

Computer Vision Summative Assignment

hfsk44

December 1, 2017

Preprocessing

Downsampling

All input images are first downsampled to a specified quality. This decrease in image size results in a significant performance improvement at the expense of some minor loss in accuracy. The upper 150 pixels are cropped out as they are highly unlikely to contain road, saving roughly 0.9s per frame.

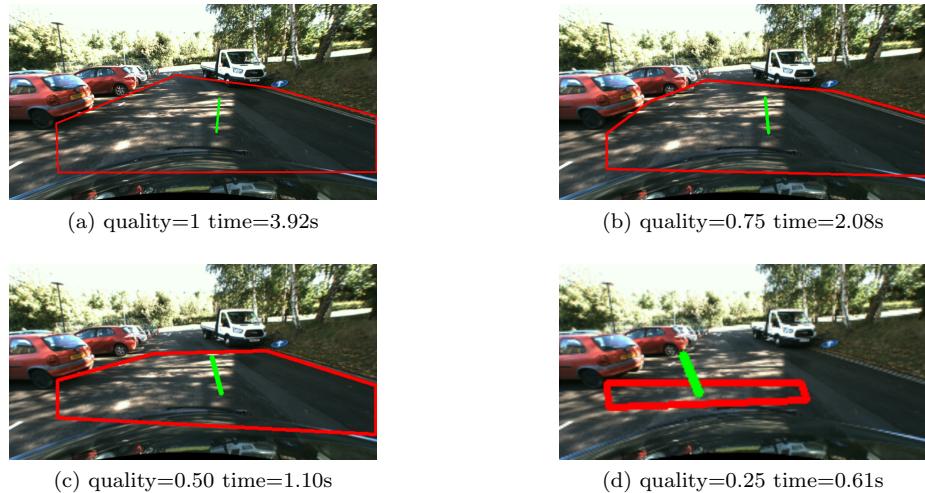


Figure 1: Results of downsampling at different rates

HSV Masking

Input from the left camera is converted to HSV and thresholded for values within the range $(40, 60, 60) \rightarrow (80, 255, 255)$ i.e. different levels of green. This resulted in a highly effective mask for removing grass and trees. Other objects such as cobbles would also fall in this range but in practice this would rarely affect results.

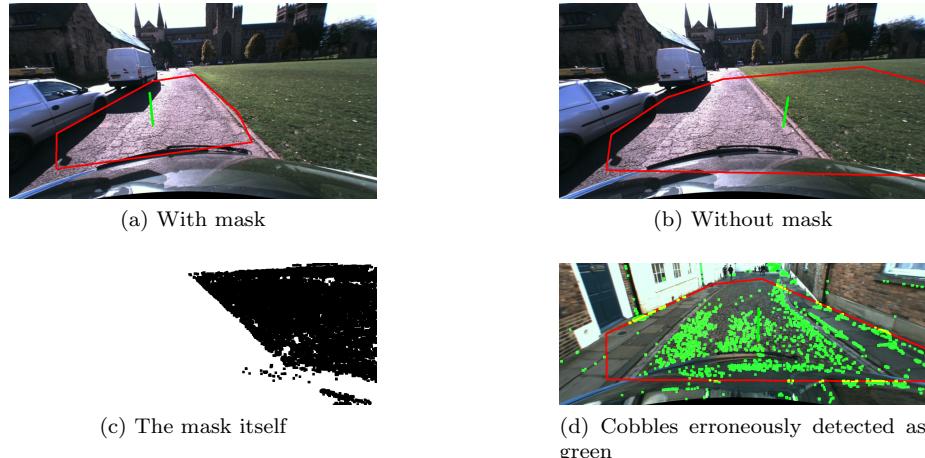


Figure 2: Example of HSV masking in use

Edge Masking

The result of a canny on the left image is used to create a mask for splitting consensus points that are separated by a strong edge. This is often effective in separating the road region from pavement and cars. The central region of the canny image is masked out in order to prevent road markings from separating the road region.

