

## Research Article

# Efficient Stereo Matching with Decoupled Dissimilarity Measure Using Successive Weighted Summation

Cheng-Tao Zhu,<sup>1</sup> Yau-Zen Chang,<sup>2</sup> Huai-Ming Wang,<sup>2</sup> Kai He,<sup>1</sup>  
Shih-Tseng Lee,<sup>3</sup> and Chung-Fu Lee<sup>2</sup>

<sup>1</sup> School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China

<sup>2</sup> Department of Mechanical Engineering, Chang Gung University, Taoyuan 33302, Taiwan

<sup>3</sup> Department of Neurosurgery, Chang Gung Memorial Hospital, Taoyuan 33305, Taiwan

Correspondence should be addressed to Yau-Zen Chang; zen@mail.cgu.edu.tw

Received 31 October 2013; Accepted 16 December 2013; Published 16 January 2014

Academic Editor: Yi-Hung Liu

Copyright © 2014 Cheng-Tao Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Developing matching algorithms from stereo image pairs to obtain correct disparity maps for 3D reconstruction has been the focus of intensive research. A constant computational complexity algorithm to calculate dissimilarity aggregation in assessing disparity based on separable successive weighted summation (SWS) among horizontal and vertical directions was proposed but still not satisfactory. This paper presents a novel method which enables decoupled dissimilarity measure in the aggregation, further improving the accuracy and robustness of stereo correspondence. The aggregated cost is also used to refine disparities based on a local curve-fitting procedure. According to our experimental results on Middlebury benchmark evaluation, the proposed approach has comparable performance when compared with the selected state-of-the-art algorithms and has the lowest mismatch rate. Besides, the refinement procedure is shown to be capable of preserving object boundaries and depth discontinuities while smoothing out disparity maps.

## 1. Introduction

Stereo vision is the technique of constructing a 3D description of the scene from stereo image pairs, which is important in many computer vision tasks such as inspection [1], 3D object recognition [2], robot manipulation [3], and autonomous navigation [4]. Stereo vision systems can be active or passive. Active techniques utilize ultrasonic transducers and structured light or laser to simplify the stereo matching problem. On the other hand, passive stereo vision based only on stereo image pairs is less intrusive and typically able to provide a compact and affordable solution for range sensing.

For passive stereo vision systems, stereo matching algorithms are crucial for correct and accurate depth estimation, which find for each pixel in one image the corresponding pixel in the other image. A 2D picture of displacements between corresponding pixels of a stereo image pair is named as a disparity map [5].

Reference [6] is an intensively cited classification of stereo matching algorithms for rectified image pairs. The paper divides most of the algorithms into four sequential parts: matching cost calculation, cost aggregation, disparity computation, and disparity refinement. Among the steps, cost aggregation determines the performance of an algorithm in terms of computational complication and correctness.

Cost aggregation can be local [7–12] or global [13–16], based on differences in the range of supporting regions or windows. Global methods assume that the scene is piecewise smooth and search for disparity assignments over the whole stereo pair [6], which requires high computational operation. The local methods, also known as window-based, typically require less memory and computation. As a result, the window-based algorithms are popular for fast disparity calculations [17].

Local methods tend to be sensitive to noise; however, and its correctness at regions with sparse texture or near depth

discontinuities relies on proper selection of window size. To overcome this problem, [7] proposed variable windows for matching calculation, while [8] proposed multiple windows to enhance correctness at regions near depth discontinuities. Nevertheless, performance of these approaches is limited, since same aggregation weights are applied over the windows.

Recent years have seen adaptive support weight approaches [9] to improve quality of disparity maps. Unfortunately, these approaches require independent support weights calculation for each pixel and dramatically increase computational complexity.

To simplify computation, [10] introduced joint histogram to reduce the search region of disparity and [11] proposed the usage of a sparse Census mask. A summed normalized cross-correlation was proposed in [12] to calculate matching cost in two stages. Segmentation and plane fitting on disparity planes [13–16] are also popular to improve accuracy of disparity, but the performance relies on correctness in both segmentation and plane fitting.

An effective local stereo matching algorithm is introduced in [18], which significantly simplifies the intensity-dependent aggregation procedure of local methods. The algorithm aggregates cost values effectively in terms of bilateral filtering by only four passes along regions, called separable successive weighted summation (SWS), eliminating iteration and support area dependency. However, the dissimilarity measures are coupled, which significantly restrict the flexibility in weighting the aggregated costs.

In this paper, we present an improved stereo matching algorithm. Similar to [18], our algorithm uses whole regions as matching primitives to assess disparity based on SWS among horizontal and vertical directions. We also use the basic metrics, such as truncated sum of absolute difference and truncated absolute gradient difference, as dissimilarity measure to provide a trade-off between accuracy and complexity.

