*Applied Mathematics, Modeling and Computer Simulation*
*C.-H. Chen et al. (Eds.)*


*doi:10.3233/ATDE221021*

# Forward-Backward Diffusion and Pruning-Based Cost Aggregation for Non-Local Stereo Matching

Jinxin XU[1], Xueling YANG, Jinbo ZUO, Jiayan MU, Zhiqiang GUAN

*724th Research Institute of CSIC, Nanjing 211153, China*

**Abstract.** Minimum spanning tree (MST) has been devised for non-local cost aggregation to solve the stereo matching problem. However, the cost aggregation is employed directly from leaf toward root node, then in an inverse pass without considering any decision rules. And a small amount of noise is also existed in stereo image pairs. Both of the limitations often lead to failure in achieving more competitive results. This paper presents a novel stereo matching algorithm using forward-backward diffusion and pruning-based cost aggregation. In "forward-backward" process, the raw image pairs are smoothened on a horizontal tree structure as well as retaining image edges sharp. During cost aggregation, the MST where a complete graph involves the whole image pixels is cut off self-adaptively when the depth edge information is referred to. Each node in this tree receives supports from all other nodes which belong to similar depth regions. Meanwhile, an enhanced edge similarity function between two nearest neighboring nodes is formulated to deal with the small-weight-accumulation problem in textureless regions. Consequently, the cost volume can be well aggregated. The proposed method is demonstrated on Middlebury v.2 & v.3 datasets and can obtain good performance in disparity accuracy compared with other five MST based stereo matching methods.

**Keywords.** Stereo matching, MST, Forward-backward diffusion, Pruning-based aggregation, Similarity function.

## 1. Introduction

Dense stereo matching, as a traditional and challenging problem in the research of computer vision, has been employed in many applications such as 3D reconstruction [1], image refocusing [2,3], etc. Stereo algorithms can be divided into two categories, namely, local and global methods[4,5]. Both of them generally implement (subsets of) the following four steps: matching cost computation, cost aggregation, disparity computation, and disparity refinement.

Global methods usually formulate the matching problem in an energy minimization framework consisting of a data term and a smoothness term. Dynamic programming (DP) [6,7], belief propagation (BP)[8] and graph cuts (GC)[9,10] are the classical global algorithms. Drawbacks exist if the smoothness assumption is infringed

[1] Corresponding Author, Jinxin XU,724th Research Institute of CSIC, Nanjing 211153, China; E-mail: 2019377631@qq.com

or if it is improperly modeled. Compared with global methods, local methods are often less time-consuming. However, their computational efficiency usually comes at the expense of reduced matching accuracy and increased sensitivity to noise[11-15].

Different from summing the matching cost over a fixed or adaptive cross-based window [14,16,17] in most cost aggregation methods, Yang [18,19] reexamined the cost aggregation problem and proposed a non-local solution, where the MST structure is built for each pixel. Mei et al. [20] proposed a segment-tree (ST) structure for non-local cost aggregation. Recently, a generic cross-scale cost aggregation framework was proposed to allow multi-scale interaction in cost aggregation [21]. Further, Yao et al. [22] improved the MST method by proposing a logarithmic transformation[23] on matching cost function and introducing depth weight into the edge weight function.

These MST based methods mentioned above perform directly from leaf nodes toward root node, then from root node toward leaf nodes without any decision rules which would contribute to better depth boundaries. In textureless regions, the color differences between two neighboring pixels usually have values close to zero. Thus many small weight edges can accumulate along a long path, and high weights will be formed inevitably, leading to the small-weight-accumulation problem [23]. In addition, a small amount of noise is also existed in stereo images.In this paper, a novel stereo matching algorithm using forward-backward diffusion and pruning-based cost aggregation is proposed. The proposed method can be competitive in several ways. First, the edge weight depends on color difference, which is vulnerable to local noise. The forward-backward smoothing can bring a different and more robust MST structure. Second, the pruning-based cost aggregation incorporates the local weights and disparity discontinuities, allowing each node to receive supports from all other nodes belonging to similar depth regions. Thus more self-adaptive aggregation results and better depth boundaries can be expected. Moreover, the enhanced edge similarity function is efficient than the iteratively color-depth weight in [20,22,23], of which the essence is to obtain a more effective weight at the expense of computational complexity. Further, the enhanced edge similarity function can decrease the percentage of error pixels in textureless regions.

