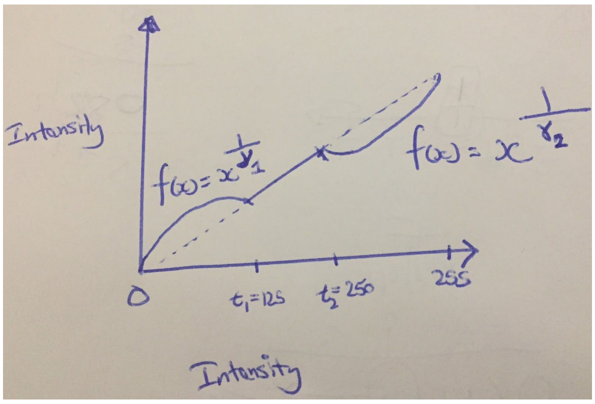
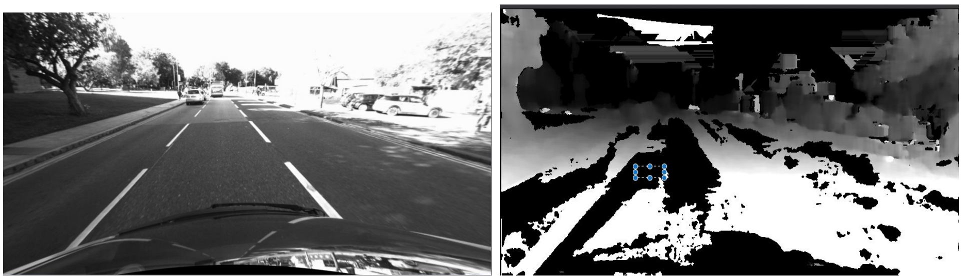
Sandeep Thapa

Computer Vision

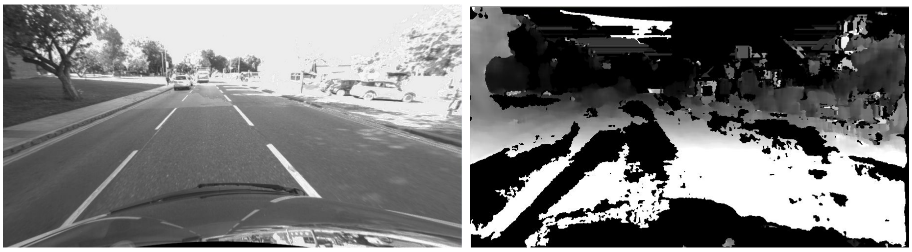
After observing the dataset, I learnt that frequently the different scenes within a frame would vary between being under and overexposed. Hence, I tried to rectify this by applying gamma correction. I used 2 different gamma values γ_1 , γ_2 to increase the intensity of dark pixels < 125 and reduce the intensity of bright pixels > 250 respectively. Pixels with intensity between 125 and 250 were mapped linearly, hence unchanged. This improved the visibility of regions that were not sufficiently illuminated,shown in B and C. However, increasing the value of γ_2 to darken the bright pixels produced fragmented textureless regions, shown in C, D and E. This increased the number of speckles and discontinuation of the disparity, shown in the right section of disparity map in C.



