

Super-resolution Reconstruction for Binocular 3D Data

Wei-Tsung Hsiao, Jing-Jang Leou, and Han-Hui Hsiao

Department of Computer Science and Information Engineering

National Chung Cheng University

Chiayi 621, Taiwan

{hwt100m, jjleou, hhh95p}@cs.ccu.edu.tw

Abstract—In this study, a super-resolution reconstruction approach for binocular 3D data is proposed. The aim is to obtain the high-resolution (HR) disparity map from a low-resolution (LR) binocular image pair by super-resolution reconstruction. The proposed approach contains five stages, i.e., initial disparity map estimation using local aggregation, disparity plane model computation, global energy cost minimization, HR disparity map composition by region-based fusion (selection), and fused HR disparity map refinement. Based on the experimental results obtained in this study, in terms of PSNR and bad pixel rate (BPR), the final HR disparity maps by the proposed approach are better than those by four comparison approaches.

Keywords—super-resolution reconstruction, low-resolution (LR) binocular image pair, high-resolution (HR) disparity map, region-based fusion (selection), global energy cost minimization.

I. INTRODUCTION

To construct 3D disparity (or depth) maps from binocular image pairs is an important issue in computer vision. To obtain high-resolution (HR) disparity maps, high quality digital cameras can be used to capture HR binocular image pairs and then compute HR disparity maps by stereo matching. However, if only low resolution (LR) binocular image pairs are available, HR disparity maps may be obtained by super-resolution reconstruction [1].

The disparity map of a binocular image pair can be determined by stereo matching [1]. For local stereo matching, the disparity of a pixel may be determined by color or intensity values within a fixed or adaptive support window [2] or by assigning adaptive support weights to different pixels based on some measure [3]. For global stereo matching, the disparities of all pixels in an image pair can be determined by using some type of global optimization [4]. Additionally, region-based stereo matching methods usually produce accurate and reliable disparity maps, whereas they are usually computationally expensive [5].

To upsample disparity (depth) maps, we may perform image super-resolution (SR) or image fusion. For image super-resolution, Schuon et al. [6] proposed a 3D depth image super-resolution approach, named LidarBoost, which combines several low-resolution noisy depth images into a high-

resolution depth image, using an optimization framework. Gevrekci and Pakin [7] proposed a 3D depth image super-resolution approach using projection onto convex sets (POCS) reconstruction. Zhu et al. [8] developed a reliable fusion approach of time-of-flight (ToF) depth and stereo geometry for high quality depth maps. Zhang et al. [9] combined ToF depth and stereo matching to get high quality depth maps using the winner-takes-all (WTA) strategy. Li et al. [10] presented a joint example-based depth map super-resolution approach. They learnt a mapping function from a set of training samples and enhanced the depth map resolution via sparse coding. Park et al. [11] performed high quality upsampling on depth maps using MRF optimization and nonlocal means (NLM) filtering to preserve fine details and local structures. Lu et al. [12] presented an MRF-based approach for depth map super-resolution and enhancement. Kopf et al. [13] described the joint bilateral upsampling (JBU) filter, a modified bilateral filter, and its applications on depth map upsampling, tone mapping, ... Garcia et al. [14] presented the pixel weighted average strategy (PWAS) for depth sensor data fusion, based on joint bilateral upsampling. Garcia et al. [15] presented an adaptive multi-lateral filter for real-time depth sensor data fusion.

In this study, an SR reconstruction approach for binocular 3D data is proposed. The proposed approach contains five stages, i.e., initial disparity map estimation using local aggregation, disparity plane model computation, global energy cost minimization, HR disparity map composition by region-based fusion (selection), and fused disparity map refinement.

This paper is organized as follows. The proposed SR reconstruction approach for binocular 3D data is described in Section II. Experimental results are presented in Section III, followed by concluding remarks.

II. PROPOSED APPROACH

A. Initial Disparity Map Estimation Using Local Aggregation

The framework of the proposed approach is illustrated in Fig. 1. The proposed approach contains five stages, i.e., initial disparity map estimation using local aggregation, disparity plane model computation, global energy cost minimization, HR disparity map composition by region-based fusion (selection), and fused disparity map refinement. The input is an LR

+ This work was supported in part by National Science Council, Taiwan, Republic of China under Grants NSC 102-2221-E-194-028-MY2 and NSC 102-2221-E-194-041-MY3.

binocular image pair and the output is the desired HR disparity map.

