MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Segmentation of Images with Insufficient Dynamical Range

Zujun Hou, How-Lung Eng, Yue Wang and Ruijing Luo
Institute for Infocomm Research, A-star, Singapore
(zhou@i2r.a-star.edu.sg)

Abstract

In this paper a method is presented to automatically detect objects from images with insufficient contrast. Using the edge map by the Canny edge detector as a reference, the corresponding edge structures in the edge map by the zero-crossing edge detector are firstly determined. Then meaningful structures are differentiated with respect to those randomly distributed edge segments and are recovered from the zero-crossing edge map. The method is validated in a realistic vision system and compared favorably with existing methods.

1. General Instructions

As a fundamental issue in computer vision and image processing, object detection aims to find semantic objects in digital images or videos. The detection can be supervised with an object model built upon a set of training data, like the Viola-Jones object detector. What is of interest here is the unsupervised object detection, which usually starts from a basic assumption: pixels associated with the same object will share similar properties (like intensity or standard deviation in a neighborhood) and pixels associated with different objects will exhibit different properties. There are two ways to apply this assumption to object detection by emphasizing the first part or the second part of the statement. An object can be detected through detecting all pixels with similar properties. In this way, object and background can be differentiated, and a partition of the image can be derived, which is named region-based image segmentation. Alternatively, one can detect the object through locating the boundaries since image properties will change at object boundary pixels. This approach is called edge detection. There have been a large number of techniques that had been proposed since 1960s to address the issue. Earlier developments treat the pixels separately, but efforts after 1980s tend to jointly consider the spatial relationship in order to yield a smooth segmentation map. For a literature review on early works, interested readers can refer to [1].

Among the recent developments, level set segmentation and graph-based segmentation are the two popular methods that have been received a number of attentions. The level set method is basically a reformulation of the active contour model [2] in the framework of level set. In spite of the increase in dimension, this formulation does not require a parameterized representation of the tracking object in the course of curve evolution. More importantly, topology change is very easy to adapt the contour towards the boundary of objects with varying shapes. These salient features make the method particularly useful in tracking interfaces and shapes [3]. Different from most active contour models that relate the object boundaries to image gradients, Chan and Vese [4] proposed a model where the stopping term is related to a particular segmentation of the image. Another advantage of this formulation is that the initial curve can be anywhere in the image. In most of traditional level set methods, it is usually require a step to periodically re-initialize the level set function to a signed distance function throughout the process of evolution in order to maintain the evolution stable. In [5], a new variational formulation was presented to force the level set function to be close to a signed distance function without the step of re-initialization.

Another type of segmentation methods that has been of intensive interest in recent years is the graph-based method, in particular due to the efficient algorithm by Boykov and Kolmogorov [6] for computing the max-flow for computer vision related graph. The method takes image pixels as graph nodes and lines linking pairs of pixels as graph edges, thus the segmentation problem can be represented in terms of a graph. Wu and Leahy introduced a minimum graph cut method for image segmentation, but the method tends to bias towards finding small components. Shi and Malik [7] proposed a normalized cuts method to address this bias issue. In spite of the good performance as reported, the method yields an NP-hard computational problem and is pretty time consuming for real-time applications. A method running in $O(m \log m)$ for m graph edges was presented in [8], where the segmentation is based on pairwise region comparison with decision following the global properties of being not too coarse and not too find according to a particular region comparison function.

Despite the tremendous efforts in object detection, it remains a challenging issue for reliably detecting objects under a wide range of variation in scene view, in particular for those computer vision systems running in outdoor and uncontrolled environment. A typical difficulty is the insufficient dynamical range for some observed objects, which would make the object appear with low contrast and barely visible even with human observation. The insufficient dynamical range could be due to several reasons. Very often it results from imperfect lighting condition, such as directional strong lighting, non-uniform weak lighting. Poor weather (like fog) or the medium (like turbulent water or some solvent) where the object is immersed could lead to the rapid decrease on the lighting transmission from objects to image plane on the one hand, and more importantly the effect of strong lighting interference due to particle reflection on the other hand. As a result, objects with different depth could exhibit remarkably different contrasts. For these images, normally it would require a good enhancement before proceeding to the step of object detection, which is another nontrivial issue. Here we present a simple yet effective method to detect the object directly.

