Homework #4

ECE661

Name:Chengzhang Zhong

Email:zhongc@purdue.edu

Concept introduction

1.1 Harris Corner detector

(a). To begin with, the first step to use a Harris Corner detector is to compute the gradient in both x and y direction. In this experiment, we choose to use Haar filter to take care of gradient. The size of Haar filter depends on a variable σ, where filter size equals to 4σ. But filter size has to be even number. For example, when $\sigma = 1$, Haar filter looks like the following:

Haar filter in horizontal: $\begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$

Haar filter in vertical: $\begin{bmatrix} -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

(b). After gradient computed, a $5\sigma * 5\sigma$ window is chosen around each pixel. The matrix C is computed as the following format:

$$\mathrm{C} = \begin{bmatrix} \sum d_x^2 & \sum d_x d_y \\ \sum d_x d_y & \sum d_y^2 \end{bmatrix}$$

(c). Check the rank of matrix C. If rank of C = 2, it is one of our candidates corner points, otherwise, we delete it from the candidate list.

(d). For every pixel, we compute the confidence score which called Corner strength:

$$\text{Corner strength} = \lambda_1\lambda_2 - \mathrm{k}(\lambda_1 + \lambda_2)^2$$

(e). In this step, we use non-maximum suppression method. The method is implemented by setting a threshold and find the local maximum of Corner strength in a certain window area. If a point satisfies both requirements above, it will be selected as an interesting point.

1.2 Establish correspondence between two images

Before the matching step, we implement Harris corner detector on both images to find interesting points. From here, we have two methods, SSD and NCC to find the correspondence between two images.

1.2.1. SSD: Sum of Squared Differences

For each interest point in each of the pair images, we define a suitable window size around the point, then SSD will be computed as following:

$$\text{SSD} = \sum_i \sum_j |f_1(i,j) - f_2(i,j)|^2$$

$f_1(i,j)$ represents the pixel value in the first image with coordinate i and j with in the suitable window size, $f_2(i,j)$ denotes the same meaning. To minimize the result amount in SSD, we take the following two steps:

1. Set a threshold to remove weak candidates, any SSD result has a value less than the threshold will be removed from the final match pair list.

2. If $\frac{minimum\ \ of\ SSD}{second\ \ minimum\ \ of\ SSD}$ is less than a certain ratio, the result will be dumped from the final match pair list.

1.2.2. NCC: Normalized Cross Correlation

Similar as SSD, we define a suitable window size around each interesting point in each image pair. NCC formula is computed as following:

$$\mathrm{NCC} = \frac{\sum_i \sum_j (f_1(i,j) - u_1)(f_2(i,j) - u_2)}{\sqrt{[\sum_i \sum_j (f_1(i,j) - u_1)^2][\sum_i \sum_j (f_2(i,j) - u_2)^2]}}$$

where u_1 is the mean inside the suitable window from image1, u_2 is the mean inside the suitable window from image2. To minimize the result amount in NCC, we take the following two steps:

1. Set a threshold to remove weak candidates, any NCC result has a value larger than the threshold will be removed from the final match pair list.

2. If $\frac{maxi\ mum\ \ of\ \ SSD}{second\ \ maximum\ \ of\ SSD}$ is less than a certain ratio, the result will be dumped from the final match pair list.

1.3 Parameter setting

σ -- scale factor in Harris corner detector, chosen as 1.0, 1.4, 1.8 and 2.2 in this experiment.

$4 * \sigma$ -- size of Haar Filter used to compute gradient

$5 * \sigma$ -- size of window used to compute C matrix

$20 * \sigma$ -- size of window chose to compute SSD

$20 * \sigma$ -- size of window chose to compute NCC

$5 * absolute\ value\ of\ minimum\ in\ SSD\ matrix$ -- threshold used to constrain SSD candidates

$0.9 * absolute\ value\ of\ maximum\ in\ NCC\ matrix$ -- threshold used to constrain NCC candidates

0.85 -- ratio used constrain SSD

1.09 -- ratio used constrain NCC

2.1.1 SIFT feature

In SIFT algorithm, we first need to find local extrema using DoG pyramid, which can be computed as:

$$\mathrm{D}(\mathrm{x}, \mathrm{y}, \sigma) = \big(\mathrm{G}(\mathrm{x}, \mathrm{y}, \mathrm{k}\sigma) - \mathrm{G}(\mathrm{x}, \mathrm{y}, \sigma) * \mathrm{I}(\mathrm{x}, \mathrm{y})\big) = \mathrm{ff}(\mathrm{x}, \mathrm{y}, \mathrm{k}\sigma) - \mathrm{ff}(\mathrm{x}, \mathrm{y}, \sigma)$$

$$\mathrm{G}(\mathrm{x}, \mathrm{y}, \sigma) = \frac{1}{2\pi\sigma^2} \mathrm{e}^{-(\mathrm{x}^2+\mathrm{y}^2)/2\sigma^2}$$

where we know $\mathrm{G}(\mathrm{x}, \mathrm{y}, \sigma)$ is a variable scale Gaussian and $\mathrm{ff}(\mathrm{x}, \mathrm{y}, \sigma)$ is Laplacian of Gaussian. And for each point in the DoG pyramid, it will need to be compared with 8 points at the same scale in 3x3 neighbor, 9 points in the next scale in 3x3 neighbor, 9 points in the previous scale in 3x3 neighbor.

To locate the extrema in the sub-pixel accuracy, we need to calculate the second derivative of $\mathrm{D}(\mathrm{x}, \mathrm{y}, \sigma)$. To do this, we need to find the Taylor series expansion first:

$$\mathrm{D}(\mathrm{x}) = \mathrm{D}(\mathrm{x}_0) + \mathrm{J}^{\mathrm{T}}(\mathrm{x}_0)\mathrm{x} + \frac{1}{2}\mathrm{x}^{\mathrm{T}}\mathrm{H}(\mathrm{x}_0)\mathrm{x}$$

where x is the incremental deviation of x_0. J is the gradient vector and H is the

Hessian matrix

$$J(x_0) = \left(\frac{\partial D}{\partial x}, \frac{\partial D}{\partial y}, \frac{\partial D}{\partial z}\right)^T_{x_0}$$

$$H(x_0) = \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial x \partial \sigma} \\ \frac{\partial^2 D}{\partial y \partial x} & \frac{\partial^2 D}{\partial y^2} & \frac{\partial^2 D}{\partial y \partial \sigma} \\ \frac{\partial^2 D}{\partial \sigma \partial x} & \frac{\partial^2 D}{\partial \sigma \partial y} & \frac{\partial^2 D}{\partial \sigma^2} \end{bmatrix}$$

To threshold out weak extrema, we can set the threshold value equals to 0.03. The remained points are named as candidate points.

To find the dominant local orientation, we need to calculate the gradient vector of the Gaussian-smoothed image $ff(x, y, \sigma)$ at the scale σ. $m(x, y)$ as gradient magnitude and $\theta(x, y)$ as gradient orientation.

$$m(x, y) = \sqrt{|ff(x+1, y, \sigma) - ff(x, y, \sigma)|^2 + |ff(x, y+1, \sigma) - ff(x, y, \sigma)|^2}$$

$$\theta(x, y) = \arctan\left(\frac{ff(x+1, y, \sigma) - ff(x, y, \sigma)}{ff(x, y+1, \sigma) - ff(x, y, \sigma)}\right)$$

In the end, for each extrema point, divide its surrounded 16x16 neighbor points into 4x4 cells. Each cell has 4x4 points. For every 16 cells, 8 bin histogram is calculated from $m(x, y)$ weighted $\theta(x, y)$ at 16 pixels in the cell. The result is a 128 elements descriptor.

2.1.2 Feature matching method

For a descriptor as SIFT, we can use Euclidean distance method to directly compare the feature vectors of interest points instead of SSD and NCC.

For Euclidean method, is simply equal to $\sqrt{SSD}$ mathematically.

$$\mathrm{SSD} = \sum_{i}\sum_{j}|f_1(i,j) - f_2(i,j)|^2$$

$f_1(i,j)$ represents the 128 element descriptor in the first image obtained with coordinate i and j, $f_2(i,j)$ denotes the same meaning. To minimize the result amount in EUC, we take the following two steps:

1. Set a threshold to remove weak candidates, any EUC result has a value less than the threshold will be removed from the final match pair list.
2. If $\frac{minimum\ \ of\ \ EUC}{second\ \ minimum\ \ of\ \ EUC}$ is less than a certain ratio, the result will be dumped from the final match pair list.

Parameter setting:

$5 * absolute\ value\ of\ minimum\ in\ EUC\ matrix$ -- threshold used to constrain EUC candidates

