Answers to questions in
Lab 1: Filtering operations

Name: Prashant Kumar　　　　　　　　　　Program: TMAIM

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:

Points of observations:

- The amplitude of all the images are same.
- The position of the coordinates Point (p,q) relative to the point Origin (0,0) in the frequency domain determines the direction of the wavelength.
- With the increasing distance between Point(p,q) and origin 0(0,0), the wavelength becomes smaller which means higher frequency of the waveform.

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 point in a fourier domain is a complex number representation having a real and imaginary part or cosine and sine wave combination. To project a point (p,q) in fourier domain to a point in spatial domain we need to use inverse fourier transformation equation which is given by the equation:

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

$$f(m,n) = \frac{1}{\sqrt{MN}} \sum_{M=0}^{M-1} \sum_{N=0}^{N-1} \hat{f}(u,v)[cos(2\pi(\frac{mu}{M} + \frac{nv}{N})) + isin(2\pi(\frac{mu}{M} + \frac{nv}{N}))]$$

Example: If we consider the point (p,q) = (5,9), then the real part of the waveform can be represented as sinusoidal curve like the one below:

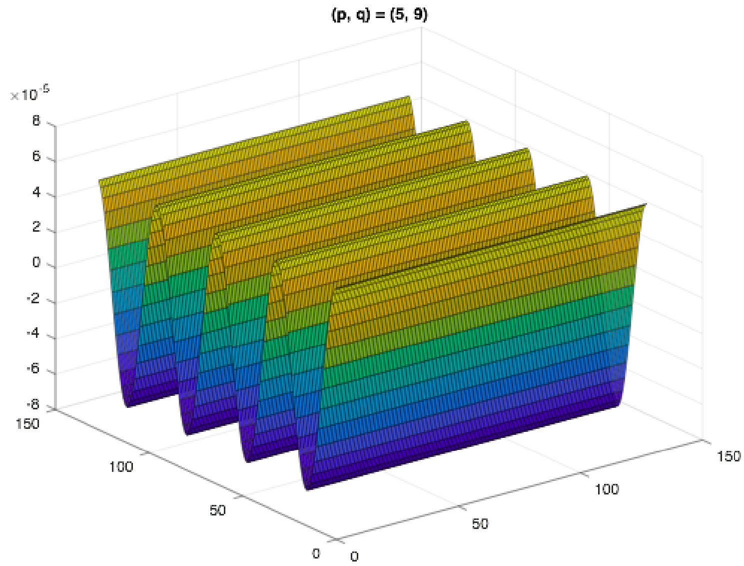


Figure 1 : Real part of the waveform for point (p,q) = (5,9)

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: Equation 4 states the Inverse Fourier Transform, which is:

$$f(m,n) = \frac{1}{\sqrt{MN}} \sum_{M=0}^{M-1} \sum_{N=0}^{N-1} \hat{f}(u,v)[cos(2\pi(\frac{mu}{M} + \frac{nv}{N})) + isin(2\pi(\frac{mu}{M} + \frac{nv}{N}))]$$

Here, the amplitude should be:

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

Where M = N = 128 and max amplitude is 1 so we get amplitude = 1/128.

Code Snippet:

```
w1 = 2 * pi * uc /sz;
 w2 = 2 * pi * vc /sz;
wavelength = 2 * pi / sqrt(w1^2 + w2^2);
amplitude = 1/sz;
```

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: The direction of travel of the waveforms in the spatial images is dictated by the position of the non-zero point (p, q) relative to the origin O(0, 0).

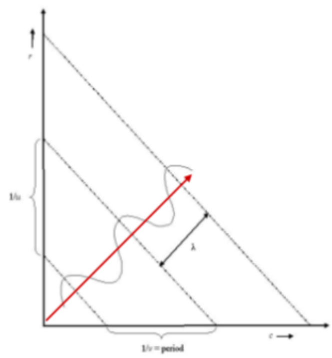


Figure 2 : Direction of the waveform

Wavelength expression

$$\lambda = \frac{2\pi}{||\omega||} = \frac{2\pi}{\sqrt{\omega 1^2 + \omega 2^2}}$$

Where $\omega_2 = 2\pi.\frac{uc}{M}$ and $\omega_2 = 2\pi.\frac{vc}{N}$ and M=N= sz which when solved gives

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

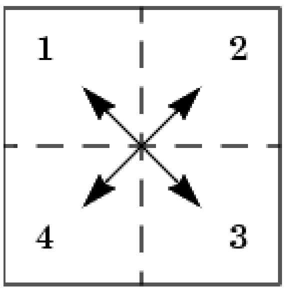
Code Snippet:

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

Assumption: p =u, q =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!

