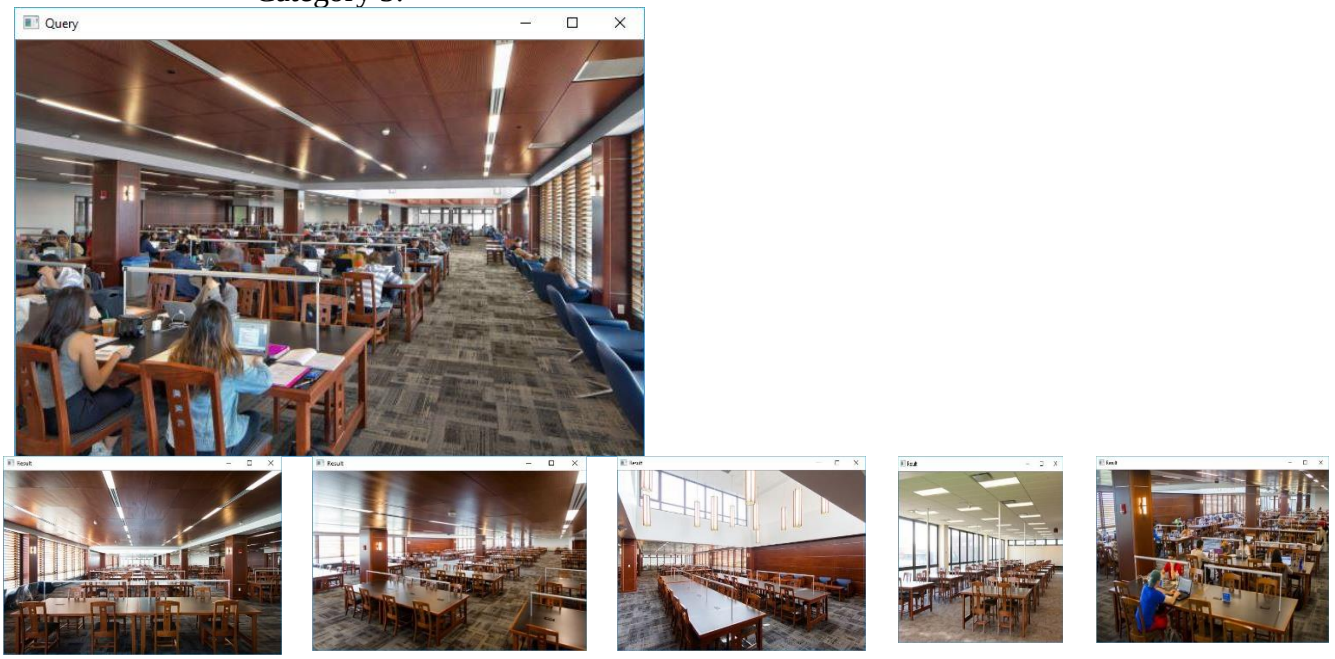




Category 4:



Category 5:



Category 1 : Lake – 4/5 images are relevant
Category 2 : Stampede bus – 5/5 images are relevant
Category 3 : Center for arts – 4/5 images are relevant
Category 4 : Davis Hall – 3/5 images are relevant
Category 5 : Library – 5/5 images are relevant

Therefore the total accuracy of the search engine is 84%

Conclusion:

The image descriptor chosen is a 3D color histogram. But we can take into account the shape, texture of the images and space relationship of the objects for better performance. Combining shape and color both in invariant fashion is a powerful combination as described by where the colors inside and outside affine curvature maximums in color edges are stored to identify objects. The identification of specific textures in an image can be achieved by modelling texture as a two-dimensional gray level variation. Co-occurrence matrix can also be used to classify textures. Apart from these, approaches like bag of words representation where we quantize local descriptors into visual words and then apply scalable textual indexing and retrieval schemes or we can use SIFT algorithm which is invariant to image scaling, transformation and rotation. It is also partially invariant to illumination changes and can present the local features of an image. Therefore, feature keypoints can be extracted more accurately by using SIFT than color, texture, shape and spatial relations feature.

The hands-on assignment is highly helpful for computer vision beginners. The steps are very well explained and provide an easy understanding. The comment section delivers better clarifications required for the assignment.

References: -

- https://staff.science.uva.nl/th.gevers/pub/overview.pdf
- https://www.pyimagesearch.com/2014/01/27/hobbits-and-histograms-a-how-to-guide-to-building-your-first-image-search-engine-in-python/