Answers to questions in Lab 1: Filtering operations

Name: Hanqi Yang　　　　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.

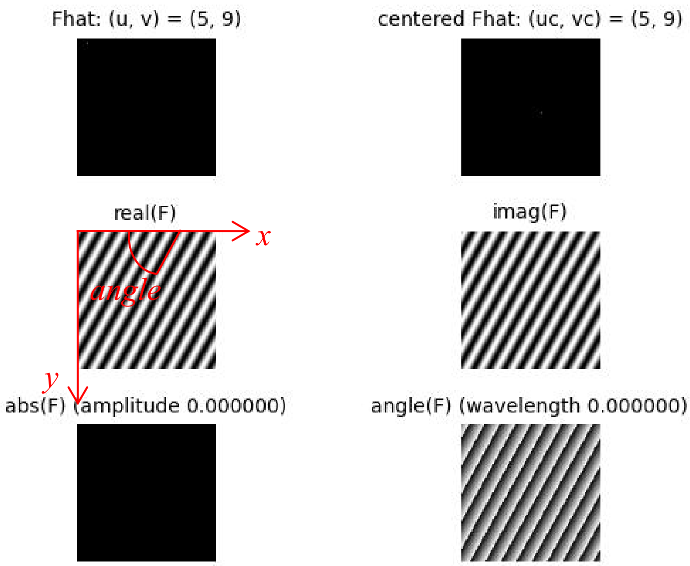
Good luck!

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:
According to the given point in frequency domain, the sine wave has different frequencies in the x and y directions respectively, which causes the sine wave to look different in sparsity and angle.

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:
A position (p,q) in the Fourier domain means the frequency variable $\omega = (\omega_1, \ \omega_2)$, where $p = \omega_1, q = \omega_2$. A point $(p,q) = (5,9)$ represents a frequency component of $\omega_1 = 5$ in x-direction and $\omega_2 = 9$ in y-direction. Therefore, this frequency position makes the sine wave have an angle of $\arctan(9/5)$ with the x-axis. The Python figure is shown below.



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:
Using the equation of inverse Fourier Transform (FT), we can derive the 2-D discrete FT as

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

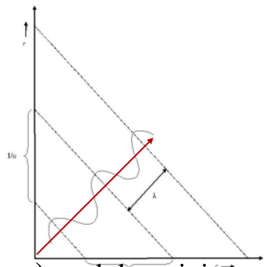
The amplitude should satisfy

$$A = \frac{1}{MN}\max\left(\left|\hat{f}(u,v)\right|\right)$$

So the code is complemented as

```
amplitude = np.max(np.abs(Fhat[u, v]))/(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:
According to the lecture notes, the wavelength can be written as

$$\lambda = \frac{1}{\sqrt{u^2+v^2}}$$

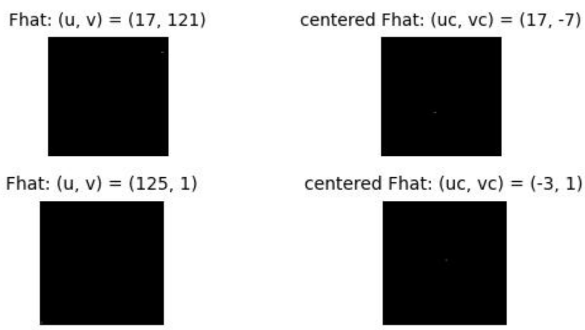
So the code is complemented as

```
wavelength = 1/np.sqrt(u**2 + v**2)
```

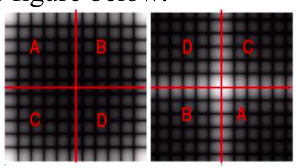
The direction of the sine wave is determined by the line between (p,q) and the origin.

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:
When we pass the point in the center and either p or q exceeds half the image size, its frequency will become negative and the negative part follows the relation $\omega' = \omega - sz = \omega - 128$, where ω' is the new frequency.



That is because the discrete FT is periodic in the image plain with period sz. So FT is symmetric around the origin as the figure below.



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

Answers:
The point (u,v) is centered by the code and transformed to (uc,vc). But this point is not used in computation but in the title of center $\hat{F}$. The function np.fft.fftshift(Fhat) will shift the zero-frequency component of $\hat{F}$ to the center of the spectrum.

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!

