CSC420 – Assignment 4

Ansh Chokshi

1004996917

UtorID - chokshi7

31th March 2023

In solving the questions in this assignment, I worked together with my classmate Uttkarsh Berwal & 1005355018. I confirm that I have written the solutions/code/report in my own words.

PART 1).

1.1} We know; $N = \frac{\log(1-P)}{\log(1-(1-e)^S)}$

Given that $1-e = 0.7$, $P = 0.995$ & $S = 4$

→ Here P is desired probability that we get in a good Sample.

→ S is number of points in a sample.

→ N is number of points in a sample.

→ e is probability that point is an outlier.

$\therefore \quad N = \frac{\log(1-0.995)}{\log(1-0.7^4)} = 19.29 \simeq 20$

. Therefore from the above calculations we need, 20 iterations for fitting a homography.

1.2} Affine only requires 3 points for matching, whereas Homography requires 4 points for matching.

Therefore Affine requires less number of inlier samples which means it requires fewer iteration overall for Affine transformation over homography.

From the equation in Part 1.1, we can check that $N = \frac{\log(1-0.995)}{\log(1-0.7^3)} \simeq 13$ [less than the iterations (20) required in 1.1]

2.1) Given $\vec{P} = \vec{P_0} + t\vec{d}$ are points on line L and

$$\vec{p} = \begin{pmatrix} wx \\ wy \\ w \end{pmatrix} = K\vec{P} = K \begin{pmatrix} x_0 + t d_x \\ y_0 + t d_y \\ z_0 + t d_z \end{pmatrix}$$

The above are fixed coordinates of the same line in the image, also

$$K = \begin{pmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{pmatrix},$$ where f is the camera focal length f (p_x, p_y) is principal point.

Next Step,

$$\vec{p} = K\vec{P} = \begin{pmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 + t d_x \\ y_0 + t d_y \\ z_0 + t d_z \end{pmatrix} = \begin{pmatrix} wx \\ wy \\ w \end{pmatrix}$$

$$= \begin{pmatrix} f x_0 + f t d_x + P_x z_0 + P_x + d_z \\ f y_0 + f t d_y + P_y z_0 + P_y + d_z \\ z_0 + t d_z \end{pmatrix} = \begin{pmatrix} wx \\ wy \\ w \end{pmatrix}$$

lets divide w on both sides, where $w = z_0 + t d_z$

$$\begin{pmatrix} (f x_0 + f t d_x + P_x z_0 + P_x t d_z)/w \\ (f y_0 + f t d_y + P_y z_0 + P_y t d_z)/w \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Now, lets move infinitely far from the camera by taking $t \to \infty$ and computing x & y.

$$x = \lim_{t \to \infty} \frac{f x + f t d_x + P_x z_0 + P_x t d_z}{z_0 + t d_z} = \boxed{\frac{f d_x + P_x}{d_x}}$$

Similarly $y = f \frac{d_y}{d_z} + P_y$.

Therefore, we conclude that the vanishing point is (x, y)

Scanned with CamScanner

2.2} Given $\vec{n}\cdot\vec{d} = 0$, $n_x d_x + n_y d_y + n_z d_z = 0$

We know the vanishing point is (x, y) from part 2.1}

$$\therefore (x, y) = \left(\frac{f d_x}{d_z} + P_x, \frac{f d_y}{d_z} + P_y \right) \quad \text{— (1)}$$

lets substitute $d_x = \left(\frac{n_y d_y + n_z d_z}{n_x} \right)$

in above equation (1)

We get, $V_x = \frac{f d_x}{d_z} + P_x = -\frac{f (n_y d_y + n_z d_z)}{n_x d_z} + P_x$

$$= -\frac{f n_y d_y}{n_x d_z} - \frac{f n_z}{n_x} + P_x$$

To show a linear relationship, i.e. a line, we also need V_y. Hence we will add and substract $P_y \frac{n_y}{n_x}$.

$$\Rightarrow V_x = \frac{-f n_y d_y}{n_x d_z} - P_y \frac{n_y}{n_x} + P_y \frac{n_y}{n_x} - \frac{f n_z}{n_x} + P_x$$

$$= -\left(\frac{f d_y}{d_z} + P_y \right) \frac{n_y}{d_x} + P_y \frac{n_y}{n_x} - \frac{f n_z}{n_x} + \frac{P_x n_x}{n_x}$$

$$= -V_y \frac{n_y}{n_x} + P_y \frac{n_y}{n_x} - \frac{f n_z}{n_x} + \frac{P_x n_x}{n_x}$$

$$\Rightarrow V_x n_x + V_y n_y = n_x P_x + P_y n_y - f n_z.$$

$\therefore$ The above takes the form of a line where V_x & V_y are variables for which the values follow this line.

