Answers to questions in Lab 1: Filtering operations

Name: Jacob Malmberg & Marcus Nystad Öhman Program: CINEK

Instructions: Complete the lab according to the instructions in the notes and respond to the questions stated below. Keep the answers short and focus on what is essential. Illustrate with figures only when explicitly requested.

Good luck!

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?

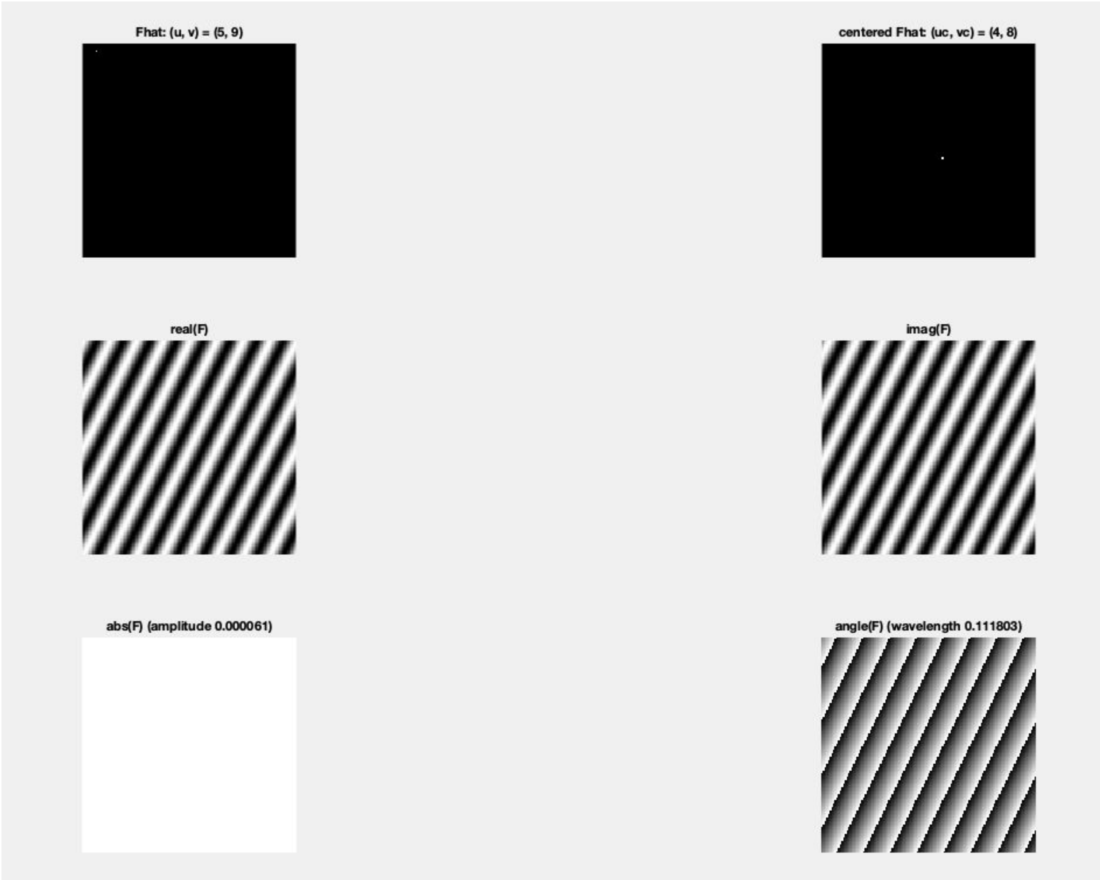
Answers:

Fhat: Shows where the point/sinuoid is 1. 0,0 the zero frequency is in the top left corner. So 5,9 is closer to the top left than 17,121.
Centered Fhat: 0,0 the zero frequency is in the middle of the picture. Due to periodic behaviour, you center it around the nearest corner to the point.
Real: Shows the direction of the wave. Real is the cos part.
Im: Sine is the sine part, which is cos + pi/2.
Angle: Shows the edges of the image. The wave goes in the same direction as the wave in the real part.
Amplitude: All the coordinates have the same magnitude (1).

