CSC420 Assignment 3

Sailing Ni

Mar 2023

1 | Laplacian of Gaussian

(1): Characteristic scale of an image of black circle on a white background
Assume radius of circle is r, so $D = 2r$.
We want to find the σ such that it maximizes the response magnitude function, which is the convolution of a LoG filter on the image.
Since the circle is black, then if $x^2 + y^2 \le r^2$, $I(x, y) = 0$; we only consider the white area outside of circle, where $x^2 + y^2 \ge r^2$.
i.e. normalized response magnitude:

$$\sigma^2 \iint_D \nabla^2 G(x, y, \sigma) I(x, y)\, dx\, dy, where\ D = \{(x, y) \in R^2 | x^2 + y^2 \ge r^2\}$$

Let $x = \rho cos\theta$, $y = \rho sin\theta$, where $\rho \in [r, +\infty), \theta \in [0, 2\pi]$

$$\begin{aligned}
F(\sigma) &= \sigma^2 \int_r^{+\infty} \int_0^{2\pi} \frac{1}{\pi\sigma^4}(\frac{\rho^2}{2\sigma^2} - 1)e^{-\frac{\rho^2}{2\sigma^2}} \rho\, d\theta\, d\rho \\
&= \frac{2\pi\sigma^2}{\pi\sigma^4} \int_r^{+\infty} (\frac{\rho^3}{2\sigma^2} - \rho)e^{-\frac{\rho^2}{2\sigma^2}}\, d\rho \\
&= \frac{2}{\sigma^2}\Big[\int_r^{+\infty} \frac{\rho^3}{2\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}}\, d\rho - \int_r^{+\infty} \rho e^{-\frac{\rho^2}{2\sigma^2}}\, d\rho\Big] \\
\int \frac{\rho^3}{2\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}}\, d\rho &= \frac{1}{2\sigma^2} \int \rho^3 e^{-\frac{\rho^2}{2\sigma^2}}\, d\rho \\
&= \frac{1}{2\sigma^2}(-\sigma^2)(\rho^2 + 2\sigma^2)e^{-\frac{\rho^2}{2\sigma^2}} + C \\
&= -\frac{1}{2}\rho^2 e^{-\frac{\rho^2}{2\sigma^2}} - \sigma^2 e^{-\frac{\rho^2}{2\sigma^2}} + C \\
\int \rho e^{-\frac{\rho^2}{2\sigma^2}}\, d\rho &= -\sigma^2 e^{-\frac{\rho^2}{2\sigma^2}} + C \\
F(\sigma) &= \frac{2}{\sigma^2}\Big[-\frac{1}{2}\rho^2 e^{-\frac{\rho^2}{2\sigma^2}} - \sigma^2 e^{-\frac{\rho^2}{2\sigma^2}} - (-\sigma^2 e^{-\frac{\rho^2}{2\sigma^2}})\Big]\Big|_r^{+\infty} \\
&= \frac{2}{\sigma^2}\Big[-\frac{1}{2}\rho^2 e^{-\frac{\rho^2}{2\sigma^2}}\Big]\Big|_r^{+\infty} \\
&= -\frac{1}{\sigma^2}\Big[\rho^2 e^{-\frac{\rho^2}{2\sigma^2}}\Big]\Big|_r^{+\infty} \\
&= \frac{r^2}{\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}}
\end{aligned}$$

To find the maximum value of $F(\sigma)$. take derivative:

$$\begin{aligned}
F'(\sigma) &= r^2[-2\sigma^{-3} e^{-\frac{\rho^2}{2\sigma^2}} + \sigma^{-2} e^{-\frac{\rho^2}{2\sigma^2}}(-\frac{r^2}{2})(-2)\sigma^{-3}] \\
&= r^2\sigma^{-3} e^{-\frac{\rho^2}{2\sigma^2}}[-2 + r^2\sigma^{-2}]
\end{aligned}$$

Since $r^2\sigma^{-3}e^{-\frac{\rho^2}{2\sigma^2}} > 0$, set $-2 + r^2\sigma^{-2} = 0$; when $\sigma = \frac{r}{\sqrt{2}} = \frac{D}{2\sqrt{2}}$, $F(\sigma)$ is at its local maximum.
When $\sigma < \frac{r}{\sqrt{2}}, F'(\sigma) > 0$; when $\sigma > \frac{r}{\sqrt{2}}, F'(\sigma) < 0$.
Therefore, $\sigma = \frac{D}{2\sqrt{2}}$ maximizes response magnitude of LoG filter.

