Accurate and Efficient Cost Aggregation Strategy for Stereo Correspondence Based on Approximated Joint Bilateral Filtering

Stefano Mattoccia, Simone Giardino, and Andrea Gambini

Department of Electronics Computer Science and Systems (DEIS)
Advanced Research Center on Electronic Systems (ARCES)
University of Bologna, Viale Risorgimento 2, 40136 - Bologna, Italy
stefano.mattoccia@unibo.it, simone.giardino@hotmail.it,
gambini00@gmail.com

Abstract. Recent local state-of-the-art stereo algorithms based on variable cost aggregation strategies allow for inferring disparity maps comparable to those yielded by algorithms based on global optimization schemes. Unfortunately, thought these results are excellent, they are obtained at the expense of high computational requirements that are comparable or even higher than those required by global approaches. In this paper, we propose a cost aggregation strategy based on joint bilateral filtering and incremental calculation schemes that allow for efficient and accurate inference of disparity maps. Experimental comparison with state-of-the-art techniques shows the effectiveness of our proposal.

1 Introduction and Related Work

Stereo algorithms aim at inferring disparity maps by processing images of the same scene from two or more cameras.. This topic was exhaustively surveyed in [10,2] and according to [10] most algorithms consist of four steps: *matching cost computation*, *cost aggregation*, *disparity computation* and *disparity refinement*.

Although cost aggregation is mandatory for local approaches in order to increase the signal to noise ratio this methodology is also frequently adopted by global (or semi-global) approaches [20,18,17,6]. An ideal (fronto-parallel) cost aggregation strategy, to deal with depth discontinuities and ambiguous regions (low textured areas, repetitive patterns, etc), should modify its support at each position according to image content to include only those points with the same (unknown) disparity. Although this behavior is far from ideal, state-of-the-art cost aggregation strategies, recently surveyed and evaluated [14,4], deploying the simple local Winner Takes All (WTA) disparity selection approach allow for obtaining disparity maps comparable to those yielded by algorithms based on schemes that use more complex reasoning. Unfortunately, the execution time is often comparable. Moreover, the results of state-of-the-art cost aggregation strategies are severely affected by image noise. In this paper we propose a cost aggregation strategy that combines the effectiveness of recent local approaches based on adaptive weights with the efficiency and the robustness of conventional local approaches. Although we assume binocular rectified stereo pairs, the proposed method can be extended in a straightforward manner to three or more rectified cameras.

Bilateral filtering is a non-iterative feature-preserving image smoothing technique [11] that due to its relevance in computer vision, computer graphics and image processing has recently gained a lot of attention [8,7,19,1]. The idea behind bilateral filtering is to jointly and independently enforce a geometric (*spatial filter*) and a color proximity constraint (*range filter*). Given an image I, the value assigned to each point p of the filtered image $\hat{I}$ is a weighted convolution with points q_i in its neighbor $S(p) \subset I$ (a square *support* region centered in p) according to weighting functions related to the spatial distance D_s between p and q_i and a distance D_c in the color space between $I(p)$ and $I(q_i)$.

$$\hat{I}(p) = \frac{\displaystyle\sum_{q_i \in S(p)} W_S(p, q_i) \cdot W_C(I(p), I(q_i)) \cdot I(q_i)}{\displaystyle\sum_{q_i \in S(p)} W_S(p, q_i) \cdot W_C(I(p), I(q_i))} \tag{1}$$