3.1} To prove : intersection of the 2D line l & l' is the 2D point $p = l \times l'$

Let the equations of lines be -

① $l : a_1 x + b_1 y + c_1 = 0$

② $l' : a_2 x + b_2 y + c_2 = 0$

In order to find the intersection of the above two lines, we have to solve the simultaneous equations above. We can do that by Isolating x from the equation of line l and substituting it in line l'.

$$x = \frac{-b_1 y - c_1}{a} \quad — ① \quad [\text{from } l_1]$$

$$\Rightarrow l' = \frac{-a_2 (b_1 y + c_1)}{a_1} + b_2 y + c_2 = 0$$

$$\Rightarrow \frac{-a_2}{a_1} b_1 y - \frac{a_2 c_1}{a_1} + b_2 y + c_2 = 0$$

$$\Rightarrow y \left(b_2 - \frac{a_2}{a_1} b_1 \right) = \frac{a_2 c_1}{a_1} - c_2$$

$$\Rightarrow y \left(\frac{a_1 b_2 - a_2 b_1}{a_1} \right) = \frac{a_2 c_1 - a_1 c_2}{a_1}$$

$$y = \frac{a_2 c_1 - a_1 c_2}{a_1 b_2 - a_2 b_1}$$

[Substitute this value of y in eq ① to find x]

Scanned with CamScanner

$$\Rightarrow x = \frac{-b_1\left(\frac{a_2c_1 - a_1c_2}{a_1b_2 - a_2b_1}\right) - c_1}{a_1}$$

$$x = \frac{-b_1\cancel{a_2}c_1 + b_1a_1c_2 - c_1a_1b_2 + c_1\cancel{a_2}b_1}{a_1(a_1b_2 - a_2b_1)}$$

$$= \frac{\cancel{a_1}}{\cancel{a_1}}\left(\frac{b_1c_2 - c_1b_2}{a_1b_2 - a_2b_1}\right) \quad \text{--- } (2)$$

Now, we have the values of both x & y. Hence we can write them as homogenous coordinates.

$$P = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} (b_1c_2 - b_2c_1)/(a_1b_2 - a_2b_1) \\ (a_2(1 - a_1c_2)/(a_1b_2 - a_2b_1) \end{bmatrix}$$

We know, scaling doesn't affect homogenous coordinates so we scale by $w = a_1b_2 - a_2b_1$

$$= \begin{bmatrix} x \cdot w \\ y \cdot w \\ w \end{bmatrix} = \begin{bmatrix} b_1c_2 - b_2c_1 \\ c_1a_2 - c_2a_1 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

Next step, $l \times l'$ $\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$

$$= \begin{bmatrix} b_1c_2 - b_2c_1 \\ c_1a_2 - c_2a_1 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

[equal]

Therefore, we prove, $p = l \times l'$

Scanned with CamScanner

3.2} To prove : Line that goes through the 2D points p & p' is $l = p \times p'$

We can write $p = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ & $p' = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$

By point slope format, $y - y_1 = m(x - x_1)$ is the equation of line passing through x_1, y_1

$m = \dfrac{y_2 - y_1}{x_2 - x_1}$ where x_2, y_2 is another point on the line.

$$\therefore (y - y_1) = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

$$\rightarrow (y - y_1)(x_2 - x_1) = (y_2 - y_1)(x - x_1)$$

$$\rightarrow y(x_2 - x_1) - y_1 x_2 + y_1 x_1 = (y_2 y_1) x - x_1 y_2 + x_1 y_1$$

$$\rightarrow y(x_2 - x_1) - (-(y_1 - y_2))x + x_1 y_2 - y_1 x_2 = 0$$

$$\rightarrow x(y_1 - y_2) + y(x_2 - x_1) + x_1 y_2 - y_1 x_2 = 0$$

$$l = \begin{bmatrix} y_1 - y_2 \\ x_2 - x_1 \\ x_1 y_2 - y_1 x_2 \end{bmatrix}$$