(2): Characteristic scale of an image of white circle on a black background

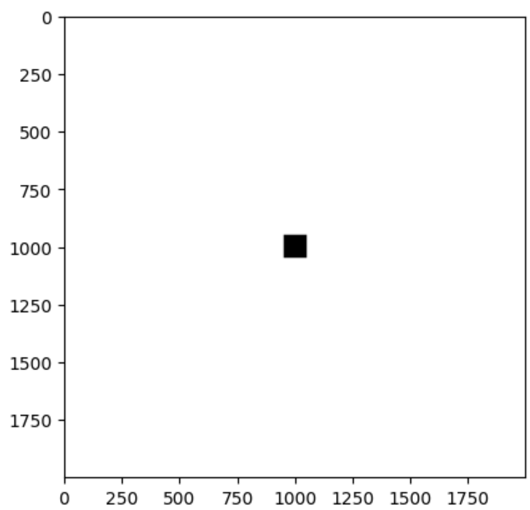
Since this time the circle is in white, if $x^2 + y^2 \geq r^2$, $I(x, y) = 0$; we only consider white area inside the circle. Set up the proof similar as part (1), but change the integral domain to $[0, r]$ and find the minimum this time.
Denote response magnitude function as $H(\sigma)$:

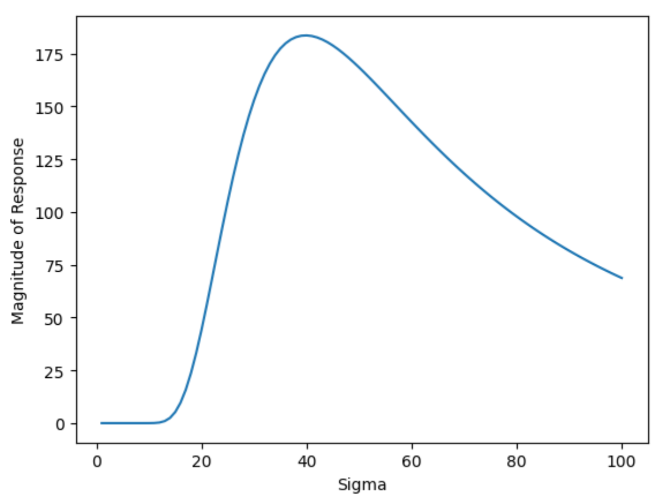
$$\begin{aligned} H(\sigma) &= -\frac{1}{\sigma^2}\left[\rho^2 e^{-\frac{\rho^2}{2\sigma^2}}\right]\Big|_0^r \\ &= -\frac{1}{\sigma^2}[r^2 e^{-\frac{\rho^2}{2\sigma^2}}] = -\frac{r^2}{\sigma^2}e^{-\frac{\rho^2}{2\sigma^2}} \\ H'(\sigma) &= -r^2[(-2)\sigma^{-3}e^{-\frac{\rho^2}{2\sigma^2}} + \sigma^{-2}e^{-\frac{\rho^2}{2\sigma^2}}(-\frac{r^2}{2}\sigma^{-3})] \\ &= r^2\sigma^{-3}e^{-\frac{\rho^2}{2\sigma^2}}[2 - \frac{r^2}{\sigma^2}] \end{aligned}$$

Compare with $F'(\sigma)$, note $H'(\sigma) = -F'(\sigma)$ holds.
When $\sigma = \frac{r}{\sqrt{2}} = \frac{D}{2\sqrt{2}}$, $H(\sigma)$ is at its minimum.
When $\sigma < \frac{r}{\sqrt{2}}, H'(\sigma) < 0$; when $\sigma > \frac{r}{\sqrt{2}}, H'(\sigma) > 0$.
Therefore, $\sigma = \frac{D}{2\sqrt{2}}$ minimizes response magnitude of LoG filter.

(3): Experiment in code

See `a3_q1ipynb` for code and solutions.
Plot a 100x100 black square in a 2000x2000 white image. Response magnitude for $\sigma \in [1, 100]$.
$\sigma = 40$ maximizes the magnitude of the response.





2 Corner Detection

(1): Compute eigenvalues of N

$$\begin{aligned}
det(N-\lambda I) &= 0 \\
det\begin{pmatrix} I_x^2-\lambda & I_xI_y \\ I_xI_y & I_y^2-\lambda \end{pmatrix} &= 0 \\
I_x^2I_y^2-\lambda(I_x^2+I_y^2)+\lambda^2-I_x^2I_y^2 &= 0 \\
\lambda^2 &= \lambda(I_x^2+I_y^2)
\end{aligned}$$

Therefore, eigenvalues of N are: $\lambda_1 = 0$, $\lambda_2 = I_x^2 + I_y^2$

(2): Prove M is positive semi-definite

By definition of positive semi-definite, want to prove:

$$Z^TMZ \geq 0 \text{ for all } Z \in \mathbf{R}^n, \text{where } M \text{ is symmetric matrix.}$$

(i). Matrix N is positive semi-definite

Matrix $N = \begin{pmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{pmatrix}$ is symmetric since $N = N^T$.

Given part (1), the two eigenvalues $\lambda_1 = 0$, $\lambda_2 = I_x^2 + I_y^2$ are both non-negative; therefore, N is positive semi-definite and we can diagonalize it to $Z^TMZ \geq 0$ for all $Z \in \mathbf{R}^n$.