The denominator acts as normalization factor and the two weighting functions W_s and W_c, respectively, assign higher values to points *closer* to the central point p and to points with color intensity *similar* to $I(p)$. Typically, the weights W_s and W_c are assigned according to Gaussian functions, respectively, with variance γ_s and γ_c. The distance between the coordinate points and between triplets in the color space are often computed according to the ℓ_2 norm. Although bilateral filtering has proven to be a very effective technique it is computationally demanding. For this reason, recently, several approximated techniques aimed at reducing its running time have been proposed [8,7]. According to a recent study [14] specifically focused on evaluating and benchmarking state-of-the-art cost aggregation strategies for stereo correspondence, algorithms belonging to the *adaptive weights* [19,12] category dramatically outperformed other approaches in terms of accuracy. In the Adaptive Weight (AW) approach [19] the weight assigned to each point within the support is obtained by applying two independent bilateral filters in the neighborhood of each potential correspondence. Given a point p_r in the reference image I_r and a potential correspondence point p_t in the target image I_t: the weights assigned to each point of the support $S(p_r, p_t)$ are computed by combining (multiplying) the weights that would be yielded by the two independent bilateral filters (with the same parameters γ_s and γ_c) applied to p_r and p_t. The cost of the correspondence $C(p_r, p_t)$ between p_r and p_t is the weighted sum of the TAD (Truncated Absolute Differences) scores within the support normalized by the weights (1).

$$C(p_r, p_t) = \frac{\displaystyle\sum_{\substack{q_{r_i} \in S(p_r) \\ q_{r_i} \in S(p_t)}} W_S(p_r, q_{r_i}) \cdot W_C(I_r(p_r), I_r(q_{r_i})) \cdot W_S(p_t, q_{t_i}) \cdot W_C(I_t(p_t), I_t(q_{t_i})) \cdot TAD(q_{r_i}, q_{t_i})}{\displaystyle\sum_{\substack{q_{r_i} \in S(p_r) \\ q_{r_i} \in S(p_t)}} W_S(p_r, q_{r_i}) \cdot W_C(I_r(p_r), I_r(q_{r_i})) \cdot W_S(p_t, q_{t_i}) \cdot W_C(I_t(p_t), I_t(q_{t_i}))} \tag{2}$$

The weighting function is Gaussian and authors use the ℓ_2 norm in the CIELAB color space for the range filters. This cost aggregation strategy provides excellent results within a WTA framework and has also been successfully adopted within global optimization frameworks [18].

However, it has been shown [12] that under certain circumstances (i.e. along depth discontinuities, low-textured regions, high-textured regions and repetitive patterns) the spatial filter embodied in [19] can lead to wrong correspondences and for these reasons a further segmentation-based constraint was introduced. Therefore, the Segment Support (SS) [12] approach when it computes the weights associated to the two distinct bilateral filters assigns weight 1 to those points belonging to the same segment of the central point and assigns the spatial weight of AW to those points outside the segment containing the central point. This method improves AW, but at the expense of almost doubling the execution time.

It is worth observing that supports for AW and SS are computed by means of a symmetric strategy that relies on both images. Unfortunately AW and SS are computationally very demanding and their execution time [14] are comparable or even worse than those required by global approaches (e.g. on Teddy, AW requires more than 18 minutes while SS requires more than 33 minutes). To reduce the computational complexity of the AW approach a simplified asymmetrical weight assignment strategy was proposed [4]. The weights were computed asymmetrically, according to the reference image only, and approximated by means of a two pass approach (the first pass along the horizontal scanline and the second pass along the perpendicular direction). These simplifications allowed for real-time GPU implementation that yields worse but reasonably accurate [4] disparity maps compared to AW. Another interesting approach based on asymmetrical weight assignment, referred to as SB, was also proposed [3]. This method appears to be a good trade-off between accuracy and computational efficiency deploying segmentation and assigning weights according to the reference image only. Finally, according to the evaluation provided in [14], among effective cost aggregation strategies Variable Windows (VW) [15] deserves particular attention. In fact, although VW is significantly less accurate than AW and SS its execution time is significantly reduced (it takes 26 seconds on Teddy). The efficient aggregation strategy deployed by VW is completely different from the adaptive weight approaches described so far since in VW the weights are always set to 1 while the size of the square support is selected according to three criteria. The best support is selected evaluating for each predefined square region the cost function and its variance. Moreover, to deal with low-textured regions, a biasing term is used to favor large windows. Massive deployment of the *integral image* technique [16] allows notable computational efficiency.

