EE 610: Image Processing

Assignment 2

Harsh Chaurasia
20B030019

Fall 2023

September 29, 2023

Problem 1

You are given the following kernel and image:

$$w = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(a) Write the matrices that will be used in computation, when the kernel w is centered at the point $(2, 3)$ of the image f. Note: the left most entry in f corresponds to $(0, 0)$. Give the output of convolution corresponding to this point.
(b) Give the convolution output for $f * w$, assuming zero-padding.

Solution.
(a) When w is centered at $(2, 3)$ of the matrix f, the following portion of f will be used for element-wise multiplication with w in order to get the final output of convolution:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Upon taking element-wise product with w, we will get

$$1 \times 1 + 2 \times 0 + 1 \times 0 + 2 \times 1 + 4 \times 0 + 2 \times 0 + 1 \times 1 + 2 \times 0 + 1 \times 0 = 4$$

as the result.
(b) For zero padding, we will require to add two additional rows and columns symmetrically around the image matrix, f. After convolving with w, we get the following matrix:

$$f * w = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 \\ 0 & 3 & 6 & 3 & 0 \\ 0 & 4 & 8 & 4 & 0 \\ 0 & 3 & 6 & 3 & 0 \\ 0 & 1 & 2 & 1 & 0 \end{bmatrix}$$

Problem 2

In the original image used to generate the three blurred images shown, the vertical bars are 5 pixels wide, 100 pixels high, and their separation is 20 pixels. The image

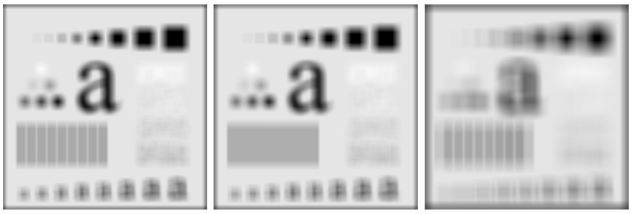


Figure 1: Two images with similar histogram

was blurred using square box kernels of sizes 23, 25, and 45 elements on the side, respectively. The vertical bars on the left, lower part of (a) and (c) are blurred, but a clear separation exists between them. However, the bars have merged in image (b), despite the fact that the kernel used to generate this image is much smaller than the kernel that produced image (c). Explain the reason for this.

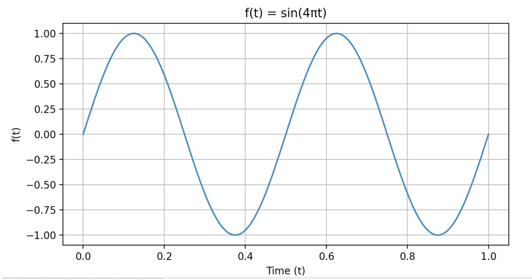
Solution.

The reason for this is quite interesting and depends on the number of pixels that make up the width of the bar and the separation between the bars.
We start by looking at the fact that the sum of the number of pixels making the width of one bar and the separation between two adjacent bars is $5 + 20 = 25$. Now, consider the kernel with size 25. When it moves one pixel to the right, it loses one column of a vertical bar. But at the same time, it picks up one column of the next vertical bar. Moreover, this always is the case as long as this kernel remains within the block of the vertical bars in the image. This leads to a uniform input to the kernel, leading to a uniformly blurred portion of the image without any separation between them.
On the other hand, this is not the case with the other kernel sizes given, i.e., 23 or 45, which leads to the clear separation between the bars being still present in the output image.

Problem 3

Consider the function $f(t) = \sin(2\pi nt)$, where n is an integer. Its Fourier transform, $F(\mu)$, is purely imaginary. Because the transform, $\overline{F}(\mu)$, of sampled data consists of periodic copies of $F(\mu)$, it follows that $\overline{F}(\mu)$ will also be purely imaginary. Draw a time-domain plot of the signal $f(t)$ and answer the following questions.
(a) What is the period of $f(t)$?
(b) What is the frequency of $f(t)$?
(c) What would the sampled function and its Fourier transform look like in general if $f(t)$ is sampled at a rate ΔT higher than the Nyquist rate?
(d) What would the sampled function look like in general if f(t) is sampled at a rate lower than the Nyquist rate?
(e) What would the sampled function look like if $f(t)$ is sampled at the Nyquist rate, with samples taken at $t = 0, \pm\Delta T, \pm 2\Delta T, ...$?

