Answers to questions in Lab 1: Filtering operations

Name: Yihan Wang   Program: TINNM

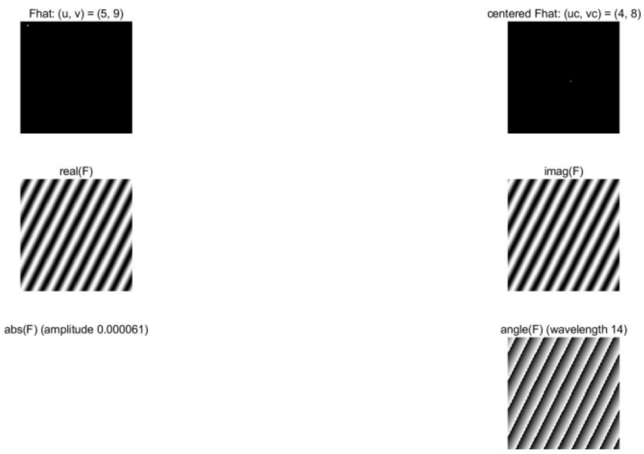
Instructions: Complete the lab according to the instructions in the notes and respond to the questions stated below. Keep the answers short and focus on what is essential. Illustrate with figures only when explicitly requested.

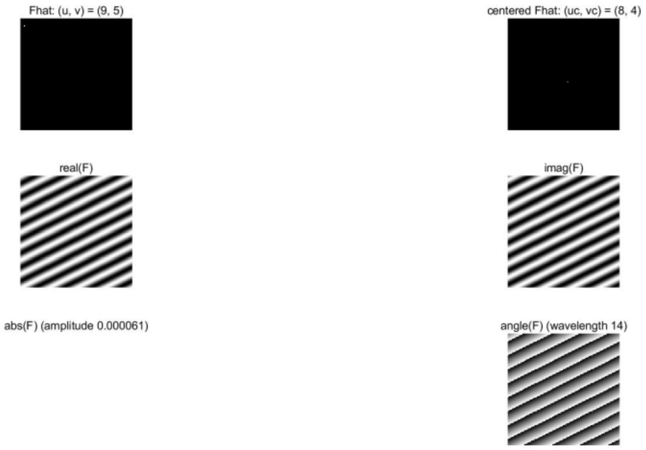
Good luck!

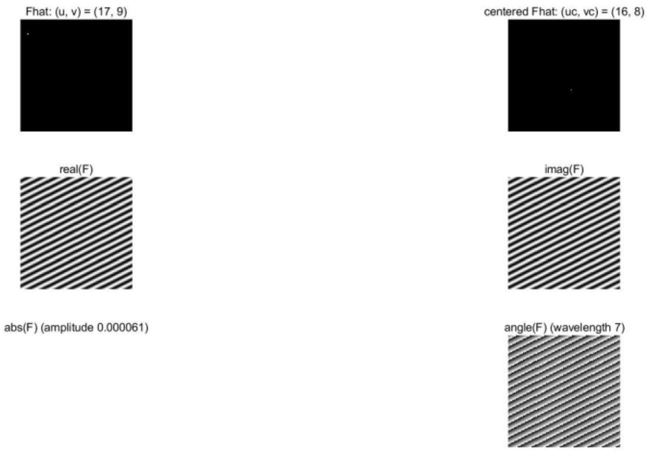
Question 1: Repeat this exercise with the coordinates p and q set to (5, 9), (9, 5), (17, 9), (17, 121), (5, 1) and (125, 1) respectively. What do you observe?

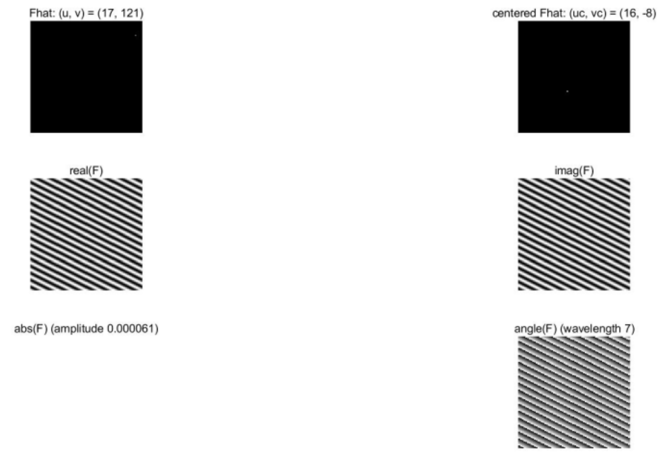
Answers:
The results are shown in **Figure 1** below. From these figures, we observe that:

1. All these settings have same amplitude. (the reason will be shown in Q3)
2. If the distance between centered q, p (i.e. uc , vc) and center (0,0) is further, the frequency will be higher and wavelength will be smaller.
3. For (q, p)=(5,1) and (125,1), the centered points (4,0) and (-4,0) are symmetry on uc=0. The real parts are the same, while image parts and angles have a shifted phase π .
   For (q, p)=(17,9) and (17,121), the centered points (16,8) and (16,-8) are symmetry on vc=0. The real, image and angle are all each other's horizontal flips.
   For (q, p)=(5,9) and (9,5), the centered points (4,8) and (8,4), as well as the real, image and angles are all on (0, 0) symmetry.
   The facts observed above show that the rotation in spatial / frequency domain will cause the rotation in another domain, and the phase will change accordingly, but the amplitude will not change.




Fhat: (u, v) = (9, 5)
centered Fhat: (uc, vc) = (8, 4)
real(F)
imag(F)
abs(F) (amplitude 0.000061)
angle(F) (wavelength 14)


Fhat: (u, v) = (17, 9)
centered Fhat: (uc, vc) = (16, 8)
real(F)
imag(F)
abs(F) (amplitude 0.000061)
angle(F) (wavelength 7)


Fhat: (u, v) = (17, 121)
centered Fhat: (uc, vc) = (16, -8)
real(F)
imag(F)
abs(F) (amplitude 0.000061)
angle(F) (wavelength 7)

