

## Research Article

# Selective Segmentation for Global Optimization of Depth Estimation in Complex Scenes

Sheng Liu,<sup>1,2</sup> Haiqiang Jin,<sup>1</sup> Xiaojun Mao,<sup>1</sup> Binbin Zhai,<sup>1</sup> Ye Zhan,<sup>3</sup> and Xiaofei Feng<sup>4</sup>

<sup>1</sup> College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China

<sup>2</sup> Key Laboratory of Visual Media Intelligent Processing Technology of Zhejiang Province, Hangzhou 310023, China

<sup>3</sup> School of Accounting, Zhejiang University of Finance and Economics, Hangzhou 310018, China

<sup>4</sup> College of Computer and Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

Correspondence should be addressed to Sheng Liu; edliu@zjut.edu.cn

Received 27 January 2013; Accepted 5 March 2013

Academic Editors: C. W. Ahn and C. Cattani

Copyright © 2013 Sheng Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a segmentation-based global optimization method for depth estimation. Firstly, for obtaining accurate matching cost, the original local stereo matching approach based on self-adapting matching window is integrated with two matching cost optimization strategies aiming at handling both borders and occlusion regions. Secondly, we employ a comprehensive smooth term to satisfy diverse smoothness request in real scene. Thirdly, a selective segmentation term is used for enforcing the plane trend constraints selectively on the corresponding segments to further improve the accuracy of depth results from object level. Experiments on the Middlebury image pairs show that the proposed global optimization approach is considerably competitive with other state-of-the-art matching approaches.

## 1. Introduction

Depth estimation from a pair of rectified stereo images is always a challenging research field in vision analysis [1, 2]. The local stereo matching methods often generate outliers in weakly textured areas, discontinuous boundaries, and occlusion areas. Consequently, the global optimization methods [3–7] are designed for more accurate depth estimating in comparison with local ones. Nevertheless, all of these above-mentioned methods neglected the segmentation information in the optimization framework.

The later global optimization methods only partially incorporated the segmentation information into a pixel-level *MRF* model [8–14]. The segmentation information was merely integrated into unary terms or pairwise terms rather than higher order terms. For instance, Wang and Lim [10] proposed a new segment-based stereo matching approach, which takes segments as graph nodes for constructing an irregular segmentation-based graph. In spite of decreasing the computation complexity immensely and showing object-level feature information clearly, it neglected the depth detail

and structure detail within the segment and accordingly resulted in the “Mosaic Effect.”

For taking full advantage of segmentation information, Kohli et al. [15] proposed a higher order term including complete detail of each segment. The *Robust P<sup>n</sup> Potts model* presented by Kohli was originally designed for segmentation applications, which is based on an assumption that the pixels inside the same segment should be label consistency. The labels are used to identify different objects for image segmentation, other than different disparities for depth estimation. So, the energy function for depth estimation cannot penalize the segment with a linear penalty which takes inconsistency pixel ratio into account. Therefore, Kohli’s approach is unable to be applied in depth estimation directly. Xie et al. [16] improved the higher order term proposed by Kohli et al. and applied it to the depth estimation successfully. The improved higher order term enforces impliedly the assumption that all the segments of the input image are regarded as various planes. Nevertheless, this assumption is unreasonable because the surfaces of objects are more likely to be irregular surfaces rather than planes in real scene.

