Saliency Detection via Statistical Non-redundancy

Aanchal Jain, Alexander Wong, Paul Fieguth

a27jain@uwaterloo.ca a28wong@uwaterloo.ca pfieguth@uwaterloo.ca

University of Waterloo

Vision and Image Processing Group
Systems Design Engineering
vip.uwaterloo.ca

Introduction:

- Saliency detection serves as an important first step in image processing and computer vision tasks such as object recognition and image segmentation.
- A novel algorithm has been proposed here for saliency detection in natural images.
- The proposed algorithm is based on the quantification of statistical non-redundancy of pixel neighbourhoods in the image.
- By modelling site neighbourhoods under a Gaussian process, the saliency of any arbitrary site can be characterized by the statistical non-redundancy of its site neighbourhood in a given image.
- Preliminary results using natural images show that the proposed method provides improved precision--recall characteristics over previous methods such as spectral residuals and spectral whitening.

Objective and Key Idea:

Automatic detection of saliency in images forms an important step in computer vision tasks such as image segmentation and object recognition. Traditional models of saliency detection are based on training, which makes them unsuitable for generalized tasks. It is desired to develop automatic methods for saliency detection in images, without the need for prior knowledge of image content. The proposed method uses a statistical modelling approach to quantify the non-redundant nature of saliency objects in the image without any prior knowledge of image content.

Statistical Non-redundancy in Images:

In natural images, salient objects of interest can be characterized by their statistical non-redundancy with respect to the remainder of a given image. The proposed method quantifies the statistical non-redundancy to divide the image into salient and non-salient regions.

Let N_i be a set of image pixels in the neighbourhood of site *i*.. Suppose that the neighbourhood around an arbitrary site *i* in an image can be seen as a realization of the neighbourhood around some other site *j:*

$$N_i = N_j + \eta_{ij}$$

where η_{ij} is a noise process between two site neighbourhoods following a distribution $P(\eta_{ij})$ Assuming this distribution to be an independent and identically distributed Gaussian process with zero mean and variance σ_i^2 the probability of an arbitrary site *k* in the neighbourhood N_i being a realization of the corresponding site in the neighbourhood N_j can be defined as:

$$P(N_i^k \mid N_j^k) = e^{\frac{-(N_i^k - N_j^k)^2}{\sigma^2}}$$

Therefore, taking into account all such sites *k* in the two neighbourhoods, one can quantify the statistical non-redundancy between two neighbourhoods as

$$P(N_i \mid N_j) = \prod_k P(N_i^k \mid N_j^k)$$

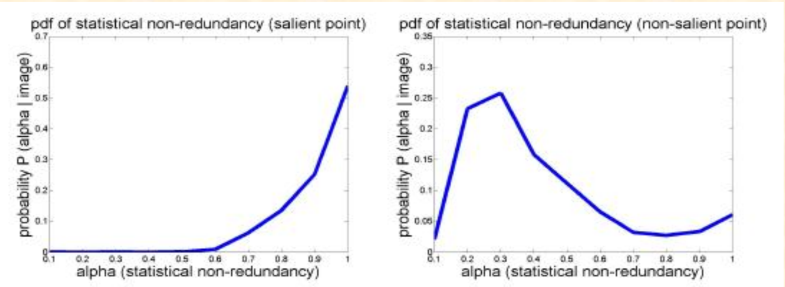
We quantify the statistical non-redundancy between two sites *i* and *j* as

$$\alpha_{ij} = 1 - P(N_i \mid N_j)$$

Saliency Map Generation:

For a particular site *i* we compute α_{ij} over all possible values of *j*, which gives us the probability of α_{ij} given the image *f* (denoted as $P(\alpha_{ij} \mid f)$). The figure below shows this distribution corresponding to two different sites in a natural image

We can see that the salient site has a higher degree of non-redundancy as compared to the non-salient site. This observation is in accordance with our expectations as the salient sites would have fewer matching neighbourhoods.



