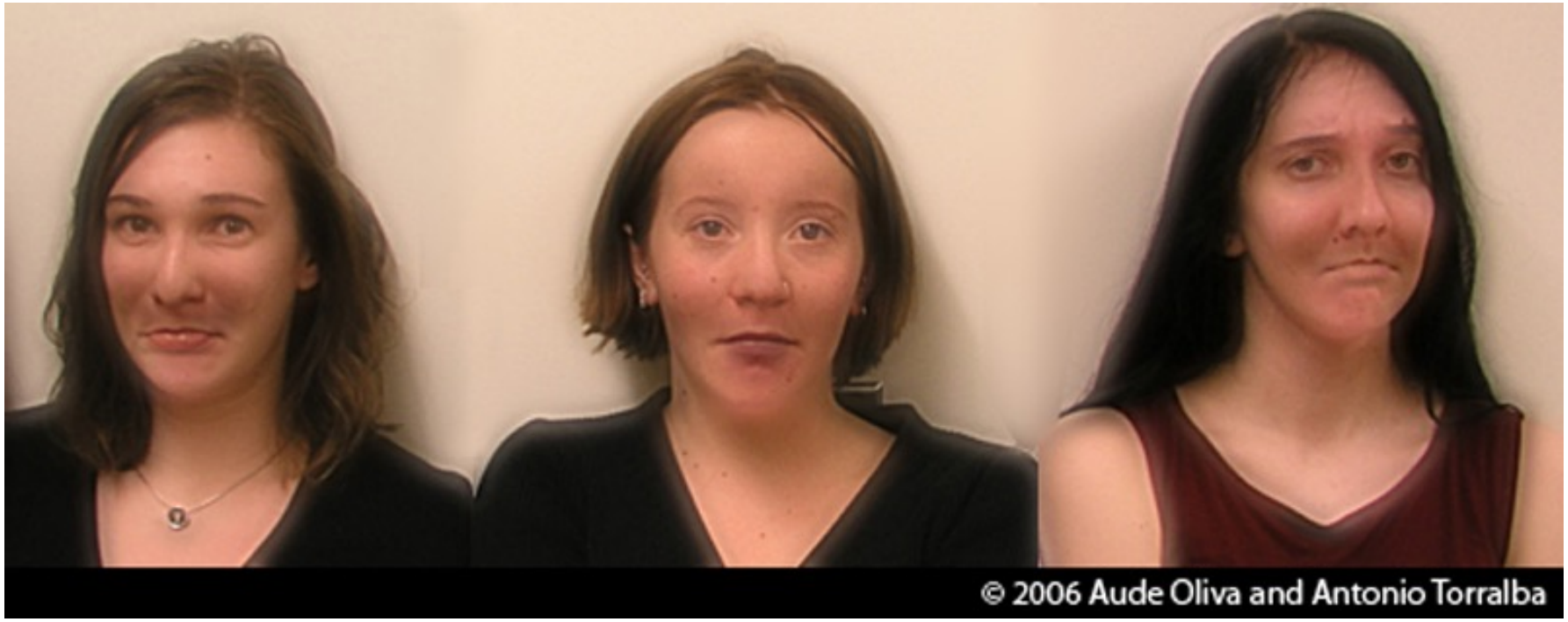
CS5670: Intro to Computer Vision

Noah Snavely

Lecture 1: Images and image filtering

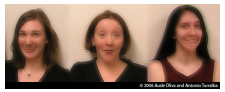


Hybrid Images, Oliva et al., http://olivalab.mit.edu/hybridimage.htm

CS5670: Intro to Computer Vision

Noah Snavely

Lecture 1: Images and image filtering

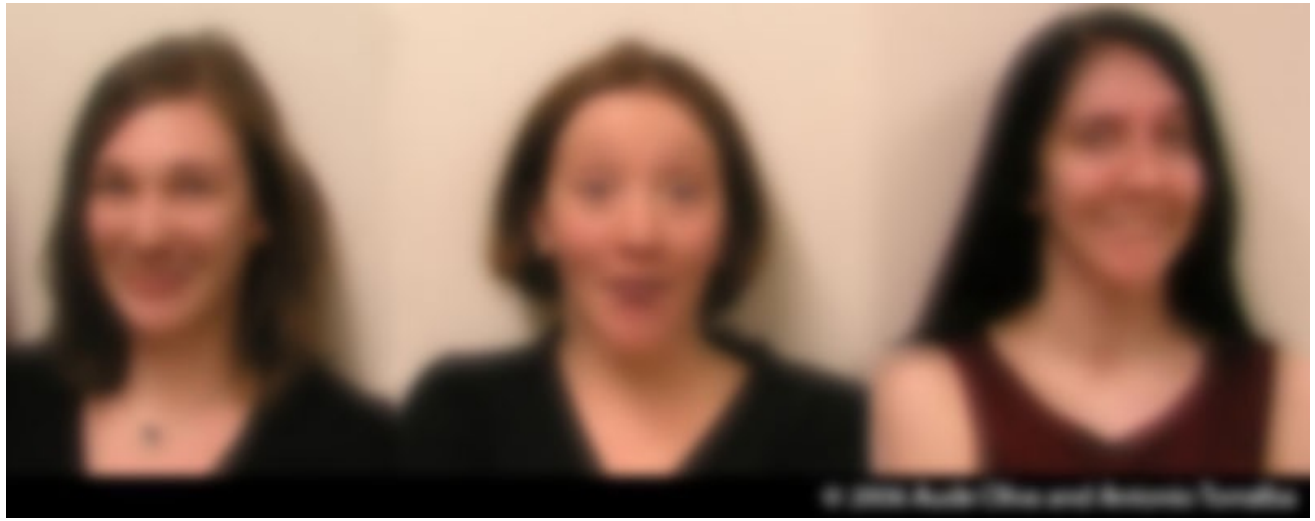


Hybrid Images, Oliva et al., http://olivalab.mit.edu/hybridimage.htm

CS5670: Intro to Computer Vision

Noah Snavely

Lecture 1: Images and image filtering



Hybrid Images, Oliva et al., http://olivalab.mit.edu/hybridimage.htm

Reading

- Szeliski, Chapter 3.1-3.3

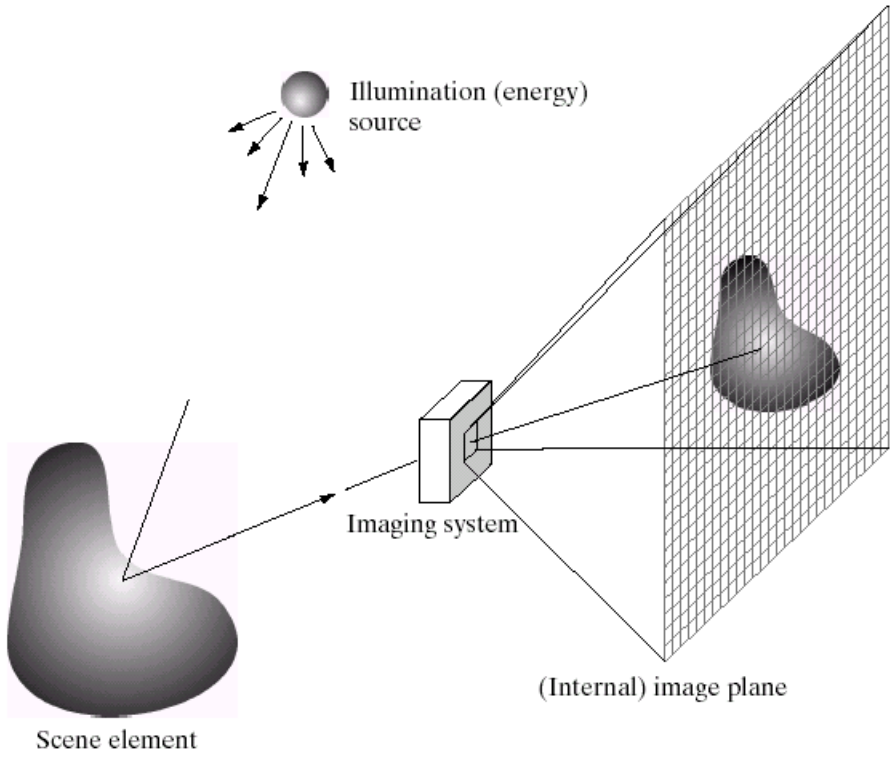
Announcements

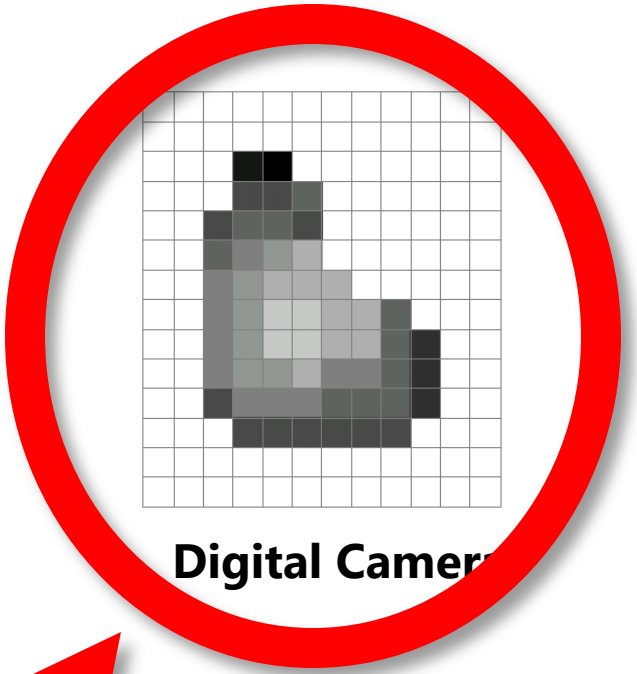
- Project 1 (Hybrid Images) will be released early next week
  - Code due Friday, Feb 10
  - This project will be done solo
  - Other projects planned to be done in groups of 2

- Project is in Python – we will provide skeleton code and instructions for setting up a Python environment for the project

What is an image?

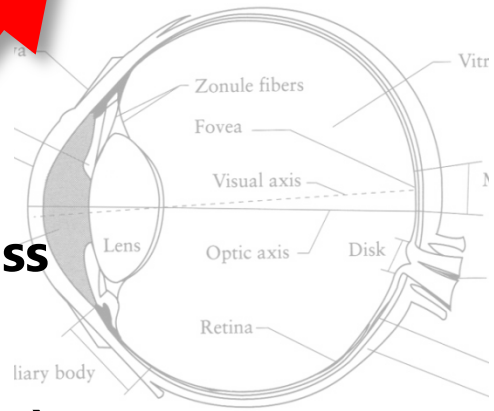
What is an image?





We'll focus on these in this class

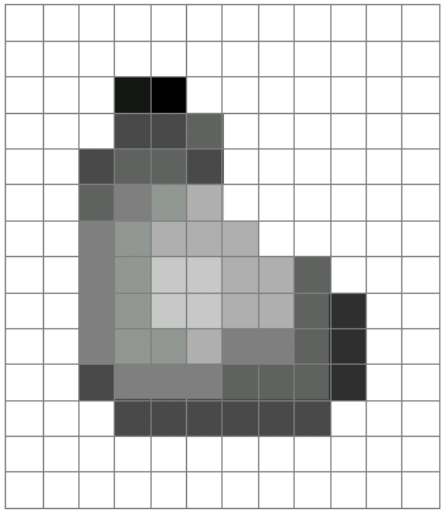
(More on this process later)



Source: A. Efros