0.85 -- ratio used constrain EUC

3. Result graph demonstration

Figure 1. Original image 1

Figure 2. Original image 2

1.1 Harris corner points of pair1 images at sigma equals to 1

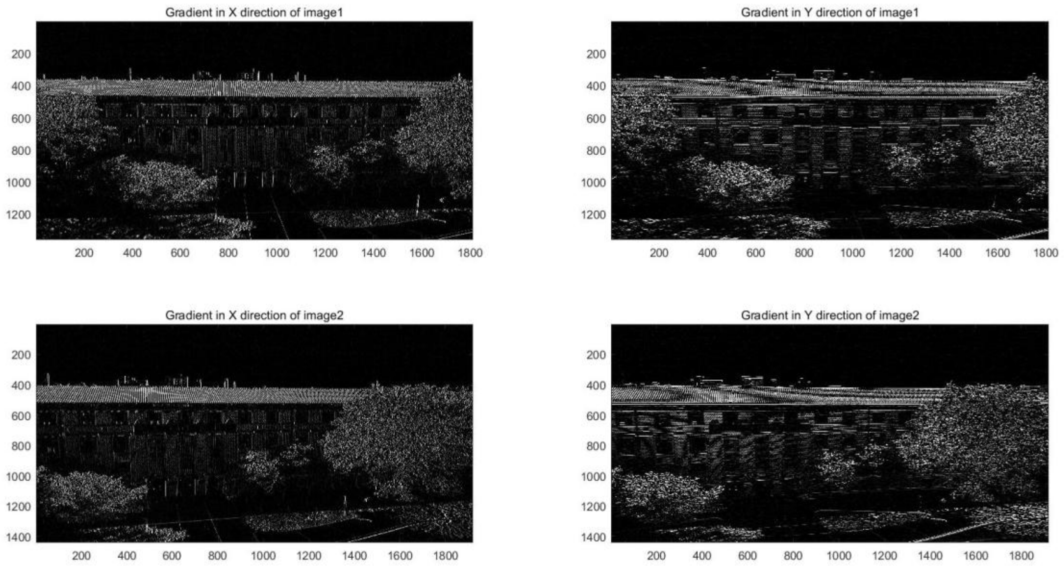


Figure 3. gradient at sigma=1

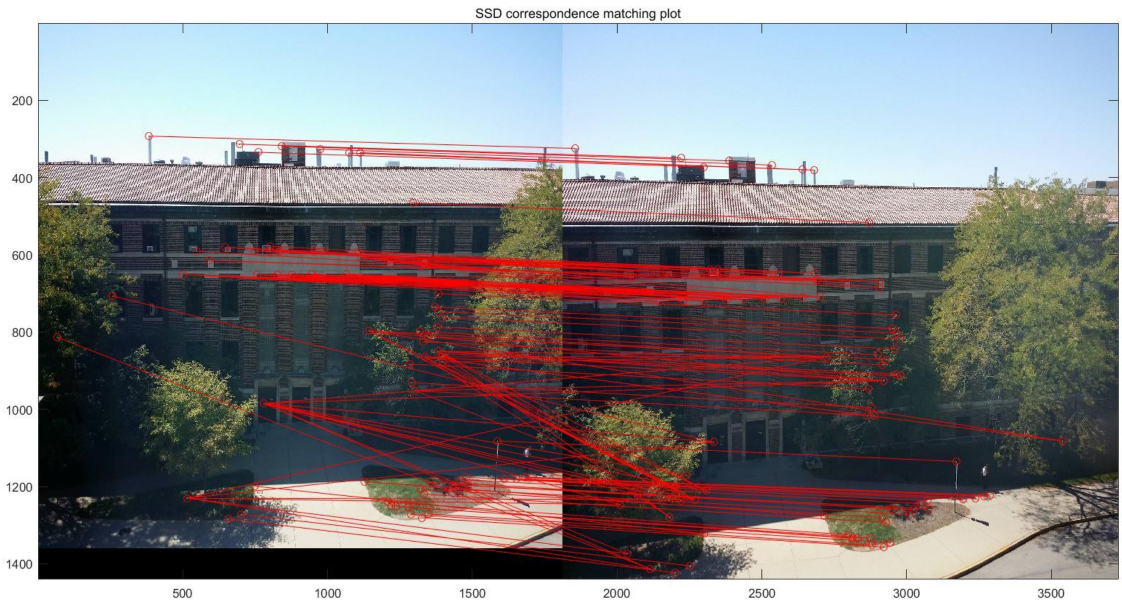


Figure 4. SSD at sigma=1

Figure 5. NCC at sigma=1

1.2 Harris corner points of pair1 images at sigma equals to 1.4

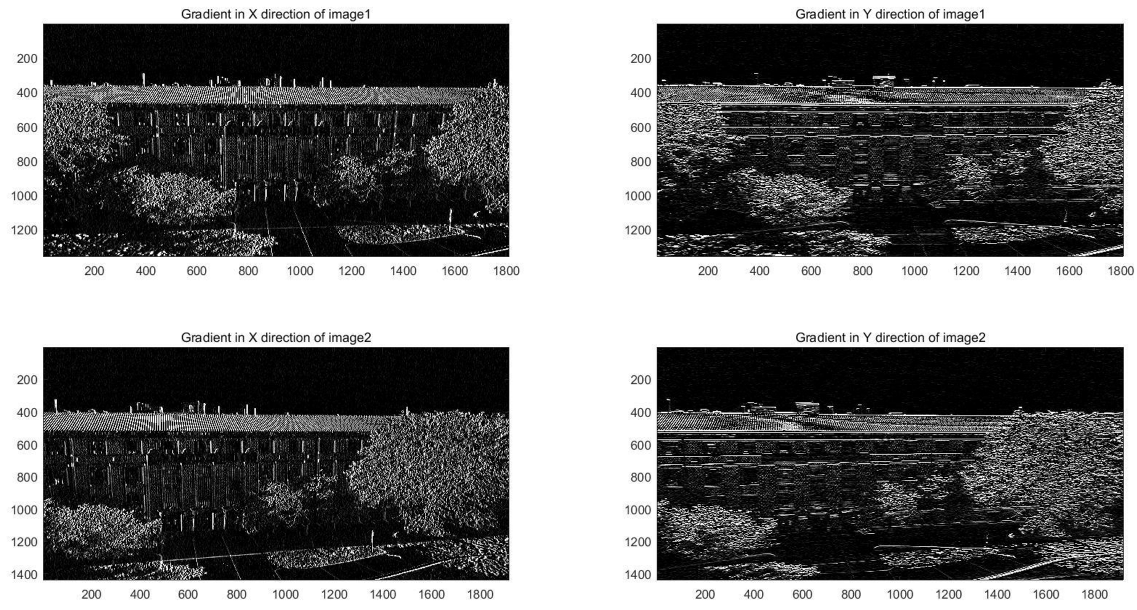


Figure 6. gradient at sigma=1.4

Figure 7. SSD at sigma=1.4

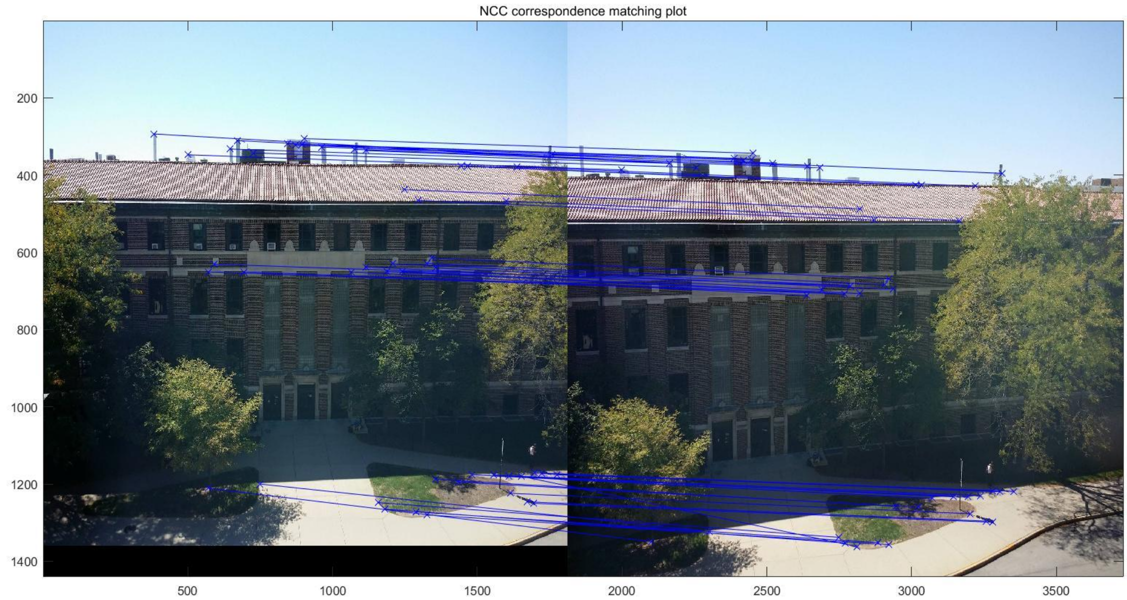


Figure 8. NCC at sigma=1.4

1.3 Harris corner points of pair1 images at sigma equals to 1.8

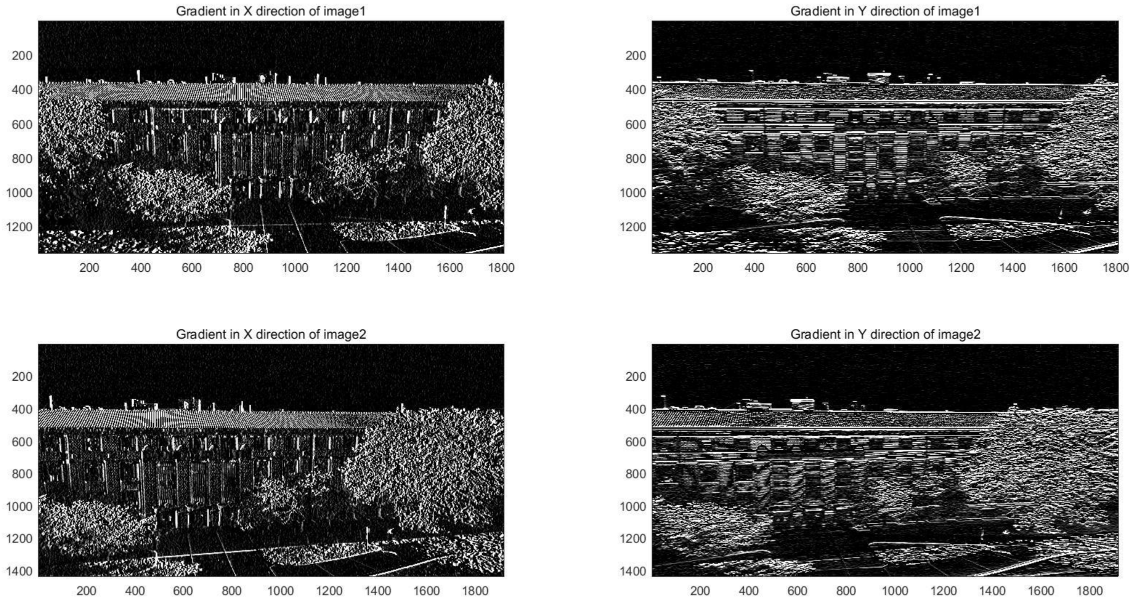


Figure 9. gradient at sigma=1.8

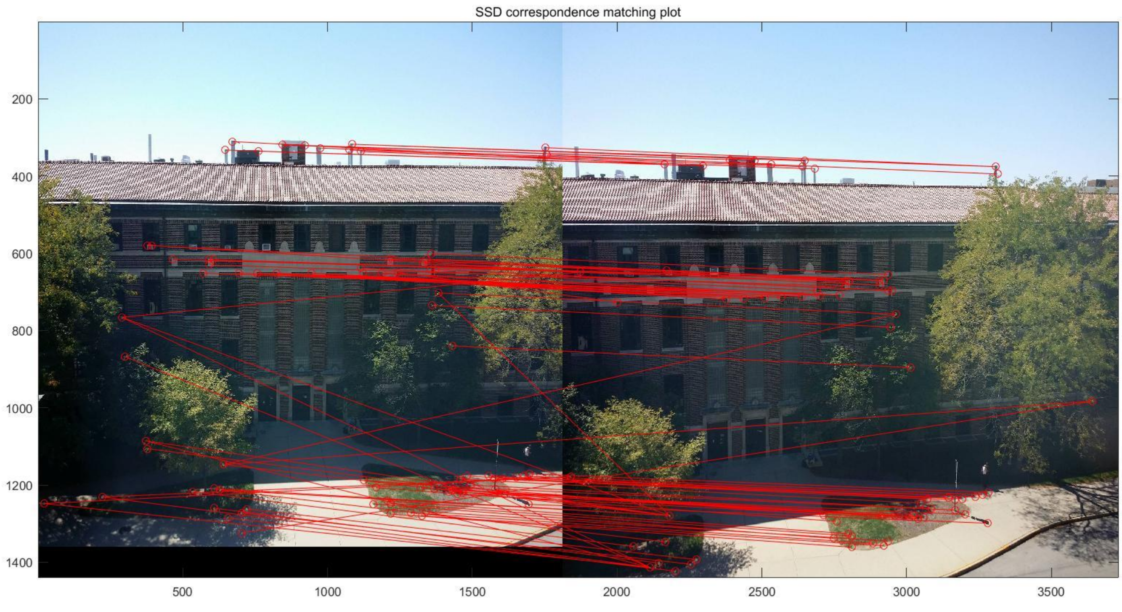


Figure 10. SSD at sigma=1.8

Figure 11. NCC at sigma=1.8

1.4 Harris corner points of pair1 images at sigma equals to 2.2

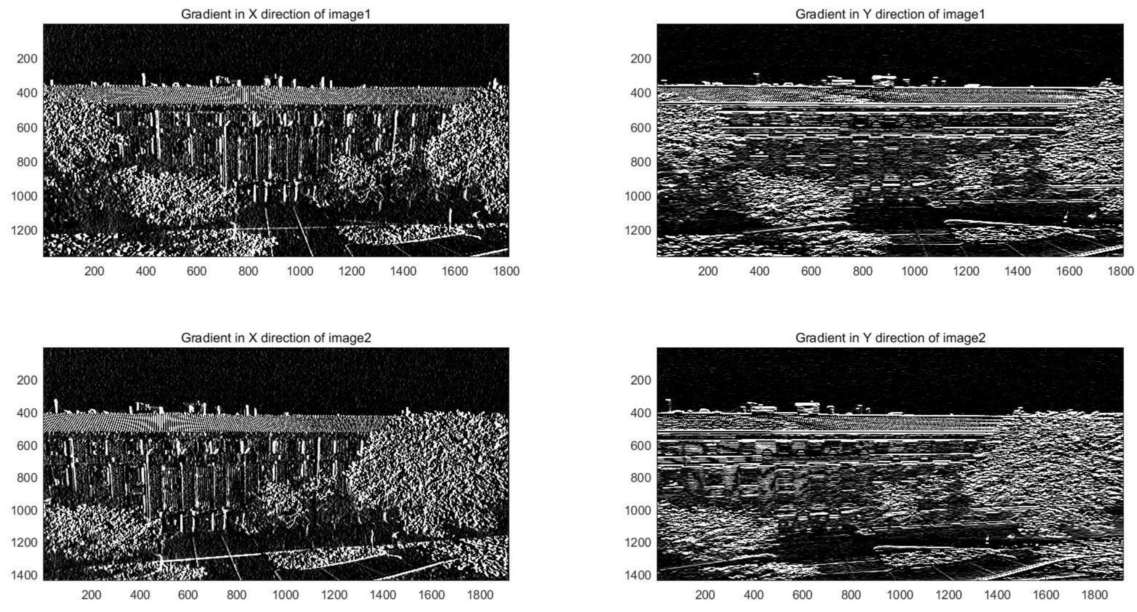


Figure 12. gradient at sigma=2.2

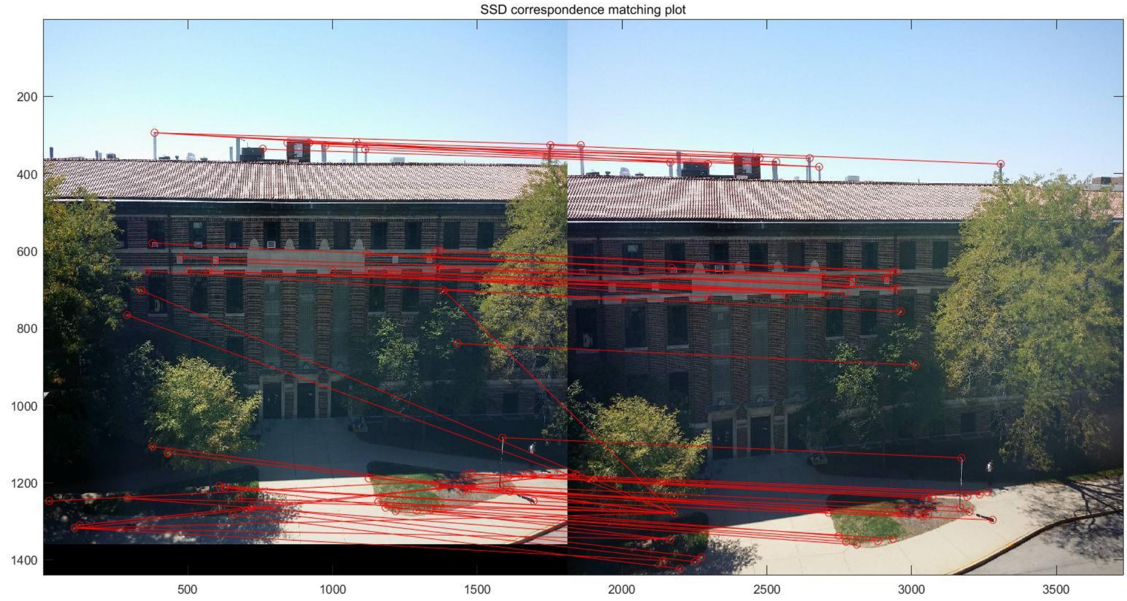


