A Dense Stereo Matching Using Two-Pass Dynamic Programming with Generalized Ground Control Points

Jae Chul Kim[1], Kyoung Mu Lee[2], Byoung Tae Choi[3], Sang Uk Lee[4]
[1,3] *Electronics and Telecommunications Research Institute, 305-350, Daejon, Korea*
[2,4] *School of Electrical Eng., ASRI, Seoul National University, 151-600, Seoul Korea*
e-mail: jaechul@etri.re.kr, kyoungmu@snu.ac.kr, btchoi@etri.re.kr, sanguk@ipl.snu.ac.kr

Abstract

A method for solving dense stereo matching problem is presented in this paper. First, a new generalized ground control points (GGCPs) scheme is introduced, where one or more disparity candidates for the true disparity of each pixel are assigned by local matching using the oriented spatial filters. By allowing "all" pixels to have multiple candidates for their true disparities, GGCPs not only guarantee to provide a sufficient number of starting pixels needed for guiding the subsequent matching process, but also remarkably reduce the risk of false match, improving the previous GCP-based approaches where the number of the selected control points tends to be inversely proportional to the reliability. Second, by employing a two-pass dynamic programming technique that performs optimization both along and across the scanlines, we solve the typical inter-scanline inconsistency problem. Moreover, combined with the GGCPs, the stability and efficiency of the optimization are improved significantly. Experimental results for the standard data sets show that the proposed algorithm achieves comparable results to the state-of-the-arts with much less computational cost.

1. Introduction

1.1. Motivation

Stereo matching is a problem to find correspondences between two or more input images. It is one of fundamental computer vision problems with a wide range of applications, and hence it has been extensively studied in the computer vision field for decades. However, there still exist some difficult inherent problems in stereo matching; for example, the presence of homogeneously textured regions, and the occlusions near the object boundaries that make the disparity assignment very difficult.

To resolve these difficulties, numerous attempts have been made to lessen the matching ambiguities by propagating the reliable matching results [4, 8, 20, 22, 23]. In these reliability-based approaches, one of the most important tasks is to select the reliably matched pixels, *i.e.* ground control points (GCPs). It is known that the false matches in GCPs could severely degrade the final matching results. On the other hand, the number of the obtained GCPs would decrease if stricter constraints are enforced for outlier removals, which in turn could lead to the lack of information needed for appropriately guiding the subsequent matching process.

The first motivation of our paper is to solve those problems of conventional GCP-based approaches. To this end, we propose the generalized ground control points (GGCPs) scheme in which unlike conventional GCP-based approaches where only reliably matched pixels are selected, multiple disparity candidates are assigned to all pixels by local matching using the oriented spatial filters. Using this scheme, the probability of false match drops remarkably, and furthermore sufficient information is always provided for dense matching, since all pixels take part in guiding the subsequent matching process without loss of reliability.

GCPs or GGCPs can be applied to various matching techniques [4, 8, 11, 22]. In this paper, GGCPs are applied to global optimization using efficient dynamic programming. In this sense, the second motivation of our paper is to develop a fast matching algorithm, while achieving the accuracy comparable to the state-of-the-arts [5, 9, 13, 19]. So, we propose a two-pass dynamic programming technique. The proposed two-pass dynamic programming is designed to resolve the inconsistency between scanlines, which is the typical problem in conventional dynamic programming. It performs optimization both along and across the

Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)
1063-6919/05 $20.00 © 2005 **IEEE**

scanlines. Furthermore, since the finite number of disparity candidates of GGCPs not only reduce the range of possible disparities to be searched, but also provide good initial points for optimization, the optimization becomes more efficient and stable.

1.2. Related works

The proposed algorithm has the workflow in which first, the local matching using spatial filters is carried out, and then the results of local matching is applied to the global optimization. This approach has been already adopted in several algorithms [1, 4, 11]. In particular, our algorithm has the similar framework to Bobick et al.'s algorithm [4], where GCPs were used together with dynamic programming. But, we propose the GGCPs as extension of the GCPs, and unlike the work of Bobick et al. where consistency between scanlines were imposed using only GCPs, we guarantee the consistency by the two-pass dynamic programming. These features bring about remarkable improvement in matching accuracy.