The contributions of this paper are as follows: 1) An effective diffusion method for preprocessing of the raw input images; 2) An novel pruning-based method for cost aggregation; 3) An enhanced edge similarity function for small-weight-accumulation problem; 4) Quantitative evaluation with several MST based methods on Middlebury *v.2 & v.3* datasets.

The following section briefly review the MST based non-local aggregation method. Following that, the proposed method is described in detail in Section 3. In Section 4, the experimental results are presented. Section 5 concludes the paper.

## 2. Non-Local Cost Aggregation

For describing the proposed method in a better way, in this section the MST based non-local cost aggregation is reviewed briefly.

### *2.1. Construction of MST structure*

First, the color image $I$ is regard as a connected, and undirected graph $G=(V, E)$, where

graph $G$ is a simple 4-connected grid, vertices $V$ are all the image pixels and edges $E$ are all the edges derived from nearest neighboring pixels. For an edge $e$ connecting neighboring pixel $s$ and $r$, the corresponding weight is defined as:

$$w(s,r) = w(r,s) = \max_{i=R,G,B} |I^i(s) - I^i(r)| \tag{1}$$

where the maximum value measured separately from the RGB color space will be chosen as the edge weight.

Second, on the basic of the edge weights and the vertices, the MST structure where edges with small weights are less likely to cross depth borders can be constructed by Kruskal algorithm.

### *2.2. Cost aggregation strategy*

Inspired by Bilateral filter method, the aggregated cost on MST structure is described as a weighted sum of matching cost:

$$C_d^A(p) = \sum_{q \in I} S(p,q) C_d(q) \tag{2}$$

where $C_d(q)$ denotes the matching cost for pixel $q$ at disparity $d$, $q$ covers all the pixels in image $I$. $C_d^A(p)$ denotes the aggregated cost. $S(p, q)$ is an edge similarity function, for denoting the similarity between pixel $p$ and $q$. $D(p, q)$ corresponds to the distance between pixel $p$ and $q$, it is the sum of weight $w$ along the path in the MST structure. $\sigma$ is a user-specified parameter used for adjusting the similarity between two pixels.

$$S(p,q) = \exp(-\frac{D(p,q)}{\sigma}) \tag{3}$$

The non-local cost aggregation can be implemented by traversing the tree structure in two sequential steps: 1) aggregate the matching cost from leaf nodes towards root node by Eqn. (4); 2) aggregate the matching cost from root node towards leaf nodes by Eqn. (5). Where $Ch(p)$ and $Pr(p)$ denote the children and parent of pixel $p$, respectively.

$$C_d^{A\uparrow}(p) = C_d(p) + \sum_{q \in Ch(p)} S(p,q) \cdot C_d^{A\uparrow}(q) \tag{4}$$

$$C_d^A(p) = S(Pr(p), p) \cdot C_d^A(Pr(p)) + (1 - S^2(Pr(p), p)) \cdot C_d^{A\uparrow}(p) \tag{5}$$

## 3. Proposed Method

In this section, the proposed method is outlined, which includes a description of (1) forward-backward diffusion, (2) pruning-based cost aggregation, (3) enhanced edge similarity function.

### *3.1. Forward-backward diffusion*

Forward-backward diffusion is a smoothing method that allows image edges to remain sharp as well as smoothening out noise in raw image pairs. First, a horizontal tree structure in which the intensities of raw images can be renewed on scanning lines is built for each pixel, such effective tree structure is introduced by Michael Bleyer[24]

for DP based stereo matching. This DP based matching method is on the basis of scanning line optimization[25] with 8 or 16 passes, which significantly improves the computational speed. Figure 1 describes the example of the horizontal tree structure which rooted at the pixel indicated by the solid yellow circle.

