DATA.ML.300 Computer Vision

Exercise Round 4

Answered to all questions (1-4)

1)

a)

Points in image space:

$$(x_0, y_0) = (-2, 1)$$
$$(x_1, y_1) = (2, 5)$$

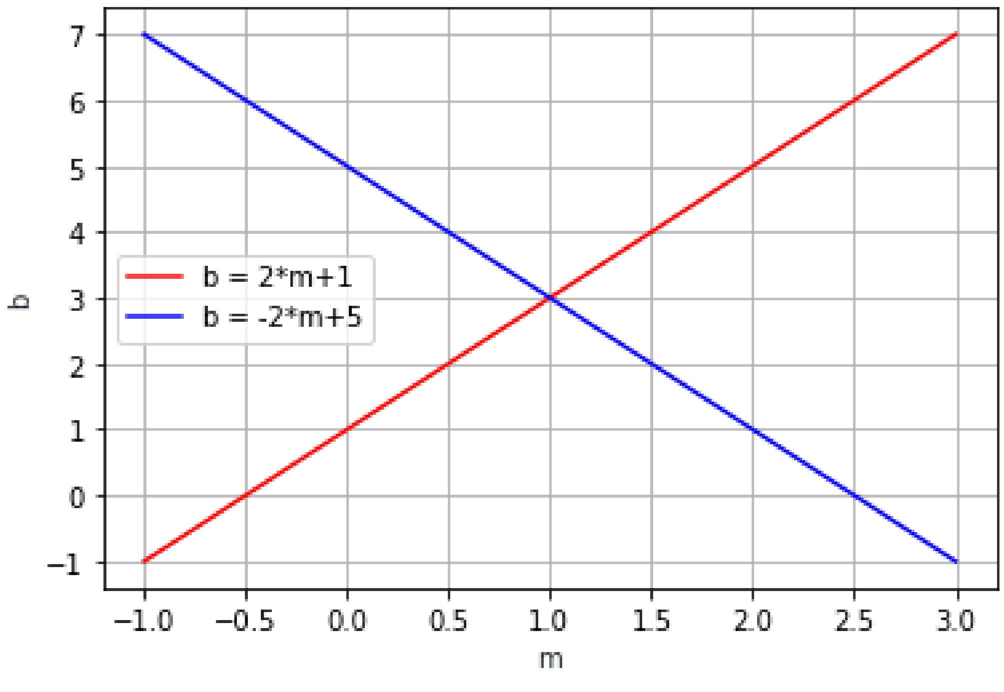
Corresponding lines in Cartesian parameter space:

$$b = 2m + 1$$
$$b = -2m + 5$$

Point of intersection of above lines in Cartesian parameter space:

$$2m' + 1 = -2m' + 5$$
$$4m' = 4 \quad , \quad b' = 2 * 1 + 1 = 3, \quad \therefore (m', b') = (1, 3)$$
$$m' = 1$$

Plotting:



b)

Points in image space:

