Signal Processing: *Image Communication* 25 (2010) 130–142

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

An asymmetric post-processing for correspondence problem

Dongbo Min [a,b], Kwanghoon Sohn [a,*]

[a] *School of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea*
[b] *Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA*

ARTICLE INFO

Article history:
Received 16 April 2009
Received in revised form
11 September 2009
Accepted 18 October 2009

Keywords:
Adaptive filtering
Asymmetric consistency check
Post-processing
Stereo matching

ABSTRACT

This paper presents a novel approach that performs post-processing for stereo matching. We improve the performance of stereo matching by performing consistency check and adaptive filtering in an iterative filtering scheme. The consistency check is only done with asymmetric information so that very few additional computational loads are necessary. The information in the valid pixels is propagated into invalid pixels through the adaptive filtering. The proposed post-filtering method can be used in various methods for stereo matching. We demonstrate the validity of the proposed method by applying it to hierarchical belief propagation and semi-global matching. The performance of the post-processing method for hierarchical belief propagation is comparable to state-of-the-art methods in the Middlebury stereo datasets. In order to verify the performance of asymmetric consistency check, we compare it with other reliability estimation methods in the proposed post-processing scheme. Moreover, in order to verify the performance of post-filtering method in noisy environment, the proposed post-filtering method is applied to the stereo images denoised by NL-means algorithm. We find that the iterative filtering scheme reduce an error which may be caused in stereo matching for the denoised images and improve the performance of stereo matching.

1. Introduction

The correspondence problem has been an important issue in the field of computer vision, and many methods have been proposed to solve this problem. An extensive review of stereo matching algorithms can be found in [1]. In order to reduce the ambiguities and uncertainties of the matching process, many algorithms have been proposed using several constraints. Generally, stereo matching algorithms can be classified into two approaches (global and local) based on the strategies used for estimation. Global approaches refer to energy models that use various constraints to reduce uncertainties of disparity maps and solve them through various minimization techniques (such as graph cut [4] and belief propagation [17,18]). Local approaches use correlations between color or intensity patterns in neighboring windows. Performance depends on how the optimal window is selected in each pixel, but finding an optimal window with an arbitrary shape and size is very difficult.

Most approaches use various constraints (such as uniqueness and color segmentation constraints) to improve stereo matching performance. The uniqueness constraint means that the corresponding points between any two given images are unique. According to this principle, each pixel must have (at the most) one disparity. A simple way of detecting incorrect correspondences with the uniqueness constraint is the cross-checking technique. Most approaches have been proposed to estimate the disparities of incorrect matched pixels by

* Corresponding author.
E-mail addresses: dbmin99@gmail.com (D. Min), khsohn@yonsei.ac.kr (K. Sohn).

0923-5965/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2009.10.001

combining the uniqueness constraint into a global optimization method. These approaches estimate the symmetric disparity fields for the left and the right images, and this causes the computational complexity to double [2,3]. Approaches that have used the color segmentation constraint generally assume that the disparity vectors vary smoothly inside homogeneous color segments and change abruptly on the segment boundaries. In this way, they use the planar disparity model inside each segment [5,6]. The color segment-based methods can produce smooth disparity fields while preserving the discontinuities resulting from the boundaries.

Various post-processing methods have also been proposed to improve stereo matching performance. These methods detect incorrect correspondences by using the cross-checking method, and then apply the interpolation technique with the disparities of the visible pixels. In previous research, Mattoccia et al. used a two-step refinement process to determine the correct disparity assignment for points that violated the cross-checking method [7]. Invalid pixels were classified into occlusion and false matches by using the depth border information and different solutions were proposed for handling occlusion and false matches. This approach used the estimated depth border information and the mean shift algorithm for the refinement of the disparity maps. Hirschmuller proposed a refinement method for semi-global matching which used intensity consistent disparity selection for untextured regions and discontinuity preserving interpolation for filling holes with color segmentation [8,19]. These existing post-processing methods have used symmetric information. They have been used to perform the cross-checking method and then assign the disparities into invalid pixels through an interpolation process. Moreover, color segmentation [9] has been used to propagate the pixel information with correct disparities into invalid pixels.

In this paper, we propose a novel approach to efficient post-processing for stereo matching. First, we do an asymmetric consistency check which detects pixels with invalid disparities by using asymmetric information only. Most conventional methods have used cross-checking methods that have caused the computational complexity to double. However, in our experiments, the costs are refined through adaptive filtering, using valid pixel information. The refinement process is performed by an iterative filtering scheme with an asymmetric consistency check and adaptive filtering. The proposed method can be used with various other methods, as long as a suitable cost function is used.

The proposed scheme is different from the joint bilateral filtering which is used to increase spatial resolution of depth image [11,12]. While they use joint bilateral filtering scheme to upsample low resolution depth maps by using the high resolution color images, we use the iterative filtering scheme with valid pixels referred to as reliable by the consistency check to improve the performance of stereo matching. The consistency check and adaptive filtering compensate each other and improve the performance.

In this paper, we extended [13], which presented the preliminary results of the proposed method. We compared the proposed consistency check with other reliability estimation methods in the proposed post-processing scheme, and performed the experiments for the stereo images denoised by NL-means algorithm to verify the performance of post-filtering method in noisy environment.

The rest of this paper is organized as follows. In Section 2, we discuss the proposed method and explain asymmetric filtering method. A brief overview of global optimization methods used in the post-processing method and the utilization for noisy images are described in Section 3. Finally, we present the experimental results and conclusions in Sections 4 and 5, respectively.

2. The proposed post-processing method

The proposed post-processing method consists of two parts, which are the consistency check and adaptive filtering. The reliability function of the disparity map is estimated by performing a consistency check, and it is used in adaptive filtering for the refinement of the disparity map, which is used for the consistency check again. Valid pixel information is propagated into invalid pixels in this process. We improve the quality of the disparity map by using the iterative filtering scheme, as shown in Fig. 1. Filtering is done in the cost domain, not in the disparity domain. This prevents the disparity from being smoothed through the filtering process. In this paper, we only use the asymmetric information for the consistency check with a minimum number of additional computational loads, that is, only the left disparity field is used. Adaptive filtering is done only with the pixels that are certified as reliable by the consistency check.

2.1. Asymmetric consistency check

Several reliability estimation methods have been proposed for improving performance of stereo matching. Most of state-of-the-art methods have used symmetric matching scheme with cross-check method. Fig. 2 shows the consistency check method which detects pixels to violate the consistency by using both left and right disparity maps. Moreover, if $d_l(p) \neq d_r(p - d_l(p))$, the disparity in pixel p can be referred to as invalid. It is possible to classify invalid pixels into occlusion and false

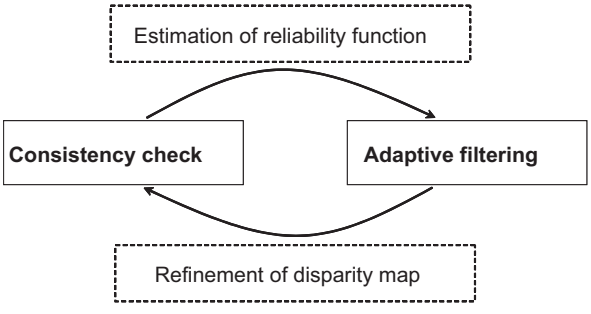

Fig. 1. Iterative filtering scheme.