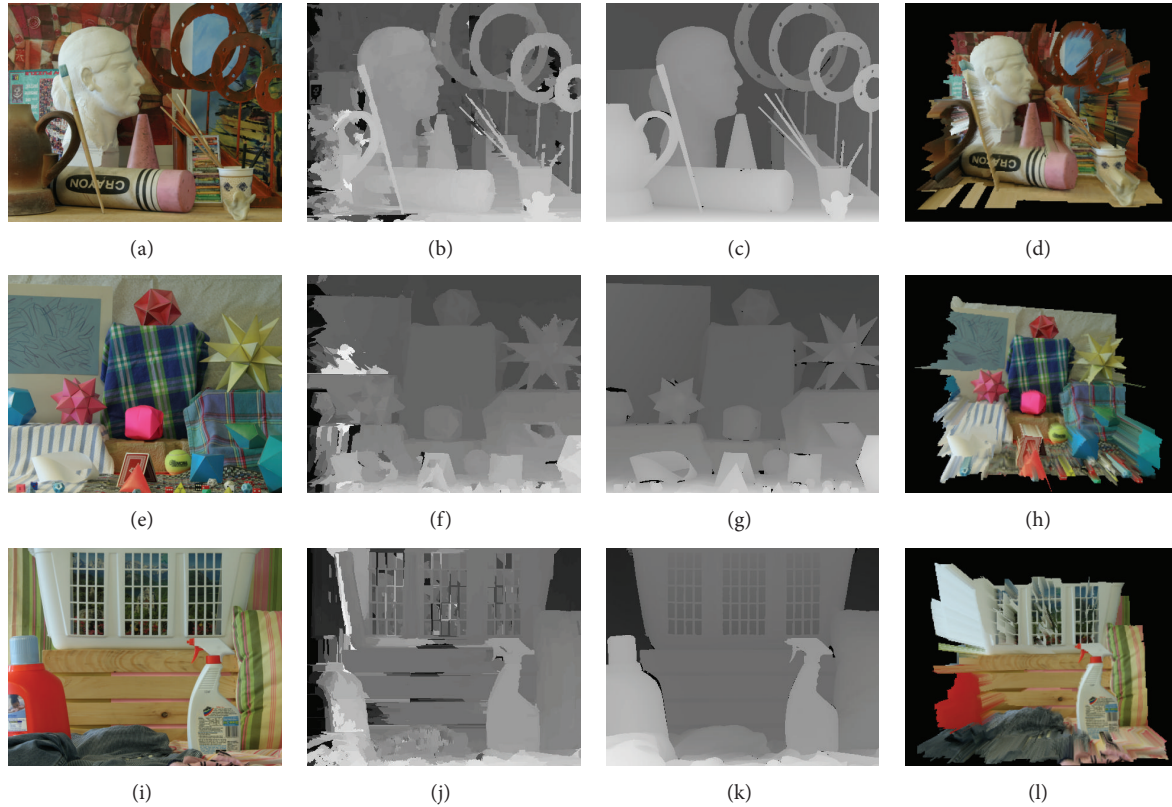


FIGURE 1: Dense depth maps for the Art, Moebius, and Laundry test sets (from top to bottom). From left to right: the input left images, our final depth maps, ground truth, and three-dimensional reconstructed results. Compared with the ground truth, our results obviously acquire most details of the scene with relatively high accuracy.

This paper proposes a segmentation-based global optimization method for the depth estimation. Our approach composed of four energy terms makes the following contributions: unlike those familiar data terms converted from local stereo matching methods directly, our data term combines a self-adapting stereo matching approach and two matching cost optimization strategies aiming at occlusion regions and border of image. Most smoothness terms only enforce a simple smoothness strategy over the whole image, which is obviously unable to satisfy the fact that different regions have varying smoothness requirements in a disparity map. Hence, our smoothness term employs a comprehensive smoothness strategy. We incorporate segmentation information in the form of higher order term and perform a selective planarity operation by enforcing a plane trend or not when facing diverse segments.

Experiment results on the stereo images in Middlebury datasets (Figure 1) have shown that our global optimization method obtains satisfactory depth results and is competitive with the state-of-the-art algorithms.

## 2. Global Optimization Method for Depth Estimation

**2.1. Algorithm Overview.** The input of our algorithm is a pair of rectified stereo images, which are used in improved local

stereo matching method based on self-adapting matching windows, color segmentation, and process of constructing smooth term. With the handling of two proposed matching cost optimization strategies, the final matching costs for the pixels are used to not only construct data term but also computer refine map. Both smooth term and segmentation term require the segmentation information produced by [17]. The proposed energy function composed of four energy terms is optimized using  $\alpha$ -expansion move algorithm [18]. The whole procedure of our algorithm is illustrated in Figure 2.

**2.2. Energy Function.** In this paper, we presented a segmentation-based global optimization approach composed of *integrated data term*, *comprehensive smoothness term*, and *selective segmentation term*. To make use of the pixel-level information more adequately, the proposed data term is not only decided by the matching costs from the local stereo matching method based on the improved self-adapting window but also mended the replacement for occlusion regions and evaluation for border of image according to two proposed optimization strategies. Due to comprehensive smoothness strategy, our smooth term is able to satisfy the smoothness requirement more fully. By fusing object-level over-segment information in our global optimization framework, we can richly utilize homogeneous information in the same segment. In addition, the selective planarity operation for segments

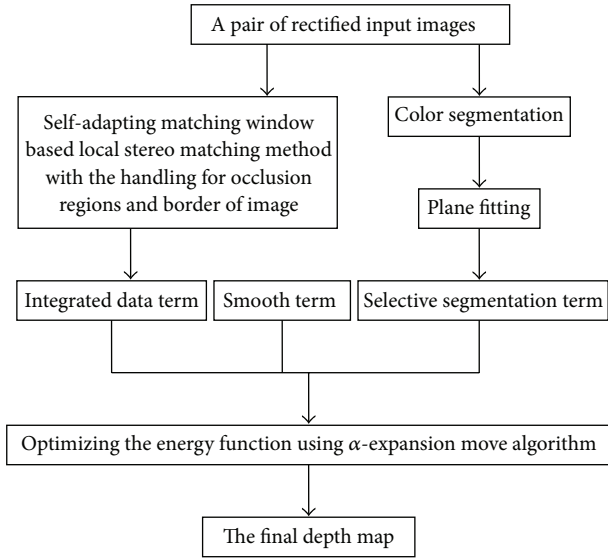


FIGURE 2: Flow chart of the proposed algorithm.

makes our segmentation term more robust. The global energy function for a unique configuration  $f$  is as follows:

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) + E_{\text{seg}}(f). \quad (1)$$

**2.3. Data Term Based on Self-Adapting Window.** In most local stereo matching methods, the fixed matching window is employed for depth estimation. Nevertheless, it is difficult to guarantee that all the pixels in a fixed window are of the same depth. Therefore, there exist amounts of outliers in weak-textured areas, discontinuous boundaries, and occlusion regions shown in Figure 3. In order to improve the accuracy of matching costs for the corresponding depths, the local stereo matching approach based on the self-adapting matching window is adopted for computing the matching costs.