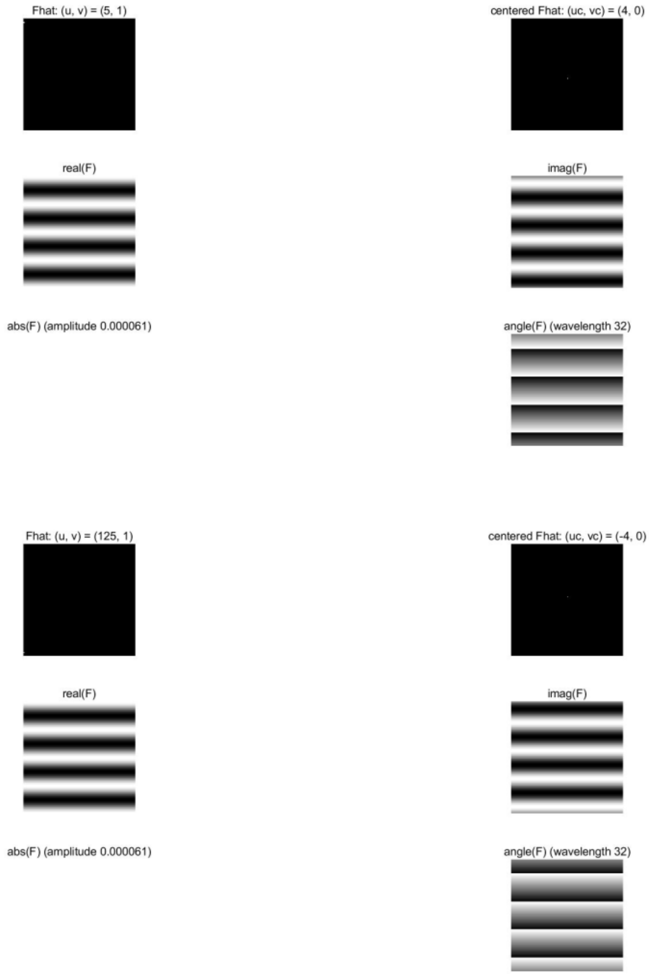


Figure1: p and q set to (5, 9), (9, 5), (17, 9), (17, 121), (5, 1) and (125, 1) respectively

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:
As indicated in Eq (4), we can derive
$F(x) = \frac{1}{N}\sum_{u,v\in[0,\ldots,N-1]} \hat{F}(u,v)e^{\frac{i2\pi(ux+vy)}{N}} = \frac{1}{N}\sum_{u,v\in[0,\ldots,N-1]} \hat{F}(u,v)(\cos\left(\frac{2\pi(ux+vy)}{N}\right) + isin(\frac{2\pi(ux+vy)}{N}))$ ①
By substituting the coordinates of a point, a sin function line corresponding to the corresponding spatial domain can be obtained.
For example, point (5,1) in Fourier domain is projected to sin waves in spatial domain shown in **Figure 2** below:

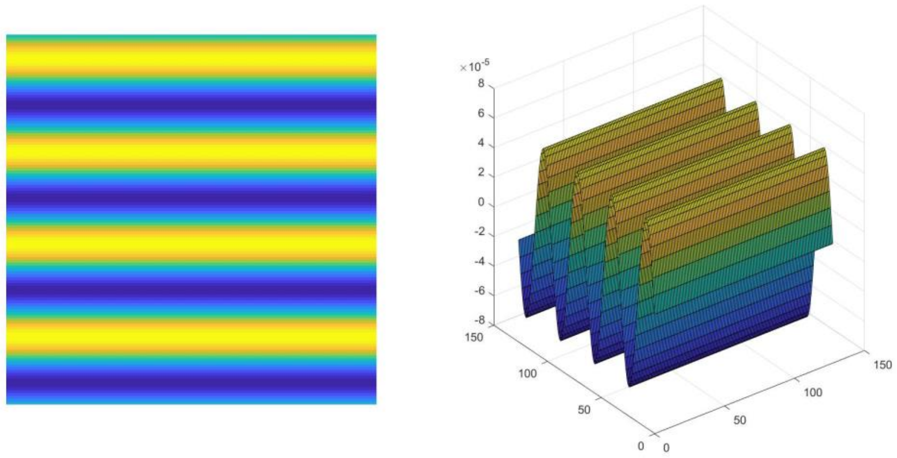


Figure 2: left: 2D image part with q=5, p=1; right: 3D image part with q=5, p=1.

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:
In this part, we define F(u,v) equals to either 0 or 1, i.e. Eq. ① is non-zero only when F(u,v)=1. Thus, the expression of amplitude can be derived from Eq.①:

$$Amplitude = \frac{1}{N}$$

However, since in Matlab the fast Fourier transform (FFT) is implemented using a factor 1 in the FFT-routine, and factor $1/N^2$ in the inverse, here the amplitude should be computed as:

$$Amplitude = \frac{1}{N^2}$$

The code is completed as:

```
amplitude =1/sz^2;
```

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:
As for the length of the sin wave, from the lecture note, we found the explicit expression of wavelength:

$$\lambda = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{\omega_1^2 + \omega_2^2}}$$

where $\omega_1 = 2\pi \frac{u}{N}$, $\omega_2 = 2\pi \frac{v}{N}$. After the centering operation, (u,v) is centered as (uc ,vc), than the wavelength equation becomes to:

$$\lambda = \frac{2\pi}{\omega} = \frac{N}{\sqrt{uc^2 + vc^2}}$$

From the above equation, we see that the further the distance between centered points and origin (0, 0) is, the shorter the wavelength will be.

As for the direction, it is determined by $2\pi(ucx + vcy)/N$, which derived from Eq.①. For (x, y) satisfying $ucx + vcy = 0$ i.e. $y = -\frac{uc}{vc}x$, this line is perpendicular to direction of the sin wave. Thus, line $y = \frac{uc}{vc}x$ is the direction of the sin wave.
The code is completed as:

```
w1=2*pi*uc/sz;
w2=2*pi*vc/sz;
wavelength =2*pi/sqrt(w1^2+w2^2);
```

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!

Answers:
After using Matlab function *fftshift* , the original origin will change from top-left of the image to the center. This means the coordinate of (p ,q) should be re-computed. Since DFT of pixels is arranged periodically, after *fftshift* the original block in the image will flip diagonally. Besides, the real, image and angle of the image also showed periodicity and repeatability. The results are shown in **Figure 3** below:

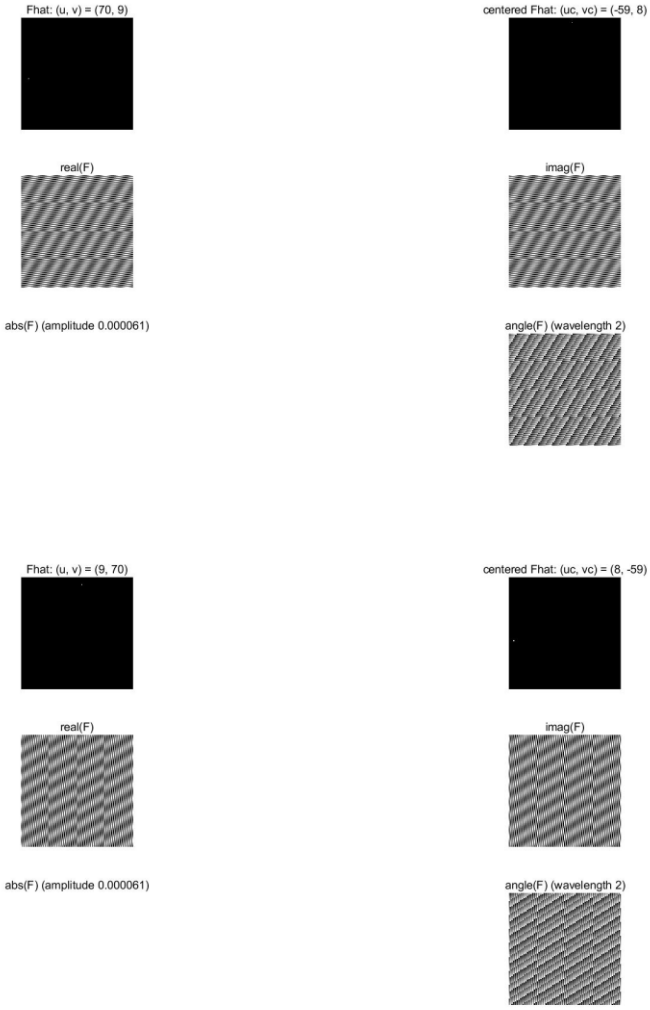


Figure 3: results when either p or q exceeds half the image size

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

Answers:

