Stereo Matching Based on Improved Matching Cost Calculation and Weighted Guided Filtering

Junxing Xu(✉), Wei He, and Zengshan Tian

School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
1791821966@qq.com

Abstract. Aiming at the problem that the existing local stereo matching algorithm has low matching accuracy in weak texture, disparity discontinuity and occlusion regions, an improved algorithm based on matching cost calculation and weighted guided filtering is proposed. The algorithm first improves the traditional gradient cost (GRAD) and Census transform, normalizes and fuses these two matching costs to form a new matching cost, then proposes a weighted guided filter based on the Kirsch operator and aggregates the matching cost, finally, the method of the winner-takes-all (WTA) is used to complete the disparity calculation, and we use the method of left and right disparity consistency and the quadratic curve interpolation to complete the disparity optimization and obtain the final disparity map. A large number of experiments prove that the proposed stereo matching algorithm has an average mismatch rate of about 5.45% relative to the standard disparity map on the test platform of Middlebury. Compared with most algorithms, proposed algorithm achieves a good matching effect.

Keywords: Census transform · Adaptive window · Kirsch operator · Weighted guided filtering

1 Introduction

Stereo matching is a process of recovering scene depth information by finding corresponding matching points of two images of the same scene under different viewing angles, and obtaining the disparity pixel by pixel. Stereo matching is widely used in 3D reconstruction [1], digital surface model generation [2], virtual reality and driverless [3].

Scharstei et al. [4] studied and summarized typical stereo matching algorithms and formed a basic stereo matching algorithm framework. The existing stereo matching algorithms are mainly divided into two categories: global stereo matching and local stereo matching algorithms. Global stereo matching algorithm is to solve the optimal solution in the global scope, although the global stereo matching algorithm has a high matching accuracy, the calculation is more complicated, so it has great limitations in practical application. Common global matching algorithms include confidence propagation [5] (BP), graph cut [6] (GC), minimum spanning tree [7] (MST), and split tree [8] (ST). The local stereo matching algorithm uses the neighborhood information in the

H. Gao et al. (Eds.): ChinaCom 2020, LNICST 352, pp. 490–504, 2021.
https://doi.org/10.1007/978-3-030-67720-6_34

window for matching, although the accuracy is poor compared to the global stereo matching, the time complexity is low, and easy to implement, so it is widely used. The steps of local stereo matching algorithm are mainly divided into four steps: matching cost calculation, matching cost aggregation, disparity calculation and disparity optimization. In recent years, the matching accuracy of some excellent local stereo matching algorithms has approached the global stereo matching algorithm, but the local stereo matching algorithm still faces two difficulties: selection of matching cost function and cost aggregation.

One difficulty of the local stereo matching algorithm is the selection of matching cost function, common matching cost functions mainly include pixel-based absolute value of gray difference (AD), gray difference square (SD), gradient-based measurement GRAD, sampling-insensitive BT algorithm [9]; Matching cost based on window mainly includes sum of absolute differences (SAD), sum of squared difference (SSD), normalized cross-correlation (NCC); Matching cost based on non-parametric transform [10] includes Census and Rank transform. GRAD is based on the gradient difference of the image, which can improve the matching accuracy of the weak texture regions to a certain extent, so that the disparity map becomes clearer, but traditional GRAD only considers the gradient value in the x direction, resulting in limited improvement in matching accuracy of weak texture regions. Census transform is an excellent stereo matching method, however, it relies heavily on the gray value of the central pixel, which makes it more sensitive to noise, and the size of support window is fixed, if the size of support window is selected too large, it will contain more texture information, which can improve the matching accuracy, however, the time complexity will also become higher; If the size of support window is too small, the texture information contained is too little, resulting in lower stereo matching accuracy, so selecting an appropriate support window is crucial. At the same time, through a lot of research, some scholars proposed that the improvement of the single measure function cannot further improve the matching accuracy, the fusion of multiple measure functions can significantly improve the matching accuracy, for example, literature [11] proposed a stereo matching algorithm combining SAD and Census transform, which achieves a good matching effect.

Another difficulty of the local stereo matching algorithm is the selection of cost aggregation method. Since the cost calculation only considers the local correlation, it is very sensitive to noise and cannot be directly used to calculate the disparity, therefore, cost aggregation is needed to improve the reliability of stereo matching. Traditional cost aggregation often uses box filtering and gaussian filtering, such filters are less difficult to implement, but the effect of maintaining the edge of the image is poor. In this regard, Yoon et al. [12] proposed to use bilateral filtering for cost aggregation, this method can effectively maintain the edge of the image, but the time complexity is relatively high. He et al. [13] proposed the use of guided filtering in cost aggregation, compared with bilateral filters, guided filtering can better maintain the edge of the image, and the time complexity is also independent of the size of window, but the traditional guided filter does not consider the different textures of different windows, it is simple to use the same regularization parameters for each window, resulting in a poor matching effect of the algorithm on regions with large texture changes.

