Multi-directional Greedy Stereo Matching

[1]Seung-Hae Baek, [1]Soon-Yong Park, [1]Soon-Ki Jung, [2]Sang-Hee Kim, [3]Jeong-Hwan Kim
[1] Department of Computer Engineering, Kyungpook National University
1370 Sankyuk-dong, Buk-gu Daegu 702-701, Korea
[2]ADD
Yuseong PO Box 35-1, Daejeon 305-600, Korea
[3]AR Vision Inc.
3-1 Doryong-dong, Yuseong-gu, Daejeon 305-340, Korea
E-mail: [1]eardrops@vision.knu.ac.kr, [1]skjung@knu.ac.kr, [1]sypark@knu.ac.kr, [2]falcon@add.re.kr, [3]kjhnew@ar-vision.com

Abstract: In this paper, we propose a new stereo matching method to reduce a memory size and handle an large image such as satellite images while its performance satisfies speed and accuracy. Our new method is based on a Multi-directional Greedy algorithm and RANSAC. First, we obtain a depth image along multi-directioanl scanlines using the Greedy algorithm. Then, we find reliable areas from the distribution of several depth images using RANSAC. Finally, we extend the reliable areas by iterating the Greedy algorithm and RANSAC several times starting with the previously obtained reliable areas and decide a final depth image. Experimental results show that our algorithm proves the possibility of stereo matching for an enormous image using low memory size while satisfying the performance from the view point of speed and accuracy.

1. Introduction

Stereo Matching is a technique to find out same image coordinates from stereo images captured from a pair of cameras or more. Many investigations in stereo matching can be categorized as two schemes in common. One is local matching and the other is global matching. Local matching consider a point which has smallest energy cost comparing to every local areas as corresponding point. Global matching consider a point which make global energy cost smallest as corresponding point. Recent investigations more focus on global matching schemes and some of them present interesting results.

Some global matching approaches such as BP(Belief Propagation), GC(Graph Cut), DP(Dynamic Programming) and SGM(Semi-Global Matching) also have weak points. BP and SGM require huge memory space and take long computation time for global energy minimization of each matching pair. A fast global matching approach such as DP algorithm yields the scan-line problem. Especially, in stereo matching of huge images such as satellite images, short computation time and proper memory space is important as much as accuracy.

In this paper, we address a stereo matching problem for an enormous data such as satellite images. Memory and speed are the most important factors to determine the performance of the method. Thus, we propose a new stereo matching algorithm to reduce a memory size with satisfying accuracy performance. A new strategy - multidirectional greedy algorithm is used to estimate depth images and then RANSAC is developed for extracting reliable regions with previously obtaiend several depth images. For the reliable regions, we iterate the greedy algorithm and RANSAC to find a final global depth image. Our approach is smilar to the SGM algorithm. It calculates the matching cost of an image pixel, accumulates the matching costs for the pixel in 8 or 16 directions, and compares the aggregated matching cost. In comparison to SGM, the proposed algorithm diffres in terrns of calculating the depth directly from multi-directional greedy algorithm. As a novel approach, we apply RANSAC to extract reliable areas from several depth images. The proposed method needs a small memory size because we store only previous position information for each matching direction, while the conventional method needs a large memory size for storing matching costs.

The organization of this paper is as follows. Section II describes the purpose of proposed algorithm and its flow. Section III explains a conventional greedy algorithm and describes the application of the modified greedy algorithm in the proposed method. Section IV and V described reliable areas extraction using RANSAC and experimental results.

2. Overview of the proposed algorithm

Proposal stereo matching algorithm has following two assumptions. First assumption is that the point which has disparity change always has intensity difference. Second assumption is that the disparity change in the textureless area is small, for instance 1 or 2 disparities. We progress our experiment with these assumptions.

The proposed algorithm's detail contents are briefed in the following sentences.

- Greedy Algorithm : The number of matching directions are 8 or 16. Processing along each matching direction, a disparity of current matching point is confirmed by comparing corresponding energies of points near to the disparity of previous matching point. But, if a current point is on an edge, the disparity can not be near to the disparity of previous matching point. So, the current point disparity should be found through all possible disparity range. As repeating to find disparity, we can get multi-directional disparities in the end.

