Answers to questions in Lab 1: Filtering operations

Benedith Mulongo, master year 2 Machine learning

September 2019

1 Tasks

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?

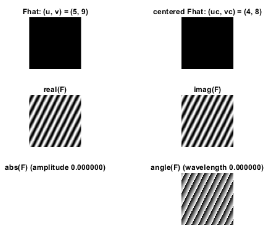

Figure 1: Result (1) for (5,9)

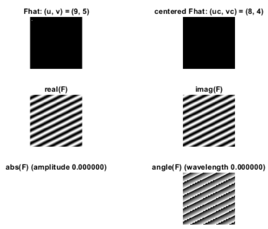

Figure 2: Result (2) for (9,5)

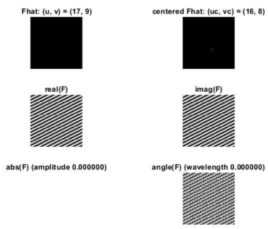

Figure 3: Result (3) for (17,9)

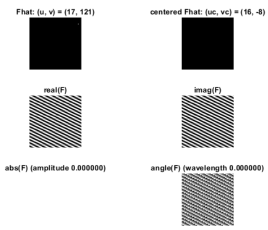

Figure 4: Result (4) for (17,121)

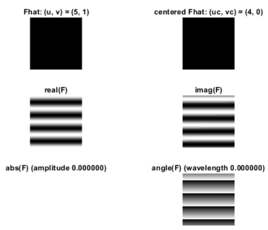

Figure 5: Result (5) for (5,1)

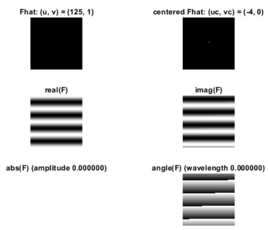

Figure 6: Result (6) for (125,1)

- **The first observation**

  The first observation is that when the point is on the first diagonal, the curve are in the form "///". When the points is near the center the wavelength is shorter and the variations is higher and vice versa.

  At the first extremity, the ifft is white and at the other extremity the ifft is in form "///"

- **The second observation**

  The first observation is that when the point is on the second diagonal, the curve are in the form "

  ∵ When the points is near the center the wavelength is shorter and the variations is higher and vice versa.

  at the first extremity the waveform are vertical and at the other the waveforms are diagonal.

- **The position and the angle**

  The position and the angle that the point forms with the x-axis play a big role. The direction of the point is perpendicular to the line in the spatial domain. And the frequency angular forms the waveform. Furthermore the number of period in the spatial domain in each axis correspond to point (v,u)

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.

The inverse discrete Fourier transform can be expressed as :

$$x(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F(k,l)\exp\left[j2\pi\Big(\frac{mk}{M}+\frac{nl}{N}\Big)\right] \tag{1}$$

Given that the matrix $F(k,l)$ is an extremely sparse matrix, all the entries are zero except for the entry (p,q). Equations can then be rewritten as follows :

For k,l such as F(k,l) = 1

$$x(m,n) = \frac{1}{MN}\exp\left[j2\pi\Big(\frac{mk}{M}+\frac{nl}{N}\Big)\right] \tag{2}$$

This is true because every entries are zeros except for the entry (k,l) where F(k,l) is equal to 1. Therefore equation (2) holds.

Using the following Euler relationship :

$$e^{jnx} = \cos(nx) + j\sin(nx) \tag{3}$$

$$e^{-jnx} = \cos(nx) - j\sin(nx) \tag{4}$$

equation (2) becomes :

$$x(m,n) = \frac{1}{MN}\left[\cos\left(2\pi\left[\frac{mk}{M}+\frac{nl}{N}\right]\right) + j\sin\left(2\pi\left[\frac{mk}{M}+\frac{nl}{N}\right]\right)\right] \tag{5}$$

Using a figure F of dimension 128X128 with

$$F(k,l) = \begin{cases} 1, & \text{if } k=5, n=10 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

We get the following figures for its IDFT and the projection of point (5,10)

