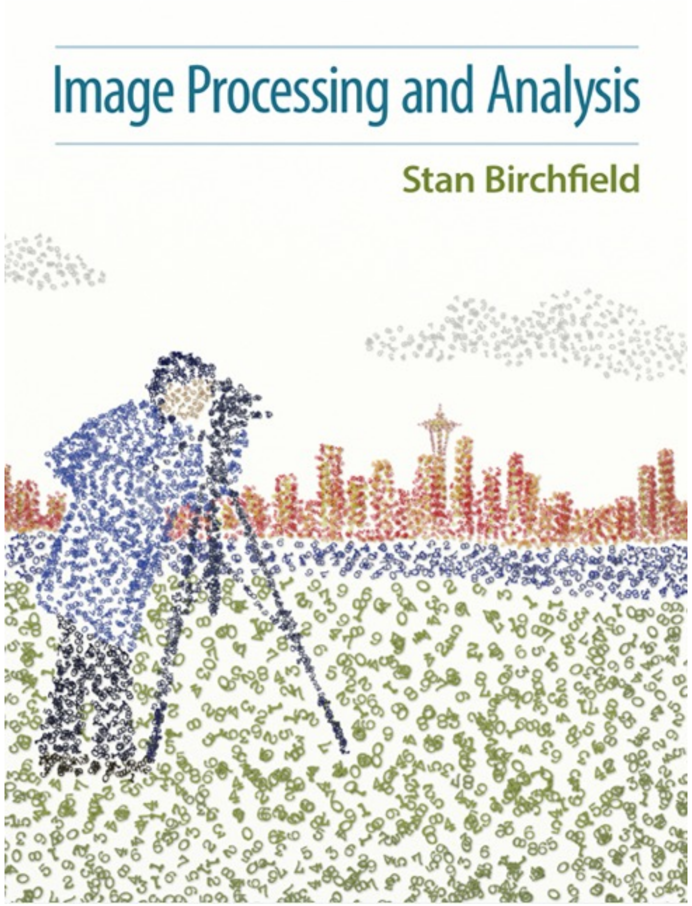
Prof. Kjersti Engan

ELE510 Image processing and computer vision

Frequency Domain processing, 2D DFT (Chap 6.3 Birchfield) 2023

University of Stavanger



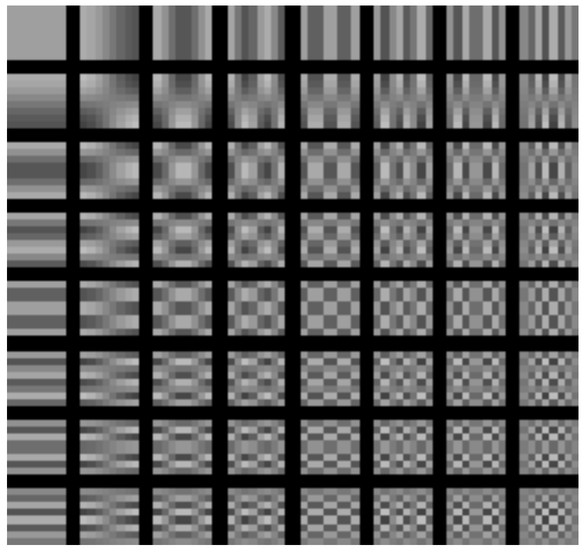
Parts in course presentations is material from Cengage learning. It can be used for teaching, and it can be share it with students on access controlled web-sites (Canvas) for use in THIS course. **But it should not be copied or distributed further in any way** (by you or anybody).

2D DFT

Three points from the topic:

1. How can we extend DFT to 2D images?
2. What is the projection – slice theorem?
3. What is Discrete Cosine Transform and how does it relate to DFT?

https://www.youtube.com/watch?v=Qm84XIoTy0s



(6.3) Discrete Fourier Transform – 2D

The 2D DFT is a natural extension of the 1D case:

DFT of image $g(k,l)$

$$\mathcal{F}(g(m,n)) = \hat{g}(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)e^{-j2\pi\left(\frac{km}{M}+\frac{ln}{N}\right)}$$

IDFT (inverse DFT)

$$g(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{g}(m,n)e^{j2\pi\left(\frac{km}{M}+\frac{ln}{N}\right)}$$

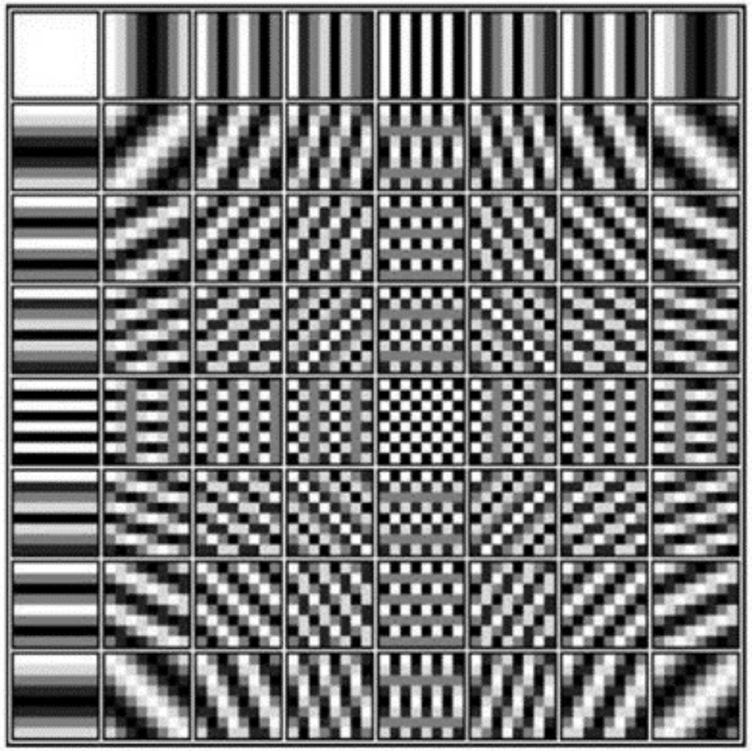
(Remember eigenimages, SVD)
Basis images for 8x8 image, DFT

$$\hat{g} = \sum_{i=1}^{N} \sum_{j=1}^{M} g_{ij} \mathbf{u}_i \mathbf{u}_j^T$$

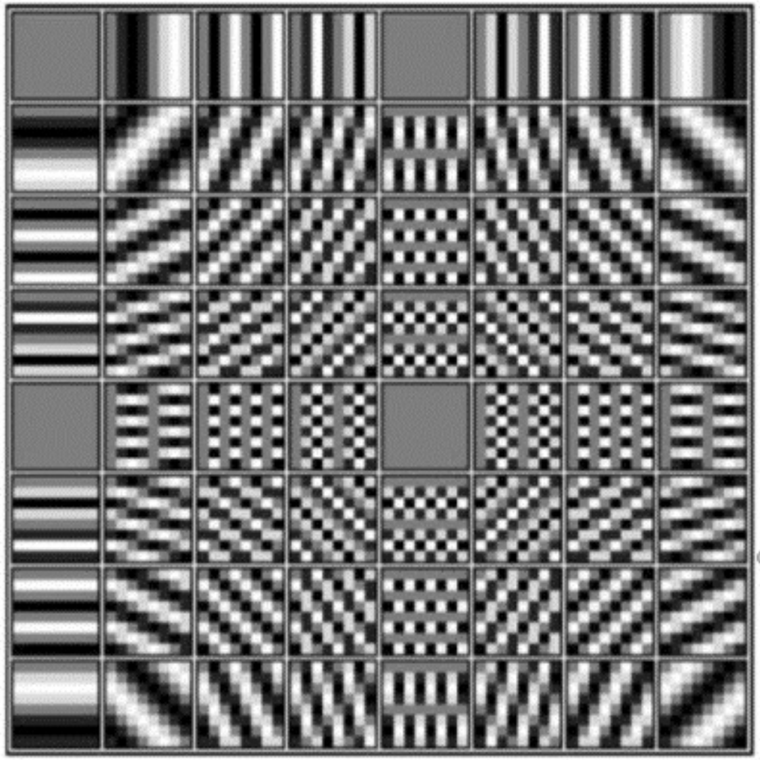
Image patch

coefficients

Basis images



Real part of basis images



Imaginary part of basis images

Discrete Fourier Transform

$$\hat{g}(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)e^{-j2\pi\left(\frac{km}{M}+\frac{ln}{N}\right)}$$

$$\hat{g}(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)e^{-j2\pi\frac{km}{M}}\, e^{-j2\pi\frac{ln}{N}}$$

Not dependent on *l*

$$\hat{g}(m,n) = \sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)\frac{1}{M}e^{-j2\pi\frac{km}{M}}\,\frac{1}{N}\, e^{-j2\pi\frac{ln}{N}}$$

$$\hat{g}(m,n) = \sum_{k=0}^{M-1}\frac{1}{M}\, e^{-j2\pi\frac{km}{M}}\,\boxed{\sum_{l=0}^{N-1} g(k,l)\,\frac{1}{N}\, e^{-j2\pi\frac{ln}{N}}}$$

1D: DFT

DFT is linear and separable!!