Many local stereo matching methods usually estimate the disparity map of a binocular image pair by window-based cost aggregation. Two main issues of local stereo matching include matching cost and matching support window. The matching cost may simply be the sum of absolute differences (SAD) or normalized cross-correlation (NCC). Here, the left initial LR disparity map $D_{L,low}^{initial}$ for the LR binocular image pair is obtained by the AD-census cost combining the absolute differences (AD) and the census transform [16-17]. The AD cost is defined as the color difference between pixel $p = (x, y)$ in LR left view $I_{L,low}$ and the corresponding pixel $p' = (x + d, y)$ in LR right view $I_{R,low}$ in RGB color space, i.e.,

$$C_{AD}(p, d) = \sum_{i \in \{R, G, B\}} |I_{L,low}^i(p) - I_{R,low}^i(p')|, \quad (1)$$

where d denotes the disparity value within the range $[d_{min}, d_{max}]$.

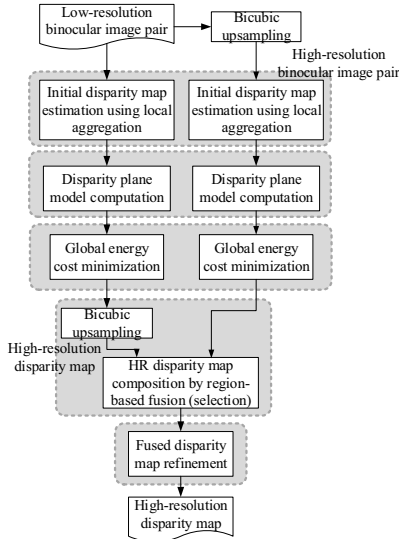


Fig. 1. The framework of the proposed approach.

The census transform encodes local image structure with relative orderings of pixel intensities. As the illustrated example shown in Fig. 2, the census transform preserves the spatial distribution of ranks by encoding the pixels in the sliding window as a binary bitstring via thresholding other pixels by the central pixel value. Note that the binary bitstring is formed by raster scanning (excluding the central pixel). The census transform $C_{CENSUS}(\cdot)$ is defined as the Hamming distance between the bitstring of pixel p and that of the corresponding pixel p' . The AD-census cost $C(\cdot)$ is defined as $C(p, d) = \min(C_{AD}(p, d), \gamma_{AD}) + \min(C_{CENSUS}(p, d), \gamma_{CENSUS})$, (2) where parameters γ_{AD} and γ_{CENSUS} are empirically set to 60 and 20, respectively.

The second main issue of local stereo matching is to determine a well-suited support window for each pixel. A support window should be large enough to capture sufficient intensity variation for handling regions of poor texture,

whereas a support window should be small enough to exclude pixels of different disparities, avoiding crossing disparity discontinuities. In this study, modified horizontal line segment [16] and cross-based cost aggregation [17] are sequentially used to obtain the adaptive support window for each pixel p . For an image, as the illustrated example shown in Fig. 3, a horizontal line segment starting at pixel r is determined by using both color and length constraints as

$$\arg \max_{i=R, G, B} |L_r^i - I^i(\bar{r})| < \tau_c, \quad (3)$$

$$dist(r, \bar{r}) < \tau_l, \quad (4)$$

where \bar{r} is the estimated termination pixel of the horizontal line segment, $dist(\cdot)$ is a distance function, parameters τ_c and τ_l are empirically set to 20 and 30, respectively, and L_r^i denoting the mean color value of the horizontal line segment L_r is defined as

$$L_r^i = \frac{\sum_{q \in L_r} I^i(q)}{length(L_r)}, \quad i \in \{R, G, B\}, \quad (5)$$

where $length(L_r)$ is the length of L_r . Note that the colors of the pixels within a horizontal line segment will be similar. The horizontal line segments of an image will be determined in a raster scanning manner.

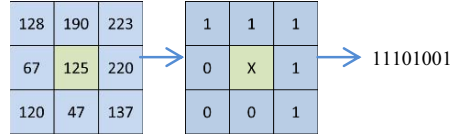


Fig. 2. An illustrated census transform example.



Fig. 3. An illustrated line segment example.

