Image Processing

6.815/6.865

Frédo Durand
A bunch of slides by Bill Freeman (MIT)
& Alyosha Efros (CMU)

- define cumulative histogram
- work on hist eq proof

- rearrange Fourier order
- discuss complex exponentials with eigenfunctions
-

Warning

- Think about final projects

Class morph

Image processing

- Filtering, Convolution, and our friend Joseph Fourier

What is an image?

- We can think of an **image** as a function, *f*,
- from R^2 to R:
  - $f(x, y)$ gives the **intensity** at position (x, y)
  - Realistically, we expect the image only to be defined over a rectangle, with a finite range:
    - $f: [a,b]\mathrm{x}[c,d] \rightarrow [0,1]$
- A color image is just three functions pasted together. We can write this as a "vector-valued" function:

$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

Images as functions

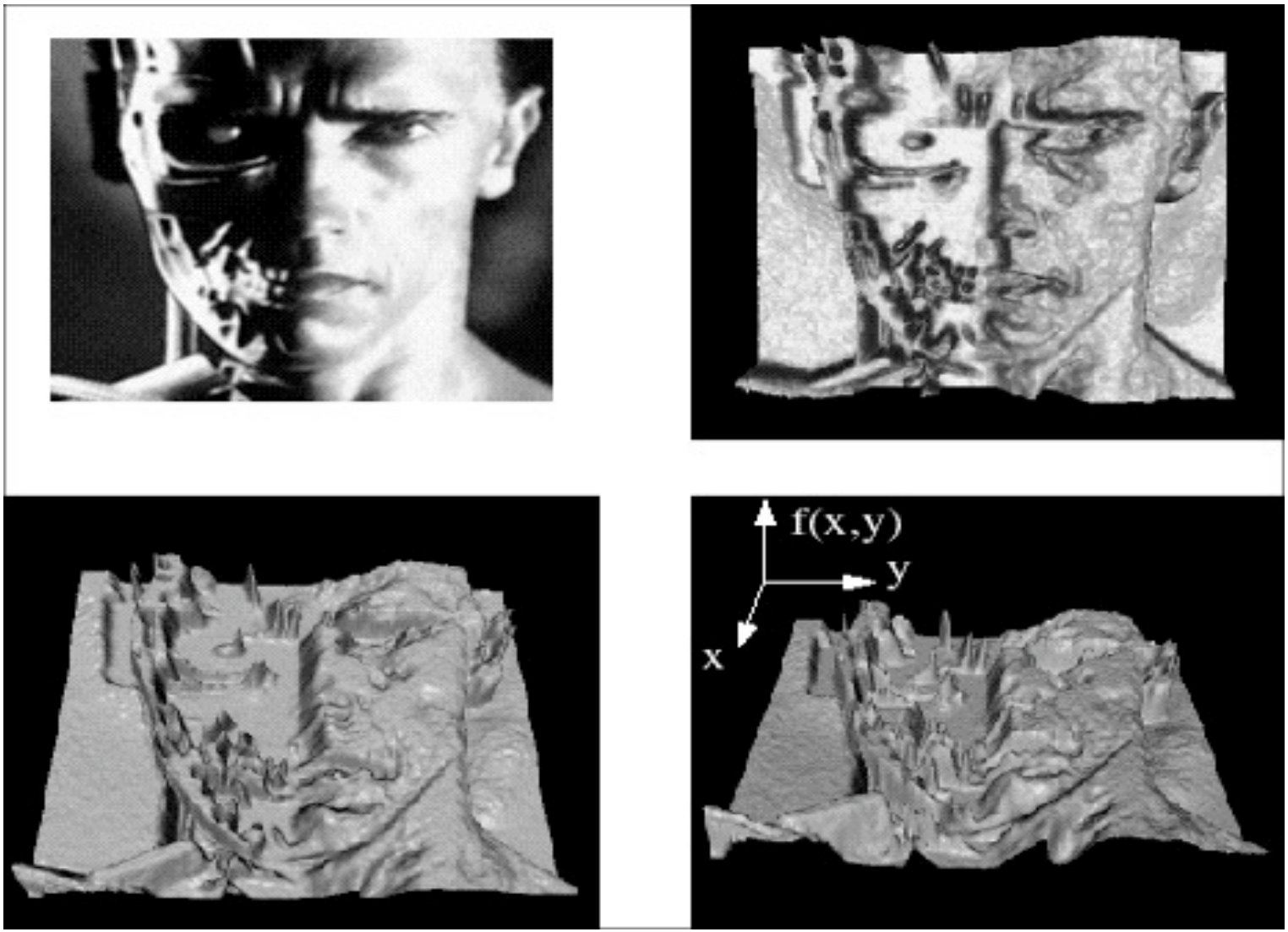
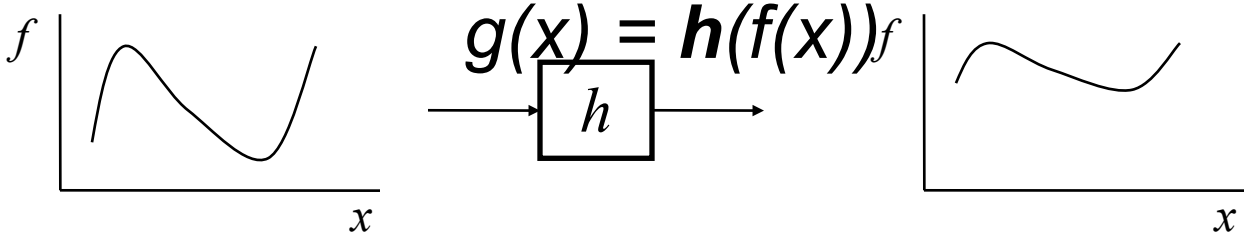


Image Processing

- image filtering: change ***range*** of image
  - 

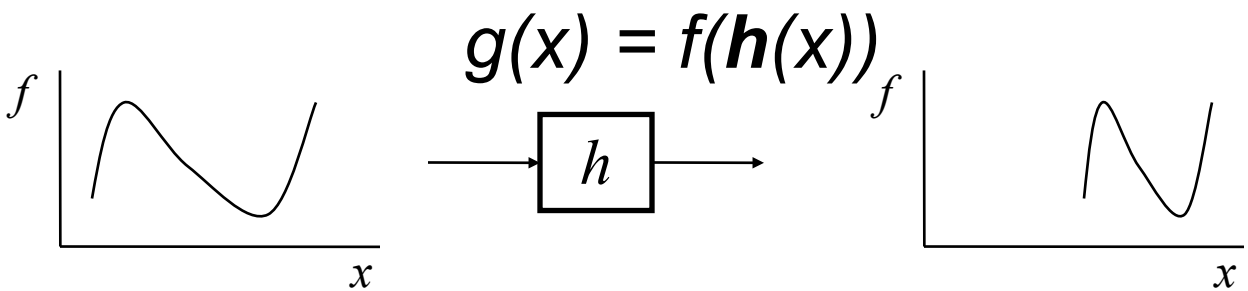
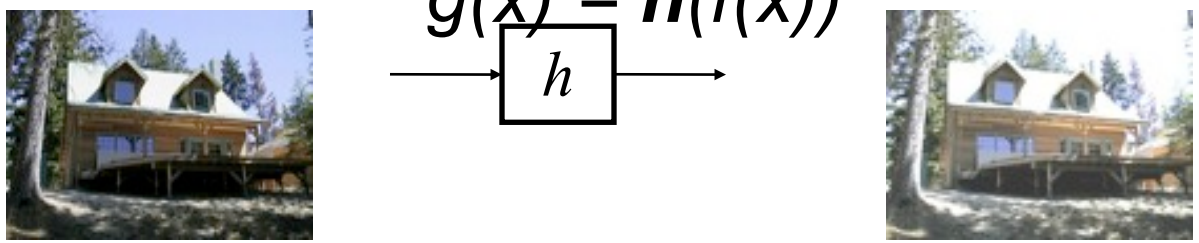
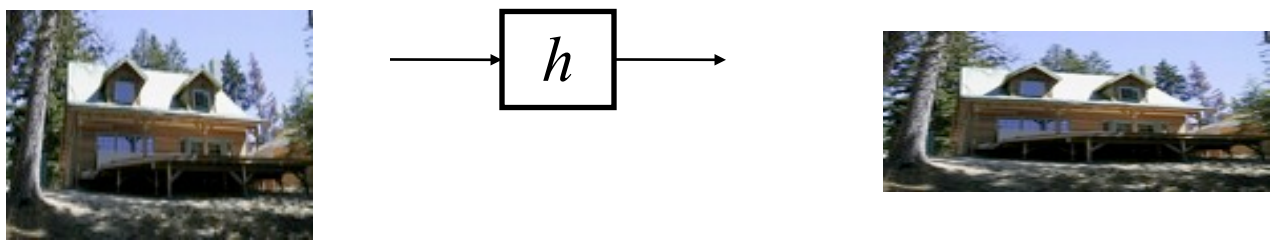
- image warping: change ***domain*** of image
  - 

Image Processing

- image filtering: change ***range*** of image
- $g(x) = \boldsymbol{h}(f(x))$



- image warping: change ***domain*** of image
- $g(x) = f(\boldsymbol{h}(x))$

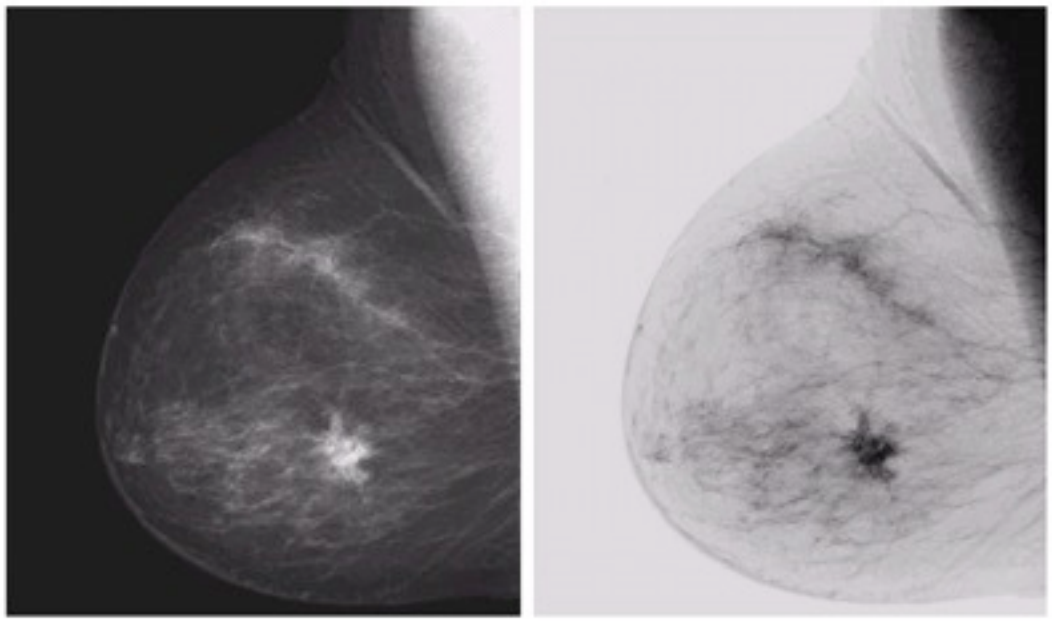


Point Processing

- The simplest kind of range transformations are these independent of position x,y:
- $$g = t(f)$$
- This is called point processing.

- **Important:** every pixel for himself – spatial information completely ignored!

Negative

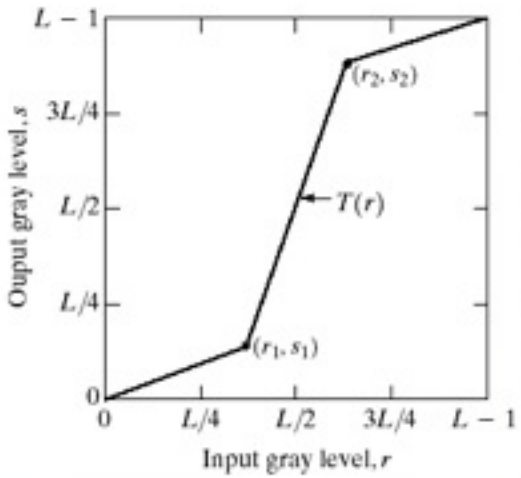


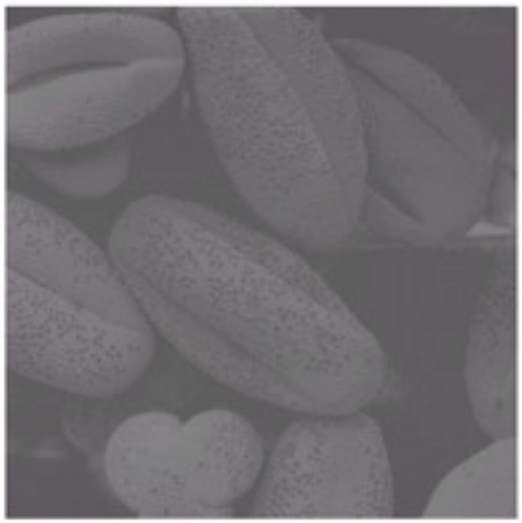
a b

FIGURE 3.4 (a) Original digital mammogram. (b) Negative image obtained using the negative transformation in Eq. (3.2-1). (Courtesy of G.E. Medical Systems.)