2   Proposed Cost Aggregation Strategy

As already pointed out by the authors [19], the most ambiguous correspondences are set by AW when the support becomes too small. Experimental results show that this behavior mainly occurs in two circumstances: a) when the supports are within regions that are highly textured b) when the support contains uniform regions with pixel intensity similar to the central pixel (not necessarily completely uniform regions). Although case a) seems intrinsically related to the method since it is likely that in highly textured regions several pixels will have different intensity when compared to the central pixel, in both cases pixels with intensity similar to the central pixel should provide the *cue* for setting unambiguous correspondences. In the two cases depicted in Figure 1, an ideal
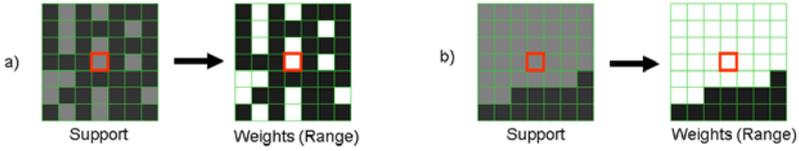
Fig. 1. Case a) and b): weights assigned by an ideal range filter

range filter[1] would assign higher weights (in white) to points with intensity similar to the central one and lower weights to the pixels with different intensity.

When higher weights are assigned to the pixels with intensity similar to the central point (combined with the weights assigned based on the same rationale by the other range filter applied in the other image), the spatial filters and the matching cost would then allow for potentially setting unambiguous correspondences. Hence, points with intensity similar to the central one play a major role in the AW approach. Analyzing the behavior of AW we observed that the origin of the discrepancy with the ideal case could be ascribed to noise and how weights are computed by means of the two independent range filters. In fact, when computing the weight for two pixels with similar color intensity (i.e. when the ℓ_2 norm $\parallel \Delta \parallel \to 0$) the exponential function embodied in the range filter becomes very sensitive to image noise since derivative is very high. Under these circumstances (i.e. when the ℓ_2 norm $\parallel \Delta \parallel \to 0$) image noise severely affects the weights assigned by means of the range filters.
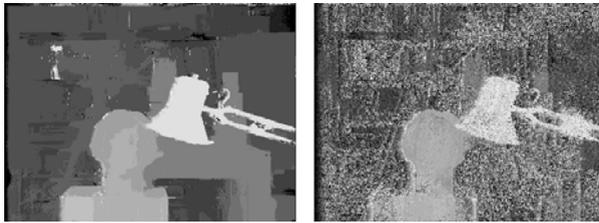

Fig. 2. Disparity maps yielded by AW: (Left) original Tsukuba stereo pair (Right) Tsukuba stereo pair corrupted by Gaussian noise ($\mu = 0$, $\sigma = 1.275$ pixels) independently applied to the three color channels

In Figure 2 we report the disparity maps yielded by AW on the original Tsukuba stereo pair and on the same stereo pair corrupted by Gaussian noise. The figure seems to confirm our analysis; the effects of noise are particularly prominent in the regions with similar color intensity. To deal with this problem the proposed cost aggregation strategy, referred to as Fast Bilateral Stereo (FBS), embodies a simple but effective noise regularization stage for the range filter. Moreover, as the major aim of this research activity was the development of an effective cost aggregation strategy that could

---

[1] For simplicity we consider greyscale images and a single range filter applied to the support of a single image.