Next, to reduce matching ambiguities, cross-based cost aggregation is employed. For pixel p , a corresponding vertical line segment is similarly determined via two constraints in Eqs. (3) and (4). As the illustrated example shown in Fig. 4, the adaptive support window is determined as the union of all horizontal line segments containing the pixels within the vertical line segment of pixel p .

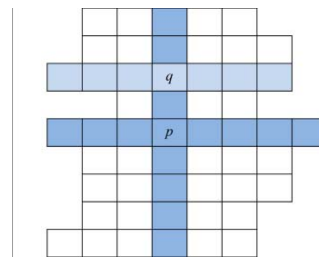


Fig. 4. An illustrated line segment example.

The aggregation matching cost $C_{aggr}(\cdot)$ of pixel p in its adaptive support window $U(p)$ is define as

$$C_{aggr}(p, d) = \sum_{q \in \mathbf{U}(p)} C(q, d). \quad (6)$$

The initial LR disparity map for LR left view $D_{L,low}^{initial}$ is defined as

$$D_{L,low}^{initial}(p) = \arg \min_d C_{aggr}(p, d). \quad (7)$$

The initial LR disparity map for LR right view can be similarly computed.

B. Disparity Plane Model Computation

In this study, a disparity plane model is used to reduce significant errors in initial disparity maps of occlusion regions and obtain sub-pixel disparity values. A disparity plane model will be determined for each segmented region. To avoid crossing disparity discontinuities, a segmented region should not include pixels of different disparities. Here, it is assumed that pixels having similar colors might have similarity disparity values. To perform disparity plane model computation on segmented regions, the mean-shift algorithm using color information [18] is employed to perform region segmentation on LR left view.

Based on the segmented regions on LR left view, the reliable pixels in the initial disparity map for LR left view will be detected to determine the disparity plane model for each segmented region. The left-right consistency check (LRC) [19] is employed to detect the reliable pixels of the initial LR disparity map. For pixel p in LR left view, the label map $L(p)$ for pixel p is defined as

$$L(p) = \begin{cases} \text{reliable, if } \alpha \leq 1, \\ \text{unreliable, otherwise,} \end{cases} \quad (8)$$

$$\alpha = \left| D_{L,low}^{initial}(p) + D_{R,low}^{initial}(x + D_{L,low}^{initial}(p), y) \right|, p = (x, y). \quad (9)$$

For each segmented region, three randomly selected reliable points can be used to determine the disparity plane model f described as

$$f: d = ax + by + c, \quad (10)$$

where d denotes the initial disparity value of reliable pixel $p=(x,y)$ and a , b , and c are three model parameters. To get an accurate disparity plane model, possible disparity plane models for each segmented region may be obtained by using randomly-sampling reliable point sets. Here, the number of reliable point sets for a segmented region is set to 100. Let \mathbf{F}_s denote 100 plane models of the 100 reliable point sets for segmented region s . To obtain the ‘‘best’’ disparity plane model f_s for segmented region s , a cost function is define as

$$C_{plane}(s, f) = \sum_{q \in s} \delta \times |f(q) - D_{L,low}^0(q)|, \quad (11)$$

where δ is a weighting function defined as

$$\begin{cases} \delta = 1, & \text{if } L(p) = \text{reliable,} \\ \delta = 0.1, & \text{otherwise.} \end{cases} \quad (12)$$

Then,

$$f_s = \arg \min_{f \in \mathbf{F}_s} C_{plane}(s, f), \quad (13)$$

and the disparity value of pixel p in segmented region s of LR left view is defined as

$$D_{L,low}(p) = f_s(p), p \in s. \quad (14)$$

C. Global Energy Cost Minimization

The obtained LR left disparity map $D_{L,low}$ can be improved by global energy cost function minimization. The refined LR left disparity map $D_{L,low}^{refined}$ for LR left view is obtained by minimizing the energy cost function

$$E(p) = E_{data}(p) + E_{smooth}(p) + E_{reliable}(p), \quad (15)$$

where $E_{data}(p)$ is the energy cost of the data term representing the color similarity measure between pixel $p(x, y)$ in LR left view and the corresponding pixel $p'(x + d, y)$ in LR right view, $E_{smooth}(p)$ is the energy cost of the smoothness term representing the disparity similarities between pixel p and its 4-connected neighboring pixels, and $E_{reliable}(p)$ is the energy cost of the reliable term denoting the disparity difference between $D_{L,low}(p)$ and $D_{L,low}^{initial}(p)$ for reliable pixel p .

