Computer Vision LAB 1

Andrej Wilczek 880707-7477
Ilian Corneliussen 950418-2438

November 2019

1 Introduction

Question 1: *Repeat this exercise with the coordinates p and q set to (5, 9), (9, 5), (17, 9), (17, 121), (5, 1) and (125, 1) respectively. What do you observe?*

Answer: We can see that when we input a value of (p,q) that is farther away from the origin of the centered $\hat{F}(uc, vc)$ we get a higher frequency in the real and imaginary plots of F. If one of the coordinates for example when (uc,vc) = (4,0) we get horizontal lines corresponding only having frequency content in one dimension. We also see for the case when (uc,vc) = (-4,0) that the amplitude is the same but the phase has been shifted 90 degrees.

Question 2: *Explain how a position (p, q) in the Fourier domain will be projected as a sine wave in the spatial domain. Illustrate with a Matlab figure.*

Answer:

$$\mathrm{f}(m,n) = \frac{1}{\sqrt{M}} \sum_{u=0}^{M-1} \left(\frac{1}{\sqrt{N}} \sum_{v=0}^{N-1} \hat{\mathrm{f}}(u,v) e^{2\pi i \frac{nv}{N}} \right) e^{2\pi i \frac{mu}{M}} = (\hat{f} \text{ is a Dirac delta}) =$$

$$= \frac{1}{\sqrt{MN}} \left(cos(2\pi(\frac{nv}{N} + \frac{mu}{M})) + isin(2\pi(\frac{nv}{N} + \frac{mu}{M})) \right)$$

From the equation above we can see that the Fourier transform decomposes an image into its sine and cosine components. Where the cosine component is just a shifted sine with the phase π. As can be visualized in the Figure 1 the real(F) looks like a cosine and the imag(F) looks like a sine.

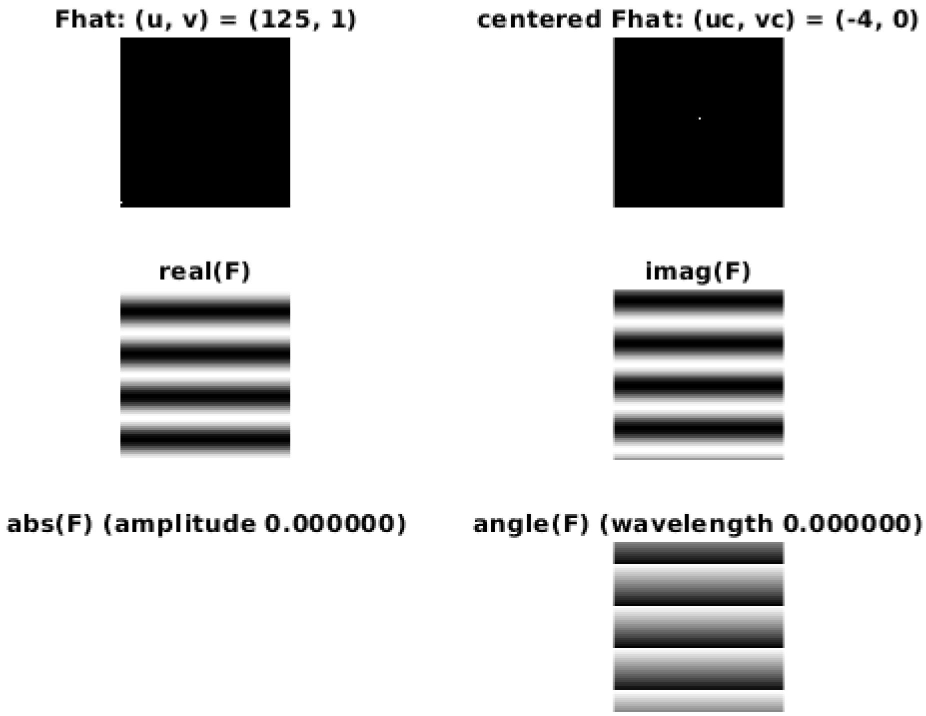

Figure 1:

Question 3: *How large is the amplitude? Write down the expression derived from Equation (4) in these notes. Complement the code (variable amplitude) accordingly.*

Answer: The amplitude can be concluded from the equation above as:

$$\text{Amplitude} = \frac{1}{\sqrt{MN}}$$

.

