DIGITAL IMAGE PROCESSING PROJECT

Saliency Filter:Contrast Based Filtering for Salient Region Detection

Ashish R (2018702004) and Deepak Kumar Singh (2018701010)

November 28, 2018

Contents

1 Abstract

The computational identification that are likely to catch the attention of a human observer is a typical perceptual research problem. The research indicate that the most influential factor in low level visual saliency is contrast. This project is based on the observation that an image can be decomposed into basic, structurally representative elements that abstract away unnecessary detail and at the same time allow for very clear and intuitive definition of contrast based saliency. The definition of contrast in previous works is based on various different types of image features, including color variation of individual pixels, edges and gradients, spatial frequencies, structure and distribution of image patches, histograms, multi-scale descriptors, or combinations thereof.

This method is based on the observation that an image can be decomposed into basic, structurally representative elements that abstract away unnecessary detail, and at the same time allow for a very clear and intuitive definition of contrast-based saliency.

2 Approach

The process goes through the following modules:

Figure 1: Execution Pipeline

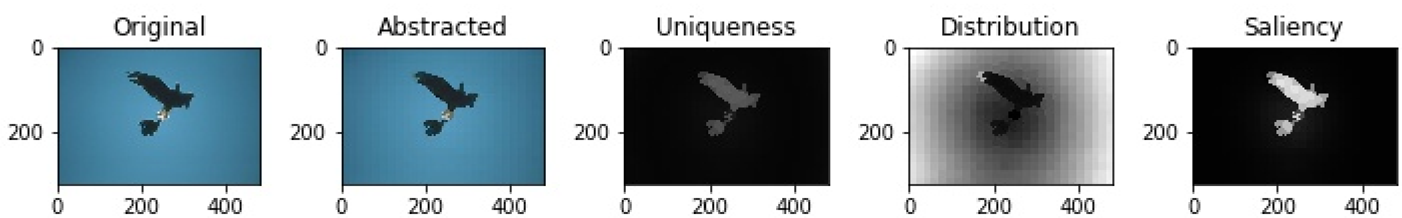


Figure 2: Results of the Pipeline

2.1 Abstraction

We aim to decompose the image into basic elements that preserve relevant structure, but abstract undesirable detail. Specifically, each element should locally abstract the image by clustering pixels with similar properties (like color) into perceptually homogeneous regions.

For the image abstraction we use an adaptation of SLIC superpixels to abstract the image into perceptually uniform regions. SLIC superpixels segment an image using K-means clustering in RGBXY space. The RGBXY space yields local, compact and edge aware superpixels, but does not guarantee compactness. For our image abstraction we slightly modified the SLIC approach and instead use K-means clustering in CIELab space.

2.2 Element Uniqueness

We evaluate how different each respective element is from all other elements constituting an image,essentially measuring the rarity of each element

Element uniqueness is generally defined as the rarity of a segment i given its position $\mathbf{p}_i$ and color in CIELab $\mathbf{c}_i$ compared to all other segments j:

$$U_i = \sum_{j=1}^{N} \|\mathbf{c}_i - \mathbf{c}_j\|^2 \cdot \underbrace{w(\mathbf{p}_i, \mathbf{p}_j)}_{w_{ij}^{(p)}}$$

By introducing $w_{ij}^{(p)} = \frac{1}{Z_i} \exp\left(-\frac{1}{2\sigma_p^2} \|\mathbf{p}_i - \mathbf{p}_j\|^2\right)$ we effectively combine global and local contrast estimation with control over the influence radius of the uniqueness operator. A local function $w_{i,j}^{(p)}$ yields a local contrast term, which tends to overemphasize object boundaries in the saliency estimation whereas $w_{i,j}^{(p)}$ yields a global uniqueness operator, which cannot represent sensitivity to local contrast variation.

σ_p controls the range of the uniqueness operator, we use $\sigma_p = 0.25$ in all our experiments and Z_i is the normalization factor ensuring $\sum_{j=1}^{N} w_{i,j}^{(p)} = 1$.

2.3 Element Distribution

Ideally colors belonging to the background will be distributed over the entire image exhibiting a high spatial variance, whereas foreground objects are generally more compact. We define the element distribution measure for a segment i using the spatial variance $\mathbf{D}_i$ of its color $\mathbf{c}_i$, i.e., we measure its occurrence elsewhere in the image. As motivated before, low variance indicates a spatially compact object which should be considered more salient than spatially widely distributed elements.

$$D_i = \sum_{j=1}^{N} \|\mathbf{p}_j - \mu_i\|^2 \underbrace{w(\mathbf{c}_i, \mathbf{c}_j)}_{w_{ij}^{(c)}}$$

$w_{ij}^{(c)} = \frac{1}{Z_i} \exp\left(-\frac{1}{2\sigma_c^2} \|\mathbf{c}_i - \mathbf{c}_j\|^2\right)$ describes the similarity of color $\mathbf{c}_i$ and color $\mathbf{c}_j$ of segments i and j, respectively, $\mathbf{p}_j$ is again the position of segment j, and $\mu_i = \sum_{j=1}^{N} w_{ij}^{(c)} \mathbf{p}_j$ defines the weighted mean position of color $\mathbf{c}_i$

The parameter σ_c controls the color sensitivity of the el- ement distribution. We use $\sigma_c = 20$ in all our experiments.