What is an image?

- A grid (matrix) of intensity values



=

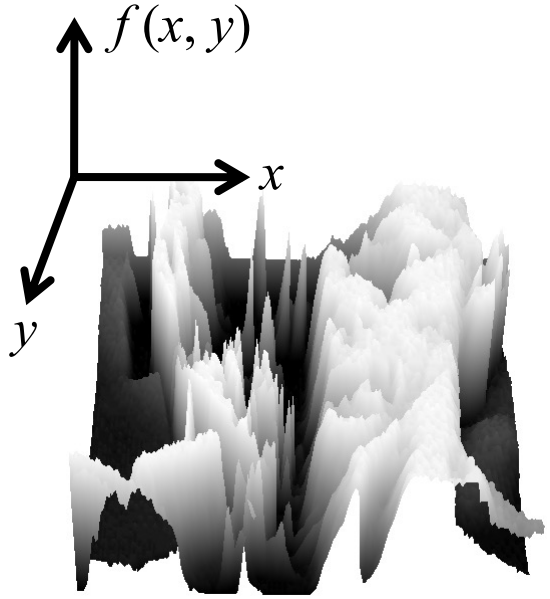
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	20	0	255	255	255	255	255	255	255
255	255	255	75	75	75	255	255	255	255	255	255
255	255	75	95	95	75	255	255	255	255	255	255
255	255	96	127	145	175	255	255	255	255	255	255
255	255	127	145	175	175	175	255	255	255	255	255
255	255	127	145	200	200	175	175	95	255	255	255
255	255	127	145	200	200	175	175	95	47	255	255
255	255	127	145	145	175	127	127	95	47	255	255
255	255	74	127	127	127	95	95	95	47	255	255
255	255	255	74	74	74	74	74	74	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255

(common to use one byte per value: 0 = black, 255 = white)

What is an image?

- Can think of a (grayscale) image as a **function** *f* from R^2 to R:
  - $f(x,y)$ gives the **intensity** at position (x,y)

snoop

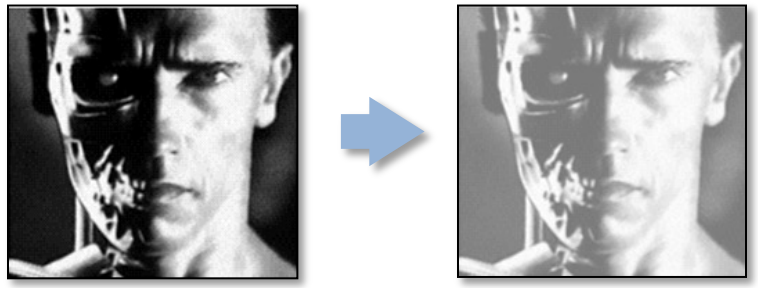


3D view

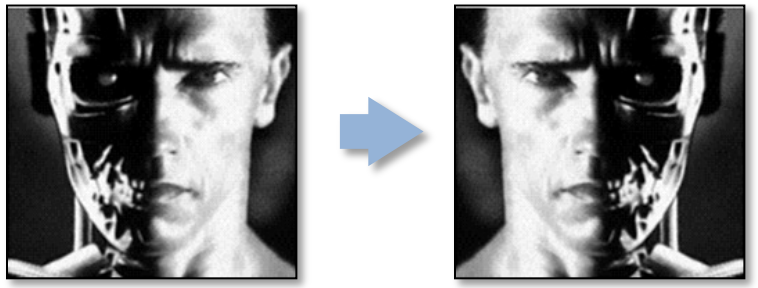
  - A **digital** image is a discrete (**sampled**, **quantized**) version of this function

Image transformations

- As with any function, we can apply operators to an image



$g(x,y) = f(x,y) + 20$



$g(x,y) = f(-x,y)$

- Today we'll talk about a special kind of operator, *convolution* (linear filtering)

