Stereo Vision Enabling Precise Border Localization Within a Scanline Optimization Framework

Stefano Mattoccia[1,2], Federico Tombari[1,2], and Luigi Di Stefano[1,2]

[1] Department of Electronics Computer Science and Systems (DEIS)
University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy
[2] Advanced Research Center on Electronic Systems 'Ercole De Castro' (ARCES)
University of Bologna, Via Toffano 2/2, 40135 Bologna, Italy
{smattoccia, ftombari, ldistefano}@deis.unibo.it

Abstract. A novel algorithm for obtaining accurate dense disparity measurements and precise border localization from stereo pairs is proposed. The algorithm embodies a very effective variable support approach based on segmentation within a Scanline Optimization framework. The use of a variable support allows for precisely retrieving depth discontinuities while smooth surfaces are well recovered thanks to the minimization of a global function along multiple scanlines. Border localization is further enhanced by symmetrically enforcing the geometry of the scene along depth discontinuities. Experimental results show a significant accuracy improvement with respect to comparable stereo matching approaches.

1 Introduction and Previous Work

In the last decades stereo vision has been one of the most studied task of computer vision and many proposals have been made in literature on this topic (see [1] for a review). The problem of stereo correspondence can be formulated as follows: given a pair of rectified stereo images, with one being the *reference* image I_r and the other being the *target* image I_t, we need to find for each point $p_r \in I_r$ its *correspondence* $p_t \in I_t$ which, due to the epipolar constraint, lies on the same scanline as p_r and within the disparity range $D = [d_{min}; d_{max}]$.

The taxonomy proposed by Scharstein and Szelinski [1] for dense stereo techniques subdivides stereo approaches into two categories: local and global. Local approaches determine the stereo correspondence for a point p_r by selecting the candidate $p_{t,d}, d \in D$ which minimizes a matching cost function $C_M(p_r, p_{t,d})$. In order to decrease the ambiguity of the scores the matching cost is not pointwise but is typically computed over a *support* which includes p_r on I_r and $p_{t,d}$ on I_t. While the support can be in the simplest cases a static squared window, notable results have been yielded by using a variable support which dynamically adapts itself depending on the surroundings of p_r and $p_{t,d}$ [2], [3], [4], [5], [6], [7], [8]. Conversely, most global methods attempt to minimize an energy function computed on the whole image area by employing a Markov Random Field model.

Y. Yagi et al. (Eds.): ACCV 2007, Part II, LNCS 4844, pp. 517–527, 2007.

Since this task turns out to be a NP-hard problem, approximate but efficient strategies such as Graph Cuts (GC) [9] and Belief Propagation (BP) [10], [11] have been proposed. In particular, a very effective approach turned out to be the employment of segmentation information and a plane fitting model within a BP-based framework [12], [13], [14].

A third category of methods which lies in between local and global approaches refers to those techniques based on the minimization of an energy function computed over a subset of the whole image area, i.e. typically along epipolar lines or *scanlines*. The adopted minimization strategy is usually based on *Dynamic Programming* (DP) or *Scanline Optimization* (SO) [15], [16], [17], [18] techniques, and some algorithms also exploit DP on a tree [19], [20]. The global energy function to be minimized includes a pointwise matching cost C_M (see [1] for details) and a smoothness term which enforces constant disparity e.g. on untextured regions by means of a discontinuity penalty π:

$$E\left(d\left(A\right)\right) = \sum_{i \in A} C_M\left(p_r^i, p_{t,d(A)}^i\right) + N\left(d\left(A\right)\right) \cdot \pi \tag{1}$$

with A being the image subset (e.g. a scanline) and N being the number of times the smoothness constraint is violated within the region where the cost function has to be minimized. These approaches achieved excellent results in terms of accuracy in the disparity maps [15] and in terms of very fast, near real-time, computational performances [17].

In order to increase robustness against outliers a fixed support (typically a 3×3 window) can be employed instead of the pointwise matching score. Nevertheless, this approach embodies all the negative aspects of a local window-based method, which are especially evident near depth discontinuities: object borders tend to be inaccurately detected.

