

Global Contrast based Salient Region Detection

Hongzhi Liu

Jun 20, 2018

Abstract

Humans routinely judge the importance of image regions and focus attention on important parts. Computationally detecting such salient image regions remains a significant goal as it allows preferential allocation of computational resources in subsequent image analysis and synthesis. Extracted saliency maps are widely used in many computer vision applications. Today, I read a thesis written by Mingming Cheng, who is in National Lab of Information Science and Technology from Tsinghua University. His team introduce a regional contrast based saliency extraction algorithm which simultaneously evaluates global contrast differences and spatial coherence. The proposed algorithm is simple, efficient, and yields full resolution saliency maps. Their algorithm consistently outperformed existing saliency detection methods, yielding higher precision and better recall rates.

1. Overview of Regulating Receptive Field

Global contrast based methods evaluate saliency of an image region using its contrast with respect to the entire image. However, for efficiency they use only luminance information, thus ignoring distinctiveness clues in other channels. Achanta *et al.* [2] propose a frequency tuned method that directly defines pixel saliency using a pixel's color difference from the average image color. The elegant approach, however, only considers first order average color, which can be insufficient to analyze complex variations common in natural images. Furthermore, these methods ignore spatial relationships across image parts which can be critical for reliable and coherent saliency detection.

In this paper, Dr. Cheng and his team focus on bottom-up data driven saliency detection using image contrast. They propose contrast analysis for extracting high-resolution, fullfield saliency maps based on the following observations [3]. Firstly, a global contrast based method, which separates a large-scale object from its surroundings, is preferred over local contrast based methods producing high saliency values at or near object edges. Then, global considerations

enable assignment of comparable saliency values to similar image regions, and can uniformly highlight entire objects. Furthermore, saliency of a region depends mainly on its contrast to the nearby regions, while contrasts to distant regions are less significant. The experiments show significant improvements over previous methods both in precision and recall rates.

2. Method of Global Contrast based Detection

2.1. Histogram Based Contrast

Based on the observation from biological vision that the vision system is sensitive to contrast in visual signal, Cheng proposes a histogram-based contrast (HC) method to define saliency values for image pixels using color statistics of the input image. Specifically, the saliency of a pixel is defined using its color contrast to all other pixels in the image *i.e.* the saliency value of a pixel I_k in image I is defined as Equation 1:

$$s(I_k) = \sum_{\forall I_i \in I} D(I_k, I_i). \quad (1)$$

where $D(I_k, I_i)$ is the color distance metric between pixels I_k and I_i in the $L^*a^*b^*$. Equation 1 can be expanded by pixel order. Hence, rearranging Equation 1 such that the terms with the same color value c_j are grouped together, they get saliency value for each color as Equation 2:

$$s(I_k) = S(c_l) = \sum_{j=1} n D(c_l, c_j). \quad (2)$$

where c_l is the color value of pixel I_k , n is the number of distinct pixel colors, and f_j is the probability of pixel color c_j in image I . However, given the strict efficiency requirement, they take the simple global approach.

In this work, the team use the full color space instead of luminance only. To reduce the number of colors needed to consider, they first quantize each color channel to have 12 different values, which reduces the number of colors to $12^3 = 1728$. Considering that color in a natural image typically covers only a small portion of the full color space, they further reduce the number of colors by ignoring less frequently occurring colors.

Table 1. Average time taken to compute a saliency map for images in the database by Achanta *et al.* [2]. Most images in the database have resolution 400×300 . Algorithms were tested using a Dual Core 2.6 GHz machine with 2GB RAM.