(ii). M is symmetric matrix

$M = \sum_x \sum_y w(x,y)N$.

Because N is a symmetric matrix and $w(x,y)$ is just a scalar, M is also a symmetric matrix.

(iIi). M is positive semi-definite

$$\begin{aligned}
Z^TMZ &= Z\sum_x\sum_y w(x,y)N(x,y)Z^T \\
&= \sum_x\sum_y w(x,y)ZN(x,y)Z^T
\end{aligned}$$

Since N is positive semi-definite, $Z^TNZ \geq 0$.

Since $w(x,y)$ is a window function, $w(x,y) \geq 0$.

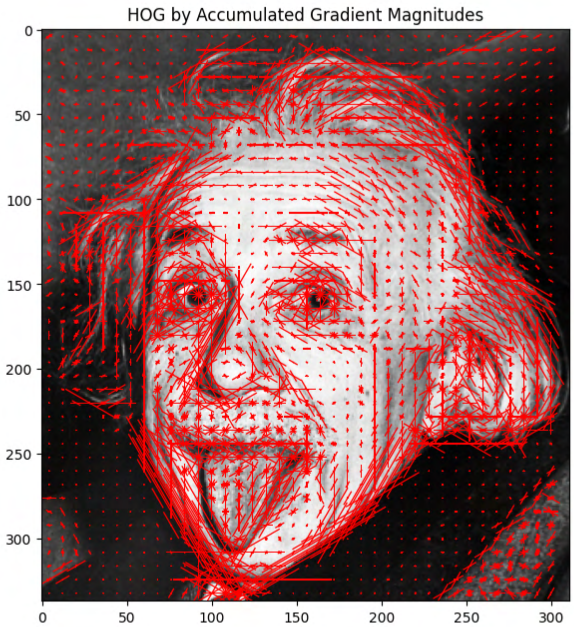
Therefore, all terms are non-negative, and $Z^TMZ \geq 0$.

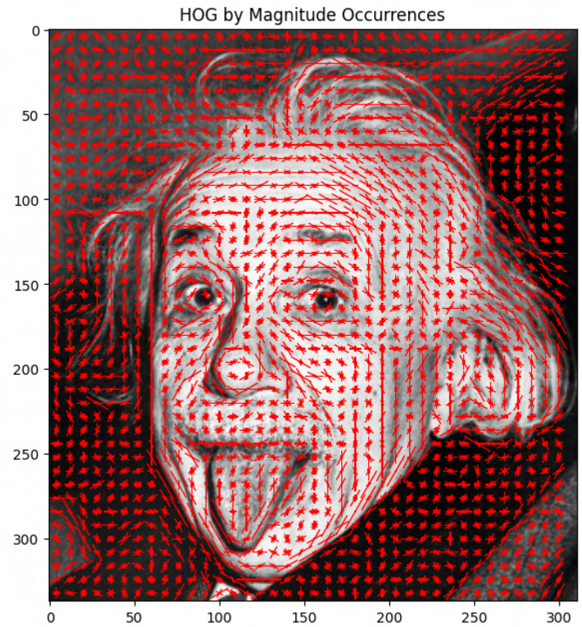
For symmetric matrix M, it satisfies $Z^TMZ \geq 0$ for all $Z \in \mathbf{R}^n$; so M is positive semi-definite.

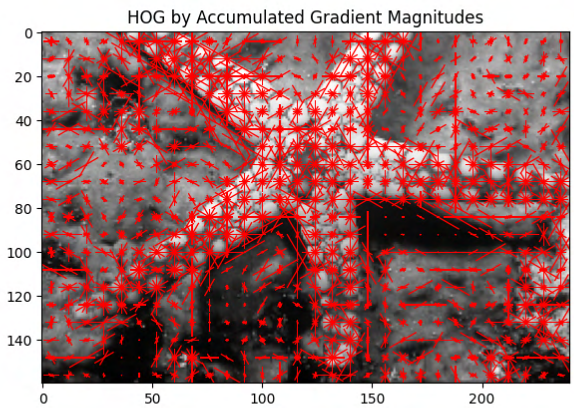
3 Histogram of oriented gradients

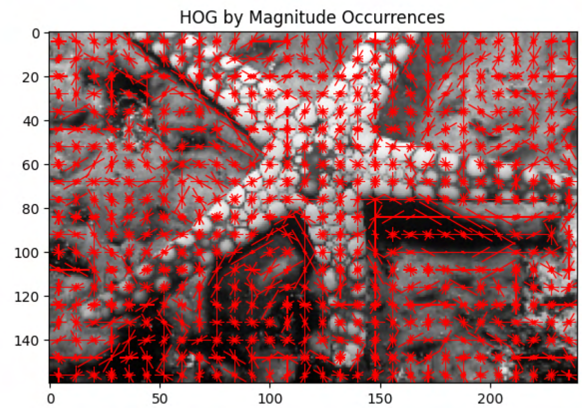
Implementation and images are in `q3` folder. See `a3_q3ipynb` for code and solutions.