Hence, a first contribution proposed by this paper is to deploy an SO-based algorithm which embodies, as matching cost C_M, a function based on a variable support. The SO framework allows to handle effectively low-textured surfaces while the variable support approach helps preserving accuracy along depth borders. In order to determine the variable support, we adopt a very effective technique based on colour proximity and segmentation [21] recently proposed for local approaches. The accuracy of the SO-based process is also improved by the use of a symmetrical smoothness penalty which depends on the pixel intensities of both stereo images. It will be shown that this approach allows to obtain notable accuracy in the retrieved disparities.

Moreover, we propose a refinement step which allows to further increase the accuracy of the proposed method. This step relies on a technique that, exploiting symmetrically the relationship between occlusions and depth discontinuities on the disparity maps obtained assuming alternatively as reference the left and the right image, allows for accurately locating borders. This is shown to be particularly useful to assign the correct disparity values to those points violating the cross-checking constraint. Finally, experimental results show that the proposed approach is able to determine accurate dense stereo maps and it is state-of-the-art for what means approaches which do not rely on a global framework.

2 The Support Aggregation Stage

The first step of the proposed technique computes matching costs based on a variable support strategy proposed in [21] for local algorithms. In particular, given the task of finding the correspondence score between points $p_r \in I_r$ and $p_{t,d} \in I_t$, during the support aggregation step each point of I_r is assigned a weight which depends on color proximity from p_r as well as on information derived from a segmentation process applied on the colour images. In particular, weight $w_r(p_i, p_r)$ for point p_i belonging to I_r is defined as:

$$w_r(p_i, p_r) = \begin{cases} 1.0 & p_i \in S_r \\ exp\left(-\frac{d_c(I_r(p_i), I_r(p_r))}{\gamma_c}\right) & otherwise \end{cases} \tag{2}$$

with S_r being the segment on which p_r lies, d_c the Euclidean distance between two RGB triplets and the constant γ_c a parameter of the algorithm. A null weight is assigned to those points of I_r which lie too far from p_r, i.e. whose distance along x or y direction exceeds a certain radius. A similar approach is adopted to assign a weight $w_t(q_i, p_{t,d})$ to each point $q_i \in I_t$. It is interesting to note that this strategy allows to ideally extract two distinct supports at every new correspondence evaluation, one for the reference image and the other for the target image.

Once the weights are computed, the matching cost for correspondence $(p_r, p_{t,d})$ is determined by summing over the image area the product of such weights with a pointwise matching score (the *Truncated Absolute Difference* (TAD) of RGB triplets) normalised by the weight sum:

$$C_{M,v}(p_r, p_{t,d}) = \frac{\sum_{p_i \in I_r, q_i \in I_t} w_r(p_i, p_r) \cdot w_t(q_i, p_{t,d}) \cdot TAD(p_i, q_i)}{\sum_{p_i \in I_r, q_i \in I_t} w_r(p_i, p_r) \cdot w_t(q_i, p_{t,d})} \tag{3}$$

3 A Symmetric Scanline Optimization Framework

The matching cost $C_{M,v}(p_r, p_{t,d})$ described in the previous section is embodied in a simplified SO-based framework similar to that proposed in [15]. Hence, in the first stage of the algorithm the matching cost matrix $C_{M,v}(p_r, p_{t,d})$ is computed for each possible correspondence $(p_r, p_{d,t})$. Then, in the second stage, 4 SO processes are used: 2 along horizontal scanlines on opposite directions and 2 similarly along vertical scanlines. The *j-th* SO computes the current global cost between p_r and $p_{t,d}$ as:

$$\begin{aligned} C^j_G(p_r, p_{t,d}) = C_{M,v}(p_r, p_{t,d}) + min(C^j_G(p^p_r, p^p_{t,d}), C^j_G(p^p_r, p^p_{t,d-1}) + \pi_1, \\ C^j_G(p^p_r, p^p_{t,d+1}) + \pi_1, c_{min} + \pi_2) - c_{min} \end{aligned} \tag{4}$$

with p^p_r and $p^p_{t,d}$ being respectively the point in the previous position of p_r and $p_{t,d}$ along the considered scanline, π_1 and π_2 being the two smoothness penalty terms (with $\pi_1 \leq \pi_2$) and c_{min} defined as:

$$c_{min} = min_i(C_G^j(p_r^p, p_{t,i}^p)) \tag{5}$$

For what means the two smoothing penalty terms, π_1 and π_2, they are dependent on the image local intensities similarly to what proposed in [22] within a global stereo framework. This is due to the assumption that often a depth discontinuity coincides with an intensity edge, hence the smoothness penalty must be relaxed along edges and enforced within low-textured areas. In particular, we apply a symmetrical strategy so that the two terms depend on the intensities of both I_r and I_t. If we define the intensity difference between the current point and the previous one along the considered scanline on the two images as:

$$\begin{aligned} \bigtriangledown\,(p_r) &= |I_r(p_r) - I_r(p_r^p)| \\ \bigtriangledown(p_{t,d}) &= |I_t(p_{t,d}) - I_t(p_{t,d}^p)| \end{aligned} \tag{6}$$

then π_1 is defined as:

$$\pi_1(p_r, p_{t,d}) = \begin{cases} \Pi_1 & \bigtriangledown(p_r) < P_{th}, \bigtriangledown(p_{t,d}) < P_{th} \\ \Pi_1/2 & \bigtriangledown(p_r) \geq P_{th}, \bigtriangledown(p_{t,d}) < P_{th} \\ \Pi_1/2 & \bigtriangledown(p_r) < P_{th}, \bigtriangledown(p_{t,d}) \geq P_{th} \\ \Pi_1/4 & \bigtriangledown(p_r) \geq P_{th}, \bigtriangledown(p_{t,d}) \geq P_{th} \end{cases} \tag{7}$$

where Π_1 is a constant parameter of the algorithm, and π_2 is defined in the same manner based on Π_2. Finally, P_{th} is a threshold which determines the presence of an intensity edge. Thanks to this approach, horizontal/vertical edges are taken into account along corresponding scanline directions (i.e. horizontal/vertical) during the SO process, so that edges orthogonal to the scanline direction can not influence the smoothness penalty terms.

Once the 4 global costs C_G are obtained, they are summed up together and a *Winner-Take-All* approach on the final cost sum assigns the disparity:

$$d_{p_r,best} = \arg\min_{d \in D}\{\sum_{j=1}^{4} C_G^j(p_r, p_{t,d})\} \tag{8}$$

4 A First Experimental Evaluation of the Proposed Approach

We now briefly show some results dealing with the use of the approach outlined so far. In particular, in order to demonstrate the benefits of the joint use of the SO-based framework with the variable support-based matching cost $C_{M,v}$, we compare the results yielded by our method to those attainable by the same SO framework using the pointwise TAD matching cost on RGB triplets, as well as by $C_{M,v}$ in a local WTA approach.

The dataset used for experiments is available at the Middlebury website[1]. Parameter set is constant for all runs. Truncation parameter for TAD in both

[1] *vision.middlebury.edu/stereo*

Table 1. Error rates using $C_{M,v}$ within the SO-based framework proposed (first row), a pointwise matching cost ($C_{M,p}$) within the same SO-based framework (second row), and $C_{M,v}$ in a local WTA approach (last row)

	Tsukuba	Venus	Teddy	Cones
	N.O. - DISC	N.O. - DISC	N.O. - DISC	N.O. - DISC
$C_{M,v}$, SO	1.63 - 6.80	0.97 - 9.03	9.64 - 19.35	4.60 - 11,52
$C_{M,p}$, SO	3.70 - 13.38	4.19 - 19.27	12.28 - 20.40	5.99 - 13.96
$C_{M,v}$, local	2,05 - 7,14	1,47 - 10,5	10,8 - 21,7	5,08 - 12,5

approaches is set to 80. For what means the variable support, segmentation is obtained by running the Mean Shift algorithm [23] with a constant set of parameters (spatial radius $\sigma_S = 3$, range radius $\sigma_R = 3$, minimum region size $min_R = 35$), while maximum radius size of the support is set to 51, and parameter γ_c is set to 22. Finally, for what means the SO framework, our approach is run with $\Pi_1 = 6, \Pi_2 = 27, P_{th} = 10$, while the pointwise cost-based approach is run with $\Pi_1 = 106, \Pi_2 = 312, P_{th} = 10$ (optimal parameters for both approaches).