$E_{data}(p)$, $E_{smooth}(p)$, and $E_{reliable}(p)$ are defined as

$$E_{data}(p) = \frac{\sum_{q \in L_p} C_{AD}(q, D_{L,low}(q))}{\text{length}(L_p)}, \quad (16)$$

$$E_{smooth}(p) = \sum_{q \in N(p)} (w_{smooth}(p, q) \times (D_{L,low}(p) - D_{L,low}(q))), \quad (17)$$

$$E_{reliable}(p) = \begin{cases} D_{L,low}(p) - D_{L,low}^{initial}(p), & \text{if } L(p) = \text{reliable,} \\ 0, & \text{otherwise,} \end{cases} \quad (18)$$

where $N(p)$ include the 4-connected neighboring pixels of pixel p and $w_{smooth}(p, q)$, a weighting function, is define as

$$w_{smooth}(p, q) = \exp\left(-\frac{\arg \max_{i \in \{R, G, B\}} |I_{L,low}^i(p) - I_{L,low}^i(q)|}{\rho}\right), \quad (19)$$

where ρ is empirically set to 10. In this study, the refined LR left disparity map $D_{L,low}^{refined}$ will be iteratively improved by global energy optimization to obtain the processed LR left disparity map $D_{L,low}^{processed}$. The number of iterations is empirically set to 20.

D. HR Disparity Map Composition by Region-based Fusion (Selection)

As shown in Fig. 1, the first initial HR left disparity map $D_{L,high}$ is obtained by performing bicubic interpolation on the processed LR left disparity map $D_{L,low}^{processed}$. On the other hand, the second HR binocular image pair is obtained by performing bicubic interpolation on the original LR binocular image pair, and the second initial HR left disparity map $D'_{L,high}$ is obtained by performing the similar steps (initial disparity map estimation using local aggregation, disparity plane model computation, and global energy cost minimization) on the second HR binocular image pair.

Based on the first and second initial HR left disparity maps $D_{L,high}$ and $D'_{L,high}$, to generate the ‘‘fused’’ left disparity map $D_{L,fusion}$, in this study region-based fusion (selection) will be performed on the two initial HR left disparity maps. Note that, using horizontal line segment structure, the line segment

structure can preserve the accuracy and continuity of the “fused” HR disparity maps. The “fused” HR left disparity map is defined as

$$D_{L,fusion}(L_j) = \arg \min_{D \in \{D_{L,high}, D_{L,low}\}} F(L_j, D), L_j \in \mathbf{L}, \quad (20)$$

where \mathbf{L} is the horizontal line segment set of HR left view, and L_j is the j -th horizontal line segment. Here, the “fusion” cost function $F(\cdot)$, including data cost $F_{data}(\cdot)$ and smoothness cost $F_{smooth}(\cdot)$, is defined as

$$F(L_j, D) = F_{data}(L_j, D) + \lambda_s \times F_{smooth}(L_j, D), \quad (21)$$

$$F_{data}(L_j, D) = \sum_{q \in L_j} C_{AD}(q, D(q)), \quad (22)$$

$$F_{smooth}(L_j, D) = \sum_{q \in L_j} S(q), \quad (23)$$

$$S(q) = \begin{cases} 1, & \text{if } D(x, y) = M(x, y-1), \\ 0, & \text{otherwise,} \end{cases} \quad (24)$$

where λ_s is empirically set to 0.1, $C_{AD}(\cdot)$ is defined in Eq. (1), the coordinate of q is (x, y) , and $M(x, y-1)$ is the disparity “mark” of position $(x, y-1)$ (either $D_{L,high}(x, y-1)$ or $D'_{L,high}(x, y-1)$).