(a) salient site (b) non-salient site

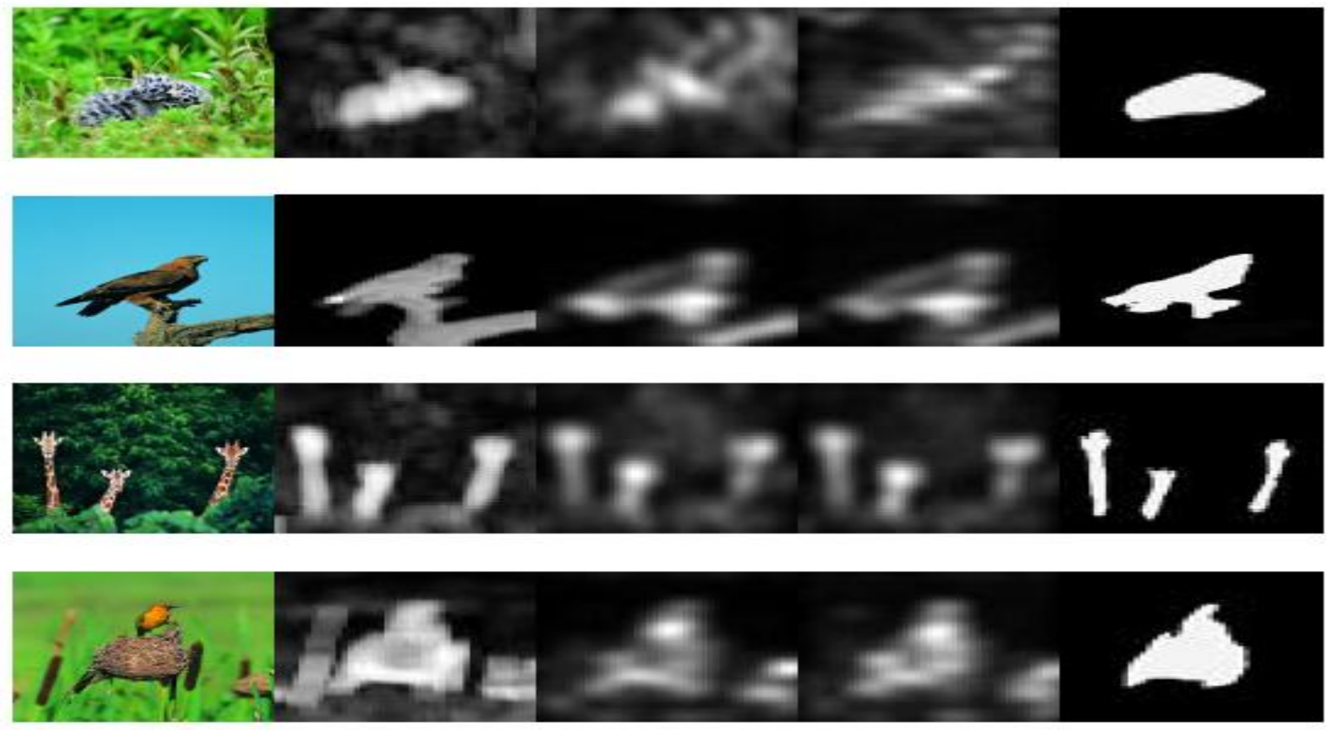
We use this fact to generate a saliency map where the intensity at each point is proportional to the statistical non-redundancy of the corresponding site in the original image. For each site *i* in the image *f*, we compute $P(\alpha_{ij} \mid f)$ based on all sites *j* in the image. We then define the overall statistical non-redundancy S_i as the expected value of α_{ij} given image *f:*

$$S_i = E(\alpha_{ij} \mid f)$$

The collective map of all S_i for all sites *i* in the image forms the saliency map of the image. The saliency map divides the image into high intensity and low intensity sites. The high intensity points in the image correspond to the salient points in the image and vice versa.