Answer: The discrete Fourier transform of the image is periodic with interval N. In the frequency domain, an image is repeated periodically in the 2D plain. The pixel value should satisfy the relation that $N + x,\ N + y = N - x,\ N - y$. As we can see, when the point moves towards the end it shifts side. The value becomes negative When the point is in the middle (1,1) the frequency is one and we see a sinusoid of frequency =1. This means that the regions of the original image exchange their position diagonally after the function fftshift is used.



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

Answers: The original image in the frequency domain visualizes the frequency within the interval $[0, 2\pi]$.The required instructions in the code in the script computes the coordinates of (p, q) after the function fftshift( ) (for centering the coordinate) is applied which displays the image of the frequency within the interval $[-\pi,\pi]$.

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: Image in spatial domain is considered to be finite like a box and hence any point lying outside the boundaries are considered to be 0 which can be mathematically represented as :

$$F(x, y) = \begin{cases} 1, & x_1 \le x \le x_2 \\ 0, & everywhere\ else \end{cases}$$

The FT of it as we know is:

$$\hat{F}(x, y) = \frac{1}{N} \sum \sum f(u, v) e^{-2\times\left(\frac{xu + vy}{N}\right)}$$

Considering the image first image F, f(u,v) =1 when its within the range {57,73} and 0 otherwise. Using the separability properties of FT and simplify it with the property of Kronecker delta function we get:

$$\mathcal{F}(F(x,y)) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cdot e^{-\frac{2\pi i(xu+yv)}{N}} =$$

$$\sum_{x=x_1}^{x_2} e^{-\frac{2\pi ixu}{N}} \sum_{y=0}^{N-1} e^{-\frac{2\pi iyv}{N}} = \sum_{x=x_1}^{x_2} e^{-\frac{2\pi ixu}{N}} \sum_{y=0}^{N-1} \mathbf{1} \cdot e^{-\frac{2\pi iyv}{N}} =$$

$$\delta(v) \cdot \sum_{x=x_1}^{x_2} \cdot e^{-\frac{2\pi ixu}{N}}$$

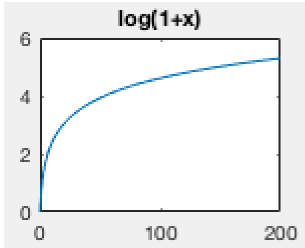
where we used the relation:

$$\mathcal{F}(\mathbf{1}) = \delta(v)$$

Since, $\delta(v)$ = 1 only when v = 0, Therefore, we can deduce that $\mathcal{F}(F(x,y))$ is non zero only when v = 0. Hence, that is why F's Fourier spectrum is concentrated in the left border.

Question 8: Why is the logarithm function applied?

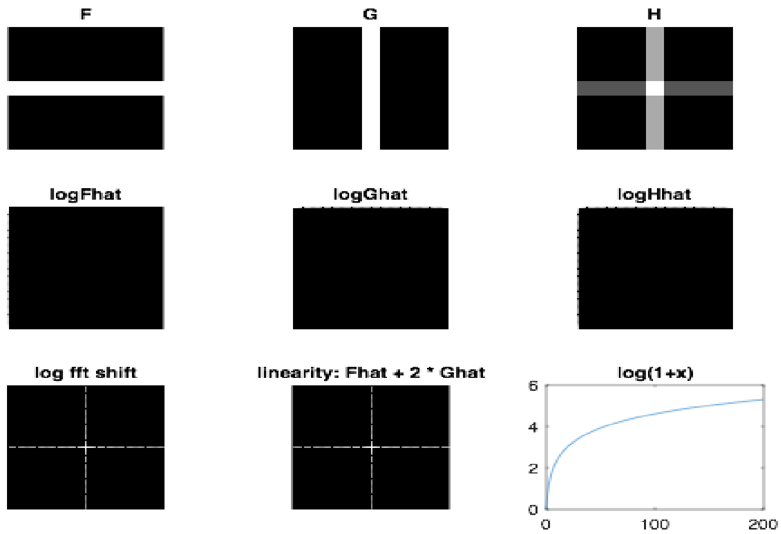
Answers: The distribution of the pixels' values concentrate on a lower range before logarithm is applied. Since logarithm transformation could expand the range of low pixel values and compress the range of high pixel values, it can redistribute the pixel values on the low range and enhance contrast in the image of frequency domain. After logarithm is used, the detail in the image becomes more visible, and the dynamic range of the pixel value become much smaller.