Figure 13. SSD at sigma=2.2

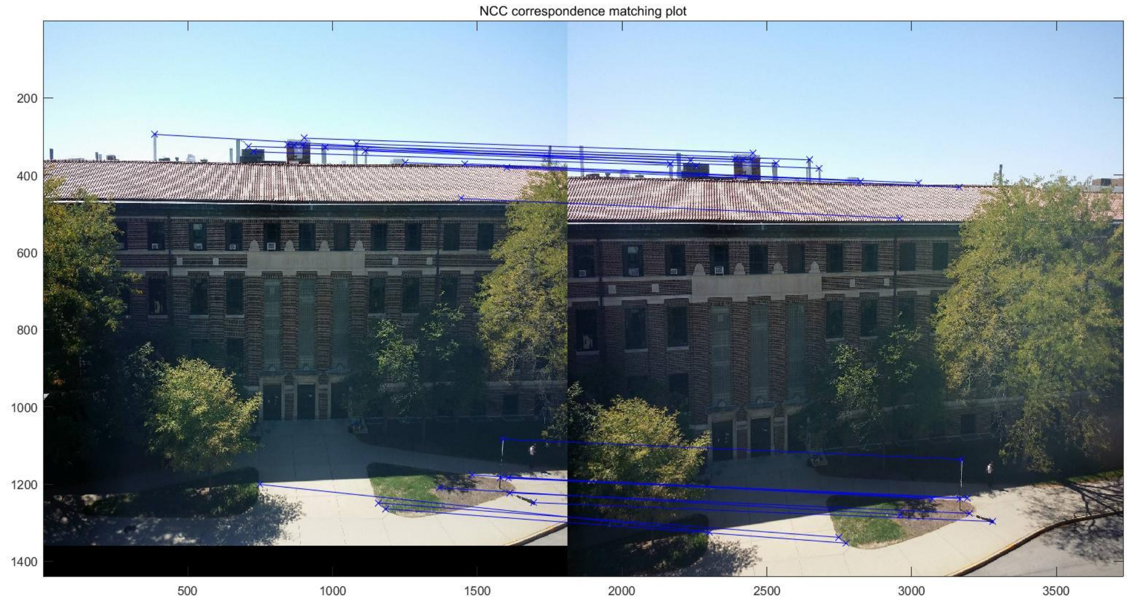


Figure 14. NCC at sigma=2.2

1.5 SIFT of pair1 images

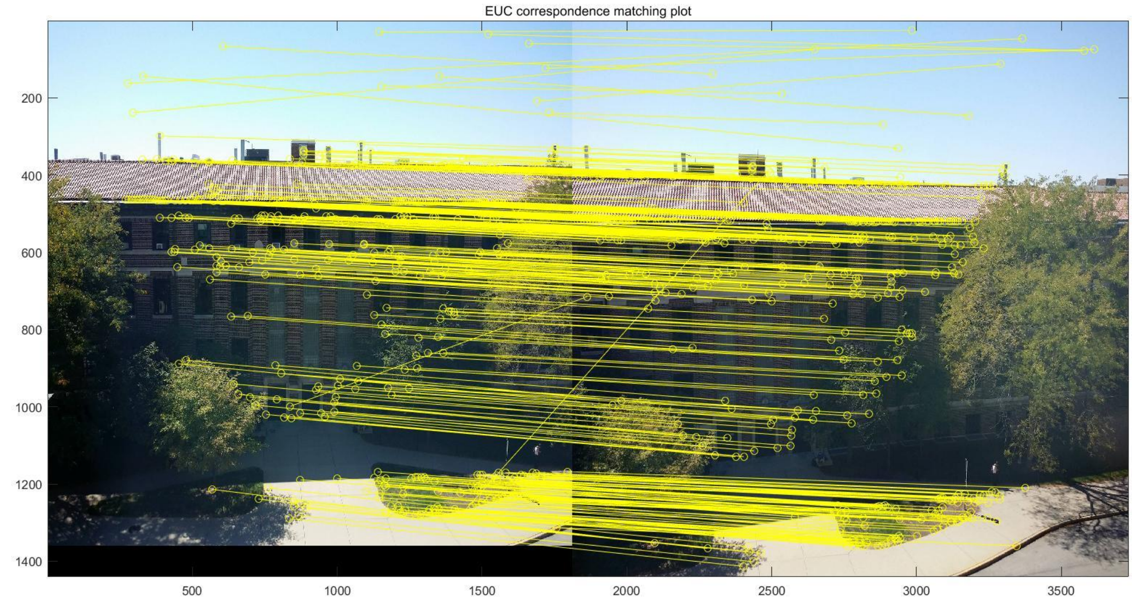
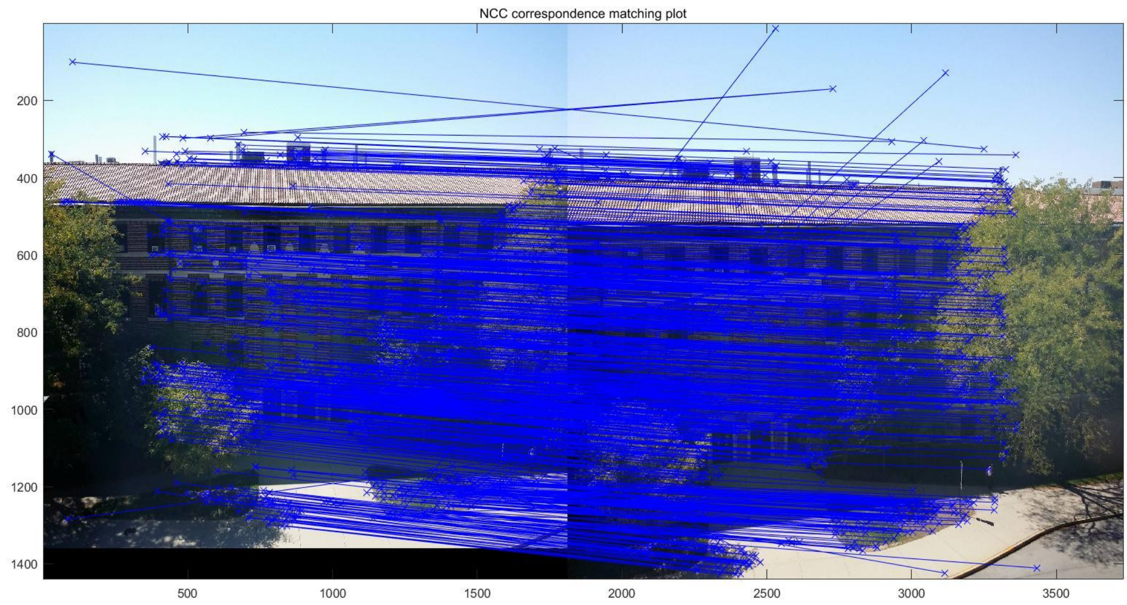


Figure 15. SIFT Euclidean result of pair1

SIFT SSD result of pair1



SIFT NCC result of pair1

1.6 Conclusion:

For this set of images, the view angle changed a little bit. For Harris corner detector, the larger the σ value, the less interest points will show up in the match result. Also, comparing the method of SSD and NCC, it looks like NCC has done a better job by making less mistakes. Moreover, a lot of interesting points on the top edges of the building are missed.

On the other hand, SIFT detected a lot more interesting points than Harris corner detector. However, SIFT also makes more mistakes, like make many points matched into one single point, false detections in the sky. Also, SIFT did a good job on detecting matched pairs on the top edges of the building.

