Computer Fraud and Security
ISSN (online): 1873-7056

Stereo Matching Algorithm Combining Image Segmentation and Improved Belief Propagation

Guangming Lu, Ping Zhang*, Yuejiao Han, Bo Song

Faculty of Environmental Arts and Architectural Engineering, Heilongjiang University of Technology, Jixi, 158100, China

**Corresponding Author.*

Abstract:

In this paper, a stereo matching algorithm combining image segmentation and improved belief propagation is presented to address the problems of poor real-time performance, low matching accuracy of weak texture regions, error matches and disparity holes. The mean shift algorithm is applied to segment the reference image and define the disparity template, which improves the real-time performance of the algorithm and the matching accuracy of the disparity flat regions. The cost value is calculated by the Sum of absolute differences algorithm to obtain the initial disparity map, then the occlusion points are detected by the left-right consistency detection and the invalid values are filled, in the disparity map acquisition. The weighted least squares method is used to calculate the parameters of disparity template, and the template is allocated by the improved belief propagation algorithm to obtain the regional optimal disparity plane. Finally, the dense disparity map is acquired by this algorithm. The result shows that the presented algorithm can greatly reduce the error match rate in weak texture regions and occlusion regions which meets the requirements of stereo matching.

Keywords: belief propagation, image segmentation, left-right consistency occlusion detection, invalid value filling, dense disparity map

INTRODUCTION

Vision is an important way for human beings to observe the surrounding environment and cognize the surrounding world intuitively. The human visual system uses the eyes to project the three-dimensional information of the surrounding environment into a two-dimensional image on the retina and converts the image information into bioelectric signals, which are transmitted to the brain by the optic nerve, and the brain analyzes and processes the image and restores the three-dimensional scene model of the objective world in the mind. Binocular stereo vision simulates the human binocular vision system by using binocular stereo camera to obtain binocular stereo image pairs, and after photo distortion correction and kernel correction of the captured original image, the matching algorithm suitable for the scene captured by the stereo image pairs is selected to establish the relationship between the matching primitives, and then restore the depth of the process to the three-dimensional coordinates [1-3]. The complete binocular stereo vision system contains five steps: camera calibration, image acquisition, feature extraction, stereo matching and 3D reconstruction [4]. Stereo matching plays the most critical and important role in the whole binocular stereo vision [5]. On the one hand, the disparity map obtained by the stereo matching algorithm directly reflects the accuracy of the depth information of the scene, which in turn affects the recovery of three-dimensional coordinate information; on the other hand, because stereo vision is a simulation of the imaging process of the human visual system, in the process of filming, the difference in the angle will certainly affect the authenticity and effectiveness of the information obtained, which determines the accuracy of the final reconstruction results. Therefore, the accuracy of stereo matching determines the quality of disparity map, depth map and the whole binocular stereo vision. Stereo matching consists of four parts: cost computation, cost aggregation, disparity computation, and disparity optimization [6], which can be divided into local algorithms and global algorithms [7-10]. The local stereo matching algorithm, i.e., selecting a pixel in a certain region of the image as the constraint region for calculating the matching cost of the pixel point, and then usually adopting the disparity selection strategy of WTA (Winner Take All, WTA) to determine the optimal disparity value of the point, which can quickly obtain the disparity map, but its matching accuracy is not high; global stereo matching algorithm, i.e., by establishing the energy function in the The global stereo matching algorithm, i.e.,by establishing data terms and smoothing terms in the energy function, transforms the matching algorithm into the problem of solving the optimal value of the global energy function, thus obtaining the dense disparity map, and the global constraint information of the image in the global algorithm has been fully applied with a stronger ability to resist local image blurring [11-15].

In recent years, DP (Dynamic Programming, DP) and BP (Belief Propagation, BP) algorithms have become hot spots in global matching algorithms by many experts and scholars because of their high matching accuracy. The DP algorithm has high computational efficiency and disparity map quality, but the matching process ignores the constraints of the disparity between each pole line, which leads to the presence of stripe defects in the disparity map. Liu and Xu [16] proposed a new dynamic programming stereo matching method, designed a new type of rectangular window to calculate the matching cost, constructed the optimization function based on the occlusion constraints, and suppressed the stripe defects by obtaining the control points on the low-resolution image, and the matching accuracy for the occlusion region and the weak texture region needs to be improved. Cheng and DA [17] added the pyramid layering technique to the traditional dynamic planning algorithm, and the low-pixel level provides the control point set for the high-pixel level to obtain the disparity map with higher accuracy. The phenomenon of horizontal streaking is effectively suppressed, but the matching accuracy in the weak texture region as well as the edge part is still unsatisfactory. With the successful application of MRF [18] (Markov Random Field, MRF) model in global stereo matching, the problem of stereo matching is converted to MRF model. Among them, the brief propagation algorithm replaces the probabilistic solution process of any node in the MRF model network with a way in which other nodes pass messages to it, the accuracy of matching is reliably improved, and the quality of the disparity map is significantly improved. The traditional brief propagation algorithm, on the other hand, has the following three shortcomings: (1) poor real-time, the message is transmitted between nodes, which makes the time complexity higher; (2) easy to produce false matching, the data term in the energy function is obtained by the AD (Absolute Difference, AD) transformation, and in the region of large gray-scale changes directly calculate the gray-scale value, the constraint term in the energy function can not be accurately and fully expressed, resulting in false matching. The constraints in the energy function can not be accurately and fully expressed, resulting in the phenomenon of mis-matching; (3) disparity null phenomenon, the smoothing term in the energy function adopts the Potts model, which does not sufficiently consider the smoothness at the band disparity jumps, so that the image in the weak texture region has a poor matching effect, which can easily lead to the emergence of the phenomenon of disparity nulls. For this reason, Liu et al. [19] proposed a stereo matching algorithm based on brief propagation, which uses watershed segmentation of the image to over-segment the image, which has good real-time performance and improved accuracy; Wang et al.[20]proposed a global stereo matching algorithm based on brief propagation, which uses the distance transform algorithm to optimize the message computation, and the bilateral graph technique to iteratively update the message using chunked Multi-scale brief propagation for updating, the speed of the algorithm is significantly improved and the occlusion region is effectively improved. However, none of them can take into account the time complexity and matching accuracy. In this paper, based on the existing global stereo matching algorithm, we propose to combine image segmentation and improved brief propagation stereo matching algorithm, which can effectively improve the shortcomings of the traditional brief propagation algorithm. The solutions are as follows: (1) message passing between segmented regions instead of message passing between pixel points, and the number of segmented regions is much smaller than the number of pixel points, which can effectively improve the efficiency of the algorithm; (2) calculation of surrogate values by the SAD (Sum of Absolute Difference , SAD) algorithm, padding of the invalid values, increasing the reliable viewpoints and thus improving the constraint terms of the energy function to reduce the phenomenon of mismatching; (3) comprehensively considering the distance and the boundary length of the smoothing term with the improved region-based brief propagation algorithm to fully utilize the pixel information of the weak texture region, which can effectively reduce the disparity null phenomenon.

