Simultaneous Depth Reconstruction and Restoration of Noisy Stereo Images using Non-local Pixel Distribution

Yong Seok Heo, Kyoung Mu Lee, and Sang Uk Lee
School of EECS, ASRI, Seoul National University,151-742, Seoul, Korea
hys@diehard.snu.ac.kr, kyoungmu@snu.ac.kr, sanguk@ipl.snu.ac.kr

Abstract

In this paper, we propose a new algorithm that solves both the stereo matching and the image denoising problem simultaneously for a pair of noisy stereo images. Most stereo algorithms employ L1 or L2 intensity error-based data costs in the MAP-MRF framework by assuming the naive intensity-constancy. These data costs make typical stereo algorithms suffer from the effect of noise severely. In this study, a new robust stereo algorithm to noise is presented that performs the stereo matching and the image denoising simultaneously. In our approach, we redefine the data cost by two terms. The first term is the restored intensity difference, instead of the observed intensity difference. The second term is the non-local pixel distribution dissimilarity around the matched pixels. We adopted the NL-means (Non Local-means) algorithm for restoring the intensity value as a function of disparity. And a pixel distribution dissimilarity is calculated by using PMHD (Perceptually Modified Hausdorff Distance). The restored intensity values in each image are determined by inferring optimal disparity map at the same time. Experimental results show that the proposed algorithm is more robust and accurate than other conventional algorithms in both stereo matching and denoising.

1. Introduction

1.1. Motivation

Stereo matching algorithms have achieved excellent advances in the last several decades [14]. Most algorithms have some common assumptions. One of these assumptions is the intensity-consistency that the corresponding intensities in two images are same. Typical stereo algorithms that use L1 or L2 intensity error as the data costs produce good results for some standard test images [1]. Unfortunately, however, considerable amount of noise could exist in real images inevitably due to various causes such as light variation, image blurring, sensor noise by optical system and photo detector [2]. In spite of the fact that these noises can influence the performance of the algorithms seriously, only a few studies can be found in the literature on the noise effects in stereo matching. And current state-of-the-art stereo algorithms fail on noisy images. This problem led us to study about the noise effects on the typical stereo algorithms. In this paper, we propose a new algorithm that solves both the stereo matching problem and the image denoising problem for a pair of noisy input images, simultaneously.

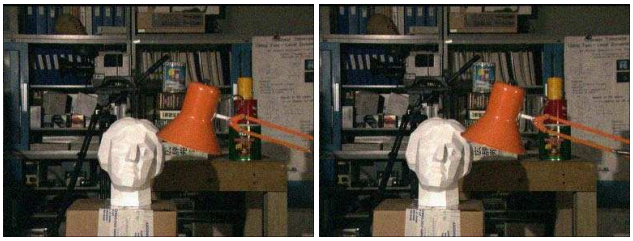
1.2. Related works

For the study on the effects of noise in stereo problem, Leclercq *et al*. [11] compared the performances of several stereo algorithms such as the Census Transform algorithm [19], window-based algorithms (such as SAD, SSD, Normalized Intensity Difference, Correlation), and the pixel-to-pixel algorithm (Birchfield and Tomasi (BT) data cost [3]) with dynamic programming (DP). They concluded that BT data cost with DP showed the best performance among the above algorithms for noisy images. Kim *et al*. [10] showed good results for severely contrast varying pairs of images using mutual information. They relaxed the constant brightness assumption and merely assume that there is a consistent relationship between corresponding scene intensities. Ogale *et al*. [12] also presented contrast-robust stereo matching algorithm using multiple frequency channels. Strecha *et al*. [15] considered the possibility of non-Lambertian effect of the same point in the scene. They proposed an Expectation Maximization based approach to solve this problem in a wide base-line stereo using multiple images with different point of views. However, they did not cover a large deviation from the Lambertian assumption.

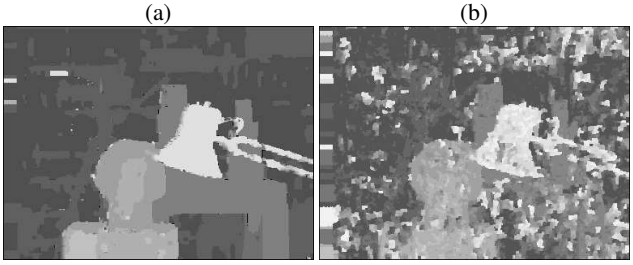
2. Standard energy minimization framework based on MAP-MRF model

The MAP-MRF (Maximum A Posteriori-Markov Random Field) is the most popular model in global energy minimization based stereo algorithms. Given a pair of rectified

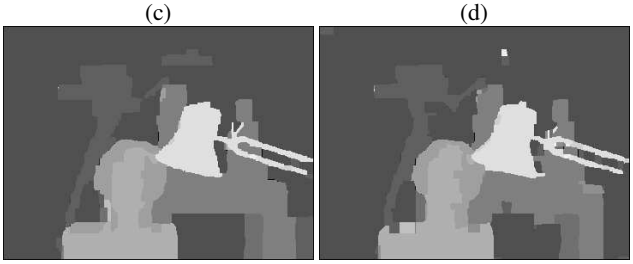
1-4244-1180-7/07/$25.00 ©2007 IEEE



(a) (b)



