An Interleaving Updating Framework of Disparity and Confidence Map for Stereo Matching

Chenbo SHI(a), Student Member, Guijin WANG(b), Member, Xiaokang PEI†, Nonmember, Bei HE†, Student Member, and Xinggang LIN†, Nonmember

SUMMARY In this paper, we propose an interleaving updating framework of disparity and confidence map (IUFDCM) for stereo matching to eliminate the redundant and interfere information from unreliable pixels. Comparison with other propagation algorithms using matching cost as messages, IUFDCM updates the disparity map and the confidence map in an interleaving manner instead. Based on the Confidence-based Support Window (CSW), disparity map is updated adaptively to alleviate the effect of input parameters. The reassignment for unreliable pixels with larger probability keeps ground truth depending on reliable messages. Consequently, the confidence map is updated according to the previous disparity map and the left-right consistence. The top ranks on Middlebury benchmark corresponding to different error thresholds demonstrate that our algorithm is competitive with the best stereo matching algorithms at present.

key words: stereo matching, message propagation, confidence map, interleaving updating

1. Introduction

Stereo matching has gained much attention over the past decades, and remains an active research topic in the community of computer vision. The central difficulty of stereo matching lies in accurately estimating the disparity map from image pair under the scenarios of smooth region and occluded areas [1].

In general, stereo matching algorithms can be classified into local-based and global-based approaches. Most of the early stereo matching algorithms are local-based, which calculates the disparity at a given pixel depending on intensity values within a finite window. But the ambiguity in textureless areas and occlusions is still residual. In recent years, some effective local-based methods were proposed to address this problem. Yoon [2] and Hosni [3] presented different weighted support windows. Humenberger [12] used plane model in segments on both color and texture images. Confidence-based Support Window (CSW) [4] combined the color similarity with reliable pixels in piecewise fitting for ambiguous pixels. However, due to the limitation of local windows, the performance of these methods is not so good as global-based methods, which treat the images as a Markov Random Field (MRF) to solve an optimization problem. Global-based algorithms handle textureless regions well but is poor on the discontinuous edges [1].

Some algorithms achieved better results on discontinuity, using color segmentation [6], matting [7] and weighted support window [9]. But most global optimization algorithms have complex model and are sensitive to parameters.

In this paper, a novel interleaving updating framework of disparity and confidence map (IUFDCM) is proposed for stereo matching, which effectively combines the local weighted support window with message propagation. We use the confidence map to denote the reliability of disparity map. Different from other algorithms which use the matching cost as messages, our algorithm updates the two maps in an interleaving manner instead. The disparity map is updated adaptively based on CSW to alleviate the effect of input parameters. Benefiting from the updated disparity map and the left-right consistence between frames, the confidence map is effectively updated to achieve adaptive neighborhood selection to help the messages propagation. Experiments show that our algorithm achieves good results both on discontinuity and smooth regions. The top ranks on Middlebury benchmark corresponding to different error thresholds sufficiently demonstrate that our algorithm is competitive to the best algorithms up to now.

2. Interleaving Framework

Generally, for most stereo matching algorithms, the basic assumption is that pixels with similar color have similar disparities. But in contrast, the pixels whose color values are different from the dominant background have more reliable messages. Our work [4] shows that the information from reliable points is more efficient than others. However, the credibility of disparity reassignment has not been evaluated effectively. The ambiguity is residual in low confidence regions. In addition, those unreliable pixels have multiple reassigned values because of the matching ambiguity while others with ground truth are more likely to keep stable. To handle this problem, our novel updating framework is presented in Fig. 1. There are 4 component parts. The first step is to initialize the disparity map and confidence map of the input left and right images, respectively. Then by utilizing the confidence map, we update the disparity map to reassign new disparity values for unreliable pixels. Furthermore, based on the previous disparity map and the left-right consistence between images, the confidence map is updated to evaluate new disparity reassignment. The process will break once it reaches the maximum iterations or the num-
ber of unreliable pixels doesn’t decrease anymore. Finally, the refinement using weighted median filter improves the performance on the discontinuity. The details will be introduced in the following sections.

3. Details of IUFDCM

3.1 Initialization of Disparity and Confidence Map

A widely used matching cost is the Sum of Absolute Difference (SAD) between pixel intensities of the left and right frames: \( \text{dis}(p, p') = |r_p - r_{p'}| + |g_p - g_{p'}| + |b_p - b_{p'}| \). The matching cost \( \text{Cost}(x, y, d) \) in a small window is applied in our initialization because of its simplicity. The best candidate is selected by Winner Take All method:

\[
D(p) = \arg \min_d \text{Cost}(x, y, d)
\]

\[
= \arg \min_d \sum_{q \in \omega} \text{dis}(L(q_x, q_y), R(q_x - d, q_y)),
\]  

where \( L, R \) are the left and right images, \( d \) is the disparity of pixel \( q \) and \( \omega \) represents the local window. Based on \( \text{Cost}(x, y, d) \), we define the confidence map \( C(p) \) by the ratio of minimum cost and second minimum cost.

\[
C(p) = \left( 1 - \frac{\min_d \text{Cost}(p_x, p_y, d)}{\min_d \text{Cost}(p_x, p_y, d)} \right) \times 100. 
\]  

The closer two minimums are, the less reliable the candidate is. We normalize the confidence value in a range from 0 to 100 which is easy for storage and display.

