The minimum spanning tree in stereo vision

Efron Manik

Department of Mathematics Education, University of HKBP Nommensen, Medan-Indonesia

efmanik@gmail.com

Abstract. Image can be presented in graph or grid graph. Segmentation is widely used in computer vision. Image segmentation is done in various approaches. One approach to image segmentation is the use of a minimum spanning tree. Almost every method of determining disparity in stereo vision uses image segmentation. The method in this paper is in accordance with the method of Segment-Based Stereo Matching using belief propagation and self-adapting dissimilarity measure with improvements for segmentation, and apply local window-based matching steps. The disparity calculation method built with minimum spanning tree approach is one alternative that can be used to calculate disparity for image pair with simple stereo configuration.

1. Introduction

Image can be presented in several ways, such as grid graph, triangular grid graph, and super grid graph like in Figure 1. The grid graph is a graph whose sides connect each pixel of digital image with four pixels around it: pixel above, down, left, and right. While the triangular grid graph is a graph that each pixel is connected to six pixels around it, namely: upper left pixel, bottom right, up, down, left, and right. Furthermore, the super grid graph is a graph that each pixel is associated with eight or all of the surrounding pixels. The multiscale graph-based segmentation (MGS) algorithm uses the popular graph-cut approach that presents the image as grid graph [1].

Segmentation is widely used in computer vision. Image segmentation is done in various approaches. Each approach has its advantages and disadvantages. For example, the mean shift method adopts a global calculation but still has weaknesses, namely: the wide segment boundary corresponding to a bandwidth value is not always a combination of segment boundary elements for smaller bandwidth value [2]. This weakness occurs because the method cannot exclude outliers data. For improved results, Peter [3] used the minimum
spanning tree to segment the image. In a tree, we can look for its sub-tree that matches another tree or segment. This tree match is widely discussed in the tree profile [4].

Many algorithms in stereo vision use image segmentation in the step section. The MGS algorithm uses the popular graph-cut approach that presents the image as grid graph [1]. In fact, almost all stereo vision uses segmentation as part of the steps in its method, among others: belief propagation methods [5, 6, 7, 8, 9]. Methods that rely on segment boundaries and metric similarity [10, 11] also use segmentation. The optimization method with global constraints [12, 13] also uses image segmentation in steps. The problem to be solved in this paper is whether the segmentation produced by the minimum spanning tree method can be used in the disparity pixel calculation step in stereo vision.

This paper consists of several sections. The second section discusses the segmentation using the tree that will be needed in the next section. The third section is the stereo vision which is the discussion of the problem posed in the first section. This paper ends with a conclusion section.

2. The Segmentation Using the Tree

The graph is used to present discrete objects and relationships between them. Image is one example of a weighted grid graph, where images can be viewed as a collection of points from pixels, and adjacent pixels linked by sides. The weight of the side can be given by the difference in the value of the intensity of the adjacent pixels. The graph of the image formed in this way can be seen in figure 2.

The following theorem or statement [14] is the theorem used for shows the singularity of image segmentation using the minimum spanning tree.

**Theorem** Let $\alpha \in \mathbb{R}$, and $S$, $T$ be the minimum spanning tree of graph $G$. Suppose that all of the weights are greater than or equal to $\alpha$ removed from the trees $S$, and $T$. Then the components connect $S_1$, $S_2$, ..., $S_p$ from trees $S$ and $T_1$, $T_2$, ..., $T_q$ of the $T$ tree will be formed. Then $p = q$ and if $V(S_i) \cap V(T_j) \neq \emptyset$ then $V(S_i) = V(T_j)$.

Suppose that graph $G$ is a graph constructed from a digital image, and the connected component of graph $G$ is viewed as a segment of the digital image. From this theorem, we can
conclude that two sets of image segments produced by different minimum spanning trees are the same.

We use the Kruskal algorithm for the graph in figure 2.b, so the tree in figure 3.a is obtained. After calculation for the tree, the average weight of all sides of the tree (μ) is equal to 8.5 and the standard deviation (σ) equals 10.9. By using a threshold value equal to μ+σ, i.e., 19.4, all sides whose weight is greater than the threshold value are removed. So that sub-trees or segments like Figure 3.b are obtained. If the original image (figure 3.c) is compared to the image of the segmentation (fig. 3.d) then the segmentation produced by the minimum spanning tree method gives a good result.

