Homework #2

Submissions: (1) PDF version of this file

(2) .ipynb file; template in the link below

COLAB TEMPLATE:

https://colab.research.google.com/drive/1vXUwb4AcX3vDvUTgzqP_1ag3aSHq7OvT?usp=sharing

(Only problem 1-3 will be graded)

Use these commands in colab to download the images.

```
!wget https://drive.google.com/uc?id=1o0UMPTyUFzX9CaQp-BwYXgkCho1Zo6yL  -O
kitty55.png
!wget https://drive.google.com/uc?id=1_-_yN30miNhzR9ZC5DHTiljH6LVq4hZz  -O
clean_cat.png
!wget https://drive.google.com/uc?id=1LEwFRI2vjSqQEd68lYwyuJ4JyJYPQX2m  -O
blurry_noisy_cat.png
!wget https://drive.google.com/uc?id=1lnO_PK81O54bLBUprBlo8-2x9smWYDcP  -O
moodeng.jpg
!wget https://drive.google.com/uc?id=1L5UFU_RyD5BAapUBc2HGl-RKtfgVGQKr  -O
moodeng2.jpg
```

1. (2 points) Apply Gaussian low pass filter in frequency domain on "Kitty55.png" image which has $M\ x\ N$ pixels. Find the minimum cutoff frequency (C) that still maintain the total image power P_T more than 99%. Where the total image power, P_T is calculated by summing the components of spectrum power at each point (u, v), for $u = 1, 2, ...\ M - 1$ and $v = 1, 2, ...,\ N - 1$

$$P_T = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} P(u, v)$$

$P(u, v)$ is the spectrum power provided in the lecture slides

α percent of the image power can be calculated from 100 x P_{T_f} / P_{T_org}, where P_{T_org} is the total image power of the original image and P_{T_f} of the filtered image

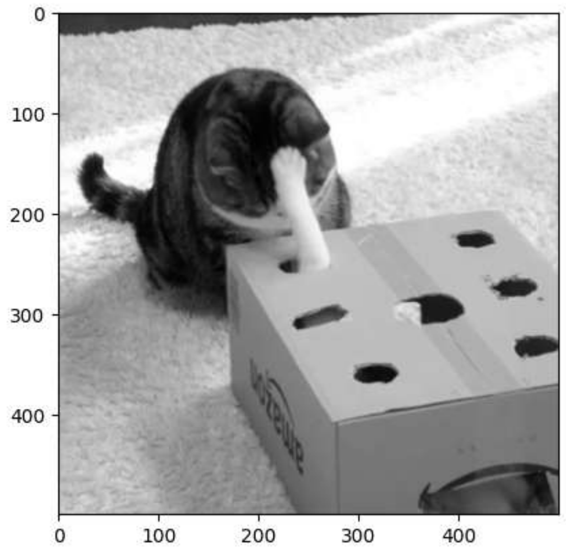
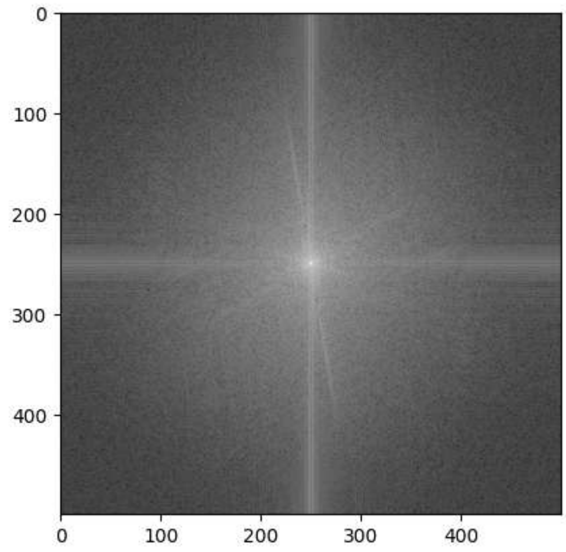
Put your results in the blank box below