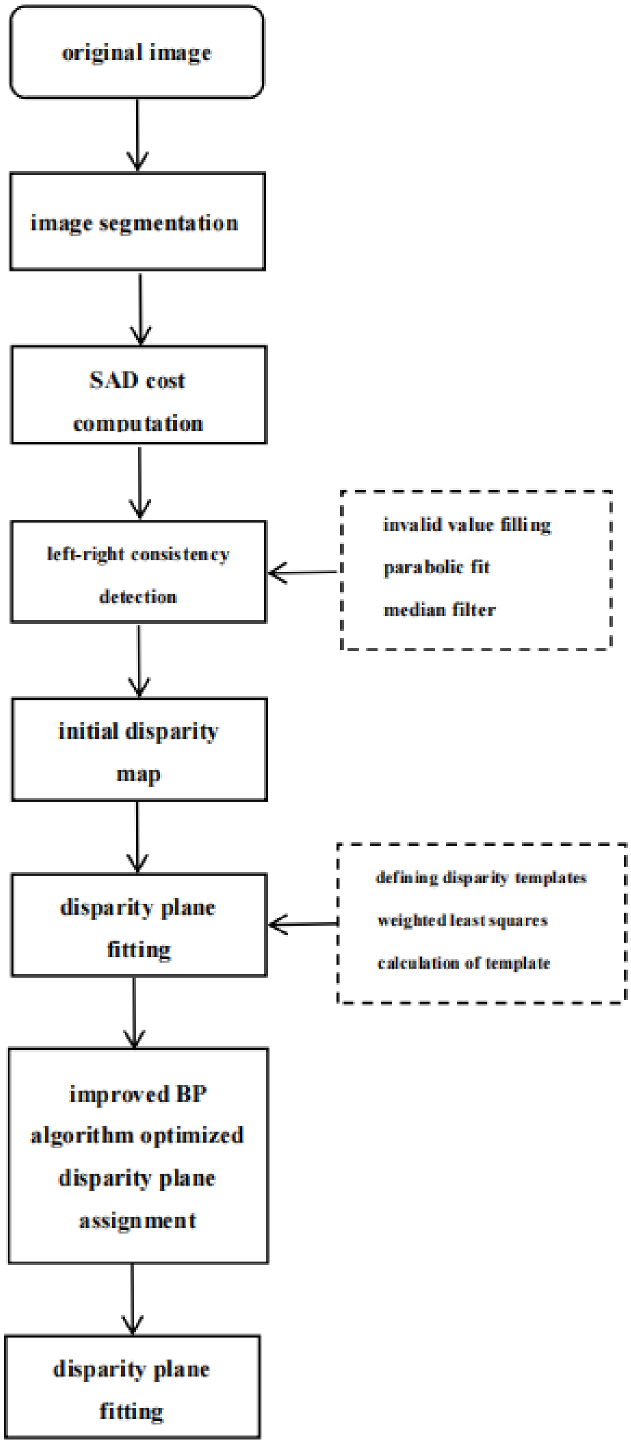


Figure 1. Flowchart of the paper

COMBINED IMAGE SEGMENTATION AND IMPROVED BRIEF PROPAGATION STEREO MATCHING ALGORITHM

In this paper, the original image is preprocessed using image segmentation theory, image segmentation is based on the relationship between the color information of the pixel and the distribution characteristics around the pixel, along the average gradient direction to find out the similarity of the color convergence points of each pixel, according to the similarity of the color of these convergence points to divide the image into different regions; and then the SAD algorithm calculates the surrogate value of each region to obtain the initial disparity map and then by the crosschecking detecting the occlusion point to get the invalid disparity value, according to the principle of polar geometry to distinguish the invalid value into occlusion point and mismatch point, and fill them reasonably and effectively; Next, define the disparity template, describe the segmentation region and use the weighted least squares method to iteratively calculate the parameters of disparity template to complete the disparity plane fitting, to obtain the disparity plane template set; Finally, use the improved region-based brief propagation algorithm is used to assign the templates to obtain the minimum disparity value to obtain a dense disparity map. By replacing pixel points by segmented regions, disparity plane template set replaces disparity search space range, and message transmission is converted to transmission between segmented regions, which reduces the computation amount of the algorithm to a certain extent and improves real-time algorithm as well as improves the matching effect of the disparity flat region and the edge portion; adopting the SAD algorithm, the left-right consistency detection and the filling of invalid value reduces the mis-matching of the gray scale region and increases the reliable viewpoints to improve the overall disparity map. The SAD algorithm, left-right consistency detection and invalid value filling can reduce the mis-matching in the gray scale region and increase the reliable viewpoints to improve the constraints of the global energy function in the weak texture region; the smoothing term in the energy function of the improved region-based brief propagation algorithm takes into account the distance of the neighboring regions and the length of the boundary to effectively inhibit the phenomenon of disparity null. The overall process is shown in Figure 1.

Acquisition of Initial Disparity Map and Disparity Value Filling

Image segmentation fuses all pixel points with similar colours in a region into the same segmentation block, which is a clustering technique that generally uses information such as colour and texture as the basis for categorisation. The assumption of image segmentation is the embodiment of the a priori knowledge of disparity map: disparity within the same segmentation block is characterised by smooth changes, and the boundary of the disparity discontinuity usually accompanies the edge of the segmentation block. In stereo matching, image segmentation can ameliorate the problems of matching accuracy that occur in weakly textured regions and in regions with depth discontinuities. The mean-shift algorithm is a nonparametric method based on density gradient ascent, which finds the target location through iterative operations, and its advantages