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.

Answers:

The wave goes from the origin towards (p,q). If the position (p,q) is far from the center, the wave has higher frequency .



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

Answers:
The amplitude is 0.000061. The amplitude can be derived as follows:

$$F(x) = \frac{1}{N^2} \sum_{u \in [0..N-1]} \hat{F}(u)\, e^{\frac{2\pi i u^T x}{N}}$$

$$= \frac{1}{N^2} \sum_{u \in [0..N-1]} \hat{F}(u) \left[\cos\left(\frac{2\pi u^T x}{N}\right) + i \sin\left(\frac{2\pi u^T x}{N}\right) \right]$$

Amplitude is $\frac{1}{N^2} \max\left(\hat{F}(u)\right)$

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.

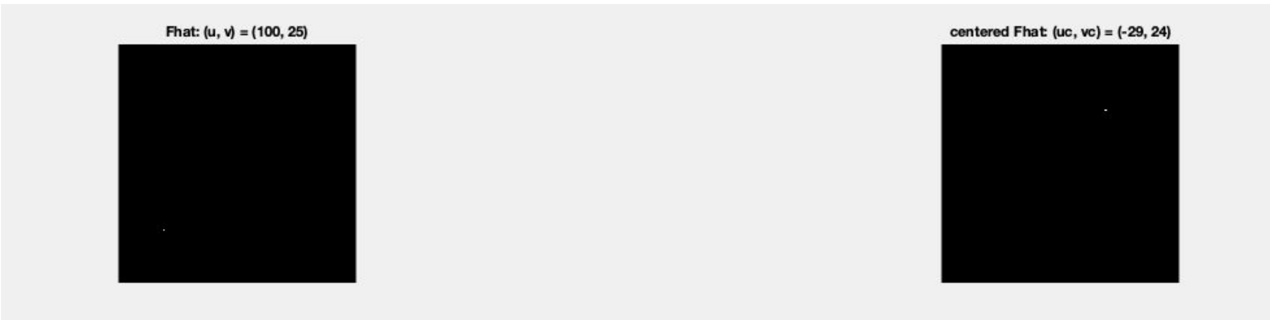
Answers:

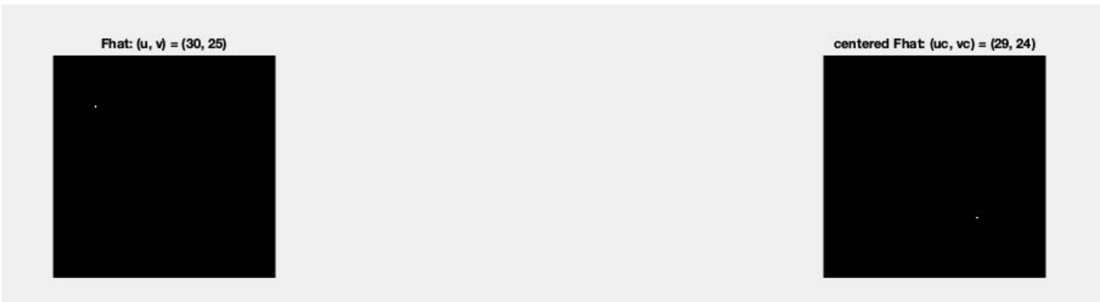
Higher frequency -> shorter wavelength. (Higher p and q) % 64 -> higher frequency.
The wave direction will point towards (p,q). The closer (p,q) are to the center of the Fourier domain the longer wave.

- The wavelength of the sinusoid is: $\lambda = \frac{1}{\sqrt{u^2+v^2}}$, where (u, v) are the frequencies along (r, c) and the periods are $1/u$ and $1/v$.

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!

Answers:





Centered Fhat shows the point p,q in the frequency domain with the zero frequency centered.
The discrete fourier transform is peridioc with the period N.