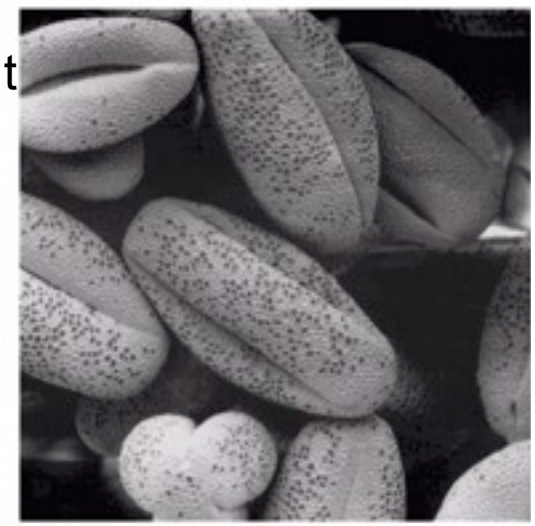
Contrast Stretching

input





output



a b
c d

FIGURE 3.10 Contrast stretching. (a) Form of transformation function. (b) A low-contrast image. (c) Result of contrast stretching. (d) Result of thresholding. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

Image Histograms

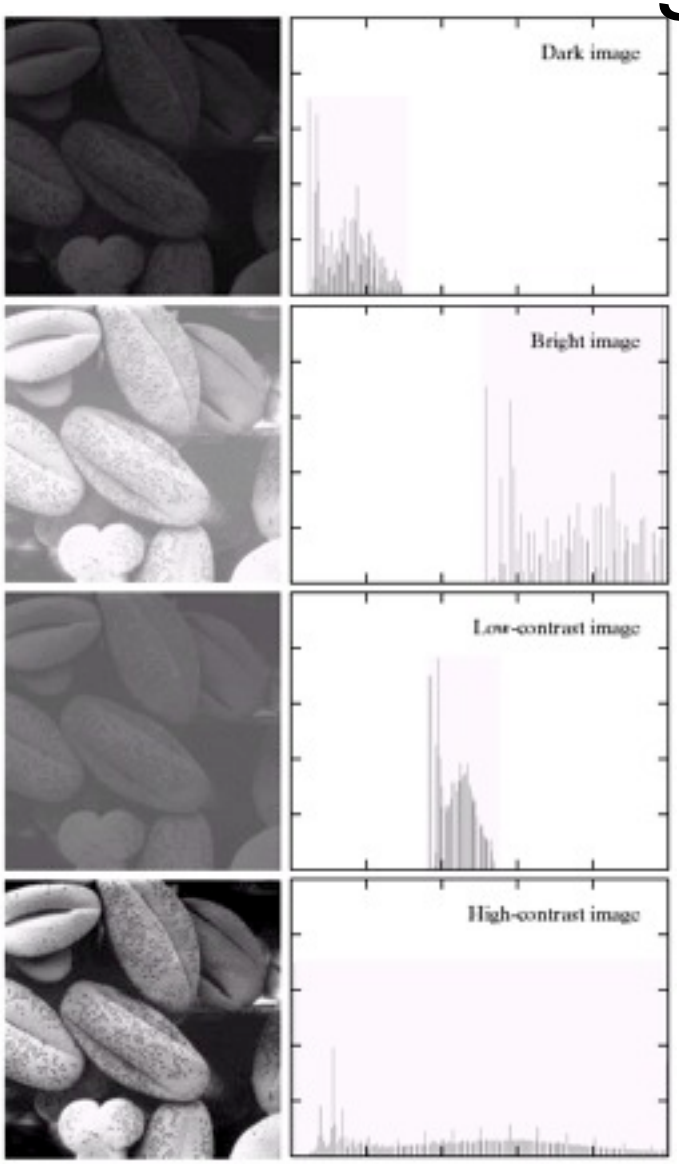
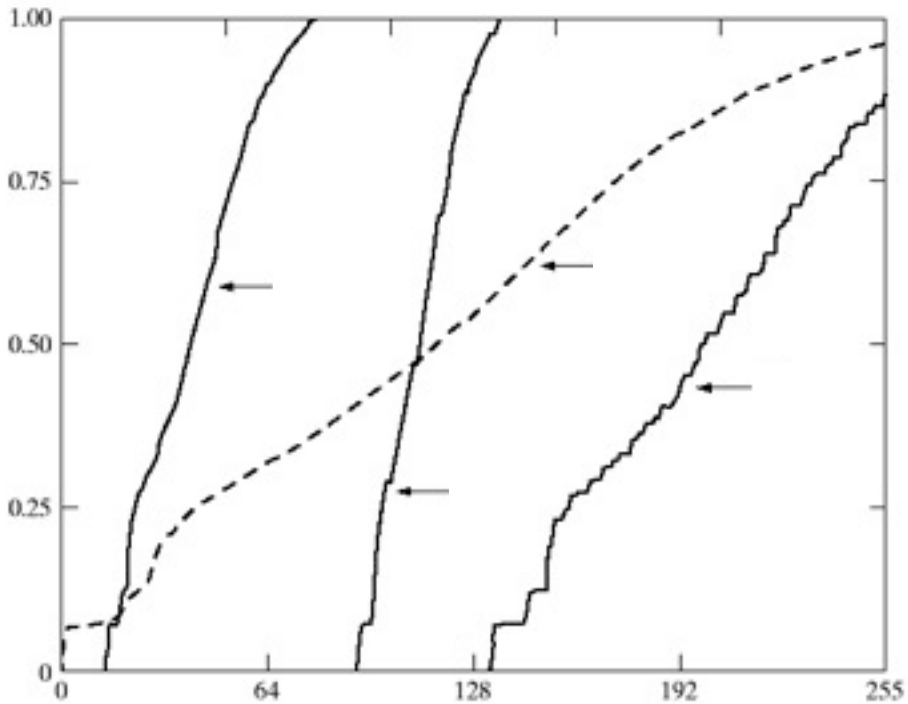


FIGURE 3.15 Four basic image types: dark, light, low contrast, high contrast, and their corresponding histograms. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

histogram H(f) = # or % pixels with value f
(implies binning of the values)

cumulative histogram:
$C(f) = \sum_{f' \le f} H(f)$
= # or % of pixels with value ≤f



Cumulative Histograms

Histogram Equalization

- point transformation: g(x)=t(f(x))
- uniform across image (t does not depend on x)
- monotonic (preserve intensity ordering)
- so that histogram of g is uniform
  - perfect uniform only possible with continuous histogram

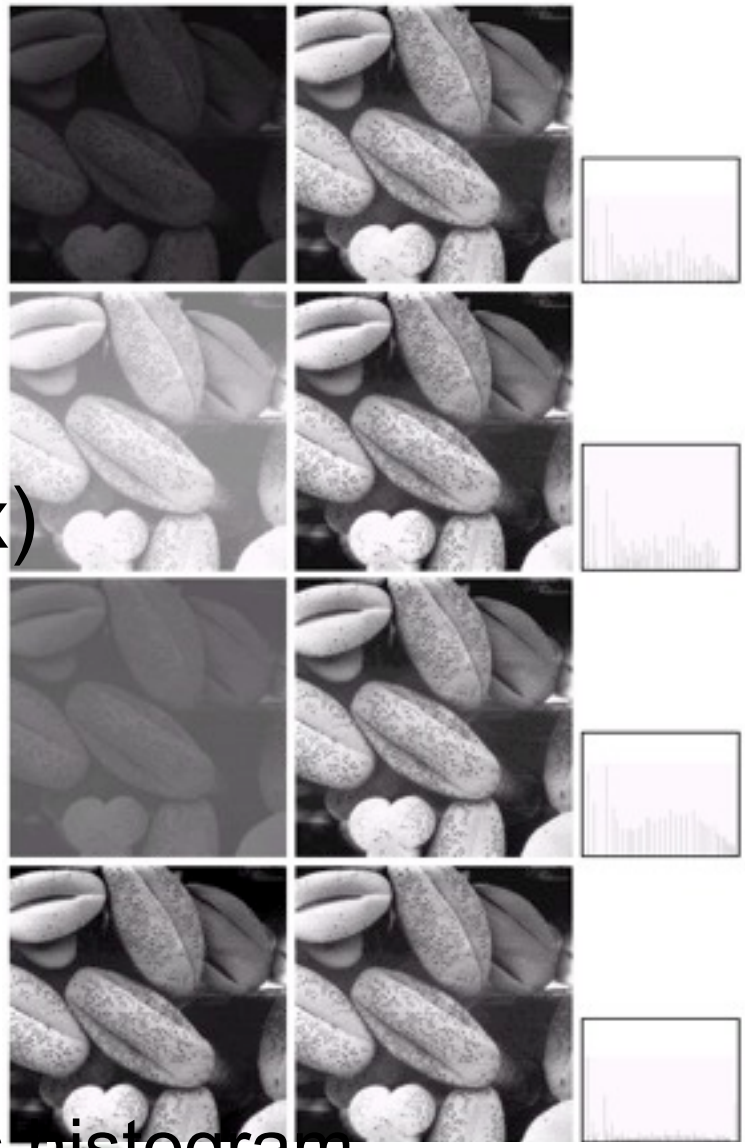
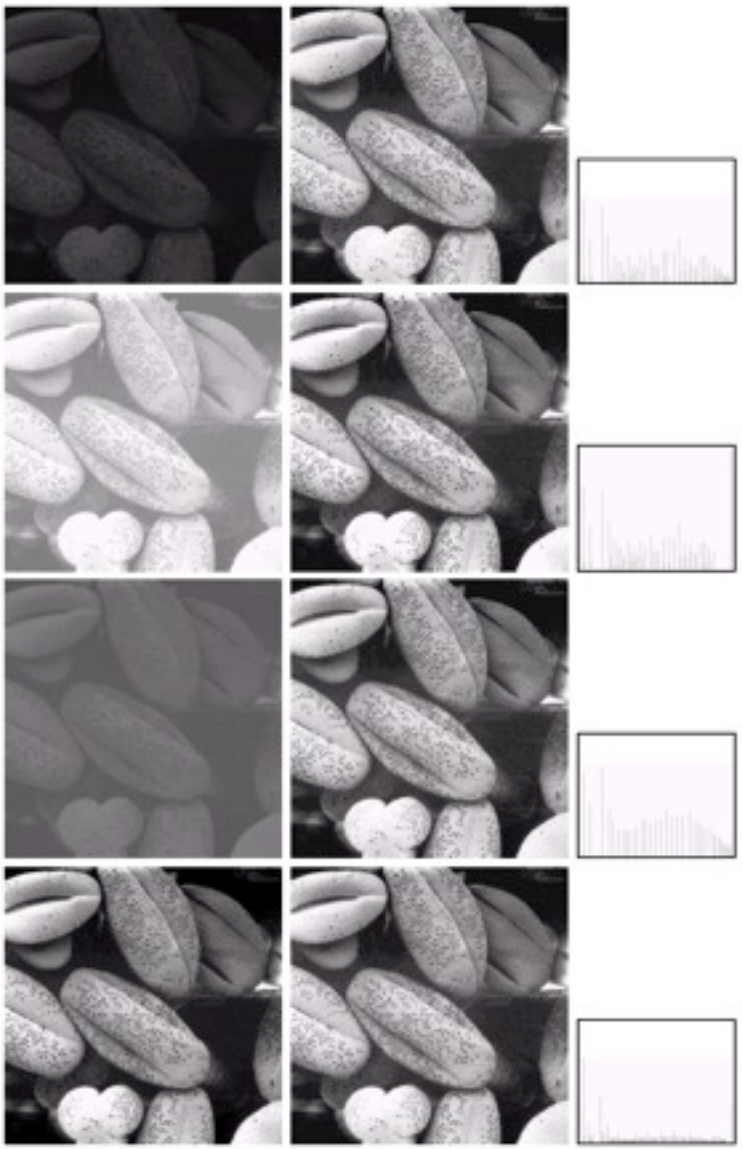


FIGURE 3.17 (a) Images from Fig. 3.15. (b) Results of histogram equalization. (c) Corresponding histograms.

Qualitative Histogram equalization

- Qualitative



Derivation

- Normalized cumulative histogram *C*: there are *C(f)*% pixels equal or darker than *f*
- In an image in [0 1] with a flat histogram, what is the greyscale value *g* so that *C(f)*% pixels are equal or darker than *f*?
  - C(f) of course!
- Therefore, histogram equalization:
  - g(x)=C(f(x))

Extension: histogram matching

- Transform image f to match histogram of f’

f

f’

result

Extension: histogram matching

- Transform image f to match histogram of f'
- $g(x)=C_{f'}^{-1}(C_f(f(x)))$
  - cumulative histogram C_f of f to get the flat case
  - inverse cumulative histogram $C_{f'}^{-1}$of f' to match that histogram
- equalization: case where f' has flat histogram and $C_{f'}^{-1}$ is identity

Questions?

Filtering

- So far we have looked at range-only and domain-only transformation
- But other transforms need to change the range according to the spatial neighborhood
  - Linear shift-invariant filtering in particular

Linear shift-invariant filtering

- Replace each pixel by a linear combination of its neighbors.
  - only depends on relative position of neighbors
- The prescription for the linear combination is called the “convolution kernel”.

10	5	3
4	5	1
1	1	7

Local image data

0	0	0
0	0.5	0
0	1	0.5

kernel

	7	

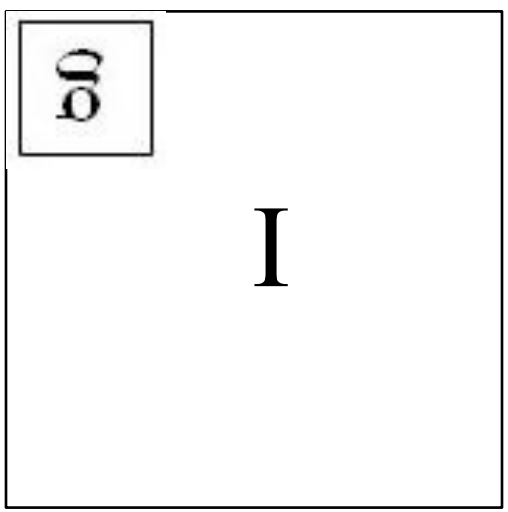
Modified image data
(shown at one pixel)

Example of linear NON-shift invariant transformation?

- e.g. neutral-density graduated filter (darken high y, preserve small y)
  $J(x,y)=I(x,y)*(1-y/ymax)$
- Formally, what does linear mean?
  - For two scalars a & b and two inputs x & y:
    $F(ax+by)=aF(x)+bF(y)$
- What does shift invariant mean?
  - For a translation T:
    $F(T(x))=T(F(x))$
  - If I blur a translated image, I get a translated blurred image

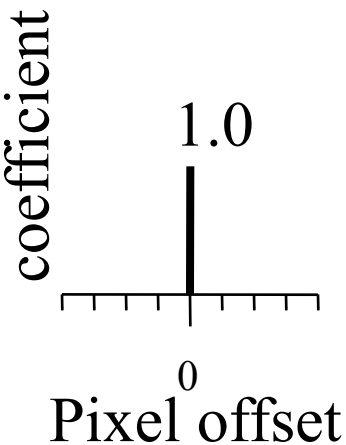
More formally: Convolution

$$f[m,n] = I \otimes g = \sum_{k,l} I[m-k, n-l]g[k,l]$$



Convolution (warm-up slide)

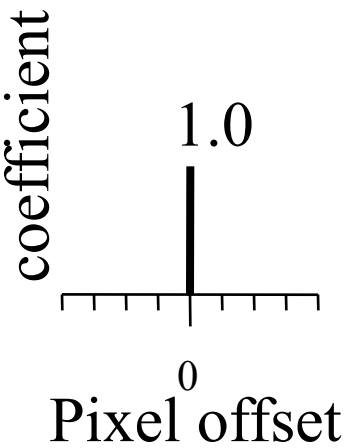
original



?