The instruction shows how to compute the new coordinate of (q, p) after using Matlab function *fftshift*, and the original image. In the original image, the frequency variable u and v are treated as being defined in the interval (0, 2π), while after the centering operation, the interval changes to (-π, π).

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 discrete Fourier transform is defined as:

$$\hat{F}(u,v) = \frac{1}{N} \sum_{x,y \in [0,\dots,N-1]} F(x,y) e^{\frac{-2\pi i(ux+vy)}{N}}$$

Calculate in single dimension and consider F as an example: since $\hat{F}(u,v) \neq 0$ only when $55<y<72$, above equation can be expressed as:

$$\hat{F}(u,v) = \frac{1}{N} \sum_{x=56}^{71} e^{\frac{-2\pi iux}{N}} \sum_{y=0}^{N-1} e^{\frac{-2\pi ivy}{N}}$$

Using the Dirac (continuous domain) and Kronecker (discrete) delta function as well as the periodicity of exponential function, we have:

$$\sum_{y=0}^{N-1} e^{\frac{-2\pi ivy}{N}} = \delta(v) = \begin{cases} 1, when\ v = 0 \\ 0, when\ v \neq 0 \end{cases}$$

This explains why these Fourier spectra concentrated to the borders of the images.

Question 8: Why is the logarithm function applied?

Answers:
It can be seen from **Figure 4-8** that before using log, the gray values are concentrated in a relatively small range, but their absolute values are very large. After using log, the absolute values of gray values become smaller, but they are distributed in a relatively large range, which can increase the contrast of the image and reduce the dynamic range, so that the details are easier to identify.

