

Binary Adaptive Semi-Global Matching based on Image Edges

Han Hu^{1,2}, Yuri Rzhanov¹, Philip J. Hatcher², R. Daniel Bergeron²

¹Center for Coastal and Ocean Mapping, University of New Hampshire, Durham, NH 03824, USA

²Department of Computer Science, University of New Hampshire Durham, NH 03824, USA

ABSTRACT

Image-based modeling and rendering is currently one of the most challenging topics in Computer Vision and Photogrammetry. The key issue here is building a set of dense correspondence points between two images, namely dense matching or stereo matching. Among all dense matching algorithms, Semi-Global Matching (SGM) is arguably one of the most promising algorithms for real-time stereo vision. Compared with global matching algorithms, SGM aggregates matching cost from several (eight or sixteen) directions rather than only the epipolar line using Dynamic Programming (DP). Thus, SGM eliminates the classical “streaking problem” and greatly improves its accuracy and efficiency. In this paper, we aim at further improvement of SGM accuracy without increasing the computational cost. We propose setting the penalty parameters adaptively according to image edges extracted by edge detectors. We have carried out experiments on the standard Middlebury stereo dataset and evaluated the performance of our modified method with the ground truth. The results have shown a noticeable accuracy improvement compared with the results using fixed penalty parameters while the runtime computational cost was not increased.

Keyword: Semi-global matching, dense matching, computer vision, 3D reconstruction, canny edges.

1. BACKGROUND

Depth information in our environment has a wide range of applications, such as land surveying, driverless assistance system and indoor navigation, etc. The depth information can be estimated through the dense matching procedure applied to two images from a stereo camera system. The dense matching is the most crucial step in the processing pipeline. Current dense matching algorithms can be basically divided into two categories: local algorithms and global algorithms. Local methods compare correspondence one point at a time, without consideration of neighboring points/measures, while global methods seek a disparity assignment that minimizes a global cost function which typically includes a data term and a smoothness term. Local methods are much faster than global methods but they usually suffer from a lack of smoothness in the final disparity map. Semi-Global Matching (SGM) as proposed by Hirschmuller [1, 2] combines the advantages of the above two methods with lower computational complexity for real-time needs given limited hardware resources and is able to achieve high precision depth estimation. Currently it is one of the most advanced and efficient dense matching algorithms which has proved to be successful in DSM generation [3] and driver assistance systems [4]. Two major research directions are being carried out in further development of this algorithm. The first direction is the optimization and acceleration of implementing SGM on different hardware architectures. This type of research focuses on the algorithm implementation on Graphics Processing Units (GPU) [5, 6] and on seeking efficiency improvement on the CPUs [7, 8]. Another research direction concentrates on improvement and evaluation of SGM regarding its accuracy and computational complexity and memory requirements. Within [1] a hierarchical approach using image pyramid was proposed to initialize and refine matching cost. The disparity of the higher level pyramid is used to refine the matching cost calculation for the lower level in order to accelerate convergence speed for higher levels. In [9] the accuracy of four different penalty functions in the cost aggregation step under two different types of matching cost calculation has been evaluated. Hirschmuller *et al.* [10] experimented with different cost calculation methods in three different stereo algorithms and concluded that hierarchical mutual information performed best for pixel-based global matching methods like SGM. Michael *et al.* [11] proposed using individual adaptive penalties for different path orientations where each path has its own weight and four penalty parameters which depend on intensity gradients (no edge selection). A large amount of data has to be considered for tuning such high numbers of parameters. In this paper, we propose an adaptive way for adjusting penalty parameters based on image edges for the reality that the image edges normally indicates disparity discontinuity. For the experiments, we consider the well-known Middlebury benchmark dataset [12] and evaluate the performance of our proposed modification by comparing the results with the ground truth.

The structure of this paper is organized as follows. The review of the original Semi-Global Matching algorithm is given in Section 2. We then shortly introduce and explain the edge-based SGM algorithm and present the experimental results and their evaluation in Section 3. The conclusion of our work and further improvements in the near future are briefly introduced in Section 4.

2. SEMI-GLOBAL MATCHING

2.1. Matching Cost Calculation with Mutual Information

Pixel Mutual Information (MI) is considered to be insensitive to recording and illumination changes [1]. The original SGM method uses MI as its pixel matching cost. It has been found out by Hirschmuller *et al.* [10] that the mutual information has better performance for most cases with SGM compared with other matching cost calculation methods like Birchfield and Tomasi (BT) interpolation [13]. MI comes from the theory of signal processing and is defined by the entropies H of the input two images I_1 and I_2 and their joint entropy H_{I_1, I_2} :

$$MI_{I_1, I_2} = H_{I_1} + H_{I_2} - H_{I_1, I_2} \quad (1)$$

The entropies are calculated from the image intensity probability distribution P :

$$H_I = - \int_0^1 P_I(i) \log P_I(i) di \quad (2)$$

$$H_{I_1, I_2} = - \int_0^1 \int_0^1 P_{I_1, I_2}(i_1, i_2) \log P_{I_1, I_2}(i_1, i_2) di_1 di_2 \quad (3)$$