Power law distribution for logarithm

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

Answers: Referring to the first row of the image below:



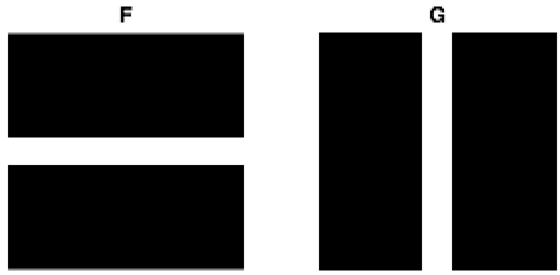
We can see here that, image H in the first row, its spectrum consists of the superposition of the spectra of images F, G. which clearly demonstrates the linearity feature of the Fourier Transform.

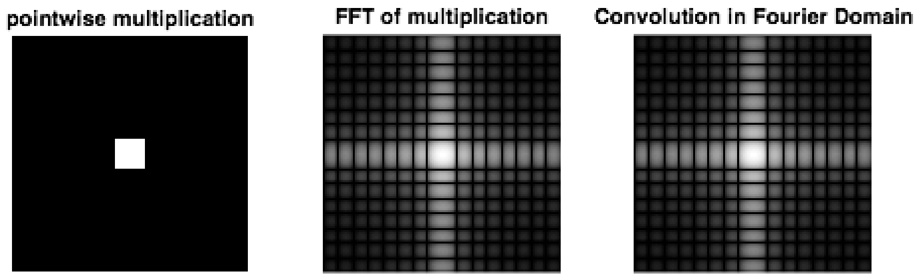
$$\mathcal{F}\left[a\, f_1(m,n) + b\, f_2(m,n))\right] = a\, \hat{f}_1(u,v) + b\, \hat{f}_2(u,v)$$

where a, b ∈ C are constants and f, g functions of real variable and F is the fourier transform which shows that if M images are combined in a linear way, their collective spectrum will consist of the same combination of their individual spectra.

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:

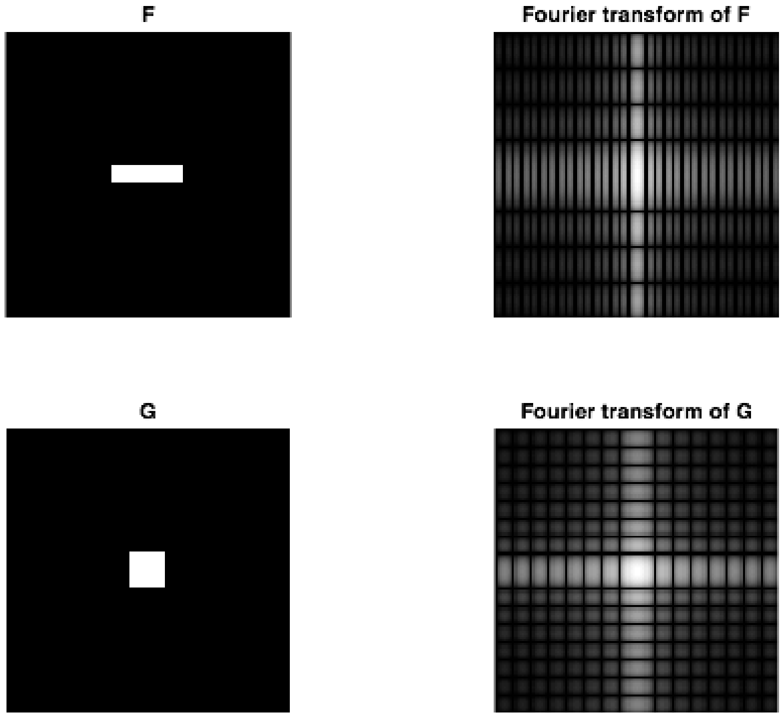




As we already know, multiplication in the spatial domain is same as convolution in the Fourier domain: $\mathcal{F}(fg) = \mathcal{F}(f) * \mathcal{F}(g)$, which we can observe from the images titled, FFT of multiplication -> fft2(F.*G) and Convolution in Fourier Domain -> conv2(Fhat,Ghat).