The main contribution of this paper is to afford a decoupled aggregation algorithm to access the stereo matching cost under the framework of SWS. The algorithm is simply yet efficient as well as robust. In addition, the resultant disparity map is in a discrete space, which is unsuitable for image-based rendering. We propose a subpixel refinement technique that employs inferior candidate disparities, rather than spatial neighbors, to smooth out discrete values in the disparity map. By this arrangement for curve-fitting, even regions near discontinuous depth can be correctly refined. Moreover, this technique increases the resolution of a stereo algorithm with marginal additional computation.

## 2. Aggregation Algorithm Design

Our stereo matching algorithm consists of three main stages. First, initial cost values are calculated based on dissimilarity measure between pixels in the reference and target images, and the costs are aggregated using the proposed method. Second, we perform initial disparity estimation by the use of a winner-takes-all minimum search based on the aggregated costs. Third, we check differences between the disparity values of corresponding pixel pairs for the existence of

obscured regions and patch them by the smallest disparity values of nearby regions. Finally, the disparity map is refined by a proposed curve-fitting procedure.

*2.1. Cost Definition.* Assuming that the image pair is rectified and horizontally aligned, two dissimilarity measures between the pixels on the reference image and the target image are used in this work.

The truncated absolute difference cost,  $e_D$ , is defined as the absolute difference between the intensity of pixels with the corresponding pixel being shifted  $d$  pixels, along horizontal direction on the reference image:

$$e_D(x, y, d) = \min \left[ \sum_{n=1}^3 |I_R(x, y, n) - I_T(x + d, y, n)|, h_D \right], \quad (1)$$

where  $I_R$  and  $I_T$  are the intensities of the pixels on reference and target images with  $n = 1, 2, 3$  corresponding to three-channel colors: R, G, and B, respectively,  $(x, y)$  are pixel coordinates, and  $h_D$  is a threshold value for  $e_D$  with  $h_D > 0$ . The use of a threshold to restrict cost value has been a well-adopted practice to reduce the effects of noise and potential mismatch in obscured regions.

Besides, the truncated absolute gradient difference cost is defined as

$$e_G(x, y, d) = \min \left[ \sum_{n=1}^3 |\nabla_x I_R(x, y, n) - \nabla_x I_T(x + d, y, n)| + \sum_{n=1}^3 |\nabla_y I_R(x, y, n) - \nabla_y I_T(x + d, y, n)|, h_G \right], \quad (2)$$

where  $\nabla_x$  and  $\nabla_y$  are the horizontal and vertical gradient operators, respectively, and  $h_G$  is a threshold value for  $e_G$  with  $h_G > 0$ .

*2.2. Cost Aggregation.* Aggregation of primary costs determines correctness and accuracy of disparity estimation [19–21]. Based on the observation that costs, such as the absolute difference and the absolute gradient difference, represent different dissimilarity characteristics, they should be independently weighted in aggregation. This section proposes a new method that is compatible with separable successive weighted summation (SWS) [18] while efficiently provides decoupled dissimilarity aggregation for robust stereo matching.

Once the  $e_D$  and  $e_G$  cost measures are obtained, as was defined in the last subsection, the aggregated cost function  $\bar{C}$  is set as the weighted sum of both measures according to a weighting factor,  $\lambda$ :

$$\begin{aligned} \bar{C}(x, y, d) &= \lambda \sum_{i=1}^m \sum_{j=1}^n W_{ix} W_{jy} e_D(i, j, d) \\ &\quad + (1 - \lambda) \sum_{i=1}^m \sum_{j=1}^n U_{ix} U_{jy} e_G(i, j, d) \\ &\triangleq \lambda \widehat{C}_D(x, y, d) + (1 - \lambda) \widehat{C}_G(x, y, d). \end{aligned} \quad (3)$$

In (3), definition of the weightings,  $W_{ix}$ ,  $W_{jy}$ ,  $U_{ix}$ , and  $U_{jy}$ , are based on the operational principles of bilateral filters [22] to dramatically reduce computation:

$$W_{ix} = \begin{cases} \prod_{k=i+1}^x \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(k, y, n) - I_T(k-1, y, n)]^2}}{-2\alpha}\right), & i < x \\ 1, & i = x \\ \prod_{k=x+1}^i \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(k, y, n) - I_T(k-1, y, n)]^2}}{-2\alpha}\right), & i > x \end{cases} \quad (4)$$

$$W_{jy} = \begin{cases} \prod_{k=j+1}^y \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(x, k, n) - I_T(x, k-1, n)]^2}}{-2\alpha}\right), & j < y \\ 1, & j = y \\ \prod_{k=y+1}^j \exp\left(\frac{\sqrt{\sum_{n=1}^3 [I_R(x, k, n) - I_T(x, k-1, n)]^2}}{-2\alpha}\right), & j > y \end{cases} \quad (5)$$

