

Object based contour detection by using Graph-cut on Stereo image

Taehoon Kang¹, Jaeseung Yu¹, Jangseok Oh¹, Yunhwan Seol¹, Kwanghee Choi¹, Mingi Kim^{1*}

¹ Korea Univ., Department of Electronics and Information Engineering,
Anam-Dong, Seongbuk-Gu, Seoul 136-701, Korea
{dreamth, yu1227, dueleldi, yhseol, wastedtime, mgkim}@korea.ac.kr

Abstract

In the last few years, computer vision and image processing techniques have been developed to solve many problems. One of them, graph cut method is powerful optimization technique for minimizing energy function. And as you know, many edge detectors are already advanced. Edge detector is widely used in computer vision to find object boundaries in images. But traditional edge detectors detect all edges that we don't want. Because we want to detect object contour, we propose that graph cut method and edge detector have to combine each other. In this paper, we describe method minimizing energy function via graph cuts and traditional edge detectors and show our result images.

1. Introduction

One of the most important techniques in computer vision is to extract object contour in images. In recent, energy minimization techniques based on graph cuts are used for applications such as image segmentation, restoration, stereo, object recognition and some others [1][2][3][4][5]. This methods gives very strong experimental results. To solve our problem, we used standard stereo geometry such as shown in Fig.1, where a pair of point $M_1 \leftrightarrow M_2$ called corresponding point.

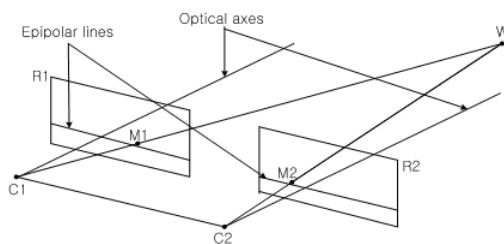


Figure 1. Standard stereo model :we can get the standard stereo images by parallel camera model

By solving corresponding problem, we acquire a new version of image that is called disparity map. The main idea of this paper is to find object contour by segmenting the disparity map rather than segmenting images. We can get a disparity map as using minimizing energy function (shown in eq(1)) via graph cuts. We can think that disparity map can be segment image similarly. The segmented image has the least information to recognize objects.

In addition, edge detectors are one of the fundamental image processing method [6][7][8][9]. They can be designed the enhance features in real images. Most

applications, such as image segmentation, region separation, and recognition use edge detection as preprocessing stage for feature extraction [6][8]. The edge means a sudden change in image intensity so that the edge may not guaranty the object boundary. With the segment image by graph cuts, the Canny edge detector [7] classifies a pixel as an edge if the gradient magnitude of pixel is larger than those of pixels at both its sides in the direction of maximum intensity change. In general images, edge detector search all edges that we don't want. Because of this reason, we will combine graph cuts method and the Canny edge detector to find object contour that we want.

The most important algorithm described in this paper is based on energy minimization because we have to find out a correct segmented image [1][2]. In this method, our focus is to find global minimum of energy function where it is NP-hard to compute global minimum [2]. We then describe two algorithms based on graph cuts, namely expansion moves and swap moves [2][4] computes efficiently a local minimum of energy function. We present only expansion moves that used in our algorithms for energy minimization.

This paper is organized as follows. In section 2 we begin with overview of energy minimization algorithms via graph cuts. We review some different edge detector and the canny edge detector that technique enhanced is introduced in section 3. We show the experimental result images in section 4 and discuss conclusion in section 5.

2. Energy minimization by graph cuts

2.1 Energy function

One of many tasks of energy minimization in computer vision is to assign a label to every pixel. We define this problem as pixel labeling problems. For motion or stereo, the labels are disparities, while for image restoration they are intensities.

In pixel labeling problems the goal is to find a labeling f that assigns each pixel a label $p \in P$ a label $f_p \in L$. This input is a set of pixels P and a set of labels L . We provide an reference Fig.2 [1][2].

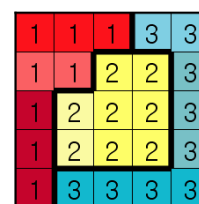


Figure 2. This figure shows an example of an different color image labeling: The color is a set of pixels P . L is

label assingning $L_p \in \{1,2,3\}$. This labeling means segment image that we want[2][4].