Question 4: *How does the direction and length of the sine wave depend on p and q? Write down the explicit expression that can be found in the lecture notes. Complement the code (variable wavelength) accordingly.*

Answer: The angular frequency, for m and n, can be written as:

$$\omega_m = \frac{2\pi u}{M}, \quad \omega_n = \frac{2\pi v}{N}$$

Where the direction of the sine wave is given by the direction of the vector,

$$\text{Direction} = \tan^{-1}\left(\frac{\omega_n}{\omega_m}\right) = \tan^{-1}\left(\frac{v}{u}\right)$$

The wavelength can be calculated as,

$$\lambda = \frac{2\pi}{\|\omega\|} = \frac{2\pi}{\sqrt{\omega_m^2 + \omega_n^2}} = \frac{N}{\sqrt{u^2 + v^2}}$$

Question 5: *What happens when we pass the point in the center and either p or q exceeds half the image size? Explain and illustrate graphically with Matlab!*

Answer: When doing fft we assume that the signal/picture to be a infinitely long singal/picture, i.e. that there are multiple copies laying next to each other. When the picture is shifted and we exceed half the window size it can be interpreted as we have enter the next picture laying next to it, or as the index $i_{new} = i_{old} - sz$, where sz is the windows size. This can be seen in Figure 2.

Question 6: *What is the purpose of the instructions following the question What is done by these instructions? in the code?*

Answer: The code is there to rescale from (u,v) to (uc,vc), i.e. putting frequency zero in the middle. In other words if for example u is in the first quadrant we subtract one from it to get zero-indexation, i.e. frewquency zero in the origin. If u is outside the first quadrant we subtract the size as well as one to get both zero-indexation and centering.

Question 7: *Why are these Fourier spectra concentrated to the borders of the images? Can you give a mathematical interpretation? Hint: think of the frequencies in the source image and consider the resulting image as a Fourier transform applied to a 2D function. It might be easier to analyze each dimension separately!*

Answer: The first two images only has frequency content in one dimension each, i.e. there is only change in the image in one direction. This results in a Fourier spectra along the border corresponding each direction. The mathematical explanation for this is that the angular frequency of one of the dimensions is equal to zero. The third image is a linear combination of the two first and since the Fourier transform is commutative the resulting spectrum is simply the sum of the two spectrum.

