FOCUS PRIOR ESTIMATION FOR SALIENT OBJECT DETECTION

Xiaoli Sun Associate Professor

Xiujun Zhang, PHD student

Shenzhen University, Shenzhen Guangdong China

xlsun@szu.edu.cn

Sept.20, 2017

Content

1. Motivation
2. Related Work
3. Our Model
4. Evaluation

Motivation

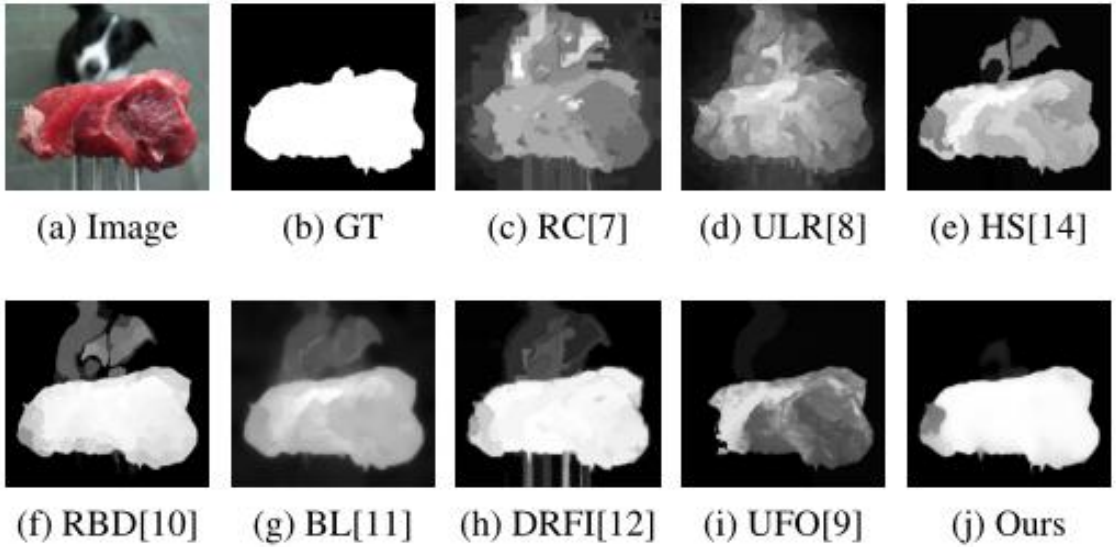
Depth of field is often used in photography.



Focus is a naturally strong indicator for the salient object detection task.

Motivation

An inspired example:



An challenging example for seven state-of-the-art methods. All results of these seven methods are not satisfied, since the focus prior is not well considered.

Focus is a strong indicator for the salient object detection, but is easily igored and not well studied.

Motivation

Our motivation:

1. Study how to estimation the focus prior for the salient object detection.

2. The focus prior should be indepent from the speicial model and easily be intergrated by the other models.

3. The model should be lightweight and run fastly.

Related Work

Defocus Map Estimation:

Elder and Zucker, TPAMI, 1998
S.J. Zhuo et al., PR, 2011
Y. Zhang et al., CVPR, 2013
Brown and Tai, ICIP, 2009
Ali Karaali et al., ICIP, 2014
.....

All are gradient based methods.

Related Work

Focusness Estimation:

P. Jiang et al., CVPR, 2013

The first to propose the focus prior for salient object detection.

Closely integrated with their model and hard to be used for the other methods.

Our Model

Dataset Construction for Evaluating:

New dataset consists of 231 images. All images have the salient objects in focus and the background in blur.

Our Model

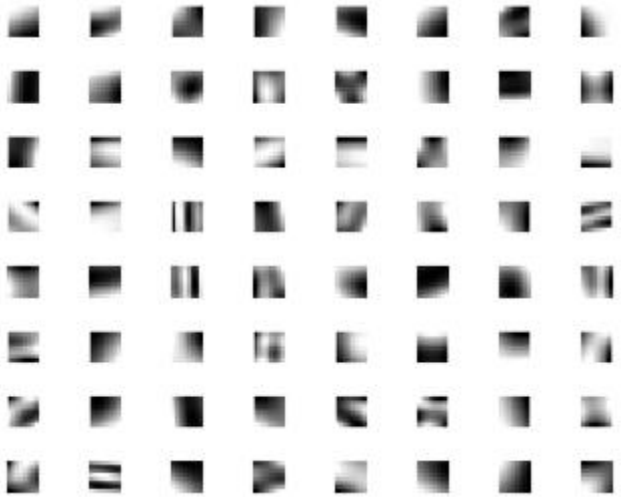
Sparse Dictionary Learning for Defocus:

$$\min_{x_i} \|x_i\|_0 \qquad (1)$$
$$s.t. \|y_i - Dx_i\|_2^2 \le \varepsilon$$

Given an input matrix $Y = \{y_1, \ldots, y_n\} \in R^{d \times n}$, each vector y_i can be represented by a set of dictionary atoms as above.

Our Model

1. A slight blurred procedure by Gaussian kernel with Sigma = 2 is imposed on images of the proposed dataset.
2. Extract overlapped image patches from them. Size of each patch is 8*8, forming an input vector y_i of d=64.
3. Our sparse dictionary D with 64 atoms is trained according to Eq.1.
4. Based on our sparse dictionary D, each image patch can be reconstructed by a few atoms together with their non-zero coefficients.