Although we cannot see it, the picture is periodic. That means that we have (0,0) in every corner. It is better to center it for visual reasons.

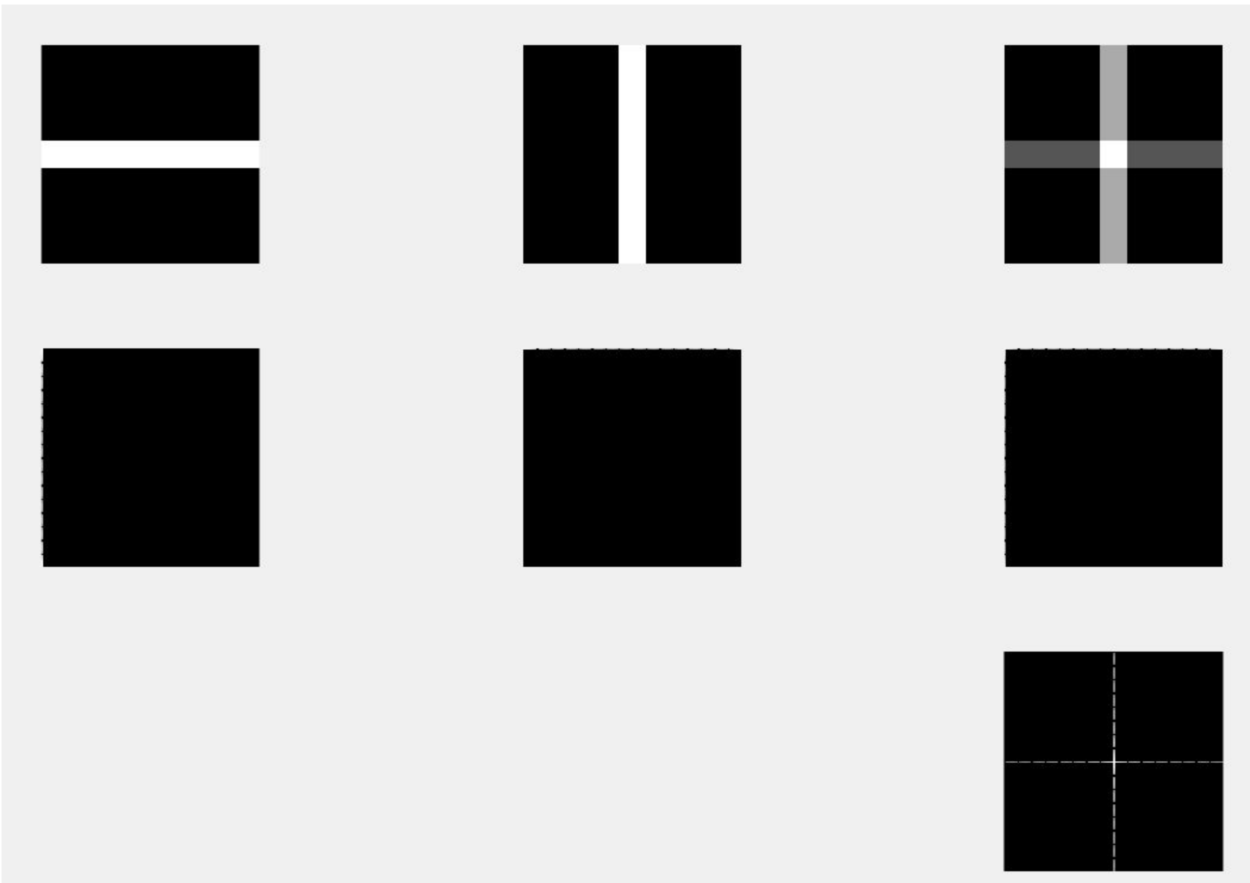
Question 6: What is the purpose of the instructions following the question *What is done by these instructions?* in the code?

Answers:

The purpose is to calculate the centered p and q values, and the **fftshift** call shifts the zero component to the centre of the spectrum. **fftshift** shifts the 1st with the 3rd quadrant, and the 2nd with the 4th quadrant.

