Assignment 2

(Report)

- Amitha
- IMT2019023

Part 1:

Image Stitching :

Differences between SIFT and SURF:

- Scale-Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF) are two robust feature extraction descriptor algorithms.
- SURF is better than SIFT in rotation invariant, blur and warp transform. SURF also produces better results when images are blurry.
- SIFT is better than SURF when the images have different scales.
- SURF is three times faster than SIFT.
- A SIFT feature is a selected image region (also called keypoint) with an associated descriptor
- SIFT computes an image pyramid by convolving the image several times with large Gaussian kernels, while SURF accomplishes an approximation of that using integral images.

FLANN matching and RANSAC:

FLANN stands for Fast Library for Approximate Nearest Neighbors. It contains a collection of algorithms optimized for fast nearest neighbour search in large datasets and for high dimensional features. It works faster than BFMatcher for large datasets. For FlannBasedMatcher, it accepts two sets of options that specify the algorithm to be used, its related parameters etc. FLANN will be much faster but will find an approximate nearest neighbour. It will find a good matching, but not necessarily the best possible one. You can play with FLANN's parameters in order to increase the precision, but it will be at the cost of slowing the algorithm.

Random sample consensus (RANSAC) is an iterative method to estimate parameters of a mathematical model from a set of observed data that contains

outliers when outliers are to be accorded no influence on the values of the estimates. Therefore, it also can be interpreted as an outlier detection method. It is a non-deterministic algorithm in the sense that it produces a reasonable result only with a certain probability, with this probability increasing as more iterations are allowed.RANSAC uses repeated random sub-sampling. A basic assumption is that the data consists of "inliers", i.e., data whose distribution can be explained by some set of model parameters, though may be subject to noise, and "outliers" which are data that do not fit the model.

Images for which Image Stitching is applied:

Image after Stitching:

Part 2:

Implement Bike vs Horse Classification:

Using Bag-of-visual words approach (SIFT/SURF + K-means + SVM/Logistic Regression/KNN).

BOVW Approach :

BOVW is used for image classification. Bag of visual words (BOVW) uses a similar concept of Bag Of Words(BOW) in NLP, but instead of words, image features are used as "words."

Image features consist of keypoints and descriptors which are identified by feature extraction descriptor algorithms such as SIFT and SURF We use the keypoints and descriptors to construct vocabularies using k-means algorithm and represent each image as a frequency histogram of features that are in the image. From the frequency histogram, later, we can find other similar images or predict the category of the image.

Procedure:

- Detecting features and extracting descriptors in an image is done by using feature extractor algorithms such as SIFT and SURF.
- Clusters are made from descriptors using the k-means algorithm. The centre of each cluster is used as a visual dictionary's vocabulary.
- Finally, we create a frequency histogram from the vocabulary in each image, as well as the frequency of the vocabularies in the image. Those histograms are our bag of visual words(BOVW).
- Next, to classify other new images we train a model with ML algorithms such as SVM, KNN, Logistic Regression.

Experiments and Results:

- The model used: BOVW
- SIFT for feature extraction.
- Classification Algorithm: SVM (Linear kernal)

'K' in K-means	Accuracy score
50	0.8888888888888888
60	0.9722222222222222
70	0.9722222222222222
80	0.9444444444444444

- The model used: BOVW
- SIFT for feature extraction.
- Classification Algorithm: KNN(K Neighbours Classifier) (n_neighbors=10)

'K' in K-means	Accuracy score
50	0.9166666666666666
60	0.9166666666666666

70	0.9444444444444444
80	0.9166666666666666

- The model used: BOVW
- SIFT for feature extraction.
- Classification Algorithm: Logistic Regression(max_iter = 5000)

‘K’ in K-means	Accuracy score
50	0.9444444444444444
60	0.9722222222222222
70	0.9722222222222222
80	0.9444444444444444

Conclusion :

Accuracy Score is best for SVM (Linear kernal) with k=60, 70 in k-means clustering and for large values of k such as above 450 the accuracy score is 1. SURF is not working with the OpenCV modules.

CIFAR 10, with 10 classes:

The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class.

I used the BOVW approach, SIFT for feature extraction, K-means for clustering and SVM (Linear Kernal) for classification.

The best accuracy so far is 0.1119 for k = 80 in k-means clustering and SVM (Linear kernal) for classification.