We can formulate a standard form of the energy function to solve this pixel labeling problem[1][2][3][4][5]. The form of energy function is

$$E(f) = E_{data}(f) + E_{smooth}(f)$$

$E_{data}(f)$ and $E_{smooth}(f)$ can be also represented as follows

$$E(f) = \sum_{p \in P} D_p(f_p) + \sum_{p,q \in N} V_{p,q}(f_p, f_q) \quad (1)$$

This function includes a variety of different concept such as first-order Markov Random Fields. $D_p(f_p)$ is a data penalty function and it means that how well assign a label f_p to a pixel p given the observed data. $V_{p,q}(f_p, f_q)$ penalizing between neighboring pixels $p, q \in N$ is spatial smoothness term. N is the set of neighboring pixels of left image. V can be separated from metric and semi-metric. For any labels $\alpha, \beta, \gamma \in L$

$$V(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta \quad (2)$$

$$V(\alpha, \beta) = V(\beta, \alpha) \quad (3)$$

$$V(\alpha, \beta) \leq V(\alpha, \gamma) + V(\gamma, \beta) \quad (4)$$

If V satisfy (2),(3) and (4), V is metric on the space of Label L . Also V is called a semi-metric if it satisfies only (2),(3)[2]. We can specify this equation by potts model as follows.

$$E(f) = \sum_{p \in L} D_p(f_p) + \sum_{p,q \in N} V_{p,q} \cdot T(f_p \neq f_q)$$

Here $T(\cdot)$ is 1 if its argument is true and 0 otherwise[1][3]. Our goal is to find labeling f for local minimum to minimize energy function(1). We can also use α -expansion move algorithm to generate labeling f . This is a strong move allowing many pixels to change their labels to α simultaneously. Given a label α , α -expansion is a move from a partition P (labeling f) to a new partition P' if $P_\alpha \subset P_\alpha$ and $P_l \subset P_l$ for any label $l \neq \alpha$ [2].

In pixel labeling problem, We have to find f that minimize energy function. To solve this problem, We can use graph cut method.

2.2 Graph cut

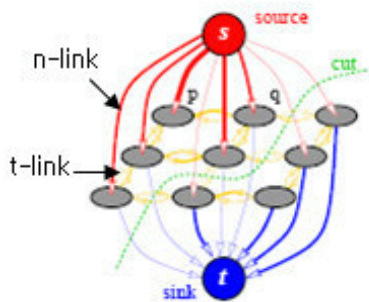


Figure 3. A cut on graph \mathcal{G} . Red and blue line is n -link. Yellow line is t -link. Edge costs are reflected by line's

thickness.

As mentioned above, graph cut is used for finding f to minimize energy function. A weighted graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ have two special nodes that are called terminals. \mathcal{V} is the set of vertices and \mathcal{E} is the set of edge. The terminals are $\{s, t\}$ that means source and sink. In this paper, the node means pixels. In Fig.3 we show a example of graph cut having two terminals (source and sink)[1].

In the graph, edge has two types called n -link and t -link. All such edges are assigned nonnegative weight and cost. n -links connect pairs of neighboring pixels in the graph. So the cost of n -link is a penalty for discontinuity and is derived from V (smoothness term) in (1). t -link connect pixels with terminals or labels the cost of t -link is a penalty for assigning the corresponding label to the pixel and is derived from D (data term) in (1).

We have to find a minimum cut that has the minimum cost among all cuts in the graph. In Fig.3 green line is a cut that means minimum cost. The cut $C = \{S, T\}$ is a partition of the vertices in \mathcal{V} into two disjoint set S and T such that $s \in S$ and $t \in T$. The minimum cut problem is derived from solving by finding a maximum flow from the source s to the sink t [1][5].

We briefly summarized to minimize Energy function by graph cut. This theory is a powerful method to find disparity map or segment image. The image is used for extracting object contour by combining edge detector.

3. Edge detector

3.1 Introduction to edge detectors

Algorithms to combine edge detection and image segmentation have been studied for many years. The edge detection is by far the most common approach for meaningful discontinuities in intensity value. The first- or second- derivative that means the gradient is useful for detecting edge on image. The edge detector performs a simple 2-D spatial gradient measurement by using vector function of 2-D gradient of image intensities[6][7].

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (5)$$

We can present a simple equation of the magnitude of the gradient[9].

$$G = \sqrt{G_x^2 + G_y^2}$$

The angle of orientation of the edge that means the direction of the gradient in given by

$$\theta = \tan^{-1} \left(\frac{G_y}{G_x} \right)$$

There are many detectors for detecting edge such as Sobel, Prewitt and Roberts detector. In Fig.4 we can detect edges through the use masks of the first- derivatives they implement[6].

Prewitt mask					
-1	-1	-1	-1	0	1
0	0	0	-1	0	1
1	1	1	-1	0	1
G_x			G_y		
Sobel mask					
-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1
G_x			G_y		
Roberts mask					
-1	0	0	-1		
0	1	1	0		
G_x		G_y			

Figure 4. Masks of some edge detector. G_x and G_y are the first derivatives of gradient vector.