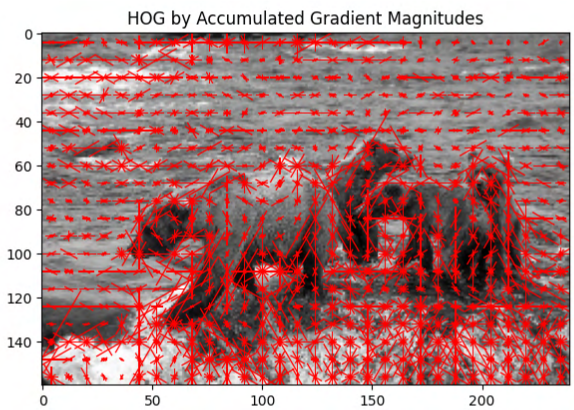
Step 1-3: Visualize HOG 3D array

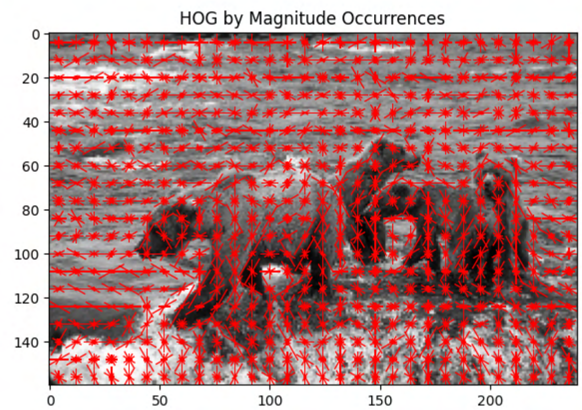






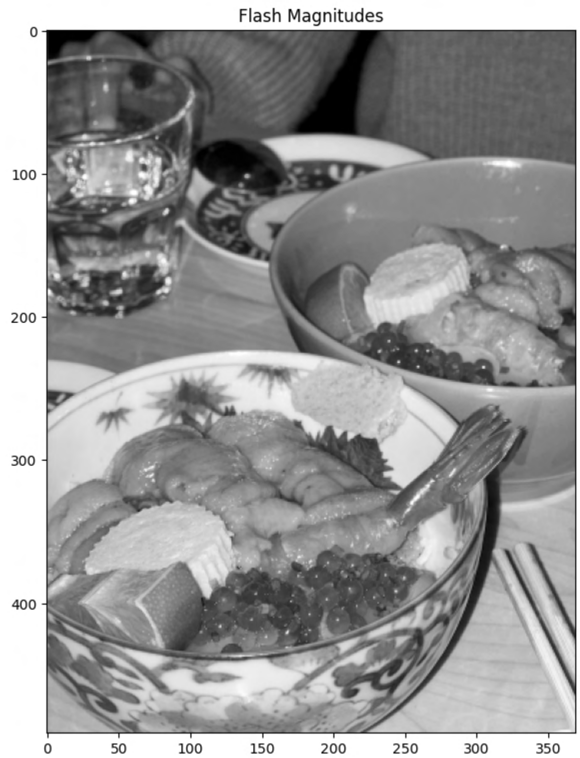


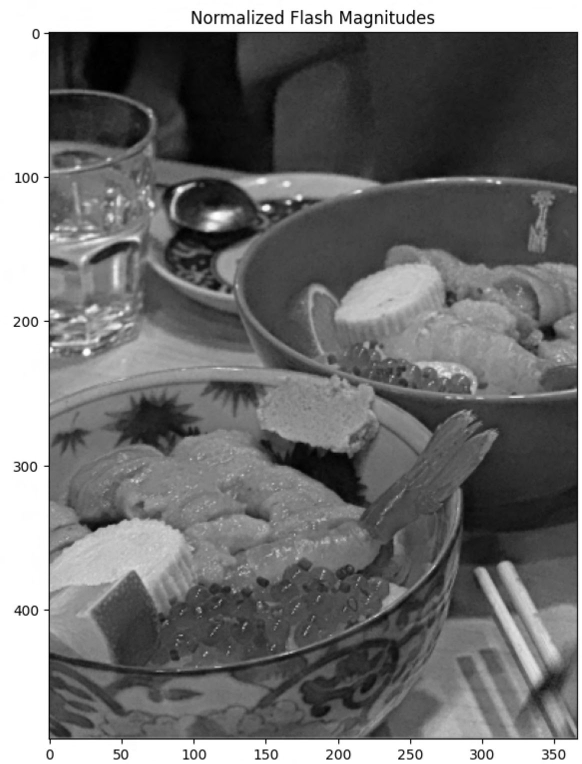




Comment:
Comparing the three pairs of images, accumulated magnitude method tends to outperform number occurrences method in object detection accuracy. Use accumulated gradient magnitudes for the remaining tasks.

