Global Contrast based Salient Region Detection

Ming-Ming Cheng[1,4,5] Niloy J. Mitra[2] Xiaolei Huang[3] Philip H. S. Torr[4] Shi-Min Hu[1]
[1]TNList, Tsinghua University, [2]UCL/KAUST [3]Lehigh University [4]Oxford University [5]Nankai University

Abstract

Automatic estimation of salient object regions across images, without any prior assumption or knowledge of the contents of the corresponding scenes, enhances many computer vision and computer graphics applications. We introduce a regional contrast based salient object extraction algorithm, which simultaneously evaluates global contrast differences and spatial weighted coherence scores. The proposed algorithm is simple, efficient, naturally multi-scale, and produces full-resolution, high-quality saliency maps. These saliency maps are further used to initialize a novel iterative version of GrabCut for high quality salient object segmentation. We extensively evaluate our algorithm using popular benchmarks and demonstrate a variety of applications.

Sample Results

Input image

Saliency maps

Saliency cut

free! C++ http://cg.cs.tsinghua.edu.cn/people/~cmm/

Core Idea: Region Based Contrast (RC)

Image Segmentation $\sigma_s^2 \to \infty$ $\sigma_s^2 \to 0.4$

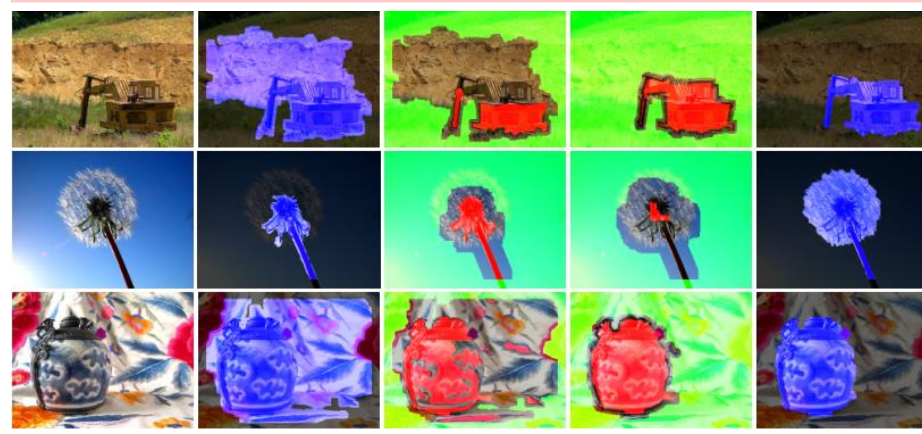
Spatial weighting

Region size

$$S(r_k) = \sum_{r_k \neq r_i} \exp\left(-\frac{D_s(r_k,r_i)}{\sigma_s^2}\right) \omega(r_i) D_r(r_k,r_i)$$

Region contrast by sparse histogram comparison.

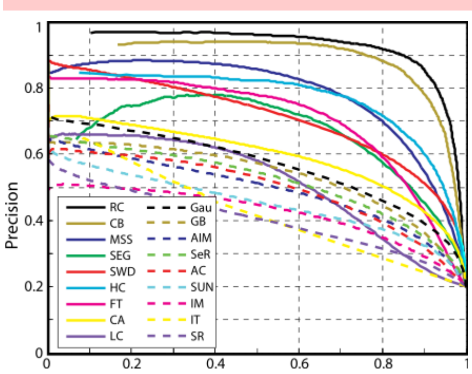
SaliencyCut: Automatic salient region extraction

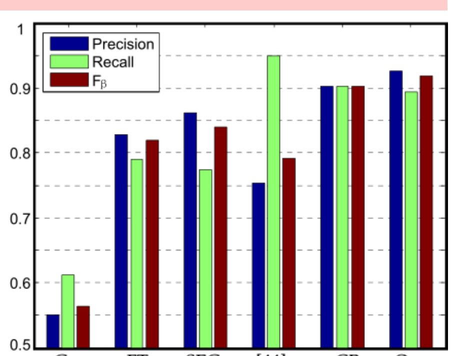


- Iterative refine: iteratively run GrabCut to refine segmentation
- Adaptive fitting: adaptively fit with newly segmented salient region