Convolution (warm-up slide)

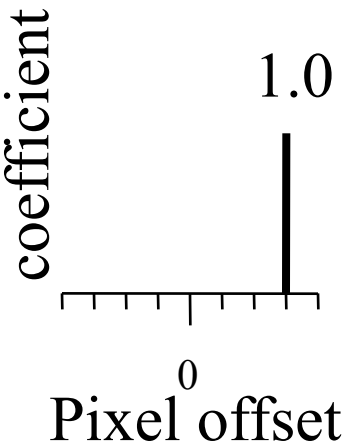
original



Filtered
(no change)

Convolution

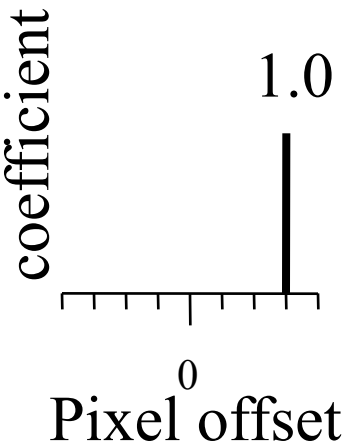
original

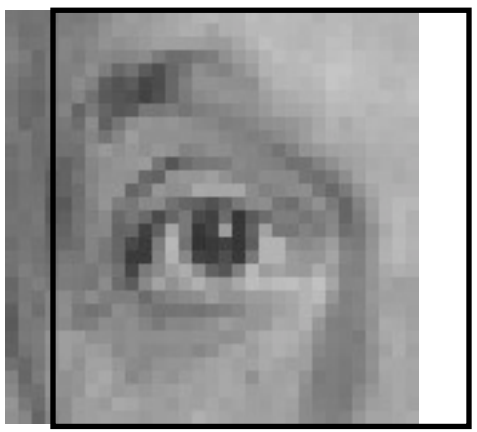


?

shift

original

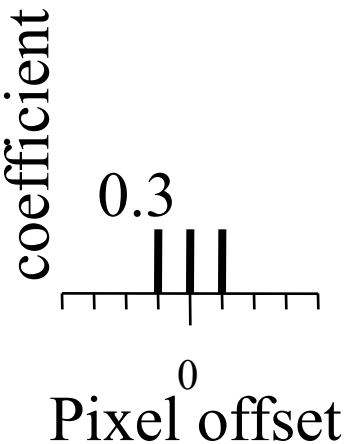




shifted

Convolution

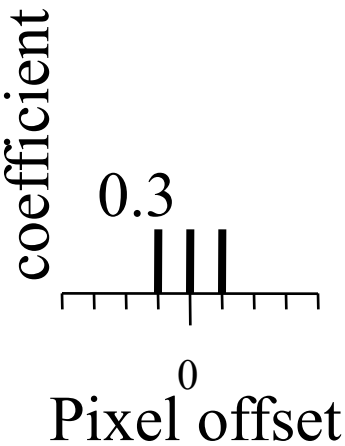
original



?

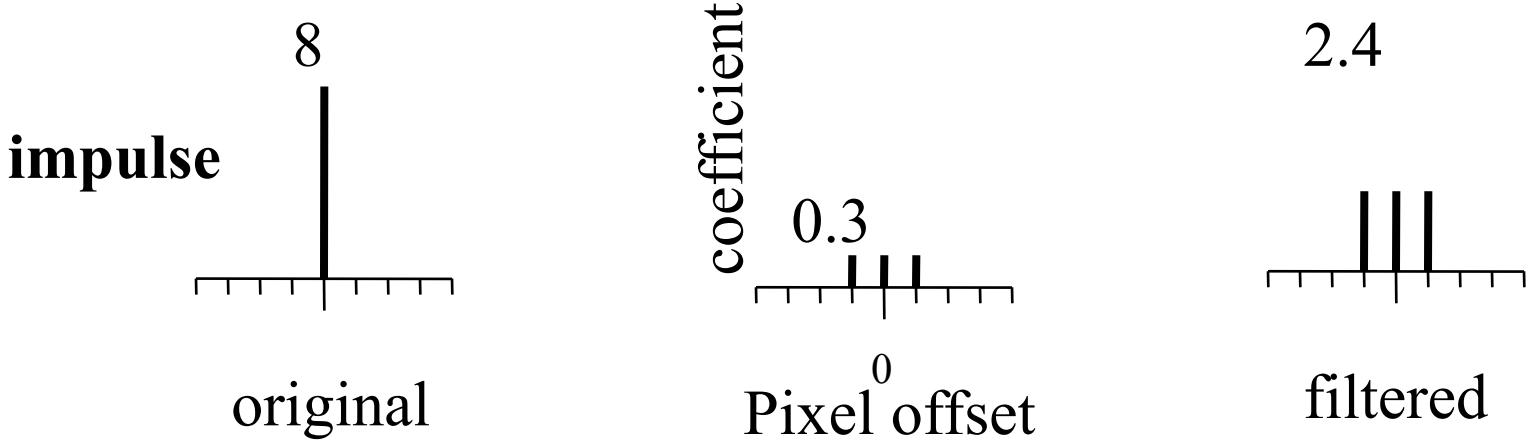
Blurring

original

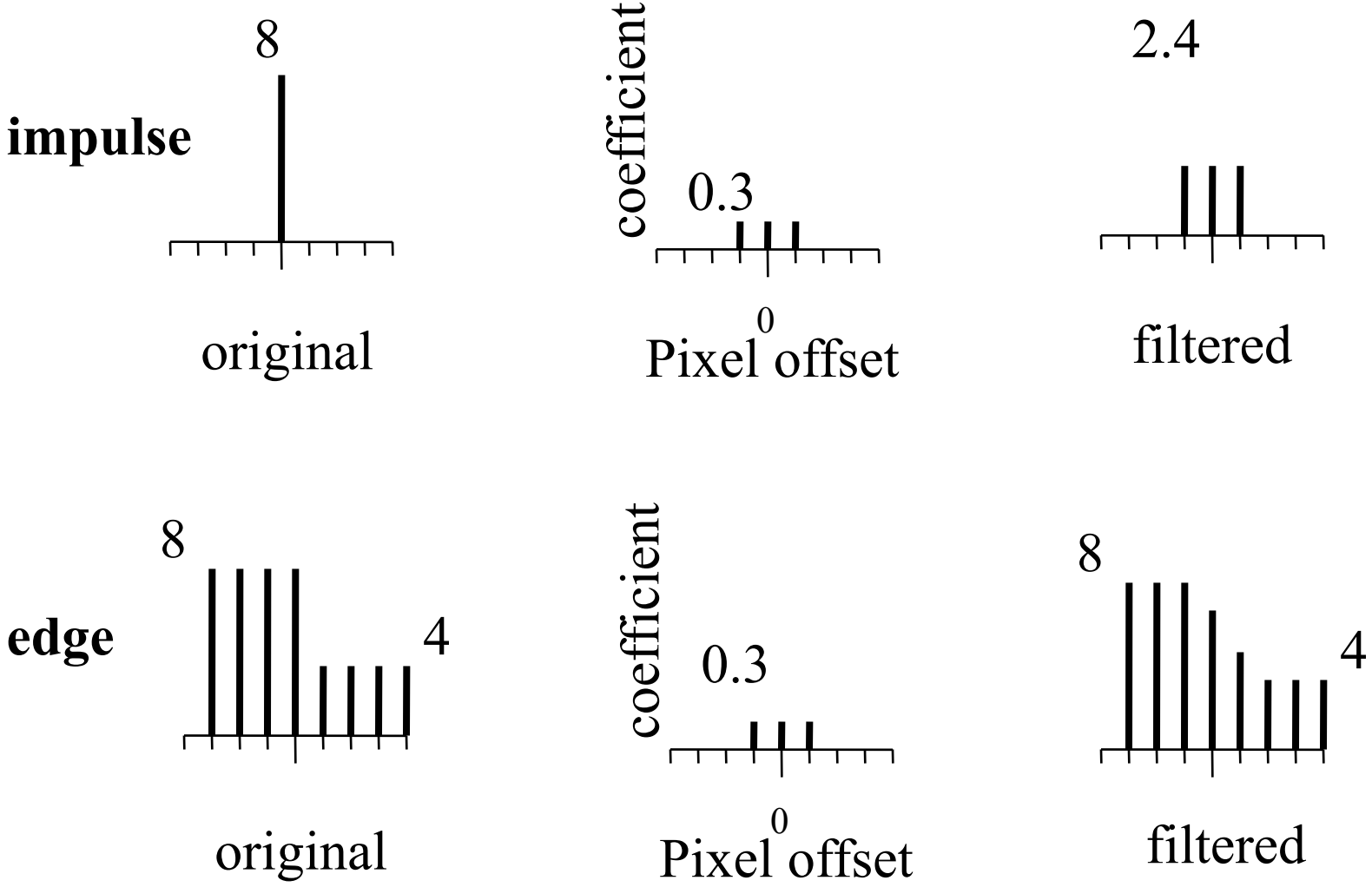


Blurred (filter applied in both dimensions).

Blur examples



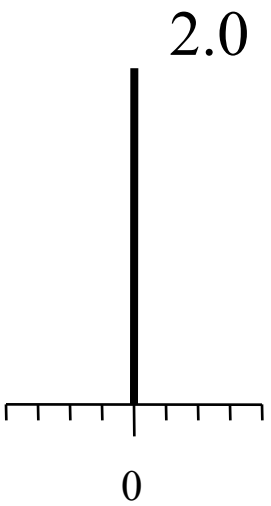
Blur examples

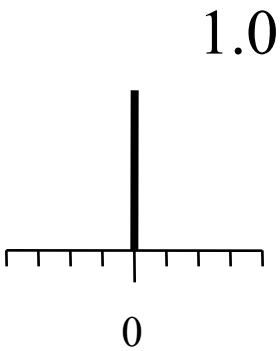


Questions?

Convolution (warm-up slide)

original

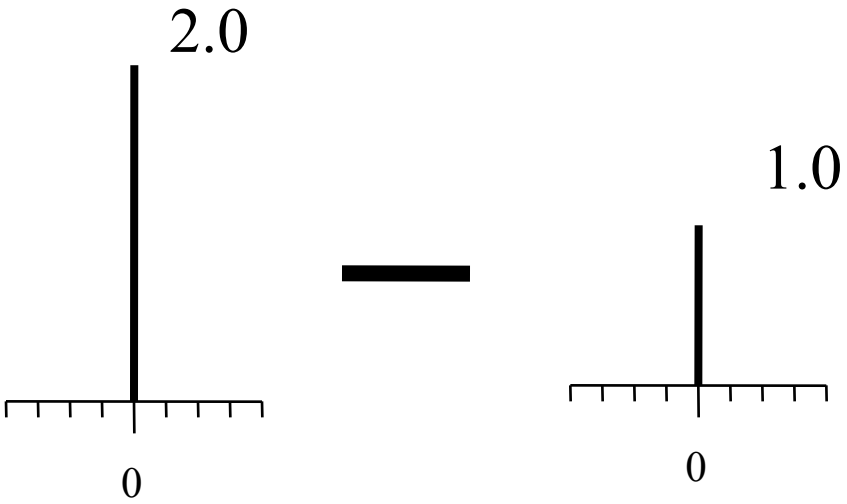


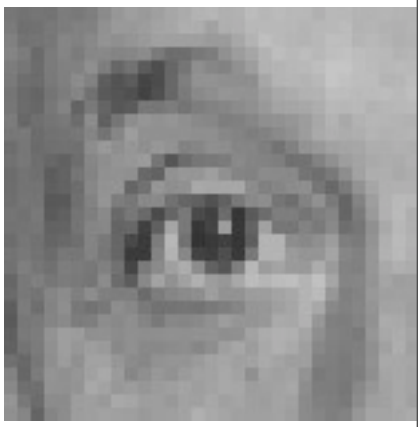


?

Convolution (no change)

original

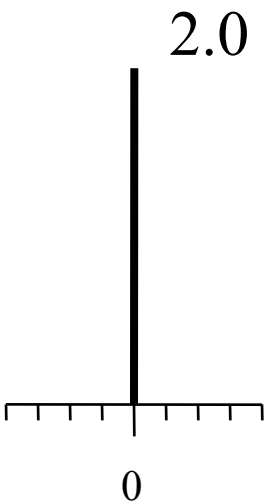


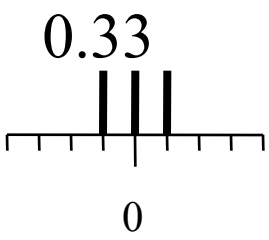


Filtered
(no change)

Convolution

original

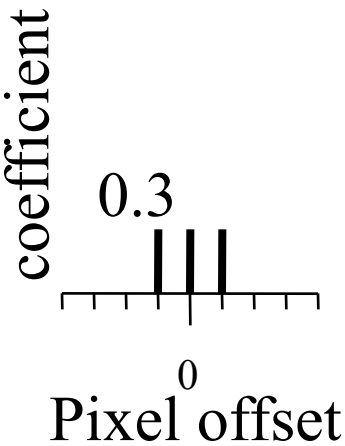




?

(remember blurring)

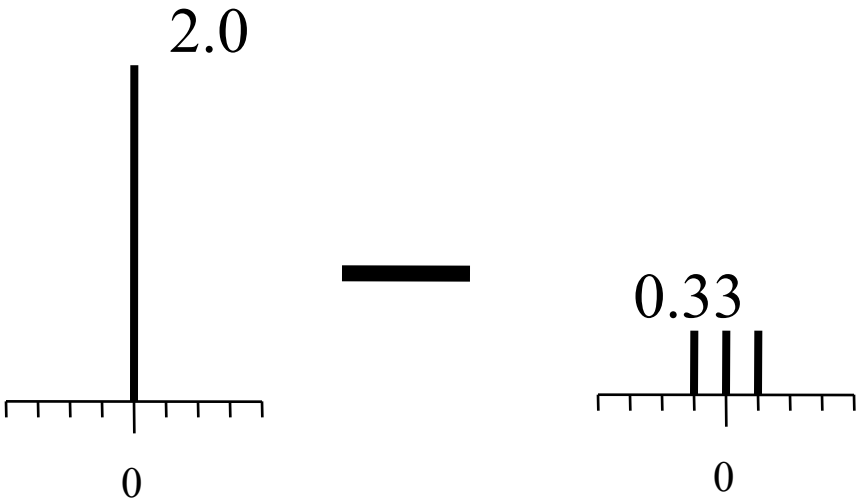
original



Blurred (filter applied in both dimensions).