Method	IT[7]	MZ[8]	GB[5]	SR[6]	FT[2]	AC[1]	CA[4]	LC[9]	HC	RC
Time(s)	0.611	0.070	1.614	0.064	0.016	0.109	53.1	0.018	0.019	0.253
Code	Matlab	C++	Matlab	Matlab	C++	C++	Matlab	C++	C++	C++

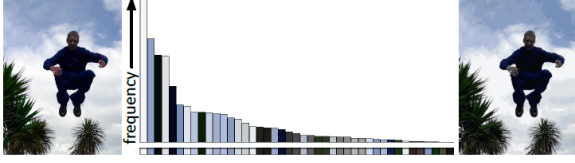


Figure 1. Given an input image (left), we compute its color histogram (middle). Corresponding histogram bin colors are shown in the lower bar. The quantized image (right) uses only 43 histogram bin colors and still retains sufficient visual quality for saliency detection.

A typical example of such quantization is shown in Fig. 1. Note that again due to efficiency requirements they select the simple histogram based quantization instead of optimizing for an image specific color palette.

In order to reduce noisy saliency results caused by such randomness, the team use a smoothing procedure to refine the saliency value for each color. They replace the saliency value of each color by the weighted average of the saliency values of similar colors (measured by $L^*a^*b^*$ distance). This is actually a smoothing process in the color feature space. Typically they choose $m = n/4$ nearest colors to refine the saliency value of color c by Equation 3:

$$S'(c) = \frac{1}{(m-1)T} \sum_{i=1}^m (T - D(c, c_i)) S(c_i). \quad (3)$$

where $T = \sum_{i=1}^m D(c, c_i)$ is the sum of distances between color c and its m nearest neighbors c_i , and the normalization factor comes from $\sum_{i=1}^m (T - D(c, c_i)) S(c_i) = (m-1)T$.

2.2. Region Based Contrast

Since directly introducing spatial relationships when computing pixel-level contrast is computationally expensive, the team introduce a contrast analysis method, region contrast (RC), so as to integrate spatial relationships into region-level contrast computation. Specifically, for any region r_k , the spatially weighted region contrast based saliency is defined as Equation 4:

$$S(r_k) = \sum_{r_i \neq r_k} \exp(-D_s(r_k, r_i) \sigma_s^2) w(r_i) D_r(r_k, r_i). \quad (4)$$

where $D_s(r_k, r_i)$ is the spatial distance between regions r_k and r_i , and σ_s controls the strength of spatial weighting.

Larger values of σ_s reduce the effect of spatial weighting so that contrast to farther regions would contribute more to the saliency of the current region. The spatial distance between two regions is defined as the Euclidean distance between their centroids. In their implementation, they use $\sigma_s^2 = 0.4$ with pixel coordinates normalized to $[0,1]$.

3. Experimental Comparisons

Cheng and his team have evaluated the results of our approach on the publicly available database provided by Achanta *et al.* [2]. They compared the proposed global contrast based methods with 8 state-of-the-art saliency detection methods. They used our methods and the others to compute saliency maps for all the 1000 images in the database. Tab. 1 compares the average time taken by each method. Our algorithms, HC and RC, are implemented in C++. For the other methods namely IT, GB, SR, FT and CA, they used the authors's implementations. For typical natural images, their HC method needs $O(N)$ computation time and is sufficiently efficient for real-time applications.

References

- [1] R. Achanta, F. Estrada, P. Wils, and S. Sstrunk. Salient region detection and segmentation. In *ICVS*, 2009. 2
- [2] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *CVPR*, 2009. 1, 2
- [3] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. M. Hu. Global contrast based salient region detection. In *CVPR*, 2011. 1
- [4] S. Goferman, L. Zelnikmanor, and A. Tal. Context-aware saliency detection. *IEEE TPAMI*, 2012. 2
- [5] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. In *NIPS*, 2006. 2
- [6] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *CVPR*, 2007. 2
- [7] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE TPAMI*, 1998. 2
- [8] Y. F. Ma and H. J. Zhang. Contrast-based image attention analysis by using fuzzy growing. In *ACM Multimedia*, 2003. 2
- [9] Y. Zhai and M. Shah. Visual attention detection in video sequences using spatiotemporal cues. In *ACM Multimedia*, 2006. 2