are smaller computational volume, fast speed, and the ability to effectively categorize each pixel in the image. The mean -shift algorithm is based on the relationship between the color information of the pixels in the image and their surrounding distribution characteristics, the precise boundary positioning of the image, to ensure the scientific validity of the region delineation, and can maintain the detail information of the disparity map to a greater extent.Mean-shift is a statistical iterative algorithm that is intuitive in nature, where the mean value of the offset of the point to be processed is computed based on the other points in the neighbourhood, the point is moved to the mean value of the offset as a new starting point, and then the movement is continued until it converges to a peak value near the point. The core of the mean drift algorithm is to cluster the points in the feature space to get the mode points, which is actually the kernel probability density estimation algorithm. As shown in Figure 2.

(a) Original image

(b) Segmentation's image

Figure 2. The segment result of tsukuba

The idea of mean-shift image segmentation is introduced into the brief propagation algorithm, the number of pixel points in the image is much larger than the number of segmented regions, so it reduces the message transmission time and complexity to a greater extent. For the traditional brief propagation algorithm in which the data items are obtained by AD transformation, it is easy to cause false matching. Therefore, after the image is segmented, the matching cost is calculated by the SAD algorithm with truncation threshold, which enhances its robustness and at the same time reduces the mis-matching in regions with large gray scale changes. The expression of the SAD algorithm is as follows:

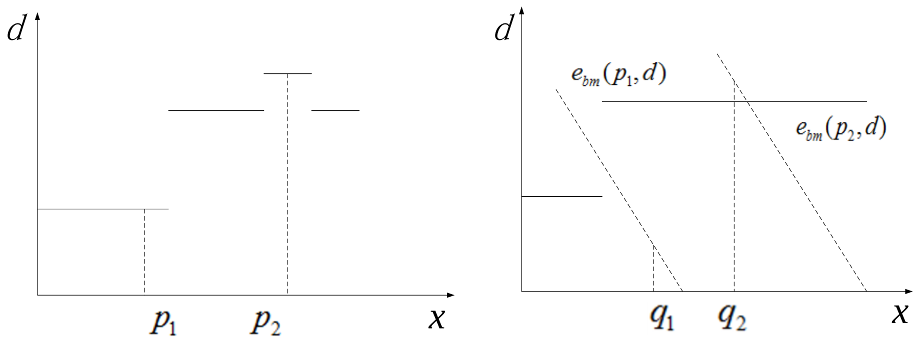
$$C_{SAD}(x,y,d) = \min\left(\sum_{(i,j)\in N(x,y)}\left|I_L(i,j) - I_R((i-d),j)\right|,\tau\right) \tag{1}$$

Let the coordinates of the point $p$be(x,y), d is the matching cost of the point, is the disparity search range($d \in [d_{\min}, d_{\max}]$), $N(x,y)$represents the neighborhood pixel of the pixel point p , $(i,j)$is the neighborhood point of the pixel point p, $I_L(i,j)$is the grayscale value of the pixel point p in the left image, $I_R((i-d),j)$is the grayscale value of the point $((i-d),j)$in the right image, and τ is the truncation threshold of the matching cost in the algorithm.(Window size of the SAD algorithm is 3*3;τ=20)

In the actual matching process, the emergence of the occlusion problem makes the accuracy of stereo matching degraded, so the solution of the occlusion problem becomes a hotspot of attention and research. The solution of the occlusion problem firstly needs to accurately detect the occlusion region, then the invalid disparity value is effectively reassessed, and the resulting disparity map is complete and accurate. The left-right consistency check is mainly used to find out the occluded regions that are present in one image but not in the other.The left and right consistency detection obtains the disparity maps of the left and right images respectively, for the point of p the left view to find its disparity is D_{bp}, the point p in the right image will be recorded as the point q, the disparity value of the point is $p - D_{bp}$, if $|D_{bp} - D_{mp}| > 1$, then it is invalid, and the point D_{bp} is the occlusion point. The disparity expression obtained is as follows.

