Vision and Perception

Image in the frequency domain

Sapienza Università di Roma

References

Basic reading:

- Szeliski textbook, Sections 3.4, 3.5

Additional reading:

The original Laplacian pyramid paper

- Burt and Adelson, “The Laplacian Pyramid as a Compact Image Code,” IEEE ToC 1983.

Overview of today's lecture

- Fourier series
- Frequency domain
- Fourier transform
- Frequency-domain filtering

The frequency domain

Is this claim true?

Jean Baptiste Joseph Fourier
(1768-1830)

The Fourier series claim (1807):

> *'<u>Any</u> univariate function can be rewritten as a weighted sum of sines and cosines of different frequencies.'*

Well, almost.

- The theorem requires additional conditions.
- Close enough to be named after him.
- Very surprising result at the time.

Is this claim true?

Jean Baptiste Joseph Fourier
(1768-1830)

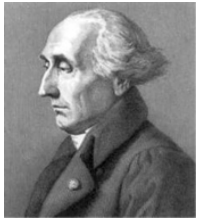
The Fourier series claim (1807):

> *'<u>Any</u> univariate function can be rewritten as a weighted sum of sines and cosines of different frequencies.'*

Well, almost.

- The theorem requires additional conditions.
- Close enough to be named after him.
- Very surprising result at the time.

Malus



Lagrange

Legendre

Laplace

The committee examining his paper had expressed skepticism, in part due to not so rigorous proofs

Fourier series

Basic building block

$$A \sin(\omega x + \phi)$$

Fourier's claim: Add enough of these to get <u>any</u> *periodic* signal you want

Basic building block

$$A\sin(\omega x + \phi)$$

$$T = \frac{2\pi}{|\omega|}$$

amplitude

sinusoid

angular frequency

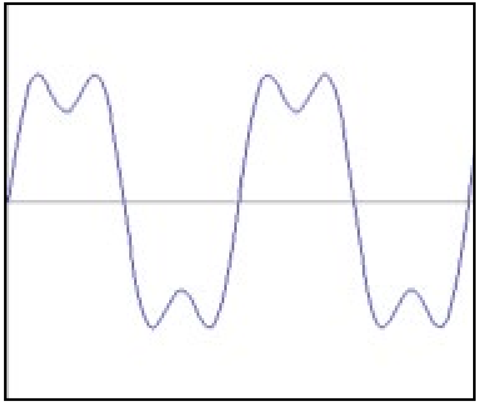
variable

phase

Fourier's claim: Add enough of these to get <u>any</u> *periodic* signal you want

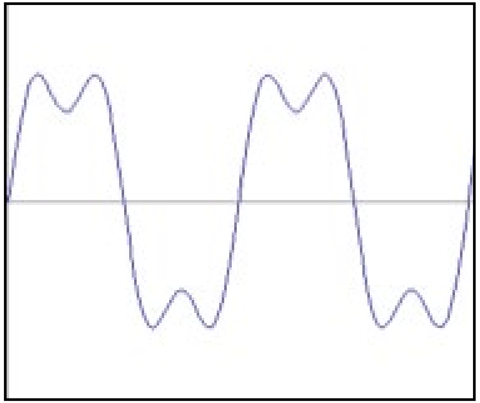
Examples

How would you generate this function?

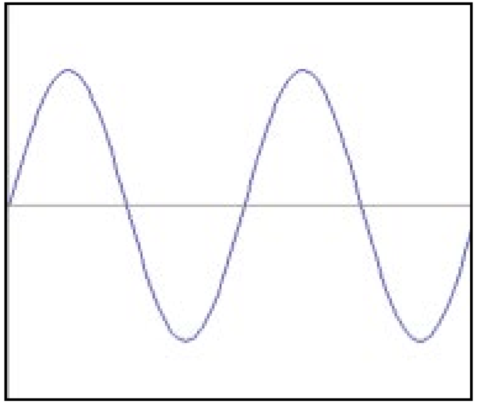
 $=$? $+$?

Examples

How would you generate this function?



=



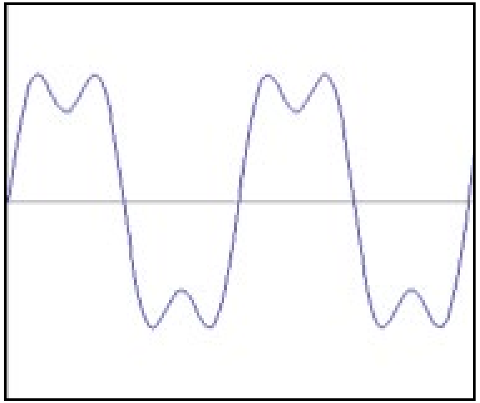
$\sin(2\pi x)$

+

?

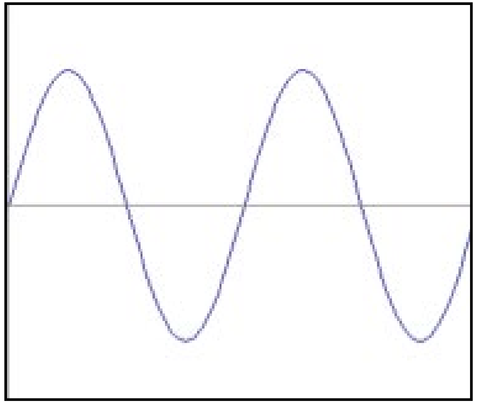
Examples

How would you generate this function?



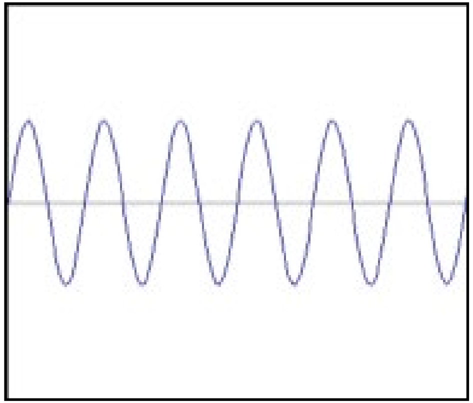
$$f(x) = \sin(2\pi x) + \frac{1}{3}\sin(2\pi 3x)$$

=



$$\sin(2\pi x)$$

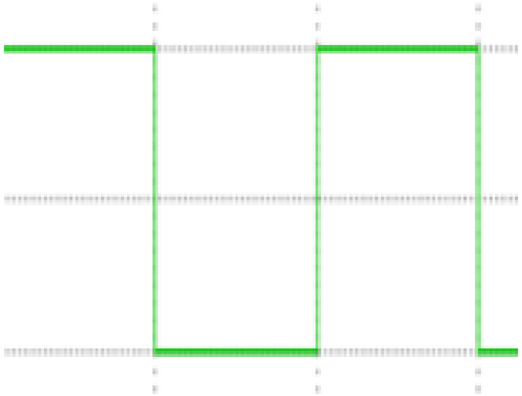
+



$$\frac{1}{3}\sin(2\pi 3x)$$

Examples

How would you generate this function?

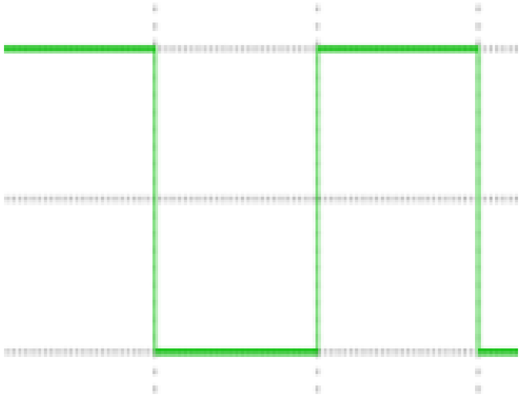


square wave

$$= \quad ? \quad + \quad ?$$

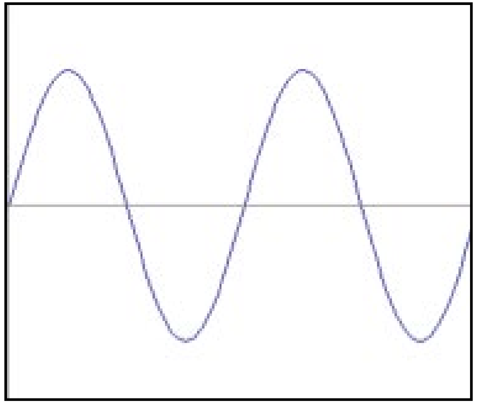
Examples

How would you generate this function?

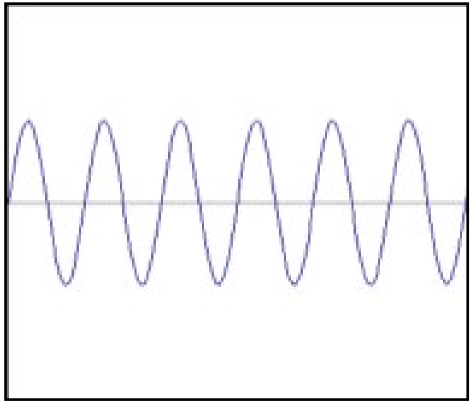


square wave

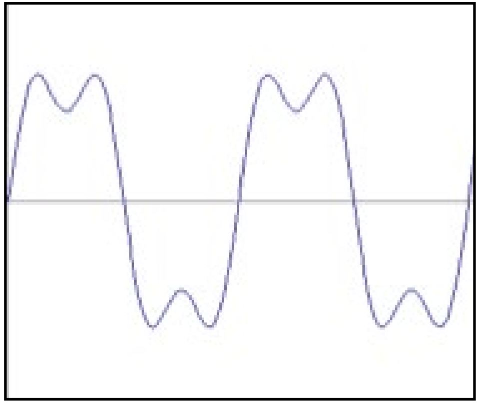
≈



+

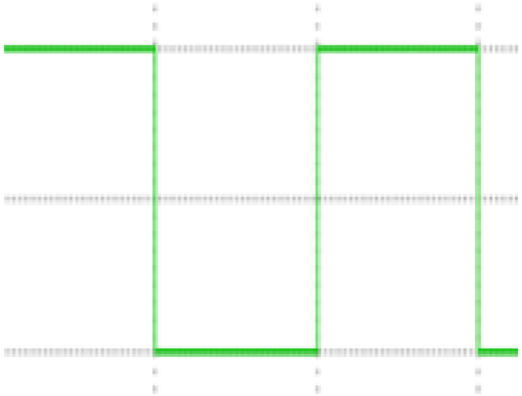


=



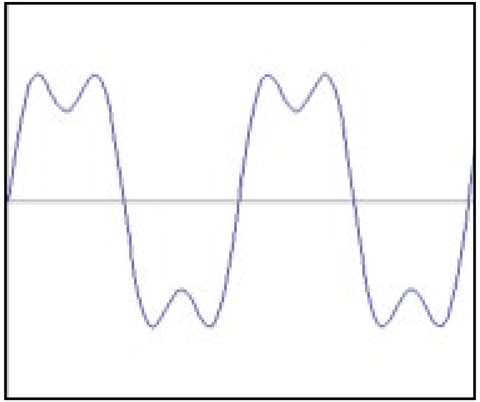
Examples

How would you generate this function?

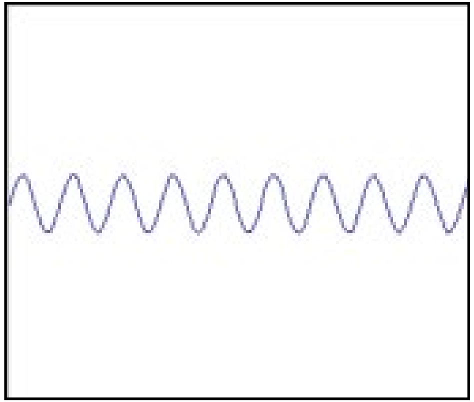


square wave

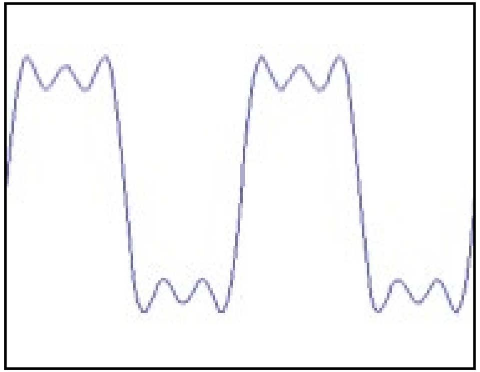
≈



+

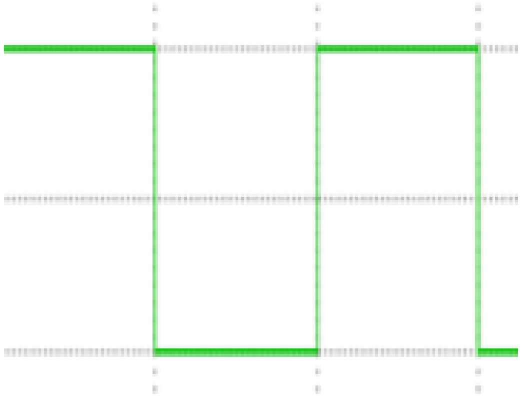


=



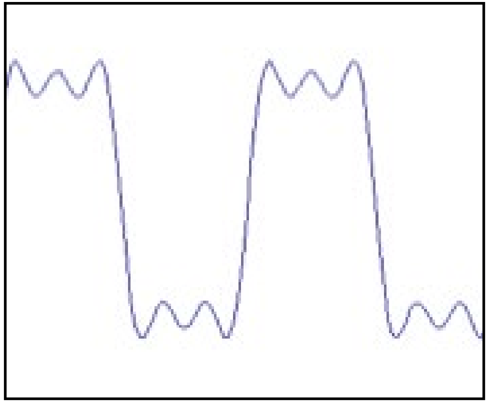
Examples

How would you generate this function?

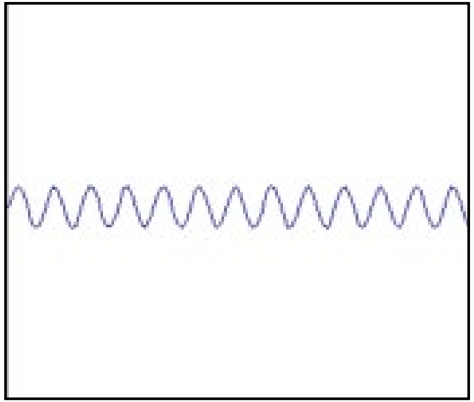


square wave

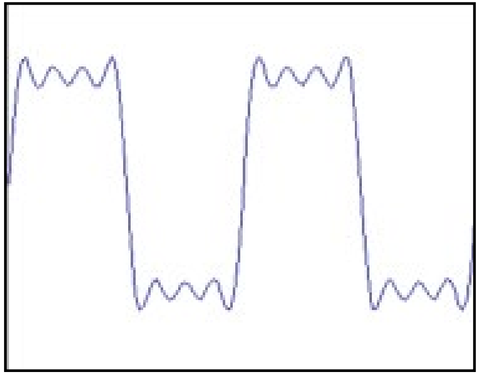
≈



+

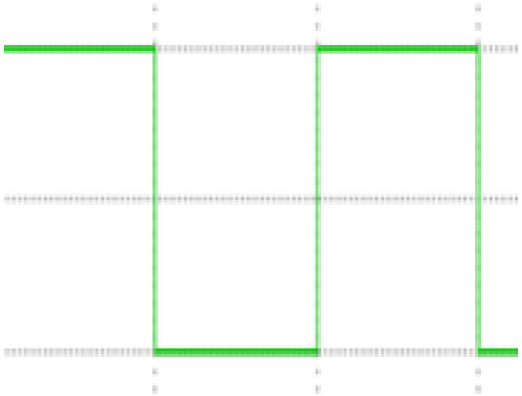


=



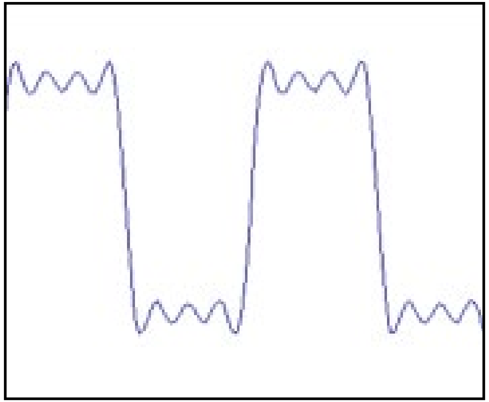
Examples

How would you generate this function?