In this paper, disparity candidates of each pixel, i.e. GGCPs, are obtained from local matching by the oriented spatial filters. These oriented filters have a few advantages over the windows commonly used in stereo matching. First, they can delineate the object boundaries more clearly. Second, even when the oriented filters are applied to the slanted plane, at least one filter among the filters with various orientations satisfies the fronto-parallel plane assumption, and therefore more accurate matching results for the slanted planes can be provided. Additionally, in order to take the best advantages of the oriented filters for stereo, it is desirable for the filters to have high resolution in orientation. To this end, we adopt the oriented rod-shaped filters, instead of the Gaussian-based filters that used commonly in conventional algorithms [10, 12]. It can be shown that the coefficients of the rod-shaped filter are more concentrated along the orientation of filter that leads to higher resolution in orientation (see figure 1). The detailed description on the rod-shaped filter will be presented in section 2.

Finally, there have been many works to solve the scanline inconsistency problem of dynamic programming [2, 3, 6, 14]. For examples, Birchfield et al. [3] conducted a post-processing using heuristics, and Cox et al. [6] locally dealt with the inconsistency problem by minimizing the discontinuities between neighboring scanlines. But, these algorithms only offered partial remedy for the inconsistency problem. The proposed algorithm carries out the two-pass dynamic programming using the scanline optimization [17] without consideration of the ordering constraint. By excluding the ordering constraint from optimization process, we can readily perform the optimization across the scanlines, by the same manner as the one used in the optimization along the scanlines. Here, we should note that our idea on the two-pass dynamic programming is inspired from the algorithm proposed by Zickler et al. [24] who applied the two-pass dynamic programming to binocular Helmholtz stereopsis. However, we adapt the two-pass dynamic programming for stereo matching. Furthermore, by combining the two-pass dynamic programming with the information from GGCPs, we can obtain a remarkably enhanced solution for inter-scanline inconsistency problem.

2. Preliminaries

For convenience, we assume that input images are rectified. Then, the correspondences between input images are represented by a univalued disparity function $d(x, y)$ with respect to a pixel (x, y) of the reference image. The disparity function can take one of integer values within the disparity ranges of the scene.

A pair of a pixel (x, y) and its disparity d generates a point (x, y, d), which constructs a 3D disparity space. An initial matching cost $C_0(x, y, d)$ measures the pixel-based error of a match at the point (x, y, d). The simplest matching cost uses absolute intensity differences between a pixel (x, y) of the reference (left) image I_1 and a pixel $(x-d, y)$ of the matching (right) image I_2, i.e. $C_0(x, y, d) = \left| I_1(x, y) - I_2(x-d, y) \right|$.

In the proposed algorithm, the rod-shaped spatial filters with N orientations are used. Examples of the filters are illustrated in Figure 1 where each filter is rotated by 15°. Generally, the rod-shaped filter which is $2l+1$ pixels long, and inclined at θ to the horizontal axis can be numerically expressed as

$$f_\theta(x, y) = \begin{cases} 1 - \left| x\sin\theta - y\cos\theta \right| & \text{if } \left| x\sin\theta - y\cos\theta \right| < 1, \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

for $|x| < l\cos\theta$ and $|y| < l\sin\theta$.

To avoid the problem that filters are across the object boundaries, we perform local matching using three filters for each orientation, where the centers of the filters are shifted to the three different positions, and only the best filtering result is taken. The shifted versions of the oriented filter are shown in figure 2. The shiftable filtering is implemented by a cascade of a single oriented filtering and a minimum (or maximum) finding for the three center points instead of three repetitive filtering for each center point. So, only little additional computation is required for the shiftable filtering. For more information on the shiftable filtering, interested readers may refer to [15, 17].

3. Local matching

In local matching, the disparity candidates for true disparity of each pixel are obtained. And these candidates are computed by the sequential operation of following three steps: preprocessing, local aggregation, and post-processing.

3.1. Preprocessing

The preprocessing aims to classify each pixel in the reference image into two groups (homogeneous group and heterogeneous group) according to the intensity variation around each pixel. To compute the intensity variation of each pixel, small-sized (for example, dimension of 3×3) Laplacian of Gaussian filter is first applied to the reference image, followed by shiftable oriented filters with N orientations in order to take account of the intensity variations along the neighborhood of the each orientation where local aggregation is performed. In the shiftable filtering, a minimum finding (*not maximum*) is used for picking the best result. Since the shiftable filtering is carried out independently for N orientations, N minimum values are assigned to each pixel as the intensity variations around it. Finally, if at least one of the N values of a pixel is over a threshold, the pixel is labeled as heterogeneous pixel; otherwise, it is labeled as homogeneous pixel. Thus, a homogeneous pixel has no significant intensity variation for any orientation along which the filtering is executed.

