Improved Weighted Median Filter with Superpixel for Disparity Refinement

Yuli Fei[1, a], Li Cao[1, b]

[1]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics,

Nanjing 211106, China;

[a]feiyuli223@163.com, [b]caoli@nuaa.edu.cn

Abstract. Stereo matching cannot get high accuracy for disparity estimations, especially at depth boundaries and textureless regions. To solve the problems, a disparity refinement method based on superpixel segmentation is proposed. A weighted median filter with superpixel information is designed. We give a penalty factor for the neighborhood pixels that are not within the same superpixel. Some experiments are done on the Middlebury dataset. The results show that the proposed method can reduce the mismatch rates around occlusion regions and textureless regions, and obtain a highly accurate disparity map.

Keywords: Stereo Matching, Disparity Refinement, Superpixel Segmentation.

1. Introduction

Stereo matching has always been a key problem in computer vision. For binocular stereo, the goal is to find corresponding points in a pair of image at the same scene. Stereo vision has a lot of important applications, including navigation, 3D reconstruction and image rendering. Most stereo matching methods take four steps: 1. cost computation; 2. cost aggregation; 3. disparity optimization 4. disparity refinement[1].

Most previous studies are about the first three steps. Global algorithms mostly focus on disparity optimization (step 3).They often get a good result, but are time-consuming. And local methods have focused on robust cost computation (step 1) and edge-aware cost aggregation (step 2) [2]. However, the impact of disparity refinement has attracted far less attention in the literature. The traditional method for disparity refinement rely on pixelwise support region, such as plane-fitting, scanline optimization and interpolation. Yang [3] proposed a non-local refinement method on a minimum spanning tree (MST) with a new non-local cost aggregation. Xie [4] proposed a refinement algorithm based on RANSAC and works well in textureless regions. Segmentation method is used with slant for estimating the disparity plane [5].

Weighted median filter is widely used in stereo matching. It is an edge-preserving filter and reduces the blur effect for disparity refinement. A brute-force implementation of WMF is time-consuming. a constant WMF algorithm is firstly proposed by Ma [6]. They discovered that the refinement step is as important as the other three step.

In this paper, we proposed a post-processing method that takes associated superpixel information into weighted median filter. Superpixel provide an effective initial information from which to use local image feature. We get a high quality superpixel in a low computation time, and compute a new refined disparity map while preserving edges. The proposed method achieves better performance than traditional WMF.

2. Disparity Refinement with Superpixel

2.1 Initial Disparity Map.

The matching cost is important in stereo matching. In this paper, the cost function we employed is based on gradient and Census transform descriptor, which is proved to be powerful for the robust optical flow computation [7] .And the cost, can be formulated as follows:

$$C(i,l) = (1-\lambda)\min(\|\nabla I(i) - \nabla I(i_l)\|, \tau_1) + \lambda \min(H(Cen(i), Cen(i_l)), \tau_2). \quad (1)$$

Here $I(i)$ denotes the intensity of pixel i . ∇ represents the gradient operator. i_l is the corresponding pixel of i in the right image with a disparity $Cen(i)$ l represents the Census transform value of pixel i. $H(\cdot,\cdot)$ calculates the Hamming distance between two values. λ is a scalar to balance the Census and gradient terms and τ_1 , τ_2 are truncation values.

Since the cost volume C is always noisy, we reformulate the cost function with weighted least square (WLS) to denoise it [8]. The cost aggregation can be formulated as:

$$\mathring{C}(i,l) = \arg\min_z \frac{1}{Z_i} \sum_{j \in N_i} K(i,j) \|z - C(j,l)\|^2 \quad (2)$$

Where N_i define a neighboring system of pixel i .$K(i,j)$ is the similarity between neighboring pixels i and j called similarity kernel [9]. $Z_i = \sum_{j \in N_i} K(i,j)$ is a normalization constant. The solution of Eq. (2) is:

$$\mathring{C}(i,l) = \frac{1}{Z_i} \sum_{j \in N_i} K(i,j) C(j,l) \quad (3)$$

Thus, different choice of cost aggregation method leads to different similarity kernel. For example, the NL method describes the kernel based on geodesic distance of two pixels in a tree structure. The bilateral filter measure the similarity by the spatial and photometric between two pixels. Rhemann [2] adopted the GF aggregation method, whose kernel is the same as guide filter [10].Then, the WTA strategy is adopted to get an initial disparity map.