2D DFT – book notation

$$G(k) = \mathcal{F}\{g(x)\} = \sum_{x=0}^{w-1} g(x) e^{-j2\pi kx/w}$$

$$g(x) = \mathcal{F}^{-1}\{G(k)\} = \frac{1}{w}\sum_{k=0}^{w-1} G(k) e^{j2\pi kx/w}$$

- The 2D DFT is a natural extension of the 1D case.

  w and h is width and height of image:

$$G(k_x, k_y) = \sum_{x=0}^{w-1}\sum_{y=0}^{h-1} g(x,y) e^{-j2\pi \mathbf{x}^{\mathrm{T}}\mathbf{f}} \qquad \text{(forward DFT)}$$

$$g(x, y) = \frac{1}{wh}\sum_{k_x=0}^{w-1}\sum_{k_y=0}^{h-1} G(k_x, k_y) e^{j2\pi \mathbf{x}^{\mathrm{T}}\mathbf{f}} \qquad \text{(inverse DFT)}$$

$$\mathbf{x} = [x \quad y]^T$$

$$\mathbf{f} = [\frac{k_x}{w} \quad \frac{k_y}{h}]^T$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{f} = \frac{k_x x}{w} + \frac{k_y y}{h}$$

Scaling factor. Sometimes (some notations) placed in forward equation, sometimes in invers equaiton, and sometimes as 1/sqrt(*) in both forward and invers. The latter gives a unitary transform ($U^{*T} \cdot U = I$)

DC component in 2d FFT

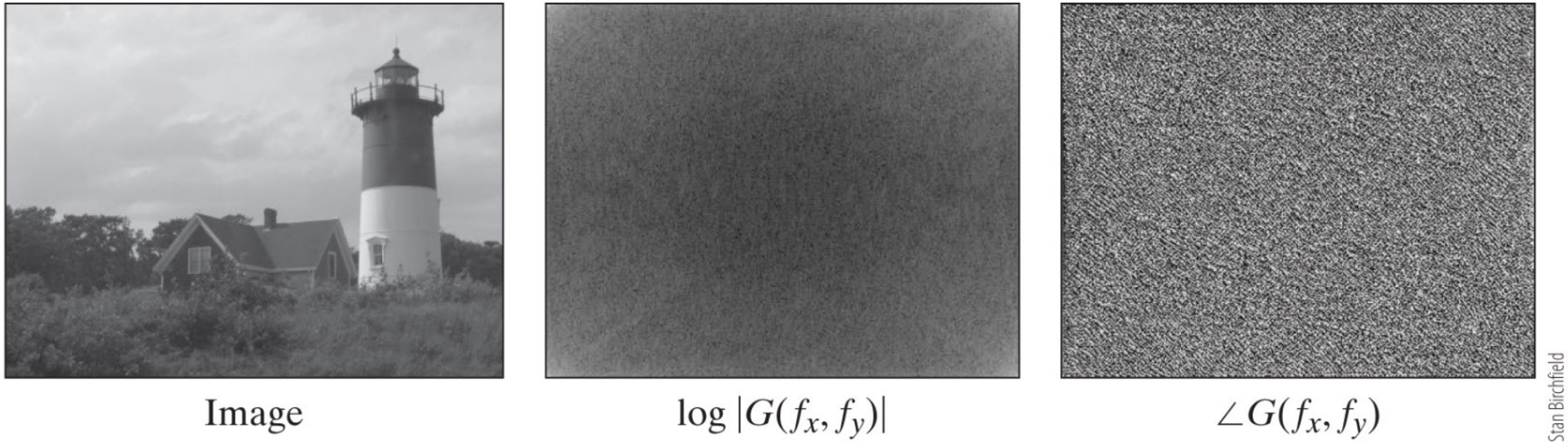
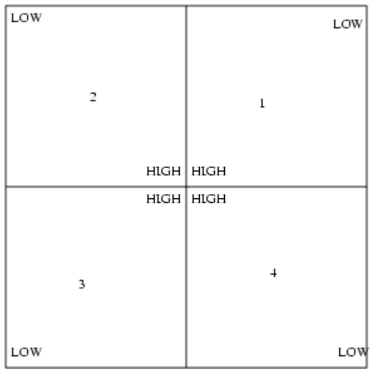
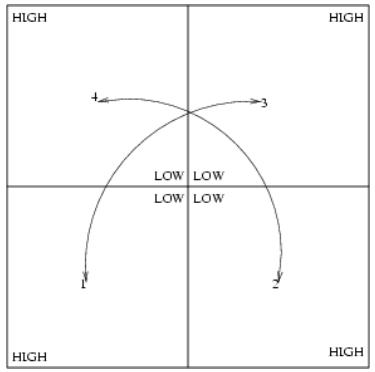


Figure 6.10 An image and its 2D DFT shown as magnitude and phase. (To increase the dynamic range of the display, the log of the magnitude is shown.) The DC component, which is the top-left corner of the magnitude, is difficult to see.





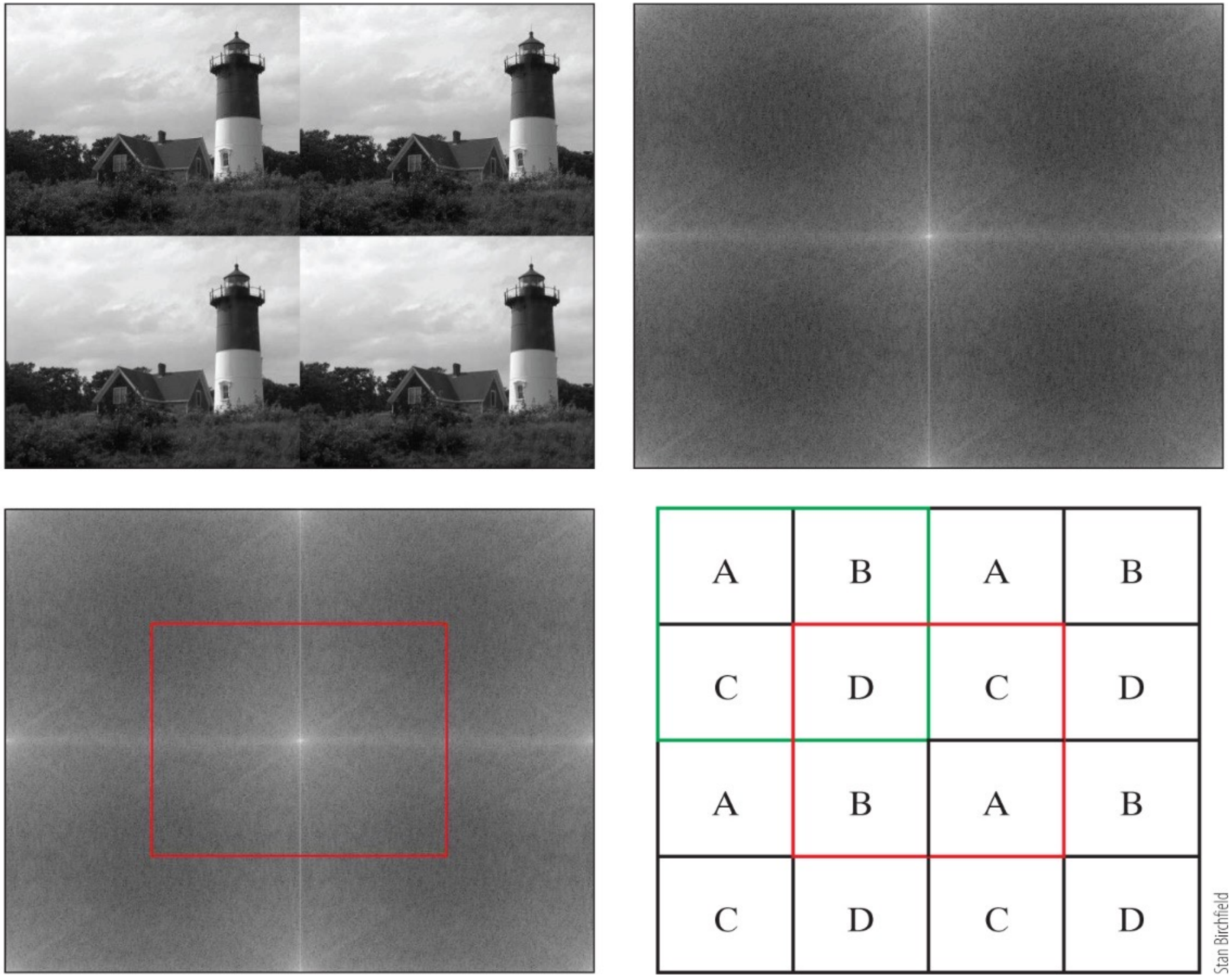
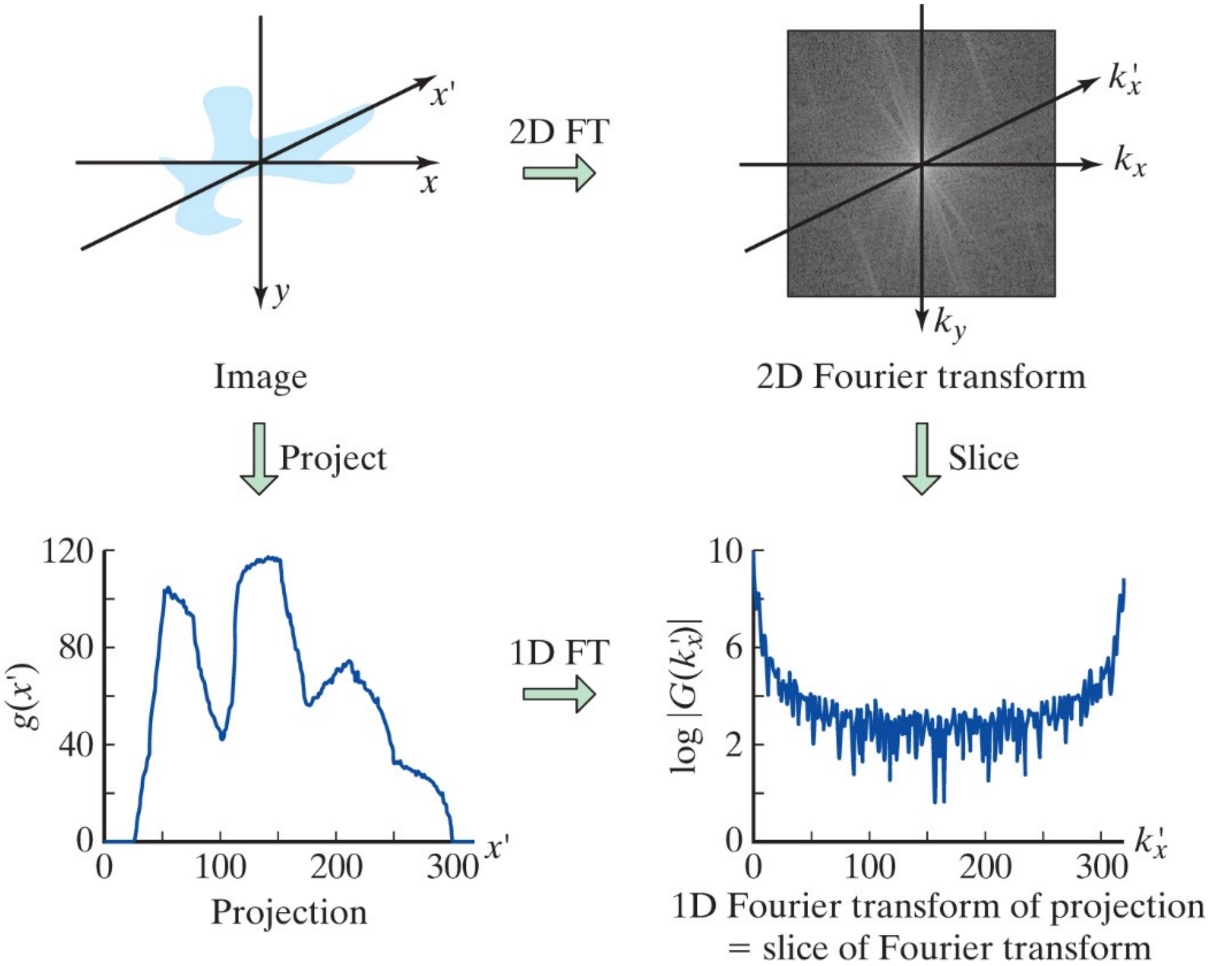


