

# Homework 2 - Applied Computer Vision

kipngeno koech - bkoech

February 4, 2025

## 1 Question 1 - Counting the number of objects in an image

To be filled

## 2 Question 2 - Finding the lines on a chessboard and Straightening the image

To be filled

## 3 Question 3 - Image Compression with Fourier Transform

There are two dimensions in which we can use to describe our image: spatial & frequency.

The spatial domain is the image itself, where we can see the pixel values of the image. The frequency domain is the representation of the image in terms of the frequency components that make up the image.

In the frequency domain, we can see the different frequencies that make up the image. Every pixel is a combination of different frequencies. The Frequency representation is a 2D array of complex numbers. The magnitude of the complex number represents the amplitude of the frequency and the angle represents the phase of the frequency.

### 3.1 Part 1 - Fourier Transform

The Fourier Transform is a mathematical tool that allows us to convert an image from the spatial domain to the frequency domain. In Python, we can use the 'numpy' library to perform a Fourier Transform on an image. Here is an example:

```
import numpy as np
import cv2

# Load the image
image = cv2.imread('image.jpg', 0)
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
fft = np.fft.fft2(gray)
fshift = np.fft.fftshift(fft)
magnitude_spectrum = 20*np.log(np.abs(fshift))
```

In the code above, we first load the image using OpenCV and convert it to grayscale. We then perform a Fourier Transform on the image using the 'np.fft.fft2' function. We then shift the zero frequency component to the center using the 'np.fft.fftshift' function. Finally, we calculate the magnitude spectrum of the image using the formula  $20 \times \log(\text{abs}(fshift))$ .

Why do we need to shift the zero frequency component to the center? The Fourier Transform of an image is usually centered at the top-left corner. Shifting the zero frequency component to the center makes it easier to visualize the frequency components of the image.

Why do we take the log of the magnitude spectrum? The magnitude spectrum is usually very large, and taking the log of it makes it easier to visualize the frequency components of the image. And also this changes from complex to real numbers.

### 3.2 Part 2 - Image Compression

Image compression is the process of reducing the size of an image file without significantly affecting its quality. One way to compress an image is to remove high-frequency components from the image. High-frequency components are usually noise or details that are not essential to the image.

So for now, we are operating in the frequency domain of an image, that means we have different frequencies that make up the image. The high-frequency components are usually located at the edges of the image, where there are sharp transitions in pixel values. By removing these high-frequency components, we can reduce the size of the image file:

```
fSorted = np.sort(np.abs(fft).flatten())
thresh = fSorted[int(np.floor((1-compress_ratio)*len(fSorted)))]
mask = np.abs(fft) > thresh
filtered = fft * mask
```

## 4 Question 4 - Data Augmentation

To be filled