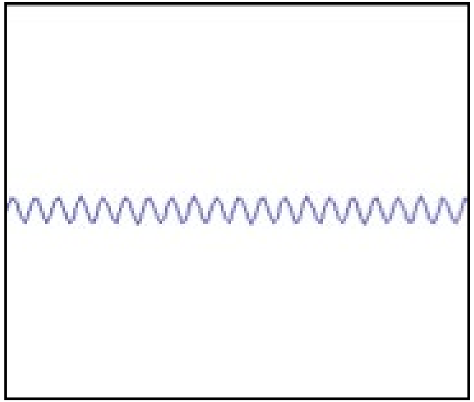


square wave

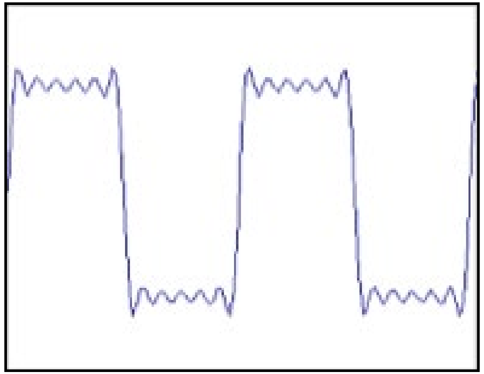
$\approx$



$+$

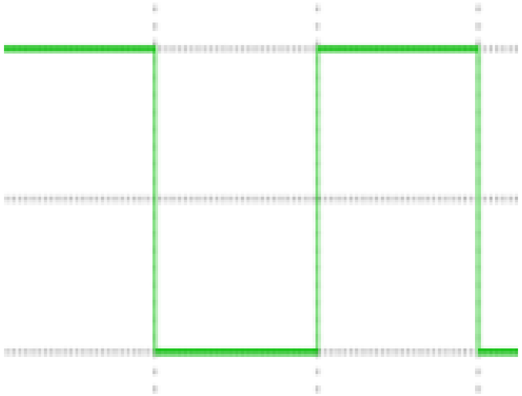


$=$



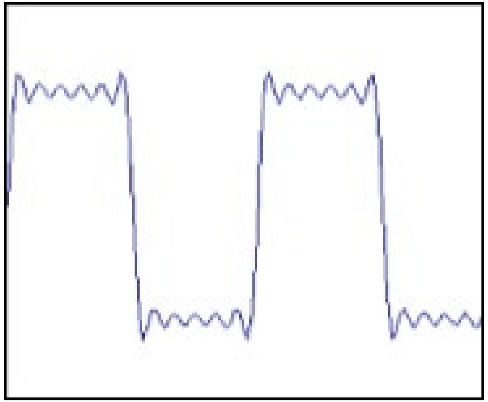
Examples

How would you generate this function?

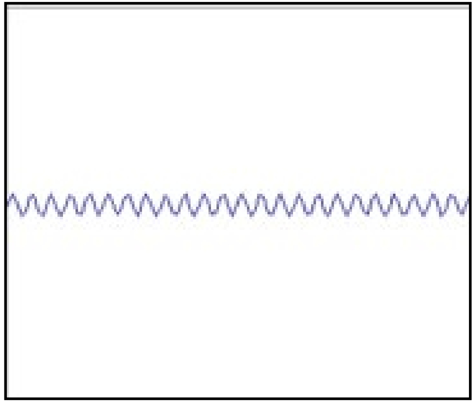


square wave

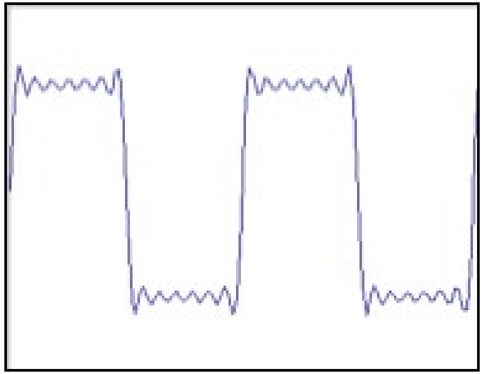
$\approx$



$+$

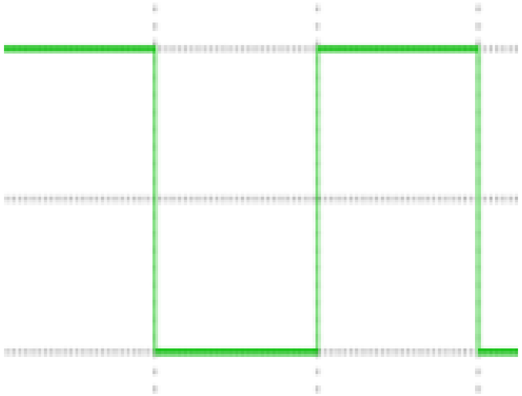


$=$



How would you express this mathematically?

Examples



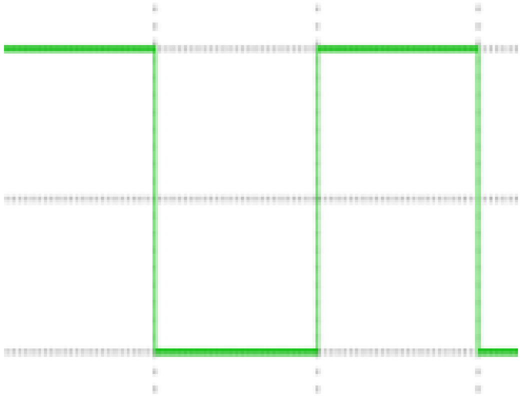
square wave

$$= \quad A\sum_{k=1}^{\infty}\frac{1}{k}\sin(2\pi kx)$$

infinite sum of sine waves

How would could you visualize this in the frequency domain?

Examples



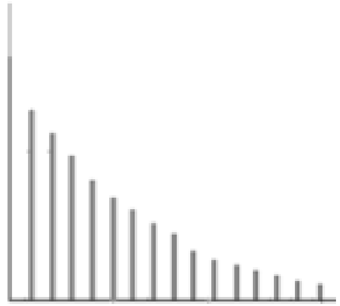
=

$$A\sum_{k=1}^{\infty}\frac{1}{k}\sin(2\pi kx)$$

square wave

infinite sum of sine waves

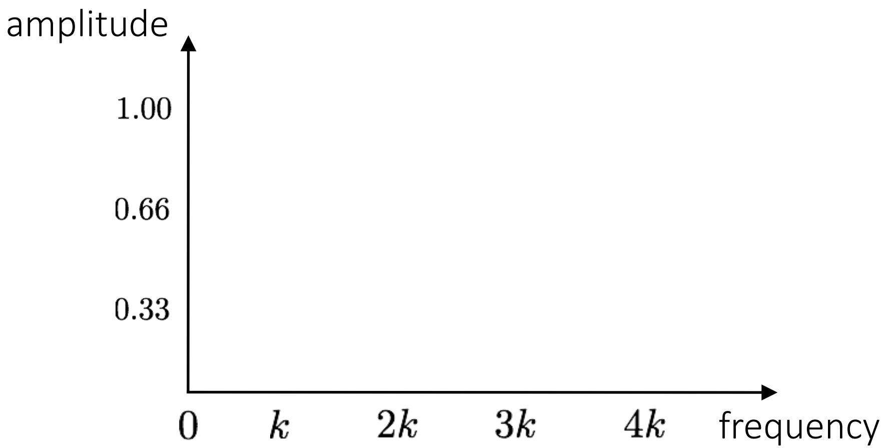
Magnitude
(amplitude)



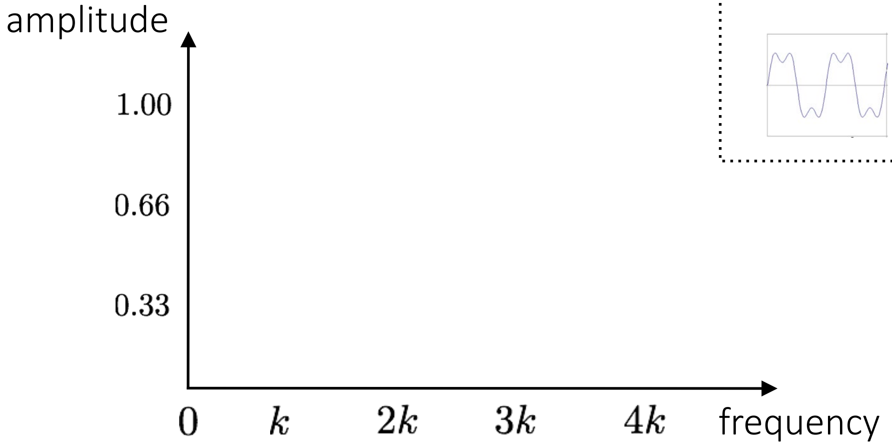
frequency

Frequency domain

Visualizing the frequency spectrum

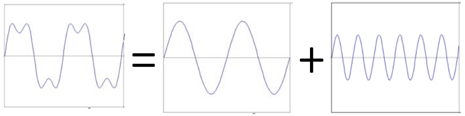


Visualizing the frequency spectrum



Recall the temporal domain visualization

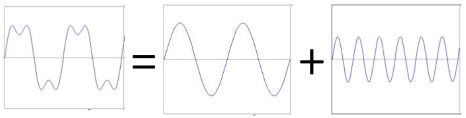
$$f(x) = \sin(2\pi kx) + \frac{1}{3}\sin(2\pi 3kx)$$

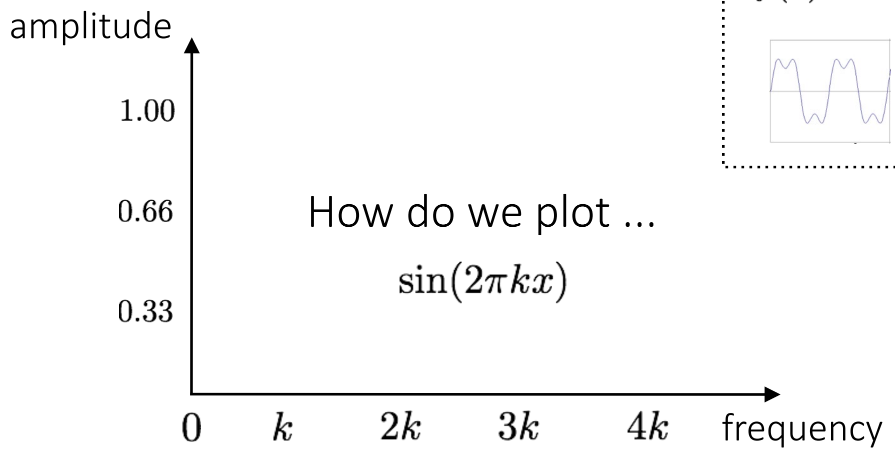


Visualizing the frequency spectrum

Recall the temporal domain visualization

$$f(x) = \sin(2\pi kx) + \frac{1}{3}\sin(2\pi 3kx)$$

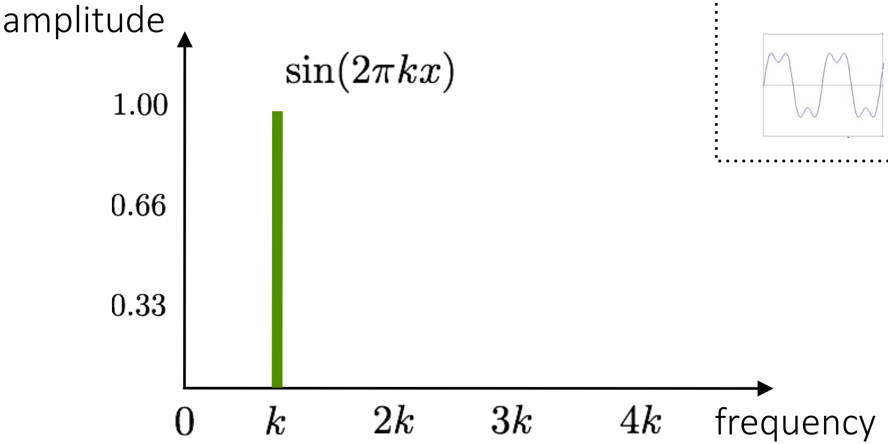




How do we plot ...

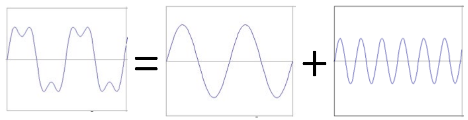
$$\sin(2\pi kx)$$

Visualizing the frequency spectrum

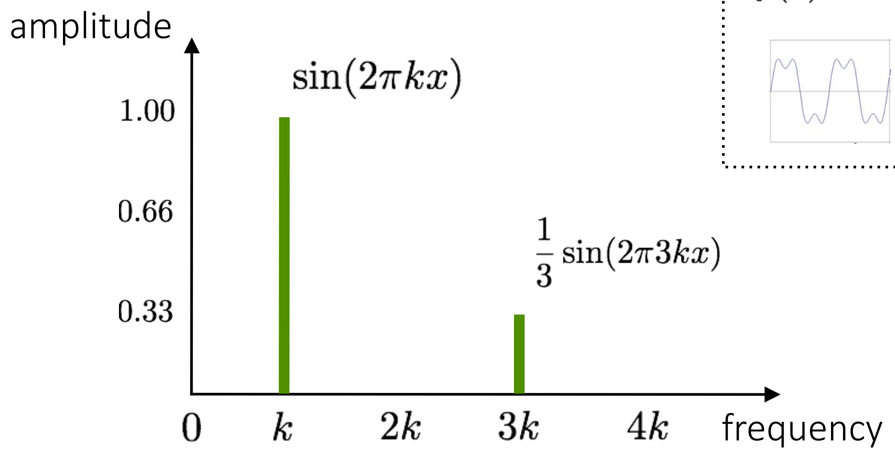


Recall the temporal domain visualization

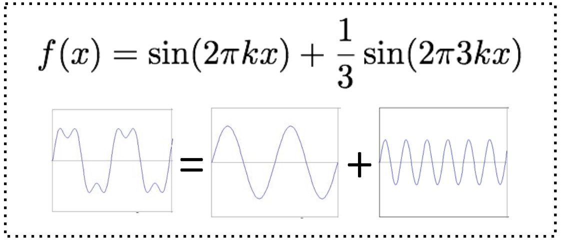
$$f(x) = \sin(2\pi kx) + \frac{1}{3}\sin(2\pi 3kx)$$



Visualizing the frequency spectrum



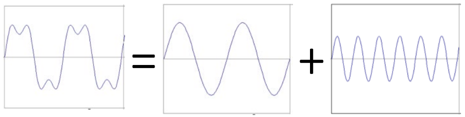
Recall the temporal domain visualization

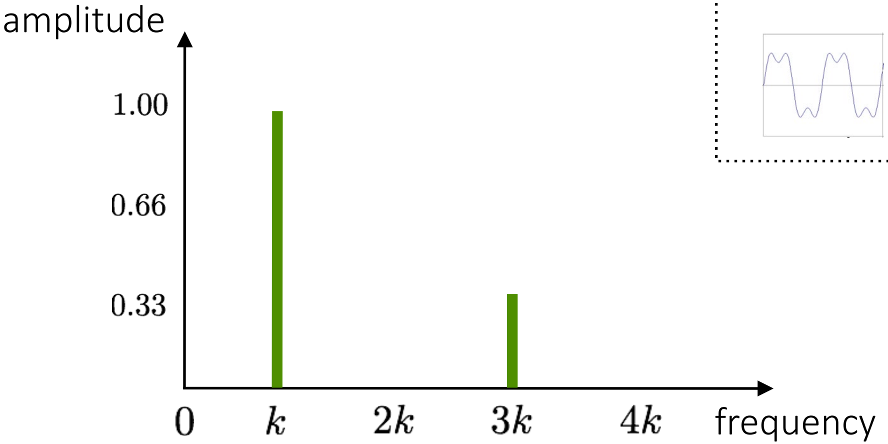


Visualizing the frequency spectrum

Recall the temporal domain visualization

$$f(x) = \sin(2\pi kx) + \frac{1}{3}\sin(2\pi 3kx)$$





not visualizing the symmetric negative part

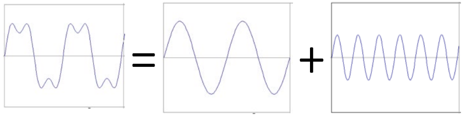
What is at zero frequency?

