# A Novel Trilateral Filter based Adaptive Support Weight Method for Stereo Matching

Dongming Chen
http://liris.cnrs.fr/membres?idn=dchen
Mohsen Ardabilian
http://liris.cnrs.fr/membres?idn=mardabil
Liming Chen
http://perso.ec-lyon.fr/liming.chen/index.htm

LIRIS UMR 5205
Ecole Centrale de Lyon
Lyon, France

The performance of local stereo matching algorithm highly depends on the support window selected, in which the cost are aggregated. A variety of cost aggregation approaches (proposed before 2008) were comprehensively analyzed in [7] and these approaches attempt to seek an optimal support window for each pixel by changing the window size, shape and center offset. The ideal optimal window should satisfy the rule that all pixels in this window lie on the same disparity with the center pixel.

Recent years have witnessed a great deal of attention focused on the Adaptive Support Weight (ASW) based methods [1, 4, 6, 8], proposed firstly by Yoon and Kweon in [8]. The ASW methods assign an adaptive weight to each pixel of the support window, depending on how it is likely to lie on the same disparity with the center pixel. Essentially, the assignment of an adaptive weight amounts to changing the support window in terms of size, shape or center offset. In ASW methods, the weight function is very important, because it directly decides the support window. The weight function proposed in [8] is based on bilateral filter. Following this pioneering work, various weight function were proposed, including in particular the segmented bilateral filter weight function [6], the geodesic weight function [1], the guided filter weight function [4]. Thus, which weight function is the most accurate one? Recently, Hosni *et al.* [2] carried out a comprehensive comparative study to fairly evaluate various weight functions while fixing the preprocessing, matching cost function and post-processing. Their conclusion is that both bilateral filter weight function [8] and guided filter weight function [4] are the best, since bilateral filter weight function performs better on the average rank while guided filter weight function produces a lower average error.

We revisit the bilateral filter weight function [8], which obeys two rules that, given a support pixel, (1) if its color is similar to the center pixel's and (2) if it is spatially close to the center pixel, it is likely to lie on the same disparity with the center pixel. Therefore, the bilateral filter weight function consists of two parts, color similarity term and spatial proximity term, defined as:

$$w_{bf}(p,q) = e^{-\frac{\Delta c_{pq}}{\gamma_c}} e^{-\frac{\Delta g_{pq}}{\gamma_g}}, \tag{1}$$

where $q$ is a pixel within the support window centered at pixel $p$. The color similarity $\Delta c_{pq}$ represents the Euclidean distance between the color of these two pixels, measured in the CIELab color space as

$$\Delta c_{pq} = \sqrt{\sum_{j\in(L,a,b)} (I_j(p) - I_j(q))^2} \tag{2}$$

and the geometric proximity $\Delta g_{pq}$ is the Euclidean distance between their coordinates $(x,y)$ as

$$\Delta g_{pq} = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}. \tag{3}$$

The parameter $\gamma_c$ and $\gamma_g$ are set by user to adjust the color similarity term and geometric proximity term respectively.

While these two rules can handle most depth ambiguities within a support window, it unfortunately fail to resolve the ambiguity induced by nearby pixels at different disparities but with similar colors as illustrated in Figure 1. (a) is the reference image and we focus on the regions within the red box, zoomed in (b). In (b), there are two nearby planks with similar colors but in different disparities. An imaginary situation is presented in (c): these two planks are substituted by one cross-shaped plank in the same disparity. The bilateral filter weight of pixel $p$ and $q$ in (b) is equal to that in (c), because their color similarity and spatial distance are the same. But obviously these two weights should not be equal because the two pixels are in the same disparity in (c) but not in (b). We can observe

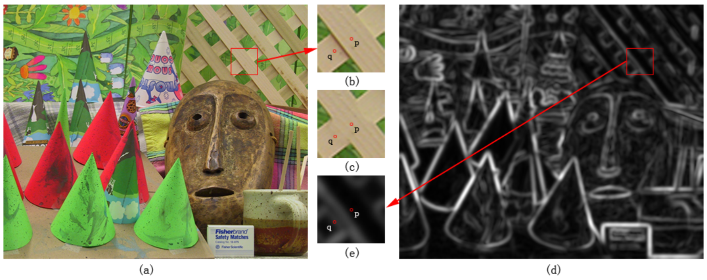


Figure 1: (a) reference image. (b) zoomed real situation. (c) zoomed imaginary situation. (d) boundary cue (e) zoomed boundary cue.

that the depth discontinuity between pixel $p$ and $q$ in (b) induces a color discontinuity between these two planks and results in color boundaries nearby, as shown in (e). Therefore, the boundary cue is helpful to obtain a more faithful weight. As a result, we propose in this paper a trilateral filter weight function which extends the bilateral filter weight function by a new boundary strength term that measures the boundary strength between two pixels, defined as

$$w_{tf}(p,q) = e^{-\frac{\Delta c_{pq}}{\gamma_c}} e^{-\frac{\Delta g_{pq}}{\gamma_g}} \left( e^{-\frac{\Delta E_{pq}}{\gamma_e}} + e^{-\frac{\Delta c_{pq}}{\gamma_c}} e^{-\frac{\Delta g_{pq}}{\gamma_g}} \right), \tag{4}$$

where $exp\left(-\Delta E_{pq}/\gamma_e\right)$ is the boundary strength term and parameter $\gamma_e$ is set by user. The boundary strength at pixel $p$ is defined as [3],

$$E(p) = \sum_{\theta} \sqrt{(I(p) * F_{\theta,odd})^2 + (I(p) * F_{\theta,even})^2}, \tag{5}$$

where $*$ denotes the convolution operator; the odd-phase filter $F_{\theta,odd}$ and even-phase filter $F_{\theta,even}$ are a pair of quadrature filters in orientation $\theta$.

Finally, we evaluated our method on the Middlebury benchmark [5], using four pairs of standard data sets, and fairly compared the proposed trilateral filter weight function with the guided filter weight function [4] and bilateral filter weight function [8]. The proposed method ranks $15^{th}$ out of 150 submissions and is the current most accurate local stereo matching algorithm on the benchmark.

[1] A. Hosni, M. Bleyer, M. Gelautz, and C. Rhemann. Local stereo matching using geodesic support weights. In *ICIP*, pages 2069–2072, 2009.

[2] A. Hosni, M. Bleyer, and M. Gelautz. Secrets of adaptive support weight techniques for local stereo matching. *CVIU*, 117:620–632, 2013.

[3] M.C. Morronea and R.A. Owens. Feature detection from local energy. *PRL*, 6:303–313, 1987.

[4] 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.

[5] D. Scharstein and R. Szeliski. Middlebury stereo evaluation - version 2. http://vision.middlebury.edu/stereo/.

[6] F. Tombari, S. Mattoccia, and L. Di Stefano. Segmentation-based adaptive support for accurate stereo correspondence. In *PSIVT*, volume 1, pages 427 – 438, 2007.

[7] F. Tombari, S. Mattoccia, L. Di Stefano, and E. Addimanda. Classification and evaluation of cost aggregation methods for stereo correspondence. In *CVPR*, 2008.

[8] K. Yoon and I. Kweon. Adaptive support-weight approach for correspondence search. *PAMI*, 28(4):650–656, 2006.