(c) (d)



(e) (f)

Figure 1. (a) and (b) are the left and right Tsukuba images contaminated by Gaussian noise with mean 0 and standard deviation 10. (c) is the disparity map of BT data cost [3] with belief propagation (BP) for the noise-free Tsukuba image pair. (d) is the disparity map of BT data cost with BP for the noisy Tsukuba image pair in (a) and (b). (e) is the disparity map of our algorithm for the noise-free Tsukuba image pair. (f) is the disparity map of our algorithm for the noisy Tsukuba image pair in (a) and (b).

input stereo images $I = \{I_L, I_R\}$, the goal is to find disparity maps $f = \{f_L, f_R\}$. Using the Bayesian rule, we have

$$p(f|I) = \frac{p(I|f)p(f)}{p(I)}. \quad (1)$$

And the optimal disparity values that maximize eq. (1) can be formulated by the following MAP framework.

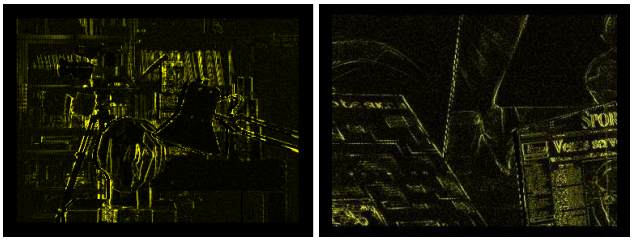
$$f_{opt} = \arg\max_f p(f|I) = \arg\max_f \frac{p(I|f)p(f)}{p(I)}, \quad (2)$$

where $p(I)$ is constant, so that we can neglect it. The relationship between energy (E) and probability (P) can be represented by the Gibbs distribution,

$$P \propto e^{-kE}. \quad (3)$$

Hence, the maximum a posteriori problem becomes equivalent to the minimization problem of following energy.

$$E(f|I) = \sum_p D_p(f_p) + \sum_p \sum_{q \in N(p), p \neq q} V(f_p, f_q), \quad (4)$$



(a) Tsukuba (b) Venus

Figure 2. Noise in the Tsukuba and Venus image

where $N(p)$ is the neighborhood pixels of p. $D_p(f_p)$ is the data cost that measures the cost when pixel p is assigned by label f_p. $V(f_p, f_q)$ is the discontinuity cost that accounts for the prior knowledge that the world consists of piecewise smooth objects. Combining these costs, therefore, the optimal disparities can be found by minimizing the total energy in eq. (4) by an inference algorithm such as belief propagation [17] and graph cuts [4].

3. Weakness of previous models for noise

Almost all pixel-to-pixel based data costs such as AD (Absolute Difference), truncated AD [16], and BT (Birchfield and Tomasi) [3] assume that corresponding intensities are consistent, on the basis of the L1 intensity difference. Various window-based methods such as SSD, SAD and adaptive window methods [9, 18], and various segment-based methods [7, 8] aggregate pixel-wise L1 or L2 differences in a proper window as the data cost. In general, noise affect the accuracy of these data costs severely as the variance of noise grows unless the noise of each pixel is eliminated. If we assume that the noise of each pixel is Gaussian, then the observed intensities in the left and right noisy images can be written as

$$I_L = S_L + n_L, I_R = S_R + n_R, \quad (5)$$

where $S = \{S_L, S_R\}$ is the noise-free intensity values, $n = \{n_L, n_R\}$ is the Gaussian noises with mean zero and standard deviations $\sigma = \{\sigma_L, \sigma_R\}$. Let us denote the intensity value at a pixel p in the left image to be $I_L(p)$. If we assume that the pixel p has true disparity f_p, then the corresponding intensity in the right image will be $I_R(p + f_p)$. The expected value of L2 error between the two intensities can be calculated by

$$\begin{aligned} &E[|I_L(p) - I_R(p + f_p)|^2] \\ &= E[|S_L(p) + n_L(p) - S_R(p + f_p) - n_R(p + f_p)|^2] \\ &= (S_L(p) - S_R(p + f_p))^2 + \sigma_L^2 + \sigma_R^2. \end{aligned} \quad (6)$$

On the other hand, the expected value of L1 error can be obtained by

$$\begin{aligned} &E\left[|I_L(p) - I_R(p+f_p)|\right] \\ &= E\left[|S_L(p) + n_L(p) - S_R(p+f_p) - n_R(p+f_p)|\right] \quad (7) \\ &\qquad = |S_L(p) - S_R(p+f_p)|\,. \end{aligned}$$

And from eq. (6) and eq. (7), the variance of L1 error becomes

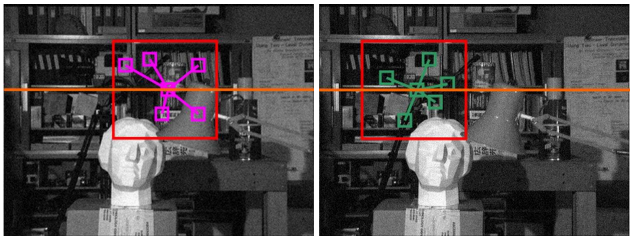
$$\begin{aligned} E[(|I_L(p) - I_R(p+f_p)| \quad &- |S_L(p) - S_R(p+f_p)|)^2] \\ &= \sigma_L^2 + \sigma_R^2. \qquad (8) \end{aligned}$$

Hence, both the mean of L2 error and the variance of L1 error increase in proportion to the noise variances. This means that when $\sigma_L \gg 0$ or $\sigma_R \gg 0$, data costs based on L1 and L2 error become no longer reliable measures for correct stereo matching. Fig. 1(c) and 1(d) show the effects of the Gaussian noise on typical stereo data cost (BT) [3] with belief propagation (BP). While this classical method is very sensitive to such noise, Fig. 1(e) and 1(f) show that our proposed algorithm is quite robust to the noise.