$$U_{ix} = \begin{cases} \prod_{k=i+1}^x \exp\left(\frac{\sqrt{\sum_{n=1}^3 [\nabla_x I_R(k, y, n) - \nabla_x I_T(k-1, y, n)]^2}}{-2\beta}\right), & i < x \\ 1, & i = x \\ \prod_{k=x+1}^i \exp\left(\frac{\sqrt{\sum_{n=1}^3 [\nabla_x I_R(k, y, n) - \nabla_x I_T(k-1, y, n)]^2}}{-2\beta}\right), & i > x \end{cases} \quad (6)$$

$$U_{jy} = \begin{cases} \prod_{k=j+1}^y \exp\left(\frac{\sqrt{\sum_{n=1}^3 [\nabla_y I_R(x, k, n) - \nabla_y I_T(x, k-1, n)]^2}}{-2\beta}\right), & j < y \\ 1, & j = y \\ \prod_{k=y+1}^j \exp\left(\frac{\sqrt{\sum_{n=1}^3 [\nabla_y I_R(x, k, n) - \nabla_y I_T(x, k-1, n)]^2}}{-2\beta}\right), & j > y. \end{cases} \quad (7)$$

The values of  $W_{ix}$  and  $W_{jy}$  increase with smaller  $e_D$ , while the values of  $U_{ix}$  and  $U_{jy}$  rise with lesser  $e_G$ . Also, the weightings are the multiples of horizontal and vertical weightings, which decrease with the increase of distances to the reference pixels. Hence, for each pixel, neighboring pixels with similar intensity have higher support during the aggregation.

The aggregation is a two-dimensional convolution. To reduce the computational complexity, each convolution is further decomposed into four one-dimensional convolutions [18]. These one-dimensional convolutions operate from left to right, from right to left, from top to bottom, and from bottom to top, respectively.

Let us take the absolute difference part of  $\bar{C}$ , denoted as  $\widehat{C}_D$  in (3), as an example. If we define the left-to-right weighted sum,  $C_D^{L2R}$ , as

$$C_D^{L2R}(x, j, d) \triangleq \sum_{i=1}^{x-1} W_{ix} e_D(i, j, d) + e_D(x, j, d), \quad (8)$$

from the definition of  $W_{ix}$ , (4):

$$W_{ix} = W_{i(x-1)} W_{(x-1)x}, \quad (9)$$

we have that  $C_D^{L2R}$  can be written in a recursive form

$$\begin{aligned} C_D^{L2R}(x, j, d) &= \sum_{i=1}^{x-2} W_{ix} e_D(i, j, d) + W_{(x-1)x} e_D(x-1, j, d) \\ &\quad + e_D(x, j, d) \\ &= W_{(x-1)x} \left[ \sum_{i=1}^{x-2} W_{i(x-1)} e_D(i, j, d) \right. \\ &\quad \left. + e_D(x-1, j, d) \right] + e_D(x, j, d) \\ &= W_{(x-1)x} C_D^{L2R}(x-1, j, d) + e_D(x, j, d). \end{aligned} \quad (10)$$

Similarly, we may define the right-to-left weighted sum,  $C_D^{R2L}$ , as

$$C_D^{R2L}(x, j, d) \triangleq \sum_{i=x+1}^m W_{ix} e_D(i, j, d) + e_D(x, j, d), \quad (11)$$

which can also be written in a recursive form

$$C_D^{R2L}(x, j, d) = W_{x(x+1)} C_D^{R2L}(x+1, j, d) + e_D(x, j, d). \quad (12)$$

We have that

$$\begin{aligned}
\widehat{C}_D(x, y, d) &= \sum_{j=1}^n \sum_{i=1}^m W_{jy} W_{ix} e_D(i, j, d) \\
&= \sum_{j=1}^n W_{jy} \left[ \sum_{i=1}^m W_{ix} e_D(i, j, d) \right] \\
&= \sum_{j=1}^n W_{jy} \left[ \sum_{i=1}^{x-1} W_{ix} e_D(i, j, d) + e_D(x, j, d) \right. \\
&\quad \left. + \sum_{i=x+1}^m W_{ix} e_D(i, j, d) \right] \\
&= \sum_{j=1}^n W_{jy} \left[ C_D^{L2R}(x, j, d) + C_D^{R2L}(x, j, d) \right. \\
&\quad \left. - e_D(x, j, d) \right] \\
&\triangleq \sum_{j=1}^n W_{jy} C_D^H(x, j, d).
\end{aligned} \tag{13}$$

Note that, in (13), we have defined

$$C_D^H(x, j, d) \triangleq C_D^{L2R}(x, j, d) + C_D^{R2L}(x, j, d) - e_D(x, j, d) \tag{14}$$

to simplify the following derivation.