E. “Fused” HR Disparity Map Refinement

To enhance the “fused” HD disparity maps, the adaptive color weighted median filter using color dissimilarity information [20] is employed. For pixel p in the “fused” HD disparity map, the disparity values of the pixels within a 11×11 window $W(p)$ centered at pixel p form a weighted disparity value histogram $hist_p$ defined as

$$hist_p(d) = \sum_{q \in W(p)} \tilde{w}(p, q), d = D_L(q), \quad (25)$$

where $\tilde{w}(p, q)$ is the normalized weighting function defined as

$$\tilde{w}(p, q) = \frac{w(p, q)}{\sum_{q \in W(p)} w(p, q)}, \quad (26)$$

where using the Laplacian kernel, $w(p, q)$ denotes the locally adaptive sorting weight defined as

$$w(p, q) = \exp\left(-\frac{C_{AD}(p, q)}{\sigma}\right), \quad (27)$$

where parameter σ is empirically set to 60. For pixel p , the bin values of the weighted disparity value histogram $hist_p(d)$ are accumulated from the minimal one. When the accumulated value exceeds 0.5, the corresponding disparity value (the median value) is decided as the final disparity value of pixel p .

III. EXPERIMENTAL RESULTS

The proposed approach has been implemented on an Intel Core i7-2700K 3.50GHz PC with 8GB main memory for Microsoft Windows platform using Microsoft visual C++ of version 2010 software develop tool. Seven binocular image pairs from Middlebury [1], namely, “Cones,” “Teddy,” “Art,” “Books,” “Dolls,” “Moebius,” and “Reindeer,” are used to evaluate the performance of the proposed approach. The magnification factors (MF) in this study are set to $m \times m$ with

$m=2$ and 4, i.e., $MF=2 \times 2$ and 4×4 . Using the original HR binocular image pairs from Middlebury as the ground truth, the LR binocular image pairs are obtained from the corresponding original HR binocular image pairs using bicubic interpolation by an MF of $(1/m) \times (1/m)$. To evaluate the performance of the proposed approach, four comparison approaches, namely, bicubic interpolation (Bicubic), joint bilateral upsampling (JBU) [13], pixel weighted average strategy (PWAS) [14], and Garcia et al.’s approach [15] are implemented. In this study, the peak signal to ratio (PSNR) and the bad pixel rate (BPR) are employed as two objective performance measures. BPR denotes the percentage of “bad” pixels in an image, i.e.,

$$BPR = \frac{1}{n} \sum_{i=1}^n (|X_i - X'_i| > T), \quad (28)$$

where T denotes the disparity error threshold (set to 1 in this study), X is the value in the HR ground truth disparity map, X' is the corresponding value in the reconstructed HR disparity map, and n is the total number of pixels in the original HR image. Here, a lower BPR indicates that the reconstructed HR disparity map has higher accuracy.

The final HR disparity maps of two binocular image pairs, “Cones” “Dolls,” by the four comparison approaches and the proposed approach with $MF=2 \times 2$ are shown in Figs. 5 and 6. The detail parts of the final HR disparity maps shown in Figs. 5 and 6 are shown in Figs. 7 and 8, respectively. Based on the final HR disparity maps obtained in this study, the visual quality of the final HR disparity maps by the proposed approach is better than those by the four comparison approaches.

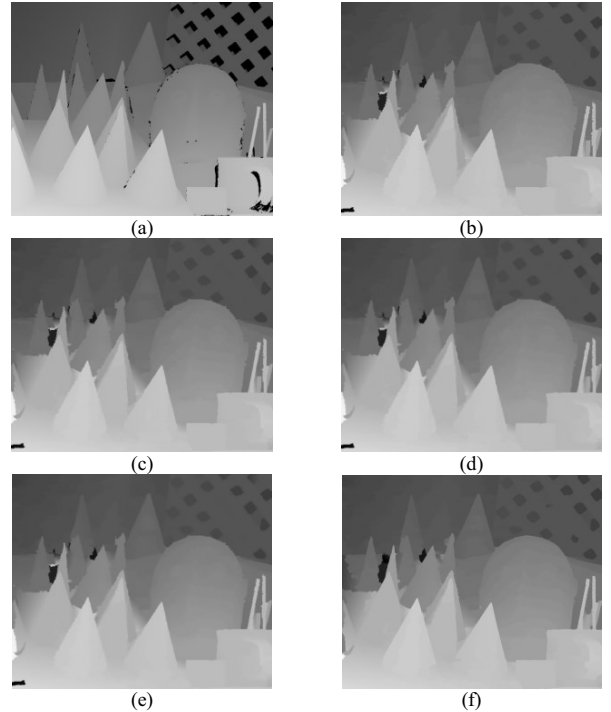


Fig. 5. The final HR disparity maps of “Cones:” (a) the ground truth; (b)-(f) the processed HR disparity maps by Bicubic, JBU [13], PWAS [14], Garcia et al.’s approach [15], and the proposed approach, respectively, with $MF=2 \times 2$.