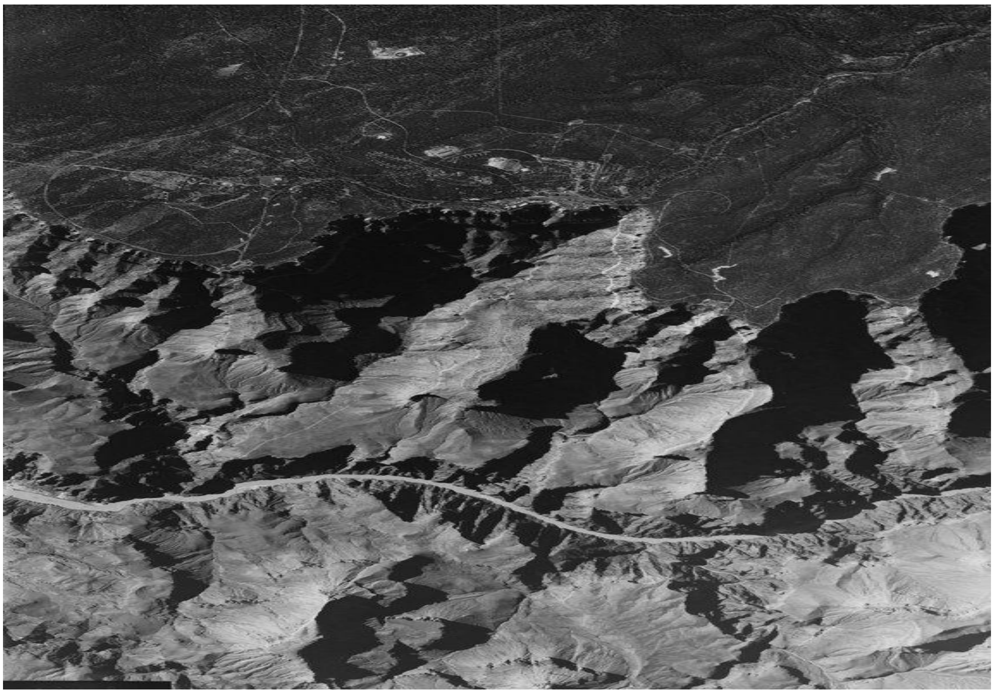


Figure 16. Original image 1

Figure 17. Original image 2

2.1 Harris corner points of pair2 images at sigma equals to 1

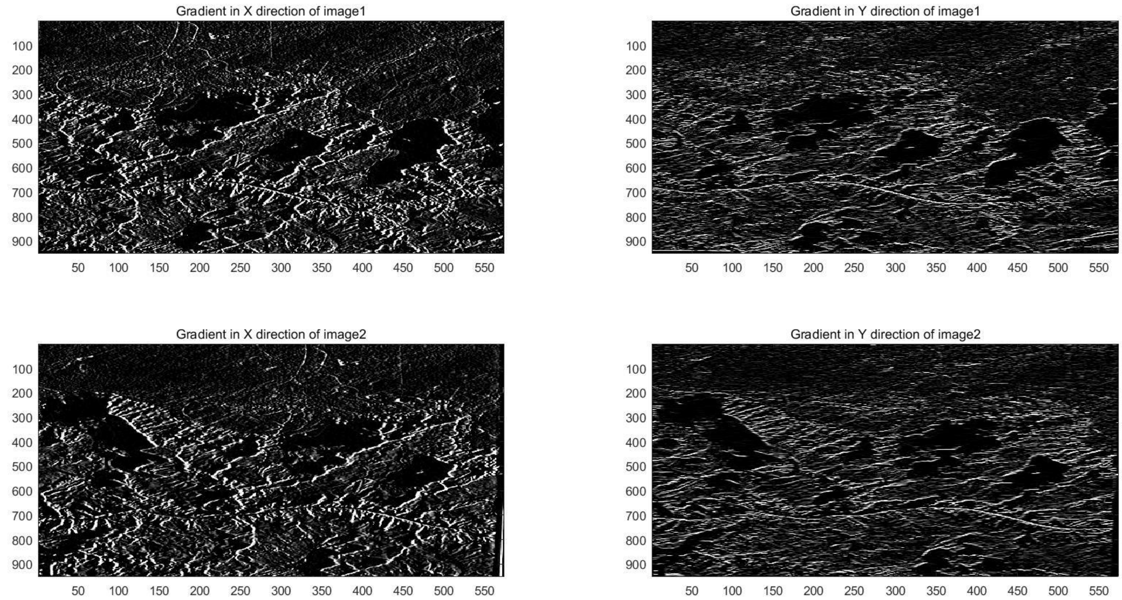


Figure 18. gradient at sigma = 1

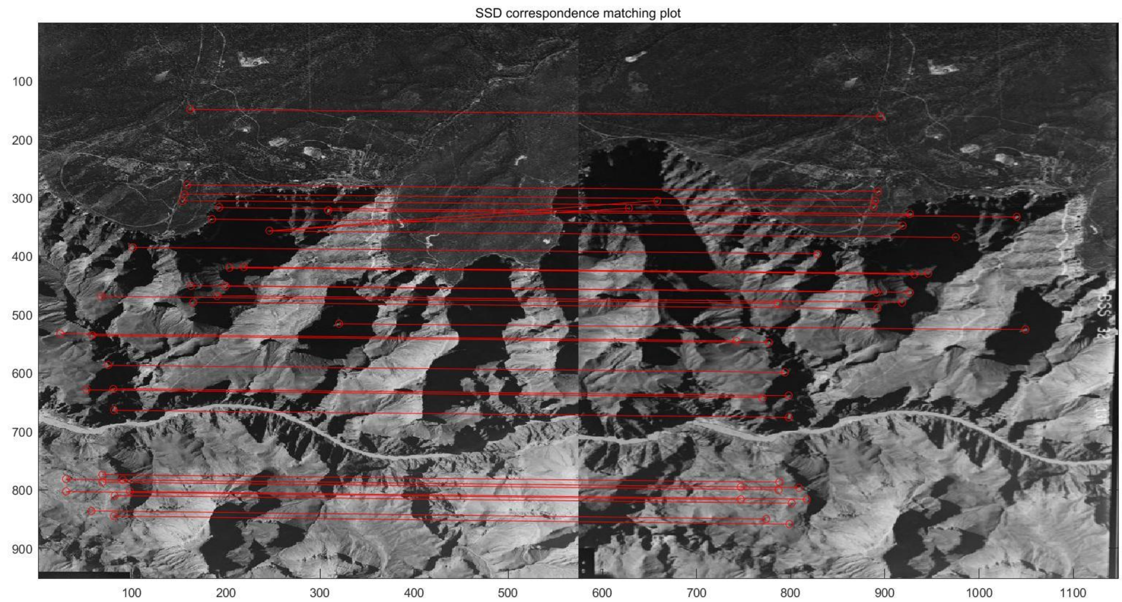


Figure 19. SSD at sigma=1

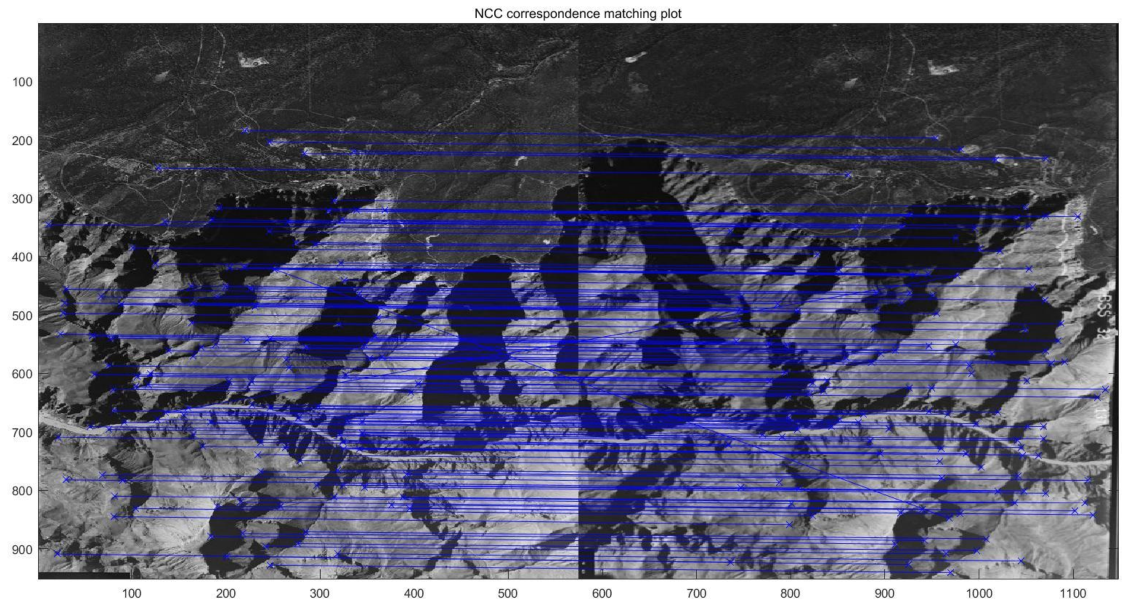


Figure 20. NCC at sigma=1

2.2 Harris corner points of pair2 images at sigma equals to 1.4

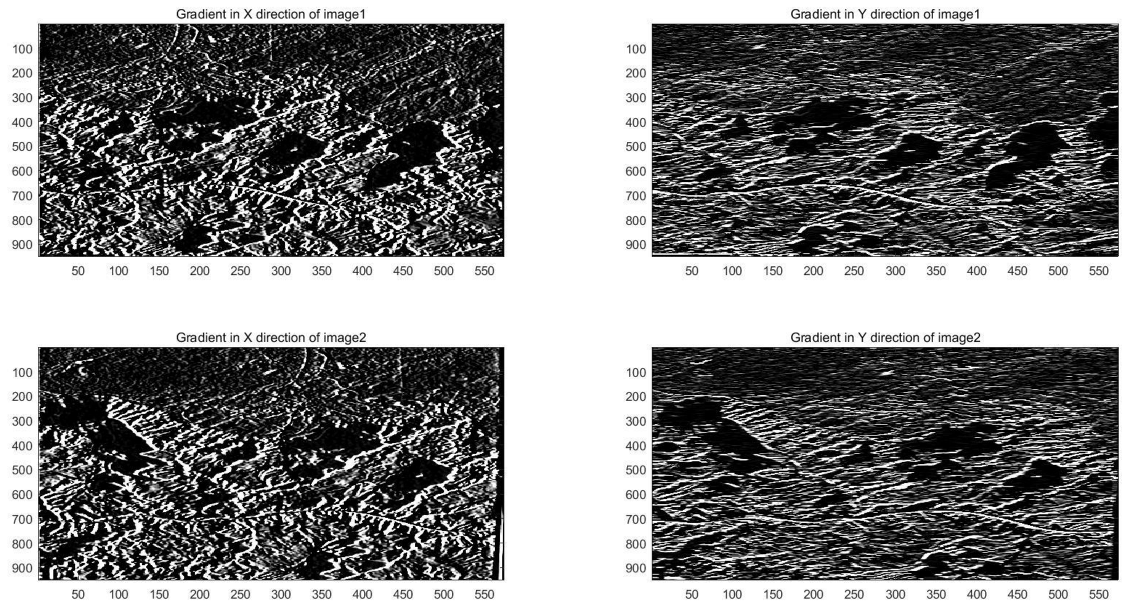


Figure 21. gradient at sigma = 1.4

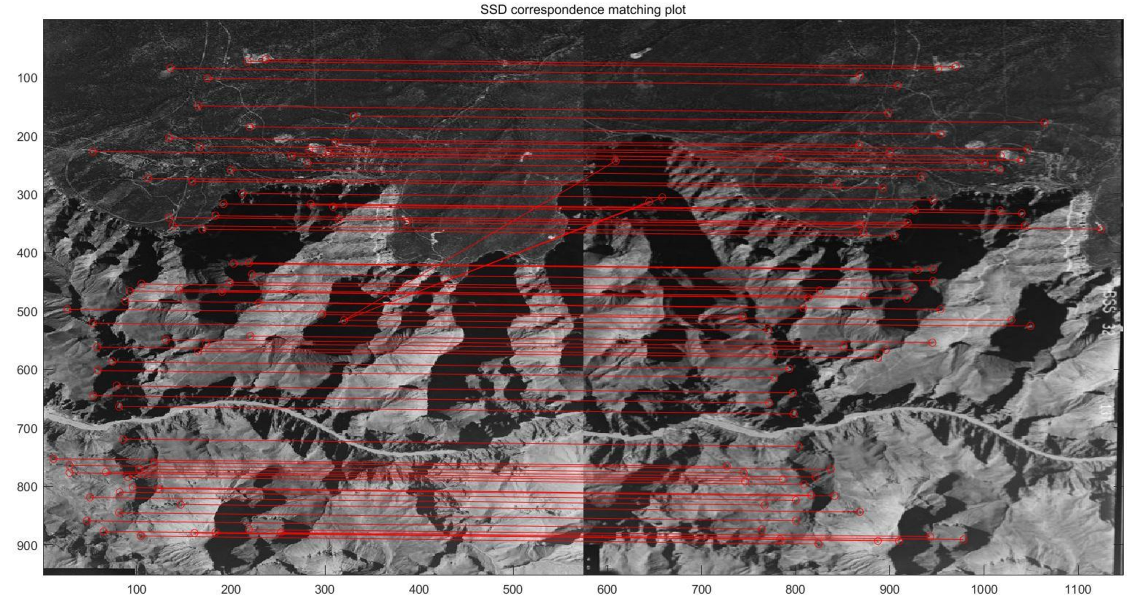


Figure 22. SSD at sigma=1.4

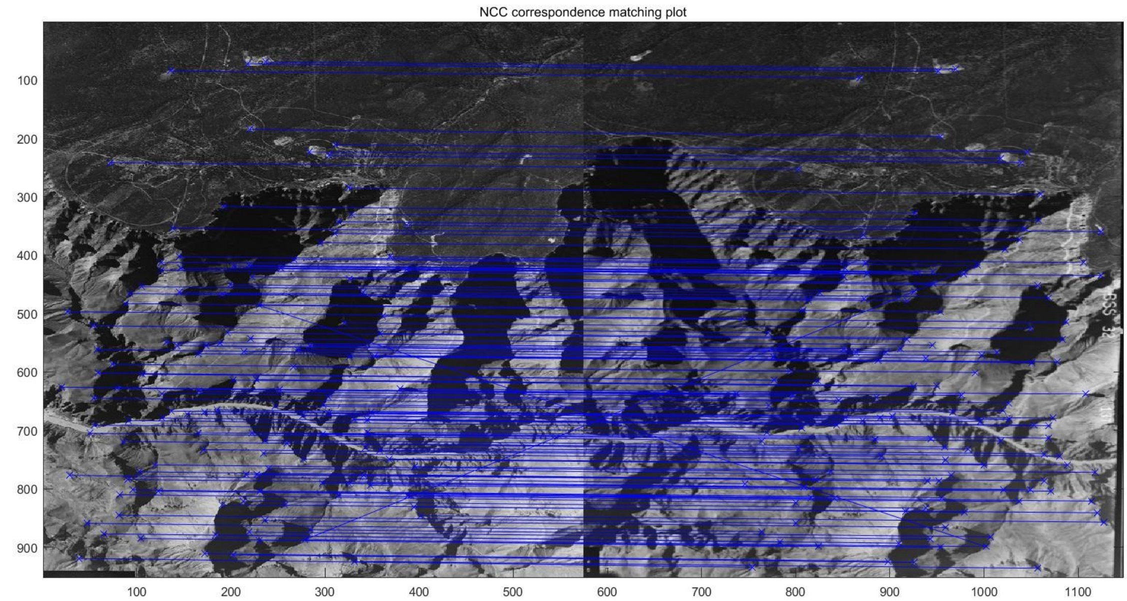


Figure 23. NCC at sigma=1.4

2.3 Harris corner points of pair2 images at sigma equals to 1.8

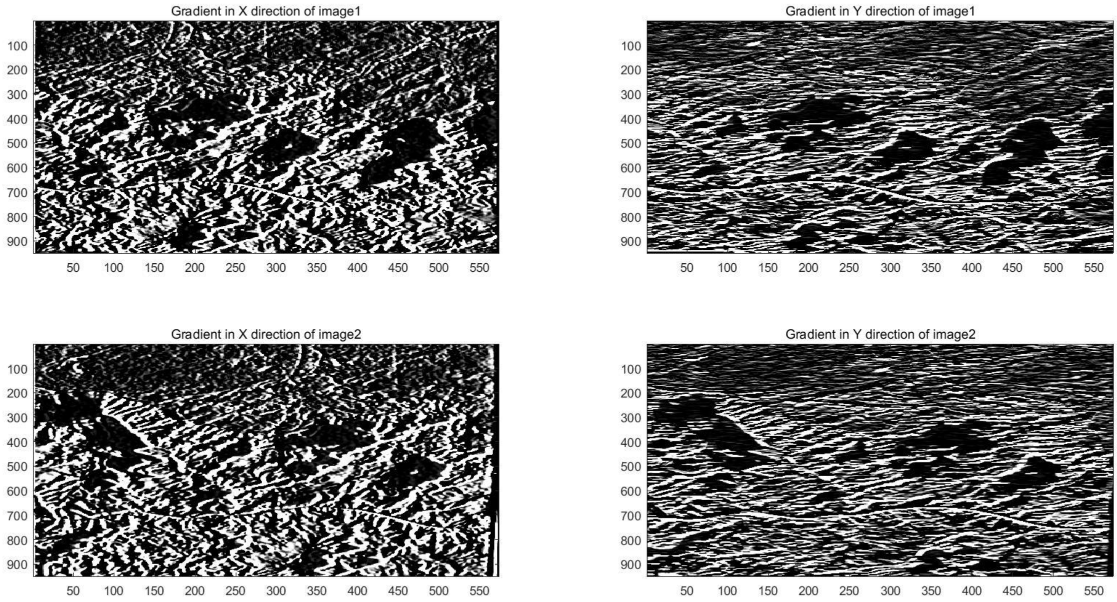


Figure 24. gradient at sigma 1.8

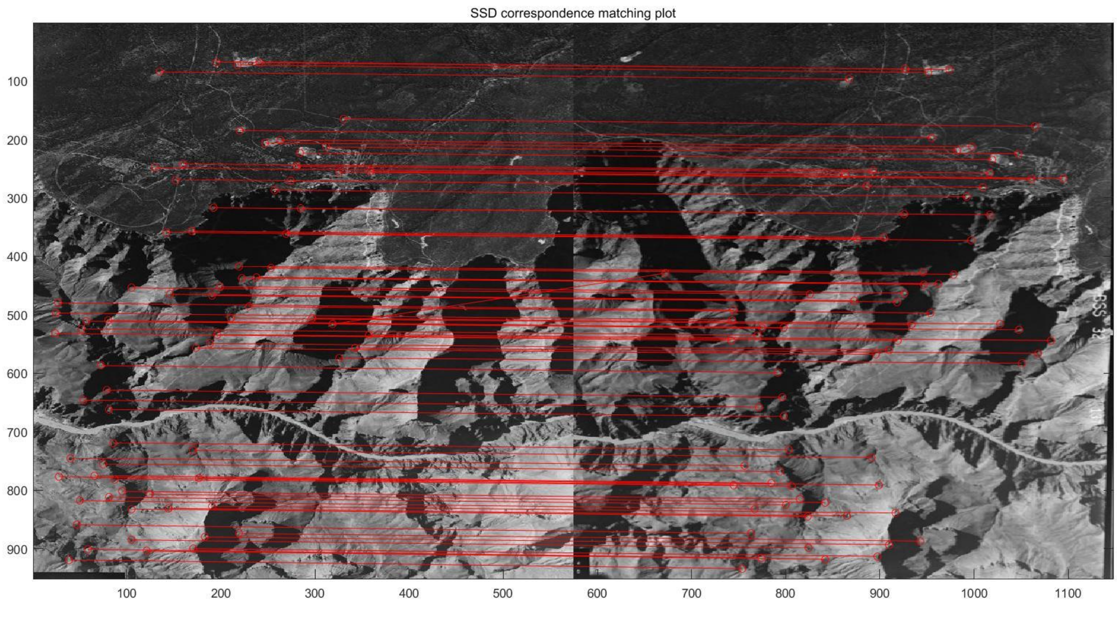


Figure 25. SSD at sigma=1.8

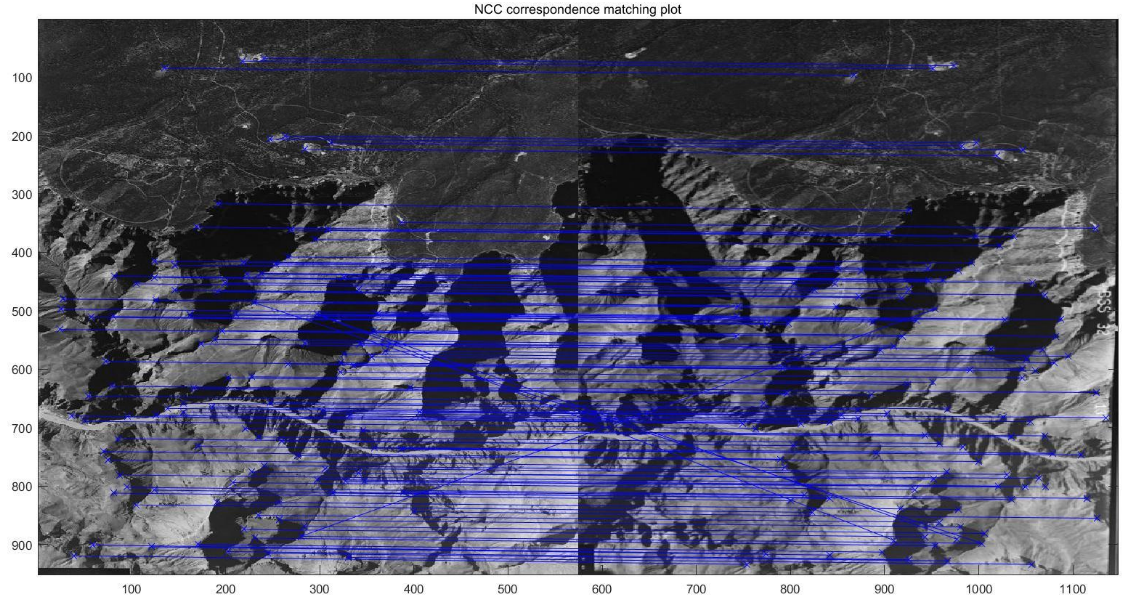