Filters

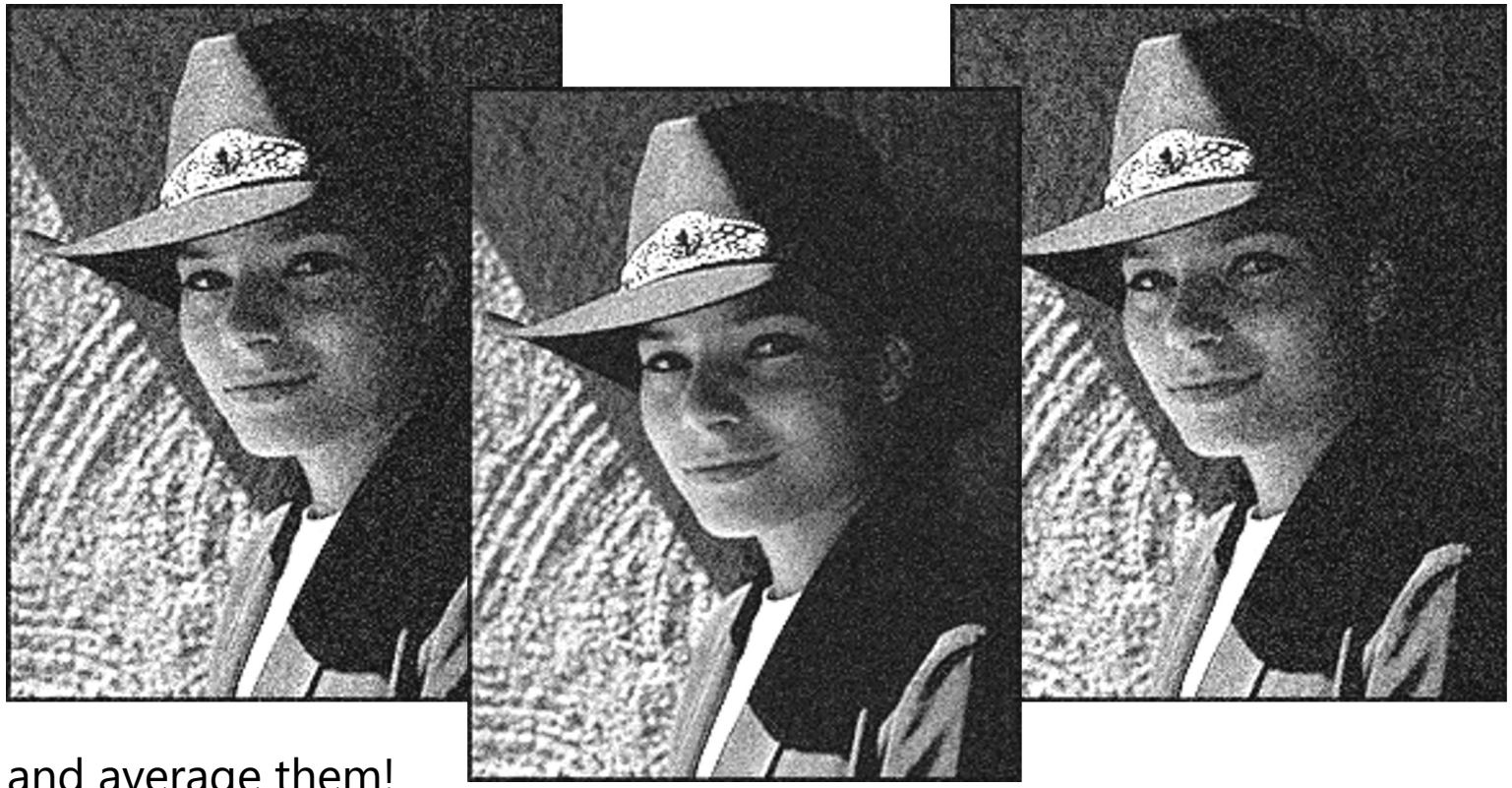
- Filtering
  - Form a new image whose pixel values are a combination of the original pixel values
- Why?
  - To get useful information from images
    - E.g., extract edges or contours (to understand shape)
  - To enhance the image
    - E.g., to remove noise
    - E.g., to sharpen and “enhance image” a la CSI
  - A key operator in Convolutional Neural Networks

Canonical Image Processing problems

- Image Restoration
  - denoising
  - deblurring
- Image Compression
  - JPEG, HEIF, MPEG, …
- Locating Structural Features
  - corners
  - edges

Question: Noise reduction

- Given a camera and a still scene, how can you reduce noise?



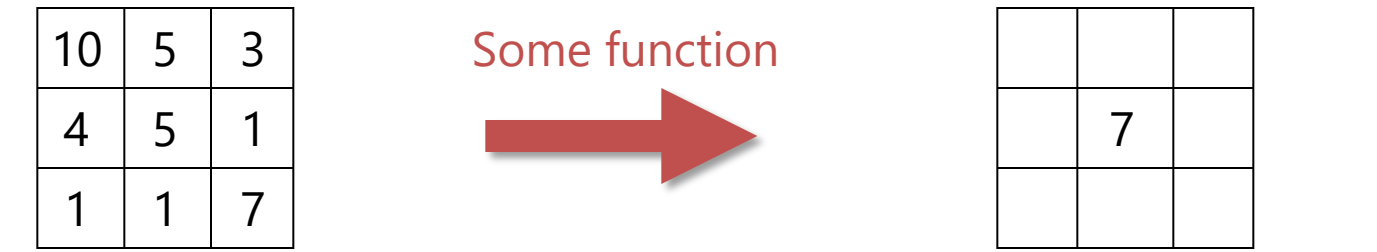
Take lots of images and average them!

What's the next best thing?

Source: S. Seitz

Image filtering

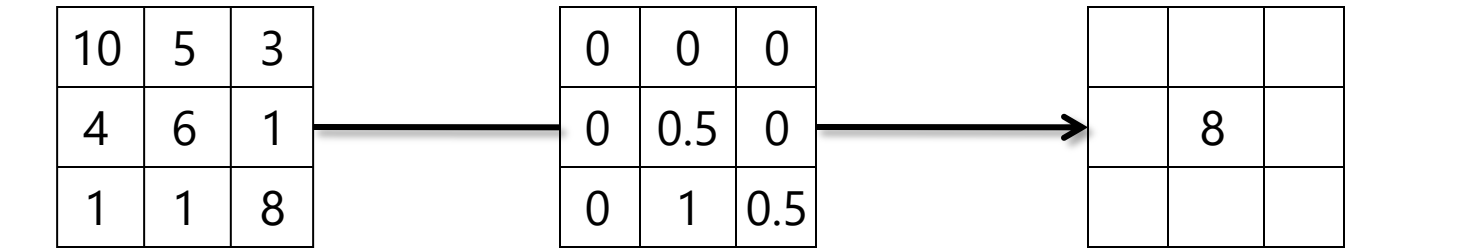
- Modify the pixels in an image based on some function of a local neighborhood of each pixel



Source: L. Zhang

Linear filtering

- One simple version of filtering: linear filtering (cross-correlation, convolution)
  - Replace each pixel by a linear combination (a weighted sum) of its neighbors
- The prescription for the linear combination is called the “kernel” (or “mask”, “filter”)



Source: L. Zhang

Cross-correlation

Let F be the image, H be the kernel (of size $2k+1 \times 2k+1$), and G be the output image

$$G[i,j] = \sum_{u=-k}^{k} \sum_{v=-k}^{k} H[u,v]F[i+u, j+v]$$

This is called a **cross-correlation** operation:

$$G = H \otimes F$$

- Can think of as a "dot product" between local neighborhood and kernel for each pixel

Convolution

- Same as cross-correlation, except that the kernel is "flipped" (horizontally and vertically)

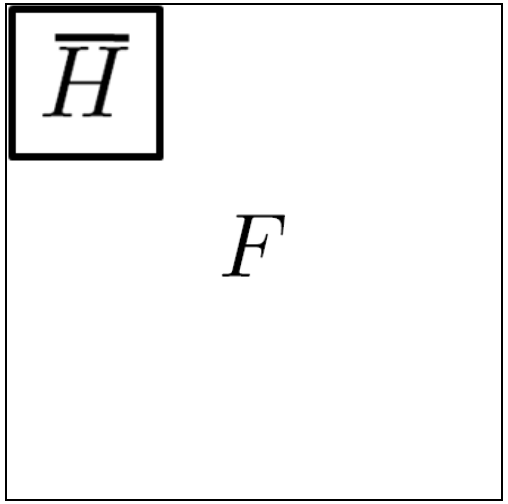
$$G[i,j] = \sum_{u=-k}^{k} \sum_{v=-k}^{k} H[u,v]F[i-u,j-v]$$