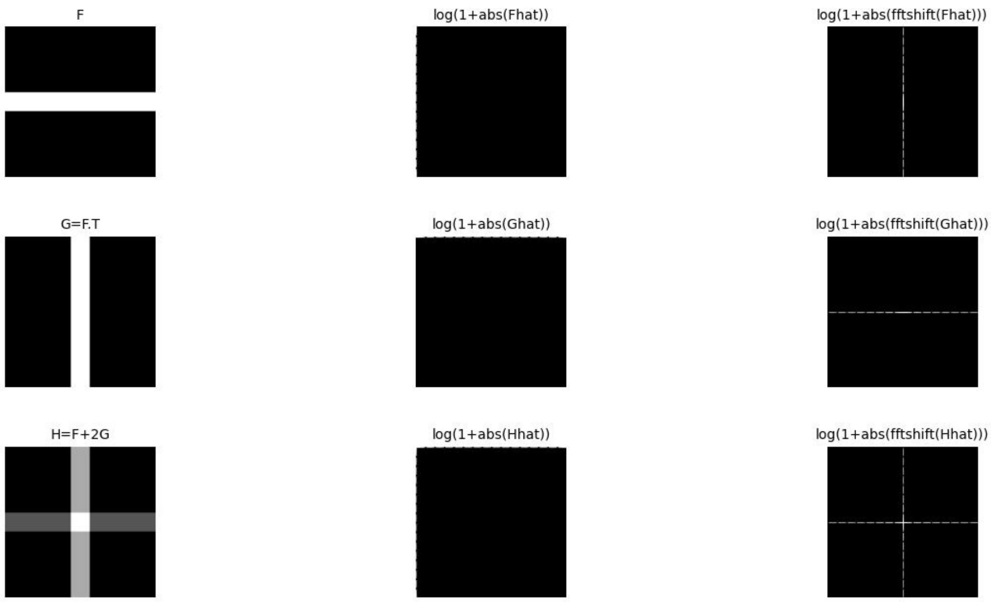
Answers:
As shown in the figure F or G, $f(u,v)=1$ when $57 \le u \le 72$, otherwise $f(u,v)=0$.
A 2-D FT can be written as

$$\hat{f}(u,v)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-2\pi i\left(\frac{xu}{M}+\frac{yv}{N}\right)}$$

As the image only changes along y-axis, this equation can be simplified as

$$\hat{f}(u,v)=\frac{1}{\sqrt{MN}}\sum_{x=57}^{72} e^{-2\pi i\frac{xu}{M}}\sum_{y=0}^{N-1} e^{-2\pi i\frac{yv}{N}}=\frac{\delta(v)}{\sqrt{MN}}\sum_{x=57}^{72} e^{-2\pi i\frac{xu}{M}}$$

Taking the figure F as an example, it only changes along y-axis and is a box function in the spatial domain, which means it will be a sinc function along y-axis in the frequency domain. These Fourier spectra are all concentrated to the borders of the images, since the origin is in the upper left corner. After we do fftshift(), the Fourier spectra will move to the center of the images, as shown in the third column of the figure below.



Question 8: Why is the logarithm function applied?

Answers:
A sinc function varies a lot between the values around zero-point and the other values, which will cause the Fourier spectra around zero-point is much more distinct than it in the other

place. The logarithm operation $\log(1+|\hat{f}|)$ will enhance the low intensity values and compress the high intensity values, although it may lead to a loss of information.

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

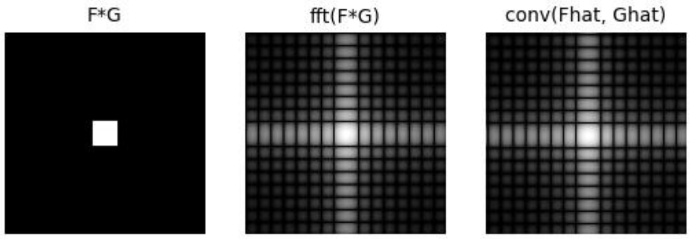
Answers:
The linear combination in the spatial domain will also lead to a linear combination in frequency domain as

$$F[af_1(x,y)+bf_2(x,y)] = a\hat{f}_1(u,v)+b\hat{f}_2(u,v)$$

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:
Multiplication in the spatial domain equals to convolution in the frequency domain.