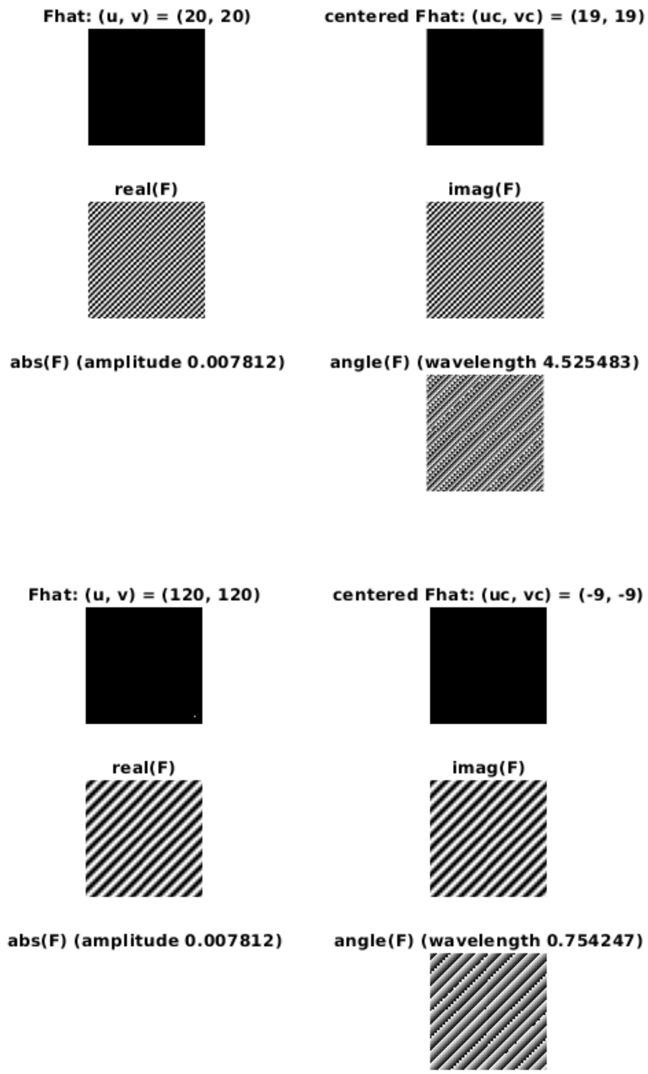

Figure 2:

Question 8: *Why is the logarithm function applied?*

Answer: The logarithm is applied to compress the gray-scale values into a smaller region so that the difference between the darkest and lightest pixels is smaller. This is because low frequencies can be very hard to see otherwise.

Question 9: *What conclusions can be drawn regarding linearity? From your observations can you derive a mathematical expression in the general case?*

Answer: As stated above (Question 7) H is a linear combination of F and G and the following holds:

$$\mathcal{F}\left[\mathrm{H}\right] = \mathcal{F}\left[\mathrm{F}(m,n) + 2\mathrm{G}(m,n)\right] = \mathcal{F}\left[\mathrm{F}(m,n)\right] + 2\mathcal{F}\left[\mathrm{G}(m,n)\right]$$

I other word the Fourier transform is a linear operation.

Question 10: *Are there any other ways to compute the last image? Remember what multiplication in Fourier domain equals to in the spatial domain! Perform these alternative computations in practice.*

Answer: A multiplication in Fourier domain is a convolution in spatial domain and the other way around therefore we can instead of doing a multiplication in spatial domain and then do the Fourier transform do a Fourier transform and then do a convolution in the Fourier domain.

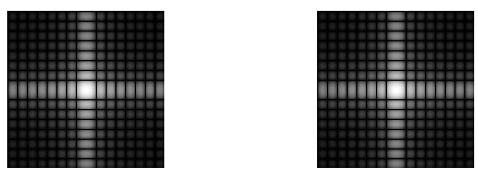

Figure 3: Original Fourier image to the left and conv2 with scaling to the right.

Question 11: *What conclusions can be drawn from comparing the results with those in the previous exercise? See how the source images have changed and analyze the effects of scaling.*

Answer: After scaling the white part of the image is stretched in one dimension and compressed in the other. This results in a lower frequency content in the stretched dimension and a higher frequency content in the compressed dimension, i.e an expansion in spatial domain is a compression in Fourier domain and vice versa. This can be seen clearly in Figure 4.

Figure 4: Original image and Fourier spectrum of image.

Question 12: *What can be said about possible similarities and differences? Hint: think of the frequencies and how they are affected by the rotation.*

Answer:
A rotation in one domain becomes a rotation in the other domain, because Fourier transform is a linear operation. If the angle of rotation is not a multiple of 90 °the poor resolution makes the edges jagged which makes the Fourier spectrum look noisy (see Figure 5).