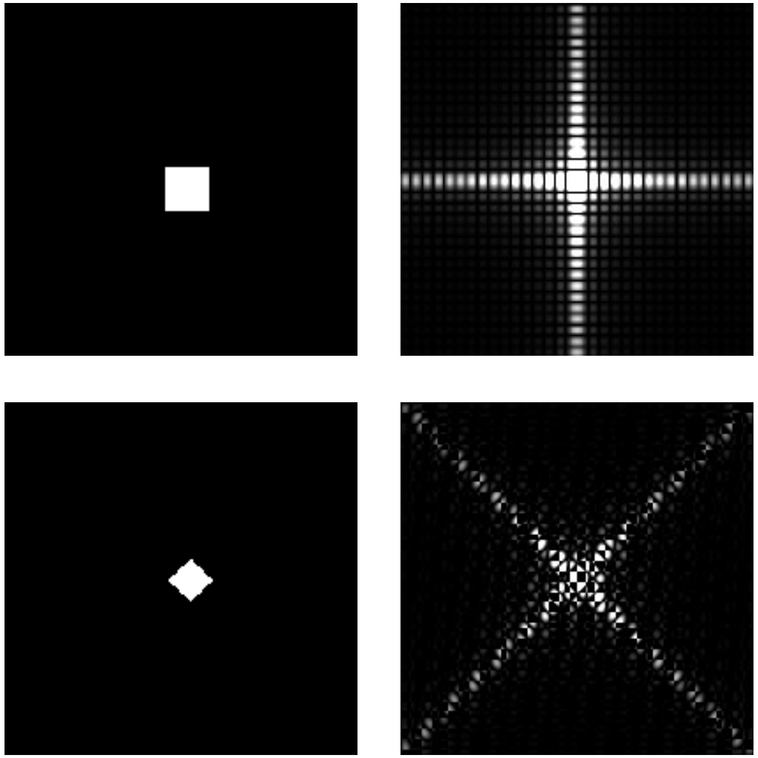
Figure 6.12 Top: The DFT treats the input as a replicated input, and produces a replicated output. Bottom: It is easier to visualize the DFT by shifting it so that the DC component is in the center, which causes no loss of information. The quadrants A, B, C, and D present in the original DFT output are also present in the shifted output, just in a different order.

Projection-slice theorem

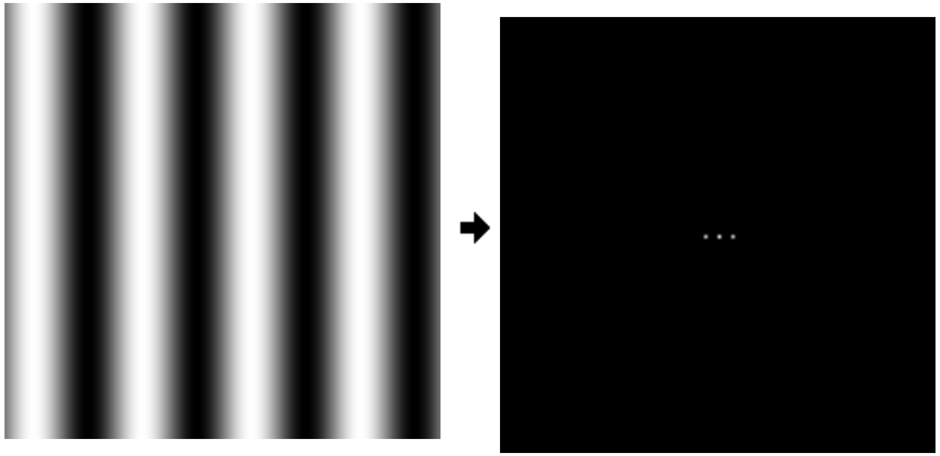


Fourier transform – more on properties

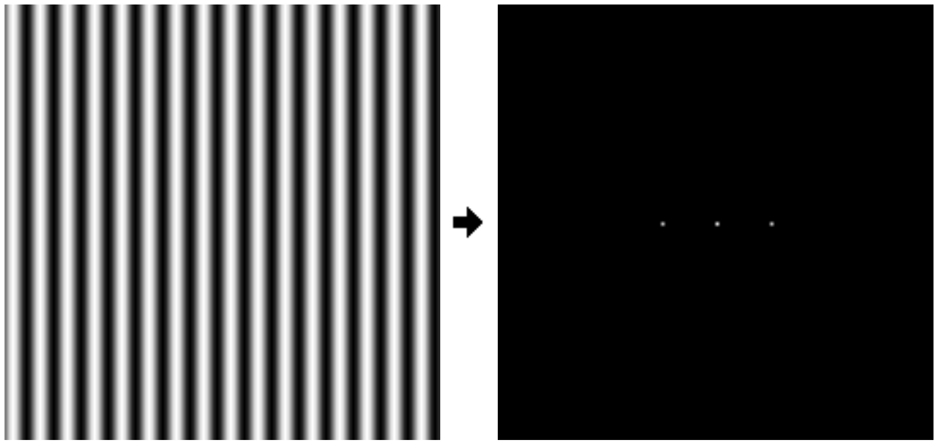
- Linear
- Separable
- Invertable
- Convolution theorem (convolution in space/time domain is multiplication in frequency domain)
- Rotation (if g is rotated by an angle, the FT of the rotated image equals the FT of the original image rotated by the same angle.)