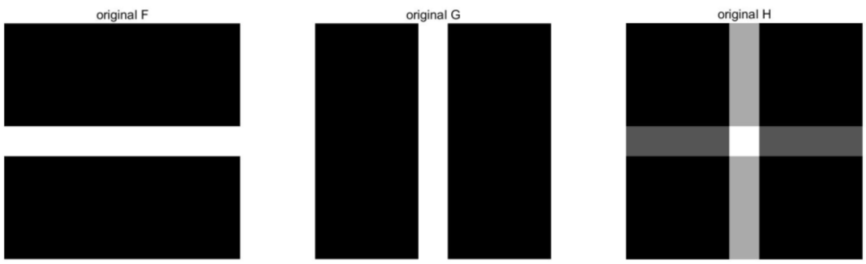


Figure 4: original images in spatial domain

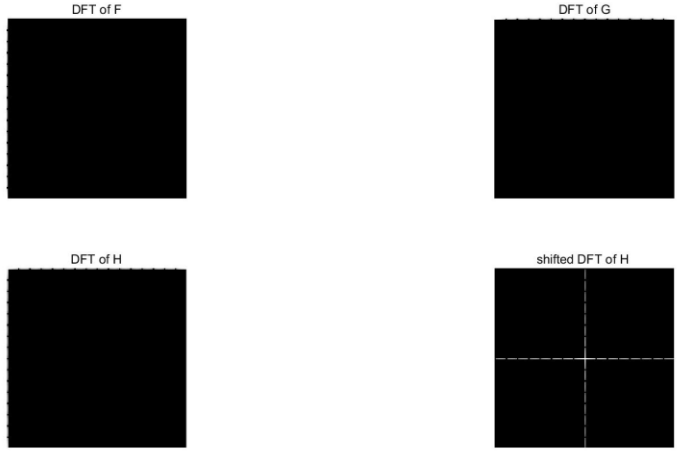


Figure 5: images in frequency domain with log function

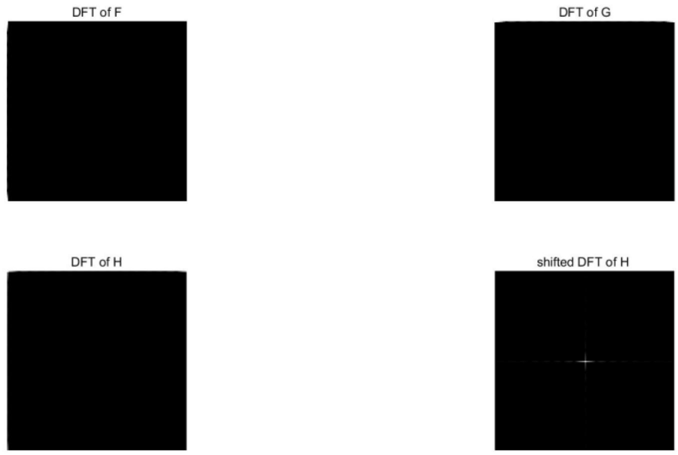


Figure 6: images in frequency domain without log function

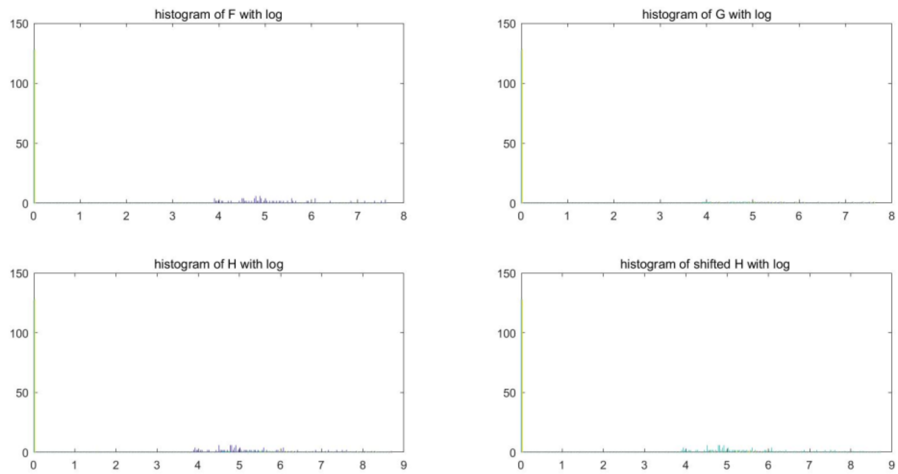


Figure 7: histogram with log function

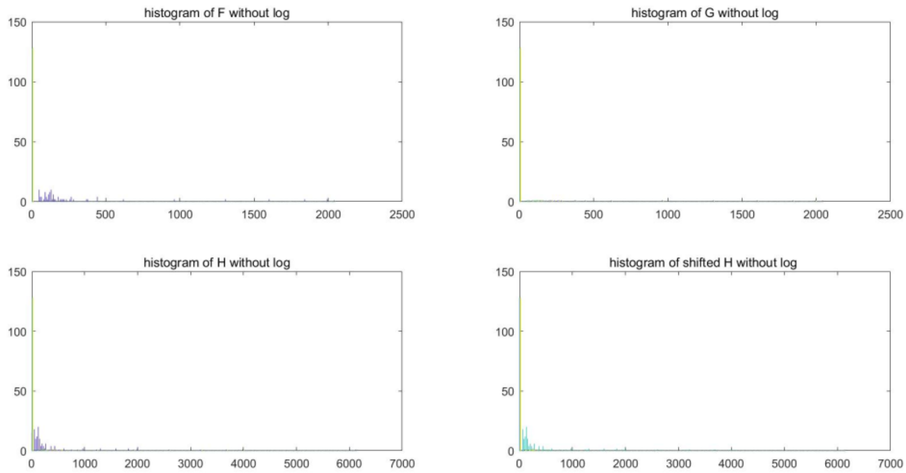


Figure 8: histogram without log function

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

Answers:
In spatial domain, H = F + 2 * G, which means H is linear combination of F and G. From figure we can see, DFT of H is also a combination of DFT(F) and DFT(G). This linearity property can be described as:

$$\mathcal{F}[a\, f_1(m,n) + b\, f_2(m,n))] \; = \; a\, \widehat{f}_1(u,v) \; + \; b\, \widehat{f}_2(u,v)$$

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:
Since multiplication in Fourier domain equals to convolution in the spatial domain:

$$\hat{f}(u) \cdot \hat{f}(v) = f(x) * f(y) \; and \; \hat{f}(u) * \hat{f}(v) = f(x) \cdot f(y)$$

We use convolution of Fhat and Ghat here to derive the same result. The results are shown in **Figure 9,10**:

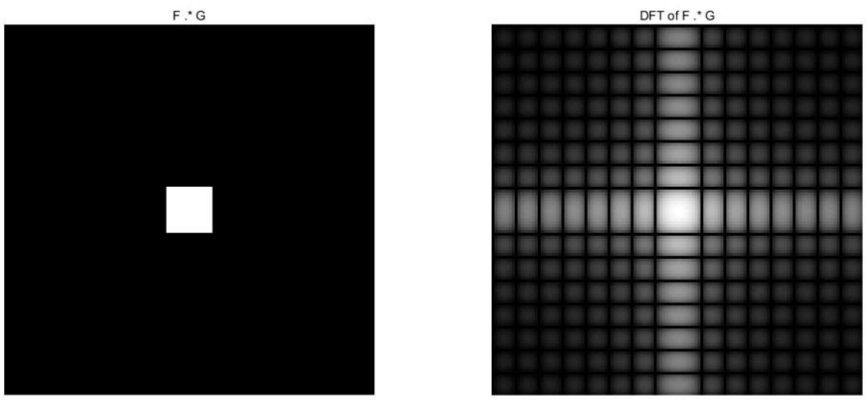


Figure 9: multiply then apply FFT

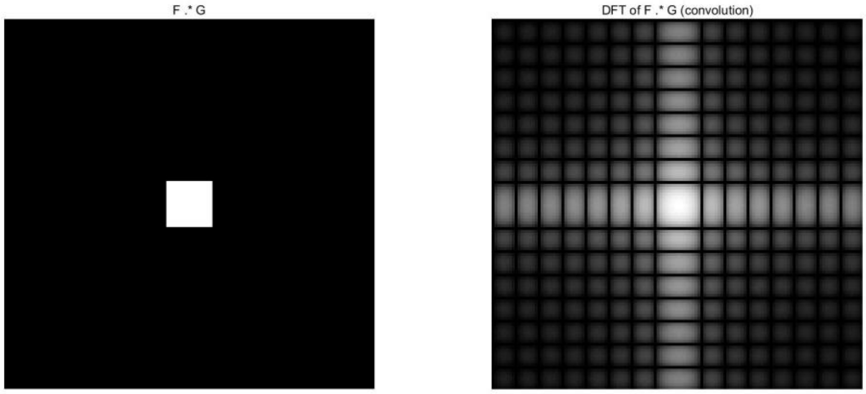


Figure 10: apply convolution operation

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:
Compare the results with those in the previous exercise, we can see from **Figure 11** there is a block with size of 16×16 in the previous source image, while a rectangular size 8×32 is in the middle of the new source image. The coordinates of them in spatial domain have relationship : $f_{pre}(x, y) = f_{new}(2\text{x}, \frac{1}{2}\text{y})$, while in frequency domain the relationship turns to $\widehat{f_{pre}}(u, v) = \widehat{f_{new}}\left(\frac{1}{2}u, 2y\right)$. From these, we derive the effect of scaling:
If in spatial domain $g(x) = f(S_1 x_1, \dots, S_n x_n)$, then in frequency domain $\hat{g}(\omega) = \frac{1}{|S_1 \dots S_n|} \hat{f}(\frac{\omega_1}{S_1}, \dots, \frac{\omega_n}{S_n})$, which means compression in spatial domain is same as expansion in Fourier domain (and vise versa).

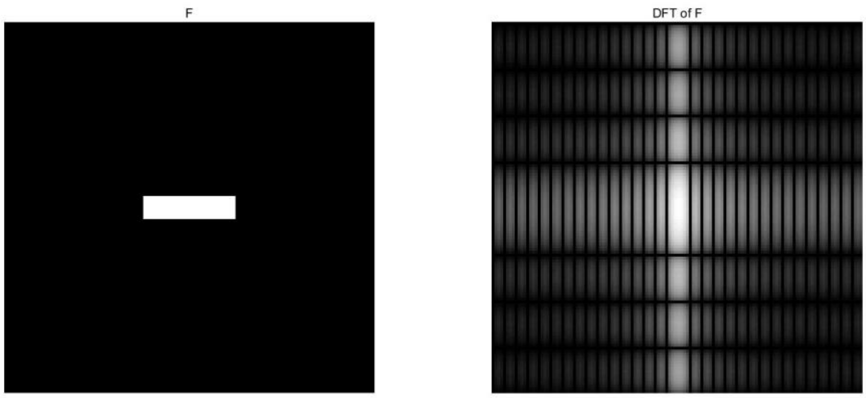


Figure 11: results after scaling

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:

From the results show in **Figure 12**, we observe that rotation of the original image rotates $\hat{f}\ (in\ the\ frequency\ domain)$ by the same angle.
We can prove this mathematically:
Recall the rotation matrix equation in Cartesian coordinate system

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = M(\theta)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xcos\theta - ysin\theta \\ xsin\theta + ycos\theta \end{pmatrix}$$

And substitute it into definition of DFT

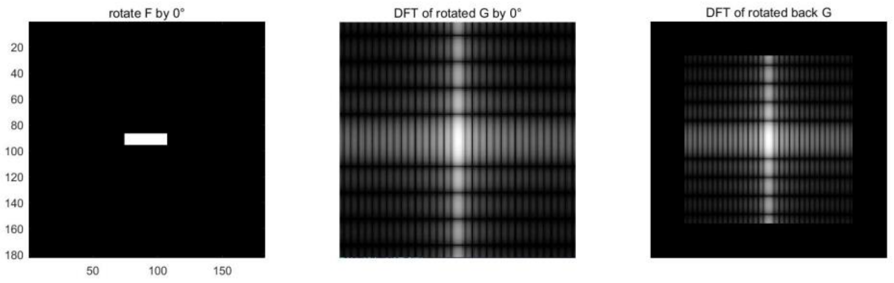
$$\hat{F}(u,v) = \frac{1}{N}\sum_{x,y\epsilon[0,\dots,N-1]} F(x,y)e^{\frac{-2\pi i(ux+vy)}{N}}$$

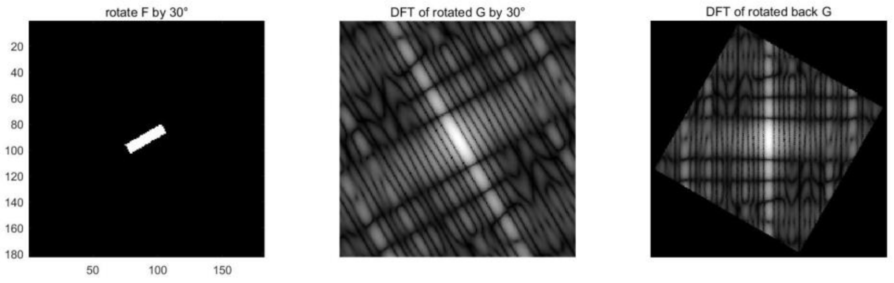
$$= \frac{1}{N}\sum_{x,y\epsilon[0,\dots,N-1]} F(x_r cos\theta - y_r sin\theta, x_r sin\theta + y_r cos\theta)e^{\frac{-2\pi i(u(x_r cos\theta - y_r sin\theta) + v(x_r sin\theta + y_r cos\theta))}{N}}$$

$$= \frac{1}{N}\sum_{x,y\epsilon[0,\dots,N-1]} F(x_r cos\theta - y_r sin\theta, x_r sin\theta + y_r cos\theta)e^{\frac{-2\pi i(x_r(ucos\theta + vsin\theta) + y_r(-usin\theta + vcos\theta))}{N}}$$

Set $u_r = ucos\theta + vsin\theta$, $v_r = -usin\theta + vcos\theta$, then above equation can be written as:

$$\hat{F}(u_r,v_r) = \frac{1}{N}\sum\nolimits_{x_r,y_r\epsilon[0,\dots,N-1]} F(x_r,y_r)e^{\frac{-2\pi i(u_r x_r + v_r y_r)}{N}} ②$$





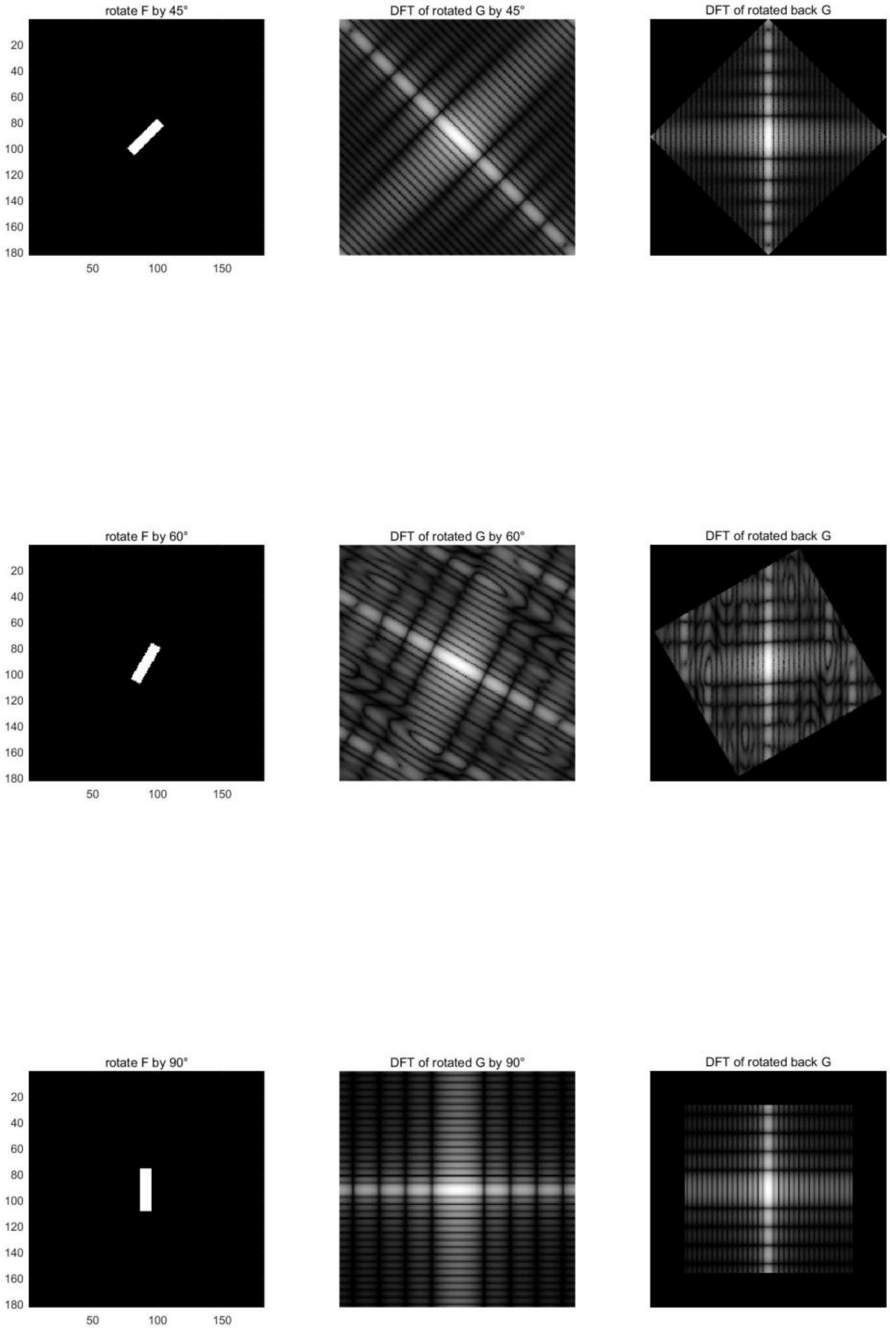


Figure 12: results after rotating 0°, 30°, 45°, 60°, 90° respectively

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

Answers:
From the result shown in **Figure 13**, we find magnitude keeps information about brightness and darkness, while phase information is about pixels' locations. Compare the second and the

third row in the figure, the intensity of corresponding pictures are the same, but the pixels with a specific intensity has different locations in the two pictures, which leads to difficulty in distinguish the objects. Then compare the first and the second row that with same phase and different magnitude, we find brightness has been influenced (pictures in second row are darker than these in first row), while locations of pixels keep the same that we can still distinguish the objects in that picture easily.