3.2. Local Aggregation

In this stage, candidates for the true disparity of each pixel are provided by locally aggregating the initial matching costs using the spatial filters. The detailed procedure for determining the disparity candidates of each pixel is as follows. First, the initial matching cost $C_0(x,y,d)$ for each pair of a pixel (x,y) and its disparity d is evaluated by

$$C_0(x,y,d)= \begin{cases} |I_1(x,y)-I_2(x-d,y)| & \text{if } (x,y)\in \textit{homogeneous} \\ |g\otimes I_1(x,y)-g\otimes I_2(x-d,y)| & \text{if } (x,y)\in \textit{heterogeneous}, \end{cases} \qquad (2)$$

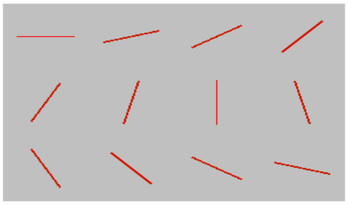

Figure 1. Examples of the rod-shaped oriented filters

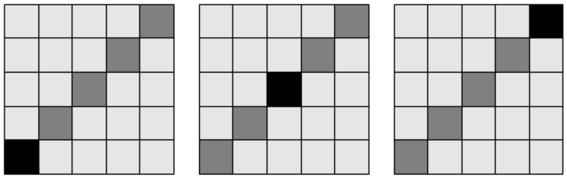

Figure 2. Diagram of the three shiftable oriented filters, where the centers of the filters are marked in black.

where g denotes the Gaussian kernel, and $g\otimes I_i (i=1,2)$ represents the convolution. Note that the initial costs of the heterogeneous pixels are computed from the smoothed input images by Gaussian filter. This low-pass filtering helps to suppress the sampling artifacts in the intensity-varying areas, which are known to be quite common in a kind of small-sized box filters like our rod-shaped filters [21]. After computing the initial cost, the aggregation of the initial matching cost is implemented using the shiftable oriented filters. However, note that in order to reduce the matching ambiguities induced by the lack of intensity variation in the homogeneous pixels, large-sized shiftable windows are additionally applied to the homogeneous pixels. In mathematical terms, the aggregation can be expressed by

$$C(x,y,d) = \sum_m \sum_n f(x-m,y-n)\cdot C_0(x-m,y-n,d), \quad (3)$$

where f denotes a 2D spatial filter. Notice that in our algorithm, the aggregation is performed for N orientations, so that N aggregated costs are assigned to each pixel-disparity pair. Of course, $N+1$ costs including the one from the shiftable windows are given to the homogeneous pixels. Finally, at each pixel the best disparity associated with the minimum cost value is selected. Since each pixel has N (or $N+1$) aggregation results, the N (or $N+1$) best disparities are stored in each pixel, and these best disparities are established as the disparity candidates of each pixel. In addition, the aggregated costs at the best disparities become the matching costs of the disparity candidates. Specially, if the same best disparity is produced from multiple filters, the smallest one among each filter's aggregated costs is assigned to the disparity. The disparities other than candidates are excluded in the subsequent matching process by setting their matching costs to be very large.

3.3. Postprocessing

To enhance the reliability of local aggregation, some heuristic methods are used; these are similar to those used by Bobick et al. [4] to identify the GCPs.

Visibility test Visibility test confirms the consistency of the bi-directional matching based on the uniqueness assumption, and mainly aims to eliminate the matching ambiguities occurred by occlusions. Let $d_{\min 1}$ be the winner-take-all disparity of a pixel (x, y) of the reference image, i.e.

$$d_{\min 1} = \arg\min_d C(x, y, d),$$

and $d_{\min 2}$ be the winner-take-all disparity of a pixel $(x - d_{\min 1}, y)$ of the matching image, i.e.

$$d_{\min 2} = \arg\min_d C(x - d_{\min 1} + d, y, d).$$

If $d_{\min 1} \neq d_{\min 2}$, the pixel (x, y) fails to pass the visibility test, and its local aggregation result is invalidated by equalizing the matching costs of all candidates at the pixel (x, y) to an arbitrary value (zero in our algorithm), which eliminates the difference between the aggregation results of the candidates. Pay attention to that the matching costs of the disparities other than the disparity candidates are kept as an initially assigned value, i.e. very large predefined value.