- Extraction of reliable area using RANSAC : This step is to find out reliable disparity regions from multi-directional disparity maps that are confirmed using the Greedy algorithm. In an image point, if all the disparities from multiple directions are equal, then we can say the disparity is correct value. In the other hand, if some disparity which are different outstandingly will be treated as an error and others are gathered to determine the disparity value. Then we consider the gathered disparity value as a reliable one. But, if the disparity values are distributed widely, we can't confirm the disparity of the point and the point remains as a hole.

- Iterative hole filling : After the Greedy algorithm and RANSAC processing, we can get reliable areas even though it does not cover the whole image. We execute the Greedy Algorithm again through several matching directions using the disparity values in reliable areas which are confirmed in previous processing. After that, we execute RANSAC to the area except the area whose disparity is already confirmed. We consider reliable area's value as confirmed disparity and fill the holes by iteratively doing these steps.

3. Extraction of a disparity image using a multi-directional Greedy algorithm

If only one matching direction is used for Greedy Algorithm, the result will not be reliable. But, if Greedy Algorithm execute through several matching directions, we can expect that right results exist from some of them.

3. 1 Error cost of correspondence

In this paper, disparity map is measured on the basis of left image. So, error cost between left point and right point which is *d* pixels far from left point is calculated in advance. To calculate the error cost, we use the normalized SAD. Before cost computation, every left pixel is classified to texture area or un-texture area using a 3x3 mask window. To classsify a pixel, the distribution of one pixel's intensity is compared to the other's around the pixel. The reason of this classification is to differ the way of using normalized SAD. The window size is set larger for an un-texture area than that in a texture area. After pixels are classified on the basis of texture densisy, normalized SAD is calculated about whole left image pixels and the expression is following.

$$E_p(d) - \Sigma_{q \in W} \left| I_{l_q} - I_{r_{q+d}} - (ave_l - ave_r) \right| \quad (1)$$

In the expression 1, I_{l_q} and $I_{r_{q+d}}$ are the intensity of pixel p in the left image and pixel q+d for the right image, and W are pixels in the window centering around pixel p. Parameter ave_l and ave_r are average intensity of left image's pixel q and right image's pixel q+d.

3. 2 Greedy algorithm

Greedy algorithm's main concept is to take best one just considering current circumstance. Usually, local stereo matching searches over all possible disparity range to find out proper disparity for each pixel. But, we can reduce the disparity search range using greedy algorithm with previous pixel's disparity. Greedy algorithm can be used in many cases about stereo matching, however, we use greedy algorithm finding out disparity through matching directions. If using just one matching direction, it's not easy to correct the error which is calculated. So, we propose a multi-directional greedy algorithm. But, more directions don't guarantee higher accuracy and we think that 8 or 16 directions are proper. Figure 1 shows that the greedy algorithm is applied to multi-directional stereo matching.

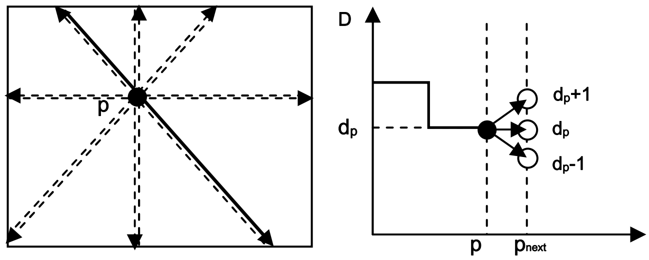

Figure 1. Matching direction and range of disparity

The left image of the figure shows directions through which the greedy algorithm progress, and the right image shows the relation between p and p_{next} that are on the same matching direction. In the right image of the figure, we use previous pixel, pixel p's disparity to calculate disparity of pixel p_{next} on the progress which calculate the disparity according to the matching direction. Possible disparity range can be from d_p+1 to d_p-1. By comparing costs of the candidates, we decide the disparity of the pixel p_{next} which has smallest $E_p(d)$ than the other candidates. At this time, if there's a change between previous and current disparity, the penalty(δ) will be paid to prevent outstanding disparity change.

$$d_{p_{next}} = \min_d \{E_{p_{next}}(d_p - 1) + \delta, E_{p_{next}}(d_p), E_{p_{next}}(d_p + 1) + \delta\} \quad (2)$$

The above equation expresses the relation between current pixel and next pixel. But, if the current pixel is on the edge as progressing the greedy algorithm according to the matching directions, we have to choose disparity *d* which has smallest $E_p(d)$ value among the candidates from all possible disparity ranges. Because in such pixel which is on the edge, disparity can be changed suddenly as stated above. In addition, we have to find out proper disparity from all possible disparity range on the start pixel of all matching directions. The Figure 2 shows some results that are caused by the greedy algorithm from 4 matching directions(left to right, right to left, up to down, down to up).