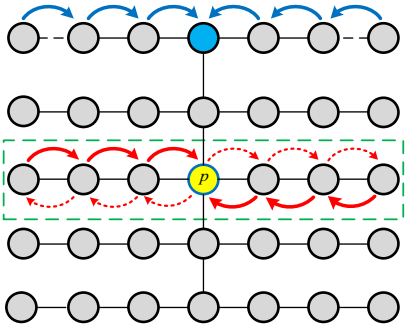



**Figure 1.** Horizontal tree structure rooted on pixel *p*.

Second, the forward-backward diffusion involves taking image $I$ with RGB color spaces as input, and applying the diffusion equation, as Equation (6):

$$I_{new}^{i}(u,v)=\sum_{r}I_{r}^{i}(u,v)-I^{i}(u,v),\ i\in R,G,B \tag{6}$$

where $I^{i}(u,v)$ is the pixel intensity at coordinate $(u,v)$ under channel $i$ of input image, $I_{r}^{i}(u,v)$ represents the result of forward or backward diffusion:

$$I_{r}^{i}(u,v)=I^{i}(u,v)+\lambda\cdot\nabla_{r}I^{i}(u,v)\cdot\exp(-\frac{|\nabla_{r}I^{i}(u,v)|}{255\cdot\omega}),\ r\in f,b \tag{7}$$

$$\nabla_{r}I^{i}(u,v)=I^{i}(u,v)-I^{i}(u,v-r) \tag{8}$$

where the constant $\lambda$ is used for maintaining the numerical stability and controlling the speed of diffusion, diffused images in this paper are obtained using a value of $\lambda=0.2$. The symbol $\nabla_{r}I^{i}(u,v)$ is the difference between two neighboring pixels under direction $r$. $(u, v\text{-}r)$ is the previous pixel coordinate of $(u, v)$ along the same direction. $f$ and $b$ indicate the forward and backward direction, respectively. $\omega$ can be a fixed value or be automatically based upon an estimate of the noise. The exponential term is small when there is large difference between neighboring pixels, especially in texture regions. Thus pixels will contribute to pixels on the other side in a minimal way and the image edges will remain largely intact after the forward-backward diffusion.

For linear time implementation, the diffusion can be performed in two passes: one pass performed from the leftmost node to the rightmost node and the forward results are stored in array $F_{LR}^{i}$, another one performs in a reverse direction and the results are stored in array $B_{RL}^{i}$, hence the forward-backward diffusion can be computed by:

$$\mathrm{I}_{new}^{i}=F_{LR}^{i}+B_{RL}^{i}-\mathrm{I}^{i},\ i\in R,G,B \tag{9}$$

The advantage of the forward-backward smoothness is that real image edges are preserved while background noise is restrained. By updating the intensity values with the forward-backward diffusion, the surface texture regions can be smoothened

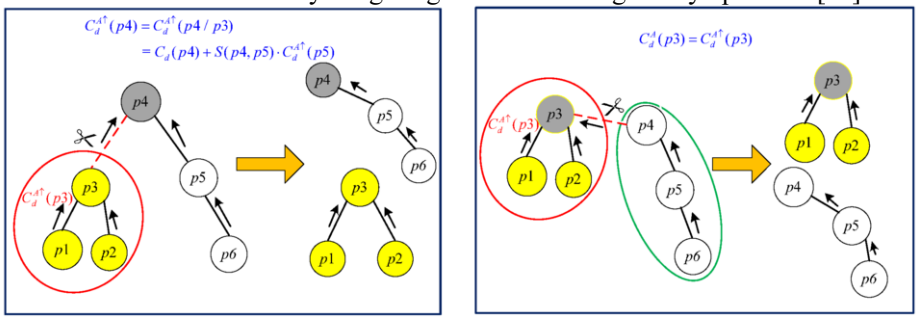
and the accuracy of disparity can be improved accordingly.