Detection of suspicious pixels Due to some reasons such as specularity or perfectly homogeneous texture, there exist pixels whose results of local aggregation cannot be trusted. We separate such pixels according to the following rules. First, if the minimum matching cost of a pixel exceeds a threshold t_1, the pixel is marked as an suspicious one. In addition, for a homogeneous pixel, if the difference between its first minimum matching cost and the second one is below a threshold t_2, that pixel is also labeled as suspicious. The aggregation results of these suspicious pixels are then invalidated by the same manner used in the visibility test except for one difference: the matching costs at all disparities are annulled without any distinction between the disparity candidates and the remainders. This is because the local aggregation result for a suspicious pixel is provided in such an unpredictable way that it is not credible that a true disparity of the pixel exists in its disparity candidates, whereas matching ambiguities of a pixel by occlusions near the depth discontinuities arise just between foreground disparity and background one, both of which are generally included in the disparity candidates of the pixel.

4. Global optimization

In this section, two-pass dynamic programming is performed for global optimization using the scanline optimization [17], rather than the typical dynamic programming enforcing the ordering constraint. By excluding the ordering constraint, the scanline optimization makes it easy to optimize across the scanlines, where it is impractical to impose the ordering constraint.

The complexity of the scanline optimization is $O(mn^2)$ for m pixels and n disparities, which is larger than $O(mn)$, the complexity of the typical dynamic programming. But, the proposed algorithm can work very efficiently because it only considers the disparity candidates of each pixel, but not all disparities within the search range.

4.1. PASS 1: Optimization along the scanlines

The optimization along the scanlines finds a path of disparities that minimizes the following energy functional,

$$\begin{aligned} \mathrm{E_h}(d(x, y_1)) = \sum_x C(x, y_1, d(x, y_1)) + \\ \sum_x \lambda(x, y_1)\rho(d(x, y_1) - d(x+1, y_1)) \end{aligned} \quad (4)$$

for a scanline y_1. In equation (4), C is the matching cost obtained from the local aggregation, ρ is an increasing function of the disparity difference between adjacent pixels evaluating the smoothness of a disparity function, and λ is a weight function.

As a ρ function, Potts model [16] has been widely used since it can handle the disparity jumps, and it is adopted in our algorithm as well. The Potts model is

$$\rho(\alpha) = \begin{cases} 0 & \alpha = 0 \\ 1 & \alpha \neq 0. \end{cases} \quad (5)$$

For homogenous pixels with valid matching costs (see section 3.3), on the other hand, modified Potts model is used to avoid the excessive smoothing in the homogeneous textured regions. The modified Potts model incorporates the disparity gradient constraint into the original Potts model, and is written as

$$\rho'(\alpha) = \begin{cases} 0 & \text{if } \alpha = 0 \\ 0.5 & \text{if } \alpha = 1 \text{ or } \alpha = \text{-}1 \\ 1 & \text{otherwise.} \end{cases} \quad (6)$$

In contrast to the original Potts model preferring the fronto-parallel planes (especially, this preference will be intensified in the homogeneous pixels where the smoothness constraint dominates the energy functional of equation 3.), the modified Potts model encourages the slanted planes by lowering the cost imposed on one-pixel sized disparity difference to be one-half. This strategy aids to diminish the influence of the smoothness constraint on the slanted surface where disparities of neighboring pixels commonly vary within a smaller range than one-pixel difference.

The weight function λ has a functional value inversely proportional to the intensity gradient to help align the disparity jumps with the intensity edges [4, 5, 7]. In our algorithm, λ is defined as

$$\lambda(x, y_1) = \begin{cases} 0.5c & \text{if } \nabla_h I_1(x, y_1) > i_2 \\ c & \text{if } i_1 < \nabla_h I_1(x, y_1) < i_2 \\ 2c & \text{otherwise,} \end{cases} \quad (7)$$

where c is a constant and $\nabla_h I_1(x, y_1)$ denotes the horizontal intensity gradient computed by 3×3 sized horizontal Sobel operator. Both i_1 and i_2 are thresholds for the intensity gradient.

For each pixel-disparity pair (x, y_1, d) on a scanline y_1, a typical scanline optimization usually proceeds to compute the minimum cost $C_h(x, y_1, d)$ required for reaching each pair, and when finally arriving at the end point x_e, the optimal path is decided as the one that gets to the $(x_e, y_1, d_{\min}(x_e))$, where $d_{\min}(x_e) = \arg\min_d C_h(x_e, y_1, d)$. Our algorithm, however, does not follow this procedure. Instead, the cost $C_h(x, y_1, d)$ is incorporated into the energy functional defined in the PASS 2 where the final optimal disparities are chosen, while the optimal disparities of the PASS 1 just bias the final disparities toward them by lowering their costs by a constant.