Table 1 shows the error rates computed on the whole image area except for occlusion (*N.O.*) and in proximity of discontinuities (*DISC*). Occlusions are not evaluated here since at this stage no specific occlusion handling approach is adopted by any of the algorithms. As it can be inferred, the use of a variable support in the matching cost yields significantly higher accuracy in all cases compared to the pointwise cost-based approach, the highest benefits being on *Tsukuba* and *Venus* datasets. Moreover, benefits are significant also by considering only depth discontinuities, which demonstrate the higher accuracy in retrieving correctly depth borders provided by the use of a variable support within the SO-based framework. Finally, benefits of the use of the proposed SO-based framework are always notable if we compare the results of our approach with those yielded by using the same cost function within a local WTA strategy.

5 Symmetrical Detection of Occluded Areas and Depth Borders

By respectively assuming as reference I_r the left and the right image of the stereo pair, it is possible to obtain two different disparity maps, referred to as D_{LR} and D_{RL}. Our idea is to derive a general method for detecting depth borders and occluded regions by enforcing the symmetrical relationship on both maps between occlusions and depth borders resulting from the stereo setup and the scene structure.

In particular, due to the stereo setup, if we imagine to scan any epipolar line of D_{LR} from left side to right side, each sudden depth decrement corresponds to an occlusion in D_{RL}. Similarly, scanning any epipolar line of D_{RL} from right side to left side, each sudden depth increment corresponds to an occlusion in D_{LR}. Moreover, the occlusion width is directly proportional to the amount of

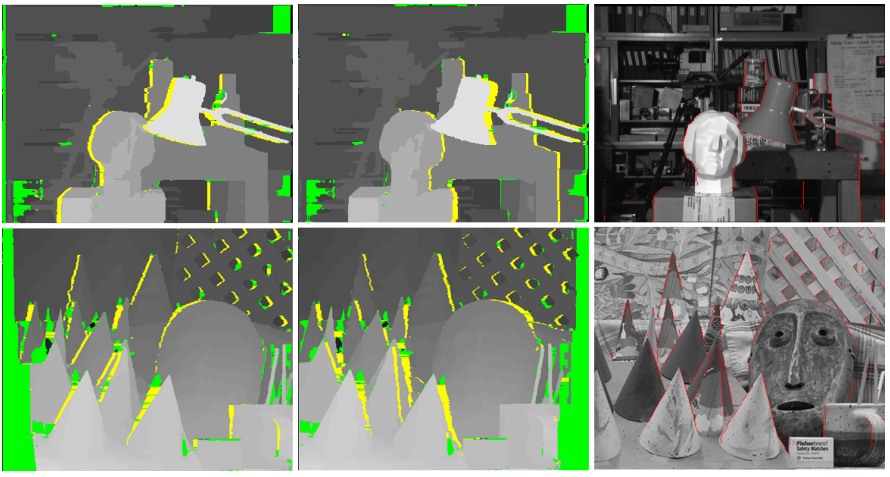

Fig. 1. Points violating (9) on D_{LR} and D_{RL} (colored points, left and center) are discriminated between occlusions (yellow) and false matches (green) on *Tsukuba* and *Cones* datasets. Consequently depth borders are detected (red points, right) [This Figure is best viewed with colors].

each depth decrement and increment along the correspondent epipolar line, and the two points composing a depth border on one disparity map respectively correspond to the starting point and ending point of the occluded area in the other map.

Hence, in order to detect occlusions and depth borders, we deploy a symmetrical cross-checking strategy, which detects the disparities in D_{LR} which violate a *weak* disparity consistence constraint by tagging as *invalid* all points $p_d \in D_{LR}$ for which:

$$|D_{LR}(p_d) - D_{RL}\left(p_d - D_{LR}\left(p_d\right)\right)| \leq 1 \tag{9}$$

and analogously detects invalid disparities on D_{RL}. Points referring to disparity differences equal to 1 are not tagged as invalid at this stage as we assume that occlusions are not present where disparity varies smoothly along the epipolar lines, as well as to handle slight discrepancies due to the different view points. The results of this symmetrical cross-checking are shown, referred to *Tsukuba* and *Cones*, on the left and center images of Fig. 1, where colored points in both maps represent the disparities violating (9). It is easy to infer that only a subset of the colored regions of the maps is represented by occlusions, while all other violating disparities denote mismatches due to outliers.