3.2 Disparity Map Updating

The initialization step calculates correct disparity for some high confidence pixels. However, a lot of pixels still need to be reassigned. We proposed the local plane model of disparity to evaluate the disparity for unreliable pixels in [4]. Only pixels whose confidence values are less than the confidence threshold \( T_s \) will be selected as centers of CSW. Different from [12], we use the color similarity and confidence to select CSW instead of color segmentation to avoid the sensitivity to parameters. Then RANSAC is employed to solve local plane fitting problem, which is robust to suppress the disturbance resulted from outliers. However, it is not easy to obtain consistent disparity values in different regions. So we adjust the parameters adaptively in the updating process. After certain iterations, there are still some low confidence pixels which cannot be correctly evaluated corresponding to current parameters. To assist the message propagation between distanced pixels, the size of support window is increased by a step of \( S_w = 2 \). Some pixels don’t have enough neighbors with similar color under current color threshold for its reassignment. It needs more reliable neighbors in a wider color range \( T_s' \) in LAB color space as following equation.

\[
T_s' = \max(T_{\text{max}}, T_{s}^{t-1} + 0.1). 
\]  

where \( t \) is iteration number and the upper threshold \( T_{\text{max}} \) is set to avoid the incorrect merger between quite different regions. The wider window size and color range help separated small regions find the corresponding suitable plane models.

More accurately, the plane model of disparity is only suitable for small local areas. If the surface is curve, our local plane fitting is piecewise fitting and is more accurate than segmentation-based algorithms [12]. The overlaps between adjacent CSWs keep good disparity smoothness as the role of soft-segmentation [5].

3.3 Confidence map Updating

The disparity updating reassign new values for pixels. The confidence should be adjusted accordingly. Pixels with ground truth are more likely to keep stable while others change because of the matching ambiguity. The updating of confidence map determines the convergence and the corresponding speed. There are two incremental terms: \( \Delta C_1 \) from the previous disparity value and \( \Delta C_2 \) from the consistency between left and right frames:

\[
C_t(p) = C_{t-1}(p) + \Delta C_1(p) + \Delta C_2(p) 
\]  

If the disparity value of the same pixel doesn’t change in the iterations, the pixel’s disparity will become more reliable. Otherwise the confidence should be reduced by a constant as a penalization. It is shown as the following equation.

\[
\Delta C_1(p) = \begin{cases} 
-c_f & \text{if } |d_t(p) - d_{t-1}(p)| > \Delta_d \\
\lambda c_f & \text{otherwise}
\end{cases} 
\]  

where \( d_t(p) \) and \( d_{t-1}(p) \) are the evaluated disparity values of iteration \( t \) and iteration \( t-1 \), respectively. \( \Delta_d \) is disparity distance tolerant. \( c_f \) is the update step and \( \lambda \) is a constant factor. The average of the confidence map should be increased to make the iteration convergent. In our implementation, \( \lambda \) can be selected in the range of 2 ~ 5. It speeds up the rate of the convergence and prevents the ambiguity in unreliable regions. On the other hand, the left right check is useful to
eliminate the ambiguity. If the corresponding pixels don’t satisfy the consistence with the disparity distance less than \( \Delta_d \), the pixel’s confidence is decreased by \( c_f \). Otherwise, the confidence will be increased by \( c_f \) if the pixel is still unreliable (less than \( T_{cf} \)).

\[
\Delta C_2(p) = \begin{cases} 
-c_f & \text{if } |D_T^f(p) - D_R^b(p)| > \Delta_d \\
c_f & \text{else if } C(p) < T_{cf} \\
0 & \text{otherwise}
\end{cases} \tag{6}
\]

Finally the confidence is limited in a range from 0 to 100. The updated confidence map is used in the CSW selection in the next disparity map updating. Although the initialization of disparity map is pixel accuracy, the local plane model is a continuous model. In the left right check and confidence map updating process, the tolerant of the disparity value \( \Delta_d \) is set to 0.75 \( \times \) scale, where scale is a factor for encoded disparity level. It is compatible both for pixel and sub-pixel accuracy.

3.4 Refinement

The left right check is widely used in the refinement [1]. We choose the minimum of the two disparity values of the corresponding pixels. The traditional median filter can deal with the unsmooth region in disparity map, but it does not keep the right value on the sharp edge. The weighted median filter is inspired by the work of Yoon [2]. The weight for each pixel is according to the color similarity. Then the first disparity whose sum of weights before that index is larger than 0.5 is the new value. The performance on the discontinuity is much better since the color similarity is considered.

