Questions

5 points each

1. Explain why the specific projective model of Eq.(1) is representative of the imaging

geometry of Figure-1?

Equation (1) represents the transformation from homogeneous coordinates in 2D (after undistortion) to 3D world coordinates. This transformation is used to convert pixel coordinates in the undistorted 2D image to corresponding 3D world coordinates in the scene. Since the camera is placed above the box and world coordinates and camera coordinates are aligned, there is no rotation or translation between the camera coordinates and world coordinates, the transformation matrix have a simplified form, with the rotation and translation components being identity matrices. The ToF camera works like a projector and camera together and provides depth information, and the transformation from 2D to 3D involves incorporating this depth information into the world coordinates.

2. How many peaks do you see in the histogram of depths within the region-of-interest?

At first, I saw four peaks, but I should have seen only two peaks because of the conveyor and box. The other two peaks were not as frequent; they were local maxima, probably caused by reflection or image noise. Therefore, I considered the more frequent peaks from the conveyor and box.

3. Does the histogram represent a bimodal distribution?

Yes. I can observe two clear peaks or modes, that represents a bimodal distribution.

4. What do each of the peaks tell you about the depth Zw of the box and the depth of

the conveyor belt from the camera?

The presence of two distinct peaks suggests that there are two primary depths in the scene—one for the box and one for the conveyor belt. The x-axis of histogram shows the depth, and the y-axis shows the frequency. The box is closer to the camera so it can be seen as a lighter color in the depth map. Also as the conveyer is larger than box we can see the frequency of it is greater. This histogram is a bimodal distribution and can be indicative of the separation or difference in depth between the two objects.

5. What is your estimate of Zw for points on the box and points on the conveyor belt

(express in centimeters)?

My estimation of depth from camera for box is 91.42493 cm and for conveyor belt is 99.339926 cm

6. What is your estimate of the height of the box (express in centimeters)?

The height of the box is achieved by difference of depths from camera and conveyor belt which is eq
ual
to 7.914996 cm

7. What is your estimate of the length of the box (express in centimeters)?

My estimate of the length of the box is 35.395010259639176 cm

8. What is your estimate of the width of the box (express in centimeters)?

My estimate of the width of the box is 28.16338799403142 cm

Provide a detailed explanation of your strategy for identifying the corners of the box, and subsequently the dimensions of the box. Clearly document what parameter values were chosen when invoking built-in functions such as histogram() or any other function you choose to use. The practice is intended to help engineers at Roadrunner Logistics build upon your findings.

I used depth map for finding the depth and corner detections. In Figure (1) it can be seen that the image is distorted and to exactly find the coordinates we need to at first undistort the images.

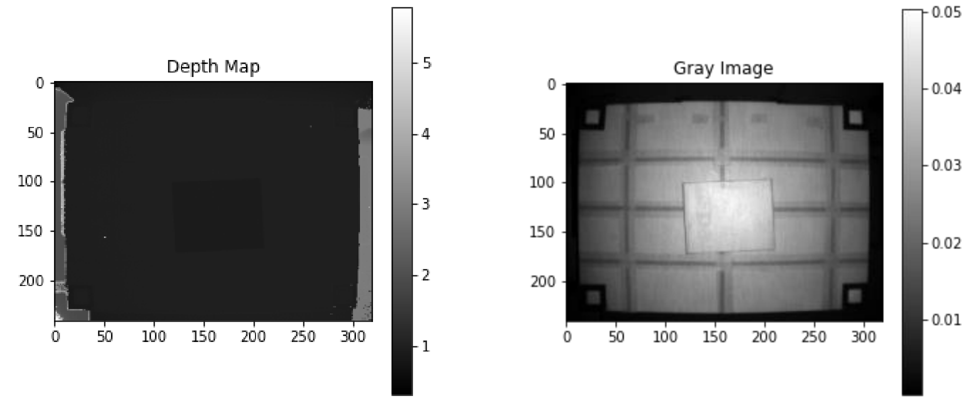

Figure 1: distorted images

I employed the camera matrix intrinsics, along with tangential and radial distortion coefficients, to perform the process of undistorting and remapping the box image. This procedure effectively eliminates distortions caused by the camera lens. The result can be seen in figure 2&3.