fit within an efficient computational framework we combined the efficiency of the traditional and local approaches for stereo correspondences [14] with a symmetric adaptive weights strategy based on two independent spatial and range filters applied on a regular block basis. Given two points $p_r \in I_r$ and $p_t \in I_t$, for which a correspondence had to be evaluated, and the associated supports $S(p_r) \subset I_r$ and $S(p_t) \subset I_t$ both of size $W \times W$ we partitioned the two supports $S(p_r)$ and $S(p_t)$ in $\frac{W}{w} \times \frac{W}{w}$ regular blocks of size $w \times w$ as shown in Figure 3[2]. At each block of the two supports S_r, S_t we independently assigned two weights according to a spatial filter and a range filter.
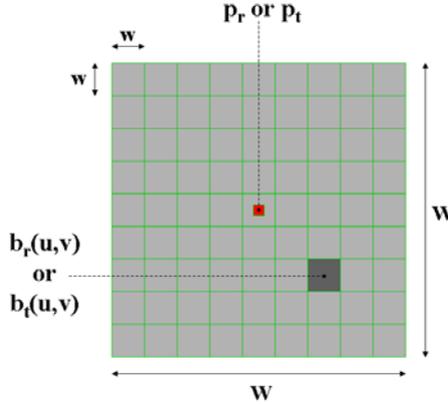

Fig. 3. Proposed partitioning scheme of the two supports $S_r \subset I_r$ and $S_t \subset I_t$

For what concerns the spatial filter, at each point within $w \times w$ block we assigned the spatial weight of the central point of the block according to the W_s function in (1). For what concerns the range filter we assigned at each $w \times w$ block a single weight according to a simple but effective strategy aimed at increasing noise robustness. Assuming that pixels within each block are spatially correlated, the average values of the intensities within the block provide a means to decrease the variance by a factor w. Clearly, when this hypothesis is not verified the averaged value induces a bias that might lead to a non perfect localization of discontinuities. However, with small w the hypothesis that pixel within the $w \times w$ block are correlated often holds and averaging is a simple, yet effective, strategy to reduce image noise. To partially deal with this issue we assigned at each block $b_r(u, v)$ (with $u \in \frac{W}{w}, v \in \frac{W}{w}$) within the support $S(p_r)$ of the reference image the weight

$$W_C\left(I_r\left(p_r\right), \overline{I}_r\left(b_r(u,v)\right)\right) = \exp(-\frac{\left\|I_r\left(p_r\right) - \overline{I}_r(b_r(u,v))\right\|}{\gamma_c}) \qquad (2)$$

with $\overline{I}(b_r(u,v))$ representing the average value of pixels within block $b_r(u,v)$. Using small w allows for reducing the variance within each block as well as the maintenance of an accurate localization with respect to the central point p_r. Experimentally we found

---

[2] Although not strictly necessary, for simplicity, $W \mod w$ is assumed to be 0.

that $w = 3$ provides optimal results on the Middlebury dataset. Obviously, the same strategy is applied to point p_t of the target image I_t.

Once we obtained the block-based weights by means of the two spatial filters and range filters, we combined the weights to obtain a symmetric block-based weighted support and computed the matching cost using a TAD cost function on a pixel basis. That is, at each block a single weight is assigned, but each point within the block is evaluated according to the pixel-wise matching cost. Similarly to (1), the matching cost is normalized by the weights. It is worth observing that cost computation and block averaging can be efficiently computed by means of incremental calculation schemes [16] typically deployed by conventional local approaches. Moreover, compared to (1), the number of range and spatial filters is intrinsically reduced by factor $w \times w$. To further reduce the computational requirements spatial and range weights are stored in look-up tables. Finally, it is worth observing, that for $w = 1$ our computational framework is equivalent to [19].

3   Experimental Results

This section aims at assessing the performance of the proposed approach within a framework specifically focused on the evaluation of state-of-the-art cost aggregation strategies [14]. The disparity maps yielded by the considered cost aggregation strategies on the four images of the Middlebury dataset [9] (Tsukuba, Venus, Teddy and Cones) were obtained by means of a simple WTA strategy without any post processing filtering and without enforcing the left-right consistency constraint. Nevertheless, although the focus here is on the evaluation of the raw cost aggregation strategy and all the considered approaches do not deal explicitly with occlusions, for completeness, we have also report the ALL parameter so as to allow a direct comparison with other approaches on the Middlebury evaluation site. We have reported in Table 1 the results obtained by FBS (parameters $W = 39$, $w = 3, \gamma_s = 14$, $\gamma_c = 23$ and TAD threshold 53), by our implementation of the AW approach (referred to as AW* with optimal parameters $W = 35$, $\gamma_s = 31$, $\gamma_c = 13$ and TAD threshold 40) and by the five top performing state-of-the-art cost aggregation strategies [12,3,19,5,15] according to [14][3].