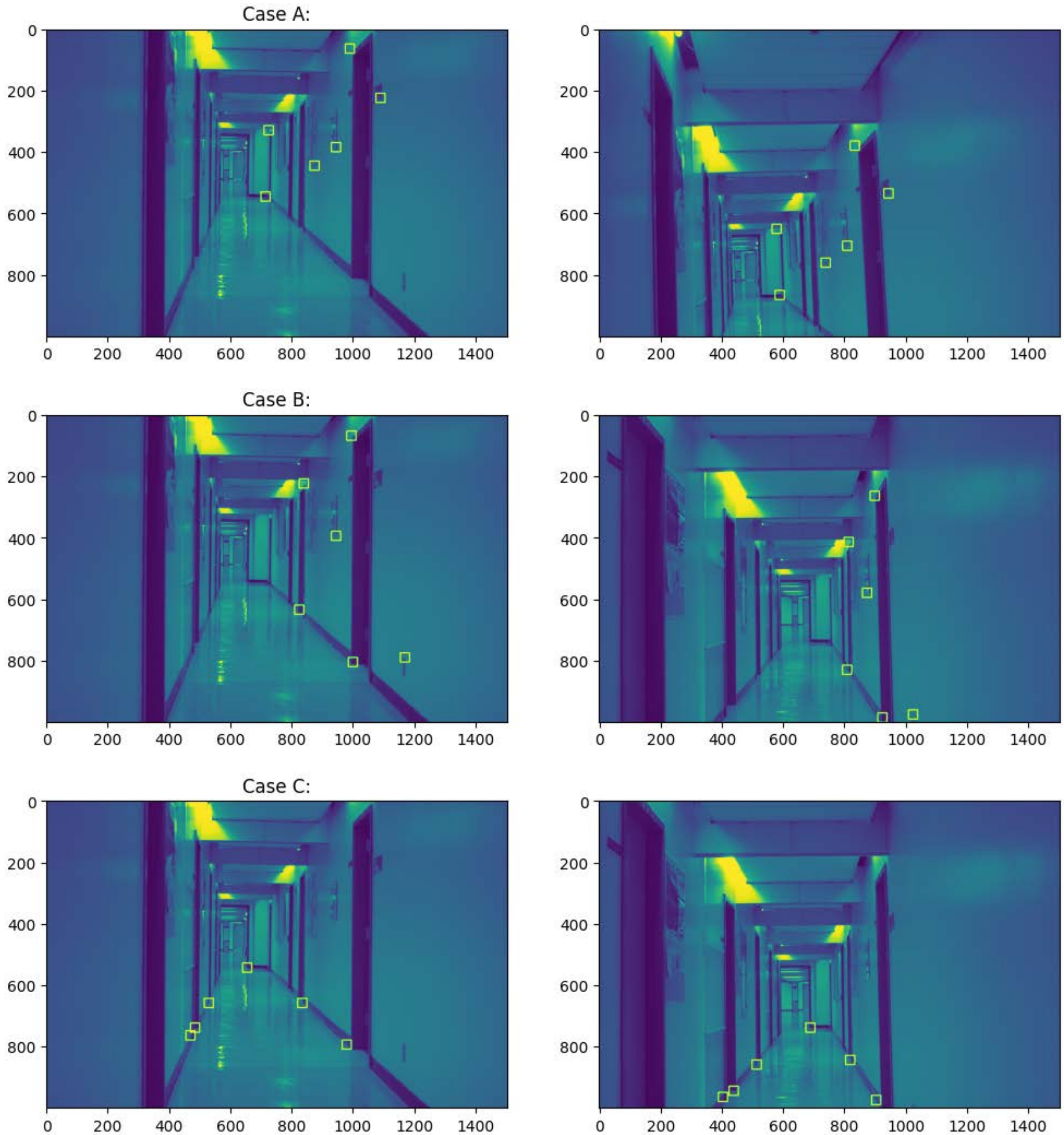
[equal]

Now, $p \times p'$ in homogenous coordinates.

$$p \times p' = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \times \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ x_2 - x_1 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} = l$$

Hence Proved !

Question 4.1 -


Case A:
Case B:
Case C:
0
200
400
600
800
0
200
400
600
800
1000
1200
1400

Question 4.2 -

Homography matrix for case A:

```
[[-2.55162336e-03 -2.15961103e-04  4.20141847e-01]
 [ 1.12669728e-04 -2.58950415e-03 -9.07447361e-01]
 [ 7.40639866e-08 -5.05755856e-08 -2.61521911e-03]]
```

This homography matrix includes a slight amount of rotation and translation. It rotates the right wall of the first image in a counterclockwise direction and moves it upward to match with the right wall of the second image.