In the typical scanline optimization, the cost $C_h(x, y_1, d)$ increases with x, but the same dimensions with respect to the x-direction are required to apply the cost to the energy functional of the PASS 2. To achieve this symmetry, we use the similar approach to the one used by Zickler et al. [24]; the cost $C_h(x, y_1, d)$ is obtained by summing $C_{h1}(x, y_1, d)$, the cost from the computation in the increasing direction of x, and $C_{h2}(x, y_1, d)$, the cost from the computation in the reverse direction.

4.2. PASS 2: Optimization across the scanlines

In this pass, the objective is to provide a final disparity path that minimizes an energy functional,

$$\mathrm{E}_v(d(x_1, y)) = \sum_y C(x_1, y, d(x_1, y)) + \sum_y C_h(x_1, y, d(x_1, y)) + \sum_y \lambda(x_1, y)\rho(d(x_1, y) - d(x_1, y+1)). \quad (8)$$

for a vertical line $x = x_1$. In equation (8), C is the matching cost by the local aggregation, identical to the one used in equation (4), and C_h is the cost obtained from the previous optimization pass, and ρ enforces a smoothness on the disparity function as before, but this time, the original Potts model is used instead of the modified Potts model on the pragmatic grounds, since it shows slightly better experimental results. A weight function λ is written as

$$\lambda(x_1, y) = \begin{cases} 0.5c & \text{if } \nabla_v I_1(x_1, y) > i_2 \\ c & \text{if } i_1 < \nabla_v I_1(x_1, y) < i_2 \\ 2c & \text{otherwise,} \end{cases} \quad (9)$$

where all parameters are identical to those of equation (7) except that the intensity gradient is computed by the 3×3 sized *vertical* (not horizontal) Sobel operator denoted by $\nabla_v I_1(x_1, y)$. Note that the energy functional of equation (8) wholly considers three factors: the local matching result, the result from optimization along the scanlines, and the smoothness of a disparity function in the vertical direction. By minimizing this functional, we can get the disparity map preserving the consistency between scanlines.

The matching process is finished with the simple sub-pixel refinement using the following rule: if the number of disparity candidates of a pixel (x, y) is two, and the disparity difference between candidates is one, then the sub-pixel refinement for the pixel (x, y) is carried out by the heuristic equation,

$$d_{\text{sub}}(x, y) = 0.75 d_0(x, y) + 0.25 d_1(x, y), \quad (10)$$

where d_0 denotes the disparity candidate selected in the optimization, and d_1 the remaining candidate that is not selected. Of course, d_{sub} represents the refined disparity in sub-pixel.

5. Experiment

We evaluated the proposed algorithm using four standard data sets, Tsukuba, Sawtooth, Venus, and Map that are provided by Sharstein and Szeliski on the web [25]. The quality metric is the percentage of error disparities deviating from the ground truth more than 1 pixel [17, 18].

Experimental setting In the experiment, 3×3 sized Laplacian of Gaussian filter with standard deviation 1.0 was used for texture analysis; 3×3 sized Gaussian filter with standard deviation 0.85 was applied to the heterogeneous pixels for removing the sampling artifacts of the oriented filters; 36 oriented filters, (i.e. rotated by 5°) which are 15 pixels long were used for local aggregation, and 11×11 sized shiftable windows were additionally used for local aggregation of the homogeneous pixels. In addition, all parameters were fixed for the four data sets. Threshold t_1, t_2 for detecting the suspicious pixels (see section 3.3) were set to 5.0 and 0.05 respectively. In equation (7) and (9), the value of constant c was given 1.0, and thresholds i_1, i_2 were set to 20 and 140 respectively. Although the above-mentioned parameter setting gives the best results, our algorithm performed well for all data sets under the significant changes of the parameters, and we believe this is due to the use of GGCPs.

Effectiveness of GGCPs To validate the effectiveness of GGCPs, two tests were executed. The first test examines the percentage of erroneous pixels that do not have the true disparity within their disparity candidates. Table 1 summarizes the result.

Table 1. Reliablity of GGCPs in terms of density of valid pixels (D) and their error rates (e), where valid pixels mean those which can provide GGCPs (see section 3.3)

	Tsukuba	Sawtooth	Venus	Map
D (%)	95.2	98.9	92.9	98.5
e (%)	0.24	0.07	0.09	0.23

For all data sets, the error rates are within 0.07%~0.24% for the valid pixels that amount to more than 90% of whole pixels. This result verifies that GGCPs can provide the enough reliable information to guide the subsequent matching process. The second test counts the number of disparity candidates that each pixel has, aiming to verify the efficiency improvement gained by GGCPs. Figure 3 plots the cumulative density function with respect to the number of candidates. For three data sets except Venus, more than 90% of pixels have the candidates less than five, and about 70% of pixels for the Venus data set. By remarkably reducing the range of possible disparities to be searched by the optimization process, GGCPs bring about the significant speed-up of the computation.