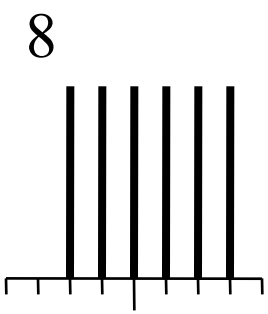
Sharpening

original

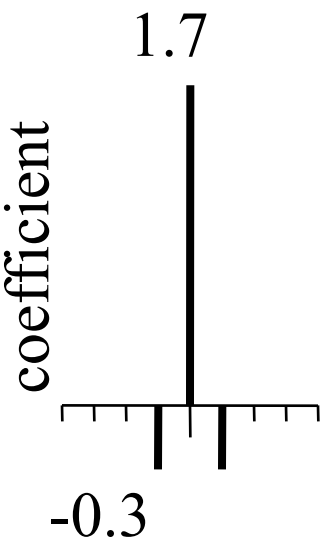


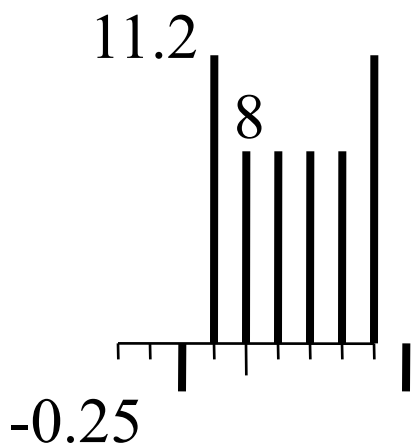
Sharpened original

Sharpening example



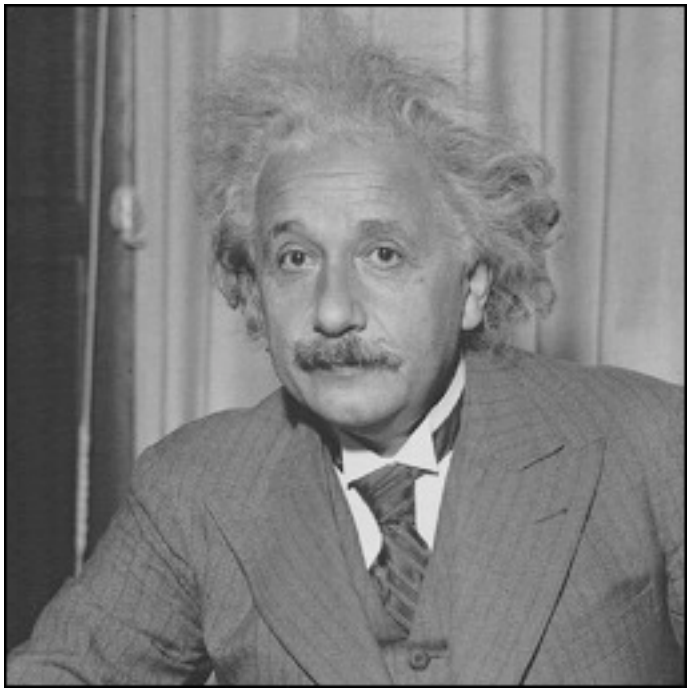
original



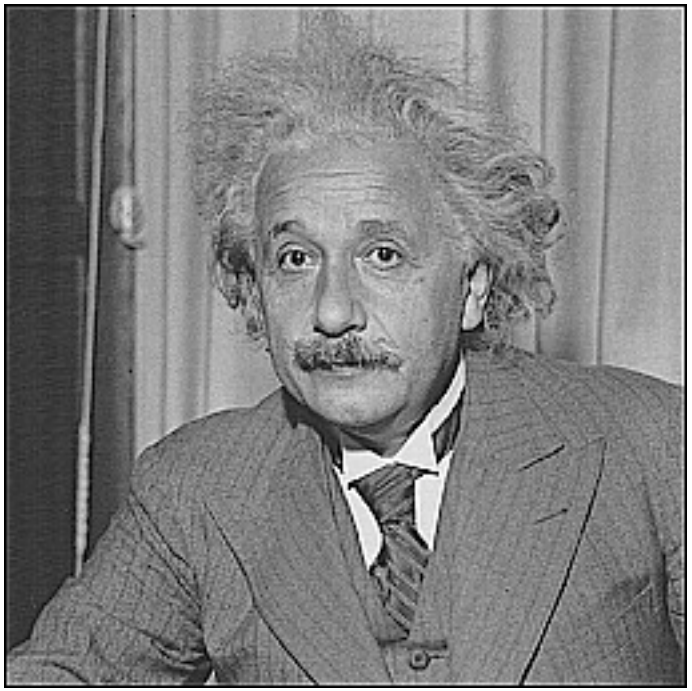


Sharpened (differences are accentuated; constant areas are left untouched).

Sharpening

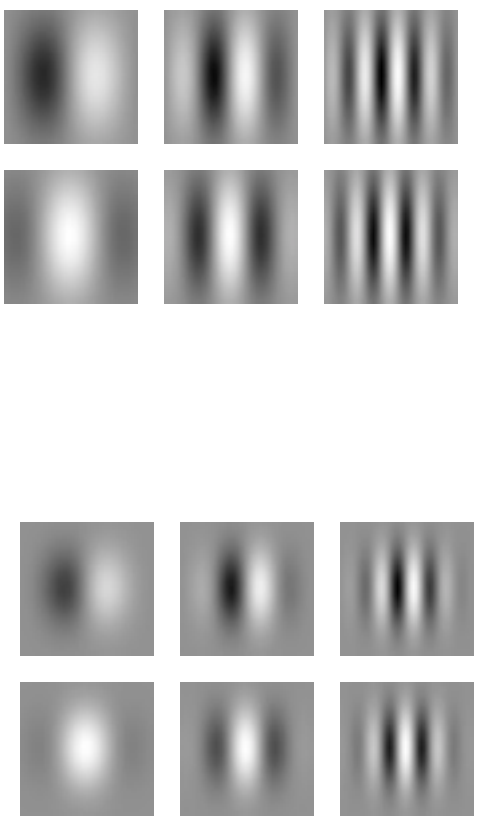


before



after

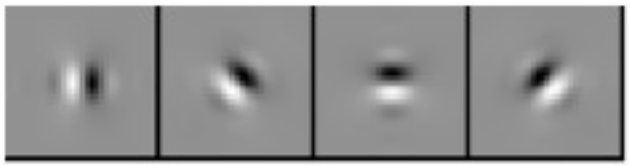
Oriented filters



Gabor filters at different scales and spatial frequencies

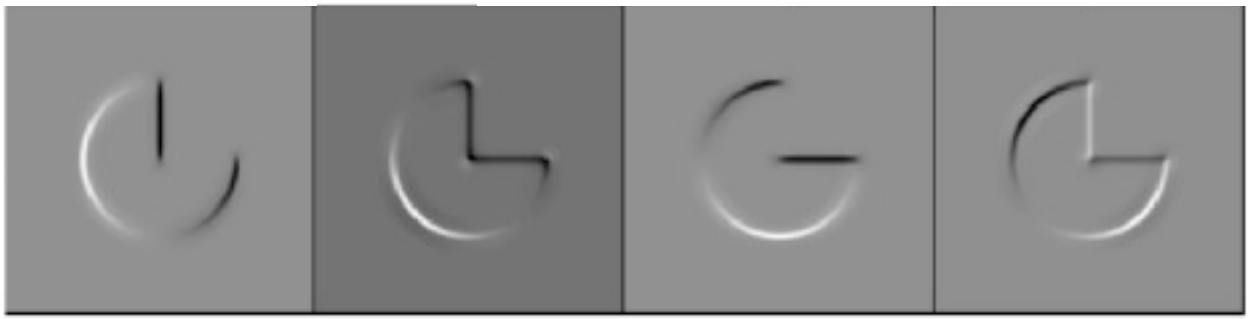
top row shows anti-symmetric (or odd) filters, bottom row the symmetric (or even) filters.

Filter Kernels



Image

Filtered images



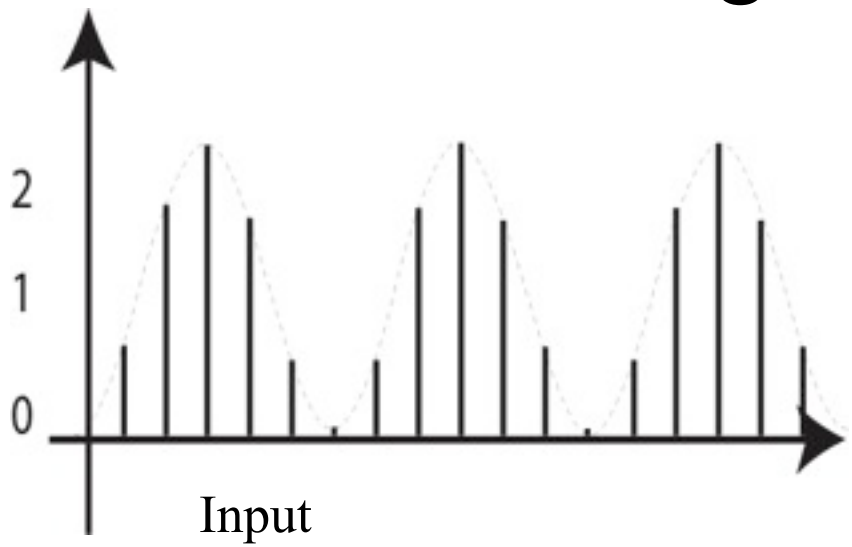
Reprinted from "Shiftable MultiScale Transforms," by Simoncelli et al., IEEE Transactions on Information Theory, 1992, copyright 1992, IEEE

Questions?

Studying convolutions

- Convolution is complicated
  - But at least it's linear
    $(f+kg)- h = f-h +k (g-h)$
- We want to find a better expression
  - Let's study functions whose behavior is simple under convolution

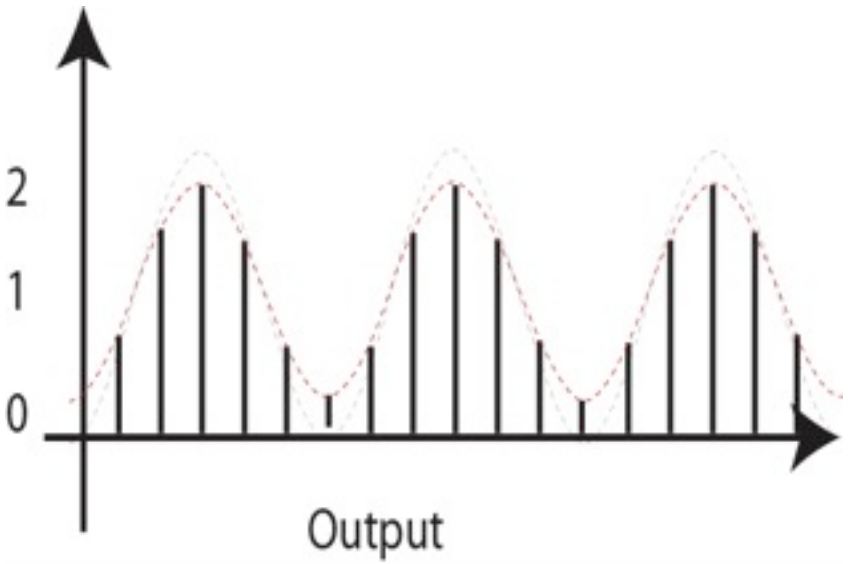
Blurring: convolution



Input

Convolution sign

Kernel



Output

Same shape, just reduced contrast!!!

This is an eigenvector
(output is the input multiplied by a constant)

Big Motivation for Fourier analysis

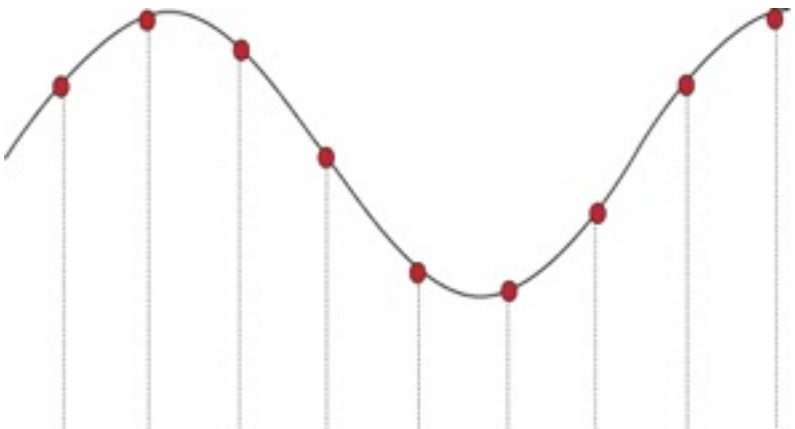
- Sine waves are eigenvectors of the convolution operator

Other motivation for Fourier analysis: sampling

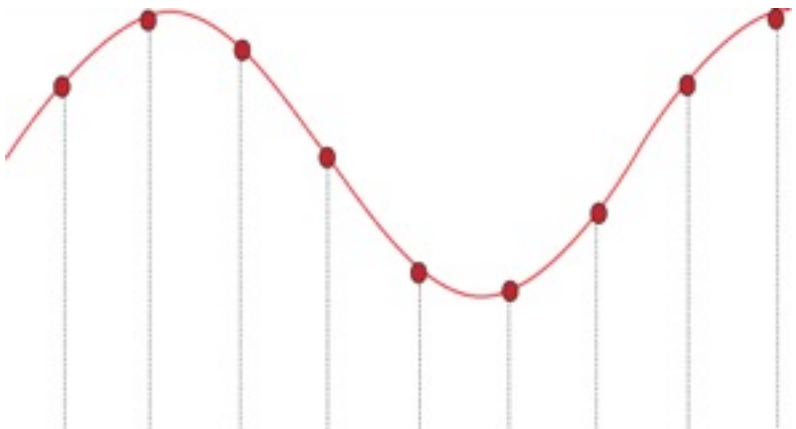
- The sampling grid is a periodic structure
  - Fourier is pretty good at handling that
  - We saw that a sine wave has serious problems with sampling
- Sampling is a linear process
  - but not shift-invariant

Sampling Density

- If we're lucky, sampling density is enough



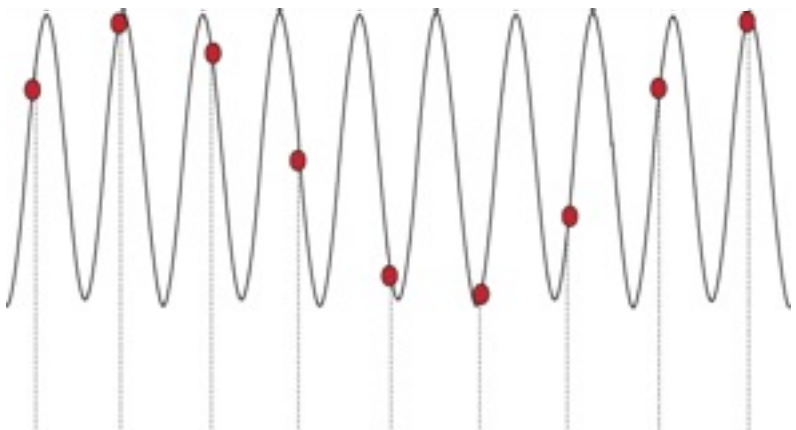
Input

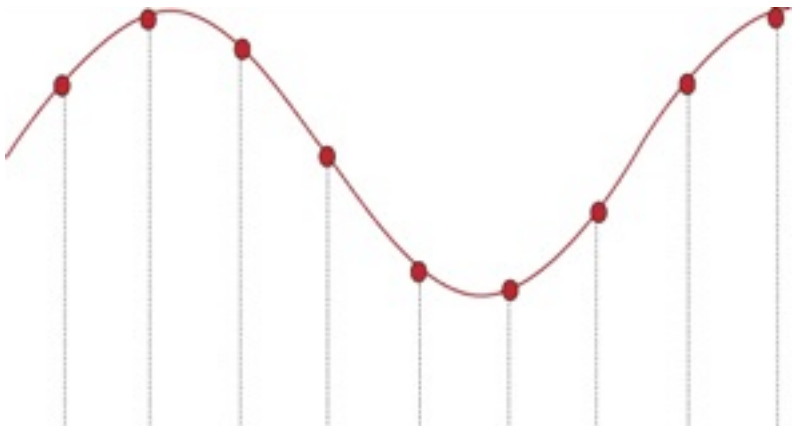


Reconstructed

Sampling Density

- If we insufficiently sample the signal, it may be mistaken for something simpler during reconstruction (that's aliasing!)





