

EXHIBIT 1036



All



ADVANCED SEARCH

Conferences > 2019 IEEE 8th Joint Internati... ?

Image Edge Based Efficient Stereo Matching

Publisher: IEEE

Cite This



PDF

<< Results | < Previous | Next >

Zhiguang Zhang ; Jianhua Zou All Authors ...

1
Cites in
Paper

158
Full
Text Views



Alerts

Manage Content Alerts
Add to Citation Alerts

Abstract

Document Sections

- I. Introduction
- II. Related Work
- III. Image Edge Based Stereo Matching
- IV. Result
- V. Conclusions

Authors

Figures

References

Citations



Downl
PDF

Abstract: In this paper, we present the image-edge based stereo matching method-an extension of the ELAS algorithm- which increases the robustness of stereo matching. The image reg... [View more](#)

► Metadata

Abstract:

In this paper, we present the image-edge based stereo matching method-an extension of the ELAS algorithm- which increases the robustness of stereo matching. The image region is divided into triangular regions by triangulation on support points. Based on image Canny edge information, our approach divides triangular regions into three groups according to whether piecewise disparity smooth is valid or not. The Bayesian method is used for piecewise disparity smooth areas, and the local stereo method is used for disparity discontinuity areas. For the disparity smooth areas, we further strengthen the influence of disparity prior by making use of color prior. Based on experimental results, our method is comparable in accuracy and efficiency.

Published in: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)

Date of Conference: 24-26 May 2019

DOI: 10.1109/ITAIC.2019.8785510



 Contents

SECTION I.

Introduction



Computer vision has been widely used in industrial applications, such as pedestrian detection [1], [2], fatigue driving detection [3], and ship type detection [4]. Accurate stereo matching is one of the most challenging research problems in the field of computer vision [5]. Pixel noise during image formation [6], texture-less regions, depth discontinuities, and occlusion are the most difficult aspects of this problem. Additionally, image resolution keeps increasing, which not only increases the number of pixels to be processed, but also increases the disparity searching space. This consumes more processing time. Real-time application is often required, such as in mobile robotics. Therefore, fast calculation algorithms are required.

It's clear that stereo matching is an ill-posed problem. For such problem, a Bayesian approach provides a way to consider it as an inference problem. For stereo matching, the target is to infer pixel disparity D based on image observation O . According to Bayes function, $P(D|O) \propto P(O|D)P(D)$, where $P(O|D)$ is the likelihood of encoding the process of forward image formation, and $P(D)$ is the prior which encodes the disparity assumptions.

Using Bayesian approach to infer disparity, the prior proposed is a key, and an efficient prior can improve the final disparity estimation. In stereo matching problems, there are robustly matched points between the left and right image, which are called 'support points' [7]. Assuming piecewise disparity smooth, support points are used to generate prior for ambiguous regions. However, piecewise disparity smooth is not always true for all image regions, such as the disparity discontinuity regions. This paper makes two major contributions: first, we assume the image is partially piecewise smooth. Making use of image edge information, we divide the image area into three groups, including piecewise disparity smooth area and disparity discontinuity area. Second, for disparity smooth area, we strengthen the prior influence by making use of color prior.

SECTION II.

Related Work

Stereo matching has been studied in computer vision for long time, and many algorithms have been made to solve the stereo problem. In the last decade especially, there was much progress made in both accuracy and computational efficiency. Generally, the existing stereo matching algorithms can be divided into two categories [8]: local stereo matching methods and global stereo matching methods. A stereo algorithm is called a global method when there is a global objective function to be optimized. Otherwise it is called a local method.

Global methods typically define a global objective energy function, which includes a data term and a smoothness term. The former sums pixel-wise matching costs, while the latter supports piecewise smooth disparity selection. The central problem for global methods is not only to define a good objective function, but also to provide an effective computing method to find local or global minimum [9], [10]. Some algorithms use more terms to penalize occlusions [11], [12], and discontinuities [13]. They also enforce a left/right or symmetric consistency between images [14], [15], [16]. The strategies to find the minimum of the global objective energy function differ. Dynamic programming (DP) approaches [17], [18] perform the optimization in one dimension for

Image Edge Based Efficient Stereo Matching

Zhiguang Zhang*, Jianhua Zou

System Engineering Institute, Xi'an Jiaotong University, Xi'an, China
zzgchina@stu.xjtu.edu.cn, jhzou@sei.xjtu.edu.cn

Abstract— In this paper, we present the image-edge based stereo matching method—an extension of the ELAS algorithm—which increases the robustness of stereo matching. The image region is divided into triangular regions by triangulation on support points. Based on image Canny edge information, our approach divides triangular regions into three groups according to whether piecewise disparity smooth is valid or not. The Bayesian method is used for piecewise disparity smooth areas, and the local stereo method is used for disparity discontinuity areas. For the disparity smooth areas, we further strengthen the influence of disparity prior by making use of color prior. Based on experimental results, our method is comparable in accuracy and efficiency.