2.2 Disparity refinement filter.

SLIC superpixel algorithm generates labels by cluster pixels based on the five-dimensional vector space $[labxy]$. And $[lab]$ is the pixel color in CIELAB color space which is more uniform for small color distances; $[xy]$ is the pixel coordinate. Euclidean distance in CIELAB color space is perceptually useful for small distance. If spatial pixel distances exceed color distance in large superpixel, they outweigh color proximity. In order to combine two distances into one measure, the measurement D can be formulated as follows [11]:

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2}$$
$$D = \sqrt{{d_c}^2 + (\frac{d_s}{S})^2 m^2} \quad (4)$$
$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

Here, m is used for weighing the importance between color and spatial. The larger m results in more compact superpixels. On the contrary, a smaller m makes superpixels adhere more tightly to boundaries. In the CIELAB color space, m can be in the range from 1 to 40. Fig.1 shows a result of SLIC superpixel segmentation which m is set to 10.

Weighted median filter is a effective method for disparity refinement. In [2] the weight of filter is the bilateral weight that suppresses the pixel with different color form center. The weight consists of distance term and color term:

$$W_p = w_d w_c \quad (5)$$

The weights w_d and w_c are assigned according to Gaussian function with the variance σ_d and σ_c , respectively. The distance between coordinate point and between triplets in color space are computed in L2 norm:

$$\begin{aligned} w_d &= G_{\sigma_d}(\|p-q\|) \\ w_c &= G_{\sigma_c}(\|I(p)-I(q)\|). \end{aligned} \tag{6}$$

We proposed a new weighted filter based on superpixel. The filter is modified the former weighted median filter (WMF) with a superpixel weight. The new weight $\overset{\frown}{W}_p$ can be decomposed into three terms based on segmentation, distance, and color similarity as follows:

$$\overset{\frown}{W}_p = w_d w_c w_s = W_p w_s \tag{7}$$

The second term is defined based on superpixel using SLIC superpixel segmentation as shown in Fig. 1. If the neighborhood that are not within the same superpixel, we give a penalty term:

$$w_s = \begin{cases} 1 & if \quad label(p) = label(q) \\ a & otherwise \end{cases} \tag{8}$$

Where $label(\cdot)$ is the superpixel label, α is the penalty factor with its value between 0 from 1. In this implementation, we set it to 0.6 from experience.

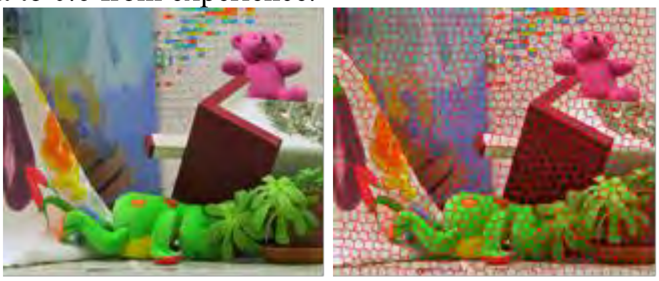
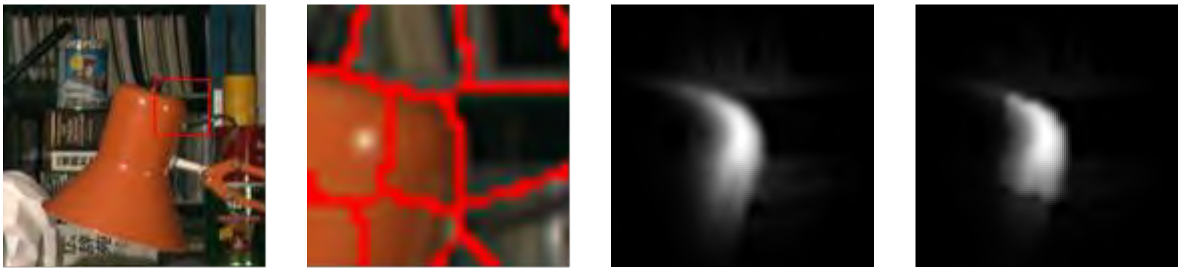


Fig. 1 Result of SLIC superpixel



(a) Part of left image (b) Result of superpixel (c) WMF kernel (d) Proposed kernel