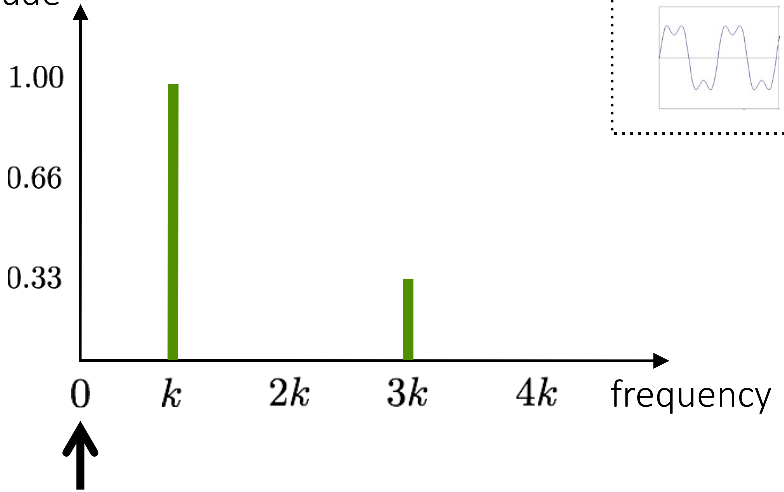
Need to understand this to understand the 2D version!

Visualizing the frequency spectrum

Recall the temporal domain visualization

$$f(x) = \sin(2\pi kx) + \frac{1}{3}\sin(2\pi 3kx)$$





not visualizing the symmetric negative part

signal average (zero for a sine wave with no offset)

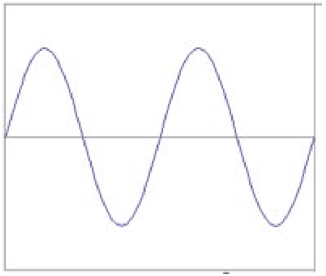
Need to understand this to understand the 2D version!

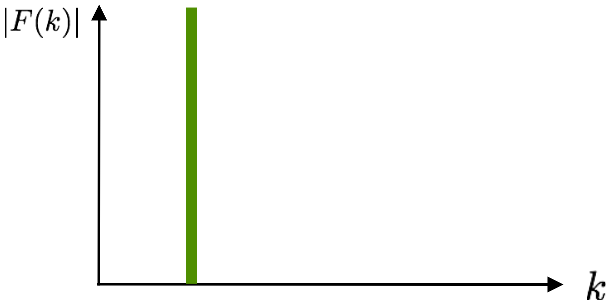
Examples

Spatial domain visualization

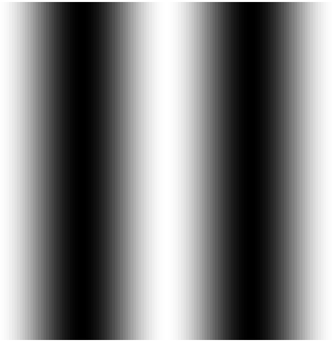
Frequency domain visualization

1D





2D

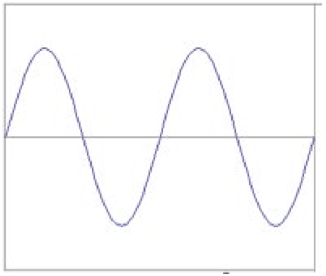


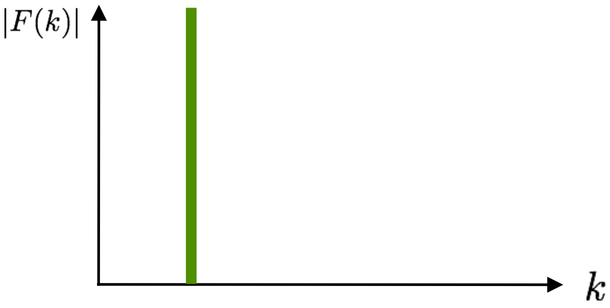
?

Examples

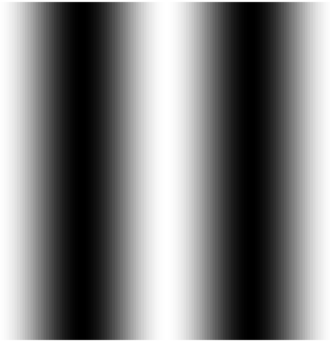
Spatial domain visualization | Frequency domain visualization

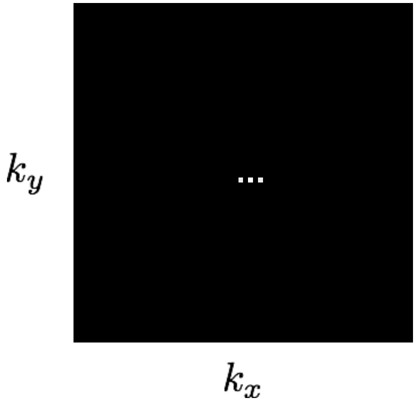
1D





2D



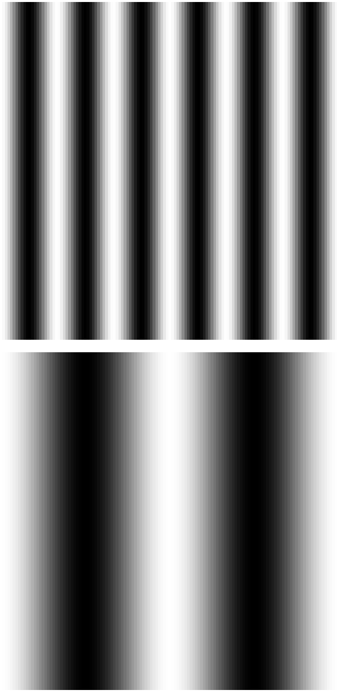


What do the three dots correspond to?

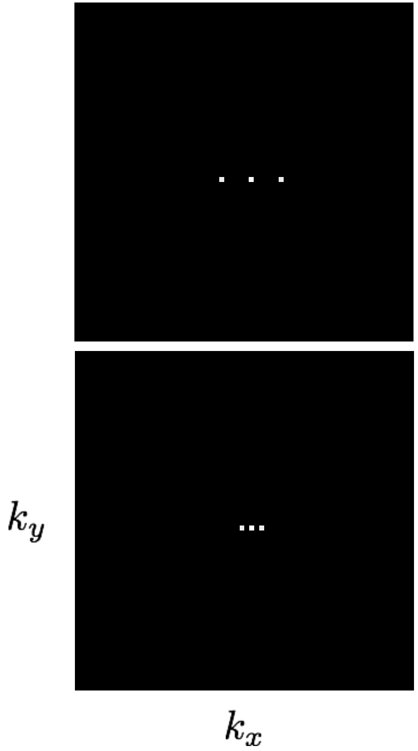
images that are pure cosines have particularly simple FT

Examples

Spatial domain visualization



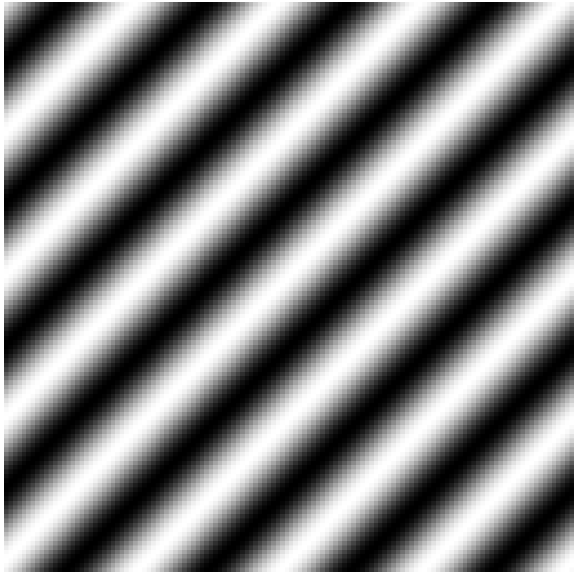
Frequency domain visualization

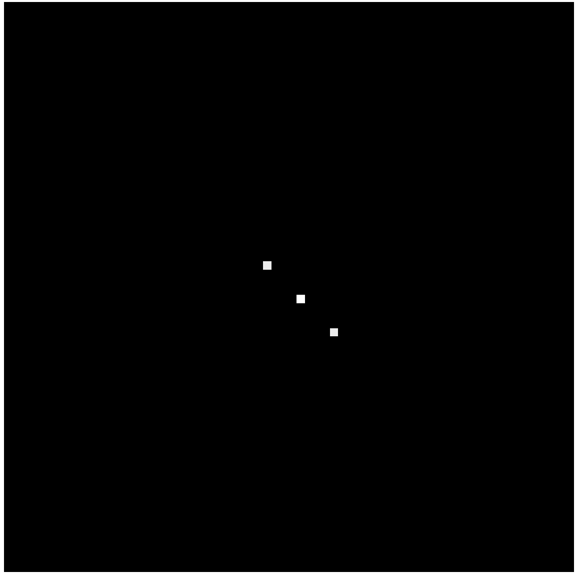


Scaling property of Fourier transform:
if we stretch a function by a factor in the time domain then squeeze the Fourier transform by the same factor in the frequency domain

Examples

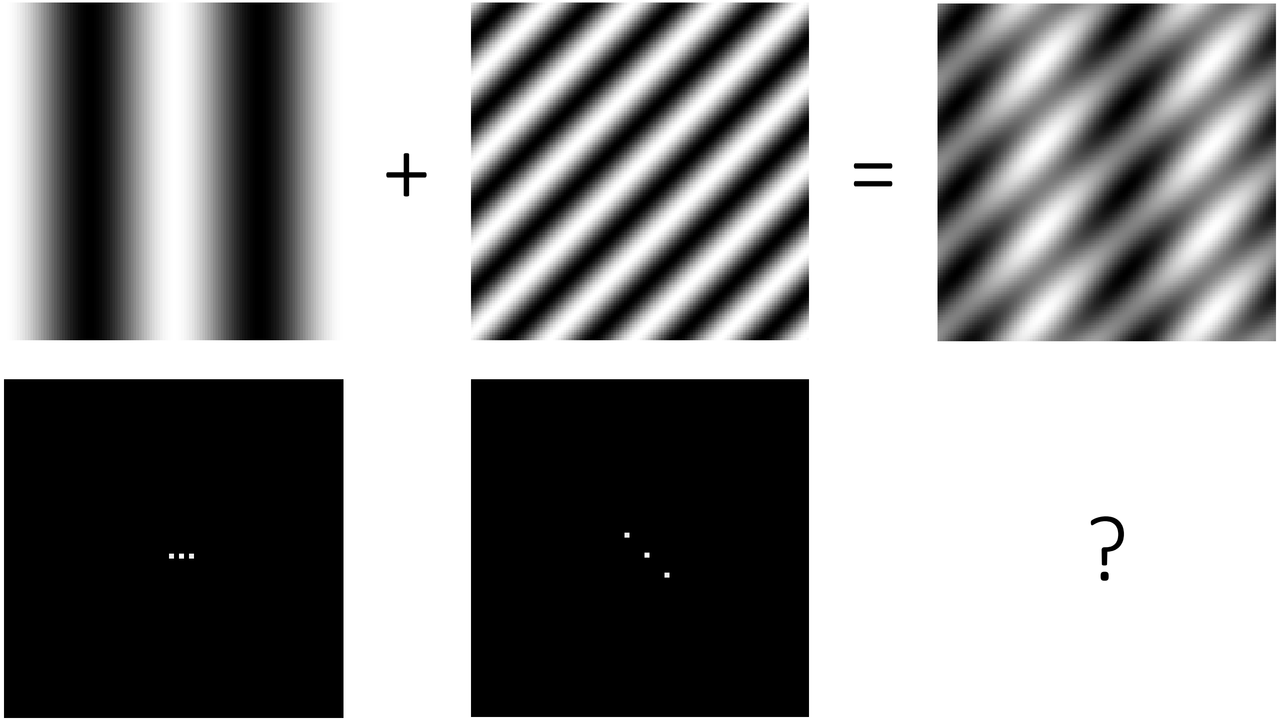
Rotation of the image results in equivalent rotation of its FT.