Experimental Results:

We have compared the performance of our algorithm (which will be denoted as NR) against spectral whitening (SW) and spectral residual (SR). The figure below shows the saliency maps produced by the mentioned methods for some natural images along with the corresponding ground truth images



original image NR SR SW ground truth

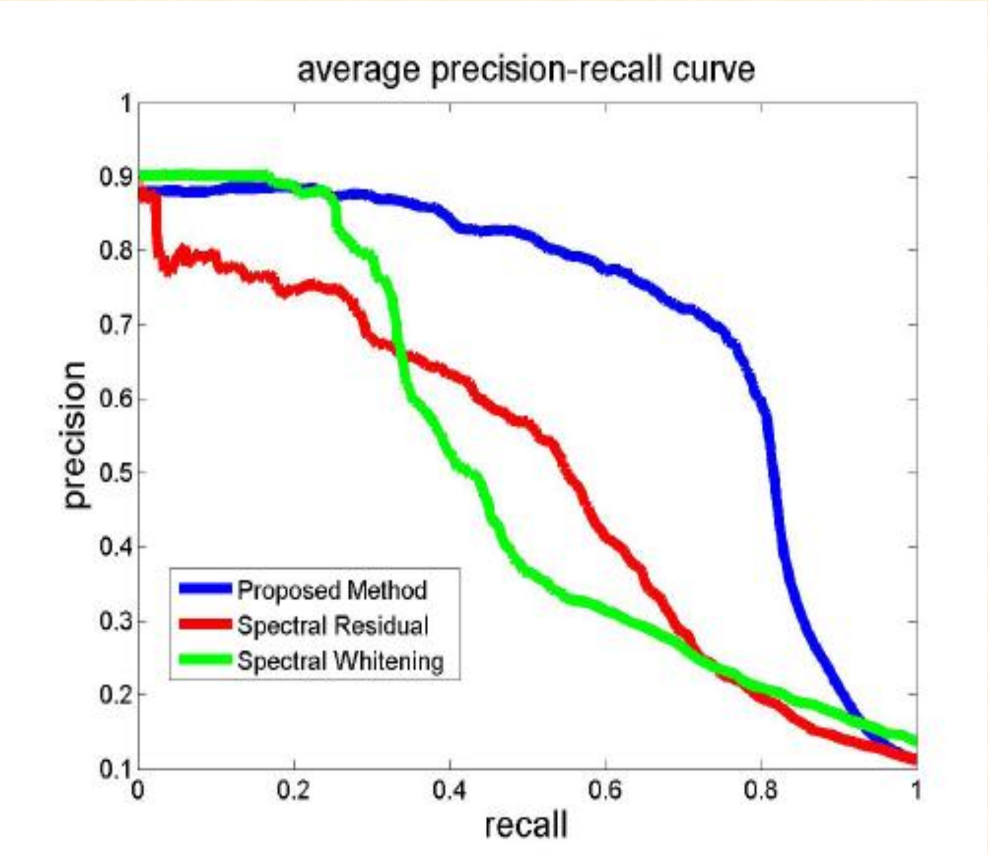
Acknowledgements:

This work is supported through funding from the Natural Sciences and Engineering Research Council of Canada.

Experimental Results (Cont.):

It can be observed that, with reference to ground truth images, the salient objects in the images are more well defined in the saliency maps obtained using the NR method when compared to other methods.

The figure below shows the average precision-recall curve taken over 20 test images. The precision-recall method shows the proposed NR method gives better performance than the SR and SW methods.



Based on these preliminary findings, the proposed NR method shows great potential as the first step in object recognition and other such computer vision and image processing tasks

Conclusions and Further Work:

- Preliminary results show that the proposed method holds strong potential for identifying salient objects of interest within an image without any prior knowledge about the underlying image content, making it a good first step in various image processing tasks.
- Future work in this direction involves the application of the proposed statistical non-redundancy based approach to a multi-scale wavelet based framework.
- Future work could involve the use of rotation-invariant filters while computing the statistical non-redundancy of the pixels in the image to give better saliency maps.

Reference:

X. Hou and L. Zhang, "Saliency Detection: A Spectral Residual Approach", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp 1-8

L.Itti and C. Koch, "Computational Modelling of Visual Attention", *Nature reviews neuroscience*, vol. 2, no. 3, pp 194-203, 2001