It is worth noting that these results differ from those published in [14] (concerned with SAD cost function) because, for fairness, we deployed the original cost functions originally proposed by the authors of each paper (see [13] section "*Original*"). Here we stress the fact that we were interested in evaluating the performance of the raw cost aggregation strategies, and the results for AW and SS available on the Middlebury evaluation site include post processing steps that are not specified. For all approaches the execution time is concerned with the Teddy stereo pair and for the proposed approach also includes initialization of look-up tables. As FBS has two parameters for the support (W and w) we used $W = 39$, similar to those deployed for the adaptive weight

---

[3] For [12,3,19,5,15] the optimal parameters found in [14] were deployed, which are available at `www.vision.deis.unibo.it/spe/data/parameters.pdf`. The execution time for SB was obtained by deploying a much faster segmentation approach compared to the results reported in [14]. For FBS we found the optimal parameters minimizing the NOCC+DISC error on the whole dataset.

Table 1. Accuracy according to the Middlebury web site [9] and (in boldface) according to [14]. The table reports the accuracy of the proposed FBS approach and the five top performing [14] state-of-the-art approaches [12,3,19,5,15]. The disparity maps tagged with symbol [†] are available in [13] - section '*Original*'. Table also reports the execution time (Intel Core Duo 2.14 GHz processor) concerned with the Teddy stereo pair.

	Tsukuba			Venus			Teddy			Cones			Time
	NOCC	ALL	DISC	NOCC	ALL	DISC	NOCC	ALL	DISC	NOCC	ALL	DISC	sec
FBS$_{39(3)}$	**2.95**	4.75	**8.69**	**1.29**	2.87	**7.62**	**10.71**	19.8	**20.82**	**5.23**	15.3	**11.34**	32
SS[†] [12]	**2.15**	4.04	**7.22**	**1.38**	3.0	**6.27**	**10.5**	19.7	**21.2**	**5.83**	16.4	**11.8**	2358
SB[†] [3]	**2.25**	2.86	**8.87**	**1.37**	2.31	**9.4**	**12.7**	20.1	**24.8**	**11.1**	19.2	**20.1**	2
AW*[19]	**3.33**	5.25	**8.87**	**2.02**	3.61	**9.32**	**10.52**	19.7	**20.84**	**3.72**	14.3	**9.37**	3226
AW[†] [19]	**4.66**	6.68	**8.25**	**4.61**	6.18	**13.3**	**12.7**	21.6	**22.4**	**5.5**	16.0	**11.9**	1221
Rel[†] [5]	**5.08**	6.94	**17.9**	**3.92**	5.5	**13.9**	**18.9**	27.0	**29.9**	**11.3**	20.7	**18.3**	803
VW[†] [15]	**3.12**	4.86	**12.4**	**2.42**	3.87	**13.3**	**17.7**	25.9	**25.5**	**21.2**	29.6	**27.3**	26

Table 2. Accuracy for FBS and adaptive weights approach (AW*) on stereo pairs corrupted by Gaussian noise ($\mu = 0$, $\sigma = 0.255, 1.275, 2.55, 12.75$ pixels) independently applied to the three color channels of the original stereo pairs.