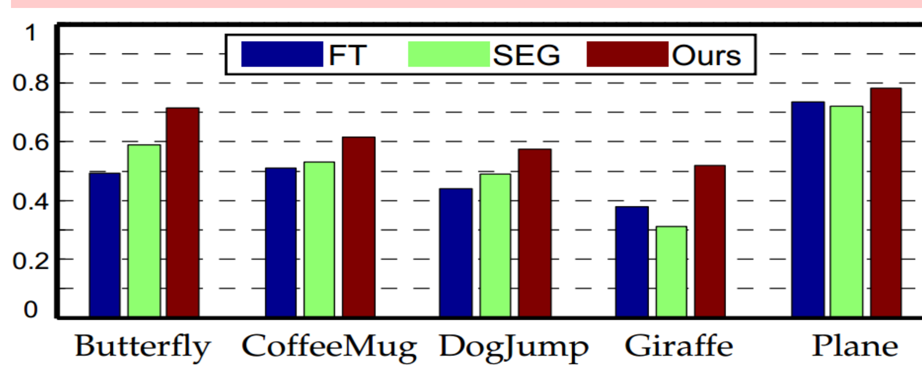
Enables automatic initialization provided by salient object detection.

Evaluation on MSRA 1000 Benchmark Dataset (Simple Images)





Challenging Benchmark: non-selected internet images



Robust Applications Design: automatically process many images + use efficient algorithms to select good results

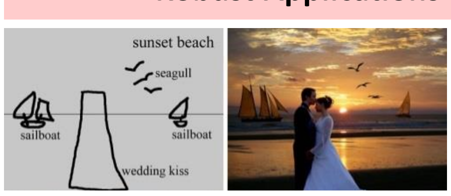


Image montage [3]

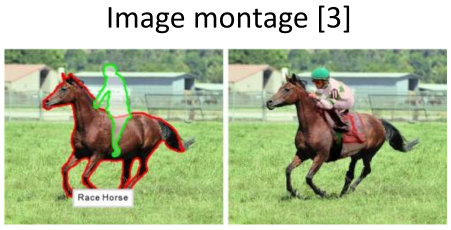
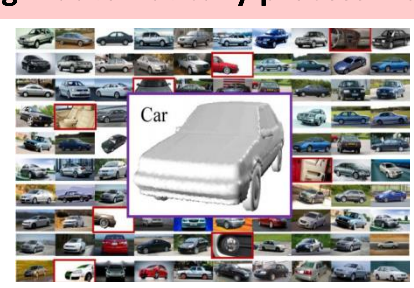


Image Manipulation [7]

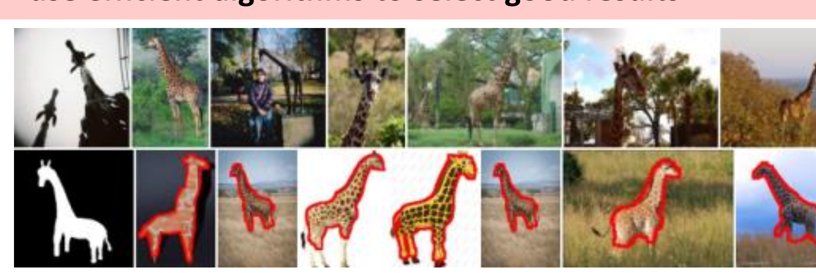
Semantic Colorization [4]



View selection [5]

Image collage [6]

Sketch Based Retrieval [1,2,8]



[1] Global Contrast based Salient Region Detection. IEEE TPAMI, 2014.
[2] SalientShape: Group Saliency in Image Collections. TVC, 2013
[3] Sketch2Photo: Internet Image Montage. SIGGRAPH Asia, 2009
[4] Semantic Colorization with Internet Images. SIGGRAPH Asia, 2011.
[5] Web-Image Driven Best Views of 3D Shapes. TVC, 2011.
[6] Arcimboldo-like Collage Using Internet Images. SIGGRAPH Asia, 2011.
[7] Data-Driven Object Manipulation in Images. Eurographics 2012.
[8] Mobile Product Search with Bag of Hash Bits and Boundary Reranking, CVPR 2012
[9] More: http://scholar.google.com/scholar?cites=9026003219213417480