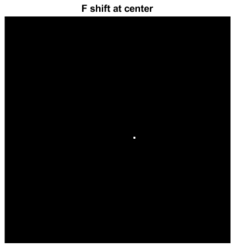

Figure 7: Point in the freq. domain

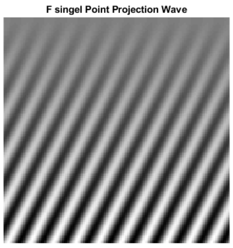

Figure 8: Projection in spatial domain 2D

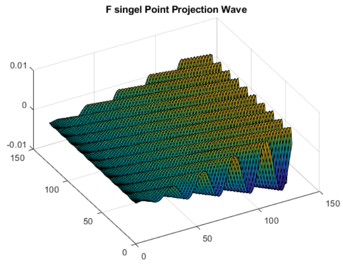

Figure 9: Projection in spatial domain 3D

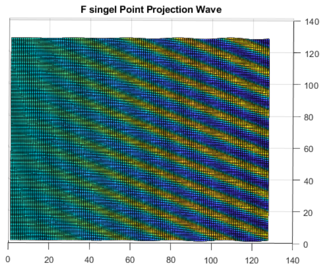

Figure 10: Projection in spatial domain as 2D

Projections of point (5,10) as a sine wave in the spatial domain.

Question 3

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

The discrete Fourier transform is defined as follows :

$$x(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F(k,l)\exp\left[j2\pi\Big(\frac{mk}{M}+\frac{nl}{N}\Big)\right]$$

Using Euler formula, we have

$$x(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F(k,l)\Big[\cos\Big(2\pi\Big[\frac{mk}{M}+\frac{nl}{N}\Big]\Big)+j\sin\Big(2\pi\Big[\frac{mk}{M}+\frac{nl}{N}\Big]\Big)\Big] \tag{7}$$

We set $t = \Big(2\pi\Big[\frac{mk}{M}+\frac{nl}{N}\Big]\Big)$ and $\int = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F(k,l)$

Then equation (7) becomes :

$$x(m,n) = \int\Big[\cos\Big(t\Big)+j\sin\Big(t\Big)\Big] \tag{8}$$

which is equivalent to :

$$x(m,n) = \int \cos\left(t\right) + j\int \sin\left(t\right) \tag{9}$$

We know that the spectrum of the Fourier is equal to :

$$|x(m,n)| = \sqrt{Re(x(m,n))^2 + Im(x(m,n))^2} \tag{10}$$

Applying equation (10) to the relation (9), we get :

$$|x(m,n)| = \sqrt{\int^2 * cos(t)^2 + \int^2 * sin(t)^2}$$

$$|x(m,n)| = \sqrt{\int^2 (cos(t)^2 + sin(t)^2)}$$

$$|x(m,n)| = \sqrt{\int^2}\sqrt{cos(t)^2 + sin(t)^2}$$

Given that the left part of the equation is equal to 1, we get the final form of the spectrum.

$$|x(m,n)| = \int$$

Replacing $\int$ we get :

$$|x(m,n)| = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F(k,l) \tag{11}$$

The amplitude is the maximum of the equation (11)

$$amplitude = max(|x(m,n)|) \tag{12}$$

In our case we have $amplitude = 1/(128^2)$ which is very small so the difference is not visible as in figure 11.

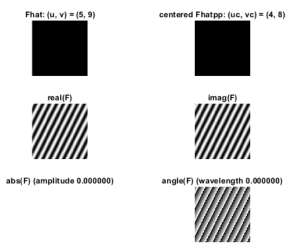

Figure 11: Point in the freq. domain

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.