Keywords—image edge; stereo matching; Bayesian method

I. INTRODUCTION

Computer vision has been widely used in industrial applications, such as pedestrian detection [1, 2], fatigue driving detection [3], and ship type detection [4]. Accurate stereo matching is one of the most challenging research problems in the field of computer vision [5]. Pixel noise during image formation [6], texture-less regions, depth discontinuities, and occlusion are the most difficult aspects of this problem. Additionally, image resolution keeps increasing, which not only increases the number of pixels to be processed, but also increases the disparity searching space. This consumes more processing time. Real-time application is often required, such as in mobile robotics. Therefore, fast calculation algorithms are required.

It's clear that stereo matching is an ill-posed problem. For such problem, a Bayesian approach provides a way to consider it as an inference problem. For stereo matching, the target is to infer pixel disparity D based on image observation O . According to Bayes function, $P(D|O) \propto P(O|D)P(D)$, where $P(O|D)$ is the likelihood of encoding the process of forward image formation, and $P(D)$ is the prior which encodes the disparity assumptions.

Using Bayesian approach to infer disparity, the prior proposed is a key, and an efficient prior can improve the final disparity estimation. In stereo matching problems, there are robustly matched points between the left and right image, which are called 'support points' [7]. Assuming piecewise disparity smooth, support points are used to generate prior for ambiguous regions. However, piecewise disparity smooth is not always true for all image regions, such as the disparity discontinuity regions. This paper makes two major contributions: first, we assume the image is partially piecewise smooth. Making use of image edge information, we divide the image area into three groups, including piecewise disparity smooth area and disparity

discontinuity area. Second, for disparity smooth area, we strengthen the prior influence by making use of color prior.

II. RELATED WORK

Stereo matching has been studied in computer vision for long time, and many algorithms have been made to solve the stereo problem. In the last decade especially, there was much progress made in both accuracy and computational efficiency. Generally, the existing stereo matching algorithms can be divided into two categories [8]: local stereo matching methods and global stereo matching methods. A stereo algorithm is called a global method when there is a global objective function to be optimized. Otherwise it is called a local method.

Global methods typically define a global objective energy function, which includes a data term and a smoothness term. The former sums pixel-wise matching costs, while the latter supports piecewise smooth disparity selection. The central problem for global methods is not only to define a good objective function, but also to provide an effective computing method to find local or global minimum [9, 10]. Some algorithms use more terms to penalize occlusions [11, 12], and discontinuities [13]. They also enforce a left/right or symmetric consistency between images [14, 15, 16]. The strategies to find the minimum of the global objective energy function differ. Dynamic programming (DP) approaches [17, 18] perform the optimization in one dimension for each scan line individually, which normally leads streaking effects. This can be avoided by tree-based DP approaches [15, 19]. A two dimensional optimization is reached by Graph Cuts [12] or Belief Propagation [14, 20]. Most of these algorithms take minutes or even hours to run on high-resolution image. Semi-global matching (SGM) [21] is a faster optimization technique which approximates 2D MRF interference by performing cost aggregation along various 1D paths in the image. SGM is faster than most global methods, but it is still slow compared to our method.

A local method consists of matching cost computation, cost aggregation, and disparity selection [8]. Local methods select the disparity with the lowest matching cost, that is, winner-takes-all [22]. The disparities are computed individually for each pixel, and the cost aggregation is calculated over a local window [8]. The choice of the size and the shape of the aggregation window is both crucial and challenging. A large window size leads to edge blurring at the depth discontinuities, while a small window size decreases the signal to noise ratio. A customized shape and size of the aggregation window [23] for each pixel significantly increase the computational time. In this case some adaptive window based local stereo matching algorithms [24] are even slower than some efficient global stereo matching algorithms, because local methods require the evaluating the full disparity space image (DSI).

Techniques that avoid evaluating the full search space have been proposed, both for global and for local methods. Methods for reducing the search space of global optimization include those outlined in research by Veksler [25], Wang et al. [26], and Yamaguchi et al. [27]; however, they still require an exhaustive DSI computation, which makes them quite slow. The seed-and-grow method [28] only explores part of the search space, therefore it produces only a partial disparity map, leaving large image regions unlabeled. Moallen et al. [29] proposed the restricted search method, which constraints the search range of an edge by directional derivative disparity. This method performs full search for the first edge-pixel, and restricted search for other edge-pixels. However, its disparity is only for edge, and its accuracy highly depends on the result of the first edge-pixels. Jaafari et al. [30] makes use of image edge information to reduce searching space in DP. The more recent highly optimized ELAS method [7] builds a triangulation over sparsely matched support points and uses this to explore the disparity search space. Finally, LPS [31] also avoids exploring the full disparity space by proposed plane generation. In result section, we compared the data of our method with the data of other methods which also avoid full range searching and demonstrate evidence that our method performs better.