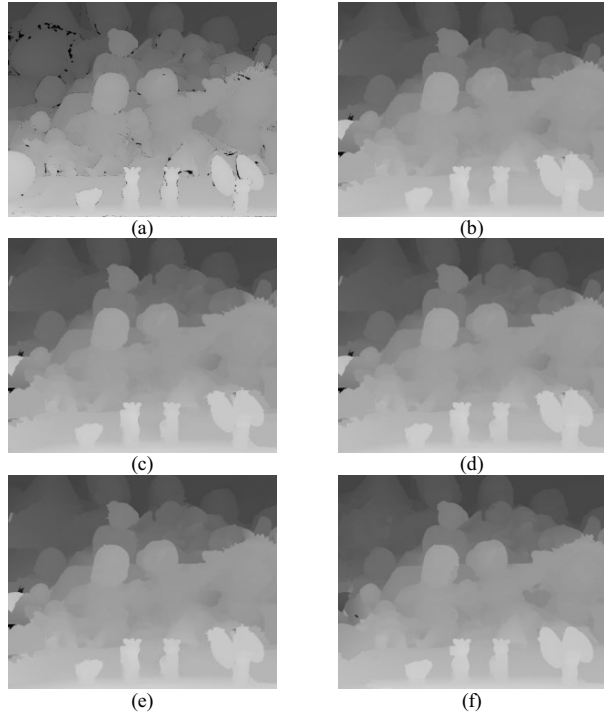


Fig. 6. The final HR disparity maps of “Dolls:” (a) the ground truth; (b)-(f) the processed HR disparity maps by Bicubic, JBU [13], PWAS [14], Garcia et al.’s approach [15], and the proposed approach, respectively, with $MF=2\times 2$.

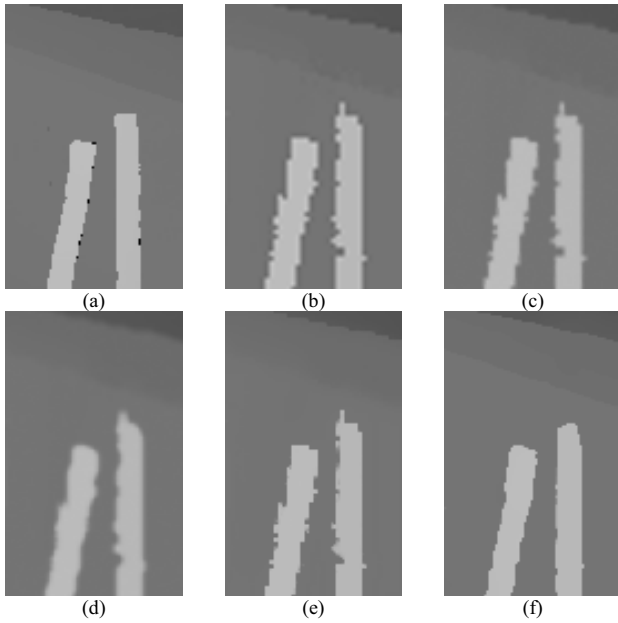


Fig. 7. The detail parts of the final HR disparity maps of “Cones:” (a) the ground truth; (b)-(f) the processed detail parts by Bicubic, JBU [13], PWAS [14], Garcia et al.’s approach [15], and the proposed approach, respectively, with $MF=2\times 2$.