In view of the above problems, in order to improve the accuracy of stereo matching, this paper proposes a stereo matching algorithm based on improved matching cost and weighted guided filter. The main contribution is: based on the traditional GRAD [14], the gradient values in the y direction and the two diagonal directions are fused to the GRAD, which increases more gradient information and makes the disparity map smoother and clearer. We replace the fixed window of traditional Census transform with adaptive window. The size of the window can be dynamically adjusted according to the standard deviation of the pixels in the window. At the same time, the method of confidence constraint is used to adjust the gray value of the center pixel in the window, and then these improved two matching costs are normalized and fused to achieve complementary advantages, which greatly improves matching accuracy; In the cost aggregation stage, a weighted guided filter based on the Kirsch operator is proposed, proposed algorithm dynamically adjusts the regularization parameter according to the different texture information in each window; Finally, the winner-takes-all (WTA) for disparity calculation, and the disparity refinement and optimization are performed by the method of left and right disparity consistency detection and the quadratic curve interpolation.

2 Matching Cost Calculation

2.1 Improved Census Transform

Census transform belongs to a kind of non-parametric transform algorithm, this algorithm does not use gray-scale information but neighborhood information for matching, so the matching accuracy of this algorithm is higher. The traditional Census transform first selects a pixel, and builds a rectangular window based on the pixel, turning the gray value of other pixels in the window except the central pixel into a bit string, the specific process is: compare the gray value of the pixel in each region with the gray value of the central pixel, if the value is less than the gray value of the central pixel, record it as 1, otherwise as 0. The formula for Census transform is as follows:

$$B_T(p) = \underset{q \in N_p}{\otimes} \xi(I(p), I(q)) \tag{1}$$

$$\xi(I(p), I(q)) = \begin{cases} 1, I(p) \geq I(q) \\ 0, I(p) < I(q) \end{cases} \tag{2}$$

Where p is the central pixel, N_p is the rectangular window with p as the center, q is the neighboring pixel of p in the window, $B_T(p)$ is the converted bit string, and $\otimes$ represents the bit-by-bit connection. The schematic diagram of Census transform is shown in Fig. 1.

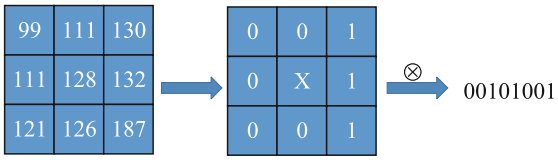

Fig. 1. Schematic diagram of Census transform

Since the Census transform is greatly affected by the central pixel and is susceptible to noise, it is very important to set the center pixel to an appropriate gray value. Some scholars proposed to use the average value of pixels in the window as a reference value, and other scholars proposed to use the median value of all pixels in the window as a reference value, but when the window is greatly affected by noise, the reference values obtained by these two methods are not effective, so an appropriate evaluation mechanism is needed to determine the reference pixel value, this paper proposes a judgment method based on confidence constraints, which improves the matching accuracy of the Census transform. The formula of confidence calculation is shown in (3)(4)(5):

$$R_{ave}(p) = |u(p) - I(p)| \tag{3}$$

$$R_{med}(p) = |m(p) - I(p)| \tag{4}$$

$$R(p) = R_{med}(p) + R_{ave}(p) \tag{5}$$

Where $u(p)$ is the average value in the rectangular window with p as the center pixel, $m(p)$ is the median value of all pixels in the rectangular window, $R_{ave}(p)$ is the mean confidence, $R_{med}(p)$ is the median confidence, and $I(p)$ is the gray value of the center pixel. The smaller the $R(p)$, the higher the confidence level of the gray value of $I(p)$ as the center pixel, which indicates that the gray value of the $I(p)$ as the center pixel of the rectangle is more reliable; When the confidence $R(p)$ is greater than a certain threshold, it is considered unreliable to use $I(p)$ as the central pixel value, and the gray value of the center pixel of the rectangle window needs to be corrected, the correction strategy is shown in (6):

$$I(p) = \begin{cases} \dfrac{u(p) + m(p)}{2} & R(p) > Th \\ I(p) & otherwise \end{cases} \tag{6}$$

When the confidence level $R(p)$ is greater than the confidence threshold Th, the gray value of the center pixel is updated to the average of $u(p)$ and $m(p)$, otherwise the gray value of the center pixel of the rectangle remains unchanged.