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:

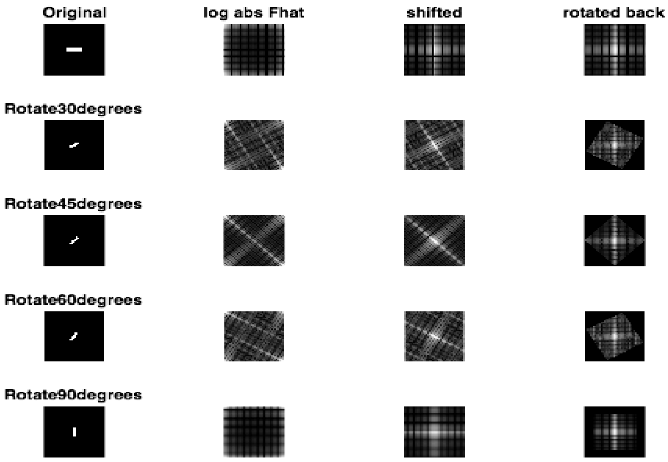


From what we can see here, the non zero region in the image F has been halved and the width has been doubled. Mathematically speaking, if we consider $f_1(x, y)$ the image " G " and $f_2(x, y)$ as a function for image " F ", we can deduce the relation:

$$f_2(x, y) = f_1\left(\frac{x}{2}, 2y\right) .$$

Recalling the scaling property of the Fourier transform, which states that compression in spatial domain is same as expansion in Fourier domain. $f(ax, by) \Leftrightarrow \frac{1}{|ab|}\hat{f}(\frac{u}{a}, \frac{v}{b})$ .We can deduce, $\hat{f_1}(u, v) = \hat{f_2}(2u, \frac{v}{2})$ which is why the number of spectrum zeros appear to be twice as many horizontally and only half vertically .

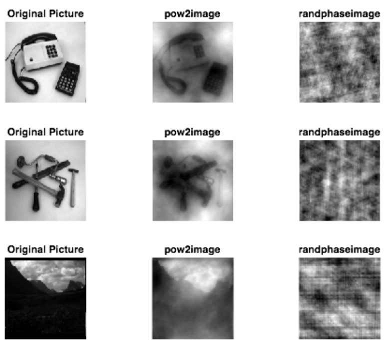
Question 12: What can be said about possible similarities and differences? Hint: think of the frequencies and how they are affected by the rotation.
Answer:



One of the properties of the Fourier Transform is Rotation, where the rotation of the original image of a function by an angle θ will result in the rotation of its Fourier transform by the same angle . However, because of the rotation, each image loses its original smoothness, due to the limited resolution and the nature of the shape of the pixels, with a degree of distortion depending on the angle of rotation. This has a direct effect on the Fourier transform of each image, as is clearly seen by distortion manifested in the wave-like patterns in the spectrum of the images rotated by 30 and 60 degrees.

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

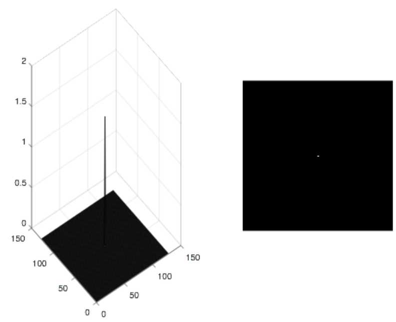
Answers:



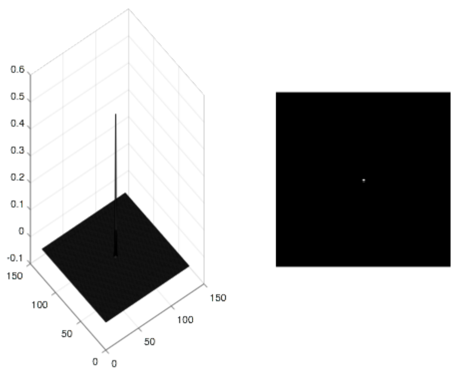
The phase is a measure of displacement all of the various sinusoids with respect to their origin. While the magnitude of the 2D Discrete Fourier transformation is an array whose components determine the intensities of the image the corresponding, phase is an array of angles that carry much of the information about where discernible objects are located in the image. The images in the second column here share the same phase as the first images, it easily distinguishable from one another for humans although we alter the magnitude. In the third column, the magnitude component is kept the same while the phase information is randomized which creates a mesh of bright and dark pixels and thus making the image indistinguishable.