The local stereo matching approaches with self-adapting matching window are based on the assumption that when pixels with similar intensity within a constrained window have similar disparity, it is necessary to produce an appropriate matching window for each pixel adaptively. In this paper, we mainly refer to the local stereo matching method proposed by Zhang et al. [22] based on self-adapting matching window. Two aspects of improvement are made on the basis of original approach: firstly, a dynamical argument mechanism of minimum window is proposed for more robust correspondence matching. Secondly, we enforce a replacement strategy for occlusion regions and a suboptimum strategy for borders of image.

Being inspired by five major approaches introduced by Egnal and Wildes [23], we present a replacement strategy to deal with the occlusion regions. Owing to the common assumption that pixels with similar intensity within a neighboring area have similar disparity, the matching costs for occlusion pixels are capable of being replaced with ones for “corresponding” pixels.

For instance,  $d(p)$  is the disparity for pixel  $p = (x_p, y_p)$  in the left input image, and  $d'(p')$  is the disparity for pixel

$p' = (x_p - d(p), y_p)$  in the right image. If  $d(p)$ ,  $d'(p')$  and  $d(p'')$  satisfy simultaneously the condition that  $d(p) > d'(p')$  and  $d'(p') \leq d(p'')$  where  $p'' = (x_p - d(p) + d'(p'), y_p)$ , we would employ a displacement strategy that the matching costs for the pixel  $p$  in left image are replaced with the one for the pixel  $p'$  in right image.

Neither estimating two disparity maps for left-right consistency check [24, 25] nor applying a simple border extrapolation step, we adopt a suboptimum strategy for the border of image. The corresponding pixel  $p'$  will locate outside the right image when  $(x_p - d(p)) < 1$ , which means that the matching cost  $C_d(p)$  cannot be achieved by making use of the corresponding pixels. In this paper, we need the suboptimum label  $d^\wedge$ ,

$$d^\wedge = \arg \min_{d \in [d_{\min}, d_{\max}], (x_p - d) > 0, d \neq d^*} C_d(p), \quad (2)$$

where  $d^*$  is the optimal label computed as follows:

$$d^* = \arg \min_{d \in [d_{\min}, d_{\max}], (x_p - d) > 0} C_d(p). \quad (3)$$

At last, we use  $C_{d^\wedge}(p)$  as the matching cost for pixel  $p$  when  $(x_p - d(p)) < 1$ . The improved local results are shown in Figure 4.

**2.4. Smooth Term Based on Comprehensive Management.** All kinds of smooth terms are presented for smoothing the coarse local results. In this paper, a new comprehensive smooth term is defined based on the similarity of color for dealing with different smoothing requirements on neighborhoods. The proposed smooth term combines the following two smooth terms.

Assume that there is a neighborhood system  $N$  on the pixel set  $P$ ,  $N \subset \{(p_1, p_2) \mid p_1, p_2 \in P\}$ , Yu et al. [7] performed the consistency of corresponding pixels and their neighbors in their smooth term as follows:

$$E_{\text{smooth}}(f) = \sum_{(p_1, p_2) \in N} \min((f(p_1) - f(p_2)), k), \quad (4)$$

where  $k$  is a constant.

Kolmogorov and Zabih [3] presented a different smooth term, which considers the color information of corresponding pixels and their neighbors. The smooth term is formulated as follows:

$$E_{\text{smooth}}(f) = \sum_{(p_1, p_2) \in N} V_{p_1, p_2} * T(f(p_1) \neq f(p_2)), \quad (5)$$

where  $V_{p_1, p_2}$  denotes a positive penalty function which imposes disparate penalties according to color differences between pixels. Suppose  $R(p)$ ,  $G(p)$ , and  $B(p)$  are the respective color components of pixel  $p$  in RGB space,

$$V_{p_1, p_2} = \begin{cases} 3\lambda, & \text{if } \max(|R(p_1) - R(p_2)|, |G(p_1) - G(p_2)|, \\ & |B(p_1) - B(p_2)|) < \varepsilon, \\ \lambda, & \text{otherwise,} \end{cases} \quad (6)$$

where  $\lambda$  is a penalty constant,  $\varepsilon$  manages a least color diversity.

