STEREO MATCHING USING MULTI-DIRECTIONAL DYNAMIC PROGRAMMING AND EDGE ORIENTATIONS

Min Chul Sung, Sang Hwa Lee and Nam Ik Cho

School of Electrical Engineering and Computer Science
Seoul National University, Seoul, S. Korea

ABSTRACT

This paper proposes a stereo matching algorithm which employs an adaptive multi-directional dynamic programming (DP) scheme using edge orientations. A new energy function is defined in order to consider the discontinuity of disparity and occlusions, which is minimized by the multi-directional DP scheme. Chain codes are introduced to find the accurate edge orientations which provide the DP scheme with optimal multi-directional paths. The proposed algorithm eliminates the streaking problem of conventional DP based algorithms, and estimates more accurate disparity information in boundary areas. The experimental results using the Middlebury stereo images [7] demonstrate that our algorithm shows better performance than previous DP based approaches.

***Index Terms*—** Stereo matching, dynamic programming, edge orientation, disparity discontinuity

1. INTRODUCTION

Stereo matching is one of the most active research topics in computer vision, which estimates depth information with two or more images obtained from different viewpoints. There have been many methods to estimate the dense disparity map as the interest in 3D modeling and systems is increased [1]. Recent works on stereo matching are related with the optimization of joint cost functions (or energy). The cost functions are derived from two factors: One is the data matching measure called likelihood term, which is the usual block matching algorithm. The other is a measurement of the correlation between neighboring disparities. The DP based stereo matching algorithms have shown better performances than the usual block matching based methods because DP can exploit the correlations of neighboring disparities. Recent related works are tree DP [3], reliability DP [4], 4-state DP [5], and two-pass DP [6]. Tree DP and reliability DP are focused on fast matching rather than high performance. 4-state DP shows good results only in the case of slanted and continuous surfaces in the stereo images.

In this paper, we propose a stereo matching algorithm which employs an adaptive multi-directional DP scheme using edge orientations. At each edge point, we decide an edge orientation among 36 possible ones using 8 directional chain codes. A new energy function is defined to consider disparity smoothness, discontinuity, and occlusions simultaneously. The directional DP which matched to the edge orientation is applied to minimize this energy function. The proposed method eliminates the streaks problem of the conventional approaches, and estimates better disparity maps at discontinuous boundaries.

The paper is organized as follows. We explain the proposed multi-directional DP scheme and energy function in Section 2. Section 3 presents how to obtain edge orientations. We show the experimental results and conclude the paper in Section 4 and 5, respectively.

2. MULTI-DIRECTIONAL DP

2.1. Multi-paths in the 3D DSI

If the image size is $N \times M$ and the disparity range is $0 \sim D$, we can imagine a 3D ($N \times M \times D$) DSI (disparity space image) for dynamic programming schemes [2]. We usually estimate disparities by finding the shortest path (or minimum cost) in the DSI. Generally, the changes of intensity tend to be small along the edge orientations. Also, disparities along the same edge orientation have more correlation than those with the different ones. Thus, it is reasonable that we select a direction of DP matched to edge orientation when we estimate disparities in boundary parts. This paper proposes a multi-directional DP method, while the conventional DP approaches are based on the scan-line direction. In the paper, we choose a path in DSI that has the same direction as edge orientation, and find the shortest path in the DSI using an energy function.

2.2. Energy function

The proposed energy function consists of data term, smoothness one for homogeneous area, discontinuity one for edge area, and occlusion one as

$$E(d_p)=O_pD(p,d_p)+E_s(d_p)+E_d(d_p)+(1-O_p)E_o(d_p,d_n), \tag{1}$$

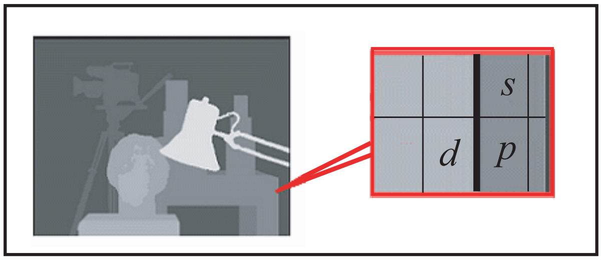


Fig. 1. Relationship with neighborhood disparities

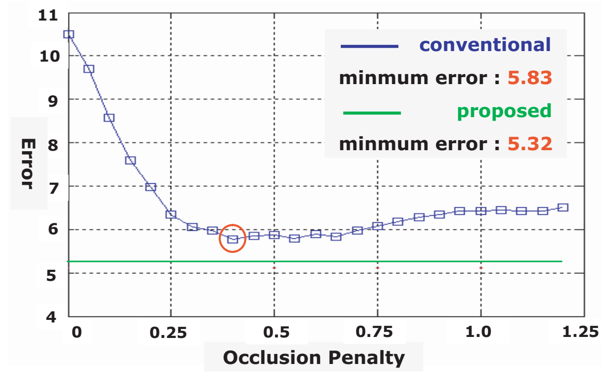


Fig. 2. Comparison of the proposed method and conventional one for assigning occlusion penalty.