Homography matrix for case B:

```
[[-9.05841959e-04  7.22670768e-05 -7.96760626e-01]
 [ 2.33949336e-04 -1.74077697e-03 -6.04288193e-01]
 [ 1.99978582e-07  1.30867768e-07 -2.09125592e-03]]
```

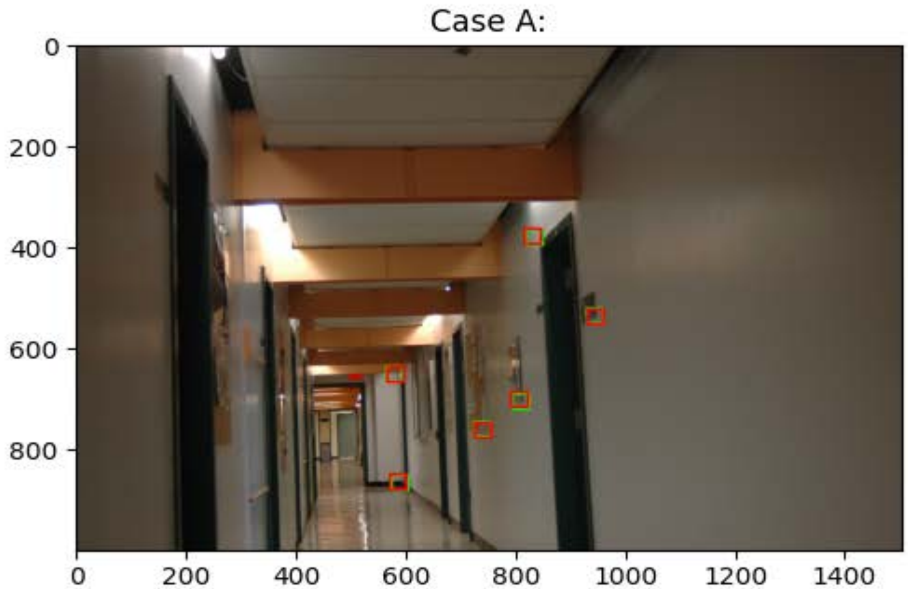
This matrix involves a significant amount of translation and a small amount of shear. The matrix translates the left wall of image 1 horizontally and upwards to align with the left wall of image 2.

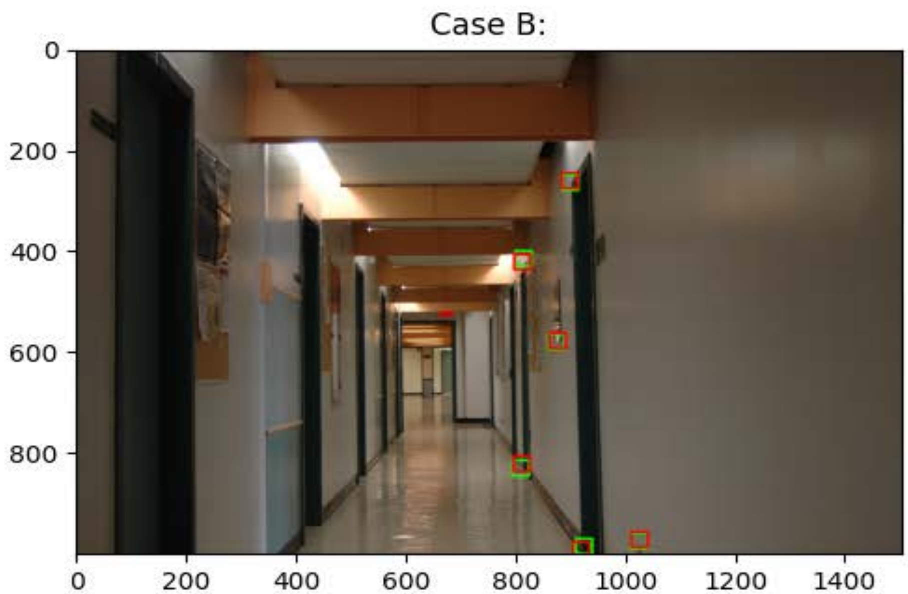
Homography matrix for case C:

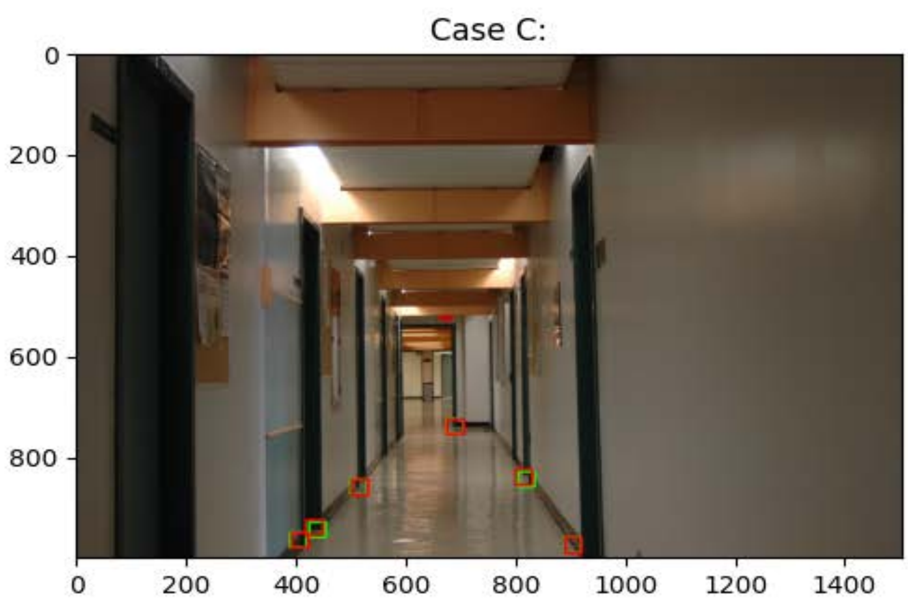
```
[[ 1.44824625e-03 -9.02148514e-04  7.38069325e-01]
 [-4.13251440e-04  1.61713178e-03  6.74717582e-01]
 [-3.61091219e-07 -1.21780345e-07  2.03940105e-03]]
```

This matrix involves a small amount of shear and translation. The matrix shears the top and bottom of image 1 slightly and translates it to align with the corresponding parts of image 2.

