Sensors & Transducers, Vol. 157, Issue 10, October 2013, pp. 272-279

Sensors & Transducers



http://www.sensorsportal.com

# Accurate Disparity Estimation Based on Integrated Cost Initialization

**[1] Haixu Liu, [2] Xueming Li**

[1] School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

[2] School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing 100876, China

[1] E-mail: liuhaixu@bupt.edu.cn

*Received: 23 August 2013 /Accepted: 25 September 2013 /Published: 31 October 2013*

**Abstract:** Disparity estimation is one of the most important research topics in computer vision. Numerous local-based approaches have been proposed to solve this problem. Among them, most state-of-the-art methods mainly focus on color information when initializing the cost volume. However, color signals are less robust and more easily affected by image noise, illumination variation and radiometric differences. In this paper, we develop a high quality disparity estimation system based on an integrated matching cost initialization algorithm. During the cost initialization step, three individual cost terms are utilized to construct the cost volume: Gradient-based Census Transform (GCT), Absolute Color Differences (ACD), and Gabor Pattern Differences (GPD). The proposed method produces impressive performance and ranks excellently in the Middlebury benchmark. Furthermore, we present that the proposed scheme is also capable for real-world outdoor scenes which contain many challenges. Both quantitative and qualitative evaluations demonstrate that our approach is currently one of the most accurate local-based stereo matching algorithms. *Copyright © 2013 IFSA.*

**Keywords:** Disparity estimation, Stereo matching, Gradient-based census transform, Absolute color differences, Gabor pattern differences.

## 1. Introduction

Disparity estimation from a stereo image pair has attracted much interest over the past few decades and it is still one of the most fundamental tasks in the research area of computer vision [1]. It is highly important in fields such as entertainment, information transfer, robotics, virtual reality systems, etc. Disparity estimation aims at estimating a dense disparity map from two original images of the same scene taken from different viewpoints. Generally, the disparity map describes the horizontal displacement between corresponding pixels located at different views. The main challenge of this task is how to accurately estimate the disparity from an image pair under different conditions including smooth surfaces, sharp boundaries, ambiguous areas and occluded regions.

A large number of studies have been explored to address this issue in the last few years. A detailed overview on these stereo matching algorithms and the state-of-the-art methods can be found in [2-4]. An integrated evaluation platform, namely Middlebury website [5], also provides a review of the latest methods. In general, these methods can be divided into two categories: global and local approaches. According to the literatures, the global methods such as Graph-Cut [6] and Belief-Propagation [7], which seek a disparity assignment through minimizing a global cost function that combines data and smoothness terms, usually produce more precise results. However, a drawback is that such global

approaches are often quite time consuming and do not work well for high-resolution images or large disparity spaces. By contrast, local-based methods are much faster and more suitable for a practical implementation, even capable for real-time task [8].

Plenty of local methods have been proposed in the literature, among these approaches, adaptive support weight algorithms like AdaptWeight [9] and CostFilter [10] currently represent the state-of-the-art local-based approaches. Results of these algorithms are comparable to those generated by global methods. The main reason of their success is to exploit the locally adaptive support weights for computing the probability that the center pixel and a neighbor pixel might have the same disparity value.

A local-based algorithm generally consists of four steps:

1) Matching cost initialization. For each pixel, a cost value is assigned to all possible disparities.

2) Cost aggregation. An aggregated cost is then calculated by an adaptive summation of the initial per-pixel cost.

3) Disparity optimization by using Winner-Takes-All strategy and

4) Disparity refinement, which aims at correcting inaccurate disparity values and handling occlusion areas.

Although matching cost initialization is the fundamental component that determines the final precision of estimated disparity map, relatively less work has been studied on this topic. Common cost initialization methods like sum of absolute or squared differences (SAD/SSD), sampling-insensitive absolute difference of Birchfield and Tomasi (BT), and truncated versions, usually calculate the intensity difference between a pixel and its corresponding one in another viewpoint. Due to they only utilize the information of color channel, these approaches are generally highly sensitive to illumination changes, image noises, radiometric differences, etc. Non-parametric matching methods such as Rank and Census [11] were introduced for being robust against outliers that occur in brightness variation areas. However, since non-parametric methods depend on the relative ordering of pixel values, they could also produce matching ambiguities in image regions with similar or repetitive local structures [12].

