Advances in Intelligent Systems Research, volume 147
International Conference on Network, Communication, Computer Engineering (NCCE 2018)

Optimization Method for Stereo Matching Based on Minimum Spanning Tree

Jingyu Yi [a)], Binwen Fan [b)] and Xiaopeng Cui

Harbin Institute of Technology, Shenzhen, China

[a)] yjy1042946923@163.com
[b)] Corresponding author: 472299471@qq.com

Abstract. This paper describes the application of minimum spanning tree algorithm in process of stereo matching and optimization of depth map. A stereo matching method for the real-time application scene of binocular vision is proposed, which the stereo matching is completed under the sample of the image and reduced the image size by the minimum spanning tree of original image. On the one side, the method reduces the amount of the image under the sample, and then uses the minimum spanning tree to form the relationship between the original image and the lower sample image. On the other side, the lower sampling image uses the minimum spanning tree to the cost aggregation, so that the cost aggregation can through the local limit. This method can generate global disparity map without mismatched black voids, and it also has a good effect on optimizing depth map obtained by other methods.

Key words: Stereo matching; Disparity refinement; MST; image up-sampling.

INTRODUCTION

Stereo matching is to restore 3D information from 2D images. The key is how to accurately match the points of left and right images. Binocular vision stereo matching is generally divided into four steps: matching cost calculation, cost aggregation, disparity selection and disparity refinement [1]. Generally, the algorithm is divided into global algorithm [2-4] and local algorithm [5-6]. The global algorithm generally constructs the global energy function and improves the image precision by minimizing the global energy, but it cannot meet the requirements of real time. Generally, there are graph cut, dynamic programming, etc. The local algorithm generally uses the domain window for the cost aggregation, and the speed can meet the real-time requirements, but it is easy to mismatch, and the image quality is generally inferior to the global algorithm. The MST (minimum spanning tree) algorithm used in this paper is similar to the local algorithm for the cost aggregation, but the cost aggregation depended on the whole tree of the image is used achieves the global algorithm effect [7].

In the paper, the MST is used in two different ways. The first is cost aggregation using in the image under lower sampling. The second is reconstructing the disparity map of original size from the small size image.

COST AGGREGATION OF STEREO MATCH USING THE MST

Matching Cost Calculation

Matching cost is a measure of the similarity between the left and right image pixels, and the lowest cost is regarded as the matching point. According to the limit conditions, the corresponding points should be on the same line, and the left point $\mathrm{p}_l = (x, y)$ corresponds to the right matching point $p_r = (x-d, y)$. The color cost of the RBG image is calculated as follows:

$$C_C = \sum_{i=R,G,B} \left| I_i(p_l) - I_i(p_r) \right| \tag{1}$$

Considering the difference of illumination between the left and right images, the color gradient characteristics which is insensitive to light is added to the cost. The gradient of the RGB three channels is as follows:

$$\begin{aligned} G_G = &\sum_{i=R,G,B} \left| \nabla_x I_i(p_l) - \nabla_x I_i(p_r) \right| + \\ &\sum_{i=R,G,B} \left| \nabla_y I_i(p_l) - \nabla_y I_i(p_r) \right| \end{aligned} \tag{2}$$

The joint matching cost based on image color and gradient features is as follows:

$$\begin{aligned} &C_d(p_l, p_r) = \left| w_c \times C_c + w_g \times C_G \right| \\ &w_c = \frac{C_c}{C_c + C_G} \\ &w_g = \frac{C_G}{C_c + C_G} \end{aligned} \tag{3}$$

The w_c and w_g are the weights of color and gradient respectively, which can reflect the importance of each feature in the current matching area.

Cost Aggregation Using the MST

Each pixel chooses the minimum cost to calculate the disparity, but the single pixel matching is easily interfered by the noise, but the cost aggregation method makes each cost more reliable. Cost aggregation refers to supporting the pixel by other related pixels. The traditional local algorithm of cost aggregation is using the domain of using pixels. However, the values can be easily interfered by poor correlation pixels such as edge location. Therefore, using MST to establish association with full image should achieve better results.

The stereo matching of MST is including three steps. The first is constructing an undirected connected graph by taking pixels as nodes and the difference between pixels as edges. The second is constructing a minimum spanning tree under the undirected connected graph. In the part, each node has a parent node, so they have three sub nodes at most.