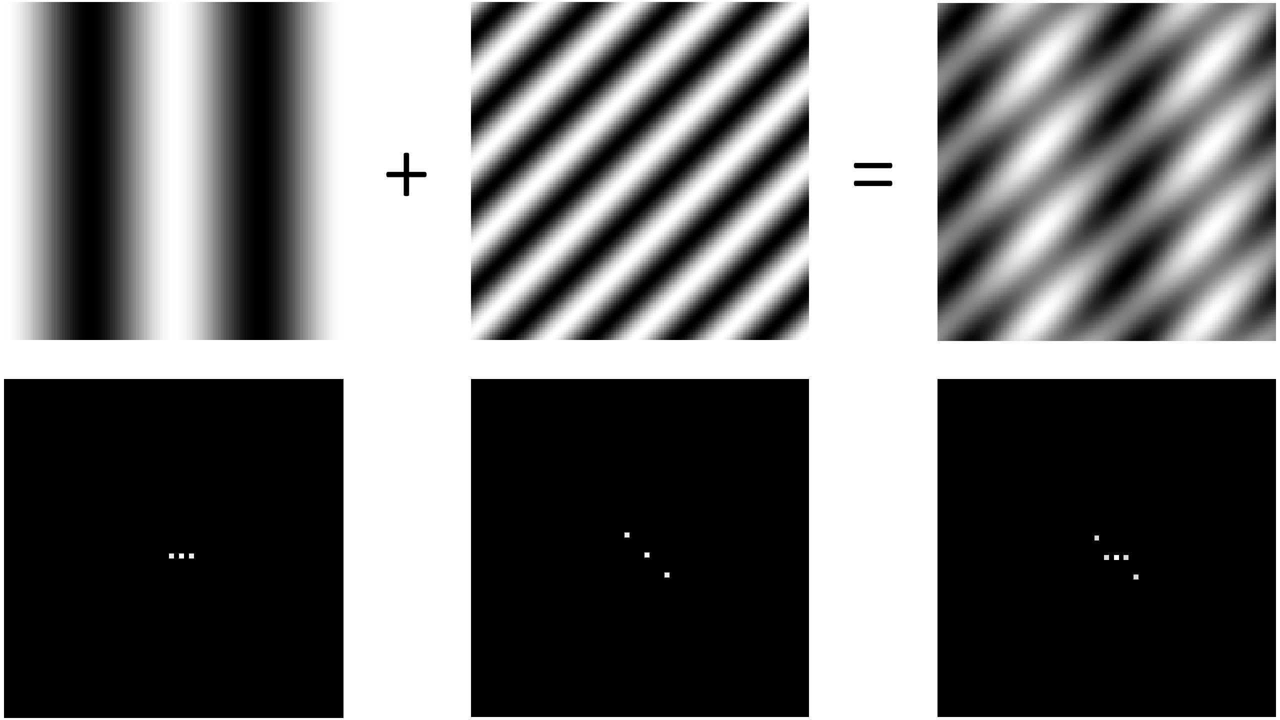
Has both an x and y components

Examples



Linearity Property

Examples



Basic building block

$$A \sin(\omega x + \phi)$$

amplitude

sinusoid

angular frequency

variable

phase

What about non-periodic signals?

Fourier's claim: Add enough of these to get <u>any</u> *periodic* signal you want!

Background

- Any function that periodically repeats itself can be expressed as the sum of sines and/or cosines of different frequencies, each multiplied by a different coefficient (Fourier series).

- Even functions that are not periodic (but whose area under the curve is finite) can be expressed as the integral of sines and/or cosines multiplied by a weighting function (Fourier transform).

Fourier transform

Fourier transform

	Fourier transform	inverse Fourier transform
continuous	$F(k) = \int_{\infty}^{-\infty} f(x) e^{-j2\pi kx} dx$	$f(x) = \int_{\infty}^{-\infty} F(k) e^{j2\pi kx} dk$
discrete	$F(k) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi kx/N}$ $k = 0, 1, 2, \ldots, N-1$	$f(x) = \sum_{k=0}^{N-1} F(k) e^{j2\pi kx/N}$ $x = 0, 1, 2, \ldots, N-1$

Where is the connection to the ‘summation of sine waves’ idea?

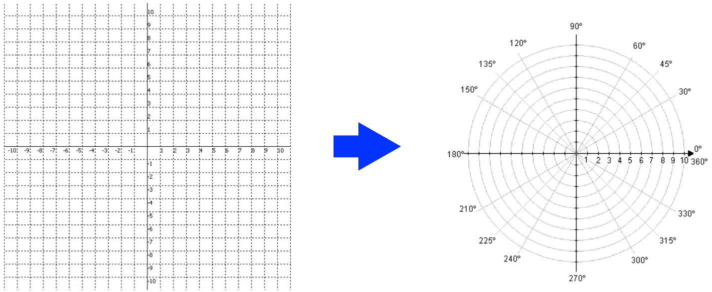
Recalling some basics

Complex numbers have two parts:

rectangular coordinates

$$R + jI$$

real imaginary



Recalling some basics

Complex numbers have two parts:

rectangular coordinates

$$R + jI$$

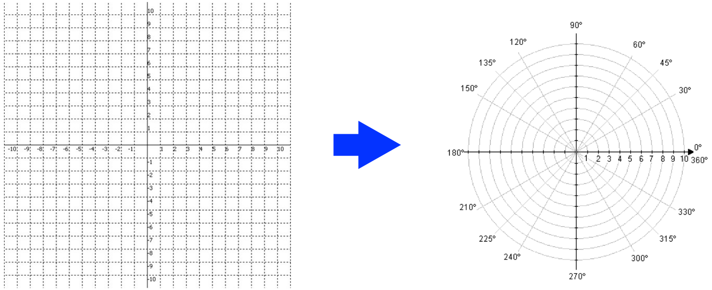
real imaginary

Alternative reparameterization:

polar coordinates

$$r(\cos\theta + j\sin\theta)$$

how do we compute these?



polar transform

Recalling some basics

Complex numbers have two parts:

rectangular coordinates

$$R + jI$$

real imaginary

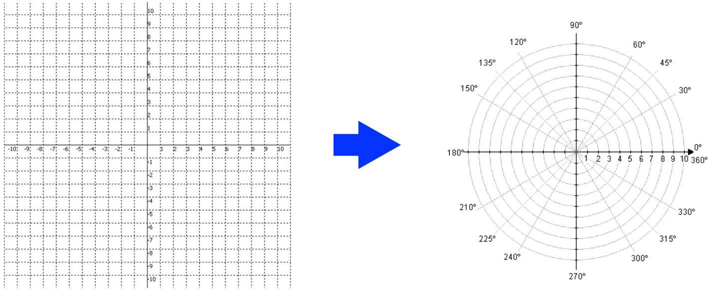
Alternative reparameterization:

polar coordinates

$$r(\cos\theta + j\sin\theta)$$

polar transform

$$\theta = \tan^{-1}\left(\frac{I}{R}\right) \quad r = \sqrt{R^2 + I^2}$$



polar transform

Recalling some basics

Complex numbers have two parts:

rectangular coordinates

$$R + jI$$

real imaginary

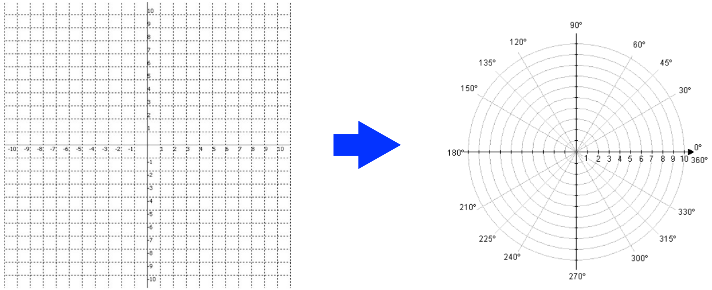
Alternative reparameterization:

polar coordinates

$$r(\cos\theta + j\sin\theta)$$

polar transform

$$\theta = \tan^{-1}(\frac{I}{R}) \quad r = \sqrt{R^2 + I^2}$$



polar transform

How do you write these in exponential form?

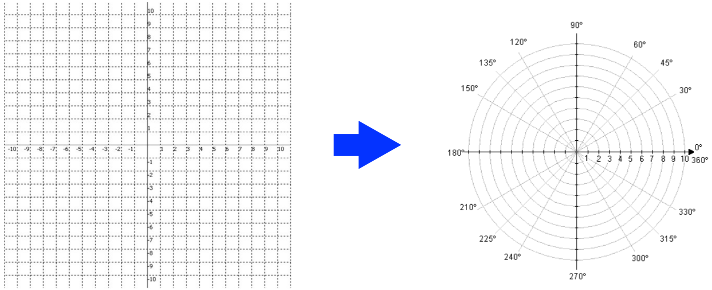
Recalling some basics

Complex numbers have two parts:

rectangular coordinates

$$R + jI$$

real imaginary



Alternative reparameterization:

polar coordinates

$$r(\cos\theta + j\sin\theta)$$

polar transform

$$\theta = \tan^{-1}(\frac{I}{R}) \quad r = \sqrt{R^2 + I^2}$$

or equivalently

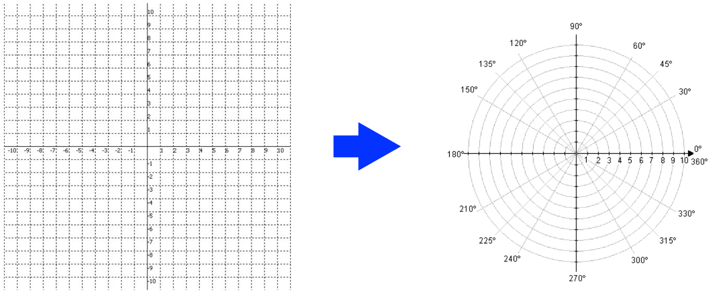
$$re^{j\theta}$$

how did we get this?

exponential form

Recalling some basics

Complex numbers have two parts:



rectangular coordinates

$$R + jI$$

real imaginary

Alternative reparameterization:

polar coordinates

$$r(\cos\theta + j\sin\theta)$$

polar transform

$$\theta = \tan^{-1}(\frac{I}{R}) \quad r = \sqrt{R^2 + I^2}$$

or equivalently

$$re^{j\theta}$$

Euler's formula

$$e^{j\theta} = \cos\theta + j\sin\theta$$

exponential form

This will help us understand the Fourier transform equations

Fourier transform

Fourier transform | inverse Fourier transform

continuous

$$F(k) = \int_{\infty}^{-\infty} f(x) e^{-j2\pi kx} dx$$

$$f(x) = \int_{\infty}^{-\infty} F(k) e^{j2\pi kx} dk$$

discrete

$$F(k) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi kx/N}$$

$k = 0, 1, 2, \ldots, N-1$

$$f(x) = \sum_{k=0}^{N-1} F(k) e^{j2\pi kx/N}$$

$x = 0, 1, 2, \ldots, N-1$

Where is the connection to the ‘summation of sine waves’ idea?

Fourier transform

	Fourier transform	inverse Fourier transform
continuous	$F(k) = \int_{\infty}^{-\infty} f(x) e^{-j2\pi kx} dx$	$f(x) = \int_{\infty}^{-\infty} F(k) e^{j2\pi kx} dk$
discrete	$F(k) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi kx/N}$ $k = 0, 1, 2, \ldots, N-1$	$\boxed{f(x) = \sum_{k=0}^{N-1} F(k) e^{j2\pi kx/N} \quad x = 0, 1, 2, \ldots, N-1}$

Where is the connection to the 'summation of sine waves' idea?

Fourier transform

Where is the connection to the 'summation of sine waves' idea?

$$f(x) = \sum_{k=0}^{N-1} F(k) e^{j2\pi kx/N}$$

Euler's formula

$$e^{j\theta} = \cos\theta + j\sin\theta$$

sum over frequencies

$$f(x) = \sum_{k=0}^{N-1} F(k) \left\{ \cos(2\pi kx) + j\sin(2\pi kx) \right\}$$

scaling parameter

wave components

2D Fourier Transform

DFT can be computed for any-dimensional input function. In particular, the 2D DFT is useful when working with images

2D Discrete Fourier transform:

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux/M+vy/N)}$$

Discrete Inverse Fourier transform:

$$f(x,y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) e^{-j2\pi(ux/M+vy/N)}$$

Spectrum and phase angle

Because the 2-D DFT is complex in general, it can be expressed in polar form:

$$F(u,v) = |F(u,v)|e^{j\,\theta(u,v)}$$

Where the magnitude

$$|F(u,v)| = \sqrt{R^2(u,v) + I^2(u,v)}$$

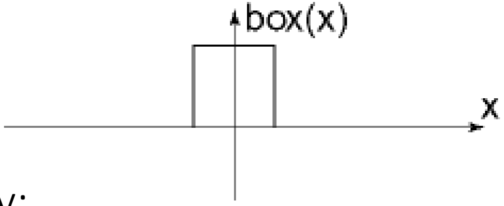
Is called the **Fourier spectrum** and

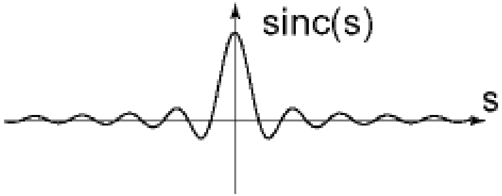
$$\theta(u,v) = \arctan\left(\frac{I(u,v)}{R(u,v)}\right)$$

Is called the **phase angle**

Fourier transform pairs

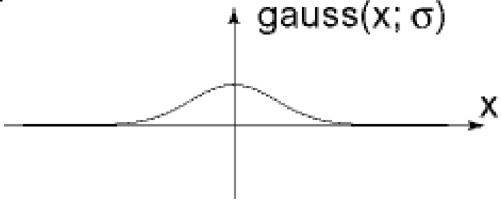
spatial domain | frequency domain

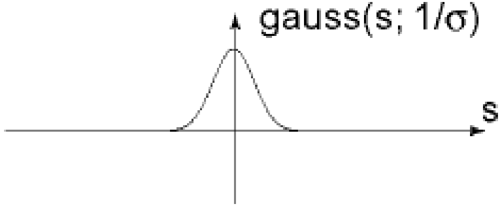


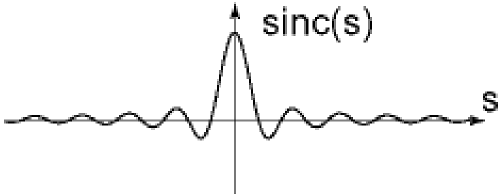


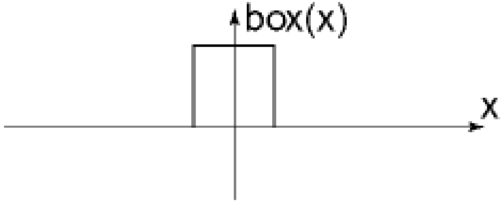
$$\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$$

Note the symmetry: duality property of Fourier transform









Computing the discrete Fourier transform (DFT)

