Self-Adaptive Matching in Local Windows for Depth Estimation

Haiqiang Jin, Sheng Liu *, Xuhua Yang and Shengyong Chen
College of Computer Science & Technology
Zhejiang University of Technology
Hangzhou 310023, China
E-mail: edliu@zjut.edu.cn

KEYWORDS
Depth Estimation, Self-Adapting Matching Window, Refining Algorithm.

ABSTRACT

This paper proposes a novel local stereo matching approach based on self-adapting matching window. We improve the accuracy of stereo matching in 3 steps. First, we integrate shape and size information, and construct robust minimum matching windows by applying a self-adapting method. Then, two matching cost optimization strategies are employed for handling both occlusion regions and image borders. Last, we perform a refinement algorithm for obtaining more accurate depth map. Experiment results on the Middlebury stereo image pairs prove that the proposed matching method performs equally well in comparison with other state-of-the-art local approaches.

INTRODUCTION

In machine vision field, depth estimation is a hot research direction all the time. The depth information is able to be applied for auto reversing system, three-dimensional scene reconstruction, obstacle avoidance, and so on. At present, the accuracy of depth results obtained by the local stereo matching methods (Xu et al. 2002; Yoon and Kweon 2006) have been approximated to the one of global approaches (Kohli et al. 2008; Bleyer et al. 2010; Bleyer et al. 2011; Wang and Lim 2010; Li and Chen 2004). And the local methods consume less time than global ones. In the beginning, all kinds of local stereo matching methods (Birchfield and Tomasi 1998; Gerrits and Bekaert 2006; Zhang et al. 2009; Lu et al. 2008; Chen et al. 2012a; Chen et al. 2012b) were proposed for depth estimation. In order to reduce the image ambiguity, early local methods (Zhang et al. 1995; Kwok et al. 2011; Scharstein et al. 2001; Stefano et al. 2004) usually made use of a fixed matching window to aggregate the support from the neighboring pixels within the matching window. Later, because of the known assumption that pixels with similar intensity within a constrained window have similar depth, the matching window was required to adapt its shape and size for the more accurate depth estimation near depth discontinuities (Chen et al. 2011; Chen et al. 2008).

In many fixed matching window based local methods, both SAD (sum of absolute difference) and SSD (sum of square difference) computed the dissimilarity between each pair of matching pixels, while NCC (normalized cross correlation) calculated the similarity. The self-adapting matching window based local approach proposed by (Ke Zhang et al. 2009) was similar with SAD, both of them adopted the absolute difference computation of color information. Nevertheless, this approach had neither considered the shape and size information of self-adapting matching window nor dealt with the situation of smallest matching window properly. In the case of the smallest matching window, the shape and size of self-adapting matching window for pixel will be shrank to the shape and size of pixel, therefore leading to the invalidation of matching window. For avoiding above situation, Ke Zhang et al. performed a minimum matching window of 3x3 for the more robust correspondence matching. Other than familiar global approaches (Wang and Lim 2010; Kolmogorov and Zabih 2001; Zhang et al. 2007) constructing an occlusion term to handle the occlusion regions, we solve the occlusion problem via using a matching cost optimization strategy.

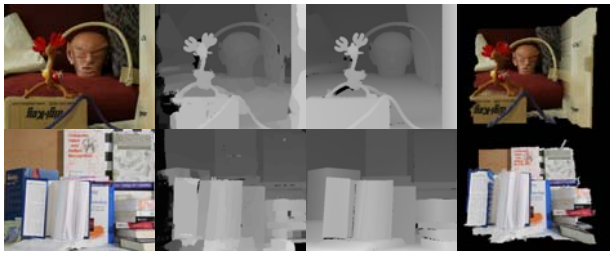

Figure 1: The Reconstructed Three-Dimensional Maps and Dense Depth Maps for the Reindeer and Dolls Stereo Datasets (from top to bottom). From left to right: the input left images, our depth maps, ground truth and reconstructed three-dimensional results. Compared with the ground truth, our depth results obviously acquire most details of the scene with relatively high accuracy.

This paper presented a novel local stereo matching approach for depth estimation. Our method mainly makes the following contributions. Based on original self-adapting local matching method, our approach not only considers the shape and size of matching window, but also improves the method constructing the minimum matching window. And both occlusion regions and border of image problems are solved rely on two matching cost optimization strategies. At last, a new refining method is proposed to calculate the final depth results.