The alternative computation can be implemented as

```
np.log(1 + np.abs(convolve2d(Fhat, Ghat, mode='same', boundary='wrap')/(128**2))
```

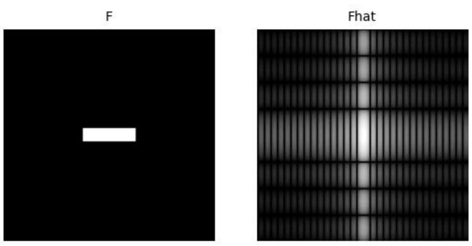


Question 11: What conclusions can be drawn from comparing the results with those in the previous exercise? See how the source images have changed and analyze the effects of scaling.

Answers:
Compare to the square in Question 10, a rectangular scales up in the x-axis and scales down in the y-axis, which will follows the relation below.

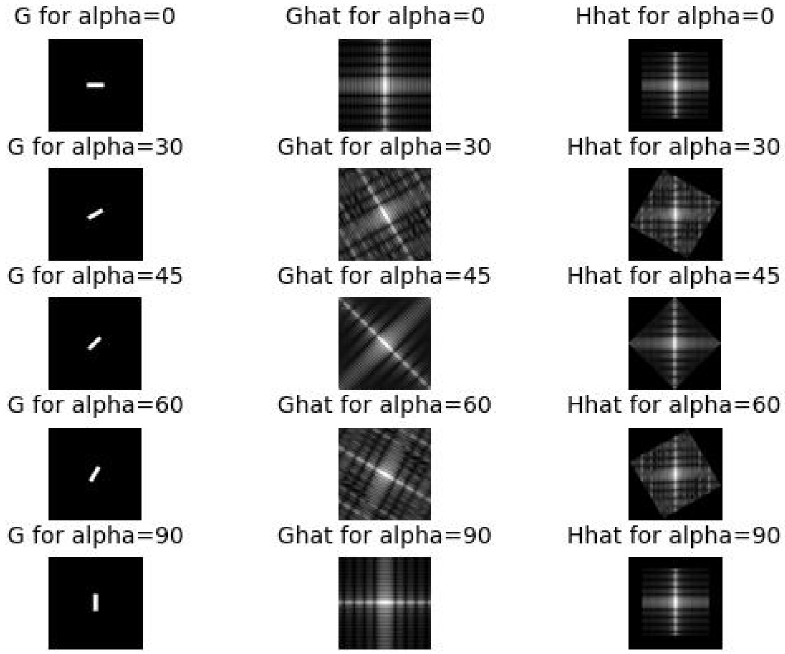
$$f_1(x,y) = f_2(S_x x, S_y y) = f_2(2x, \frac{1}{2}y)$$

$$\hat{f}_1(u,v) = \frac{1}{|S_x S_y|}\hat{f}_2\left(\frac{u}{S_x},\frac{v}{S_y}\right) = \frac{1}{1}\hat{f}_2\left(\frac{u}{2},\frac{v}{\frac{1}{2}}\right) = \hat{f}_2\left(\frac{u}{2},2v\right)$$



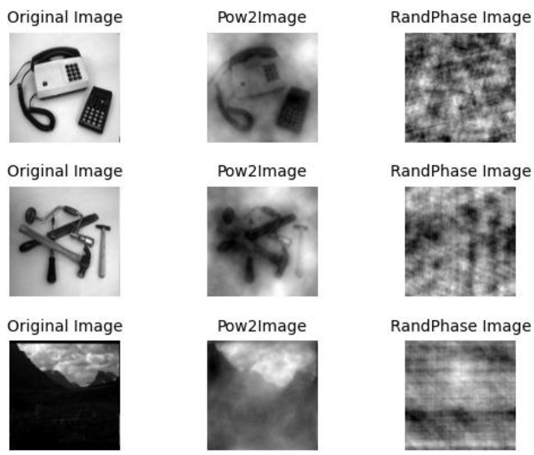
Question 12: What can be said about possible similarities and differences? Hint: think of the frequencies and how they are affected by the rotation.

Answers:
Rotation in the spatial domain will cause a rotation in the frequency domain by the same angle and the magnitude as well as the phase will not change.
But we can see some waves in the Fourier spectra when alpha=30 and 60 due to the poor resolution.



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

Answers:
Phase defines how waveform is shifted along its direction and magnitude defines how large the waveform is.
The second column has the same phase but different magnitude compare to the first column. We can still recognize what object or scene is in the picture. But the third column has the same magnitude but different phase compare to the first column. So we can’t recognize anything.