$$D_p = \begin{cases} D_{bp} & |D_{bp} - D_{mp}| \le 1 \\ D_{inv} & otherwise \end{cases} \tag{2}$$

After the left-right consistency detection, the disparity value of the invalid point can be obtained, and a reasonable assignment of the invalid value can improve the completeness of the disparity map. The invalid points can be categorized into two types: occlusion points and mismatch points, as shown in Figure 3. From the geometric principle of pairwise polarity, it can be seen that the epipolar line of the occlusion point P_1 passes through the intermittent position caused by the occlusion, and does not intersect with the curve of disparity function. On the contrary, the epipolar line of the mis-matched point P_2 intersects the disparity function curve. With the above method, any invalid point in the disparity map can be scientifically identified [16].



(a) Occlusion Points (b) Mismatch Points

Figure 3. Diagram of identifying occlusion and mismatch points

Disparity Plane Fitting

In order to increase the reliable viewpoints so that the constraints of the global energy function are more accurate and stringent, the concept of parabolic fitting is first introduced to further optimize the initial disparity maps of each segmented region before the disparity plane fitting. Since disparity is continuous in the real world, parabolic fitting is used to achieve subpixel resolution. The parabolic fitting method is based on the principle that a parabola is formed by neighbouring integer disparity values, with the subpixel disparity being the lowest point of the parabola.The disparity value of the disparity map obtained through the above steps is of integer level, and after the subsequent stereoscopic image is recovered from the depth information of the pixels, in order to ensure the continuity of the formation of the depth map and to ensure that the disparity map conforms to the real physical situation, the parabolic fitting method is adopted to satisfy the above requirements while ensuring that the sub-pixel level disparity maps are obtained, and finally, the median filtering is utilized to optimize the disparity map. The parabolic fitting schematic is shown in Figure 4:

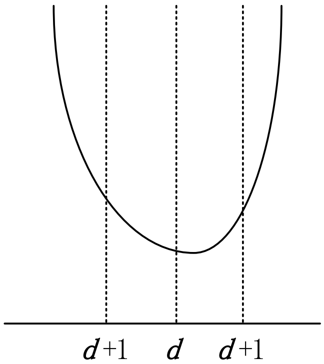


Figure 4. Schematic diagram of parabola fitting

In real spatial scenes, continuous regions with similar colors usually correspond to smooth surfaces, therefore, after obtaining the initial disparity of each segmented region of the segmented image, the disparity of each region can be approximated by a smooth disparity model, and generally this disparity model is taken as a plane called the disparity plane, which is expressed by the equation as:

$$d(x, y) = ax + by + c \tag{3}$$

Where a, b, c,is the template parameter and d(x, y) is the disparity value at the pixel point (x, y). After left-right consistency detection, disparity value filling and disparity optimization, the number of reliable viewpoints is increased, and then the initial template parameters are solved by the weighted least squares method, and several iterations are performed until the parameters converge. The planar parameters of all reliable regions are calculated sequentially to constitute the disparity planar template set. Since the mean--shift image segmentation algorithm segments the image based on the color and spatial information of the pixel points, the same scene plane in the image may be divided into several different regions. The significance of defining the disparity template function is to determine the disparity value that minimizes the matching cost of the segmentation region, and then group the plane templates with the same disparity value together in order to reduce the amount of subsequent computation, and thus reduce the complexity of the algorithm. The minimum cost disparity for the segmented region is found by summing the disparity cost of all valid pixels in that segmented region.

Traditional Brief Propagation Algorithm

The traditional brief propagation algorithm first models the image, and then assigns the pixel points in the image correspondingly to the nodes on the Markov random field, and finds the value of the quantity to be observed that satisfies the maximum joint probability distribution by iteratively reasoning about the probabilities among the nodes. That is, the probabilistic solution of a node in a MRF network is described as a kind of process in which other nodes pass messages to it. The energy function of the brief propagation algorithm consists of a data term and a smoothing term as shown in equation (4).

$$E(d) = \sum_p D(p, d_p) + \sum_{(i,j)} V(d_i, d_j) \tag{4}$$

Where the first term D(p,d_p) is the data term of the energy function and the second term V(d_i,d_j) is the smoothing term of the energy function, where i,j is the neighboring pixel point. The message passing between nodes is shown in Figure 5.

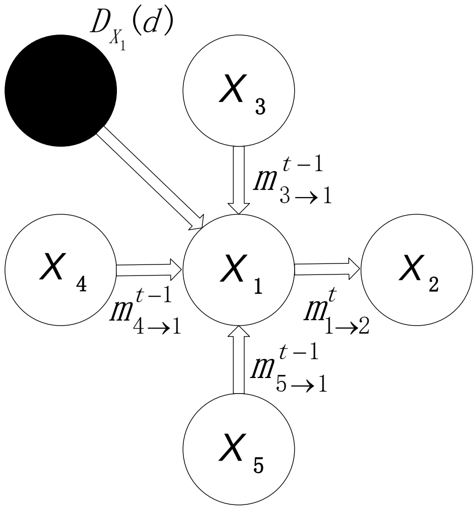


Figure 5. Node Transfer Model Between Neighbors

In Figure 5, X_1, X_2, X_3, X_4, X_5 are the neighbor nodes of the point X_1, and the message is delivered at each possible disparity d of the node ($d \in [d_{min}, d_{max}]$, $[d_{min}, d_{max}]$, is the set of possible disparityes), During the t-th iteration,and the message delivered by the node X_1 to its neighbor nodes X_2 during the first iteration when the disparity is as in Eq. (5):

$$m_{1\to2}^{t}(d) = \min\left(D_{X_1}(d) + V_d(X_1, X_2) + \sum_{X_N \in (X_1)\setminus X_2} m_{X_N \to X_1}^{t-1}(d)\right) \tag{5}$$