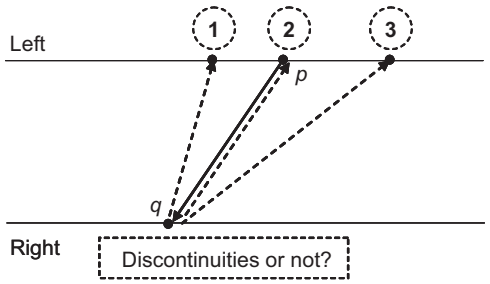

Fig. 2. Symmetric consistency check. The disparity value is valid only in case 2, and it is only possible to classify the invalid disparities into occlusion and false matches in cases 1 and 3 by using the depth discontinuities of the right disparity map.

matches by determining whether there are depth discontinuities in right image, since the depth discontinuities in one image correspond to occlusions in the other image. Some methods use the reliability which defines how distinct matching cost for the estimated disparity is [14,15]. Merrell et al. proposed the reliability function which was estimated by computing the likelihood that the estimated disparity does not have the lowest cost due to lack of texture and noise [15].

In the proposed method, the reliability function is asymmetrically estimated by doing a consistency check. The left disparity map is used in the consistency check, since the invalid pixels are detected with asymmetric information only. We determine a candidate set of invalid pixels, and do not discriminate occlusion and false matches. The same method is applied to both occlusion and false matches. This is different from conventional methods which handle occlusion and false matches in a different manner [7,8,19]. Although some valid pixels may be contained in the candidate set of invalid pixels, this problem can be solved by using the proposed method. For the asymmetric consistency check, we use both geometric and photometric constraints. To determine whether a pixel is valid or not, we evaluate the disparity values of the neighboring pixels. We define the binary reliability function $R(i)$ as follows:

$$R(p) = \begin{cases} 1 & \text{if } d(p) \text{ is valid,} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where d represents the disparity of the pixel. For estimating the reliability function, we describe the function $S_r(j)$ as a set of pixels in the right image:

$$S_r(j) = \{i | i - d(i) = j \text{ all } i \text{ with } 0 \le i \le W-1\},$$

where i and j represent the x coordinates of the left and the right images, respectively. W represents the width of the image. When there are multiple matching points at the pixels in the other image, that is, $\#(S_r(j)) > 1$, the pixel with the largest disparity value among $S_r(j)$ is considered valid and the remaining pixels are considered invalid. This is true only if the valid pixels have reliable disparities. Fig. 3 shows two forms of the asymmetric consistency check. We use the photometric constraint to evaluate the reliability of the disparities in valid pixels. The costs incurred by invalid pixels are generally larger than those

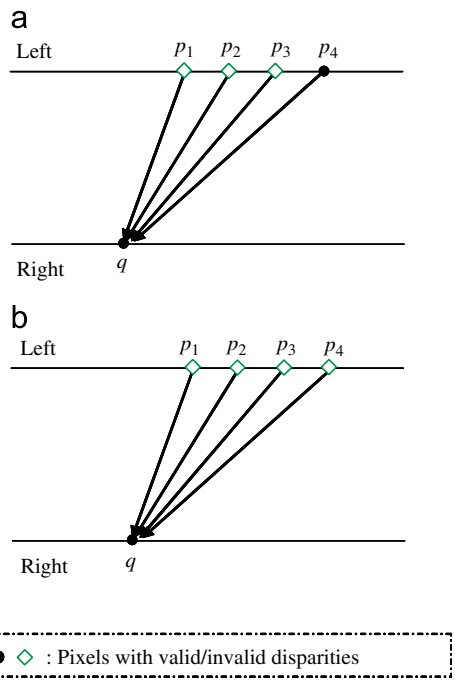

Fig. 3. Asymmetric consistency check. When $\#(S_r(q)) > 1$, (a) $Cost(p_4, d(p_4)) < Cost(p_i, d(p_i))$ for all $i = 1, 2, 3$ and (b) $Cost(p_4, d(p_4)) \ge Cost(p_i, d(p_i))$ for any $i = 1, 2, 3$.

of valid pixels. If the cost at the pixel, which is determined as valid pixels by using geometric constraints, is not smaller than that of the remaining invalid pixels, we cannot guarantee that the valid pixels are reliable. Therefore, all the pixels in $S_r(j)$ are referred to as invalid, that is, $R(i) = 0$ for all $i \in S_r(j)$, as shown in Fig. 3(b). Here, various cost functions are used to determine the reliability of the pixels: the aggregated cost function $E_L(p, d_p)$ computed by local methods such as the shiftable window [10] and the adaptive support window [16], or the cost function $E_G(p, d_p)$ computed by global methods such as belief propagation [17,18] and dynamic programming [8,19] as follows:

$$E_G(p, d) = E_L(p, d) + \lambda \sum_{q \in N(p)} S(d_p, d_q), \tag{2}$$

where S represents the smoothness term in the global method.

2.2. Adaptive filtering

Given the reliability function R_l of the disparity map, we are able to refine costs with the adaptive filtering technique. This process is based on the assumption that the costs would vary smoothly, except at the object boundaries. Fig. 4 shows the costs estimated by belief propagation and refined with the proposed filtering method for the 'Tsukuba' image, when $d = 0$. From this, we use an iterative nonlinear filtering method with a weight function w as follows:

$$E^{k+1}(p, d_p) = \frac{\sum_{m \in N(p)} R_l(m) w(p, m) E^k(m, d_p)}{\sum_{m \in N(p)} R_l(m) w(p, m)}, \tag{3}$$

Fig. 4. Costs estimated by belief propagation and refined with the proposed adaptive filtering method for the 'Tsukuba' image.

where $w(p,m)$ means the weighting function in the support window $N(p)$. This is similar to bilateral filtering [20], which is an intuitive nonlinear filtering method that smoothes images while preserving edges. The cost function $E(p,d_p)$ can generally be initialized to $E_L(p,d_p)$ or $E_G(p,d_p)$ in Eq. (3).

One reason for slowing down the convergence in Eq. (3) is that the updated components in each pixel are used only after one iteration is complete. We compensate for this problem by using the updated components of each pixel intermediately. Since the adaptive filtering is done in the raster scan order, we divide a set of neighboring pixels $N(p)$ into two parts: the causal part $N_c(p)$ and the non-causal part $N_n(p)$. Eq. (3) appears as follows, based on this relationship:

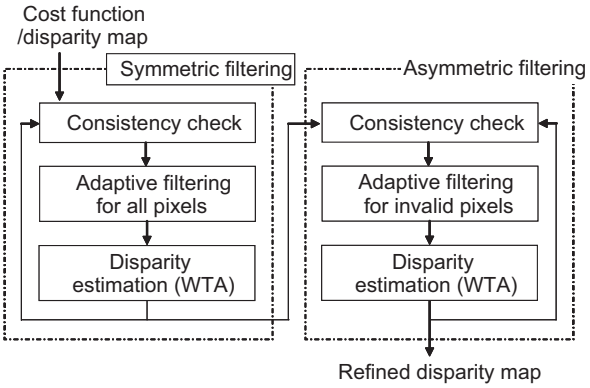

Fig. 5. Proposed post-filtering method.

$$E^{k+1}(p,d_p)=\frac{\sum_{m\in N_c(p)}R_l(m)w(p,m)E^{k+1}(m,d_p)+\sum_{m\in N_n(p)}R_l(m)w(p,m)E^{k}(m,d_p)}{\sum_{m\in N(p)}R_l(m)w(p,m)}. \tag{4}$$