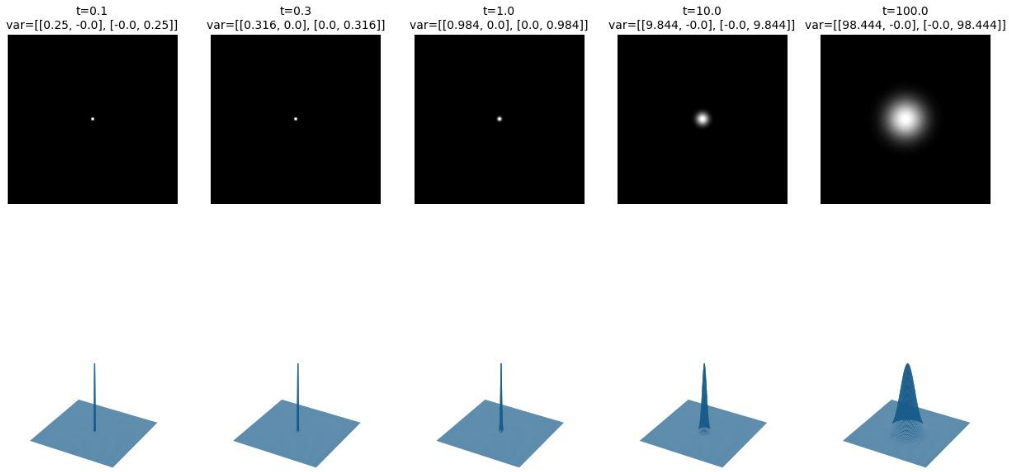


Question 14: Show the impulse response and variance for the above-mentioned t-values. What are the variances of your discretized Gaussian kernel for t = 0.1, 0.3, 1.0, 10.0 and 100.0?

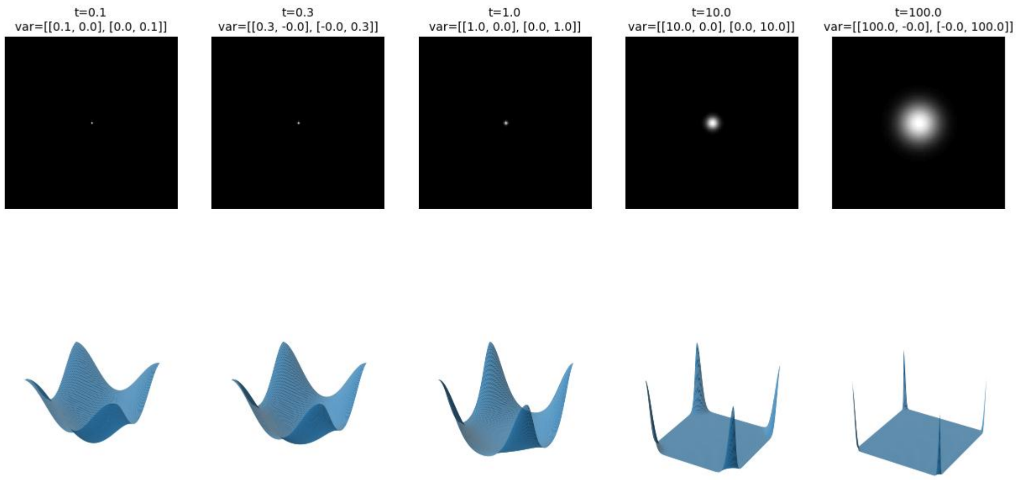
Answers:

t	Variances of discretized Gaussian kernel
0.1	0.250
0.3	0.316
1.0	0.984
10.0	9.844
100.0	98.444

Result of gaussfft.py



Result of discgaussfft.py (ideal case)



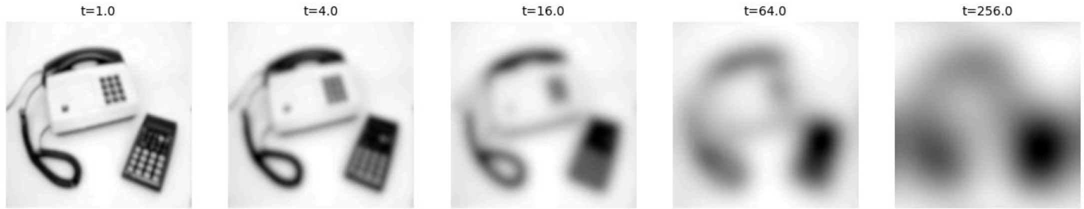
Question 15: Are the results different from or similar to the estimated variance? How does the result correspond to the ideal continuous case? Lead: think of the relation between spatial and Fourier domains for different values of t.