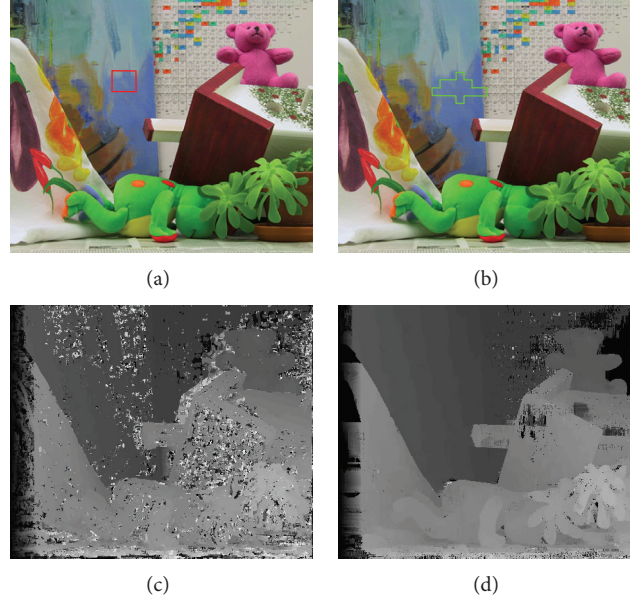


FIGURE 3: Comparison of local stereo matching methods with fixed matching window and self-adapting matching window for the Teddy (from left to right). Top row: the fixed matching window is marked by red, and the self-adapting matching window is marked by green (from left to right). Bottom row (from left to right): the results by NCC with the fixed matching window, and the results by proposed local stereo matching method with self-adapting matching window. In the NCC case, a mass of obvious outliers occurred in weak-textured regions, discontinuous boundaries, and occlusion areas. The proposed local method has achieved much better results.

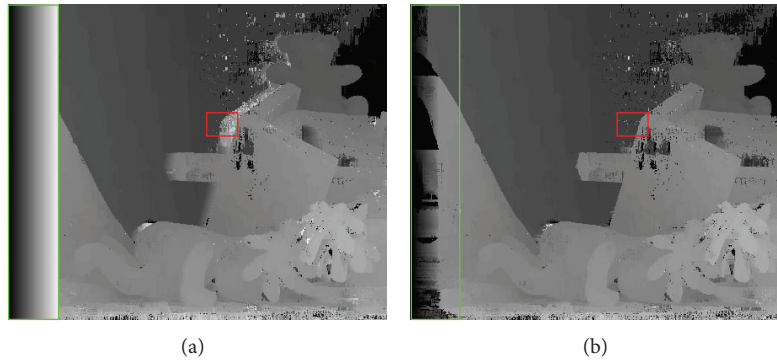


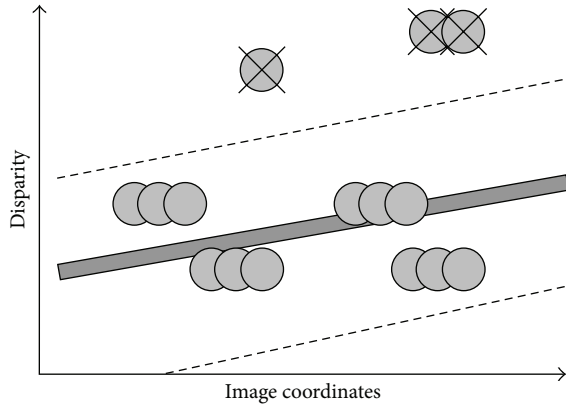
FIGURE 4: Comparison between the local depth result without occlusion region and border of image (ORBI) handling and the one with ORBI handling for the Teddy (from left to right). Left column: the local depth results without ORBI handling. Right column: the local depth results with ORBI handling. The red frame demonstrates the comparison in occlusion region, while the green frames denote the comparison in border of image. The results show that ORBI handling makes the matching costs for the corresponding depths more reliable.

Nevertheless, the smoothness on the boundaries between two adjacent objects will influence the accuracy of the final disparity map. So, we only need to perform the smooth operation in the segments. Compositing the above two kinds of smoothness terms, we propose a new hierarchical smoothness strategy in the identical segment. The new smoothness term is as follows:

$$E_{\text{smooth}} = \sum_{(p_1, p_2) \in N, S(p_1)=S(p_2)} V'_{p_1, p_2} * \min((f(p_1) - f(p_2)), k), \quad (7)$$

where  $S(p)$  is the identification of segment to which the pixel  $p$  belongs,  $V'_{p_1, p_2}$  denotes a new penalty function which enforces a different penalty on the basis of color differences:

$$V'_{p_1, p_2} = \begin{cases} 8\lambda & \text{if } \max(|R(p_1) - R(p_2)|, |G(p_1) - G(p_2)|, |B(p_1) - B(p_2)|) < \varepsilon_1, \\ 2\lambda & \text{if } \max(|R(p_1) - R(p_2)|, |G(p_1) - G(p_2)|, |B(p_1) - B(p_2)|) < \varepsilon_2, \\ \lambda & \text{if } \max(|R(p_1) - R(p_2)|, |G(p_1) - G(p_2)|, |B(p_1) - B(p_2)|) < \varepsilon_3, \\ \sigma & \text{otherwise,} \end{cases} \quad (8)$$



⊗ Deflected pixel

FIGURE 5: Sketch map for deflected pixels.

where  $\sigma$  is a penalty constant,  $\varepsilon_1$ ,  $\varepsilon_2$ , and  $\varepsilon_3$  are several color diversities and  $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$ .