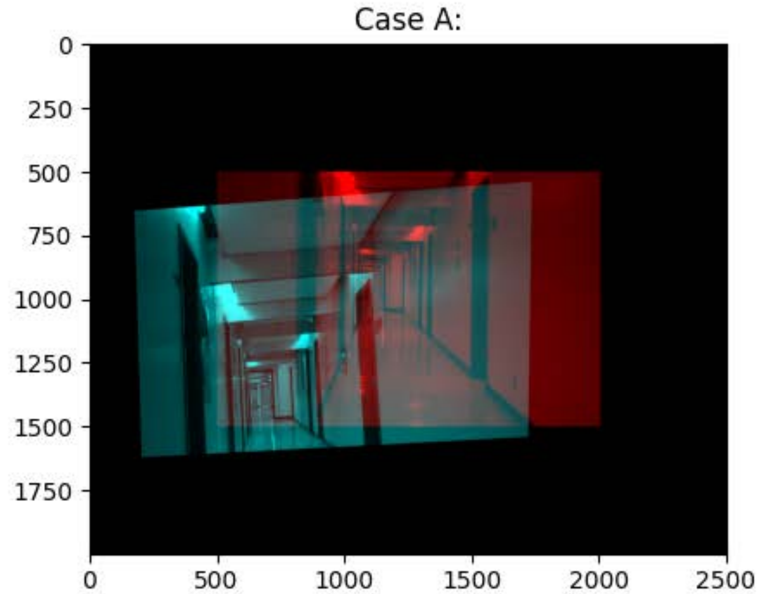
Question 4.3 -

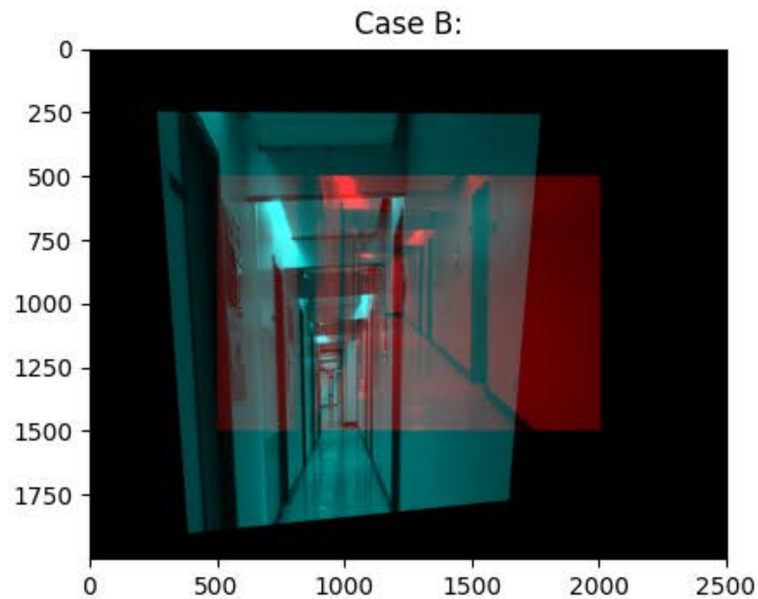

Case A:
0
200
400
600
800
0
200
400
600
800
1000
1200
1400

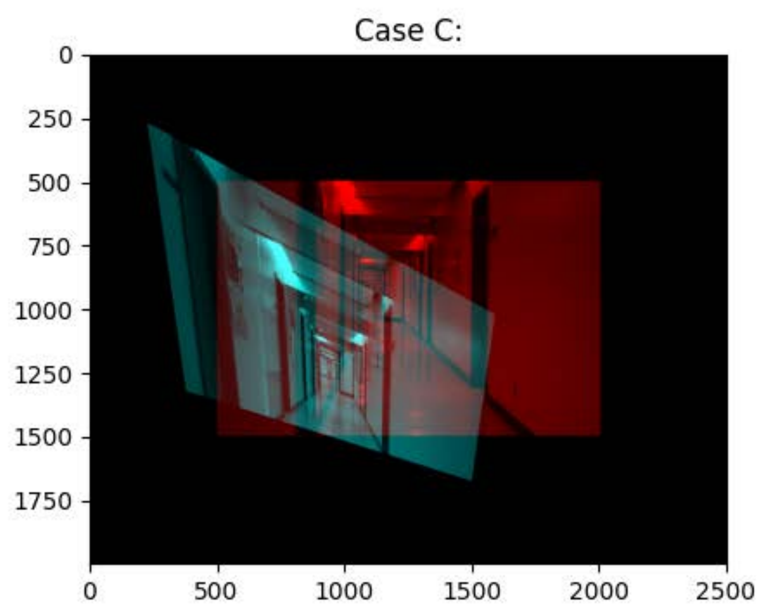

Case B:
0
200
400
600
800
0
200
400
600
800
1000
1200
1400


Case C:
0
200
400
600
800
0
200
400
600
800
1000
1200
1400

Question 4.4 -


Case A:
0
250
500
750
1000
1250
1500
1750
0
500
1000
1500
2000
2500


Case B:
0
250
500
750
1000
1250
1500
1750
0
500
1000
1500
2000
2500


Case C:
0
250
500
750
1000
1250
1500
1750
0
500
1000
1500
2000
2500

Case A : The blue and green pixels are rotated counterclockwise compared to the red pixels, indicating that the camera was rotated counterclockwise in image 2 compared to image 1.

Case B: The blue-green pixels are proportionally larger than the red pixels, indicating that image 2 is a magnified version of image 1, and scaling transformation has been applied to the image.