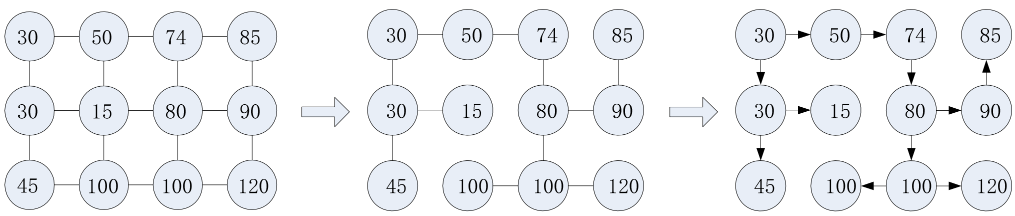

FIGURE 1. The process of MST

The third is recalculating the value of each node. In the MST, the smallest adjacent points of the gradient are first connected, so the points connected are the most closely related points, and the mutual support between these points is the most effective. Because of the parent-child relationship in trees, we can use all the descendants and ancestor nodes of the center node to support the node by the top-down and bottom-up way, just as figure 2. We can judge the degree of association between the two points by the weight of edges. The global relationships of trees will also associate unrelated blocks. In order to solve this problem, a threshold can be set. When the correlation degree of two pixels is lower than a certain value, we think that they are in different blocks, and the cost of matching is no longer transmitted to each other.

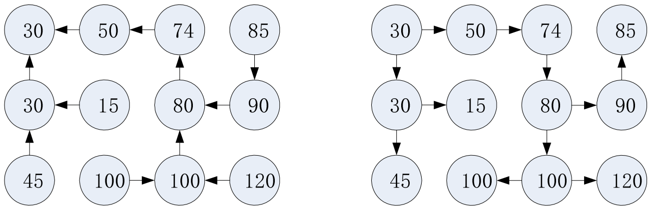

FIGURE 2. The top-down or bottom-up way of cost aggregation

Disparity Selection

After cost aggregation, the matching cost between pixels is more reliable. And then, using the Winner-take-all algorithm to select the pixels whose cost is minimum as matching pixels. The disparity map and depth map of the image can be calculated from the position difference between the left and right images.

DISPARITY UP-SAMPLING BASE ON MST

The computation of the depth map for a large original image will be very large and the computation speed of each frame in the video stream will be very slow. So, using small size image to compute the disparity map and restore to a large size one should be a great idea.

The first is sampling the original image to a small one and using the image of small size to build the disparity map of small size. It is supposed that the down sampling ratio is 2 and the disparity map is based on the left image, just as shown in figure 3. The pixel of disparity map and left small size image must be one-by-one correspondence. And the pixels of left small size image are come from the original image. The second is inserting the pixels to the corresponding position of the original left image and setting the unknown position to 0, and then using the relationship of tree to rebuild the disparity of 0 positions. Because of the MST, the 0 positions can easily find the most similar point and use the similar point to rebuild itself.

This method can not only be used to restore depth map with up-sampling, it can also be used for depth map optimization of disparity. There always are some black holes in stereo matching disparity map, presenting the palace of mismatched. When using the MST to rebuild the disparity map, the black holes are considered as 0 positions, so the values of black holes will be replaced by reliable values of tree.

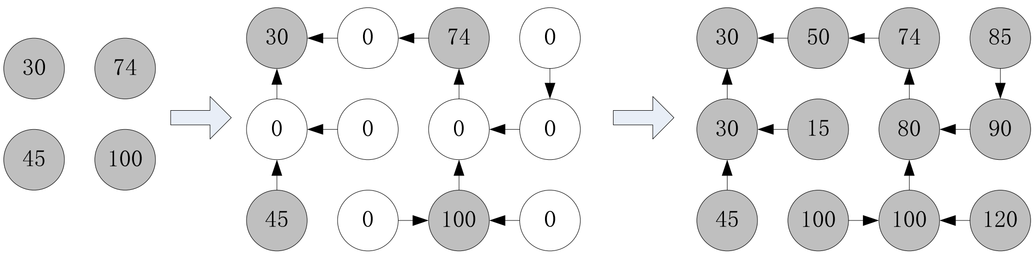

FIGURE 3. Rebuild the disparity map of original size

EXPERIMENT RESULT

We test the MST stereo matching algorithm in human body detection, as shown in figure.4. And then we compare the MST method to SGBM.