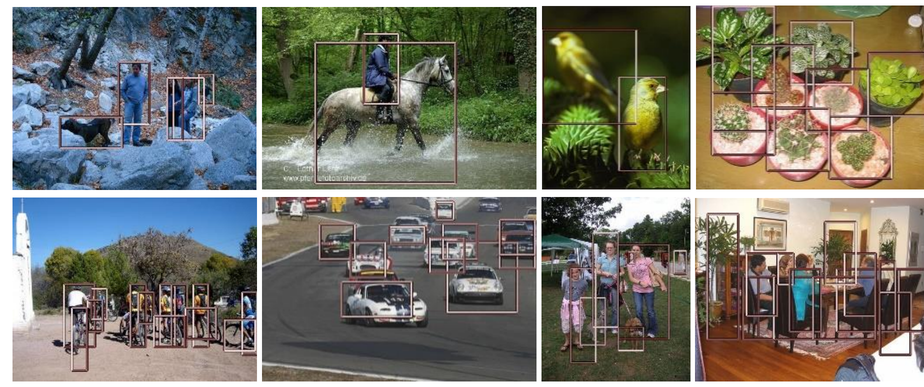
BING: Binarized Normed Gradients for Objectness Estimation at 300fps

Ming-Ming Cheng[1] Ziming Zhang[2] Wen-Yan Lin[3] Philip H. S. Torr[1]
[1]Oxford University, [2]Boston University [3]Brookes Vision Group

Abstract

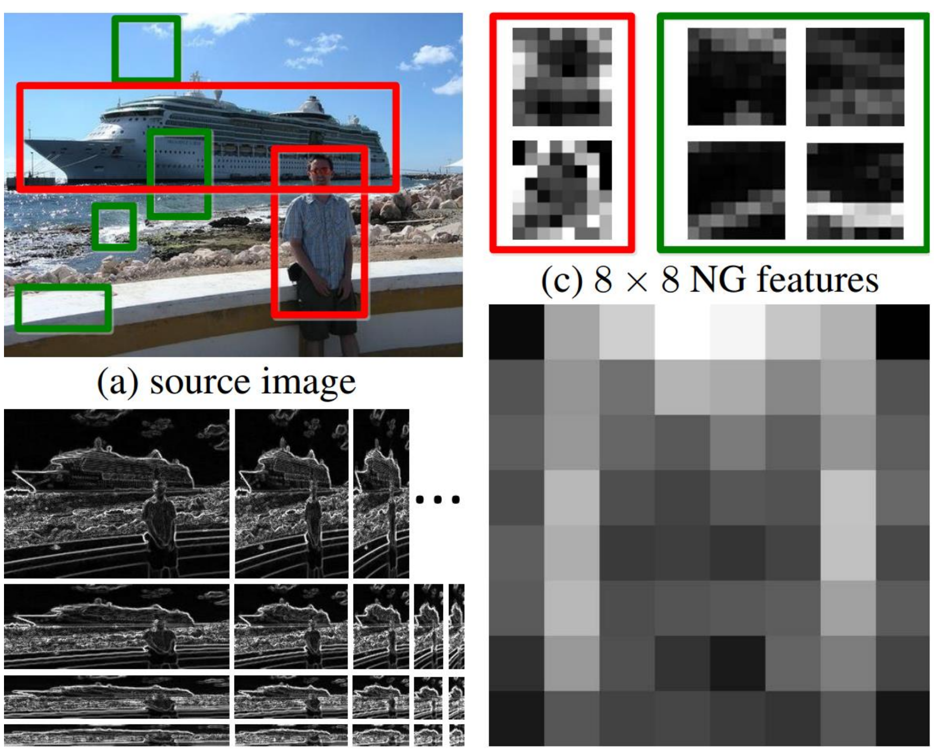
Training a generic objectness measure to produce a small set of candidate object windows, has been shown to speed up the classical sliding window object detection paradigm. We observe that generic objects with well-defined closed boundary can be discriminated by looking at the norm of gradients. Based on this observation, we propose to use a binarized normed gradients (BING) for efficient objectness estimation. Experiments on the PASCAL VOC 2007 dataset show that our method efficiently (300fps on a single laptop CPU) generates a small set of category-independent, high quality object windows, yielding 96.2% detection rate (DR) with 1,000 proposals. Increasing the numbers of proposals and color spaces for computing BING features, our performance can be further improved to 99.5% DR.

Sample Results (true positives)



free! C++ http://mmcheng.net/bing/ FAQ FORUM

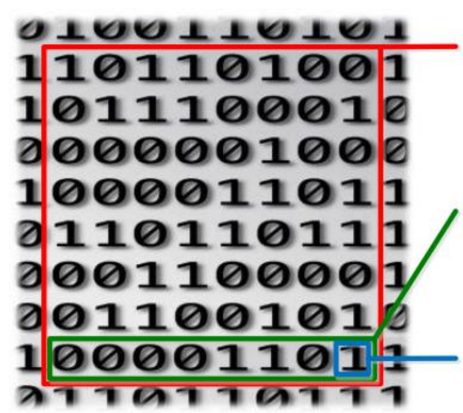
Normed gradients (NG) and objectness



(a) source image (b) normed gradients maps (c) 8 × 8 NG features (d) learned model $\mathbf{w} \in \mathbb{R}^{8\times 8}$

Although object (red) and non-object (green) windows present huge variation in the image space (a), in proper scales and aspect ratios where they correspond to a small fixed size (b), their corresponding normed gradients, i.e. a NG feature (c), share strong correlation. We learn a single 64D linear model (d) for selecting object proposals based on their NG features.