Computing the discrete Fourier transform (DFT)

$$F(k) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi kx/N}$$ is just a matrix multiplication:

$$\boldsymbol{F} = \boldsymbol{W}\boldsymbol{f}$$

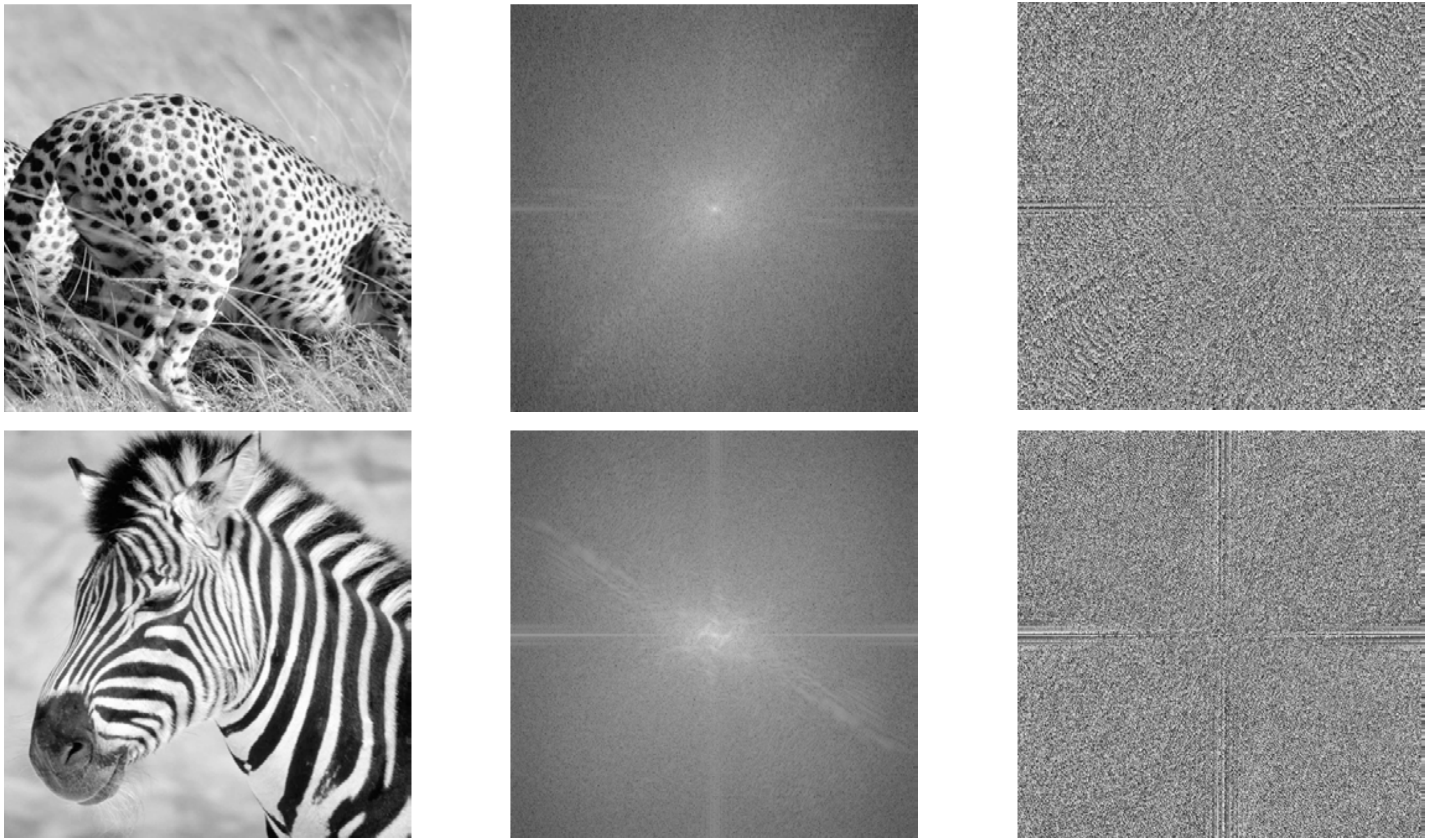
- Basically a matrix-vector product:

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & W_N^3 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & W_{N6} & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \times \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{bmatrix}$$

$$(W_N = e^{-j2\pi/N})$$

In practice this is implemented using the *fast Fourier transform* (FFT) algorithm.

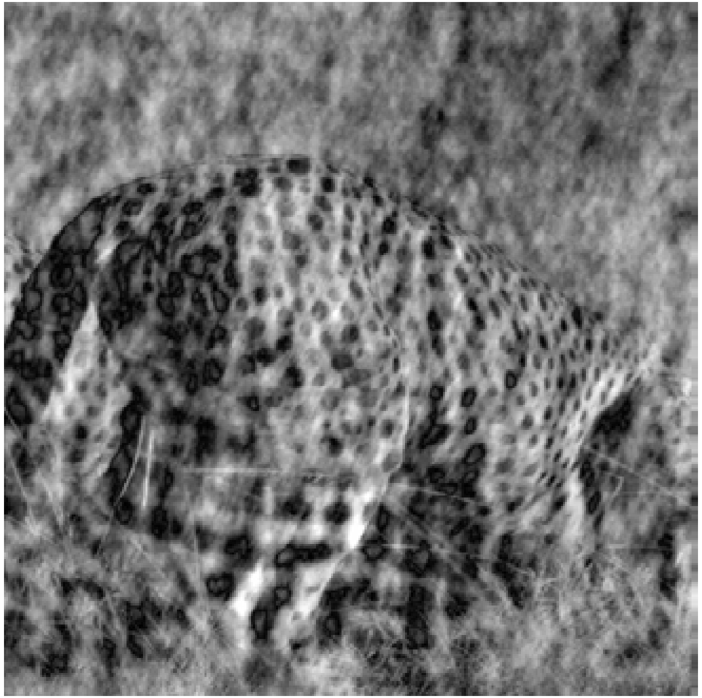
Fourier transforms of natural images



original amplitude phase

Fourier transforms of natural images

Image phase matters!



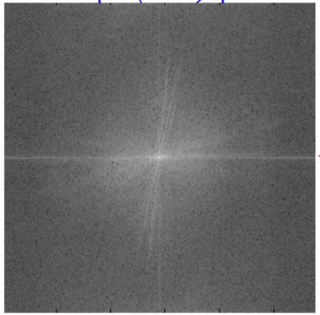
cheetah phase with zebra amplitude

zebra phase with cheetah amplitude

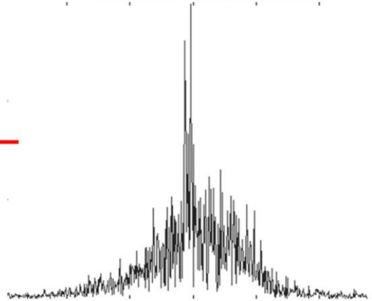
Discrete Fourier Transform - Visualization

$f(x,y)$

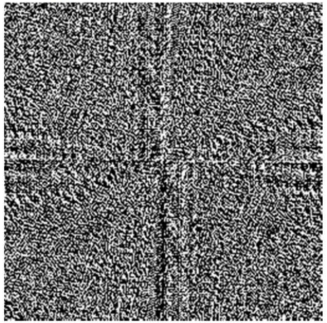
$|F(u,v)|$



cross-section



$phase\ F(u,v)$



- $|f(u,v)|$ generally decreases with higher spatial frequencies
- phase appears less informative

Slide by A. Zisserman

Frequency-domain filtering

Why do we care about all this?

The convolution theorem

The Fourier transform of the convolution of two functions is the product of their Fourier transforms:

$$\mathcal{F}\{g * h\} = \mathcal{F}\{g\}\mathcal{F}\{h\}$$

The inverse Fourier transform of the product of two Fourier transforms is the convolution of the two inverse Fourier transforms:

$$\mathcal{F}^{-1}\{gh\} = \mathcal{F}^{-1}\{g\} * \mathcal{F}^{-1}\{h\}$$

Convolution in spatial domain is equivalent to multiplication in frequency domain!

What do we use convolution for?

Convolution for 1D continuous signals

Definition of linear shift-invariant filtering as convolution:

$$(f * g)(x) = \int_{-\infty}^{\infty} f(y) g(x - y) dy$$

filtered signal

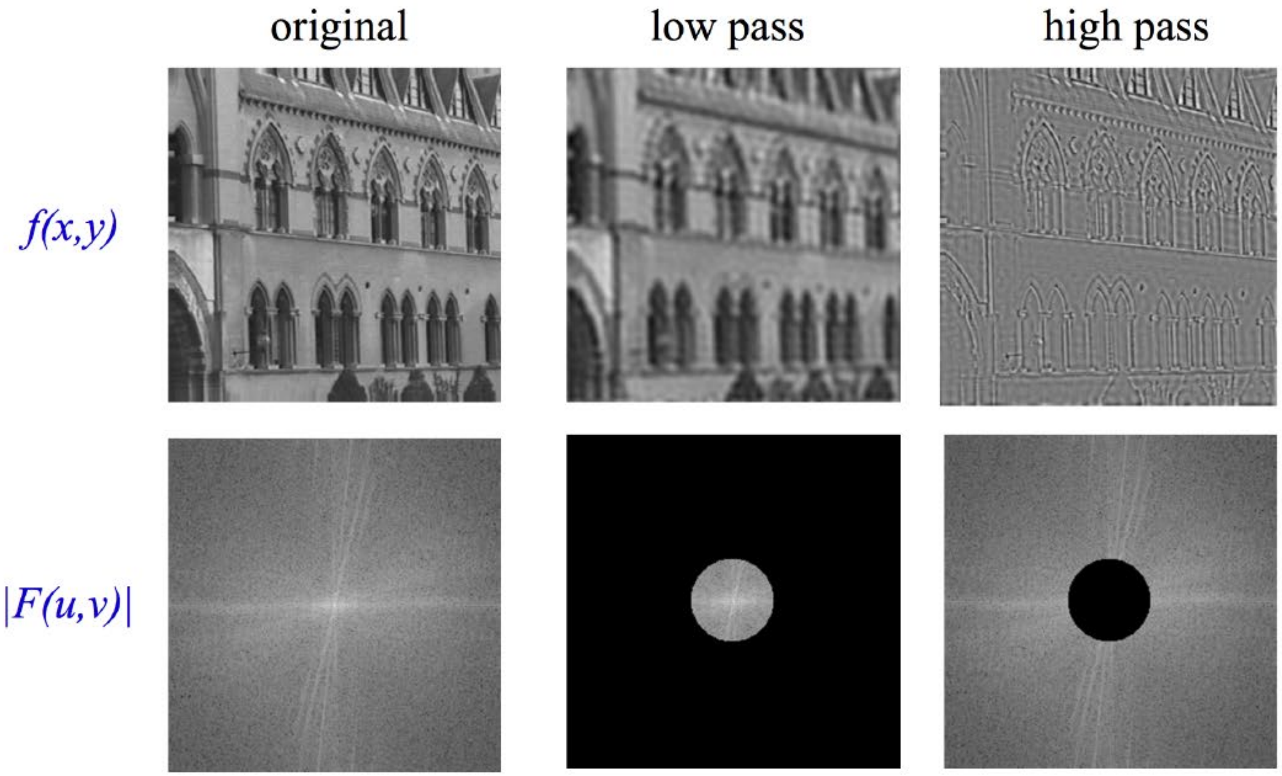
filter

input signal

Why implement convolution in frequency domain?

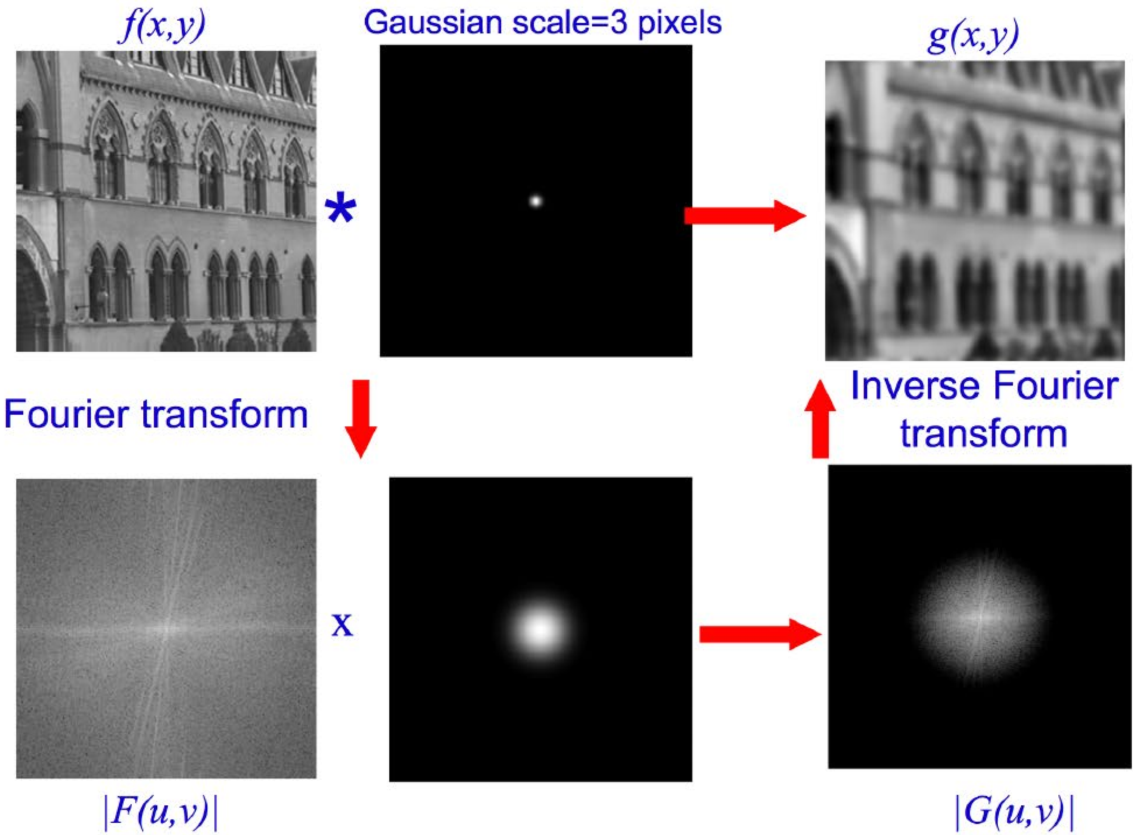
Using the convolution theorem, we can interpret and implement all types of linear shift-invariant filtering as multiplication in frequency domain.

Image Filtering in the Frequency Domain



Slide by A. Zisserman

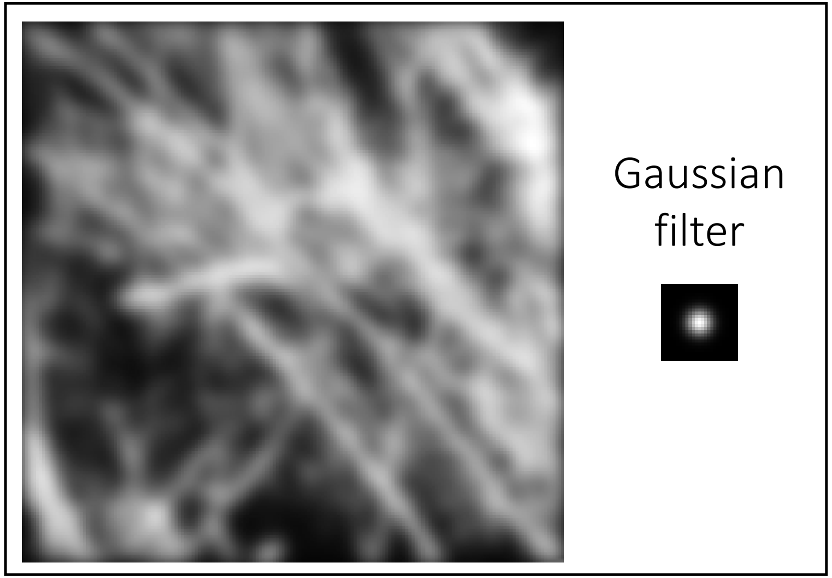
Blurring in the Time vs Frequency Domain