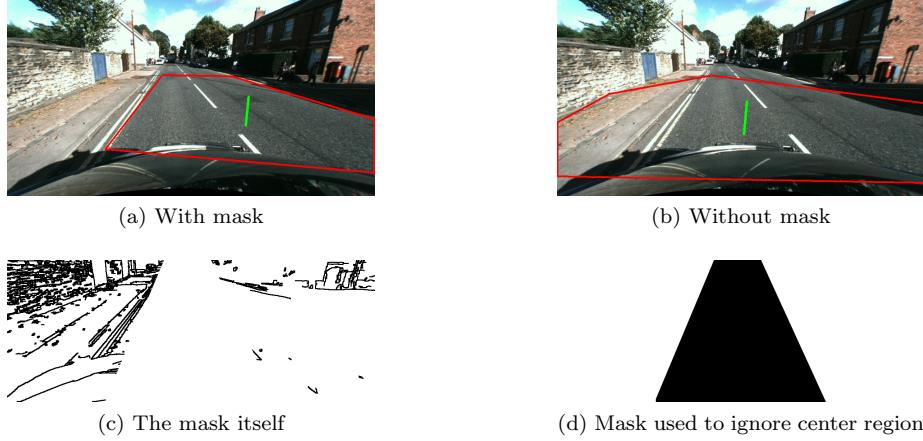


Figure 3: Example of edge masking in use

Contrast Stretching

When calculating disparity, pixel values of the greyscale input images are raised to the power 0.75. This helps prevent sharp changes in disparity along the edge of shadows on the road surface.

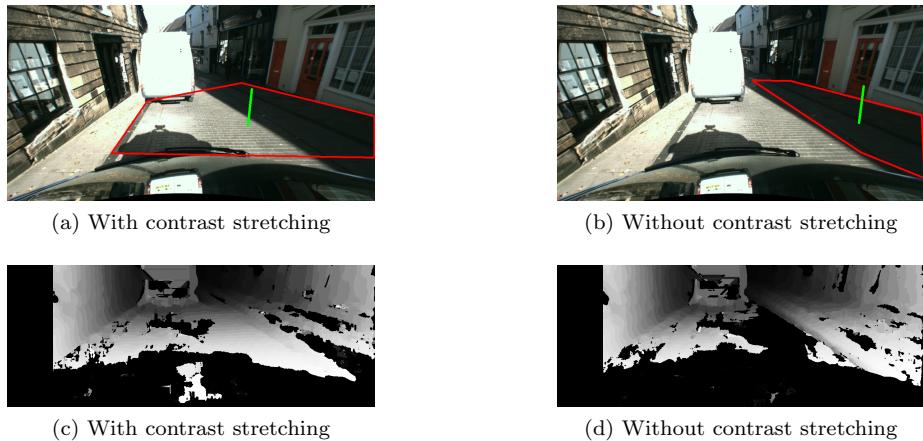


Figure 4: Example of contrast stretching in use

Disparity Cleanup

Navier-Stokes inpainting is used to fill holes in the disparity image, improving results significantly as the majority of these holes occur on the road surface. Motion blur will often create sharp spikes in disparity, preventing a good plane from being found. Filtering out disparity values above 50 eliminates this issue. A pre-defined mask is applied to remove areas that are highly likely to not be a road e.g. sky and the car bonnet.

