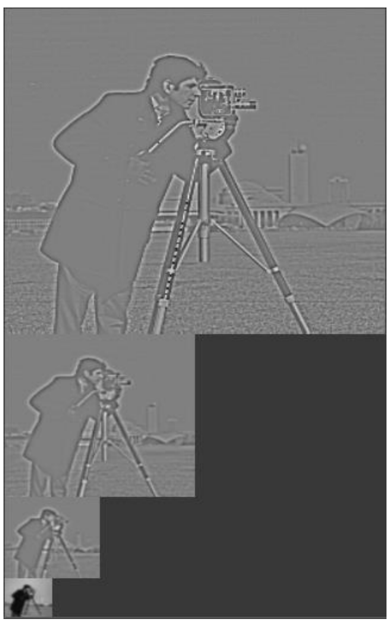
COMP90086 – Computer Vision – Assignment 1

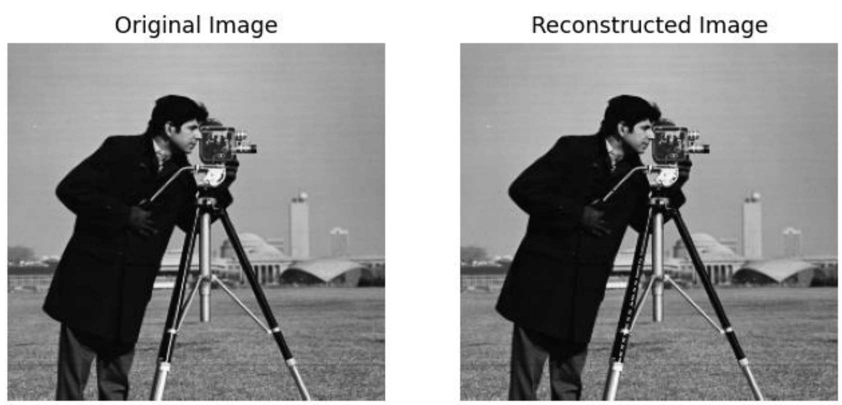
Arya Araban – 1439683 – aaraban@student.unimelb.edu.au

---

1. Laplacian image pyramid:

I have implemented functions to construct the Laplacian Pyramid of an image and to reconstruct the image from its Laplacian pyramid, based on the provided pseudocode. For example, the following figures show the four-level Laplacian Pyramid of an image, along with a comparison of the original image and the image reconstructed from this pyramid.





As shown in the above figure, the original image and the reconstructed image are highly identical.

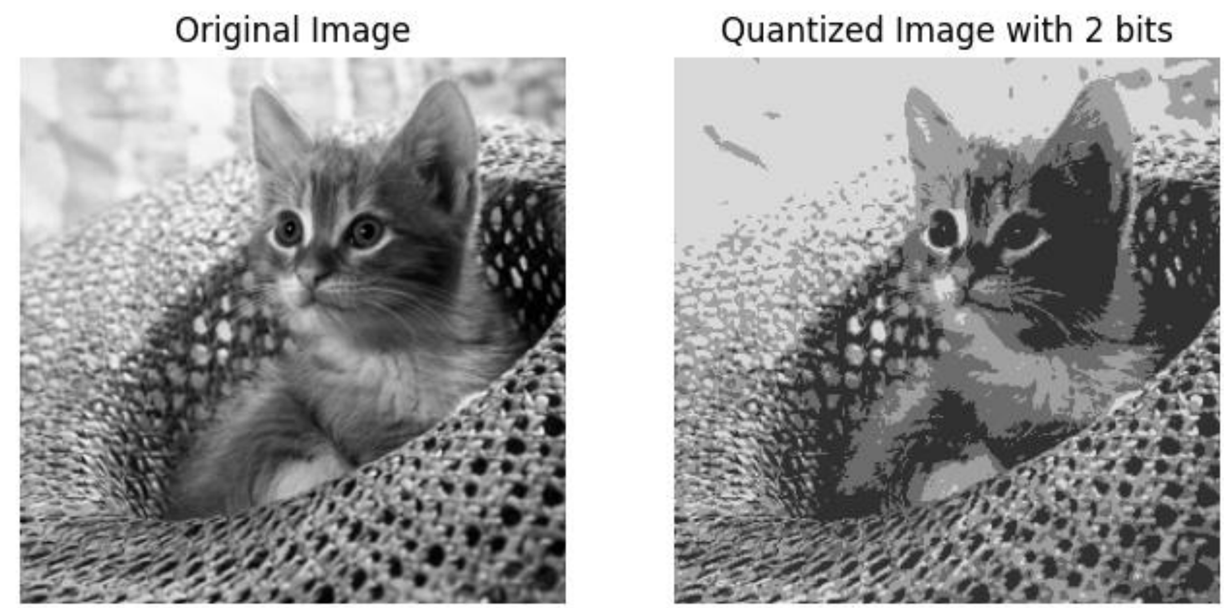
2. Quantization and compression:

I have developed a quantization function that adheres to the given specifications, ensuring that the number of pixels per bin is roughly equal across all bins:

```
Pixels per bin:
Bin 0: 16247
Bin 1: 16451
Bin 2: 16247
Bin 3: 16591
```

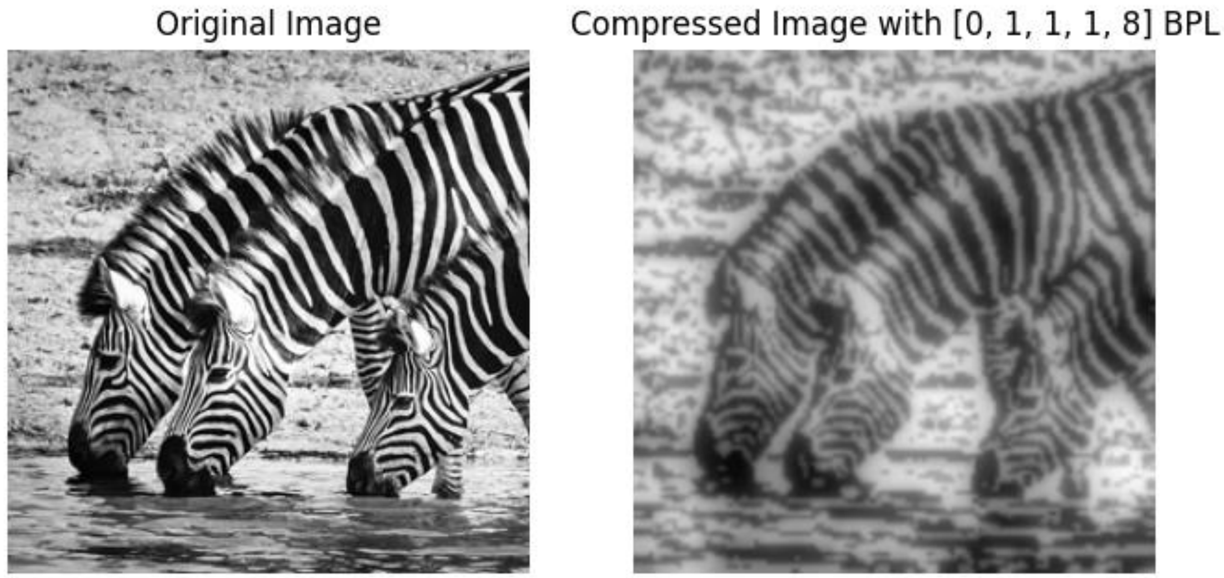
Followed by replacing all pixels within a given bin with the mean pixel value in that bin.

As an example, we can see the output of quantizing an image with 2 bits, (Note that this corresponds to quantizing the image to 4 levels).