This is called a **convolution** operation:

$$G = H * F$$

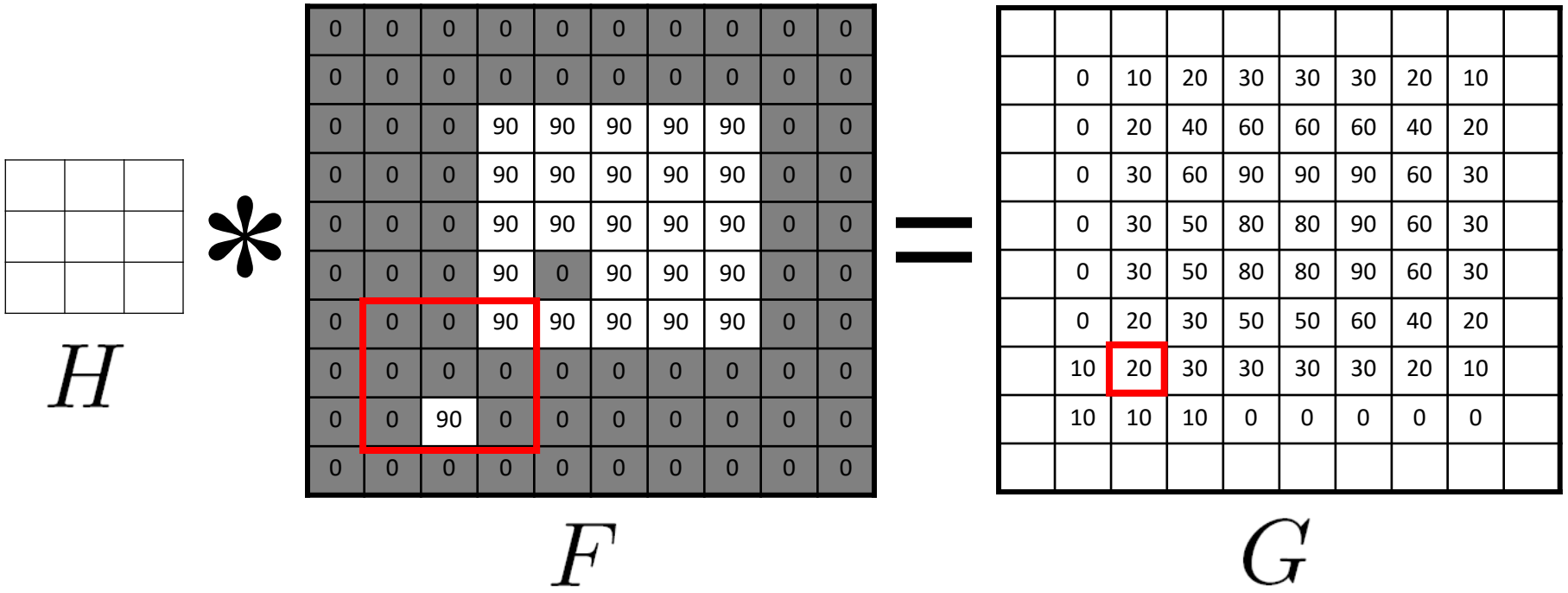
- Convolution is **commutative** and **associative**