2.4 Saliency Assignment

The two above contrast measures are defined on a per-element level. In a final step, we assign the actual saliency values to the input image to get a pixel-accurate saliency map.

We start by normalizing both uniqueness U_i and distribution D_i to the range [0::1]. We assume that both measures are independent, and hence we combine these terms as follows to compute a saliency value S_i for each element:

$$S_i = U_i \cdot \exp(-k \cdot D_i)$$

In practice we found the distribution measure D_i to be of higher significance and discriminative power. Therefore, we use an exponential function in order to emphasize D_i.In all our experiments we use k = 6 as the scaling factor for the exponential. As the final step, we need to assign a final saliency value to each image pixel, which can be interpreted as an upsampling of the per-element saliency S_i. However, naive up-sampling by assigning S_i to every pixel contained in element i carries over all segmentation errors of the abstraction algorithm. Instead we adopt an idea proposed in the context of range image up-sampling [2] and apply it to our framework. We define the saliency $\tilde{S}_i$ of a pixel as a weighted linear combination of the saliency S_j of its surrounding image elements

$$\tilde{S}_i = \sum_{j=1}^{N} w_{ij} S_j$$

By choosing a Gaussian weight $w_{ij} = \frac{1}{Z_i} \exp\left(-\frac{1}{2}\left(\alpha \|\mathbf{c}_i - \mathbf{c}_j\|^2 + \beta \|\mathbf{p}_i - \mathbf{p}_j\|^2\right)\right)$, we ensure the up-sampling process is both local and color sensitive. Here alpha and beta are parameters controlling the sensitivity to color and position. We found alpha = 1/30 and beta = 1/30 to work well in practice.

3 Work Performed

3.1 Experiments

Various test images were taken from The Berkeley Segmentation Dataset and Benchmark [3] and from our own phone camera and were tested on various parameter settings.