The rest of the paper is structured as follows. Section 2 details the object detection method, which works automatically for image with sufficient or insufficient dynamical range. Experiments as well as comparison with state-of-the-arts are presented in Section 3. Finally the paper is concluded in Section 4.

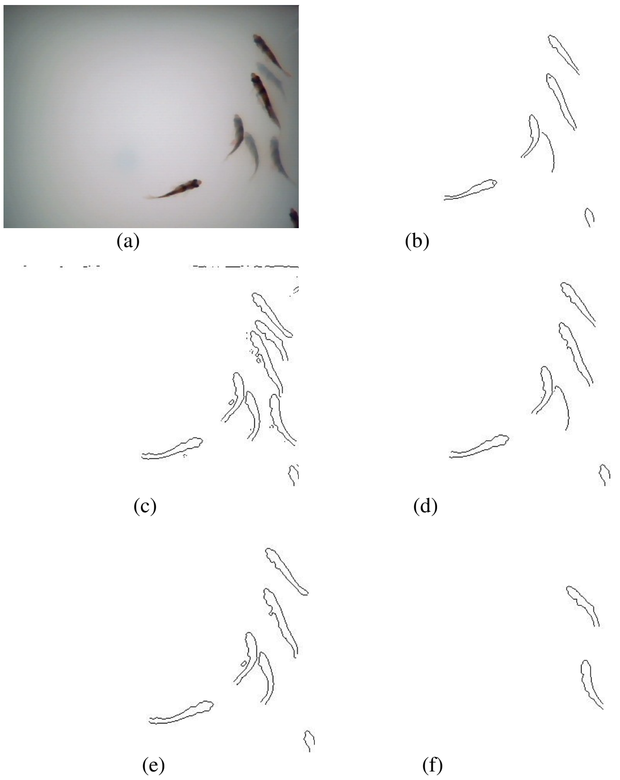

Figure 1. Illustration of object detection based on analysis of the response from the Canny and the zero-crossing edge detector, (a) original image, (b) edge image by the Canny method, (c) edge image by the zero-crossing method, (d) edge structures by zero-crossing corresponding to those by Canny, (e) edges by zero-crossing are picked up if connected to those in (d), (f) detected objects with poor contrast.

2. Method

As human vision system is more sensitive to illumination difference, it is natural to employ the edge information for differentiating objects from the background. Most edge detection methods are based on the response of image gradient, majority of which utilize the gradient magnitude information for edge structure inference. Large values of the gradient magnitude would indicate the presence of edge structure and small values would be likely corresponding to smooth surface. A typical example is the Canny edge detector [9] which detects edges as much as possible using two thresholds to pick up those faint edges connected to the stronger ones. However, the gradient magnitude is particularly sensitive to the image contrast variation. For edges in a region with good contrast, the response of gradient magnitude would be sufficiently large. But for edges in a region with poor contrast, the response of gradient magnitude would become fairly small. Even equipped with the hysteresis thresholding, the Canny detector still has difficulty in detecting the edges in low-contrast regions. For example, for the image in Fig. 1(a), the Canny detector successfully detects most of the edges corresponding to the fish located on the right side, but evidently misses the edges corresponding to a few fish on the right side where the contrast is low (as shown in Fig. 1(b)).