In the vertical direction, if we define the top-to-bottom and bottom-to-top weighted sums,  $C_D^{T2B}$  and  $C_D^{B2T}$ , as

$$\begin{aligned}
C_D^{T2B}(x, y, d) &\triangleq \sum_{j=1}^{y-1} W_{jy} C_D^H(x, j, d) + e_D(x, y, d) \\
C_D^{B2T}(x, y, d) &\triangleq \sum_{j=y+1}^n W_{jy} C_D^H(x, j, d) + e_D(x, y, d),
\end{aligned} \tag{15}$$

we have that

$$\widehat{C}_D(x, y, d) = C_D^{T2B}(x, y, d) + C_D^{B2T}(x, y, d) - e_D(x, y, d). \tag{16}$$

In calculating (16), both  $C_D^{T2B}$  and  $C_D^{B2T}$  are recursively obtained as

$$\begin{aligned}
C_D^{T2B}(x, y, d) &= W_{(y-1)y} C_D^{T2B}(x, y-1, d) + e_D(x, y, d), \\
C_D^{B2T}(x, y, d) &= W_{y(y+1)} C_D^{B2T}(x, y+1, d) + e_D(x, y, d).
\end{aligned} \tag{17}$$

Hence, the first part of the aggregated cost,  $\widehat{C}_D$ , can be efficiently calculated by (10), (12), and (17).

With similar procedure, we have that

$$\widehat{C}_G(x, y, d) = C_G^{T2B}(x, y, d) + C_G^{B2T}(x, y, d) - e_G(x, y, d), \tag{18}$$

where

$$C_G^{T2B}(x, y, d) \triangleq \sum_{j=1}^{y-1} U_{jy} C_G^H(x, j, d) + e_G(x, y, d), \tag{19}$$

$$C_G^{B2T}(x, y, d) \triangleq \sum_{j=y+1}^n U_{jy} C_G^H(x, j, d) + e_G(x, y, d)$$

with

$$\begin{aligned}
C_G^H(x, j, d) &= C_G^{L2R}(x, j, d) + C_G^{R2L}(x, j, d) - e_G(x, j, d), \\
C_D^{L2R}(x, j, d) &\triangleq \sum_{i=1}^{x-1} W_{ix} e_D(i, j, d) + e_D(x, j, d), \\
C_D^{R2L}(x, j, d) &\triangleq \sum_{i=x+1}^m W_{ix} e_D(i, j, d) + e_D(x, j, d).
\end{aligned} \tag{20}$$

These terms can all be written in the following recursive form:

$$\begin{aligned}
C_G^{T2B}(x, y, d) &= U_{(y-1)y} C_G^{T2B}(x, y-1, d) + e_G(x, y, d), \\
C_G^{B2T}(x, y, d) &= U_{y(y+1)} C_G^{B2T}(x, y+1, d) + e_G(x, y, d), \\
C_G^{L2R}(x, j, d) &= U_{(x-1)x} C_G^{L2R}(x-1, j, d) + e_G(x, j, d), \\
C_G^{R2L}(x, j, d) &= U_{x(x+1)} C_G^{R2L}(x+1, j, d) + e_G(x, j, d).
\end{aligned} \tag{21}$$

**2.3. Disparity Computation.** In the last section, the matching cost is aggregated through weighted summation over the entire image, and the disparity which provides minimum cost is assigned to the corresponding pixel. That is, the assigned disparity for a pixel in the reference image,  $D_{\text{ref}}$ , is the one with the minimum aggregated matching cost:

$$\bar{d}_{\text{Ref}}(x, y) = \arg \min_{d \in [0, 1, \dots, d_{\text{max}}]} \bar{C}(x, y, d), \tag{22}$$

where  $[0, 1, \dots, d_{\text{max}}]$  is the disparity search space with  $d_{\text{max}}$  being the maximum possible disparity value.

The initial disparity map normally contains obscured outlier regions. The disparities at these regions are significantly different when they are found from different reference images. If  $\bar{d}_{\text{Ref}}^{\text{Left}}$  and  $\bar{d}_{\text{Ref}}^{\text{Right}}$  are the disparities with the left and right images as reference images, respectively, we can apply the Left-Right Consistency Check (LRC) [23] to determine if a pixel is located at an obscured region:

$$\begin{aligned}
\left| \bar{d}_{\text{Ref}}^{\text{Left}}(x, y) - \bar{d}_{\text{Ref}}^{\text{Right}}\left(x - \bar{d}_{\text{Ref}}^{\text{Left}}(x, y), y\right) \right| &> 1, \\
\left| \bar{d}_{\text{Ref}}^{\text{Right}}(x, y) - \bar{d}_{\text{Ref}}^{\text{Left}}\left(x + \bar{d}_{\text{Ref}}^{\text{Right}}(x, y), y\right) \right| &> 1.
\end{aligned} \tag{23}$$

Once the obscured regions are found, occlusion handling [24] can be applied to patch them with the smallest disparity values of nearby regions, and the corresponding costs are also assigned for disparity refinement, as to be used in the next subsection.

**2.4. Disparity Refinement.** The disparity map obtained by the method proposed in the last section is discrete, since the disparity search space is an integer set. We propose a smoothing technique for the disparity refinement in this subsection.

Considering that the initial disparity has the smallest aggregation cost in the potential solution space, we may interpolate for refined value by fitting data sets to upward curves. Besides, rather than directly using the neighboring disparity for refinement, we use both the costs and disparity in the curve-fitting.