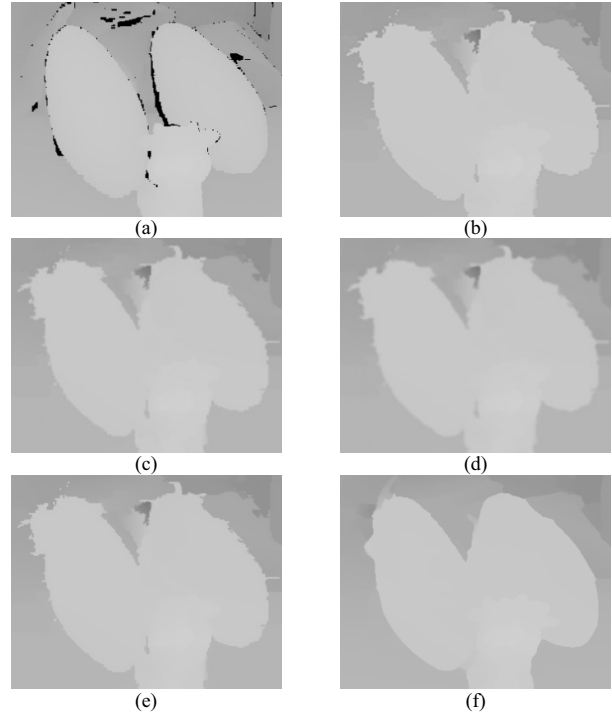


Fig. 8. The detail parts of the final HR disparity maps of “Dolls:” (a) the ground truth; (b)-(f) the processed detail parts by Bicubic, JBU [13], PWAS [14], Garcia et al.’s approach [15], and the proposed approach, respectively, with $MF=2\times 2$.

In terms of PSNR (dB) and the bad pixel rate (BPR) (%), the performance comparisons between the four comparison approaches, namely, Bicubic, JBU, PWAS, Garcia et al.’s approach, and the proposed approach with $MF=2\times 2$ and 4×4 for the seven binocular image pairs are listed in Tables 1 and 2. The average BPR and PSNR values of the final HR disparity maps of the seven binocular image pairs by the proposed approach are better than those by the four comparison approaches.

IV. CONCLUDING REMARKS

In this study, a super-resolution reconstruction approach for binocular 3D data is proposed. The aim is to obtain the high-resolution (HR) disparity map from a low-resolution (LR) binocular image pair by super-resolution reconstruction. The proposed approach contains five stages, i.e., initial disparity map estimation using local aggregation, disparity plane model computation, global energy cost minimization, HR disparity map composition by region-based fusion (selection), and fused HR disparity map refinement. Based on the experimental results obtained in this study, in terms of PSNR and bad pixel rate (BPR), the final HR disparity maps by the proposed approach are better than those by the four comparison approaches.

REFERENCES

- [1] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *Int. Journal of Computer Vision*, vol. 47, no. 1-3, pp. 7–42, Jun. 2002.

- [2] O. Veksler, "Fast variable window for stereo correspondence using integral images," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2003, pp. 556–561.
- [3] L. De-Maezta, A. Villanueva, and R. Cabeza, "Near real-time stereo matching using geodesic diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 410–416, Feb. 2012.
- [4] Q. Yang, L. Wang, R. Yang, H. Stewenius, and D. Nister, "Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 3, pp. 492–504, Mar. 2009.
- [5] Z. H. Wang and Z. G. Zheng, "A region based stereo matching algorithm using cooperative optimization," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.
- [6] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, "LidarBoost: depth superresolution for ToF 3D shape scanning," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2009, pp. 343–350.
- [7] M. Gevrekci and k. Pakin, "Depth map super resolution," in *Proc. of IEEE Int. Conf. on Image Processing*, 2011, pp. 3449–3452.
- [8] J. Zhu et al., "Reliability fusion of time-of-flight depth and stereo geometry for high quality depth maps," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 7, pp. 1400–1414, July 2011.
- [9] J. Zhang, L. H. Wang, D. X. Li, and M. Zhang, "High quality depth maps from stereo matching and ToF camera," in *Proc. of IEEE Int. Conf. on Soft Computing and Pattern Recognition*, 2011, pp. 68–72.
- [10] Y. Li, T. Xue, L. Sun, and J. Liu, "Joint example-based depth map Super-resolution," in *Proc. of IEEE Int. Conf. on Multimedia and Expo*, 2012, pp. 152–157.
- [11] J. Park et al., "High quality depth map upsampling for 3D-TOF cameras," in *Proc. of IEEE Int. Conf. on Computer Vision*, 2011, pp. 1623–1630.
- [12] J. Lu, D. Min, R. S. Pahwa, and M.N. Do, "A revisit to MRF-based depth map super-resolution and enhancement," in *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 2011, pp. 985–988.
- [13] J. Kopf, M. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," *ACM Trans. on Graphics*, vol. 26, no. 3, pp. 96:1–5, July, 2007.
- [14] F. Garcia et al., "Pixel weighted average strategy for depth sensor data fusion," in *Proc. of IEEE Int. Conf. on Image Processing*, 2010, pp. 2805–2808.
- [15] F. Garcia et al., "A new multi-lateral filter for real-time depth enhancement," in *Proc. of IEEE Int. Conf. on Advanced Video and Signal-Based Surveillance*, 2011, pp. 42–47.
- [16] X. Sun et al. "Stereo matching with reliable disparity propagation," in *Proc. of Int. Conf. on Imaging, Modeling, Processing, Visualization and Transmission*, 2011, pp. 132–139.
- [17] X. Mei et al. "On building an accurate stereo matching system on graphics hardware," in *Proc. of IEEE Int. Conf. on Computer Vision*, 2011, pp. 467–474.
- [18] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: color image segmentation," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 1997, pp.750–755.
- [19] J. Jachalsky, M. Schlosser, and D. Gandolph, "Confidence evaluation for robust, fast-converging disparity map refinement," in *Proc. of IEEE Int. Conf. on Multimedia and Expo*, 2010, pp. 1399–1404.
- [20] X. Pei, G. Wang, C. Shi, B. He, and X. Lin, "Effective refinement for accurate dense stereo matching using in-situ color-weighted median filter," in *Proc. Of IEEE Int. Conf. on Cyber Technology in Automation, Control, and Intelligent Systems*, 2012, pp.424–428