Recap: motivation for sine waves

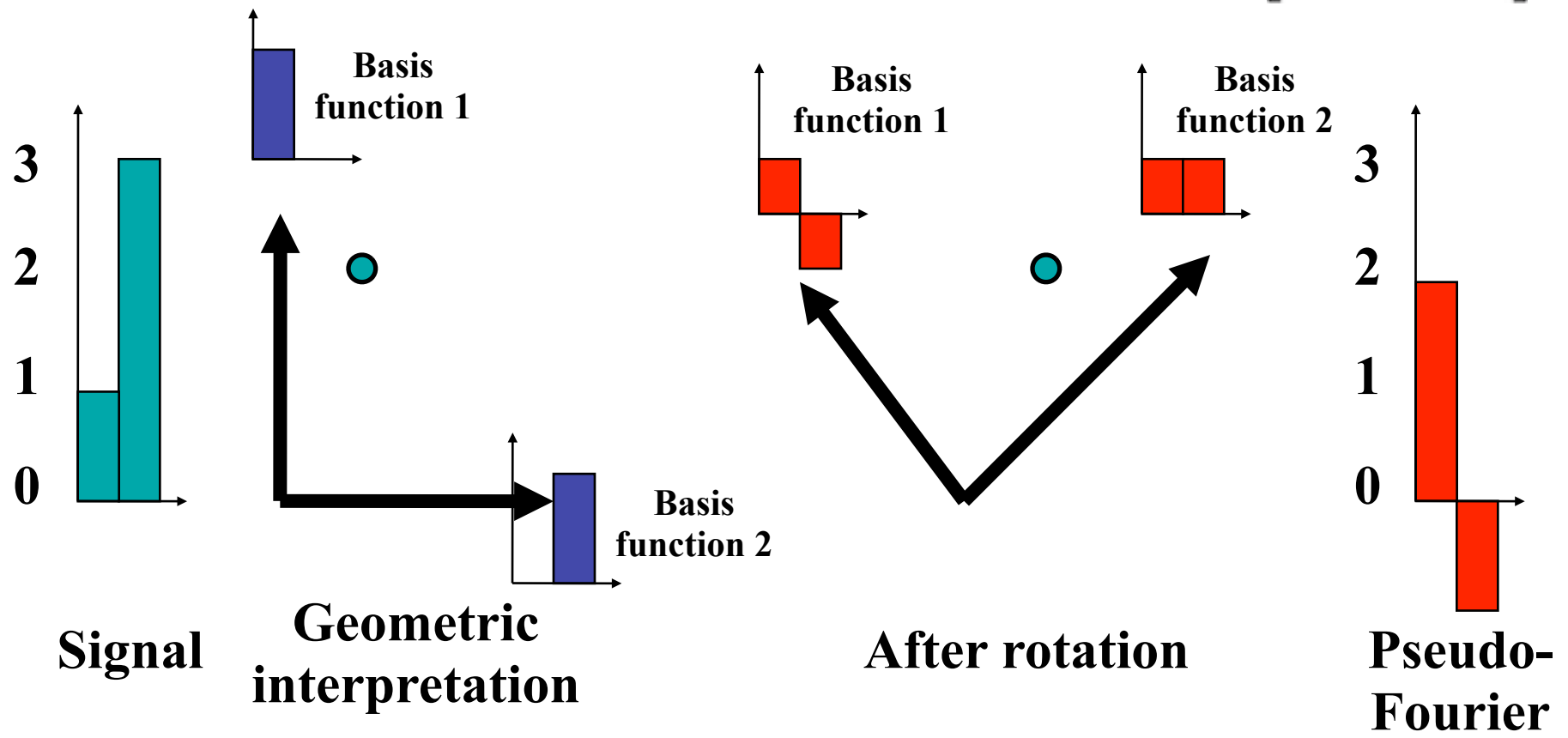
- Blurring sine waves is simple
  - You get the same sine wave, just scaled down
  - The sine functions are the eigenvectors of the convolution operator
- Sampling sine waves is interesting
  - Get another sine wave
  - Not necessarily the same one! (aliasing)

If we represent functions (or images) with a sum of sine waves, convolution and sampling are easy to study

Questions?

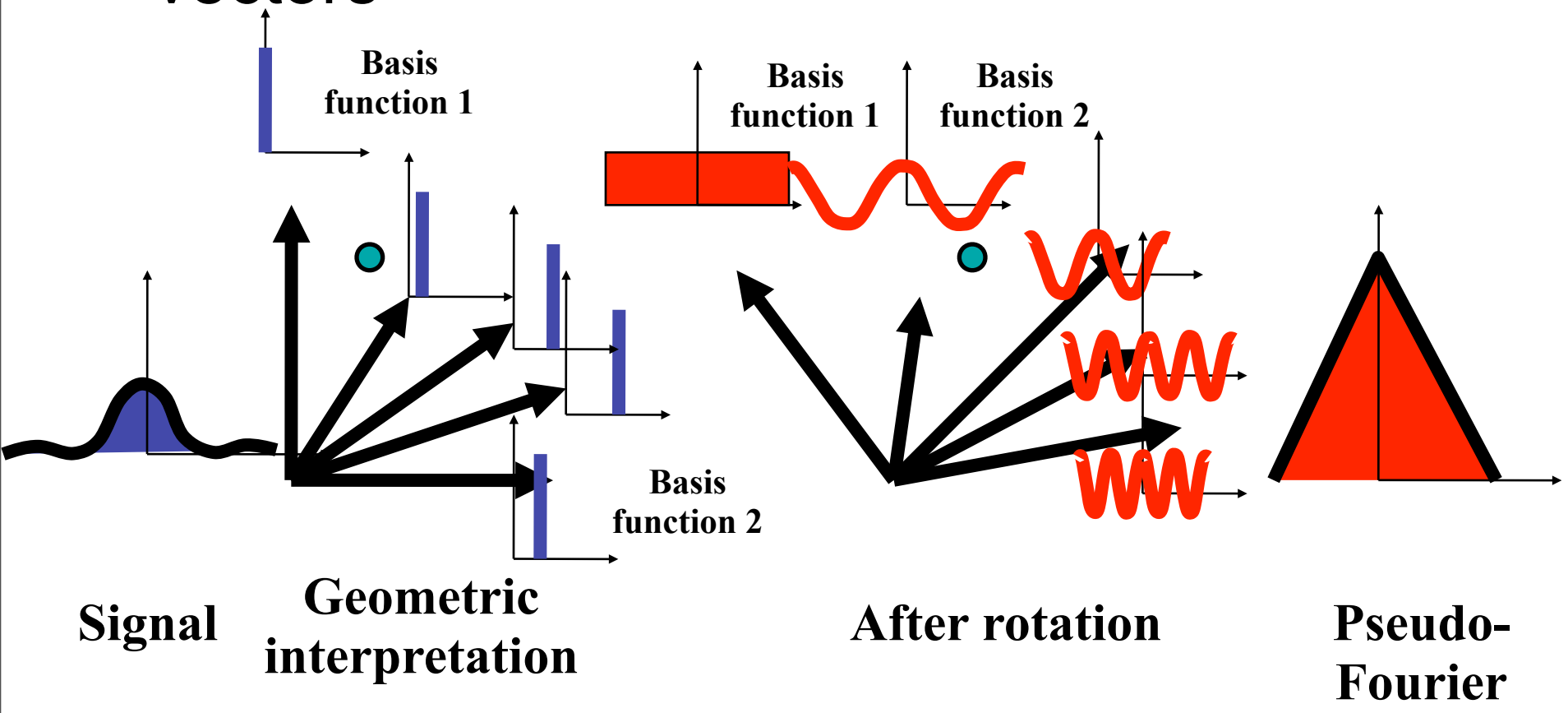
Fourier as change of basis

- Shuffle the data to reveal other information
- E.g., take average & difference: matrix $\begin{bmatrix} 0.5 & 1 \\ 0.5 & -1 \end{bmatrix}$



Fourier as change of basis

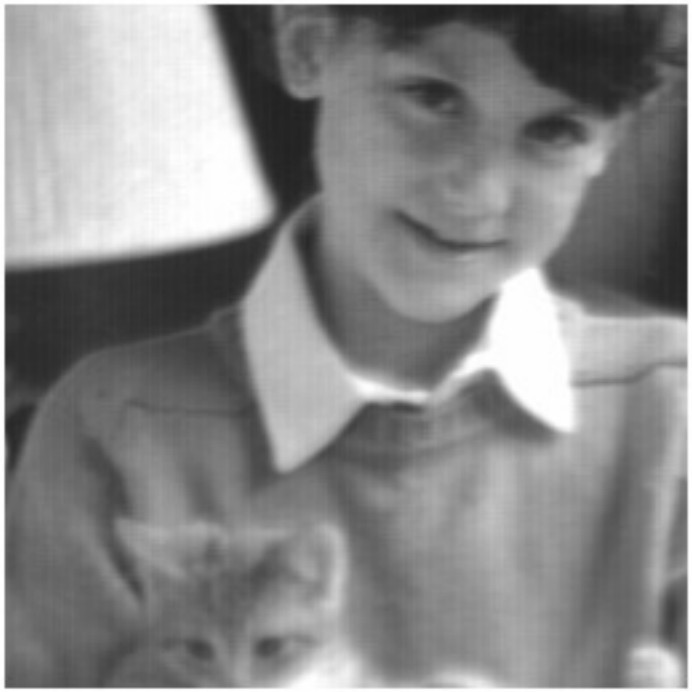
- Same thing with infinite-dimensional vectors

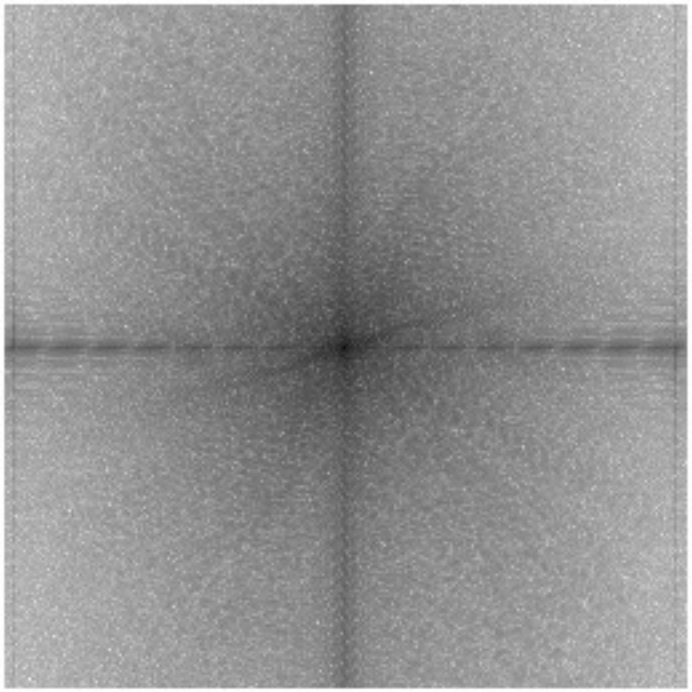


Question?

Fourier as a change of basis

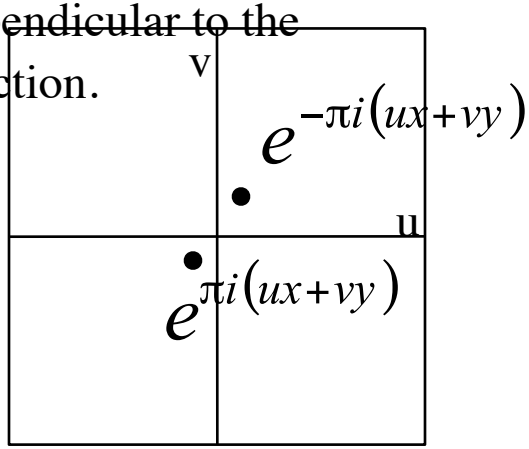
- Discrete Fourier Transform: just a big matrix

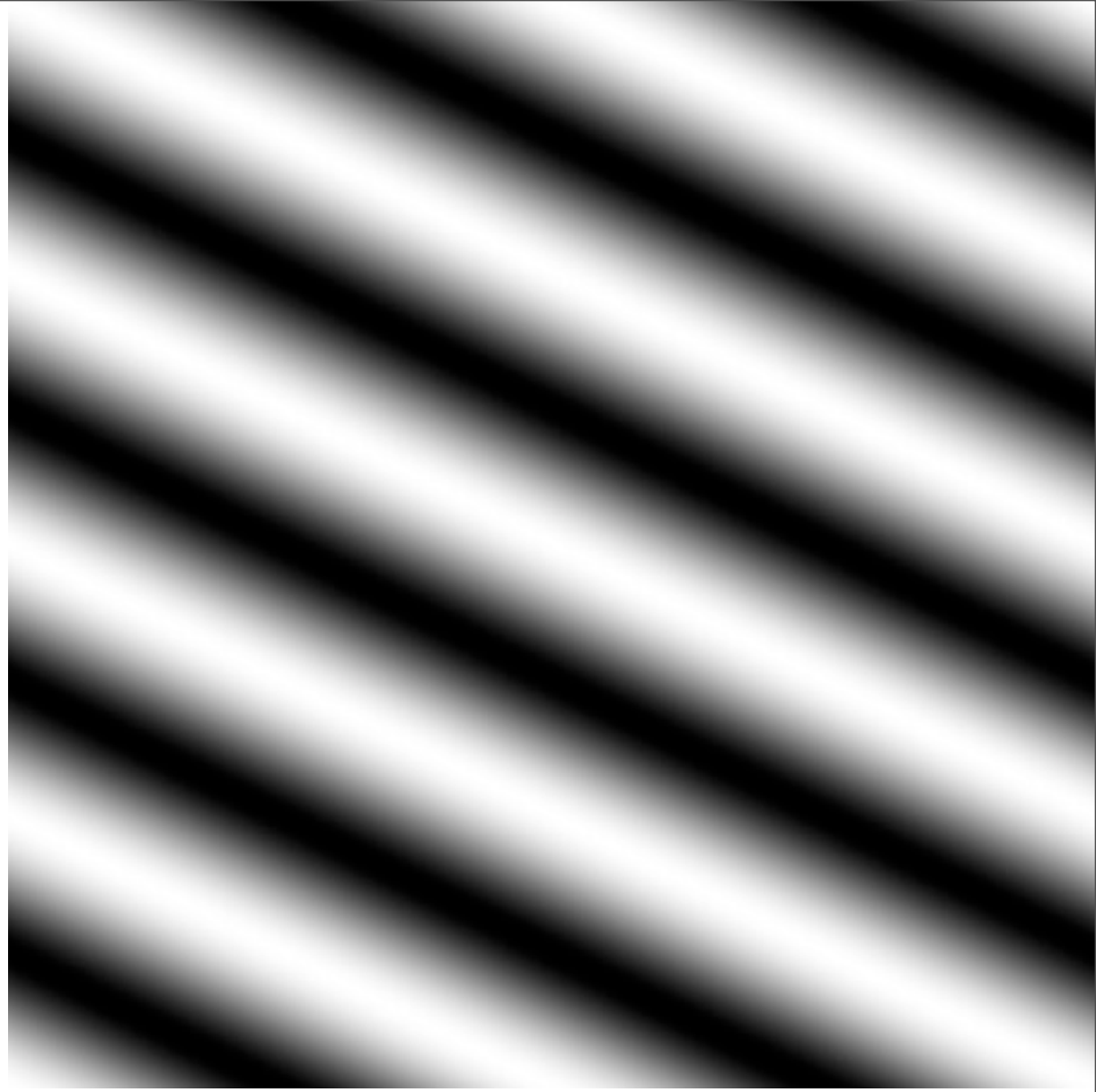




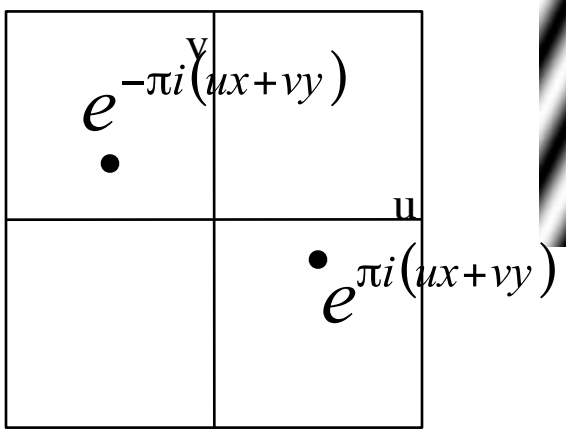
http://www.reindeergraphics.com

To get some sense of what basis elements look like, we plot a basis element --- or rather, its real part --- as a function of x,y for some fixed u, v. We get a function that is constant when (ux+vy) is constant. The magnitude of the vector (u, v) gives a frequency, and its direction gives an orientation. The function is a sinusoid with this frequency along the direction, and constant perpendicular to the direction.

