Computer Vision

ASSIGNMENT 3

Chinmay Dhawan
I61711015 | C.DHAWAN@STUDENT.MAASTRICHTUNIVERSITY.NL

TABLE OF CONTENTS

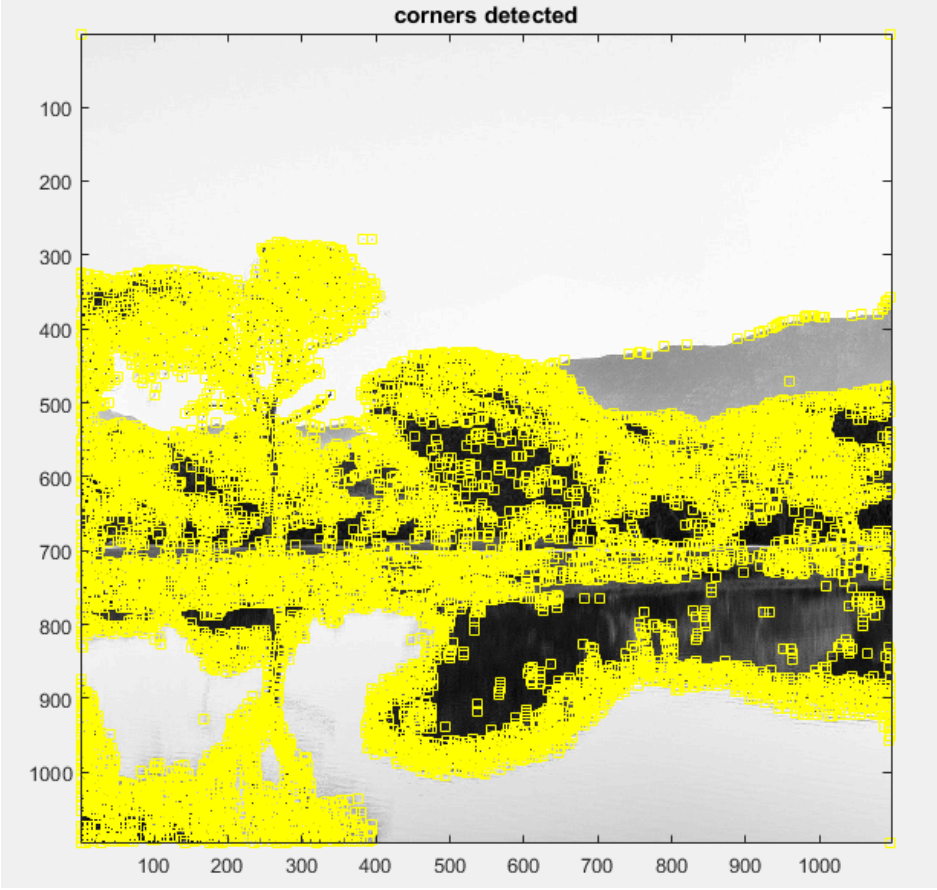
PART 1: PANAROMA

KEY POINTS EXTRACTION/DESCRIPTION AND VARIOUS PARAMETERS:

Key points are extracted using two approaches i.e., Harris and SIFT to make additional comparisons. Harris key points are extracted using different parameter values of standard deviation and radius. Both parameters greatly influence the number of type and especially number of key points extracted from the images.

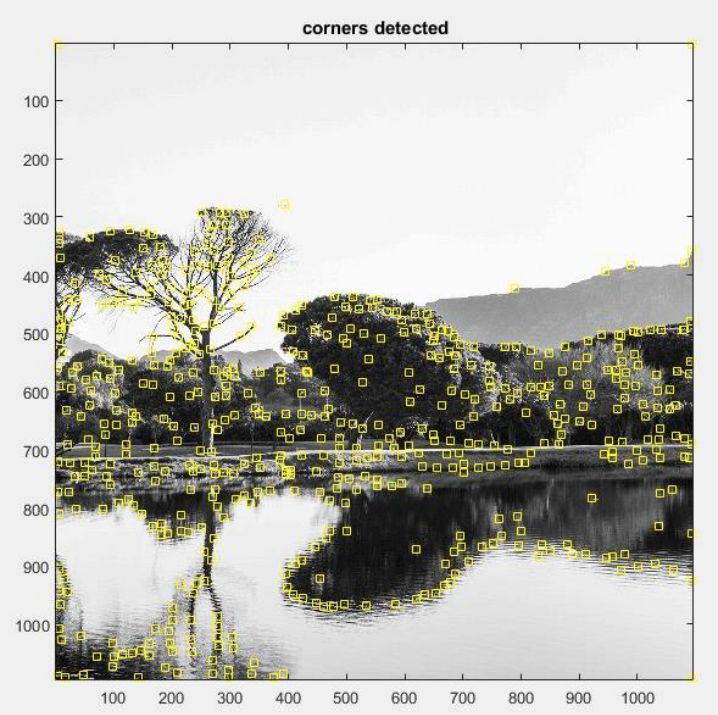
The key points detected using Harris corner detection algorithm are described using SIFT descriptors with each key point represented using 128 elements describing 16 histograms for a particular frame/keypoint.

The SIFT key-points are also extracted as mentioned before using default vl_sift parameters and the described in the same way.



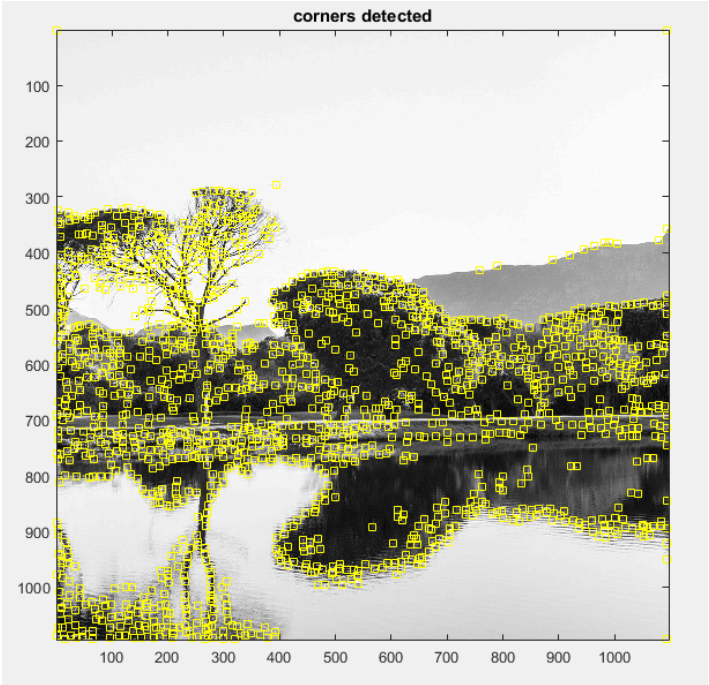
Features extracted using standard deviation of 1, window of radius 1 and threshold of R as 1000.