$$(x_0, y_0) = (-2, 1)$$
$$(x_1, y_1) = (2, 5)$$

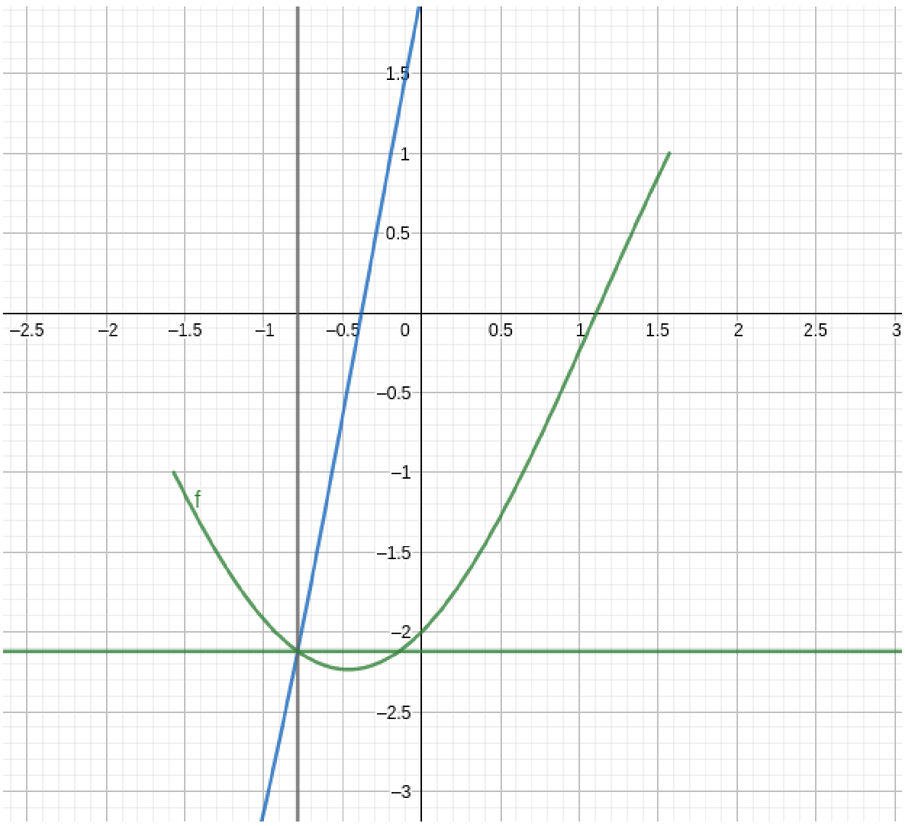
Sinusoids in Polar coordinate parameter space:

$$\rho = -2cos\theta + sin\theta$$
$$\rho = 2cos\theta + 5sin\theta$$

Point of intersection of above sinusoids in Polar coordinate parameter space:

$$\begin{aligned} -2cos\theta' + sin\theta' &= 2cos\theta' + 5sin\theta' \\ -4cos\theta' - 4sin\theta' &= 0 \\ cos\theta' + sin\theta' &= 0 \\ \theta' &= -\frac{\pi}{4} \end{aligned} \quad , \quad \begin{aligned} \rho' &= 2cos\theta' + 5sin\theta' \\ &= 2x - 5x \\ &= -3x \\ &= -3cos\theta' \\ &= -3cos\frac{\pi}{4} \end{aligned} \quad , \therefore (\theta', \rho') = (-\frac{\pi}{4}, -3cos\frac{\pi}{4})$$

Plotting ($x = \theta$, $y = \rho$):

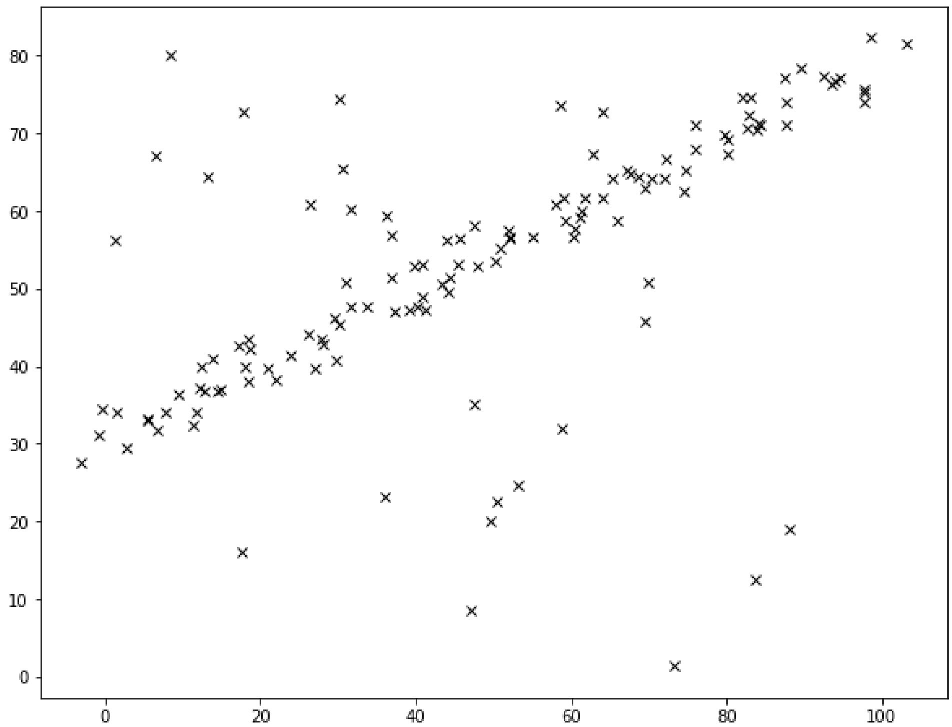


c)