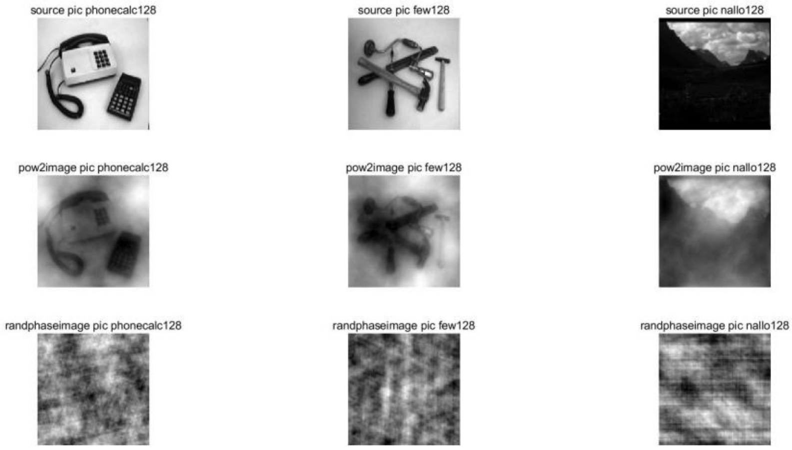


Figure 13: information contained in magnitude and phase

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?

Answers:
The 2D and 3D illustration of impulse response are shown in **Figure 14, 15**, and the results show that Gaussian filter with higher variance (larger t) will has wider impulse graphic.

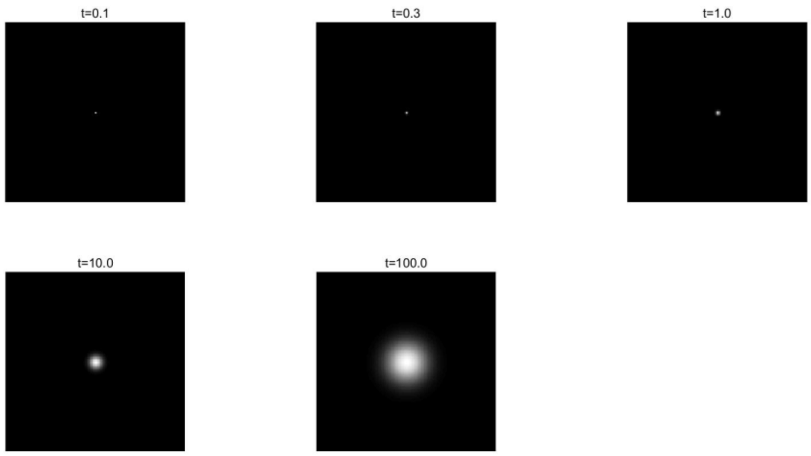


Figure 14: 2D illustration of impulse

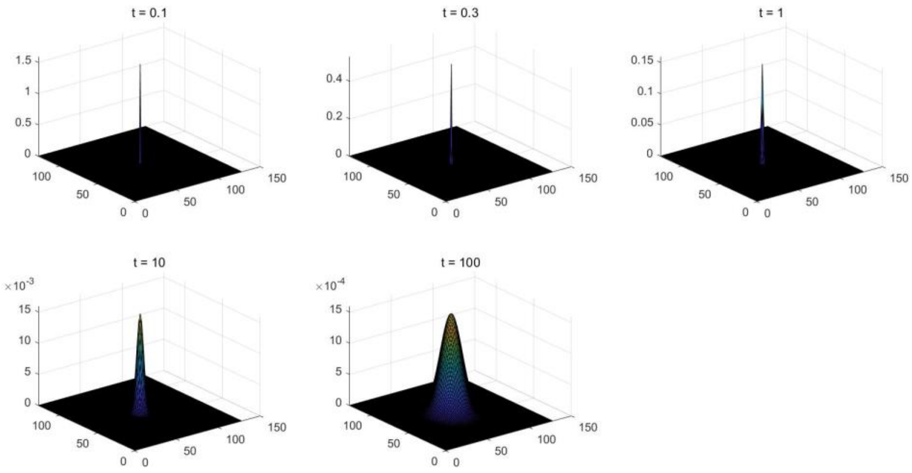


Figure 15: 3D illustration of impulse

The variance is shown in the table below:

t	variance
0.1	$\begin{pmatrix} 0.0133 & 0 \\ 0 & 0.0133 \end{pmatrix}$
0.3	$\begin{pmatrix} 0.2811 & 0 \\ 0 & 0.2811 \end{pmatrix}$
1.0	$\begin{pmatrix} 1.0000 & 0 \\ 0 & 1.0000 \end{pmatrix}$
10.0	$\begin{pmatrix} 10.0000 & 0 \\ 0 & 10.0000 \end{pmatrix}$
100.0	$\begin{pmatrix} 100.0000 & 0 \\ 0 & 100.0000 \end{pmatrix}$

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:
Estimated variance (using function ***discgaussfft***):

t	variance
0.1	$\begin{pmatrix} 0.1000 & 0 \\ 0 & 0.1000 \end{pmatrix}$
0.3	$\begin{pmatrix} 0.3000 & 0 \\ 0 & 0.3000 \end{pmatrix}$
1.0	$\begin{pmatrix} 1.0000 & 0 \\ 0 & 1.0000 \end{pmatrix}$
10.0	$\begin{pmatrix} 10.0000 & 0 \\ 0 & 10.0000 \end{pmatrix}$
100.0	$\begin{pmatrix} 100.0000 & 0 \\ 0 & 100.0000 \end{pmatrix}$

Ideal continuous case variance:

$$C\big(g(\cdot,\cdot;t)\big) = t\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

t	variance
0.1	$\begin{pmatrix}\mathbf{0.1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0.1}\end{pmatrix}$
0.3	$\begin{pmatrix}\mathbf{0.3} & \mathbf{0}\\ \mathbf{0} & \mathbf{0.3}\end{pmatrix}$
1.0	$\begin{pmatrix}\mathbf{1.0} & \mathbf{0}\\ \mathbf{0} & \mathbf{1.0}\end{pmatrix}$
10.0	$\begin{pmatrix}\mathbf{10.0} & \mathbf{0}\\ \mathbf{0} & \mathbf{10.0}\end{pmatrix}$
100.0	$\begin{pmatrix}\mathbf{100.0} & \mathbf{0}\\ \mathbf{0} & \mathbf{100.0}\end{pmatrix}$

Compare the results with estimated variance and ideal continuous variance, we find they are different when t<1.0, and basically same when t>=1.0. when t<1.0, the smaller the t value is, the larger the difference will be. This shows that when t<1.0, it is more likely to have non-gaussian distribution.