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!

Answers:



The Fourier spectra is along the image border because we only have waves in that direction. That is, for F we only have waves in the ‘y’ direction, for ‘G’ we only have waves in the ‘x’ direction’, and H has waves in both directions.

The zero frequency component is at top left, therefore the spectras are concentrated to the border. If we use **fftshift** we can concentrate to the center, as shown above.

Math:

Thinking of the frequencies in the source image, we can see that there are only vertical waves in the image. That means that in the Fourier domain, only when v = 0 we have non-zero values. It becomes sort of a dirac delta function in that direction.

$$\delta(v) = \begin{Bmatrix} 1, v = 0 \\ 0, v \neq 0 \end{Bmatrix}$$

Looking at the 2d DFT, the inner sum becomes the dirac delta function.

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

becomes

$$\hat{f}(u, v) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} e^{-2\pi i \frac{mu}{M}} * \delta(v)$$

Question 8: Why is the logarithm function applied?

Answers: The logarithm function is applied to enhance the lowest frequency. The difference between the highest and lowest frequency becomes smaller by using the log transform, that is the dynamic range is lowered. That way, we can see more frequencies in the plot.

https://homepages.inf.ed.ac.uk/rbf/HIPR2/pixlog.htm

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

Answers: L (F + 2*G) = L(F) + L(2*G).

That is, the first row and the second row gives the same sum, since 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.

Answers: Multiplication in the spatial domain is the same as convolution in the Fourier domain and vice versa.

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwjos5nwtL3eAhWKDywKHQsDDMYQFjAAegQICRAB&url=https%3A%2F%2Fwww.mathworks.com%2Fmatlabcentral%2Fanswers%2F87413-how-do-i-have-to-normalize-fft-and-ifft-when-i-am-using-the-convolution-theorem&usg=AOvVaw2kdEyuzthY_9dfEF3mMtBz

---

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.

Answers: The source image has been compressed y-wise and expanded x-wise. The effects shown in the fourier domain is the opposite: compression leads to higher frequencies and expansion leads to fewer high frequencies.

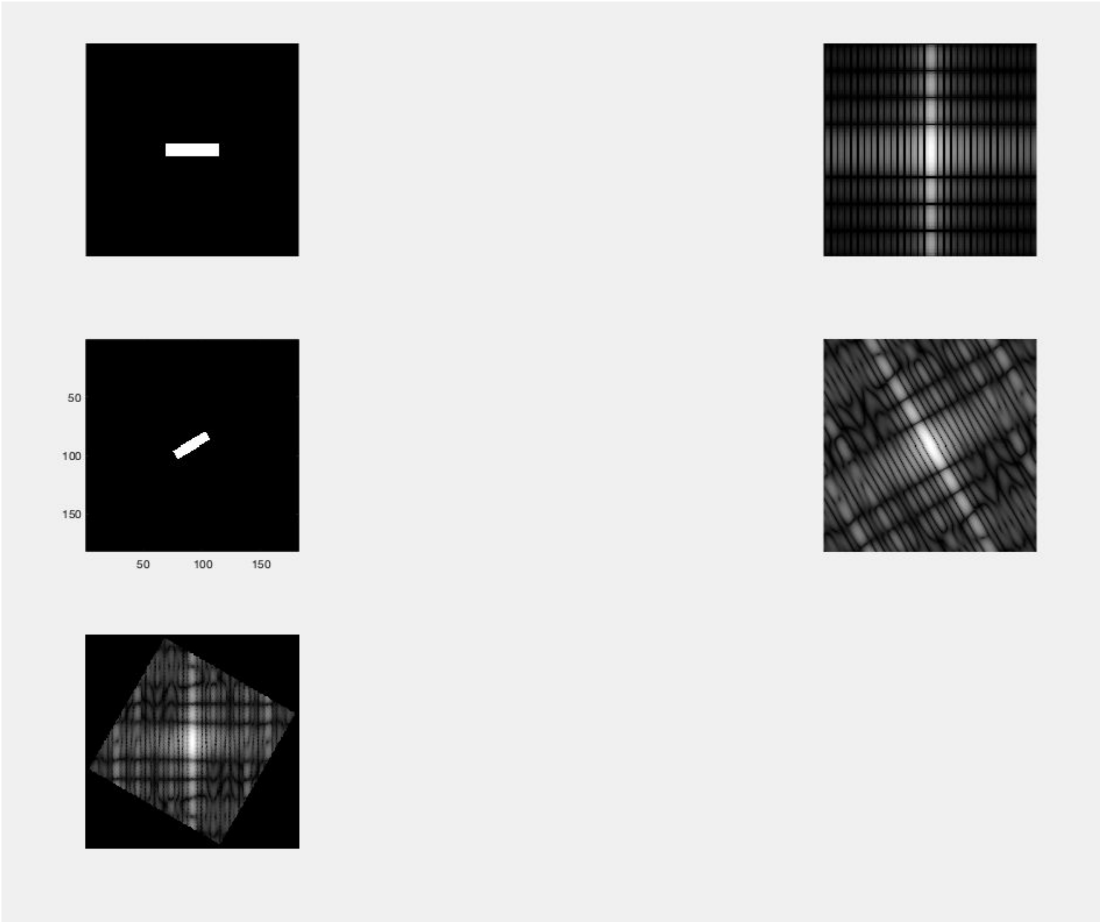
Question 12: What can be said about possible similarities and differences? Hint: think of the frequencies and how they are affected by the rotation.

