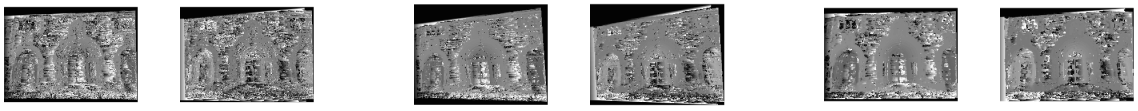
Report: Stereo Matching

Rishabh Singh: 19-953-793

November 15, 2019

1 Winner-takes-all

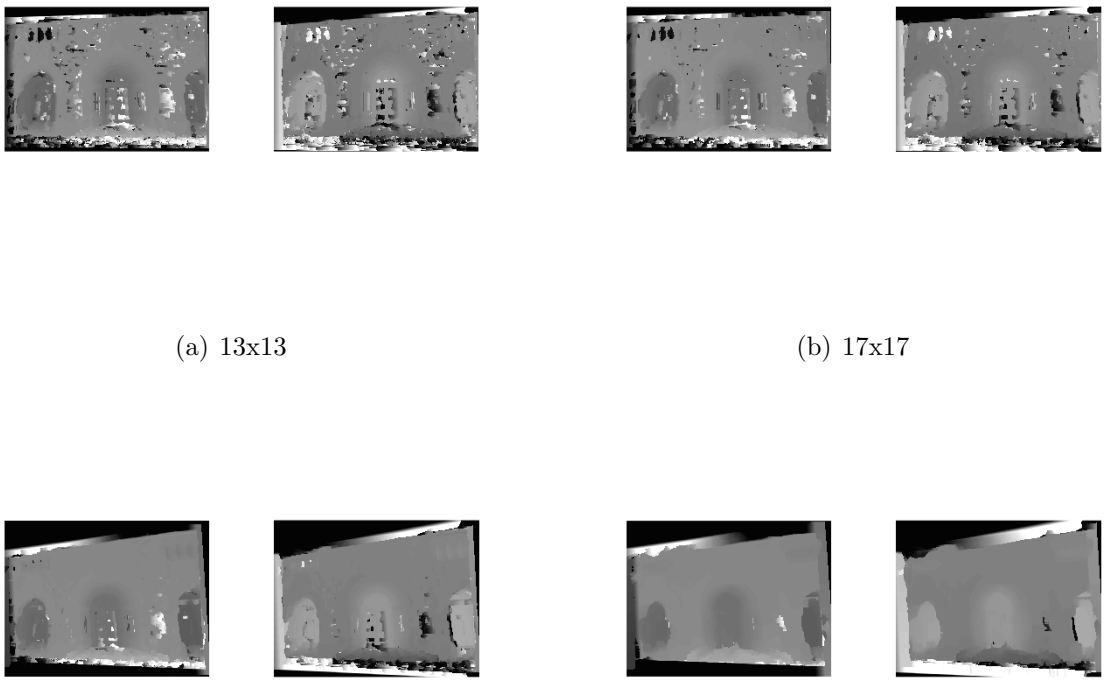
The given images were scaled to a lower resolution in order to have faster computation. Each image was accompanied with the associated camera parameters. The images were rectified. Rectification makes the search of corresponding points easier as the points lie on a horizontal scan-line. The disparity range for this part was kept as a constant value from -40:40. In order to compute the disparity map, one of the images was shifted horizontally by the given disparity value and the SSD value was computed in a window of size NxN at each pixel. The value of N was varied and the corresponding disparity maps were generated as shown in Fig.1:



(a) 3x3 (b) 5x5 (c) 7x7

Figure 1: Disparity map for winner-takes-all for different window sizes

On increasing the size of window from 3 to 7, it was observed that the disparity maps were losing the details. Although the effect was not very pronounced, it was hypothesized that the map is becoming smoother as the noise is getting reduced. In order to validate this, the window size was increased to from 13 to 41 as shown in Fig2. The obtained disparity maps are highly smooth with negligible noise.



