SALIENT OBJECT DETECTION VIA OBJECTNESS MEASURE

Sai Srivatsa R and R Venkatesh Babu
Video Analytics Lab, SERC, IISc, Bangalore, India
http://val.serc.iisc.ernet.in

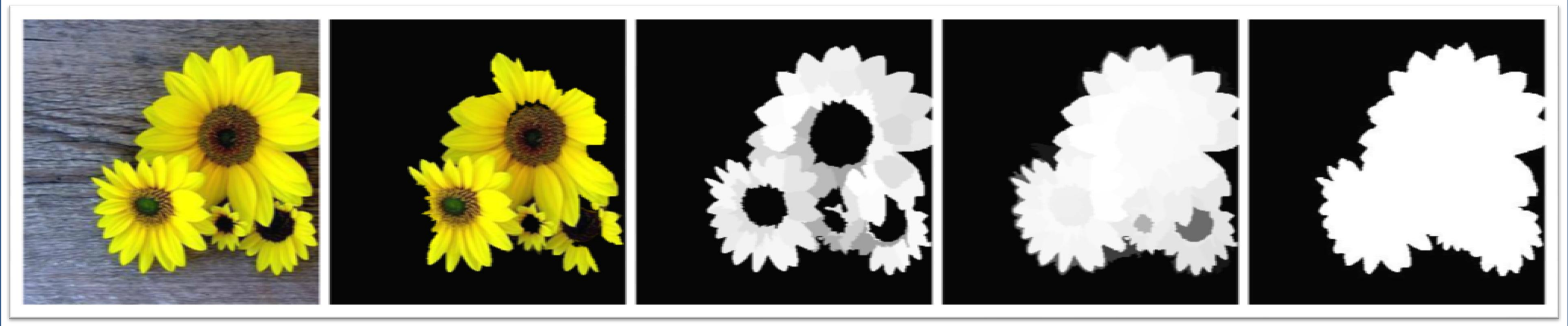
Problem Definition

- Human visual system has an innate ability to process only relevant regions of a scene while discarding the rest.
- Saliency detection algorithms aim at computational approach for detecting these conspicuous object regions in an image

Overview

- ❑ Most of the existing approaches exploit contrast prior and background prior to detect salient objects. However background prior is fragile and is prone to failure
- ❑ In our work, instead of using background cues, we estimate the foreground regions in an image by generating objectness proposals and refining them for finally obtaining smooth and accurate saliency maps
- ❑ We propose a novel saliency measure called *'foreground connectivity'* which determines how tightly a pixel or a region is connected to the estimated foreground
- ❑ Scores assigned to superpixels by the proposed foreground connectivity measure are integrated into a saliency optimization framework to obtain the final map
- ❑ We extensively evaluate the proposed approach on MSRA and CSSD datasets and demonstrate that the results obtained are better than the existing state of the art approaches

Algorithm



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

Illustration of main phases of our algorithm. (a) Input Image. (b) Thresholded Foreground Map. (c) Weights based on foreground connectivity (d) Saliency Map after Optimization (e) Ground truth

Objectness Map and Superpixels:

❑ Object proposal using modified version of BING [1] where object windows are scored by Laplacian of Gaussian (LoG) like filter applied over 8x8 normed gradient.

❑ Given object windows with scores $s_1, s_2, s_3 \ldots s_k$, compute:

Pixelwise Objectness Score
$$PixObj(p) = \sum_{i=1}^{k} s_i G_i(x, y)$$

where G_i is a Gaussian kernel placed on the center of i^{th} object proposal window.

❑ Segment image into superpixels and assign a score equal to the sum of its constituent pixel scores.

❑ Objectness map is obtained by adaptive thresholding

Foreground Connectivity :

❑ Construct graph with super-pixels as nodes where two adjacent superpixels are connected by an edge of weight equal to Euclidean distance of the mean *Lab* values.