Fig. 2 shows the noisy pixels in Tsukuba and Venus images provided in [1]. These noisy pixels are estimated by warping the left image to the right image using the ground truth disparity map and then computing absolute differences of corresponding intensities. If two corresponding pixel intensities are different, we consider that these pixels have noise. We note that most noisy pixels exist at the high frequency region, and this is the reason why algorithms employing the BT data cost [3] that is insensitive to image sampling by using the linearly interpolated intensity function showed relatively better results than others on those original data set. However, if considerable amount of noise is added in the images, even this data cost no longer guarantees reliable results as shown in Fig. 1(c) and 1(d).

4. Proposed algorithm

Note that if we know the exact correspondences between input noisy images, we can restore the true intensity values more precisely using multiple samples. On the other hand, if we have noise-free images, then the performance of stereo matching algorithm will be improved. In general, however, it is a "chicken-and-egg problem" to solve both the image denoising and stereo matching problem, although human eye can do this. In this section, we describe a new algorithm that solves the stereo matching problem and image denoising problem for two noisy input images simultaneously. The novelty of this algorithm is on devising a new data cost that is robust to noise. It consists of two terms: The first term is the restored intensity difference as a function of the disparity. And the second term is the dissimilarity of non-local supporting pixel distributions around the matched pixels. We adopt NL-means (Non Local-means) algorithm [5] for the restoration of intensity values, and the PMHD (Perceptually Modified Hausdorff Distance) [13] for the calculation of dissimilarity of support pixel distributions. The total energy incorporating the proposed data cost and a symmetric discontinuity cost [18] is minimized using belief propagation. The restored intensity values in each image are determined by inferring the optimal disparity map, and the denoised images are used for inferring the more correct disparity map iteratively.



(a) (b)

Figure 3. The distributions of support point sets for the Tsukuba image pair

4.1. Restored intensity difference

Most local image details (in texture or color) occur multiple times in an image. This redundancy and self-similarity of an image can be used to eliminate noise. The NL-means algorithm replaces the noisy value by a weighted average of all the pixels in an image. The weight depend on the similarity between the neighborhoods of pixels. In this paper, we extended the NL-means algorithm to denoising stereo images under the stereo matching framework.

Let us denote the restored intensity of $I_L(p)$ and $I_R(p+f_p)$ to be $S_{L,p}(f_p)$ and $S_{R,p+f_p}(f_p)$, respectively. First, we find a point set $T = \{t | t \in M_p, t \neq p\}$ in the left image that have similar neighborhoods with that of the pixel p in the $M \times M$ large search window M_p as shown in Fig. 3(a). The dissimilarity between pixel p and each pixel t in the point set T is defined by

$$d(v(N_p), v(N_t)) = G_\sigma * |(v(N_p)) - (v(N_t))|^2\,, \qquad (9)$$

where N_p and N_t are $m \times m$ small comparison windows centered at pixel p and t, respectively. $v(N_k)$ denotes the 1-D vector of N_k. G_σ is a Gaussian kernel with standard deviation of σ. The pixels with similar neighborhood to $v(N_p)$ have larger weight. And this weight is defined by

$$w(p,t) = \frac{1}{Z(p)} \exp\left(-\frac{d(v(N_p), v(N_t))}{h^2}\right), \qquad (10)$$

where $Z(p)$ is the normalization factor. And $d(\cdot,\cdot)$ is the Euclidean distance between two vectors. For the computational convenience, by sorting the pixels in the search window M_p according to the weights, we select and use only

the first N highly weighted pixels among them. We denote this subset as $T_L(p)$ and call it as the *support point set* for the pixel p in the left image. To fully utilize the redundancy in two images, we warp all pixels in $T_L(p)$ to the right image according to f, and denote the corresponding set in the right image as $T_{WR}(p)$ and call it as the *warped support point set*. If two pixels $p \in I_L$ and $p + f_p \in I_R$ are the exact corresponding pixels, the restored value $S_{L,p}(f_p)$ using both $T_L(p)$ and $T_{WR}(p)$ would have less noise than using $T_L(p)$ only. Thus, let us define the restored intensity at $p \in I_L$ to be the weighted average of all the pixels in both $T_L(p)$ and $T_{WR}(p)$ as follows.

$$S_{L,p}(f_p) = \frac{1}{2} \sum_{t_l \in T_L(p)} w(p, t_l) I_L(t_l) + \frac{1}{2} \sum_{t_r \in T_{WR}(p)} w(p, t_r) I_R(t_r). \tag{11}$$