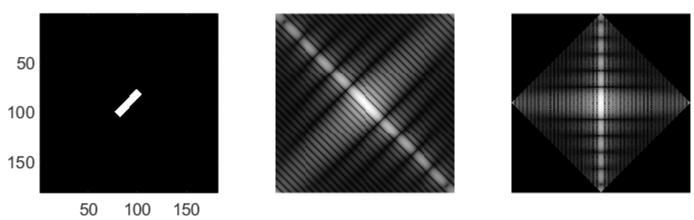

Figure 5:

Question 13: *What information is contained in the phase and in the magnitude of the Fourier transform?*

Answer:
The magnitude contains information of the amplitude of the waveforms, this corresponds to what gray-levels are on either side of an edge in the image. The Phase contains information about edges in the image, it defines how a waveform are shifted along its direction. As can be seen in Figure 6 the phase has more impact when trying to identify what is in the picture. When we change the magnitude we can still see that it is a telephone but when the phase in changed we cannot see the telephone anymore.

Figure 6: Original image, changed magnitude, changed phase.

Question 14: *Show the impulse response and variance for the above mentioned t-values. What are the variances of your discretized Gaussian kernel for t = 0.1, 0.3, 1.0, 10.0 and 100.0?*

Answer: The impulse response can be seen in Figure 7 and also the variance for each t value.

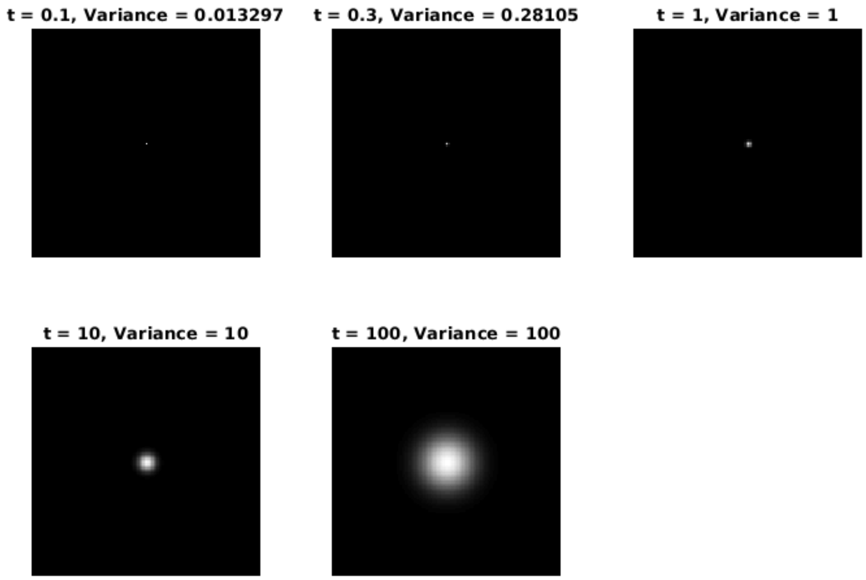

Figure 7: Impulse response for different t values.

Question 15: *Are the results different from or similar to the estimated variance? How does the result correspond to the ideal continuous case? Lead: think of the relation between spatial and Fourier domains for different values of t.*

Answer: The variance is equal to the expected variance i.e. the ideal continuous case (the t-value) in all cases when $t \geq 1$. For the two cases when $t \leq 1$ the variance differ from the expected, this is due to the filter becoming non-Gaussian for values of t ≤ 1, see Figure 8.

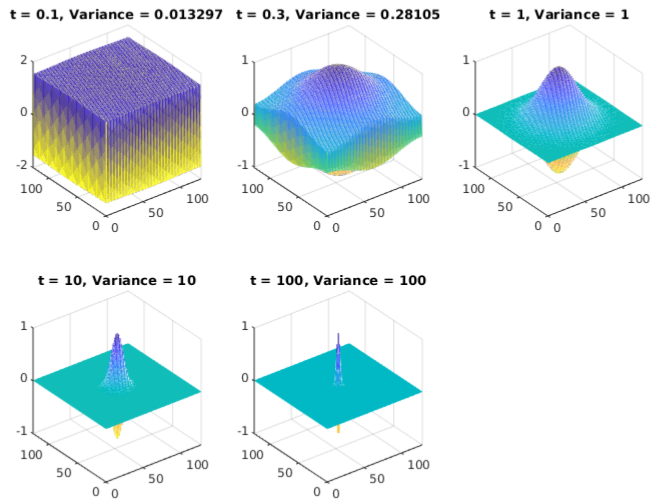

Figure 8: Fourier spectrum of G.

