CS4186 Assignment 2 Report – Research of Disparity Map computation

ZHANG Deheng 55199998

1 Introduction

Disparity map is an important feature for depth estimation problem.

$$z = \frac{bf}{x_l - x_r} \ (1)$$

In equation (1), z represents the depth (distance to the camera), b represents the baseline (distance of two cameras), f represents the focal length of the two cameras, and $x_l - x_r$ represents the disparity on the image plane. We can find that the disparity is inversely proportional to the depth. However, we need to match the features between two images to calculate the disparity. Epipolar constraint can limit the searching range of the feature from a plane to a line, but it is still not computational efficient. Rectification provides better epipolar lines, which are horizontal lines parallel to each other. In this case, we only need to find corresponding epipolar scanline in the right image for each pixel in the left image.

This project aims to implement an algorithm to generate disparity map for rectified image pairs. And the algorithm is assessed by the Peak Signal to Noise Ratio (PSNR) [1], which is defined as:

$$MSE = \frac{\sum_{M,N}[I_1(m,n) - I_2(m,n)]^2}{M * N}, PSNR = 10 * \log_{10}\left(\frac{R^2}{MSE}\right) (2),$$

where R is the maximum fluctuation in the input image data type (255). In this project, I implemented Semiglobal Matching (SGBM) [2] based on OpenCV [3]. As a result, the algorithm can generate reasonable disparity map and the average PSNR for the three images reaches 27.21.

2 Methodology

In this project, I use SGBM [2] to match the pixels and compute the disparity. The pipeline of SGBM is presented

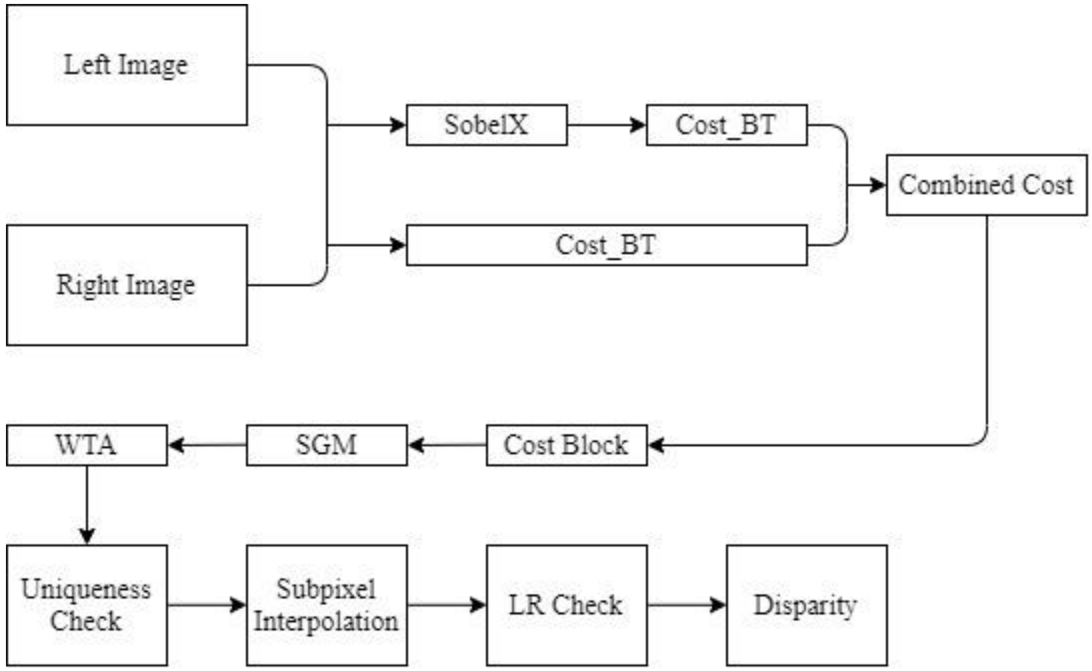

Figure 1: SGBM Pipeline

in Figure 1.

The images are filtered by the Sobel to extract the gradient information, which is defined as:

$$I'(x,y) = 2[I(x+1,y) - I(x-1,y)] + I(x+1,y-1) - I(x-1,y-1) + I(x+1,y+1) + I(x-1,y+1). (3)$$

Then the image is filtered by a threshold p:

$$I'' = \begin{cases} 0, I' \in (-\infty, -p) \\ I' + p, I' \in [-p, p] \\ 2 * p, I' \in (p, +\infty) \end{cases} (4).$$

After that, the cost of left and right image is defined as a combination of Birchfield and Tomasi metric (equation (5)) for original image I and preprocessed image I''. For each pixel, the block cost is defined by an aggregation of pixel value in a window. After that, the SGM is applied to minimize the energy function in global scale. Then the disparity map is computed by "winner take all" strategy (WTA) and checked by uniqueness, ordering, and smoothness constraint. After that, I used fast global image smoother [4] implemented by OpenCV [5] to post-process the generated disparity map.