Similarly and symmetrically, the restored value of pixel $p + f_p \in I_R$ is defined by

$$S_{R,p+f_p}(f_p) = \frac{1}{2} \sum_{t_r \in T_R(p+f_p)} w(p + f_p, t_r) I_R(t_r) + \frac{1}{2} \sum_{t_l \in T_{WL}(p+f_p)} w(p + f_p, t_l) I_L(t_l). \tag{12}$$

Now, the difference of the two corresponding restored intensity values at $p \in I_L$ and $p + f_p \in I_R$ with given f_p can be defined by

$$\Delta S(p, f_p) = \left| S_{L,p}(f_p) - S_{R,p+f_p}(f_p) \right|. \tag{13}$$

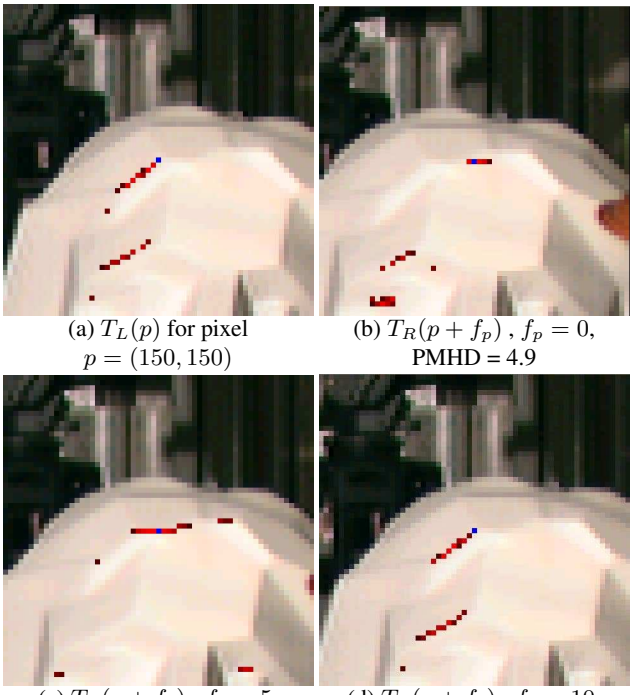
We note that if f_p is correct, then

$$\Delta S(p, f_p) \approx |S_L(p) - S_R(p + f_p)| \approx 0.$$

Thus, the minimization of the restored intensity difference in eq. (13) can lead us to achieve both image denoising and correct stereo matching.

4.2. Non-local pixel distribution dissimilarity

In the previous subsection, we have constructed the *support point set* $T_L(p)$ for pixel $p \in I_L$ and $T_R(p + f_p)$ for pixel $p+f_p \in I_R$ that have non-local pixel distributions. We note that if f_p is correct, then the geometric configurations of $T_L(p)$ and $T_R(p+f_p)$ will be similar. Thus, the similarity of the geometric configurations of these two support point sets also can be a good measure for how well the two pixels p and $p + f_p$ correspond to each other. In general, Hausdorff distance (HD) is known to be an effective and popular metric for the dissimilarity measure between two point sets. However, the limitation of the conventional HD is that it considers the distance between two points only regardless



(a) $T_L(p)$ for pixel $p = (150, 150)$ (b) $T_R(p + f_p)$, $f_p = 0$, PMHD = 4.9 (c) $T_R(p + f_p)$, $f_p = 5$, PMHD = 12.7 (d) $T_R(p + f_p)$, $f_p = 10$, PMHD = 0.3

Figure 4. (a) The distribution of support points for pixel $p = (150, 150)$ in the Tsukuba left image, and (b),(c),(d) show the distributions of support points and the corresponding PMHD values for pixel $p + f_p$ in the Tsukuba right image when $f_p = 0, 5, 10$, respectively. At correct disparity $f_p = 10$, the geometric configurations of $T_L(p)$ and $T_R(p + f_p)$ becomes almost the same and the corresponding PMHD value is minimum.

of the weight or mass of each point. Thus, in our case, since each pixel in a support point set has its own weight, the conventional HD does not work properly. Recently, PMHD (Perceptually Modified Hausdorff Distance) [13] has been proposed which accounts for the weight of each point in the sets. Adopting this PMHD measure, we can calculate the dissimilarity between two weighted support point sets of

$$\begin{aligned} T_L(p) &= \{(t_l, w(p, t_l)) | l = 1, ..., N_l\}, \\ T_R(p + f_p) &= \{(t_r, w(p + f_p, t_r)) | r = 1, ..., N_r\}. \end{aligned} \tag{14}$$

by

$$\begin{aligned} &D_H(T_L(p), T_R(p + f_p)) \\ &= Max\{d_H(T_L(p), T_R(p + f_p)), \\ &\qquad d_H(T_R(p + f_p), T_L(p))\}, \end{aligned} \tag{15}$$

where $d_H(\cdot, \cdot)$ is the directed PMHD defined by

$$\begin{aligned} &d_H(T_L(p), T_R(p + f_p)) \\ &= \frac{\sum_l \left[w(p, t_l) \times \min_r \frac{d(t_l, t_r)}{\min(w(p,t_l), w(p+f_p,t_r))} \right]}{\sum_l w(p, t_l)}, \end{aligned} \tag{16}$$