Assuming that  $\bar{C}(x, y, \bar{d})$  is the cost value corresponding to disparity  $\bar{d}$  at  $(x, y)$  on the reference image, we denote  $\bar{C}_{\bar{d}} \triangleq \bar{C}(x, y, \bar{d})$ , such that  $\bar{C}_{\bar{d}+1} \triangleq \bar{C}(x, y, \bar{d} + 1)$  and  $\bar{C}_{\bar{d}-1} \triangleq \bar{C}(x, y, \bar{d} - 1)$ , to simplify the following presentation.

Firstly, the disparity-cost sets around the pixel,  $(\bar{d} - 1, \bar{C}_{\bar{d}-1})$ ,  $(\bar{d}, \bar{C}_{\bar{d}})$ , and  $(\bar{d} + 1, \bar{C}_{\bar{d}+1})$ , are fitted to a hyperbolic function

$$\varphi = f(\xi) = a_1\xi + \frac{a_2}{\xi} + a_3, \text{ where} \quad (24)$$

$$a_1, a_2, a_3 \in \mathbb{R}, \quad a_1 > 0, \quad a_2 > 0.$$

As the minimum value of the curve is located at  $\bar{d}_1^* = \sqrt{a_2/a_1}$ , we have that

$$\bar{d}_1^* = \sqrt{\frac{(\bar{d}^3 - \bar{d})(\bar{C}_{\bar{d}+1} + \bar{C}_{\bar{d}-1} - 2\bar{C}_{\bar{d}})}{\bar{d}(\bar{C}_{\bar{d}+1} + \bar{C}_{\bar{d}-1} - 2\bar{C}_{\bar{d}}) + \bar{C}_{\bar{d}+1} - \bar{C}_{\bar{d}-1}}}. \quad (25)$$

Secondly, an upward parabola equation

$$\theta = g(\zeta) = b_1\zeta^2 + b_2\zeta + b_3, \text{ where} \quad (26)$$

$$b_1, b_2, b_3 \in \mathbb{R}, \quad b_1 > 0, \quad b_2 < 0$$

is used to fit other disparity-cost sets around the pixel,  $(\bar{d} - 2, \bar{C}_{\bar{d}-2})$ ,  $(\bar{d}, \bar{C}_{\bar{d}})$ , and  $(\bar{d} + 2, \bar{C}_{\bar{d}+2})$ . As the minimum value of the parabola is located at  $\bar{d}_2^* = -b_2/2b_1$ , we have

$$\bar{d}_2^* = \bar{d} - \frac{\bar{C}_{\bar{d}+2} - \bar{C}_{\bar{d}-2}}{\bar{C}_{\bar{d}+2} + \bar{C}_{\bar{d}-2} - 2\bar{C}_{\bar{d}}}. \quad (27)$$

The averaged value,  $\bar{d}^* = (\bar{d}_1^* + \bar{d}_2^*)/2$ , is then used as the refined disparity.

### 3. Experimental Study

In contrast to the approach of [18], the proposed algorithm applied independent weights to the mismatch measures while

TABLE 1: Comparison of percentage of miss-matches among the proposed algorithm and [18] in nonoccluded regions.

Image set	Algorithm	
	Proposed	InfoPerm [18]
Tsukuba	3.53%	5.99%
Venus	0.34%	1.32%
Teddy	6.09%	6.20%
Cones	3.60%	4.58%

preserving comparable computational efficiency. Below, we rewrite the aggregated cost function, (3), for comparison:

$$\begin{aligned} \bar{C}(x, y, d) = & \lambda \sum_{i=1}^m \sum_{j=1}^n W_{ix} W_{jy} e_D(i, j, d) \\ & + (1 - \lambda) \sum_{i=1}^m \sum_{j=1}^n U_{ix} U_{jy} e_G(i, j, d). \end{aligned} \quad (28)$$

The aggregated cost function of [18] is equivalent to

$$\begin{aligned} C'(x, y, d) = & \sum_{i=1}^m \sum_{j=1}^n \bar{W}_{ix} \bar{W}_{jy} [\mu e_D(i, j, d) + (1 - \mu) e_{\text{CENSUS}}(i, j, d)], \end{aligned} \quad (29)$$

where  $e_{\text{CENSUS}}$  is the Census measure [25, 26]. It is clear from the comparison between (28) and (29) that the proposed algorithm (28) enables separate weightings on different measures in the aggregation.

A performance comparison between the proposed method and the algorithm of [18], denoted as InfoPerm [18], using the Middlebury stereo test bench [27, 28] is presented in Figure 1. In the computation, the parameters of the proposed algorithm,  $h_D$ ,  $h_G$ ,  $\alpha$ ,  $\beta$ , and  $\lambda$ , are selected as 22, 38, 32, 23, and 0.6, respectively.