4. Experiments

The proposed algorithm has been implemented using C language under Windows OS. As for evaluation protocol, our algorithm is evaluated on the a popular standard evaluation platform presented by Scharstein [1]. Parameters in CSW module are set to the values as mentioned in [4]. Other parameters mentioned above are set as follows. \( T_{cf} = 20 \), \( c_f = 5 \). All test images use the same parameters.

The convergence rate is a crucial factor concerned in our algorithm. Figure 2 show the updating results of proposed method on cones image. The green and blue curves represent the error rates corresponding to different error threshold (1.0 and 0.5, respectively). The red curve depicts the total time cost along with iterations. The updating disparity maps show that the number of unreliable pixels reduces quickly along with iterations. The error rate converges to a stable status in less than 10 iterations (the yellow dash line). It proves that our algorithm achieves the good convergence in both pixel and sub-pixel accuracy. The time cost is linear with the number of unreliable pixels. As the error converges, our algorithm spends about 30 seconds on Cones image as shown in Fig. 2 and about 10 s, 12 s and 40 s on Tsukuba, Venus and Teddy images respectively.

On the Middlebury benchmark, our algorithm ranks top 2 corresponding with both error threshold 1.0 and 0.5 among all algorithms. The error rate and rank in all pixels under threshold 1.0 are presented in Table 1. The average percent of bad pixels is 4.25% and the performance on each image is competitive. Especially on Tsukuba and Venus images, our algorithm achieves the lowest error rate among all algorithms.

Figure 3 shows the comparison results between this paper and [4]. In red rectangles, inaccurate pixels in [4] are rectifies by the interleaving updating process. Our algorithm presents better results on discontinuous edge and corner in Tsukuba and Venus images. The ambiguity of the triangular occluded region in Venus image is eliminated by the weighted median filter. The slant regions on the left border of Teddy and Cones images are difficult to find correct matches by local information [4]. With our message propagation scheme, the disparity values in these regions are correctly estimated.

5. Conclusion and Future Work

In this paper, a novel interleaving updating framework of disparity and confidence map (IUFDCM) is proposed, combining weighted support window and message propagation.
In order to alleviate the effect of input parameters, the disparity map is updated adaptively in CSW. Furthermore, the updating of confidence map is designed to evaluate the reliability of disparity reassignment. Experiments show that our algorithm is robust corresponding to different error thresholds and competitive to the state of art best algorithms. However, the color based initialization is not suitable for gray images. In the future work, we will improve the initialization and accelerate the iteration process.

References

[1] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” IJCV, vol.47, no.1, pp.7–42, 2002.
[2] K.-J. Yoon and I.S. Kweon, “Stereo matching with the distinctive similarity measure,” Proc. ICCV, pp.1–7, 2007.
[3] A. Hosni and M. Bleyer et al., “Local stereo matching using geodesic support weights,” Proc. ICIP, pp.2093–2096, 2009.
[4] C. Shi and G. Wang, X. Pei, B. He, and X. Lin, “Stereo matching using local plane fitting in confidence-based support window,” IEICE Trans. Inf. & Syst., vol.E95-D, no.2, pp.699–702, Feb. 2012.
[5] M. Bleyer, C. Rother, and P. Kohli, “Surface stereo with soft segmentation,” CVPR, pp.1570–1577, 2010.
[6] A. Klaus, M. Sormann, and K. Karner, “Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure,” ICPR 2006.
[7] M. Bleyer, M. Gelautz, and C. Rother et al., “A stereo approach that handles the matting problem via image warping,” CVPR, pp.501–508, 2009.
[8] http://vision.middlebury.edu/stereo/
[9] Q. Yang and L. Wang et al., “Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling,” IEEE Trans. Pattern Anal. Mach. Intell., vol.31, no.3, pp.492–504, March 2009.
[10] X. Mei, X. Sun, and M. Zhou et al., “On building an accurate stereo matching system on graphics hardware,” GPUCV 2011.
[11] Z. Wang and Z. Zheng, “A region based stereo matching algorithm using cooperative optimization,” CVPR, pp.1–8, 2008.
[12] M. Humenberger, T. Engelke, and W. Kabinger, “A census-based stereo vision algorithm using modified Semi-Global Matching and plane fitting to improve matching quality,” CVPRW, pp.77–84, 2010.
[13] X. Sun and X. Mei et al., “Stereo matching with reliable disparity propagation,” 3DIM 2011.
[14] L. Xu and J. Jia, “Stereo matching: An outlier confidence approach,” ECCV 2008.
[15] Q. Yang and R. Yang et al., “Spatial-depth super resolution for range images,” CVPR, pp.1–8, 2007.