Answers:

The result is different form the estimated variance. When the variance is below 1.0, the different result is due to the discretization of a small amount of pixels. When the variance is larger and equal to 1.0, the difference is due to the improper code writing.

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?

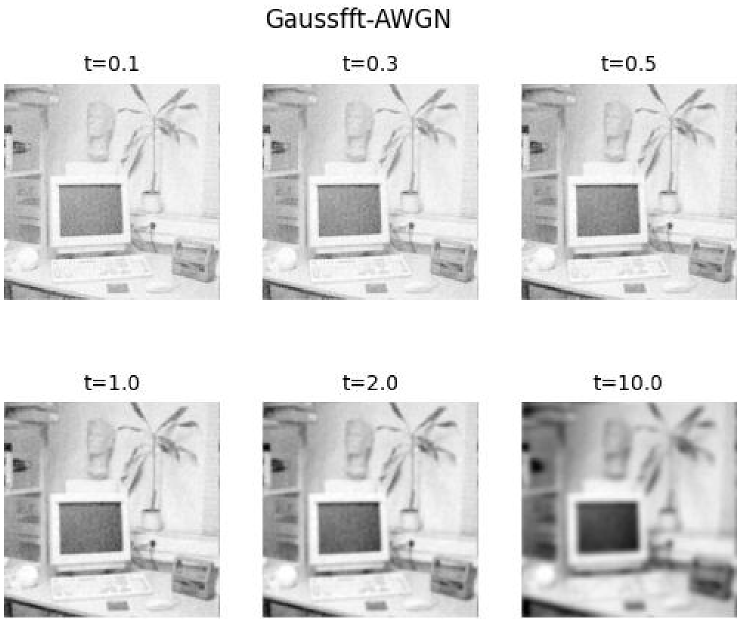
Answers:
The higher the t-value, the more blurred the image will be. A higher t means a lower cut-off frequency of the Gaussian kernel, which means the component with high frequency will be lost, such as edges and corners.

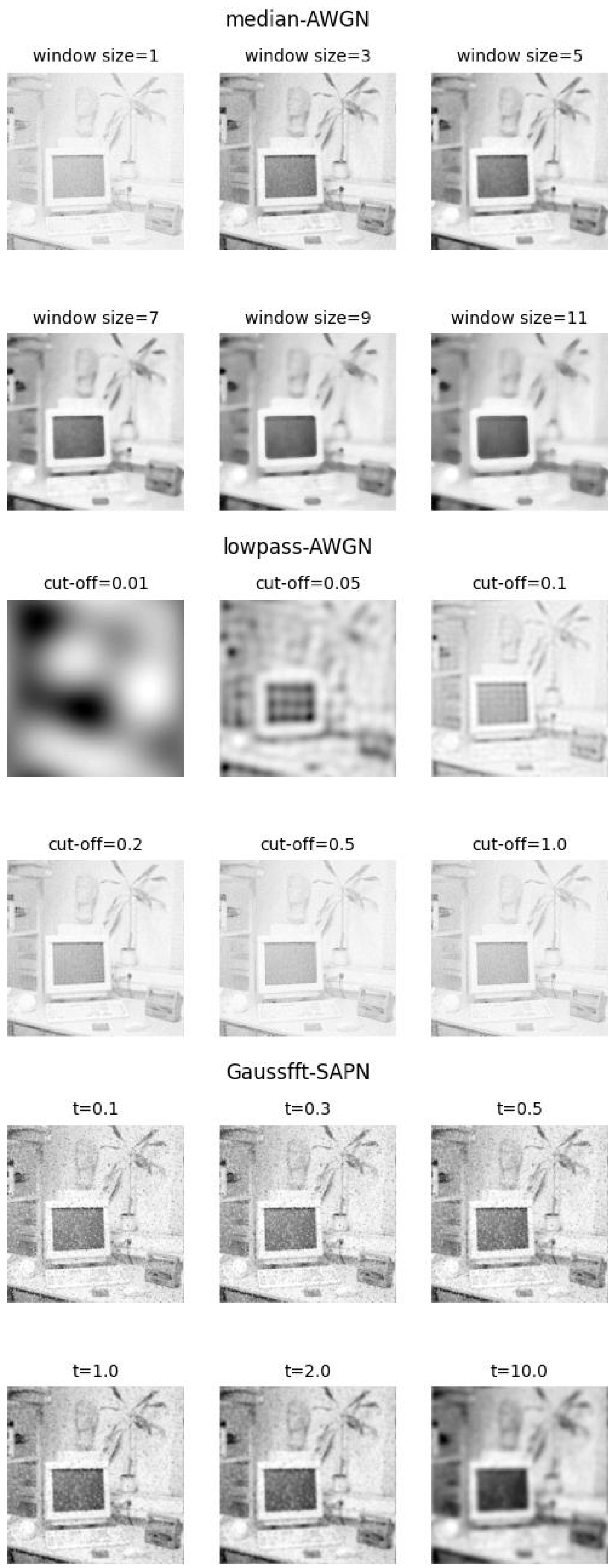


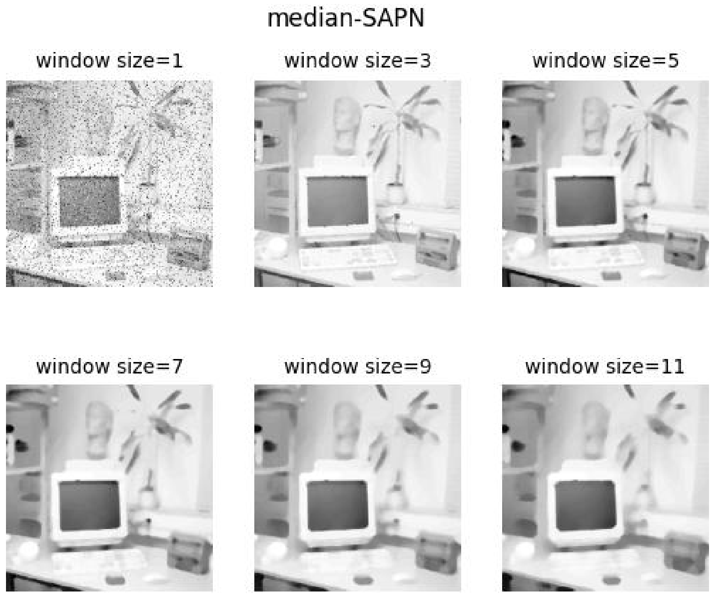