	Noise	Tsukuba			Venus			Teddy			Cones		
	σ	NOCC	ALL	DISC	NOCC	ALL	DISC	NOCC	ALL	DISC	NOCC	ALL	DISC
FBS$_{39(3)}$	0.255	**7.58**	9.32	**14.83**	**14.13**	15.48	**17.83**	**15.42**	24.00	**26.13**	**6.67**	16.60	**13.48**
AW*	0.255	**21.34**	22.94	**25.65**	**24.61**	25.81	**27.54**	**19.98**	28.16	**30.71**	**7.30**	17.55	**14.68**
FBS$_{39(3)}$	1.275	**34.26**	35.44	**36.81**	**45.15**	45.97	**41.58**	**47.85**	53.09	**54.00**	**31.11**	38.42	**38.34**
AW*	1.275	**52.93**	53.81	**49.89**	**56.08**	56.73	**54.74**	**51.28**	56.23	**60.19**	**35.72**	42.80	**41.38**
FBS$_{39(3)}$	2.55	**53.94**	54.70	**54.52**	**68.13**	68.57	**62.87**	**68.56**	71.62	**69.43**	**56.50**	61.03	**59.57**
AW*	2.55	**66.32**	66.87	**63.07**	**74.76**	75.13	**70.26**	**73.56**	76.19	**77.73**	**64.55**	68.36	**65.21**
FBS$_{39(3)}$	12.75	**71.78**	72.15	**74.82**	**85.69**	85.84	**83.30**	**92.30**	93.00	**91.31**	**91.69**	92.41	**91.20**
AW*	12.75	**78.61**	77.50	**78.81**	**87.95**	88.09	**85.10**	**94.65**	95.14	**94.17**	**94.07**	94.59	**93.01**

approach, and $w = 3$. Therefore, our proposal in Table 1 is referred to as FBS$_{39(3)}$. With respect to the NOCC and DISC errors, we noticed that our proposal has an accuracy comparable to the best performing cost aggregation strategies, SS and AW (both implementations) being close in most cases to the results yielded by the best one (SS). It is also worth pointing out that SS and AW run is minutes on the Teddy stereo pair (e.g. SS is almost 40 minutes, AW is about 20 minutes) while FSB takes only 32 seconds. The table also shows that FBS significantly outperforms the accuracy of the VW and Rel approaches on the whole dataset. For SB, we noticed that, in most cases, it is outperformed by FBS (in particular, on Cones, NOCC and DISC errors are about 50 %). Nevertheless, SB is significantly faster than FSB (it takes only 2 seconds on Teddy).