$$d_l(x_l, x_r) = \min\ |I_l(x_l) - \hat{I}_r(x)|$$

$$d_r(x_l, x_r) = \min\ |\hat{I}_l(x) - I_r(x_r)|$$

$$d(x_l, x_r) = \min\{d_l(x_l, x_r), d_r(x_l, x_r)\}\,(x_l, x_r\ \mathit{are\ columns\ along\ the\ same\ scanline})(5)$$

3 Result Analysis

During the experiment, I tuned hyper-parameters [3] including:

- minDisparity: Minimum possible disparity value.
- numDisparities: Maximum disparity minus minimum disparity
- blockSize: Window size.
- P1 and P2: Smoothness parameters.
- Disp12MaxDiff: Maximum allowed difference in the disparity check.
- PreFilterCap: The threshold p in SobelX step.
- uniquenessRatio: Margin in percentage by which the best (minimum) computed cost function value should "win" the second best value to consider the found match correct
- speckleWIndowSize: Maximum size of smooth disparity regions to consider their noise speckles and invalidate.
- speckleRange: Maximum disparity variation within each connected component.

The PSNR result is shown as below:

The Peak-SNR value is 24.1103

The Peak-SNR value is **33.1090**

The Peak-SNR value is 24.4220

The computed disparity map is shown in Figure 2. We can observe that the disparity result is acceptable. However, it still has limitation. For example, in the Reindeer disparity map, there is a black region (which means negative or zero disparity is assigned to this region). In the Art disparity map, the disparity values of pen and circle pixels are blurred.

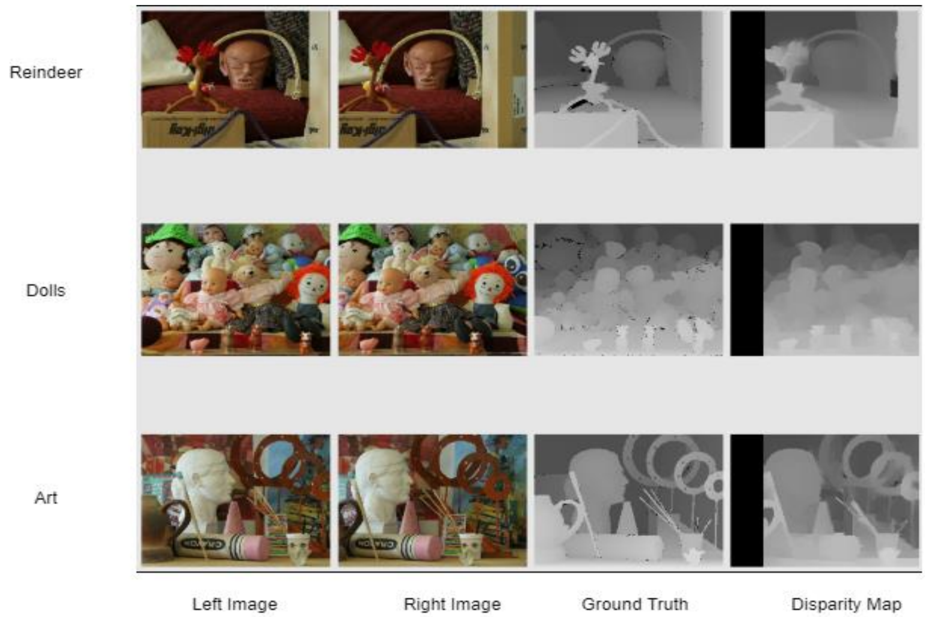

Figure 2: Disparity Result

4 File Structure

The *pred* folder includes the disparity result for three images (Art, Dolls, and Reindeer), as well as the text file of the PSNR result. The *code* folder includes the implementation of the algorithm. The *data* and *gt* folder contain three images and their disparity ground truth, respectively.

5 Conclusion

As a conclusion, we can find that SGBM is an elegant algorithm, which can efficiently estimate the disparity based on rectified image pairs. The algorithm uses dynamic programming to include global information for each window and selects the best match to minimize global energy function. I am very surprised that an algorithm without deep neural network can have such good result.

References

[1] MathWorks, "MathWorks Help Center," MathWorks, [Online]. Available: https://www.mathworks.com/help/vision/ref/psnr.html. [Accessed May 2021].

[2] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* p. 30(2):328–341, 2008.

[3] OpenCV, "cv::stereo::StereoBinarySGBM Class Reference," [Online]. Available: https://docs.opencv.org/3.4/d1/d9f/classcv_1_1stereo_1_1StereoBinarySGBM.html. [Accessed May 2021].

[4] S. C. J. L. B. H. K. S. a. M. N. D. Dongbo Min, "Fast global image smoothing based on weighted least squares," *Image Processing, IEEE Transactions on,* p. 23(12):5638–5653, 2014.

[5] OpenCV, "cv::ximgproc::FastGlobalSmootherFilter Class Reference," [Online]. Available: https://docs.opencv.org/3.4/d0/de3/citelist.html#CITEREF_Min2014. [Accessed May 2021].