3.2 Canny edge detector

The Canny edge detector developed by John Canny [7] is the optimal detector used around the world. This edge detector is based on detecting at the zero-crossing of the second directional derivative of the smoothed image. The procedure of the Canny edge detector was as follows [6]:

1. We can use an appropriate 2-D Gaussian filter to make smoothed image.
2. In smoothed image we have to compute the gradient direction and magnitude at each point. An edge point is defined that strength of the point is locally maximum in the direction of the gradient.
3. Perform non-maximal suppression. The edge point determined above rise to ridges in the gradient magnitude image. Tracking the top of the ridges and set to zero all pixels that are not on the ridge top to make a thin line in the output image.
4. The ridge pixels are thresholded by using $T1$ and $T2$ with $T1 < T2$. If Ridge pixels have greater values than $T2$, it is called 'strong'. If Ridge pixels have values between $T1$ and $T2$, it is called 'weak'.

As procedure above, Canny edge detector introduced threshold values eliminate insignificant edges. It produces an output image as tracking discontinuities of intensity. In equation (5) we can think that the segment image by graph cut is $f(x, y)$. The application of this technique is applied by our object contour detecting.

4. Results

To make an experiment of our method the stereo image pairs acquiring by using graph cuts algorithm are used for extracting object contour efficiently. According to algorithm minimizing energy function by graph cuts, we can get a output image (segment image) shown in Fig.5.

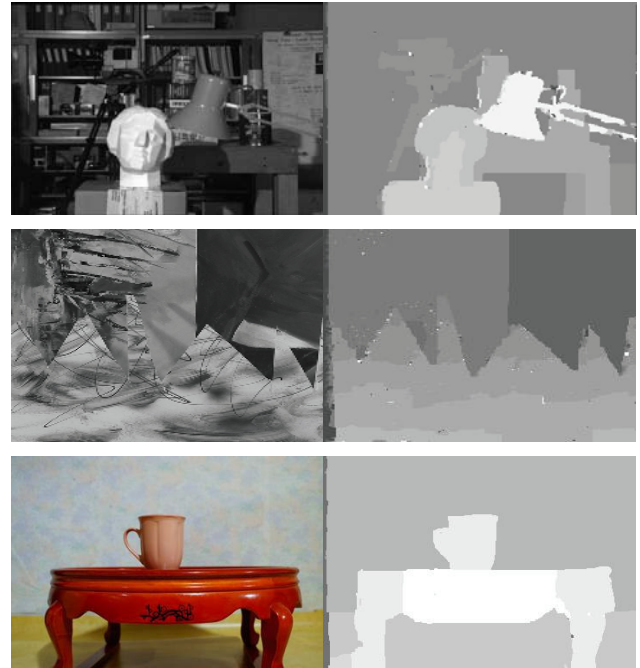


Figure 5. Output images by graph cuts :Left images are input images of stereo model. Right images are output images by graph cuts.

In Fig.5 the first head images from the University of Tsukuba have a size of 384x288. The second sawtooth images are 'synthetic' image which was taken from the Middlebury Stereo Vision Page. The size of this image is 217x190. The third cup image made by us has a size of 361x240. With each segment images in Fig.5 we applied some edge detector of section 3. The code of this algorithm is available on the web from the homepage of Kolmogorov.

<http://www.adastral.ucl.ac.uk/~vladkolm/software.html>

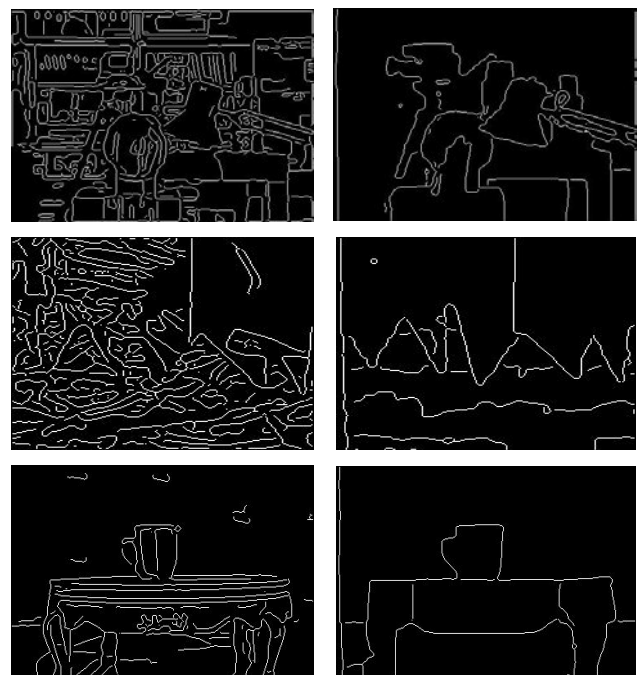


Figure 6. The results of the Canny edge detector. The left is the results of original left image of stereo pairs. The right is the result images of segmented images.