(a) 13x13

(b) 17x17

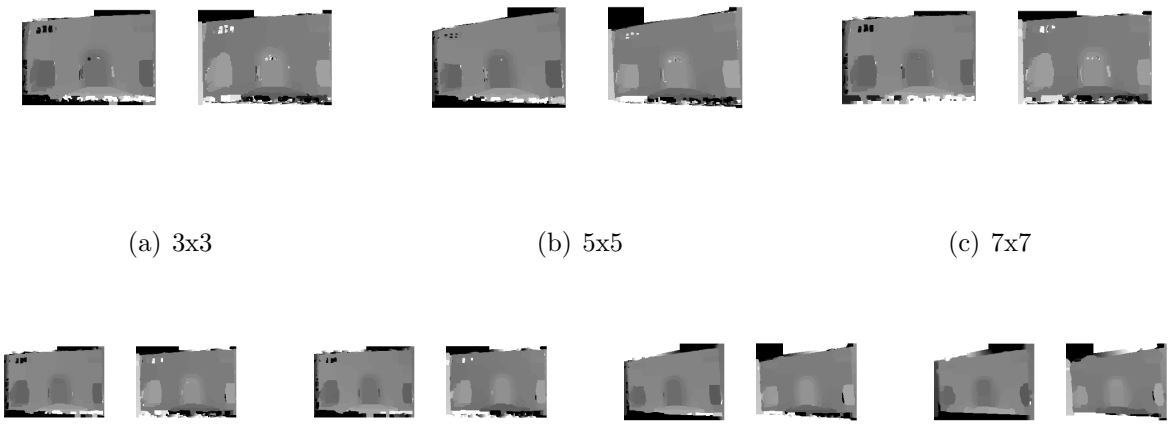
(c) 21x21

(d) 41x41

Figure 2: Disparity map for larger window sizes

2 Graph Cut

The given framework for Graph cut algorithm was utilised to generate the disparity map of given image pair. The data costs were obtained by implementing the function $diffsGC$. The smoothness term was the same as provided in the code-framework. The disparity maps thus obtained for different window sizes are shown in Fig.3: When compared with the above method, Graph Cut produces smoother disparity maps even for smaller window sizes. For larger window sizes, say 41, the disparity maps of Graph cut and winner-takes-all are more or less same.



(a) 3x3 (b) 5x5 (c) 7x7

(d) 13x13 (e) 17x17 (f) 21x21 (g) 41x41

Figure 3: Disparity map obtained from Graph Cut algorithm

3 Texture 3D models

Meshlab was utilised to generate the textured 3D models from the disparity maps generated from both the algorithms. The outputs are shown in Fig.4 and Fig.5. The texture models for winner-takes-all are usually poorly generated with lots of missing details for smaller windows. This improves as the window size is increased. The models for Graph cut are smoother and gain more details on increasing window size.

4 Automatic Disparity Range

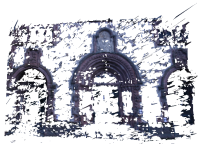
Instead of using the hard-coded disparity range $-40 : 40$ as provided in the framework, one can think of learning this range for optimal performance on a smaller search grid of disparity values. Two methods were explored in this regard:

4.1 Manual Point Clicking

Using the *getClickedPoints* function provided in the framework, corresponding points were clicked manually. Those points were chosen which can give sufficient information for depth range in the provided images. This also depends on inherent human bias of what we perceive as a region of more depth. For instance, the door areas seems to be within the building, hence depth value should be high there. Based on such assumptions, the corresponding points were clicked and the absolute value of the difference of the x-coordinates of each matching pair was taken. Since the image pair is rectified before, matched points should lie on a horizontal scan-line, hence the difference in x-coordinates were considered. The range

(a) 3x3

(b) 5x5



(c) 7x7

(d) 13x13

(e) 17x17

(f) 21x21

(g) 41x41

Figure 4: 3D texture models obtained from Winner-takes-all for different window sizes

.

was set to maximum value obtained thus, in a symmetric manner (positive and negative direction). The results for two different scenarios of choosing points are shown in Fig 6 and 7:

4.2 Automated matching pairs

Another way of automating the process can be automating the process of finding the point correspondence. SIFT features were extracted from both rectified images. VL-feat functions were used to obtain matches between extracted interest points. RANSAC based matching based on fundamental matrix estimation was used to eliminate outlier matches and keep only the significant matches in order to reduce noise in the disparity calculation. This is based on the knowledge of the previous assignments. The inlier points thus obtained were used to estimate the disparity range as described in the above section. In one particular iteration, the disparity range obtained was $-6 : 6$ based on RANSAC matching the the results are shown in Fig.8.