Running time Table 2 reports the running time obtained on a Pentium IV 2.4GHz PC. Through the reference to Y. Wei et al.'s paper [22] mentioning that their algorithm should be one of the fastest among the state-of-the-arts, we argue that our algorithm may be the fastest among the state-of-the-arts since our algorithm is faster than their method.

Table 2. Running Time in seconds.

	Tsukuba	Sawtooth	Venus	Map
pixels	110,592	164,920	166,222	61,344
disp.	15	21	21	30
time	4.4	11.8	11.1	4.9

Matching Accuracy Figure 4 shows the final disparity maps computed using the two-pass dynamic programming. For illustrating the enhanced inter-scanline consistency by the two-pass algorithm, the intermediate results obtained by single optimization along the scanlines (PASS 1) are shown as well. Due to the use of GGCPs, the disparity maps provided by the PASS 1 show better inter-scanline consistency than the conventional single-pass dynamic programming, but we still have scanline inconsistency problem. However, this inconsistency is fully removed in the final disparity maps. Here, acute observers would see a few vertical streaks in the final disparity maps. But, these streaks don't have an impact on the matching accuracy since they don't produce gross errors more than 1 pixel. The overall evaluation and comparison on the standard platform [25] is presented in Table 3. Our algorithm ranks the 5th out of 32 algorithms, and has little difference from the top-ranked algorithm.

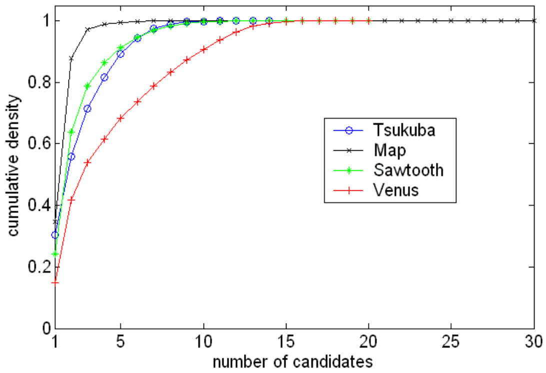

Figure 3. Cumulative density function with respect to the number of disparity candidates

1063-6919/05 $20.00 © 2005 IEEE

6. Conclusion

This paper presented a new two-pass dynamic programming technique combined with generalized ground control points (GGCPs) for a dense stereo matching. The main contribution of our work is twofold. First, we introduced a new scheme of GGCPs, which guarantee to provide a sufficient number of starting points needed for guiding the subsequent matching process without loss of reliability, improving the previous GCP-based approaches where the number of the selected control points tends to be inversely proportional to their reliability. Second, we resolved the inconsistency between scanlines, which was the typical problem in dynamic programming, by employing the two-pass dynamic programming technique combined with GGCPs. Experiment results on the standard data sets showed that our algorithm achieved a high accuracy ranked among the state-of-the-arts with much less computational cost.

Acknowledgement

J.C. Kim and B.T. Choi acknowledge the support by the Korean Ministry of Information & Communications (MIC) under the project No. 2004/S/003/5210. K.M. Lee and S.U. Lee were supported in part by the MIC under the Information Technology Research Center Support Program.

References

[1] M. Agrawal and L. Davis, "Window-based, discontinuity preserving stereo", In *Proc. Conf. on Computer Vision and Pattern Recognition*, vol.1, pp. 66-73, 2004.

[2] P.N. Belhumeur, "A binocular stereo algorithm for reconstructing sloping, creased, and broken surfaces in the presence of half-occlusion", In *Proc. Int. Conf. on Computer Vision*, pp. 431-438, 1993.

[3] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo", *International Journal of Computer Vision,* 35(3), pp. 269-293, 1999.

[4] A.F. Bobick and S.S. Intile, "Large occlusion stereo", *International Journal of Computer Vision,* 33(3), pp. 181-200, 1999.

[5] Y. Boykov, O. Veksler and R. Zabih, "Fast approximate energy minimization via graph cuts", IEEE *Trans. On Pattern Analysis and Machine Intelligence*, vol.23, No.11, pp. 1222-1239, 2001.