Figure 4 shows the disparity maps for the considered approaches. From the figure, we noticed that, compared to AW, our approach allows for reducing several errors (region in front of the camera in Tsukuba, slanted surfaces in Venus, roof in Teddy). Depth maps yielded by FBS are, in general, less noisy compared to AW. Although not reported here
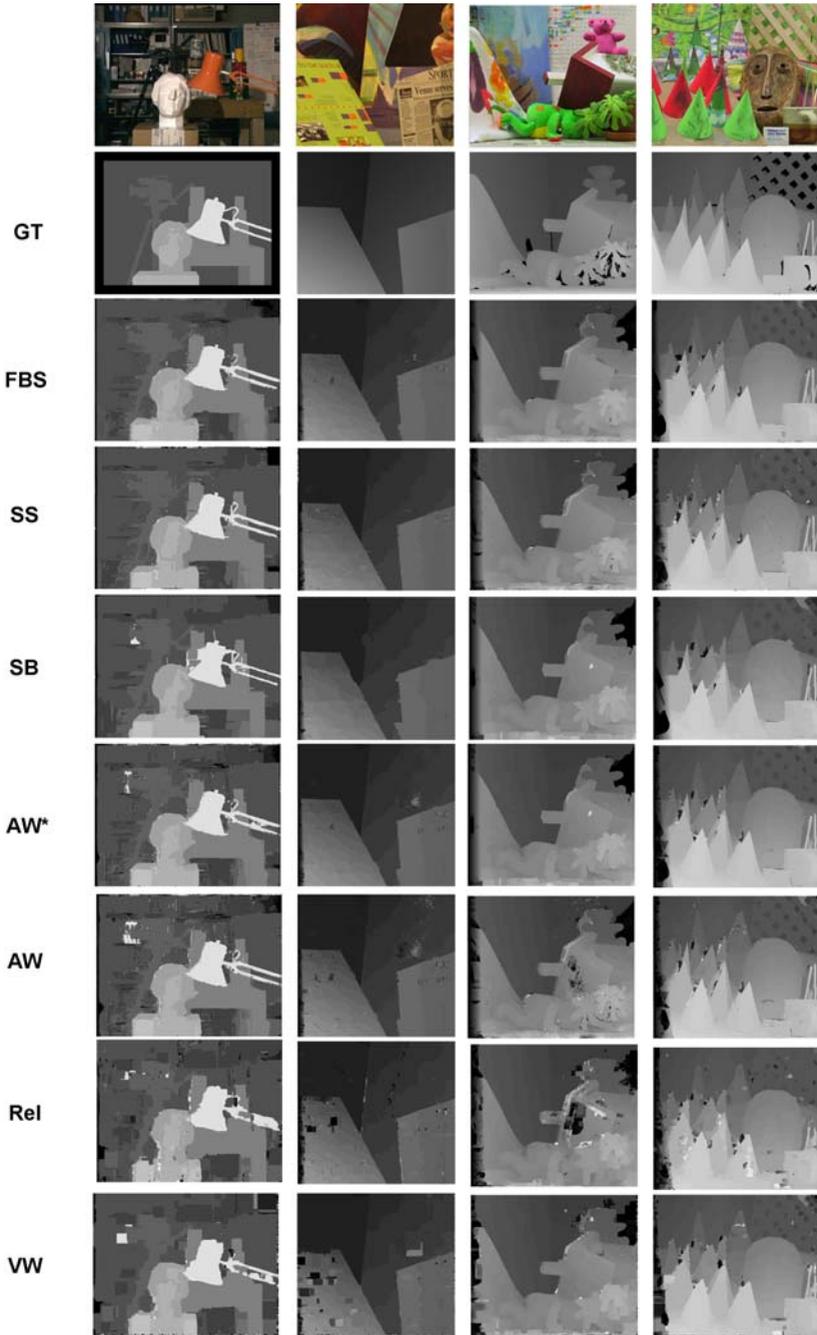
Fig. 4. From top to bottom: reference image, groundtruth (GT), disparity map yielded by the proposed FBS approach, SS [12], SB [3], our implementation of AW [19], original implementation of AW [19], Rel [5] and VW[15].

for the lack of space[4], increasing w decreases accuracy but improves efficiency: in fact ($W = 39$) on Teddy FBS takes 14 sec (4 sec on Tsukuba) with $w = 5$, 9 sec (2 sec on Tsukuba) with $w = 7$ and 5 sec (1 sec on Tsukuba) with $w = 9$. This highlights an interesting behavior of the proposed approach: by modifying w one can trade accuracy for speed and vice versa. This might be interesting in certain applications (e.g. *robot picking*, face detection and recognition) where accurate disparity maps are required only when objects are close to the camera.

To prove the effectiveness of the noise reduction technique embodied in our proposal, we report in Table 2 the results obtained by our proposal and by our implementation of the adaptive weights technique (AW*) on stereo pairs corrupted by Gaussian noise (mean value $\mu = 0$ and variance (in pixels) $\sigma = 0.255, 1.275, 2.55, 12.75$) independently applied to the three color channels of the original stereo pairs. Although both approaches have poor results with higher noise levels, Table 2 reports that for the whole dataset the noise regularization step embodied in the range filter calculation of our proposal is always notably more effective than adaptive weights. For both approaches we deployed the optimal parameters described for Table 1.

4   Conclusions

We have proposed a cost aggregation strategy that combines the efficiency of traditional local algorithms with the accuracy of state-of-the-art approaches. The weight computation strategy proposed deploys a simple, but effective, noise regularization step that allows for improving the accuracy of the original AW approaches and, exploiting efficient incremental calculation schemes, for obtaining a disparity map at a small fraction of the time required by state-of-the-art approaches. Experimental results within a framework specifically aimed at evaluating the performance of state-of-the-art cost aggregation strategies for stereo correspondence confirm the effectiveness of our proposal. Future work is aimed at exploiting the block-based framework proposed to deal with photometric distortions that typically arise in real applications. We are also interested in deploying the cost aggregation strategy proposed within global or semi-global optimization frameworks.