Fourier examples



The center dot is the average DC value. The other two dots represent the perfect sine wave that the Fourier Operator found in the image. There are exactly 4 cycles across the image, and as a result two frequency points are 4 pixels away from the center DC value.
The two points represent the single wave with two different descriptions; one with a negative direction and phase compared to the other.

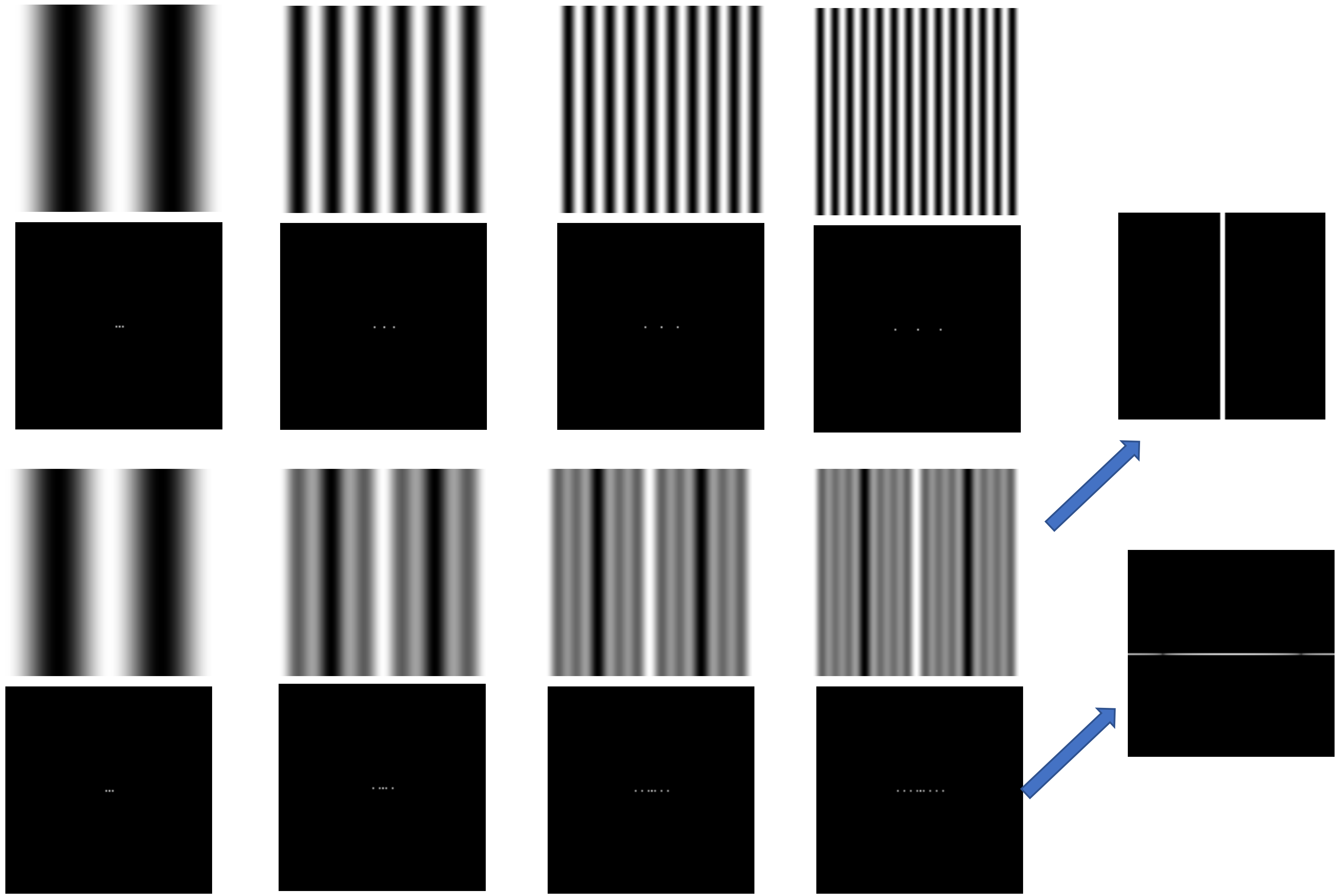


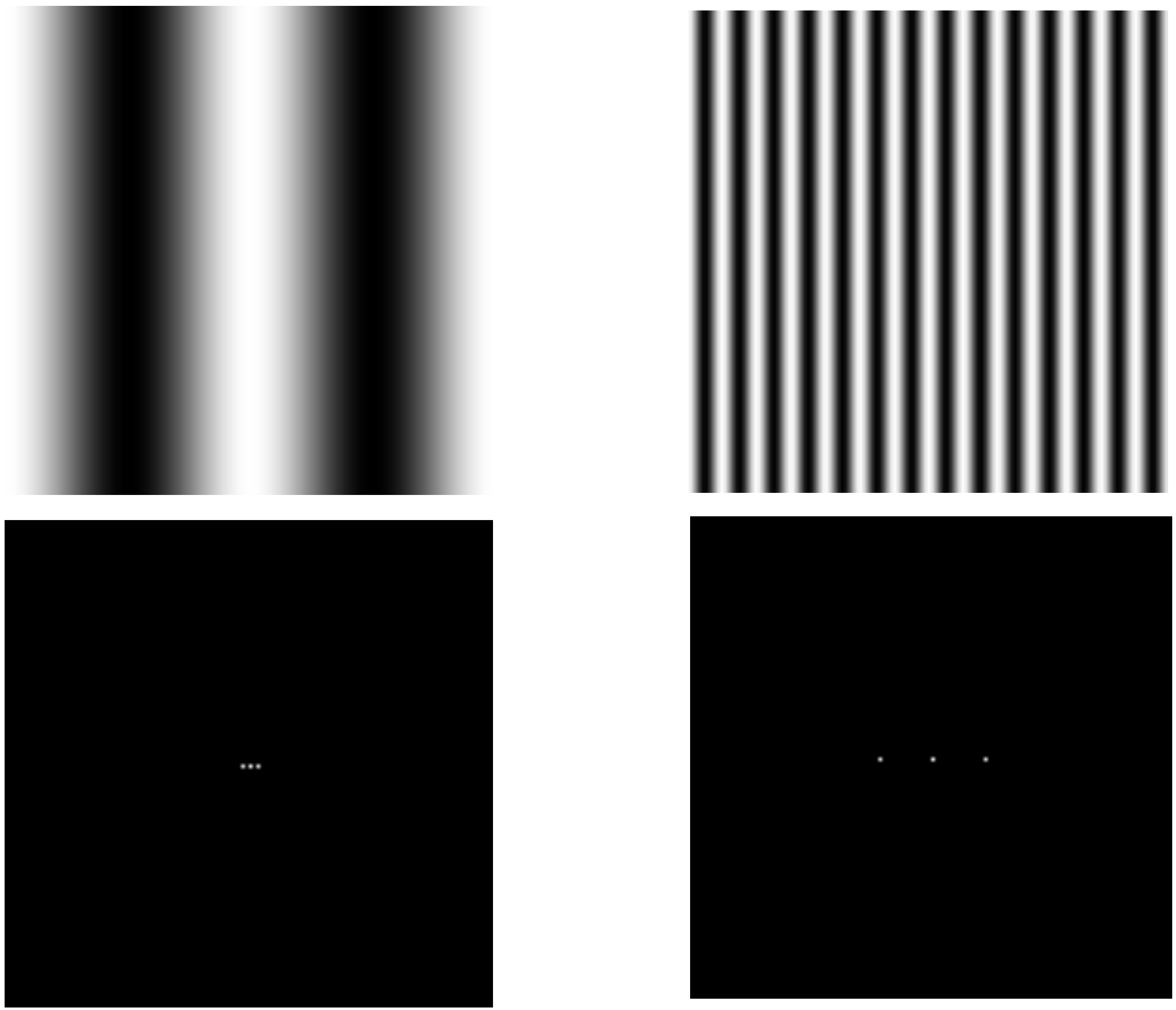
Higher frequency = smaller wavelength.
In the Fourier specter the frequency points will be further away from the DC value.