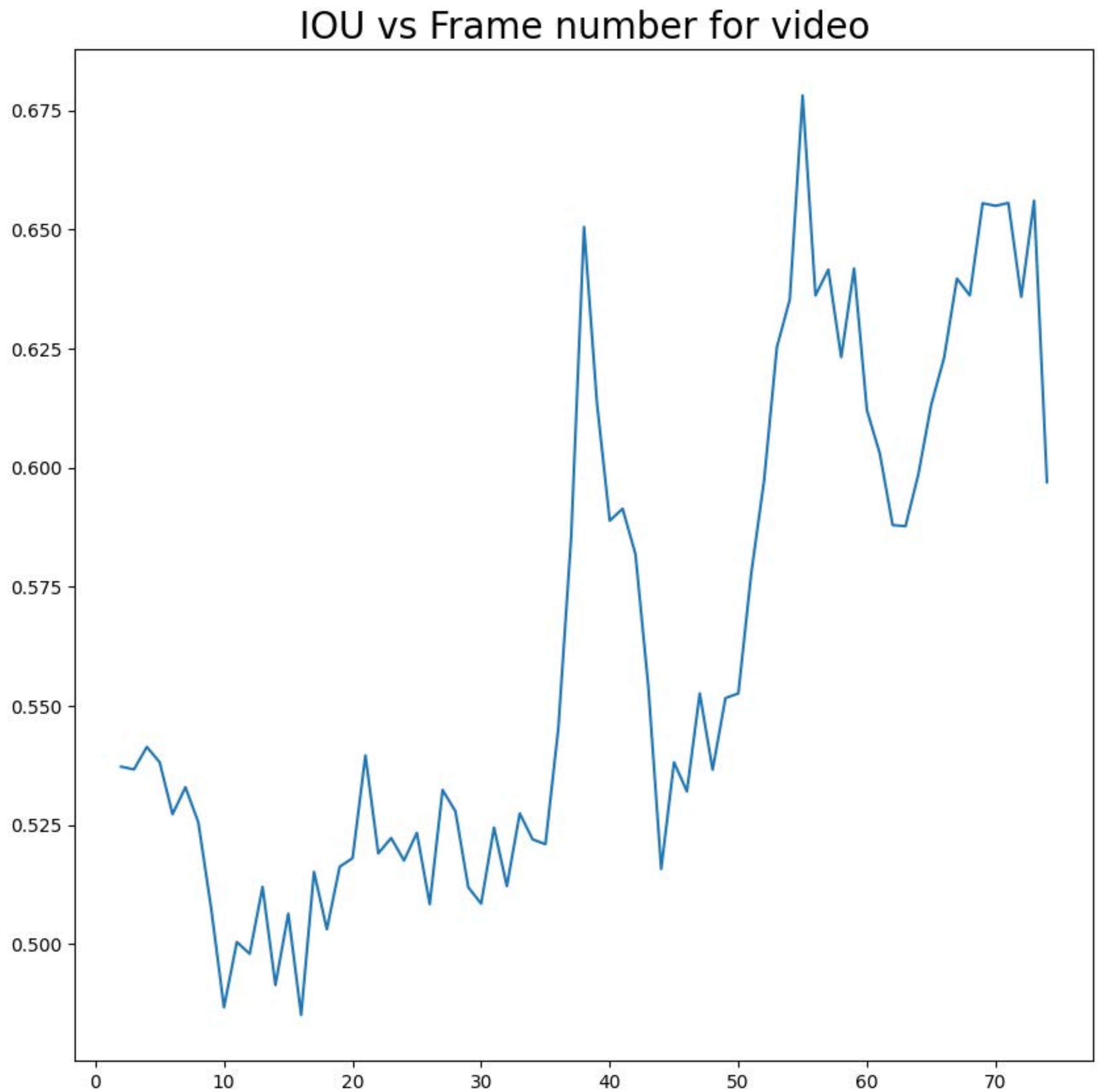
Case C: The blue-green image displays evidence of a shear transformation, demonstrated by the shift in viewing point from left to right and the change in the camera's direction and orientation.

The floor and wall exhibit different light-reflecting properties, with the floor reflecting more light than the wall. Specifically, the tiles that make up the floor reflect a significant amount of light due to their glossy surface, resulting in specular reflections. On the other hand, the right wall appears more Lambertian, indicating a less glossy and more diffuse surface that reflects light in multiple directions.

[The code is provided in the end of Q4.ipynb]

Question 5.1 -

Above is the detected face on the first frame using Viola-Jones detector.



Above is the IOU over-time from the 2nd frame to the last.

Red - Tracked bounding box
Green - Detected bounding box

- The above is the larger IOU (iou> 0.65). It is evident that there is a significant amount of similarity between the two boxes, implying that the Viola Jones detector is proficient at recognizing the face.