The smooth terms perform different smoothness strategies inside the segments according to the diverse color differences of neighborhoods.

**2.5. Segmentation Term of Selective Planarity.** In this paper, we use the segmentation information to construct the segmentation term for further improving the accuracy of depth estimation. Our segmentation term is different from the higher order term presented by Kohli et al. The higher order term in [15] was originally designed for image segmentation, according to the assumption that the pixels in the same segment should share the same label. However, depth estimation is more likely to satisfy the assumption that the pixels in the same region follow the same distribution such as plane distribution or surface distribution; in other words, the pixels in the same segment could have multiple labels other than only a single label. So, directly making use of Kohli's higher order term for depth estimation is unreasonable.

Obviously, the surface distribution is more representative than the plane distribution because the objects in real scene are more likely composed of irregular surfaces rather than planes. Nevertheless, in this paper the plane distribution is adopted with considering its lower computation complexity and more commonly approximate representativeness. The segments obtained by [17] are further divided into many more subsegments using certain plane distribution. The plane distribution is achieved by plane fitting for the local results. And all the pixels in each subsegment are more likely to share the same label.

Not all the segments are appropriate to enforce the plane distributions. If the plane distribution is employed roughly in those segments which are unable to be represented by plane, the worse influences on resulting depth map would occur.

In this paper, before performing the plane distributions in the segments, we employ a segment classify procedure for every segment using a proposed plane-judge approach as shown in Figure 5.

For instance, the pixel  $p$  is judged as deflected when it meets the condition that  $|f(p) - d^*(p)| > \Phi$ , where  $d^*(p)$  is the disparity value for the pixel  $p$  after plane fitting using the local depths, and  $\Phi$  is a constant that controls the planarity quality of segments.  $p_S$  is a pixel set for all pixels in the segment  $S$ ,  $N_d(f(p_S))$  denotes the number of deflected pixels in the segment  $S$ ,  $N(p_S)$  denotes the number of pixels in the segment  $S$ , and  $\mu \in (0, 1)$  controls the planarity level of the "planar" segment. If  $N_d(f(p_S)) > N(p_S) * \mu$ , we would not construct a homologous segmentation term for the segment  $S$ . Otherwise, the segmentation term would be constructed using the *Robust  $P^n$  Potts model*.

The segmentation function  $E_{seg}(f)$  using the *Robust  $P^n$  Potts model* is defined as

$$E_{seg}(f) = \begin{cases} N_i(f(p_S)) \frac{1}{Q} \gamma_{max} & \text{if } N_i(f(p_S)) \leq Q, \\ \gamma_{max} & \text{otherwise,} \end{cases} \quad (9)$$

where  $N_i(f(p_S))$  denotes the number of pixels in the segment  $S$  not taking the dominant label,  $\gamma_{max}$  is the maximum value of label inconsistency cost, and  $Q$  is the truncation parameter controlling the rigidity of segmentation function. The *Robust  $P^n$  Potts model* proposed by Kohli et al. [15] is shown in Figure 6.

Concrete constructing procedure of segmentation term for each segment is shown in Algorithm 1.

The segmentation terms enforce the plane trends into the segments which can be represented by plane approximately.

**2.6. Energy Minimization Process Based on Graph Cuts.** In order to minimize the global energy function by graph cut, all energy terms of this energy function must be submodular according to [26]. In the light of additive principle, if every term in energy function is submodular, the whole global energy function will be submodular. The unary term, such as data term, is always submodular. The pairwise term, namely, smooth term, also is submodular since it satisfies the inequality  $E^{i,j}(0, 0) + E^{i,j}(1, 1) \leq E^{i,j}(0, 1) + E^{i,j}(1, 0)$ . And from the definition of *Robust  $p^n$  Potts model*, the segmentation term does satisfy the definition of the submodularity on  $F^N$  ( $N \geq 3$ ) [27], if and only if all its projections on two variables are submodular.

According to [26], the segmentation terms can be transformed into sum of pairwise terms:

$$E_{seg} = \min_{m_0, m_1} \left( r_0 (1 - m_0) + \theta_d m_0 \sum_{i \in c_d} w_i (1 - t_i) + r_1 m_1 + \theta_\alpha (1 - m_1) \sum_{i \in c} w_i t_i - \delta \right). \quad (10)$$

Finally, the global energy function is minimized by utilizing the minimum cut on the graph as shown in Figure 7. The minimum cut can be calculated very efficiently using the  $\alpha$ -expansion move algorithm [18].