As the standard deviation increases, the features become more define and less in number as shown below.

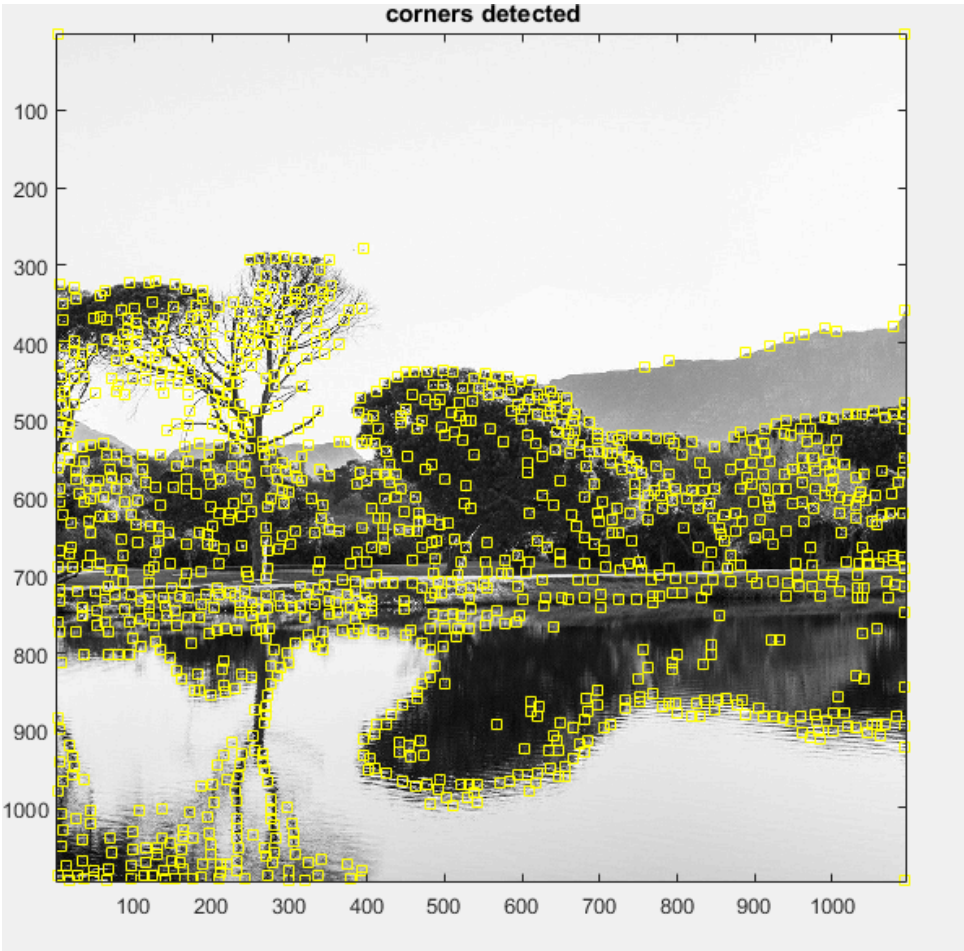


Features extracted using standard deviation of 5, window of radius 1 and threshold of R as 1000.

The same can be said for window size, the radius or the neighborhood for which the gradients are analysed for corner detection algorithm.



2421 features using standard deviation 3 and radius = 1



2421 features using standard deviation 3 and radius = 3

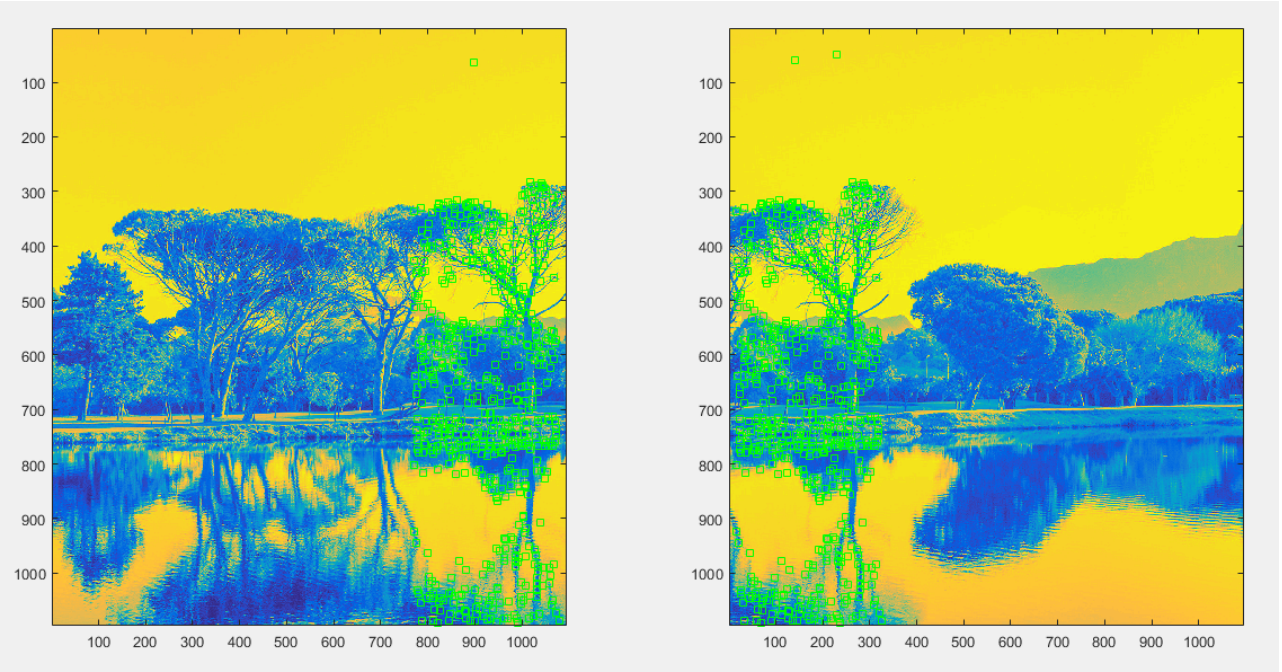
Overall, it is observed that increasing the radius and standard deviation results in less number of features, but they are more defined, but of course increasing these parameters beyond a certain limit will have a negative effect resulting in too few features.