[6] I.J. Cox, S.L. Hingorani, S.B. Rao, and B.M. Maggs, "A maximum likelihood stereo algorithm", *Computer Vision and Image Understanding*, 63(3), pp. 542-567, May 1996.

[7] P. Fua, "A parallel stereo algorithm that produces dense depth maps and preserves image features", *Machine Vision and Applications*, pp. 35-49, 1993.

[8] M.L. Gong and Y.H. Yang, "Fast stereo matching using reliability-based dynamic programming and consistency constraints", In *Proc. Int. Conf. on Computer Vision*, pp. 610-617, 2003.

[9] L. Hong, G.Chen, "Segment-based stereo matching using graph cuts, In *Proc. Conf. on Computer Vision and Pattern Recognition*, vol.1, pp. 74-81, 2004.

[10] D. Jones and J. Malik, "A computational framework for determining stereo correspondence from a set of linear spatial filters", In *Proc. European Conf. on Computer Vision*, pp. 395-410, 1992.

[11] S.B. Kang, R. Szeliski, "Extracting view-dependent depth maps from a collecting of images", *International Journal of Computer Vision*, 58(2), pp. 139-163, July 2004.

[12] M. Kass, "Computing visual correspondence", In *Proc. Image Understanding Workshop*, pp. 54-60, June 1983.

[13] V. Kolmogorov and R. Zabih, "Multi-camera scene reconstruction via graph cuts". In *Proc. European Conf. on Computer Vision*, vol.3, pp. 82-96, 2002.

[14] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming", IEEE *Trans. On Pattern Analysis and Machine Intelligence,* 7(2), pp. 139-154, 1985.

[15] M. Okutomi, Y. Katayama and S. Oka, "A simple stereo algorithm to recover precise object boundaries and smooth surfaces", *International Journal of Computer Vision,* 47(1/2/3), pp. 261-273, 2002.

[16] R.B. Potts, "Some generalized order-disorder transitions", In *Proc. Camb. Phil. Soc.*, vol. 48, pp. 106-109, 1952.

[17] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", *International Journal of Computer Vision*, 47(1/2/3), pp. 7-42, April-June 2002.

[18] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light", In *Proc. Conf. on Computer Vision and Pattern Recognition*, vol.1, pp. 195-202, 2003.

[19] J. Sun, H.Y. Shum and N.N. Zheng, "Stereo matching using belief propagation", IEEE *Trans. On Pattern Analysis and Machine Intelligence*, vol.25, No.7, pp .787-800, July 2003.

[20] R. Szeliski and D. Scharstein, "Symmetric sub-pixel stereo matching", In *Proc. European Conf. on Computer Vision*, vol.2, pp. 525-540, 2002.

[21] R. Szeliski and D. Scharstein, "Sampling the disparity space image", IEEE *Trans. On Pattern Analysis and Machine Intelligence*, vol.25, No.3, pp. 419-425, 2004.

[22] Y. Wei and L. Quan, "Region-based progressive stereo matching", In *Proc. Conf. on Computer Vision and Pattern Recognition*, vol.1, pp. 106-113, 2004.

[23] Z. Zhang and Y. Shan, "A progressive scheme for stereo matching", *Second European Workshop on 3D Structure from Multiple Images of Large-Scale Environments (SMILE2000)*, pp. 68-85, 2000.

[24] T.E. Zickler, J. Ho, D.J. Kriegman, J. Ponce, and P.N. Belhumeur, "Binocular Helmholtz stereopsis", In *Proc. Int. Conf. on Computer Vision*, vol.2, pp. 1411-1417, 2003.

[25] http://www.middlebury.edu/stereo/.

Table 3. Performance comparison table (incomplete) from the Middlebury Stereo Vision Page [25]. Error percentages are calculated over three different areas in the image, classified as untextured(untex), discontinuous(disc), and the entire image(all). Algorithms are in order of their ranking. Our algorithm ranked the 5th out of 32 algorithms. Subscript number is ranking in each column.