### 3.2. Pruning-based cost aggregation

The original MST algorithm only takes two-pass cost aggregation on a tree structure directly. However, the depth boundaries would be blurred on account of lacking 3D cues at depth discontinuities. Accordingly, the pruning-based decision rule incorporates the local weights and disparity discontinuities, and each node can receive supports from all other nodes belonging to similar depth regions. Therefore, more self-adaptive aggregation results and better depth boundaries can be expected. The following two claims for the pruning-based cost aggregation are defined as:

**Claim 1.** Let $T_r$ denote a sub-tree with root node $r$ of which the parent node is $s$. There are two cases need to be considered. In one case, when node $r$ is a pixel in disparity discontinuities, the connection between node $s$ and $r$ is cut off and node $s$ will not receive any supports from $T_r$. The latter alternative is that the supports aggregated from $T_r$ to $s$ is the sum value of : 1) the aggregation supports from node $r$ to $s$. 2) $S(s, r)$ times the supports from $r$'s sub-trees to $r$.

According to Claim 1, the matching cost is aggregated the first time from leaf nodes to root node as shown in figure 2(a) where $p4$ indicated by solid gray circle is the root node, and $p5$ is the leaf node which does not locate at disparity discontinuities. $p3$ indicated by solid yellow circle is the leaf node at disparity discontinuities, where the pruning is needed to be performed. Since $p5$ does not locate at disparity discontinuities, the cost aggregation is directly obtained from Eqn. (4). Instead, the link between node $s$ and $r$ will be cut off and the aggregated cost at node $p3$ will no longer contribute to node $p4$, thus the aggregated cost value $C_d^{A\uparrow}(p4)$ actually contains supports except from node $p3$ and its sub-tree. As disparity discontinuities are more likely to occur at image edges, in this paper the disparity discontinuities are described by image edges obtained using Canny operation [26].




(a) from leaf to root (b) from root to leaf

**Figure 2.** Two pruning-based cost aggregation steps.

Let $C_d^{A\uparrow}$ denote the aggregated cost volume and $Ch(p)$ denote the children of node $p$, then at each node $p$,

$$C_d^{A\uparrow}(p)=\begin{cases}C_d(p)+\sum\limits_{q\in Ch(p)} S(p,q)\cdot C_{\mathrm{d}}^{A\uparrow}(q), & q\notin edge\\ C_d(p), & q\in edge\end{cases} \tag{10}$$

**Claim 2.** Let $T_r$ denote a sub-tree with root node *r* of which the parent node is *s*, and $C_d^A(s)$ denote all supports received by node *s* in the first aggregation. There are also two cases need to be considered. In one case, when node *r* locates at disparity discontinuities, the supports *r* received from nodes other than $T_r$ is null. Alternatively, the supports *r* received from nodes other than $T_r$ is $S(s,r)\cdot[C_d^A(s)-S(s,r)\cdot C_d^{A\uparrow}(s)]$.

According to Claim 2, the aggregated cost volume is then aggregated the second time from root node to leaf nodes as shown in figure 2(b), where *p*3 indicated by solid gray circle is the root node as well as a node locates at disparity discontinuities. Since in the first aggregation, *p*3 and its sub-tree contribute no supports to root node *p*4, it is unnecessary for *p*4 to provide any supports to *p*3 in a reverse direction. However, when *p*3 does not belong to disparity discontinuities, the supports *p*3 received from its children nodes is the same way as described in Eqn. (5). Then the aggregated cost volume $C_d^A(p)$ for any node *p* from its parent *Pr*(*p*) can be described as follows:

$$C_d^A(p)=\begin{cases}S(Pr(p),p)\cdot C_d^A(Pr(p))+(1-S^2(Pr(p),p))\cdot C_d^{A\uparrow}(p), & p\notin edge\\ C_d^{A\uparrow}(p), & p\in edge\end{cases} \tag{11}$$

Hence, the whole pruning aggregation processes can be separated into two steps:1) aggregation from leaf to root node using Eqn. (10); 2) aggregation from root node to leaf nodes using Eqn. (11).

### *3.3. Enhanced edge similarity function*

In this section, an enhanced edge similarity function between two nearest neighboring nodes is proposed for suppressing the impact of this problem in textureless regions.