Question 16: *Convolve a couple of images with Gaussian functions of different variances (like t = 1.0, 4.0, 16.0, 64.0 and 256.0) and present your results. What effects can you observe?*

Answer: When applying the filter, we can see that for higher t values the picture gets more and more blurred, see Figure 9-11. This is because the filter kernel is larger for large t-values which gives more smoothing.

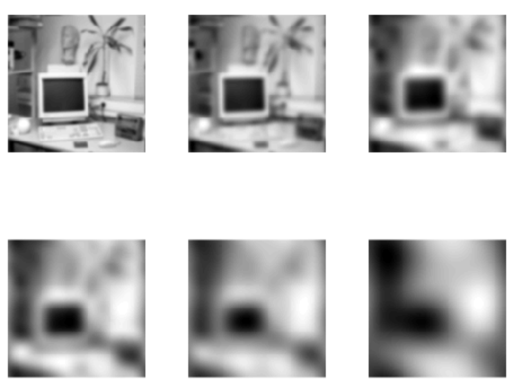

Figure 9: Hackerman.

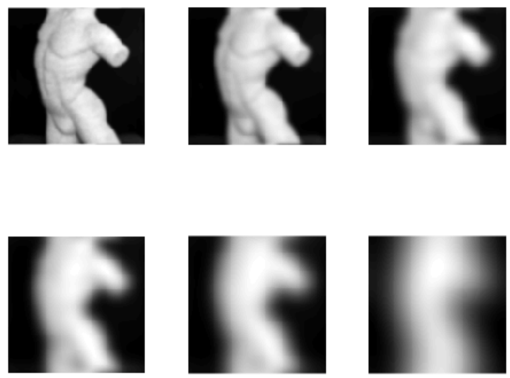

Figure 10: Maleman.

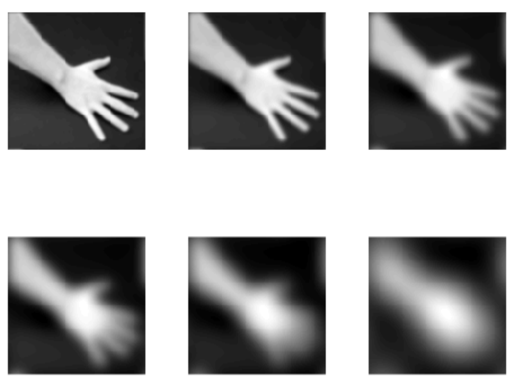

Figure 11: Manhand.

Question 17: *What are the positive and negative effects for each type of filter? Describe what you observe and name the effects that you recognize. How do the results depend on the filter parameters? Illustrate with Matlab figure(s).*

Answer: For the Gaussian filter a higher variance corresponds to a greater smoothing effect thus a too high variance means that the image is very blurry. Therefore, it's a trade-of between removing the noise and keeping the image sharp. Regarding the median filter a larger height and width removes more noise but gives an increased painting effect. The cut-off frequency of the ideal low-pass filter determines what frequencies are filter out. A low cut-off frequency results in less information being preserved in the image but with a high cut-off a lot of the noise remains in the image as noise often has high frequency content.