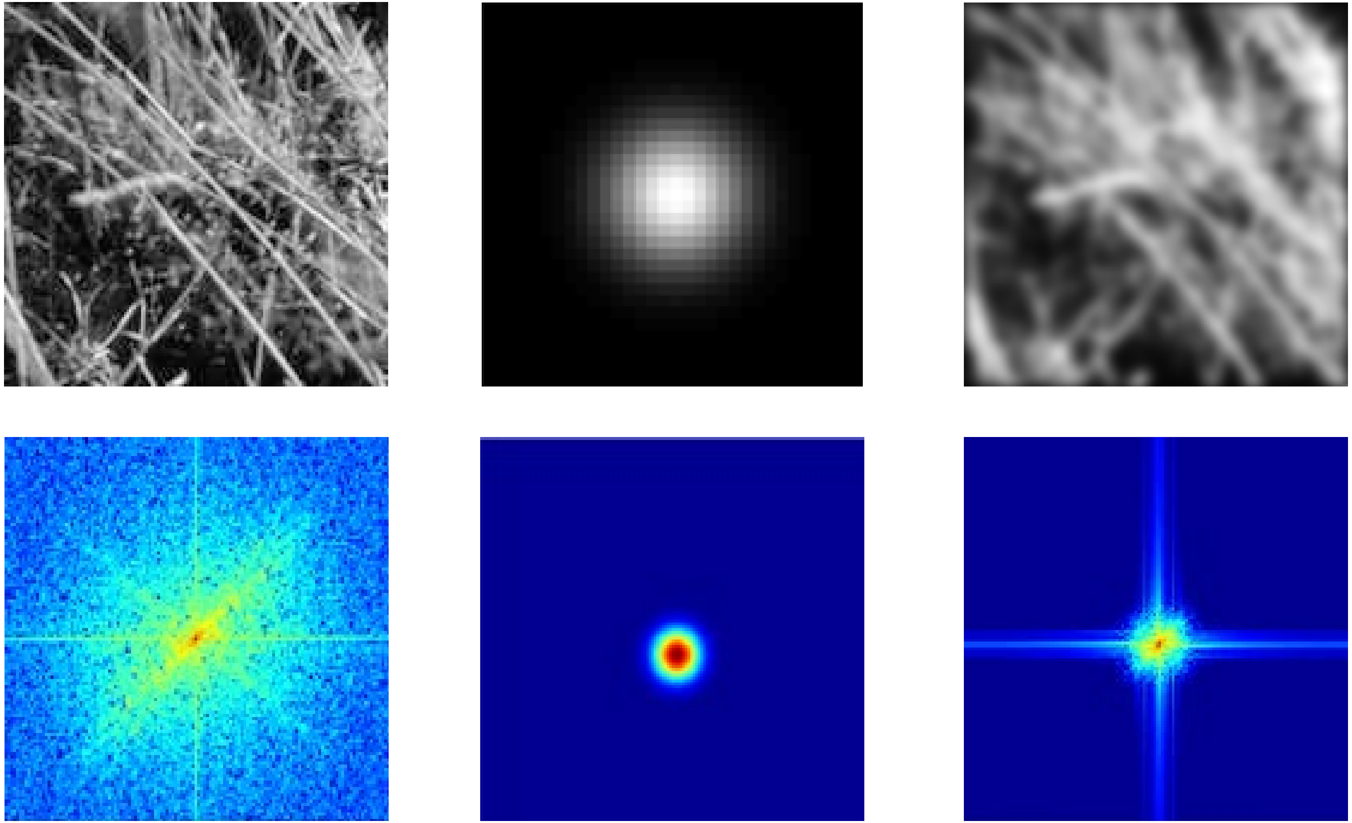
Example by A. Zisserman

Revisiting blurring

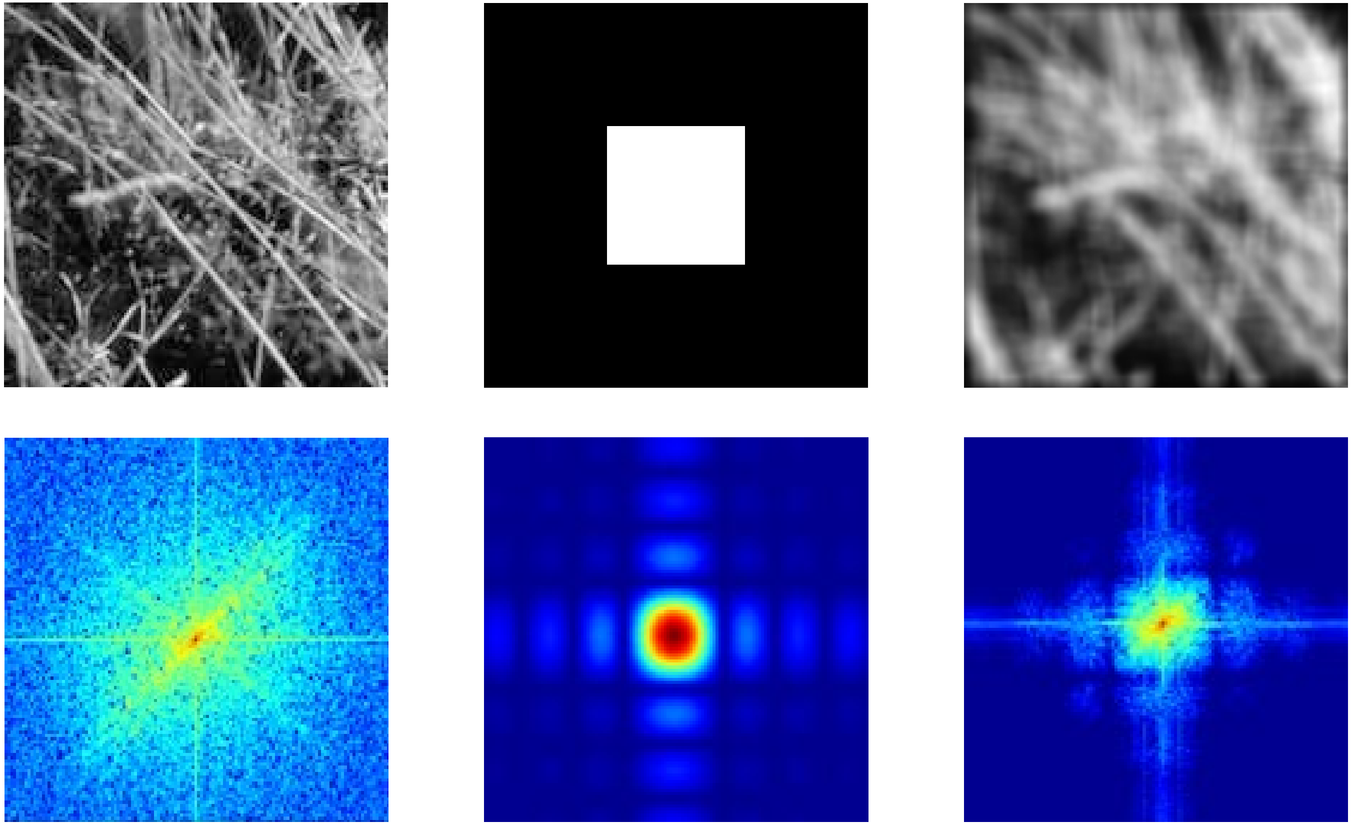
Why does the Gaussian give a nice smooth image, but the square filter give edgy artifacts?



Gaussian blur



Box blur

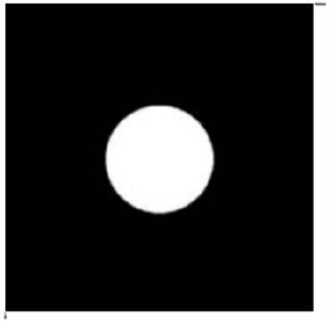


Ideal Low-pass filter

An ideal low-pass filter ILPF is defined by:

$$H(u,v) = \begin{cases} 1 & \text{if } D(u,v) \le D_0 \\ 0 & \text{if } D(u,v) > D_0 \end{cases}$$

The point of transition between H(u, v) = 1 and H(u, v)= 0 is called the **cutoff frequency**



Ideal low pass filter

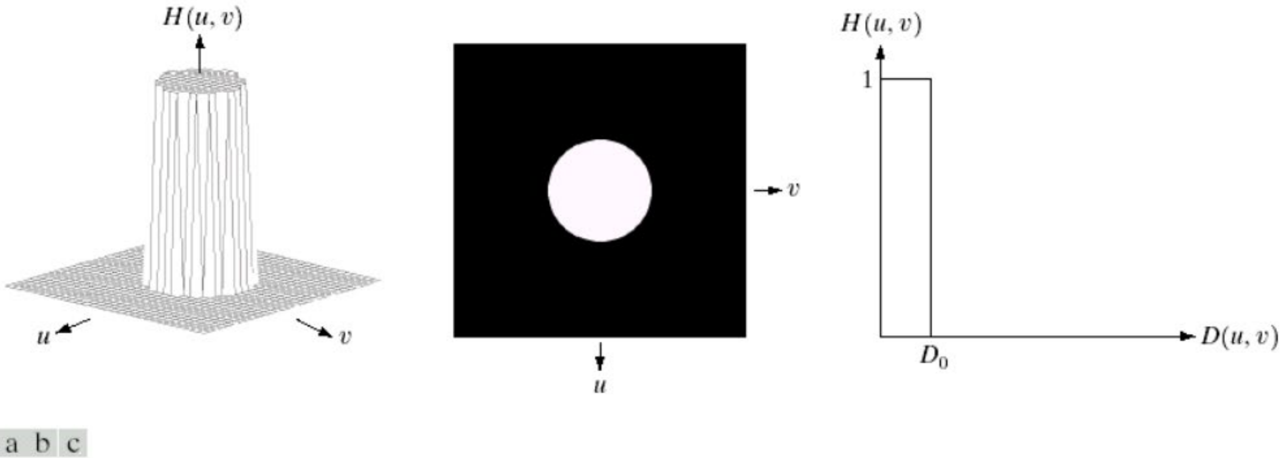
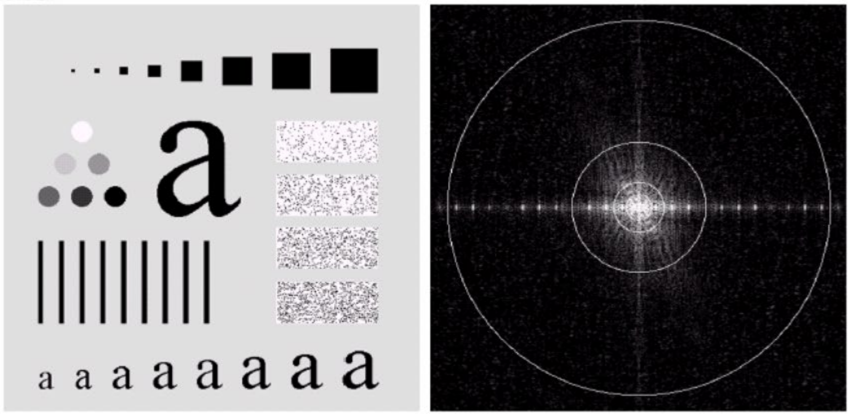


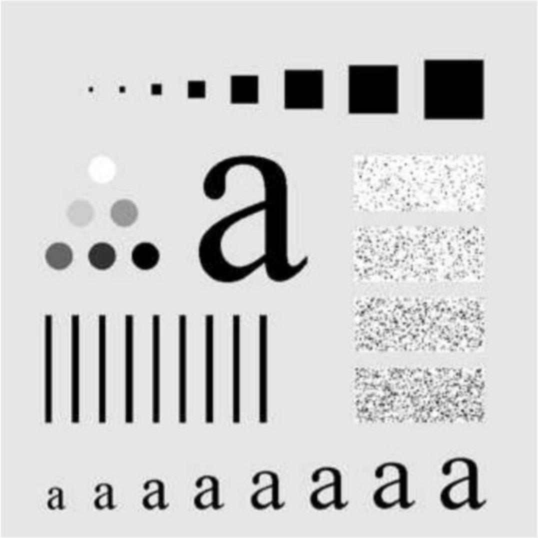
FIGURE 4.10 (a) Perspective plot of an ideal lowpass filter transfer function. (b) Filter displayed as an image. (c) Filter radial cross section.

Idel low pass filter

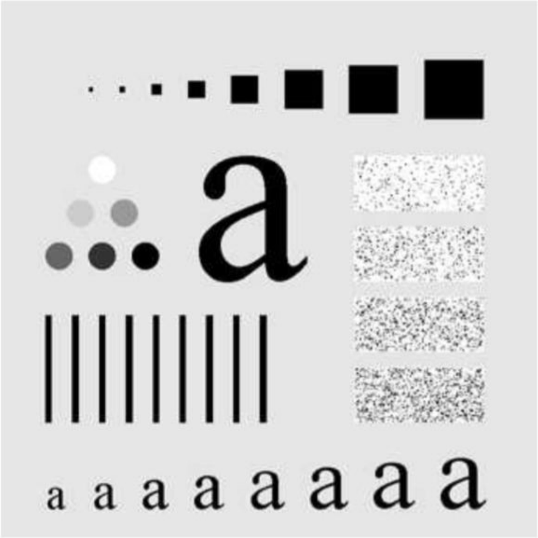


a b

FIGURE 4.11 (a) An image of size 500 × 500 pixels and (b) its Fourier spectrum. The superimposed circles have radii values of 5, 15, 30, 80, and 230, which enclose 92.0, 94.6, 96.4, 98.0, and 99.5% of the image power, respectively.



ILPF with cutoff frequency =60



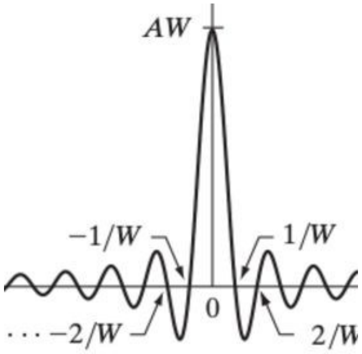
ILPF with cutoff
frequency =30

ILPF

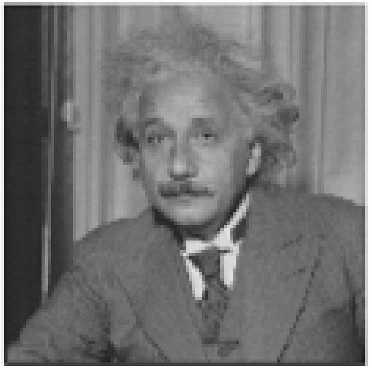
The blurring and ringing properties of ILPFs can be explained using the convolution theorem:

Because a cross section of the ILPF in the frequency domain looks like a box filter, a cross section of the corresponding spatial filter has the shape of a sinc.

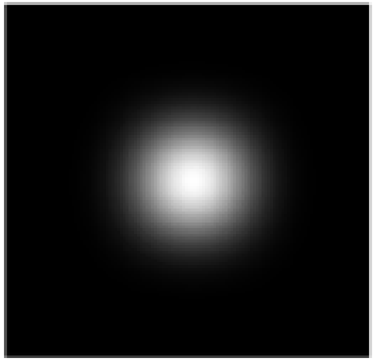
Convolving a sinc with an impulse copies the sinc at the location of the impulse. The sinc center lobe causes the blurring, while the outer lobes are responsible for ringing.

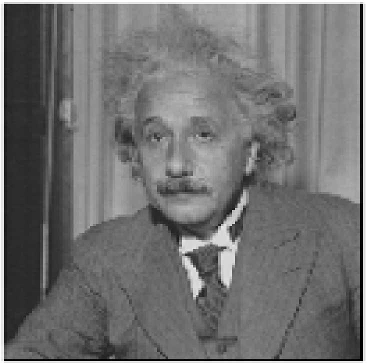


More filtering examples

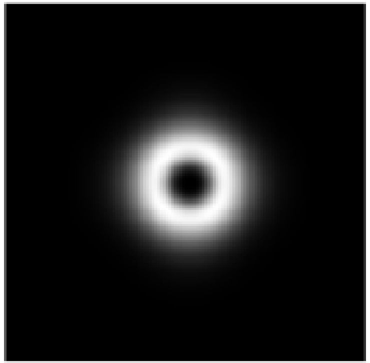


?