In Figure 1, and hence the following presentations, disparities with errors larger than 0.5 disparity levels are named as mismatches, which are denoted in gray at non-occluded regions and black for occluded regions. The percentages of mismatches are further calculated and summarized in Table 1, which shows that the proposed algorithm outperforms InfoPerm [18] in correctness at non-occluded regions in the benchmark tests.

In addition to InfoPerm [18], several state-of-the-art methods, such as SNCC [12], HistAggr2 [10], RTCensus [11], InfoPerm [18], AdaptWeight [9], FeatureGC [13], ObjectStereo [15], and AdaptingBP [16] were also implemented on the Middlebury stereo test bench [25, 26] for a complete performance comparison. Among them, SNCC [12], HistAggr2 [10], RTCensus [11], InfoPerm [18], and AdaptWeight [9] are local stereo match algorithms, while FeatureGC [13], ObjectStereo [15], and AdaptingBP [16] are global techniques.

Figure 2 shows a comparison of the disparity maps between RTCensus [11], SNCC [12], AdaptWeight [9] and the proposed approach, where all the disparity maps have

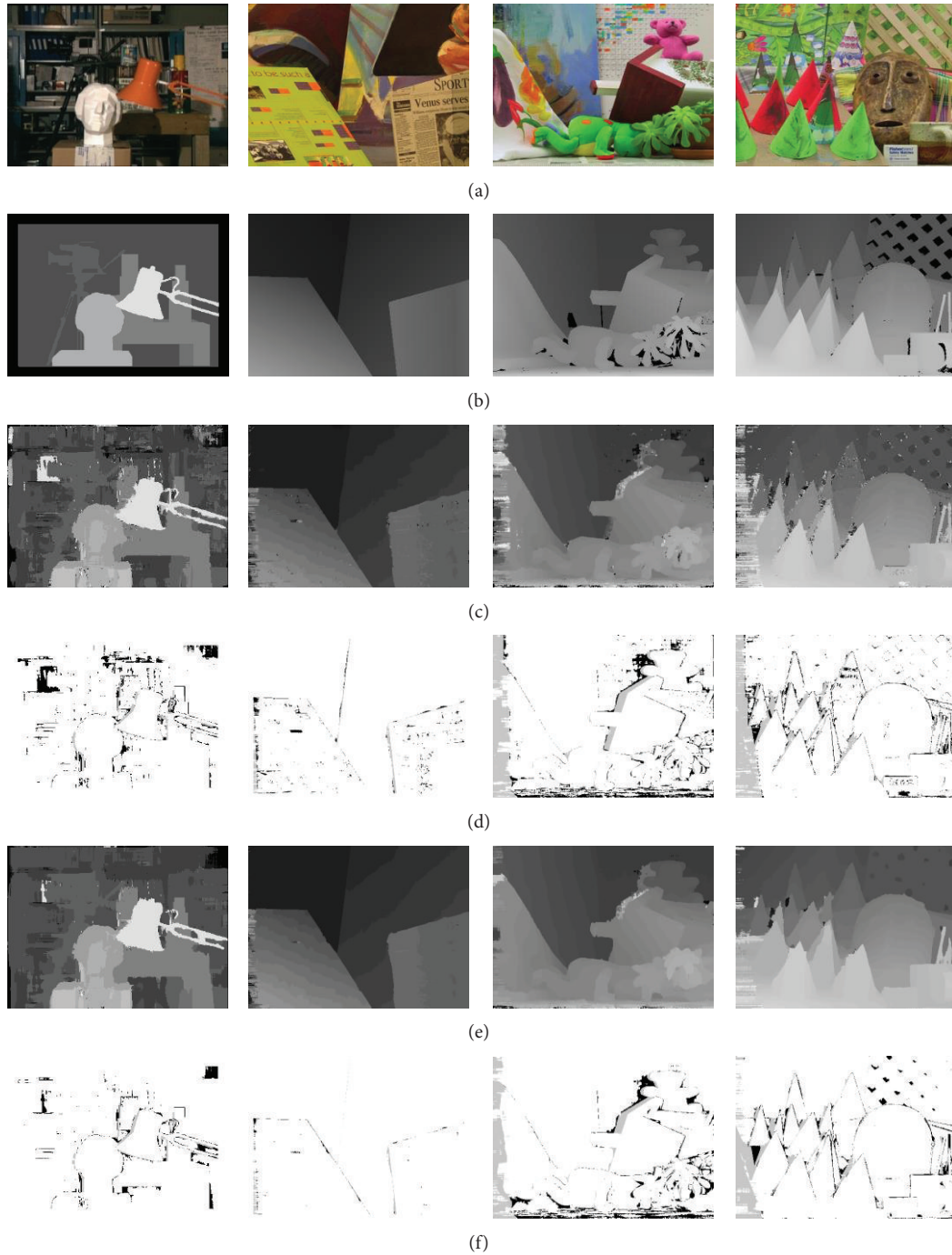


FIGURE 1: A comparison of disparity estimation performance between [18] and the proposed algorithms. Top-to-bottom: (1) first row: images (Tsukuba, Venus, Teddy, and Cones) from Middlebury stereo database, (2) second row: ground truth disparity maps, (3) third row: disparity maps by the InfoPerm [18] algorithm, (4) fourth row: mismatches of the InfoPerm [18] algorithm, (5) fifth row: disparity maps by the proposed algorithm, and (6) sixth row: mismatches of the proposed algorithm.

been refined. The results show that the proposed method has comparable performance with these state-of-the-art methods, and the refinement strategy, introduced in Section 2.4, is able to preserve clear boundaries.