III. IMAGE EDGE BASED STEREO MATCHING

In this section, we describe our approach to improve the prior. In ELAS [7], assuming piecewise smooth disparities, robustly matched ‘support points’ contain prior information to estimate the disparities for the remaining ambiguous pixels. The coordination of support points is used to generate 2D mesh through Delaunay triangulation. We find that support points can provide correct disparities estimation for piecewise smooth area, however, they normally provide improper disparities estimation for disparity discontinuity area. In our work, making use of image edge information, we research the disparities estimation based on support points. Triangular regions are divided into three groups according to whether the triangle area is piecewise smooth disparities or not. The triangular areas in the image are shown in Fig.1. We also find that image regions are normally planar when the color of these regions is similar. In such a case, we can strengthen the influence of disparity prior, because it is normally the real disparity. Our method proceeds as follows: we start with support points generation and edge information extraction, then we describe the triangle area division in detail. Finally, we show the mathematical derivation of the Bayesian approach for the stereo problem. We assume rectified input images for simplicity, so the vertical coordinate of correspondence points is same in both images.

A. Support Points Generation

The support points are the pixel regions, which can be robustly matched between image pairs due to their unique texture in the horizontal line. There are methods which can obtain stable correspondences between image pairs [32, 33, 34]. In this work, both images are processed with horizontal and vertical Sobel filter with windows size 9×9 . For each pixel, the descriptor is a vector of 16 elements, which is shown in Fig.2.

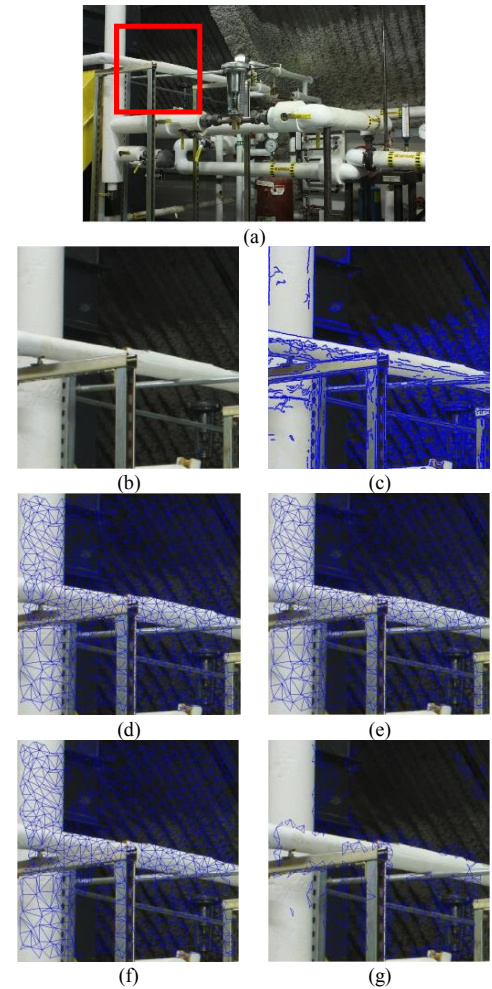


Fig. 1. Demonstration of triangular area generation and classification. (a) The left image of the Middlebury 2014 Pipes. (b) The region defined by the red rectangle in (a). (c) Canny edges of selected image region. (d) All the triangular areas of selected image region. (e) All the triangular regions which contains edge information. (f) All the triangular areas which can be considered as disparity smooth. (g) All the triangular areas which are considered as disparity discontinuity.

To balance the efficiency and accuracy in support points searching, the step-size is 5, and the disparity search range is half the image width. To remove the ambiguous matches, we compare the matching ratio between the best match and the second-best match. The matched point will be removed if its matching ratio is larger than a fixed value threshold τ . To ensure the support points are robustly matched, we perform a consistency check, which compares the disparity from left-to-right and right-to-left, to filter the outliers.

B. Triangle Division Based on Edge Information

The whole image is divided into triangular 2D mesh through Delaunay triangulation on support points. ELAS assumes piecewise smooth disparity, but this assumption is not always true. It works well on a plane area but constricts the disparity search range for a disparity discontinuities area.