Convolution



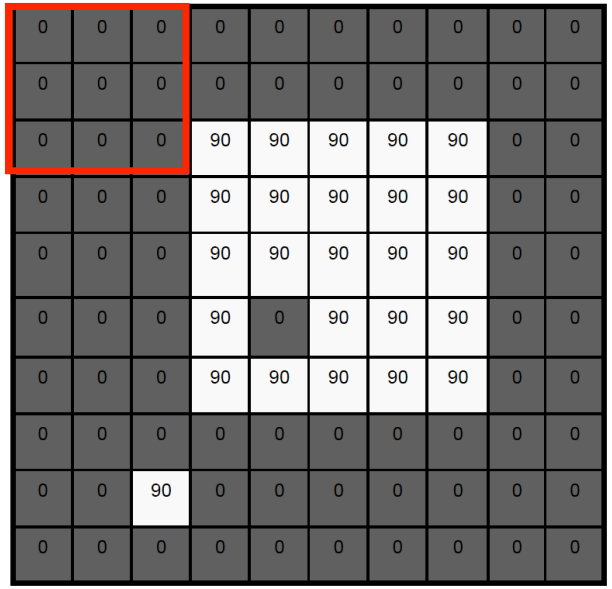
Adapted from F. Durand

Mean filtering

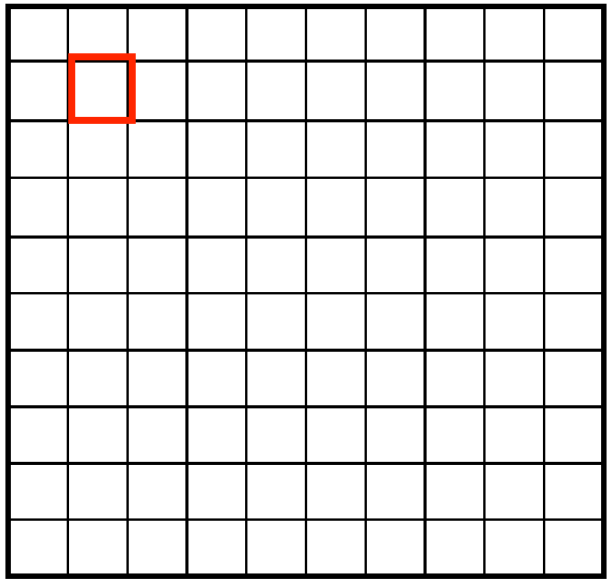


Mean filtering/Moving average

$F[x, y]$

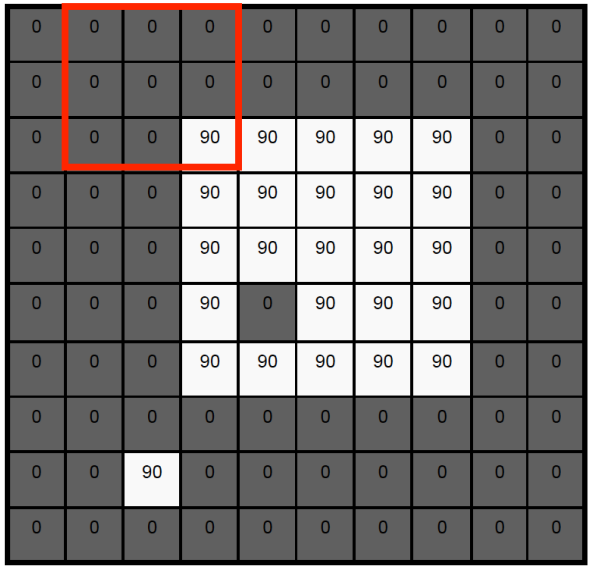


$G[x, y]$

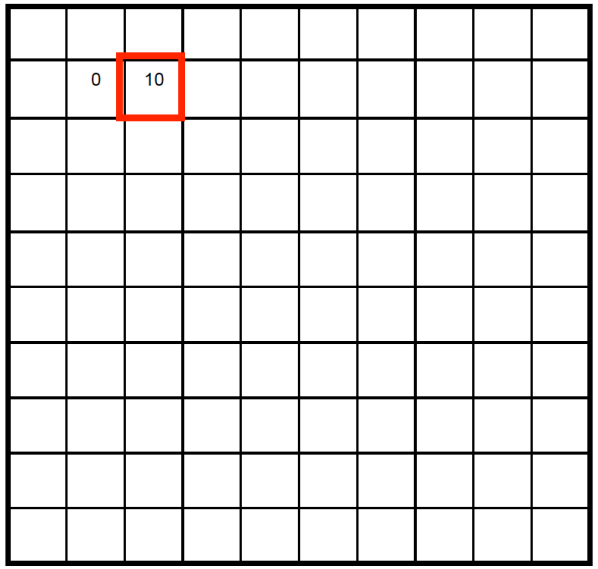


Mean filtering/Moving average

$F[x, y]$

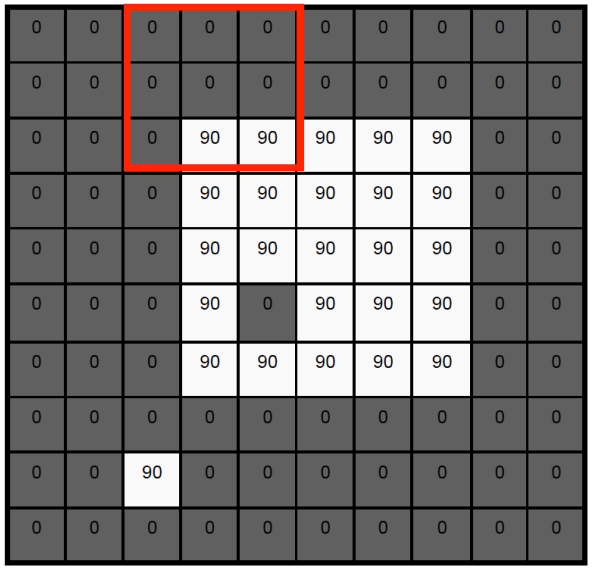


$G[x, y]$

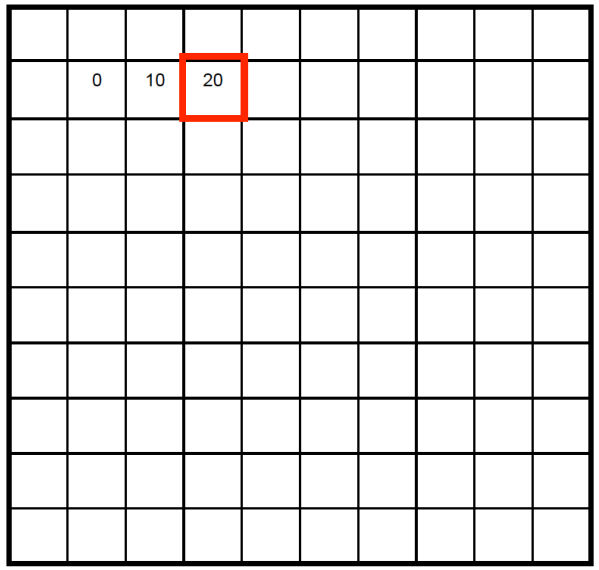


Mean filtering/Moving average

$F[x, y]$

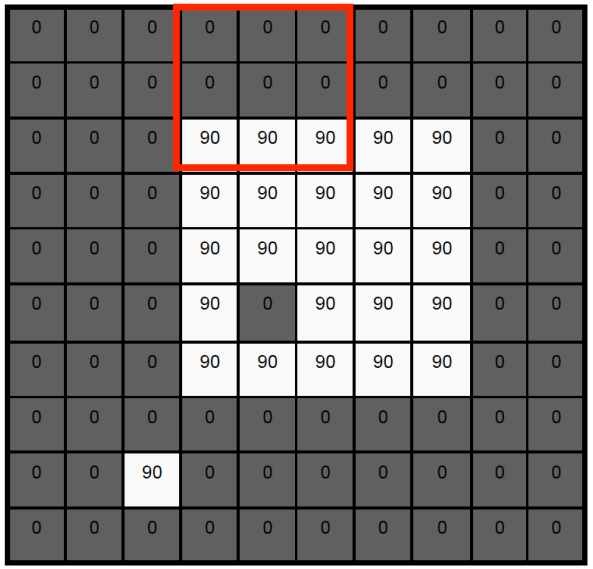


$G[x, y]$

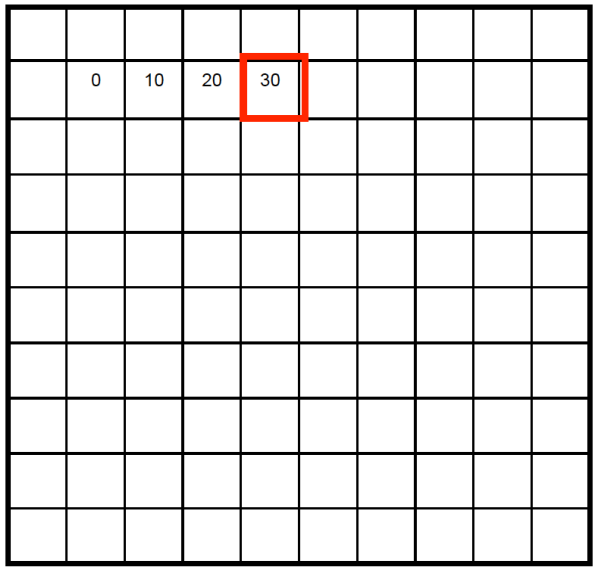


Mean filtering/Moving average

$F[x, y]$

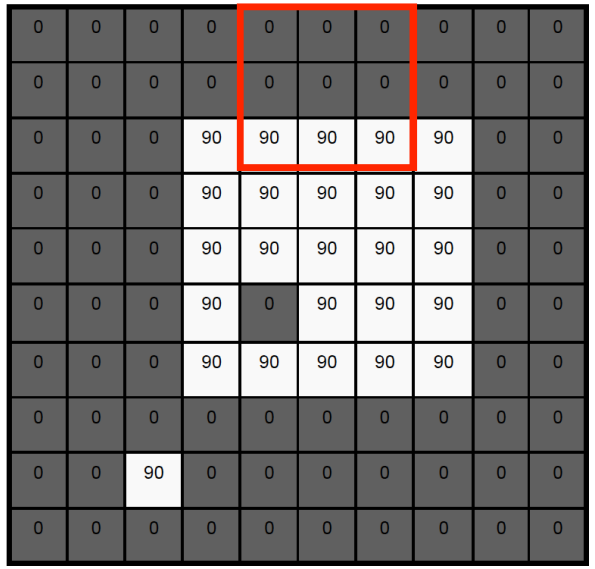


$G[x, y]$

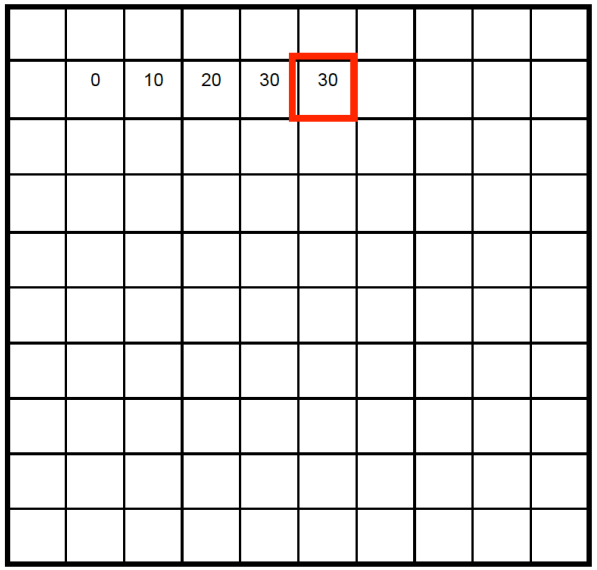


Mean filtering/Moving average

$F[x, y]$



$G[x, y]$



Mean filtering/Moving average

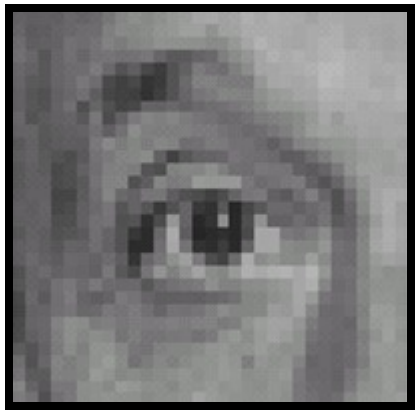
$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$G[x, y]$

	0	10	20	30	30	30	20	10	
	0	20	40	60	60	60	40	20	
	0	30	60	90	90	90	60	30	
	0	30	50	80	80	90	60	30	
	0	30	50	80	80	90	60	30	
	0	20	30	50	50	60	40	20	
	10	20	30	30	30	30	20	10	
	10	10	10	0	0	0	0	0	

Linear filters: examples



Original

$$*\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

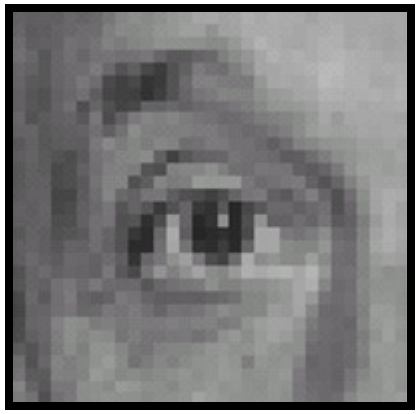
Source: D. Lowe

slido

What image operation does filtering with this kernel perform?
([0 0 0; 0 1 0; 0 0 0])

ⓘ Start presenting to display the poll results on this slide.