where $d(t_l, t_r)$ is the Euclidean distance between two points. Using PMHD, the support point distribution dissimilarity between the pixel p in the left image and $p + f_p$ in the right image can be defined by

$$\Delta G(p, f_p) = D_H(T_L(p), T_R(p + f_p)). \quad (17)$$

Fig. 4 shows the variation of the support points distribution according to the disparity. Fig. 4(a) represents the set of support points $T_L(p)$ for the pixel at $p = (150, 150)$ in the left image. Fig. 4(b)-(d) show the corresponding sets of supporting points for the pixel $p + f_p$ in the right image according to the disparity value f_p. As we can see from these figures, at the correct disparity $f_p = 10$, the geometric configuration between $T_L(p)$ and $T_R(p + f_p)$ becomes almost the same and the corresponding PMHD value gets its minimum value.

4.3. Data cost incorporating restored intensity difference and support pixel distribution dissimilarity

Now, let us combine the restored intensity error in eq. (13) and the support points distribution dissimilarity in eq. (17) for our new data cost as follows.

$$D_p(f_p) = -\ln((1 - e_d) \quad \exp(-(\frac{\Delta S(p,f_p)}{\sigma_s} + \frac{\Delta G(p,f_p)}{\sigma_g})) + e_d), \quad (18)$$

where we adopted the robust function [17] to make it more robust to outliers.

4.4. Discontinuity cost

In this paper, we employed following discontinuity cost between neighboring pixels p and q [18].

$$V(f_p, f_q) = \begin{cases} 0 & if\ f_p = f_q \\ \rho_{f_p}(\Delta C_l, \Delta C_r) & otherwise \end{cases} \quad (19)$$

$$\rho_{f_p}(\Delta C_l, \Delta C_r) = \begin{cases} P_l \times P_r \times s & if\ \Delta C_l < T, \Delta C_r < T \\ P_l \times s & if\ \Delta C_l < T, \Delta C_r \geqslant T \\ P_r \times s & if\ \Delta C_l \geqslant T, \Delta C_r < T \\ s & otherwise \end{cases} \quad (20)$$

where ΔC_l is the magnitude of intensity gradient between p and q in the left image and ΔC_r is that between $p + f_p$ and $q + f_q$ in the right image. T is the threshold value for intensity gradients ΔC_l and ΔC_r to impose discontinuity cost. P_l, P_r and s are the penalty terms that used for determining the total discontinuity cost according to the intensity gradient values ΔC_l and ΔC_r. The penalty increases as the intensity gradient decreases. The intensity gradient is calculated by the restored intensity values that rely on the disparity.

4.5. Global energy modelling

Now, the global energy can be defined by

$$E(f) = E_{data}(f) + E_{discontinuity}(f), \quad (21)$$

where

$$E_{data}(f) = \sum_p D_p(f_p), \quad (22)$$

and

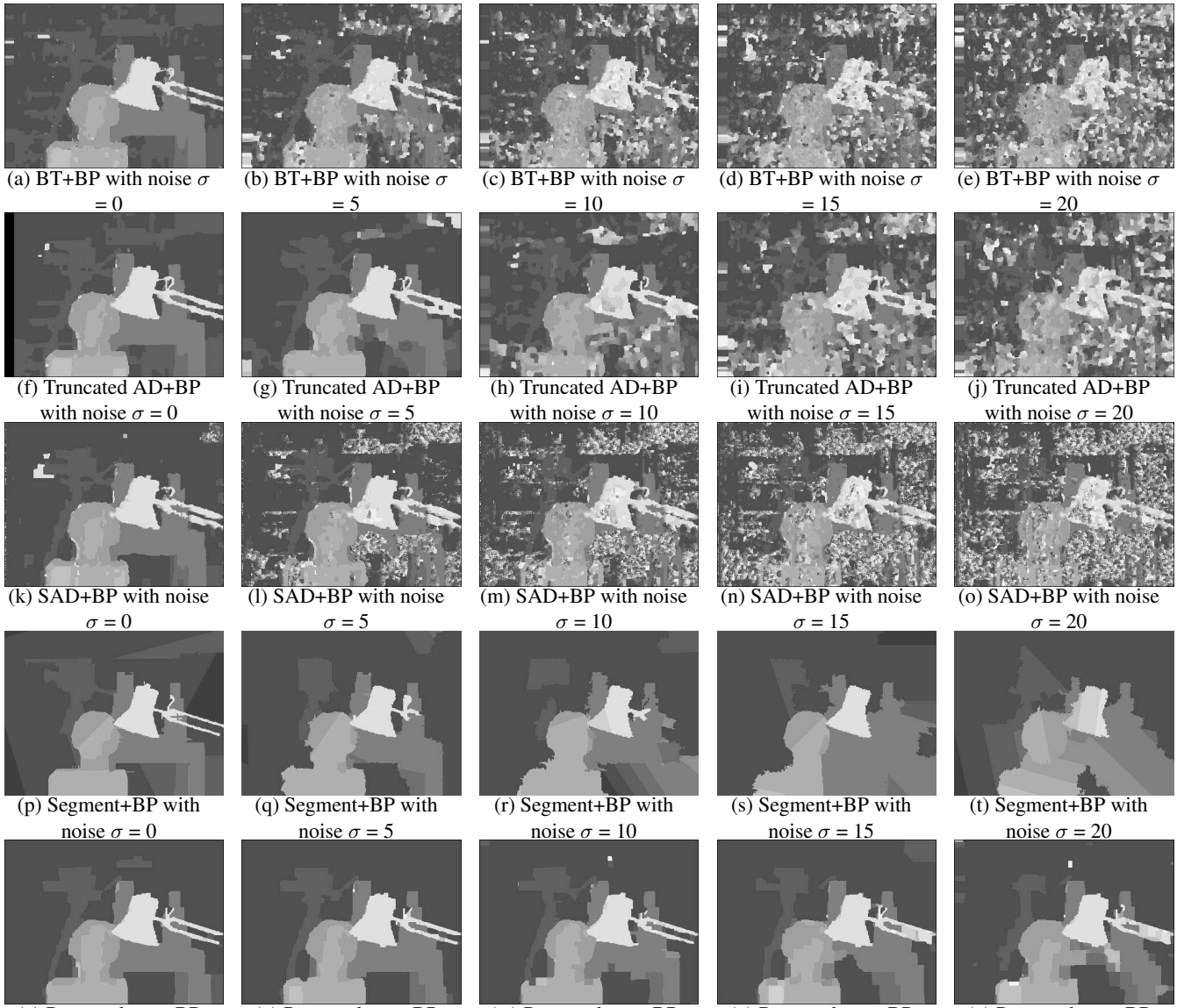
$$E_{discontinuity}(f) = \sum_p \sum_{q \in N(p)} V(f_p, f_q). \quad (23)$$