Algorithm	Tsukuba			Sawtooth			Venus			Map	
	all	untex	disc	all	untex	disc	all	untex	disc	all	disc
Segm.-based CG	1.23 3	0.29 2	6.94 4	0.30 3	**0.00** 1	3.24 3	**0.08** 1	**0.01** 1	**1.39** 1	1.49 20	15.46 25
Segm.+glob.vis.	1.30 5	0.48 5	7.50 6	**0.20** 1	**0.00** 1	**2.30** 1	0.79 4	0.81 5	6.37 8	1.63 22	16.07 27
Layered	1.58 8	1.06 10	8.82 9	0.34 4	**0.00** 1	3.35 4	1.52 11	2.96 20	2.62 3	0.37 10	5.24 10
Belief prop	**1.15** 1	0.42 3	**6.31** 1	0.98 11	0.30 15	4.83 8	1.00 7	0.76 4	9.13 14	0.84 17	5.27 11
Our method	**1.53** **7**	**0.66** **7**	**8.25** **8**	**0.61** **7**	**0.02** **7**	**5.25** **9**	**0.94** **5**	**0.95** **6**	**5.72** **7**	**0.70** **15**	**9.32** **16**

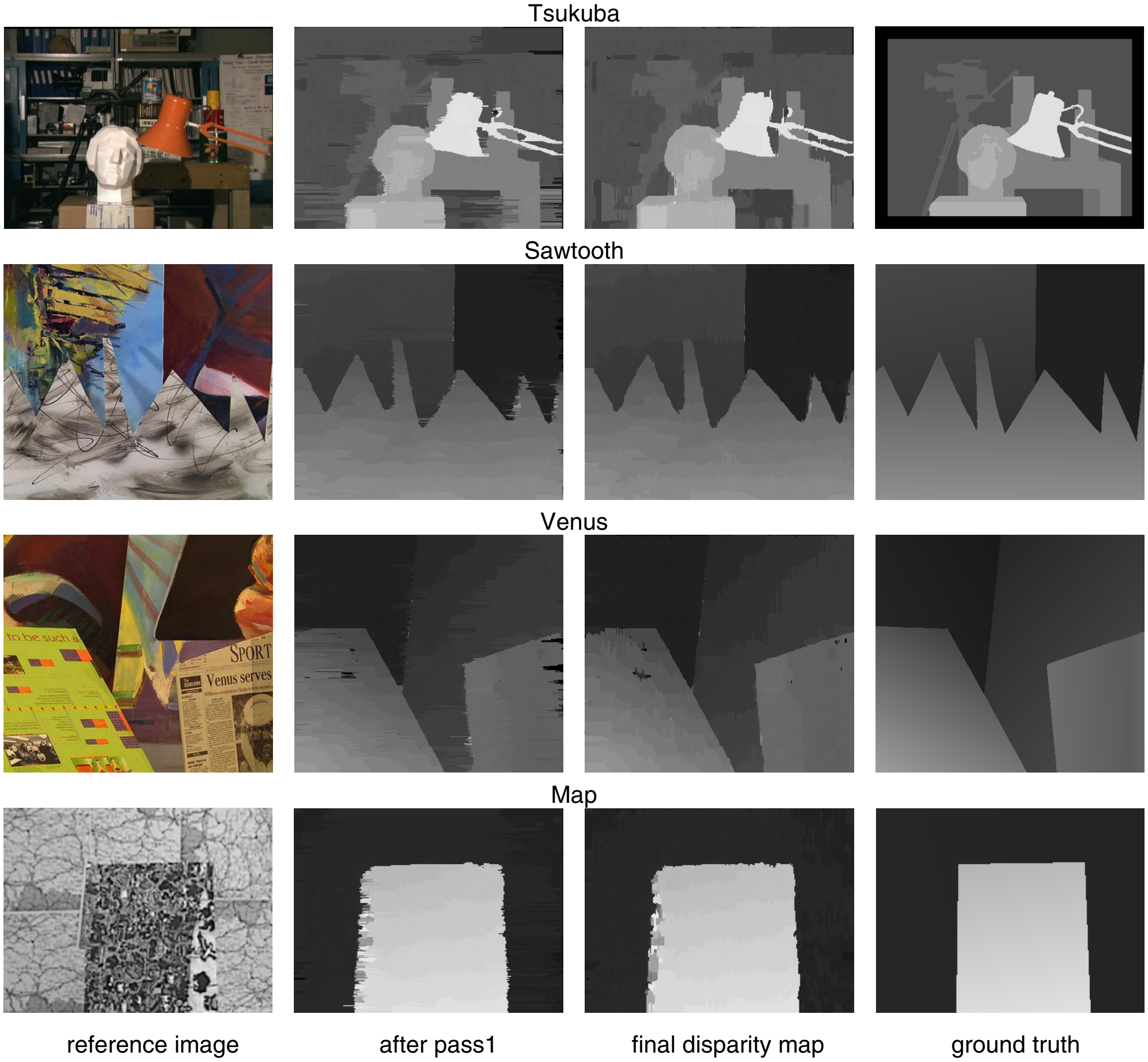

Figure 4. Disparity maps produced by our method.