In this paper, we explore to develop a novel and robust cost initialization strategy which can integrate the advantages of various cost computation schemes. We propose a *Gradient-based Census Transform (GCT)* cost term for keeping the illumination invariability. Meanwhile, *Absolute Color Differences (ACD)* is adopted to reduce the matching ambiguities in repetitive areas. To further improve the accuracy of the initial cost, we also present an effective cost term, namely, *Gabor Pattern Differences (GPD)*, which is particularly appropriate for texture representation and discrimination. Experimental results demonstrate that the proposed system outperforms most existed local-based algorithms.

The remainder of this paper is structured as follows. Section 2 presents the detailed explanation of the proposed scheme. After that, Section 3 illustrates the experimental results tested on both Middlebury platform and the real-world outdoor scenes. Finally, conclusions and future work are discussed in Section 4.

## 2. Technical Details

The flowchart of the proposed algorithm is illustrated in Fig. 1.

The whole system consists of four steps: cost computation, cost aggregation, disparity selection and disparity refinement. In the following sections, we will go through each individual step in detail.

### 2.1. Cost Computation

Cost computation (i.e. cost initialization) is not only the groundwork for the whole progress but also vital to the final accuracy of stereo matching. Therefore, in this paper, an integrated and robust cost function composed of three special cost terms is explored for cost initialization.

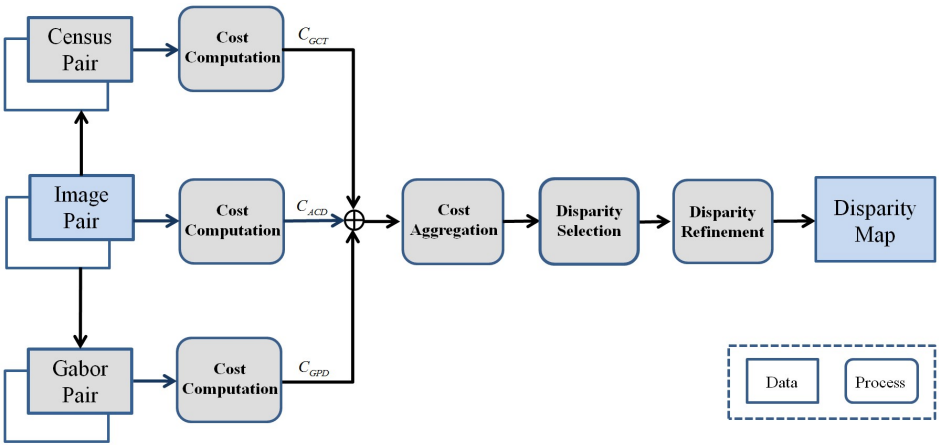



**Fig. 1.** Flowchart of the proposed algorithm.

The first cost term is *Gradient-based Census Transform (GCT)*. Census [11] is a non-parametric method, it encodes local structures with a bit-string calculated by the comparison of pixel intensities, rather than the intensities themselves. Consequently, it is robust to illumination variation and image noise. Unlike the conventional approaches which calculate census on the original images, in this paper, instead, we implement the transformation on the gradient images of the horizontal direction.

To be specific, for a pixel $p$, we define $\nabla_H(I(p))$ as its intensity of horizontal gradient image, and specify $N_p$ as the set of neighbor pixels in a square window of radius $r$ surrounding $p$, the transforms $\xi(p,q)$ is depended on the comparison between the intensities of two pixels:

$$\underset{p\in\nabla_H(I),q\in N_p}{\xi(p,q)} = \begin{cases} 1 & if\ \nabla_H(I(p)) < \nabla_H(I(q)) \\ 0 & otherwise \end{cases} \quad (1)$$

The census transform can then be presented as a binary vector:

$$S(p) = \otimes\xi(p,q), \quad (2)$$

where $\otimes$ denotes the act of concatenation. After that, the cost value $C_{GCT}$ at pixel $p$ with disparity $d$ is defined as the Hamming distance of two bit strings:

$$C_{GCT}(p,d) = \sum_{i=1}^{N} S_{left}(p)[i] \oplus S_{right}(p-d)[i], \quad (3)$$

where $S_{left}$ and $S_{right}$ depict the bit string in the left and right image respectively. $N$ is the length of the string, and $\oplus$ indicates XOR operation.