Figure 26. NCC at sigma=1.8

2.4 Harris corner points of pair2 images at sigma equals to 2.2

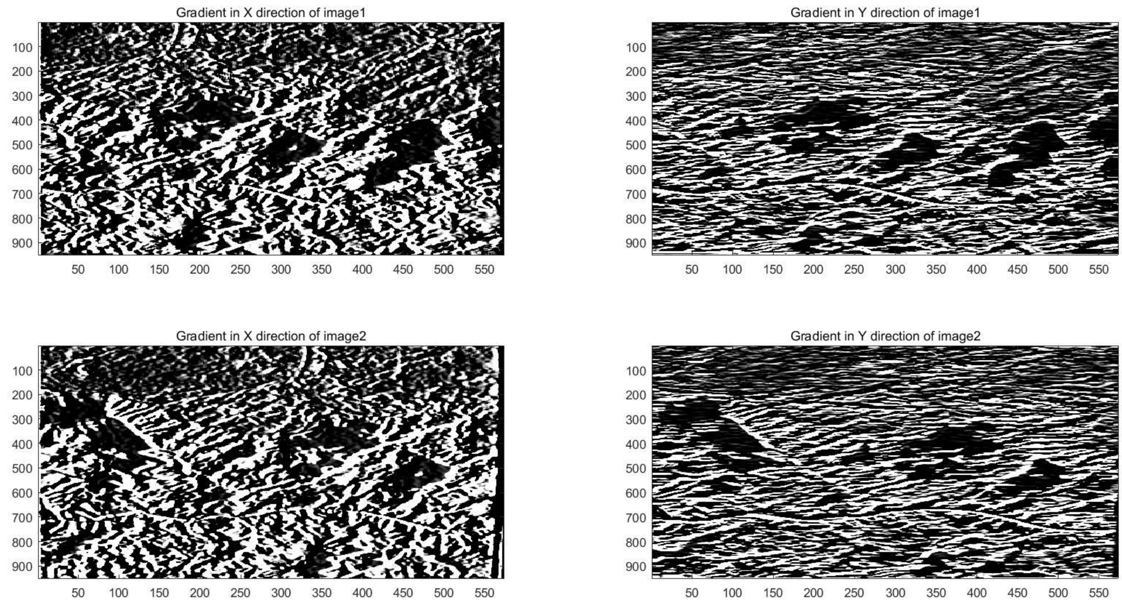


Figure 27. gradient at sigma 2.2

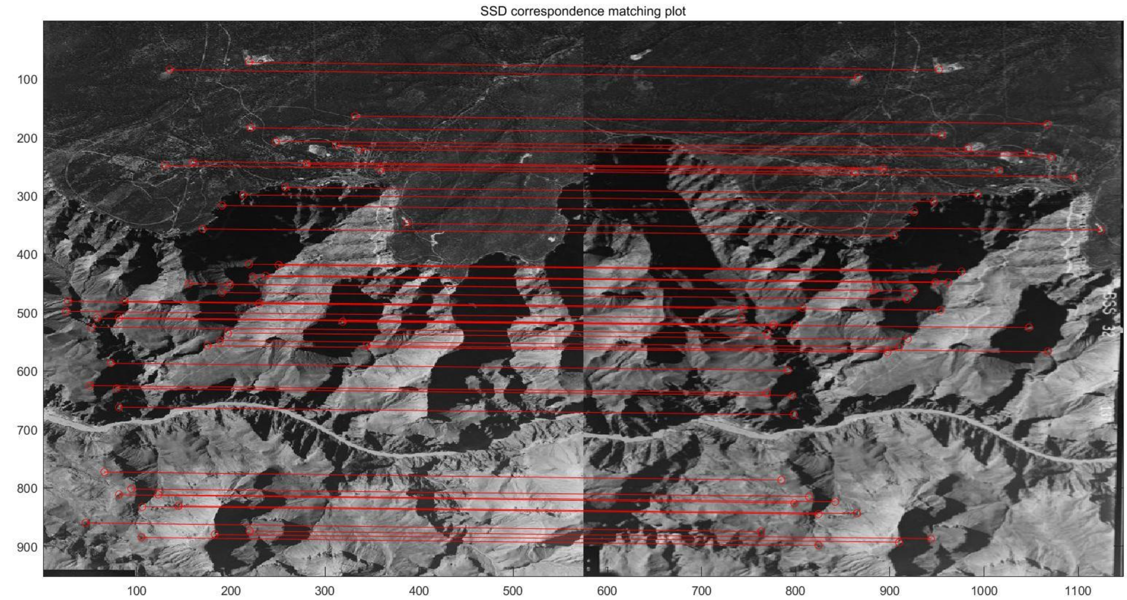


Figure 28. SSD at sigma=2.2

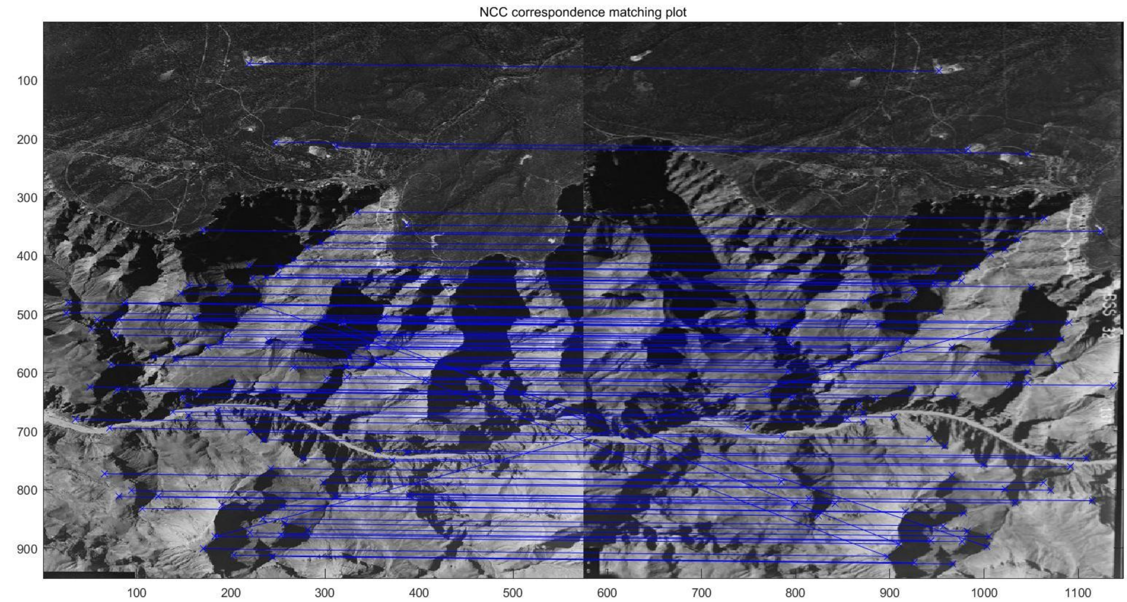


Figure 29. NCC at sigma=2.2

2.5 SIFT result of pair2

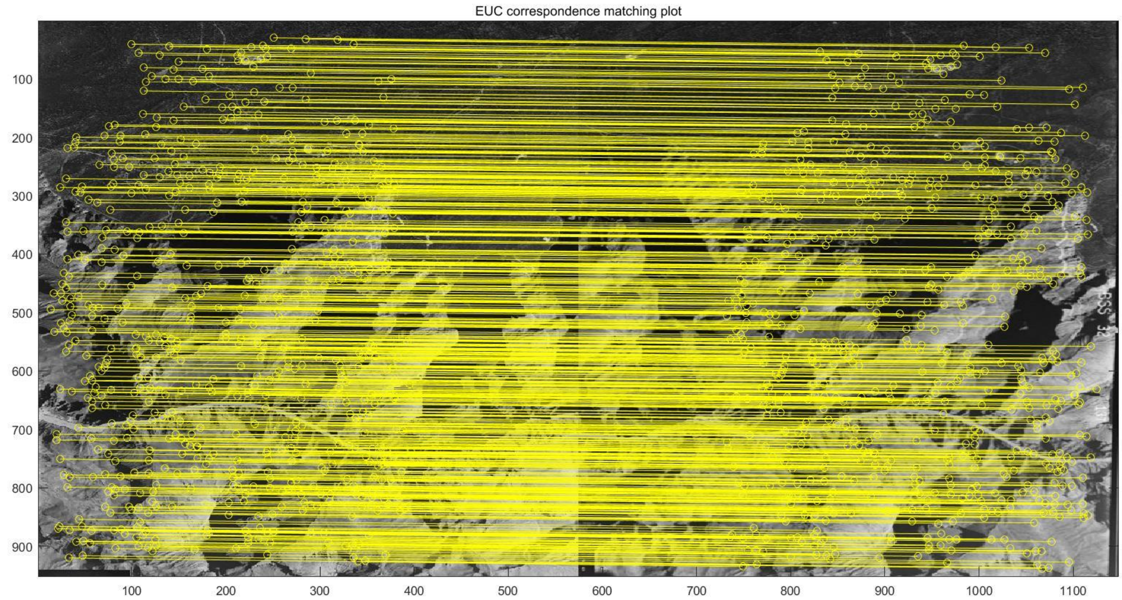
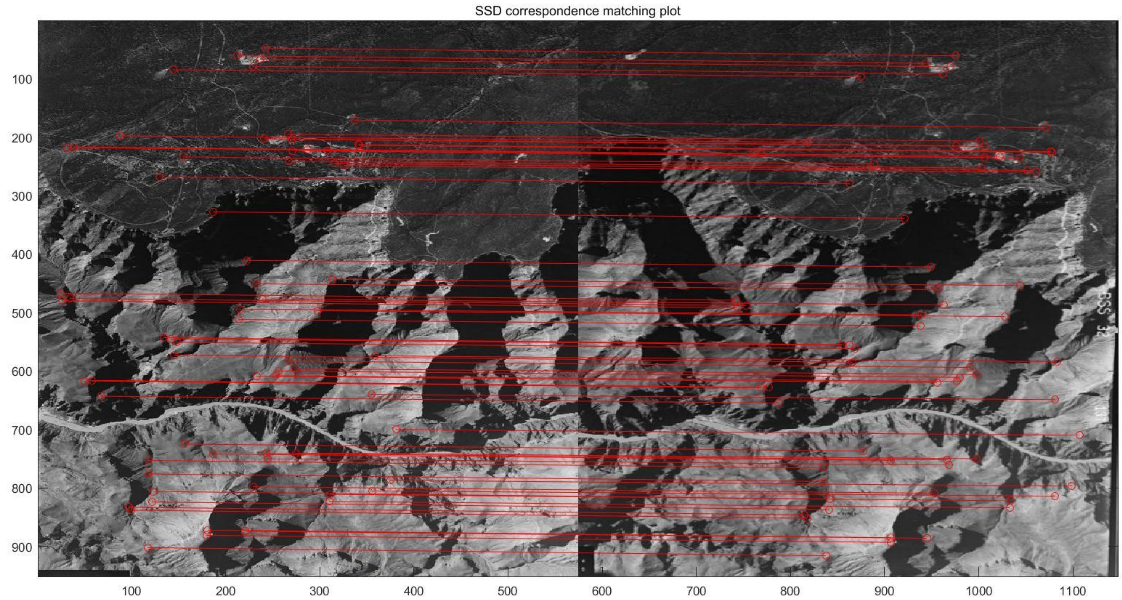
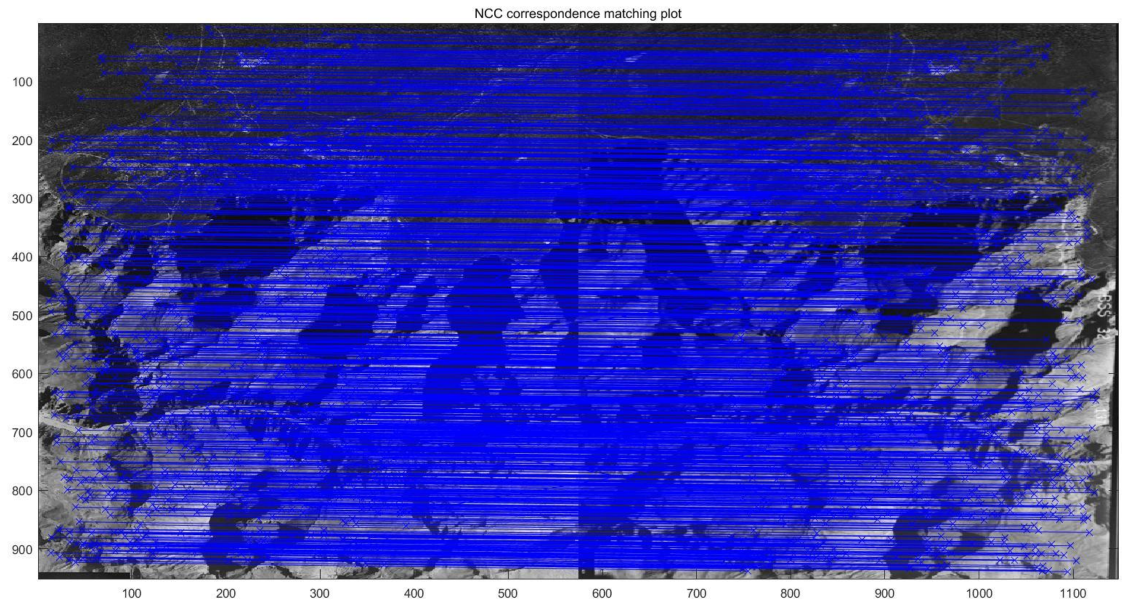


Figure 30. SIFT Euclidean of image pair2



SIFT SSD of image pair2



SIFT NCC of image pair2

2.6 Conclusion:

For this set of images, the view angle doesn't change too much as well. For Harris corner detector, the larger the σ value, the less interest points will show up in the match result. For the method of SSD and NCC, NCC always contains more interesting points than SSD.

For SIFT, it detector much more matched pairs than Harris corner. Also, it made more mistakes.

