An Adaptive Window Stereo Matching Based on Gradient

Yan He
College of Mechanical and Electronic Engineering
Shanghai Normal University
Shanghai, China
abche111@163.com

Pei Wang and Jie Fu
College of Mechanical and Electronic Engineering
Shanghai Normal University
Shanghai, China
peiwang@shnu.edu.cn

***Abstract*—The key of local matching algorithm is the selection of similarity measure function and window's size, this paper through the two aspects to do improvement. Due to the traditional adaptive window algorithm is relatively complex, this paper extracted a relatively simple adaptive window algorithm, that is to say the window size based on the gradient points, and adopt a new measure function instead of the SAD function, not only to reduce noise effects but also to improve matching accuracy. In addition, in order to get a better disparity map, we adopt a regional seed propagation method which based on smooth constraint and a hypothesis that adjacent pixels with similar color should have the same disparity. Experiments show that the algorithm is compared with the traditional matching algorithm of adaptive window, computational complexity is reduced and the matching accuracy improved.**

Keywords- stereo matching; adaptive window; disparity map

I. INTRODUCTION

Stereo matching problem is one of the key problems of computer vision. All over the world, many scholars in this field studied this issue. Among them, *Scharstein* and *Szeliski* [1] do some research and evaluation for typical stereo matching technology, which can be divided into global algorithm and local algorithm. In general, the global algorithm accuracy is higher, but the calculation is large and parameter set complex. Local algorithm has higher efficiency and easy to implement, but the problem of how to determine the window's size and shape adaptively cause its hard to achieve high matching accuracy. In order to solve this problem, many scholars put forward typical local algorithms, the improved local algorithm in general can be divided into two categories [2]:

The first kind of algorithm is mainly focus on selecting the optimal window as support window from the given multiple windows [3-4], or point by point to select support window's size and shape adaptively [5-6]. To some extent, these methods improved the precision of disparity map in the area where it's not easy to match, suppresses the foreground fattening effect, but due to the shape of the window with the fixed or limited, lack of flexibility, so it is difficult to adapt to the changing of the structure of image, matching error rate are still high, disparity edge is not clear enough.

The second method [7-10, 15] is mainly to adopt different weighted strategies for pixels in constant support windows. In Paper [7], according to each pixel in the given window with matching pixel colors and geometric relations proposed an adaptive weighting method, which can effectively structure match cost, greatly reduce the matching ambiguity, the disparity map can compete with global optimization result, but the computation is big, this method could not reflect the efficient advantages of local algorithm. *Federico* carefully analysis of the deficiency of the adaptive weighting method [8], using color segmentation information improves the weighting function, improve the accuracy of the algorithm, but further increased the operation cost.

Therefore, under the condition of the trade-offs between efficiency and precision, this paper proposes adaptive windows of all different shapes and sizes based on different gradient. Compared with traditional methods, its computational complexity is reduced, but the matching precision is improved. In order to get a better performance, we use the method of seed propagation method, and to introduce a new similar measure function which has good noise resistance. The experiments show that the algorithm can improve the quality of image matching and compared with the traditional algorithm, computational complexity is reduced.

II. THE PROPOSED ALGORITHM

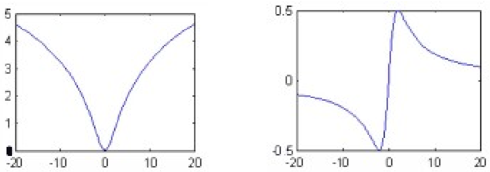
A. Similar measure function

In the local matching algorithm, for commonly used similarity measure function as SAD, SSD can produce error matching, when the pixel grayscale distribution is different from normal points or there is a big noise in the image. In order to solve this problem, here introducing a similar measure function which is based on $\rho_\sigma(n)$ function, its expressions and its derivative function of the mathematical expressions as follows:

$$\rho_\sigma(n) = \log\left[1+\frac{1}{2}\left(\frac{n}{\sigma}\right)^2\right] \quad (1)$$

$$\psi_\sigma(n) = \frac{2n}{2\sigma^2+n^2} \quad (2)$$

In which $n(x,y) = I_L(x,y) - I_R(x+d,y)$, σ denotes the scale parameter. The curve of the function and its derivative when σ =2 as follows:



a. the curve of function   b. the curve of function derivative

Figure 1. The curve of function $\rho_\sigma(n)$ and its derivative

The figure shows that $\rho_\sigma(n)$ function can make the noise almost does not affect the quality of the image, so we can get very good treatment effect. In this paper, we use a new similar measure function to instead SAD function,