The discrete Fourier transform is defined as follows :

$$x(m,n) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k,l) \exp\left[j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)\right]$$

The wave length is defined as :

$$\lambda = \frac{2\pi}{||w||} = \frac{2\pi}{\sqrt{w_1^2 + w_2^2}}$$

we have the following in our case :

$$w = \begin{bmatrix} 2\pi\frac{p}{M} \\ 2\pi\frac{q}{N} \end{bmatrix}$$

The wavelength can be then written as follow :

$$\lambda = \frac{1}{\sqrt{\frac{p^2}{M^2} + \frac{p^2}{N^2}}}$$

Assuming that $M = N$, we get the following formula for the wavelength :

$$\lambda = \frac{M}{\sqrt{p^2 + q^2}}$$

The phase angle is defined as :

$$\phi(w_1, w2) = \arctan\frac{Im(w_1, w2)}{Re(w_1, w2)}$$

In our case, we we only have one point equal to 1, the discrete four transform can be writen as :

$$x(m,n) = \frac{1}{MN}\Big[\cos\left(w^T[m,n]\right) + j\sin\left(w^T[m,n]\right)\Big] \tag{13}$$

where $w = \begin{bmatrix} 2\pi\frac{p}{M} \\ 2\pi\frac{q}{N} \end{bmatrix}$

The phase is :

$$\phi(w_1, w2) = \arctan\Big[\tan\left(w_1 m + w_2 n\right)\Big]$$

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!

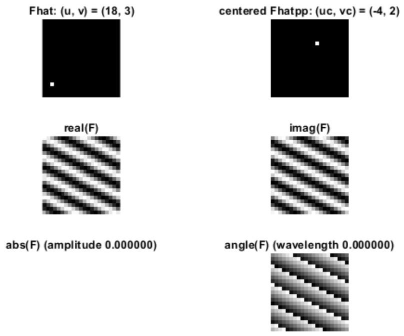

Figure 12: Point (18,3)

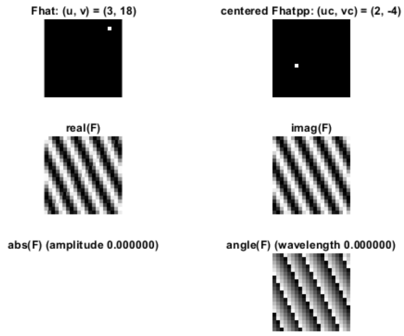

Figure 13: Point (3,18)

What happens is that if p exceeds half of the image size, the point is map to the first quadrant. A swap between the third quadrant and the first quadrant occurs as shown in figure 12.

In the second case when q exceeds half of the image size, the point is map to the third quadrant. A swap between the third quadrant and the first quadrant occurs as shown in figure 12.

In general there is a swap between the first quadrant and the third quadrant and between the second quadrant and the fourth. That means we change the axis from $[0, M]$ to $[-\frac{M}{2}, \frac{M}{2}]$

Question 6

What is the purpose of the instructions following the question What is done by these instructions? in the code?

The fftshift function changes the axis from $[0, M]$ to $[-\frac{M}{2}, \frac{M}{2}]$. The purpose of the instruction is to calculate the new location of the point (p,q) after the shift to the center. and we use -1 in order to compensate the fact that Matlab begins indexing at point 1 compared to other programming languages.

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!

The discrete Fourier transform is defined as :

$$X(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m,n) \exp\Big[-j2\pi\Big(\frac{mk}{M} + \frac{nl}{N}\Big)\Big] \tag{14}$$

We can use the separability property in order to write equation (25) as :

$$X(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m,n) \exp\Big[-j2\pi\Big(\frac{mk}{M}\Big)\Big] \exp\Big[-j2\pi\Big(\frac{nl}{N}\Big)\Big] \tag{15}$$

Given the example defined in Matlab we have, a regtangular form where $x(m,n) = 1$ in the range $m \in [57, 72]$. We can then rewrite (15) as :

$$X(k,l) = \sum_{m=57}^{72} \sum_{n=0}^{N-1} \exp\Big[-j2\pi\Big(\frac{mk}{M}\Big)\Big] \exp\Big[-j2\pi\Big(\frac{nl}{N}\Big)\Big]$$

$$X(k,l) = \sum_{m=57}^{72} \exp\Big[-j2\pi\Big(\frac{mk}{M}\Big)\Big] \sum_{n=0}^{N-1} \exp\Big[-j2\pi\Big(\frac{nl}{N}\Big)\Big] \tag{16}$$

$$\exp\Big[-j2\pi\Big(\frac{nl}{N}\Big)\Big] = \begin{cases} 1, & \text{if } l = 0 \\ 0, & \text{if } l \neq 0 \end{cases} \tag{17}$$