Question 17: What are the positive and negative effects for each type of filter? Describe what you observe and name the effects that you recognize. How do the results depend on the filter parameters? Illustrate with Matlab figure(s).

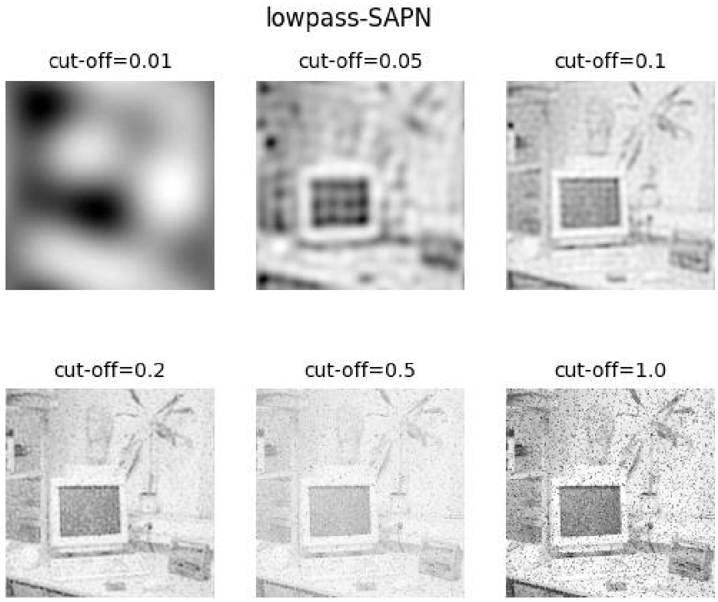
Answers:
This figure shows the comparison between the original image and two noisy images.

For additive white Gaussian noise (AWGN), Gaussian smoothing has the best result. Gaussian smoothing restores more detail to the original image relative to the other two filters. A higher variance in Gaussian smoothing, a higher window size in median filter and a lower cut-off frequency in low-pass filter will cause a blurrier image.