		1		
2		3		4
	5	6,7	8	
9		10		11
		12		

	13	
14		15
	16	

Fig. 2. Locations of 16 element of the feature vector on X-Sobel(left) and Y-Sobel(right) gradient images. The candidate support point is at the center. The Sobel response are normalized and shifted to be 8 Bits in size and have values between 0 and 255.

The Canny edge detector approach is used to generate image edge. Based on image edge information we divide triangles into three groups: 1) triangular areas that contain no edge information; 2) triangular areas that contain edge information, however, the edge is not due to disparity discontinuity, but due to color discontinuity. Triangular areas and corresponding support points are in same plane; and 3) triangular areas that contain edge, and the edge is a disparity discontinuity area.

For the triangles which contain edge information, if the disparities difference between the triangular vertex are within threshold α , they are classified into Group 2; otherwise, they are put into Group 3.

For Groups 1 and 2, the assumption that the image region is piecewise smooth disparities is true, and they are processed similar with ELAS. For Group 3, they are usually disparity discontinuity areas, the disparity estimation from support points will constrict the disparity search range. A larger search range and local method are used for such triangular areas, which contain rich texture.

C. Model of Proposed Method

With the stereo image pair, we can obtain the support points and edge information. $O = \{o_1, \dots, o_N\}$ is defined as a set of image observations. Each observation, $o_n = (u_n, v_n, f_n)^T$, is formed as the concatenation of its image coordinates, $(u_n, v_n) \in \mathbb{N}^2$, and a feature vector, $f_n \in \mathbb{R}^Q$. We denote $o_n^{(l)}$ and $o_n^{(r)}$ as the observations in the left and right image respectively. $S = \{s_1, \dots, s_M\}$ is defined as a set of robustly matched support points. Each support point, $s_m = (u_m, v_m, d_m)^T$, is defined as the concatenation of its image coordinates, $(u_m, v_m) \in \mathbb{N}^2$, and its disparity, $d_m \in \mathbb{N}$. $E = \{e_1, \dots, e_P\}$ is defined as a set of image edge points. Each edge point, $e_p = (u_p, v_p, i_p)$, is defined as the concatenation of its image coordinates, $(u_p, v_p) \in \mathbb{N}^2$, and edge information, $i_p \in \{0, 1\}$. Without loss of generality, in the following we consider the left image as the reference image. $T = \{t_1, \dots, t_Q\}$ is defined as triangles, which are computed by Delaunay triangulation on support points S . Each triangle, $t_q = \{s_{q1}, s_{q2}, s_{q3}\}$, is defined as a set of three support points, which are vertices of the triangle.

We assume that the observations $o_n^{(l)}$, $o_n^{(r)}$ and support points S , edge information E are conditionally independent given their disparities d_n , the joint distribution can be factorized as:

$$p(d_n, o_n^{(l)}, o_n^{(r)}, S, E) \propto p(d_n | S, E, o_n^{(l)}) p(o_n^{(r)} | o_n^{(l)}, d_n) \quad (1)$$

where $p(d_n | S, E, o_n^{(l)})$ is the prior and $p(o_n^{(r)} | o_n^{(l)}, d_n)$ is the image likelihood.

According to triangle division and disparity threshold α , the prior $p(d_n | S, E, o_n^{(l)})$ is defined as:

$$p(d_n | S, E, o_n^{(l)}) \propto \begin{cases} p(d_n | S, o_n^{(l)}), & \text{if } e_p \notin t_q \vee |d_{qi} - d_{qj}| < \alpha \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Making use of the color information of support points, and with the assumption piecewise disparity smooth, the prior is defined as:

$$p(d_n | S, o_n^{(l)}) \propto \begin{cases} \gamma + \exp\left(-\frac{(d_n - \mu_d(S, o_n^{(l)}))^2}{2\sigma^2}\right) \times \rho, & \text{if } |d_n - \mu| < 3\sigma \vee d_n \in N_S \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where ρ is defined as color prior, we build color prior with a similar method as for disparity prior.

$$\rho = \exp\left(-\frac{(1 - \mu_c(S, o_n^{(l)}))^2}{2\sigma^2}\right) \quad (4)$$

This includes a uniform part and a Gaussian part. The estimated disparity $\mu_d(S, o_n^{(l)})$, $\mu_c(S, o_n^{(l)})$ in the Gaussian part depends on the Delaunay triangulation on the support points and the pixel coordinates, and N_S is the set of all support point disparities in a small neighborhood around $(u_n^{(l)}, v_n^{(l)})$. For time efficiency, we do not calculate the neighborhood for each pixel, instead, we divide image region into a 20×20 block size, and support points disparities within each block are calculated for pixels within that block. To calculate the computing efficiency, all disparities farther than 3σ from the estimated disparity are excluded. The condition $d_n \in N_S$ enables the prior to locally extend its range to better handle disparity discontinuities in place where the linearity assumption might be violated.