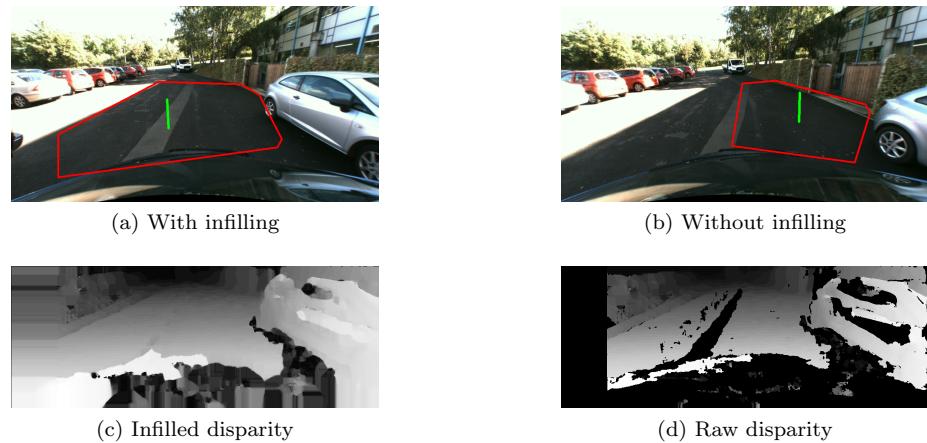


Figure 5: Example of disparity infilling in use

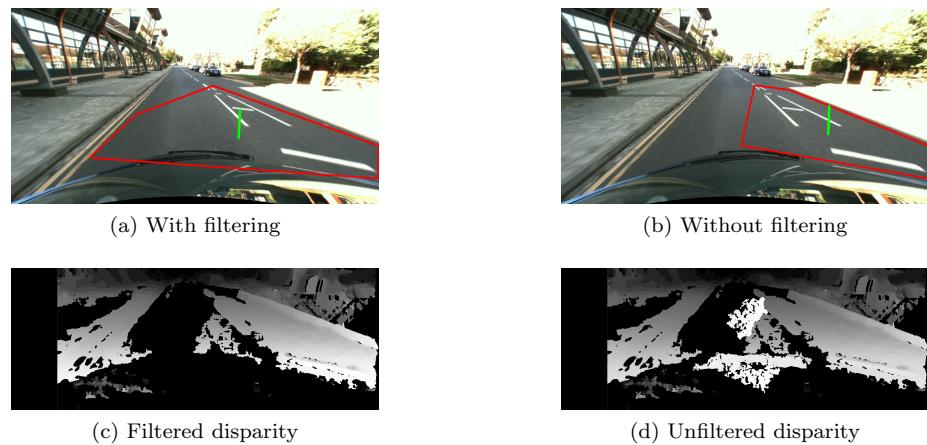


Figure 6: Example of high disparity value filtering in use

3D RANSAC

Disparity Projection

Within the projection from disparity values to 3D points, the Z value is used as a coefficient in the calculations for X and Y instead of Zmax. This gives a smoother distribution of points in the Z axis, increasing the chance of finding a large plane significantly.



Figure 7: Results of using Z instead of Zmax

RANSAC

200 iterations of a 3D RANSAC are performed, where the consensus points are those within a distance of 0.1 from the randomly chosen plane. The plane with the highest number of consensus points is taken to be the best estimate.

Any plane with a y coefficient less than 0.5 is immediately rejected, this is effective in preventing vertical planes such as the sides of cars and houses from being considered.

After all 200 iterations have been performed, an average is taken between the best plane found for the current frame and its predecessor. If there are more consensus points to the average plane then it is returned instead. This often results in a slight improvement in results for the equivalent of one more RANSAC iteration.



(a) With y coefficient rejection

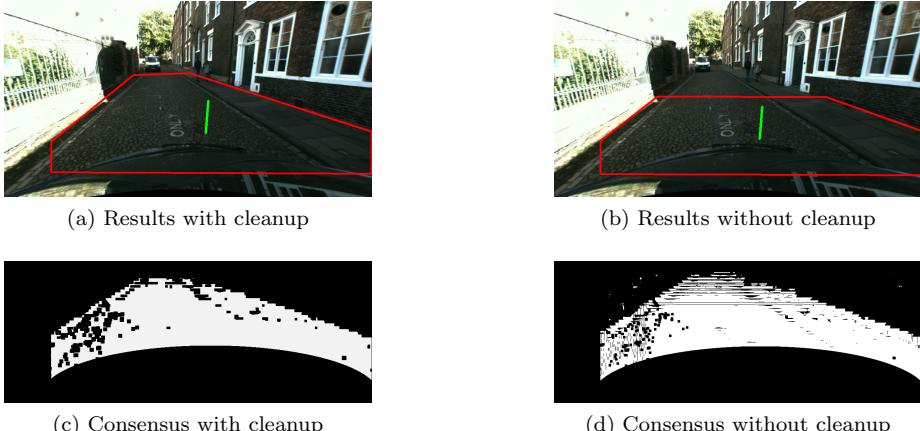
(b) Without y coefficient rejection

Figure 8: Results of rejecting RANSAC planes based on y coefficient

Postprocessing

Consensus Cleanup

The consensus points returned from RANSAC are projected back to 2D, where they are converted to a binary image. Morphology functions can then be used to remove banding effects in the consensus image caused by the discrete nature of the disparity image. Erosion is also performed to remove small areas of consensus separated from the main region.



(a) Results with cleanup

(b) Results without cleanup

(c) Consensus with cleanup

(d) Consensus without cleanup

Figure 9: Results of using morphology functions to clean up consensus

Contours

Contours of the consensus image are then taken and the road region is defined as the contour with the largest area. The convex hull of this contour is found and approximated to a polygon. This produces more consistent road boundaries at

the expense of occasionally ignoring objects on the road's surface such as people and traffic cones.