Kim *et al.* [14] transformed the entropy calculation into discrete space using Taylor expansion. As a result, the joint entropy is calculated as a sum of data terms that depend on corresponding intensities of a pixel p :

$$H_{I_1, I_2} = \sum_p h_{I_1, I_2}(I_{1p}, I_{2p}) \quad (4)$$

$$h_{I_1, I_2}(I_{1p}, I_{2p}) = -\frac{1}{n} \log \left(P_{I_1, I_2}(i, k) \otimes g(i, k) \right) \otimes g(i, k) \quad (5)$$

The single image entropy is calculated by the following equation

$$H_I = \sum_p h_I(I_p) \quad (6)$$

$$h_I(i) = -\frac{1}{n} \log \left(P_I(i) \otimes g(i) \right) \otimes g(i) \quad (7)$$

Where P is the intensity distribution, n is the number of total correspondences and g denotes Gaussian convolution.

The resulting definition of MI is hence

$$MI_{I_1, I_2} = \sum_p mi_{I_1, I_2}(I_{1p}, I_{2p}) \quad (8)$$

$$mi_{I_1, I_2}(I_{1p}, I_{2p}) = h_{I_1}(i) + h_{I_2}(i) - h_{I_1, I_2}(i, k) \quad (9)$$

Therefore, the matching cost based on MI is defined as

$$C_{MI}(p, d) = -mi_{I_1, I_2}(I_{bp}, I_{mq}) \quad (10)$$

Where q is the corresponding pixel in match image of p in base image with disparity d

$$q = e_{bm}(p, d) \quad (11)$$

Since the calculation of the joint intensity distribution requires an initial disparity map to warp the match image towards the base image, the SGM uses an iterative computation strategy where the initial disparity map is assigned randomly.

2.2. Cost Aggregation

Traditionally, the 1D energy $E(D)$ of a disparity map D is calculated using the following equation

$$E(D) = \sum_p \left(C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{p \in N_p} P_2 T[|D_p - D_q| > 1] \right) \quad (12)$$

The first data term is the sum of matching cost for all pixels p . The second term adds a constant penalty P_1 for all the neighboring pixels q of pixel p if the disparity of q is different from the disparity of p by 1. The third data term adds a larger constant value P_2 for all the neighboring pixels q if the disparity difference between p and q is larger than 1. The problem of stereo matching is formulated as a problem of finding the disparity image D that minimizes the energy function $E(D)$. This global minimization problem is NP-complete and can be efficiently solved using Dynamic Programming (DP) [15]. However, it is well known that this minimization along separate epipolar lines is causing the “streaking problem” [15] due to independent processing between image rows. SGM solves this problem by aggregating matching cost from many different directions.

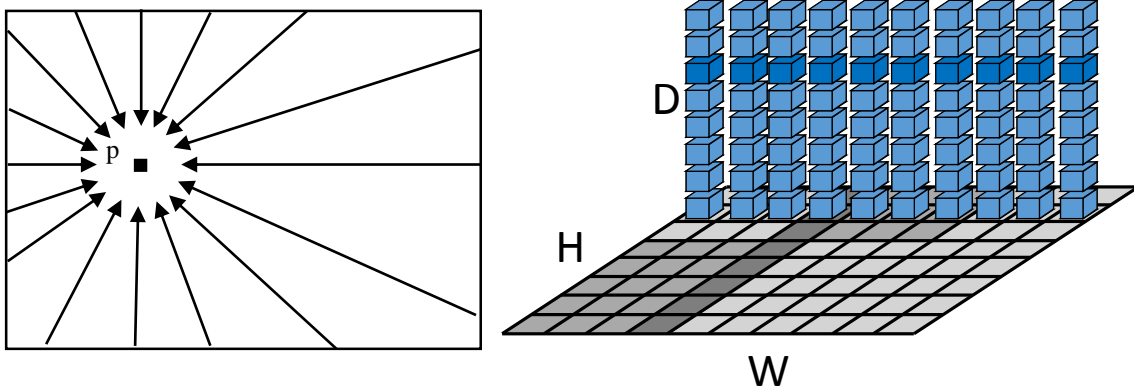


Figure 1 Cost aggregation. Left: 16 paths cost aggregation at a pixel p in 2D image space. Right: illustration of horizontal path cost structure on a single image row.