The detailed minimization process is as shown in Algorithm 2.

```

Input: segment S, constant  $\Phi$ , ratio  $\mu$ 
Initialize  $c \rightarrow c = 0$ 
Do
  If  $(|f(p) - d^*(p)| > \Phi)$ 
     $c \rightarrow c = c + 1$ 
  end If
While  $p \in S$ 
 $N_d(f(p_s)) \rightarrow N_d(f(p_s)) = c$ 
If  $(N_d(f(p_s)) > N(S)^* \mu)$ 
  Reject Constructing Segmentation Term
Else
  Constructing Segmentation Term using the Robust Pn Potts Model.

```

ALGORITHM 1: Constructing procedure of segmentation term.

```

Input: Labeling  $f$ , label sets  $L$ 
Initialize  $finish \rightarrow finish = 1$ 
Repeat
  For  $(\alpha \in L)$ 
    Using one  $\alpha$ -expansion of  $f$ , find  $f^* = \operatorname{argmin} E(f')$  among  $f'$ 
    If  $(E(f^*) < E(f))$ 
       $f \rightarrow f = f^*$ 
       $finish \rightarrow finish = 0$ 
    end If
  end For
Until  $finish = 1$ 
Return  $f$ 

```

ALGORITHM 2: Energy minimization process by graph cut.

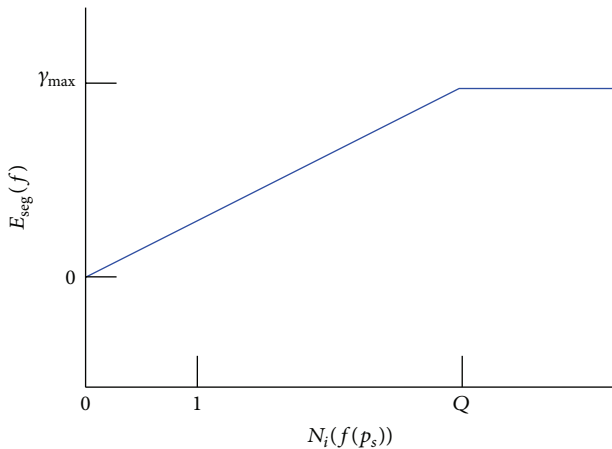


FIGURE 6: Behavior of the *Robust P<sup>n</sup> Potts function*. The figure shows how the higher order cost of the *Robust p<sup>n</sup> Potts function* changes with the number of pixels in the segment not taking the dominant label.

### 3. Experiment

Our program is tested by a personal computer with a 2.20 GHz AMD Dual-Core CPU. All data sets are from [28–31].

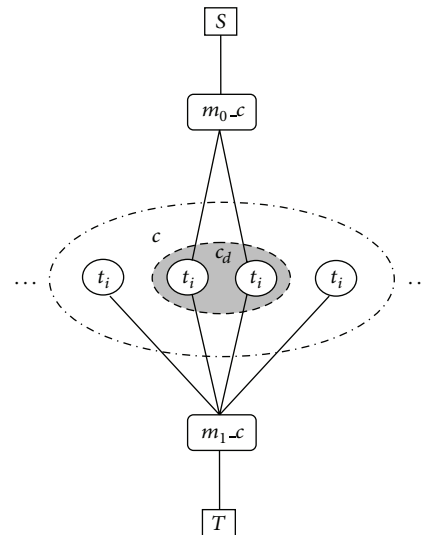


FIGURE 7: The graph for segmentation terms. S is source, T is the sink, and c represents clique; only two auxiliary nodes, namely,  $m_0$  and  $m_1$  are needed for each clique.

For the Middlebury stereo datasets with four stereo test pairs, that is, Tsukuba, Venus, Teddy, and Cones, Table 1 summarizes the quantitative performance of our method and