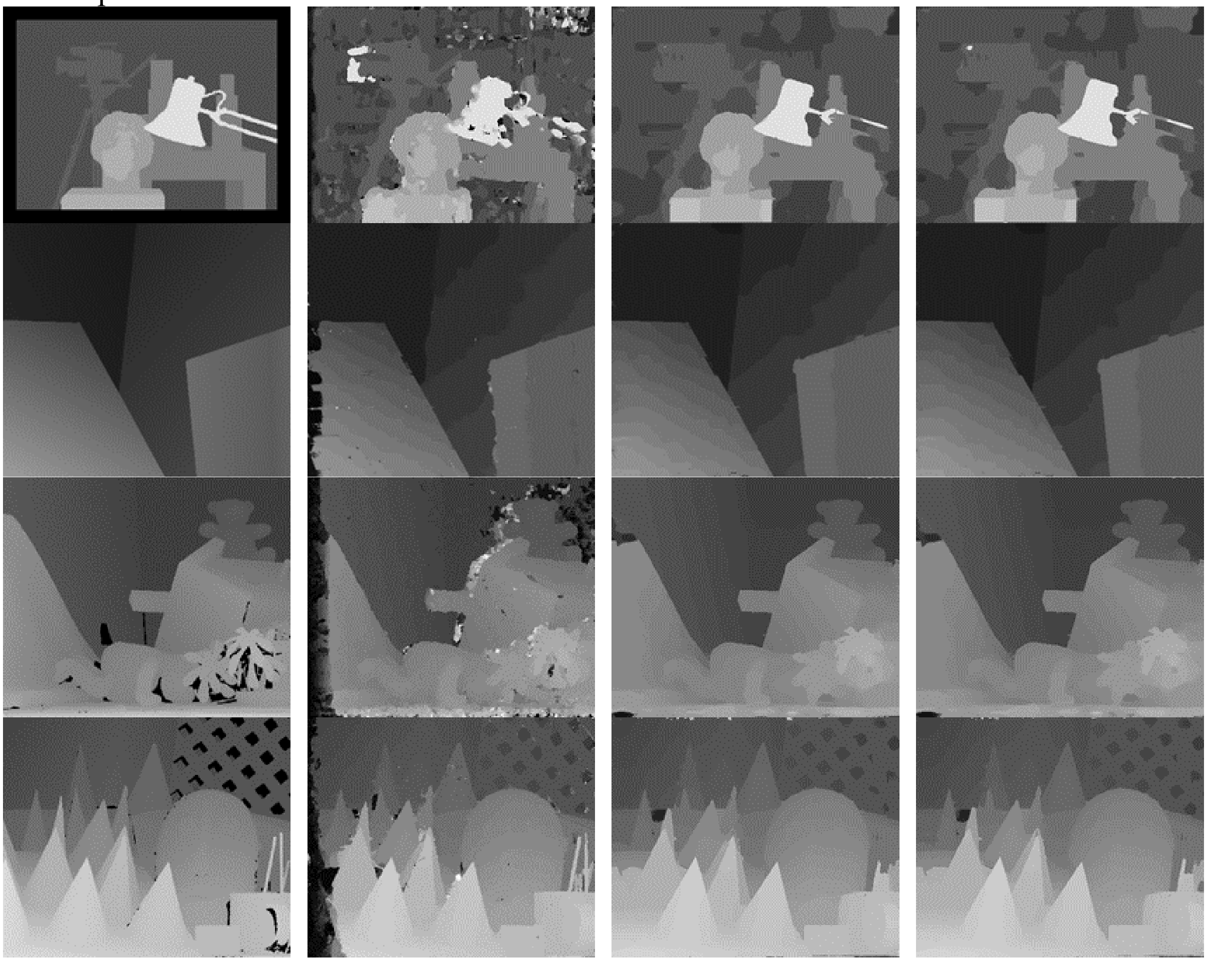
Fig. 2 Weighted filter kernel weights comparision: WMF and proposed method

The traditional weighted median filter is edge-preserving. With appropriate parameters, superpixel segmentation can adhere tightly to image boundaries. So our proposed method which is combined with WMF and superpixel segmentation is also edge-preserving. Weighted median filter is widely used in disparity refinement, but it’s still not performing well, especially in textureless region or occlusive area. Our proposed method adds the superpixel information to the traditional WMF. Figure 2 (c) and (d) give kernel’s visual comparisons between WMF and proposed method. Figure 2 (b) shows the details of superpixel segmentation with the zoom-in region. We assume that the mismatching point’s depth is closer to points in the same superpixel. As we can see, our proposed method decreases the weights outside the superpixel to which the center pixel belongs. So the weights gather in the superpixel region. That makes the outliers points refine to the exact disparity.

3. Experiment Result and Analysis

In this section, we compare the performance of the proposed method with traditional WMF method on the dataset provided by Middlebury. Four pairs image (Tsukuba, Venus, Teddy, Cones) are tested and evaluated with Middlebury benchmark. Table 1 and Table 2 gives the quantitative result which means the percentage of bad matching pixels in three subsets of an image. “All” represents pixels in nonoccluded and half-occluded regions, “Noc” represents pixels in nonoccluded regions, and “Disc” represents the visible pixels near the occluded regions.

Evaluations for WMF and the proposed method are tested with two different cost aggregation method: box filter and guide filter. And the error thresholds is set as 1. For visual comparison, we present the disparity results of the data sets in Fig.3 which cost aggregation method is box filter. Furthermore, before the refinement, left-right consistency check (LRC) is adopted to detect the mismatch points.