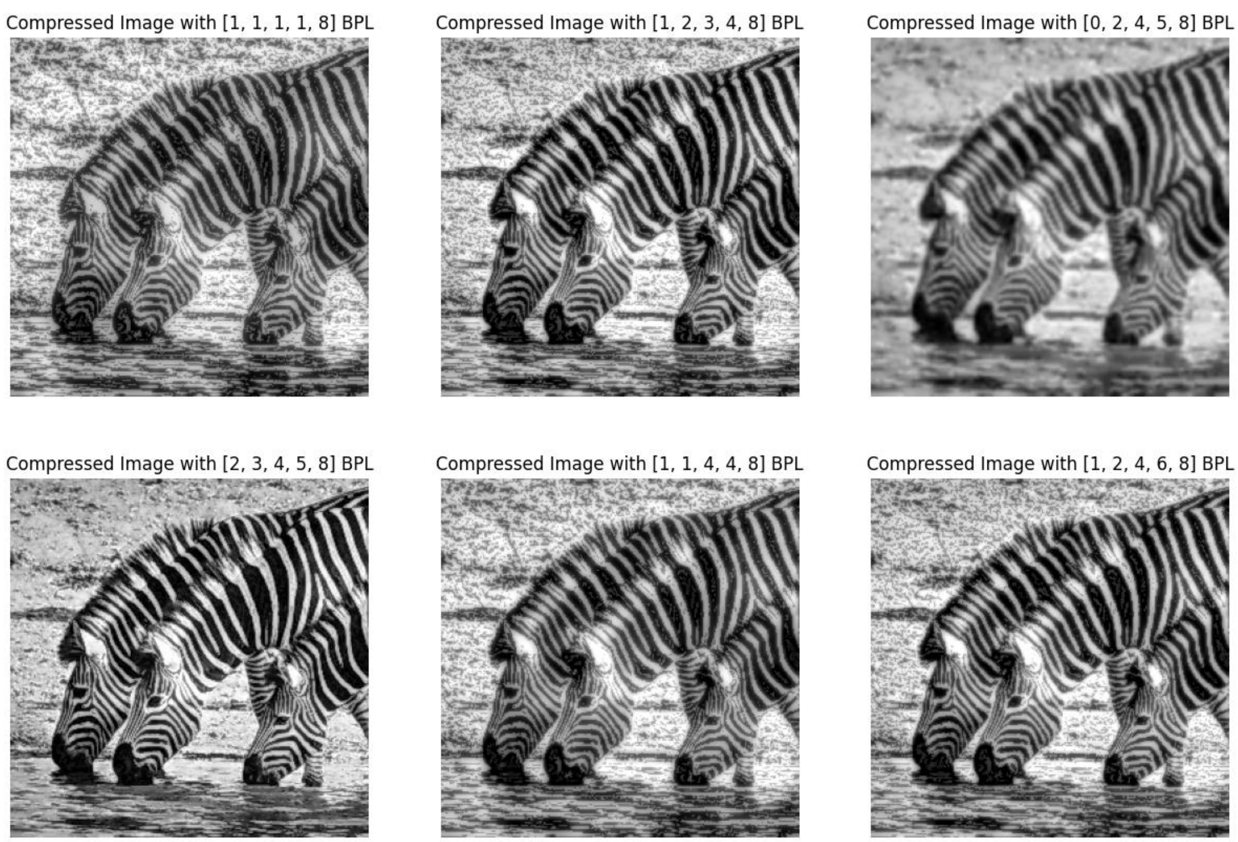


Following this, I created a function that compresses an image by generating its Laplacian pyramid and quantizing each level with a specified number of bits, except for the final level where we always keep unchanged, as specified in the instructions of the assignment specification.

As an example, we can observe the outcome of compressing an image using a five-level Laplacian pyramid. In this compression process, we allocate the following number of bits for each level: [0, 1, 1, 1, 8] (read from left to right).



We now try various bit value combinations for compressing this image using its five-level Laplacian pyramid:



The upper-left compression, where we allocate only 1 bit for each level of the pyramid except the final one, seems to result in a lower quality image. This can be observed from the grainy appearance of the image and the lack of smoothness in the edges.