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:
The results are shown in **Figure 16**. From the results we derive that as the value of t increases, the image becomes more and more blurred. The reason for this phenomenon is that when the value of t increases, the cut-off frequency of the Gaussian filter becomes smaller, that is to say, more high-frequency components are discarded. However, the high frequency component contains more edge information, which makes the amount of information in the remaining pictures less, and the edges are difficult to distinguish, so the pictures are more fuzzy.


t = 1
t = 4
t = 16

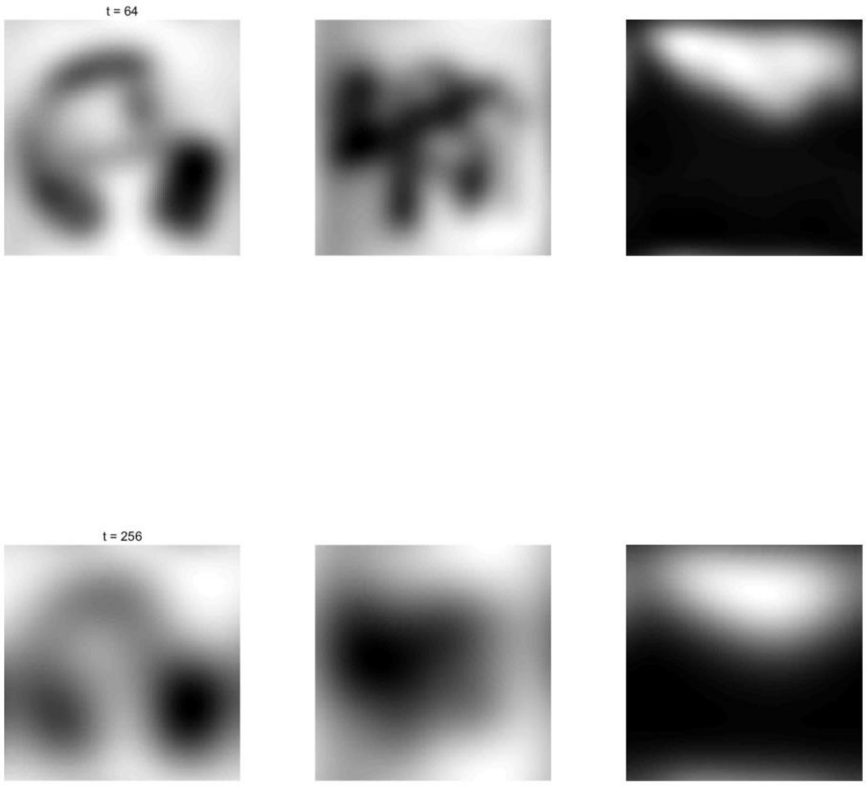


Figure 16: original images & convolution of a couple of images with Gaussian functions of different variances

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:
The results are shown in **Figure 17-20**. From the results we find Gaussian filter has best performance when handling Gaussian noise, while median filter preforms best with pepper-and-salt noise.

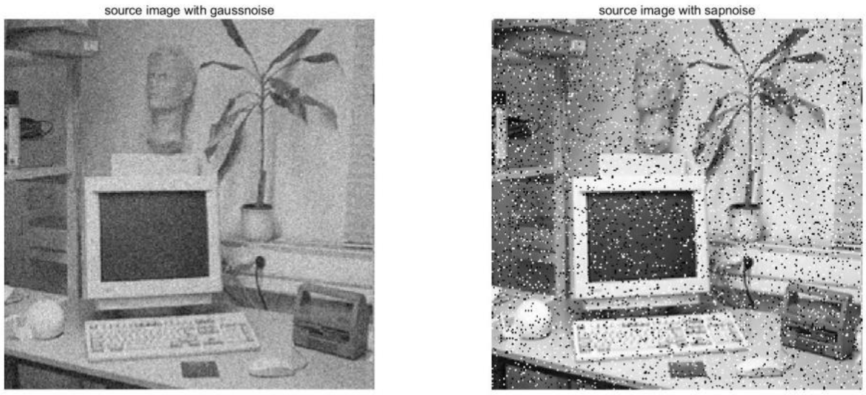


Figure 17: original images (left with Gauss noise and right with pepper-and-salt noise)

Figure 18: Gaussian filter

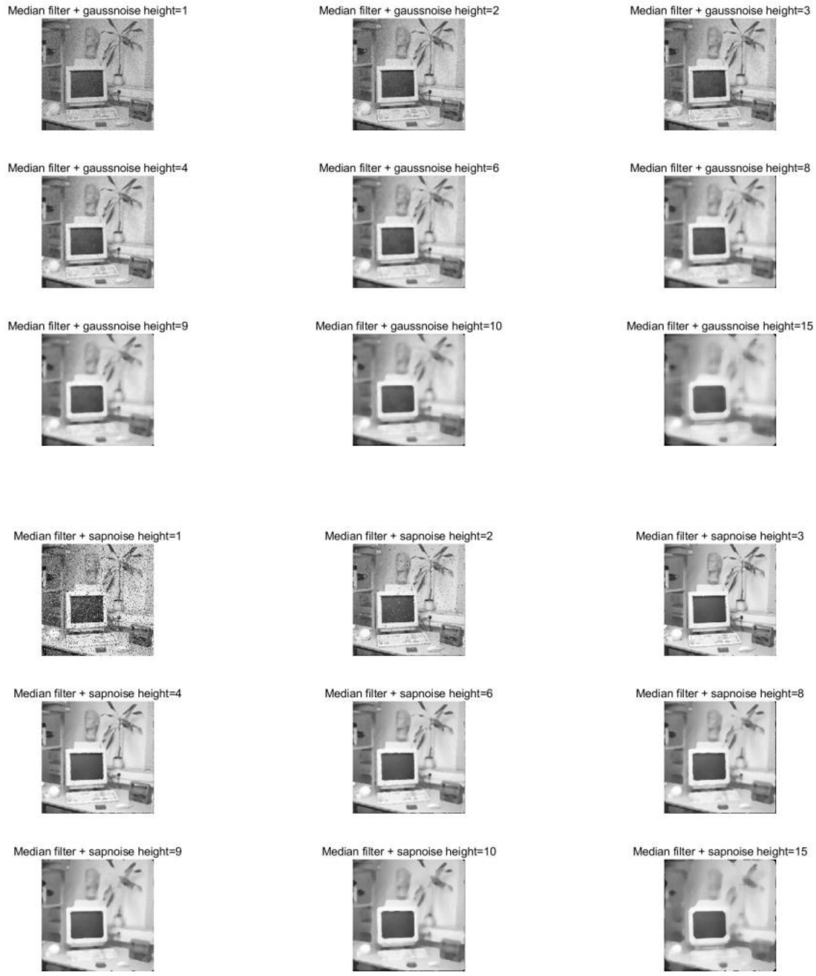
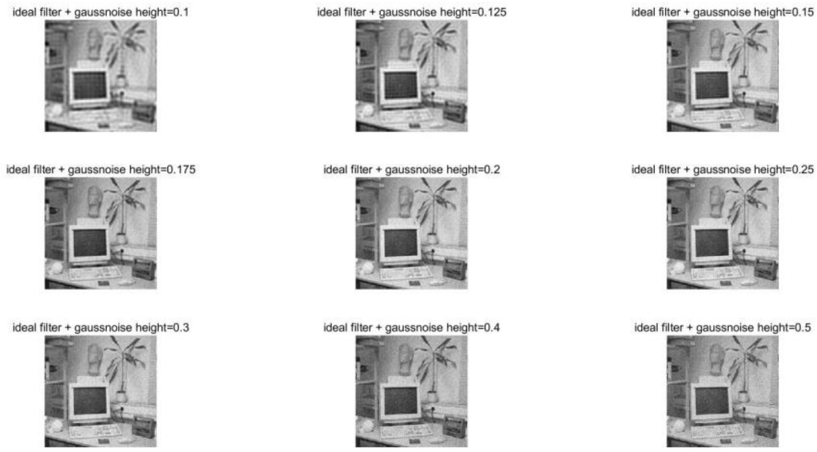