$\mu(S, o_n^{(l)})$ is defined as a piecewise linear function, which interpolates the disparities using support points as triangle vertices. For each triangle, the plane is defined as:

$$\mu_q(o_n^{(l)}) = a_q u_n + b_q v_n + c_q \quad (5)$$

where q is the index of the triangle, which the pixel (u_n, v_n) belongs to, and $o_n = (u_n, v_n, f_n)^T$ is an observation. For each triangle, the plane parameters (a_q, b_q, c_q) are obtained by solving a linear plane problem. Therefore, the proposed prior, μ , is a linear interpolation between support point disparities.

The image likelihood is modeled using a constrained Laplace distribution, which is robust against outlier:

$$p(o_n^{(r)} | o_n^{(l)}, d_n) \propto \begin{cases} \exp(-\beta \|f_n^{(l)} - f_n^{(r)}\|_1), & \text{if } \begin{pmatrix} u_n^{(l)} \\ v_n^{(l)} \end{pmatrix} = \begin{pmatrix} u_n^{(r)} + d_n \\ v_n^{(r)} \end{pmatrix} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where $f_n^{(l)}$, $f_n^{(r)}$ are feature vectors in the left and right images respectively, and β is a constant. The if-condition ensures that correspondences are located on the same epipolar line and matched through the disparity d_n .

A prior and an image likelihood are proposed for stereo matching. We maximize a-posteriori (MAP) estimation to compute the disparities:

$$\begin{aligned} d_n^* &= \arg \max p(d_n | o_n^{(l)}, o_1^{(r)}, \dots, o_N^{(r)}, S, E) \\ &= \arg \max p(d_n | S, E, o_n^{(l)}) p(o_1^{(r)}, \dots, o_N^{(r)} | o_n^{(l)}, d_n) \\ &= \arg \max p(d_n | S, E, o_n^{(l)}) \prod_{j=1}^N p(o_j^{(r)} | o_n^{(l)}, d_n) \end{aligned} \quad (7)$$

Maximizing the posterior probability is equivalent to minimizing the negative logarithmic energy. To find the winning disparity, we minimize the energy shown in (8).

$$\begin{aligned} E(d) &= \beta \|f_n^l - f_n^r(d)\|_1 - \log(\rho) - \\ &\log \left[\gamma + \exp \left(-\frac{(d_n - \mu(S, o_n^{(l)}))^2}{2\sigma^2} \right) \right], \\ &\text{if } |d_n - \mu| < 3\sigma \vee d_n \in N_S \end{aligned} \quad (8)$$

For the triangles, which are disparity discontinuity areas, local method is used to compute disparity. Because piecewise disparity smooth is not true, the prior cannot provide proper information for disparity calculation. For these areas, enlarged disparity searching range is used. We use half image width for the maximum disparity searching range.

IV. RESULT

We evaluated our algorithm on the Middlebury Stereo Evaluation Dataset-Version 3 from 2014 [35]. Our algorithm performs better with sparse datasets, because we didn't use an advanced post-processing method for refinement. In the following analysis, we focus on the sparse evaluation. In Middlebury benchmark, sparse disparity map leaves some pixels unlabeled, typically in half-occluded or texture-less regions. This new Middlebury dataset is more challenging compared to the previous datasets. It includes images taken in industrial environments and images taken in indoor environments. The resolution is up to 3000x2000 pixels on the full resolution images. We implement the proposed algorithm using C++ under a desktop PC with Intel Core i5-6500 processor 3.20GHz and 8GB RAM. All results are single core performance. We selected 4 popular algorithms and compared them with our method on Middlebury dataset. Table 1 shows the Middlebury evaluation tables for the sparse dataset. We chose test sparse as set and nonocc as mask. The error rate of our method is lower than that of ELAS, LPS and TSGO [10]. SGM shows a lower error value than our method, but it is a slower method. Tables 2 shows a comparison of the algorithms running time. Our algorithm is slower than ELAS, but faster than other three algorithms. Compared with ELAS, we implemented edge detection method in our method, which takes more time. Table 3 shows the comparison result for all the Middlebury metrics on the test sparse data set. Our algorithm outperforms ELAS on accurate but underperforms it on time.

TABLE I. PERFORMANCE EVALUATION ON THE MIDDLEBURY DATASET, BAD 2.0, % ERROR, SPARSE.

