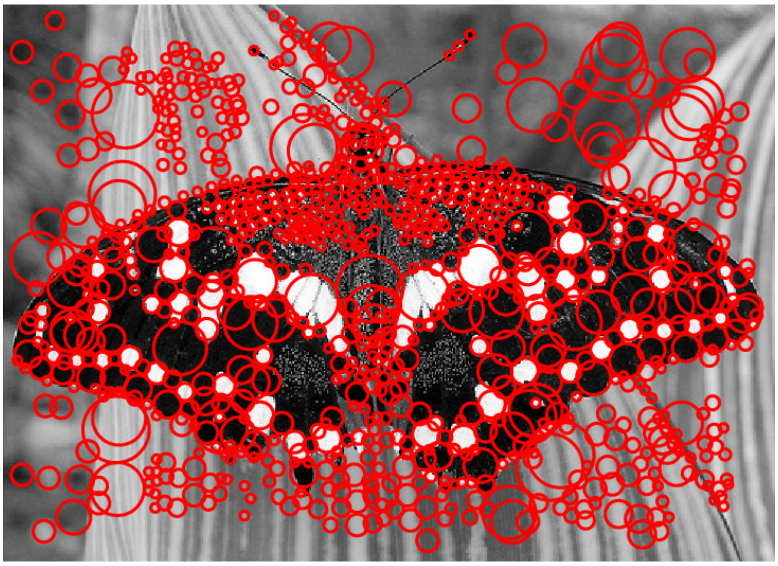
Assignment 1 (Due: Apr. 18, 2021)

1. (Math) In our lectures, we mentioned that matrices that can represent isometries can form a group. Specifically, in 3D space, the set comprising matrices $\{\mathbf{M}_i\}$ is actually a group, where
$\mathbf{M}_i = \begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{4\times4}$, $\mathbf{R}_i \in \mathbb{R}^{3\times3}$ is an orthonormal matrix, $\det(\mathbf{R}_i)=1$, and $\mathbf{t}_i \in \mathbb{R}^{3\times1}$ is a vector.
Please prove that the set $\{\mathbf{M}_i\}$ forms a group.
Hint: You need to prove that $\{\mathbf{M}_i\}$ satisfies the four properties of a group, i.e., closure, associativity, existence of identity element, and existence of inverse element for each group element.

2. (Math) Gaussian function is
$$G(x,y;\sigma) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$
The scale-normalized Laplacian of Gaussian (LOG) is
$$LoG = \sigma^2\nabla^2 G$$
Please verify that Difference of Gaussian (DOG)
$$DoG = G(x,y;k\sigma) - G(x,y;\sigma)$$
can be a good approximation of LoG.

3. (Math) In the lecture, we talked about the least square method to solve an over-determined linear system $A\mathbf{x}=b, A \in \mathbb{R}^{m\times n}, \mathbf{x} \in \mathbb{R}^{n\times1}, m>n, rank(A)=n$. The closed form solution is $\mathbf{x} = \left(A^T A\right)^{-1} A^T b$. Try to prove that $A^T A$ is non-singular (or in other words, it is invertible).

4. (Programming) Get two images, taken from the same scene but with scale transformations. Detect the scale invariant points on the two images. You can use the center of the circle to indicate the spatial position of the point and use the radius of the circle to indicate the characteristic scale of the point, just like the following example.



5. (Programming) Get two images I_1 and I_2 of our campus and make sure that the major parts of I_1 and I_2 are from the same physical plane. Stitch I_1 and I_2 together to get a panorama view using scale-normalized LoG (or DoG) based interest point detector and SIFT descriptor. You can use OpenCV or VLFeat.