Experimental results on the Middlebury data sets in figure 1 have shown that the proposed local approach is

Proceedings 27th European Conference on Modelling and Simulation ©ECMS Webjørn Rekdalsbakken, Robin T. Bye, Houxiang Zhang (Editors)
ISBN: 978-0-9564944-6-7 / ISBN: 978-0-9564944-7-4 (CD)

able to obtain satisfactory depth maps and is competitive with the state-of-the-art algorithms.

PROPOSED LOCAL STEREO APPROACH

Algorithm overview

The rough procedure of proposed local stereo approach is divided into three steps: First, the raw matching costs are computed relying on the improved local method. Second, we improve the raw matching costs using two matching cost optimization strategies. Third, the final depth map is obtained using a new refining approach. The whole process of our method is illustrated in figure 2.

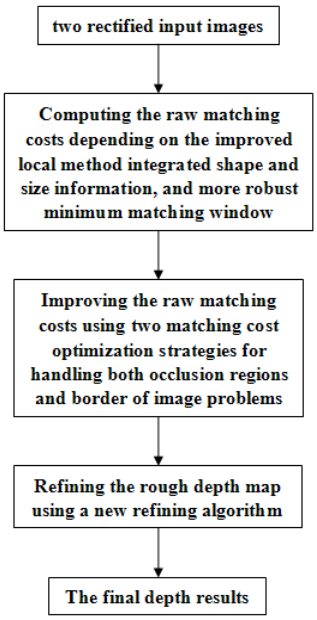

Figure 2: Flow Chart of Our Algorithm.

Improved Local Method based on Self-Adapting Window

In general local stereo matching methods, a fixed matching window is predefined for depth estimation. The matching window for certain pixel is to aggregate the support from neighboring pixels with the same depth within the matching window, but it is not certain that all the pixels in the fixed matching window have the same depth. For example, in the depth result based on NCC with a fixed window, as shown in Fig. 3, there are a mass of noisy depths in weak-textured regions, fuzzy depths in discontinuous boundaries and depths at occlusion areas. Obviously, the proposed approach based on self-adapting window achieves more accurate results shown in figure 3.

Based on the assumption that pixels with similar color within a constrained window have similar depth, 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. 2009) based on self-adapting matching window. Three improvements are made on the basis of original approach. Firstly, we add the shape and size information of matching window for each pixel, which will further improve the reliability of matching costs. Secondly, a dynamical argument strategy for minimum matching window is presented for more robust correspondence matching. Thirdly, we enforce a replacement strategy for occlusion regions and a suboptimum strategy for border of image.

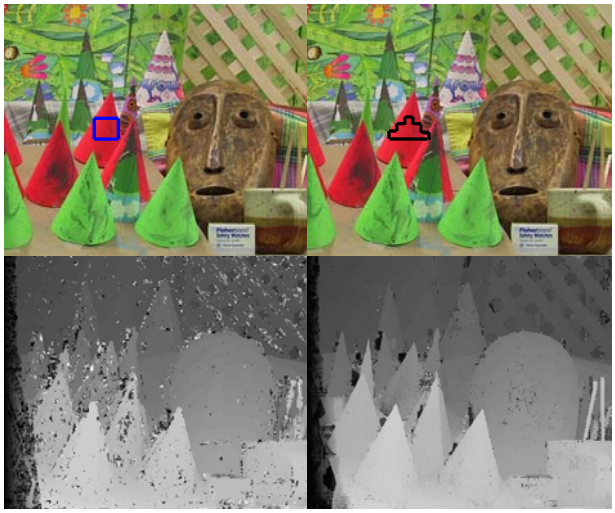

Figure 3: Comparison of Depth Results between Fixed Matching Window and Self-Adapting Matching Window for the Cone (from left to right). Top row: fixed matching window marked by blue, self-adapting matching window marked by black (from left to right). Bottom row: the depth map of NCC with fixed matching window, the depth result of proposed approach with self-adapting matching window (from left to right).

The proposed approach is roughly divided into the following four steps. First, we will determine a self-adapting cross for every pixel in stereo images. Then, the self-adapting window for every pixel is estimated using the cross in stereo images. Thirdly, the matching costs are computed using the self-adapting window. Last, we handle the occlusion regions and border of image.