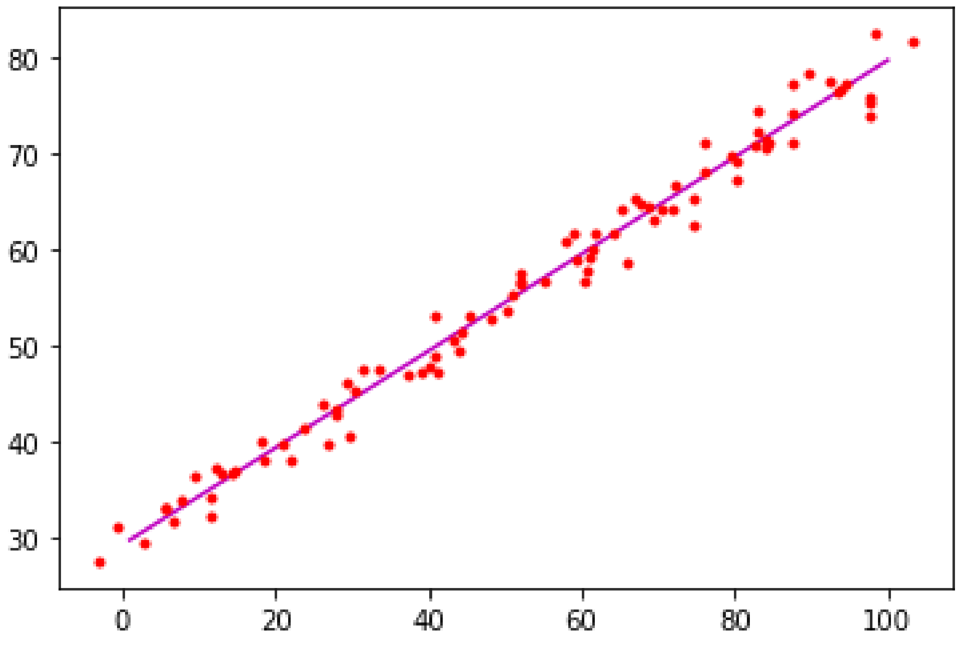
In polar coordinate case, the other parameter (theta) is now bounded.

2.

Original points:



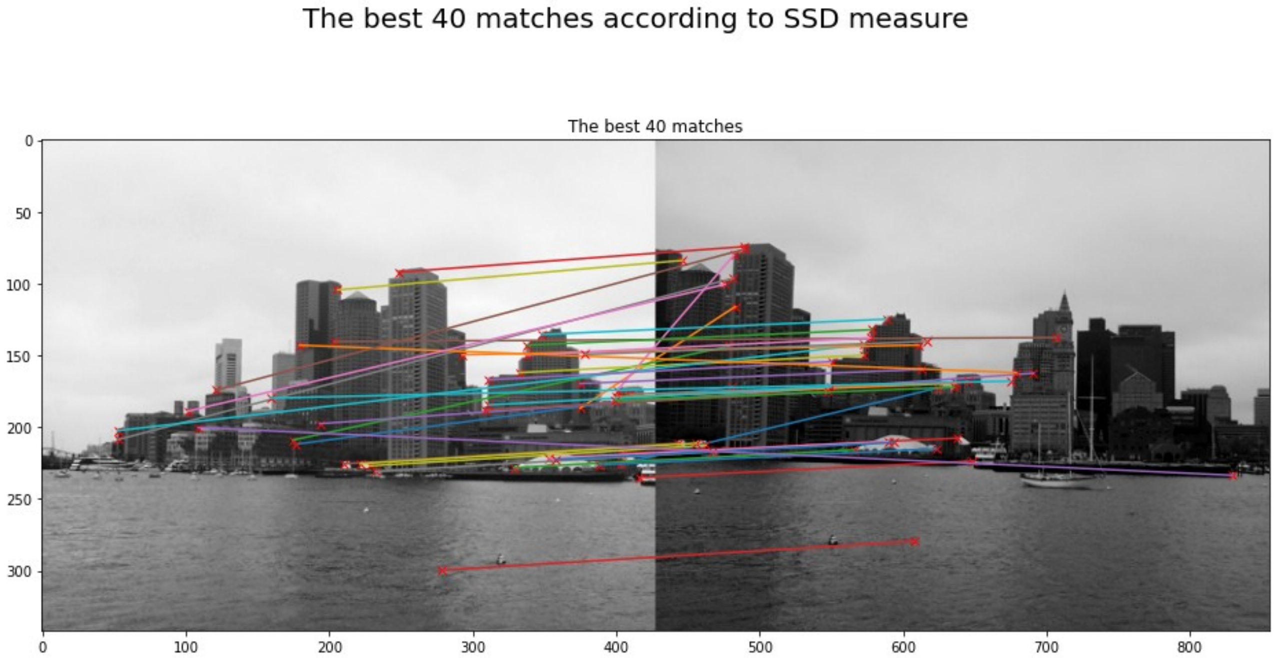
Line with best RANSAC-fit and corresponding inliers:



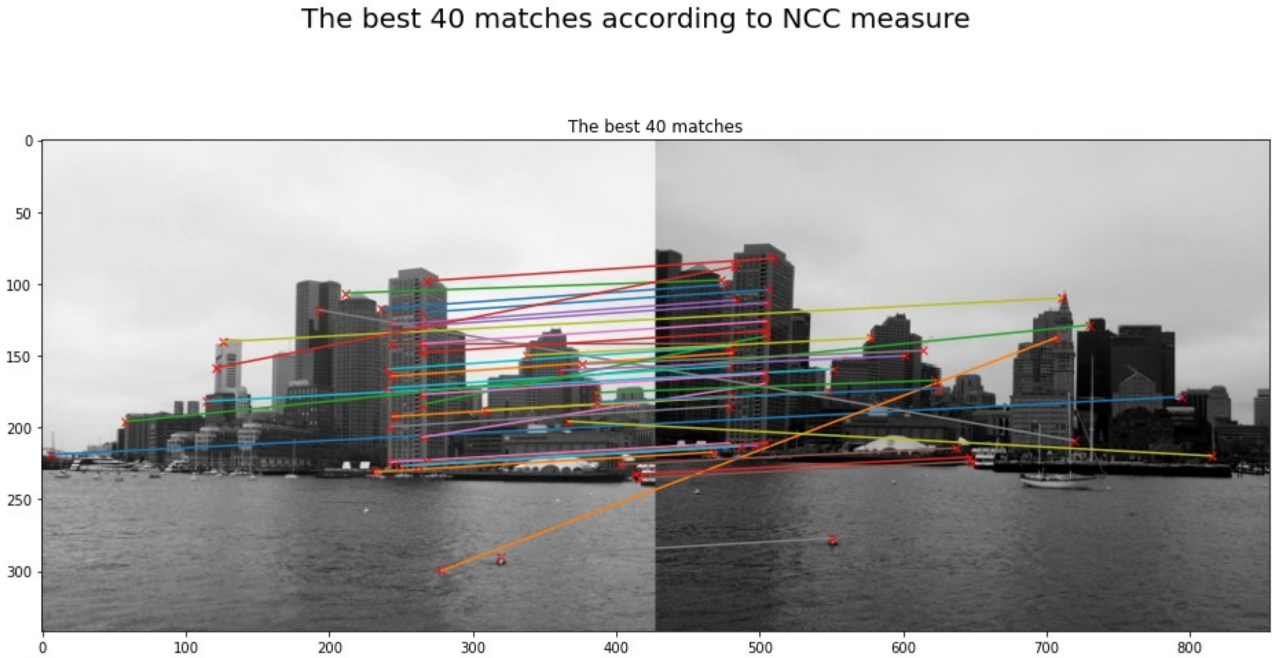
3.