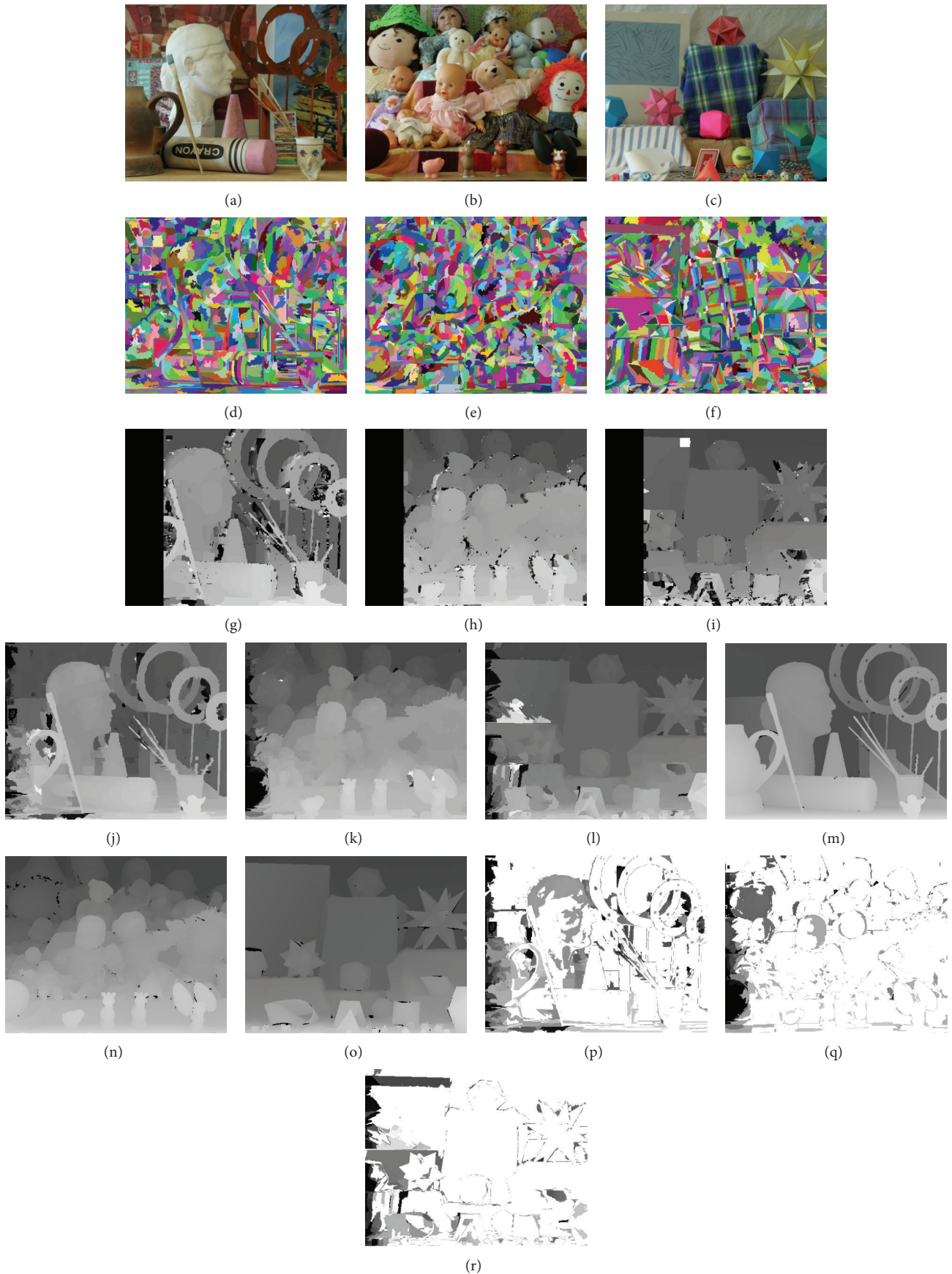


FIGURE 8: The comparison of final depth maps for the Art, Dolls, and Moebius stereo datasets (from top to bottom). First row: the input left images. Second row: the color segmentation results. Third row: the depth results of regular graph cut. Fourth row: the final depth map of our global optimization method. Fifth row: ground truth. Sixth row: "bad pixel" map of matching results.

TABLE 1: Quantitative evaluation results (bad pixels percentage) of different stereo matching methods for the Tsukuba, Venus, Teddy, and Cones stereo test pairs.

Algorithm	Tsukuba	Venus	Teddy	Cones	Average percent of bad pixels
DCBGrid [19]	5.16	1.23	10.8	9.48	6.67
Our method	2.04	1.58	12.60	12.70	7.23
BioPsyASW [20]	4.91	3.41	14.10	11.30	8.43
CSBP [21]	3.84	2.52	17.30	14.20	9.47
Regular GC	4.43	6.56	39.80	59.00	27.45

those of other stereo matching methods, roughly in descending order of overall performance. The comparisons with other approaches show that our global optimization method is fairly competitive with those state-of-the-art approaches.

For sake of declaring the generality of our global optimization method, abundant other stereo image pairs from Middlebury datasets are adopted for depth estimation. Figure 8 illustrates that our global optimization method still achieves satisfactory performance on other stereo images.

#### 4. Conclusion and Discussion

Obviously, the local stereo matching methods based on self-adapting matching window have obtained more outstanding results than fixed matching window based ones. After applying the two proposed matching cost optimization strategies, the local depth results are more accurate in occlusion areas and borders of image. The smooth term makes the surface of segments more close to the real objects. The higher order term, namely, the proposed selective segmentation term, which introduces the plane trend constraint selectively, further enhances the accuracy at object level. In a word, our global optimization method has achieved good performance on Middlebury stereo datasets.

#### Acknowledgments

This work was supported by the National Natural Science Foundation of China (61173096, 61103140, and 60605013), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (12YJC630281), and the Science and Technology Department of Zhejiang Province (2012R10052, R1110679, Y1090592, Y1110688, Y1100824, and Y1110882).

#### References

[1] C. Cattani, R. Badea, S. Chen, and M. Crisan, "Biomedical signal processing and modeling complexity of living systems," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 298634, 2 pages, 2012.