The reasons we utilize GCT instead of the original Census method are as follows:

1) GCT is less sensitive to radiometric differences caused by exposure differences, varying lighting, vignetting, etc.

2) For image regions with repetitive structures, GCT could alleviate the matching ambiguities more effectively.

3) GCT is more sensitive to object edges due to the feature of gradient, this characteristic can guarantee a better edge-preserving result.

The second cost term we employed is the *Absolute Color Differences (ACD)*. ACD is adopted due to the assumption that the neighboring pixels with homogeneous color share the similar disparity. Accordingly, this cost term is defined as the average value of absolute color differences in all three channels:

$$C_{ACD}(p,d) = \frac{1}{3}\sum_{i=R,G,B}\left|I_i^{Left}(p) - I_i^{Right}(p-d)\right| \quad (4)$$

The third cost term we proposed is the *Gabor Pattern Differences (GPD)*. As a matter of fact, Gabor feature [13] has several intrinsic advantages for the application of stereo matching:

1) It is robust to illumination changes and image noise.

2) It is particularly appropriate for texture discrimination and representation.

3) Gabor filter is a linear filter which can be operated sufficiently fast.

4) The orientation of Gabor filter can be easily controlled. (For the rectified images whose motion is purely along the x-axis, it is only necessary to utilize the information extracted from the horizontal direction.) These superiorities make it become a kind of potential and ideal stereo matching cost indicator.

Generally, the kernel of Gabor filter can be written as:

$$G(x,y,\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right) \quad (5)$$

$$\begin{cases} x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases} \quad (6)$$

In Eq. (5) and Eq. (6), $\lambda$ depicts the wavelength of the sinusoidal factor, $\theta$ represents the orientation of the normal to the parallel stripes of a Gabor function, $\psi$ is the phase offset, $\sigma$ is the sigma of the Gaussian envelope and $\gamma$ is the spatial aspect ratio.

In order to construct the GPD term, we first apply this Gabor filter to extract the feature $G_H(I)$ from original image $I$ in the horizontal direction, then the cost of GPD term $C_{GPD}(p,d)$ for pixel $p$ at disparity $d$ can be specifically expressed as:

$$C_{GPD}(p,d) = \left|G_H(I_{left}(p)) - G_H(I_{right}(p-d))\right| \quad (7)$$

Finally, the whole matching cost function $C_{all}$ is summarized by merging three different cost terms:

$$C_{all}(p,d) = \sum_{i=GCT,ACD,GPD}\min(\rho(C_i(p,d),\lambda_i),T_i), \quad (8)$$

where $T_i$ is the truncation parameter, $\rho(C,\lambda)$ is the robust function of variable $C$:

$$\rho(c,\lambda) = 1-\exp(-\frac{C}{\lambda}) \quad (9)$$

The value of each cost term is normalized to the range of [0, 1] by parameter $\lambda$, this strategy can guarantee that each term appropriately contributes to the whole cost function. The result of cost initialization is shown in Fig. 2 (b).

(a) (b) (c) (d) (e)

**Fig. 2.** Results generated in each step.
(a) The original color image; (b) Result of matching cost computation (disparity=1);
(c) Result of cost aggregation (disparity=1); (d) Disparity map (before refinement); (e) Disparity map (after refinement).

## 2.2. Cost Aggregation

This step aggregates the initial cost value of each pixel over a support region to preserve the borders of objects and reduce the matching ambiguity for further disparity selection. Fig. 3 depicts the whole process of cost aggregation.

Inspired by the previous work [10], in our proposed scheme, we utilize the guided filter [14] to implement cost aggregation. The main reason we choose guided filter is that it is a non-approximate linear-time algorithm whose computational complexity is independent of the filtering kernel size. This advantage implies that it can be performed sufficiently fast, even is capable of real-time task. To be more precise, the result of aggregation $C'_{all}(p)$ at pixel $p$ can be calculated by the initial cost map $C_{all}(p)$ and the value of guidance image $I$ :

$$C'_{all}(p)=\frac{1}{|\omega|}\sum_{k:p\in\omega_k}(a_k I(p)+b_k) \tag{10}$$

$$a_k=(\Sigma_k+\epsilon \mathrm{U})^{-1}(\frac{1}{|\omega|}\sum_{p\in\omega_k} I(p)C_{all}(p)-\mu_k\overline{C}_{all}(k)), \tag{11}$$