The total energy $E(f)$ can be minimized by an optimization algorithm such as belief propagation. Note that the restored intensity values in each image are determined simultaneously by inferring the optimal disparity map using eqs.(11) and (12).

5. Experimental results

In our experiments, we fixed all the parameters of the proposed algorithm such that $M_p = (61 \times 61)$, $N_p = (3 \times 3)$, $N = 200$, $\sigma_s = 6.0$, $\sigma_g = 100$, $e_d = 0.01$. And for the discontinuity cost, we set $P_r = P_t = 2.0$, $T = 8.5$, and $s = 1.8$. We have tested and compared the performance of our algorithm with others such as BT data cost with BP (BT+BP), truncated AD data cost with BP (Truncated AD+BP), SAD (5 by 5 window) data cost with BP (SAD+BP), mean-shift segmentation [6] and plane fitting based data cost (similar to [7, 8]) with BP (Segment+BP) on the standard test image set in [1]. We added Gaussian noises with mean 0 and various standard deviation σ to this test image set with the same amount of noise in both left and right images.

Fig. 5 compares the performance of various stereo algorithms with our algorithm for Tsukuba image by varying the noise variance. And Fig. 6 shows the quantitative comparison results (the percentage of bad pixels) of the reconstructed disparities on several other test stereo images. We see from those figures that the proposed algorithm produces the best results in both qualitative and quantitative evaluations, while most algorithms based on L1 and L2 error suffer seriously as the noise variance gets increased. Interestingly, the segmentation-based BP algorithm gives better performance among others. We think that this is due to the somewhat robust property of the mean-shift and plane-fitting procedures to noise in the segmentation step. Fig. 7 shows more disparity results by our algorithm on Venus, Teddy and Cone images with varying amount of noise. In order to verify the effectiveness of the simultaneous reconstruction and denoising of noisy stereo images by our method, we have compared our results with those of a naive approach in which the image denoising was performed first



(a) BT+BP with noise σ = 0 (b) BT+BP with noise σ = 5 (c) BT+BP with noise σ = 10 (d) BT+BP with noise σ = 15 (e) BT+BP with noise σ = 20

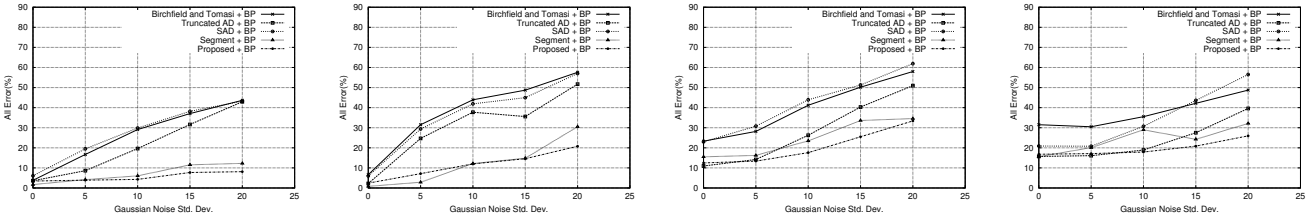
(f) Truncated AD+BP with noise σ = 0 (g) Truncated AD+BP with noise σ = 5 (h) Truncated AD+BP with noise σ = 10 (i) Truncated AD+BP with noise σ = 15 (j) Truncated AD+BP with noise σ = 20

(k) SAD+BP with noise σ = 0 (l) SAD+BP with noise σ = 5 (m) SAD+BP with noise σ = 10 (n) SAD+BP with noise σ = 15 (o) SAD+BP with noise σ = 20