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?

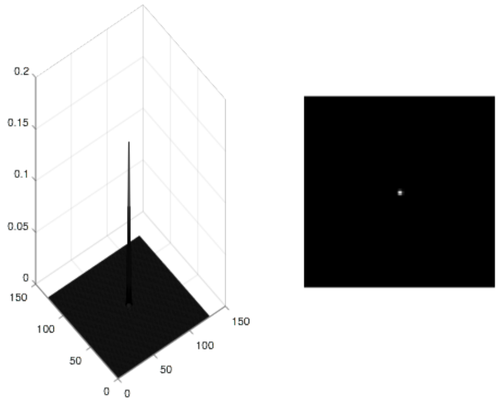
Answers:



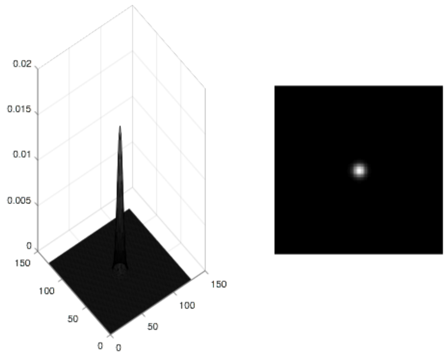
Impulse response t=0.1



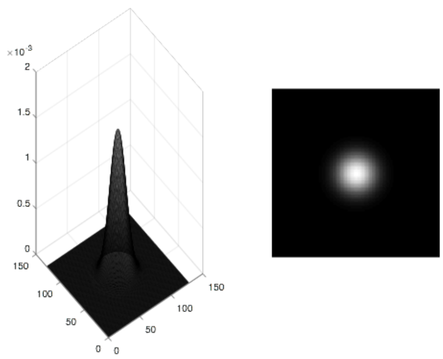
Impulse response t=0.3



Impulse response t = 1



Impulse response t = 10



Impulse t =100

t	Σ Discretized Gaussian Kernel
0.1	$\begin{bmatrix} 0.0133 & \\ & 0.0133 \end{bmatrix}$

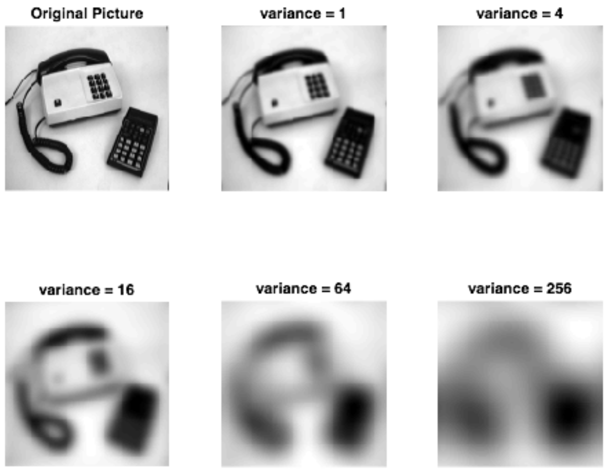
0.3	$\begin{bmatrix} 0.2811 & \\ & 0.2811 \end{bmatrix}$
1.0	$\begin{bmatrix} 1.000 & \\ & 1.000 \end{bmatrix}$
10.0	$\begin{bmatrix} 10.000 & \\ & 10.000 \end{bmatrix}$
100.0	$\begin{bmatrix} 100.000 & \\ & 100.000 \end{bmatrix}$

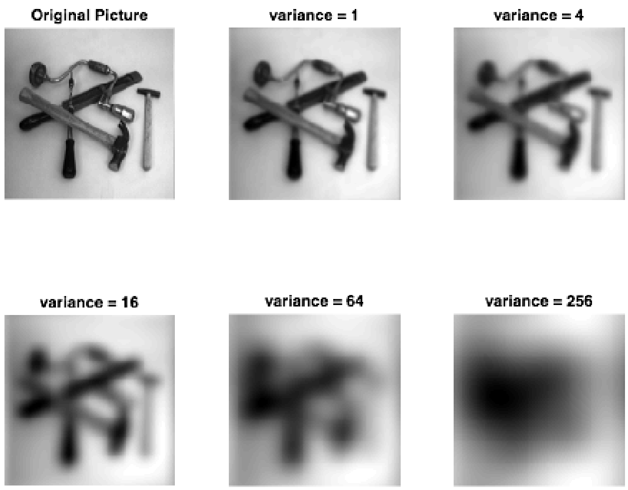
Question 15: Are the results different from or similar to the estimated variance? How does the result correspond to the ideal continuous case? Lead: think of the relation between spatial and Fourier domains for different values of t.