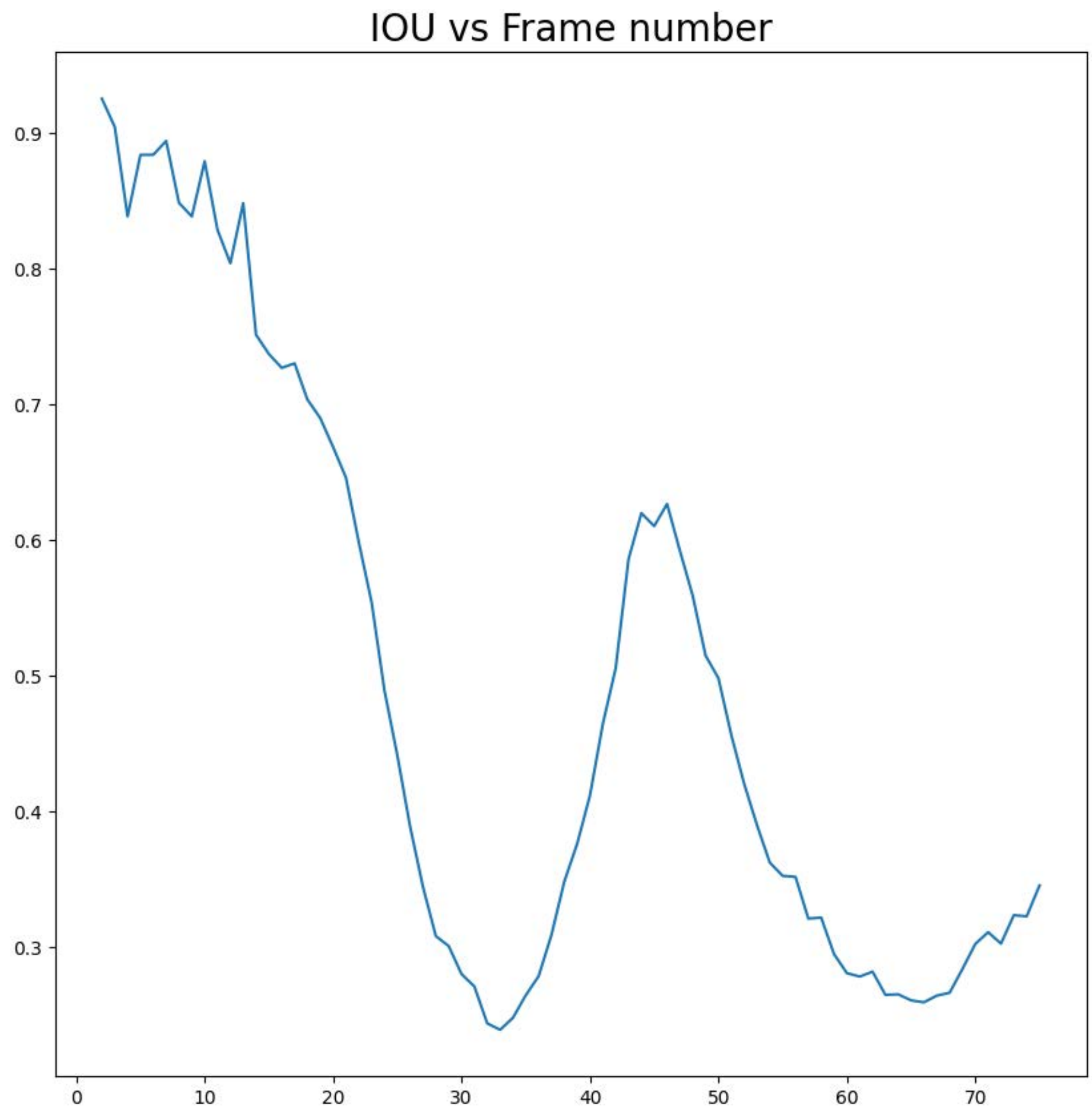
- The above is the smaller IOU (iou< 0.5). It is evident that there is low variance in this method. The lowest IOU is 0.48 and highest is around 0.78.

Percentage of frames in which the IoU is larger than 50% - 94.5%

There are several factors that could contribute to the mean shift algorithm's inaccurate detection of Mbappe's face in certain frames. For example, variations in lighting, shadows, and occlusion could all affect the algorithm's ability to accurately track the face. Additionally, the mean shift algorithm's reliance on color histograms for object tracking may not be sufficient in cases where the object of interest, in this case, Mbappe's face, has a similar color or hue to the background.

Viola Jones detector outperforms the mean shift algorithm in detecting Mbappe's face in the video frames, even in cases where the IOU is relatively small.

Question 5.2 -



Above is the IOU over-time from the 2nd frame to the last using the second method.

Red - Tracked bounding box
Green - Detected bounding box

- The samples presented above demonstrate the effectiveness of the method in producing high levels of accuracy for both the detected and tracked bounding boxes, as evidenced by the significantly large intersection over union (IOU) values.

- One limitation of this method is its high variance, which can result in inaccuracies in the detected and tracked bounding boxes, as demonstrated in the examples above where the intersection over union (IOU) values were very low.

Percentage of frames in which the IoU is larger than 50% - 40.5%