Another important factor that affects the matching accuracy of Census transform is the size of the support window. If the size of support window is smaller, the amount of information contained will be less, matching accuracy will be lower; If the size of window is larger, time complexity will become higher. Therefore, this paper proposes an adaptive window method, which achieves a good matching effect, the core idea is:

when the pixel standard deviation in the support window is large enough, the size of the current window is used directly, otherwise the window size is increased to obtain more information, thus achieving the balance of time complexity and matching accuracy. The specific process is shown in Fig. 2:

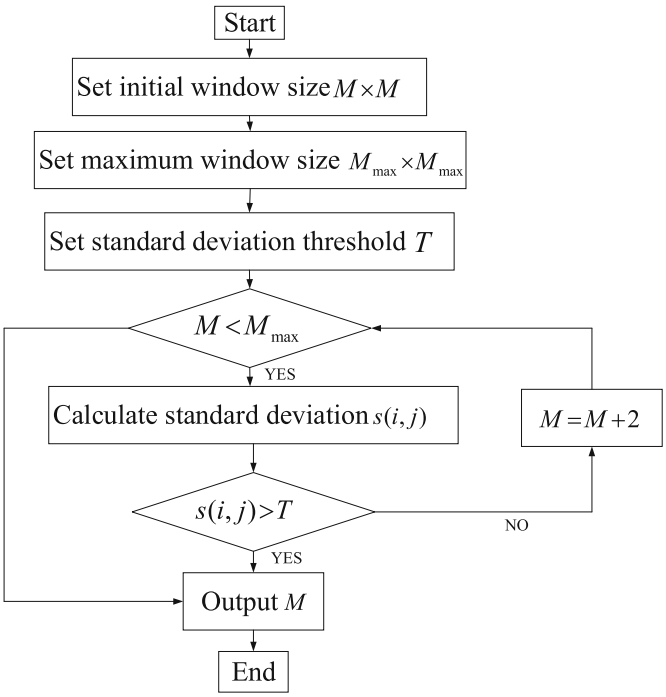

Fig. 2. Flow chart of adaptive window algorithm

The final improved Census transform is as follows:

$$B_T^{'}(p) = \underset{q \in N_p}{\otimes} \xi(I^{'}(p), I(q)) \tag{7}$$

$$\xi(I'(p), I(q)) = \begin{cases} 1, I'(p) \geq I(q) \\ 0, I'(p) < I(q) \end{cases} \tag{8}$$

Where $B_T'(p)$ is the bit string obtained by the improved Census transform, $I'(p)$ is the modified pixel reference value, and the meanings of the remaining symbols are the same as those in (1) and (2). When the disparity is d, the Hamming distance calculation formula of the two bit strings is:

$$H_h(p,d) = \min(Ham\min g(C_{CTL}(p), C_{CTR}(p-d)),\ T_{cen}) \tag{9}$$

Where $H_h(p,d)$ is the matching cost of Census transform, $C_{CTL}(p)$ is the Hamming value of the left image at pixel p, $C_{CTR}(p-d)$ is the Hamming value of the right image at pixel $p-d$, and T_{cen} is the threshold value of the Census transform, and we can come to a conclusion: the smaller the Hamming distance between two pixels, the higher the matching accuracy, otherwise the lower the matching accuracy. In order to further reduce the interference of noise, the matching cost is normalized:

$$C_{cen}(p,d) = 1 - \exp(-H_h(p,d)) \tag{10}$$

2.2 Enhanced Gradient Cost

In order to further obtain a high-precision disparity map, this paper improves the traditional GRAD algorithm, traditional GRAD only considers the gradient value in the x direction, which leads to a limited improvement in the matching accuracy of the weak texture regions, in this regard, this paper proposes an improved GRAD algorithm, the improved algorithm fuses the gradient information in the y and the two diagonal directions, the calculation formula is (11).

$$\begin{cases} grad_x = \partial G/\partial x \\ grad_y = \partial G/\partial y \\ grad_{45^\circ} = \partial G/\partial_{45^\circ} \\ grad_{135^\circ} = \partial G/\partial_{135^\circ} \end{cases} \tag{11}$$

$$grad = \sqrt{(grad_x)^2 + \left(grad_y\right)^2 + (grad_{45^\circ})^2 + (grad_{135^\circ})^2} \tag{12}$$