$$b_k=\overline{C}_{all}(k)-a_k^{\mathrm{T}}\mu_k, \tag{12}$$

where $I$ depicts the reference image in stereo matching operation, $|\omega|$ is the pixel number of $I$ in a window $\omega_k$ centered at pixel $k$, $\Sigma_k$ is the 3×3 covariance matrix of $I$ in $\omega_k$, $\epsilon$ is a threshold for smoothing, U is an 3×3 identity matrix, while $\mu_k$ and $\overline{C}_{all}(k)$ present the mean of $I$ in $\omega_k$ and the mean of $C_{all}$ in $\omega_k$, respectively. The result of cost aggregation is illustrated in Fig. 2 (c).

## 2.3. Disparity Selection

To select the best candidate for estimating disparity value $W_{pq}=k\cdot\exp(-(\frac{\Delta c_{pq}}{\gamma_c}+\frac{\Delta s_{pq}}{\gamma_s}))$ at each pixel $p$, the Winner-Takes-All (WTA) strategy is adopted:

$$D(p)=\arg\min_{d\in S_d} C'_{all}(p,d), \tag{13}$$

where $S_d=\{d_{\min},\cdots,d_{\max}\}$ is the set of all possible disparity values. The disparity selection result can be seen in Fig. 2 (d).

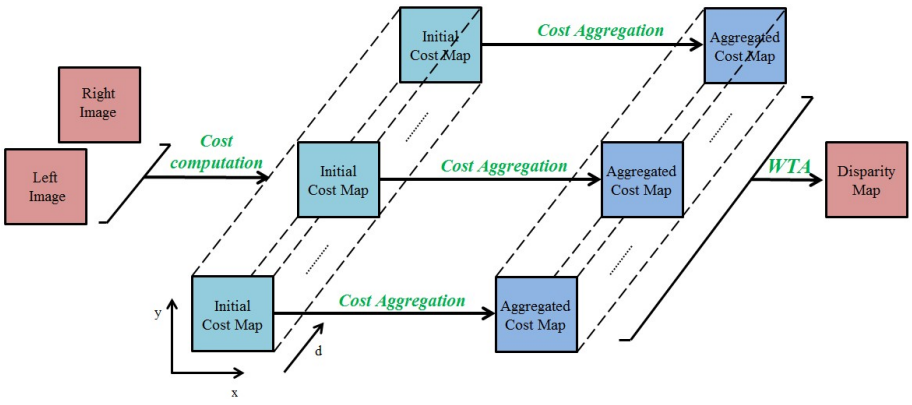



**Fig. 3.** The process of cost aggregation.

### 2.4. Disparity Refinement

In the previous step, WTA selects a minimum cost value for each pixel, however, the disparity maps obtained usually contain numerous errors in the occluded regions. To handle this problem, we implement the left-right consistency checking. That is to say, for each pixel $p$ in the left image, we also calculate the disparity value of its matching point $p'$ in the right view. If $\left|D_{left}(p)-D_{right}(p')\right|>1$, pixel $p$ is defined as an invalid pixel. This consistency check can effectively detect the occluded pixels as well as the mismatched ones.

For each invalid pixel $p$, as illustrated in Fig. 4, the disparity values of its left nearest valid pixel $p_l$ and right nearest valid pixel $p_r$ are computed and denoted as $D(p_l)$ and $D(p_r)$, respectively. Then the invalid pixel $p$ is filled by the lower value of $D(p_l)$ and $D(p_r)$, namely, $D(p)=min\left(D(p_l),D(p_r)\right)$.

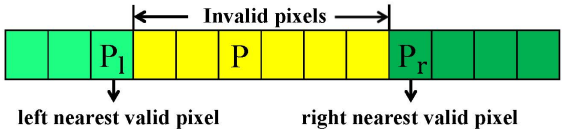



**Fig. 4.** Demonstration for dealing with invalid pixels.

Selecting the lower disparity is motivated by the fact that the occluded pixels primarily belong to the background objects.