Figure 31. Original image1

Figure 32. Original image2

3.1 Harris corner points of pair3 images at sigma equals to 1

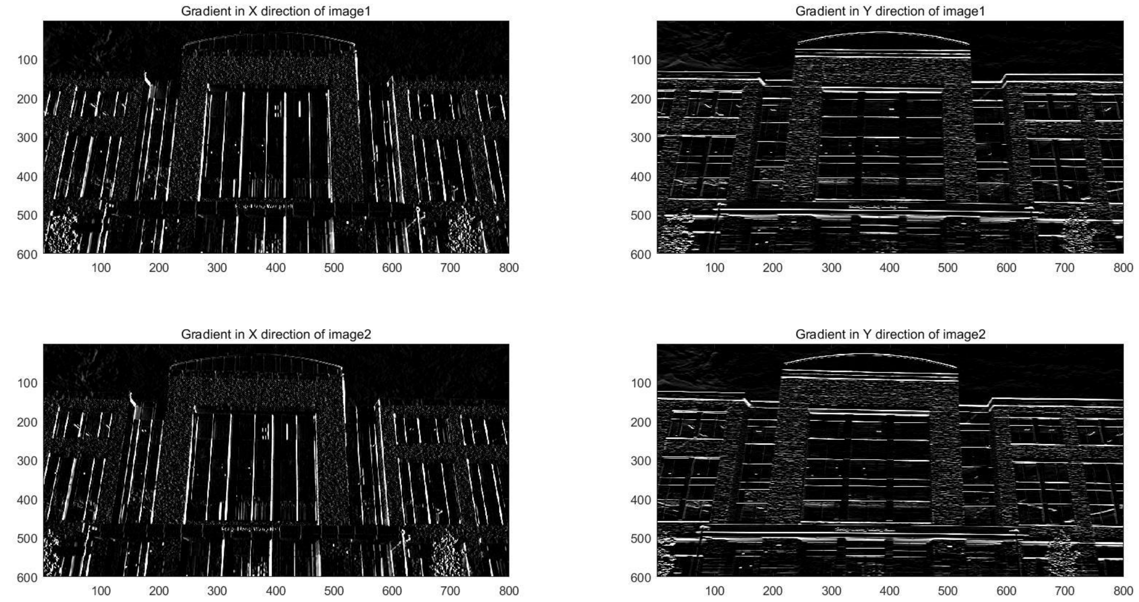


Figure 33. gradient at sigma 1

Figure 34. SSD at sigma=1

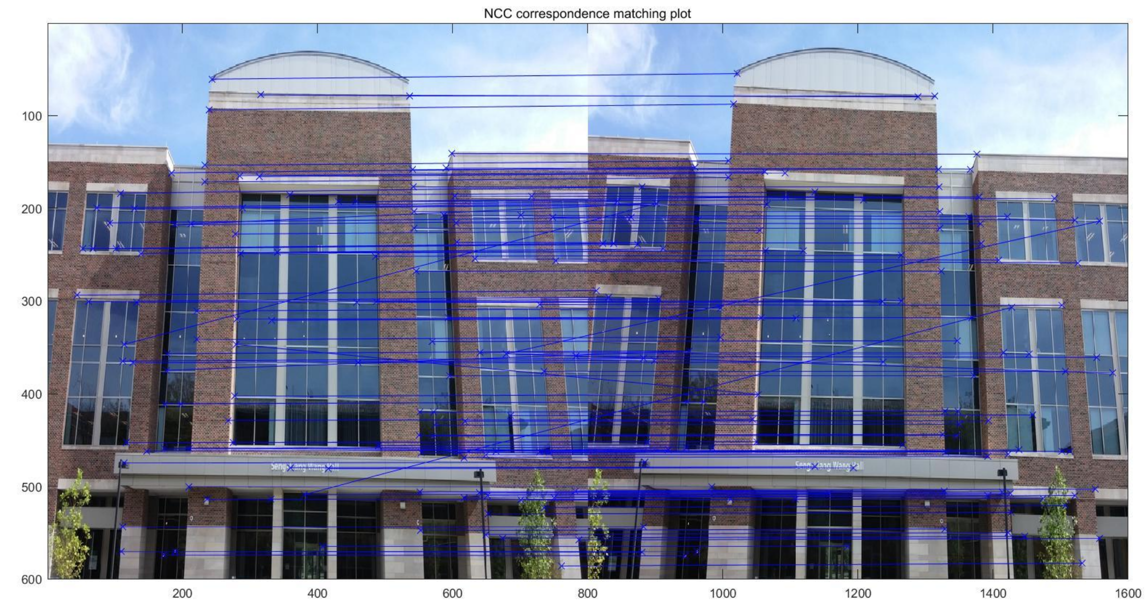


Figure 35. NCC at sigma=1

3.2 Harris corner points of pair3 images at sigma equals to 1.4

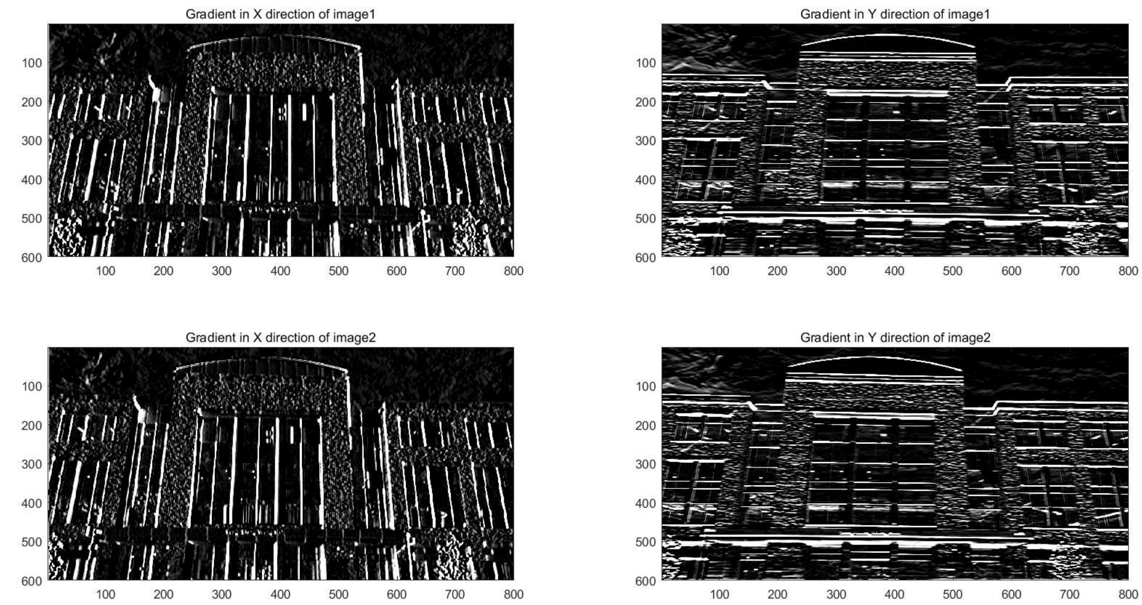


Figure 36. gradient at sigma 1.4

Figure 37. SSD at sigma=1.4

Figure 38. NCC at sigma=1.4

3.3 Harris corner points of pair3 images at sigma equals to 1.8

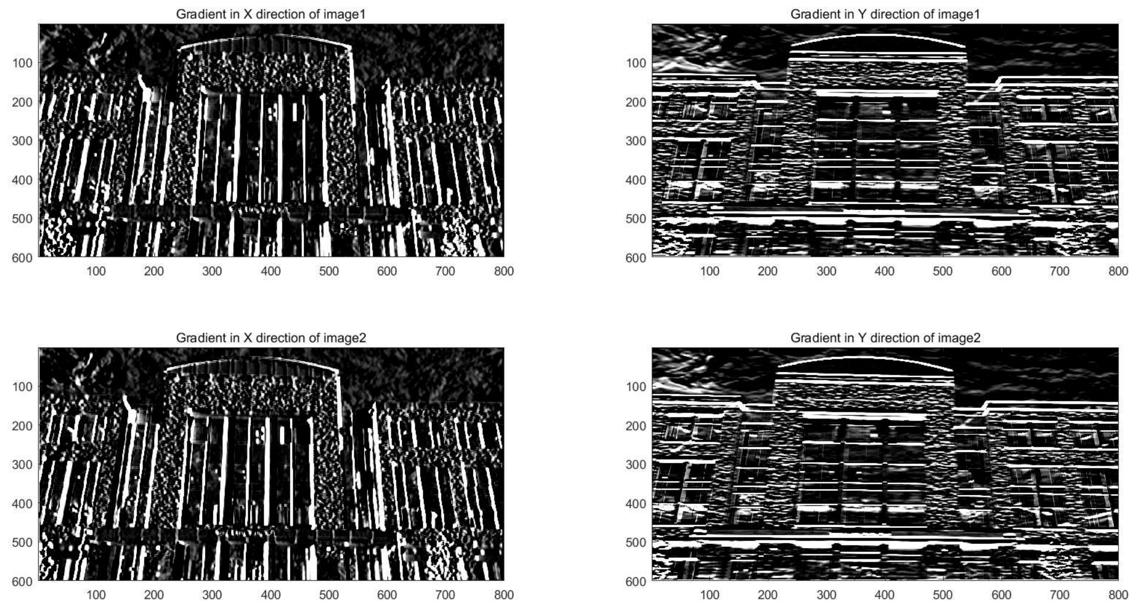


Figure 39. gradient at sigma 1.8

Figure 40. SSD at sigma=1.8

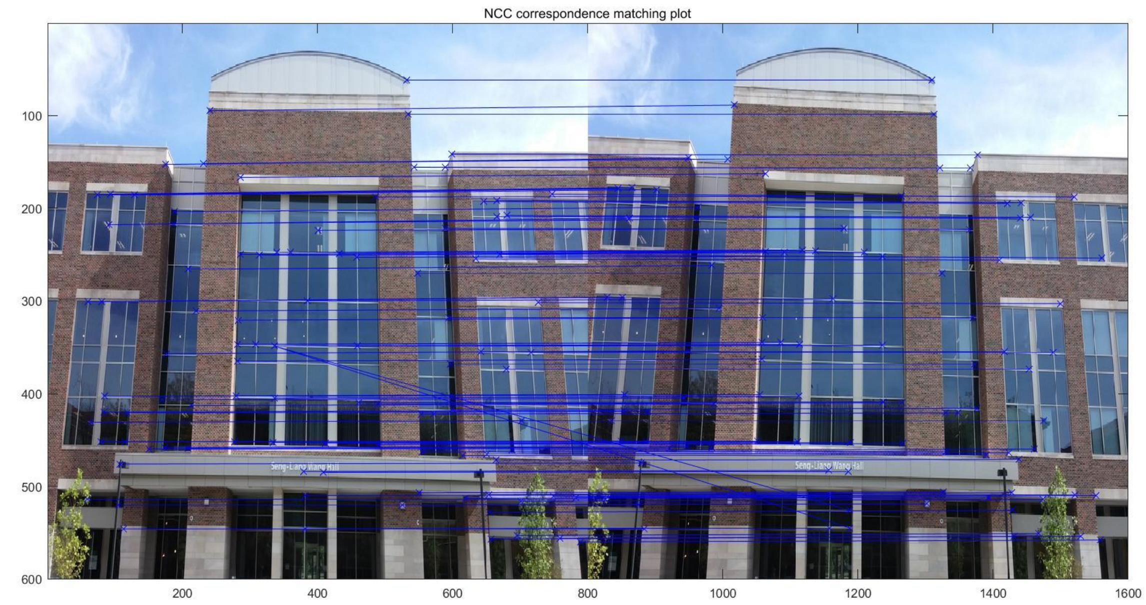


Figure 41. NCC at sigma=1.8

3.4 Harris corner points of pair3 images at sigma equals to 2.2

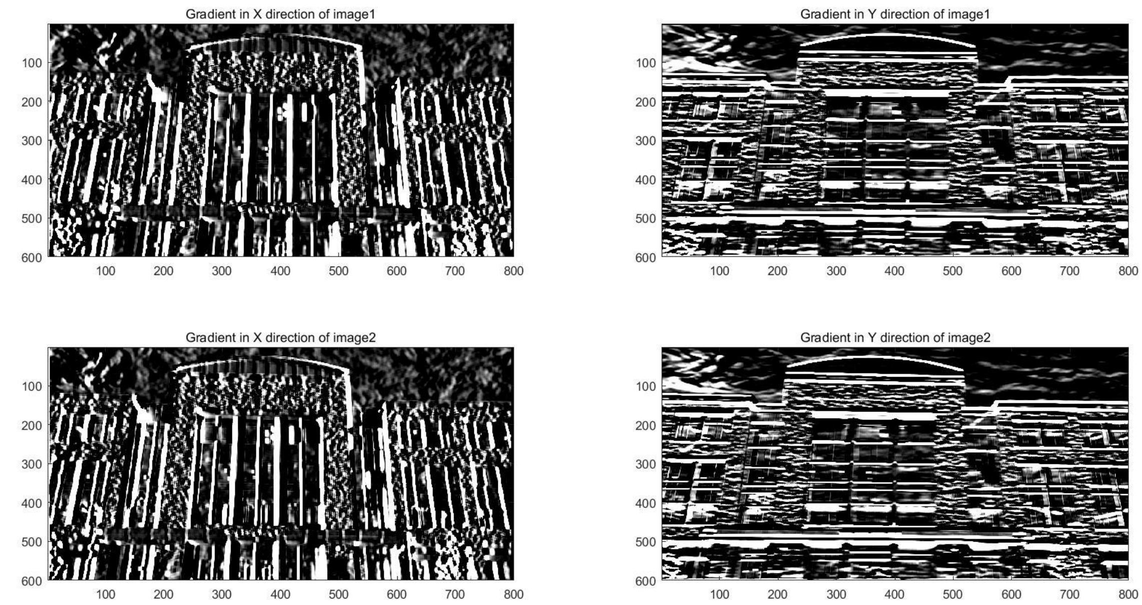


Figure 42. gradient at sigma 2.2

Figure 43. SSD at sigma=2.2

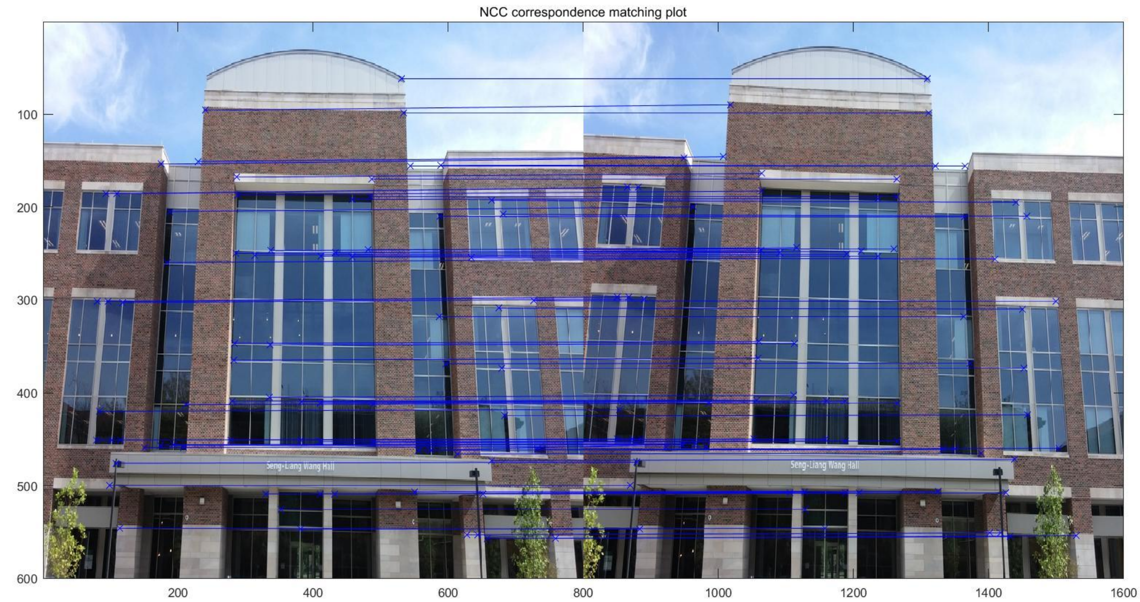


Figure 44. NCC at sigma=2.2

3.5 SIFT

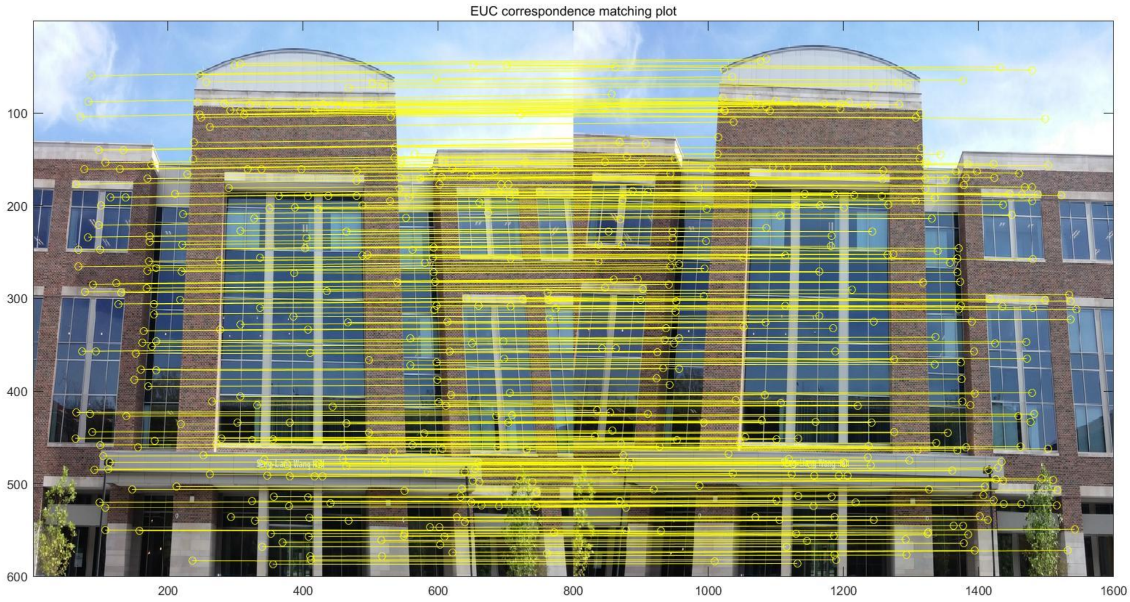
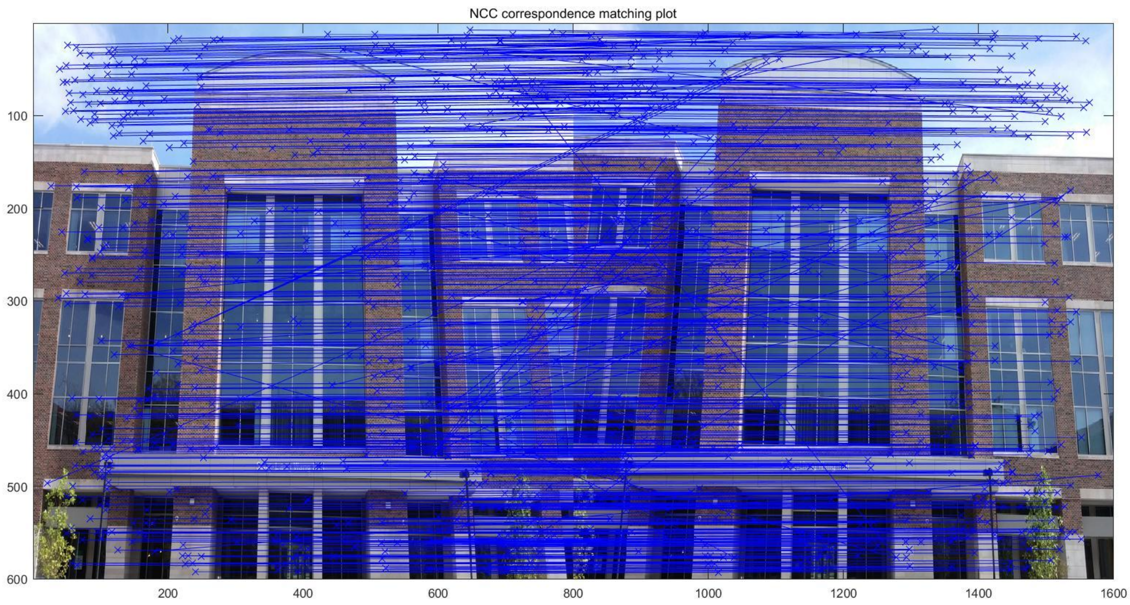


Figure 44. SIFT Euclidean of image pair 3

SIFT SSD of image pair 3



SIFT NCC of image pair 3

3.6 Conclusion:

For this set of images, the view angle doesn't change too much as well. For Harris corner detector, the larger the σ value, the less interest points will show up in the match result. For the method of SSD and NCC, NCC always contains more interesting points than SSD. And the overall accuracy of Harris corner is very high

As usual, SIFT detection result contains much more interesting pairs as well as making mistakes. Such as mismatch and false detections on the sky.