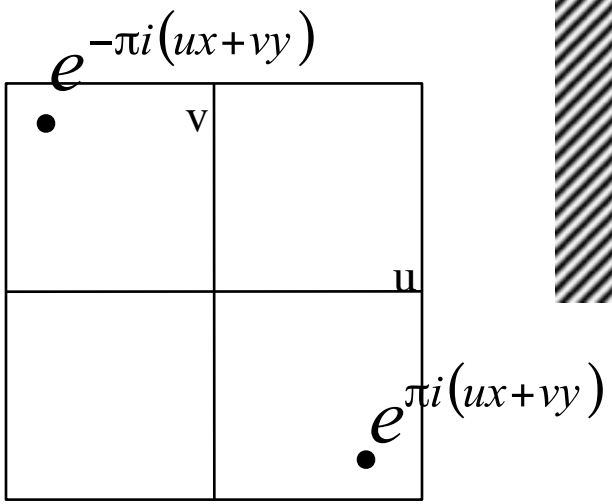




Here u and v are larger than in the previous slide.



And larger still...



Question?

Other presentations of Fourier

- Start with Fourier series with periodic signal
- Heat equation
  - more or less special case of convolution
  - iterate -> exponential on eignevalues

Motivations

- Insights & mathematical beauty
- Sampling rate and filtering bandwidth
- Computation bases
  - FFT: faster convolution
  - E.g. finite elements, fast filtering, heat equation, vibration modes
- Optics: wave nature of light & diffraction

Questions?

The Fourier Transform

- Defined for infinite, aperiodic signals
- Derived from the Fourier series by "extending the period of the signal to infinity"
- The Fourier transform is defined as

$$X(\omega) = \frac{1}{\sqrt{2\pi}} \int x(t) e^{-j\omega t} dt$$

- $X(\omega)$ is called the spectrum of $x(t)$
- It contains the magnitude and phase of each complex exponential of frequency ω in $x(t)$

The Fourier Transform

- The inverse Fourier transform is defined as

$$x(t) = \frac{1}{\sqrt{2\pi}} \int X(\omega) e^{j\omega t} d\omega$$

- Fourier transform pair

$$x(t) \Leftrightarrow X(\omega)$$

- $x(t)$ is called the *spatial domain* representation
- $X(\omega)$ is called the *frequency domain* representation

Beware of differences

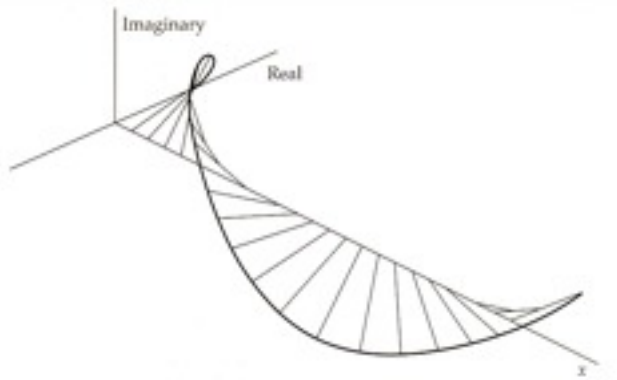
- Different definitions of Fourier transform
- We use

$$X(\omega) = \frac{1}{\sqrt{2\pi}} \int x(t) e^{-j\omega t} dt$$

- Other people might exclude normalization or include 2π in the frequency
- X might take ω or $j\omega$ as argument
- Physicist use *j*, mathematicians use *i*

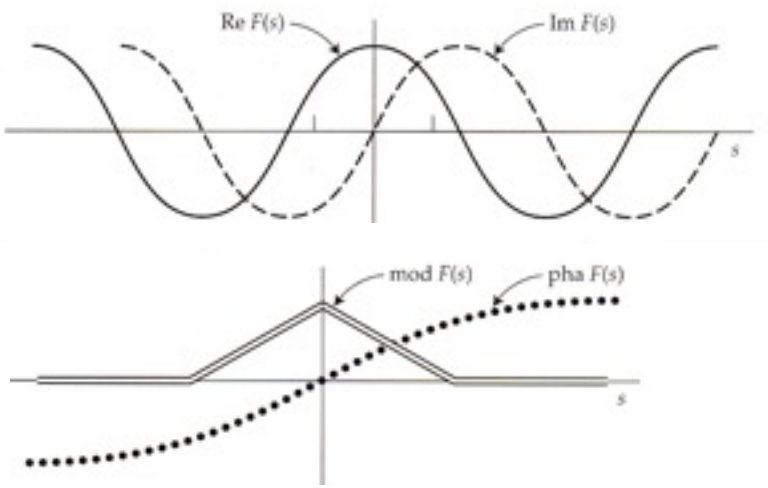
Phase

- Don’t forget the phase! Fourier transform results in complex numbers

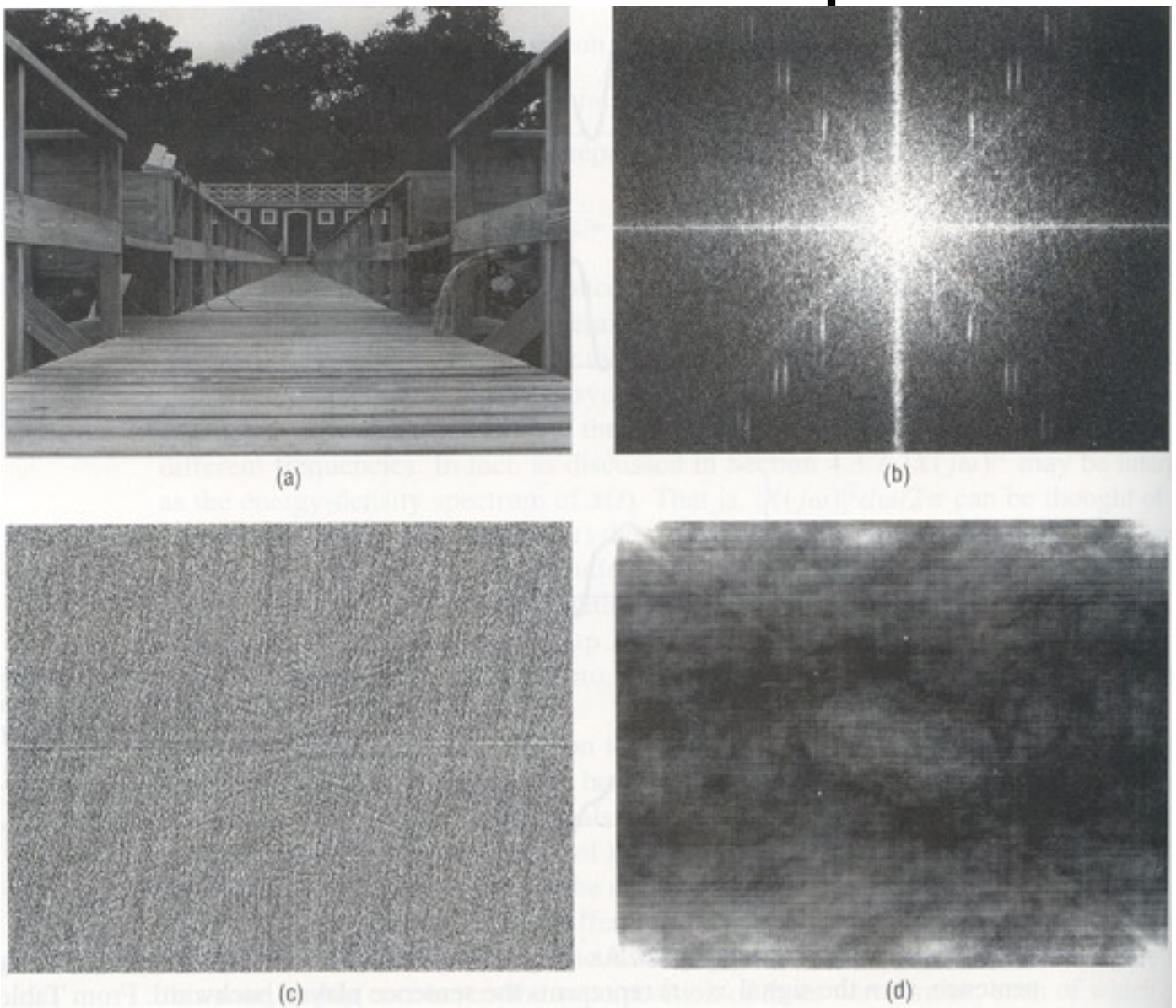


- Can be seen as sum of sines and cosines

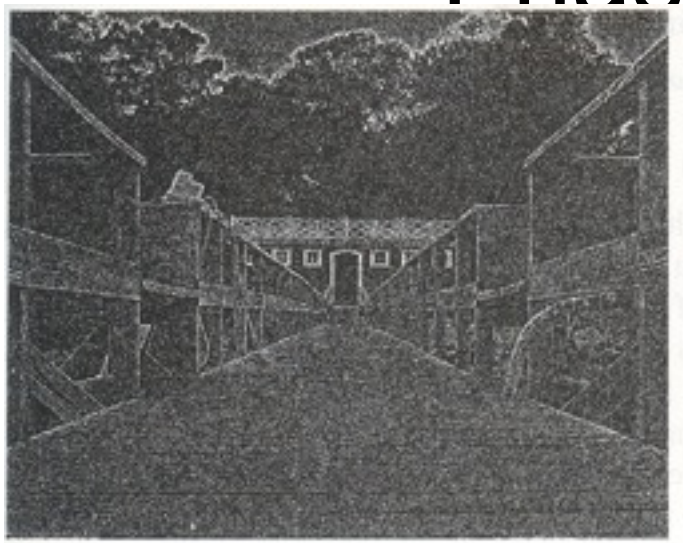
- Or modulus/phase



Phase is important!



Phase is important!



(e)

(f)

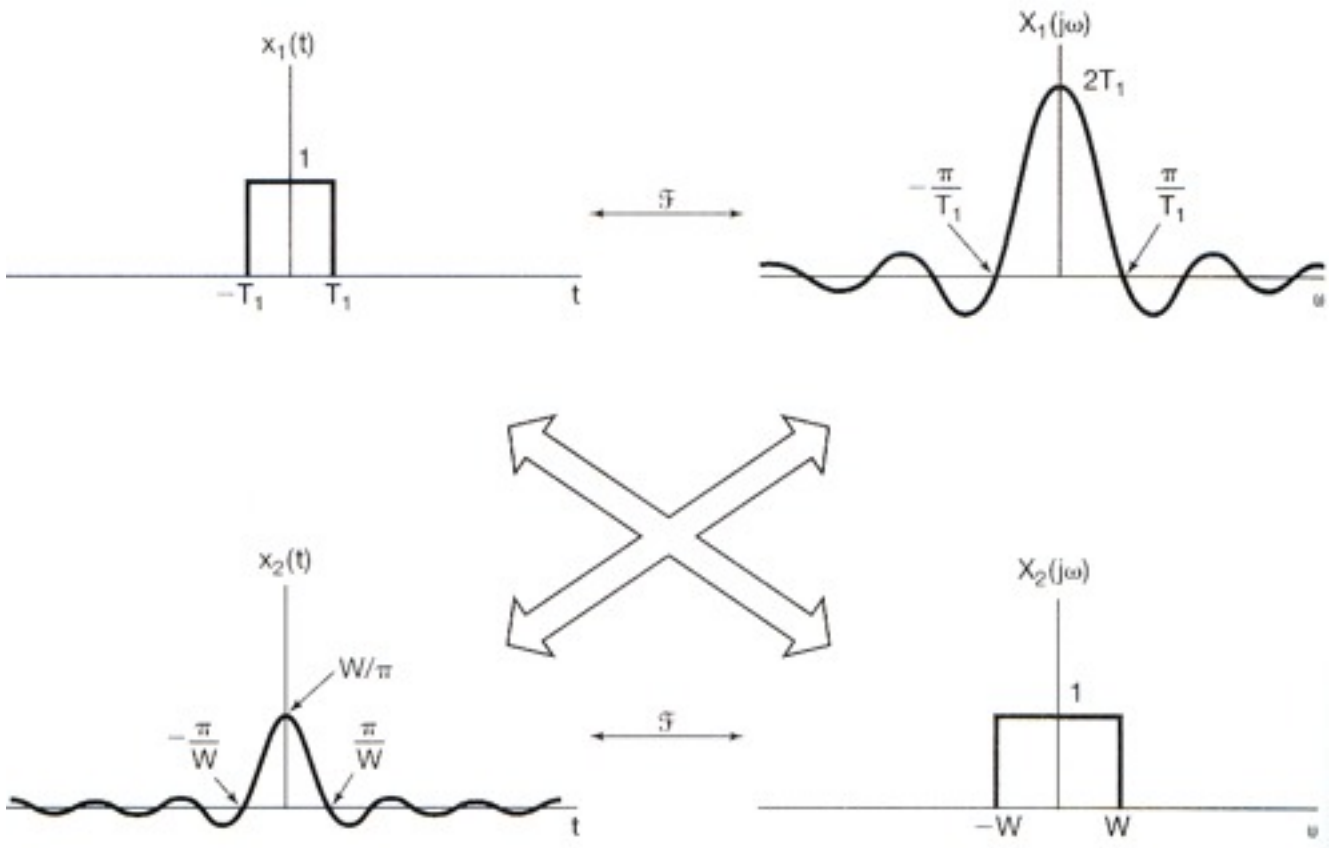
(g)

Figure 6.2 (a) The image shown in Figure 1.4; (b) magnitude of the two-dimensional Fourier transform of (a); (c) phase of the Fourier transform of (a); (d) picture whose Fourier transform has magnitude as in (b) and phase equal to zero; (e) picture whose Fourier transform has magnitude equal to 1 and phase as in (c); (f) picture whose Fourier transform has phase as in (c) and magnitude equal to that of the transform of the picture shown in (g).