$$NSAD(d) = \sum_{xy} \rho_\sigma\left(I_L(x,y) - I_R(x+d,y)\right) \qquad (3)$$

The disparity value: $d = \arg\min\left(NSAD(d)\right)$

B. Adaptive Window

In local matching, how to choose the size of the window is very important. On the one hand, the window must large enough to contain abundant information, because the small window can not accurate characterization of the image gray level change and reduce the precision of matching results. On the other hand, too large window will cause the distortion of image projection by the large span of disparity. This will not be able to accurately determine the matching position. So based on the advantages and disadvantages, local matching requires that the selected window must be with strong ability to adapt the shape of image. So in this paper, we propose adaptive windows of all different shapes and sizes based on different gradient.

Firstly, a certain pixel $p(i,j)$ in the image should calculate its gradient $G(i,j)$, horizontal gradient $G_h(i,j)$ and vertical gradient $G_v(i,j)$ .In this paper, the choice of window is as follows:

1. When $G(i,j) > \alpha$ and $G_h(i,j) - G_v(i,j) > \beta$, choose 3*9 as the matching window size;

2. When $G(i,j) > \alpha$ and $G_v(i,j) - G_h(i,j) > \beta$, choose 9*3 as the matching window size;

3. When $G(i,j) > \alpha$ and not in the first two conditions, choose 3*3 as the matching window size;

4. When $G(i,j) \le \alpha$, choose 9*9 as the matching window size.

C. Seed Point Detection

In this step, we obtain highly reliable seed pixels from the original disparity map, it as a good starting point of spread process. By detecting the pixel of constraint conditions of consistency and belief evaluation to determine whether it's a seed point s_0, conditions as:

1. $D_l(s_0) = D_r(s_0 - (D_l(s_0), 0))$
2. $C(s_0, d) > \lambda_c C(s_0, D_l(s_0))$, $\forall d \ne D_l(s_0)$

The first condition means that seeds should satisfy the detection of the consistency between left view and right view. The second condition is the belief detection on seeds if their matching cost is obviously smaller than other pixels. Besides λ_c is the threshold value for evaluating the reliability of disparity. In this experiment, we set λ_c=0.8.

D. Seeds Growth

Firstly, execute scan line propagation algorithm [11], which based on a hypothesis, namely in every segment of neighboring pixels with similar color and similar disparities. Secondly, use the method of regional seed growth. In addition to the boundary, disparity map should change slowly no abrupt change, so we use disparity smooth constraint.

In scan line propagation algorithm [11], according to color similarity, create a line segment along the scan line direction in each pixel of the stereo image pair. For an unseeded pixel p, we first search the nearest seed pixels (s_1, s_2) within its line segment from left and right. If only one seed point is found in the segment, $D_l(s_1)$ or $D_l(s_2)$ can simply replace $D_l(p)$, there is no more reliable pixel can be used for disparity estimation. If two available seed point are found, the propagation strategy as follows:

1. $D_l(p) = \min(D_l(s_1), D_l(s_2))$. If p is in occluded regions, it fails in the left-right consistency check.
2. $D_l(p) = \min(D_l(s_1), D_l(s_2))$. If p is near depth discontinuous, $|D_l(s_1) - D_l(s_2)| > \alpha \cdot d_{\max}$, $d_{\max}$ denotes the maximum disparity level. p is in depth discontinuous area.
3. In all other conditions, $D_l(p)$ is replaced by the linearly interpolated results of $D_l(s_1)$ and $D_l(s_2)$.

While pixel p is updated, it is immediately marked as seed point. If no seed point was found in its segment area, p remain unchanged. Then we search the nearest unseeded points from p 's left mark as s_l, right mark as s_r, up mark as s_u, and down mark as s_d. And growth strategy as follows:

1. If the gradient of p is zero and the disparities in three or four directions are the same, $D_l(p)$ can be replaced by the same disparity of the three or four directions. This step is based on the disparity smoothness constraints.

2. If the color difference between p and its left seed s_l is within a certain threshold, we can use the disparity of s_l to instead $D_l(p)$. In the same way, if its right seed s_r, up seed s_u or down seed s_d also satisfy this condition, the disparity of the satisfied seed can replace $D_l(p)$. This step is based on the assumption that neighboring pixels with similar colors should have similar disparities in each line segment.

Similarly, once update the pixel p, it is marked as seed point immediately. After executing seed growth process, the unseeded pixels updated with the smallest disparity values of recent seed point. After the above processing, the black part of the disparity map will be removed, and the part of the zero disparity is replaced by their nearest nonzero disparity value. In order to get better effect of disparity map, in this paper, we will do precise refinement of disparity.