KEY-POINT MATCHING:

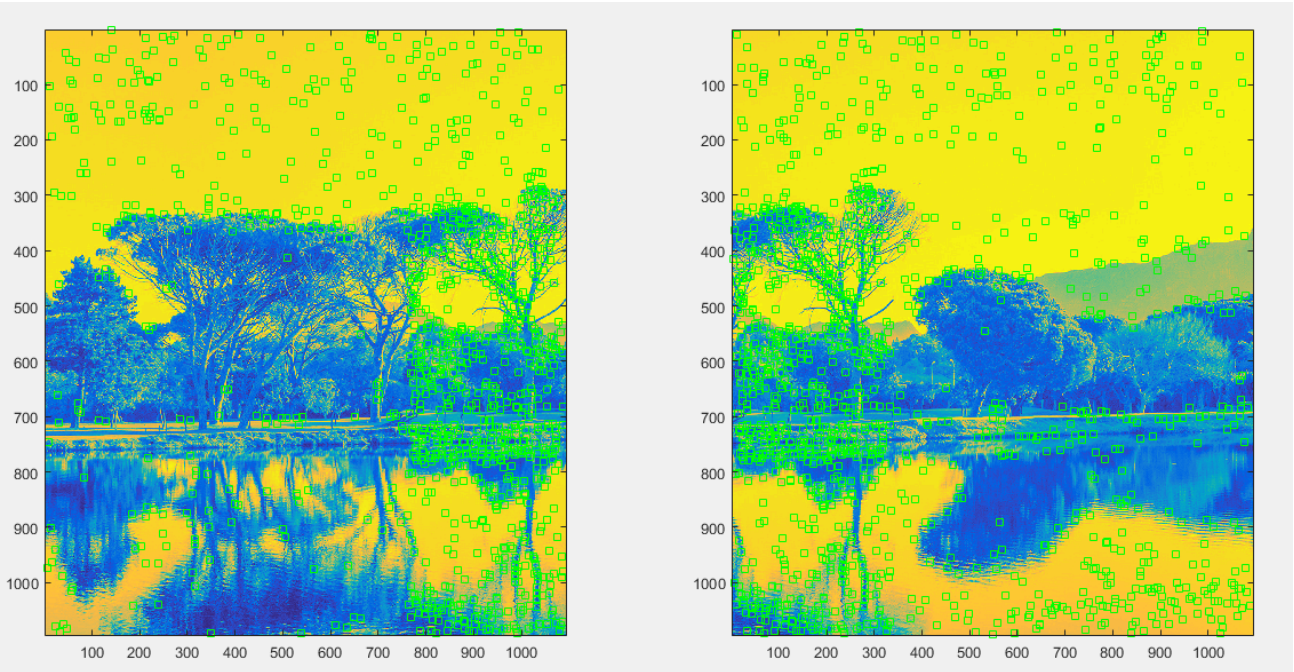
Key-points are matched using the descriptors given by SIFT algorithm. They are first transformed/normalized to their equivalent z-scores, which brings the mean of whole distribution to zero and each point is represented by how much standard deviation does it lie from the mean.

The Euclidian distances between the transformed descriptors is then calculated and all the points below a certain threshold are discarded. The threshold is determined by looking at the data and analyzing centroid points like mean, median, modes, max and min distances among the points. So, it will vary for different images based on their similarity and key point matchings.

A threshold of 1 is used for the current schema, which provides the key points which are almost similar to each other. This threshold proved to work best for SIFT key points but for Harris key points, a much higher threshold was needed.



Corresponding SIFT Key-points that match with each other after applying threshold = 1



Corresponding SIFT Key-points that match with each other after applying threshold = 6

MODEL FITTING:

An affine transformation is assumed and Random Sampling Consensus(RANSAC) is used to find the case best model parameters. It works be finding the model which agrees with maximum number of inliers which makes sense in our case as one would like to make a transformation based on majority of agreeing data points instead of an average transformation which can be found by running least squares on all the data set.

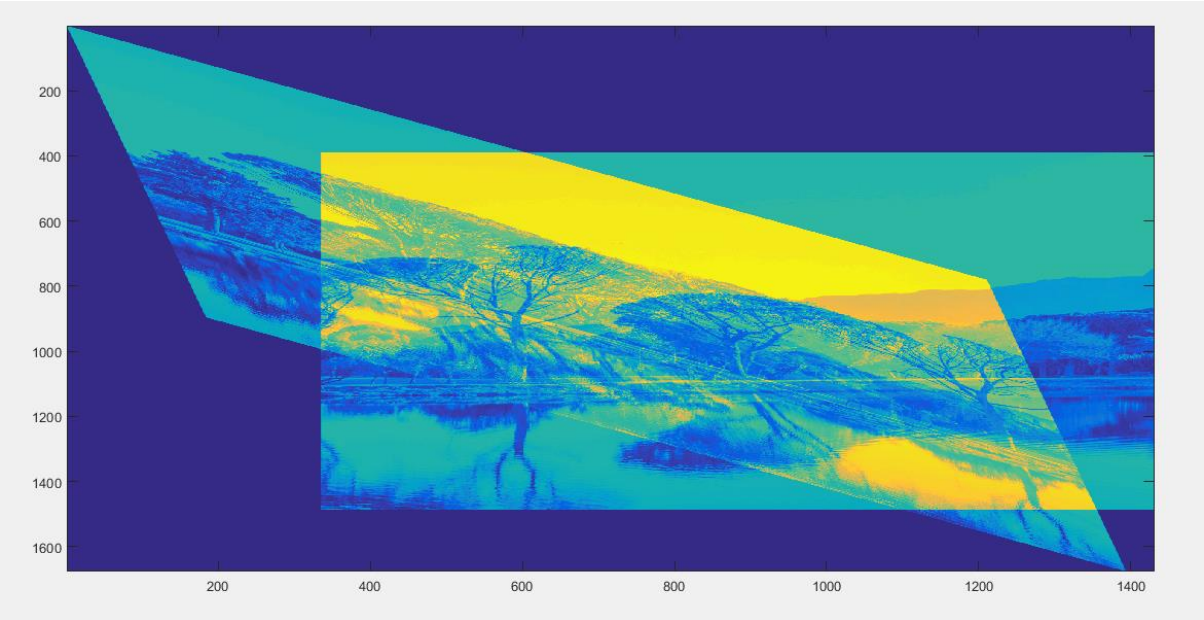
A sample size of 3 is chosen as 3 pairs of points (6 overall) would be the minimum number of points required to find the parameters of affine transformation (DOF = 6). The process is repeated again and again, and the transformation which agrees with most of the points is stored.

The agreement is calculated by computing squared distance between the point obtained by applying transformation on a point from image 1, and the corresponding matching point in image 2. If the difference is less than a threshold (0.9 for the current implementation), then the corresponding points are considered as inliers.