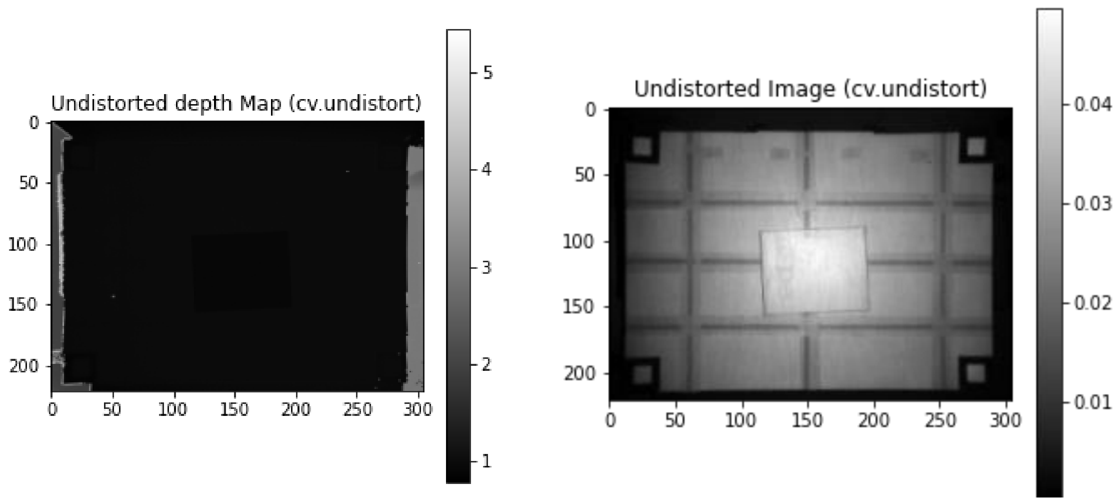

Figure 2: Undistort

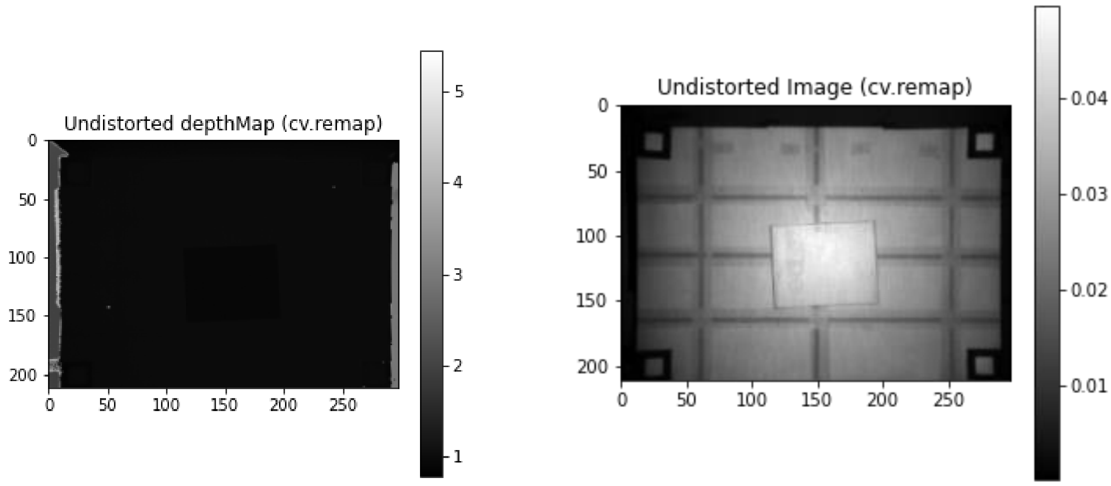

Figure 3: remap

Then we cropped the image to focus on the part we are interested in (Figure 4)

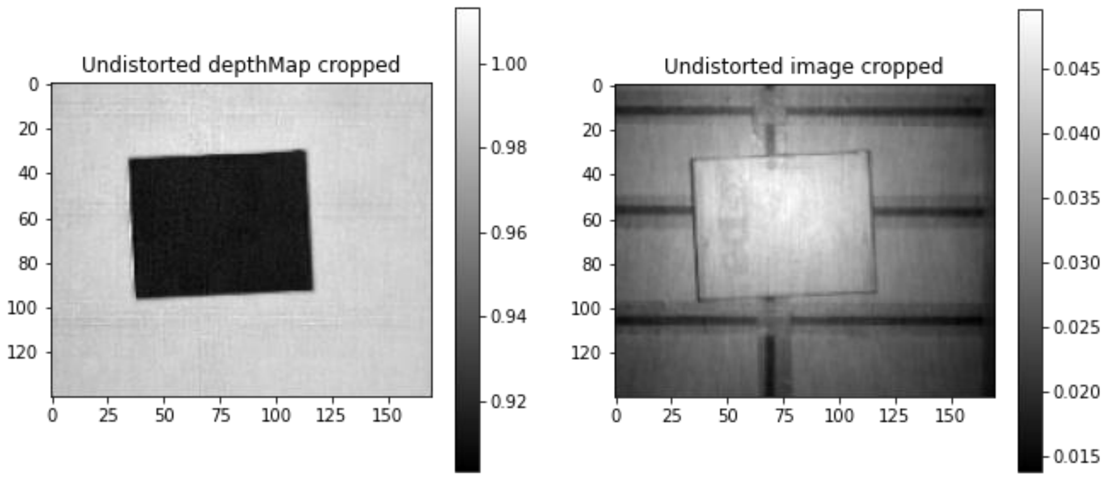

Figure 4: cropped to focus on the part we are interested in

Now I calculate the histogram and local maxima function to find the depth. As we have box and conveyor and each of them have uniform depth in the image, we expect to have 2 major peaks (Figure 5).