Based on Eqn. (3), the enhanced edge similarity function is defined as follows:

$$S_e(p,q)=S(p,q)\cdot\exp\left(-\frac{S(p,q)}{\mathrm{Z}}\right) \tag{12}$$

where $S_e(p,q)$ denotes the enhanced edge similarity function. In textureless regions, the edge similarity function $S(p,q)$ is close to 1, however the proposed edge similarity function $S_e(p,q)$ can obtain a relatively small value, which significantly implies that the weights in textureless regions are magnified and are no longer zero. Besides, in texture regions where the color differences are large, the value of the exponential term in Eqn. (12) tends to be 1, thus color differences in these regions will not be magnified.

In [20,22,23], a hybrid MST structure combining the color and disparity distance is proposed to obtain a more effective weight, which is expressed as follows:

$$w_H(s,r)=(1-\alpha)\,|D(s)-D(r)|+\alpha\,|I(s)-I(r)| \tag{13}$$

where the $\alpha$ is a parameter for balancing the relative contributions of color and disparity, *D* denotes the initial disparity map, and a value of $\alpha$=0.4 is assumed.

Either using the iterative MST algorithm to redefine the MST structure by Eqn. (13) or using the enhanced edge similarity function, the value of the weight are optimized in essence. However, extra computational complexity is inevitable by using Eqn. (13)

because the MST structure is needed to be built at least two times.

## 4. Experimental results

Five typical nonlocal cost aggregation methods are computed with the proposed method: MST [18], Segment-Tree (ST-2) [20], Cross-Scale (CS-MST) [21], Weighted Cost Propagation with Smoothness Prior (WCPSP) [27], and Iterative Color-Depth (MST-CD2) [22]. All of these methods are based on MST structure. The experiments are carried on the Middlebury *v.2 & v.3* stereo data sets[28-29].

### *4.1. Evaluation on Middlebury v.2 dataset*

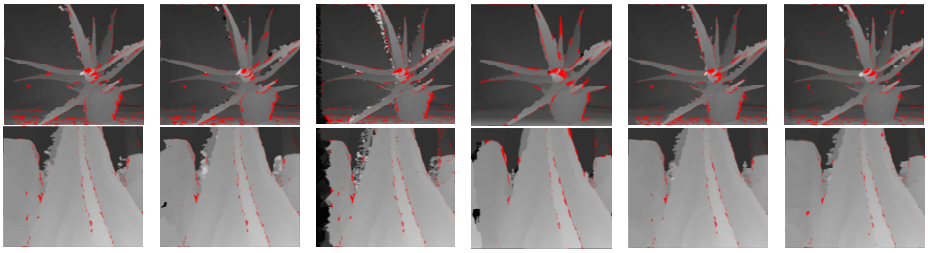
#### *4.1.1 Qualitative evaluation*

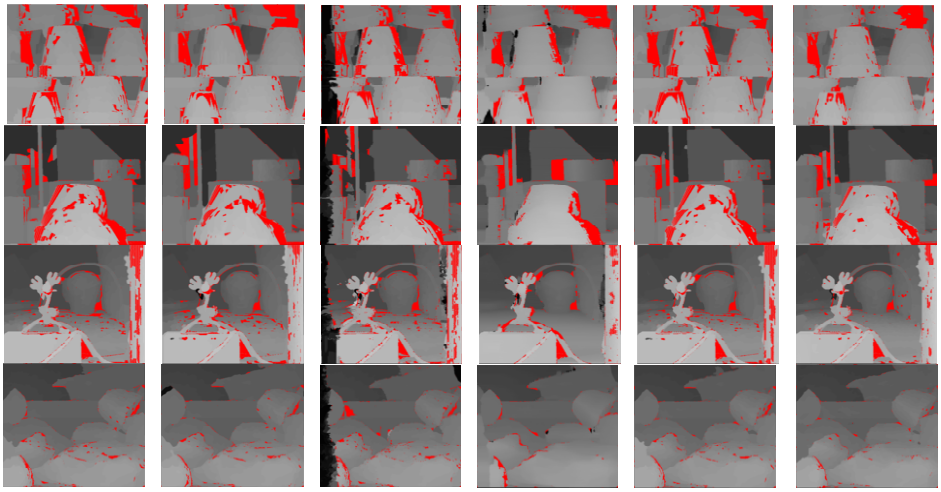
Six representative images pairs (*Aloe*, *Cloth4*, *Flowerpots*, *Lampshade1*, *Reindeer* and *Rocks1*) are selected to show the superior performance of the proposed method visually. The results are shown in figure 3, where the error pixels in non-occluded regions are marked in red. It can be observed from the disparity maps that the proposed method achieves more accurate disparity maps and reliable image boundaries especially in low-textured regions.