Figure 45. Original image 1

Figure 45. Original image 2

4.1.Harris corner detector at sigma 1

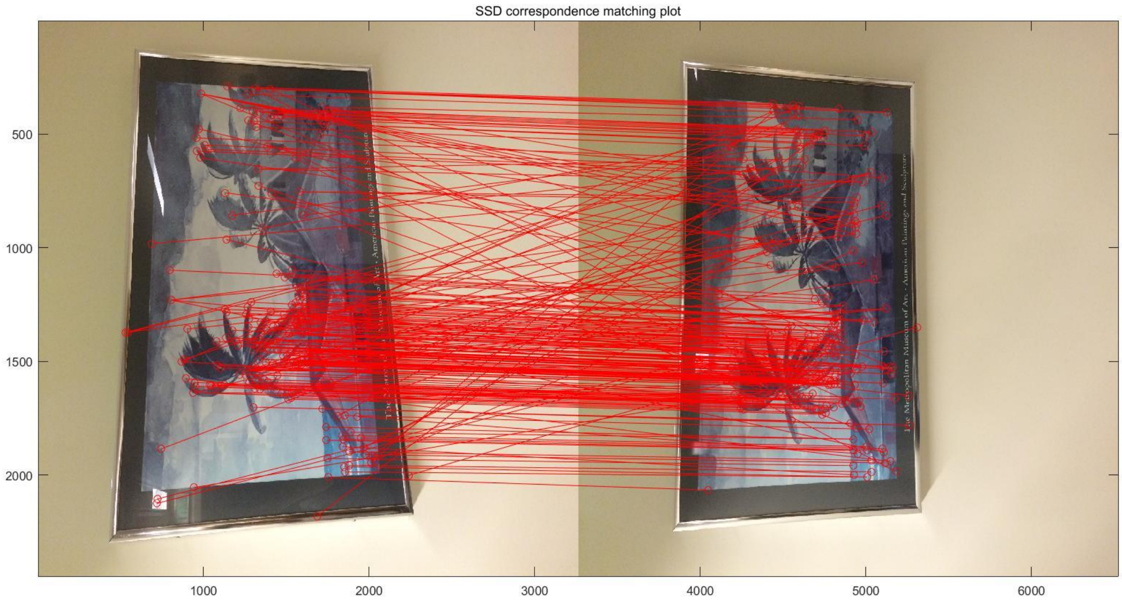


Figure 46. SSD at sigma = 1

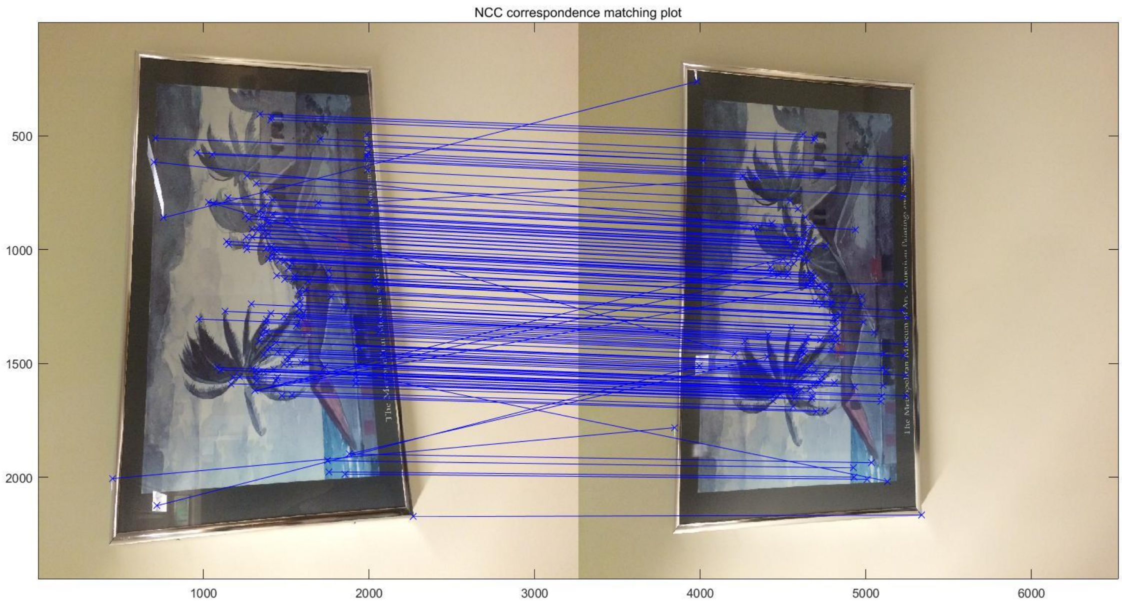


Figure 47. NCC at sigma = 1

4.2.Harris corner detector at sigma 1.4

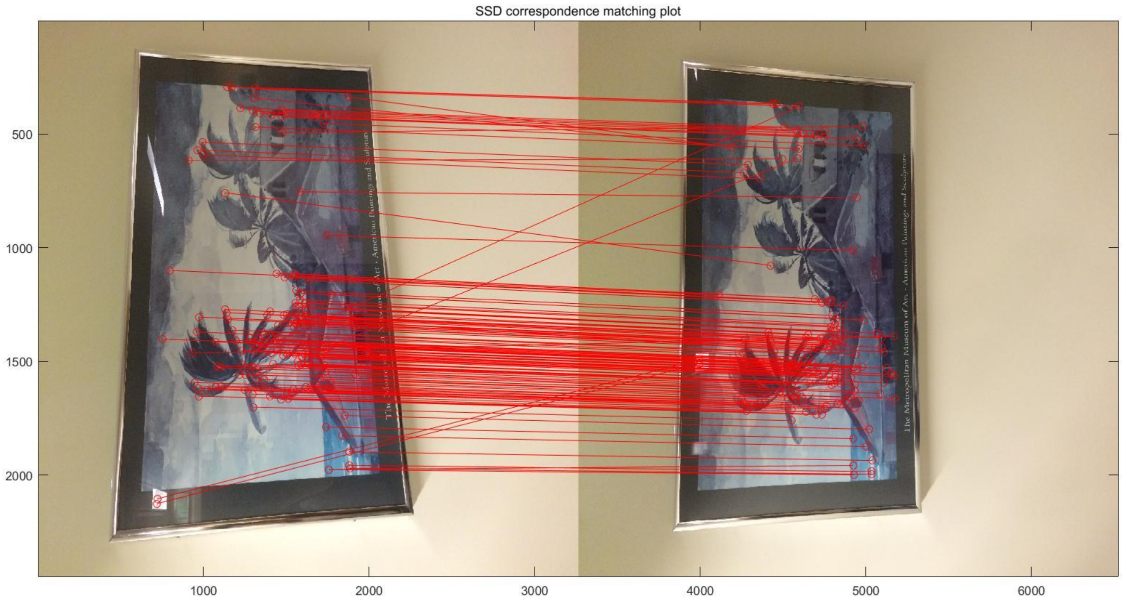


Figure 48. SSD at sigma = 1.4

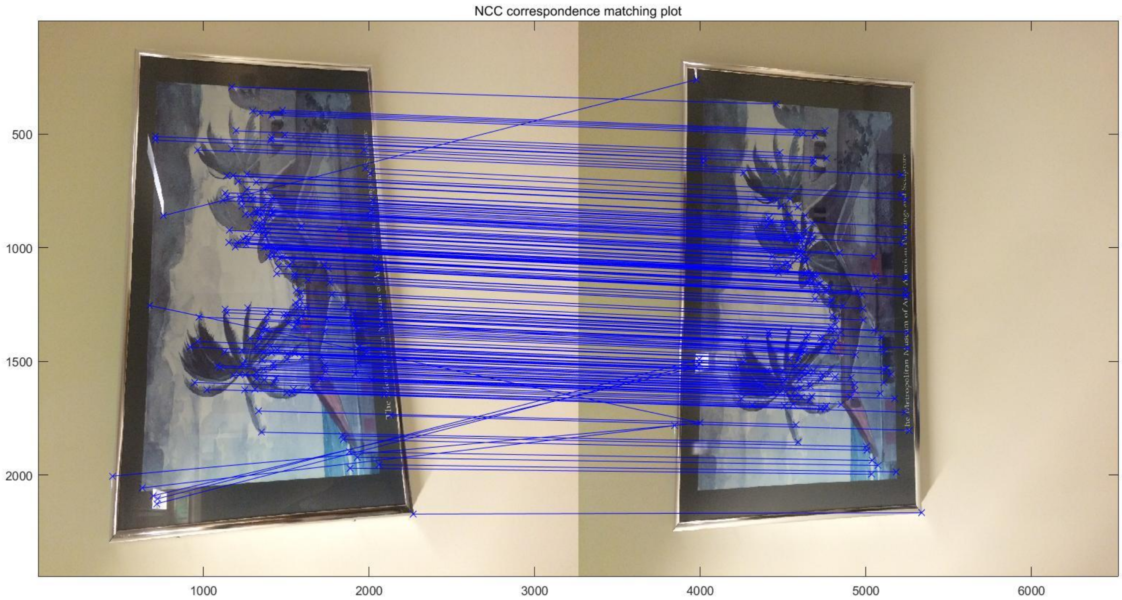


Figure 49. NCC at sigma = 1.4

4.3.Harris corner detector at sigma 1.8

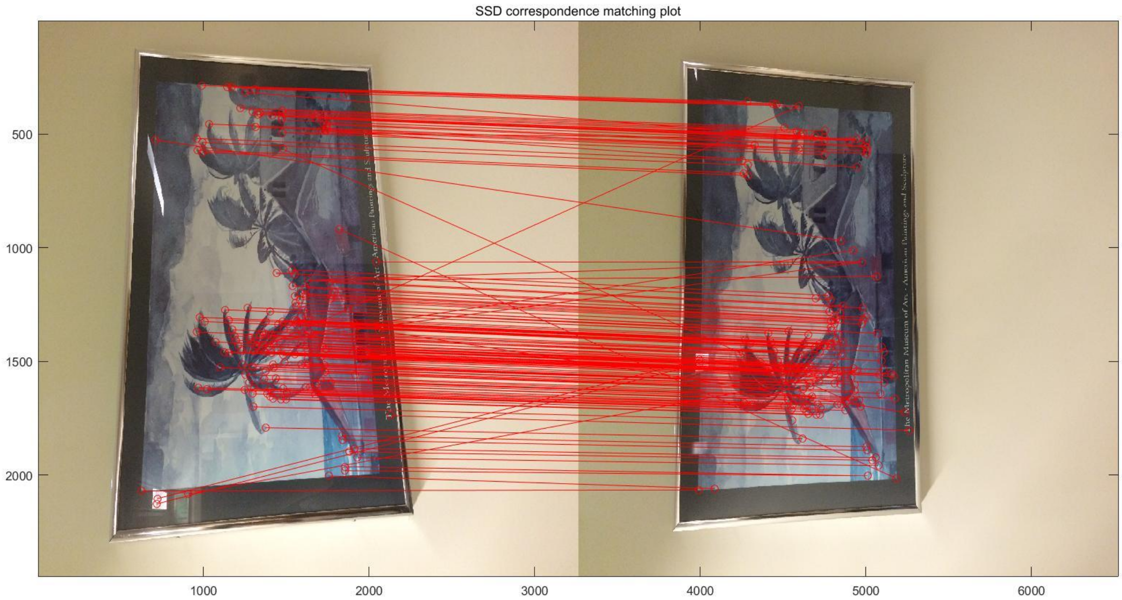


Figure 50. SSD at sigma = 1.8

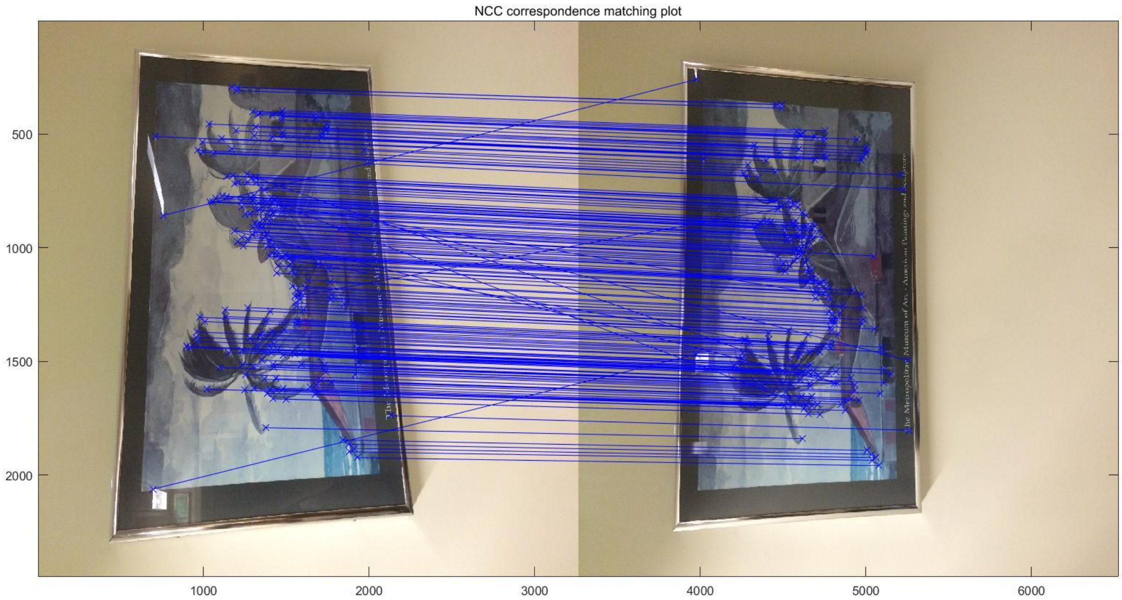


Figure 51. NCC at sigma = 1.8

4.4.Harris corner detector at sigma 2.2

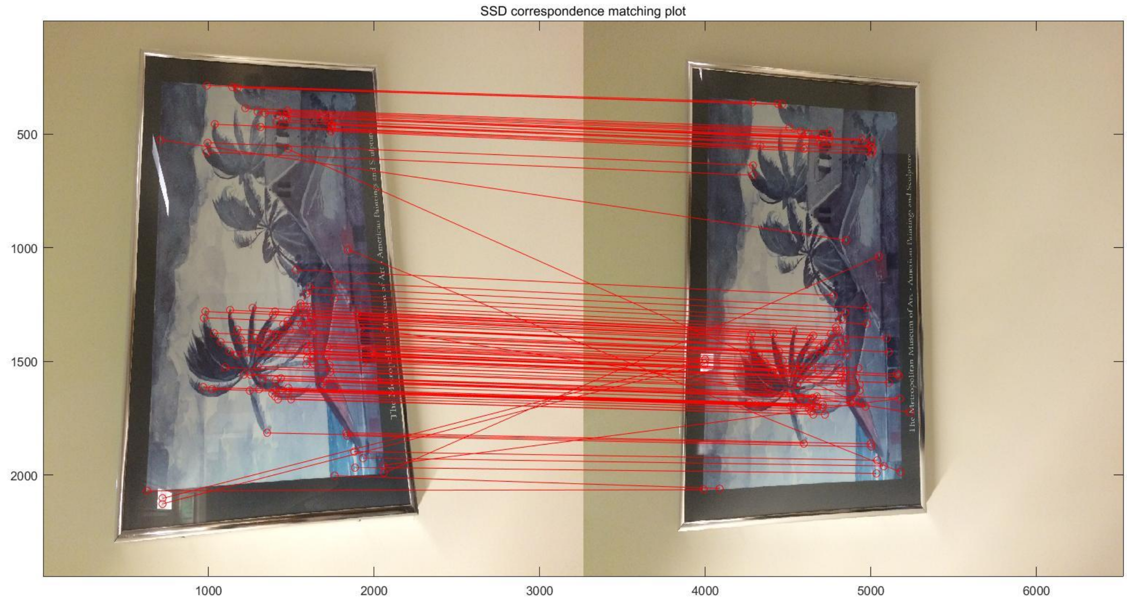


Figure 52. SSD at sigma = 2.2