Answers:

Differences: Along the 'diagonals' the frequencies are different, especially the high frequencies. The entire frequency spectrum is shifted N degrees.

Similarites: same structure.

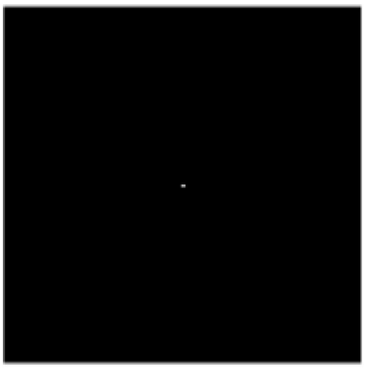
The image is slightly distorted because artifacts smaller than a pixel is introduced, because the rotated image can no longer be repesented perfectly on the grid. To represent these artifacts, you get more of the high frequencies.



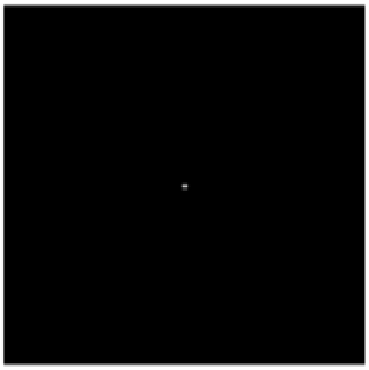
Question 13: What information is contained in the phase and in the magnitude of the Fourier transform?

Answers: The phase contains information about the edges/”shape”, where the edges are. The magnitude shows the “strength” of the various frequencies in the picture, how grey the picture is on either side of the edge.

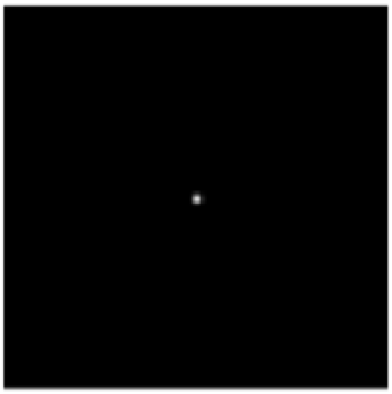
Question 14: Show the impulse response and variance for the above-mentioned t-values. What are the variances of your discretized Gaussian kernel for t = 0.1, 0.3, 1.0, 10.0 and 100.0?