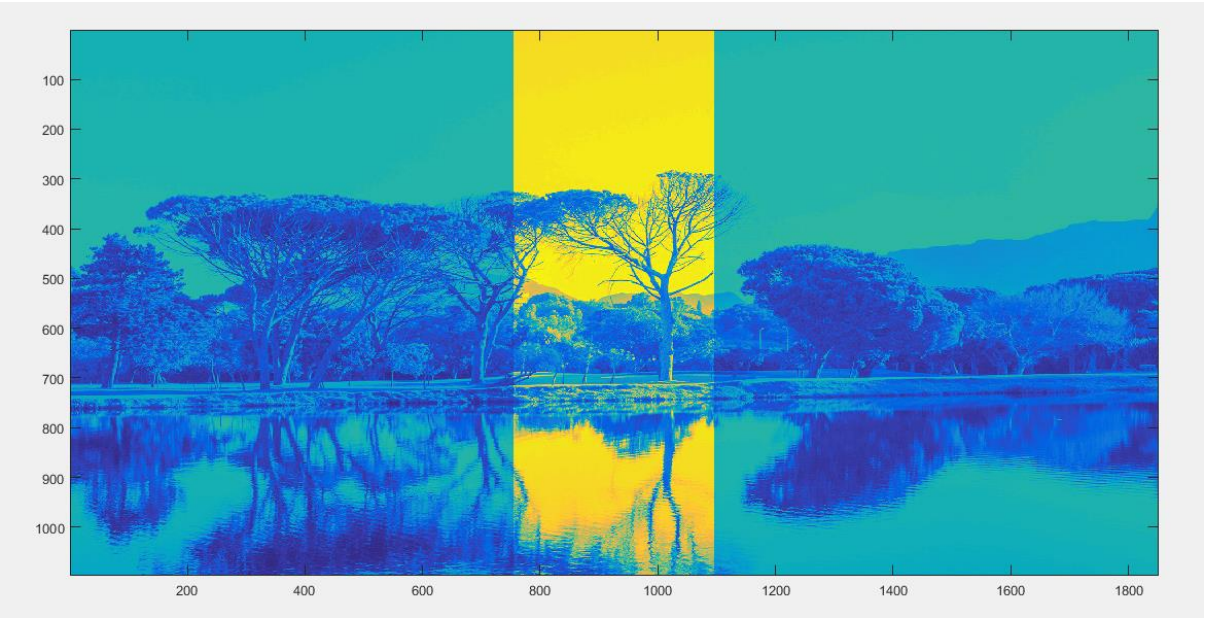
Finally, the average transformation is computed on all the inliers by using least squares method.

MORPHING THE IMAGES:

At the end image warping is done by aligning matching points of image 1 and image 2 on top of each other and the final panoramic result is obtained.



Images combined (Harris key points)



Images combined (SIFT key points)

Images combined (Harris key points)

As you can observe from above, SIFT key points proved to work much more effectively than Harris.

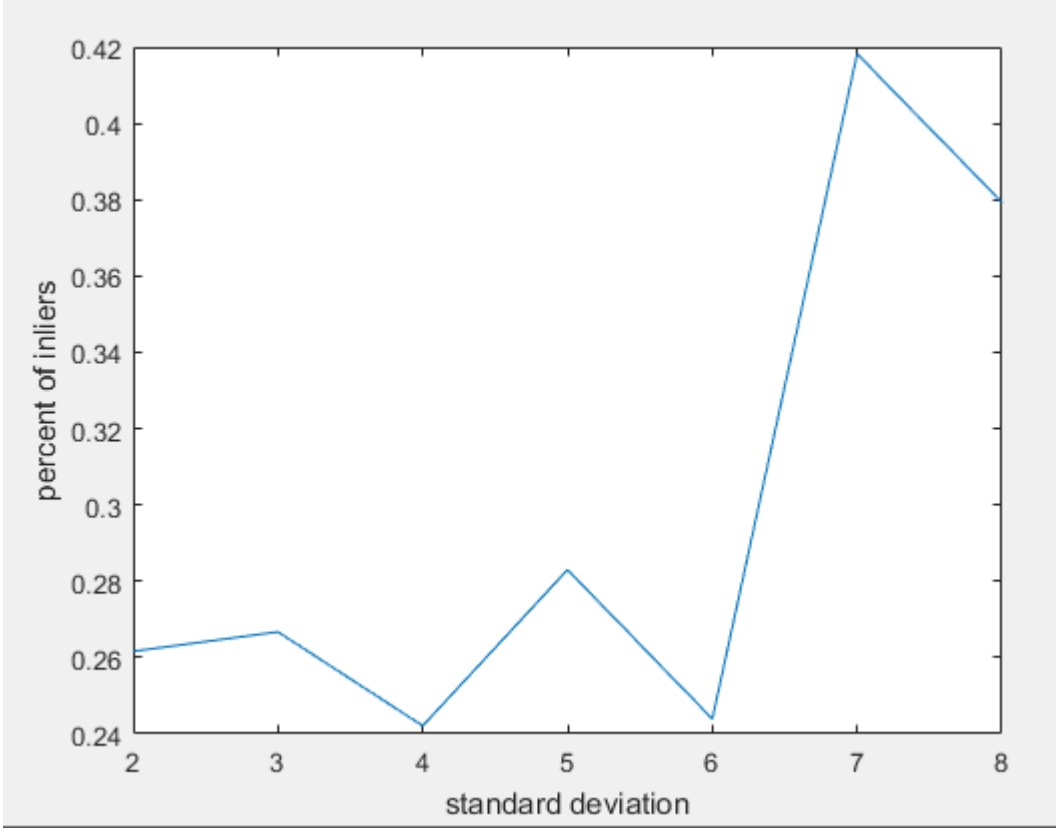
INLIERS/OUTLIERS ANALYSIS:

Harris:

The percent of inliers obtained after RANSAC are plotted against various parameters of Harris detectors. According to below figures a standard deviation of 7 resulted in greatest number of inlier percentage but this can't be directly linked with increased performance.

It also depends on the image being used as the the 'ideal' standard deviation for good features will depend on how gradients vary in the image.