[2] Q. Guan, B. Du, Z. Teng et al., "Bayes clustering and structural support vector machines for segmentation of carotid artery plaques in multicontrast MRI," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 549102, 6 pages, 2012.

[3] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," in *Proceedings of the 8th IEEE International Conference on Computer Vision*, vol. 2, pp. 508–515, July 2001.

[4] P. Zhang, Y. Xu, X. Yang, and L. Traverson, "Multi-scale Gabor phase-based stereo matching using graph cuts," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '07)*, pp. 1934–1937, July 2007.

[5] A. Zureiki, M. Devy, and R. Chatila, "Stereo matching using reduced-graph cuts," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 1, pp. 1237–1240, September 2007.

[6] Y. Liu, X. Lin, X. Chen, and L. Hu, "Different labels in energy minimized via graph cuts for stereo matching," in *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL '08)*, pp. 455–459, Qingdao, China, September 2008.

[7] L. Yu, Q. Liao, and Z. Lu, "A novel method using kde and graph cut in stereo matching," in *Proceedings of the IEEE International Workshop on Imaging Systems and Techniques (IST '09)*, pp. 151–154, May 2009.

[8] M. Bleyer, C. Rother, and P. Kohli, "Surface stereo with soft segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1570–1577, June 2010.

[9] M. Bleyer, C. Rother, P. Kohli, D. Scharstein, and S. Sinha, "Object stereo—joint stereo matching and object segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 3081–3088, June 2011.

[10] D. Wang and K. B. Lim, "A new segment-based stereo matching using graph cuts," in *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT '10)*, vol. 5, pp. 410–416, Chengdu, China, July 2010.

[11] L. Hong and G. Chen, "Segment-based stereo matching using graph cuts," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 1, pp. 74–81, July 2004.

[12] S. Chen, H. Tong, and C. Cattani, "Markov models for image labeling," *Mathematical Problems in Engineering*, vol. 2012, Article ID 814356, 18 pages, 2012.

[13] S. Chen, H. Tong, Z. Wang et al., "Improved generalized belief propagation for vision processing," *Mathematical Problems in Engineering*, vol. 2011, Article ID 416963, 12 pages, 2011.

[14] S. Chen, Y. Wang, and C. Cattani, "Key issues in modeling of complex 3D structures from video sequences," *Mathematical Problems in Engineering*, vol. 2012, Article ID 856523, 17 pages, 2012.

[15] P. Kohli, L. U. Ladický, and P. H. S. Torr, "Robust higher order potentials for enforcing label consistency," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.

[16] Y. Xie, N. Liu, and S. Liu, "Stereo matching using sub-segmentation and robust higher-order graph cut," in *Proceedings of the International Conference on Digital Image Computing Techniques and Applications (DICTA '11)*, Noosa, Australia, December 2011.

[17] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.

[18] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern*



- Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [19] C. Richardt, D. Orr, I. Davies, A. Criminisi, and N. Dodgson, “Real-time spatiotemporal stereo matching using the dual-cross-bilateral grid,” in *Computer Vision—ECCV 2010*, pp. 510–523, 2010.
- [20] L. Nalpantidis and A. Gasteratos, “Biologically and psychophysically inspired adaptive support weights algorithm for stereo correspondence,” *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 457–464, 2010.
- [21] Q. Yang, L. Wang, and N. Ahuja, “A constant-space belief propagation algorithm for stereo matching,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1458–1465, Urbana, Ill, USA, June 2010.
- [22] K. Zhang, J. Lu, and G. Lafruit, “Cross-based local stereo matching using orthogonal integral images,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 1073–1079, 2009.
- [23] G. Egnal and R. P. Wildes, “Detecting binocular half-occlusions: empirical comparisons of five approaches,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1127–1133, 2002.
- [24] K. J. Yoon and I. S. Kweon, “Adaptive support-weight approach for correspondence search,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 650–656, 2006.
- [25] F. Tombari, S. Mattoccia, and L. D. Stefano, “Segmentation based adaptive support for accurate stereo correspondence,” *Advances in Image and Video Technology*, vol. 4872, pp. 427–438, 2007.
- [26] V. Kolmogorov and R. Zabih, “What energy functions can be minimized via graph cuts?” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.
- [27] D. Freedman and P. Drineas, “Energy minimization via graph cuts: settling what is possible,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 2, pp. 939–946, San Diego, Calif, USA, June 2005.
- [28] D. Scharstein and R. Szeliski, “High-accuracy stereo depth maps using structured light,” in *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 195–202, June 2003.
- [29] D. Scharstein and C. Pal, “Learning conditional random fields for stereo,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, Minneapolis, Minn, USA, June 2007.
- [30] H. Hirschmüller and D. Scharstein, “Evaluation of cost functions for stereo matching,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, June 2007.
- [31] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *International Journal of Computer Vision*, vol. 47, no. 1–3, pp. 7–42, 2001.