a)

Matches using SSD-measure:



Matches using NCC-measure:



b)

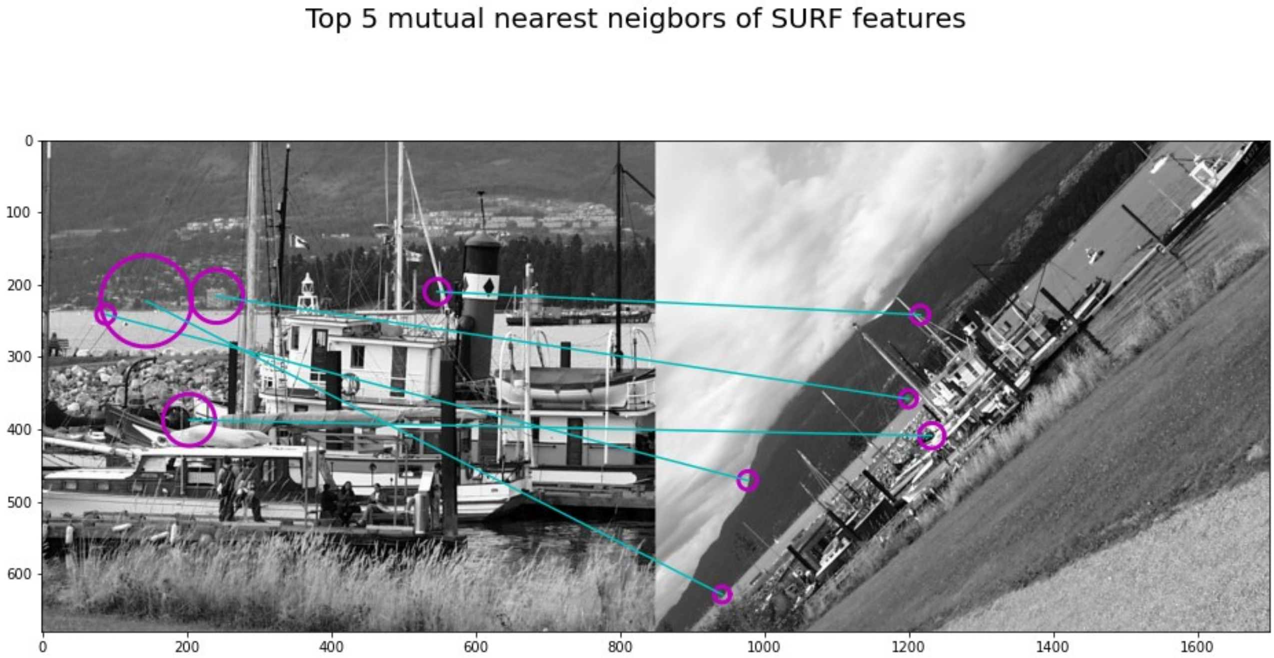
First of all, it is quite difficult to strictly put one of the methods above the other, since in this problem, the performance can be estimated in multiple ways (and because the plots are messy). For example, if we look only the left part of the left image, the corner matching with NCC-measure does better job since there are fewer keypoints assigned in this area (this area doesn't appear in the right picture at all). Correspondingly, on the right side of the right image, there aren't notable difference in methods in the terms of the counts of mispositioning. The second merit that might be good to evaluate is the diversity of the (correct) pairings. In method using SSD, there is visible cluster of keypoints on the location of the medium-large building (about at the center of the right image). There is also similar cluster in the output produced by the method using NCC, but the concentration is at the location of the foremost tall building. In both cases, these clusters might have too tight variation to present any valuable information compared to the same cluster with a smaller amount of keypoints included. However, I think that the method with NCC measure covers (correctly) the shared part of the both images (center) better compared to the SDD measure. By this brief analysis, I would say that the method with NCC similarity measure performs better.

