Technical Report on Salient Object Detection and Segmentation

Chuan Sun  
University of KansasLawrence, KS  
chuansun@ku.edu

*Abstract*—This electronic document is a “live” template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. *\*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract*. (*Abstract*)

Keywords—component, formatting, style, styling, insert (iniy words)

# Introduction

## Salient Object Detection

Salient object detection is a technique to detect and highlight salient objects, as well as to suppress the background of an image by computing the saliency map. Much progress has been made in this topic recently due to its practical application in computer vision and image processing.

With the rapid development of neural networks and deep learning, the supervised learning-based methods have gained great popularity and performance. While the unsupervised methods still provide much value and advantage in terms of computational efficiency.

A prevailing algorithm in unsupervised salient object detection is based on the assumption that the background regions are typically connected to the borders of an image, so the salient objects could be determined by constructing a distance map indicating the distance between a pixel and a set of background seed pixels, which are typically set as some pixels along the image boundaries. The *Minimum Barrier Distance*(MBD) method is one of the robust methods for measuring a pixel’s distance to image boundaries, which constructs a distance map by computing the path cost function defined as



where is a path on image in which consecutive pixels are adjacent, and is the pixel value of path . The induced distance map has been quite robust to pixel value fluctuations and noises, while it could be computationally costly in practice.

A fast MBD method using raster scan was further put forward to accelerate the iteration process, which works by a raster scan pass and an inverse raster scan pass that compute the cost function defined as

. 2

where denotes the path to pixel x passing through y, are the highest pixel values on and are the lowest pixels values. FastMBD continuously updates the highest and lowest pixel values on all possible paths to a pixel through its 4-adjacent neighbors.

## Image Segmentation and Thresholding Techniques

Image segmentation is the technique that aims at partitioning an image into several homogeneous subregions, to facilitate further meaningful analysis. It has been widely adopted in areas like pattern recognition and object localization.

The thresholding method is a commonly used algorithm in image segmentation, that selects a threshold to split the pixel intensity histogram, and performs a transformation denoted as

. 3

where the constant T is a *global thresholding* parameter. Several optimizations of global thresholding include: the Otsu’s method that chooses the threshold according to between-class variance of pixel intensity level probabilities; using image smoothing mask before thresholding; involving edge detection in computing the histogram for better separateness.

Going further from global thresholding, the *variable thresholding* generates better performance and robustness by adopting moving averages. Pixels are visited through a certain direction so that local thresholds are generated based on a moving average of last visited pixels. A threshold is computed for each pixel by calculating the mean and standard deviation in a moving neighborhood according to (4) and (5).

4

5

where a and b are constants and are the std and mean of the moving neighbor. Variable thresholding achieves great performance boost for the segmentation on noisy or shaded images.

Some other segmentation methods tend to focus on other characteristic of a regions. For instance, region growing selects several initial seeds, and gradually group neighboring pixels into the region if they’re similar to the region seed.

While texture based segmentation finds the closings and openings of texture patterns in an image, and detects a segmentation boundary. Color based segmentation focuses on the color channel of interest by computing the mean and standard deviation of pixel channel values, and detect segmentation according to the pixels falling in the channel value interval. Practically texture based methods could be combined with color based methods for images where textures co-vary

# Approach

Various segmentation approaches have been attempted here to segment the salient object from the saliency maps generated by the FastMBD algorithm.

## Global Thresholding

The global thresholding method splits the image intensity histogram according to a tunable parameter. The Otsu’s method is applied here for determining the global threshold.

## Variable Thresholding

The variable thresholding excels at handling noisy images especially with shaded background. In this scenario of saliency detection, though the sample images may not present a strong noise nor shadiness, the variable thresholding method serves as a comparison benchmark in order to obtain a deeper understanding on the algorithm mechanism.

## Edge Detection with Region Filling

This is a quite different method than thresholding methods tried in this project, which utilizes the output of edge detection algorithm for further region analysis to segment the salient object. The idea came from the observations on the saliency maps together with the original images, that it is quite common for the salient object to contain some pale areas other than its strong main body sometimes. These areas still have discernible edges, and this is why the method is considered to be worth a try.

The method goes through several steps as below.

* Apply edge detection on the saliency map. Here the Canny detection is used, as shown in Fig. 1.

A picture containing woman, man, standing

Description automatically generated

Fig 1: The original saliency map compared with after canny detection with low and high threshold set to be 0.1 and 0.2.

* Since the Canny edge detector applies non-maxima suppression, the generated edges look rather thin. Thus a thickening on the edges is applied (displayed in Fig. 2) using image dilating on both x and y axis, to stretch the detected edges bolder for further process. The width of line dilation is set at a rather thick level to encapsulate all the surrounded regions we want.

A picture containing light, man

Description automatically generated

Fig 2: The saliency edges after being dilated in both x and y axis with both width set at 7.

* Now that strong bold edges are available for the salient object, we try to fill in the internal regions with the same value as its edge to make it a segment, shown in Fig. 3. An algorithm that fills these holes is applied, which recognizes the pixels that cannot be reached from the image borders.

A picture containing standing, looking, dark, man

Description automatically generated

Fig 3: Filling the internal regions encapsulated by the edges.

The effect of region filling is highly impacted by the edge dilation. If the widths are set to be rather small, there could be breaches in the edges that impede some regions to be filled, as illustrated in Fig. 4.

A picture containing light

Description automatically generated

Fig 4: The impact of line dilation on region filling. The middle one is dilated with a width of 7 and the right one is with a width of 3.

* Clear the excessive edges connected with the salient object, which are some connected edges or dots. The 8-connectivity is used to remove both adjacent and diagonal connections.
* The salient main body is quite well extracted. While there are still some separate edges scattering around. So a simple erosion is applied to remove these scattered edges and dots.
* Finally the saliency map generated by FastMBD is segmented into a salient region, and we will outline the region we generate in the original image, and compare it with the ground truth in the next chapter.

# Experimental Results

The FastMBD algorithm is first applied on all the 200 images for generating saliency maps. A comparison is made here for a Gaussian filter before applying FastMBD. Fig. 5 shows some samples of FastMBD output with and without Gaussian filter.

A blurry photo of a fire place in a dark room

Description automatically generated

Fig 5: Saliency maps generated by FastMBD. Left hand side images are without any filter; right hand side images are applied with Gaussian filter first.

It could be observed that the saliency maps with Gaussian filter contain a fade out effect for the background cues near the salient object, which is a result of the Gaussian filter that mitigates some noises.

Based on the generated saliency maps, two segmentation methods are applied here. The basic global thresholding serves as a baseline here for comparison, and the region filling on edge detection is the main method being focused and analyzed during the experiment.

Two highlighting boundaries are generated for each original image, with one for the segmentation output and the other for the ground truth, as displayed in Fig. 6.

A picture containing grass, photo, different, dog

Description automatically generated

Fig 6: Samples of segmentation outputs compared with the ground truths. Red rectangles outline the segmentation output; green rectangles outline the ground truth.

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**