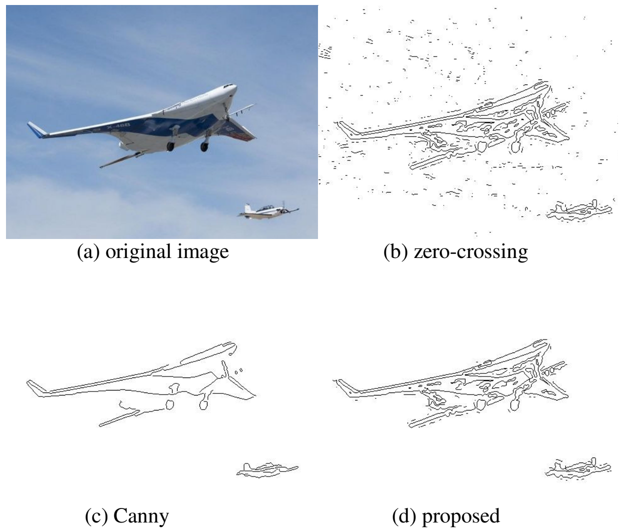

Figure 2. Object detection based on analysis of the response from the Canny and the zero-crossing edge detector, (a) original Plane image, (b) edge image by the zero-crossing method, (c) edge image by the Canny method, (d) final edge structures.

To deal with this difficulty, one can resort to another type of edge detector, which filters the image with second order difference and seeks the zeros crossing from positive to negative or vice versa along the gradient direction as the edges [10]. As the zero-crossing property is invariant to the strength of step edge, this edge detector is able to find all the edges regardless of changes in contrast. For instance, as shown in Fig. 1(c), the zero-crossing detector successfully finds most edges related to the fish (either good contrast or low contrast). However, the zero-crossing edge detector is very sensitive to random noise. To suppress the impact of noise, the implementation of zero-crossing edge detector usually incorporates a step of post-processing by requiring that the amplitude of the candidate edge response should be higher than a threshold. Nevertheless, it is still clear visible the zero-crossing result in Fig. 1(c) contains more fragmented segments. Another example is given in Fig. 2, where the original Plane image is shown in (a) and the result of zero-crossing edge detection is in (b). From Fig. 2(b), one can observe that the edge structure attributed to the plane is immersed in the sea of fragmented segments which correspond to very fine scale variations in the original image. Some of these variations are connected to meaningful objects, but majority is of little meaning for perception.

In order to extract useful information from the zero-crossing result, we firstly try to identify those structures that are common with the Canny result. To that end, a direct comparison between these two results is performed. Because both methods involve some filtering before the detection, there may be some position difference on the located edges. To tolerant this position shift, a distance threshold, T_{dist}, is introduced. If an edge pixel in the zero-crossing result appears in the T_{dist} neighborhood of an edge pixel in the Canny result, these two edge pixels will be considered as related to the same structure. Fig. 1(d) displays the initially detected common structure by this process, which is generally reasonable. Next, this result is refined by picking up the edge pixels in the zero-crossing result that are connected to the preliminarily detected common structure. For the example in Fig. 1, the refined common structure is shown in Fig. 1(e).

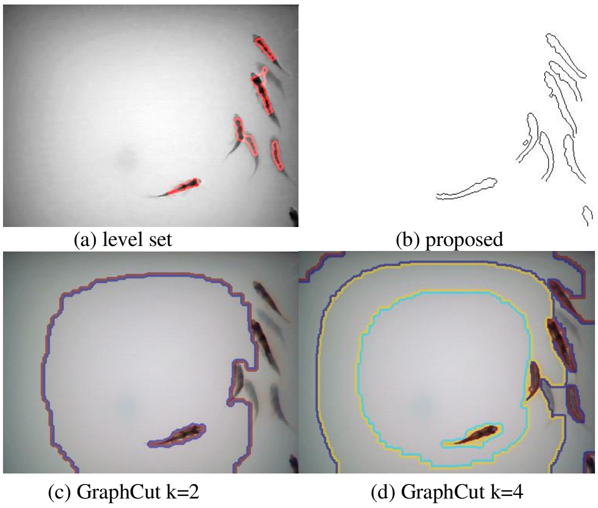

Figure 3. Object detection results for the example image in Fig. 1: (a) level set method with 1000 iterations, (b) proposed method, (c) graph cut method with 2 clusters and (d) graph cut method with 4 clusters.