Eq. (5) consists of three terms, the first one $D(.)$ is a data term that represents the cost of matching assigned to a pixel point at possible disparity d , independent of neighboring points. Binocular stereo vision utilizes the calculation of the matching cost as a measurement parameter of the similarity of the matching pixel points, the traditional brief propagation algorithm utilizes the AD transformation as the initial matching cost function, when the AD transformation is carried out, the point to be matched on the left image of the image pairs is $p(x, y)$, and its grayscale is $I_L(x, y)$, and the directional disparity is d pixel point $q(x, y)$ on the right image is x, and its grayscale is $I_R\left((x-d), y\right)$then the matching cost of the two points p, q is as shown in Equation (6):

$$C(p, q) = I_L(x, y) - I_R\left((x-d), y\right) \tag{6}$$

$V(.)$is the smoothing term, which represents the neighboring pixel discontinuity and the amount of penalty at possible disparity values d, while the smoothing term of the global energy function of the brief propagation algorithm uses the Potts model with the following expression:

$$V\left(d_p, d_q\right) = \begin{cases} 0 & d_p = d_q \\ g_f(\Delta f) & else \end{cases} \tag{7}$$

function $g_f(.)$ is a segmented function with gradient as the variable, where the function f represents the gradient between neighboring pixel points p and points q in the image, and the specific expression of the function $g_f(.)$is shown in equation (8):

$$g_f(\Delta f) = \begin{cases} p * s & \Delta f < T \\ s & else \end{cases} \tag{8}$$

The above equation p, s, T All are threshold parameters, which p are taken as constants throughout the stereo matching process, where s is used to constrain the change at small gradient, is used to smooth constrain the change at large gradient, the last term in equation (4) indicates the sum of messages delivered from pixels in the neighborhood of the point X_1 (except the point X_2) in the process of (t-1)-th iterations, $N(X_1)$ denotes the neighborhood of the point X_1. After t-th iterations, the brief information of the pixel point at the time is shown in equation (9):

$$d_{X_1}^{*} = \operatorname{argmin} b_{X_1}(d)\,;\, d \in [d_{\min}, d_{\max}] \tag{9}$$

After the above iterative optimization, the information transfer is updated to finally generate a global disparity map.

Global Assignment of Disparity Plane Based on Improved Brief Propagation Algorithm

The data term in the traditional brief propagation algorithm is obtained by the AD transform, which is more real-time, but in the region where the grayscale of the left and right images change a lot, the pixel grayscale difference is directly calculated, which is easy to cause the phenomenon of mis-matching; the smoothing term is optimized by the Potts model, which doesn't fully take into account the smoothing at the disparity jumps, so the image is poorly matched in the region of weak texture, and it is easy to cause the disparity hollow phenomenon occurs. In this paper, the algorithm is to take the segmented region as the matching primitive and redefine the global energy function. In terms of data term, the SAD algorithm calculates the matching surrogate value of each region, which enhances its robustness and reduces the mis-matching in regions with large gray scale changes; in terms of smoothing, the image segmentation algorithm segments the image based on the color and spatial information of the pixels, and the disparity jump corresponds to the edge of the segmentation region, so that the problem of calculating the optimal disparity plane of each region replaces the problem of calculating the optimal disparity plane of each pixel directly in the traditional brief propagation algorithm. The problem of calculating the optimal disparity plane of each region is replaced by the problem of calculating the optimal disparity plane of each pixel in the traditional brief propagation algorithm, and the segmented regions are penalized by a smoothing term that takes into account the distances of the neighboring regions as well as the length of the boundaries, so that more stringent smoothing constraints are carried out to reduce the emergence of the phenomenon of disparity null. The global energy function is redefined as in equation (10):

$$E(f) = E_{data}(f) + E_{smooth}(f) \tag{10}$$

where $E_{data}(f)$, $E_{smooth}(f)$denotes the data term and the smoothing term, the data term is used to denote the overall matching cost and the smoothing term is used to denote the disparity discontinuity penalty (constraint) between neighboring templates. The data term and smoothing term are expressed as follows:

$$E_{data}(f) = \sum C(S_i U_i) \tag{11}$$

$$E_{smooth}(f) = \begin{cases} \sum_{S_i S_j} L_{(i,j)} \mu\left(U_i \neq U_j\right) & L_{(i,j)} \neq 0 \\ \sum_{S_i S_j} B\omega & L_{(i,j)} = 0 \end{cases} \tag{12}$$