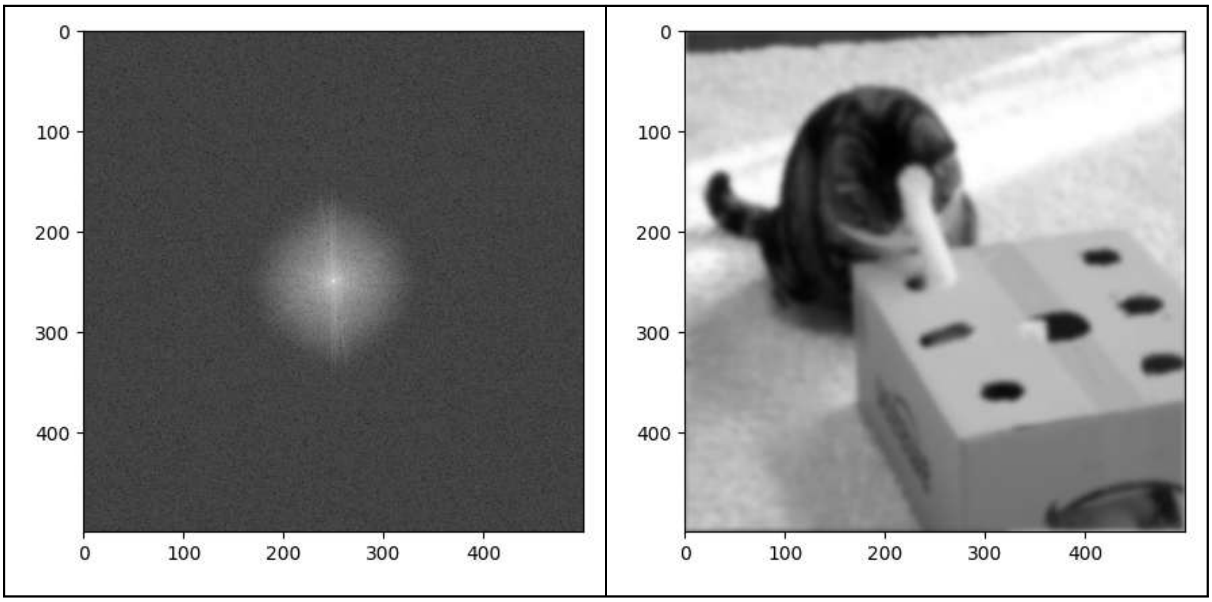
Cutoff frequency (C) =

28.135

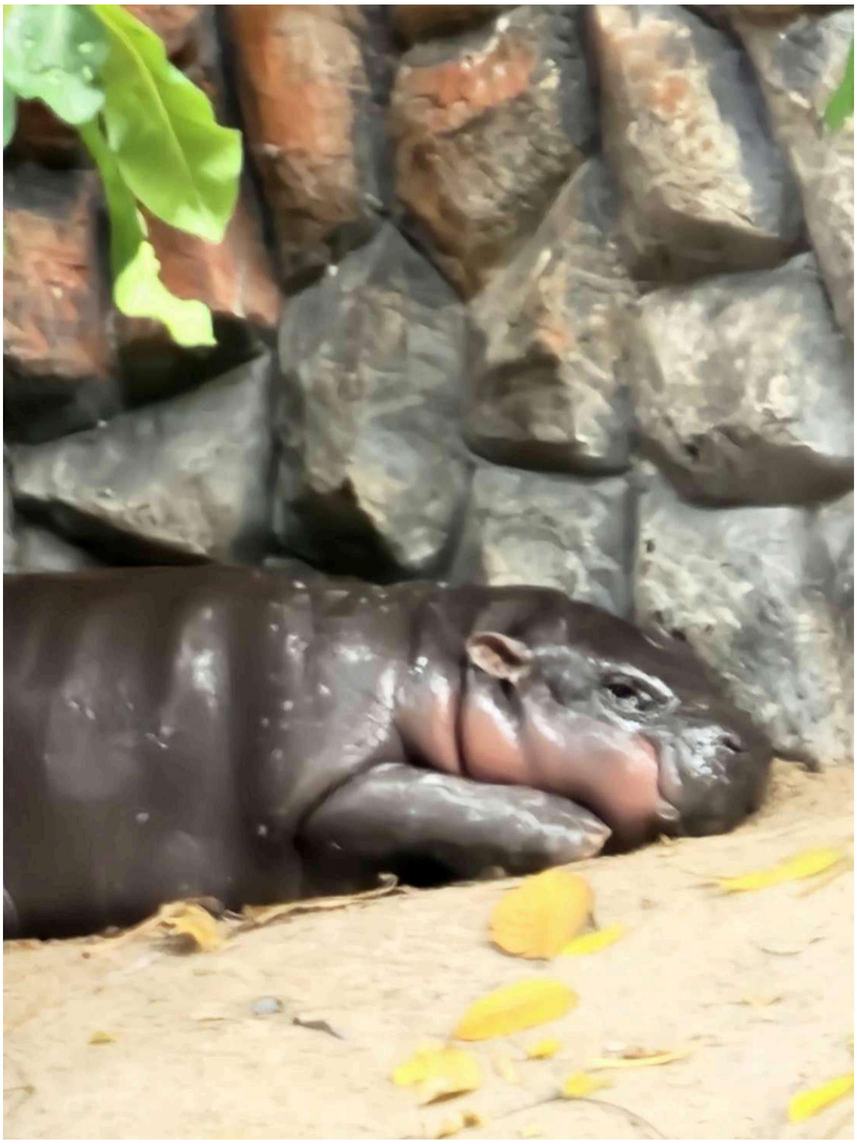
α =

0.9900000

Original Image ("kitty55.png")	Fourier Spectrum of the original image
	
Fourier Spectrum of the filtered image	Filtered images (P_T > 99%)

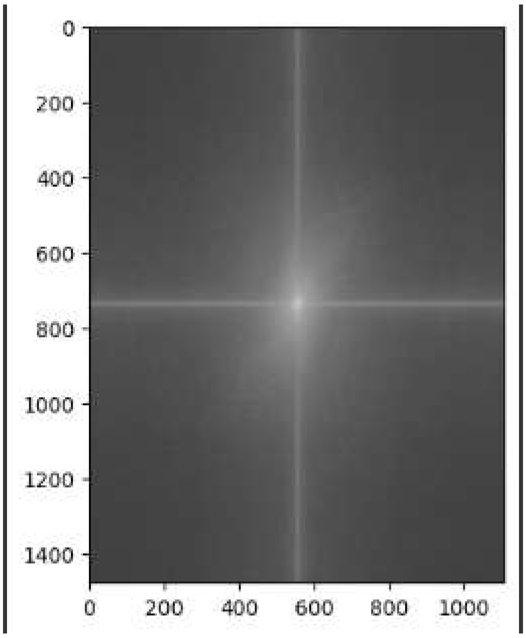


2. Using only the concepts and techniques from Lecture 4-6, analyze the given image (moodeng2.jpg) and describe how it can be improved. Show your improvement process step by step, applying the appropriate methods discussed in these lectures. For each step, explain what you are doing and why this step helps enhance the image, and discuss the changes observed in the image after each step.