3. Stereo Vision
Stereo vision is one area of vision computers that are growing very rapidly. Stereo vision learns how to determine the three dimensional coordinates of the digital image pair taken from a pair of cameras with simple stereo configurations. The three dimensional coordinates of each pixel of an object are calculated based on the disparity of each pixel pair between the left and right images.

The Middlebury homepage (http://cat.middlebury.edu/stereo/) provides pairs of images: Tsukuba, Venus, Teddy, and Cones with simple stereo configuration. There are three categories of regions that are calculated, namely: non-occluded pixels, near occluded regions, and all pixels. The best method in this race is the method of "Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure" [5]. This method
calculates the disparity of each pixel with several steps: homogenous regions in reference image (segmentation), apply local window-based matching, extract a set of disparity planes, and approximate optimal disparity plane assignment.

In this paper we will create a visualization of Cones images in the graph, only 137×137 pixels, which are: lines 71 – 207 and columns 61 – 197, which can be seen in Figure 4. The left and right image will be converted into grid graph. Furthermore, the minimum spanning tree for each graph will be determined. After that the mean (μ) and standard deviation (σ) weights of all graph-forming sides corresponding to the left digital image are calculated. The linear combination of μ and σ is the selected threshold value. All sides of the minimum spanning tree of the graph corresponding to the left digital image whose weight is greater than the threshold value will be removed. This deletion will form the sub-trees on the left graph and the sub-trees also on the right graph.

![Image from Left Cones](image1)

![Image from Right Cones](image2)

**Figure 4.** Paired image in simple stereo configuration

One advantage of image segmentation using a minimum spanning tree is that the larger segment boundary corresponding to a threshold value is a collection of the boundaries of the smaller segments corresponding to a smaller threshold value. This can be seen in figure 5 for the minimum spanning tree of the graph corresponding to the left image (left cones) with the threshold value μ and the threshold value μ+σ. All sub-tree boundaries in graph 5.a must be part of the sub-tree boundary of the graph of figure 5.b

![Sub Tree with Threshold Value μ](image3)

![Sub Tree with Threshold Value μ+σ](image4)

**Figure 5.** Trees for different threshold
The image pair on figure 4 is presented in the trees as in Figure 6 for a threshold value. For each of the same rows in the left graph and right graph, we will look for pairs of each sub tree in the left graph with the subtree most similar to the subtree on the right graph. The method will be performed in accordance with the "Segment-Based Stereo Matching Using Belief Propagation and Self-Adapting Dissimilarity Measure" method with improvements for segmentation, and apply local window-based matching steps.

The segmentation used for the left image and the right image is the minimum spanning tree method. This method will be better because in the process of segmentation this method does not include outlier data. In addition, the segment boundary corresponding to the larger threshold value is a combination of the boundaries of the segments corresponding to the smaller threshold value.

Step of "apply local window-based matching" will be made better because determining the local window is not always square. But the local window is the intersection of the square of the local window and the segmentation created in the previous step. But we must also limit the minimum number of pixels from the intersection of the local window. So the "apply local window-based matching" step will not include outlier data.

For the same row, we will look for pairs of each sub-tree on the left graph with the subtree on the right graph using the improved "apply local window-based matching" step. The improvements made to this method are in the first and second steps of the four steps used in this method. The result of this repair method will be better than the original method.

Suppose that pixels on the 89th line and the 74th column of the left image will be searched for the pair of the pixel on the right image. The mean value of the difference in pixels intensity of the corresponding right and left rectangles along the 89th row of the right image is calculated. So the result is obtained in Figure 7, where the blue dashed line is the result of the use of square and the red line is the result of intersection of the square and the segment. Both methods result in the pixel on the 89th and the 74th row on the left image corresponding to the pixel of the right image on the 89th row and 52th column. So the disparity pixel is equal to 22 i.e 76 minus 52. But the result of the use of intersection of the square and the segment is better than just using the square as shown in Figure 7. It can be seen that the minimum value
of the average corresponding pixel intensity difference is equal to 23.16 for square and 5.29 for intersection of the square and the segment. So the disparity calculation method built with minimum spanning tree approach is one alternative that can be used to calculate disparity for image pair with simple stereo configuration.