Where $C(S_i U_i)$ denotes the matching cost after the template U assigned for the i-th region, and the i-th region and the j -th region are neighboring regions. In order to be able to scientifically and accurately obtain the optimal disparity plane of each segmented region, the distance of the neighboring regions, and the length of the region's own boundary are accounted for in the smoothing term of the energy function. $L_{(i,j)}$ on behalf of the distance of neighboring regions, the value of μ which is determined by the distance of neighboring regions (positive relationship), when, $U_i = U_j$, μ=1.$B = B_{(i,j)}/B_j$, B_j for the length of the S_j region's own boundary, $B_{(i,j)}$ for the region S_j and the neighboring regions S_i constitute the length of the common boundary, the value of ω which depends on the proportion of the length of the boundary (inversely proportional to the relationship). Since the disparity jump is the edge of the segmented region, the improved smoothing term can make full use of the information of the segmented region, focusing on the smooth constraints at the disparity jump, and the disparity null phenomenon is effectively suppressed, so that the quality of the disparity map is further improved. After establishing the new global energy function, the improved region-based brief propagation algorithm is used to find the disparity plane assignment corresponding to the minimum energy function. The information passed from the region$S_i$to the neighboring regionsS_j at the t-th iteration is denoted by $m^t_{S_i \to S_j}\left(f(S_j)\right)$ the information update rule between different regions as shown below:

$$m^t_{S_i \to S_j}\left(f(S_j)\right) = \min\left[C\left(S_i, f(S_i)\right) + V\left(S_i, S_j\right) + \sum_{S_k \in N(S_i) \backslash S_j} m^{t-1}_{S \to S_i}(f(S_i))\right] \tag{13}$$

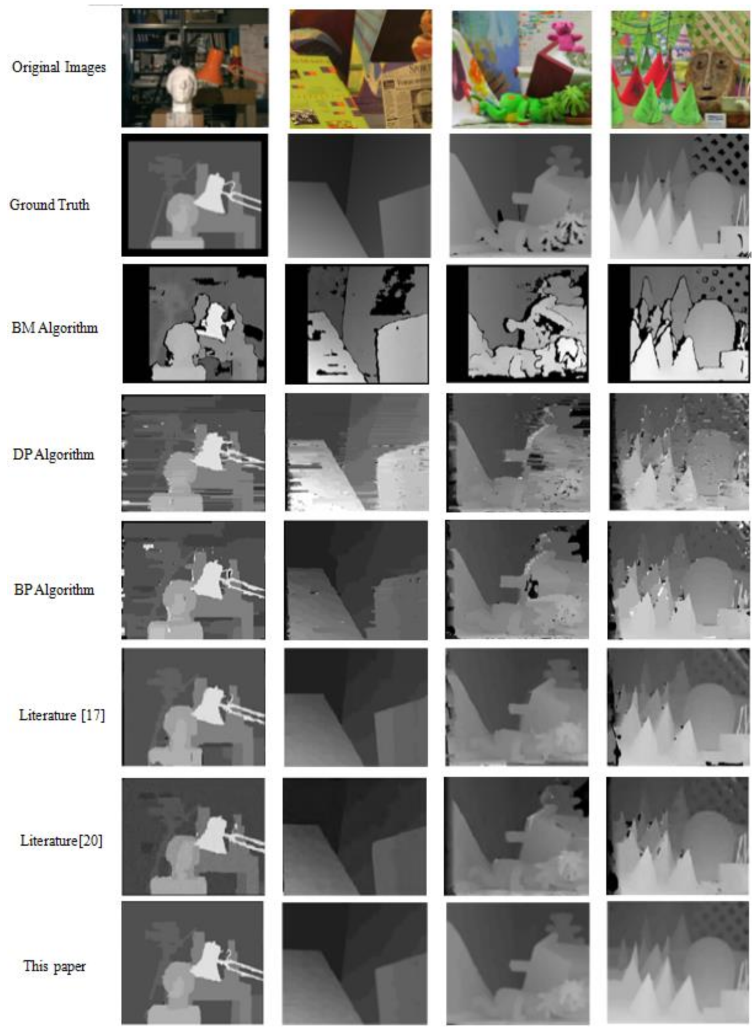
where $C\left(S_i, f(S_i)\right)$ is a data term indicating the matching cost of region S_i when the disparity plane is $f(S)$. $V\left(S_i, S_j\right)$ is an inter-region smoothing term indicating the amount of disparity discontinuity penalty between region S_i and region S_j. $N(S_i)\backslash S_j$ denotes the set of all neighboring regions of region S_i except S_j. After T iterations of information transmission, the brief level of region S_i is obtained:

$$b_{S_i}\big(f(S_i)\big) = C\big(S_i, f(S_i)\big) + \sum_{S\in N(S_i)} m_{S\rightarrow S_i}^{T}\big(f(S_i)\big) \tag{14}$$

According to the principle of brief propagation algorithm, the plane corresponding to the smallest brief level is the optimal disparity plane for region S_i:

$$f^*(S_i) = \mathrm{argmin} b_{S_i}\big(f(S_i)\big) \tag{15}$$

Based on the obtained optimal disparity plane parameters of each region, the final dense disparity map can be obtained by calculating the optimal disparity value d of the corresponding pixel point using equation (9).



(a) Tsukuba (b) Venus (c) Teddy (d) Conesde

Figure 6. Comparison of results of different algorithms

EXPERIMENTAL ANALYSIS

In order to verify the feasibility of this paper's algorithm, four image pairs (Tsukuba, Venus, Teddy, Conesde) were selected from the standard stereo images provided by the Middlebury stereo matching algorithm test platform to test this paper's algorithm, and the original images (left) and their corresponding standard disparity maps are shown in rows 1 and 2 of Figure 6. The algorithm of this paper was implemented in the integrated environment of Matlab 2016a. The block matching algorithm (BM) in local stereo matching, dynamic programming matching algorithm (DP) in global stereo matching, brief propagation algorithm (BP) in global stereo matching, a global stereo matching algorithm based on brief propagation, and a study of hierarchical stereo matching algorithms based on dynamic programming are compared with the results of this paper's algorithms, and the experimental results correspond to the columns of columns 3 to rows 8 of Figure 6, respectively .