For salt-and-pepper noise (SAPN), median smoothing has the best result.




median-AWGN
window size=1
window size=3
window size=5
window size=7
window size=9
window size=11
lowpass-AWGN
cut-off=0.01
cut-off=0.05
cut-off=0.1
cut-off=0.2
cut-off=0.5
cut-off=1.0
Gaussfft-SAPN
t=0.1
t=0.3
t=0.5
t=1.0
t=2.0
t=10.0





Effects of filters:

- Gaussian filter:

  Positive: Good at reducing AWGN. Smooth the image.

  Negative: Bad at reducing SAPN. Blur the edges.
- Median filter:

  Positive: Good at reducing AWGN especially SWAP. Preserve edges well when reduce the noise.

  Negative: Eliminate small details such as dots and lines in images.
- Ideal low-pass filter:

  Positive: Good at reducing high frequency noise.

  Negative: In actual images, high-frequency components often imply details. Cause ringing effect.

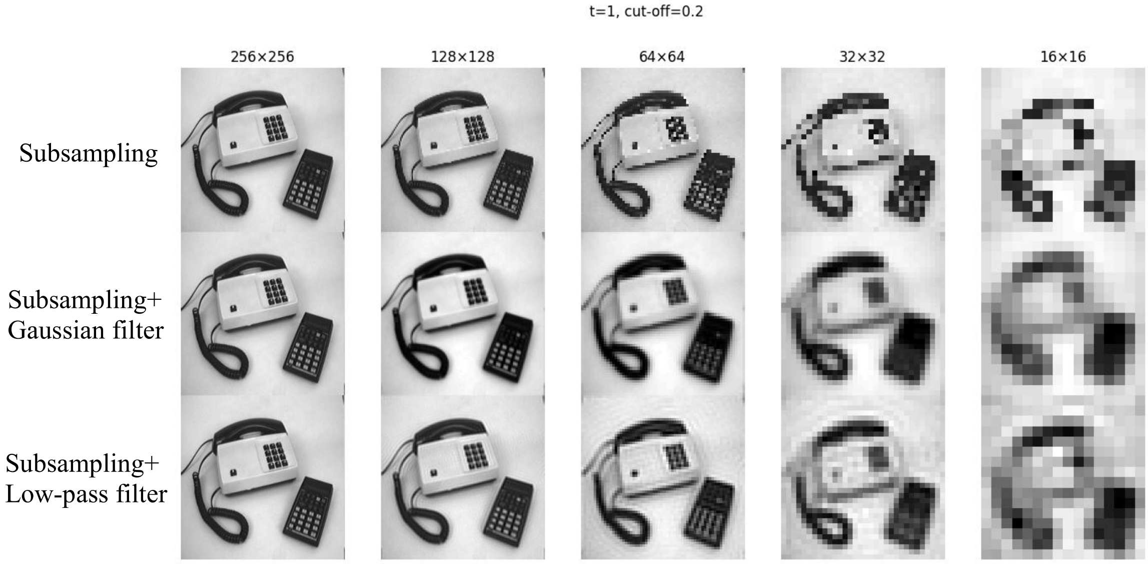
Question 18: What conclusions can you draw from comparing the results of the respective methods?

Answers:

Gaussian filter: This is also a kind of low-pass filter since they have similar working principle. The increase of variance equals to the decrease of cut-off frequency in low-pass filter, the details in the image will be eliminated and then we can see a blurrier image.
Median filter: This filter preserves the edges well since it takes the median value in the window instead of mean value. And it also work well with SWAP.
Ideal low-pass filter: It has almost no application in real life due to its single application scenario. And it works worst among the three filters in the two cases.

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 of Gaussian filter and low-pass filter in the last two iteration have much smoother edges than that of subsampling. But there is obvious ringing effect using low-pass filtering.



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:
According to the Nyquist theorem, the sampling frequency should be greater than the Nyquist Rate, which is twice the maximum frequency in the signal. If the Nyquist theorem is not satisfied, the aliasing will happened.
If we subsampling the original image without any pre-processing, the Nyquist Rate will be very high and we will get a bad result of the sampled image. A Gaussian filter or a low-pass filter will cut off the high frequency component, resulting a drop of the Nyquist Rate, which means the sampling rate may reach or approach the Nyquist Rate and the information loss will decrease.