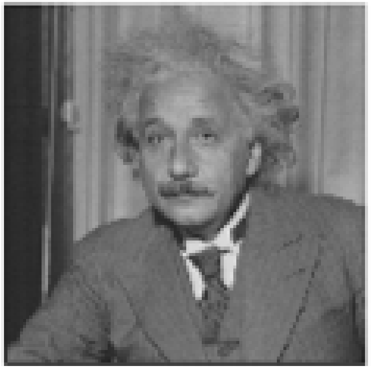


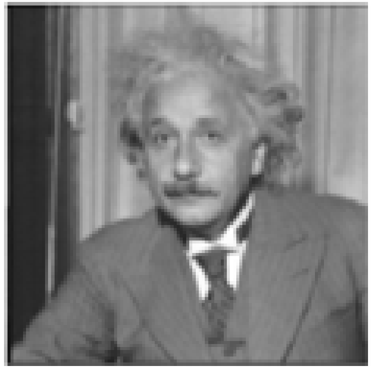
?



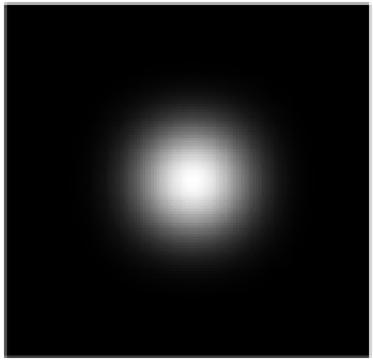
filters shown in frequency-domain

More filtering examples

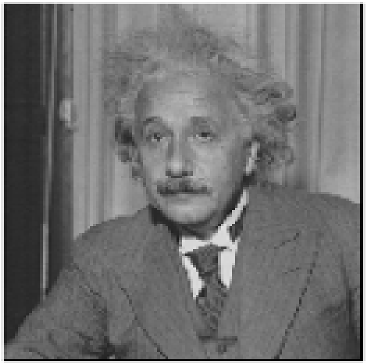


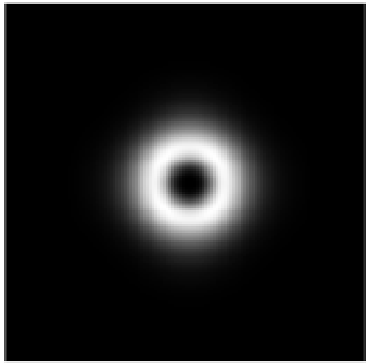


low-pass



band-pass



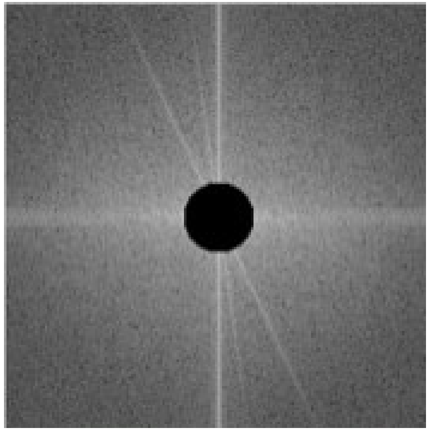


filters shown in frequency-domain

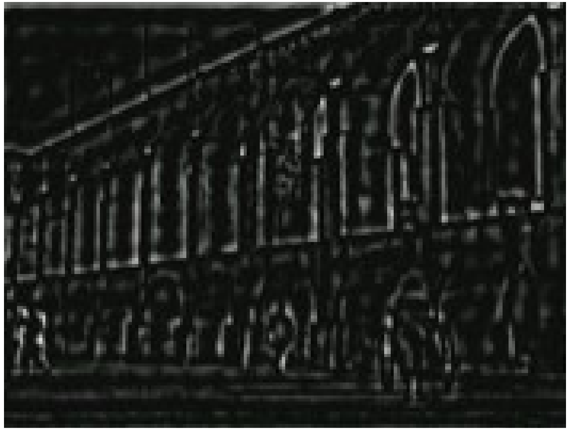
More filtering examples

?

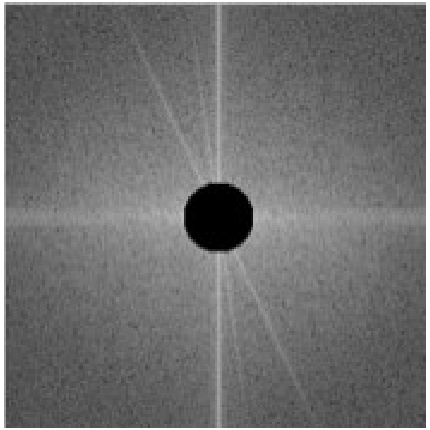
high-pass



More filtering examples

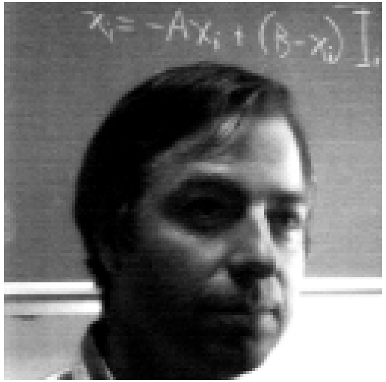


high-pass

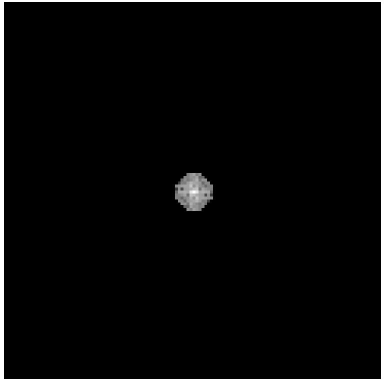


More filtering examples

original image

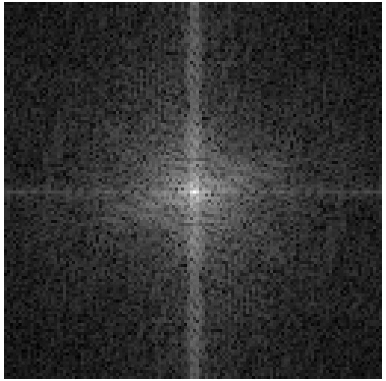


low-pass filter



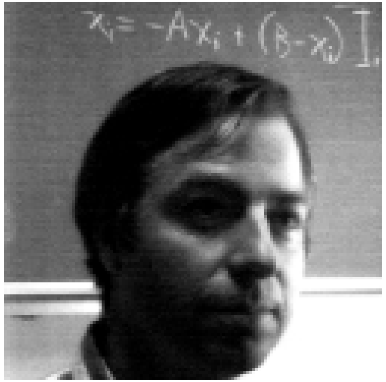
?

frequency magnitude

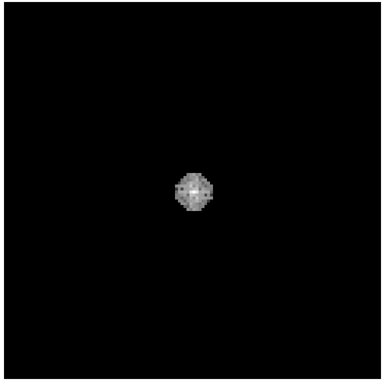


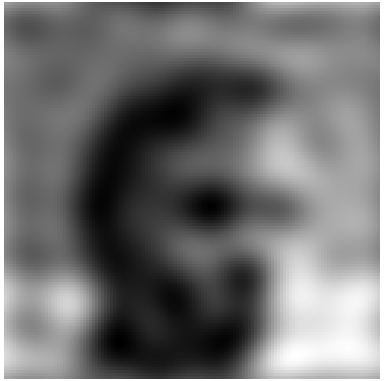
More filtering examples

original image

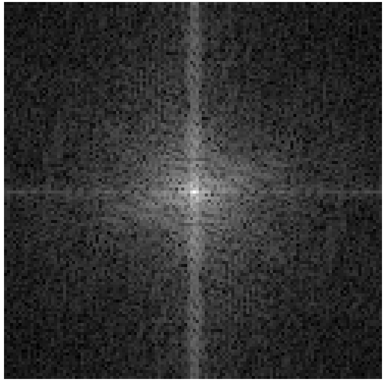


low-pass filter



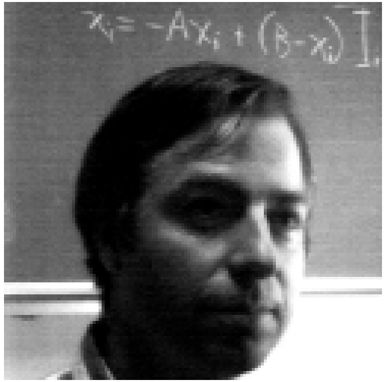


frequency magnitude

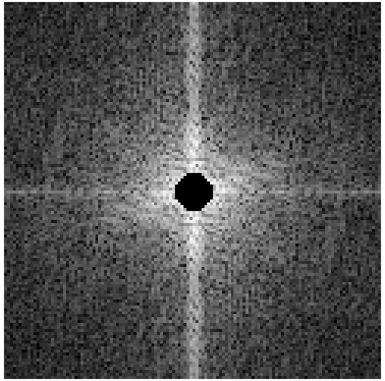


More filtering examples

original image

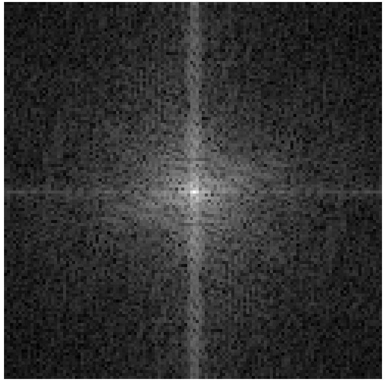


high-pass filter



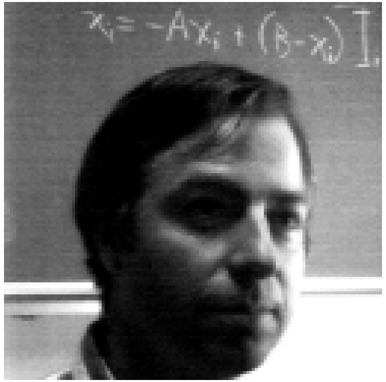
?