Algorithm	Avg	Bad 2.0														
		<i>Adir</i>	<i>ArtL</i>	<i>Jadep</i>	<i>Motor</i>	<i>MotorE</i>	<i>Piano</i>	<i>PianoL</i>	<i>Pipes</i>	<i>Playrm</i>	<i>Playt</i>	<i>PlaytP</i>	<i>Recyc</i>	<i>Shelves</i>	<i>Teddy</i>	<i>Vintge</i>
SGM(Q)	13.7	12.8	9.84	13.3	11.9	10.9	15.1	13.7	10.3	19.3	25.3	15	13.8	28.5	5.74	17.8
OURS(Q)	21.5	23.4	16.4	17.3	23.4	23.2	23.4	23.4	18.7	27	26.5	21.3	25.1	28.7	12.7	23
LPS(H)	23.2	11.4	32.3	21.7	5.72	68.5	19	50.2	8.89	25.9	20.9	18	16.5	39.7	6.96	26.6
ELAS(Q)	28.4	29.3	19.1	26.3	25.2	27.8	33.5	38	22	38.2	32.8	27.3	32.1	43.6	18.4	34.8
TSGO(F)	31.3	27.3	12.3	53.1	23.5	25.7	33.4	54.5	22.5	49.6	45	27	24.2	52.2	13.3	57.5

TABLE II. PERFORMANCE EVALUATION ON THE MIDDLEBURY DATASET, TIME(S)/MP, SPARSE.

Algorithm	Avg	Time per Megapixel														
		<i>Adir</i>	<i>ArtL</i>	<i>Jadep</i>	<i>Motor</i>	<i>MotorE</i>	<i>Piano</i>	<i>PianoL</i>	<i>Pipes</i>	<i>Playrm</i>	<i>Playt</i>	<i>PlaytP</i>	<i>Recyc</i>	<i>Shelves</i>	<i>Teddy</i>	<i>Vintge</i>
SGM(Q)	0.38	0.37	0.42	0.4	0.38	0.38	0.38	0.38	0.39	0.39	0.41	0.38	0.37	0.38	0.36	0.4
OURS(Q)	2.92	2.84	2.91	3.91	2.82	2.82	2.71	2.71	2.95	2.98	3.02	2.86	2.69	2.73	2.61	3.46
LPS(H)	3.13	3.04	2.58	4.97	2.83	2.83	2.78	2.86	3.06	3.26	3.1	3.05	2.81	2.59	2.49	5.55
ELAS(Q)	7.14	5.15	5.45	28.1	4.52	4.67	5.44	5.08	4.87	5.78	5.48	5.46	5.3	5.53	6.56	5.54
TSGO(F)	11.4	11	21.8	7.22	11.6	11.8	10.7	11	12	4.44	12	12.4	10.4	9.9	11	8.24

TABLE III. PERFORMANCE EVALUATION ON THE MIDDLEBURY DATASET, SPARSE, OURS VS ELAS.

Algorithm	Percentage of wrong matches averaged for all images. Set: sparse, Mask: nonocc											
	<i>Bad0.5</i>	<i>Bad1</i>	<i>Bad2</i>	<i>Aerr</i>	<i>rms</i>	<i>A50</i>	<i>A90</i>	<i>A95</i>	<i>A99</i>	<i>Sec</i>	<i>Sec/mp</i>	<i>Sec/gd</i>
OURS(Q)	65.3	45.7	21.5	9.49	3.9	13.2	1.22	5.3	14.7	65.	0.91	2.92
ELAS(Q)	73	53.3	28.4	14.1	4.9	15	1.34	8.73	25.3	72.	0.12	0.38

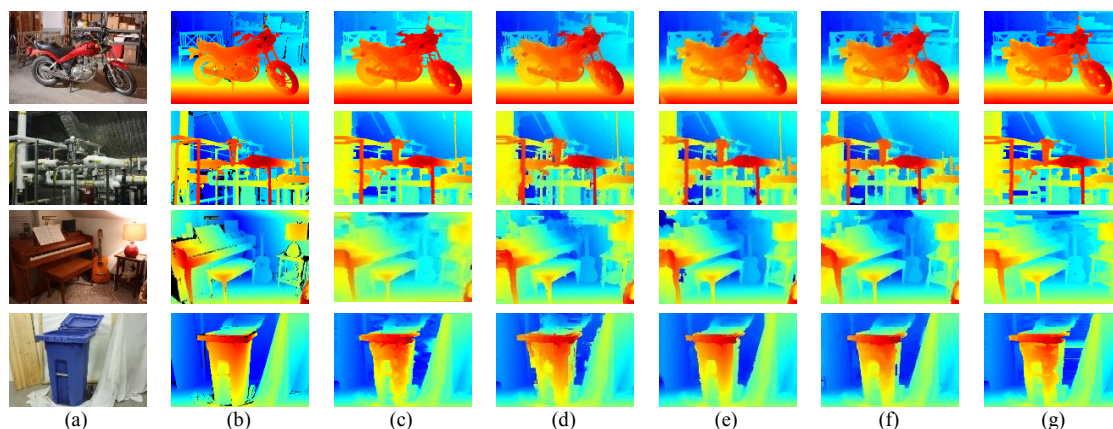


Fig. 3. Results on a subset of Middebury2014 test images with ground truth (Motorcycle, Pipes, Piano, and Recycle). (a) Left image, (b) disparity ground truth, (c) result of ELAS, (d) result of LPS, (e) result of SGM, (f) result of TSGO, (g) result of our method.