Linear filters: examples



Original

$*$

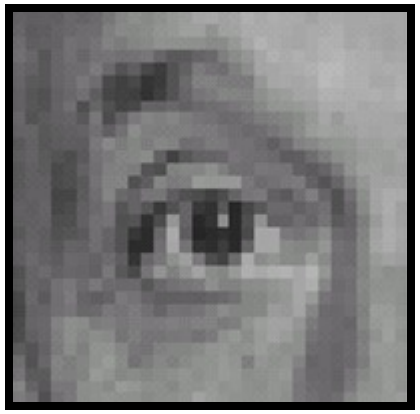
0	0	0
0	1	0
0	0	0

$=$

Identical image

Source: D. Lowe

Linear filters: examples



Original

*

0	0	0
1	0	0
0	0	0

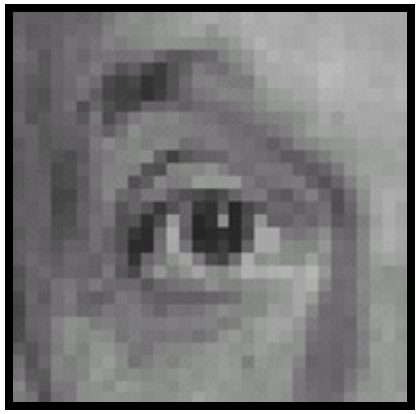
Source: D. Lowe

slido

What image operation does filtering with this kernel perform? ([0 0 0; 1 0 0; 0 0 0])

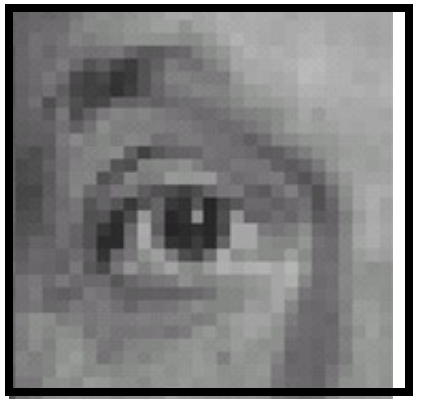
ⓘ Start presenting to display the poll results on this slide.

Linear filters: examples



Original

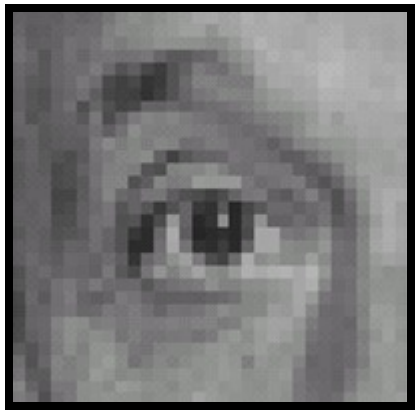
$$*\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} =$$



Shifted left by 1 pixel

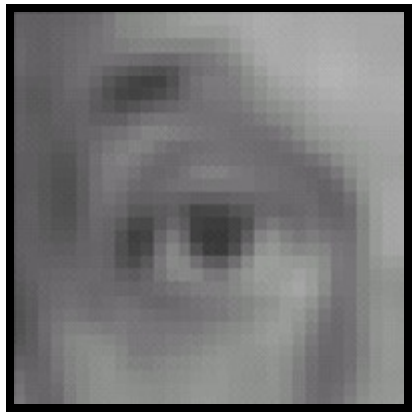
Source: D. Lowe

Linear filters: examples



Original

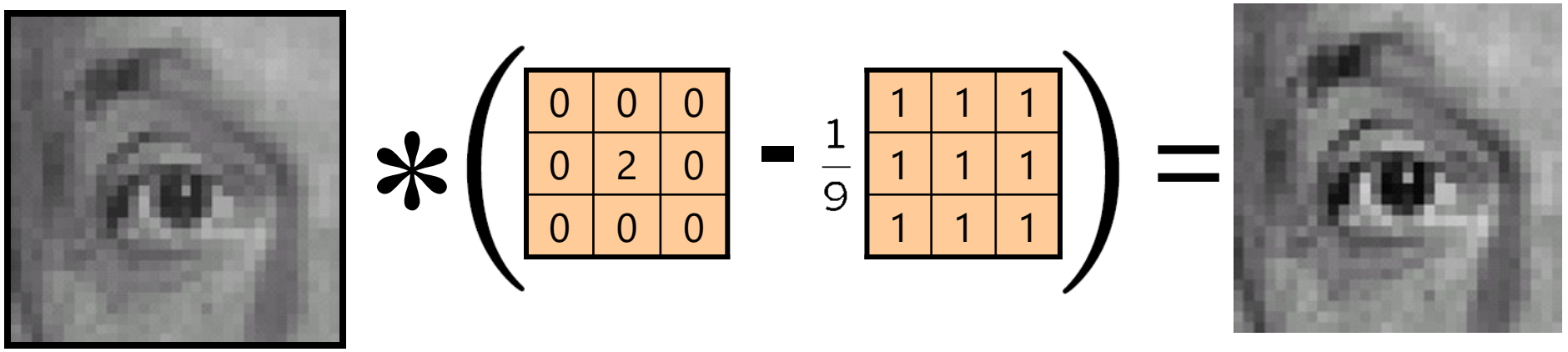
$$* \quad \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad =$$



Blur (with a mean filter)