Solution.
Here is the time-domain plot of the signal $f(t)$ for $n = 2$:



(a) The period of $f(t)$, T, can be found out by using the relation:

$$\begin{aligned} f(t) &= f(t+T) \\ \sin(2\pi nt) &= \sin(2\pi n(t+T)) \\ &= \sin(2\pi nt + 2\pi nT) \\ &= \sin(2\pi nt)\cos(2\pi nT) + \cos(2\pi nt)\sin(2\pi nT) \end{aligned}$$

For this to hold, $\sin(2\pi nT)$ needs to be 0 and hence, simultaneously, $\cos(2\pi nT)$ will become 1. This will happen at

$$T = \frac{1}{n}$$

(b) The frequency of a periodic function is the reciprocal of its period. Therefore, the frequency of $f(t)$, ω, will be given by the following equation:

$$\omega = \frac{1}{T} = n$$

(c) Sampling $f(t)$ at a rate of ΔT is done as:

$$f(t) = f(t)s_{\Delta T}(t) = \sum_{m=-\infty}^{\infty} f(t)\delta(t - n\Delta T)$$

Here, if ΔT is higher than the Nyquist rate, the samples acquired will be usable in order to extract a single unique period from the transform of the samples, $\tilde{F}(\mu)$. Moreover, this prevents *aliasing*, i.e., no high-frequency components in the signal fold into lower-frequency components due to insufficient sampling. This happens as a consequence of the *sampling theorem*. The Fourier transform will comprise of sequences of non-overlapping replicas of the original spectrum, all centred at multiples of $1/\Delta T$.

(d) If sampling is done at a rate lower than the Nyquist rate, *aliasing* will inevitably be introduced. This means that the high-frequency components occurring at half the sampling frequency will appear as low-frequency components in the sampled signal. This will lead to a misrepresentation of the original signal. In addition to loss of details, reconstructing this signal accurately from a Fourier transformation of the samples will not be possible and an entirely different signal will be recovered.

(e) In this case, if $f(t)$ is sampled at the Nyquist rate, with the samples being taken at $t = 0, \pm\Delta T, \pm 2\Delta T, \ldots$, the samples will just be a sequence of zeros. This can be seen as:

$$x[n] = \sin(2\pi nt/2t) = \sin(n\pi) = 0$$

This will not allow us to reconstruct the original sine wave using these samples. This is why the sampling frequency needs to be *higher than* the Nyquist rate.

Problem 4

The image on the left in the figure below consists of alternating stripes of black/white, each stripe being two pixels wide. The image on the right is the Fourier spectrum of the image on the left, showing the dc term and the frequency terms cor- responding to the stripes. (Remember, the spectrum is symmetric so all components, other than the dc term, appear in two symmetric locations.)
(a) Suppose that the stripes of an image of the same size are four pixels wide. Sketch what the spectrum of the image would look like, including only the dc term and the two highest-value frequency terms, which correspond to the two spikes in the spectrum above.
(b) Why are the components of the spectrum limited to the horizontal axis?
(c) What would the spectrum look like for an image of the same size but having stripes that are one pixel wide? Explain the reason for your answer.
(d) Are the dc terms in (a) and (c) the same, or are they different? Explain.

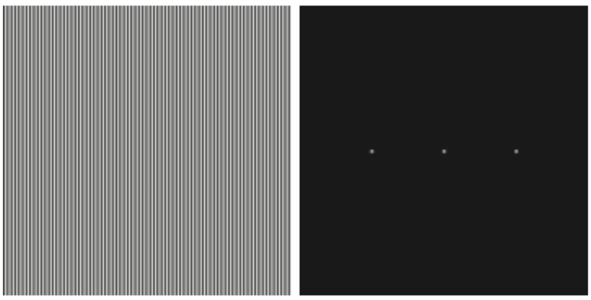


Figure 2: (Left) Image with two pixel wide stripes and (Right) corresponding Fourier spectrum

Solution.
(b) The components of the spectrum are limited to the horizontal axis. This is due to the following reasons:
1. The black and white stripes are arranged horizontally, exhibiting a distinct spatial frequency in the horizontal direction.
2. The magnitude spectrum of the Fourier transform is symmetric about the center of the frequency domain along the horizontal and vertical axes.