After identification of the common structure, it is ready to find the possible structure detected by the zero-crossing method but missing in the Canny result. As aforementioned, the main difficulty lies in the existence of the fragmented segments due to noise or less meaningful objects. Basic idea in mining possible objects from these segments is as follows. Usually, edge segments from objects would be more coherently located, and those from noise or less meaningful objects would be more randomly distributed. Bearing this in mind, a simple size filter could help to remove those small and randomly distributed segments. For objects with insufficient dynamical range, the detected boundaries are quite likely consisted of fragmented segments. Thus, a morphological dilation could be useful to link these fragmented object segments. The selection of the size threshold is recommended to associate with the size of objects in the common structure. Note that the common structure will be under the same morphological operation before size filtering. The filtered result is depicted in Fig. 1(f) where two fish with low contrast are perfectly detected.

Before proceeding to the next section, the major steps in the proposed object detection method are summarized as follows:

(1) to apply the zero-crossing and the Canny edge detector to the input image and obtain two preliminary edge maps;
(2) to find the common edge structures;
(3) to find the meaningful edge structures missing in the Canny edge map;
(4) to yield the final edge map by combining the edge maps in (2) and (3).

3. Experiments

The method has been evaluated in a realistic vision system where a group of fish in a fish tank was monitored through a camera and the fish behavior is studied using computer vision techniques for water contamination assessment and real-time toxicity detection. A main problem in the development of the vision system is that the fish could appear less visible in the captured image due to the higher level of turbidity in the water (e.g., from rain storm). Two examples are shown in Figs. 1(a) and 4(a) respectively, where it can be seen that some fish swimming closer to the bottom of fish tank appears pretty dim.

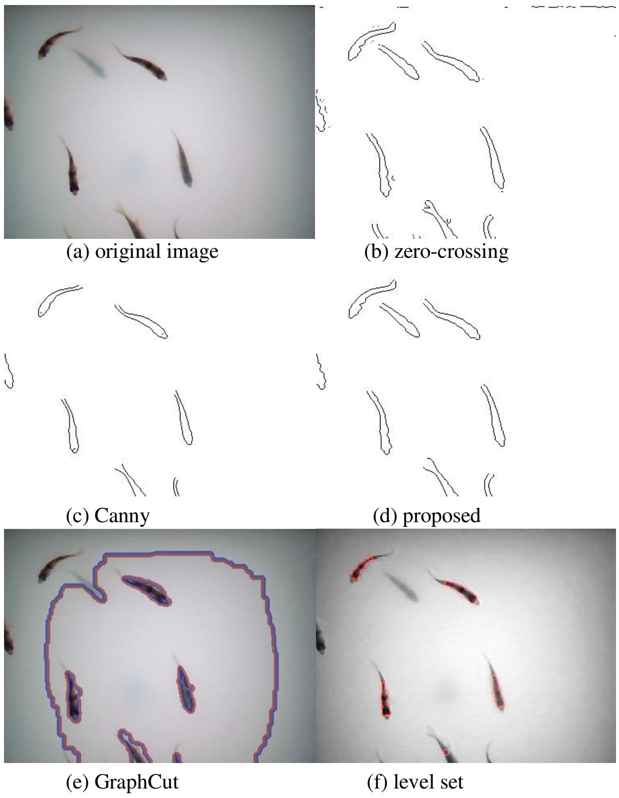

Figure 4. Object detection results for another fish example image: (a) original image, (b) the zero-crossing edge detector, (c) the Canny edge detector, (d) proposed method, (e) graph cut method with 2 clusters and (f) level set method with 1000 iterations.