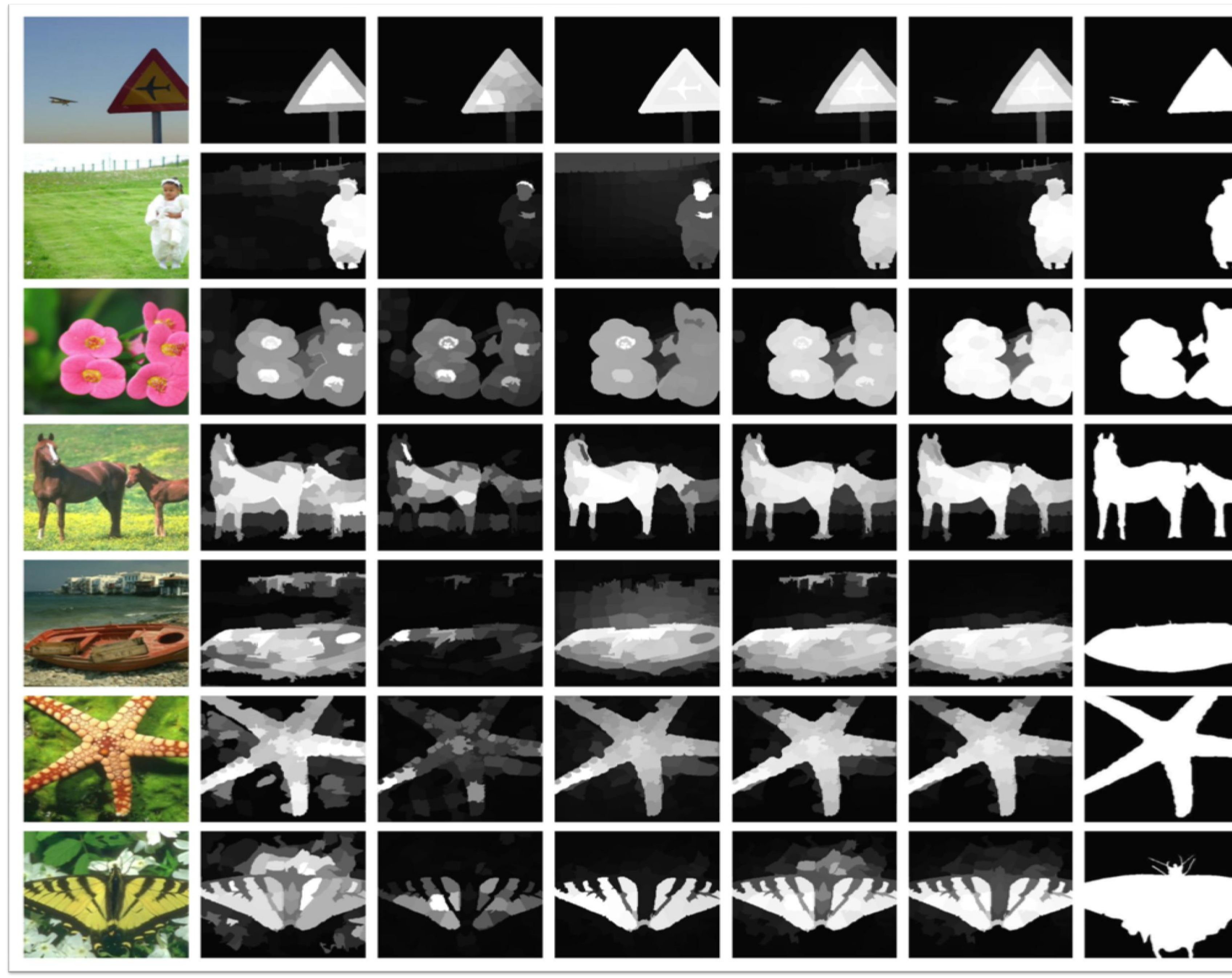
❑ Foreground connectivity of a superpixel region *R* is now defined as

$$FG(R) = \frac{\sum_{k=1}^{N} d(R, R_k).\delta(R_k)}{\sum_{k=1}^{N} d(R, R_k).(1 - \delta(R_k))}$$

where $d(R, R_k)$ is the shortest distance between two superpixels R and R_k. $\delta(R_k)$ is 1 if the superpixel is estimated as FG in thresholded map, 0 otherwise.