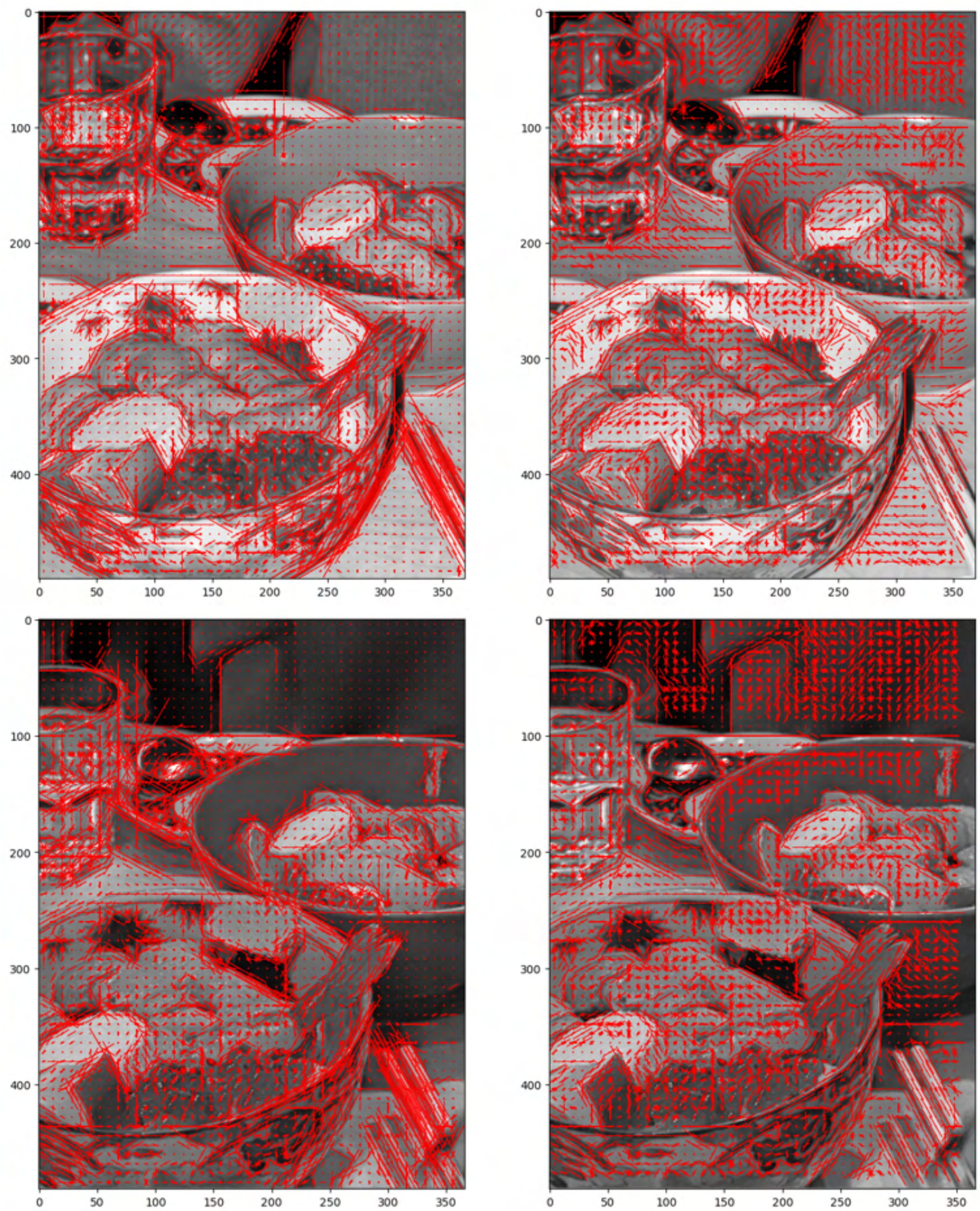
Step 4: Block Normalization on Flash vs Non-flash Images
Below are the grayscale version of flash and non-flash image.





Perform L2 normalization on each 2x2 blocks for both images, and visualize arrays generated before and after normalization for comparisons.

Visualization:
- To compare the results with and without the normalization, I visualized two pairs of comparing pictures for both flash and non-flash pictures.
- After normalization, there are (m-1)*(n-1) cells of 2x2 blocks; each contains 24 entries, corresponding to the 6 direction bins each of the 4 cells contains.
- For each 2x2 block's entries, the first 6 entries are the normalized magnitudes of the original 8*8 cells for the top-left cell among the four.
- Therefore, we will use these first 6 normalized entries to visualize each cell. Due to reduced size, the last horizontal and vertical cells are ignored in visualization since there are only (m-1)*(n-1) 2x2 blocks.