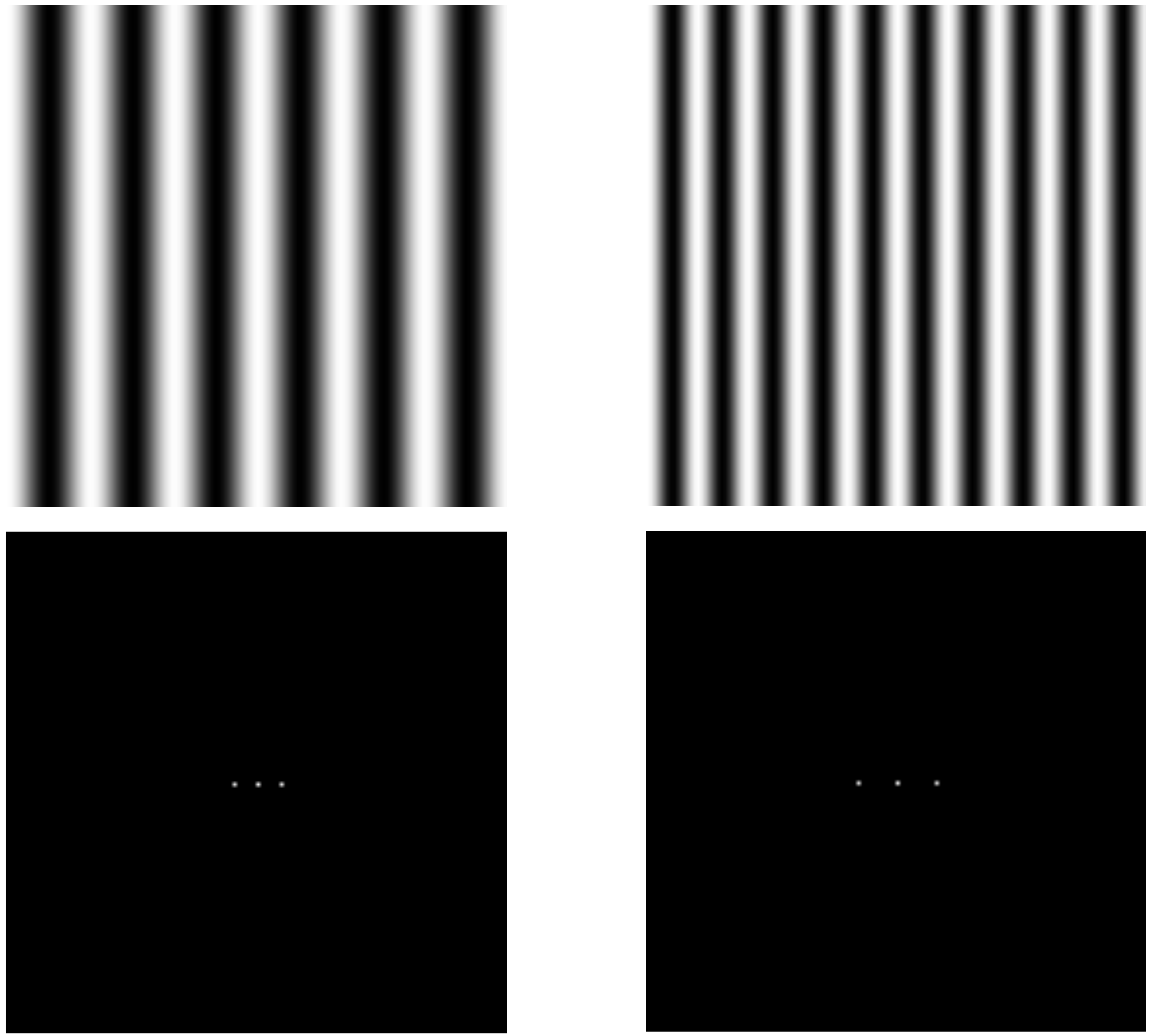
http://www.imagemagick.org/Usage/fourier/#fft
http://cns-alumni.bu.edu/~slehar/fourier/fourier.html

Above; sine
1, 3, 5, 7
(harmonics)

Below; 1,
(1+3),
(1+3+5),
(1+3+5+7)