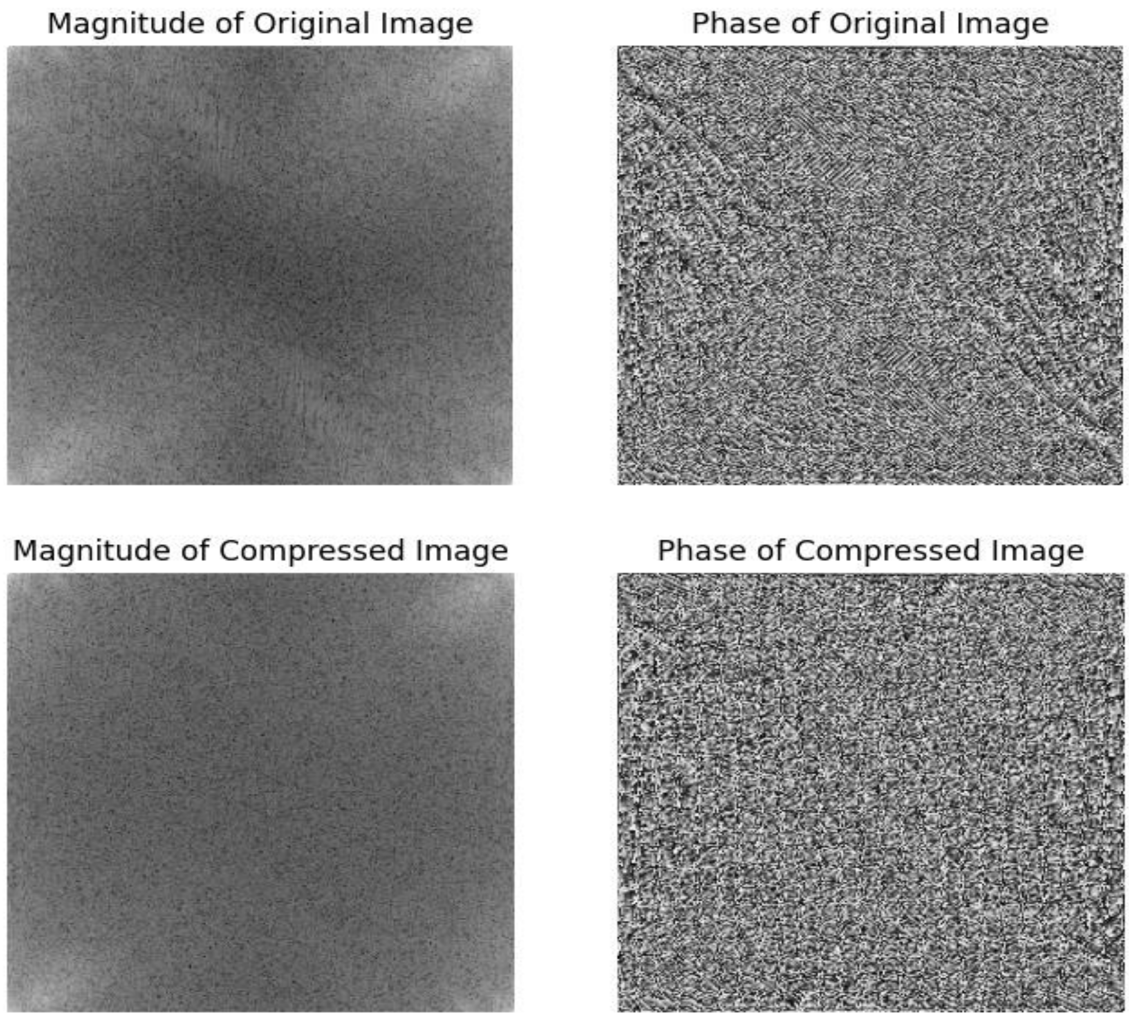
The upper-middle compression, with bit allocations of [1, 2, 3, 4, 8] per level, closely resembles the original image. Despite using significantly fewer bits than the original image, it achieves a high level of similarity. To put it into perspective, if we consider the original image to be 256x256 pixels, this compressed image would utilize only 116,736 bits, which corresponds to approximately 22% of the total bits in the original image.