#### *4.1.2 Quantitative evaluation*

The quantitative evaluation results of the all 31 Middlebury v.2 stereo pairs in non-occluded regions are shown in table 1. It can be seen that MST has the worst results both in the average accuracy and rank, and the performance of CS-MST is slightly better than MST. Compared to MST and CS-MST, the performance has been significantly improved in other four methods, especially the proposed method, which achieves a tremendous advance and ranks 1 on 19/31 image pairs. WCPSP also has competitive performance and ranks 2 both in average errors and ranks. The proposed method show better accuracy performance than WCPSP for most of the image pairs.

The average time of MST, ST-2, CS-MST, WCPSP, MST-CD2 and the proposed method are shown in the last rows of table 1, respectively. It can be seen that MST is the most efficient method among the six methods while the proposed method is a bit slower than MST mainly because of the forward-backward diffusion. The overall runtime cost of the proposed method does not increase obviously in contrast to MST and is even shorter than ST-2, which ranks 3 among the six methods.





(a) (b) (c) (d) (e) (f)

**Figure 3.** Disparity maps on *Aloe*, *Cloth4*, *Flowerpots*, *Lampshade1*, *Reindeer* and *Rocks1* (from top to bottom row) by six different stereo matching algorithms. (a) MST. (b) ST-2. (c) CS-MST. (d) WCPSP. (e) MST-CD2. (f) Proposed method. The error pixels in non-occluded regions are marked in red for each disparity map and the error threshold is 1.0 pixel.

**Table 1.** Accuracy evaluation on all 31 Middlebury v.2 stereo pairs by six algorithms.

| | MST | ST-2 | CS-MST | WCPSP | MST-CD2 | Proposed |
|---|---|---|---|---|---|---|
| **Avg.Error** | $10.79_6$ | $9.17_3$ | $10.01_5$ | $8.69_2$ | $9.78_4$ | $\mathbf{8.04_1}$ |
| **Avg.Rank** | $4.71_6$ | $3.68_4$ | $4.68_5$ | $2.97_2$ | $3.26_3$ | $\mathbf{1.71_1}$ |
| **Avg.Time(s)** | $\mathbf{0.80_1}$ | $1.16_3$ | $4.17_6$ | $3.61_5$ | $2.28_4$ | $0.90_2$ |

### *4.2. Evaluation on Middlebury v.3 dataset*

The accuracy evaluation results of the 10 representative image pairs from Middlebury *v.3* data set in non-occluded regions are shown in table 2, where the five algorithms mentioned above and the proposed method are evaluated. It can be seen from table 2 that the proposed method outperforms the MST on all of the image pairs and performs the best in *Adirondack, ArtL*, *Jadeplant*, *Motorcycle*, *MotorcycleE* and *Pipes* over other five methods. The average percentage of error pixels and the average rank of the proposed method still ranks 1 on the image pairs. It is worth noting that in table 2 the average errors of MST-CD2 ranks 4, but its average rank is second only to the first. From statistical data, it can be seen that MST-CD2 also outperforms the MST on the ten image pairs and the performance of this method is stable. The average time of MST, ST-2, CS-MST, WCPSP, MST-CD2 and the proposed method on the ten image pairs are shown in the last rows of table 2, respectively. MST is still the most efficient method among the six methods. The overall runtime of the proposed method ranks 2 among the six methods.