Where G is the gray scale of the image, $grad_x$ is the gradient of the image in the x direction, $grad_y$ is the gradient in the y direction, $grad_{45^\circ}$ and $grad_{135^\circ}$ are the gradients in the two diagonal directions, and $grad$ is the image gradient, then the gradient cost is:

$$grad_g(p,d) = \min\left(grad_L(p) - grad_R(p-d),\ \ T_g\right) \tag{13}$$

Where $grad_g(p,d)$ s the gradient cost when the disparity value of pixel p is d, $grad_L(p)$ is the gradient of the left image at pixel p, $grad_R(p-d)$ is the gradient of the right image at pixel $p-d$, and T_g is the gradient threshold. The normalized similarity measure function is:

$$C_g(p,d) = 1 - \exp(-grad_g(p,d)) \tag{14}$$

2.3 Matching Cost Fusion

In order to achieve the complementary advantages of the two measurement functions, the two measurement functions are weighted and fused, the fused matching costs are as follows:

$$C(p,d) = \lambda_1 C_{cen}(p,d) + \lambda_2 C_g(p,d) \tag{15}$$

Where λ_1 and λ_2 are the weight values of the improved Census transform and the enhanced gradient cost, and $\lambda_1 + \lambda_2 = 1$.

3 Cost Aggregation

After calculating the matching cost of the image, it is necessary to perform cost aggregation on the matching window. The essence of cost aggregation is filtering, which is used to filter out the noise introduced in the matching cost calculation. Among all the methods of cost aggregation, the guided filtering is an excellent one, The process of guiding the filtering algorithm is as follows: suppose I is the guide image, usually the left image is the guide image, the input image is p, U is the filtered image, the filtered linear conversion model is:

$$U_i = a_k I_i + b_k, \quad \forall\, i \in w_k \tag{16}$$

Where k is the center pixel in the window, w_k is a rectangular window centered on pixel k, U_i and I_i are the corresponding points before and after the guided filtering. a_k and b_k are the linear coefficients, a_k and b_k can be obtained by minimizing (17).

$$E(a_k, b_k) = \sum_{i \in w_k} \left[(a_k I_i) + b_k - p_i \right)^2 + \varepsilon a_k^2] \tag{17}$$

Where ε is the regularization parameter, usually $\varepsilon > 0$, and the larger the value, the smoother the filtered image; The smaller the value, the better the edge of filtered image. The solution of (17) can be obtained by the least square method.

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i p_i - u_k \overline{p}_k}{\sigma_k^2 + \varepsilon} \tag{18}$$

$$b_k = \overline{p}_k - a_k u_k \tag{19}$$

Where u_k and σ_k^2 represent the mean and variance of the guide image I in the window w_k; $|w|$ represents the total number of pixels in the window w_k; $\overline{p}_k$ is the average gray value of the input image p in the window w_k.

However, the traditional guided filtering algorithm does not consider the difference in pixel texture of different windows, and each window uses the same regularization parameters, resulting in limited improvement in matching accuracy. Aiming at the

problem, Li et al. [15] improved the guided filtering, and proposed a weighted guide filtering, which defines the pixel variance in the window as a weighted factor, and adaptively adjusts the regularization parameters. However, defining the variance as a weighted factor does not highlight the edges of the image very well, so this paper proposes to define the Kirsch operator [16] as a weighted factor, the Kirsch operator can detect the edges in the image very well, in the edge regions of the image, a larger amplitude response can be obtained, in the smooth region of the image, the amplitude response is close to 0.

The edge weight of Kirsch operator is defined as:

$$\Gamma_I(p) = \frac{1}{N}\sum_{p'=1}^{N}\frac{KIR(p)}{KIR(p')} \tag{20}$$

Where $KIR(p)$ is the edge value of p corresponding to Kirsch operator, I is the guide image, N represents the total number of pixels, (18) can be replaced with (21).

$$a_k = \frac{\frac{1}{|w|}\sum_{i\in w_k} I_i p_i - u_k \overline{p}_k}{\sigma_k^2 + \frac{\varepsilon}{\Gamma_I}} \tag{21}$$

4 Disparity Calculation

The final matching cost obtained after cost aggregation is $C'(p,d)$. We use the WTA strategy to select the minimum matching cost correspondence for each pixel as the initial disparity value.

$$d_p = \arg\min_{d\in D}[C'(p,d)] \tag{22}$$

Where D represents all possible disparity values.