Hence, after cross-checking the two disparity maps D_{LR} and D_{RL}, it is possible to discriminate on both maps occluded areas from incorrect correspondences (respectively yellow and green points on left and center image, Fig. 1) by means of application of the constraints described previously. Then, putting in correspondence occlusions on one map with homologous depth discontinuities in the

other map, it is possible to reliably localize depth borders generated by occlusions on both disparity maps (details of this method are not provided here due to the lack of space). Right images on Fig. 1 show the superimposition of the detected borders referred to D_{LR} (in red color) on the corresponding grayscale stereo image. As it can be seen, borders along epipolar lines are detected with notable precision and very few outliers (detected borders which do not correspond to real borders) are present.

Fig. 2. The reliability of assigning disparities to points violating the strong cross-checking (10) along depth borders (green points, left) is increased by exploiting information on depth borders location (red points, center) compared to a situation where this information is not available (right)

6 Refinement by Means of Detected Depth Borders and Segmentation

Depth border detection is employed in order to determine the correct disparity values to be assigned to points violating cross-checking. In particular, a two-step refinement process is now proposed, which exploits successively segmentation and depth border information in order to fill-in, respectively, low textured areas and regions along depth discontinuities.

First of all, the following *strong* cross-checking consistency constraint is applied on all points of D_{LR}:

$$D_{LR}(p_d) = \begin{cases} D_{LR}(p_d) & D_{LR}(p_d) \neq D_{RL}\left(p_d - D_{LR}(p_d)\right) \\ invalid & otherwise \end{cases} \tag{10}$$

The first step of the proposed refinement approach employs segmentation information in order to fill-in regions of D_{LR} denoted as invalid after application of (10). In particular, for each segment extracted from the application of the Mean Shift algorithm, a disparity histogram is filled with all valid disparities included within the segment area. Then, if a unique disparity value can be reliably associated with that segment, i.e. if there is a minimum number of valid disparities in the histogram and its variance is low, the mean disparity value of the histogram is assigned to all invalid points falling within the segment area. This allows to

correctly fill-in uniform areas which can be easily characterized by mismatches during the correspondence search.

As this first step is designed to fill-in only invalid points within uniform areas, then a second step allows to fill-in the remaining points by exploiting the previously extracted information on border locations, especially along depth border regions which usually are not characterized by uniform areas. In particular, the assigned disparity value for all invalid points near to depth discontinuities is chosen as the minimum value between neighbours which do not lie beyond a depth border. This allows to increase the reliability of the assigned values compared to the case of no information on borders location, where e.g. the minimum value between neighbouring disparities is selected, as shown in Fig. 2.

7 Experimental Results

This section shows an experimental evaluation obtained by submitting on the Middlebury site the results yielded by the proposed algorithm. The parameter set of the algorithm is constant for all runs and is the same as for the experiments in Sec. 4. As it can be seen from Table 2, our algorithm (*SO+border*), which ranked *4th* (as of May 2007), produces overall better results compared to [16], which employs a higher number of scanlines during the SO process, and also compared to the other SO and DP-based approaches and most global methods, for higher accuracy is only yielded by three BP-based global algorithms. Obtained disparity maps, together with corresponding reference images and groundtruth are shown in Fig. 3 and are available at the Middlebury website. The running time on the examined dataset is of the order of those of other methods based on a variable support [21], [2] (i.e. some minutes) since the majority of time is required by the local cost computation, while the S.O. stage and the border refinement stage only account for a few seconds and are negligible compared to the overall time.