References

1. Ansar, A., Castano, A., Matthies, L.: Enhanced real-time stereo using bilateral filtering. In: 3DPVT 2004, pp. 455–462 (2004)
2. Brown, M.Z., Burschka, D., Hager, G.D.: Advances in computational stereo. IEEE Trans. Pattern Anal. Mach. Intell. 25(8), 993–1008 (2003)
3. Gerrits, M., Bekaert, P.: Local stereo matching with segmentation-based outlier rejection. In: Proc. CRV 2006, p. 66 (2006)
4. Gong, M., Yang, R.G., Liang, W., Gong, M.W.: A performance study on different cost aggregation approaches used in real-time stereo matching. Int. Journal Computer Vision 75(2), 283–296 (2007)

---

[4] Additional experimental results are available at:
`www.vision.deis.unibo.it/smatt/fast_bilateral_stereo.htm`

5. Kang, S.B., Szeliski, R., Chai, J.: Handling occlusions in dense multi-view stereo. In: Proc. CVPR 2001, pp. 103–110 (2001)
6. Mattoccia, S., Tombari, F., Di Stefano, L.: Stereo vision enabling precise border localization within a scanline optimization framework. In: Yagi, Y., Kang, S.B., Kweon, I.S., Zha, H. (eds.) ACCV 2007, Part II. LNCS, vol. 4844, pp. 517–527. Springer, Heidelberg (2007)
7. Paris, S., Durand, F.: A fast approximation of the bilateral filter using a signal processing approach. Int. Journal Computer Vision 81(1), 24–52 (2009)
8. Porikli, F.M.: Constant time o(1) bilateral filtering. In: CVPR 2008, pp. 1–8 (2008)
9. Scharstein, D., Szeliski, R.: http://vision.middlebury.edu/stereo/
10. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. Journal Computer Vision 47(1/2/3), 7–42 (2002)
11. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: ICCV 1998, pp. 839–846 (1998)
12. Tombari, F., Mattoccia, S., Di Stefano, L.: Segmentation-based adaptive support for accurate stereo correspondence. In: Mery, D., Rueda, L. (eds.) PSIVT 2007. LNCS, vol. 4872, pp. 427–438. Springer, Heidelberg (2007)
13. Tombari, F., Mattoccia, S., Di Stefano, L., Addimanda, E.: Classification and evaluation of cost aggregation methods for stereo correspondence, www.vision.deis.unibo.it/spe/SPEHome.aspx
14. Tombari, F., Mattoccia, S., Di Stefano, L., Addimanda, E.: Classification and evaluation of cost aggregation methods for stereo correspondence. In: CVPR 2008, pp. 1–8 (2008)
15. Veksler, O.: Fast variable window for stereo correspondence using integral images. In: Proc. CVPR 2003, pp. 556–561 (2003)
16. Viola, P., Jones, M.J.: Robust real-time face detection. Int. Journal of Computer Vision 57(2), 137–154 (2004)
17. Wang, L., Liao, M., Gong, M., Yang, R., Nister, D.: High-quality real-time stereo using adaptive cost aggregation and dynamic programming. In: Proc. 3DPVT 2006, pp. 798–805 (2006)
18. Yang, Q., Wang, L., Yang, R., Stewénius, H., Nistér, D.: Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. IEEE Trans. on Pattern Analysis and Machine Intelligence 31(3), 492–504 (2009)
19. Yoon, K.J., Kweon, I.S.: Adaptive support-weight approach for correspondence search. IEEE Trans. PAMI 28(4), 650–656 (2006)
20. Yoon, K.J., Kweon, I.S.: Stereo matching with symmetric cost functions. In: Proc. CVPR 2006, vol. 2, pp. 2371–2377 (2006)