E. *Disparity precise refinement*

It can be divided into two steps for this part:

1) Vertical vote:

Because the disparity computation depend on horizontal line segments, horizontal striates can be reduced by this way. We construct a ballot box $H_p(d)$ for pixel p with $D_{\max}+1$ bins, $D_{\max}$ is the maximum disparity level. Then we collect votes in a vertical line segment which in the range between $p-(0,N/2)$ and $p+(0,N/2)$. For any pixel in the line, if the color difference $D_c(p,q)$ is less than threshold τ, the vote of disparity $D_l(q)$ increase by 1. And the final result of $D_l(q)$ use $D_l(p)=\arg\max_d H_p(d)$ to update. After the 1D vote processing, most horizontal streaking artifacts in large areas can be reduced.

2) Median filtering:

In order to get more reliable disparity map, we take the median filtering method to eliminate isolated point. In this paper, we adopt 2D square median filter in this paper.

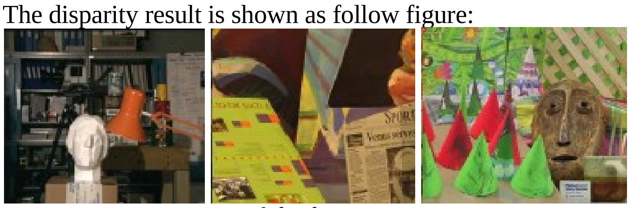
III. EXPERIMENT RESULTS

In order to verify the effectiveness of the algorithm, the algorithm on MATLAB 7 platform test and analysis. Experiment of image and real disparity map come from Middlebury dataset. Table 1 shows the three kinds of the algorithm's error matching rate. From the table we can know that the result of this algorithm is similar to the other two algorithms in literature [9] and literature [12], but the complexity of this algorithm is significantly lower than them. And we can see that this algorithm has a great superiority in the result of "Tsukuba".

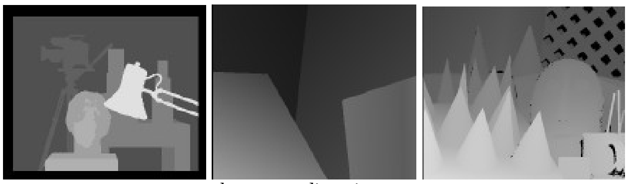
TABLE I. ERROR MATCHING RATE FOR MIDDLEBURY

Algorithm	Tsukuba	Venus	Cones
In this paper	3.12%	1.56%	13.52%
In literature[9]	5.10%	1.62%	12.30%
In literature[12]	6.08%	2.48%	10.20%

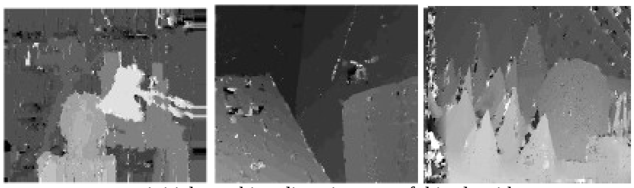
The disparity result is shown as follow figure:



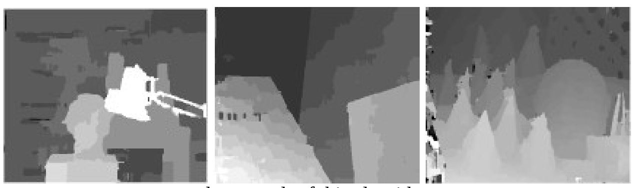
a. left reference view;



b. true disparity map



c. initial matching disparity map of this algorithm

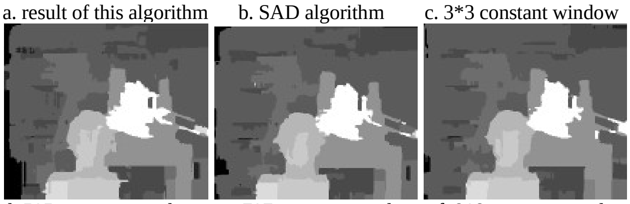


d. result of this algorithm

Figure 2. Results for Tsukuba, Venus and Cones scene;

In order to verify the complexity of the algorithm, on the one hand, we change the adaptive window to constant window (3*3, 5*5, 7*7, 9*9). On the other hand, we change the similarity measure function to SAD. Then compare them, result as the figure 3.

a. result of this algorithm b. SAD algorithm c. 3*3 constant window



d. 5*5 constant window e. 7*7 constant window f. 9*9 constant window

Figure 3. Matching results for different matching windows

Figure 4 shows the matching error rate of different windows and different similarity measure function. In figure 5, we can compare the running time of 'Tsukuba' and 'Venus'. Therefore, this algorithm is better than the constant window and SAD algorithm, it can get more accurate disparity map. On the processing speed, except the processing time is longer than the small window algorithm like 3*3 and 5*5, this algorithm cost the same amount of time as big window algorithm. Through the above analysis, we can know that the algorithm is not complicated calculations to estimate the window size. The gradient decides the size and shape of the window. And the new similarity measure function instead of SAD function can get better parallax effect.