3. Since the alternating stripes have a prominent spatial frequency in the horizontal direction, the Fourier spectrum will show peaks along the horizontal axis in the frequency domain.
(c) Upon making the stripes one pixel wide, the spatial frequency goes up, and the Fourier image will contain additional higher-frequency components. This will also lead to more dots in the image.

Problem 5

We are interested in finding resolution of the sensor that we need in an industry. They have cloth patterns with repeating black-and-white vertical stripes, with each having a width of 2 cm. Each image captured of a cloth piece will have 250 such stripes. Ignore any other distortions caused due to image capture. You are to choose the imaging sensor so that you appropriately sample the image and capture it without any loss. What should be minimum resolution of the sensor that need to be used to capture this image? Following parts will help you answer this.
(a) What is the frequency of the vertical stripe (pair) in cycles/cm?
(b) What is the required minimum sampling frequency in cycles/cm?
(c) What is the required resolution for capturing a cloth piece?

Solution.
In order to prevent aliasing, we need to sample at a rate higher than the Nyquist rate. In this case, it corresponds to the width of a single stripe.
The total width of the pattern is: $2cm \times 250 = 500cm$.
(a) The frequency of the vertical stripe pair:

$$\begin{aligned} f &= \frac{1}{\lambda} \\ &= \frac{1}{2cm} \\ &= 0.5 \text{cycles/cm} \end{aligned}$$

(b) The minimum sampling frequency is the Nyquist rate, i.e., at least twice the highest frequency component present.

$$\begin{aligned} f_s &\geq 2 \times \text{Highest frequency} = 2 \times 0.5 \text{cycles/cm} \\ f_s &\geq 1 \text{cycle/cm} \end{aligned}$$

(c) Resolution can be calculated in pixels per cm as

$$R = \frac{\text{Total number of pixels}}{500cm}$$

Depending on the total number of pixels, we can find an appropriate resolution.

Problem 6

First-order derivatives can be approximated by the spatial differences $g_x = \frac{\partial f(x,y)}{\partial x} = f(x+1,y) - f(x,y)$ and $g_y = \frac{\partial f(x,y)}{\partial y} = f(x,y+1) - f(x,y)$
(a) Find the equivalent filter transfer functions $H_x(u,v)$ and $H_y(u,v)$ in the frequency domain.
(b) Show that these are highpass filter transfer functions.

Solution.
(a) To find the equivalent filter transfer functions $H_x(u,v)$ and $H_y(u,v)$ in the frequency domain, we use the Fourier transform. Let's start with the Fourier transform of g_x:

$$\begin{aligned} G_x(u,v) &= \mathscr{F}\{g_x\} \\ &= \mathscr{F}\{f(x+1,y) - f(x,y)\} \\ &= e^{-2\pi iu}F(u,v) - F(u,v) \\ &= F(u,v)(e^{-2\pi iu} - 1) \end{aligned}$$

This gives us

$$H_x(u,v) = e^{-2\pi iu} - 1$$

Similarly, it can be shown that

$$H_y(u,v) = e^{-2\pi iv} - 1$$

(b) A high-pass filter is one which essentially allows high-frequency components while blocking low-frequency components. Let's analyse these transfer functions to check if this is true for them.
At zero frequency, i.e., $u = 0$ and $v = 0$, both $H_x(u,v)$ and $H_y(u,v)$ become -1, which means they attenuate the DC-component.
At higher frequencies, i.e., when $u \neq 0$ and $v \neq 0$, the magnitudes of the exponentials in the transfer functions become 1, so the magnitude of the transfer functions is always greater than 1.
The observation that these transfer functions attenuate the low-frequency components and allow the high-frequency components implies that these are indeed *high-pass filters*.