Finally, a weighted median filter is utilized on the locations of invalid pixels to eliminate the horizontal artifacts. In order to preserve the object boundaries and generate the sharp disparity map, the bilateral filter is selected, and the weight is defined as:

$$W_{pq}=k\cdot\exp(-(\frac{\Delta c_{pq}}{\gamma_c}+\frac{\Delta s_{pq}}{\gamma_s})), \tag{14}$$

where $k$ is a normalization factor, $\Delta c_{pq}$ and $\Delta s_{pq}$ depict the color difference and the spatial distance between pixel $p$ and $q$, respectively. $\gamma_c$ and $\gamma_s$ control the color and spatial similarity, respectively. In our experiment, we set $\gamma_c=0.16$, $\gamma_s=7$, and the window size = (17 × 17).

Note that all the pixels which have survived in the left-right consistency checking are not affected by the whole refinement process, this helps to rebuild the object boundaries and construct the sharp disparity map without substantial increase in computation overheads. The final refined disparity map is demonstrated in Fig. 2 (e).

## 3. Experiments and Analysis

### 3.1. Evaluation on Middlebury Platform

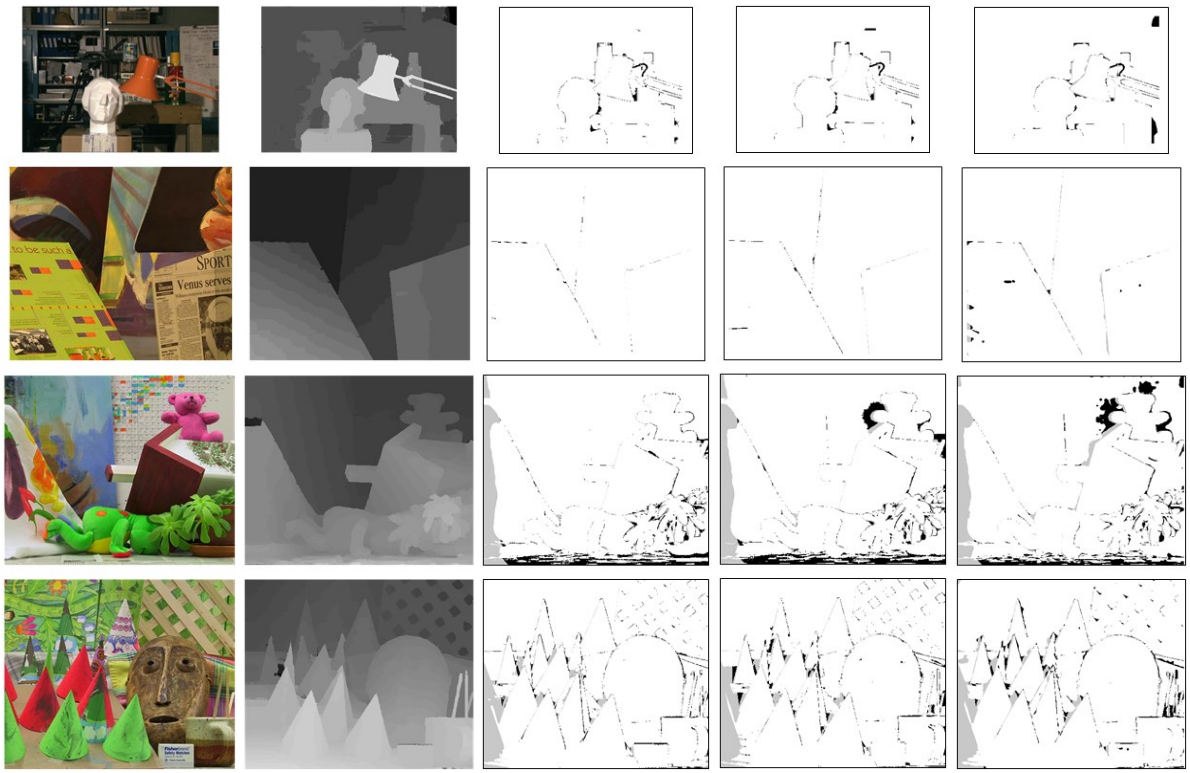
In order to assess the performance of our algorithm, we compare our scheme with some classic local-based algorithms. For fair comparison, all algorithms are tested on the Middlebury website [5], which is a popular standard evaluation platform for various stereo matching approaches. This platform provides 4 stereo pairs as well as their ground truth disparity maps. Note that these test images are challenging due to their complex object boundaries and low textured regions. For our approach, the parameters are set as follows: $\{\lambda,\theta,\psi,\sigma,\gamma,\lambda_{GCT},\lambda_{ACD},\lambda_{GPD},T_{GCT},T_{ACD},T_{GPD},\epsilon\}=\{3,\ 3\pi/2,\ 0,\ 1.5,\ 1,\ 32,\ 40,\ 0.18,\ 0.008,\ 0.025,\ 0.018,\ 0.01^2\}$. These parameters have been found empirically.