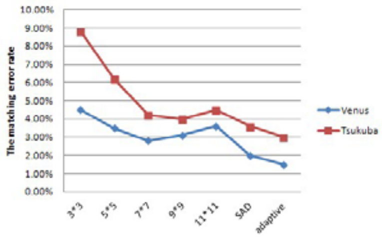


Figure 4. The matching error rate of different windows

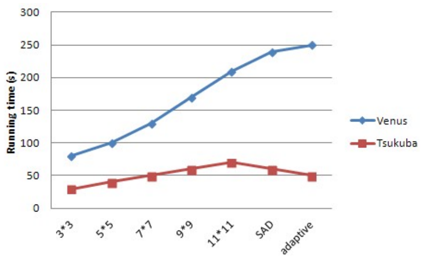


Figure 5. Running time of different windows

IV. CONCLUSIONS

This paper proposed a simple matching algorithm of adaptive window, which is based on the gradient to select the different size and shape of the window. Especially focused on the horizontal or vertical concentration gradient region, we use the narrow width and broad height window in the horizontal gradient concentrated concentration areas, and use the broad width and narrow length in the horizontal gradient concentration areas. Compared with the traditional adaptive window matching algorithm, the algorithm is improved significantly lower computational complexity and accuracy. In addition, we propose a new similar measure function to replace SAD function, making the algorithm have a certain ability to resist noise. After get the initial disparity map, we adopt the seed region propagation method to obtain more accurate results. The method is based on the assumption that neighboring pixels of similar color should have the same disparity and the smooth constraint of disparity. The experiment indicated that, the algorithm has higher precision and computation complexity is greatly reduced.

ACKNOWLEDGMENT

This work is Supported by the Program of Shanghai Normal University (DZLA-7001-12-002006), Supported by Shanghai Normal University (SK201127).

REFERENCES

[1] Scharstein D, Szeliski R. A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms[J]. International Journal ofComputer Vision,2002,47(112/3):7~42

[2] Zhang K, LU J B, Lafruit G. Cross–based local stereo matching using orthogonal integral images[J]. IEEE Transactions on Circuits and Systems for VideoTechnology, 2009, 19(7):1073-1079.

[3] Xiuzhi Zhou, Wen G J,Wang R SH. Fast stereo matching using adaptive window[J]. Chinese Journal of Computer, 2006, 29(3): 473-479.

[4] Fusiello A, Roberto V. Efficient stereo with multiple windowing[A]. IEEE International Conference on Computer Vision and Pattern Recognition[C], San Juan, PR: IEEE, 1997:858-863.

[5] Velsler O. Stereo correspondence with compact windows via minimum ratio cycle[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(12): 1654-1660.

[6] Velsler O. Fast variable window for stereo correspondence using integral image[A]. IEEE International Conferenceon Computer Vision and Pattern Recognition[C], Madison, USA, 2003:556-561.

[7] Yoon K J, Kweon I S. Locally adaptive support-weight approach for visual correspondence search[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4):924-931.

[8] Tombari F, Mattoccia S, Stefano L D. Segmentation based adaptive support for accurate stereo correspondene[A]. IEEE Pacific Rim Symposium on Video and Technology[c], Santiago, CHILE:IEEE, 2007: 427-438.

[9] R. Brockers. Cooperative stereo matching with color-based adaptive local support. In: X. Jiang and N. Petkov(eds.) CAIP 2009. LNCS 5702, pp. 1019–1027. Springer-Verlag Berlin Heidelberg 2009.

[10] A.Fusiello, V.Roberto. Efficient stereo with multiple windowing, IEEE Conference on Computer Vision and Pattern Recognition, 1997:858-63.

[11] X. Sun, X. Mei, S. Jiao, M. Zhou, and H. Wang. Stereo matching with reliable disparity propagation. 3DIMPVT 2011.

[12] Brockers, R.Hund, M., Mertsching, B.: Stereo vision using cost-relaxation with 3d support regions. In: Proc. Image and Vision Comp. New Zealand (IVCNZ), pp. 96–101 (2005)