

Computer Vision 1: Assignment 2

(Due date: 21.11.2022)

Task 1: Template matching (programming)

Correlation filters respond most strongly to regions of images that look like the filter itself. This property is used for template matching. In template matching, we cross correlate a template with an image. The template is another image that is a prototype for what we want to localize in an image. The cross-correlation is the strongest where the image looks most like the template.

In this task, we will use template matching to match an image of a ship captured from a webcam stream to a wider panoramic view of the area. The matching result could potentially later be used for other applications, for example for estimating the heading of the ship.

- Download the template and panorama images, and convert them to grayscale.
- Use the function `skimage.feature.match_template` to compute the normalized correlation between two images. Find the coordinates where the cross-correlation is the strongest. Visualize the matching result by drawing a rectangle with the size of the template image over the panorama image. You can for example use `matplotlib.pyplot.Rectangle` for this - see the API documentation for how to draw the rectangle.
- Flip the template horizontally. Verify that the matching is no longer working by visualizing it.

Task 2: Edge detection (programming)

Noise negatively affects edge detection. To examine the effect of this in practice, we apply edge detection on a noisy image and a smoothed image.

- Load the image `woman.png` provided. Note that it is already grayscale.
- Using `skimage.util.random_noise`, add Gaussian distributed noise to the image with a variance of 0.01. This gives you a noisy image N .
- Apply `skimage.filters.gaussian` to filter N . Use $\sigma = 1.0$. This gives you a smoothed image S .

Next, apply the Sobel filter as follows.

- Use the function `skimage.filters.sobel`. This function applies the two Sobel kernels (horizontal and vertical) on the image, and then computes the root of the squared sum of the results.
- Visualize the result of applying the function on the noisy image N , and on the smoothed image S .

Suppose the output of the `sobel` function on the noisy image is F_n and the output on the smoothed image is F_s .

- Draw histograms of intensity values in F_s and F_n and use them to select two threshold values t_n and t_s for edge detection.

- Create a binary mask by applying the logical operations $F_n > t_n$ and $F_s > t_s$ in NumPy. The parts of the image where the threshold is exceeded (there is an edge) are indicated by true values. Visualize the binary masks.
- Tune your threshold values until the detected edges show the outline of the woman's face.
- Did you notice it is almost impossible to tune the threshold for the noisy image to show just the outline of the face? Why is edge detection improved by applying smoothing?

Task 3: Image pyramids (programming)

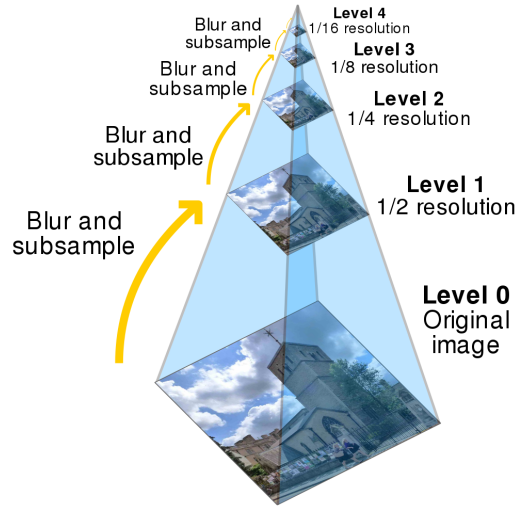


Figure 1: A Gaussian image pyramid with four layers. Image by user “Cmglee”, Wikipedia.

Recall the Gaussian image pyramid as shown in Figure 1. In this task, we use Gaussian pyramids to calculate a saliency map, following the procedure illustrated in Figure 2. The approach generates two Gaussian pyramids, a center pyramid and a surround pyramid. Subtracting two layers from these pyramids corresponds to applying a Difference of Gaussian filter. This is useful to detect contrasts in images, which is the basis for saliency computation.

First read the documentation of the function `skimage.transform.pyramids.pyramid_gaussian` online. Load the image `visual_attention.png` from Moodle, and convert it to grayscale. Then, calculate the center and surround pyramids. Use the following parameters:

- Use `max_layers=4` to obtain five layers as shown in Figure 1.
- Center pyramid: $\sigma_1 = 9$,
- Surround pyramid: $\sigma_2 = 16$.

We shall denote layer j of the center pyramid by C_j and layer j of the surround pyramid by S_j . Visualize each of the layers in both pyramids.

Next, calculate *on-off contrast pyramid* and the *off-on contrast pyramid*, shown in the middle part of Figure 2, as follows:

- For the on-off contrast pyramid, the j th layer is given by $C_j - S_j$
- For the off-on contrast pyramid, the j th layer is given by $S_j - C_j$
- Important: use `np.clip` to clip values from both subtractions to the range $[0, 1]$.

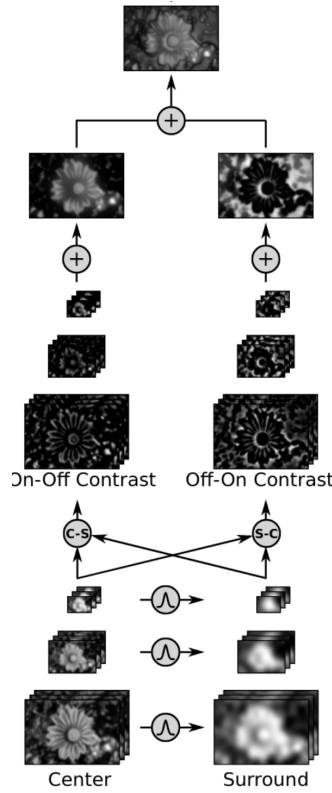


Figure 2: Saliency computation using two difference-of-Gaussian pyramids

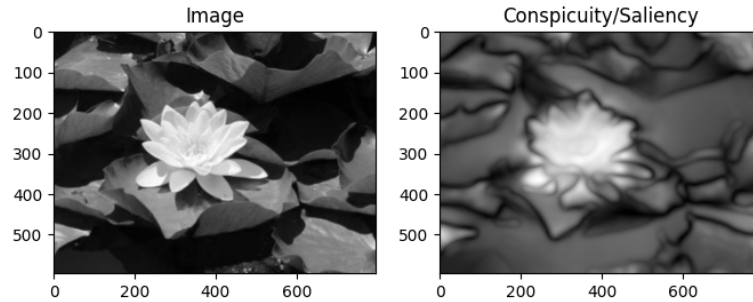


Figure 3: The input image converted to grayscale (left), and the final conspicuity map (right).

- Visualize each of the layers in both pyramids.

For both the on-off pyramid and the off-on pyramid calculate the feature maps shown in the top part of Figure 2 after the first plus operators as follows:

- Upsample each layer $j \geq 1$ to the same size as layer $j = 0$. Use `skimage.transform.resize` which will perform interpolation to upsample.
- Average across all layers pixelwise to obtain the feature map.
- Visualize the feature map.

Finally, calculate the conspicuity map or saliency map corresponding to the top of Figure 2 by averaging the two feature maps (on-off and off-on) pixelwise. Visualize the final conspicuity map. The result should look similar to Figure 3 (right).

What advantage does using this approach based on image pyramids have over the integral image method from assignment sheet 1?