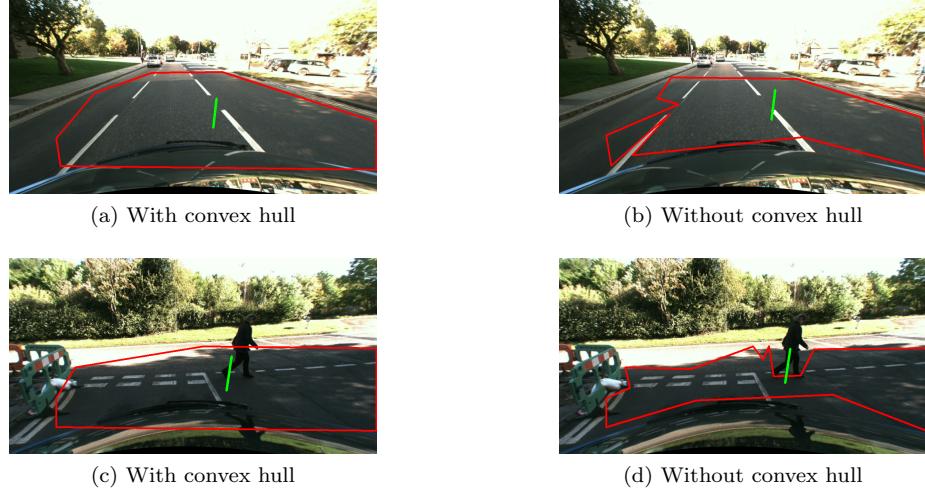


Figure 10: Results of using convex hull to approximate road region

Results

Using a random sample of 100 frames, the identified road regions were evaluated in terms of their *coverage* and *purity* where coverage describes the proportion of available road region that was found and purity describes the proportion of the identified region that was in fact road.

Score	Coverage	Purity
1	No road found	Contains no road
2	$\approx 25\%$ of road found	Majority non-road
3	$\approx 50\%$ of road found	Large areas of non-road
4	$\approx 75\%$ of road found	Small areas of non-road
5	$\approx 100\%$ of road found	Contains only road

		Coverage					Total
		1	2	3	4	5	
Purity	1	1	0	0	0	0	1
	2	1	0	0	2	4	7
	3	1	1	2	8	6	18
	4	0	2	7	7	16	32
	5	0	2	10	15	15	42
Total		3	5	19	32	41	

From this sample, the mean score was 4.03 for coverage and 4.07 for purity. This indicates that on average, a large proportion of road will be successfully identified with only a small area of the identified region being a false positive.

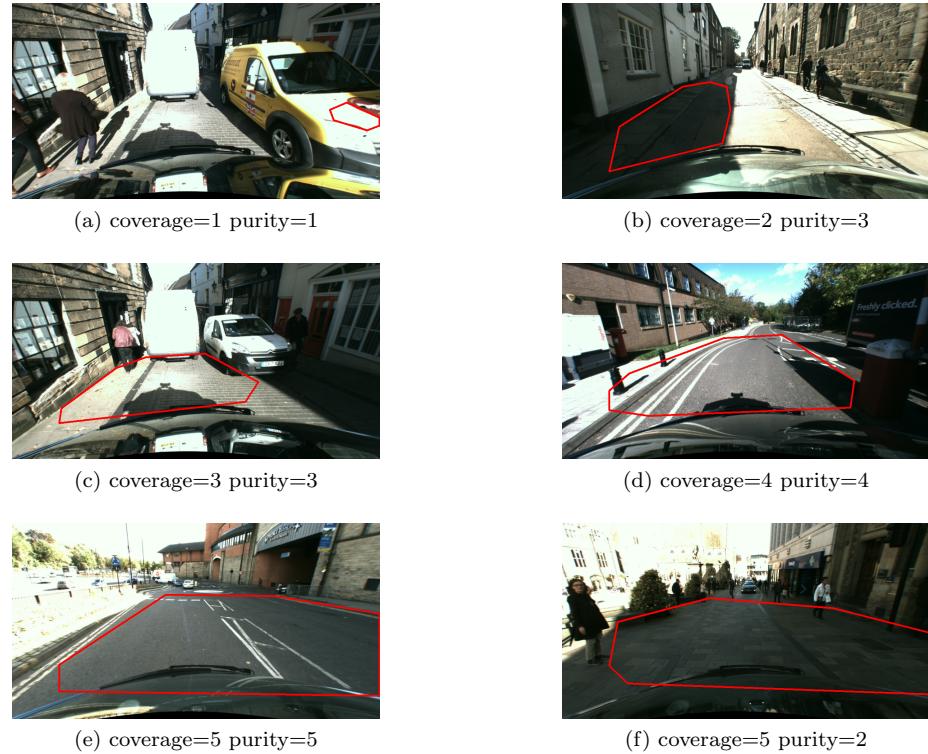


Figure 11: Sample of result images used in evaluation