Table 2. Disparity error rates and rankings obtained on Middlebury website

	Rank	Tsukuba N.O.-ALL-DISC	Venus N.O.-ALL-DISC	Teddy N.O.-ALL-DISC	Cones N.O.-ALL-DISC
AdaptingBP [12]	♯1	1.11-1.37-5.79	0.10-0.21-1.44	4.22-7.06-11.8	2.48-7.92-7.32
DoubleBP [10]	♯2	0.88-1.29-4.76	0.14-0.60-2.00	3.55-8.71-9.70	2.90-9.24-7.80
SymBP+occ	♯3	0.97-1.75-5.09	0.16-0.33-2.19	6.47-10.7-17.0	4.79-10.7-10.9
SO+border	**♯4**	**1.29-1.71-6.83**	**0.25-0.53-2.26**	**7.02-12.2-16.3**	**3.90-9.85-10.2**
Segm+visib [13]	♯5	1.30-1.57-6.92	0.79-1.06-6.76	5.00-6.54-12.3	3.72-8.62-10.2
C-SemiGlob [16]	♯6	2.61-3.29-9.89	0.25-0.57-3.24	5.14-11.8-13.0	2.77-8.35-8.20
RegionTreeDP [19]	♯10	1.39-1.64-6.85	0.22-0.57-1.93	7.42-11.9-16.8	6.31-11.9-11.8

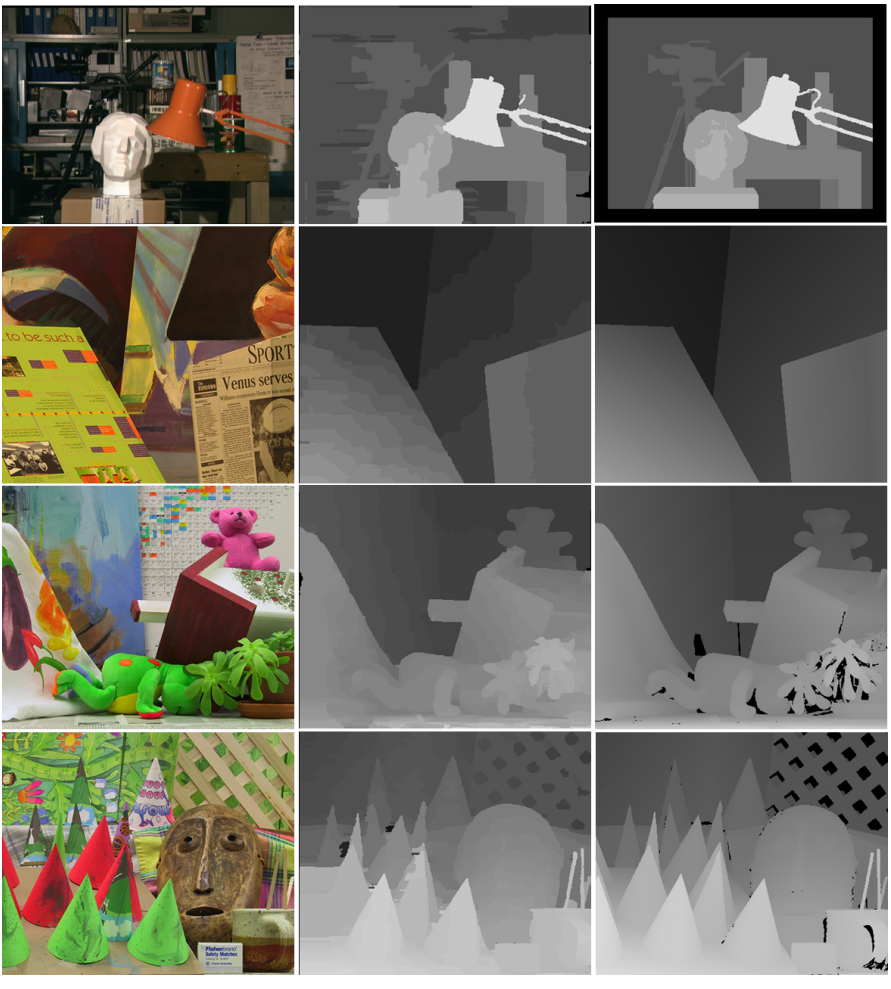

Fig. 3. Disparity maps obtained after the application of all steps of the proposed approach

8 Conclusions

A novel algorithm for solving the stereo correspondence problem has been described. The algorithm employs an effective variable-support based approach in the aggregation stage together with a SO-based framework in the disparity optimization stage. This joint strategy allows for improving the accuracy of both SO-based and local variable-support based methods. Further improvements are obtained by embodying the disparity refinement stage with border information and segmentation, which allows our proposal to outperform all DP and SO-based approaches as well as most global approaches on the Middlebury dataset.