5 Disparity Refinement

There are many mismatched points in the initial disparity map obtained by the WTA method, so post-processing optimization is needed to reduce the number of mismatched points. Post-processing optimization methods mainly include left and right consistency detection, occlusion processing, weighted median filtering and sub-pixel refinement [17, 18]. The way to detect mismatched points through left and right consistency detection is as follows.

$$|d_L(p) - d_R[p - d_L(p)]| < 1 \tag{23}$$

Where $d_L(p)$ is the disparity value of pix p in the left view, and $d_R[p - d_L(p)]$ is the disparity value of pix $p - d_L(p)$ in the right view. If the left and right disparity value are not equal, it is considered as an anomaly point, and at the same time, the anomaly points are divided into occlusion points and mismatched points according to the principle of epipolar geometry [19].

Different interpolation methods are used for the occlusion points and mismatched points obtained by the left and right consistency detection. For the occlusion pixels, since the occlusion regions are usually located in the background of the image, the disparity value from the non-occlusion region pixels is used to interpolate the occlusion point. We search the effective pixel closest to the occlusion pixel in 8 directions, and use its disparity value as the disparity value of the current occlusion point; For mismatched pixels, the disparity value of the effective pixel with the most similar pixel color is used as the disparity value of the mismatched pixel in 8 directions, and the weighted median filter [20] is used to smooth the interpolated disparity map. In order to reduce discontinuity of the disparity map caused by discrete disparity, sub-pixel estimation is performed based on binomial polynomial interpolation [21]. For each pixel, the best sub-pixel interpolation value is:

$$d_{sub} = d - \frac{S(p, d_+) - S(p, d_-)}{2[S(p, d_+) + S(p, d_-) - 2S(p, d)]} \tag{24}$$

Where d is the disparity value after pixel p weighted median filtering; $S(p, d_+)$ and $S(p, d_-)$ are the cost aggregation of pixel p when the disparity is $d+1$ and $d-1$, respectively.

6 Experimental Results and Analysis

In order to evaluate the matching effect of the proposed algorithm, this paper uses the data set provided by the Middlebury [22] experimental platform for experiments. The simulation experiments are performed on a computer with a CPU of Intel Core i3-3210M, 2.5 GHz, and 8G of memory, and the programming environment is Matlab R2016a. The parameter settings involved in the experiment are shown in Table 1, the following dates are obtained through many experiments.

Table 1. Parameter settings

Parameter	Value	Parameter	Value	Parameter	Value
Th	6	T_{cen}	0.06	T	5.23
M	0.39	T_g	0.35	λ_2	0.55
$M_{\max}$	11	λ_1	0.45	ε	0.012

6.1 Matching Cost Calculation Experiment

In order to verify the effectiveness of the improved matching cost, this paper selects two cost functions to compare with the proposed cost function, these two cost functions are SAD + GRAD [23] and AD + Census [17]. The experimental images are 4 sets of stereo matching image pairs (Cones, Tsukuba, Teddy, Venus). Figure 3 is the result of different cost functions experiments, the experiments do not undergo any post-processing, Table 2 is error matching rates of different algorithms for different images, Fig. 4 is the histogram of the average mismatch rate of each cost function.

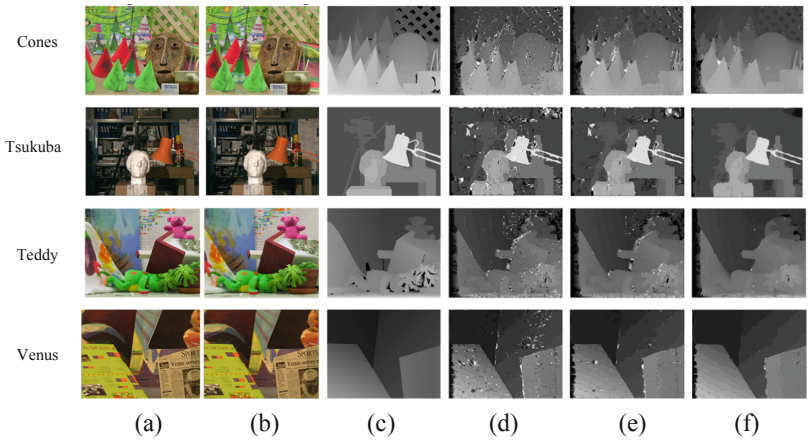

Fig. 3. Disparity map of different cost functions. (a) left image; (b) right image; (c) real disparity map;(d) SAD + GRAD;(e) AD + Census; (f) Proposed

Table 2. Error matching rates of different cost functions for different images %

Algorithm	Cones	Tsukuba	Teddy	Venus
SAD + GRAD	27.165	23.898	24.554	22.441
AD + Census	19.145	17.984	19.689	18.356
Proposed	13.895	11.546	14.262	13.654

As can be seen from Fig. 3, compared with the matching cost of SAD + GRAD and AD + Census, the proposed cost function shows a better matching effect in four different pairs of stereo matching images, and it can be seen from Table 2 and Fig. 4 that the proposed cost function has a lower mismatch rate, which proves the effectiveness of the proposed cost function.