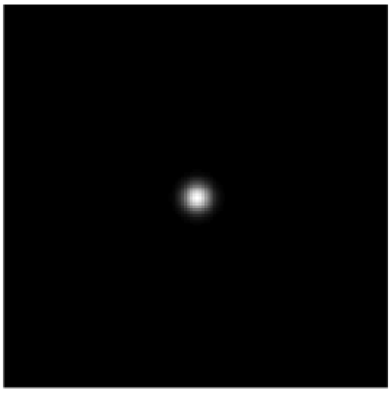
t = 0.1



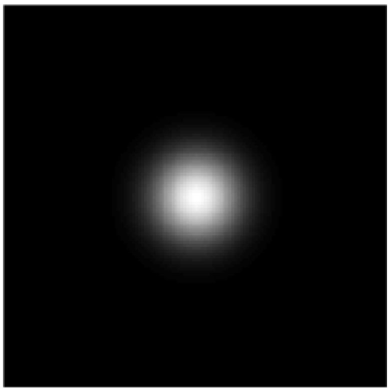
t = 0.3



t =1



t = 10



T=100

Answers:

T is variance. Gaussian sampling = put a grid on the gauss function in 2d! That is, the filter is mountain-like with its mean in the middle of the filter!

Variances, all diagonal 2x2 matrix with elements on the diagonal:

t=	est variance
0.1	0.0133
0.3	0.2811
1	1
10	10
100	100

Higher t value => more squished out image. Always a circle.

When t < 1, the Gaussian is like a spike. That means, it is smaller than a pixel.

A spike in the spatial domain becomes an extremely wide curve in the fourier domain. That means that not everything fits in the Fourier spectra, and some is cut off. When we then estimate the variance, it is not equal to t because some of the values are ‘cut off’.

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.

Answers:

When $t < 1$, the Gaussian is like a spike. That means, it is smaller than a pixel.

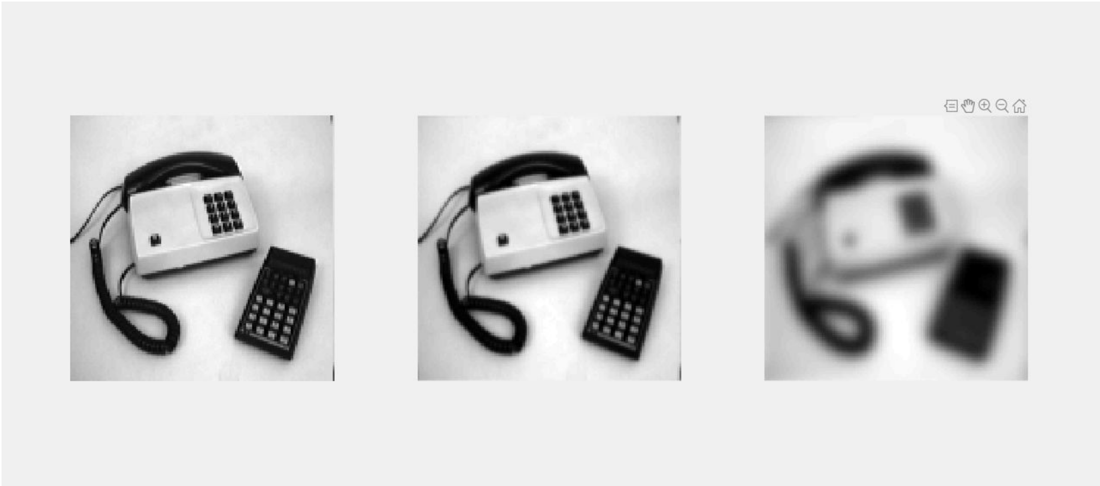
A spike in the spatial domain becomes an extremely wide curve in the fourier domain. That means that not everything fits in the Fourier spectra, and some is cut off. When we then estimate the variance, it is not equal to t because some of the values are ‘cut off’.