The complete comparison of the mismatch rates between these algorithms is summarized in Table 2. In the table,

“nonocc.” denotes the pixels in the non-occluded region, and “disc.” represents the discontinuous but visible pixels near the occluded regions. According to Table 2, the proposed algorithm outperforms AdaptWeight [9] in all of the mismatch evaluations, although it seems inferior to AdaptWeight [9] in generating sharp boundary, as shown in Figure 1.



FIGURE 2: Top-to-bottom: (1) color views (Tsukuba, Venus, Teddy, and Cone), (2) ground truth disparity maps, (3) disparity maps via RTCensus [11], (4) disparity maps via SNCC [12], (5) disparity maps via AdaptWeight [9], and (6) disparity maps via the proposed approach.

It is also interesting to note that the local stereo matching algorithms, such as RTCensus [11] and SNCC [12], outperform other algorithms for the Teddy and Cones image pairs. Nevertheless, the global stereo matching algorithms, such as FeatureGC [13] and AdaptingBP [16], perform better for the Tsukuba and Venus image pairs. This observation indicates that both the local and global approaches are case-sensitive. However, the proposed approach has comparable

performance in all of the cases and has the lowest mismatch rate in the benchmark evaluation.

#### 4. Conclusion

Stereo matching algorithms are crucial for correct and accurate depth estimation in passive stereo vision systems. A stereo matching algorithm processes rectified stereo image

TABLE 2: Comparison of mismatch percentage among the proposed algorithm and several representative algorithms.

Algorithms	Tsukuba			Venus			Teddy			Cones			Average percent of mismatch
	nonocc.	all	disc.	nonocc.	all	disc.	nonocc.	all	disc.	nonocc.	all	disc.	
SNCC [12]	11.3	12.3	27.5	2.35	3.23	15.4	10.6	15.2	28.6	4.71	11.1	13.2	13.0
HistAggr2 [10]	15.2	15.7	16.4	6.44	7.00	12.5	11.5	18.2	24.4	7.90	13.7	14.2	13.6
RTCensus [11]	12.9	14.1	28.1	3.67	4.63	17.8	11.4	18.6	27.7	5.54	11.8	15.9	14.4
InfoPerm [18]	25.7	26.2	21.2	8.64	9.34	15.0	15.0	22.1	29.2	7.68	15.1	15.1	17.5
AdaptWeight [9]	18.1	18.8	18.6	7.77	8.40	15.8	17.6	23.9	34.0	14.0	19.7	20.6	18.1
FeatureGC [13]	8.22	8.86	13.3	4.58	4.73	10.1	14.7	17.0	32.5	11.5	18.0	21.0	13.7
ObjectStereo [15]	16.4	16.8	16.1	2.56	2.67	7.69	19.6	22.7	30.3	16.3	20.7	19.7	16.0
AdaptingBP [16]	19.1	19.3	17.4	4.84	5.08	7.84	12.8	16.7	26.3	7.02	13.2	14.0	13.6
Proposed	9.12	10.2	15.0	1.04	1.72	8.35	11.2	17.0	26.8	6.36	12.6	15.1	11.2

nonocc.: the pixels in the nonoccluded region; disc.: the visible pixels near the discontinuous regions.

pairs to generate the disparity map, which is used to calculate the depth image (z-map), and hence the 3D point cloud in camera coordinates. For practical applications, the algorithms should require less computational resources and provide precise disparity maps.

In this paper, we proposed an efficient stereo matching algorithm and a refinement strategy for the disparity maps. The algorithm effectively aggregates cost values in terms of bilateral filtering by only four passes along regions, which is able to provide a decoupled dissimilarity measure aggregation while preserving computational efficiency. Besides, the refinement strategy is a simple application of the aggregated costs that use both the costs and disparity in the curve-fitting, rather than directly using the neighboring disparity for refinement.

Experimental results using the Middlebury stereo test bench [25, 26] show that the algorithm has comparable performance with the state-of-the-art algorithms and outperforms the representative algorithms in the overall mismatch rate.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This paper was sponsored by Chang Gung Memorial Hospital, Chang Gung University, and the National Science Council, Taiwan, under Contract nos. CMRPD1C0021, CMRPD2C0051, NSC 100-2221-E-182-008, NSC 101-2221-E-182-006, and NSC 102-2221-E-182-073 and the National Science Foundation of China, under Contract no. 61271326.