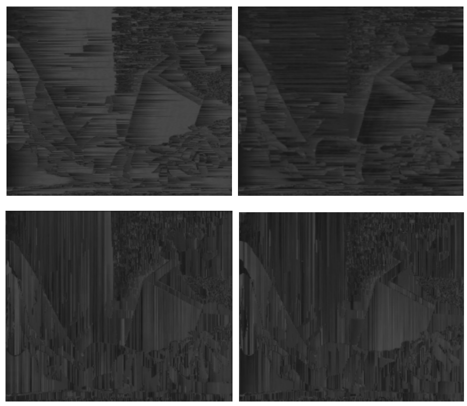

Figure 2. Result of 4 directions' Greedy algorithm to Teddy image

4. Iterative extraction and expansion of the reliable area

After confirming several disparity images using the multi-directional greedy algorithm, we obtain the final disparity image from these images. If all disparity values which are obtained from the algorithm are equal, then it's an ideal case. But, actually, the disparity vlaues are different. In this case, we use RANSAC to find out correct disparity value of a pixel, which estimates correct disparity from all of them. After obtaining the disparity from reliable areas, and we iterate the greedy algorithm and RANSAC in a couple of times and to expand reliable areas.

4. 1 Extraction of reliable area using RANSAC

RANSAC is used to remove the outlier which is not in the acceptable scope and considered as noise when sample's distribution is not flat. We use the left image to operate RANSAC to find realiable iamge areas. In this paper, we consider the disparity of a pixel as reliable value when the ratio that is the percentage of suitable value to all is bigger than a pre-dxefined value as follows.

$$D_p = \left\{ \mathrm{Ave}(S_i) \,\middle|\, \frac{N(S_i)}{M} > RANSAC\_ratio \right\}$$

Where, $S_i = \{ d_j \,|\, |d_i - d_j| \leq \mathrm{Difference\_threshold} \}$ (3)

Ave(S) : average of set S' elements ,i=1,…,M

In Equation 3, i and j are matching directions. If the greedy algorithm is perfomred according to M matching directions, then M disparity values are obtained in pixel p. At this time, S_i is the set of the value d_j whose difference to d_i is smaller and equal than threshold(about 1 pixel). If there is a set S_i whose elements number ratio to M is bigger and equal than RANSAC_ratio(user defined), then the average of the set S_i's elements is confirmed to a pixel p's disparity value.
In this process, the result obey to parameter RANSAC_ratio and Difference_threshold. If the value of Difference_ threshold is small, then correct result is expected in high probability, but the resulting image has many holes because of small number of reliable areas. In an other way, the bigger threshold value brings some incorrect results but small nuber of holes because of large reliable areas. RANSAC_ratio is in opposition to Difference_threshold. If the value is large, then correct result is expected but many holes are occurred, otherwise the result is not so much correct but have small number of holes. But, RANSAC_ratio has different value according to texture's existence. The reason why it has difference value is because untexture area need more reliability.

4. 2 Processing of greedy algorithm from reliable area

When the reliable areas are obtained using RANSAC, it's not easy to obtain disparity value over whole image. So, in this paper, we propose that the greedy algorithm progress again through M matching directions from reliable areas which are already confirmed. After M directional disparity images are obtained, the reliable areas are obtained again using RANSAC about unreliable areas which are not confirmed in the previous step.

4. 3 Iterative expansion of the reliable area

If a hole is still left on an image except reliable areas, then we can fill the hole with reliable area by the greedy algorithm and RANSAC. Iterating these processes expand the reliable areas. After 3 or 4 times iterations, the remaining holes are not filled any more. We can see the results of the reliable area's expansion by steps in Figure 3. The result of the last step is a final disparity image.