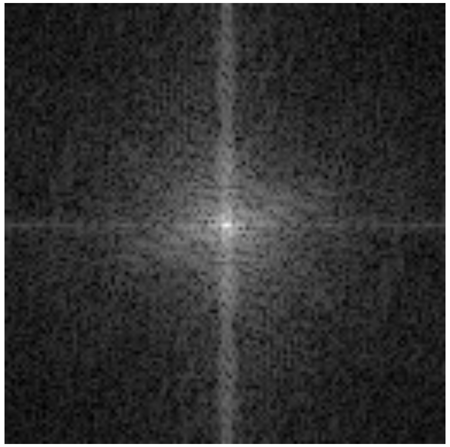




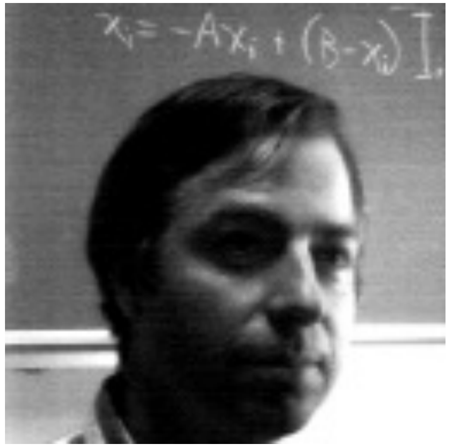


Image

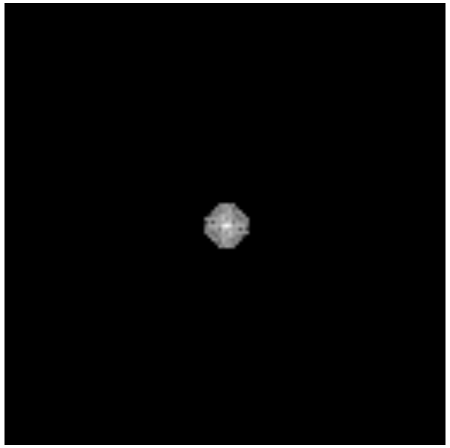
Fourier transformed



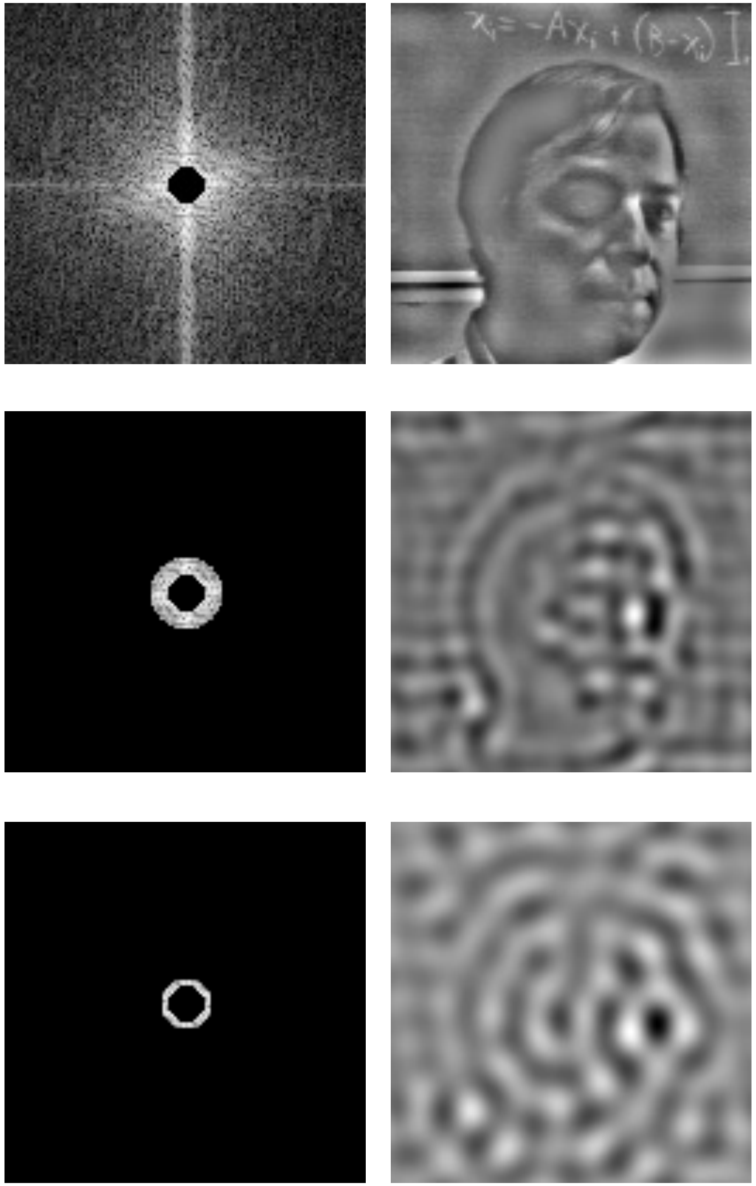
invers Fourier transformed

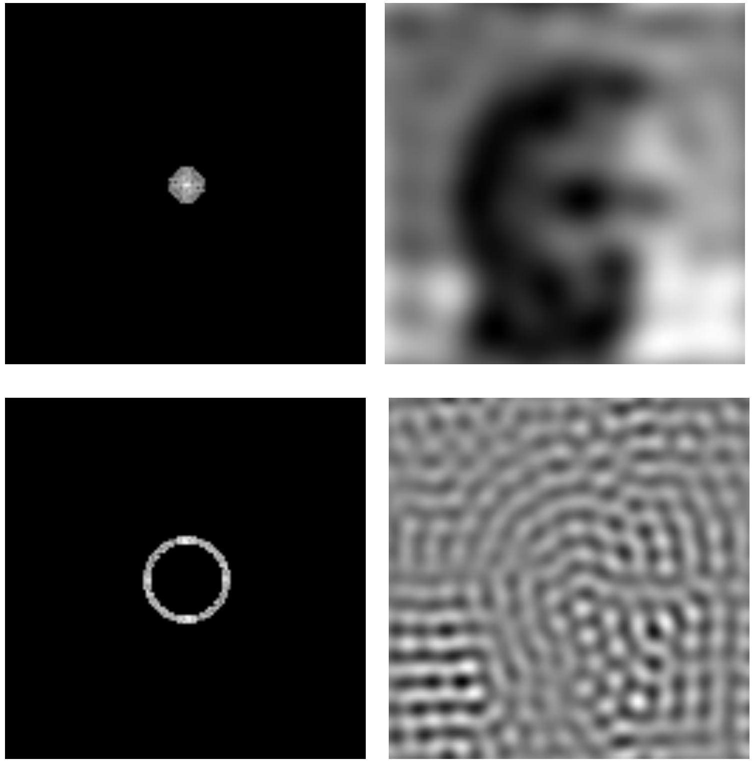


Lowpass filtered and invers transformed.

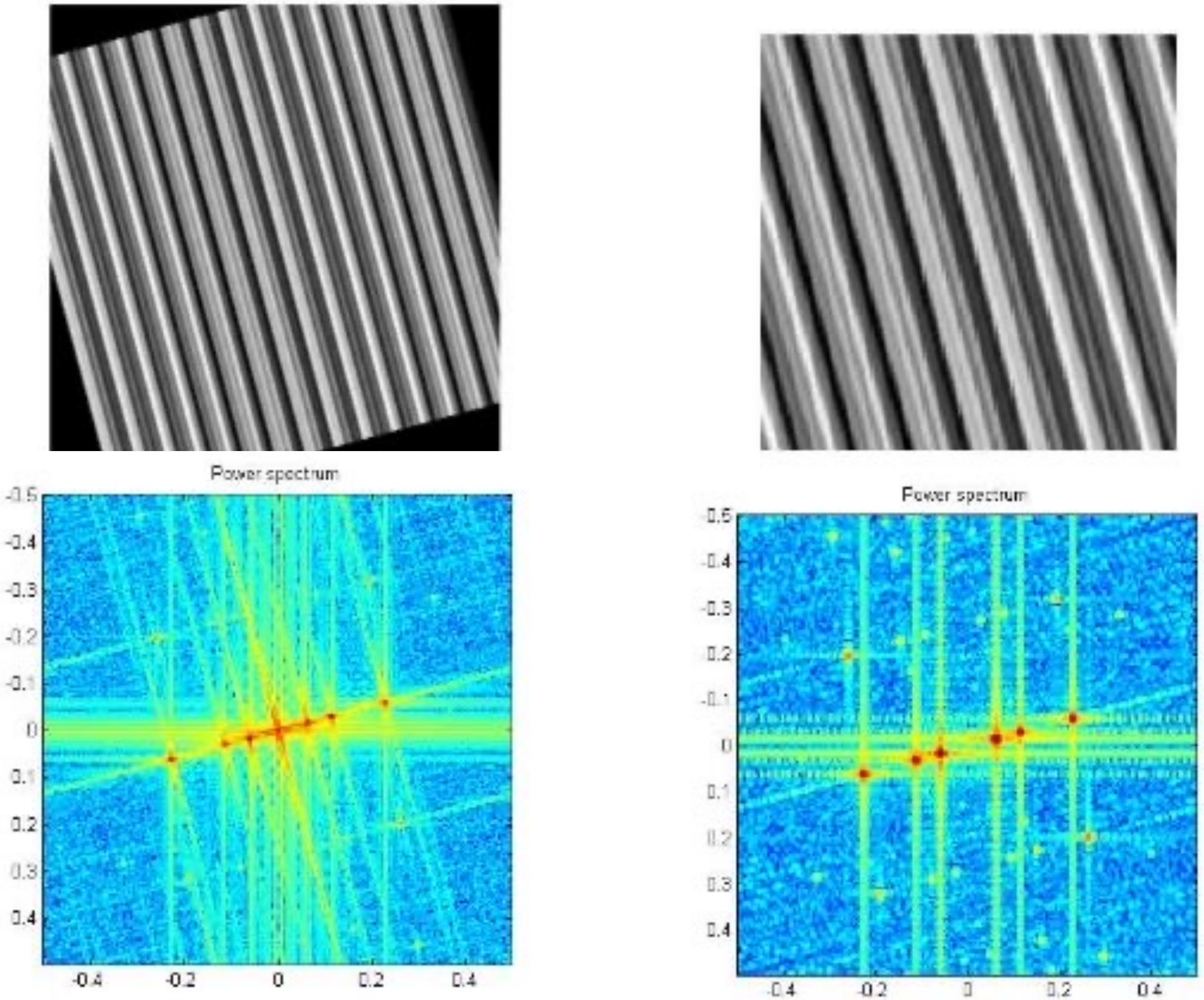


From: http://cns-alumni.bu.edu/~slehar/fourier/fourier.html





From:
http://cns-alumni.bu.edu/~slehar/fourier/fourier.html


Power spectrum
Power spectrum

Fourier transform is generally complex

- We need to consider both magnitude and phase!
- In general for a complex number:

$$c = a + jb = |c|e^{j\theta}$$

$$|c| = \sqrt{a^2 + b^2} \qquad \theta = \tan^{-1}(\frac{b}{a})$$

More on DFT is complex

DFT of image

image

$$G(k_x, k_y) = \mathcal{F}[g(x,y)] = real\{\mathcal{F}[g(x,y)]\} + j \cdot im\{\mathcal{F}[g(x,y)]\}$$

$$Phase : arg(\mathcal{F}[g(x,y)) = \tan^{-1}(\frac{im\{\mathcal{F}[g(x,y)]\}}{real\{\mathcal{F}[g(x,y)]\}})$$

Phase

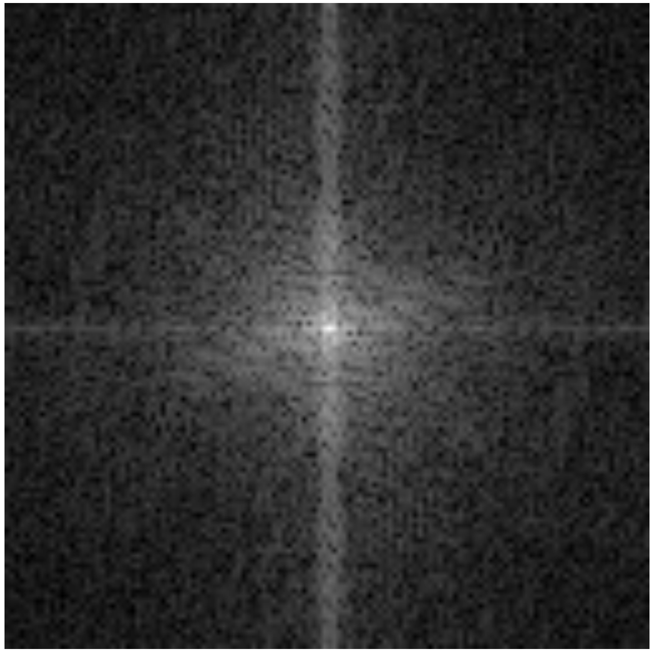
Magnitude

$$|\mathcal{F}[g(x,y)| = \sqrt{im\{\mathcal{F}[g(x,y)]\}^2 + real\{\mathcal{F}[g(x,y)]\}^2}$$