## References

- [1] Á. González, M. Á. Garrido, D. F. Llorca et al., "Automatic traffic signs and panels inspection system using computer vision," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 485–499, 2011.
- [2] F. Tombari, F. Gori, and L. di Stefano, "Evaluation of stereo algorithms for 3D object recognition," in *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops '11)*, pp. 990–997, Barcelona, Spain, November 2011.
- [3] K. Ohno, K. Kense, E. Takeuchi, and Z. Lei, "Unknown object modeling on the basis of vision and pushing manipulation," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO '11)*, pp. 1942–1948, Phuket, Thailand, December 2011.
- [4] K. Schauwecker and A. Zell, "On-board dual-stereo-vision for autonomous quadrotor navigation," in *Proceedings of the International Conference on Unmanned Aircraft Systems (ICUAS '13)*, pp. 333–342, Atlanta, Ga, USA, May 2011.
- [5] V. Murino, U. Castellani, and A. Fusiello, "Disparity map restoration by integration of confidence in Markov random fields models," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '01)*, pp. 29–32, Thessaloniki, Greece, October 2001.
- [6] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1–3, pp. 7–42, 2002.
- [7] Y. Boykov, O. Veksler, and R. I. Zabí, "A variable window approach to early vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1283–1294, 1998.
- [8] H. Hirschmüller, P. R. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," *International Journal of Computer Vision*, vol. 47, no. 1–3, pp. 229–246, 2002.
- [9] K.-J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 650–656, 2006.
- [10] D. Min, J. Lu, and M. N. Do, "Joint histogram based cost aggregation for stereo matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, pp. 2539–2545, 2013.
- [11] M. Humenberger, C. Zinner, M. Weber, W. Kubinger, and M. Vincze, "A fast stereo matching algorithm suitable for embedded real-time systems," *Computer Vision and Image Understanding*, vol. 114, no. 11, pp. 1180–1202, 2010.
- [12] N. Einecke and J. Eggert, "A two-stage correlation method for stereoscopic depth estimation," in *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications (DICTA '10)*, pp. 227–234, Sydney, Australia, December 2010.



- [13] G. Saygili, L. van der Maaten, and E. A. Hendriks, "Feature-based stereo matching using graph cuts," in *Proceedings of ASCI-IPA-SIKS Tracks, IPA Fall Days/ICT.OPEN*, Veldhoven, The Netherlands, November 2011.
- [14] Z.-F. Wang and Z.-G. Zheng, "A region based stereo matching algorithm using cooperative optimization," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.
- [15] M. Bleyer, C. Rother, P. Kohli, D. Scharstein, and S. Sinha, "Object stereo Joint stereo matching and object segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 3081–3088, Washington, DC, USA, June 2011.
- [16] A. Klaus, M. Sormann, and K. Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, pp. 15–18, Hong Kong, August 2006.
- [17] R. Yang and M. Pollefeys, "Multi-resolution real-time stereo on commodity graphics hardware," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '03)*, vol. 1, pp. I/211–I/217, Washington, DC, USA, June 2003.
- [18] C. Cigla and A. A. Alatan, "Efficient edge-preserving stereo matching," in *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVWorkshops '11)*, pp. 696–699, Barcelona, Spain, November 2011.
- [19] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, 2012.
- [20] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Computer Vision—ECCV 2010*, vol. 6311 of *Lecture Notes in Computer Science*, pp. 1–14, Springer, Heidelberg, Greece, 2010.
- [21] L. De-Maeztu, S. Mattoccia, A. Villanueva, and R. Cabeza, "Linear stereo matching," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 1708–1715, Barcelona, Spain, November 2011.
- [22] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the 6th IEEE International Conference on Computer Vision (ICCV '98)*, pp. 839–846, Bombay, India, January 1998.
- [23] P. Fua, "Combining stereo and monocular information to compute dense depth maps that preserve depth discontinuities," in *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI '91)*, vol. 2, pp. 1292–1298, Sydney, Australia, 1991.
- [24] L. D. Stefano, M. Marchionni, and S. Mattoccia, "A fast area-based stereo matching algorithm," *Image and Vision Computing*, vol. 22, no. 12, pp. 983–1005, 2004.
- [25] R. Zabih and J. Woodfill, "Non-parametric local transforms for computing visual correspondence," in *Computer Vision—ECCV 1994*, *Lecture Notes in Computer Science*, pp. 151–158, Springer, 1994.
- [26] N. Y.-C. Chang, T.-H. Tsai, B.-H. Hsu, Y.-C. Chen, and T.-S. Chang, "Algorithm and architecture of disparity estimation with mini-census adaptive support weight," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 6, pp. 792–805, 2010.
- [27] "Middlebury stereo vision database," <http://vision.middlebury.edu/stereo/>.
- [28] F. Tombari, S. Mattoccia, L. D. Stefano, and E. Addimanda, "Classification and evaluation of cost aggregation methods for stereo correspondence," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