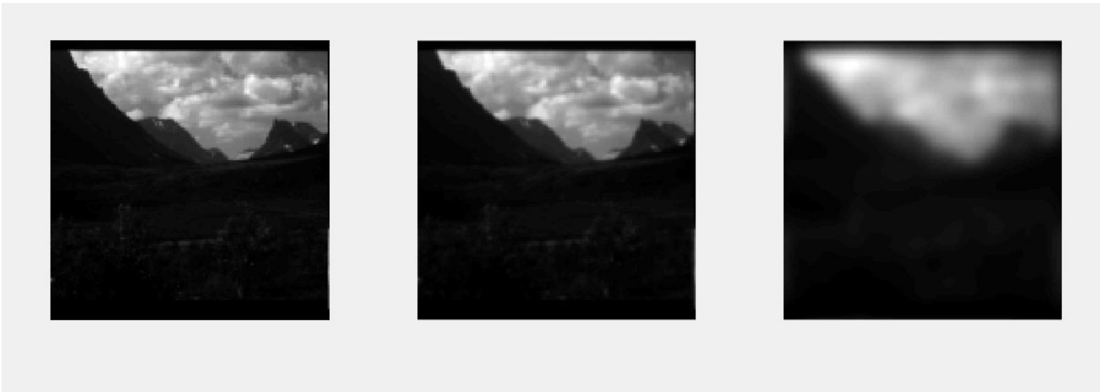
Question 16: Convolve a couple of images with Gaussian functions of different variances (like t = 1.0, 4.0, 16.0, 64.0 and 256.0) and present your results. What effects can you observe?

Answers:

Higher variance = more smoothing! Smaller the t-value, the larger the cutoff frequency and milder the filtering.

fft(gauss) is black in the corners and white in the middle, looks exactly like a mountain. So when we convolute, we multiply the fft(image) with the fft(gauss) elementwise. So in the edges the product of gauss and image is 0 / close to zero => high frequencies are removed. The picture gest smoother.





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).

Answers:

sap = salt and pepper noise.
add = additive noise. white gaussian noise. It changes the graylevel of each picture.

Gaussian smoothing: Gaussian smoothing removes higher frequencies and smoothes the image. The positive is that the smoothing removes the **add** noise, but it also smoothes the edges. When it comes to **sap** noise, t = 1 removes most of the noise and keeps the edges! t does not need to be large to remove the noise. Slightly more effective for sap.

Median filtering:

Effetively removes sap noise and add noise with a 10x10. However the image looks like a painting. Artifacts around the edges.

Ideal low pass filtering:

Can remove **sap** noise by setting a very large cutoff and still capture the edges/gist of the image. Not effective on **add** noise. With a low cutoff =0.1 there is noticeable ringing and blurring.

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

Answers:
sap noise easier to remove. add noise is hard to get rid of without removing edges or turning the image into a drawing. See median filter middle row.

Ideal Low pass filter / ringing:

Ideal low pass filter cuts off all frequencies above a certain threshold D. In the spatial domain, this looks like a sinc function. What happens is that we get white lines in the black areas, and black lines in the whiter areas.

