Gaze Fixation Density Maps / Wooding maps

Context

In this lab, we will will compute the Gaze Fixation Density Maps / Wooding maps from the gaze fixation points available in the MexCulture142 dataset.

GFDM / Wooding maps

A Gaze Fixation Density Map (GFDM) is obtained by convolving the fixation map by an isotropic bi-dimentional Gaussian Function:

$$S_g(X) = [\frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} (\sum_{m=1}^{M_{fix}} \delta(X - x_{f(m)}))] * G_\sigma(X) \tag{1}$$

where:

- X is a vector representing the spatial coordinates
- $x_{f(m)}$ is the spatial coordinate of the m^{th} visual fixation
- M_{fix} is the number of visual fixation for the i^{th} observer
- N_{obs} is the number of observers
- $\delta(.)$ is the Kronecker symbol, $\delta(t) = 1$ if $t = 1$, otherwise $\delta(t) = 0$

For each gaze record, we compute a partial saliency map by applying a two-dimensional Gaussian centered on the fixation. Then these partial saliency maps are summed to get a global saliency map. Finally, this global map is normalized by its maximum value.

Wooding proposed to fix the Gaussian σ propagation at an angle α of 2° (based on an imitation of the functioning of the fovea of the human eye which covers an area of 1.5° to 2° of the diameter in the center of the retina). The Gaussian reflects the projection of the fovea on the screen, so the Gaussian spread σ is defined as:

$$\sigma = R.D.tan(\alpha) \tag{2}$$

with R the vertical resolution of the screen in pixels per mm and D about three times the height of screen. Gaze fixations were acquired on a screen of resolution

1920x1200, of height 325mm, and with eyes at a distance D of 900mm, and images were displayed rescaled to screen resolution, preserving the aspect ratio.

The partial saliency map $S(I, m)$ for the image I and fixation $m = (x_{0_m}, y_{0_m})$ is computed according to the following equation:

$$S(I, m) = Ae^{-\left(\frac{(x-x_{0_m})^2}{2\sigma_x^2}+\frac{(y-y_{0_m})^2}{2\sigma_y^2}\right)} \tag{3}$$

where $\sigma_x = \sigma_y = \sigma$ and $A = 1$.

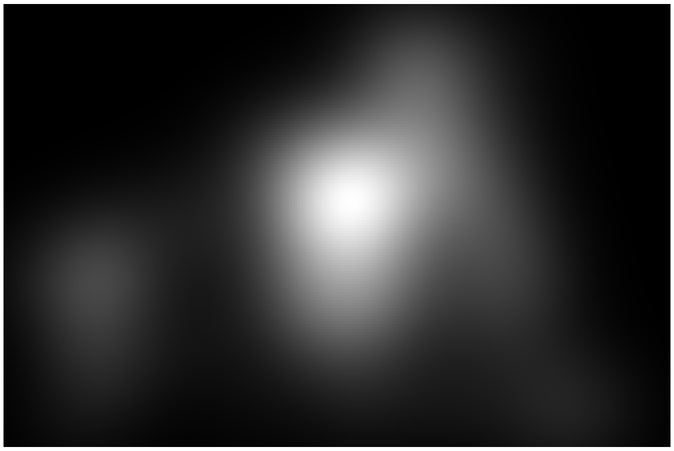
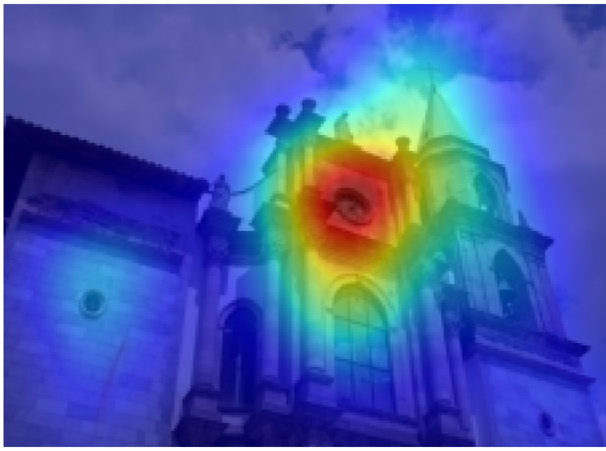
MexCulture142 dataset

We will use the MexCulture142 dataset. This dataset contains 284 images of 142 Mexican monuments. They are classified in 3 styles: Prehispanic, Colonial and Modern. The images are in different resolutions.

The dataset is available at CREMI in: */net/ens/DeepLearning/DLCV2024/MexCulture142*.
The fixations are available in: */net/ens/DeepLearning/DLCV2024/MexCulture142/fixations*
The *official* GFDM are available in:
/net/ens/DeepLearning/DLCV2024/MexCulture142/gazefixationsdensitymaps
They may be used for comparison purposes.

Work

You should compute the saliency maps and save them as grayscale images, and projected color mapped on the original image. Here is an original image and the correpsonding two saliency maps images.

The following libraries are useful to that:

- `numpy` to handle the feature maps
- `Pillow` to handle the generated images and `blend` them
- `matplotlib` to handle the colormaps

You should submit an archive with your work on Moodle. This archive should contain your python code and a small report. You should show some images of saliency maps as grayscale images and projected on the original image with a color map.