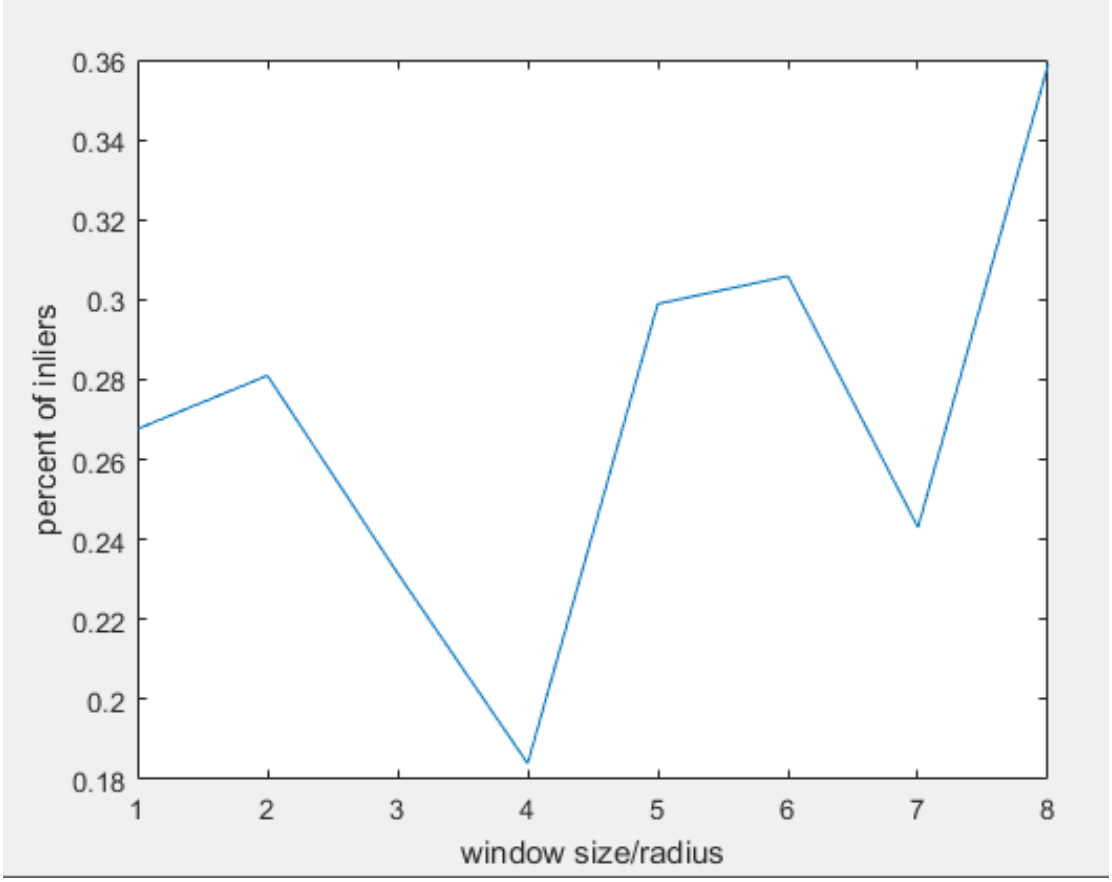
Window size is fixed at 2.



Inliers vs std

The below table shows number of inliers in row 1 and total number of features in row 2 as standard deviation varies from 2 to 8.