This is done through summing the costs of all 1D minimum cost paths that end in the pixel p at disparity d as illustrated in Figure 1 (left). The matching cost $L_r(p, d)$ of pixel p at disparity d along one particular path is calculated recursively by the following equation:

$$L_r(p, d) = C(p, d) + \min \left(\begin{aligned} &L_r(p-r, d), \\ &L_r(p-r, d-1) + P_1, \\ &L_r(p-r, d+1) + P_1, \\ &\min L_r(p-r, i) + P_2 - \min L_r(p-r, k) \end{aligned} \right) \quad (13)$$

The first term is the matching cost as it is in the energy function $E(D)$. $p-r$ is the previous pixel along the path. The last term is subtracted from the aggregated cost to avoid number overflow and this term is the same for all the disparities of pixel p . The final cost of pixel p at disparity d is the sum of all costs from all paths.

$$S(p, d) = \sum_r L_r(p, d) \quad (14)$$

2.3. Disparity Computation

After computing the matching cost cube, the disparity of a pixel p is determined by selecting the disparity that corresponds to the minimum cost from its disparity search range, that is $\min_d S[p, d]$. Hence the disparity image that corresponds to the base image D_b is obtained. The disparity image that corresponds to the match image D_m can also be determined from the same costs as well by traversing the epipolar line that corresponds to the pixel q of the match image. For sub-pixel disparity accuracy, a quadratic curve is fitted using neighboring costs next to the disparity that has the minimum cost. The occlusions and false matches can be determined by performing a consistency check between D_b and D_m . This consistency check enforces the uniqueness constraint by permitting one to one mapping only.

Summarizing, the algorithm steps for the SGM computation are:

1. Cost computation
2. Cost aggregation
3. Minimum cost determination
4. Sub-pixel interpolation
5. Median filter
6. Right-to-left consistency check

3. PROPOSED BINARY ADAPTIVE SGM AND EXPERIMENTS

The penalty parameters P_1 and P_2 used in original SGM are fixed constants during the aggregation process. We propose to incorporate a binary adaptive property into the original Semi-Global Matching algorithm based on image edges extracted by edge detectors. The edges used in this paper are defined to be points where there is a boundary (or an edge) between two image regions. In another words, we propose to use a smaller penalty parameter at image edges and a larger penalty at non-edge areas. This is reasonable since image edges which are typically object boundaries are more likely to indicate depth discontinuity. Thus, a smaller penalty should be used at image edges in order to give more freedom and to allow disparity changes. However, it should be also noted that even when edges are detected because of object texture rather than depth discontinuity, the penalty parameters will not affect the disparity estimation since the disparity with the minimum matching cost will remain minimum regardless of what penalty parameters are used there. Hence in our implementation, we have in total three different penalty parameters P_1, P_2 and P_3 where $P_3 > P_2 > P_1$. In the cost aggregation step, P_1 is added to the cost as it does in the original algorithm for neighboring pixels whose disparity changes a little bit (1 pixel). For the neighboring pixels whose disparity changes more than 1 pixel, P_3 is added to the matching cost if the pixel is located at image edges, otherwise the smaller penalty P_2 is added to the matching cost.

We performed our experiments on the standard Middlebury stereo datasets, specifically, “cones”, “teddy”, and “tsukuba”. We implemented the SGM algorithm in 8 paths rather than 16 paths. Only the median filter is applied after the initial disparity map is obtained from minimum cost selection. The consistency check and sub-pixel interpolation are not implemented because we only expect the qualitative justification of the proposed binary adaptive SGM. We use the popular Canny edge detector to extract image edges [16]. The performance of the proposed algorithm is evaluated by comparison of the results with the ground truth to see how much accuracy improvement could be achieved under these conditions. The parameters used in our experiment are $P_1 = 5, P_2 = 7, P_3 = 8$. We estimated the Root-Mean-Squared (RMS) error using the following equation

$$R = \left(\frac{1}{N} \sum_{(x,y)} |d_C(x,y) - d_T(x,y)|^2 \right)^{\frac{1}{2}} \quad (15)$$

Figure 2 shows the original image, the Canny edge image, the ground truth, the disparity map obtained through the original SGM and the disparity map obtained through our modified SGM. Table 1 presents the disparity map evaluation results for the original SGM and for our modified SGM. The accuracy of the disparity estimation is increased by about 6 to 7 percent as can be seen in the table. Note that this improvement does not require additional computations or memory cost when compared to the original SGM. We believe that with both consistency and sub-pixel processing added to the final disparity optimization, the accuracy improvement will become even more significant since the base RMS error becomes smaller.

	Original SGM	Binary Adaptive SGM	Accuracy Improvement
RMS(cone)	6.494	6.10	6.16%
RMS(teddy)	6.405	6.01	6.25%
RMS(tsukuba)	1.32	1.22	7.58%

Table 1 Disparity map evaluation results.