where O_p is the binary indicator of occlusion. d_p is the disparity of pixel p, d_n are neighboring disparities with respect to edge information. For example, if a pixel is on the vertical orientation edge, the upper neighboring pixel of p becomes s, and the left pixel across the edge is the discontinuity pixel d as shown in Fig. 1. The correlation of neighboring disparities is considered by the smoothness and discontinuity terms. The data term $D(p, d_p)$ is a cost in the DSI. We apply adaptive window sizes to compute $D(p, d_p)$ considering edge information. A small size window is used for edge regions and a large size one for non-edge regions.

The smoothness term $E_s(d_p)$ is modeled with disparities in the homogeneous region as below

$$E_s(d_p) = |d_p - d_s|, \tag{2}$$

where s is a pixel in the homogeneous region. And the discontinuity cost is defined as

$$E_d(d_p) = \begin{cases} \lambda, & \text{if } d_p = d_d \\ \frac{\lambda}{|d_p - d_d|}, & \text{otherwise.} \end{cases} \tag{3}$$

where λ is a weighting constant. The discontinuity term is modeled for discontinuous pixels across the edge to have different disparities. The more different is the disparities between discontinuity pixels, the more increases the denominator in Eq. (3). And the increase of denominator plays a role in reducing the cost $E_d(d_p)$.

Finally, $E_o(d_p, d_n)$ is the occlusion penalty using neighboring disparities d_n. Occlusion area is partially visible in the stereo images. It is difficult to estimate the accurate disparity in the occlusion area because there is no matching point between stereo images. Instead of the usual constant value, we assign penalty values adaptively based on the relationship among adjacent disparities. We define a penalty function which reflects correlations of neighboring disparities as below

$$E_o(d_p, d_n) = min \frac{|d_p - d_s|}{|d_p - d_d|}. \tag{4}$$

Eq. (4) states that the adaptive occlusion cost is a minimum ratio of disparity differences. This cost makes the disparities discontinuous across the edge and smooth in the homogeneous region in spite of no matching. The data term $D(p, d_p)$ is replaced with the occlusion cost when the data term is too large due to no matching. The interchange of costs is implemented by the binary indicator of occlusion O_p. Fig. 2 is a graph for the comparison of constant occlusion cost and adaptive one. In the case of constant cost, the error of stereo matching is computed for various occlusion costs from 0 to 1.25. The error is changed with respect to the occlusion costs, and the minimum error is 5.83. In the case of our method, the error is measured for adaptive costs. The proposed method has a constant estimation error 5.32 for the minimum error. This result shows that the proposed occlusion cost reduces the errors and makes the estimation performances stable.

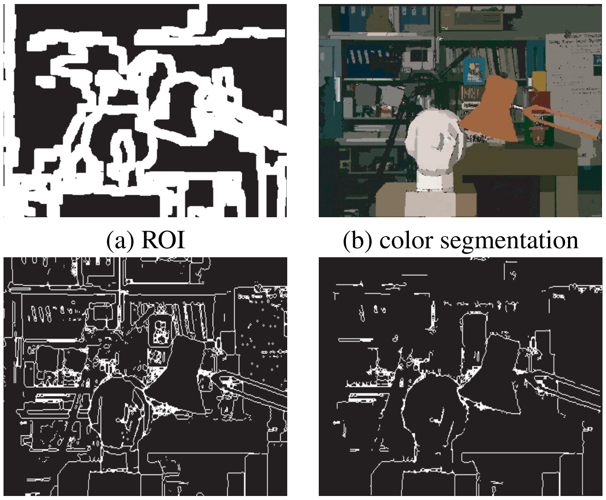
3. EDGE INFORMATION

Since we apply edge orientations to multi-directional DP, edge refinement and decision of edge orientations are very important in the proposed algorithm.