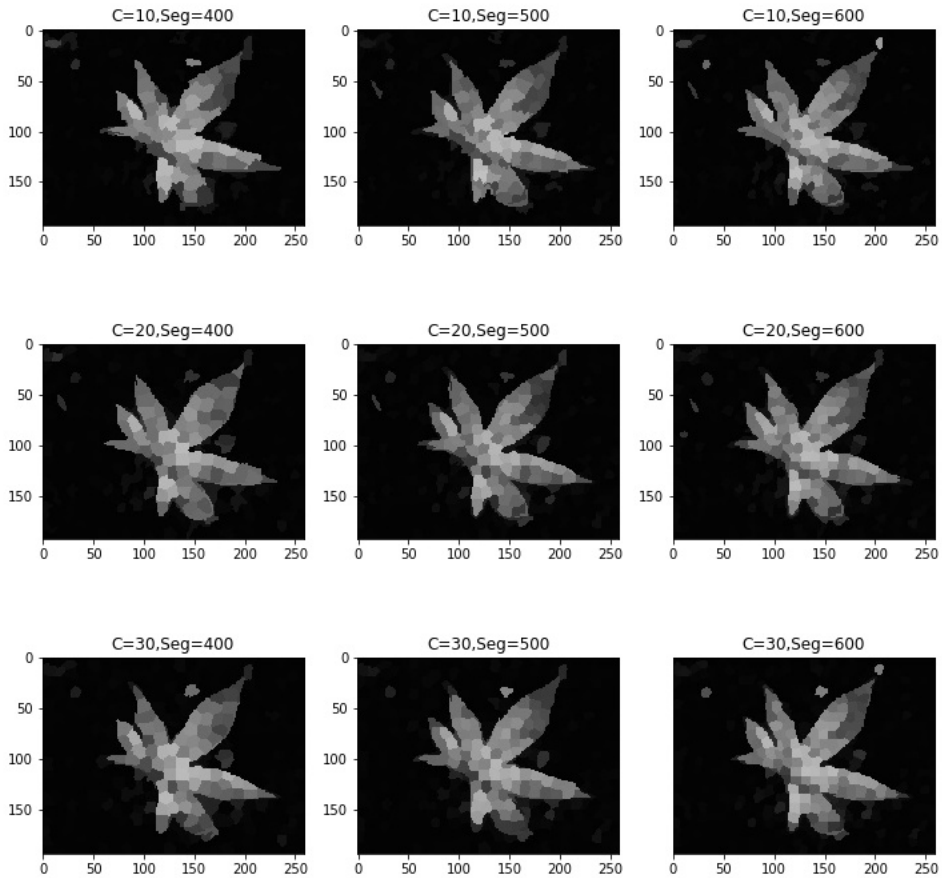


Figure 3: Results for varying Compactness and Segments parameters

The following table shows the grid search performed for various parameters.The table below explains the time taken for execution with various values of Compactness and Segments used in the SLIC implementation of Super Pixels.

Compactness/Segments	400	500	600
10	0.32	0.43	1.3
20	1.3	1.48	1.13
30	0.42	0.5	1.16

Table 1: Timing Analysis (minutes) for Compactness and Number of Segments for super-pixels

The value of K decides the weight of contribution of **Uniqueness** and **Distribution**, higher the value of K assigns lesser weight for **Distribution**

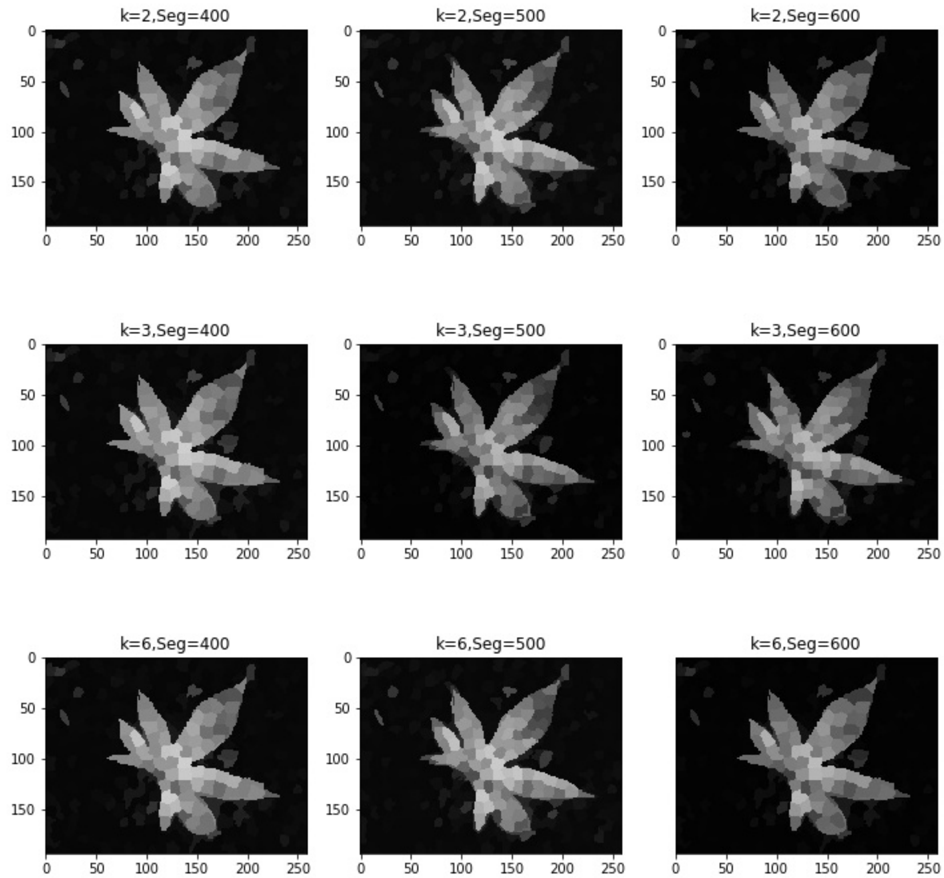


Figure 4: Results for varying K and Segments parameters

K/Segments	400	500	600
2	0.37	0.53	1.15
3	0.38	0.47	1.12
6	0.40	0.49	1.15

Table 2: Timing Analysis (minutes) for K and Number of Segments of super-pixels

3.2 Softwares & Libraries

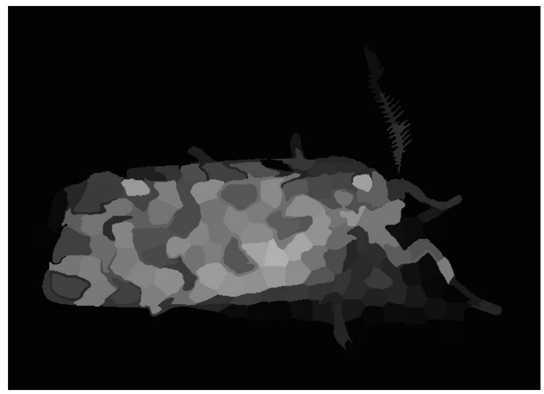
- Jupyter Notebook
- python3
- numpy
- skimage
- matplotlib
- cv2
- sys