![Result of: square intersection of the square and the segment](image)

**Figure 7.** Curve of average value of difference of pixel intensity of left and right image

4. Conclusions

The segmentation used for the left image and the right image is the minimum spanning tree method. This method will be better because in the process of this segmentation method does not include outlier data. Step of "apply local window-based matching" will be made better because determining the local window is not always square. But the local window is the intersection of the square of the local window and the segmentation created in the previous step. So the "apply local window-based matching" step will not include outlier data.

For the same row, we will look for pairs of each sub-tree on the left graph with the subtree on the right graph using the improved "apply local window-based matching" step. The improvements made to this method are in the first and second steps of the four steps used in this method. The result of this repair method will be better than the original method. So the disparity calculation method built with minimum spanning tree approach is one alternative that can be used to calculate disparity for digital image pair with simple stereo configuration.

References

[1] Kallasi, F., D.L. Rizzini, F. Oleari, and J. Aleotti. (2015). Computer Vision in Underwater Environments: a Multiscale Graph Segmentation Approach. *Proceeding of the OCEANS 2015-Genova*. IEEE. p. 1-6.

[2] Manik, E., 2010, “Pengaruh Bandwidth Terhadap Segmentasi Citra Digital Dengan Menggunakan Mean Shift,” VISI, 18 (1), pp. 43-49.

[3] Peter, S.J. (2010). Minimum Spanning Tree-based Structural Similarity Clustering for Image Mining with Local Region Outliers. *IJCA* Vol. 8(6), p. 0975-8887.

[4] Manik, E., S. Suwilo, Tulus, and O. S. S. Sitompul. (2017). On the 5-Local Profiles of Trees. *IOP Conf. Series: Materials Science and Engineering* 300.

[5] Klaus, K., M. Sormann and K. Karner. (2006). Segment-based Stereo Matching Using Belief Propagation and a Self-adapting Dissimilarity Measure. *Proceedings of the IEEE International Conference on Computer Vision*. p. 15-18.

[6] Yang, Q., L. Wang, R. Yang, H. Stewenius, and D. Nister. 2006. Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. p. 492-504.

[7] Yang, Q., R. Yang, J. Davis, and D. Nistér. (2007). Spatial-depth super resolution for range images. *Proceedings of the CVPR*. p. 1-8.
[8] Sun, J., Y. Li, S.B. Kang, and H.-Y. Shum. (2005). Symmetric Stereo Matching for Occlusion Handling. *Proceedings of the CVPR*.

[9] Zitnick, L. and S.B. Kang. (2007). Stereo for Image-Based Rendering Using Image over-Segmentation. *International Journal of Computer Vision*, 75(1): 49-65.

[10] Mattoccia, S., F. Tombari, and L. Di Stefano. (2007). Stereo vision enabling precise border localization within a scanline optimization framework. *Proceedings of the Asian Conference on Computer Vision*. Springer Berlin Heidelberg. p. 517-527.

[11] Yoon, K.-J., and I. S. Kweon. (2007). Stereo Matching with the Distinctive Similarity Measure. *Proceedings of the ICCV*. p. 1-7.

[12] Bleyer M., and M. Gelautz. (2004). A Layered Stereo Algorithm Using Image Segmentation and Global Visibility Constraints. In: *Image Processing of ICIP’04*. p. 2997-3000.

[13] Hirschmüller, H. (2006). Stereo Vision in Structured Environments by Consistent Semi-Global Matching. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)* (Vol. 2, p. 2386-2393).

[14] Manik, E., S. Suwilo, Tulus, and O. S. S. Sitompul. (2017). The Uniqueness of Image Segmentation Generated by Different Minimum Spanning Tree. *Global Journal of Pure and Applied Mathematics*, 13(7), pp. 2975–298