Figure 2 Experiment Results. Top to bottom row: image, Canny edges, ground truth, SGM depth map and depth map obtained by the reported algorithm.

4. CONCLUSION AND FUTURE WORK

In this paper, we propose to use binary penalty parameters based on the image edges extracted by the Canny algorithm instead of a single penalty parameter in the Semi-Global Matching algorithm. We experimented with this improvement on the standard Middlebury stereo dataset and evaluated its performance. We have found that this modification increases the accuracy of the disparity map estimation by approximately 6 to 7 percent without an increase of the computational cost. The implementation of this modification is straightforward and can be integrated into other SGM variants regardless of hardware configurations since it doesn't influence the parallelization of the original method.

The Canny detector only indicates presence or absence of sharp gradients in the brightness image with a single parameter threshold. However, there exist segmentation algorithms which produce a whole hierarchy of edges depending on their strength. In particular the algorithm proposed in [17] segments an image hierarchically thus attributing different strengths to different edges. This allows us to introduce a sequence of penalty parameters instead of single P_3 . This is a straightforward modification and we will experiment with this in the near future. Another possible improvement is associated with taking into account individual color channels. According to estimates in [18] this may provide up to 10% more edges compared to grayscale imagery. In general, edge detection in color images is a well-researched subject (see [19] for a review). We also plan to experiment with more combinations of these penalty parameters in the future in order to optimize their values for different types of images.

ACKNOWLEDGMENTS

This project was funded by the NOAA Joint Hydrographic Center grant NA05NOS4001153.

REFERENCES

- [1] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," *Pattern Analysis and Machine Learning*, 30(2), 328-341 (2008).
- [2] H. Hirschmuller, "Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information," *Proc. CVPR 2*, 807-814 (2005).
- [3] S.Gehrke, *et al.*, "Semi-Global Matching: An Alternative to LiDAR for DSM Generation?" *Remote Sensing and Spatial Information Sciences*, 38(B1) (2010).
- [4] S. Hermann and R. Klette, "Iterative Semi-Global Matching for Robust Driver Assistance Systems," *Proc. ACCV*, 465-478 (2013).
- [5] I. Ernst and H. Hirschmuller, "Mutual Information based Semi-Global Stereo Matching on the GPU," *Proc. Advances in Visual Computing*, 228-239 (2008).
- [6] J. Gibson and O. Marques, "Stereo Depth with a Unified Architecture GPU," *Proc. CVPRW'08*, 1-6 (2008).
- [7] R. Spangenberg, *et al.*, "Large Scale Semi-Global Matching on the CPU," *Intelligent Vehicles Symposium Proceedings*, 195-201 (2014).
- [8] S. K. Gehrig and C. Rabe, "Real-time Semi-Global Matching on the CPU," *CVPRW*, 86-92 (2010).
- [9] C. Banz, P. Pirsch, H. Blume, "Evaluation of Penalty Functions for Semi-Global Matching Cost Aggregation," *Proc. ISPRS 39*, B3 (2012).
- [10] H. Hirschmuller and D. Scharstein, "Evaluation of Cost Functions for Stereo Matching," *Proc. CVPR*, 1-8 (2007).
- [11] M. Michael *et al.*, "Real-time Stereo Vision: Optimizing Semi-Global Matching," *Intelligent Vehicles Symposium (IV)*, 1197-1202 (2013).
- [12] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *International journal of computer vision* 47(1-3), 7-42 (2002).
- [13] S. Birchfield and C. Tomasi, "A Pixel Dissimilarity Measure That Is Insensitive to Image Sampling," *Pattern Analysis and Machine Intelligence*, 20(4), 401-406 (1998).
- [14] J. Kim, V. Kolmogorov, R. Zabih, "Visual Correspondence Using Energy Minimization and Mutual Information," *Proc. Computer Vision*, 1033-1040 (2003).
- [15] S. Birchfield and C. Tomasi, "Depth Discontinuities by Pixel-to-Pixel Stereo", *International Journal of Computer Vision* 35(3), 269-293 (1999).

- [16] J. Canny, "A Computational Approach To Edge Detection", *Pattern Analysis and Machine Intelligence* (6), 679-698 (1986).
- [17] L. Prasad and A. N. Skourikhine, "Vectorized image segmentation via trixel agglomeration," *Pattern Recognition*, 39(4), 501-514 (2006).
- [18] C. L. Novak and S .A. Shafer, "Color Edge Detection", *Proceedings DARPA Image Understanding Workshop Vol. 1*, 35-37 (1987).
- [19] A. Mittal, S. Sofat, E. Hancock, "Detection of Edges in Color Images: A Review and Evaluative Comparison of State-of-the-Art Techniques", *Proc of Third Intl Conference AIS 2012*, 250-259 (2012).