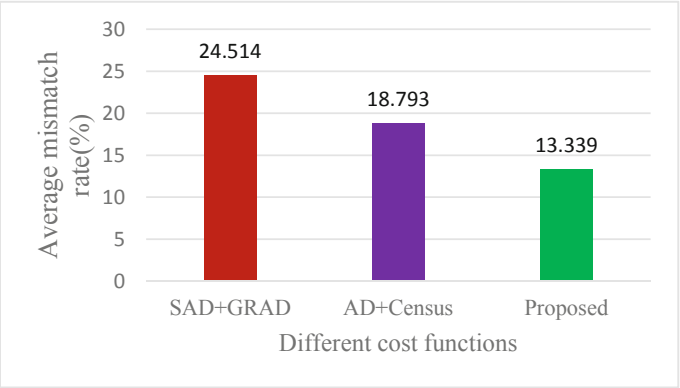

Fig. 4. Average mismatch rate of different cost functions

6.2 Cost Aggregation Experiment

In order to verify that the proposed weighted guided filtering has an excellent effect in maintaining edges, this paper compares literature [15] with the proposed weighted guided filter. The experimental images are 2 sets of stereo matching image (Sawtooh, Barn2). Figure 5 shows the experimental results of the proposed weighted guided filter and literature [15] in 2 different stereo matching image pairs. Figure 5 (a) is the left view of the stereo matching image pair, (b) is the right view of the stereo matching image pair, (c) is the real disparity map, (d) is the disparity map of literature [15], (e) is the disparity map of the proposed weighted guided filter.

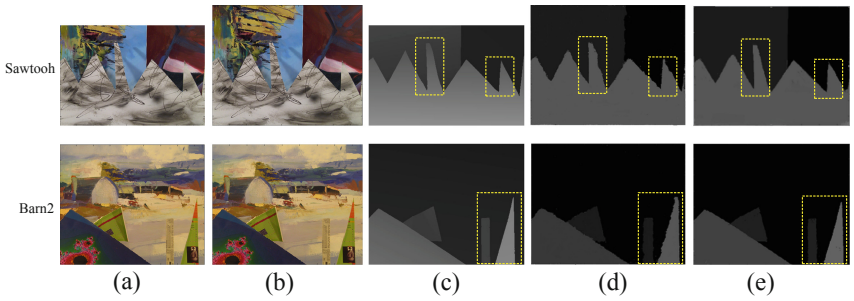

Fig. 5. Disparity map of different cost aggregation algorithms. (a) left view; (b) right view; (c) real disparity map; (d) literature [15]; (e) proposed weighted guided filter

As shown in Fig. 5, in the regions marked by the rectangular frame, the weighted guided filter proposed in this paper has a better effect in maintaining edges than the literature [15], which proves the effectiveness of the improved weighted guided filter proposed in this paper.

6.3 Algorithm Comparison Analysis

In order to further verify the superiority of the proposed algorithm, four common stereo matching algorithms are selected to perform experiments on four sets of stereo matching image pairs (Venus, Cones, Tsukuba, Teddy). These four algorithms include semi-global stereo matching [23] (SGM), cross-scale minimum spanning tree [14] (CS-MST), block matching [24] (BM), BP algorithm. The experimental simulation diagrams are shown in Fig. 6, the error matching rates of different algorithms for different images are shown in Table 3, Table 4, and where ***n-occ*** represents the ratio of mismatched pixels in non-occluded regions, ***all*** represent the total pixel mismatch rate, ***disc*** represents the mismatch rate in the disparity discontinuity regions, Fig. 7 is the histogram of the average mismatch rate of each stereo matching algorithm.