In order to further evaluate the matching effect of stereo matching algorithms in different regions, the Middlebury testbed compares different matching algorithms by three indexes, namely, the false matching rate of nonoccultation region (nonocc), the zone false matching rate of the whole image (all), and the false matching rate of the disparity discontinuity region (disc), respectively. The specific data of this paper's algorithm is shown in Table 1, and the comparison between this paper's algorithm and other algorithms is shown in Table 2. Meanwhile, the computing time of this paper's algorithm and traditional BP algorithm is shown in Table 3.

Table 1. The results of this algorithm in middlebury

nonocc	Sukua all	disc	nonocc	Venus all	disc	nonocc	teddy all	disc	nonocc	cones all	disc
1.87	2.06	4.97	0.98	1.45	4.16	4.92	7.83	9.75	4.25	11.73	10.88

Wherein, the occluded region indicates an occluded portion of the image, and the disparity discontinuity region indicates that the disparity values between neighboring pixels in the image cannot satisfy a certain threshold.

Table 2. The comparison of mismatch rates of algorithms

Comparison of algorithms	Nonocc (%)			
	Tsukuba	Vennus	Teddy	Cones
DP Algorithm BP Algorithm	4.12 2.24	10.10 11.60	14.00 21.00	10.50 10.60
Literature [17]	2.13	3.12	5.47	7.70
This Paper	1.87	0.98	4.92	4.25

Table 3. The comparison of time of algorithms

Comparison of algorithms	Time(s)			
	Tsukuba	Vennus	Teddy	Cones
BP Algorithm	8.24	15.60	83.93	85.60
This Paper	7.07	11.98	43.42	54.25

The values in columns 2-4 of Table 2 indicate the percentage of mis-matched points in the non-occluded region (nonocc) of the 4 pairs of test images using the various algorithms, respectively. Combining the results shown in Figure 6 and the data in Table 2, the disparity optimization part of the block matching algorithm can not be carried out effectively in the whole image, and can not obtain the complete edge part, compared with the algorithm in this paper: the improved global energy function can obtain a higher matching accuracy while taking into account the image segmentation method to obtain the complete disparity map of the edges; the matching algorithm of the DP has the problem of optimizing the disparity only on a single scan line, which leads to the problem of optimizing the disparity on a single scan line, resulting in a higher matching accuracy. The DP matching algorithm has the problem of disparity optimization only on a single scan line, which leads to the constraints within the scan line only can be effectively utilized, while the constraints between the scan lines are not sufficiently used, and thus the stripe trailing phenomenon occurs. Compared with the algorithm in this paper: combining the BP algorithm for image segmentation, the information between nodes is iteratively updated instead of the constraints within the scan lines to avoid the phenomenon of horizontal streaking; the literature [17] adds a hierarchical model to the traditional DP algorithm, dividing the pixels into high and low pixel hierarchies, with the low pixel hierarchies providing the set of control points for the high pixel hierarchies, and a similarity measure function combining a kind of the color information of the pixel with the gradient information The matching cost is calculated and matched cost filtering is applied. The matching accuracy of edge region and weak texture region can be effectively improved, and the phenomenon of horizontal stripe is obviously improved and the real-time performance is better. At the same time, the two-way DP algorithm is used in the process of stereo matching, which can reduce the false matching of some pixels. However, the hierarchical model divides pixels into high and low pixel levels and reduces the number of reliable pixel points, which reduces the constraints of the DP energy function and the two-way DP algorithm exacerbates the mis-matching rate of the occluded region and the edge part, resulting in the overall matching accuracy is still unsatisfactory; in the literature [20], the use of the distance transformation method to redefine the iterative message updating method in the BP algorithm can significantly reduce the computational complexity, but it can also significantly reduce the computational complexity and real-time matching. significantly reduce the computational complexity, but it ignores the truncation value in the smoothing term in the global energy function, which will produce disparity null phenomenon. At the same time, the combination of bilateral graph technique and multi-scale brief propagation can effectively improve the running efficiency of the algorithm, however, the group iteration between nodes and the layering of images can cause the matching effect of disparity map edge part to be weakened. Finally, the use of left-right consistency detection to find invalid disparity values but only assigning values to the occluded points will reduce the true objectivity of the disparity map; combining the algorithmic results in Figure 6 and the running time in Table 3, it can be obtained that the combination of image segmentation and improved BP stereo matching algorithm proposed in this paper compared with the traditional BP algorithms firstly utilizes the theory of mean -shift image segmentation to preprocess the image, and the segmentation region between the Message transfer between segmentation regions instead of transmission between nodes, significantly improving real-time; SAD algorithm(Window size of the SAD algorithm is 3*3; $\tau = 20$) to get the initial disparity map by the left-right consistency detection of the occlusion point detection and based on this different types of invalid values(invalid values: $\left|D_{bp} - D_{mp}\right| > 1$,occlusion point: $\left|D_{bp} - D_{mp}\right| \leq 1$) for different ways to fill the SAD transform to get the energy function of the data items can effectively reduce the phenomenon of mis-matching phenomenon, occlusion phenomenon, to increase the reliability of the point of view so as to strengthen the global energy function ($Ui = Uj$, $\mu = 1$) in the weak texture area The use of the smoothing term to replace the Potts model, which takes into account the distance between neighboring regions and the proportion of the length of the common boundary, can effectively inhibit the emergence of disparity voids. That is, compared with the traditional BP, literature [17], and literature [20], this paper has better processing effect for weak texture regions and edge regions, and the disparity map is more complete. The smoothing effect of the algorithm results can show that the occluded region is filled with more reasonable and effective disparity, disparity voids are effectively suppressed, which is more in line with the characteristics of the actual scene.