By running the iteration scheme, the cost function E is regularized with the weighted neighboring pixel cost. In the proposed method, the weighting function w used in Eq. (4) is classified as the symmetric weighting function w_S and the asymmetric weighting function w_A. Before defining the weighting functions w_S and w_A, we describe the following Gaussian weight function w_l and w_r in the CIE-Lab color space as follows:

$$w_i(p,m)=\exp\left(-\left(\frac{C_i(p,m)}{2r_c^2}+\frac{S(p,m)}{2r_s^2}\right)\right),\quad i=l,r,$$

$$C(p,m)=(L_p-L_m)^2+(a_p-a_m)^2+(b_p-b_m)^2,$$

$$S(p,m)=(p-m)^2. \tag{5}$$

By using the above equation, w_S and w_A can be expressed as follows:

$$w_S(p,m)=w_l(p,m)w_r(p_d,m_d),$$

$$w_A(p,m)=w_l(p,m), \tag{6}$$

w_S is symmetric in the sense that the weighting function is computed using both the left and the right images, while w_A is computed using only the left image. Fig. 5 shows the overall process for the proposed method. Given the cost function and the disparity map, the costs are refined in the adaptive filtering process and the disparities are selected by the WTA (Winner-Takes-All) method iteratively. Symmetric and asymmetric filtering means that adaptive filtering is done with symmetric and asymmetric weighting functions w_S and w_A. In the symmetric filtering process, adaptive filtering is done for all the pixels. The symmetric weighting function is available only when the corresponding pixels for both images are visible, since it uses both the left and right images. Therefore, asymmetric filtering is performed to refine the costs of the invalid pixels. Reasonable costs are assigned to invalid pixels through asymmetric filtering.

Adaptive filtering is sequentially performed. After adaptive filtering is performed on an invalid pixel, the pixel becomes valid. In other words, $R(p)=1$. Therefore, the invalid pixels filtered with the valid pixels are used as valid pixels in Eq. (4) again. The information in the valid pixels is propagated to estimate the costs of the invalid pixels. Fig. 6 shows the pseudo-codes of the symmetric and asymmetric filtering methods.

Symmetric filtering

for all $p \in I_l$

Compute $E^{k+1}(p, d)$ for all possible d using Eq. (4)

$R(p) = 1$

Asymmetric filtering

for all $p \in I_l$

if $R(p) = 0$

Compute $E^{k+1}(p, d)$ for all possible d using Eq. (4)

$R(p) = 1$

Fig. 6. Pseudo-codes for adaptive filtering in the proposed method.

3. Validity of the proposed post-filtering method

We demonstrate the usefulness of the proposed method by showing the improvement of performance when using several algorithms. In addition, the proposed method is also shown to improve the quality of the disparity maps in noisy environments.

3.1. Application to global methods

The proposed method uses several optimization techniques. In this paper, we confirm the validity of the proposed method by applying it to hierarchical belief propagation (HBP) [17] and semi-global matching (SGM) [19]. In this section, we present a brief overview of HBP and SGM.

3.1.1. Hierarchical belief propagation

The loopy belief propagation method works by passing messages between a pixel and its 4-connected neighborhoods. It gathers the information from the observed data and messages of the neighboring pixels and updates the pixel's message by iteratively optimizing the message as follows:

$$m_{pq}^{k}(d_q) = \min_{d_p}\left(S(d_p, d_q) + E_L(p, d_p) + \sum_{s \in N(p)/q} m_{sp}^{k-1}(d_p)\right), \tag{7}$$

where m_{pq}^{k} is the message that a pixel p sends to a neighboring pixel q at the k th iteration and $N(p)/q$ means a set of neighboring pixels for p except q. After K iterations, we compute the belief (or global cost function in this paper) of pixel p as follows:

$$E_G(p, d_p) = E_L(p, d_p) + \sum_{s \in N(p)} m_{sp}^{K}(d_p). \tag{8}$$

The disparity value that minimizes the cost function E_G for each pixel is selected. One problem of loopy belief propagation is that a number of iterations are necessary to pass the message over a large distance in the 4-connected grid graph. In previous research [17], the multiscale approach is used to solve this problem. Hierarchical belief propagation works in a coarse-to-fine manner. The final value in the coarse level is used as the initial value in the fine level, and it reduces the number of iterations required for propagation of information at a large distance. After we compute the global cost E_G with HBP, it is used to determine the reliability of the disparity in the asymmetric consistency check.

3.1.2. Semi-global matching

The semi-global matching (SGM) method computes the global cost in the SO (scanline optimization) framework. It calculates $E_G(p, d)$ efficiently along 1D paths from eight or 16 directions towards each pixel [19]. The cost of a pixel p with disparity d from the direction r is defined as follows:

$$L_r(p, d_p) = E_L(p, d_p) - \min_k L_r(p - r, k) + \min\begin{pmatrix} L_r(p - r, d_p), L_r(p - r, d_p - 1) + P_1 \\ L_r(p - r, d_p + 1) + P_1, \min_i L_r(p - r, i) + P_2 \end{pmatrix} \tag{9}$$

where P_1 and P_2 represent a penalty which is added when the difference of the disparity value is 1 or more. The second term in Eq. (9) does not affect subsequent calculations, and just prevent L from monotonically increasing along each path. The global cost $E_G(p, d_p)$ is computed by summing the costs $L_r(p, d_p)$ along the paths from all directions r:

$$E_G(p, d_p) = \sum_r L_r(p, d_p). \tag{10}$$

The disparity map is obtained by selecting the disparity value that corresponds to the minimum cost of $E_G(p, d)$. In the SGM, we use a 4-pass scanline optimization to compute the global cost function, which is different from the process that is used in [19]: 2 along horizontal scanlines and two along vertical scanlines. Although the results of the SGM differ from those in [19], it does not matter since we are focused on improving the performance of stereo matching algorithms via the proposed method.

3.2. Robustness in noisy environments

Noise has proven to be one of the most serious problems in the fields of computer vision, and this fact has influenced the performance of many algorithms. A number of methods have been proposed to remove noise and restore the original image. Many stereo algorithms may also be affected by noise. Recently, some researchers [21] proposed an algorithm that simultaneously performs both stereo matching and image denoising for a pair of noisy stereo images. This method introduced a data cost that is robust to noise with a restored intensity difference and a non-local pixel distribution dissimilarity around the matched pixels. An NL-means algorithm (non-local-means) algorithm [22,23], which means non-local and nonlinear averaging of the pixels in an image, was used for restoring the intensity value as a function of disparity. The stereo matching and image denoising problems were solved simultaneously in the iterative framework, based on the assumption that these

problems could compensate for each other. In other words, denoising could be done more precisely with multiple samples provided by the correct correspondence. Also, stereo matching performance was improved in the noise-free images.

In this section, we evaluate the proposed method by applying it to noisy stereo images. We restore the intensity value with the NL-means algorithm but perform denoising and stereo matching independently, which is different from previous research [21]. In general, the performance of denoising can be improved when more samples in two or more images are provided. However, this does not mean that we need the correct corresponding points [23]. Correspondence methods generally suffer from an ambiguity of matching (the so-called aperture problem). In denoising, the aperture problem just means that there are many blocks in the other images similar to the given one in the current image. Hence, singling out one specific block in the other images is an unnecessary and probably harmful step [23]. We denoise a pixel in the given image with the NL-means algorithm by involving the similar pixels in the other images with spatial and temporal similarities. Once the denoised images are obtained with the noisy left and right images, they are used in stereo matching and the proposed modules.