3.1. Edge refinement

The edge information we need is the boundary area of objects. Other edges such as letters and textures on the same plane are unnecessary since they are not related with the disparity discontinuities. General edge detectors are very sensitive to noise and include unnecessary edges. In this paper, we apply color segmentation scheme [8] and region of interest (ROI) to exclude noisy and unnecessary edge information. The proposed process to refine the edges is as follows.

The first step is to roughly find the object boundaries by the bidirectional stereo matching. We estimate the disparity maps changing the reference images, and find an edge map using simple edge detection process followed by median filtering. The second step is to define the ROI. The edge map obtained from the first step is rough and incorrect due to the



(a) ROI (b) color segmentation

(c) edges from result (b) (d) refined edges

Fig. 3. Edge refinement process

limitation of stereo matching. We set the ROI around the edges. The union of $m \times m$ window at each edge pixel is defined as the ROI. The third step for edge refinement is to find edges from color segmentation result. The edges of color segmentation result are finally selected when they are in the ROI obtained from the second step. The orientations of refined edges are determined and used for multi-directional DP. Fig. 3 shows the results of edge refinement process.

3.2. Edge orientation

Most conventional methods for edge orientations extract various directional edges using 3×3 compass mask [9]. However, these methods have to decide the optimal threshold, and show errors in the case of slanted edges except for the multiple angles of 45 degrees. The slanted edges are composed of diagonal, vertical and horizontal line segments in the pixel coordinates as shown in Fig. 4. Human supervisors can perceive that the vertical and horizontal line segments are not independent, but the pixel coordinates can not perceive this fact. This causes the errors in finding orientations of general slanted edges. Therefore, we need some proper processes to classify edge orientations in detail.

In this paper, we present a new scheme to obtain edge orientations using chain codes. First, we label edge pixels using 8 directional chain codes and classify edges into three types such as single, link, and branch. The single is a point which has no connected points. The link is a straight line, and the branch is connected edges with two or three links. For each link and branch, we trace starting and ending pixels along the edge using assigned chain codes. The chain codes help us to find exact starting and ending pixels on the edges. We calculate the gradient with the starting and ending pixels

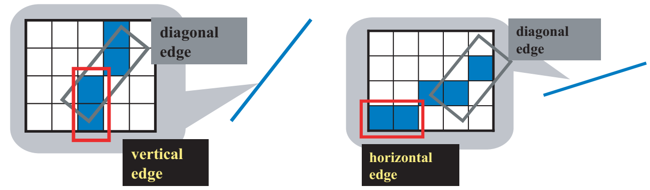
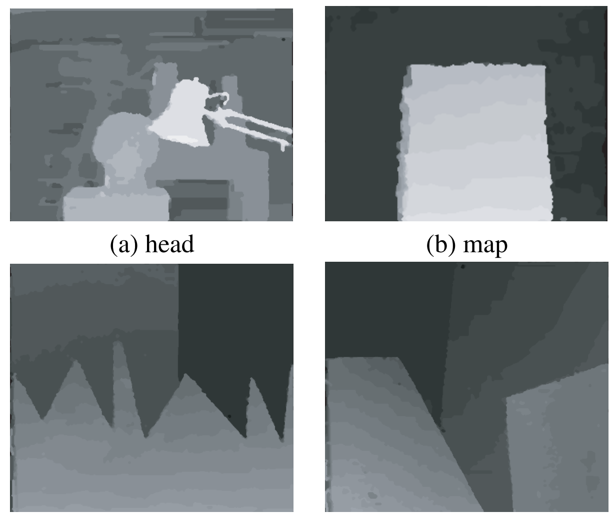


Fig. 4. Composition of diagonal edge



(a) head (b) map

(c) sawtooth (d) venus

Fig. 5. Estimated disparity maps by the proposed method.

as

$$\begin{aligned} \theta &= \tan^{-1}\left(\frac{dy}{dx}\right), \\ \text{where } dy &= |y_1 - y_2|,\ dx = |x_1 - x_2|. \end{aligned} \tag{5}$$