Questions?

Duality

Up to details (such as factors of 2π or signs):
if function a is the Fourier transform of b, then b is the Fourier transform of a
For example, the Fourier transform of a box is a sinc,
and the Fourier transform of a sinc is a box.



Duality

Any theorem that involves the primal and Fourier domains is also true when swapping the two domains.

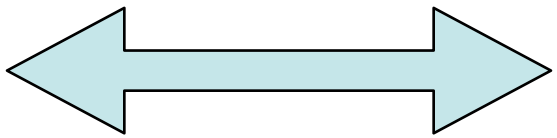
e.g. shift theorem:

Primal		Fourier
$f(x+a)$	$\Leftrightarrow$	$e^{-2\pi i a \omega} F(\omega)$
$e^{-2\pi i a x} f(x)$	$\Leftrightarrow$	$F(\omega+a)$

Duality

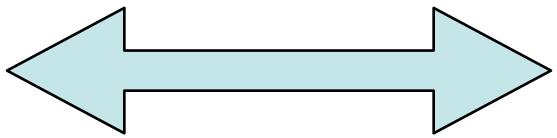
Any theorem that involves the primal and Fourier domains is also true when swapping the two domains.

e.g. scaling theorem:

Primal		Fourier
$f(ax)$		$1/a\ F(x/a)$
$1/a\ f(x/a)$		$F(\omega a)$

Convolution/Modulation

A convolution in one domain is a multiplication in the other one

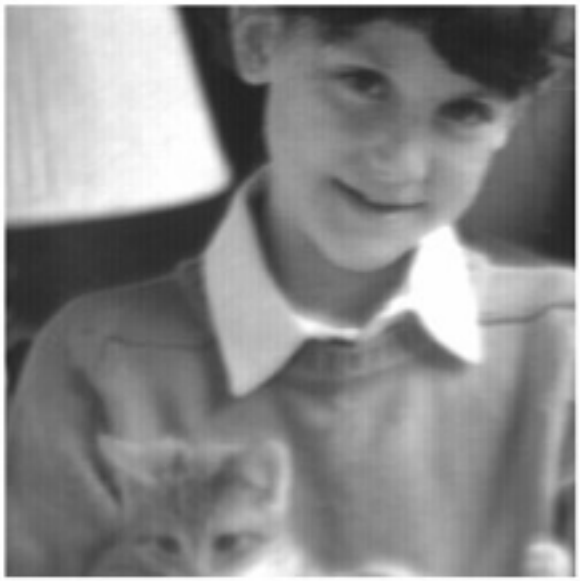
Primal		Fourier
$f \otimes g$		FG
fg		$F \otimes G$

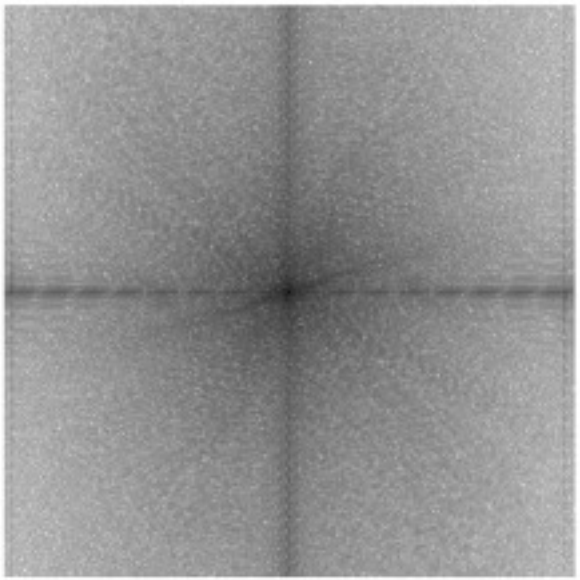
Recall that Fourier bases are eigenvectors of the convolution

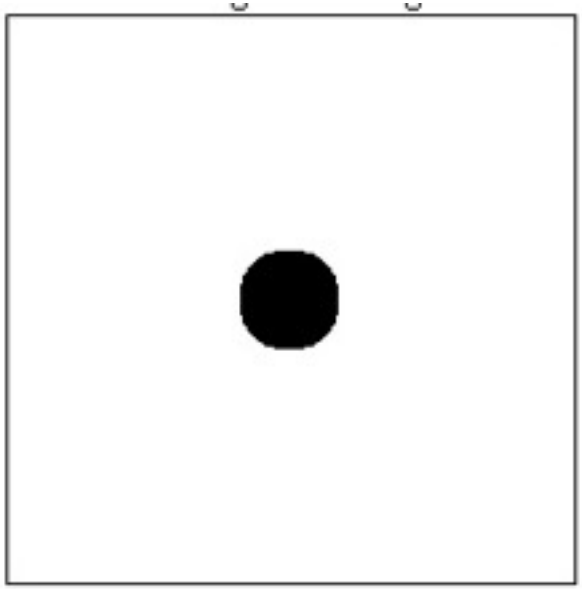
Questions?

Low pass

http://www.reindeergraphics.com



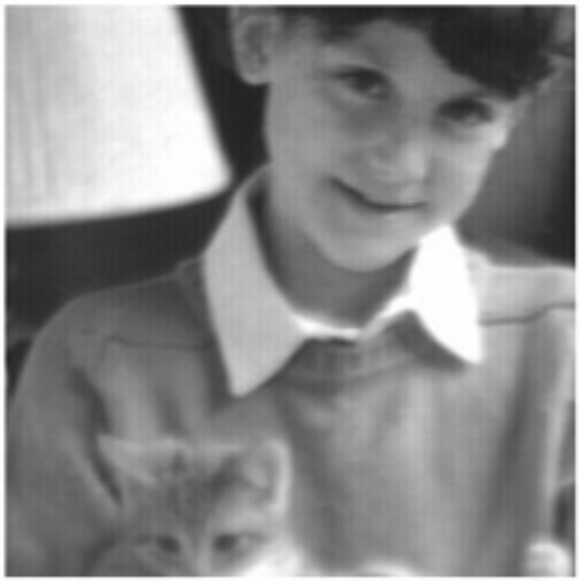


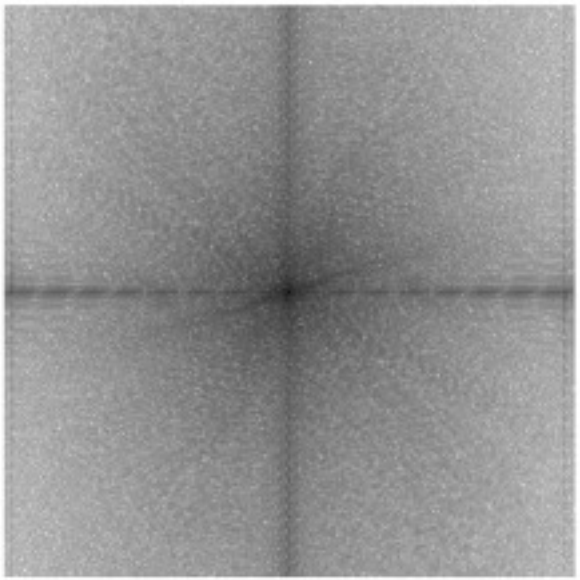


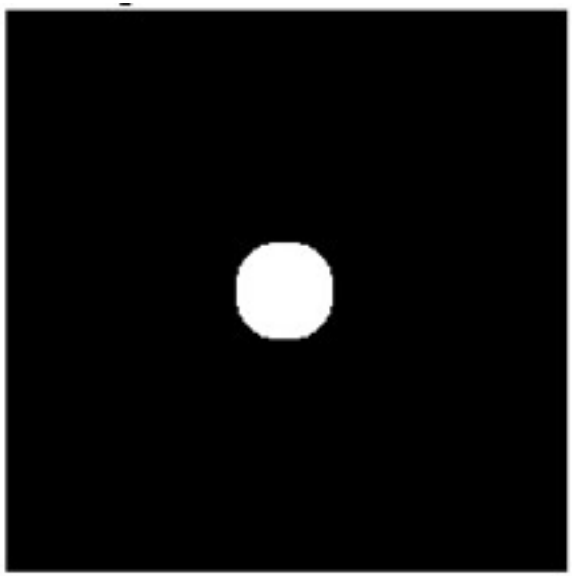
black means 1,
white means 0

High pass

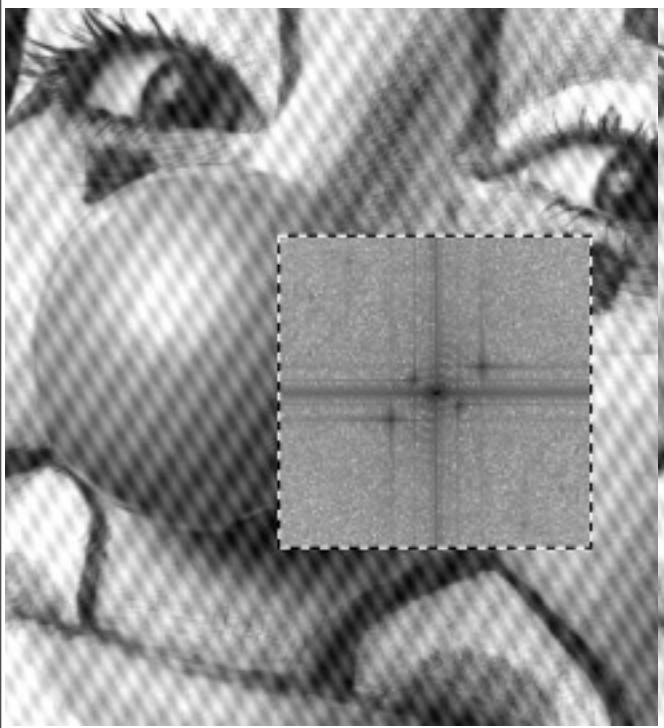
http://www.reindeergraphics.com

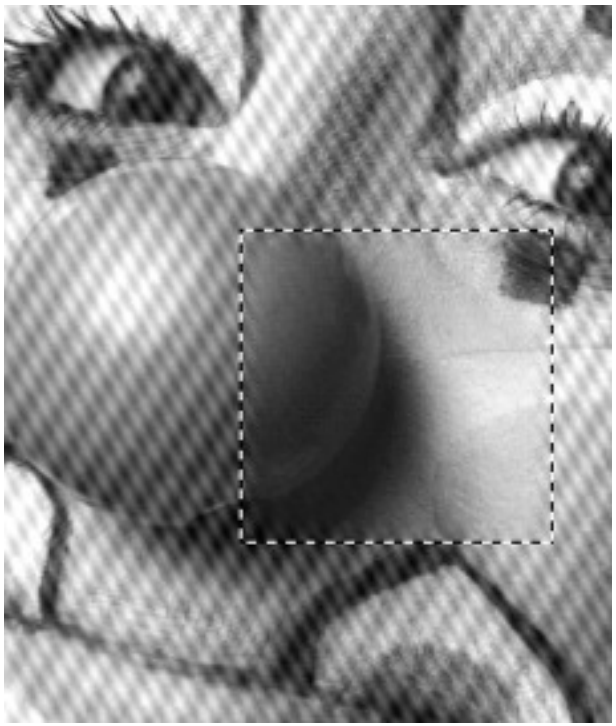






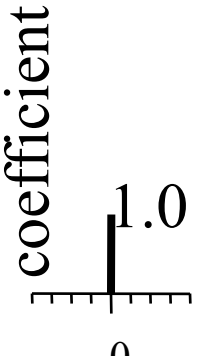
Filtering in Fourier domain





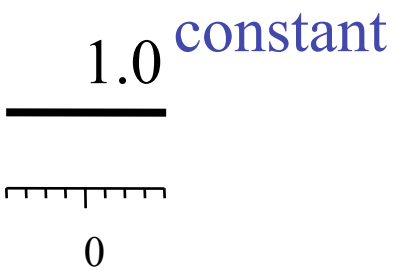
Analysis of our simple filters

original



Filtered
(no change)

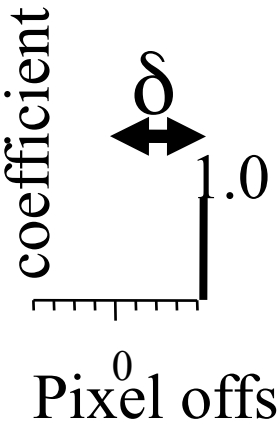
spectrum: F(ω)=1

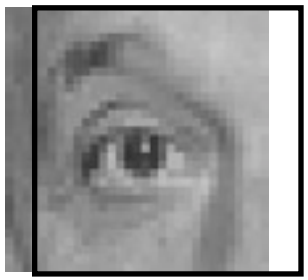


(yes, I am now using the definition without 1/sqrt(2pi)

Analysis of our simple filters

original

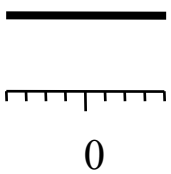




shifted

spectrum:

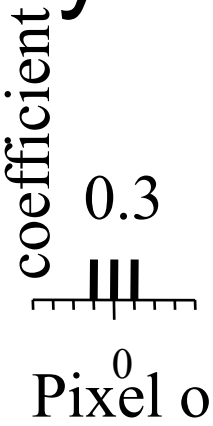
$$F(\omega) = e^{-2\pi j \omega \delta}$$



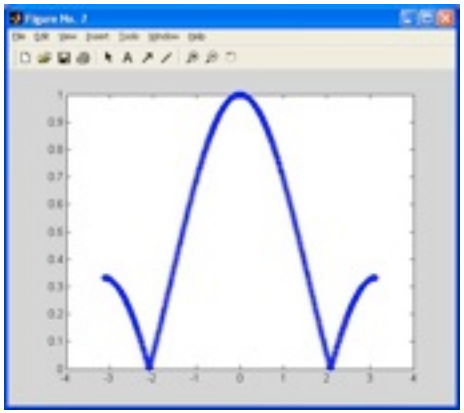
Constant magnitude, linearly shifted phase

Analysis of our simple filters

original

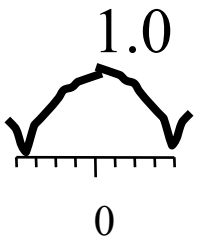


blurred



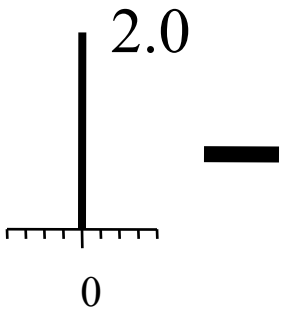
Low-pass filter

spectrum:
$F(\omega)=\mathrm{sinc}(\omega)$
$=\sin(\omega)/\omega$

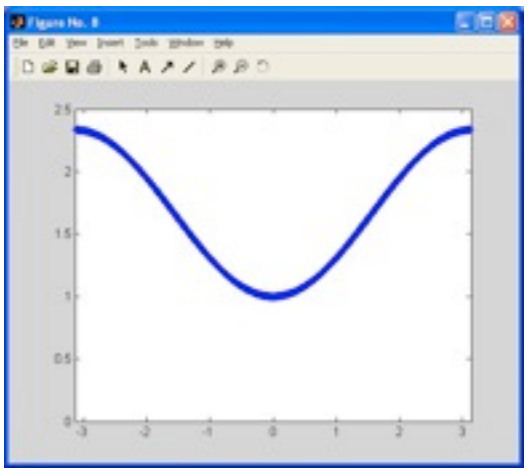


Analysis of our simple filters

original

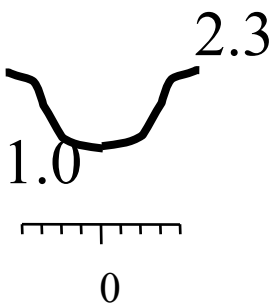


sharpened



high-pass filter

spectrum:
$F(\omega)=2-\text{sinc}(\omega)$



Convolution versus FFT

- 1-d FFT: O(NlogN) computation time, where N is number of samples.
- 2-d FFT: 2N(NlogN), where N is number of pixels on a side
- Convolution: K N^2, where K is number of samples in kernel
- Say N=2^{10}, K=100. 2-d FFT: 20 2^{20}, while convolution gives 100 2^{20}