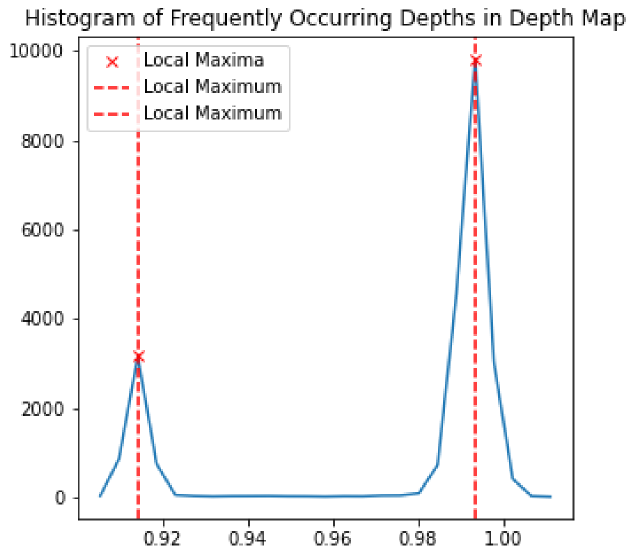

Figure 5: the depth histogram, the x-axis shows the depth and the y-axis shows the frequency

Based on our histogram, we can see that the points [0.9142493, 0.99339926] have the most frequency in our histogram which is implied that they are related to box and conveyor. As the box is nearer to the camera and its area is less than the conveyor (less frequent) we can understand that the first peak belonges to the box and the second one belongs to the conveyor. In this way we calculate the depth and if we differentiate them I can find the height of box (7.914996 cm).

I now apply corners, lines and edge detection strategies to find the corners. At first, I applied a gaussi an filter to remove the noise and have a smoother image. Then I normalized the image and got Harris corner detection and with thresholding I removed the corners that does not have so many votes. I fin d this points as my corners:(113, 31),(112, 32),(35, 35),(36, 35),(115, 92),(116, 92),(39, 95), (39, 96) a s it can be seen, every two points of the above coordinates are very close to each other (Figure 6)

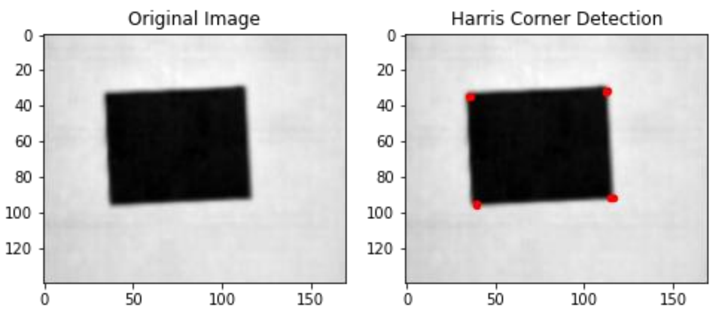

Figure 6: Harris Corner detection

I also tried canny edge detections and Hough line transform with non-maxima suppression and I got:

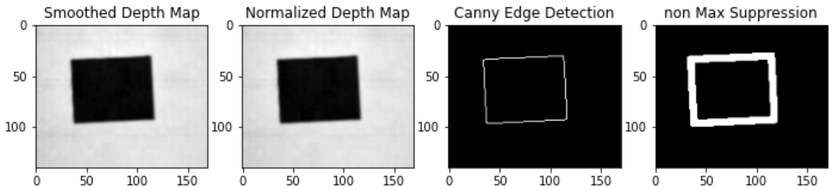

Figure 7: canny edge detection and Hough line transform wih non-maximum suppression.

As the canny and Hough line transform have noises and they generally give us edges information we h ave more candidates in this case. Therefore, I preferred to continue using the corner candidates achie ved from Harris corner detection. Then I calculate the world coordinates for every corner candidate t hat I had using equation 1 and camera matrix for undistorted image and then I calculate the distance between corners and remove distances less than 0.01 meter and the greatest distances which were t he diameter of rectangular. Then I got the average of two remaining series and found the width and l ength of the box accordingly.