frequency magnitude

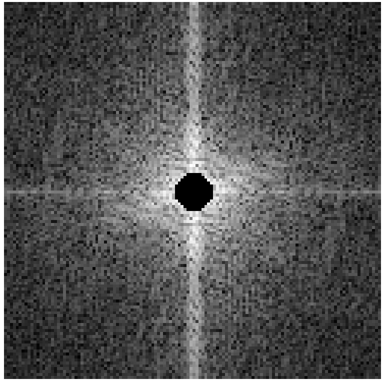


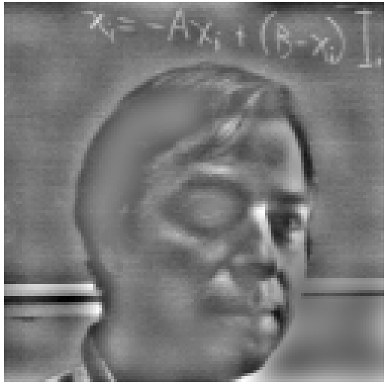
More filtering examples

original image

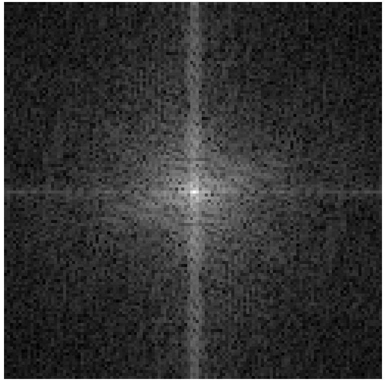


high-pass filter



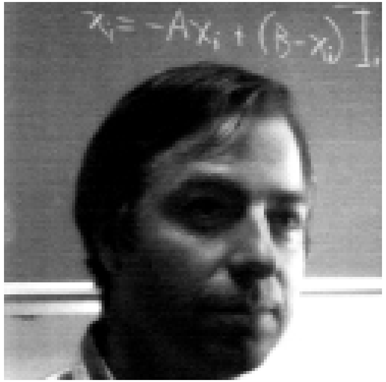


frequency magnitude

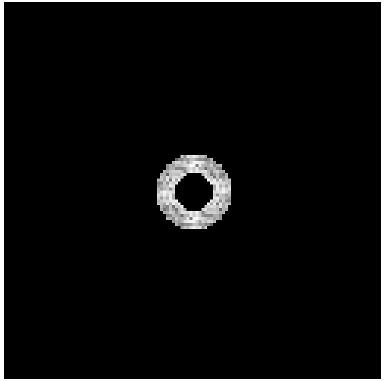


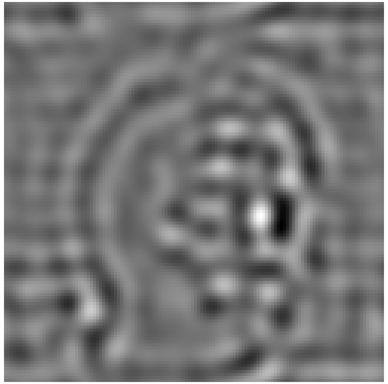
More filtering examples

original image

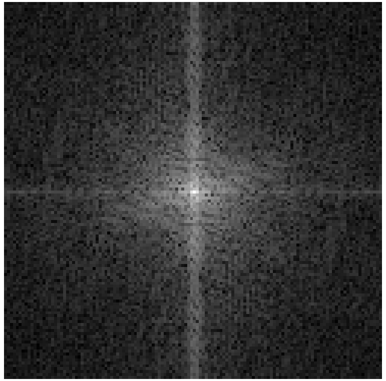


band-pass filter



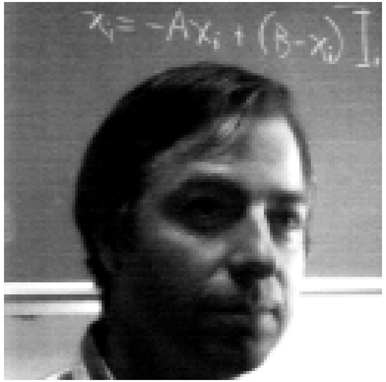


frequency magnitude

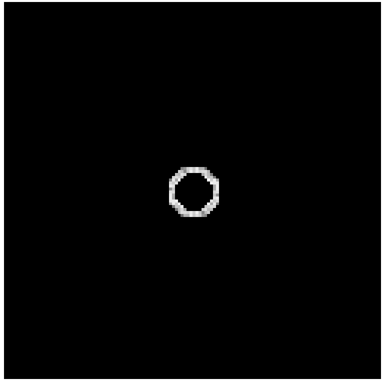


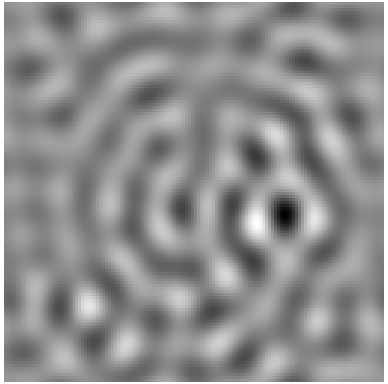
More filtering examples

original image

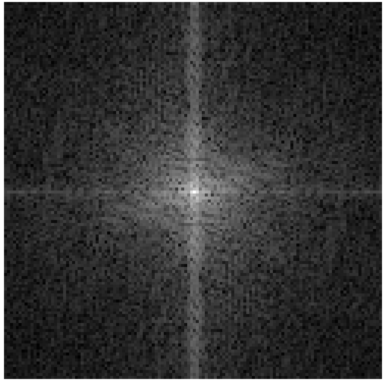


band-pass filter



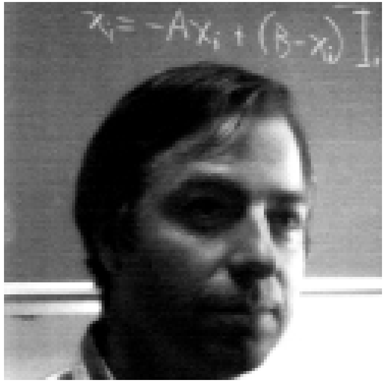


frequency magnitude

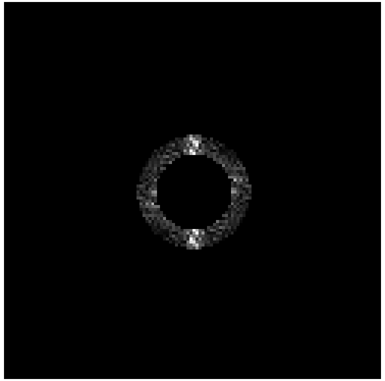


More filtering examples

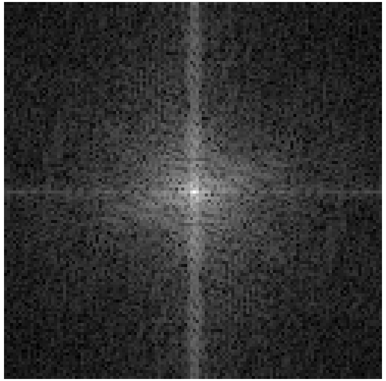
original image



band-pass filter

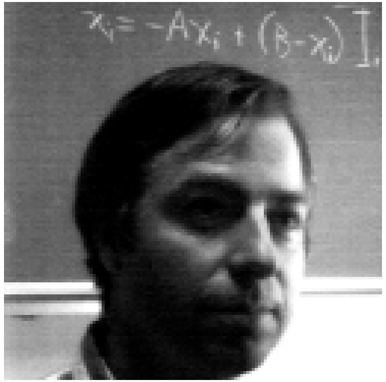


frequency magnitude

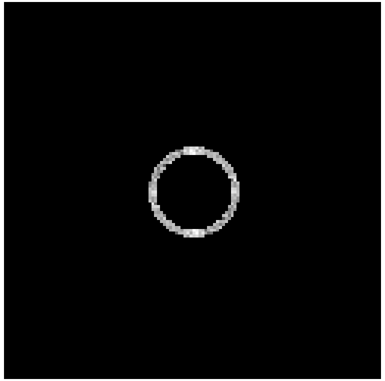


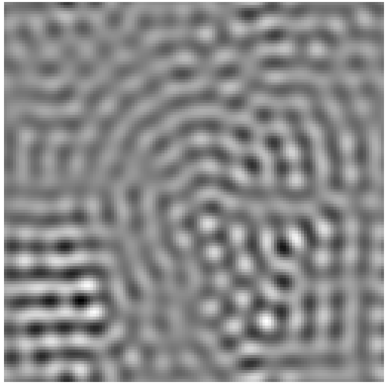
More filtering examples

original image

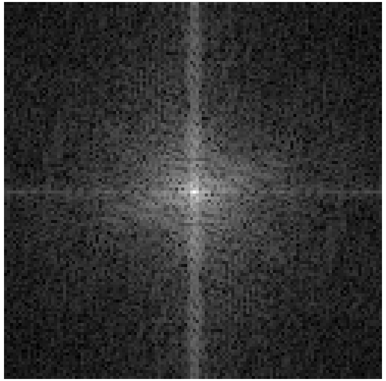


band-pass filter

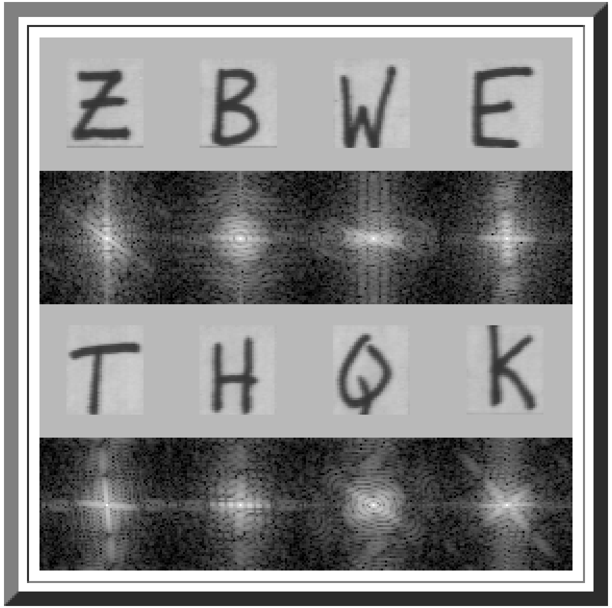




frequency magnitude



Recognizing character



Revisiting sampling

The Nyquist-Shannon sampling theorem

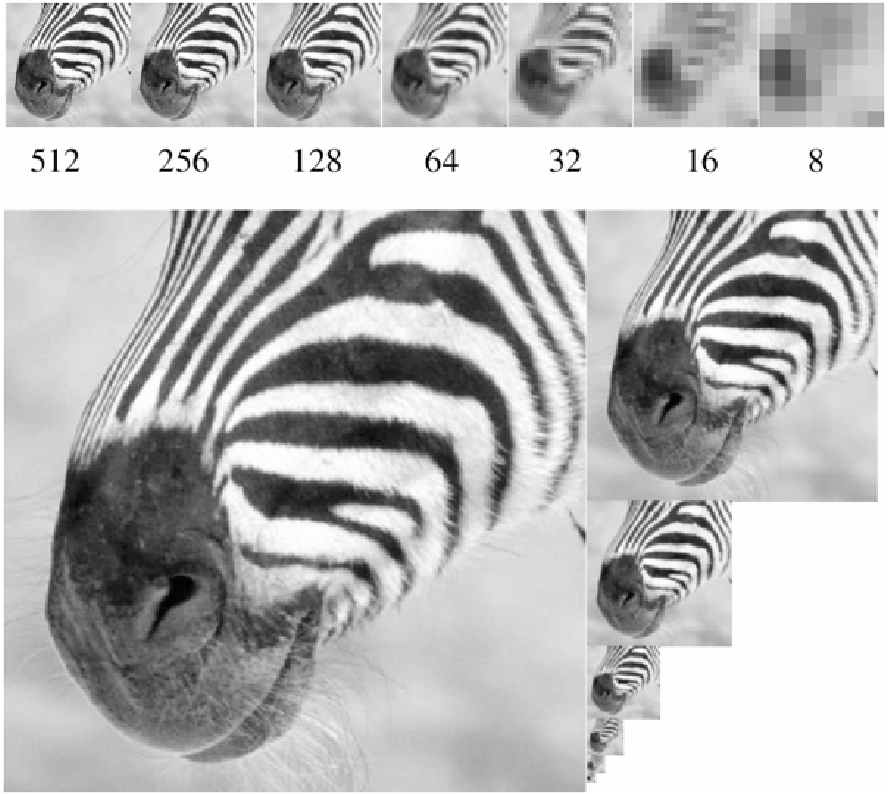
A continuous signal can be perfectly reconstructed from its discrete version using linear interpolation, if sampling occurred with frequency:

$$f_s \geq 2 f_{\max}$$

This is called the Nyquist frequency

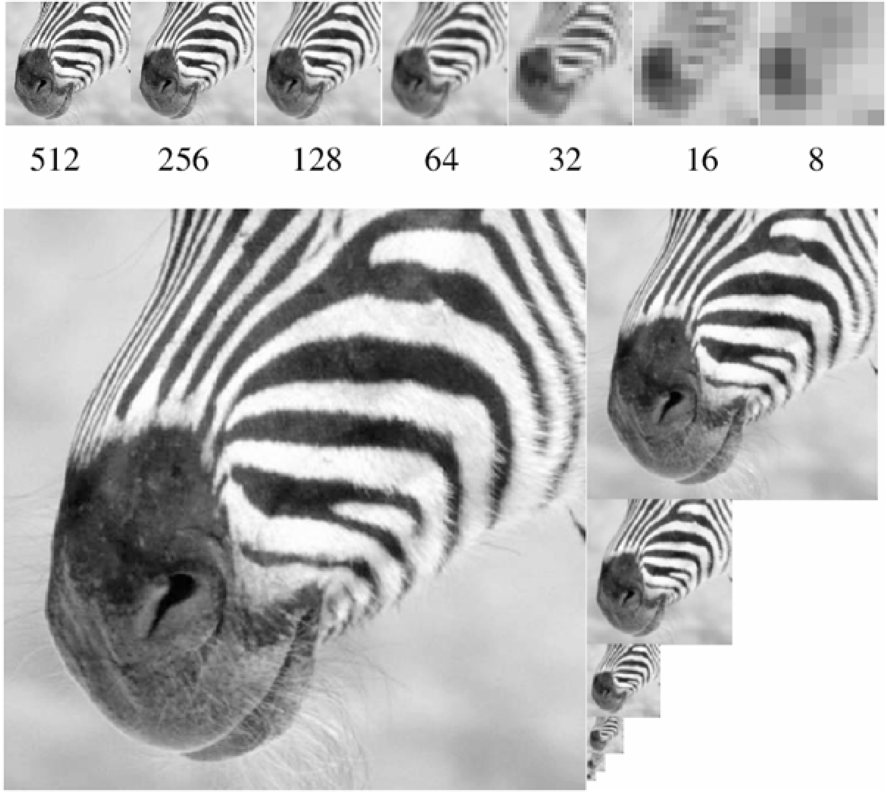
Equivalent reformulation: When downsampling, aliasing does not occur if samples are taken at the Nyquist frequency or higher.

Gaussian pyramid



How does the Nyquist-Shannon theorem relate to the Gaussian pyramid?

Gaussian pyramid



How does the Nyquist-Shannon theorem relate to the Gaussian pyramid?

- Gaussian blurring is low-pass filtering.
- By blurring we try to sufficiently decrease the Nyquist frequency to avoid aliasing.

How large should the Gaussian blur we use be?

- The cut-off frequency of the Gaussian filter is proportional to the standard deviation of the filter in the frequency domain
- The range is equal to double the standard deviation. In general we use a mask three times the standard deviation

Frequency-domain filtering in human vision

"Hybrid image"

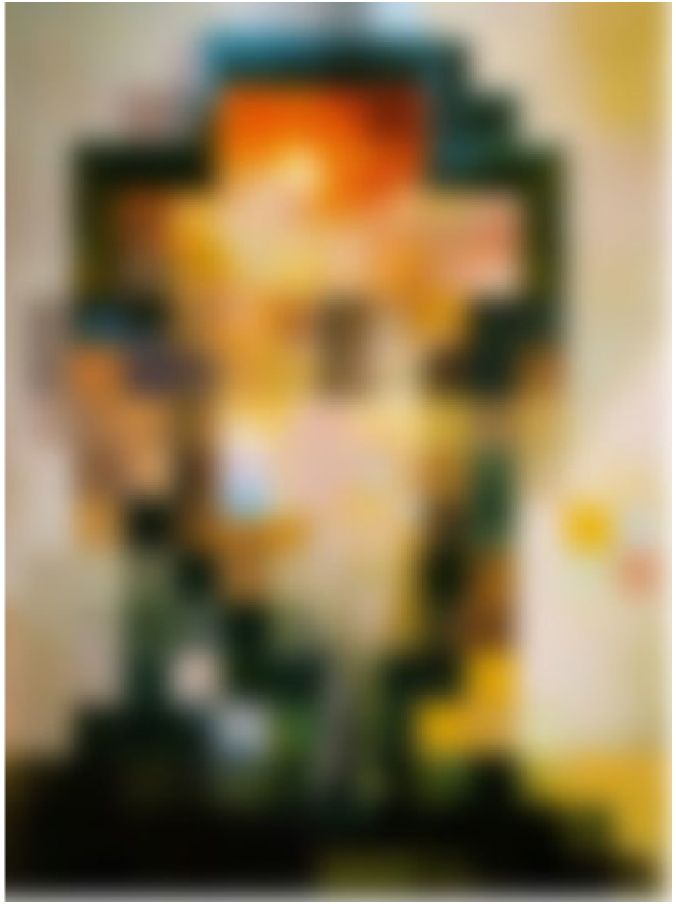
Aude Oliva and Philippe Schyns

Frequency-domain filtering in human vision

Gala Contemplating the Mediterranean Sea Which at Twenty Meters Becomes the Portrait of Abraham Lincoln (Homage to Rothko)

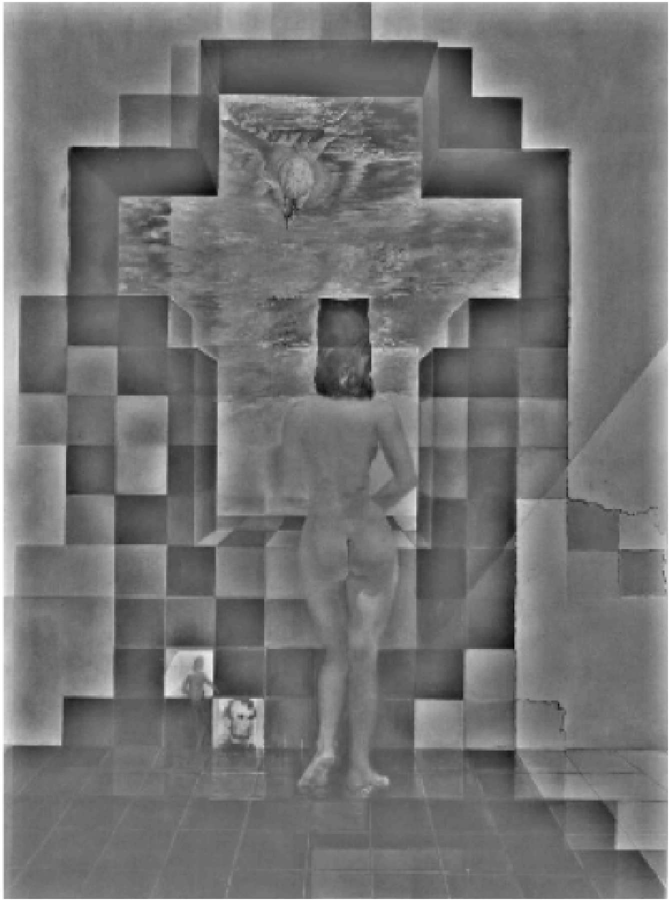
Salvador Dali, 1976

Frequency-domain filtering in human vision



Low-pass filtered version

Frequency-domain filtering in human vision



High-pass filtered version

Acknowledgement: some slides and material from Bernt Schiele, Mario Fritz, Michael Black, Bill Freeman, Fei-Fei, Justin Johnson, Serena Yeung, R. Szelisky, Fabio Galasso, Ioannis Gkioulekas