Step 1. Determining an upright cross for every pixel in stereo images.

$\{h_p^-, h_p^+, v_p^-, v_p^+\}$ are adopted to represent the left, right, up and bottom arm length for the pixel p respectively. A color similarity computation is performed for a consecutive set of pixels which reside on the left horizontal side of the pixel p, L is the preset maximum arm length which controls the size of maximum matching window for the pixel p. I_c denotes the intensity of corresponding color component and τ controls the confidence degree of color similarity. The algorithm for acquiring $\{h_p^-, h_p^+, v_p^-, v_p^+\}$ is as shown in table 1.

In the light of the attained quadruple $\{h_p^-, h_p^+, v_p^-, v_p^+\}$, $H(p)$ and $V(p)$ are attained for each pixel p in stereo

images. *H(p)* stands for the horizontal integral of the pixel *p* as well as the vertical integral of the pixel *p* represented by *V(p)*, and they jointly decide the self-adapting cross for the pixel p shown in figure 4.

Table 1: Estimation of $\{h_p^-, h_p^+, v_p^-, v_p^+\}$.

Algorithm for estimating $\{h_p^-, h_p^+, v_p^-, v_p^+\}$

Input: length L, constant τ, augment τ_{arg}
Initialize $h_p^- \rightarrow h_p^- = 0$, $h_p^+ \rightarrow h_p^+ = 0$,
$v_p^- \rightarrow v_p^- = 0$, $v_p^+ \rightarrow v_p^+ = 0$
Repeat
For i = 1 to L
If $(\max_{c\in\{R,G,B\}}(\lvert I_c(p) - I_c(p_i)\rvert) > \tau)$
Break
end If
end For
$h_p^- \rightarrow h_p^-$ = i-1
For i = 1 to L
If $(\max_{c\in\{R,G,B\}}(\lvert I_c(p) - I_c(p_i)\rvert) > \tau)$
Break
end If
end For
$h_p^+ \rightarrow h_p^+$ = i-1
For i = 1 to L
If $(\max_{c\in\{R,G,B\}}(\lvert I_c(p) - I_c(p_i)\rvert) > \tau)$
Break
end If
end For
$v_p^- \rightarrow v_p^-$ = i-1
For i = 1 to L
If $(\max_{c\in\{R,G,B\}}(\lvert I_c(p) - I_c(p_i)\rvert) > \tau)$
Break
end If
end For
$v_p^+ \rightarrow v_p^+$ = i-1
$\tau \rightarrow \tau = \tau + \tau_{arg}$
Until $(h_p^- + h_p^+) > T_{arms}$ *and* $(v_p^- + v_p^+) > T_{arms}$

$$\begin{cases} H(p) = \{(x,y) \mid x \in [x_p - h_p^-, x_p + h_p^+], y = y_p\} \\ V(p) = \{(x,y) \mid x = x_p, y \in [y_p - v_p^-, y_p + v_p^+]\} \end{cases} \quad (1)$$

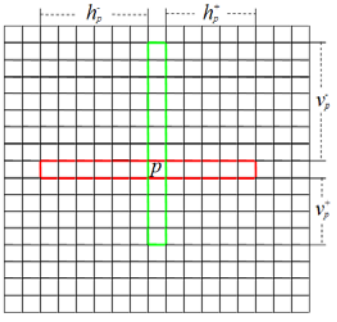

Figure 4: Sketch Map for *H(p)* & *V(p)*. Red frame represents *H(p)*, the length of *H(p)* is $h_p^- + h_p^+ + 1$. And green frame represents *V(p)*, the length of *V(p)* is $v_p^- + v_p^+ + 1$.

Step 2. Estimating the Self-adapting Window for every pixel using the Cross in stereo images.

Given the self-adapting cross for each pixel, we can readily construct a self-adapting matching window *U(p)* for the pixel *p*. The key process is to model the matching window *U(p)* as an area integral of multiple horizontal integrals *H(q)*, sliding along the vertical segment V(p) of the pixel p,

$$U(p) = \bigcup_{q \in V(p)} H(q) \quad (2)$$

where q is a pixel located on the vertical integral V(p).