Comment:

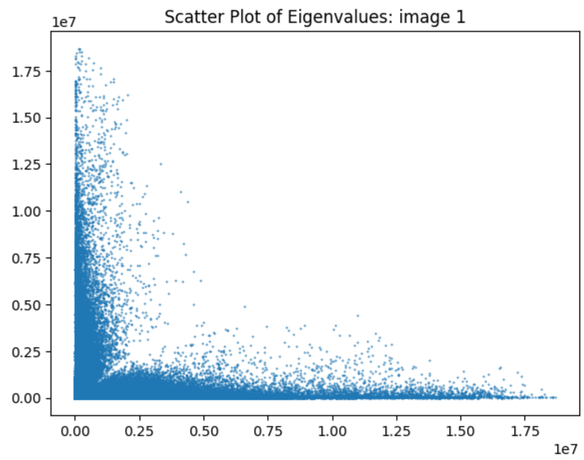
- Overall, visual comparisons suggest normalization of HOG in this case is not very beneficial. - After normalization, the resultant magnitudes are more invariant in illumination and shadowing changes. The block normalization process has a trade-off:
- On the one hand, normalization reduces the effects of local variation in the texture of same objects. (e.g. after normalization, accuracy for the chopsticks on the bottom-right corner has improved).
- On the other hand, normalization can lead to loss of finer details for large magnitudes and added noise. (e.g. plain areas like bowl are harder to detect due to introduced noise).
- Compare normalization results on flash and non-flash images, flash image seems to have slightly better performance and less noise due to its higher contrast.

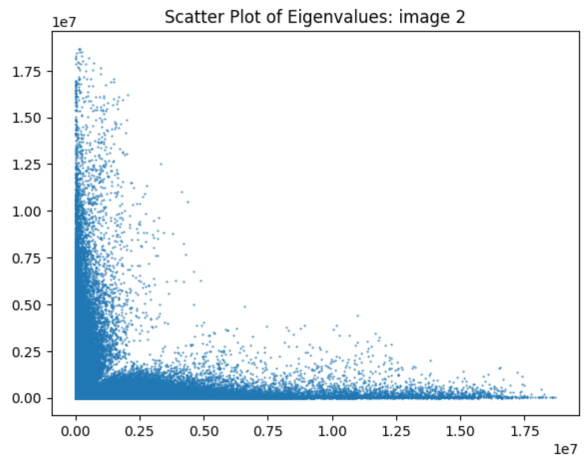
4 Corner Detection

Implementation and images are in `q4` folder. See `a3_q4ipynb` for code and solutions.

Step 1-2: Scatter Plot of eigenvalues

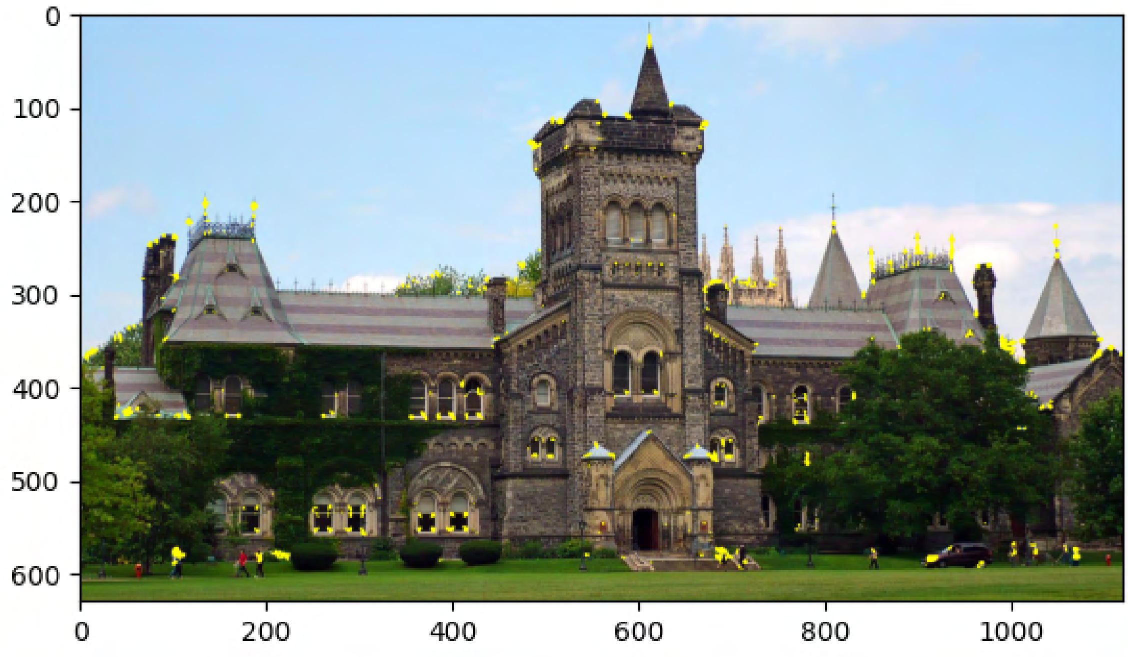
Use $\sigma = 2$ to calculate second moment matrix for each graph. Scatter plots are shown below:



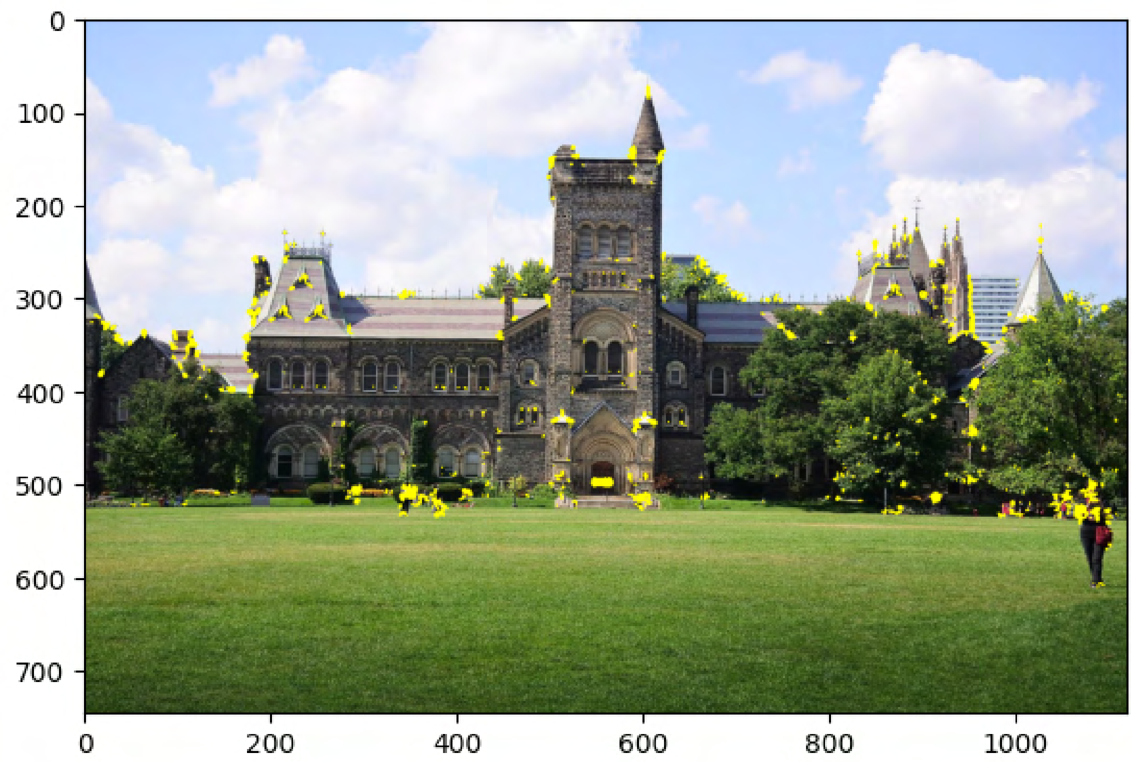


Step 3: Pick threshold and detect corners.

For image1, use threshold of 6e5.



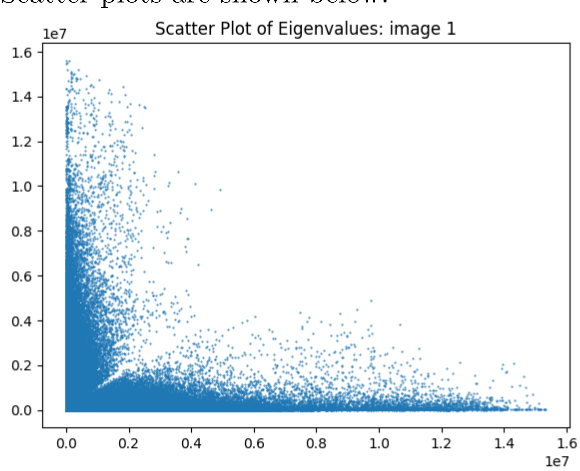
For image2, use threshold of 8e5.

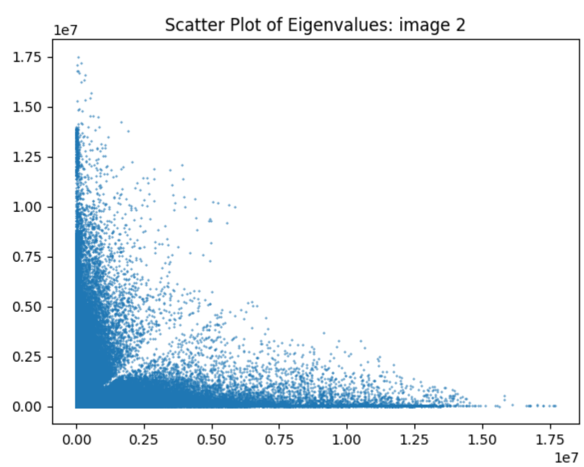


Step 4: Different Gaussian kernel.