3.2.1. Denoising with non-local means filtering

This section offers a simple explanation of the denoising algorithm with NL-means filtering for stereo images. For a more detailed explanation, please refer to [22,23]. In the experiments, we denote $u(i)$ and $r(i)$ as the observed noisy and denoised images. The restored intensity is obtained by using the weighted average of all the pixels in an image I as follows:

$$r(i) = \sum_{j \in I} w(i,j)u(j), \tag{11}$$

where the weights $w(i,j)$ mean the similarity of pixels i and j, which is defined with the Gaussian averaging of the amount of similarity between the neighborhoods of two pixels:

$$w(i,j) = \frac{1}{Z(i)} e^{-((G*|u(N_i)-u(N_j)|^2)(0))/h^2},$$

$$(G * |u(N_i) - u(N_j)|^2)(0) = \int_t G(t)|u(i+t) - u(j+t)|^2 \, dt, \tag{12}$$

where G represents a Gaussian kernel and $Z(i)$ represents the normalizing factor. h represents a filtering parameter, and N_i means a set of neighboring pixels of a pixel i. The weight $w(i,j)$ becomes larger when the neighborhoods in the Gaussian window around i and j are similar.

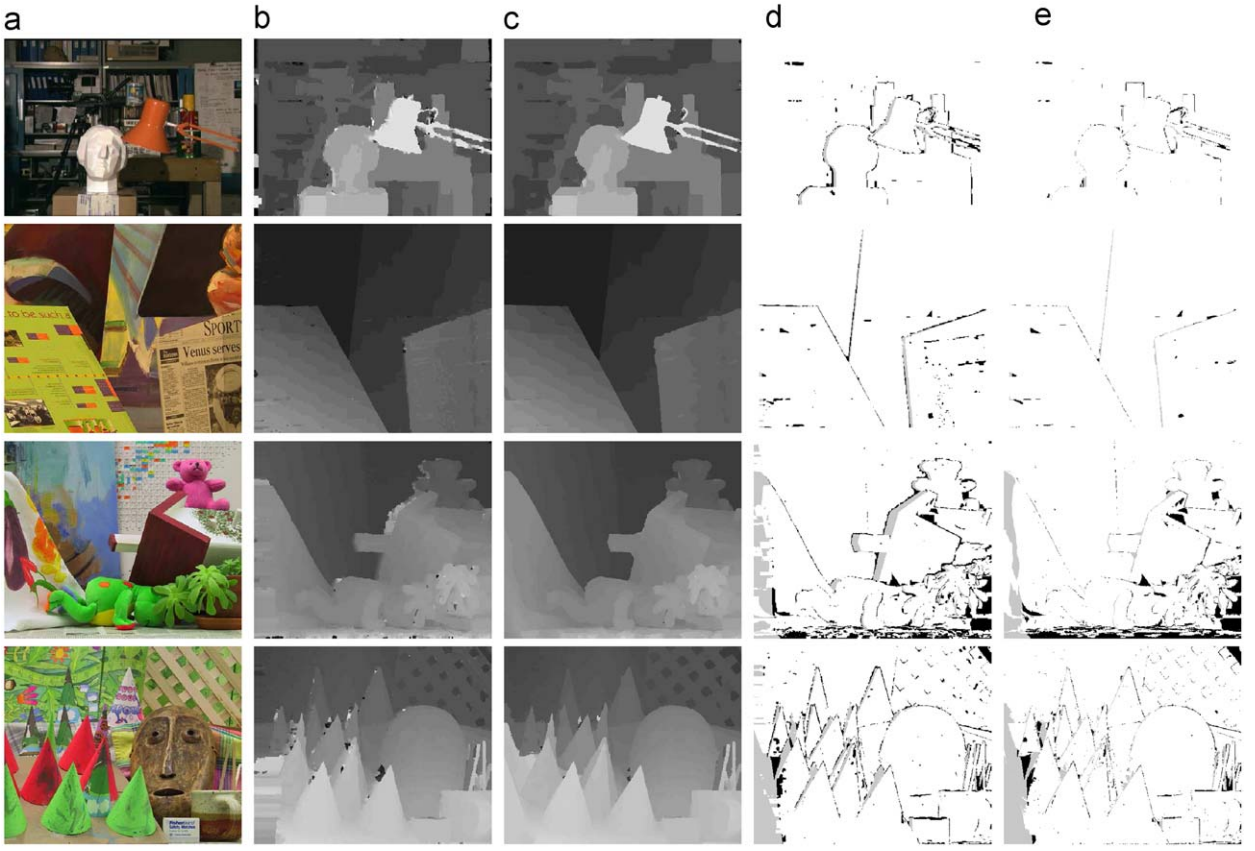

Fig. 7. Results for (from top to bottom) the 'Tsukuba', 'Venus', 'Teddy', and 'Cone' image pairs: (a) original images, (b) HBP, (c) HBP + post-processing, (d) error map of HBP, and (e) error map of HBP + post-processing.

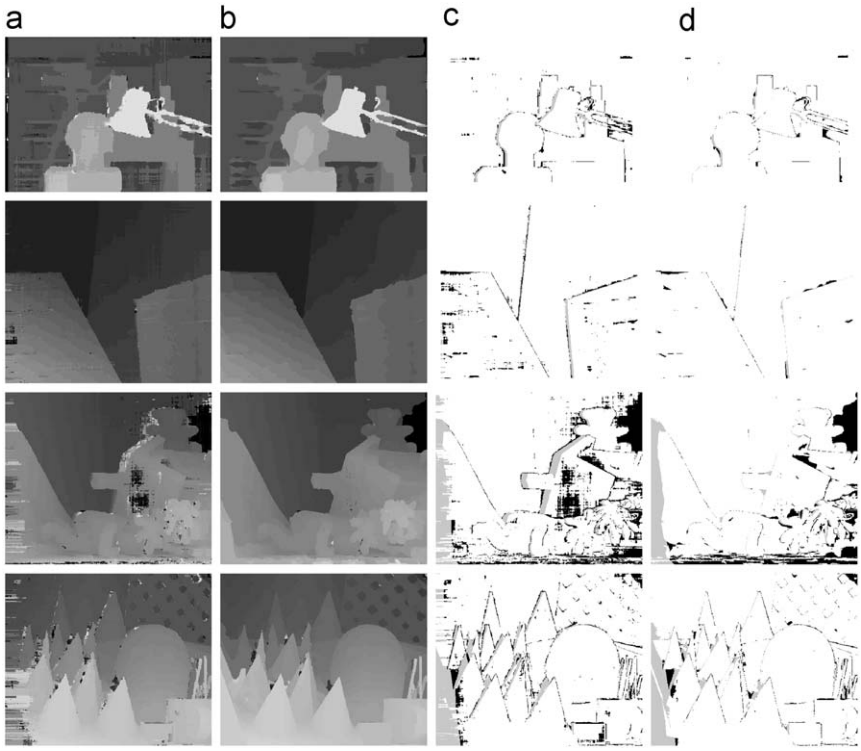

Fig. 8. Results for (from top to bottom) the 'Tsukuba', 'Venus', 'Teddy', and 'Cone' image pairs: (a) SGM, (b) SGM + post-processing, (c) error map of SGM, and (d) error map of SGM + post-processing.

Table 1
Objective evaluation for the proposed method with the Middlebury test bed.