Moodeng2.jpg

จากรูปภาพสามารถรู้สึกได้ว่ามีจุดขาวที่ติตาพอสมควรทั้งบนตัวหมูเด้งและหิน มองว่าเป็น S&P noise ดังนั้นจึงจะมีการใช้ median filter นอกจากนี้แล้วยังรู้สึกภาพนี้ค่อนข้างเบลอ จึงว่าแผนจะทำการ sharpening ด้วย high pass filter

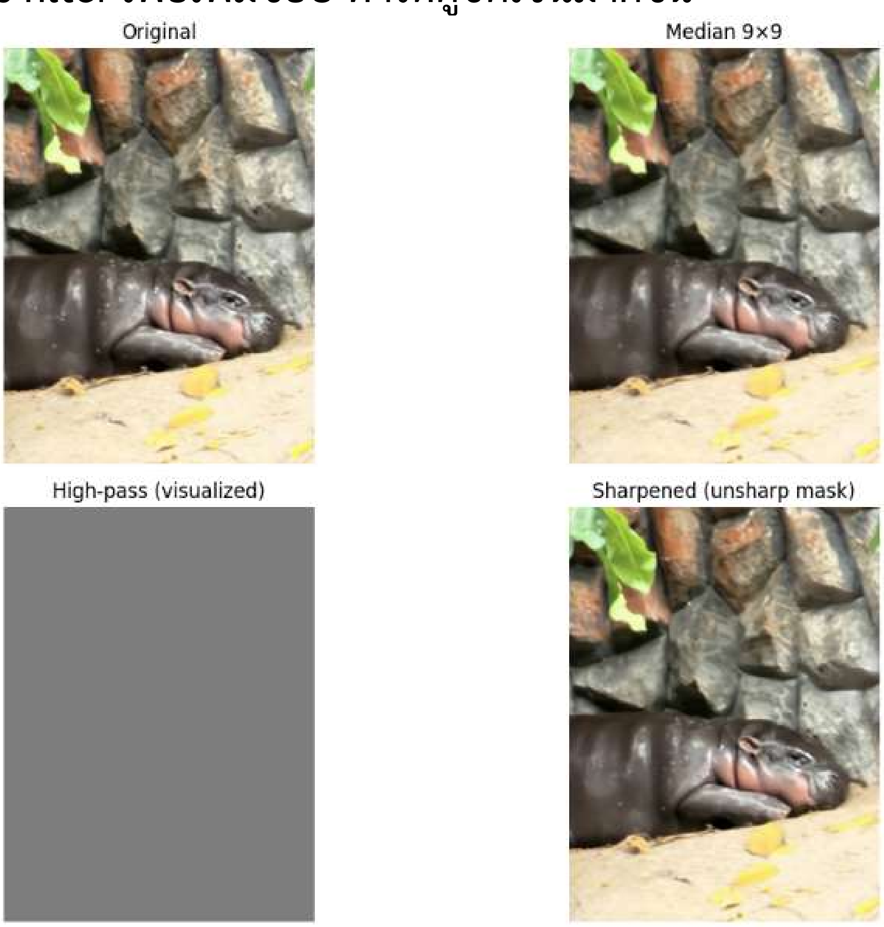


จาก Freq domain เห็นว่า ไม่ได้มี Strong periodic spike อะไร ดังนั้นขั้นตอนอย่าง notch/band reject จึงไม่ได้จำเป็น

แผนที่วางไว้จึงเป็นประมาณนี้ median filter -> high pass filter

Median filter เพื่อลดจุดสว่าง

High pass filter เพื่อเพิ่มขอบ ทำให้ดูชัดเจนมากขึ้น



3. Write a paragraph (300-400 words) to summarize three image applications using frequency analysis in your own words and add the references trustworthy references from sources like IEEE, Elsevier, Springer, ACM, and other reputable academic publishers. (excluded from the total word count).

Frequency analysis is crucial for image processing by transforming data into special type of data where information can be more efficiently represented, analyzed, and manipulated.

One well-known application is image watermarking, where hidden information is embedded within the frequency coefficients of an image to achieve both imperceptibility and robustness. By placing watermarks in mid- or low-frequency bands using methods such as DFT, DCT, or wavelet, the embedded data becomes resistant to common distortions such as compression or noise while remaining invisible to human. This has made frequency-domain watermarking an essential tool for authentication.