Replacing (16) in equation (17), we get :

$$X(k,l) = \begin{cases} \sum_{m=57}^{72} \exp\Big[-j2\pi\Big(\frac{mk}{M}\Big)\Big], & \text{if } l = 0 \\ 0, & \text{if } l \neq 0 \end{cases} \tag{18}$$

Equation (18) shows exactly why the Fourier spectra are concentrated to the borders. We observe that the Fourier spectra is only defined when $l = 0$, which implies that everything will be concentrated at the border.

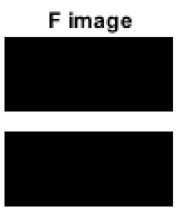

Figure 14: Spatial domain of F

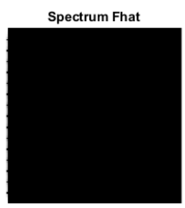

Figure 15: Frequency domain of F

Question 8

Why is the logarithm function applied?

The logarithm is use in order to reduce the range of $|X(k,l)|$ with the purpose of increasing the details for visualization purposes. Figure (16) shows how slowly the logarithm function increases, therefore it is used here in order to reduce the range and increase the details.

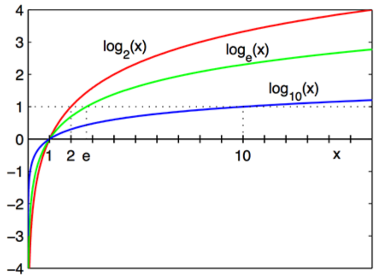

Figure 16: Logarithm function

Question 9

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

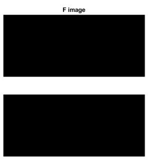

Figure 17: F image

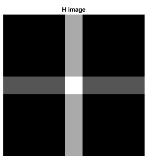

Figure 18: G image

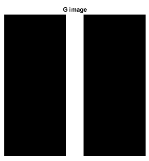

Figure 19: H image

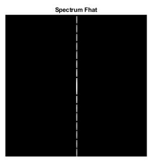

Figure 20: F spect

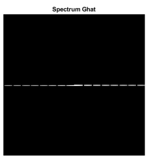

Figure 21: G spectrum

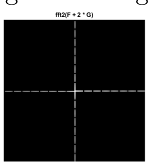

Figure 22: fft(F + 2G)

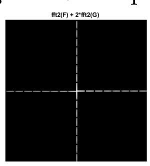

Figure 23: fft(F) + 2fft(G)

We can conclude that the Fourier transform is a linear operation which can be expressed mathematically as :

$$\mathcal{F}\big[ax(m,n)+by(m,n))\big] = a\mathcal{F}\big[X(k,l)\big] + b\mathcal{F}\big[Y(k,l)\big] \tag{19}$$

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.

The multiplication in the Fourier domain is the convolution in the spatial domain. Which can summarizes as follows :

$$\mathcal{F}\big[x(m,n) \circledast y(m,n)\big] = \mathcal{F}\big[X(k,l)\big] \odot \mathcal{F}\big[Y(k,l)\big] \tag{20}$$

$$\mathcal{F}\big[[x(m,n) \odot y(m,n)\big] = \frac{1}{NM}\big[X(k,l) \circledast Y(k,l)\big] \tag{21}$$

In short another way of computing the last image is to perform the convolution in the spatial domain.