(p) Segment+BP with noise σ = 0 (q) Segment+BP with noise σ = 5 (r) Segment+BP with noise σ = 10 (s) Segment+BP with noise σ = 15 (t) Segment+BP with noise σ = 20

(u) Proposed cost+BP with noise σ = 0 (v) Proposed cost+BP with noise σ = 5 (w) Proposed cost+BP with noise σ = 10 (x) Proposed cost+BP with noise σ = 15 (y) Proposed cost+BP with noise σ = 20

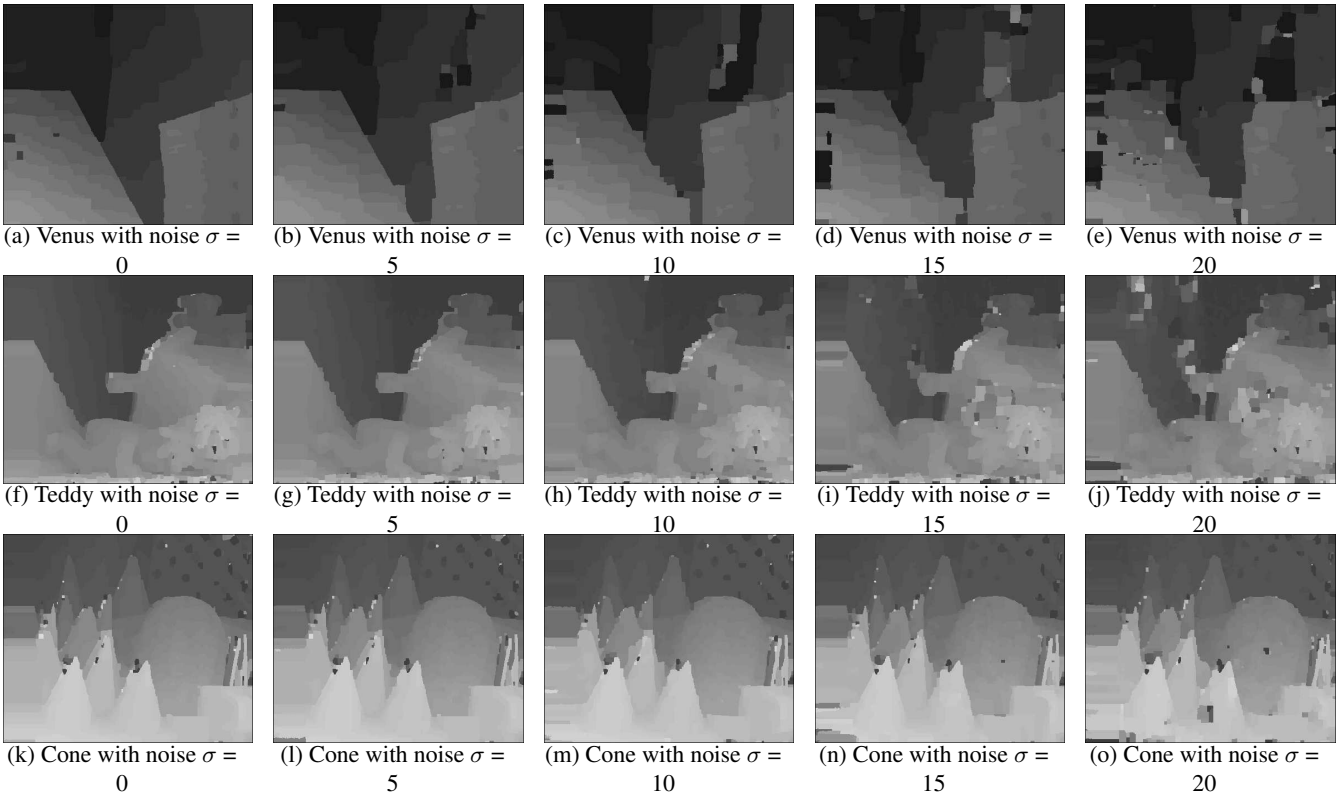
Figure 5. Results of test stereo algorithms on noisy Tsukuba image pair with varying standard deviation of noise σ.



(a) Tsukuba (b) Venus (c) Teddy (d) Cone

Figure 6. Quantitative comparison of the stereo matching results on several test images according to the noise standard deviation σ

and then stereo matching followed. We used the Tsukuba images contaminated by Gaussian noise with mean 0 and standard deviation 10 as the test images. The PSNR of both left and right noisy images were 28.38dB. These noisy images were filtered independently by the NL-means algorithm, and after denoising, the PSNR of the left and right



(a) Venus with noise σ = 0 (b) Venus with noise σ = 5 (c) Venus with noise σ = 10 (d) Venus with noise σ = 15 (e) Venus with noise σ = 20

(f) Teddy with noise σ = 0 (g) Teddy with noise σ = 5 (h) Teddy with noise σ = 10 (i) Teddy with noise σ = 15 (j) Teddy with noise σ = 20

(k) Cone with noise σ = 0 (l) Cone with noise σ = 5 (m) Cone with noise σ = 10 (n) Cone with noise σ = 15 (o) Cone with noise σ = 20

Figure 7. Reconstructed disparity maps of noisy Venus, Teddy and Cone images by the proposed algorithm