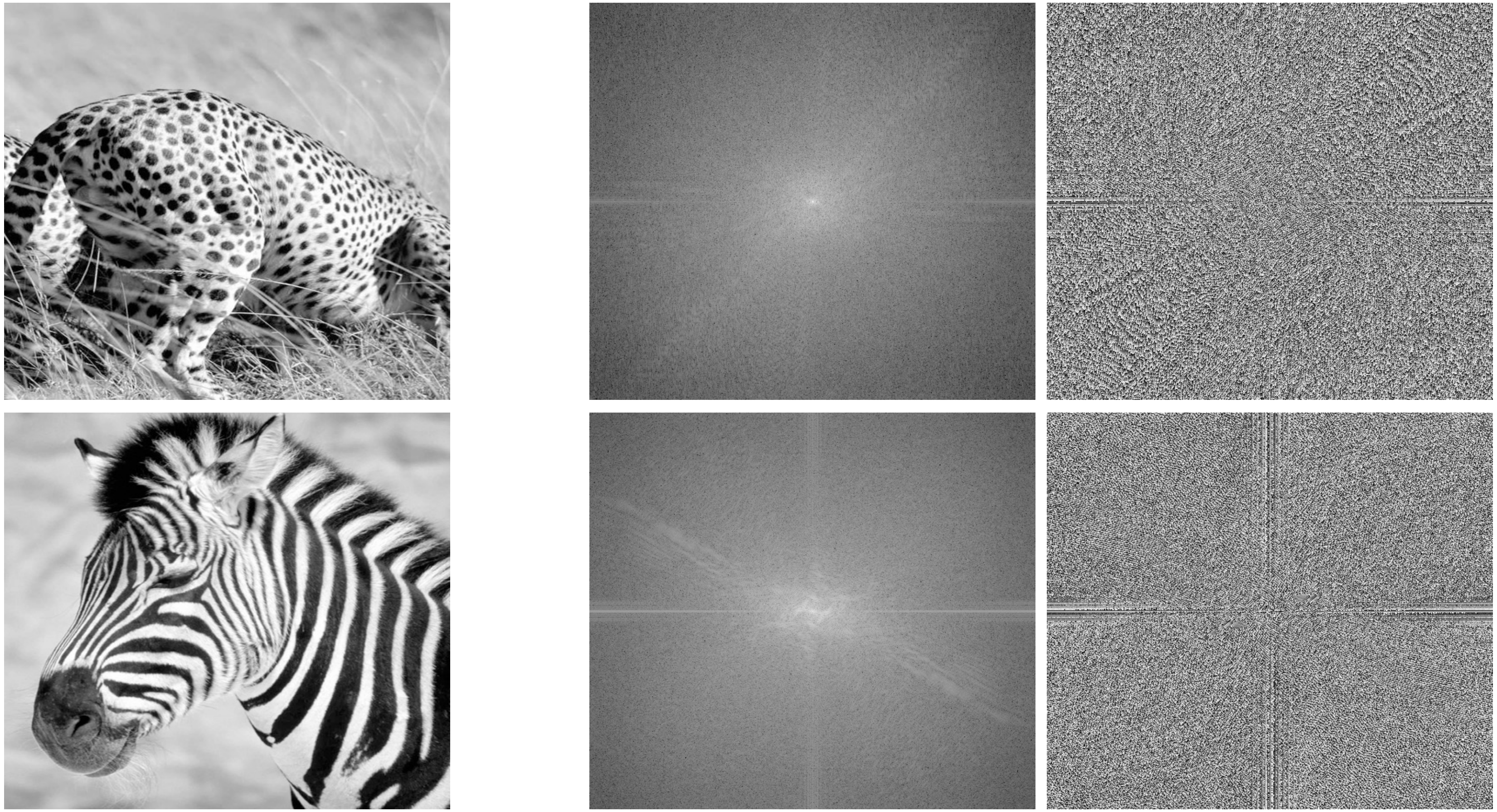
$$|\mathcal{F}[g(x,y)| = \sqrt{\mathcal{F}[g(x,y)]\mathcal{F}[g(x,y)]^*}$$

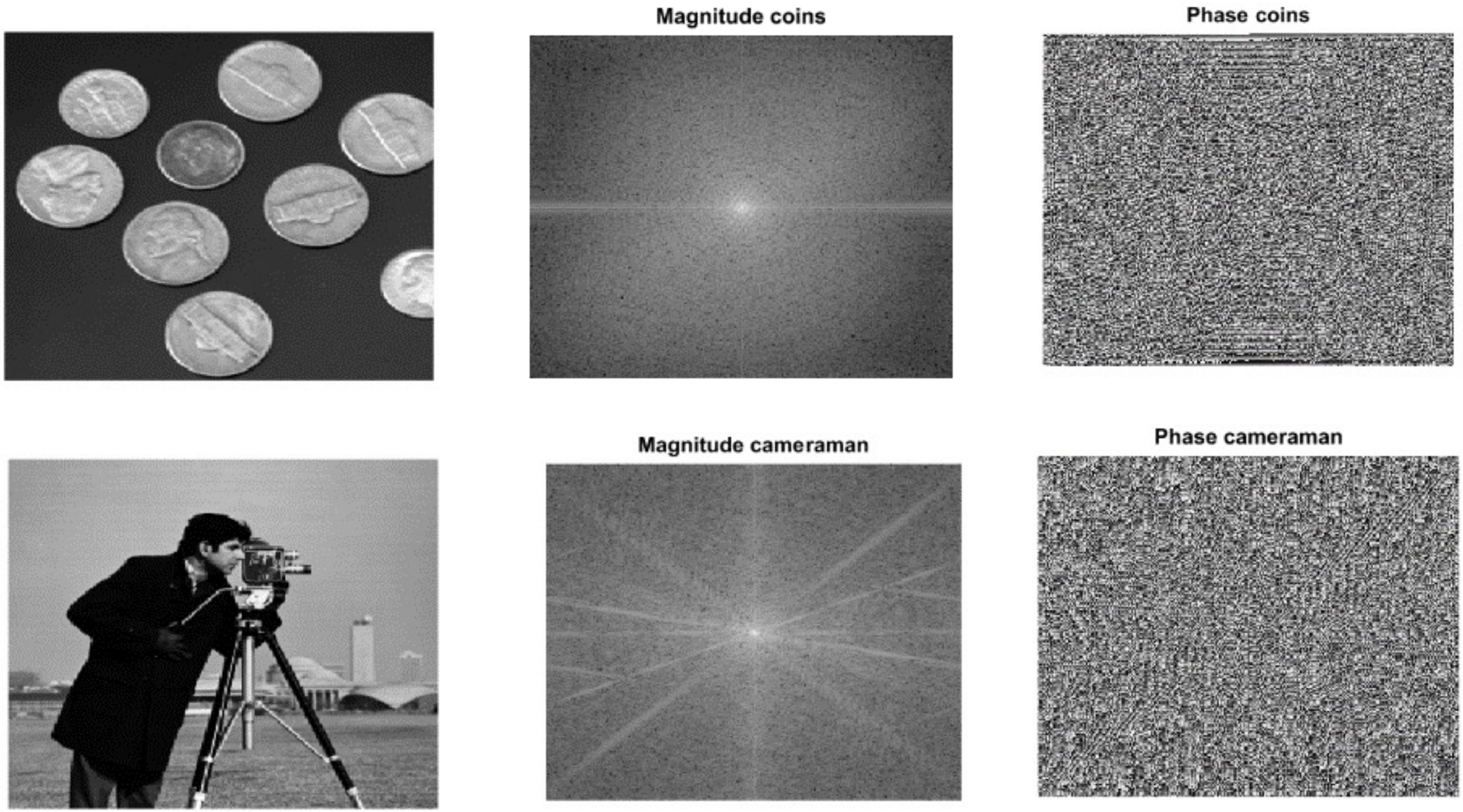
More on DFT is complex

- In previous examples we looked at the DFT magnitude



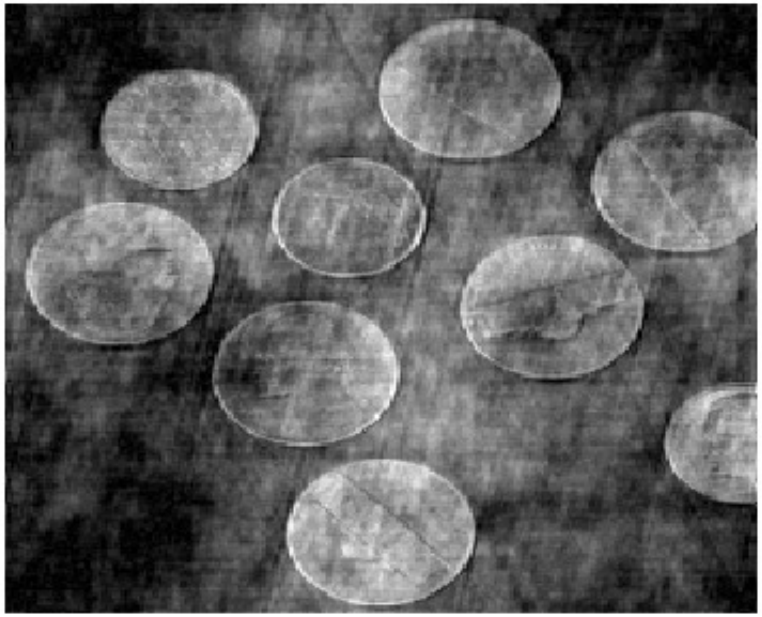
Examples, magnitude and phase




Magnitude coins
Phase coins
Magnitude cameraman
Phase cameraman

The phase is surpringsingly important!

magnitude from cameraman, phase from coins



magnitude from coins, phase from cameraman

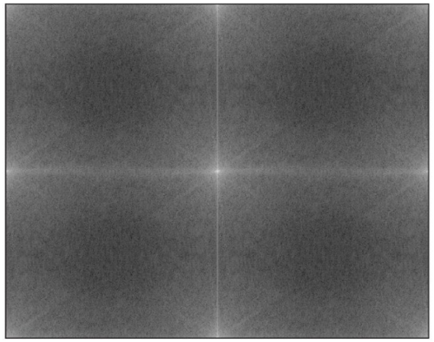
Power spectral density - Periodogram

- How is the power (or energy) of the signal/image distributed over the different frequencies?
- Often seen as magnitude of DFT squared, called Periodogram:

$$\boxed{P(i,j) = \left|G(k_x\,, k_y)\right|^2}$$

Parseval's theorem:

The energy of a signal in the time domain equals the energy of the transformed signal in the frequency domain.