In Matlab we have :

```
Cfull = conv2(fft2(F),fft2(G))/(128*128);
Cfull = Cfull(1:128,1:128);
```

Listing 1: convolution in spatial domain

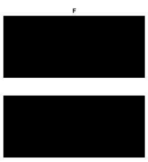

Figure 24: F image

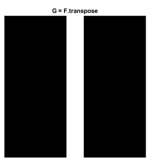

Figure 25: G image

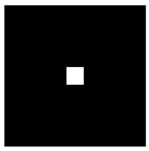

Figure 26: F multiplied by G

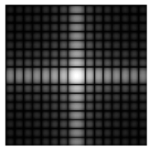

Figure 27: fft(F x +G)

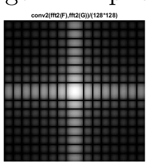

Figure 28: G spectrum

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.

The scaling property of the Fourier transform can be summarized as follows :

$$\mathcal{F}\big[f(ax,by)\big] = \frac{1}{|ab|}\mathcal{F}\big[\frac{k_x}{a}, \frac{k_y}{b}\big] \tag{22}$$

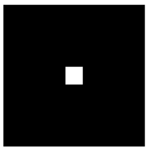

Figure 29: Square image

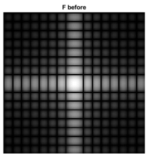

Figure 30: Square spectrum

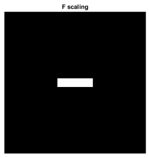

Figure 31: Rectangle image

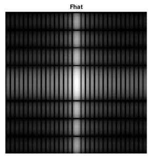

Figure 32: Rectangle spectrum

What we can say is that an expansion in the spatial domain results to a compression and vice versa.

Question 12

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

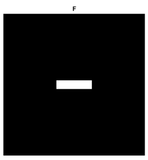

Figure 33: F image

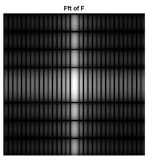

Figure 34: F spectrum

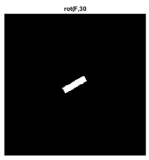

Figure 35: rotate(F, 30)

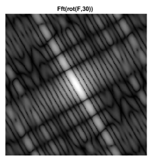

Figure 36: Fft(rotate(F, 30))

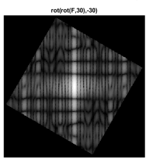

Figure 37: rot(fft(rot(F, 30)),-30)

The rotation property is given by :

$$\mathcal{F}\big[f(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)\big] = F\big[(X\cos\theta + Y\sin\theta, -X\sin\theta + Y\cos\theta)\big] \tag{23}$$

When an image is rotated in the spatial domain, the image is also rotated in the frequency domain. The magnitude is the same but the phase its direction changes. The image are a little bit different because the rotation is not perfect. The rotated image is more periodic than the original image.

Question 13

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

The Fourier transform can be written as :

$$X(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x(m,n)\exp\Big[-j2\pi\Big(\frac{mk}{M}+\frac{nl}{N}\Big)\Big] \tag{24}$$

which can be written in polar form :

$$X(k,l) = |X(k,l)|\exp\Big[j\Phi\Big] \tag{25}$$

Where :

- **Magnitude**
  $|X(k,l)| = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x(m,n)$ $|X(k,l)| = \sqrt{Re(X(k,l))^2 + Im(X(k,l))^2}$
- **Phase**
  $\Phi(u,v) = \arctan\frac{Im(X(k,l)}{Re(X(k,l)}$

Figure 38: Img1

Figure 39: Magn1.

Figure 40: Img2

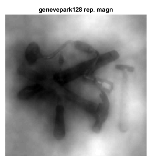

Figure 41: Magn2.

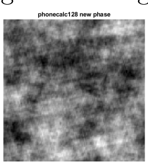

Figure 42: Img1 phase

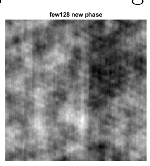

Figure 43: Img2 phase

When we change the magnitude of the image, we still can distinguish some similarity between the original image and its new version.

When we use a new phase the image is completely indistinguishable so we can say that he magnitude contains the information of the power or "intensity" of the frequency for a particular component. However the phase decides which frequency component are present or not in the image.

> **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?