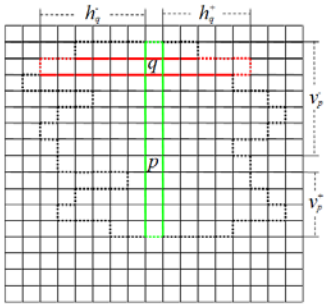

Figure 5: Self-Adapting Matching Window for the Pixel *p*. Green frame represents *V(p)*, and red frame represents *H(q)*. Dotted box stands for the self-adapting region.

Step 3. Computing the Matching Costs using the Self-adapting Window.

We symmetrically consider both self-adapting matching window *U*(p) and *U′*(*p′*) depending on the pixel *p* and *p′* respectively so as to achieve reliable matching cost aggregation. Here, $p' = (x_p - d, y_p)$ is the corresponding pixel in the right image for $p = (x_p, y_p)$ with depth *d* in the left image. The matching cost $C_d(p)$ between the pixel *p* and *p′* is computed as follows:

$$C_d(p) = \frac{1}{\|U_d(p)\|} * \sum_{t \in U_d(p)} e_d(t) * (\log\theta + 1) \quad (3)$$

where $U_d(p) = \{(x,y) \mid (x,y) \in U(p)\}$, $(x-d, y) \in U'(p')$, $\theta = \|U(p)\| / \|U_d(p)\|$, $e_d(t)$ denotes the raw matching cost for the pixel *t* with depth *d* and $U_d(p)$ is the combined matching window which only contains these valid pixels. $\|U_d(p)\|$ denotes the number of pixels in $U_d(p)$, used for normalizing the aggregated matching cost $\sum_t \in U_d(p) e_d(t)$. The raw matching cost is computed from a pair of corresponding pixels, for example, the matching cost of *t* in the left image and *t′* in the right image with the depth value *d* is computed as

$$e_d(t) = \min\left(\sum_{c \in \{R,G,B\}} \lvert I_c(t) - I_c'(t') \rvert, T \right) \quad (4)$$

where T controls the truncation limit of the matching cost.

Step4. Handling the Occlusion Regions and Border of Image.

Being inspired by five major approaches introduced by (Egnal and Wildes 2002), 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 depth, the matching costs for occlusion pixels are capable of being replaced with ones for "corresponding" pixels.

For instance, $d(p)$ is the depth for pixel $p = (x_p, y_p)$ in left image, and $d'(p')$ is the depth for pixel $p' = (x_p - d(p), y_p)$ in 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 ones for the pixel p' in left image are replaced with ones for the pixel p′ in right image.

Neither estimating two depth maps for left-right consistency check (Yoon and Kweon 2006; Tombari et al. 2007) nor applying a simple border extrapolation step, we adopt a suboptimum strategy for 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 can not be achieved by making use of the corresponding pixels.

$$\hat{d} = \arg\min_{d\in[d_{\min},d_{\max}],(x_p-d)>0,\, d\neq d^*} C_d(p) \tag{5}$$

where $\hat{d}$ is the suboptimum label we need, d^* is the optimal label is computed as follows:

$$d^* = \arg\min_{d\in[d_{\min},d_{\max}],(x_p-d)>0} C_d(p) \tag{6}$$

At last, we use $C_d(p)$ as the matching cost for pixel p when $(x_p - d(p)) < 1$. The handling for border of image will work aftering the global optimization process.

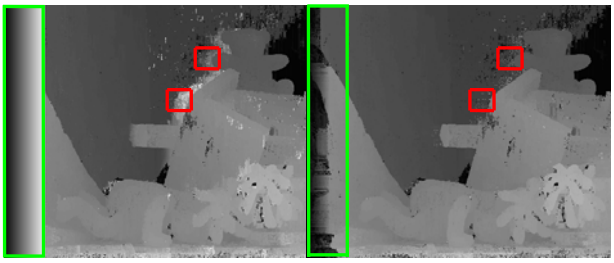

Figure 6: Comparing between the Rough Result without ORBI (Occlusion Regions and Border of Image) Handling and the Rough Result with ORBI Handling for the Teddy (from left to right). Left column: the rough result without ORBI handling. Right column: the rough result with ORBI handling. Occlusion regions are marked by red frames, border of image is marked by green frame.