Our learned Sparse Dictionary

Our Model

Focus prior estimation map:

$$f(\mathrm{P}_i) = \|x_i\|_0 \qquad (2)$$

Here, P_i denotes the i-th patch. The focus strength for the i-th patch, denoted as $f(P_i)$, is defined as the number of non-zero elements in x_i .

Our Model

Enhancement by Objectness Analysis:

It is observed that when glancing at a scene, objects are much easier to abstract the attention of human than the flat regions.

Use the objectness proposal method in [21] to enhance the focus prior map.

$$\text{Focus}(q) = f(q) \times o(q)$$

Our Model

(a) (b) (c)

Fig. 4. An illustration of our method. (a) original images, (b) raw results of focus prior maps, (c) enhanced results by objectness analysis.

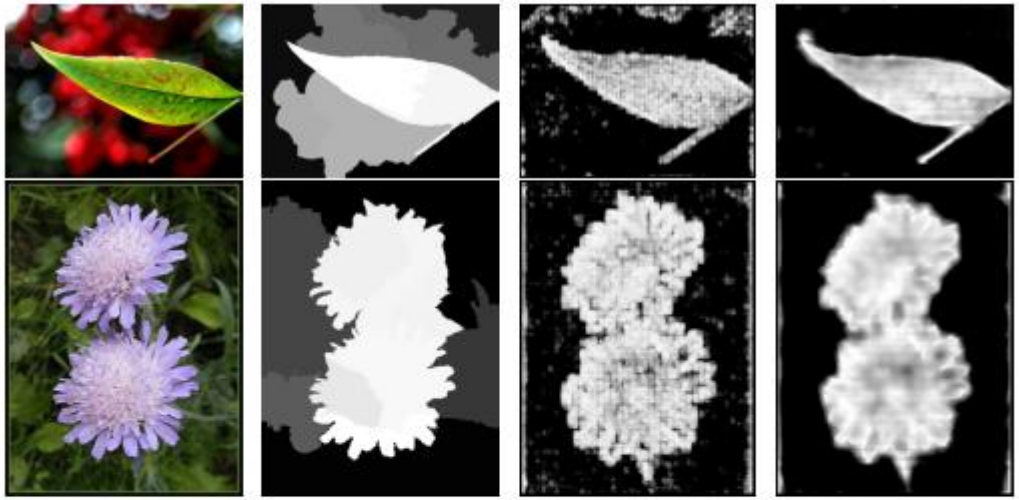
Evaluation

- Dataset
  - ASD
  - The proposed dataset

- Evaluation Category
  - Comparisons with the Other Focus Map for Salient Object Detection
  - Improvement to State-of-the-art Methods for Salient Object Detection

Evaluation

Comparisons with the Other Focus Map for Salient Object Detection:

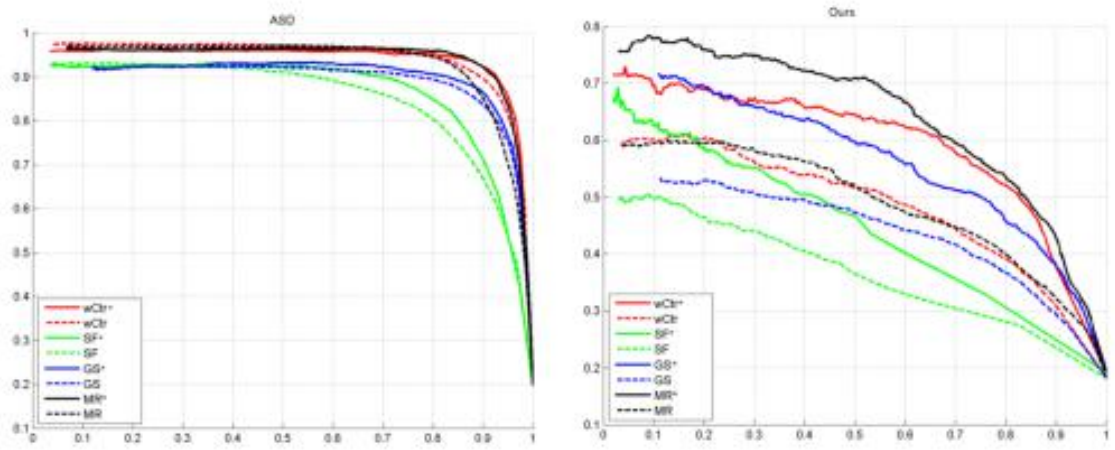


(a) (b) (c) (d)

Fig. 6. Comparison on the focus prior maps. (a) original images; (b) results of focusness estimation in [9]; (c) original focus prior maps of our proposed method; (d) enhanced results by objectness analysis.

Evaluation

Improvement to State-of-the-art Methods for Salient Object Detection:



(a) ASD dataset (b) Our dataset

Fig. 7. Precise-recall curves comparison. In both subfigures, wCtr*, SF*, GS*, MR* are the results of wCtr[10], SF[22], GS[13] and MR[23] methods integrating the proposed focus prior map, respectively.

Conclusions

1. We have proposed a novel method of estimating the focus prior map for any given images.

2.The sparse representation has been used to learn the defocus dictionary and the non-zero number of the coefficients is used to describe the focus strength of each patch.

3.Object level analysis is introduced to boost the performance, since the focus prior is meaningful to the objects in focus.

Thank You!