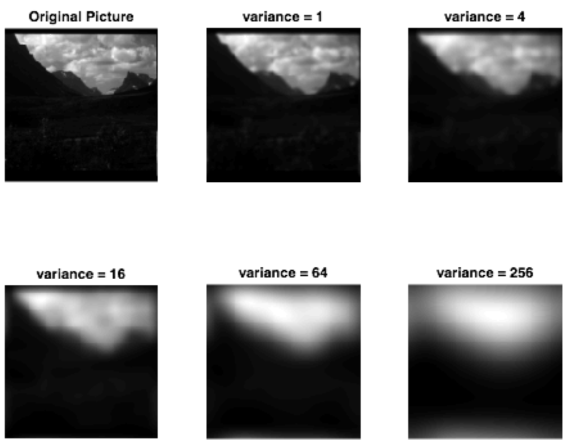
Answers: For the value of variance t below 1, the value of covariance and the estimated value has some degree of difference whereas when t is more than 1 it is aligned to the estimated value. It is because that when the variance is small, the values of the Gaussian filter kernel after sampling have a tendency to become non-Gaussian. The smaller the variance t, the fewer points sampled from the region within the standard deviations of the mean.

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:





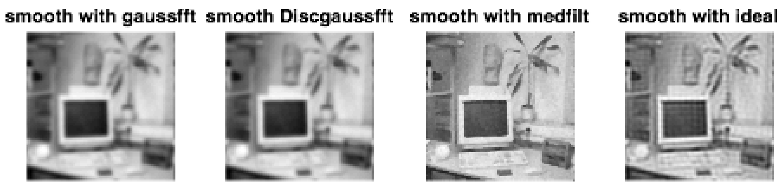


As the variance t increases, the images become more blur. This is reasonable because the high the variance t, the lower the cut-off frequency in the frequency domain for the Gaussian function. The lower the cut-off frequency is, the more components with high frequency, such as edges and corners, are lost. That's why we can observe that the images are becoming more blur as the value of t goes up. Hence, as t increases, the less details, that is, regions of higher frequencies, are preserved.

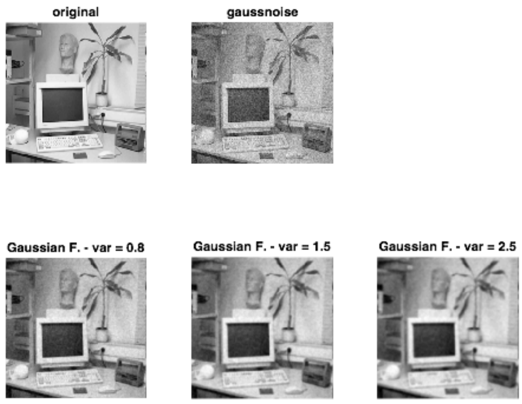
Question 17: What are the positive and negative effects for each type of filter? Describe what you observe and name the effects that you recognize. How do the results depend on the filter parameters? Illustrate with Matlab figure(s).

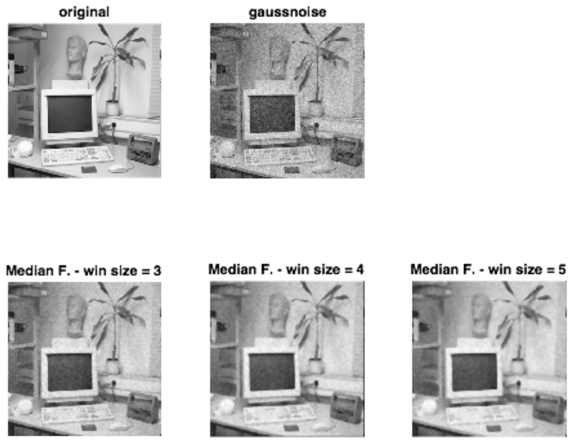
Answers:

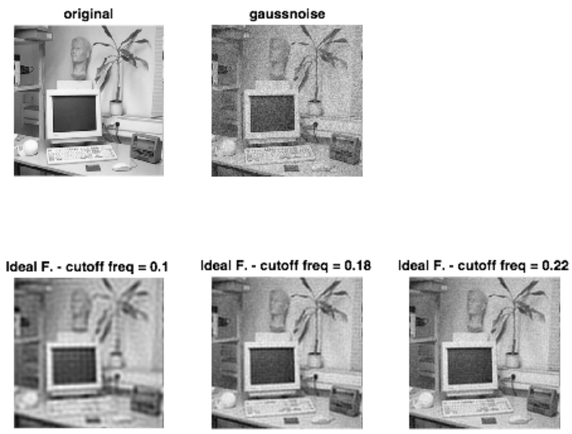
Smoothing of noisy data



Smoothing of Gaussian noisy data - Gaussian, median and ideal low-pass filters

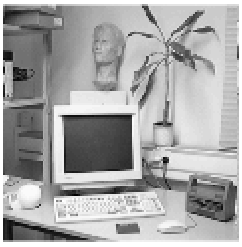






Smoothing of SAP noisy data - Gaussian, median and ideal low-pass filters

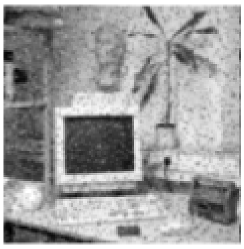
original



sapnoise

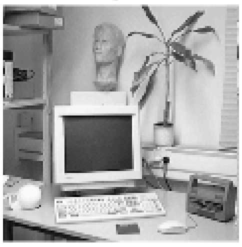
Gaussian F. - var = 0.8

Gaussian F. - var = 2



Gaussian F. - var = 5

original

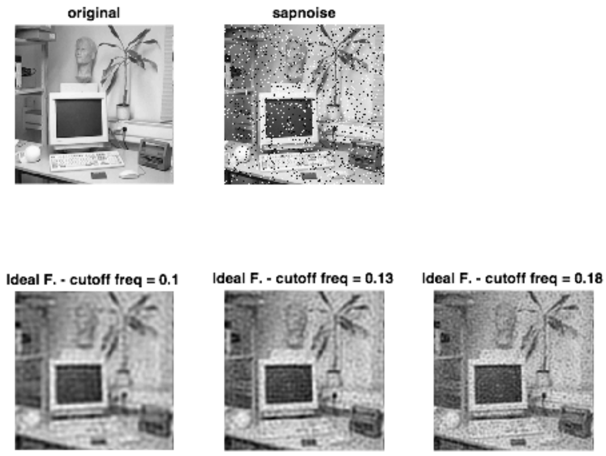


sapnoise

Median F. - win size = 3

Median F. - win size = 4

Median F. - win size = 5



<u>Gaussian:</u>

Removes “high-frequency” components from the image (low pass filter) which is used to blur images and remove detail and noise.More effective at smoothing images.

Positive effect is the removal of noise.

The negative effect is that important features get blurred. The effects we can see in all cases are blurring. We also observe that is really good at removing white noise but challenged by sap. It reduces sap but does not eliminate completely.Parameter: **σ**,The higher σ used, we observe more blurring and increased loss of detail.This is due to the fact that larger values of σ produce a wider peak in the gaussian distribution.

<u>Median:</u>

Used for noise reduction.Much better in removing sap noise due to the fact that neighboring pixels contribute to the calculation of a value (neighboring pixels are ranked based on their brightness), thus extreme values are rejected. Parameter: **window size** :Median performs good in removing both types of noise but it is much more effective in sap. In both cases though, using bigger windows results in great loss of detail. A positive effect is how well it removes sap noise.

A negative effect is the extreme blurring of sudden changes in important features or sharp changes (since they also correspond to high frequencies).It is also computationally expensive due to the sorting of pixels.

<u>Ideal:</u>

The ideal low-pass discards all frequencies above a threshold F, i.e. this filter passes low frequencies so image becomes blurred. Perhaps the positive of this filter is the fact that it is quite simple to implement?

The drawback of this method is that it tends to create images with ”ringing” at sharp boundaries. Ringing occurs due to the fact that the ideal low pass filter function is rectangular. The observation that the application of the low-pass filter is equivalent to the convolution of the image with the sinc function provides an explanation for this phenomenon.

Parameter: cut-off frequency. (0.2, 0.5, 0.9)

As the cut off frequency decreases, the image becomes more blurred, noise reduces.