(a) Ground truths (b) Initial disparity (c) WMF refinement (d) Proposed method

Fig. 3 Result of refinement using WMF and the proposed method in Middlebury data sets.

Table 1 Evaluation for WMF and the proposed method with box filter aggregation

Algorithm	Tsukuba			Venus			Cone			Teddy		
	Noc	All	Disc	Noc	All	Disc	Noc	All	Disc	Noc	All	Disc
WMF	3.05	3.22	15.36	0.20	0.37	2.15	4.28	9.39	12.04	6.48	11.71	17.13
Our method	3.01	3.21	15.04	0.18	0.30	2.08	3.87	9.05	11.10	6.43	11.69	17.01

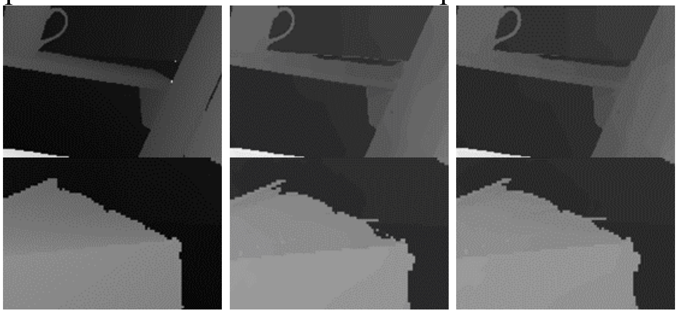
Table 2 Evaluation for WMF and the proposed method with guide filter aggregation

Algorithm	Tsukuba			Venus			Cone			Teddy		
	Noc	All	Disc	Noc	All	Disc	Noc	All	Disc	Noc	All	Disc
WMF	2.74	2.96	12.67	0.26	0.41	2.65	4.43	10.20	12.23	6.52	11.63	17.05
Our method	2.76	3.00	12.39	0.18	0.31	2.24	4.32	10.09	12.04	6.34	11.56	16.78

As Table 1 and 2 show, our proposed method outperforms WMF with marginal improvement. Especially, in the depth discontinuity areas, our method gets a much smaller error rate. It is well suited to removing error points in disparity discontinuity region. For the different cost aggregation algorithm, the proposed refinement method achieved a better performance. On the other hand, when using our refinement method, even the simple box filter can reach an impressive improvement. From the Fig. 3

we can see, the proposed obtain a better disparity on occluded region and removes the noise near the discontinuity areas.

We also evaluated our method on the 2014 Middlebury Data Set, as we can see in the Fig. 4 which is the detail comparison between WMF and ours. Due to the superpixel attach tightly to the boundary of the object, the proposed method effectively separate each objects with different depth. At the boundary regions, the proposed method matches better compared with the WMF.



(a) Ground-truth (b) WMF (c) Proposed

Fig. 4 Detail comparison of ground truth, WMF and proposed method on 2014 Middlebury Data Set.

4. Conclusion

In this paper, we propose an effective disparity refinement algorithm with superpixel information. We used weighted least square (WLS) at cost aggregation step to reduce the noise and get an initial disparity map. Then, we adopted SLIC superpixel segmentation in the reference image. So we use the superpixels which contain edge information and pixel’s similarity to compute a new depth map. Compared with traditional WMF, the proposed method achieves better performance. Particularity, at the discontinuity area, our method gets a better edge-preserving result.

Acknowledgments

This work was supported by National Natural Science Foundation of China (U1633105).

References

[1]. D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47:7–42, 2002.

[2]. C.Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual correspondence and beyond. In CVPR, pages 3017–3024, 2011.

[3]. Q. Yang. A non-local cost aggregation method for stereo matching. In CVPR, 2012.

[4]. Y. Xie , S. Zeng , L. Chen. A Novel Disparity Refinement Method Based on Semi-Global Matching Algorithm. IEEE International Conference on Data Mining Workshop, 2015.

[5]. A. Klaus, M. Sormann, K. Karner. Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure. Pattern Recognition, pages 3:15-18,2006

[6]. Z. Ma, K. He, Y. Wei, J. Sun, and E. Wu. Constant time weighted median filtering for stereo matching and beyond. In ICCV, 2013.

[7]. D. Hafner, O. Demetz, and J. Weickert. Why is the census transform good for robust optic flow computation? In SSVM, 2013.

[8]. K. Zhang, Y. Fang, D. Min, and L. Sun. Cross-scale cost aggregation for stereo matching. In CVPR, 2014.

[9]. P. Milanfar. A tour of modern image filtering: new insights and methods, both practical and theoretical. IEEE Signal Processing Magazine, 2013.

[10]. K. He, J. Sun, and X. Tang. Guided image filtering. In ECCV, 2010.