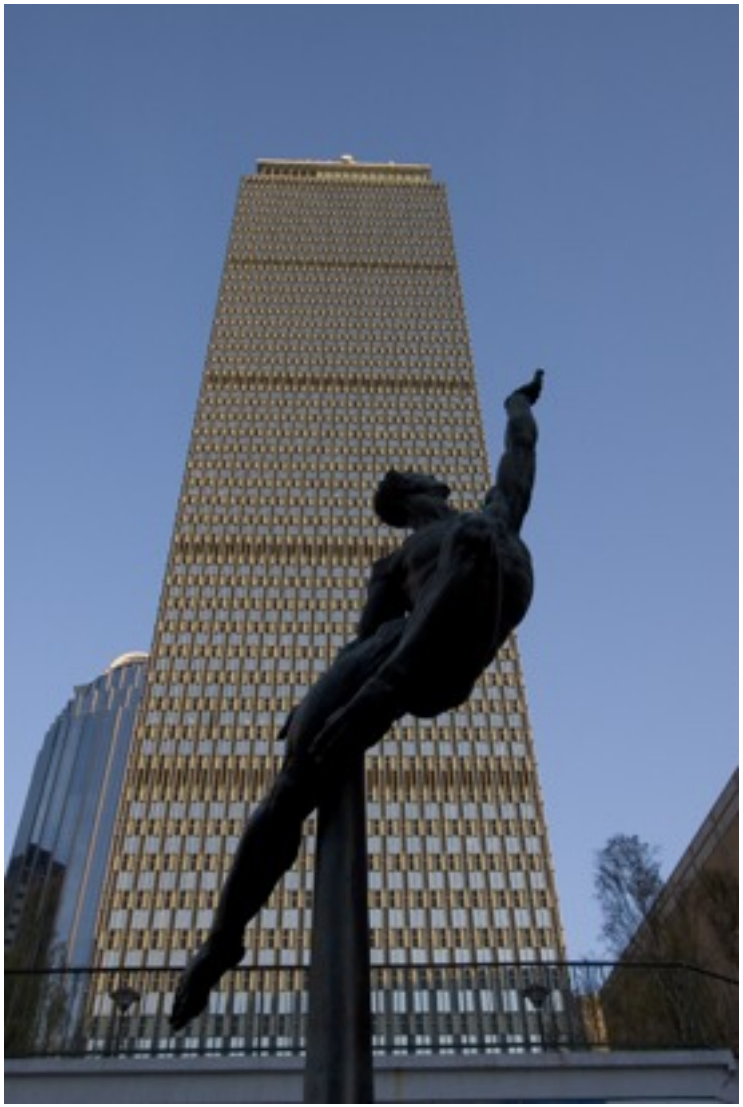
Words of wisdom

- **Careful with the FFT:**
  it assumes a cyclic signal
- Oftentimes, the answer you get mostly shows wraparound artifacts
- Proper windowing might be needed to analyze the frequency content of an image
  - e.g. multiply function by a smooth function that falls off away from the center so that the boundary is zero

Questions?

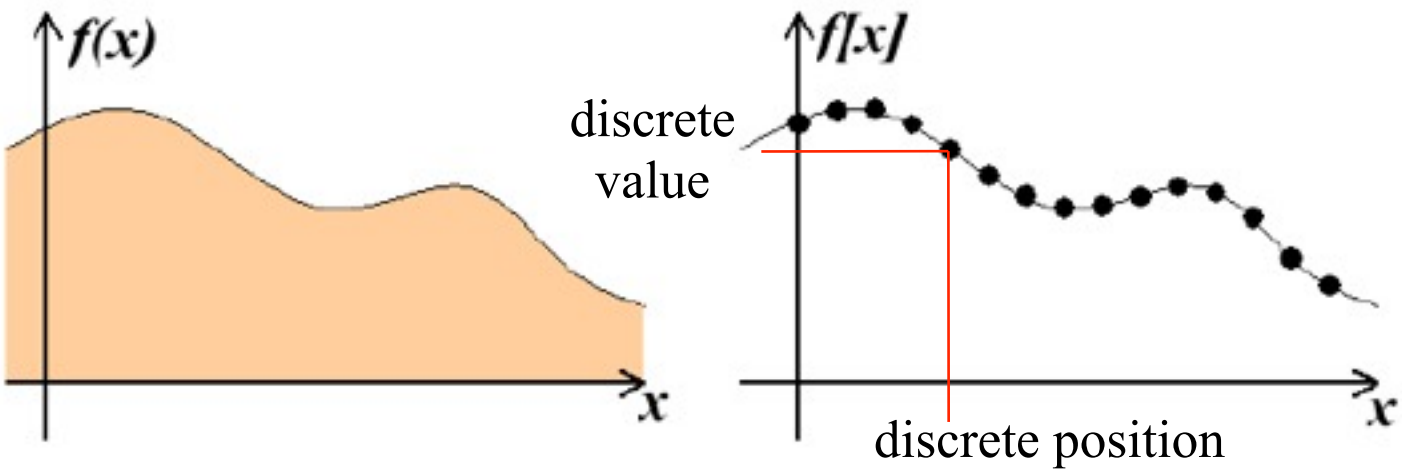
Sampling and aliasing

In photos too

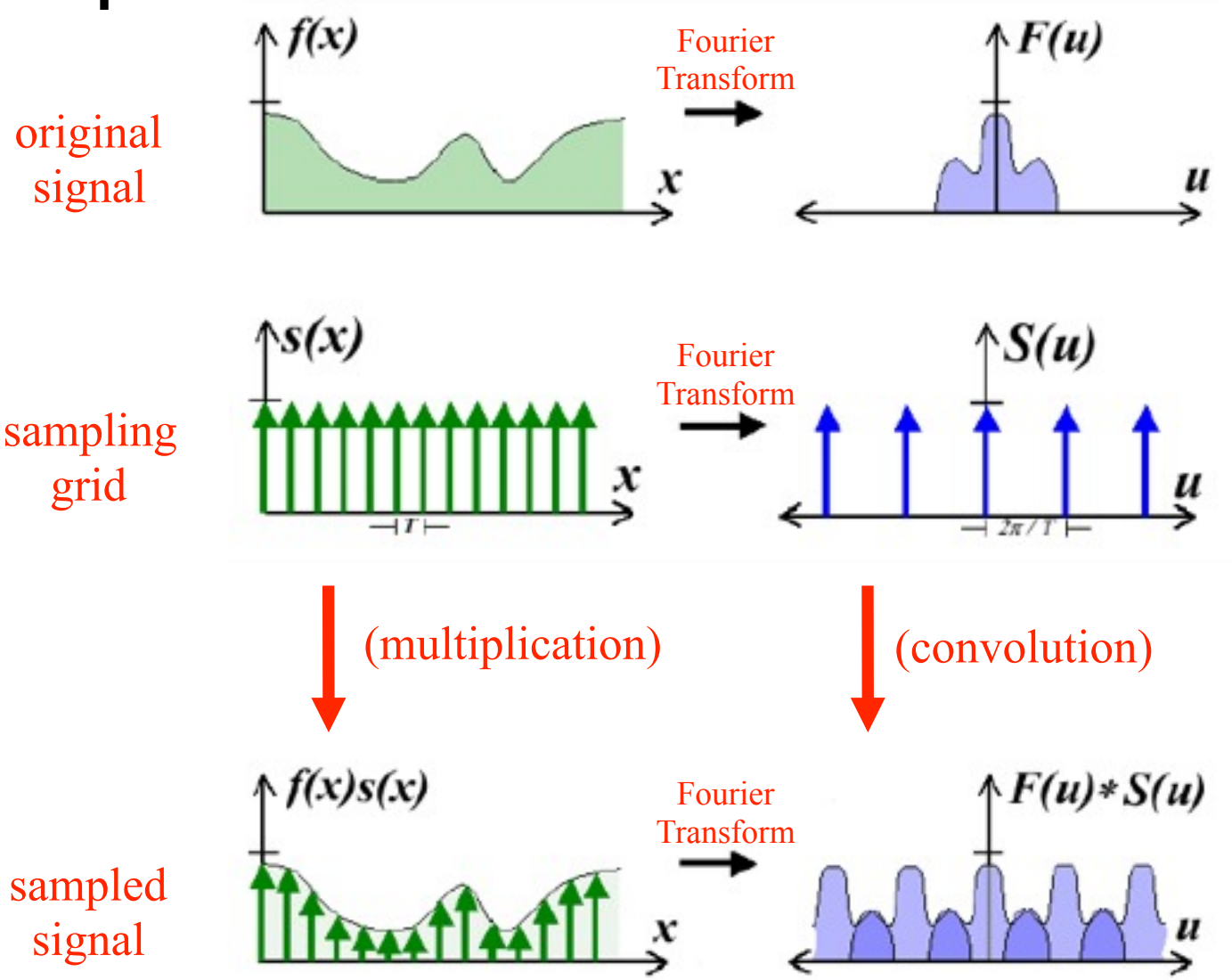


More on Samples

- In signal processing, the process of mapping a continuous function to a discrete one is called *sampling*
- The process of mapping a continuous variable to a discrete one is called *quantization*
- To represent or render an image using a computer, we must both sample and quantize
  - Now we focus on the effects of sampling and how to fight them



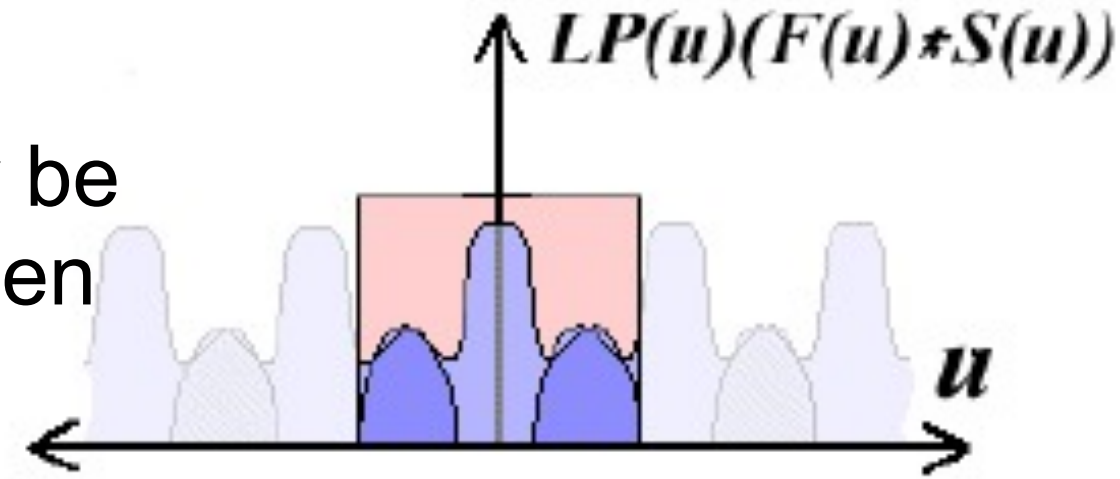
Sampling in the Frequency Domain



Reconstruction

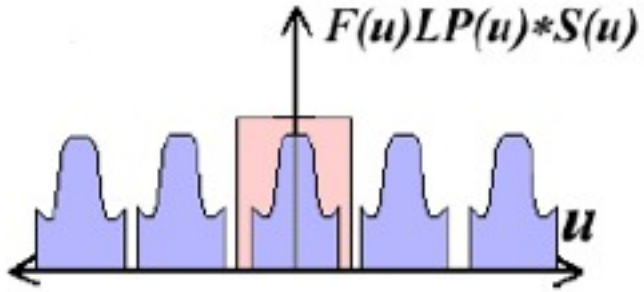
- If we can extract a copy of the original signal from the frequency domain of the sampled signal, we can reconstruct the original signal!

- But there may be overlap between the copies.

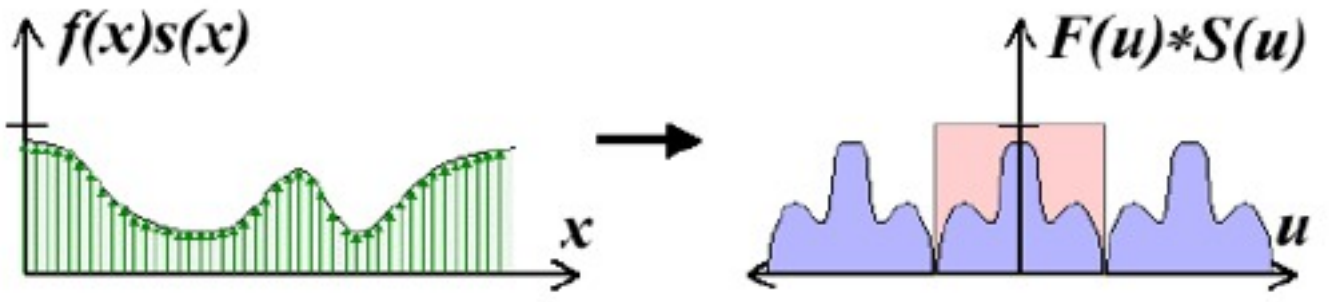


Guaranteeing Proper Reconstruction

- Separate by removing high frequencies from the original signal (low pass pre-filtering)



- Separat



- If we can't separate the copies, we will have overlapping frequency spectrum during reconstruction → *aliasing*.

Sampling Theorem

- When sampling a signal at discrete intervals, the sampling frequency must be *greater than twice* the highest frequency of the input signal in order to be able to reconstruct the original perfectly from the sampled version (Shannon, Nyquist, Whittaker, Kotelnikov, Küpfmüller)

Final project brainstorming

Fredo Durand
MIT EECS 6.815/6.865

Final project

- Groups of 1 or 2
- Proposal due soon (with last pset)
- Deliverables: report + small presentation

Your ideas?

Some ideas

- Use CHDK to provide new features to Canon compact cameras
- Use flickr API to do something creative
- Explore different types of gradient reconstructions
- Improve time lapse
- Handle small parallax in panoramas
- Exploit flash/no-flash pairs
- Editing with images+depth (e.g. from stereo)
- Smart color to greyscale
- Face-aware image processing
- Sharpening out-of-focus images using other pictures from the sequences
- Application of morphing/warping
- Motion without movements and automatic illusions