Refining Algorithm

Although the raw result obtained by proposed approach after the "Winner-Take-All" is of relative good accuracy, there existed many noisy depth areas as shown in figure 7. Therefore we still require further refining the raw result for more accurate result. This paper utilizes a new refining algorithm to achieve this goal. Figure 7 demonstrates that the refining algorithm eliminates a mass of outliers.

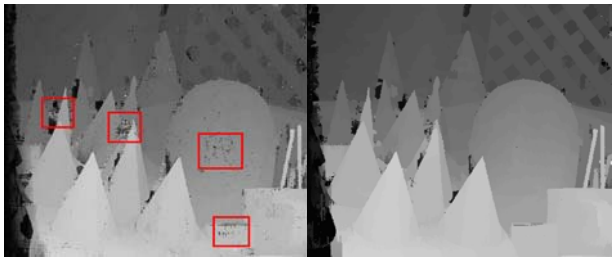

Figure 7: The Comparison between the Depth Maps (before and after refining) for the Cone. It is obvious that the refining method eliminates amounts of noisy depths. Left: the raw result before refining with numerous noisy regions marked by red frame. Right: the result after refining.

Table 2: Refining the Rough Depth Map

Algorithm for refining the rough depth map

Input: Input image P, minimum depth d_{min},
maximum depth d_{max}
Initialize $cost \rightarrow cost = 0$
Do
For $d = d_{min}$ to d_{max}
Do
If $(d'(p)==d)$
$cost \rightarrow cost=cost+1/\bar{C}_{p'}(d)$
end If
While $p' \in U(P)$
$\Phi_p(d) = cost$
$cost \rightarrow cost = 0$
end For
While $p \in P$
Do
For $d = d_{min}$ to d_{max}
If $(\Phi_p(d)>\Phi_p(d\_max))$
$d\_max=d$
end If
end For
$d'(p) = d\_max$
While $p \in P$

In this paper, we mainly refer to the refining approach proposed by (Lu et al. 2008) with a local high-confidence voting scheme. They make a statistic function $\varphi_p(d)$ for the number of pixels with the depth d in the self-adapting neighbourhood $U(p)$ of the pixel p, the maximum of φ_p corresponds to a statistically optimal depth value d^*_p. Accordingly, the depth of the pixel p

Table 3: 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
VarMSOH(Ben-Ari and Sochen 2010)	3.60	0.49	10.10	8.20	5.60
BioPsyASW(Nalpantidis and Gasteratos 2010)	4.91	3.41	14.10	11.30	8.43
Our Method	5.47	4.05	13.00	11.50	8.51
CSBP(Yang et al. 2010)	3.84	2.52	17.30	14.20	9.47
Regular GC	4.43	6.56	39.80	59.00	27.45