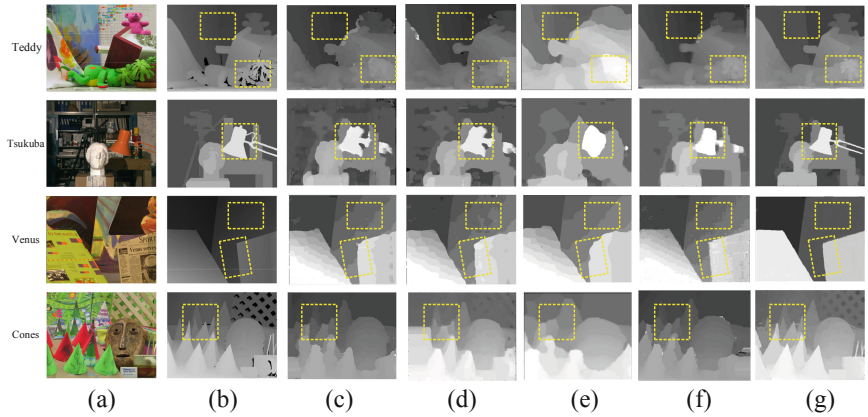

Fig. 6. Disparity map of different algorithms. (a) Left image (b) Real disparity map (c) CS-MST (d) SGM (e) BM (f) BP (g) Proposed

Table 4. Error matching rates of different algorithms for different images%

Algorithm	Venus			Cones		
	n-occ	all	disc	n-occ	all	disc
CS-MST	8.654	10.562	14.236	10.445	12.886	15.234
SGM	11.565	14.784	17.889	16.223	18.442	23.896
BM	15.447	17.845	21.246	22.445	25.221	30.568
BP	4.245	6.231	9.556	6.315	9.564	11.235
Proposed	3.562	5.556	8.233	4.225	5.253	7.565

Table 3. Error matching rates of different algorithms for different images%

Algorithm	Teddy			Tsukuba		
	n-occ	all	disc	n-occ	all	disc
CS-MST	6.235	9.254	11.556	8.369	10.546	13.456
SGM	7.453	10.545	13.563	10.656	13.895	16.235
BM	20.231	25.235	30.231	25.336	29.256	33.445
BP	5.562	7.236	9.235	5.356	6.745	8.754
Proposed	4.535	5.652	8.123	4.114	5.852	7.565

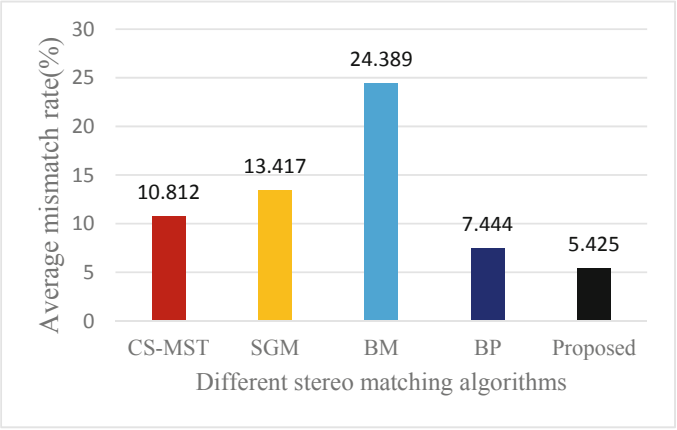

Fig. 7. Average mismatch rate of different stereo matching algorithms

As can be seen from Fig. 6, compared with SGM, CS-MST and other algorithms, the proposed algorithm shows good matching accuracy in weak texture, disparity discontinuity and occlusion regions. As shown in the Cones test chart, the part marked by the rectangular frame is a disparity discontinuity regions, the proposed algorithm achieves a smoother disparity in this regions. The part marked by rectangles in the Teddy test chart are the weak texture and the occlusion regions, the disparity of proposed algorithm in these regions is more clear and smooth, and in Venus test chart, the proposed algorithm is better in the edge regions. Similarly, the algorithm of this paper is better than other algorithms on the test charts of Tsukuba, which proves the effectiveness of proposed algorithm.

As can be seen from Table 3, Table 4 and Fig. 7, compared with other algorithms, proposed algorithm has better matching effect in multiple image regions. The average mismatch rate of CS-MST, SGM, BM, BP and proposed algorithm is 10.812%, 13.417%, 24.389%, 7.444%, 5.425%, respectively, which shows the superiority of proposed algorithm.

7 Conclusion

Based on the existing local stereo matching algorithm, this paper proposes a stereo matching algorithm with higher matching accuracy. In the matching cost calculation stage, two major improvements are made. ① the gradient information in the *y* direction and the two diagonal directions is fused to the traditional GRAD, which improves matching accuracy of weak texture regions. ② Based on the traditional Census transform, a confidence constraint method is proposed to modify the central pixel value in the support window, at the same time, the size of the support window is dynamically adjusted by the method based on gray standard deviation, then these two cost functions are weighted and fused, which further improves the matching accuracy. In the cost aggregation stage, a weighted guided filtering algorithm based on Kirsch operator is proposed, which maintains the edges of the image. A large number of experimental results show that compared with other algorithms, proposed algorithm can effectively enhance the matching accuracy of weak texture, disparity discontinuity and occlusion regions.