Algorithm	*Tsukuba*			*Venus*			*Teddy*			*Cone*		
	Nonocc	All	Disc	Nonocc	All	Disc	Nonocc	All	Disc	Nonocc	All	Disc
AdaptingBP [5]	1.11	1.37	5.79	0.10	0.21	1.44	4.22	7.06	11.8	2.48	7.92	7.32
DoubleBP [24]	0.88	1.29	4.76	0.14	0.60	2.00	3.55	8.71	9.70	2.90	9.24	7.80
AdaptOvrSegBP [28]	1.69	2.04	5.64	0.14	0.20	1.47	7.04	11.1	16.4	3.60	8.96	8.84
HBP + PostFil	1.12	1.63	5.44	0.48	1.08	2.51	7.65	11.4	18.1	3.46	10.6	8.79
DistinctSM [29]	1.21	1.75	6.39	0.35	0.69	2.63	7.45	13.0	18.1	3.91	9.91	8.32
CostAggr + occ [25]	1.38	1.96	7.14	0.44	1.13	4.87	6.80	11.9	17.3	3.60	8.57	9.36
OverSegmBP [26]	1.69	1.97	8.47	0.51	0.68	4.69	6.74	11.9	15.8	3.19	8.81	8.89
EnhancedBP [27]	0.94	1.74	5.05	0.35	0.86	4.34	8.11	13.3	18.5	5.09	11.1	11.0
SGM + PostFil	1.44	2.25	7.08	0.68	1.25	6.98	9.86	14.6	19.5	3.02	9.47	7.90
Semi-Glob	3.26	3.96	12.8	1.00	1.57	11.3	6.02	12.2	16.3	3.06	9.75	8.90
RealtimeBP	1.49	3.40	7.87	0.77	1.90	9.00	8.72	13.2	17.2	4.61	11.6	12.4
HBP	2.35	4.49	11.0	1.62	2.72	11.3	8.42	14.3	21.6	5.14	13.4	12.9
SGM	2.59	4.71	11.9	2.22	3.41	15.6	13.8	20.1	22.7	5.74	14.1	13.4

Fig. 9. Intermediate results of the proposed method for HBP: Results for (a) HBP, (b), (c) invalid map and refined disparity map created by symmetric filtering, (d), (e) invalid map and refined disparity map created by asymmetric filtering. We found that the consistency check and the adaptive filtering process compensated each other and improved performance in the iterative filtering scheme.

In practical implementation, the pixels in the large search window are used instead of all the pixels in image I for reduction of computational complexity in Eq. (11).

Denoising for the stereo images is performed by NL-means filtering with both the left and the right images. We define the denoised left image as the weighted average of the pixels in the large search windows $T_l(i)$ and $T_r(i)$ of the left and the right images as follows:

$$r_l(i) = \frac{1}{Z(i)} \left\{ \sum_{j \in T_l(i)} e^{-(G * |u_l(N_i) - u_l(N_j)|^2)/h^2} u_l(j) + \sum_{k \in T_r(i)} e^{-(G * |u_l(N_i) - u_r(N_k)|^2)/h^2} u_r(k) \right\}, \quad (13)$$

$$T_l(i) = \{(m,n) | -N_x \le m \le N_x, -N_y \le n \le N_y\},$$

$$T_r(i) = \{(m,n) | -(N_x + SR) \le m \le N_x, -N_y \le n \le N_y\},$$

where u_l and u_r represent the left and right images, respectively. The large search window $T_r(i)$ of the right image is defined by the search range used in stereo matching. The denoising for the right image is done in a similar manner.

4. Experimental results

We evaluated the performance of the proposed method and compared it with state-of-the-art methods in the Middlebury test bed [30]. Since our main goal was to propose an efficient post-processing method, we also evaluated the performance by using post-processing methods for the HBP and the SGM. For objective evaluation, we measured the percentage of bad matching pixels (where the absolute disparity error was greater than 1 pixel). The measurement was computed for three

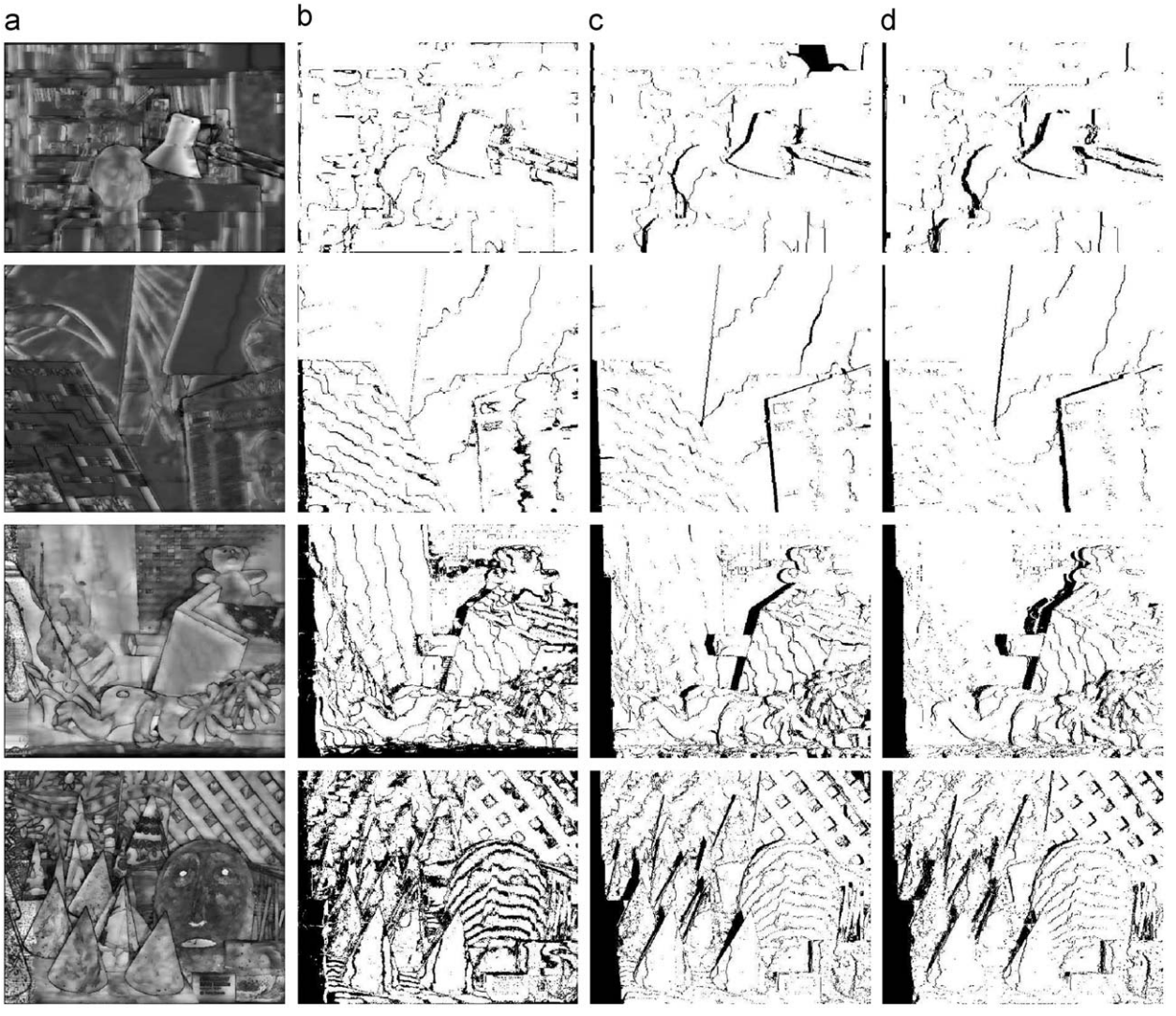

Fig. 10. Reliability maps estimated with costs of HBP for (from top to bottom) the 'Tsukuba', 'Venus', 'Teddy', and 'Cone' image pairs: (a) Merrell's method [15], (b) distinctness function (Eq. (14)), (c) cross-check method, and (d) proposed method.