The subjective evaluation results are illustrated in Fig. 5. The original four color images (left view) are shown in Fig. 5 (a). The disparity maps generated by our approach are illustrated in Fig. 5 (b). We can see that the proposed algorithm could obtain smooth disparities on the surface of objects while preserving the discontinuity property at the boundaries.

Furthermore, we also present the error maps of our method in Fig. 5 (c). These error maps are calculated according to the ground truth maps, pixels whose disparity error greater than one are displayed in black. Compared with other classic local-based methods (shown in Fig. 5 (d) and Fig. 5 (e)), the proposed method can generate much clearer object boundaries as well as less error regions. Moreover, the bad pixels in our results are eliminated well both in low texture areas and occluded regions.

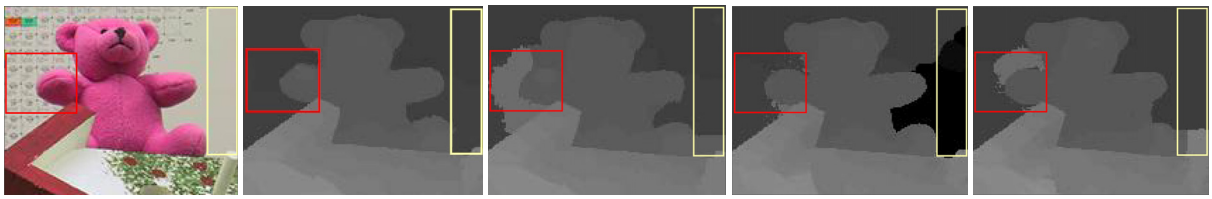
To further verify the effect of the proposed scheme, some close-up disparity results on Teddy and Cones image pair are presented in Fig. 6 and Fig. 7. It is apparent that the proposed approach is adaptive for both textureless areas (e.g., wall of the Teddy data set) and object boundaries (e.g., arms of the teddy bear). Meanwhile, for the repetitive regions (e.g., background of the cup), the proposed method is also qualified.

For the objective evaluation, all of the disparity results are compared with the ground truth maps pixel by pixel. The assessment results are shown in Table.1. The performance of each method is mainly evaluated by the error rate of the whole image. Consequently, according to the report of Middlebury website, our approach obtains satisfactory results ranking 18th out of over 150 (global and local) competitive algorithms, even better than CostFilter [15], which is one of the best local-based methods so far.



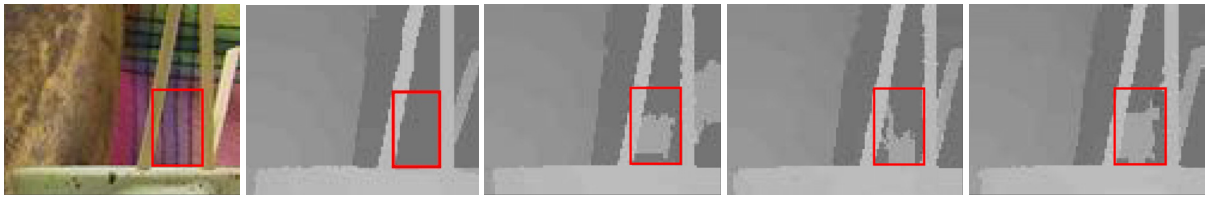
(a) Left images | (b) Disp. (Ours) | (c) Error (Ours) | (d) Error (CostFilter [15]) | (e) Error (RBF [16])

**Fig. 5.** Subjective evaluation of different local-based methods.
(a) Left images; (b) Disparity maps generated by proposed approach;
(c)-(e) Error maps of different schemes, (Error pixels larger than one are shown in black in the error maps;
while gray pixels correspond to errors in occluded areas).



(a) Color image | (b) ours | (c) CostFilter [15] | (d) VSW [17] | (e) AdaptWeight [9]

**Fig. 6.** Close-up disparity results on Teddy image pair. Our method performs well for both textureless regions
and object boundaries. (Invalid pixels are shown in black.)