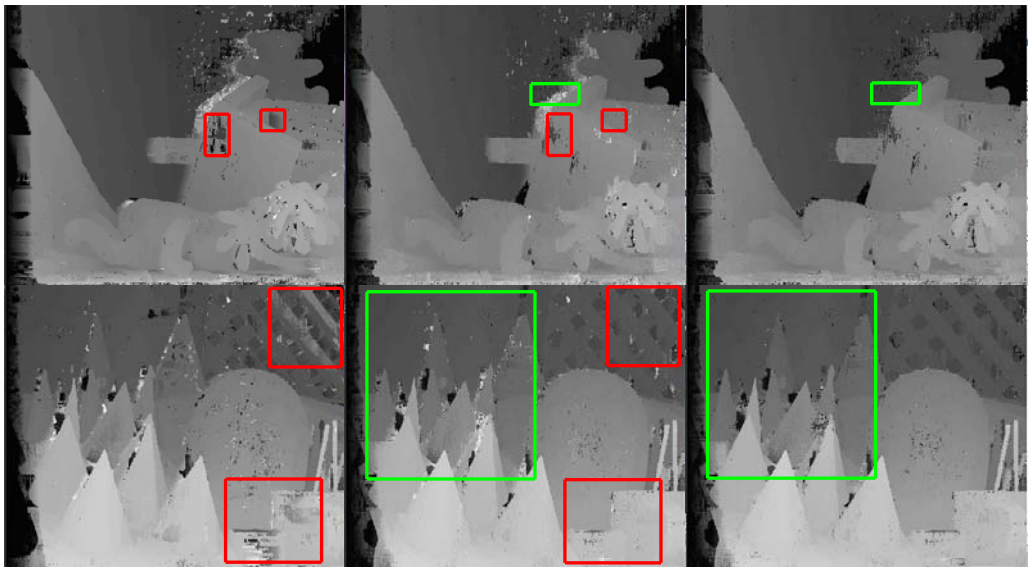

Figure 8: The Evolutive Local Results for the Middlebury Stereo Datasets Teddy and Cones (from top to bottom). First column: the original depth results in (Zhang et al. 2009). Second column: the depth results with integrating size information and improved minimum matching window. Third column: the final local results using the proposed local matching approach. Red frame indicates that the results of second column are more accurate than these of first column. And green frame shows that the results of third column are more competitive in occlusion regions than ones of second column.

after refining, d^*_p, is decided as

$$d_p^* = \arg\max_d \varphi_p(d), \quad d \in [d_{\min}, d_{\max}] \tag{7}$$

After the local matching approach proposed, the matching cost $C_d(p)$ for the pixel p is obtained, and then we can achieve the local depth estimate $d(p)$. Taking advantage of the matching cost $C_d(p)$, we compute a weighted average of matching costs $\bar{C}_p$ for the pixel p in its self-adapting neighbourhood $U(p)$, is followed as

$$\bar{C}_p(d) = \frac{\sum_{p' \in U(p)} C_d(p')}{\|U(p)\|}, \quad d \in [d_{\min}, d_{\max}] \tag{8}$$

where $p' \in U(p)$ is a pixel in the self-adapting neighborhood $U(p)$, $\|U(p)\|$ denotes the total number of pixels in the self-adapting matching window $U(p)$. The algorithm for computing final depth map $d'(p)$ is as shown in table 2.

EXPERIMENTS

In this paper, the performance on weakly textured regions has been mended in some way by integrating shape and size information and improving the minimum matching window. And, a new process has been presented to handle the occlusion pixels. In figure 8, the fence parts of Cones are restored obviously in red frame. And, occlusion areas outside the eave of Teddy are more accurate in green frame after proposed occlusion handling.

All of test sets are from the Middlebury stereo datasets (Scharstein et al. 2001; Scharstein and Szeliski 2003; Scharstein and Pal 2007; Hirschmuller and Scharstein 2007). For the standard Middlebury image pairs with four test pairs, i.e., Tsukuba, Venus, Teddy, and Cones, Table 3 summarizes the quantitative performance of our method and those of other stereo matching methods, roughly in descending order of overall performance. The final depth results prove that our approach is able to competitive with those state-of-the-art approaches.

CONCLUSION

Our local stereo matching approach adopting the shape-size information and the more robust minimum matching window has more excellent performance on weak-textured regions. Meanwhile, the conventional occlusion regions and border of image problems are solved quite successfully depending on two matching cost optimization strategies. And after the handling of new refinement method, the raw local depth result is transformed into the final depth map with high accuracy. In general, our local approach has been proved to be able to obtain good performance on the Middlebury stereo test sets.

REFERENCES

Ben-Ari, R. and N. Sochen. 2010. “Stereo matching with Mumford-Shah regularization and occlusion handling”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.32, No.11, 2071-2084.

Birchfield, S. and C. Tomasi. 1998. “A Pixel Dissimilarity Measure That Is Insensitive to Image Sampling”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20, No.4, 401-406.

Bleyer, M.; C. Rother; and P. Kohli. 2010. “Surface Stereo with Soft Segmentation”. In *Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition*. 1570-1577.

Bleyer, M.; C. Rother; P. Kohli; D. Scharstein; and S. Sinha. 2011. “Object Stereo — Joint Stereo Matching and Object Segmentation”. In *Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition*. 3081-3088.

Chen, S.; H. Tong; C. Cattani, "Markov models for image labeling", Mathematical Problems in Engineering, Vol. 2012, Article ID 814356, 2012, 18 pages. doi:10.1155/2012/814356.

Chen, S.Y.; Z.J. Wang, "Acceleration Strategies in Generalized Belief Propagation", IEEE Transactions on Industrial Informatics, Vol. 8, No. 1, 2012, pp. 41-48.