TABLE I. IN TERMS OF PSNR (DB), PERFORMANCE COMPARISONS OF THE SEVEN MIDDLEBURY BINOCULAR IMAGE PAIRS BETWEEN THE FOUR COMPARISON APPROACHES AND THE PROPOSED APPROACH WITH $MF=2\times 2$ AND 4×4 .

MF	Approaches	Cones	Teddy	Art	Books	Dolls	Moebius	Reindeer	Average
2×2	Bicubic	26.79	32.81	23.43	27.47	30.31	23.51	26.94	27.32
	JBU[13]	27.01	33.08	23.63	27.84	30.47	23.59	27.23	27.55
	PWAS[14]	27.00	33.10	23.69	27.48	30.84	23.60	27.27	27.57
	Garcia et al. [15]	26.77	32.80	23.62	27.56	30.33	23.53	27.22	27.40
	Proposed	27.18	35.10	24.50	26.69	32.09	25.14	27.84	28.36
4×4	Bicubic	27.49	29.45	23.01	25.81	32.80	23.22	26.88	26.95
	JBU[13]	27.72	29.63	23.25	26.10	32.73	23.37	27.24	27.15
	PWAS[14]	27.87	29.77	23.37	26.38	32.96	23.33	27.41	27.30
	Garcia et al. [15]	27.86	29.80	23.42	26.10	32.78	23.30	27.48	27.25
	Proposed	28.22	32.41	24.27	29.51	33.18	24.87	28.42	28.73

TABLE II. IN TERMS OF BAD PIXEL RATE (BPR) (%), PERFORMANCE COMPARISONS OF THE SEVEN MIDDLEBURY BINOCULAR IMAGE PAIRS BETWEEN THE FOUR COMPARISON APPROACHES AND THE PROPOSED APPROACH WITH $MF=2\times 2$ AND 4×4 .

MF	Approaches	Cones	Teddy	Art	Books	Dolls	Moebius	Reindeer	Average
2×2	Bicubic	11.05	13.64	16.36	11.91	14.69	13.34	11.95	13.28
	JBU[13]	13.95	15.51	18.40	14.20	14.68	17.06	14.38	15.45
	PWAS[14]	15.05	16.35	19.82	14.68	16.42	19.67	15.76	16.82
	Garcia et al. [15]	13.16	15.20	19.05	13.67	14.49	16.98	14.22	15.31
	Proposed	7.64	10.26	13.05	11.13	9.85	10.76	6.72	9.92
4×4	Bicubic	30.30	27.14	19.02	19.60	19.54	19.81	17.95	21.91
	JBU[13]	33.08	29.19	19.57	23.96	24.13	24.11	23.77	25.40
	PWAS[14]	33.09	29.85	20.09	24.41	25.10	26.91	23.59	26.14
	Garcia et al. [15]	32.34	30.25	18.68	24.83	27.78	24.27	23.86	26.00
	Proposed	9.78	13.60	12.99	11.24	12.40	12.27	10.34	11.93