**Table 2.** Accuracy evaluation on image pairs from Middlebury v.3 by six algorithms.

| Stereo Pairs | MST | ST-2 | CS-MST | WCPSP | MST-CD2 | Proposed |
|---|---|---|---|---|---|---|
| Adirondack | $12.85_6$ | $11.72_3$ | $12.52_4$ | $10.67_2$ | $12.82_5$ | $\mathbf{9.13_1}$ |
| ArtL | $10.01_3$ | $10.22_4$ | $14.80_6$ | $10.87_5$ | $9.99_2$ | $\mathbf{9.50_1}$ |
| Jadeplant | $16.91_3$ | $20.07_4$ | $20.78_5$ | $26.52_6$ | $16.43_2$ | $\mathbf{13.36_1}$ |
| Motorcycle | $4.14_3$ | $4.43_4$ | $6.48_6$ | $5.23_5$ | $3.87_2$ | $\mathbf{3.72_1}$ |
| MotorcycleE | $7.24_2$ | $7.79_4$ | $10.70_6$ | $8.09_5$ | $7.58_3$ | $\mathbf{6.81_1}$ |
| Piano | $20.15_5$ | $20.97_6$ | $17.60_2$ | $\mathbf{17.50_1}$ | $19.29_3$ | $19.81_4$ |
| Pipes | $10.17_4$ | $9.73_3$ | $11.65_5$ | $12.69_6$ | $9.61_2$ | $\mathbf{9.23_1}$ |
| Playtable | $35.04_5$ | $33.25_4$ | $\mathbf{19.33_1}$ | $31.48_3$ | $33.54_6$ | $30.95_2$ |
| Recycle | $6.94_5$ | $6.60_3$ | $6.96_6$ | $\mathbf{4.55_1}$ | $6.44_2$ | $6.70_4$ |
| Vintage | $38.74_6$ | $32.98_3$ | $34.87_4$ | $\mathbf{25.79_1}$ | $36.61_5$ | $29.63_2$ |
| **Avg.Error** | $16.22_6$ | $15.78_5$ | $15.57_3$ | $15.34_2$ | $15.62_4$ | $\mathbf{13.88_1}$ |
| **Avg.Rank** | $4.2_5$ | $3.8_4$ | $4.5_6$ | $3.5_3$ | $3.2_2$ | $\mathbf{1.8_1}$ |
| **Avg.Time(s)** | $\mathbf{2.21_1}$ | $3.11_3$ | $10.93_5$ | $11.24_6$ | $5.35_4$ | $2.44_2$ |

## 5. Conclusion

In this paper, a stereo matching algorithm using forward-backward diffusion and pruning-based cost aggregation is proposed. The proposed method is developed with the forward-backward diffusion and the pruning-based cost aggregation as well as the enhanced edge similarity function. The proposed method has some advantages. First, the raw stereo image pairs are smoothened on a weighted horizontal tree structure with "forward-backward" process, it allows depth edges to remain sharp while smoothening out noise. Second, the pruning method is used to cut off the MST structure self-adaptively so that every node receives supports only from similar depth regions. For the sake of suppressing the impact of small-weight-accumulation problem in textureless regions, an enhanced edge similarity function between two nearest neighboring nodes is formulated. Experimental results show that the proposed method could achieve better matching accuracy with a minor cost of increased execution time.

In the future, the proposed method is expecting further improvements, focusing on how to exact a more effective depth edge for pruning-based aggregation and how to reduce the complexity of the disparity refinement for high-resolution image pairs.

## References

[1] A. Kuhn, D. Scharstein, H. Mayer, A tv prior for high-quality scalable multi-view stereo reconstruction, Int. J. Comput. Vis. 124 (1) (2017) 1-16.

[2] T.-E. Bishop, P. Favaro, The light field camera: extended depth of field, aliasing, and superresolution, IEEE Trans. Pattern Anal. Mach. Intell. 34 (5) (2012) 972-986.

[3] Y. Li, K. Huang, L. Claesen, High-quality view interpolation based on depth maps and its hardware implementation, in: Proceedings of the International Conference on Field Programmable Logic and Applications, 2017, pp. 1-6.

[4] D. Scharstein, R. Szeliski, A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms, Int. J. Comput. Vis. 47 (1-3) (2002) 7-42.