Chen, S.Y.; Y.F. Li, "Determination of Stripe Edge Blurring for Depth Sensing", IEEE Sensors Journal, Vol. 11, No. 2, Feb. 2011, pp. 389-390.

Chen, S.Y.; Y.F. Li; J.W. Zhang, "Vision Processing for Realtime 3D Data Acquisition Based on Coded Structured Light", IEEE Transactions on Image Processing, Vol. 17, No. 2, Feb. 2008, pp. 167-176.

Egnal, G. and R.P. Wildes. 2002. “Detecting Binocular Half-Occlusions: Empirical Comparisons of Five Approaches”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No.8(Aug), 1127-1133.

Gerrits, M. and P. Bekaert. 2006. “Local Stereo Matching with Segmentation-based Outlier Rejection”. In *Proceedings of the 3rd Canadian Conference on Computer and Robot Vision*. 66.

Hirschmuller, H. and D. Scharstein. 2007. “Evaluation of cost functions for stereo matching”. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 1-8.

Kohli, P.; L. Ladicky; and P. Torr. 2008. “Robust Higher Order Potentials for Enforcing Label Consistency”. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 1-8.

Kolmogorov, V. and R. Zabih. 2001. “Computing Visual Correspondence with Occlusions using Graph Cuts”. In *Proceedings of Eighth IEEE International Conference on Computer Vision*. Vol.2, 508-515.

Kwok, N.M.; X. Jia, et al., "Visual impact enhancement via image histogram smoothing and continuous intensity relocation", *Computers & Electrical Engineering*, Vol. 37, No. 5, Sep. 2011, pp. 681-694.

Li, H. and G. Chen. 2004. “Segment-based Stereo Matching Using Graph Cuts”. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol.1, 74-81.

Lu, J.; G. Lafruit; and F. Catthoor. 2008. “Anisotropic local high-confidence voting for accurate stereo correspondence”. In *Proceedings of SPIE-IS&T Electronic Imaging*. Vol.6812.

Nalpantidis, L. and A. Gasteratos. 2010. “Biologically and psychophysically inspired adaptive support weights algorithm for stereo correspondence”. *Robotics and Autonomous Systems*.

Scharstein, D.; R. Szeliski; and R. Zabih. 2001. “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms”. In *Proceedings of IEEE Workshop on Stereo and Multi-Baseline Vision*. 131-140.

Scharstein, D. and R. Szeliski. 2003. “High-accuracy stereo depth maps using structured light”. In *Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol.1, 195-202.

Scharstein, D. and C. Pal. 2007. “Learning conditional random fields for stereo”. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 1-8.

Stefano, L.D. ; M. Marchioni et al. 2004. “A fast area-based stereo matching algorithm”. *Image and vision computing*, No.22, 983-1005.

Tombari, F.; S. Mattoccia; and L.D. Stefano. 2007. “Segmentation based adaptive support for accurate stereo correspondence” In *Proceedings of Pacific-Rim Symp. Image Video Technol*. 427–438.

Wang, D. and K.B. Lim. 2010. “A New Segment-based Stereo Matching using Graph Cuts”. In *Proceedings of 2010 3rd IEEE International Conference on Computer Science and Information Technology*. Vol.5, 410-416.

Xu, Y.; D. Wang; T. Feng; and H.Y. Shum. 2002. “Stereo computation using radial adaptive windows”. In *Proceedings of 16th International Conference on Pattern Recognition*. Vol.3, 595-598.

Yang, Q.; L. Wang; and N. Ahuja. 2010. “A constant-space belief propagation algorithm for stereo matching”. In *Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition*. 1458-1465.

Yoon, K.J. and I.S. Kweon. 2006. “Adaptive support-weight approach for correspondence search”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28, No.4 (Apr), 650-656.

Zhang, K.; J. Lu; and G. Lafruit. 2009. “Cross-Based Local Stereo Matching Using Orthogonal Integral Images”. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.19, No.7, 1073-1079.

Zhang, Z.; R. Deriche et al. 1995. “A robust technique for matching two uncalibrated images fllmugh the recovery of the unknown epipolar geometIy”. *Artificial Intelligence*, Vol.78, 87-119.

Zhang, P.; Y. Xu; X. Yang; and L. Traversoni. 2007. “Multi-Scale Gabor Phase-Based Stereo Matchingusing G raph Cuts”. In *Proceedings of 2007 IEEE International Conference on Multimedia and Expo*. 1934-1937.