In the upper-right compression, where no bits are allocated to the initial level of the pyramid, we observe that the resulting image exhibits a high level of noise and lacks fine details. Therefore, this compression approach is deemed undesirable.

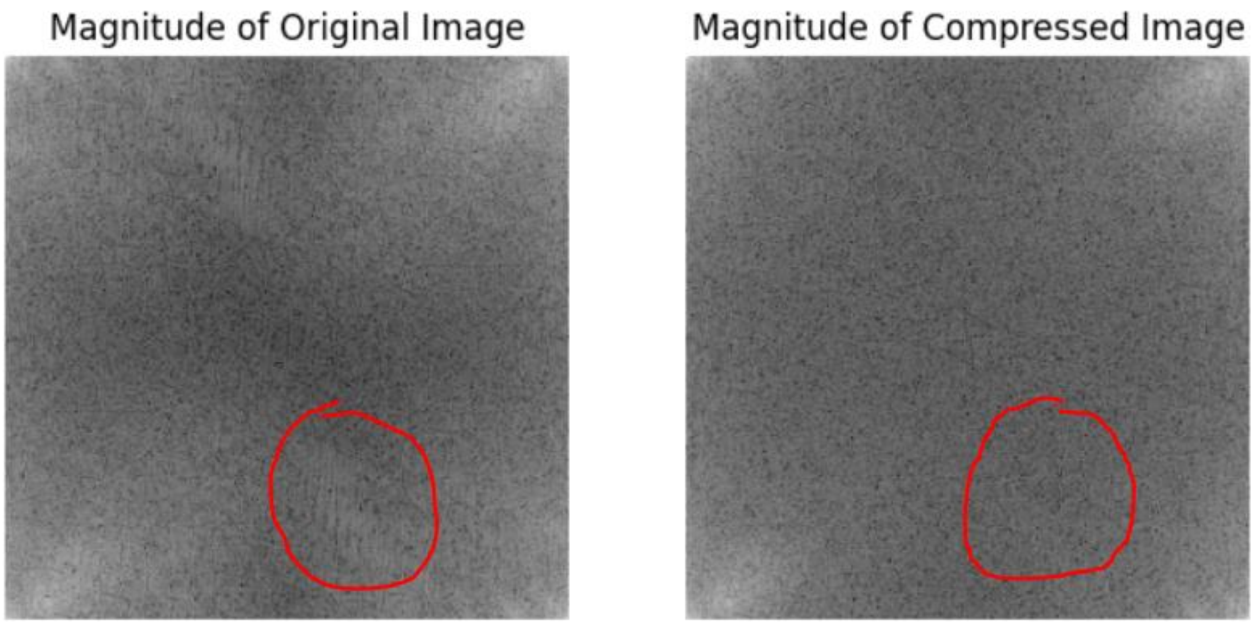
Among the final three compressed images, all show favorable outcomes. Notably, the bottom-left compression exhibits the cleanest appearance. However, due to its utilization of a significantly higher number of bits compared to the upper-middle compression, we will opt to analyze the upper-middle compression in the subsequent analysis.

3. Evaluation in the frequency domain:

After applying Fourier transformations to both the original Zebra image and its corresponding compressed image, we obtain the magnitude and phase components for each of them as follows:



When it comes to visually comparing the phase components of images, it can be challenging to interpret their significance. The phase information of images is generally difficult to comprehend due to the abstract nature of it. However, by examining the magnitude components (transformed into logarithmic space for visualization purposes), we can observe notable differences resulting from the image compression process.



Specifically, in the highlighted area of interest, we can observe that the original image exhibits intricate details. However, in the corresponding region of the compressed image, the level of detail appears to be diminished. The compressed image appears "smoothed out" in comparison, lacking the finer details that are present in the original image.

This discrepancy observed in the magnitude components suggests that the compression of the image has led to a degradation of fine details, consequently affecting the visual quality of the compressed image in specific regions. It is important to note that as we saw, this impact is less discernible in the spatial domain compared to the frequency domain.