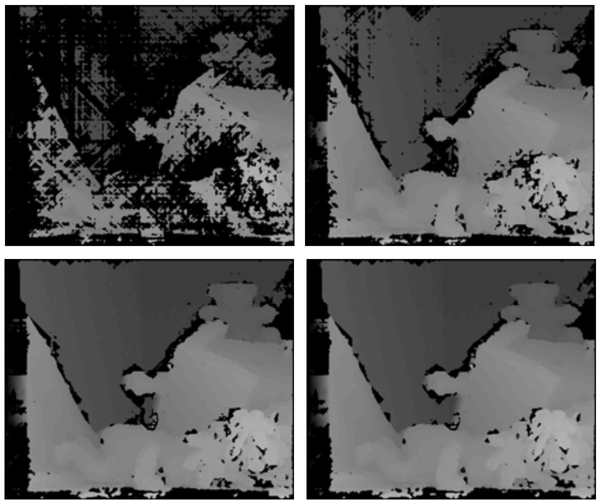

Figure 3. Results of each stage, 1~4

5. Experimental results

In this paper, we perform experiments using the middlebury stereo images, Teddy, Venus, Cones, and other stereo images which are Tsukuba and a part of satellite image. The stereo matching is progress on the basis of left image and uses the information of gray intensity, not color information. When the experiment is done with middle-bury images and Tsukuba, the window size of error cost on texture area is 3x3 and that on untexture area is 5x5. Ransac_ratio is 0.4 on texture area and 0.6 on untexture area. Figure 5 shows the results after 4 times iteration. Though the results have some holes, yet the accuracy is as good as the results of SGM whose window size of error cost is 3 and aggregation directions are 8. But, the amount of memory usage in our method is much smaller than that of SGM and the processing time is shorter in factor of 1/2. From experiemnts, we can see that our method becomes more effective as the disparity search range is bigger. When the experiment is progressed with satellite image, the window size of error cost is 3x3 to 13x13 and Ransac_ratio is 0.7 to 0.3 with iteratively expanding the reliable area. In Figure 4, the result of satellite image has a comparatively correct building's disparity, though there are some errors in some image regions.

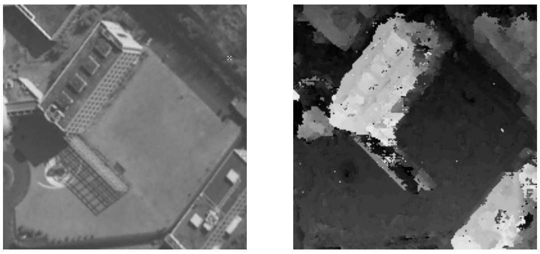

Figure 4. Application to the satellite image and results

6. Conclusion

We propose a novel global stereo matching algorithm which is fast and needs small memory, but the efficiency is comparative to existing stereo matching algorithms. We obtain M disparity images using the greedy algorithm through M matching directions, and reliable areas from M disparity images using RANSAC. The final disparity image is obtained by iteratively using the greedy algorithm and RANSAC to expand reliable area to whole image area.
Our method, different to existing method, doesn't need to search whole disparity range, because of greedy algorithm. The disparity search range is just between previous disparity ±1. This narrow disparity search range reduces processing time and amount of system memory.
In this paper, our novel method offers accuracy as good as SGM's. And the processing time is reduced about 1/2 times. The system resource memory is relatively smaller than SGM'. Therefore, it's possible to obtain depth image from huge size of stereo imagea. Disparity holes left in the final image are problems to solve. In addition, we have to do more experiments to obtain better result from the satellite stereo images which are difficult to solve.

Acknowledgment

This research has been supported by funding from ADD and AR Vision Corp., Korea.

References

[1] Heiko Hirschmüller, "Stereo Vision in Structed Environments by Consistent Semi-Global Matching" IEEE Conference on Computer Vision and Pattern Recognition (CVPR), New York, NY, USA, June 17-22, 2006
[2] R. Neapolitan and K. Naimipour "Foundations of ALGORITHS using C++ pseudocode" jones and Bartlett, 2nd edition, 1997, pp.134-175.
[3] A. Klaus, M. Sormann, and K. Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," in International Conference on Pattern Recognition, 2006.
[4] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," in Proceedings of the Sixth IEEE International Conference on Computer Vision, Mumbai, India, January 1998, pp. 1073–1080.
[5] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," in International Conference for Computer Vision, vol. 2, 2001, pp. 508–-515.
[6] Q. Yang, L. Wang, R. Yang, H. Stewenius, and D. Nister. "Stereo matching with color-weighted correlation, hirarchical belief propagation and occlusion handling" IEEE Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 17-22 June 2006.
[7] Heinrichs, M., Rodehorst, V., Hellwich, O."Efficient Semi-Global Matching for Trinocular Stereo"PIA07(185).
[8] Heiko Hirschmüller, "Stereo Processing by Semi-Global Matching and Mutual Information" IEEE Transaction on Pattern Analysis and Machine Intelligence 2007

Figure 5. Comparing our method with SGM