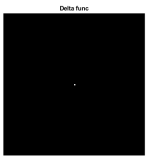

Figure 44: Delta fnc

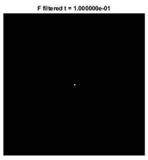

Figure 45: $t = 0.1$

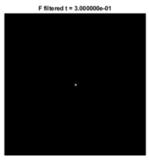

Figure 46: $t = 0.3$

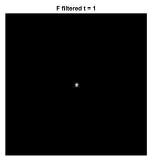

Figure 47: $t = 1$

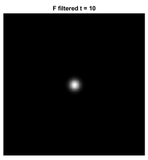

Figure 48: $t = 10$

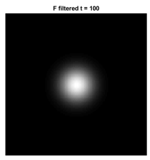

Figure 49: $t = 100$

The impulse response for different t-values

The different co-variance for t = 0.1, 0.3, 1, 10, 100 are:

$$w = \begin{bmatrix} 0.0133 & 0.0000 \\ 0.0000 & 0.0133 \end{bmatrix}$$

$$w = \begin{bmatrix} 0.2811 & 0.0000 \\ 0.0000 & 0.2811 \end{bmatrix}$$

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

$$w = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

$$w = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}$$

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.

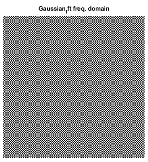

Figure 50: fft(gauss = 0.1)

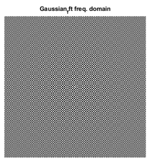

Figure 51: fft(gauss = 0.3)

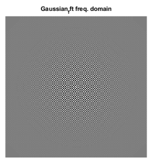

Figure 52: fft(gauss = 1)

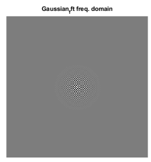

Figure 53: fft(gauss = 10)

The Fourier transform for Gaussian with variance t = 0.1, 0.3, 1, 10

The results are quite similar to the estimated variance for t ¡ 1 and identical for $t \geqslant 1$

For higher variance for the Gaussian in the spatial domain results to a Gaussian with smaller variance in the frequency domain. which make the result more similar to Delta for higher t-values

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?

Figure 54: Orig

Figure 55: t = 1

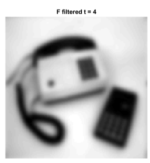

Figure 56: t = 4

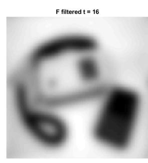

Figure 57: t = 16

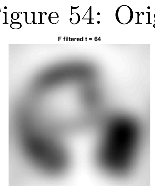

Figure 58: t = 64

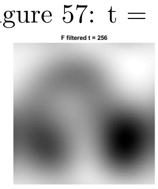

Figure 59: t = 256

The Gaussian blurring for different variance t = 1, 4, 16, 64, 256

The blurring effect or smoothing effect is amplified when the variance is increased. Higher frequency is increased.

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).

Ideal low pass filter :

1. Salt noises

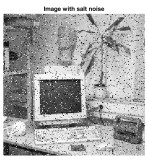

Figure 60: Ideal F.

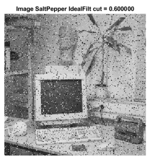

Figure 61: Ideal F.

Figure 62: Ideal F.

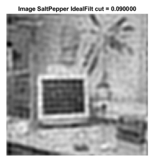

Figure 63: Ideal F.

2. Gaussian noises

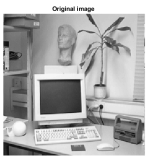

Figure 64: Ideal F.

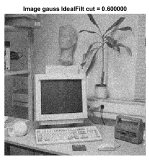

Figure 65: Ideal F.

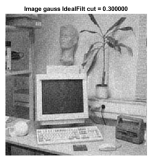

Figure 66: Ideal F.

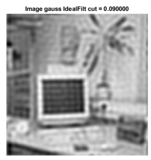

Figure 67: Ideal F.

- **Advantages**

  The pros of the Ideal filter is that it is good to filter out high frequency outside the center

- **Disadvantages**

  The cons is the ringing effect it transfers to the image. There is no good performance for the Gaussian and Salt and pepper noises.

Median filter :

1. Salt noises

Figure 68: Ideal F.