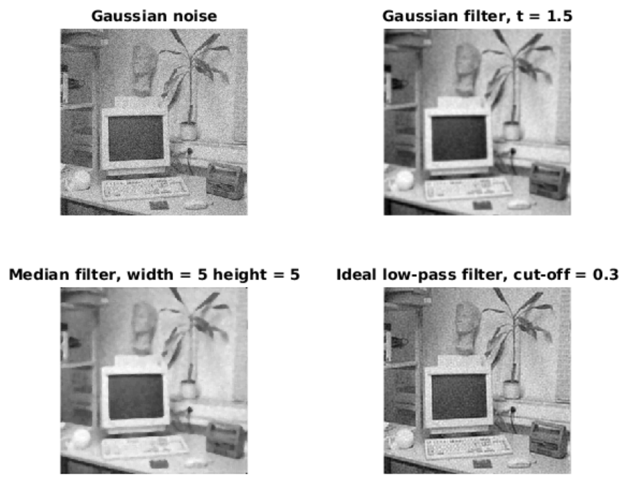

Figure 12: Gaussian noise with filters.

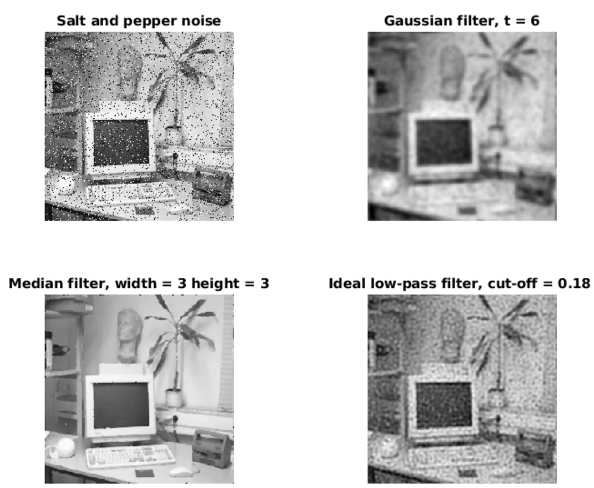

Figure 13: Sap noise with filters.

Question 18: *What conclusions can you draw from comparing the results of the respective methods?*

Answer: The Gaussian filter does mainly removes high frequency content which can make the image look blurry, because the information of edges is in the higher frequencies. The median filter preserves the edges while smoothing the image. The ideal low-pass filter cuts of higher frequencies and preserves the low frequency content entirely. As can be observed in Figure 12 and Figure 14, the Gaussian filter works best on the Gaussian noise and the median filter works best on the salt and pepper noise.

Question 19: *What effects do you observe when subsampling the original image and the smoothed variants? Illustrate both filters with the best results found for iteration $i = 4$.*

Answer: When the original image is subsampled we loose every other pixel which leaves spots without information in the image, this is equivalent with lowering the sampling frequency.

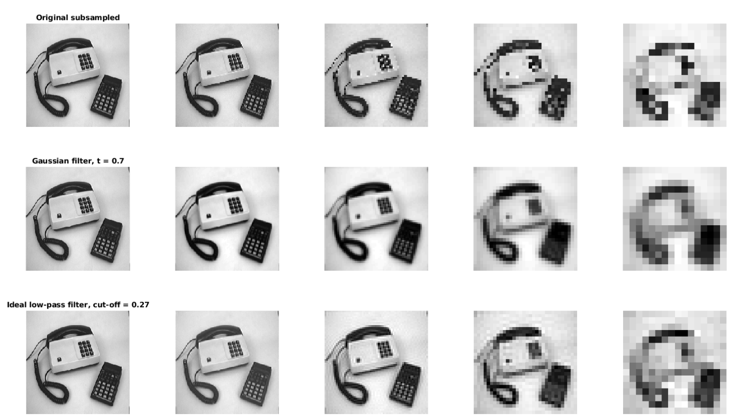

Figure 14:

Question 20: *What conclusions can you draw regarding the effects of smoothing when combined with subsampling? Hint: think in terms of frequencies and side effects.*

Answer: When we smooth the images before subsampling we reduce the high frequency content and therefore the effects of subsampling are not as strong as we have less high frequency content. The effects of aliasing are lessened due to removing these higher frequencies.