1215	291	92	45	20	23	11
4644	1091	380	159	82	55	29

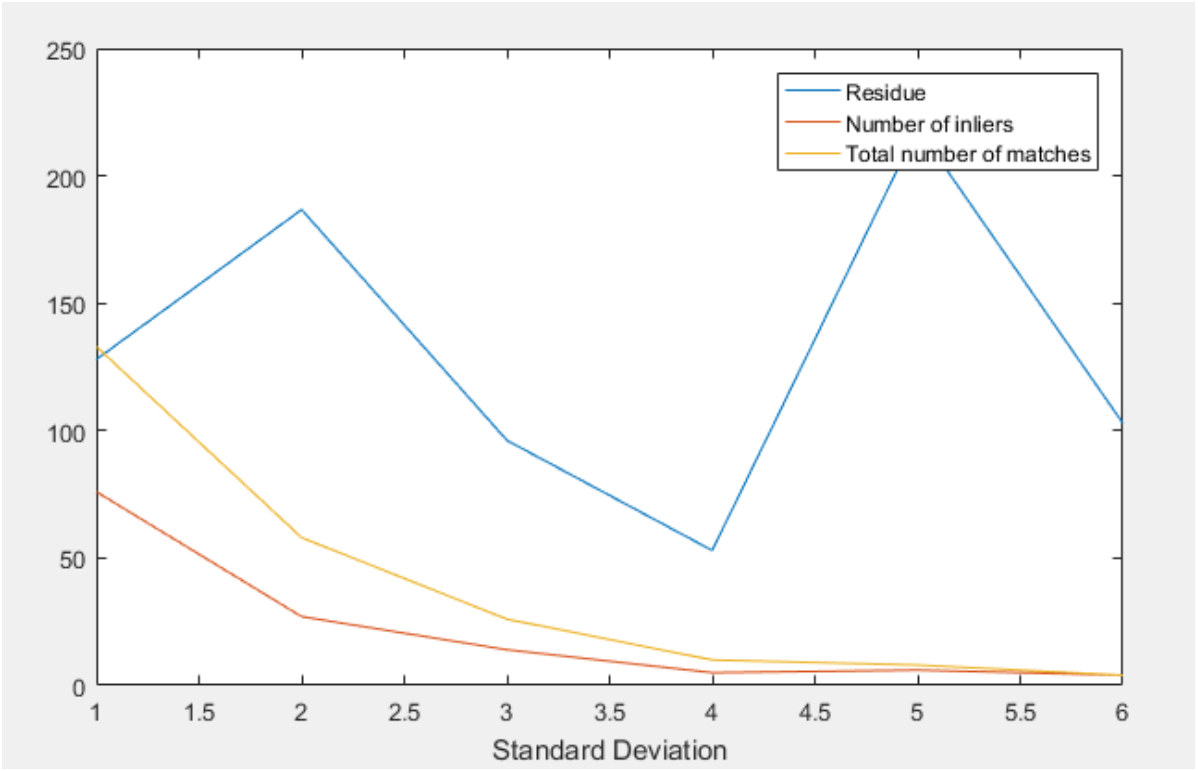


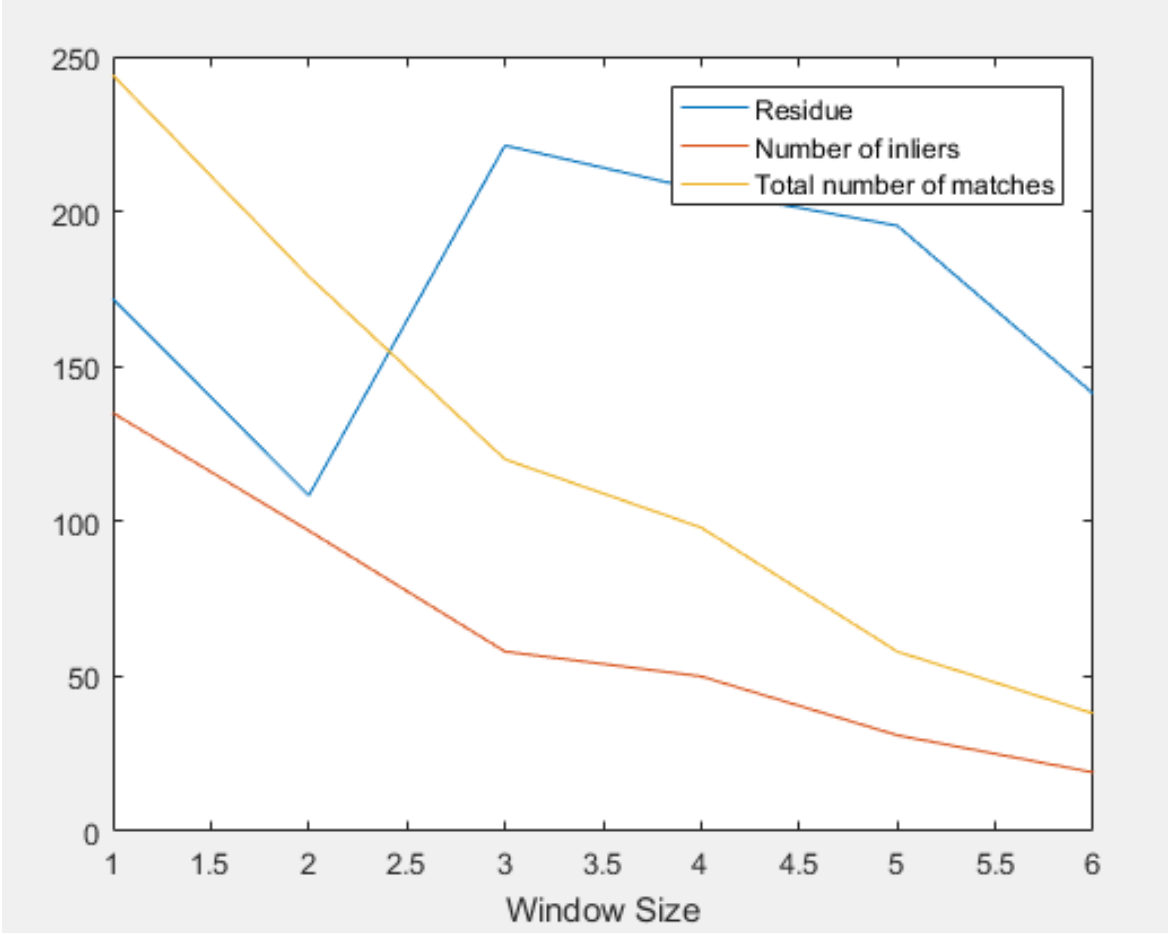
Inliers vs std

The below table shows number of inliers in row 1 and total number of features in row 2 as standard window size varies from 1 to 8.

1616	1305	777	388	423	301	157	160
6038	4644	3355	2109	1415	984	646	446

For a different image, the analysis is shown as below for various standard deviations and window sizes. Here, threshold for matching corresponding pairs is used as 4 and threshold the points being counted as inlier is 30.

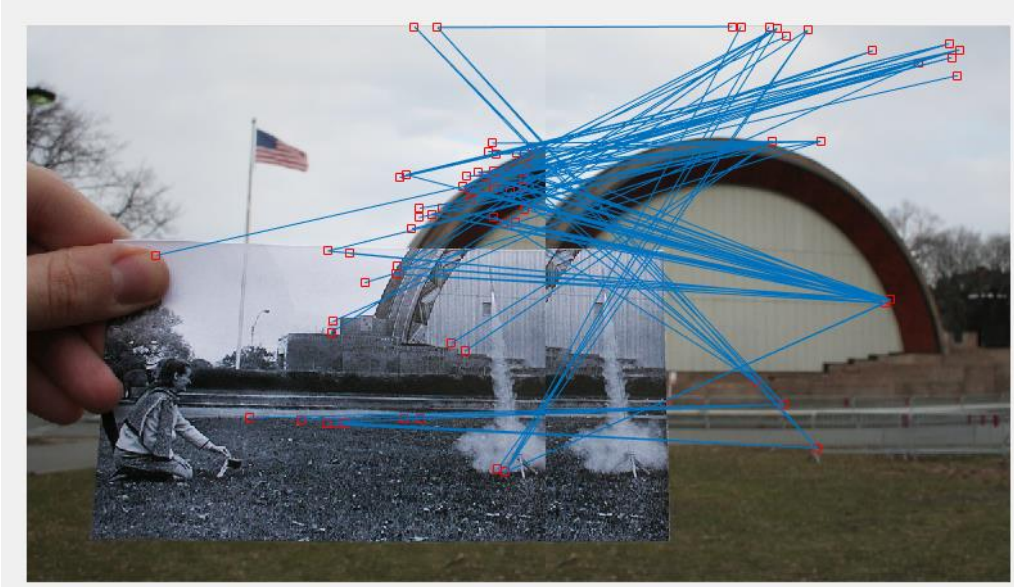

Residue
Number of inliers
Total number of matches
250
200
150
100
50
0
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
Standard Deviation


Residue
Number of inliers
Total number of matches
250
200
150
100
50
0
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
Window Size

Problem of Combining Harris with SIFT Descriptors:

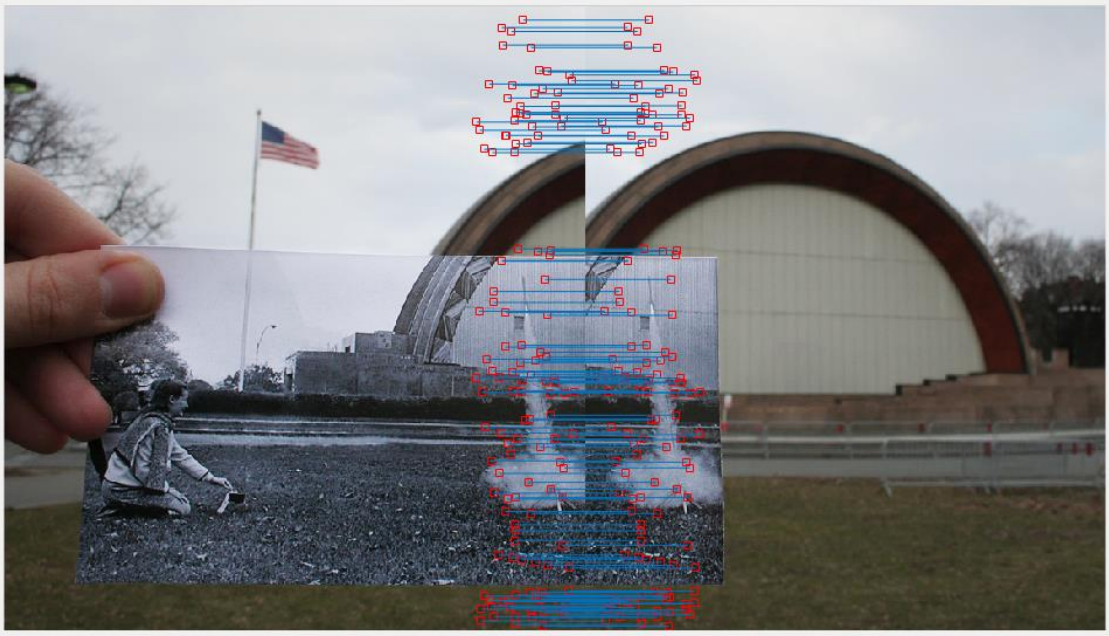
The harris keypoints detected may match very well between the two images using SIFT descriptors, but for complicated images, the parameters need to be tuned manually by analyzing the data. For example, the value of scale of SIFT may have a relationship with the patch size used for Harris. For badly tuned parameters, the features may not match at all. For example, for a scale of 5 and window size 4 using standard deviation 2 and threshold as 30, following matches are obtained:

Matching using Harris key-points



Matching using Harris key-points

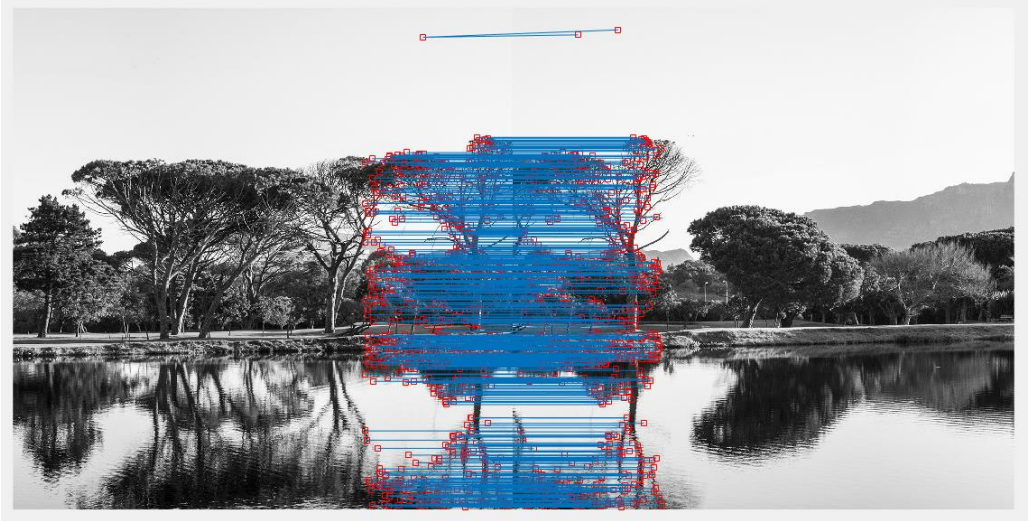
While SIFT features are able to match very well as shown below:



SIFT:

For SIFT key points, the results achieved were far better with almost 97% of pairs being detected as inliers. The matching between inliers is shown below along with the total matches and number of inliers.

For the below particular image 726 points were detected as inliers out of total 730 points which were matched between the two images. Of course, a good value of thresholds are required which can be detected by looking at the data.



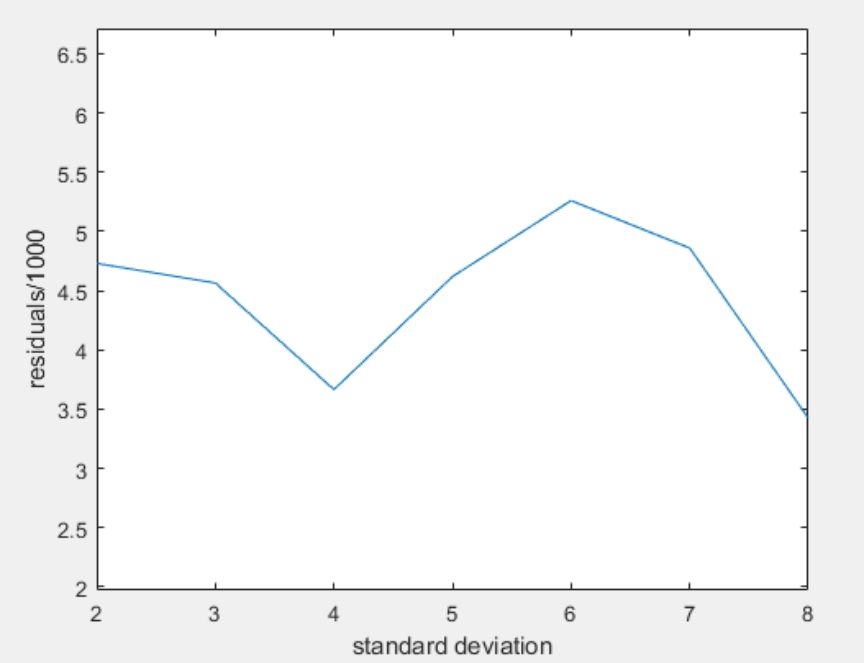
SENSITIVITY ANALYSIS 1:

A sensitivity analysis on point matching is done by varying the sigma and window size parameters for Harris detector. The threshold for deciding the matched pairs and the threshold used for counting a pair of points as inlier is also kept fixed. A plot of residuals against standard deviation is plotted while keeping all other parameters fixed.

Residuals are calculated by summing the squared distances between transformed points from image 1 to image 2 and corresponding actual points in image 2. Residuals are only calculated for the inliers obtained.

The below table shows residuals as standard deviation varies from 2 to 8.

4.7296e+03	4.5650e+03	3.6648e+03	4.6186e+03	5.2597e+03	4.8595e+03	3.4238e+03

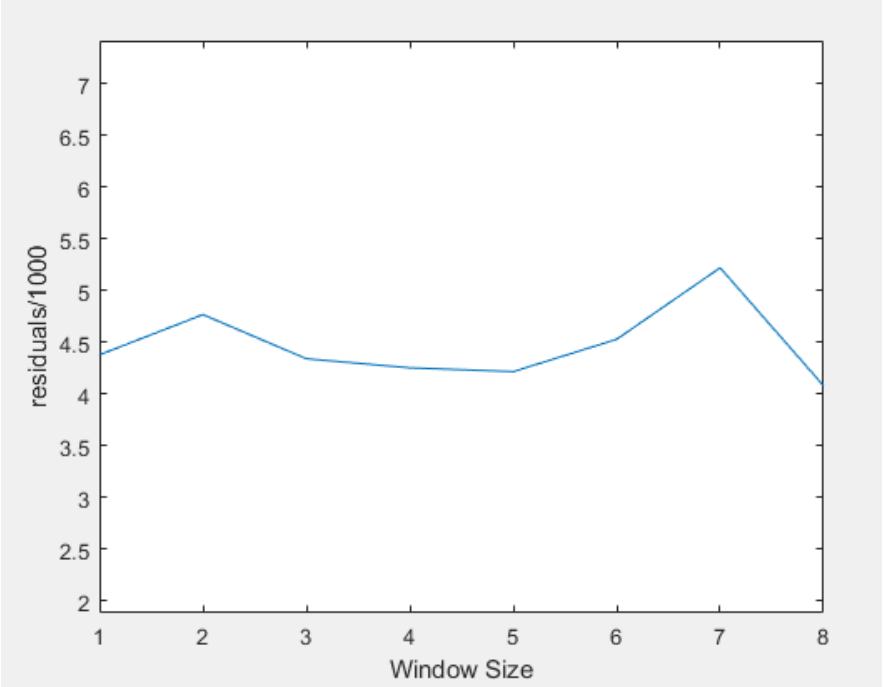


Residuals vs. Std

A can be seen in figure above, a standard deviation of 4 and 8 results in low residue