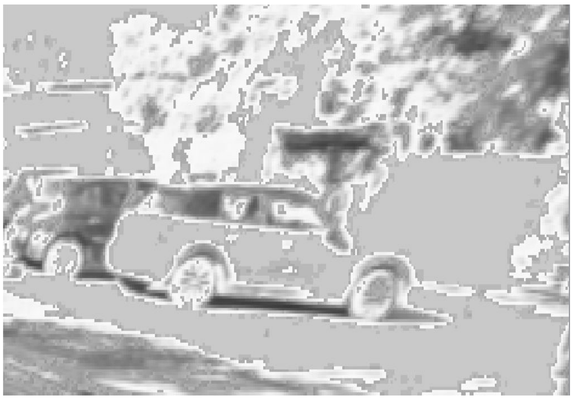
A)Function used to remap intensity



B) Original Gray Scale Image L and Disparity Map

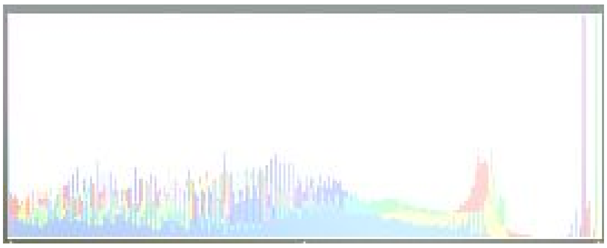


C) Gamma Corrected Image L γ_1 = 1.5 , γ_2 = 1/5

D) Zoomed in Image L $\gamma_1 = 1.5$, $\gamma_2 = 1/10$ E) Zoomed in Image L $\gamma_1 = 1.5$, $\gamma_2 = 1/20$

To solve the problem of irregular illumination while maintaining texture, I experimented with histogram equalization HE. The varying illumination in different parts of a frame meant the distribution of pixel values varied throughout the image. Hence, HE was unable to enhance the contrast of regions which were significantly brighter than the rest of the image. This is validated by the largely concentrated spike in an in the far right of the image G.Furthermore, noise and wrong colour dots and patches were introduced, circled in black in F.

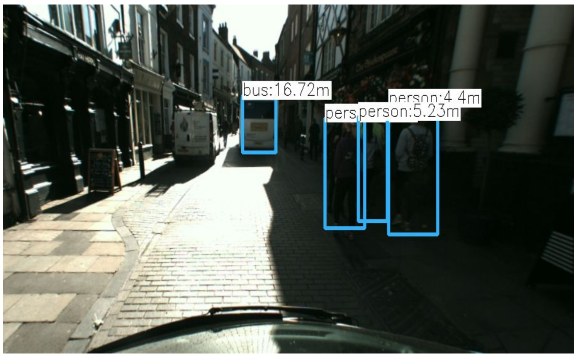
 

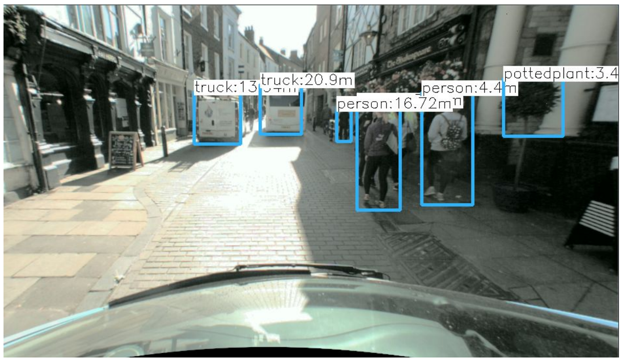
F) Image L after Histogram Equalization G) Histogram of Image L after Histogram Equalization (all 3 channels)

Therefore, I used adaptive histogram equalization AHE to individually perform HE at each pixel using the overlapping local NxN neighbourhood. The texture of the road, road marks and the trees are better defined. This is because such homogeneous regions are being remapped into more pixel intensity level, hence resulting in more texture. This lead to the generation of disparity map with increased area of disparity continuation and reduced noise, shown in H. It also improved the number of objects detected within the dark region of a frame using YOLO, show in J.
To further remove noise from the input images, I applied bilateral filtering with a standard deviation of 10 for spatial and intensity parameters.

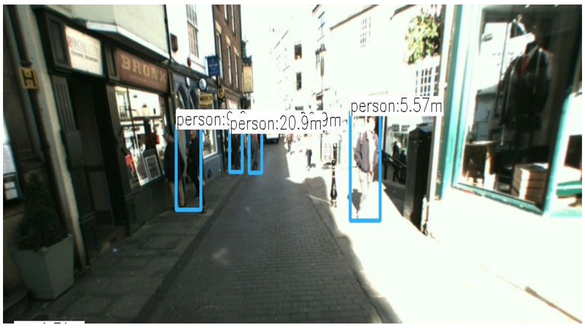
 

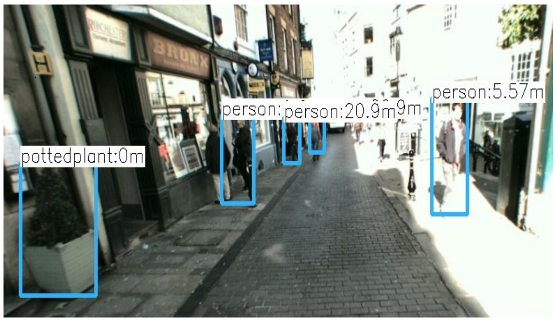
H) Image L after AHE 16x16 neighbourhood I) Image L disparity map after AHE16x16 neighbourhood





