Image Saliency: From Intrinsic to Extrinsic Context report

Paper's Authors

Meng Wang, Janusz Konrad, Prakash Ishwar, Kevin Jing and Henry Rowley

Advisor

Dr. Maryam Abedi

Student

Mohammad Shahpouri

October
2022

Contents

List of Figures

List of Tables

List of Equations

1 Saliency estimation algorithm

They introduce two methods, global and local. Extrinsic saliency estimation is what, global method is addressed and local method is called intrinsic saliency estimation.

1.1 Extrinsic saliency estimation

1.1.1 KNN image retrieval

K-nearest neighbors of input image are retrieved by algorithm of [1] using the image distance function of [2].

1.1.2 Image warping

SIFT flow algorithm [3] is applied for warping (alignment) each image by its neighbors retrieved by KNN search.

1.1.3 Anomaly estimation

The saliency $S(x, y)$ at point (x, y) of the input image $Q(x, y)$ as the absolute error $|I_{w,i}(x, y) - Q(x, y)|$ averaged across all k warped neighbors $I_{w,i}(x, y), i = 1, \dots, k$ and normalized to the range $[0, 1]$. Mathematically:

$$\begin{aligned} E(x,y) &:= \frac{1}{k}\sum_{i=1}^{k} |I_{w,i}(x,y) - Q(x,y)| \\ S(x,y) &:= E(x,y) / \max_{x,y} E(x,y) \end{aligned} \tag{1}$$

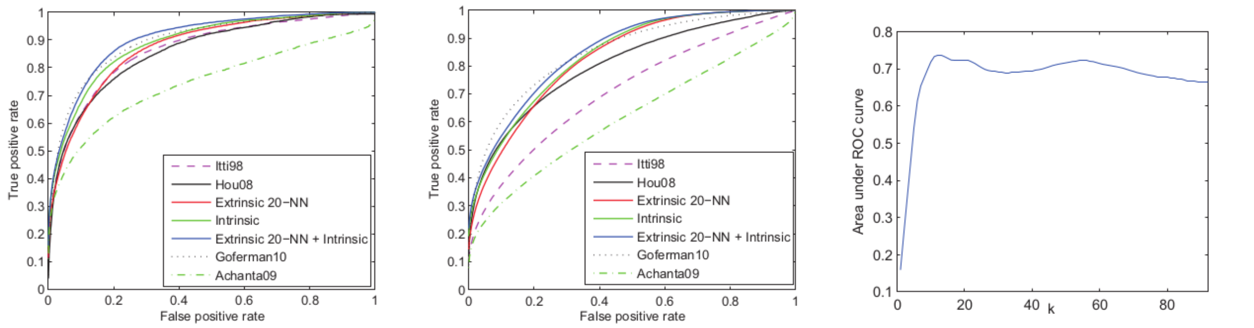
A sufficiently large and diverse dictionary will be well-approximated salient regions.

1.2 Intrinsic saliency estimation

The idea is to make 16×16 non-overlapping patches of each image. Then for each block which are tested for saliency, all blocks in dictionary, $k =$ dictionary-size, are used, warp them and calculate the difference and determine the saliency as 1.1.2 and 1.1.3. The saliency map of the input image is the composition of the saliency maps of the non-overlapping blocks.

2 Experimental results

The proposed methods are evaluated on two datasets. Berkeley segmentation database (BSDB) [4] and MSRA salient object database [5].

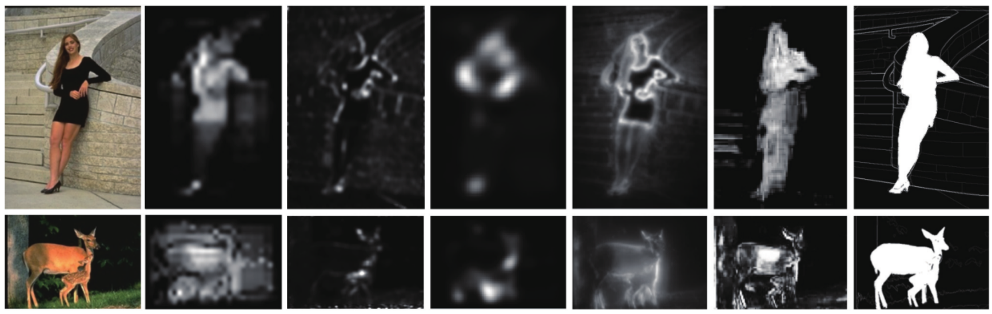


(a) BSDB (b) MSRA-B (c) Area under curve for extrinsic (BSDB)

Figure 1. Comparison of average performance of various methods on: (a) BSDB, and (b) MSRA-B data sets, and (c) area under the curve for our extrinsic method on BSDB data set (red curve in part (a)) as a function of k.

Table 1. Area under the curves from Fig. 1(a-b).

Method	BSDB	MSRA
Achanta09	0.7442	0.6743
Itti98	0.8641	0.6967
Hou08	0.8579	0.7934
Goferman10	0.8957	0.8437
Extrinsic 20-NN	0.8728	0.8289
Intrinsic	0.8881	0.8389
Extrinsic 20-NN + Intrinsic	0.9042	0.8515



(a) Input (b) Itti98 (c) Hou07 (d) Hou08 (e) Goferman *et al.* (f) Our extrinsic (g) Ground truth

Figure 2. Comparison of saliency estimation results for various methods with our extrinsic approach for ($k = 20$) on the BSDB data set.

References

[1] A. Torralba, R. Fergus, and Y. Weiss, "Small codes and large image databases for recognition," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008. 1

[2] D. Grangier and S. Bengio, "A discriminative kernel-based approach to rank images from text queries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 8, pp. 1371–1384, 2008. 1

[3] C. Liu, J. Yuen, A. Torralba, J. Sivic, and W. T. Freeman, "Sift flow: Dense correspondence across different scenes," in *Computer Vision – ECCV 2008* (D. Forsyth, P. Torr, and A. Zisserman, eds.), (Berlin, Heidelberg), pp. 28–42, Springer Berlin Heidelberg, 2008. 1

[4] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 2, pp. 416–423 vol.2, 2001. 1

[5] T. Liu, J. Sun, N.-N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007. 1