[5] F. Tombari, S. Mattoccia, L.-D. Stefano, Classification and evaluation of cost aggregation methods for stereo correspondence, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-8.

[6] Y.-S. Heo, Two-step mutual information-based stereo matching, Electron. Lett. 52 (14) (2016) 1225-1227.

[7] J.-K. Suhr, H.-G. Jung, Dense Stereo-Based Robust Vertical Road Profile Estimation Using Hough Transform and Dynamic Programming, IEEE Trans. Intell. Transp. 16(3) (2015) 1528-1536.

[8] Q. Yang, W. Liang, R. Yang, Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling, IEEE Trans. Pattern Anal. Mach. Intell. 31 (3) (2009) 492-504.
[9] A. Arranz, A. SáNchez, M. Alvar, Multiresolution energy minimisation framework for stereo matching, IET Comput. Vis. 6 (6) (2012) 425-434.
[10] M. O'Byrne, V. Pakrashi, F. Schoefs, A stereo matching technique for recovering 3D information from underwater inspection imagery, Comput-aied Civ. Inf. 33 (3) (2018) 193-208.
[11] H. Hirschmüller, Stereo processing by semiglobal matching and mutual information, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2) (2007) 328-341.
[12] S. Perri, P. Corsonello, G. Cocorullo, Adaptive census transform: a novel hardware-oriented stereovision algorithm, Comput. Vis. Image Und. 117 (1) (2013) 29-41.
[13] Y. Qu, J. Jiang, X. Deng, Robust local stereo matching under varying radiometric conditions, IET Comput. Vis. 8 (4) (2013) 263-276.
[14] X. Mei, X. Sun, M. Zhou, On building an accurate stereo matching system on graphics hardware, in: Proceedings of the IEEE Conference on Computer Vision, 2012, pp. 467-474.
[15] C. Rhemann, A. Hosni, M. Bleyer, Fast cost-volume filtering for visual correspondence and beyond, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 3017-3024.
[16] K.-J. Yoon, I.-S. Kweon, Adaptive support-weight approach for correspondence search, IEEE Trans. Pattern Anal. Mach. Intell. 28 (4) (2006) 650-656.
[17] K. Zhang, J. Lu, G. Lafruit, Cross-based local stereo matching using orthogonal integral images, IEEE Trans. Circ. Syst. Vid. 19 (7) (2009) 1073-1079.
[18] Q. Yang, A non-local cost aggregation method for stereo matching, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 1402-1409.
[19] Q. Yang, Stereo matching using tree filtering, IEEE Trans. Pattern Anal. Mach. Intell. 37 (4) (2015) 834-846.
[20] X. Mei, X. Sun, W. Dong, Segment-tree based cost aggregation for stereo matching, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 313-320.
[21] K. Zhang, Y. Fang, D. Min, Cross-scale cost aggregation for stereo matching, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1590-1597.
[22] P. Yao, H. Zhang, Y. Xue, Iterative color-depth MST cost aggregation for stereo matching, in: Proceedings of the IEEE International Conference on Multimedia and Expo, 2017, pp. 1-6.
[23] X. Huang, Y.-j. Zhang, An O (1) disparity refinement method for stereo matching, Pattern Recognit. 55 (2016) 198-206.
[24] M. Bleyer, M. Gelautz, Simple but effective tree structures for dynamic programming-based stereo matching, in: Proceedings of the Third International Conference on Computer Vision Theory and Applications, 2008, pp. 415-422.
[25] H. Hirschmüller, Stereo processing by semiglobal matching and mutual information, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2) (2007) 328-341.
[26] C. John, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 30 (1986) 679-698.
[27] Q. Yang, Local smoothness enforced cost volume regularization for fast stereo correspondence, IEEE Signal Proc. Let. 22 (9) (2016) 1429-1433.
[28] D. Scjarstein and R. Szeliski, Middlebury stereo evaluation, 2002. Available < http://vision.middlebury.edu/stereo/eval/>.
[29] P. Yao, H. Zhang, Y. Xue, AGO: accelerating global optimization for accurate stereo matching, in: Proceedings of the International Conference on Multimedia Modeling, 2018, pp. 67-80.