(a) Color image | (b) ours | (c) CostFilter [15] | (d) VSW [17] | (e) AdaptWeight [9]

**Fig. 7.** Close-up disparity results on Cones image pair. Our method produces
better disparity result at repetitive structures.

**Table 1.** Ranks of local-based algorithms on Middlebury.

| Algorithm | Rank | Error (All) | | | | Av.E |
|---|---|---|---|---|---|---|
| | | Tsukuba | Venus | Teddy | Cones | |
| Ours | 1 | 1.83 | 0.36 | 10.3 | 7.85 | 5.12 |
| CostFilter [15] | 2 | 1.85 | 0.39 | 11.8 | 8.24 | 5.55 |
| RecursiveBF [16] | 3 | 2.51 | 0.88 | 12.1 | 8.91 | 5.68 |
| VSW [17] | 4 | 1.88 | 0.81 | 13.3 | 8.85 | 6.29 |
| AdaptWeight [9] | 5 | 1.85 | 1.19 | 13.3 | 9.79 | 6.67 |
| DCBGrid [18] | 6 | 7.26 | 1.91 | 17.2 | 11.9 | 10.9 |
| SSD+MF [2] | 7 | 7.07 | 5.16 | 24.8 | 19.8 | 15.7 |

Note that, the error rate of our method is much smaller than others for all the four test cases. Additionally, the Average percent of Error pixels (Av.E), which is calculated based on all test areas (including occluded regions, non-occluded regions and whole image), can be effectively reduced to 5.12 by utilizing our algorithm, that also outperforms other competitors.

### 3.2. Outdoor Results

To test the outdoor performance and the robustness of our algorithm, we implement our scheme on the images captured from two complex outdoor scenes with a consumer stereo cellphone (HTC EVO 3D). The original images and corresponding disparity maps calculated by our method are presented in Fig. 8. The results show that our approach performs very well and calculates a sharp disparity map, although the input images are without any rectification before stereo matching, and contain many challenges including complicated patterns (e.g., leaves), ambiguous regions (e.g., shadows), specular surfaces (e.g., fitness equipments) as well as illumination changes.

## 4. Conclusions and Future Work

In this paper, we proposed a high quality local-based algorithm for stereo matching. Unlike the existed methods which only use the color and gradient information, we introduced a novel and robust cost function using three special cost terms (namely, Gradient-based Census Transform (GCT), Absolute Color Differences (ACD) and Gabor Pattern Differences (GPD)) into the cost initialization process. Experimental results tested on both Middlebury platform and the outdoor scenes demonstrate that our approach outperforms most existed local-based methods.

For the future research, we would like to accelerate our algorithm with graphics hardware, this implementation may lead to real-time performance and will be useful for more applications.