Source: D. Lowe

Linear filters: examples

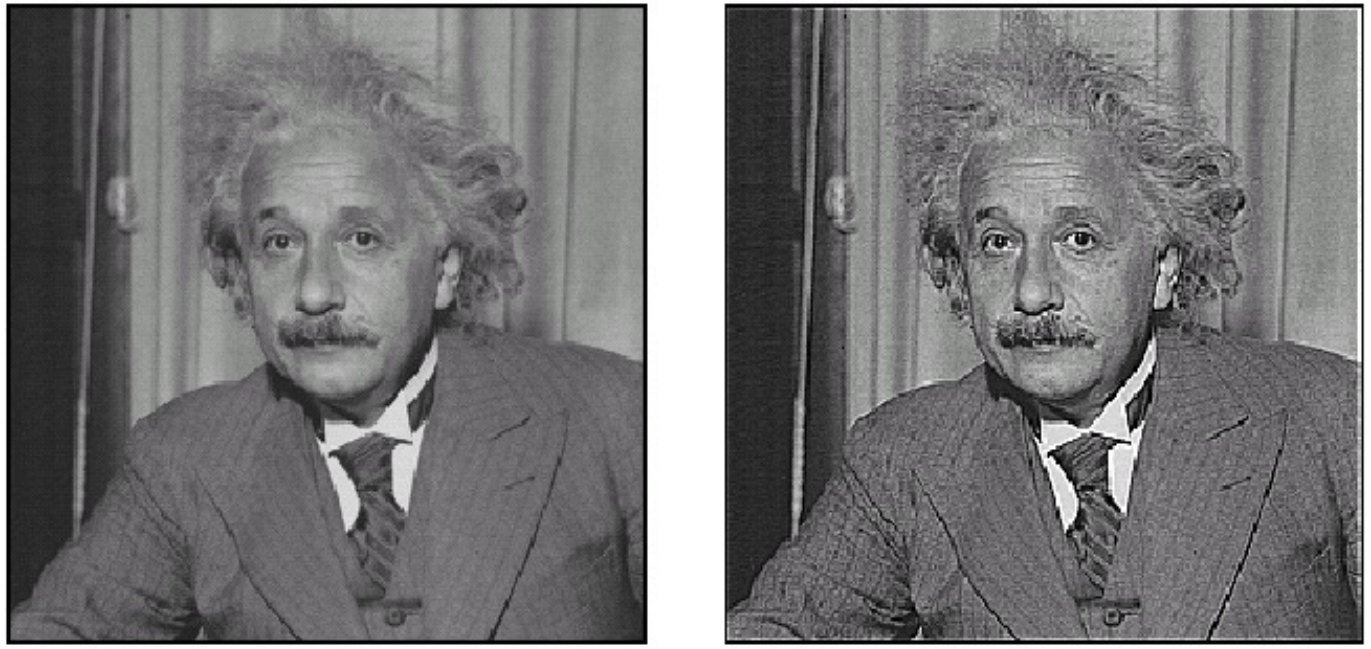


Original

Sharpening filter
(accentuates edges)