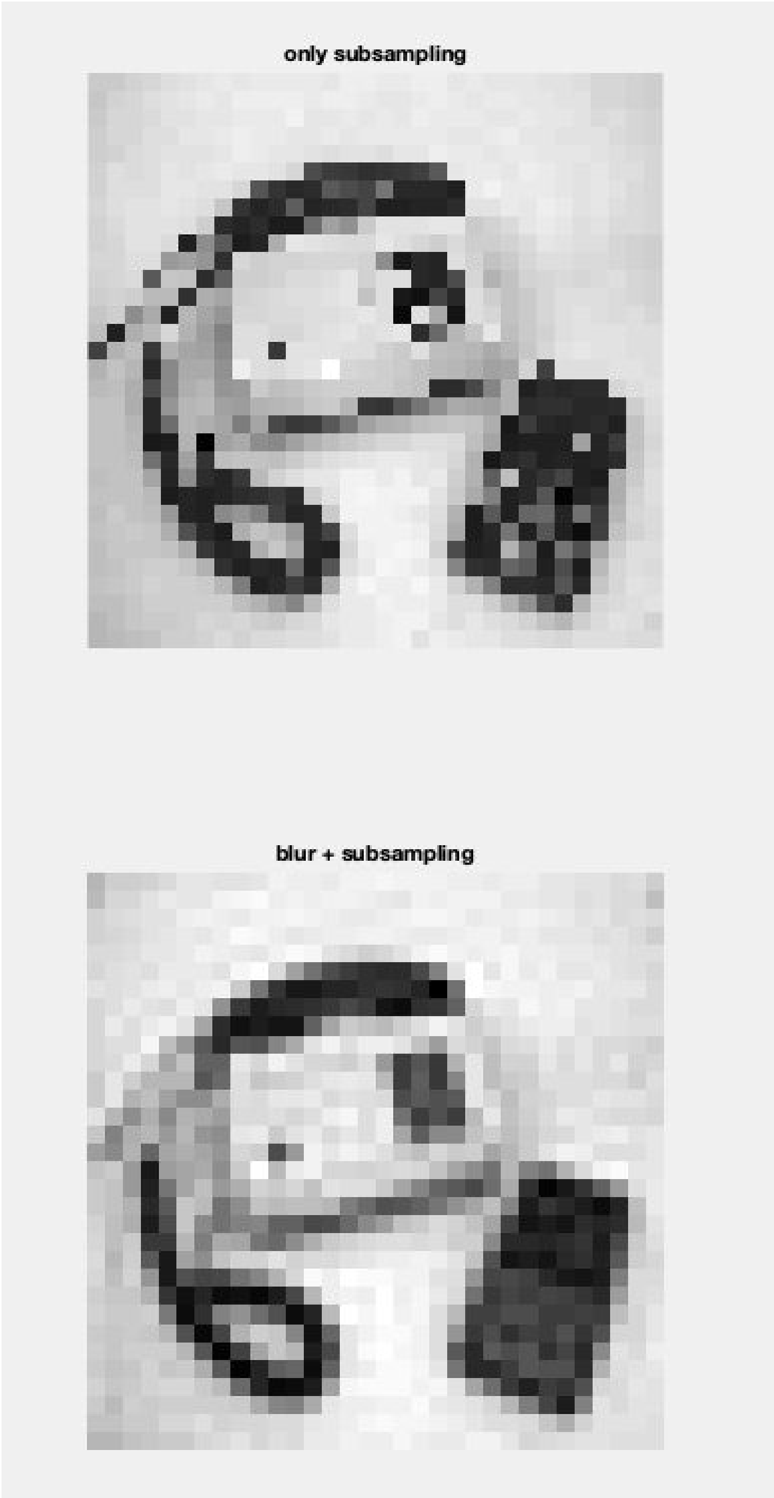
Question 19: What effects do you observe when subsampling the original image and the smoothed variants? Illustrate both filters with the best results found for iteration i = 4.

Answers:

Gauss: Subsampling and not smoothing leads to artifacts because of aliasing. By smoothing before subsampling, the image does not degrade as fast. Smoothing removes edges which solely subsampling does not, but subsampling removes the finer structures such as the numpad. By blurring before subsampling, the result is only a little aliasing.

Ideal low pass filtering:

If the cutoff is set too low rippling occurs. By setting the cutoff to 0.25, less ringing is achieved and the result can be considered to ‘better’ than for cutoff 0.1. The image is “smoother” than if we just subsample.



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.

Answers:

The result is ‘better’, that is more pleasing to the eye if we blur before we subsample.

The picture:

When we subsample in this case, the max frequency is cut in half because we remove every other pixel. The high frequencies that get cut off are **mirrored** in low frequency land. If we blur with a for instance gaussian filter, we can remove the high frequencies before subsampling and thus avoid the artifacts (and the mirroring). The artifacts come from the removal of the high frequencies.