

Computer Vision Assignment Report

Optimisation for disparity

Smoothing Filters (Gaussian, Bilateral, Weighted Least Squares (WLS))

Smoothing noise out of the input image / disparity map helps to reduce the influence noise has on our distance output. I have compared the effects of Gaussian noise removal, and edge preserving filters and have found that the WLS filter performs the best.

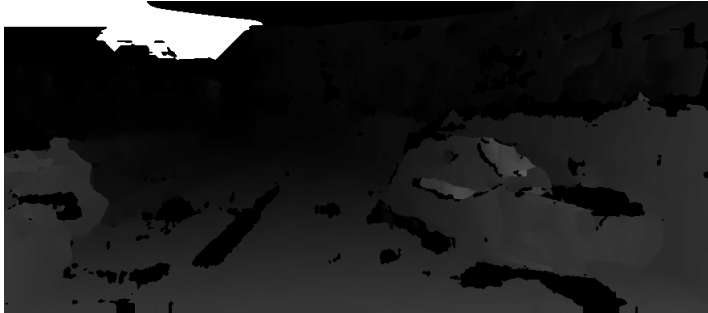


Disparity image with no filter



Disparity image with Gaussian Blur

In this case we see a slight increase in gaps – not ideal.



Disparity image with Bilateral Filter

Edge preservation gives a cleaner image around edges of objects.

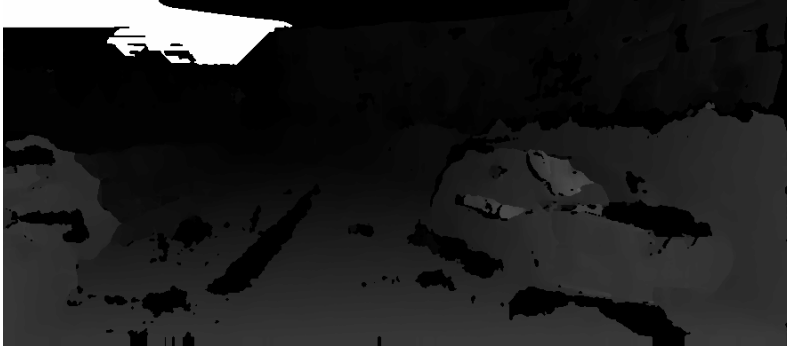


Disparity image with WLS Filter

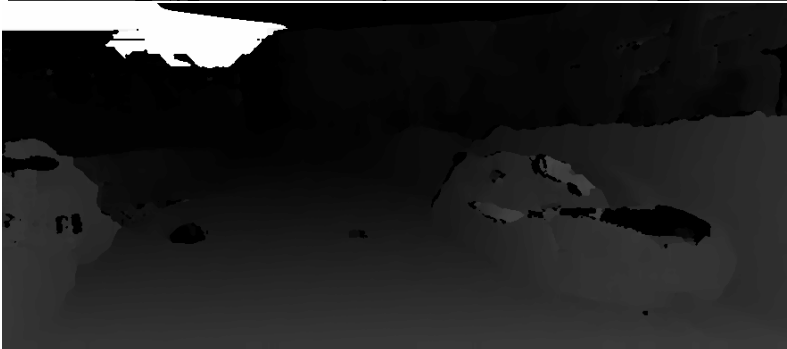
A very smooth disparity image. The ideal choice.

Histogram Equalisation

The disparity image contains noise introduced by irregular lighting in the image. By equalising the histogram of the image, we hope to reduce the irregularity and obtain a smoother image. In practice, histogram smoothing is an effective approach.



Disparity image with no correction



Disparity image with histogram equalisation

A lot of noise has been removed by this technique. In some areas, we have improved the edge detection.

Thresholding

Not all bounding boxes will be tight to the detected object, therefore will contain background information that can affect our distance value. Ideally, we remove the background information, and a method to do this is to threshold our image. I use the idea of Otsu Thresholding, that adaptively thresholds an image based on the centre of two histogram peaks – the theory being that the two peaks represent the foreground and background information and we remove the background.



No Thresholding

Classic example of a bounding box that contains a lot of background (car on the right). Notice how the distance value is further away than it should be (~3.5m).



Otsu Threshold Image

With Otsu thresholding against a clear background, we filter out the background info, obtaining a more accurate distance value.



Otsu Threshold Image w/ Erroneous Box

Observe the rightmost person having the same distance value as the car. This is due to Otsu thresholding the wrong peak due to the similarity of the background to the foreground – a flaw of this technique.

Mean vs Median vs Percentile

We face a multitude of choices calculating the final distance value from our disparity values. The first choice would be to take the mean; however, this is not robust against outlier values. The median is more robust against outlier values and so tends to be the better choice compared to the mean. Another choice would be to take a percentile of the values. We test the 25th and 75th percentile i.e. the lower and upper regions of the disparity map. In conclusion, the performance of the 75th percentile is the best in comparison.

Inference time: 315.24 ms



Mean of Values – Front Car

Taking the mean of the box has included background values – skewing the distance to be further away.

Inference time: 308.06 ms



Median of Values – Front Car

The median is more robust to outliers, resulting in a more accurate value.

Inference time: 316.59 ms



25th Percentile of Values – Front Car

Taking the 25th percentile has taken the background values (hedge) instead of the foreground – resulting in a poor distance estimate.

Inference time: 314.86 ms



75th Percentile of Values – Front Car

Taking the 75th percentile has taken the foreground values, giving the best estimate of the four averages.