In our algorithm, we distinguish discontinuities area based on edge information, and use the local method for the calculation. Compared with ELAS, it is obvious that our method conserves better discontinuity information in disparity image (Fig. 3). For a flat area, our method shows better consistence than ELAS. This is because our algorithm strengthens disparity prior based on color prior. LPS often makes blur edges (for instance the Motorcycle). TSGO loses thin objects (for instance the Pipes).

V. CONCLUSIONS

In this paper, we have proposed an approach to divide image region into groups based on extracted image edge information and strengthen disparity prior based on color prior. The experimental results are better than those of other popular algorithms. However, while using the local method, it is very difficult to obtain the correct disparity searching range. Even though we have enlarged the disparity searching space for the disparity discontinuity areas, the searching range is still partially valid for some image regions because the true disparity of these pixels may be outside the range. In our method, the disparity relationship between neighbor pixels is not considered, this relationship will be considered in future work.

ACKNOWLEDGMENT

This research was funded by the National Nature Science Foundation of China (Grant No. 91746111, Grant No. 71702143, Grant No. 71731009, Grant No. 71732006), Ministry of Education & China Mobile Joint Research Fund Program (No. MCM20160302), Shanghai Pujiang Program (Grant No. 5PJ060).

REFERENCES

- [1] Leibe, B., Seemann, E., and Schiele, B., 2005, "Pedestrian Detection in Crowded Scenes," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 878–885 vol. 1.
- [2] Jia, L. H., Zou, J. H., and Che, K., 2011, "Human Identification Based on 3D Tracking Trajectory of Head Vertex," *Acta Autom. Sin.*, **37**(1), pp. 28–36.
- [3] Mi, C., Liu, H., Huang, Y., Mi, W., and Shen, Y., 2016, "Fatigue Alarm Systems for Port Machine Operators," *Asia Life Sci.*, **25**(1), pp. 31–41.
- [4] Mi, C., Shen, Y., Mi, W., and Huang, Y., 2015, "Ship Identification Algorithm Based on 3D Point Cloud for Automated Ship Loaders," *J. Coast. Res.*, **73**, pp. 28–34.
- [5] Ren, M., Liu, R., Hong, H., Ren, J., and Xiao, G., 2017, "Fast Object Detection in Light Field Imaging by Integrating Deep Learning with Defocusing," *Appl. Sci.*, **7**(12), p. 1309.
- [6] Jia, L., and Liu, Y., 2008, "A Novel Thresholding Approach to Background Subtraction," *2008 IEEE Workshop on Applications of Computer Vision*, pp. 1–6.
- [7] Geiger, A., Roser, M., and Urtasun, R., 2011, "Efficient Large-Scale Stereo Matching," *Proceedings of the 10th Asian Conference on Computer Vision - Volume Part I*, Springer-Verlag, Berlin, Heidelberg, pp. 25–38.
- [8] Scharstein, D., and Szeliski, R., 2002, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *Int. J. Comput. Vis.*, **47**(1–3), pp. 7–42.
- [9] Sun, J., Zheng, N.-N., and Shum, H.-Y., 2003, "Stereo Matching Using Belief Propagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, **25**(7), pp. 787–800.
- [10] Mozerov, M. G., and Weijer, J. van de, 2015, "Accurate Stereo Matching by Two-Step Energy Minimization," *IEEE Trans. Image Process.*, **24**(3), pp. 1153–1163.
- [11] Bleyer, M., and Gelautz, M., 2004, "A Layered Stereo Algorithm Using Image Segmentation and Global Visibility Constraints," *2004 International Conference on Image Processing, 2004. ICIP '04.*, pp. 2997–3000 Vol. 5.
- [12] Kolmogorov, V., and Zabih, R., 2001, "Computing Visual Correspondence with Occlusions Using Graph Cuts," *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, pp. 508–515 vol.2.
- [13] Li, J., Shi, W., Deng, D., Jia, W., and Sun, M., 2012, "Bayesian Stereo Matching Method Based on Edge Constraints," *Int J Adv Comput Technol*, **4**(22), pp. 36–47.
- [14] Yang, Q., Wang, L., Yang, R., Stewénius, H., and Nistér, D., 2009, "Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation, and Occlusion Handling," *IEEE Trans. Pattern Anal. Mach. Intell.*, **31**(3), pp. 492–504.
- [15] Lei, C., Selzer, J., and Yang, Y.-H., 2006, "Region-Tree Based Stereo Using Dynamic Programming Optimization," *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, pp. 2378–2385.
- [16] Sun, J., Li, Y., Kang, S. B., and Shum, H.-Y., 2005, "Symmetric Stereo Matching for Occlusion Handling," *2005 IEEE Computer*