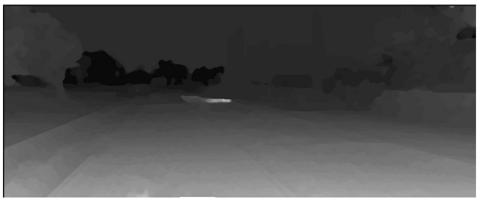
J) Original Image : 4 Object detected After AHE: 6 Object detected





K) Original Image: 4 Objected detected After AHE: 5 Object detected

Despite the improvement, the results of AHE still contained gaps. Therefore, I investigated using the weighted least square filter WLS. It generates 2 disparity map from left and right images then combine them to produce a single gapless disparity map, shown in images below. I then applied a median filter of window size 9x9 over the resulting disparity map to remove incorrect disparities.

Left Image: disparity map after AHE and WLS Filter, Speckles is circled in yellow. Right image: output after applying the median filter

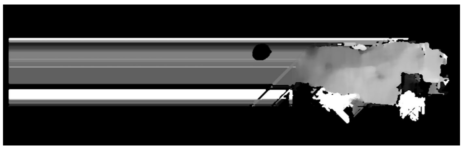
I further investigated with object segmentation to help bring focus on a particular object by ignoring the background while removing external noise in the image's background to construct a disparity map with higher accuracy.

To perform object segmentation I used a pre-trained Mask-RCNN model. The outputs of Mask-RCNN is a mask M_{class} of size 15x15 for each object in the image and the spatial position S of the object relative to the input image. M_{class} describes the set of pixels in the region R occupied by the object. The spatial position is in the form of 4 coordinates of a rectangle in which the object lives. I then overlaid M_{class} on an empty black image of the same size as the input image at position S to generate mask M_{final}. Similarly, for each object I updated M_{final} by overlaying respective M_{class}. As M_{final} is of the same size as the input image, I was able to overlay M_{final} on the input image to separate objects from the background.

My generated disparity map using the semi-global block matching SGBM algorithm with the 2 rectified segmented stereo images were highly inaccurate in many frames. False streak lines shown in Figure A were a common occurrence in many frames. Furthermore, I observed that the disparity map of many images failed to hold a boundary outline/shape of the objects.

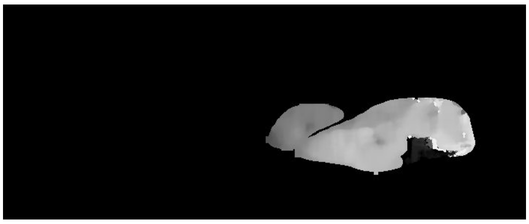
Right Segmented Image R

Left Segmented Image L



The unscaled disparity of image R and L. Figure A

Through trial and error, I was able to overcome this problem by first generating disparity map with non-segmented stereo images as input to SGBM algorithm then overlaying M_{final} over the disparity map to cut out disparity values within the region of the object. This gave me access to streakless and precise disparity values within the region of interest, shown in Figure B & C. The



The unscaled disparity of image R and L.Figure B

To find the distance to an object I experimented with 2 different methods to approximate the disparity of an object;1) taking the disparity value at the centre of the region of interest ROI, 2) Taking average disparity value of the region of interest. Both methods contained some margin of error in the generated result. However, I observed that in many cases, method 1 would lead to larger distance estimation for relatively nearby objects. This scenario is extremely dangerous in real life, as the car could assume nearby pedestrian as being far and cause an accident. Thus, in my final implementation, I used method 2.