(a) BT+BP using denoised images (b) Truncated AD+BP using denoised images (c) SAD+BP using denoised images (d) Segment+BP using denoised images (e) Our proposed cost+BP

Figure 8. Comparison of "independent denoising and stereo matching schemes" (a)-(d) and the proposed algorithm (e)

images increased to 33.01 dB and 33.04 dB, respectively. Then, to these denoised images, conventional stereo algorithms were applied. Fig. 8(a)-(d) illustrate the results of the independent denoising and stereo matching procedures by BT+BP, Truncated AD+BP, SAD+BP and Segment+BP, respectively. By comparing with the result of our method that incorporates image denoising in stereo matching framework as shown in Fig. 8(e), we can convince the validity of our approach over the sequential approach. Moreover, regarding the performance of image denoising, the denoised images using our method had also even higher PSNR values (left one was 33.78 dB and right one was 33.79 dB) than those of the denoised images using each single image only. Table 1 compares the PSNR values of denoised (left) images by the NL-means algorithm using single image with those

(a) Original (b) NL-m (c) Ours

Figure 9. (a) input left image (noise σ = 20), (b) denoised image by the NL-means method, (c) denoised image by proposed method

by proposed method. Fig. 9 shows the denoised images for the Tsukuba image for noise $\sigma = 20$. We notice that the proposed method restores the stereo images reliably even under severe noise conditions.

noise σ	Tsukuba			Venus			Teddy			Cone		
	Input	NL-m	Ours	Input	NL-m	Ours	Input	NL-m	Ours	Input	NL-m	Ours
5	34.23	34.19	34.51	34.17	31.8	33.42	34.2	32.5	32.6	34.2	31.7	31.7
10	28.38	32.01	33.78	28.20	32.3	32.83	28.23	31.85	32.31	28.2	31.11	31.4
15	25.02	30.19	32.03	24.78	29.81	31.53	24.7	29.48	31.5	24.7	28.90	30.4
20	22.68	28.40	30.06	22.38	26.63	29.67	22.3	26.36	29.56	22.2	25.96	28.56

Table 1. PSNRs [dB] of the input left noisy image (Input) and denoised left images by the conventional NL-means algorithm using single image (NL-m) and those by the proposed algorithm (Ours)

6. Conclusion

In this paper we have addressed the simultaneous reconstruction of disparity and image denoising problem for noisy stereo images, and proposed a novel algorithm to solve it. In the MAP-MRF framework, we defined new data cost that consisted of the restored intensity error term and the support point distribution dissimilarity term. We have generalized the NL-means algorithm to stereo images for denoising and employed PMHD for the measure of dissimilarity between support point sets. Comparative experimental results on standard test images demonstrated that the proposed algorithm achieved better results in not only stereo matching, but also image denoising than conventional approaches.

Acknowledgements

This work was supported in part by the ITRC program by Ministry of Information and Communication and in part by the Agency for Defense Development, through the Image Information Research Center, at KAIST, Korea.

References

[1] http://cat.middlebury.edu/stereo/.

[2] F. Alter, Y. Matsushita, and X. Tang. An intensity similarity measure in low-light conditions. In *Proc. of European Conference on Computer Vision*, 2006.

[3] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(4):401–406, 1998.

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.

[5] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[6] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2001.

[7] Y. Deng, Q. Yang, X. Lin, and X. Tang. A symmetric patch-based correspondence model for occlusion handling. In *Proc. of International Conference on Computer Vision*, 2005.

[8] L. Hong and G. Chen. Segment-based stereo matching using graph cuts. In *Proc. of Computer Vision and Pattern Recognition*, 2004.

[9] T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window: Theory and experiments. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16(9):920–932, 1994.

[10] J. Kim, V. Kolmogorov, and R. Zabih. Visual correspondence using energy minimization and mutual information. In *Proc. of International Conference on Computer Vision*, 2003.

[11] P. Leclercq and J. Morris. Robustness to noise of stereo matching. In *Proc. of International Conference on Image Analysis and Processing*, 2003.

[12] A. S. Ogale and Y. Aloimonos. Robust contrast invariant stereo correspondence. In *Proc. of IEEE Conf. on Robotics and Automation*, 2004.

[13] B. G. Park, K. M. Lee, and S. U. Lee. A new similarity measure for random signatures: Perceptually modified hausdorff distance. In *Proc. of Advanced Concepts for Intelligent Vision Systems*, 2006.

[14] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, 2002.

[15] C. Strecha, R. Fransens, and L. V. Gool. Wide-baseline stereo from multiple views: a probabilistic account. In *Proc. of Computer Vision and Pattern Recognition*, 2004.

[16] J. Sun, Y. Li, S. B. Kang, and H.-Y. Shum. Symmetric stereo matching for occlusion handling. In *Proc. of Computer Vision and Pattern Recognition*, 2005.

[17] J. Sun, N.-N. Zheng, and H.-Y. Shum. Stereo matching using belief propagation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(7):787–800, 2003.

[18] K. Yoon and I. Kweon. Stereo matching with symmetric cost functions. In *Proc. of Computer Vision and Pattern Recognition*, 2006.

[19] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proc. of European Conference on Computer Vision*, 1994.