Figure 19: Median filter



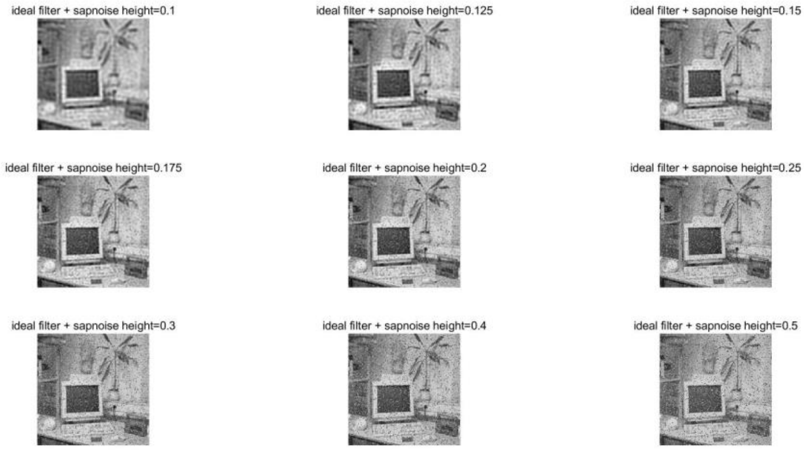


Figure 20: Ideal low-pass filter

The advantages and disadvantages of these three filters are shown in the table below:

filter	advantages	disadvantages
Gaussian	1. good at handling Gaussian noise 2. much more smoothing than other filters	1. images get more blurred when variance becomes larger, i.e. easy to lose edge information 2. not good at handling pepper-and-salt noise.
Median	1. good at handling pepper-and-salt noise 2. can eliminate local extreme value 3. better for keeping high frequency information	1. images look like paintings when windows size becomes larger. 2. The content in the image becomes illegible
Ideal low-pass	1. Simple principle, easy implementation and simple use	1. The image is prone to distortion and frequency leakage (ripple effect) 2. Not smooth enough 3. The performance of both kinds of noise is not very good

As for how the results depend on the filter parameters:

filter	effect of parameters
Gaussian	images get more blurred when variance becomes larger
Median	images look more like paintings and it's harder to distinguish objects when windows size becomes larger.
Ideal low-pass filter	images becomes more distorted and blurred when cut-off frequency goes smaller, and the impact of frequency leakage is more significant. When the cut-off frequency is too high, the image contains too much noise

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

Answers:
Gaussian filter makes images go blurring when variance becomes larger, and it is good at remove Gaussian noise. Median filter makes images more like paintings with parameters go bigger, and it performs better when handling pepper-and-sail noise with better ability to retain high frequency information. Ideal low-pass filter shows the ideal filtering effect, and cuts off the frequency completely according to the needs. However, due to the ripple effect, the filtering effect is not very good for the observable image.
This shows that in practical application, different filters should be selected according to the desired effect, and some information can be discarded when necessary to obtain better visual effect.

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:
The results are shown in **Figure 21**, and we can see when subsampling the smoothed variants are still smoother than the original image.

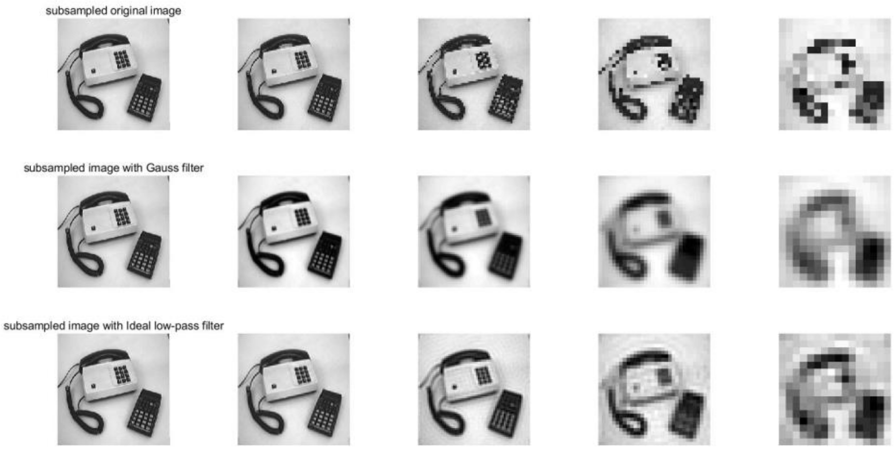


Figure 21: subsampling with Gaussian and ideal low-pass filter

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:
Smoothing before sampling can effectively reduce the loss of high frequency components, that is to say, more information can be retained. This can be explained by the smoothing principle and sampling theorem: after using the above two filters for smoothing filtering, part of the high-frequency components are discarded, so the maximum frequency of the image becomes smaller. According to the sampling theorem, the sampling frequency should be greater than or equal to half of the maximum frequency (also known as Nyquist frequency). Therefore, compared with the original image, the smoothed image can meet the sampling frequency with a lower frequency, so the information loss can also be reduced.