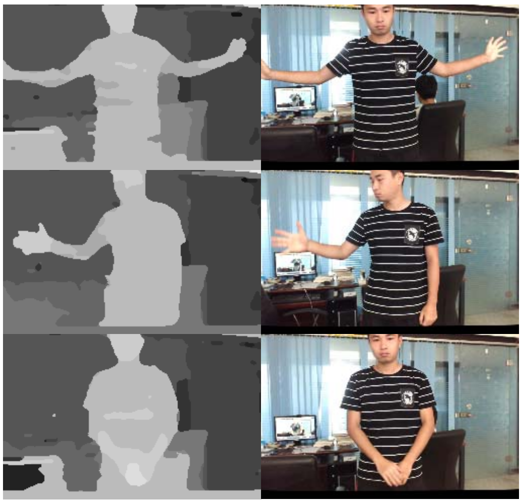

FIGURE 4. The disparity map under MST

The result is shown in the figure 5, that the SGBM can get a pretty map reflecting the distance of the frame, but the details are not elaborate enough, and the edge of human body have black holes. The MST can get more detail images and do not produce black holes.

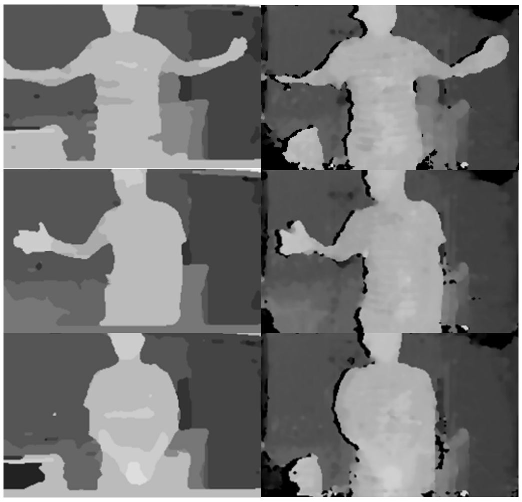

FIGURE 5. Compare the MST to SGBM (the left is MST, the right is SGBM)

In the test of optimizing the disparity map using MST algorithm, we reconstruct the disparity map calculated by SGBM and fill the black holes. The results show that the algorithm can preserve disparity value of SGBM algorithm and effectively fill black holes. Compared with the traditional Gauss filter and the median filtering, the MST is more effective.

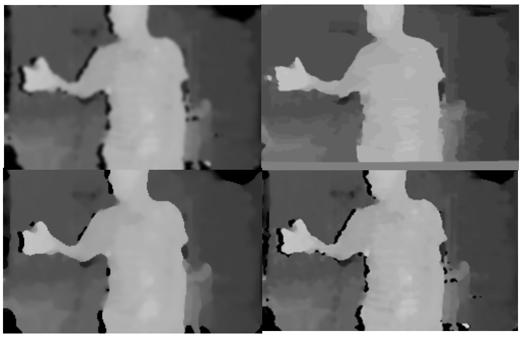

FIGURE 6. Disparity refinement (the upper left is original image, the upper right is MST image, the lower left is median filter image, the lower right is gauss filter image)

CONCLUSION

In this paper, the MST algorithm is applied to stereo matching and disparity optimization of binocular stereo vision. The first is to use MST aggregating cost, breaking through the domain window limit of the traditional local algorithm in cost aggregation and establishing the association of all the global pixels using MST. The algorithm has achieved good results. Compared with the commonly used SGBM, it can produce finer disparity map and avoid building mismatch area.

Then the MST is applied to disparity optimization. Aiming at the mismatch black hole part of SGBM algorithm, MST can fill it by reliable nodes in the tree of image. It can be applied to the optimization of most local stereo matching algorithms in mismatch areas.

ACKNOWLEDGMENTS

This work was supported by Special funds for the development of strategic emerging industries in Shenzhen (no. JCYJ20150625142543448).

REFERENCES

1. Scharstein D, Szeliski R. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms[J]. International Journal of Computer Vision, 2002, 47(1-3):7-42.
2. Y. Boykov, O. Veksler, and R. Zabih. Fast approx-imate energy minimization via graph cuts. PAMI, 23(11):1222–1239, 2001.
3. Sun J, Shum H Y, Zheng N N. Stereo matching using belief propagation[C]//Computer Vision—ECCV 2002. Berlin/Heidelberg: Springer, 2003: 787-800.
4. Hirschmuller H. Stereo Processing by Semiglobal Matching and Mutual Information[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 30(2):328-341.
5. Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. International Journal of Computer Vision, 2002, 47(1): 7-42.
6. Yoon K J, Kweon I S. Locally adaptive support weight approach for visual correspondence search[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 2: 924-931.
7. Yang Q. A non-local cost aggregation method for stereo matching[C]// Computer Vision and Pattern Recognition. IEEE, 2012:1402-1409.