Overall, the performance difference in this case can be explained by the differences of the images. Here the images differ from each other basically only based of the brightness (intensity of the pixels), where the rightmost picture is darker. There aren't any other notable differences, such as rotation, translation or zooming. There is slight shifting effect, but in this it case doesn’t favour either of the methods. Therefore it is reasonable to only focus on the intensity difference, in which case the NCC measure is more robust. SSD-measure can be easily fooled by the intensity changes since it measures the similarity of the templates only based on their pixel-intensities. However, the scale changes in pixels doesn’t affect much to the NCC-measure, since the scaling does not generally affect to the correlations. This can be noted if we consider the NCC-measure as a cosine of the angle between the image patch vectors (after subtracting their corresponding means). Scaling the other vector doesn’t affect to the angle between them.

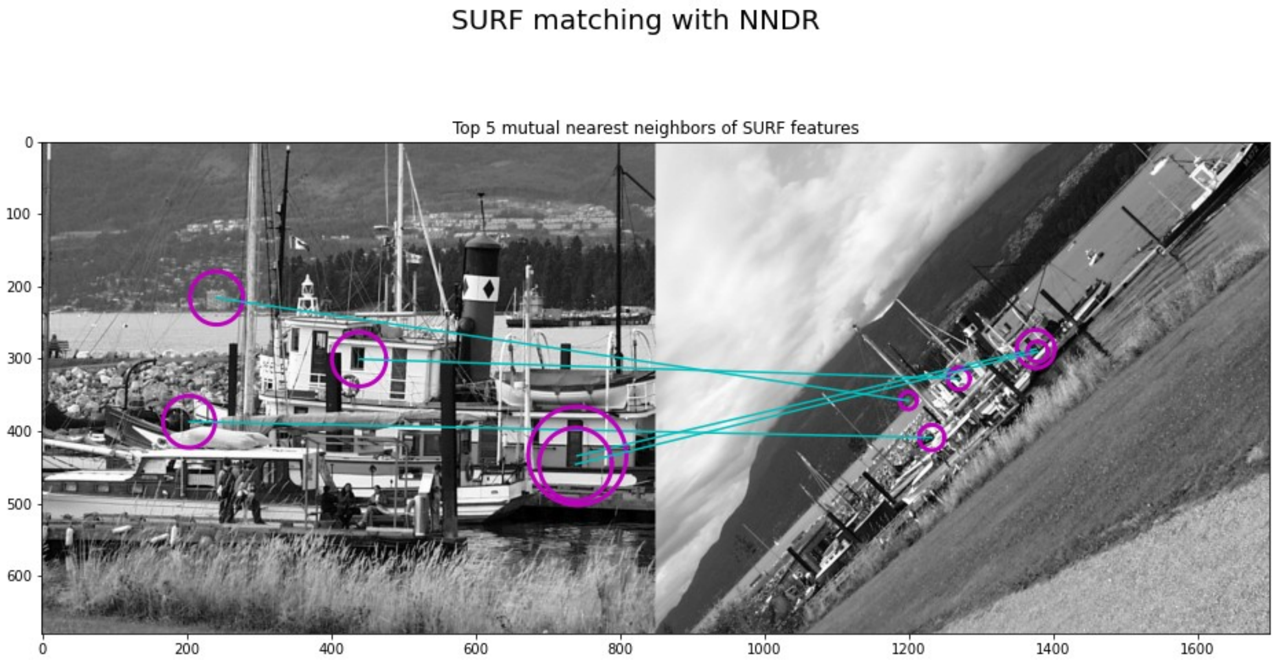
4.

a)

Nearest neighbors based on euclidean distance:



Nearest neighbors based on euclidean distance ratio (NNDR):



By the NNDR measure, it seems that all top 5 keypoints are correctly matched. In the case of nearest neighbor distance there were only 2 correct matches. Therefore the NNDR seems to work better in this case.

b)

There are several benefits using SURF over Harris corners. First of all, Harris corner detection, as the name says, can only detect come keypoints (corners). In addition to keypoint extraction, SURF (and SIFT) can however also produce more sophisticated keypoint descriptions (which are also invariant to image transformations) compared to Harris method’s raw pixel values. Instead of SIFT-method’s corner detector (blob detector) SURF uses corner detection methods as the Harris method, but also provides the descriptors, so I guess it can be thought to be overall more sophisticated tool. SURF is also more robust to image transformations, such as scaling and rotation. However, SURF is computationally more expensive to apply