The ringing artifacts increase as the amount of high frequencies are removed.

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

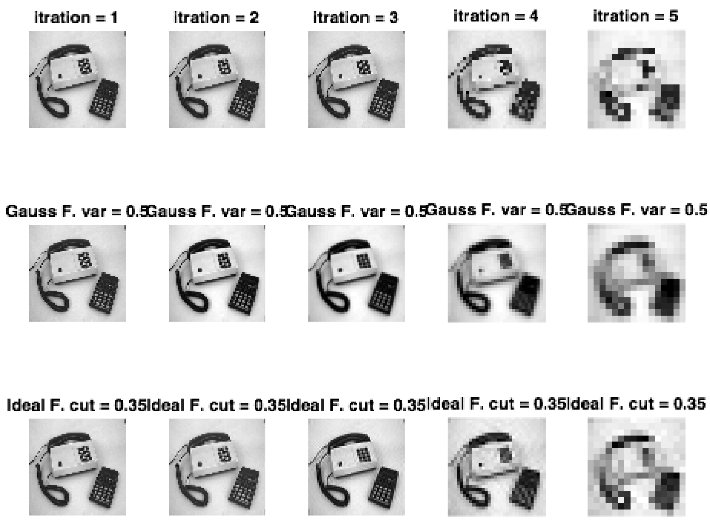
Answers: With regard to the Gaussian filter, we observed that when its variance increases, so does the blurring. This is reasonable since in the frequency domain the variance of the filter is the inverse of the one we use. That means that the higher the variance in the spatial domain, the lower the variance in the frequency domain. The lower the variance in the frequency domain, the more high frequencies are discarded. Hence only noise of relatively lower frequency remains present in the image.

The ideal low-pass filter can be seen as a perfect disc with radius r = D0 in the frequency domain. In the spatial domain, this filter is represented however by two components. An intense component in the origin and a component comprised by concentric circles around the first component. The first component is responsible for the blurring and the second one for the ringing effect. The lower the cut-off frequency, the higher the ringing effect is spread, that is, the ringing effect of each pixel reaches longer distances. Since a multiplication of the Fourier transform of two signals in the frequency domain is equivalent of a convolution in the spatial domain, the above latter component affects the filtered image in a way such that noise is not only maintained but also transformed. For instance, salt- and-pepper noise in the sap image is enhanced and magnified, resulting in large freckle-like shapes in the image.

In contrast to the two low-pass filters, the median filter is a nonlinear filter whose operations are centred only on a neighbourhood. As the filter works with the median of a neighbourhood of pixels, and not the mean, it can directly remove the effect that outliers have in images, that is, small regions of pixels affected by noise. The above two reasons are the reasons why this filter is so successful in removing 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.

Answers:



Sub-sampling results in pixels of bigger size, hence a neighbourhood of pixels are compacted into one, and the different values of all those pixels are lost. Hence, information is lost and shapes become coarser. If the sub-sampling occurs at a lower frequency than the Nyquist frequency then the image's characteristics are distorted irrevocably. The first thing noticed in this exercise is that there is an qualitative information loss balance: it is possible to smooth the sub-sampled images to a higher degree than the original images. Smoothing the original images in that degree results in information loss, just as the one introduced when sub-sampling. Since the two filters used here are low-pass filters, we can see that the outline in the image remains fairly accurate, even a higher variances, or lower cut-off frequencies. With regard to smoothing the sub-sampled versions of the original images, the same effects as given above the Gaussian filter introduces blurring and the ideal low-pass filter introduces ringing. The higher the resolution of the image, the higher the level of details preserved with regard to blurring, and the coarser the ringing effect.

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: From the above question we could conclude that smoothing before subsampling could prevent from losing information of the image. This can be explained by the sampling theorem. To lower the information loss during the process of the subsampling, the sampling frequency should be above the Nyquist frequency, which is half the maximum frequency. If we first smooth the image by Gaussian filter or low-pass filter, the maximum frequency will decrease, which causes a reduction of the Nyquist frequency. This means that the relatively low sampling frequency could also satisfy the requirement of the sampling theorem. In this way, the information loss will be decreased.

Laboration 1 in DD2423 Image Analysis and Computer Vision

891214-T572 Prashant Kumar
……………………………………………………………………...

Student's personal number and name (filled in by student)

....
Approved on (date) Course assistant