References

1. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. Jour. Computer Vision 47(1/2/3), 7–42 (2002)
2. Yoon, K., Kweon, I.: Adaptive support-weight approach for correspondence search. IEEE Trans. PAMI 28(4), 650–656 (2006)
3. Boykov, Y., Veksler, O., Zabih, R.: A variable window approach to early vision. IEEE Trans. PAMI 20(12), 1283–1294 (1998)
4. Gong, M., Yang, R.: Image-gradient-guided real-time stereo on graphics hardware. In: Proc. 3D Dig. Imaging and Modeling (3DIM), Ottawa, Canada, pp. 548–555 (2005)
5. Hirschmuller, H., Innocent, P., Garibaldi, J.: Real-time correlation-based stereo vision with reduced border errors. Int. Jour. Computer Vision 47(1-3) (2002)
6. Xu, Y., Wang, D., Feng, T., Shum, H.: Stereo computation using radial adaptive windows. In: ICPR 2002. Proc. Int. Conf. on Pattern Recognition, vol. 3, pp. 595–598 (2002)
7. Gerrits, M., Bekaert, P.: Local stereo matching with segmentation-based outlier rejection. In: CRV 2006. Proc. Canadian Conf. on Computer and Robot Vision, pp. 66–66 (2006)
8. Wang, L., Gong, M., Gong, M., Yang, R.: How far can we go with local optimization in real-time stereo matching. In: 3DPVT 2006. Proc. Third Int. Symp. on 3D Data Processing, Visualization, and Transmission, pp. 129–136 (2006)
9. Kolmogorov, V., Zabih, R.: Computing visual correspondence with occlusions via graph cuts. In: ICCV 2001. Proc. Int. Conf. Computer Vision, vol. 2, pp. 508–515 (2001)
10. Yang, Q.e.a.: Stereo matching with color-weighted correlation, hierachical belief propagation and occlusion handling. In: CVPR 2006. Proc. Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 2347–2354 (2006)
11. Sun, J., Shum, H., Zheng, N.: Stereo matching using belief propagation. IEEE Trans. PAMI 25(7), 787–800 (2003)
12. Klaus, A., Sormann, M., Karner, K.: Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In: ICPR 2006. Proc. Int. Conf. on Pattern Recognition, vol. 3, pp. 15–18 (2006)
13. Bleyer, M., Gelautz, M.: A layered stereo matching algorithm using image segmentation and global visibility constraints. Jour. Photogrammetry and Remote Sensing 59, 128–150 (2005)
14. Tao, H., Sawheny, H., Kumar, R.: A global matching framework for stereo computation. In: ICCV 2001. Proc. Int. Conf. Computer Vision, vol. 1, pp. 532–539 (2001)
15. Hirschmuller, H.: Accurate and efficient stereo processing by semi-global matching and mutual information. In: CVPR 2005. Proc. Conf. on Computer Vision and Pattern recognition, vol. 2, pp. 807–814 (2005)
16. Hirschmuller, H.: Stereo vision in structured environments by consistent semi-global matching. In: CVPR 2006. Proc. Conf. on Computer Vision and Pattern recognition, vol. 2, pp. 2386–2393 (2006)
17. Gong, M., Yang, Y.: Near real-time reliable stereo matching using programmable graphics hardware. In: CVPR 2005. Proc. Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 924–931 (2005)
18. Kim, J., Lee, K., Choi, B., Lee, S.: A dense stereo matching using two-pass dynamic programming with generalized ground control points. In: CVPR 2005. Proc. Conf. on Computer Vision and Pattern Recognition, pp. 1075–1082 (2005)

19. Lei, C., Selzer, J., Yang, Y.: Region-tree based stereo using dynamic programming optimization. In: CVPR 2006. Proc. Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 2378–2385 (2006)
20. Deng, Y., Lin, X.: A fast line segment based dense stereo algorithm using tree dynamic programming. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 201–212. Springer, Heidelberg (2006)
21. Tombari, F., Mattoccia, S., Di Stefano, L.: Segmentation-based adaptive support for accurate stereo correspondence. Technical Report 2007-08-01, Computer Vision Lab, DEIS, University of Bologna, Italy (2007)
22. Yoon, K., Kweon, I.: Stereo matching with symmetric cost functions. In: CVPR 2006. Proc. Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 2371–2377 (2006)
23. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Trans. PAMI 24, 603–619 (2002)