References

1. Shen, S.: Accurate multiple view 3D reconstruction using patch-based stereo for large-scale scenes. IEEE Trans. Image Process. **22**(5), 1901–1914 (2013)
2. Hamzah, R.A., Rahim, R.A., Rosly, H.N.: Depth evaluation in selected region of disparity mapping for navigation of stereo vision mobile robot. In: 2010 IEEE Symposium on Industrial Electronics and Applications (ISIEA), Penang, pp. 551–555 (2010)
3. Irijanti, E., Nayan, M.Y., Yusoff, M.Z.: Fast stereo correspondent using small-color census transforms. In: Proceedings of the 4th International Conference on Intelligent and Advanced Systems, pp. 685–690 (2012)
4. Scharstein, D., Szeliski, R., Zabih, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In: Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001), Kauai, HI, USA, pp. 131–140 (2001)
5. Sarkis, M., Diepold, K.: Sparse stereo matching using belief propagation. In: 2008 15th IEEE International Conference on Image Processing, San Diego, CA, pp. 1780–1783 (2008)
6. Kolmogorov, V., Zabih, R.: Computing visual correspondence with occlusions using graph cuts. In: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, Vancouver, BC, Canada, vol. 2, pp. 508–515 (2001)
7. Yang, Q.: A non-local cost aggregation method for stereo matching. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, pp. 1402–1409 (2012)
8. Mei, X., Sun, X., Dong, W., Wang, H., Zhang, X.: Segment-tree based cost aggregation for stereo matching. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, pp. 313–320 (2013)
9. Birchfield, S., Tomasi, C.: Depth discontinuities by pixel-to-pixel stereo. In: Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), Bombay, India, pp. 1073–1080 (1998)
10. Zabih, R., Woodfill, J.: Non-parametric local transforms for computing visual correspondence. In: 3rd European Conference on Computer Vision, vol. 2, pp. 151–158 (1994)

11. Zhou, J., Ying, W., Meng, L.: A new stereo matching algorithm based on adaptive weight SAD algorithm and Census algorithm. Bull. Survey. Map. **0**(11), 11–15 (2008)
12. Yoon, K.-J., Kweon, I.S.: Adaptive support-weight approach for correspondence search. IEEE Trans. Pattern Anal. Mach. Intell. **28**(4), 650–656 (2006)
13. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35** (6), 1397–1409 (2013)
14. Zhang, K., et al.: Cross-scale cost aggregation for stereo matching. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, pp. 1590–1597 (2014)
15. Li, Z., Zheng, J., Zhu, Z., Yao, W., Wu, S.: Weighted guided image filtering. IEEE Trans. Image Process. **24**(1), 120–129 (2015)
16. Mani, D.S., Nagaraju, C.: Face recognition based on kirsch compass kernel operator. In: 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, pp. 1322–1324 (2017)
17. Mei, X., Sun, X., Zhou, M., Jiao, S., Wang, H., Zhang, X.: On building an accurate stereo matching system on graphics hardware. In: 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, pp. 467–474 (2011)
18. Hosni, A., Rhemann, C., Bleyer, M., Rother, C., Gelautz, M.: Fast cost-volume filtering for visual correspondence and beyond. IEEE Trans. Pattern Anal. Mach. Intell. **35**(2), 504–511 (2013)
19. Hirschmuller, H.: Stereo processing by semiglobal matching and mutual information. IEEE Trans. Pattern Anal. Mach. Intell. **30**(2), 328–341 (2008)
20. Ma, Z., He, K., Wei, Y., Sun J., Wu, E.: Constant time weighted median filtering for stereo matching and beyond. In: 2013 IEEE International Conference on Computer Vision, Sydney, NSW, pp. 49–56 (2013)
21. Yang, Q., Yang, R., Davis, J., Nister, D.: Spatial-depth super resolution for range images. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, pp. 1–8 (2007)
22. Yang, Q., Wang, L., Yang, R., Stewénius, H., Nistér, D.: Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling. IEEE Trans. Pattern Anal. Mach. Intell. **31**(3), 492–504 (2009)
23. Hirschmuller, H.: Accurate and efficient stereo processing by semi-global matching and mutual information. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, vol. 2, pp. 807–814 (2005)
24. Hamzah, R.A., Hamid, A.M.A., Salim, S.I.M.: The solution of stereo correspondence problem using block matching algorithm in stereo vision mobile robot. In: 2010 Second International Conference on Computer Research and Development, Kuala Lumpur, pp. 733–737 (2010)