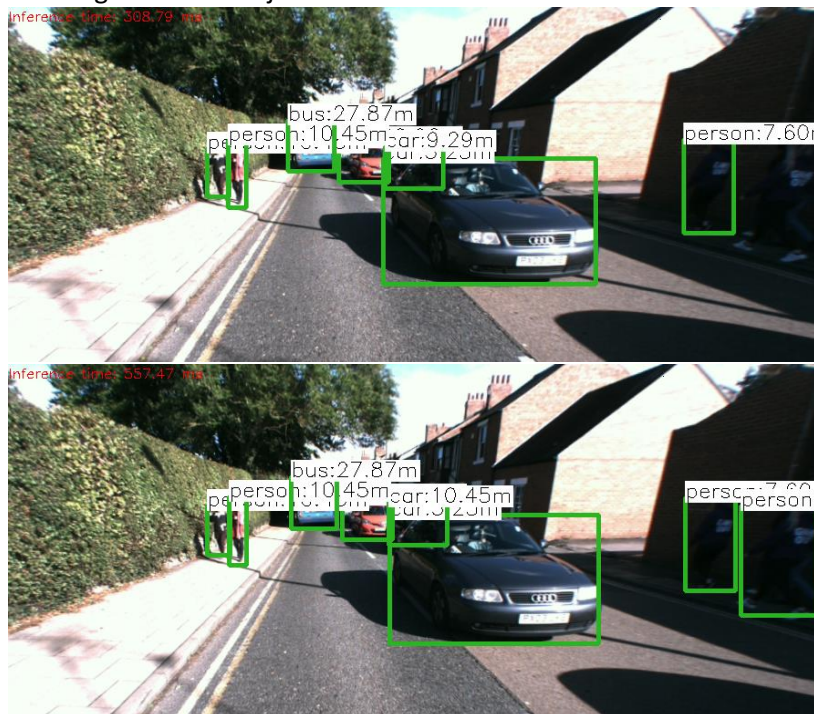
Optimisation for object detection

A basic optimisation for the object detection is to crop useless information out of the input image, namely the bonnet of the car. I have also opted to crop out the left side of the image where the dead zone for disparity exists – as this area is not of interest to us (we cannot obtain a distance here).

One of the harder scenarios for object detection is when objects of interest are in sections of the image that are poorly lit, so I have attempted contrast enhancement to mitigate the issue.

Contrast Enhancement via “Contrast Limited Adaptive Histogram Equalization (CLAHE)”

By converting the image to the HSV colour space, we can isolate the luminance channel V and perform CLAHE on it in order to improve the background contrast of an image while minimizing foreground information loss due to over-brightening. This exposes features in the background, allowing for better object detection.



No CLAHE image

CLAHE image

We have detected an extra person located in darkness.



No CLAHE image



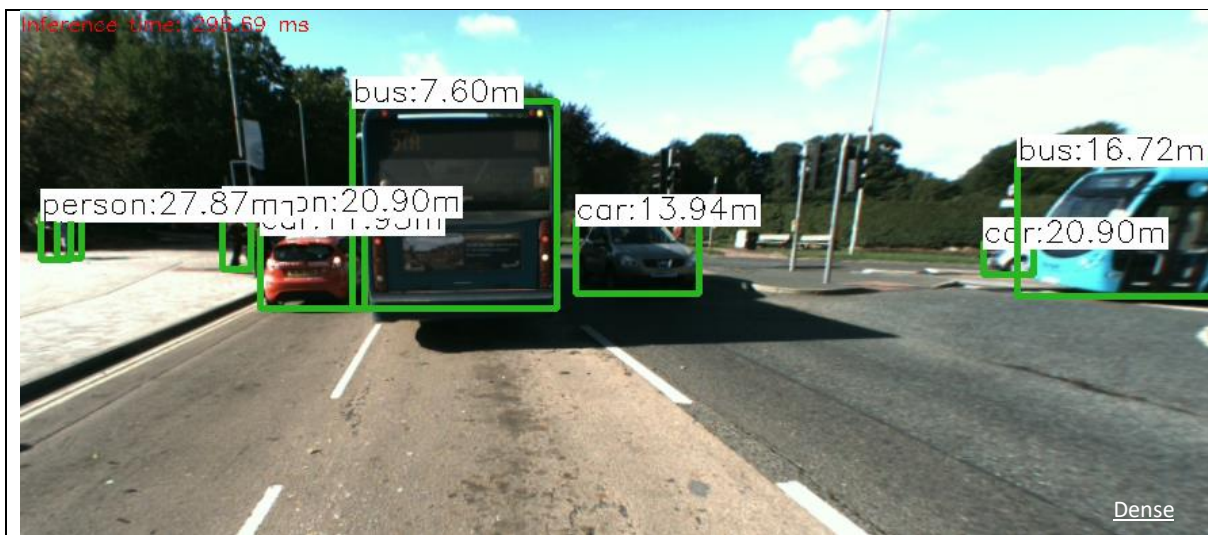
CLAHE image

We have lost a person in the bright area of the image due to over-brightness.

Sparse vs Dense Stereo

I have completed a basic implementation of sparse stereo imaging for disparity. To fairly compare the technique to dense stereo, we will do a comparison with all the optimisations mentioned earlier disabled, and we take the median of the disparity values in the bounding box.







In terms of accuracy, dense performs better than sparse in this implementation. However, given the feature point maps, sparse does have the potential to perform well with some optimisations.

Evaluation of the System

Overall, the system performs well in good conditions – images that are well lit, little noise etc. The system can handle densely populated areas reasonably well. The system can mitigate dark conditions using CLAHE at the expense of performance in extreme brightness. The system doesn't handle occluded objects very well, and the detection of people at a large distance could be improved.