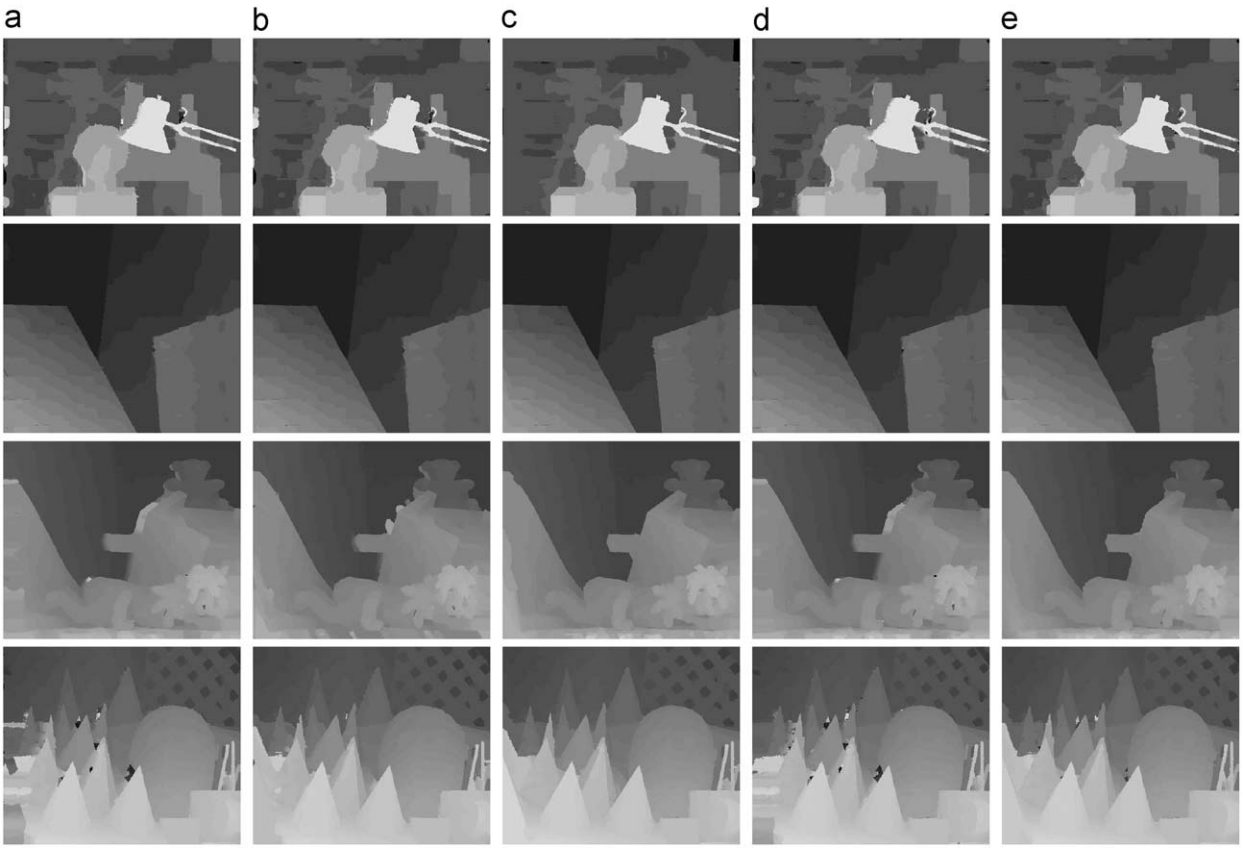

Fig. 11. Subjective evaluation of HBP+post-processing for (from top to bottom) the 'Tsukuba', 'Venus', 'Teddy', and 'Cone' image pairs: (a) Merrell's method [15], (b) distinctness function (Eq. (14)), (c) cross-check method, (d) without consistency check, and (e) proposed method.

Table 2
Objective evaluation of the post-processing for HBP with various reliability estimation methods.

Algorithm	Avg. rank	*Tsukuba*			*Venus*			*Teddy*			*Cone*		
		Nonocc	All	Disc	Nonocc	All	Disc	Nonocc	All	Disc	Nonocc	All	Disc
Merrell's method	21.0	1.26	2.43	6.42	0.47	1.19	2.91	8.89	14.4	21.6	3.68	11.3	9.34
Dist. func.	21.7	1.12	2.58	5.67	0.57	1.50	4.19	9.53	14.8	23.4	4.08	11.1	10.1
Cross-check	12.4	1.20	1.58	5.95	0.53	0.81	1.92	7.24	12.0	18.3	3.31	7.66	8.42
Without check	21.6	1.26	3.16	6.11	0.61	1.61	4.35	8.23	14.1	20.2	3.8	11.8	9.68
Proposed method	14.1	1.12	1.63	5.44	0.48	1.08	2.51	7.65	11.4	18.1	3.46	10.6	8.79

Table 3
Processing times of the post-processing for HBP with various reliability estimation methods.

Algorithm	*Tsukuba* (s)	*Venus* (s)	*Teddy* (s)	*Cone* (s)
Merrell's method	3.3	5.6	15.4	15.5
Dist. func.	3.0	5.3	15.1	15.2
Cross-check	5.9	10.7	30.2	30.1
Without check	2.9	5.4	15.2	15.3
Proposed method	3.1	5.5	15.3	15.5

subsets of an image: nonocc (the pixels in the non-occluded regions), all (the pixels in both the non-occluded and half-occluded regions), and disc (the visible pixels near the occluded regions). The proposed method was tested using the same parameters for all the test images. The two parameters in the weighting function were $r_c = 8.0$, $r_s = 8.0$, and the size of the window for adaptive filtering was 11×11 pixels. In Fig. 5, the number of iterations for symmetric and asymmetric filtering was 1. We obtained sufficient improvements after only one iteration. In the NL-means denoising method, the large search window was 31×31 pixels ($N_x, N_y = 31$) and N_i was 3×3 pixels.

Figs. 7 and 8 show the results of the proposed method for the HBP and the SGM. The proposed method yielded accurate results for the discontinuity, occluded, and textureless regions. Table 1 shows that the proposed method yielded comparable performance with state-of-the-art methods, even though it performed post-

processing using only asymmetric information. We found that the proposed method improved the performance of the HBP in the nonocc, all, and disc regions significantly.

Fig. 9 shows the intermediate results of the proposed method for the HBP. Given the cost function and the disparity map, the invalid pixels were estimated with an asymmetric consistency check. We found that the consistency check and the adaptive filtering method compensated each other and improved the performance in the iterative filtering scheme. As shown in Fig. 5, the costs of all the pixels were refined with the valid pixel information only in the symmetric filtering process, and the valid pixel information was propagated into invalid pixels in the asymmetric filtering process. Note that the filtering was performed in the cost function, not in the disparity map. In the iterative filtering scheme, the costs

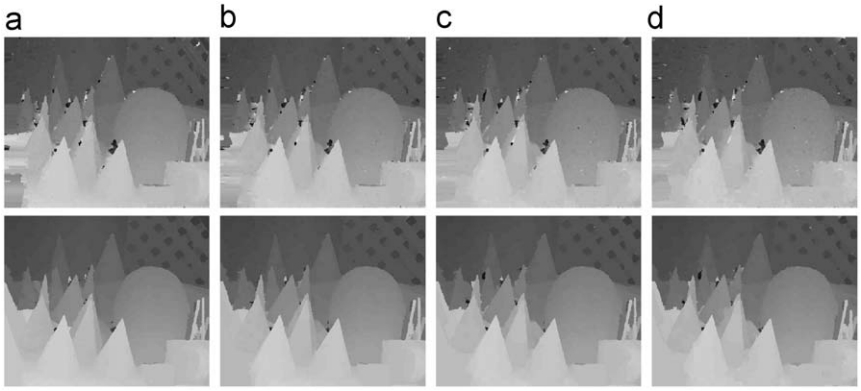

Fig. 12. Refinement by the proposed post-processing process in the noisy 'Cone' images: Results for (from top to bottom) HBP and HBP+PostFil. AWGN (additive white Gaussian noise) with mean 0 and various standard deviation σ was added to the standard stereo image pairs. The results of the post-filtering method for HBP are shown only. The disparities of invalid pixels were replaced with reasonable values and the discontinuity localization on the object boundaries was improved in the iterative filtering scheme: (a) $\sigma = 5$, (b) $\sigma = 10$, (c) $\sigma = 15$, and (d) $\sigma = 20$.