For comparison with state-of-the-arts, the level set without re-initialization [5] and the graph cut method [6] are implemented. The codes of both methods can be downloaded from their websites, where the parameters are set as default.

For the example image in Fig. 1(a), the detection results are shown in Fig. 3, where (a) is by the level set method after 1000 iterations (the curve is stable after 800 iterations), (b) the proposed method, (c) and (d) the graph cut method with 2 and 4 clusters respectively. It can be seen that the level set method successfully captures most fish swimming near the surface of the water. The graph cut method detects 1 fish with 2 clusters and fails in detecting 2 fish with 4 clusters. Comparably, the proposed method almost finds the contour for every fish. Similar pattern can be observed from Fig. 4.

4. Concluding remarks

In this study we present a method to automatically detect objects from images with insufficient dynamical range. The method starts from an analysis on the edge maps derived from the zero-crossing and the Canny edge detector. Randomly distributed fragmented edge segments are removed and meaningful edge structures are recovered from the zero-crossing edge map. The method has been evaluated in a realistic vision system and it turns out that the proposed method is compared favorably with state-of-the-arts.

(a) original image

(c) enhanced image

Figure 5. An example on enhancement using the knowledge of object detection by the proposed method

It should be pointed out that it is not necessary for the proposed method to work for images with insufficient dynamical range only. Instead, it can directly be applied to good contrast images. For example, the example image shown in Fig. 2, where (d) is the edge structures detected by the proposed method. For these images, the detected extra edge structures will simply be empty.

Another point we would like to highlight is that for object detection of images with insufficient dynamical range most methods need a step of enhancement before the detection. Here we demonstrate the possibility to directly detect the image based on some prior knowledge of human vision processing and the intensity variation resistant property of the zero-crossing edge detector. Through this process, the structures in an image can be detected. In addition, these structures can be differentiated with respect to contrast. The structures detected by the Canny method usually have good contrast, and the contrast for the structures recovered from the zero-crossing edge map is generally insufficient. This type of knowledge could in turn help to design more effective algorithm for image enhancement. One example is given in Fig. 5 and the detail will be reported elsewhere.

References

[1] N. R. Pal and S. K. Pal: "A Review on Image Segmentation Techniques", Pattern Recognition, vol 26, no. 9, pp. 1277-1294, 1993.

[2] M. Kass, A. Witkin and D. Terzopoulos: "Snake: Active Contour Models", Int'l J. Computer Vision, vol 1, no. 4, pp. 321-331, 1988.

[3] S. Osher and J. A. Sethian: "Fronts Propagating with Curvature-Dependent Speed: Algorithms based on Hamilton-Jacobi Formulations", J. Computational Physics, vol. 79, no. 1, pp. 12-49, 1988.

[4] T. F. Chan and L. A. Vese: "Active Contours without Edges", IEEE Trans. Image Processing, vol. 10, no. 2, pp. 266-277, 2001.

[5] C. Li, C. Xu, C. Gui and M. D. Fox: "Level Set Evolution Without Re-Initialization: A New Variational Formulation", CVPR 2005.

[6] Y. Boykov and V. Kolmogorov: "An Experiment Comparison of Min-Cut/Max-Flow Algorithm for Energy Minimization in Vision", IEEE Trans. Patt. Anal. Mach. Intell., vol. 26, no. 9, pp. 1124-1137, 2004.

[7] J. Shi and J. Malik: "Normalized Cuts and Image Segmentation", IEEE Trans. Patt. Anal. Mach. Intell., vol. 22, no. 8, pp. 888-905, 2000.

[8] P. F. Felzenszwalb and D. P. Huttenlocher: "Efficient Graph-Based Image Segmentation", Int'l J. Computer Vision, vol. 59, no. 2, pp. 167-181, 2004.

[9] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698, 1986.

[10] R. Haralick, Digital step edges from zero crossing of second directional derivatives. IEEE Trans. on Pattern Analysis and Machine Intelligence, 6(1): 58–68, 1984.