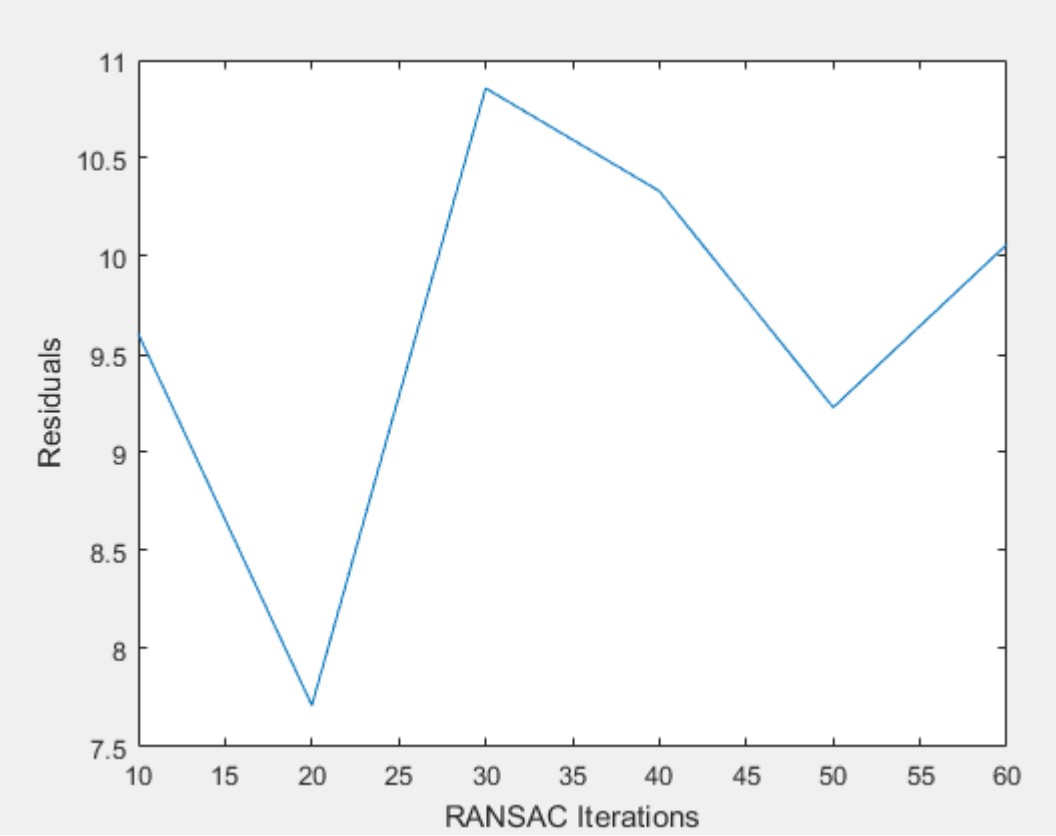
The below table shows residuals as window size varies from 1 to 8.

4.3749e+03	4.7639e+03	4.3369e+03	4.2493e+03	4.2127e+03	4.5251e+03	5.2167e+03	4.0750e+03



Residuals Vs. Window size

A small window size of 3 proved to work best



Residuals vs RANSAC iterations

Increasing the number of iterations should result in low residue, but since the algorithm is random in nature, hence the above deviations in plot can be explained.

The below table shows residuals as window size varies from 1 to 8.

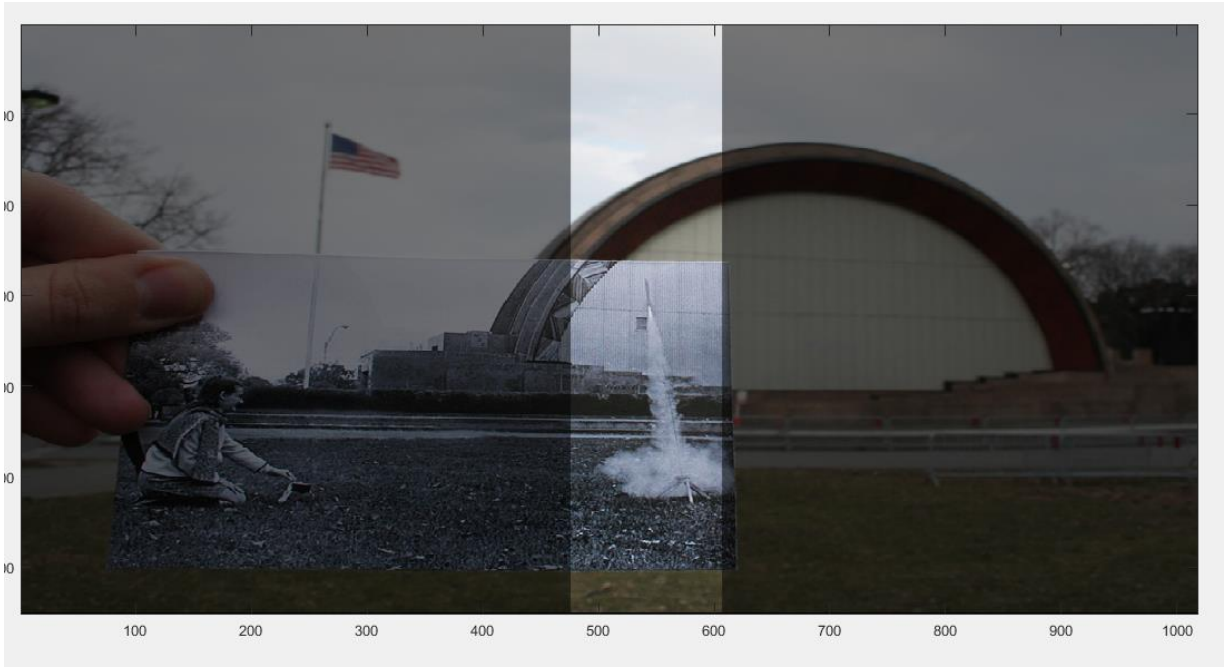
9.6066e+03	7.7118e+03	1.0857e+04	1.0332e+04	9.2297e+03	1.0061e+04

BONUS IMAGES:

The following images are combined using 'SIFT' parameter of the function.


100
200
300
400
500
600
700
800
100
200
300
400
500
600
700
800
900
1000
1100


100
200
300
400
500
600
700
800
100
200
300
400
500
600
700
800
900
1000
1100


100
200
300
400
500
600
700
800
900
1000