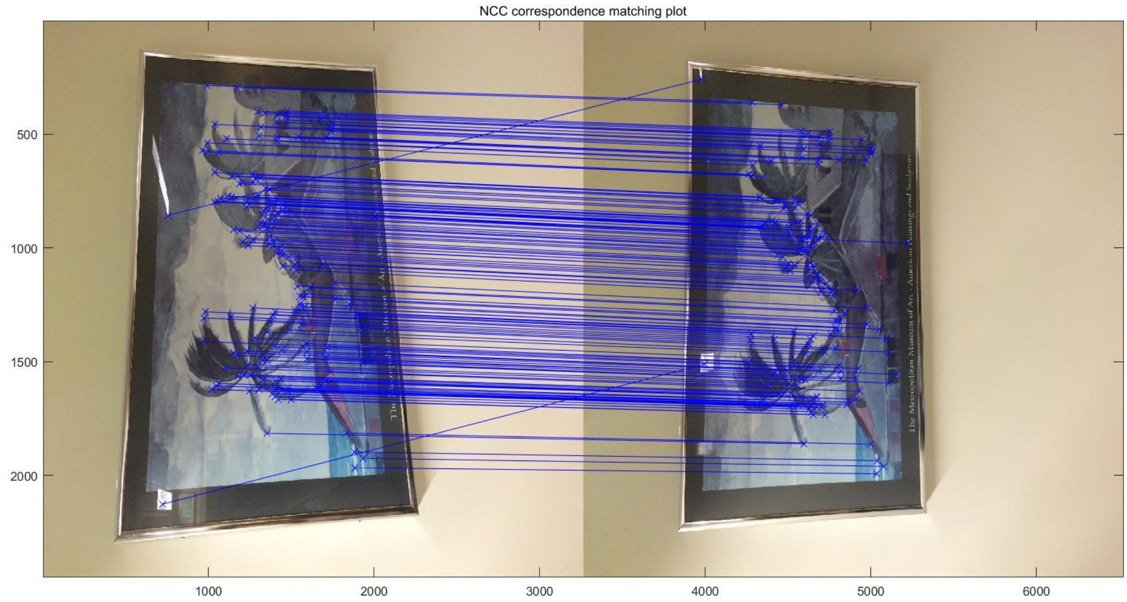


Figure 53. NCC at sigma = 2.2

4.5.SIFT

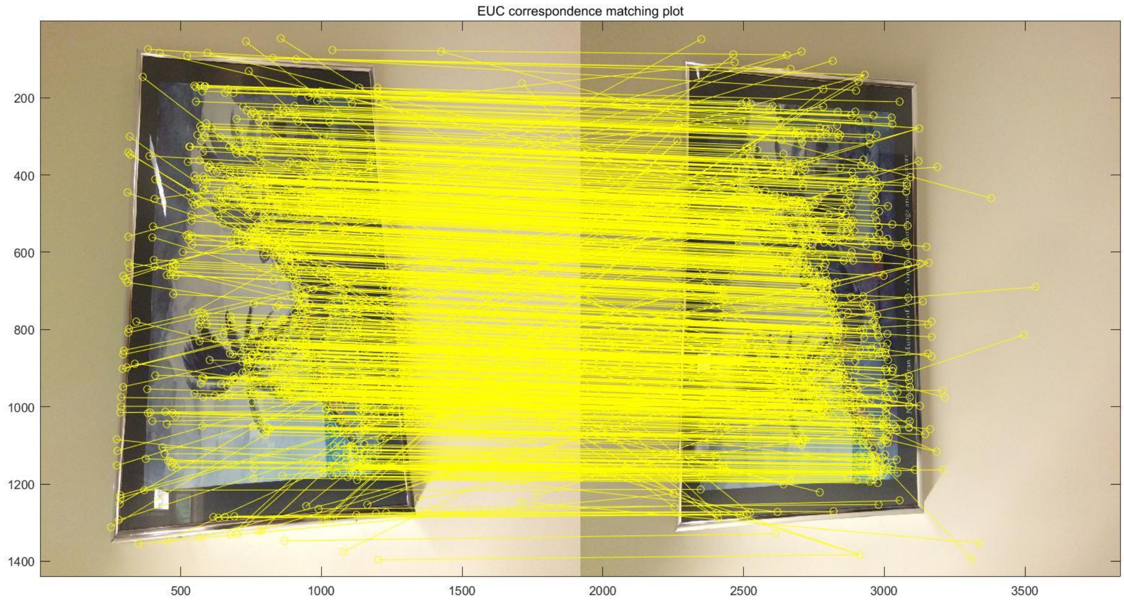
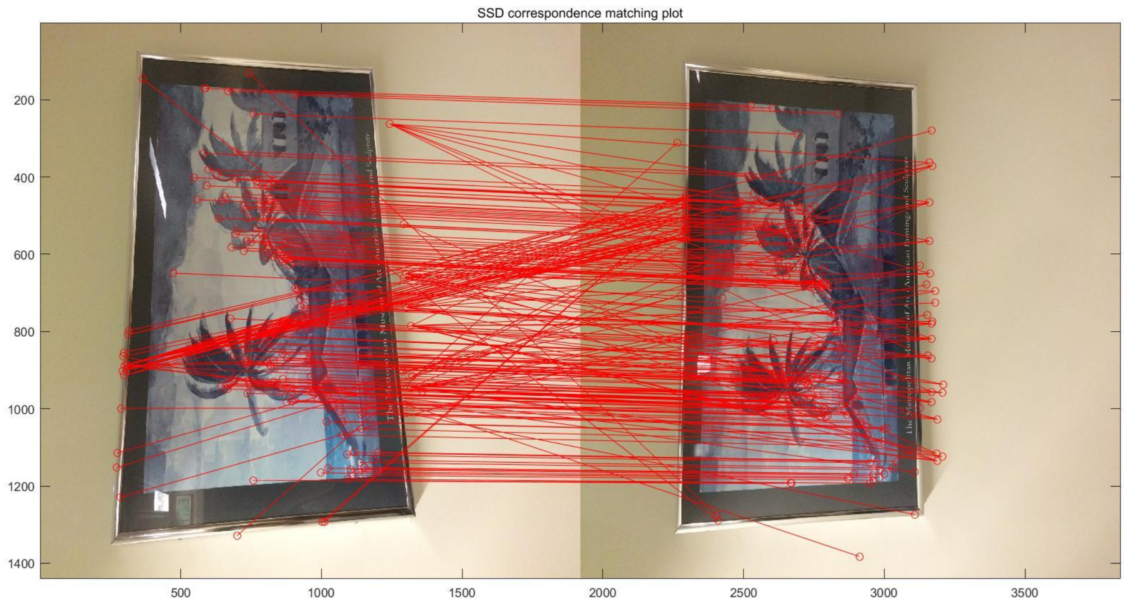
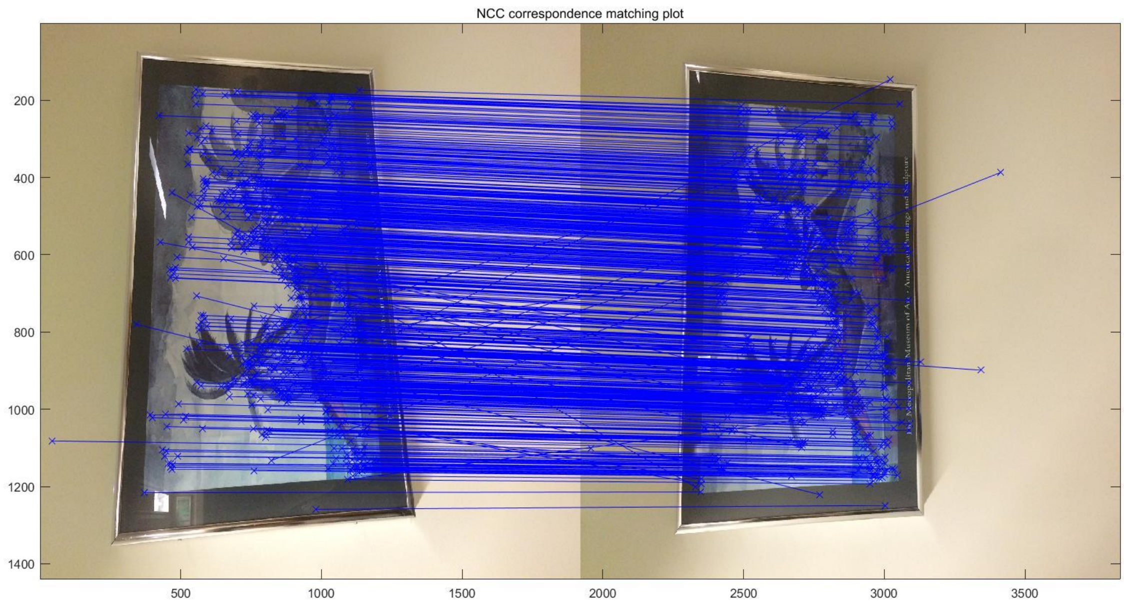


Figure 54. SIFT Euclidean result



SIFT SSD result



SIFT NCC result

4.6 Conclusion:

For my own set of image pairs, the viewing angle doesn't change too much as well. For Harris corner detector, the larger the σ value, the less interesting points will show up in the match result. For the method of SSD and NCC, both these two methods have similar performance. But NCC turns out to have more interesting points with less mistakes.

As usual, SIFT detection result contains much more interesting pairs as well as making mistakes.