5 Discussion

- Winner-takes-all is a local method where depth of each pixel is estimated independently, while Graph cut is a global method where the depth surfaces are refined as a whole by enforcing smoothness constraints.

- Winner-takes-all relies heavily on the matches of corresponding interest points. Thus a common source of error is the error in matching of pixel owing to occlusions, illumination and viewpoint changes etc.

(a) 3x3

(b) 5x5

(c) 7x7

(d) 13x13

(e) 17x17

(f) 21x21

(g) 41x41

Figure 5: 3D texture models obtained from Graph Cut method for different window sizes

.

(a) points

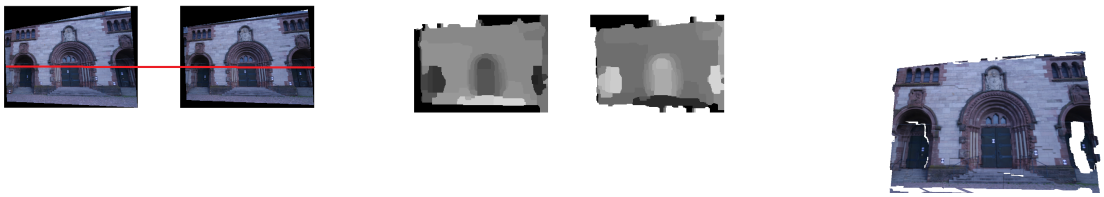
(b) Disparity map

(c) Texture model

Figure 6: Results obtained from manually clicking corresponding points for Graph Cut method (5x5 window)

.

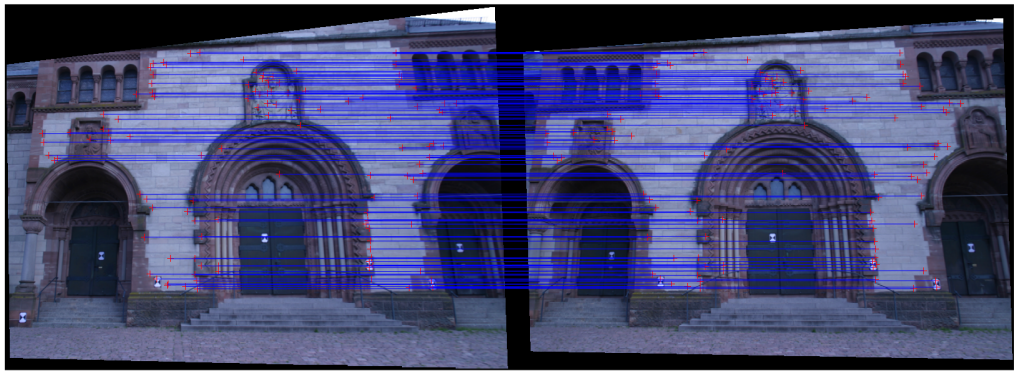
- Smaller window size for computing SSD at a pixel gives more details in the disparity map, but is plagued with noise too. Larger window on the other hand gives much smoother disparity map with lesser noise, but the intricate details go missing.
- The disparity maps obtained from Graph cut method were in general smoother than the ones obtained from winner-takes-all. The output of winner-takes-all is highly noisy and having finer details for lower window sizes, but it tends to become smoother on increasing window size. Thus the outputs of both the algorithms are equivalent for higher window sizes.
- The 3D texture models obtained from Winner-takes-all are perforated and grainy. As the window size is increased, we can observe a gradual increase in the amount of textured components present in each model. More information is present in higher



(a) points (b) Disparity map (c) Texture model

Figure 7: Corresponding points chosen on a strict horizontal line for Graph Cut method (5x5 window)

.



(a) SIFT-RANSAC matching

(b) Disparity map (c) Texture model

Figure 8: Points chosen based on SIFT feature matching

.

(a) hard-coded (b) manual clicked (c) SIFT matches

Figure 9: The amount of information in the disparity maps is more when the disparity range is automated

.

window models

- 3D texture models obtained from Graph Cut are usually smooth and have sufficient information/ textured components present. Increasing window size doesn't have a drastic effect as in case of winner-takes-all, but we can observe some minor improvements such as filling of holes in higher window sizes.
- Automating the disparity range, either through manually clicking or SIFT matching improves the disparity maps over hard-coded disparity range. (Fig.9)