We apply the Canny edge detector to output image by graph cuts in order to find object boundary. Fig.6 shows the difference of the result in the same condition of parameter. In this Fig.6 we use the same parameter of $T1=0.05$ and $T2=0.1$. As mentioned above the $T1$, $T2$ means threshold values. The parameters was obtained by repeated experiments and can adopt the optimized results. The left images have edges of all objects. We don't need all edges when we detect or track objects. Hence, we readjust parameters, $T1$ and $T2$ to show the result of object boundary we want.

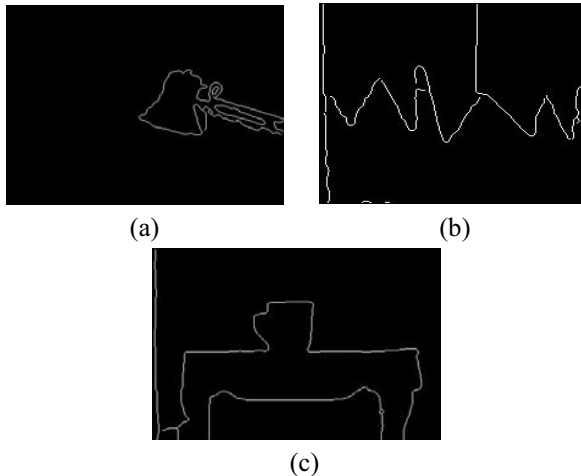


Figure 7. The developed result of the Canny edge detector. It is the more specified object contour that we want.

The main results of our experiment are illustrated in Fig.7. The Canny edge detector computes a gradient direction and magnitude for each pixel of the segmented image. We tried several times to find an appropriate boundary of object. We can compare our results of the right images of Fig.6 and the results of Fig.7. Fig.7 shows the better performance than Fig.6. The resetted values of parameter are $T1=0.2$ and $T2=0.7$ in (a), $T1=0.05$ and $T2=0.3$ in (b). In (c) the parameter values are $T1 = 0.1$ and $T2 = 0.511$. The algorithm is powerful enough to detect object contour in images as changing parameter values.

5. Conclusions

In this paper, We assume the acquired images satisfy the standard stereo geometry. However, the image we acquire may not satisfy the geometry. Thus we are preparing the image rectification stage. We have shown that combining graph cuts algorithm and edge detection are useful algorithms for detecting object contour. Our approach is based on the constraint that the images are standard stereo. We suggest that the use of the Canny

edge detector performs good results for finding object boundary. Our results show the general boundary of the objects. However, the contour extracted by our segmented image of Fig.6 is not matched with the exact boundary. In addition, our final goal is to extract boundaries of each object we want to find. For instance, if we want to find only the boundary of camera in Tsukuba image, our algorithm is limited. To improve the more accurate boundary, our future works will use snake algorithms of active contour model[10]. In snake algorithm, we can consider the boundaries of our results as initial points. We believe that our algorithm will get a powerful result if active contour algorithm is inserted.

Acknowledgements

This research was supported by a grant from Strategic National R&D Program of Ministry of Commerce, Industry and Energy.

References

- [1] Boykov, Y., Kolmogorov, V., "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", *In IEEE Transactions on PAMI*, Vol. 26, No. 9, pp. 1124-1137, 2004
- [2] Boykov, Y., Veksler, O.Zabih, R., "Fast Approximate Energy Minimization via Graph Cuts", *Proc. IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-123, 2001
- [3] V. Kolmogorov and R. Zabih, "Computing Visual Correspondence with Occlusions via Graph Cuts," *Proc. Int'l Conf. Computer Vision*, vol. 2, pp. 508-515, 2001.
- [4] Olga Veksler. "Efficient Graph-based Energy Minimization Methods in Computer Vision." PhD thesis, Cornell University, July 1999.
- [5] V. Kolmogorov and R. Zabih. "What energy functions can be minimized via graph cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147-159, Feb. 2004.
- [6] Rafael C. Gonzales, Richard E. Woods, "Digital Image Processing, second edition", Prentice Hall, 2002, chapter 10.
- [7] J.Canny, "A computational approach to edge detection,"*IEEE Trans. PAMI*, vol. 8, no. 6, pp. 679-698,1986.
- [8] D. Marr and E. Hildreth, "Theory of Edge Detection", *Proc. of the Royal Society of London B*, Vol. 207, pp. 187-217, 1980.
- [9] Ali, M. and Clausi, D.A. (2001) "Using the Canny edge detector for feature extraction and enhancement of remote sensing images", *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, Sydney, Australia, July 9 - 13
- [10] C. Xu and J. Prince. "Snakes, shapes, and gradient vector flow", *IEEE Transactions on Images Processing*, 7(3):359-369, 1998.