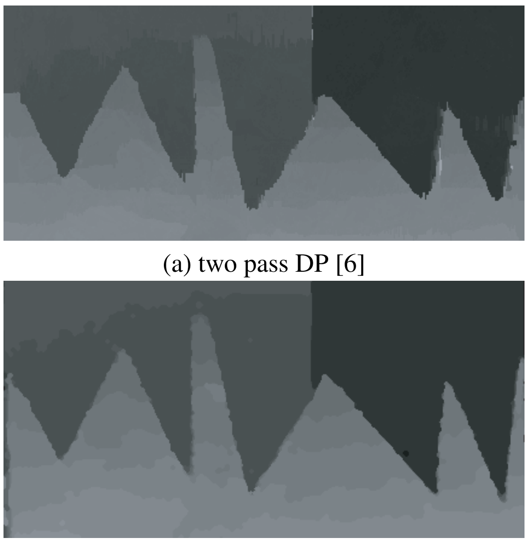
At each edge point, we decide an edge orientation among 36 possible ones using 8 directional chain codes. Then we apply the directional DP whose path is matched to the edge orientation. The proposed energy function in Section 2 is calculated and minimized by the multi-directional DP.

4. EXPERIMENTAL RESULTS

The proposed method is evaluated using the stereo image database of Middlebury stereo site [7]. Test stereo images are head, map, sawtooth, and Venus. Fig. 5 shows the estimated disparity maps by the proposed method. As we can see in Fig. 5, the sharp boundaries of disparity maps are obtained along the real edge orientations. These results mean that the proposed multi-directional DP and energy function are well defined to estimate the disparity maps at the boundaries. Table 1

Table 1. Performance comparison from the Middlebury site.

Matching error(all)				
Algorithm	Head	Sawtooth	Venus	Map
Our method	**1.54**	**1.11**	**0.80**	**0.76**
Two pass DP [6]	1.53	0.61	0.94	0.70
Reliability DP [4]	1.36	1.09	2.35	0.55
Tree DP [3]	1.77	1.44	1.21	1.45
4-state DP [5]	4.70	1.43	1.18	0.30



(a) two pass DP [6]

(b) proposed method

Fig. 6. Comparison of the proposed method and two pass DP for sawtooth image.

shows all error results of various DP methods from the Middlebury stereo site. As shown in Table 1, our method shows good estimation results in overall performances. Our method is ranked at the 2nd position in Table 1, but the estimated disparity maps are better than those of two pass DP [6]. Fig. 6 shows the boundaries of estimated disparity maps by the two pass and proposed methods. As shown in Fig. 6, the disparity boundaries by the proposed method are better and closer to the truth than two pass algorithm. And the proposed method shows the best performance on the untextured errors as shown in Table 2. According to various experiments, the proposed algorithm outperforms other DP-based methods in discontinuous and smooth regions. Some overall errors will be reduced by the parameter optimization in the energy function.

5. CONCLUSIONS

In this paper, we have proposed a new stereo matching technique using multi-directional dynamic programming (DP) based on edge orientations. The proposed method estimates the disparity maps by using the optimal directional DP. The direction of DP is based on the edge orientations which is extracted by chain codes. We have also defined an energy function with data, smoothness, discontinuity, and occlusion terms. The energy function is minimized in the multi-direction DP. Experimental results on image database of the Middlebury stereo site [7] show that our algorithm provides better performances than the previous DP based approaches.

Table 2. Performance comparison of untextured part error.

Matching error(untex.)			
Algorithm	Head	Sawtooth	Venus
Our method	**0.41**	**0.07**	**0.45**
Two pass DP [6]	0.66	0.02	0.95

6. REFERENCES

[1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithm," *IJCV,* vol. 47, no. 1, pp. 7-42, April 2002.

[2] A. F. Bobick and S. S. Intille, "Large occlusion stereo," *IJCV,* vol. 33, no. 3, pp. 1-20, 1999.

[3] O. Veksler, "Stereo correspondence by dynamic programming on a tree," *Proc. CVPR,* vol. 2, pp. 384-390, 2005.

[4] M. Gong and Y. H. Yang, "Fast stereo matching using reliability-based dynamic programming and consistency constrains," *Proc. ICCV,* vol. 2, 2003.

[5] A. Criminisi, J. Shotton, A. Blake, C. Rother, and P. H. S. Torr, "Efficient dense-stereo with occlusions and new view synthesis by four state DP for gaze correction," *IJCV,* online published June. 2006.

[6] C. Kim, K. M. Lee, B. T. Choi, and S. U. Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points," *Proc. CVPR,* vol. 2, pp. 1075-1082, June 2005.

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

[8] D. Comaniciu and P. Meer, "Robust analysis of feature spaces:Color image segmentation," *Proc. CVPR,* pp. 750-755, June 1997.

[9] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 2nd ed., Prentice Hall Inc., 2002.