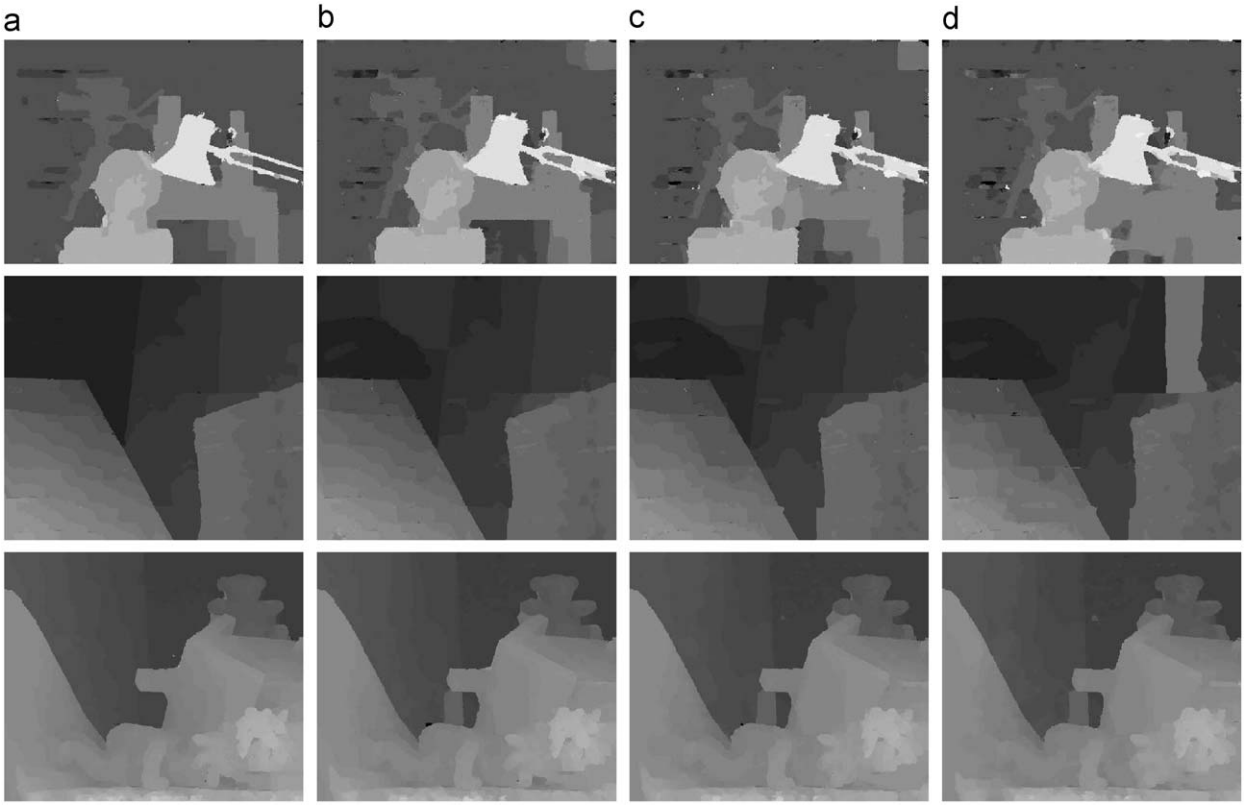

Fig. 13. The results of post-processing for HBP in the other noisy stereo images: (from top to bottom) the 'Tsukuba', 'Venus', and 'Teddy' image pairs: (a) $\sigma = 5$, (b) $\sigma = 10$, (c) $\sigma = 15$, and (d) $\sigma = 20$.

of all the pixels were refined by adaptive filtering with an invalid pixel map.

In order to verify the performance of asymmetric consistency check, we used several reliability estimation methods into the proposed post-processing scheme: (1) Merrell's method [15], (2) reliability function computed by the rate between the first minimum cost E_1 and second minimum cost E_2 (in this paper, we call this a distinctness function), (3) cross-check method, and (4) without consistency check. The distinctness function is computed as follows:

$$\left|\frac{E_1 - E_2}{E_2}\right|. \tag{14}$$

If it is above a threshold, the reliability function $R(p)$ is set to 1 and otherwise 0. Merrell [15] defined the reliability function as the inverse of the sum of the likelihood that the minimum disparity d_0 does not have the lowest cost for all possible disparities in Eq. (15). It defines how sharp the cost of estimated disparity d_0 is:

$$R(p) = \left(\sum_{d \neq d_0} e^{-(E(p,d)-E(p,d_0))^2/\sigma_r^2}\right)^{-1}. \tag{15}$$

We used symmetric matching scheme to perform the proposed post-processing with cross-checking. Invalid pixels were estimated by cross-checking for left and right disparity maps, and then the adaptive filtering was done for both left and right costs.

Figs. 10 and 11 compare the asymmetric consistency check with other reliability estimation methods in the post-processing scheme for HBP. Fig. 10 shows the results of reliability estimation for standard test image pairs. Merrell's method computes the reliability function with continuous value between 0 and 1, and brighter value means that its pixel is more reliable. As many invalid pixels as possible should be included in the candidate set. Although some valid pixels may be contained in the candidate set of invalid pixels, it can be solved by using the proposed post-processing method. Fig. 11 shows the results of post-processing method by using several reliability estimation methods. The effects of consistency check on the iterative filtering scheme are shown. We could find that the post-processing method with the distinctness function and Merrell's method did not address the problems in the occluded pixels. Without consistency check, the information of valid pixels is not propagated into the valid pixels as shown in Fig. 11(d). There are no significant improvements through the post-processing scheme, especially in the occluded region. The objective evaluation results for all image pairs are shown in Table 2. Although the results by cross-check method was slightly better than those by the asymmetric

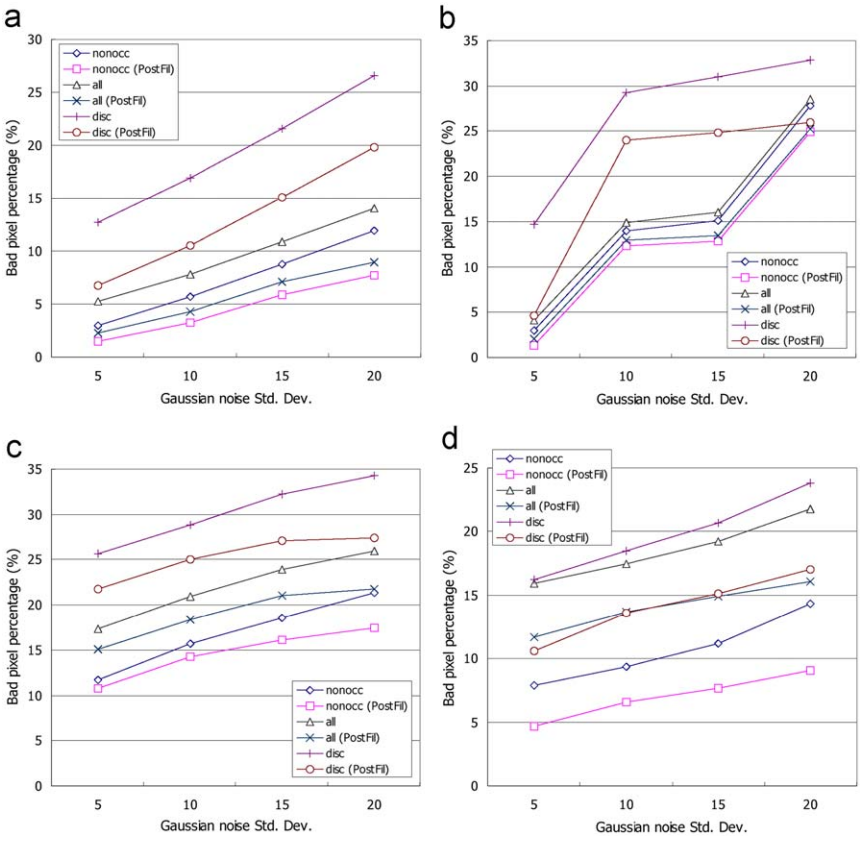

Fig. 14. Objective comparison of the proposed post-processing method for HBP in the noisy stereo images: the (a) 'Tsukuba', (b) 'Venus', (c) 'Teddy', and (d) 'Cone' image pairs. The proposed method improved the disparity maps significantly in the nonocc, all, and disc regions.

