**International Journal of Engineering Research and Development**
**E-ISSN**: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com
Volume 7, Issue 3 (May 2013), PP. 78-81

# Super Active Visual Segmentation

Sasi Kumar A.V[1], M.S.Balamurugan[2]
1. PG Scholar, Sri Shakthi Institute of Engineering and Technology, Coimbatore, Tamil Nadu.
2. Assist.Prof./ECE, Sri Shakthi Institute of Engineering and Technology, Coimbatore.

**Abstract**: **-** Attention is an integral part of the human visual system and has been widely studied in the visual attention literature. The human eyes fixate at important locations in the scene, and every fixation point lies inside a particular region of arbitrary shape and size, which can either be an entire object or a part of it. Using that fixation point as an identification marker on the object, we propose a method to segment the object of interest by finding the "optimal" closed contour around the fixation point in the polar space, avoiding the perennial problem of scale in the Cartesian space. The proposed segmentation process is carried out in two separate steps: First, all visual cues are combined to generate the probabilistic boundary edge map of the scene; second, in this edge map, the "optimal" closed contour around a given fixation point is found. Having two separate steps also makes it possible to establish a simple feedback between the mid-level cue (regions) and the low-level visual cues (edges). In fact, we propose a segmentation refinement process based on such a feedback process. Finally, our experiments show the promise of the proposed method as an automatic segmentation framework for a general purpose visual system.

**Keywords: -** Fixation-based segmentation, polar space, cue integration, visual attention. Region based segmentation.

## I. INTRODUCTION

To see an object in the scene, we look at it. The human visual system makes a series of fixations at various salient locations while it observes the scene and tries to makes sense of it. See Fig. 1.1 for the fixations made by an observer while looking at a still image, as recorded by an eye tracker. The eyes cannot be static. They have to keep moving to a new location either voluntarily or involuntarily.

**Overview**

The proposed segmentation framework that takes as its input a fixation (a point location) in the scene and outputs the region containing that fixation. The fixated region is segmented in terms of the area enclosed by the "optimal" closed boundary around the fixation using the probabilistic boundary edge map of the scene (or image). The probabilistic boundary edge map, which is generated by using all available visual cues, contains the probability of an edge pixel being at an object (or depth) boundary. The separation of the cue handling from the actual segmentation step is an important contribution of our work, because it makes segmentation of a region independent of types of visual cues that are used to generate the probabilistic boundary edge map. The proposed segmentation framework is a two step process: first, the probabilistic boundary edge map of the image is generated using all available low level cues (section 3.2); second, the probabilistic edge map is transformed into the polar space with the fixation as the pole (section 3.3), and the path through this polar probabilistic edge map (the blue line in Fig.3.4g) that "optimally" splits the map into two parts (Fig.3.4f) is found. This path maps back to a closed contour around the fixation point. The pixels on the left side of the path in the polar space correspond to the inside of the region enclosed by the contour in the Cartesian space, and those on the right side correspond to the outside of that region. So, finding the optimal path in the polar probabilistic edge map is a binary labeling problem and graph cut is used to find this globally optimal solution to this binary problem (section 3.4)

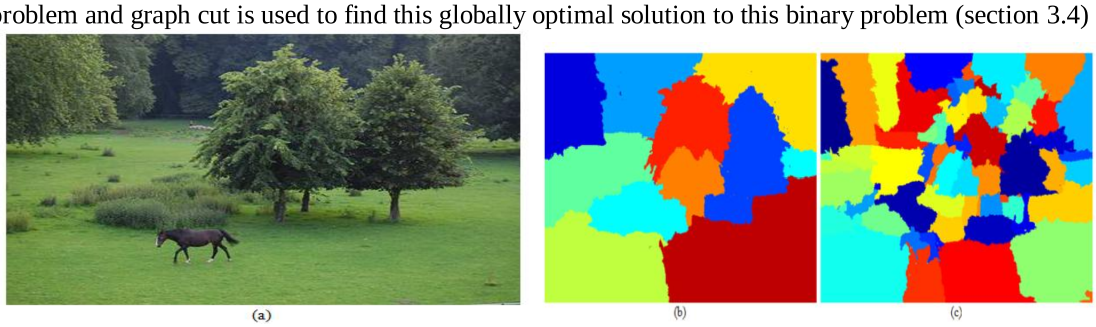


Figure 1.2: Segmentation of a natural scene in (a) using the Normalized Cut algorithm [78] for two different values of its input parameter (the expected number of regions) 10 and 60 are shown in (b) and (c) respectively.