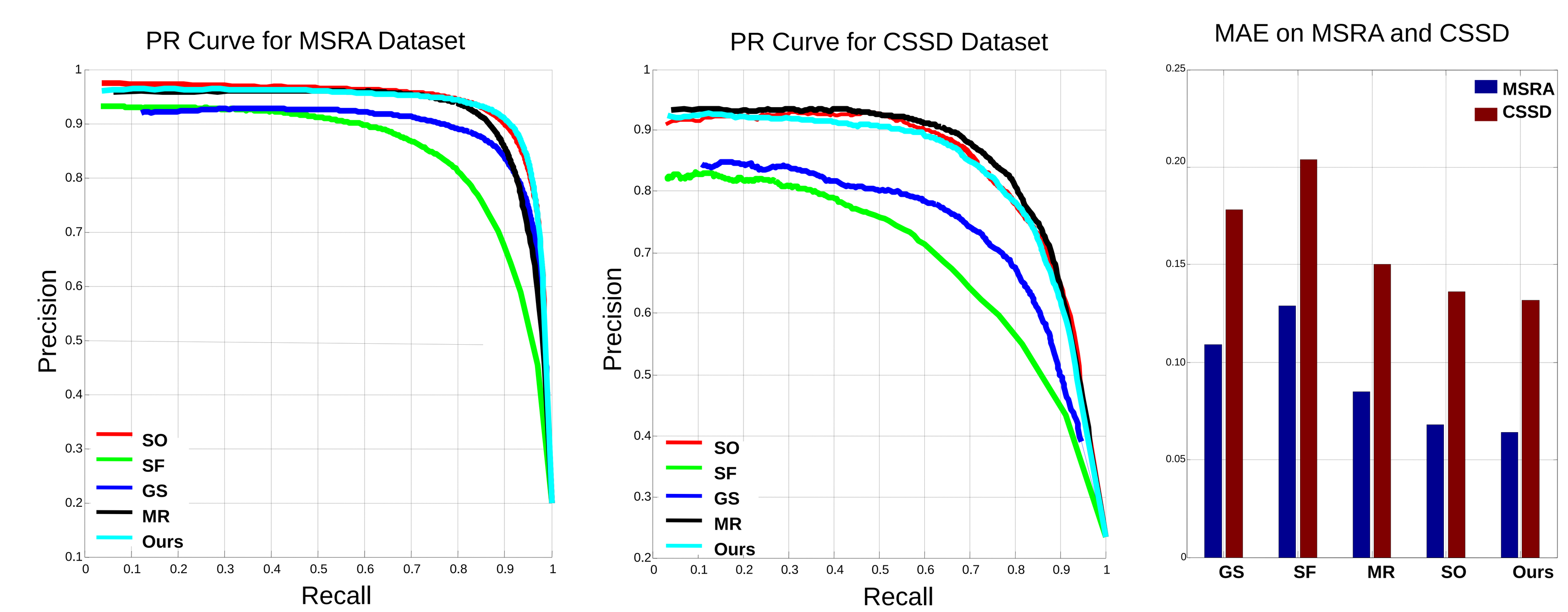
❑ Final saliency map is obtained by incorporating these foreground and background weights in the saliency optimization framework of Wei *et al.* [2]

Results and Evaluation



Input Image GS[3] SF[4] MR[5] SO[2] Ours Ground-truth

Visual comparison of various Saliency Maps



PR curves and Mean Absolute Error on MSRA Dataset and CSSD Dataset.

References

[1] Cheng et al., "BING: Binarized normed gradients for objectness estimation at 300fps," CVPR, 2014
[2] Wei et al., "Saliency optimization from robust background detection," CVPR, 2014.
[3] Zhu et al., "Geodesic saliency using background priors," ECCV, 2012.
[4] Perazzi et al., "Saliency filters: Contrast based filtering for salient region detection," CVPR, 2012
[5] Lu et al., "Saliency detection via graph-based manifold ranking," CVPR, 2013