4 Successful Results

Input

Output

Input



Output

Input

Output

Input

Output

Input

Output

Input

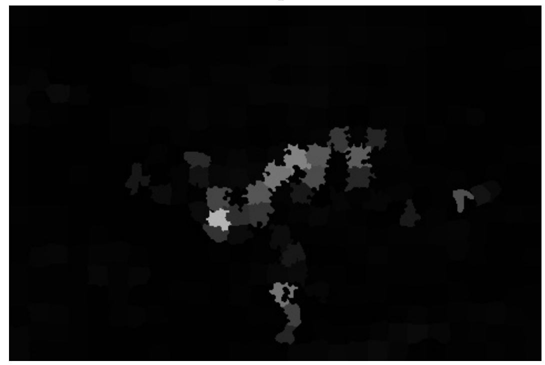
Output

5 Failure Results

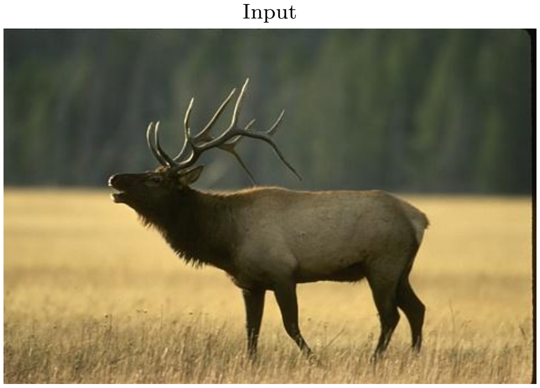
Input

Output

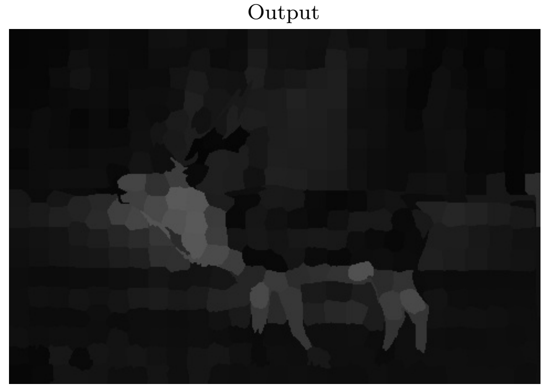
Input



Output



Input



Output

6 Analysis of Results

From the experiments performed on various test images , as shown above, it is found that the Salient Region Detection, the foreground and background must have good contrast difference.It is also found that, if there are components of the foreground that look similar to background, then the results of the Saliency Image is affected.

Further Improvement that could give better results are, trying to exploit the variance and other statistical properties of the independent segments generated by K means clustering, to decide the variance values in the **Uniqueness** and **Distribution** assignment.

Presence of the shadow of the foreground in the background also affects the results.

6.1 Mean Absolute Error

For a more balanced comparison ,we also evaluate the mean absolute error (MAE) between the continuous saliency map S (prior to thresholding) and the binary ground truth GT

$$MAE = \frac{1}{W \times H} \sum_{x=1}^{W} \sum_{y=1}^{H} |S(x,y) - GT(x,y)|$$

The MAE was found to be 0.46 for our test with a binary ground truth image.

7 Github link

https://github.com/deepakksingh/Contrast-Based-Filtering-for-Salient-Region-Detection

8 Task Assignment

Concept Understanding Concept of Super Pixels Element Distribution Tuning of Parameters for Element Distribution	Ashish
Uniqueness Assignment Understanding Permutohedral Lattice Embedding Tuning of Parameters for Element Uniqueness Applying Permutohedral Lattice for Element Distribution	Deepak
Saliency Assignment Testing and Improvement PPT and Report	Both

9 References

References

[1] *Perazzi, F., Krahenbuhl, P., Pritch, Y., Hornung, A. (2012, June). Saliency filters: Contrast based filtering for salient region detection. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 733-740). IEEE*

[2] *J. Dolson, J. Baek, C. Plagemann, and S. Thrun. Upsampling range data in dynamic environments. [In CVPR, pages 1141–1148, 2010].*

[3] The Berkeley Segmentation Dataset and Benchmark
`http://www-cs-faculty.stanford.edu/~uno/abcde.html`

[4] SLIC based Superpixel Segmentation
`https://jayrambhia.com/blog/superpixels-slic`