CONCLUSION

This paper proposes a stereo matching algorithm that combines image segmentation and improved BP. On the basis of binocular stereo vision theory, the region segmentation information of the image is obtained by scientific and effective image segmentation method, the initial disparity value of each segmented region is calculated by SAD algorithm, and the occlusion point is detected using left-right consistency detection, and the invalid value is filled on this basis. Combined with the weighted least squares iteration to obtain reliable disparity plane parameters, the improved region-based brief propagation algorithm is used to assign templates to the disparity planes to obtain a dense disparity map. The algorithm is tested with the images in Middlebury database, and the experimental results show that the algorithm has considerable effectiveness, universality, and meets the needs of stereo matching, and the following conclusions are obtained: (1) message passing between segmented regions instead of transmission between nodes can effectively reduce the computational complexity to improve real-time and improve the matching effect of the edge region; (2) effective filling of the mismatched points and occluded points can be achieved by using the improved region-based brief propagation algorithm to obtain dense disparity map; (3) The optimal disparity planes in each region are obtained from the template allocation of the improved region-based BP algorithm, and the appearance of disparity nulls can be effectively suppressed. It can lay a solid foundation for subsequent military target detection, medical diagnosis and visual navigation.

Meanwhile, given the limited ability and energy of the authors, the algorithm in this paper still has deficiencies, and the following three aspects will be explored in depth in the subsequent research. (1) the real-time, sensitivity to noise aspect of the algorithm still needs to be improved; (2) the algorithm in this paper is only experimented for the dataset provided by the Middlebury platform, and no relevant experiments are conducted on remote sensing images; and (3) the recovery of 3D scene is completed for disparity map.

ACKNOWLEDGMENT

This work was supported by Heilongjiang Provincial Natural Science Foundation of China (LH2023D023)

REFERENCES

[1] Xiao Y Q, Liu D H, Sun P. Research progress of image stereo matching. Measurement and Control Technology, 2009(8):1-5+10.
[2] Tankovich V, Hne C., Fanello S., Zhang Y., Izadi S, & Bouaziz S. Hitnet: hierarchical iterative tile refinement network for real-time stereo matching, 2020.
[3] Urtasun R, Hu R, Wang S, Duggal S, & Ma W C. DeepPruner: Learning Efficient Stereo Matching via Differentiable PatchMatch, 2019.
[4] Lim Jaeseung, Lee Sankeum. Patchmatch-Based Robust Stereo Matching Under Radiometric Changes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(5):1203-1212.
[5] Henderson P, & Ferrari V. Learning single-image 3d reconstruction by generative modelling of shape, pose and shading. International Journal of Computer Vision, 2019(8).
[6] Scharstein D, Szeliski R.A taxonomy and evaluation of dense two-frame stereo correspondence Algorithms. International Journal of Computer Vision, 2002, 47(1-3):7-42.
[7] Zhang K, Lu J, Lafruit G. Cross-Based localstereo matching using orthogonal integral images. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(7):1073-1079.
[8] Hosni A, Bleyer M, Gelautz M, et al. Local stereo matching using geodesic support weights//IEEE International Conference on Image Processing, 2010:2093-2096.
[9] Hirschmuller H. Stereo processing by semi global matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 30(2):328-341.
[10] Emst I, Heiko Hirschmuller. Mutual information based semi-global stereo matching on the GPU// ISVC08.Springer, Berlin, Heidelberg, 2008:228-239.
[11] Lu G M, Wang J X. Semi-global stereo matching algorithm combined with image segmentation. Remote Sensing Information, 2020, 35(06):85-91.
[12] Henderson P, Ferrari V. Learning single-image 3d reconstruction by generative modelling of shape, pose and shading. International Journal of Computer Vision, 2019(8).
[13] Kim J S, Park C H, Lee D W. Block-based stereo matching using image segmentation. The Journal of Korean Institute of Communications and Information Sciences, 2019, 44(7), 1402-1410.
[14] Lee Y, Yoon K. Deep learning based disparity map estimation using stereo vision for uav. Transactions of the Korean Institute of Electrical Engineers, 2020, 69(5), 723-728.
[15] Sébastien D, Bouvry A, Leemans V, Dumont B, & Mercatoris B. Imaging wheat canopy through stereo vision: overcoming the challenges of the laboratory to field transition for morphological features extraction. Frontiers in Plant Science, 2020, 11, 1.
[16] Xu Y, Liu X J. Research on dynamic planning stereo matching algorithm in computerized binocular vision. Journal of Southwest Normal University (Natural Science Edition), 2020, 45(09), 118-123.
[17] Cheng S P, Da F P. Research on hierarchical stereo matching algorithm based on dynamic planning. Journal of Instrumentation, 2016, 37(07), 1665-1672.
[18] Pal N R, Pal S K. A review on image segmentation techniques. Pattern Recognit, 1993, 26(9):1277-1294.
[19] Liu Z Y, Zhou B, Zhang X H, Liu C Y. A stereo matching algorithm based on brief propagation. Automation and Instrumentation Table, 2010(01), 111-113.
[20] Wang Z, Zhao X, Yan Q H, Hu H Y. A global stereo matching algorithm based on brief propagation. Journal of Beijing Architecture University, 2015, 31(04):58-64.