A second important application is feature extraction and saliency detection, which analyzing localized frequency features to identify the most informative parts of an image. Techniques such as Gabor and log-Gabor filtering capture textures and edges at different orientations and scales, producing more stable descriptors under variations in lighting, rotation, and scale

Finally, compression and efficient representation is another application where frequency analysis is widely used. JPEG uses block-based DCT to represent image blocks in frequency domain. Low frequencies (which carry most visual content) are preserved with high precision; high frequencies are quantized more aggressively or dropped. This yields lossy compression with limited perceptual degradation. Wavelet transforms (such as in JPEG2000) provide multi-resolution decomposition; they capture both approximate (low frequency) and detail (high frequency) components, enabling progressive decoding and better preservation of edges.

References
M. Barni, F. Bartolini, and A. Piva, *“Improved Wavelet-Based Watermarking Through Pixel-Wise Masking,”* **IEEE Transactions on Image Processing**, vol. 10, no. 5, pp. 783–791, May 2001.

A. K. Jain and F. Farrokhnia, *“Unsupervised texture segmentation using Gabor filters,”* **IEEE Transactions on Pattern Analysis and Machine Intelligence**, vol. 12, no. 5, pp. 561–573, May 1991.

A. Skodras, C. Christopoulos, and T. Ebrahimi, *“The JPEG 2000 Still Image Compression Standard,”* **IEEE Signal Processing Magazine**, vol. 18, no. 5, pp. 36–58, Sept. 2001.

Practice (Do not submit)

4. Problem: Restoring blurry and noisy image

You are provided with a blurred and noisy of cat image blurry_noisy_cat.png and a clean reference image clean_cat.png. You have to restore the image using the **Wiener Filter** and compute the **Structural Similarity Index (SSIM)** between the restored image and the clean reference image.

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

with:

- μ_x the average of x;
- μ_y the average of y;
- σ_x^2 the variance of x;
- σ_y^2 the variance of y;
- σ_{xy} the covariance of x and y;
- $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ two variables to stabilize the division with weak denominator;
- L the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel} - 1$);
- $k_1 = 0.01$ and $k_2 = 0.03$ by default.

SSIM Overview:

The **Structural Similarity Index (SSIM)** is a perceptual metric that quantifies the similarity between two images. It considers changes in luminance, contrast, and structure to measure how close the restored image is to the original clean image. SSIM values range from -1 to 1, where:

- 1 indicates perfect similarity.
- 0 indicates no similarity.
- Negative values indicate dissimilarity.

(Hint: this metric is available inside skimage.metrics)

Blurry Noisy Cat

Clean Cat Image

Show how to restore the image using Wiener Filter from noise and blur effects and display the result (don't worry if the output is not perfect, just select the best one in your thought),

Structural Similarity Index (SSIM) of the restored image

5. **WAVELETS AND MULTIRESOLUTION PROCESSING**

5.1 You are provided with Nong Moo Deng image moodeng.jpg. Your task is to perform **multi-level wavelet decomposition** using the **Discrete Wavelet Transform (DWT)** and analyze the different layers of decomposition. The DWT breaks the image into four sub-bands: Approximation, Horizontal details, Vertical details, and Diagonal details. You will progressively decompose the image into three levels, visualizing the components at each level.

Your Wavelet Decomposition Level & Thresholding :

Original Image **Level J**	Horizontal Detail Image	Vertical Detail Image	Diagonal Detail Image

Level J-1			
Level J-2			
Level J-3			

5.2 For level J-1, remove the right half of the three components - horizontal, vertical and diagonal details - of the ‘moodeng’ image and then apply inverse Discrete Wavelet Transform to reconstruct the original-sized image with a blur on the right half. You can blur the right side of the image using a grayscale image, but you will receive a 1-point bonus for producing an RGB output.

Explain your method and provide the output results