The main contributions of the work described in this thesis are as follows;

• The proposed automatic method to segment an object (or region) given a fixation on that object (or region) in the scene/image. Segmenting the region containing a given fixation point is a well-posed binary labeling problem in the polar space, generated by transforming the probabilistic boundary edge map from the Cartesian space to the polar space with the fixation point as pole. In the transformed polar edge map, lengths of the possible closed contours around the fixation point are normalized (section 3.1) and thus the segmentation results are not affected by the scale of the fixated region. The proposed framework does not depend upon any user input to output the optimal segmentation of the fixated region.

• Since we carry out segmentation in two separate steps, it provides an easy way to incorporate feedback from the current segmentation output to influence the segmentation result for the next fixation by just changing the probabilities of the edge pixels in the probabilistic boundary edge map (see chapter 7 for how it is used to generate a multi-fixation framework). Also, using noisy motion and stereo cues to only modify the boundary probabilities of the static monocular edges provides better localization of the region boundaries while tracing actual depth boundaries around any fixation point in the scene.

## II. SEGMENTING A FIXATION REGION

Segmenting a fixated region is equivalent to finding the "optimal" closed contour around the fixation point. This closed contour should be a connected set of boundary edge pixels (or fragments) in the edge map. However, the edge map contains both types of edges, namely, boundary (or depth) and internal (or texture/intensity) edges. In order to trace the boundary edge fragments through the edge map to form the contour enclosing the fixation point, it is important to be able to differentiate between the boundary edges from the non-boundary (e.g. texture and internal) edges.

**Polar space is the key**

The gradient edge map of the disc, shown in Fig. 3.1b, has two concentric circles. The big circle is the actual boundary of the disc whereas the small circle is just the internal edge on the disc. The edge map correctly assigns the boundary contour intensity as 0.78 and the internal contour 0.39 (the intensity values range from 0 to 1). The lengths of the two circles are 400 and 100 pixels. Now, the cost of tracing the boundary and the internal contour in the Cartesian space will be respectively 88 = (400 × (1 - 0.78)) and 61 = (100 × (1 - 0.39)). Clearly, the internal contour costs less and hence will be considered optimal even though the boundary contour is the brightest and should actually be the optimal contour. In fact, this problem of inherently preferring short contours over long contours has already been identified in the graph cut based approaches where the minimum cut usually prefers to take "short cut" in the image.

**Probabilistic boundary edge map by combining cues**

In this section, we carry out the first step of the segmentation process: generating the probabilistic boundary edge map using all available visual cues. There are two types of visual cues on the basis of how they are calculated: 1) static monocular cues, that come from just a single image; 2) stereo and motion cues that need more than one image to be computed.

## III. RESULTS AND SIMULATIONS

**Segmentation Accuracy**

Our dataset is a collection of 20 videos with an average length of seven frames and 50 Stereo pairs with respect to their ground-truth segmentation.

Table 4.1 shows that after adding motion or stereo cues with color and texture

For videos F-measure

With Motion 0.95 ± 0.01

Without Motion 0.62 ± 0.02

For stereo pairs

With Stereo 0.96 ± 0.02

Without Stereo 0.65 ± 0.02

**Table 4.1:** The performance of our segmentation for the videos and the stereo pairs.

**Figure 4.1:** Row 1-3: a moving camera and stationary objects. Row 4: an

Image from a stereo pair. Row 5: a moving object (car) and a stationary camera. Column 1: the original images with fixations (the green "X"). Column 2: Our segmentation results for the fixation using static monocular cues only. Column 3: Our segmentation results for the same fixation after combining motion or stereo cues with static monocular cues.