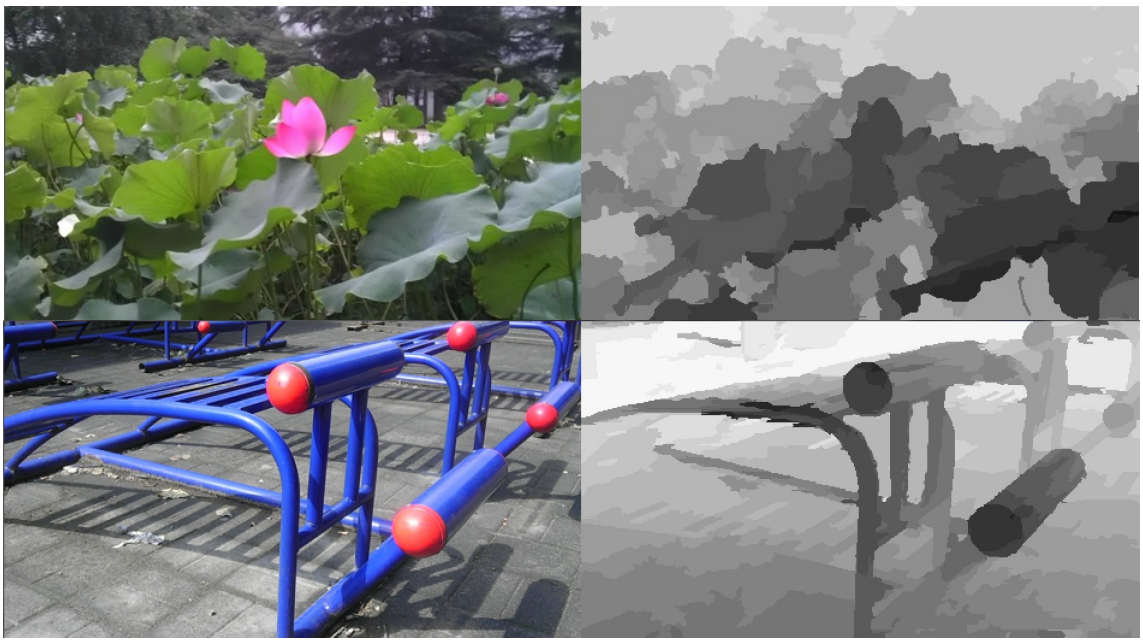


**Fig. 8.** Outdoor Results. Left images captured with a consumer stereo cellphone and the corresponding disparity maps generated by our approach.

## References


[1]. F. Mroz and T. P. Breckon, An empirical comparison of real-time dense stereo approaches for use in the automotive environment, *EURASIP Journal on Image and Video Processing*, Vol. 2012, No. 1, 2012, pp. 1–19.

[2]. D. Scharstein and R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, *International Journal of Computer Vision (IJCV)*, Vol. 47, 2002, pp. 7–42.

[3]. M. Z. Brown, D. Burschka and G. D. Hager, Advances in computational stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol. 25, No. 8, 2003, pp. 993–1008.

[4]. R. Szeliski, Computer vision: algorithms and applications, Springer, 2011.

[5]. Middlebury website. (http://vision.middlebury.edu/stereo/).

[6]. V. Kolmogorov and R. Zabih, Multi-camera scene reconstruction via graph cuts, in *Proceedings of the European Conference of Computer Vision (ECCV)*, 2002, pp. 82–96.

[7]. A. Klaus, M. Sormann and K. Karner, Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure, in *Proceedings of the International Conference on Pattern Recognition (ICPR)*, Vol. 3, 2006, pp. 15–18.

[8]. M. Humenberger, C. Zinner, M. Weber, W. Kubinger and M. Vincze, A fast stereo matching algorithm suitable for embedded real-time systems, *Computer Vision and Image Understanding (CVIU)*, Vol. 114, No. 11, 2010, pp. 1180–1202.

[9]. K. J. Yoon and I. S. Kweon, Adaptive support-weight approach for correspondence search, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol. 28, No. 4, 2006, pp. 650–656.

[10]. A. Hosni, C. Rhemann, M. Bleyer, C. Rother and M. Gelautz, Fast cost-volume filtering for visual correspondence and beyond, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol. 35, No. 2, 2013, pp. 504–511.

[11]. R. Zabih and J. Woodfill, Non-parametric local transforms for computing visual correspondence, in *Proceedings of the European Conference of Computer Vision (ECCV)*, 1994, pp. 151–158.

[12]. H. Hirschmuller and D. Scharstein, Evaluation of stereo matching costs on images with radiometric differences, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol. 31, No. 9, 2009, pp. 1582–1599.

[13]. D. Gabor, Theory of communication. part 1: The analysis of information, *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, Vol. 93, No. 26, 1946, pp. 429–441.

[14]. K. He, J. Sun and X. Tang, Guided image filtering, in *Proceedings of the European Conference of Computer Vision (ECCV)*, 2010, pp. 1–14.

[15]. C. Rhemann, A. Hosni, M. Bleyer, C. Rother and M. Gelautz, Fast cost-volume filtering for visual correspondence and beyond, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 3017–3024.

[16]. Q. Yang, Recursive bilateral filtering, in *Proceedings of the European Conference of Computer Vision (ECCV)*, 2012, pp. 399–413.

[17]. W. Hu, K. Zhang, L. Sun, J. Li, Y. Li and S. Yang, Virtual support window for adaptive-weight stereo matching, in *Proceedings of the Conference on Visual Communications and Image Processing (VCIP)*, 2011, pp. 1–4.

[18]. C. Richardt, D. Orr, I. Davies, A. Criminisi and N. Dodgson, Real-time spatiotemporal stereo matching using the dual-cross-bilateral grid, *in Proceedings of the European Conference of Computer Vision (ECCV)*, 2010, pp. 510–523.