Figure 69: Ideal F.

Figure 70: Ideal F.

Figure 71: Ideal F.

2. Gaussian noises

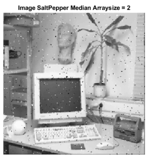

Figure 72: Ideal F.

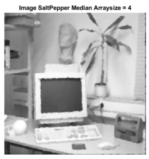

Figure 73: Ideal F.

Figure 74: Ideal F.

Figure 75: Ideal F.

- **Advantages**

  No influence from outliers. It works well with salt and pepper noise

- **Disadvantages**

  Bad for random noises without peak and sink. Higher filter size NXM increase the blurring effect.

Gaussian filter :
1. Salt noises

Figure 76: Ideal F.

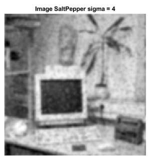

Figure 77: Ideal F.

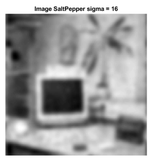

Figure 78: Ideal F.

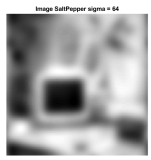

Figure 79: Ideal F.

2. Gaussian noises

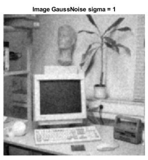

Figure 80: Ideal F.

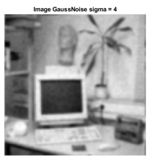

Figure 81: Ideal F.

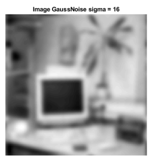

Figure 82: Ideal F.

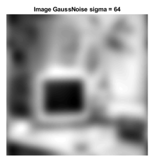

Figure 83: Ideal F.

- **Advantages**

  Works well for Gaussian noises. Good for smooth divergence attenuation. No ringing effect

- **Disadvantages**

  No good for other types of noises such as salt and pepper noises.

Question 18

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

Gaussian :

Increasing the variance increasing the blurring effect without noise however the image is blurred. Its performance is best for Gaussian noises

Median :

Best performance for alt and pepper noises. Has the best performance on the average.

Ideal low-pass :

Not good performance on the overall. It is good for removing high frequencies based on the cutoff.

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.

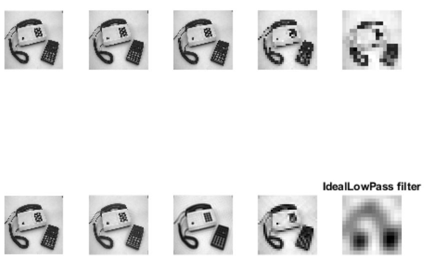

Figure 84: Sub-sampling Ideal F.

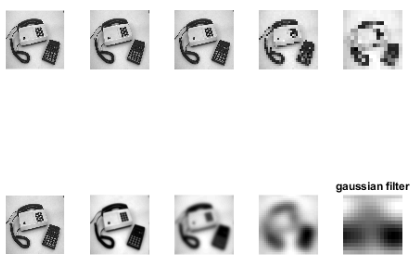

Figure 85: Sub-sampling Gauss

But looking at the effect of different filters, we can observe that the ideal filter has a better effect in that aspect it keeps much of the information contained in the original image even after sub-sampling. However the Gaussian filter renders the image more blurrier by ignoring the fine details in the original image. However the image obtained with the Gaussian filter it more "clean" no aliasing effect, smoother to look at.

When using sub-sampling alone without smoothing or filtering beforehand the resulting image is of poor quality and indistinguishable even for 4 iterations. The filter enhance the image quality.

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.

When we use image sub-sampling the resulting image is of poor quality, because the sampling rate is nit high enough to capture the details in the image. High frequencies are transformed to lower frequencies provoking the aliasing effect. In order to avoid the aliasing effect, the sampling rate must be at least two ties the maximum frequency of the image (Nyquist rate, at least tow samples per cycle). When using the low-pass filtering, it attenuates the high frequencies and enhances the low-frequencies within a certain area, which has the effect of improving the image quality.

References