Remember: DFT assumes implicit that the input signal/image is periodic. This gives false edges

Often better to use a window / bellshape to smooth the effect of edges before looking at the power spectrum density P(i,j)

This is called windowed Periodogram

There are many other methods to estimate power spectrum – not covered here.

2D DFT – some observations

- Offers a rich representation of an image
- It is a global transform – input the entire image
- Can apply the DFT in local windows. This leads to what is called Gabor functions

- DFT is in gerneral complex – which can be a disadvantage.
- If the image is symmetric about both axes, the DFT values are real (not complex). This leads to the definition of Discrete Cosine Transform (DCT)

DFT vs Discrete Cosine Transform (DCT)

- DFT is genereally complex
  - Symmetric extension would be real
  - This idea is used when defining the DCT

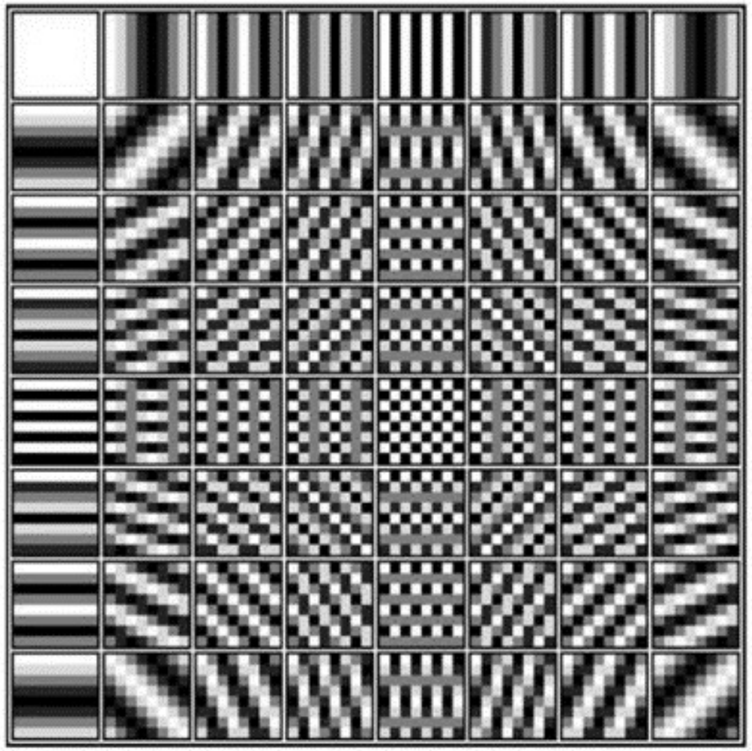
- DFT implicitly treat the signal/image as if it was one period of a periodic signal/image

- **Discrete Cosine Transform (DCT)** is a variant – DCT implicitly treat the signal/image as if it was one period of an mirrowed/symmetric extension of the signal/image – becomes real (not complex)

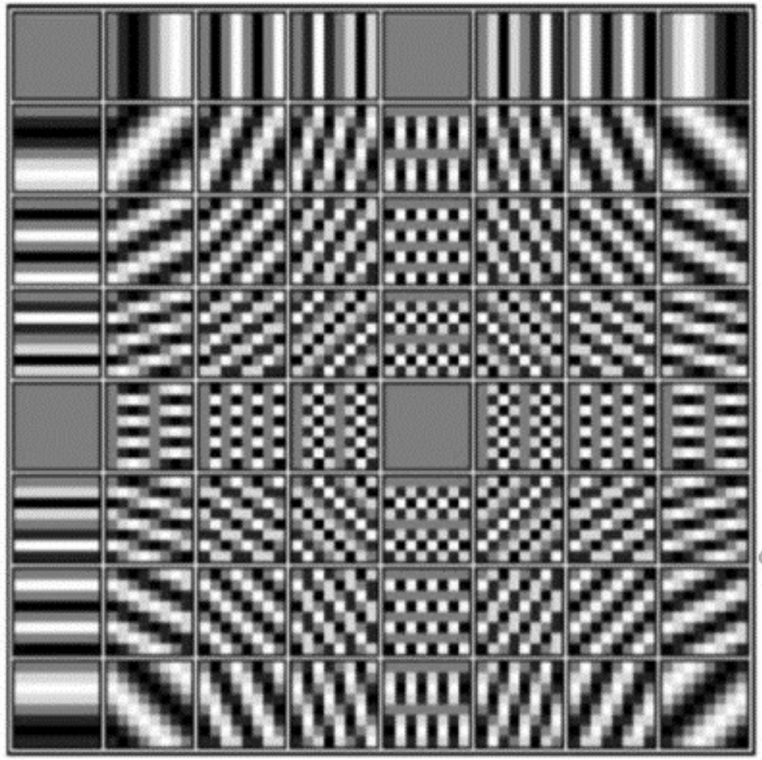
- DCT is much used in compression (jpeg, mpeg) and many other things

(Remember eigenimages, SVD)
Basis images for 8x8 image, DFT

$$\hat{g} = \sum_{i=1}^{N} \sum_{j=1}^{M} g_{ij} \mathbf{u}_i \mathbf{u}_j^T$$



Real part of basis images

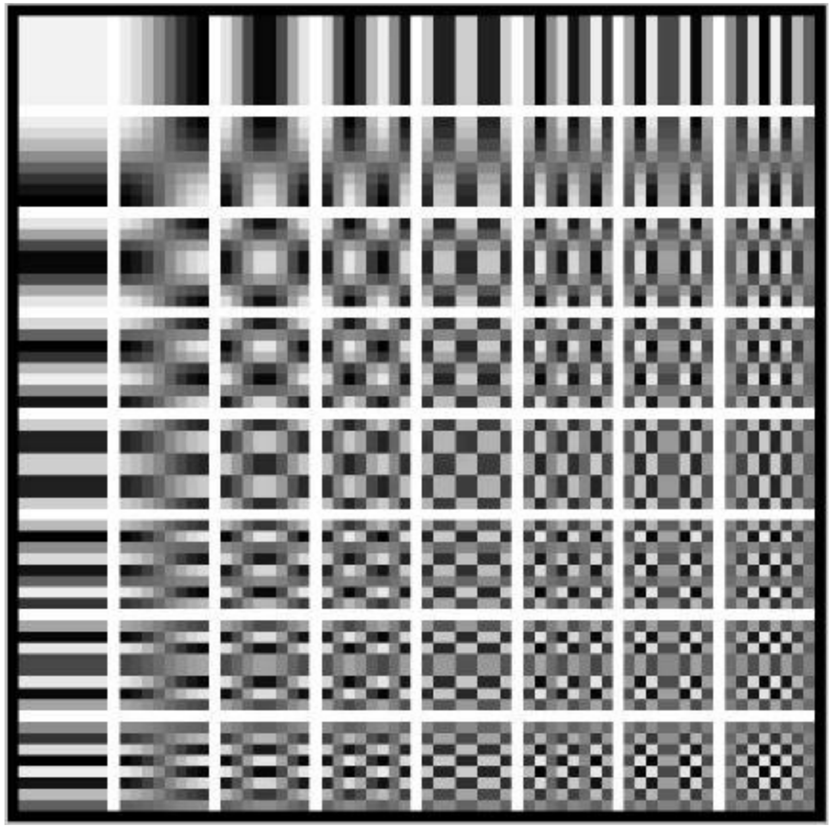


Imaginary part of basis images

Discrete Cosine Transform (DCT) - Basis images for 8x8 image (patch)

Can build any image patch as a linear combination of these basis images.

DCT is a unitary transform. It is not complex, thus it is a orthogonal transform.



$$\hat{g} = \sum_{i=1}^{N} \sum_{j=1}^{M} g_{ij} \mathbf{u}_i \mathbf{u}_j^T$$

Image patch

Basis images

coefficients

2D DFT

Three points from the topic:

1. How can we extend DFT to 2D images?
   ✓ Straight forward. Its separable.

2. What is the projection –slice theorem?
   ✓ Project all intensity values down on a line with an angle in space domain -> 1D DFT of this signal corresponds to a line with same angle in DFT domain

3. What is Discrete Cosine Transform and how does it relate to DFT?
   ✓ DCT implicitly treat the signal/image as if it was one period of an mirrowed/symmetric extension of the signal/image. Gives a real (not complex) transform, similar properties as DFT. Used in ex. JPEG and MPEG