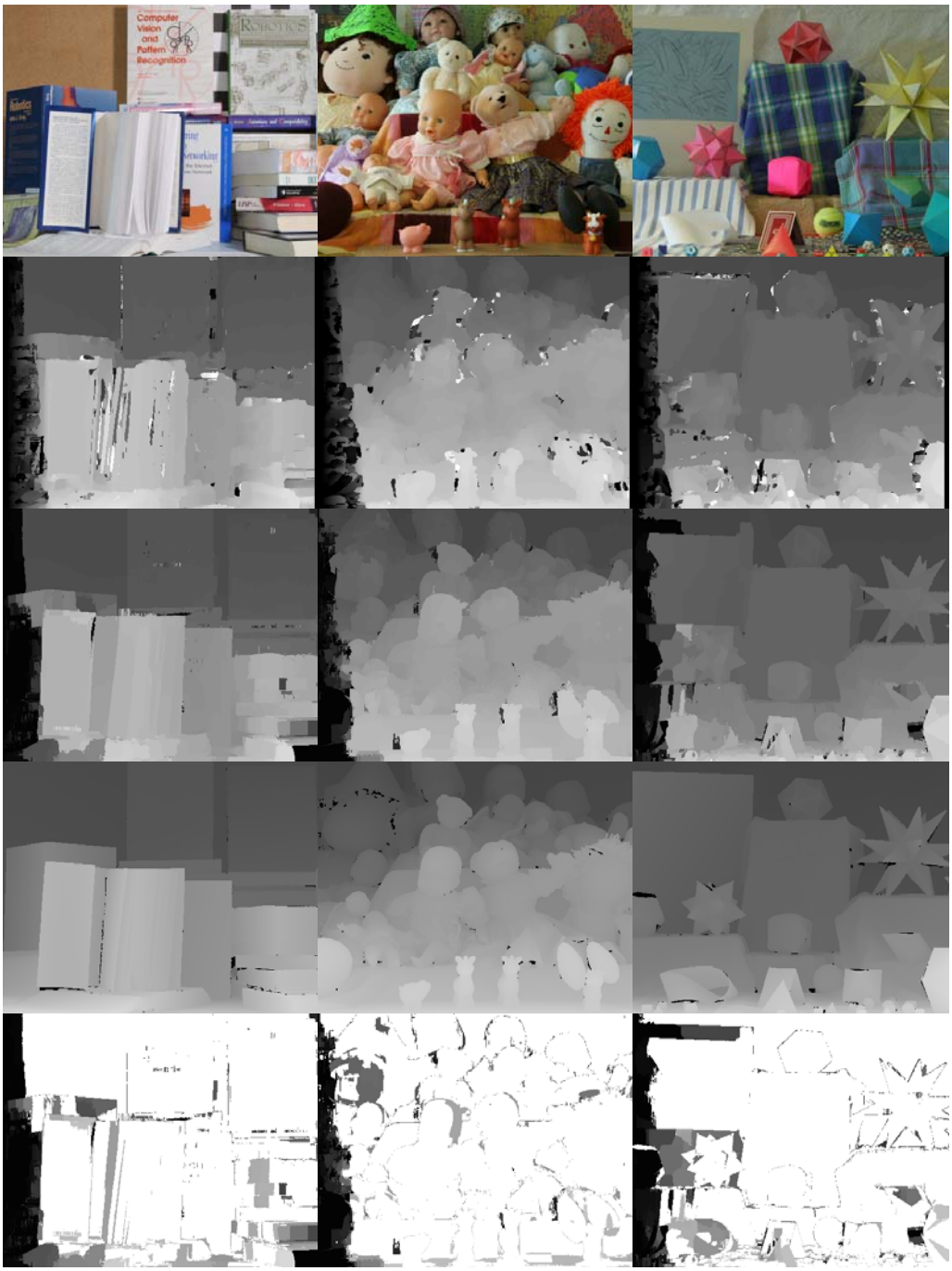

Figure 9: The Comparison of Final Depth Results for the Books, Dolls, and Moebius Stereo Image Pairs (from left to right). Rows from top to bottom: the input left images, results of NCC, final depth results by proposed method, ground truth and "bad pixel" matching results.