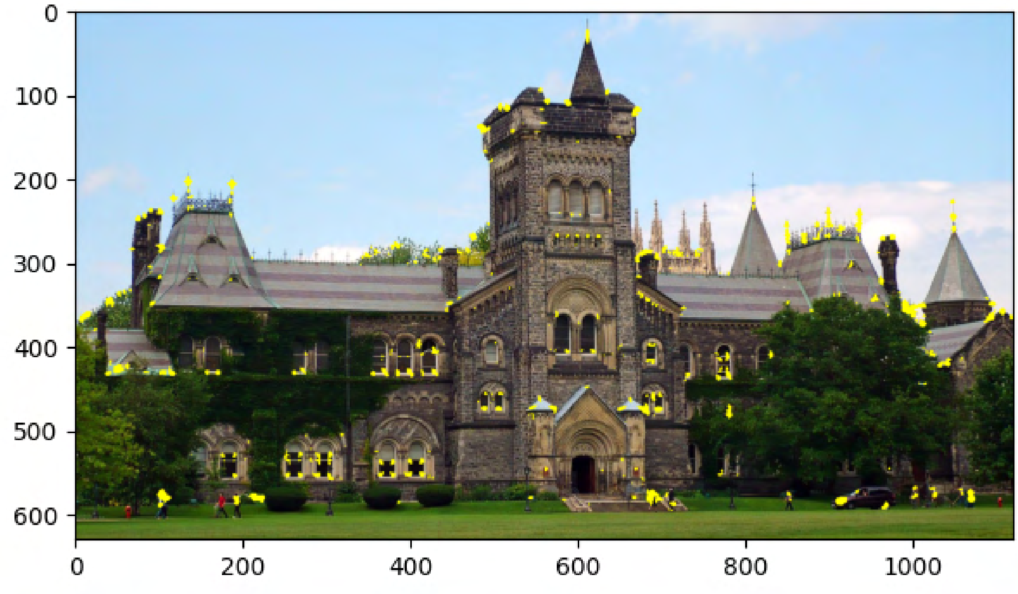
Instead of $\sigma = 2$, this time we use sigma of 30 for the Gaussian kernel.

Scatter plots are shown below:

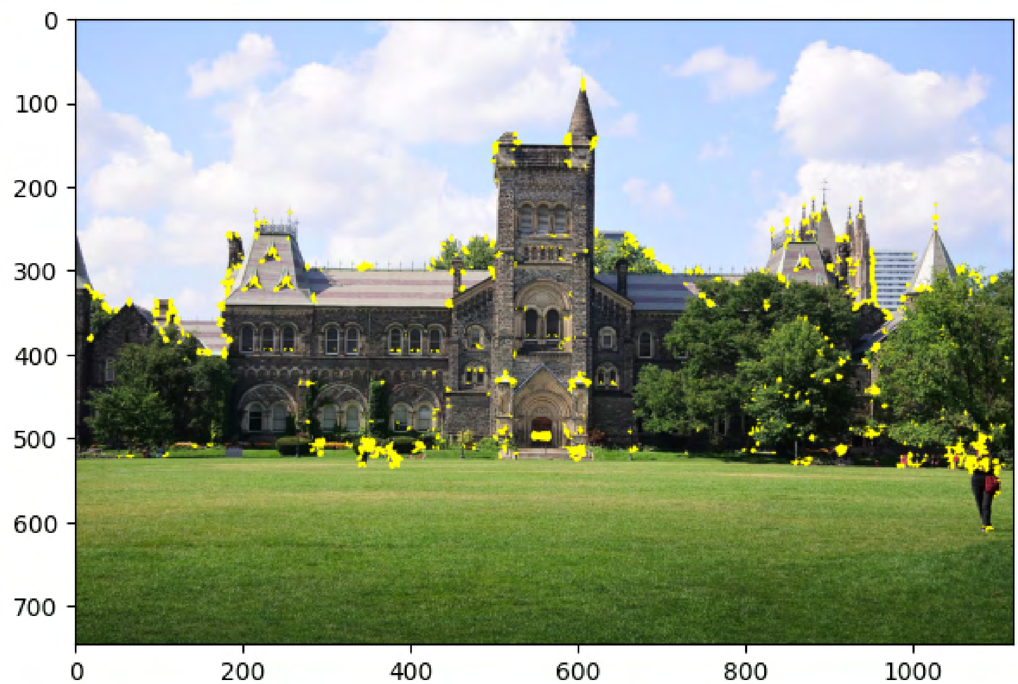




For image1, same threshold of 6e5.



For image2, same threshold of 8e5.



Comment:

- Change sigma to 30, which is 15 times of its original value of 2.
- For each image, we observe more corners are detected and detection areas are larger.
- Eigenvalues in the scatter plot are more spread out with this larger sigma value.
- Since eigenvalues determine whether a region is an edge, a corner or flat, with a larger sigma, more value points are classified as "corner", using the same threshold.