**Region merging algorithm:**

The goal of image segmentation is to partition an image into a certain number of pieces which have coherent features (color, texture, etc.) and in the mean while to group

The meaningful pieces together for the convenience of perceiving. First, regions carry on more information in describing the nature of objects. Second, the number of primitive regions is much fewer than that of pixels in an image and thus largely speeds up the region merging process.

Figure 4.2: 1.original image 2. Splitting part

By using region merging and splitting algorithm the accuracy of the paper is improved.

**Stability Analysis**

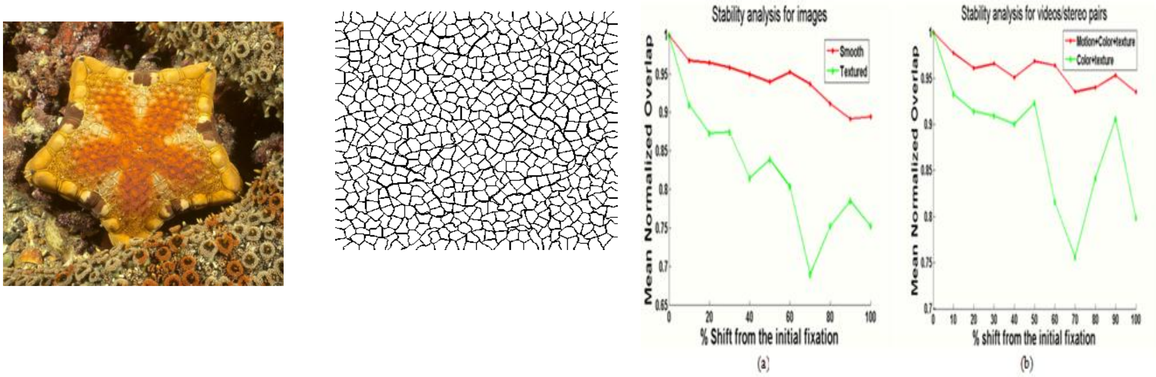


**Figure 4.3:** Stability Analysis of region segmentation with respect to the fixation Locations.

## REFERENCES

[1]. R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency Tuned Salient Region Detection," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2009.

[2]. S. Alpert, M. Galun, R. Basri, and A. Brandt, "Image Segmentation by Probabilistic Bottom-Up Aggregation and Cue Integration," Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2007.

[3]. P. Arbelaez and L. Cohen, "Constrained Image Segmentation from Hierarchical Boundaries," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 454-467, 2008.

[4]. S. Bagon, O. Boiman, and M. Irani, "What Is a Good Image Segment?A Unified Approach to Segment Extraction," Proc. 10th European Conf. Computer Vision, pp. 30-44, 2008.

[5]. A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, "Interactive Image Segmentation Using an Adaptive GMMRF Model," Proc. European Conf. Computer Vision, pp. 428-441, 2004.

[6]. Y.Y. Boykov and M.P. Jolly, "Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in n-d Images," Proc. Eighth IEEE Int'l Conf. Computer Vision, pp. 105-112, 2001.

[7]. Y.Y. Boykov and V. Kolmogorov, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 9, pp. 1124-1137, Sept. 2004.

[8]. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, High Accuracy Optical Flow Estimation Based on a Theory for Warping, pp. 25-36. Springer, 2004.

[9]. N.D.B. Bruce and J.K. Tsotsos, "Saliency, Attention, and Visual Search: An Information Theoretic Approach," J. Vision, vol. 9, no. 3, pp. 1-24, 2009.
[10]. M. Cerf, J. Harel, W. Einha¨user, and C. Koch, "Predicting Human Gaze Using Low-Level Saliency Combined with Face Detection," Proc. Neural Information Processing Systems, 2008.
[11]. E. Craft, H. Schu¨ tze, E. Niebur, and R. von der Heydt, "A Neural Model of Figure-Ground Organization," J.Neurophysiology, vol. 6, no. 97, pp. 4310-4326, 2007.
[12]. P. Dimitrov, C. Phillips, and K. Siddiqi, "Robust and Efficient Skeletal Graphs," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 417-423, 2000.