Binary normed gradients (BING)



$\mathbf{b}_{k,i,x,y} \in \{0,1\}^{8\times 8}$
shorthand: $\mathbf{b}_{x,y}$ or $\mathbf{b}_{k,l}$

$\mathbf{r}_{k,i,x,y} \in \{0,1\}^{8}$
shorthand: $\mathbf{r}_{x,y}$ or $\mathbf{r}_{k,l}$

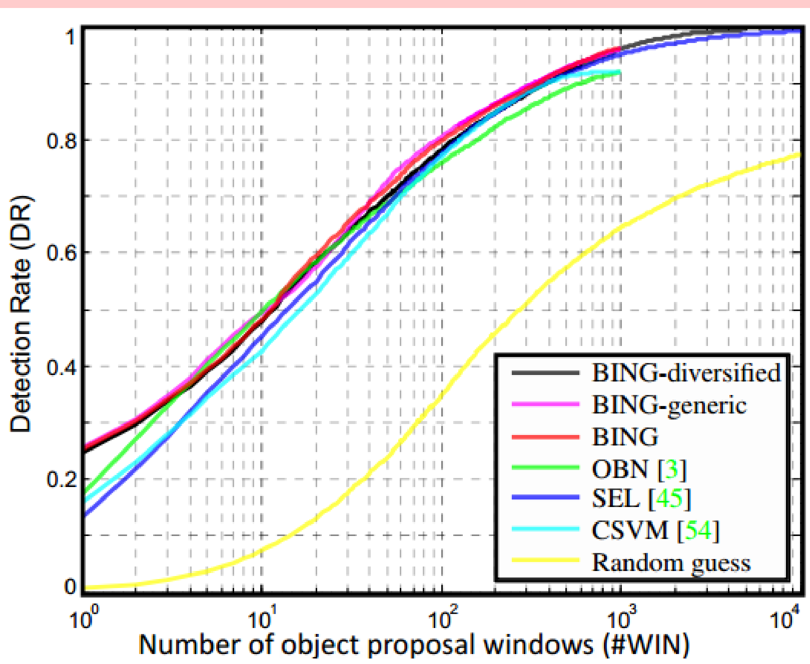
$b_{k,i,x,y} \in \{0,1\}$
shorthand: $b_{x,y}$

Illustration of variables: a BING feature $\boldsymbol{b}_{x,y}$, its last row $r_{x,y}$ and last element $b_{x,y}$. We can use a single atomic variable (int64 and byte) to represent a BING feature and its last row, enabling efficient feature computation (Alg. 2).

Algorithm 2 Get BING features for $W \times H$ positions.

Comments: see Fig. 2 for illustration of variables
Input: binary normed gradient map $b_{W\times H}$
Output: BING feature matrix $\mathbf{b}_{W\times H}$
Initialize: $\mathbf{b}_{W\times H} = 0, \mathbf{r}_{W\times H} = 0$
for each position (x,y) in scan-line order **do**
  $\mathbf{r}_{x,y} = (\mathbf{r}_{x-1,y} \ll 1) \mid b_{x,y}$
  $\mathbf{b}_{x,y} = (\mathbf{b}_{x,y-1} \ll 8) \mid \mathbf{r}_{x,y}$
end for

Experimental results on Challenging PASCAL VOC benchmark



Method	[22]	OBN [3]	CSVM [57]	SEL [48]	Our BING
Time (seconds)	89.2	3.14	1.32	11.2	**0.003**

Table 1. Average computational time on VOC2007.

	BITWISE			FLOAT		INT,BYTE	
	SHIFT	|, &	CNT	+	×	+, −	min
Gradient	0	0	0	0	0	9	2
Get BING	12	12	0	0	0	0	0
Get score	0	8	12	1	2	8	0

Table 2. Average number of atomic operations for computing objectness of each image window at different stages: calculate normed gradients, extract BING features, and get objectness score.