consistency check, the processing time of cross-check method was nearly double of that of asymmetric consistency check as shown in Table 3.

Fig. 12 shows a subjective evaluation for the proposed post-filtering process in the noisy 'Cone' images. We added AWGN (additive white Gaussian noise) with a mean of 0 and various standard deviation σ (5−20) amounts to the standard stereo image pairs. Fig. 12 shows the results of the post-filtering method for the HBP only. There were still some problems in the disparity maps of the HBP, although stereo matching was done in the stereo images denoised by NL-means filtering. We found that the disparities of invalid pixels were replaced with reasonable values and the discontinuity localization on the object boundaries was improved in the iterative filtering scheme.

The other results in the noisy stereo images are shown in Fig. 13. We found that the proposed post-filtering method also improved the quality of the disparity maps in noisy situations. Fig. 14 shows an objective comparison of the proposed post-filtering method for noisy stereo images. We evaluated the performance of the proposed method by measuring the percentage of bad matching pixels of the disparity maps which were obtained by the HBP and the HBP+PostFil. We found that the proposed method improved the disparity maps significantly in the nonocc, all, and disc regions.

5. Conclusions

In this paper, we have proposed an asymmetric post-processing method for stereo matching. The iterative filtering scheme improves stereo matching performance by filtering the costs with an invalid pixel map. Since the asymmetric information is used for estimating the invalid pixel map, a minimum number of additional computational loads are necessary for the consistency check. Valid pixel information is propagated into the invalid pixels by adaptive filtering. The proposed method can be used in various applications. The experimental results show that the proposed post-filtering method improved the performance of the HBP and the SGM, and especially the performance of the post-processing method for the HBP is shown to be comparable to state-of-the-art methods in the Middlebury stereo datasets. We compared the asymmetric consistency check with other reliability estimation methods in the post-processing scheme. The proposed method also improved stereo matching performance for noisy stereo images. The iterative filtering scheme with valid pixels significantly reduces the number of errors when the disparity maps are estimated from the denoised images.

Acknowledgments

This work was financially supported in part by the MKE, Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA, and was partially supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

References

[1] D. Scharstein, R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, International Journal of Computer Vision 47 (1–3) (April 2002) 7–42.
[2] J. Sun, Y. Li, S. Kang, H. Shum, Symmetric stereo matching for occlusion handling, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2005, pp. 399–406.
[3] J.Y. Chang, K.M. Lee, S.U. Lee, Stereo matching using iterative reliable disparity map expansion in the color-spatial-disparity space, Pattern Recognition 40 (12) (December 2007) 3705–3713.
[4] V. Kolmogorov, R. Zabih, Computing visual correspondence with occlusions using graph cuts, in: Proceedings of IEEE International Conference on Computer Vision, 2001, pp. 508–515.
[5] A. Klaus, M. Sormann, K. Karner, Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure, in: Proceedings of IEEE International Conference on Pattern Recognition, 2006, pp. 15–18.
[6] M. Bleyer, M. Gelautz, A layered stereo algorithm using image segmentation and global visibility constraints, in: Proceedings of IEEE International Conference on Image Processing, 2004, pp. 2997–3000.
[7] S. Mattoccia, F. Tombari, L. Stefano, Stereo vision enabling precise border localization within a scanline optimization framework, in: Proceedings of Asian Conference on Computer Vision, 2007, pp. 517–527.
[8] H. Hirschmuller, Stereo vision in structured environments by consistent semi-global matching, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 2386–2393.
[9] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 603–619.
[10] S.B. Kang, R. Szeliski, C. Jinxjang, Handling occlusions in dense multi-view stereo, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2001, pp. 103–110.
[11] Q. Yang, R. Yang, J. Davis, D. Nister, Spatial-depth super resolution for range images, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[12] J. Kopf, M. Cohen, D. Lischinski, M. Uyttendaele, Joint bilateral upsampling, ACM Transactions on Graphics 26 (3) (2007).
[13] D. Min, J. Oh, K. Sohn, Asymmetric post-processing for stereo correspondence, in: Proceedings of IEEE International Conference on Pattern Recognition, 2008.
[14] B.K. Muhlmann, D. Maier, J. Hesser, R. Manner, Calculating dense disparity maps from color stereo images, an efficient implementation, International Journal of Computer Vision 47 (1–3) (2002) 79–88.
[15] P. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J.-M. Frahm, R. Yang, D. Nister, M. Pollefeys, Real-time visibility-based fusion of depth maps, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[16] K. Yoon, I. Kweon, Adaptive support-weight approach for correspondence search, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (4) (April 2006) 650–656.
[17] P.F. Felzenszwalb, D.P. Huttenlocher, Efficient belief propagation for early vision, International Journal of Computer Vision 70 (1) (2006) 41–54.
[18] J. Sun, N. Zheng, H. Shum, Stereo matching using belief propagation, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (7) (2003) 787–800.
[19] H. Hirschmuller, Stereo processing by semi-global matching and mutual information, IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (2) (2008) 328–341.
[20] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Proceedings of IEEE International Conference on Computer Vision, 1998, pp. 839–846.
[21] Y. Heo, K.M. Lee, S.U. Lee, Simultaneous depth reconstruction and restoration of noisy stereo images using non-local pixel distribution, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1–8.
[22] A. Buades, B. Coll, J. Morel, A non-local algorithm for image denoising, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2005, pp. 60–65.

[23] A. Buades, B. Coll, J. Morel, Nonlocal image and movie denoising, International Journal of Computer Vision 76 (2) (2008) 123–139.
[24] Q. Yang, L. Wang, R. Yang, H. Stewenius, D. Nister, Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 2347–2354.
[25] D. Min, K. Sohn, Cost aggregation and occlusion handling with WLS in stereo matching, IEEE Transactions on Image Processing 17 (8) (2008) 1431–1442.
[26] L. Zitnick, S.B. Kang, Stereo for image-based rendering using image oversegmentation, International Journal of Computer Vision 75 (1) (2007) 49–65.
[27] S. Larsen, P. Mordohai, M. Pollefeys, H. Fuchs, Temporally consistent reconstruction from multiple video streams using enhanced belief propagation, in: Proceedings of IEEE International Conference on Computer Vision, 2007.
[28] Y. Taguchi, B. Wilburn, L. Zitnick, Stereo reconstruction with mixed pixels using adaptive over-segmentation, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[29] K. Yoon, I. Kweon, Stereo matching with the distinctive similarity measure, in: Proceedings of IEEE International Conference on Computer Vision, 2007.
[30] 〈http://vision.middlebury.edu/stereo〉.