- [17] Birchfield, S., and Tomasi, C., 1998, “Depth Discontinuities by Pixel-to-Pixel Stereo,” *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pp. 1073–1080.
- [18] Van Meerbergen, G., Vergauwen, M., Pollefeys, M., and Van Gool, L., 2002, “A Hierarchical Symmetric Stereo Algorithm Using Dynamic Programming,” *Int J Comput Vis.*, **47**(1–3), pp. 275–285.
- [19] Veksler, O., 2005, “Stereo Correspondence by Dynamic Programming on a Tree,” *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 384–390 vol. 2.
- [20] Klaus, A., Sormann, M., and Karner, K., 2006, “Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure,” *18th International Conference on Pattern Recognition (ICPR '06)*, pp. 15–18.
- [21] Hirschmuller, H., 2008, “Stereo Processing by Semiglobal Matching and Mutual Information,” *IEEE Trans Pattern Anal Mach Intell*, **30**(2), pp. 328–341.
- [22] Hirschmüller, H., Innocent, P. R., and Garibaldi, J. M., 2002, “Real-Time Correlation-Based Stereo Vision with Reduced Border Errors,” *Int. J. Comput. Vis.*, **47**(1–3), pp. 229–246.
- [23] Hosni, A., Bleyer, M., and Gelautz, M., 2013, “Secrets of Adaptive Support Weight Techniques for Local Stereo Matching,” *Comput Vis Image Underst.*, **117**(6), pp. 620–632.
- [24] Zhang, K., Lu, J., and Lafruit, G., 2009, “Cross-Based Local Stereo Matching Using Orthogonal Integral Images,” *IEEE Trans. Circuits Syst. Video Technol.*, **19**(7), pp. 1073–1079.
- [25] Veksler, O., 2006, “Reducing Search Space for Stereo Correspondence with Graph Cuts,” *British Machine Vision Conference 2006, Edinburgh, Uk, September*, pp. 709–718.
- [26] Wang, L., Jin, H., and Yang, R., 2008, “Search Space Reduction for MRF Stereo,” *Proceedings of the 10th European Conference on Computer Vision: Part I*, Springer-Verlag, Berlin, Heidelberg, pp. 576–588.
- [27] Yamaguchi, K., Hazan, T., McAllester, D., and Urtasun, R., 2012, “Continuous Markov Random Fields for Robust Stereo Estimation,” *Proceedings of the 12th European Conference on Computer Vision - Volume Part V*, Springer-Verlag, Berlin, Heidelberg, pp. 45–58.
- [28] Cech, J., and Sara, R., 2007, “Efficient Sampling of Disparity Space for Fast and Accurate Matching,” *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8.
- [29] Moallem, P., and Faez, K., 2002, “Fast Edge-Based Stereo Matching Algorithm Based on Search Space Reduction,” *Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing*, pp. 587–596.
- [30] El jaafari, I., El Ansari, M., and Koutti, L., 2016, “Fast Edge-Based Stereo Matching Approach for Road Applications,” *Signal Image Video Process.*, **11**.
- [31] Sinha, S. N., Scharstein, D., and Szeliski, R., 2014, “Efficient High-Resolution Stereo Matching Using Local Plane Sweeps,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1582–1589.
- [32] Bobick, A. F., and Intille, S. S., 1999, “Large Occlusion Stereo,” *Int. J. Comput. Vis.*, **33**(3), pp. 181–200.
- [33] Hu, X., and Mordohai, P., 2010, “Evaluation of Stereo Confidence Indoors and Outdoors,” *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1466–1473.
- [34] Sára, R., 2002, “Finding the Largest Unambiguous Component of Stereo Matching,” *Proceedings of the 7th European Conference on Computer Vision-Part III*, Springer-Verlag, Berlin, Heidelberg, pp. 900–914.
- [35] Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nescic, N., Wang, X., and Westling, P., 2014, “High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth,” *GCPR*, X. Jiang, J. Hornegger, and R. Koch, eds., Springer, pp. 31–42.