Source: D. Lowe

Sharpening

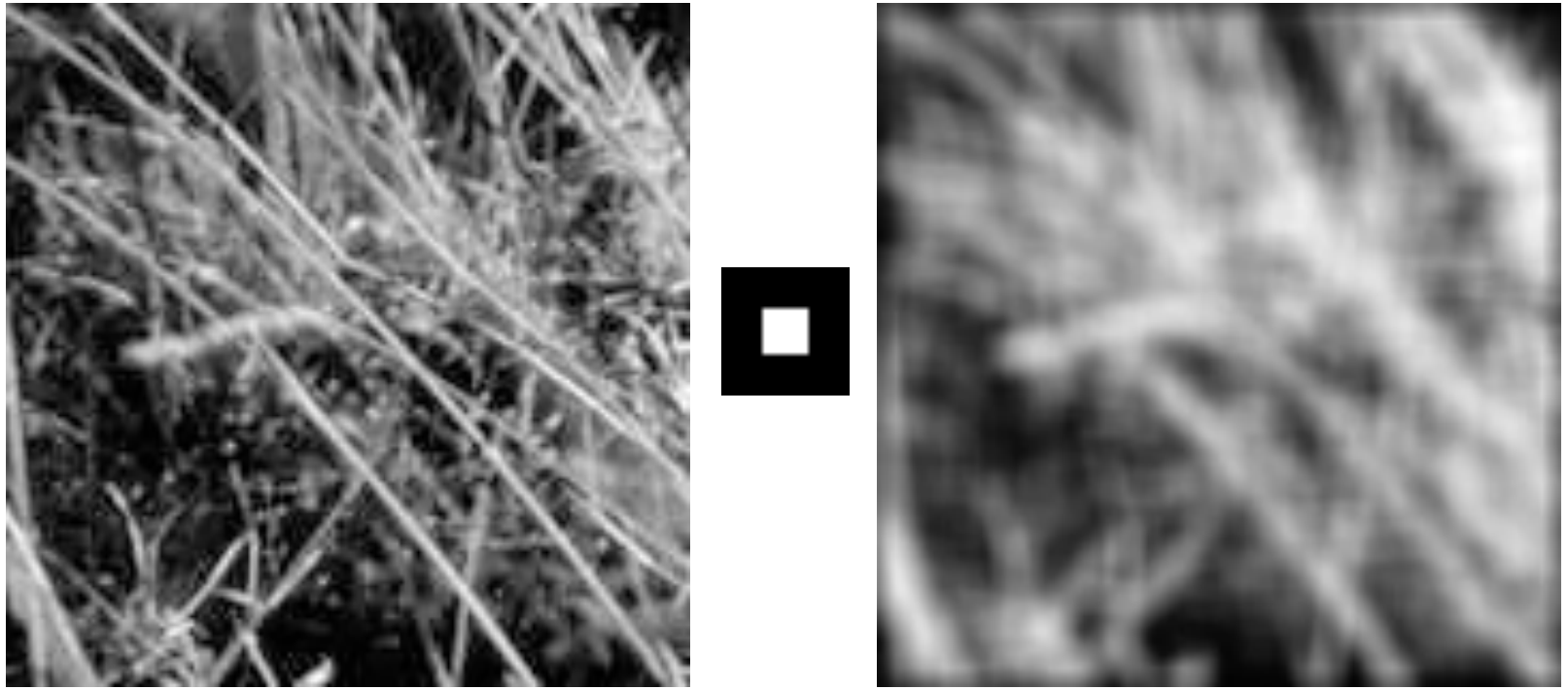


before

after

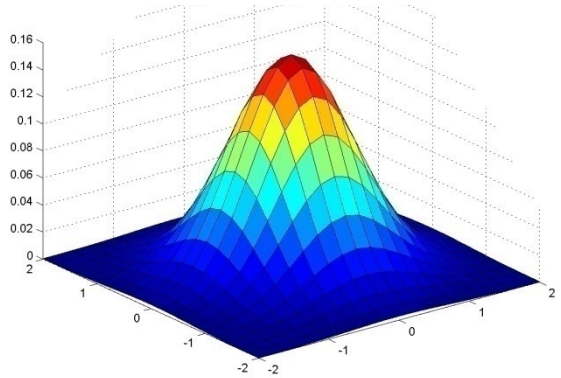
Source: D. Lowe

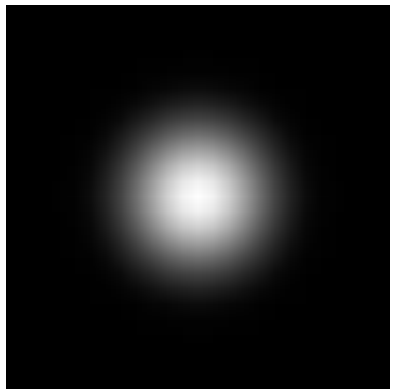
Smoothing with box filter revisited



Source: D. Forsyth

Gaussian kernel

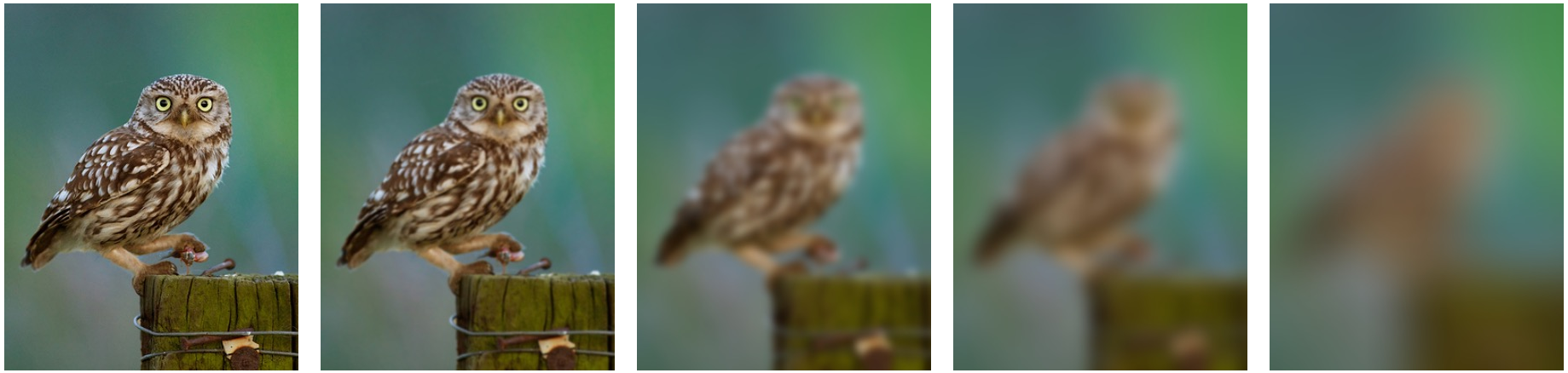




$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

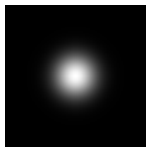
Source: C. Rasmussen

Gaussian filters



σ = 1 pixel

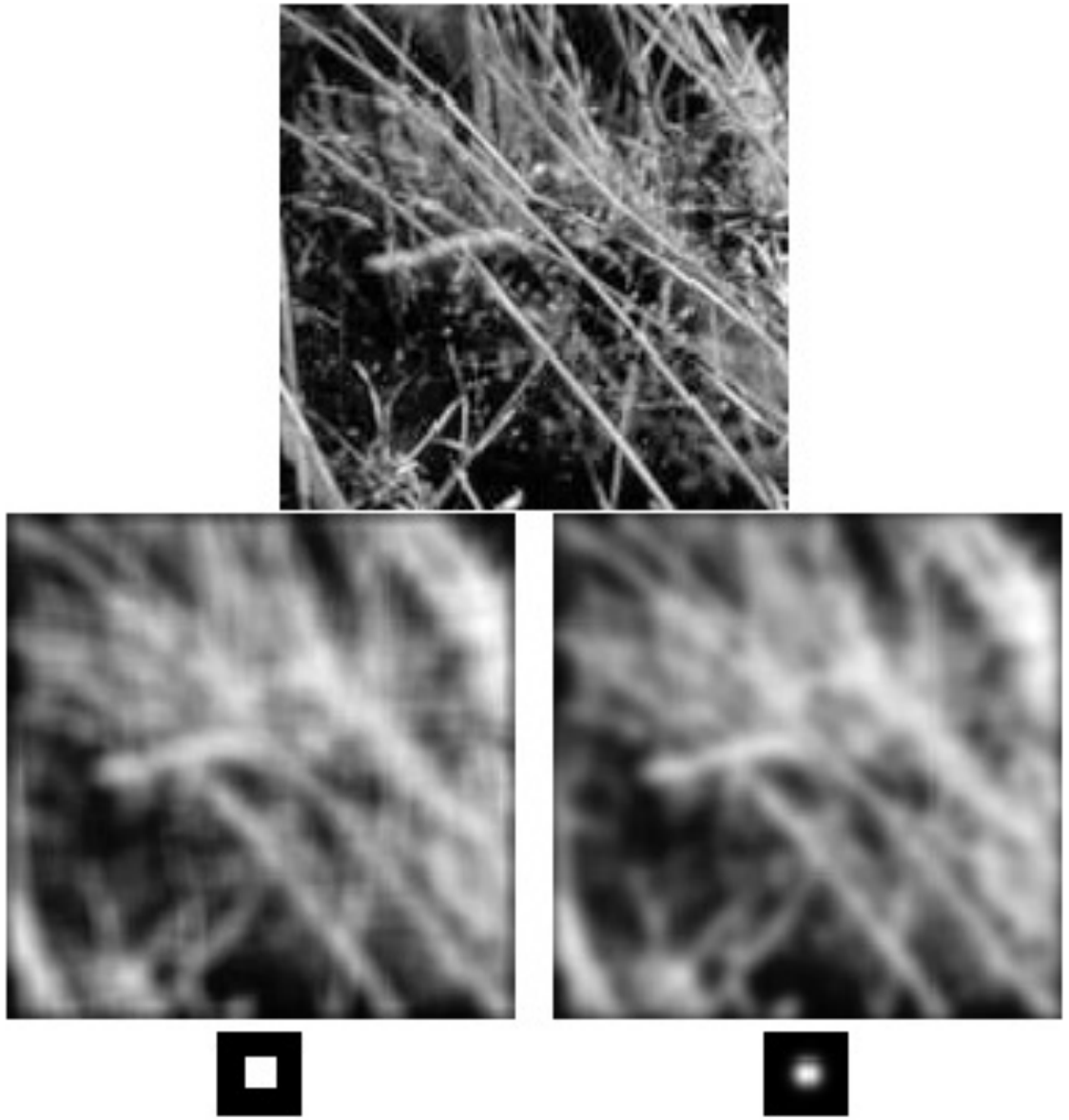
σ = 5 pixels



σ = 10 pixels

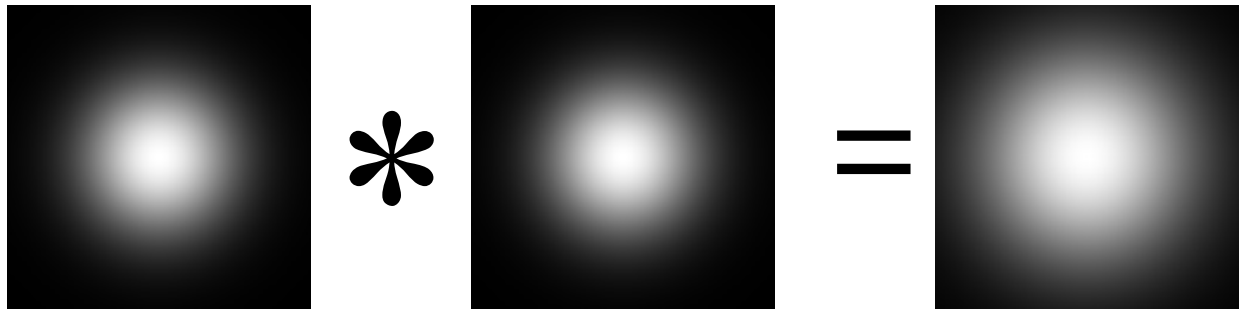
σ = 30 pixels

Mean vs. Gaussian filtering



Gaussian filter

- Removes “high-frequency” components from the image (low-pass filter)
- Convolution with self is another Gaussian



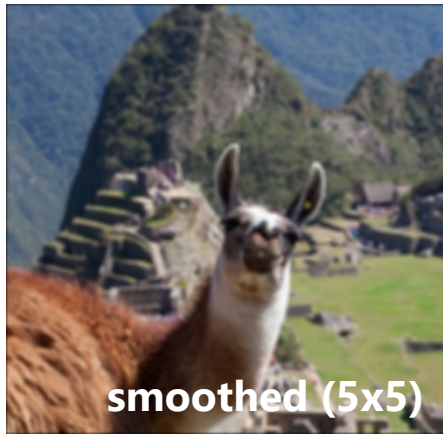
– Convolving twice with Gaussian kernel of width σ
= convolving once with kernel of width $\sigma\sqrt{2}$

Source: K. Grauman

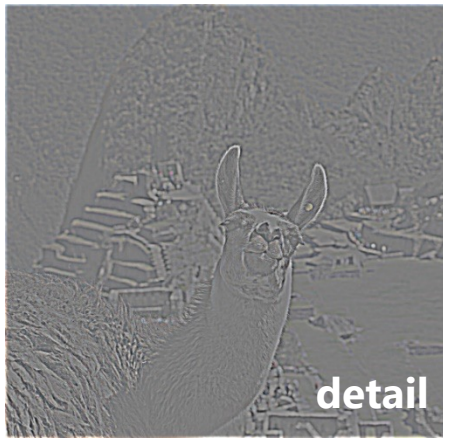
Sharpening revisited

- What does blurring take away?

−



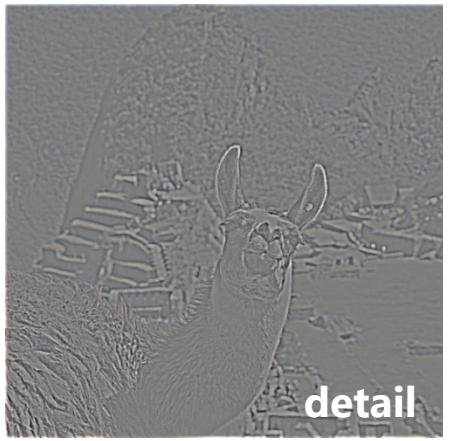
=



(This “detail extraction” operation is also called a ***high-pass filter***)

Let’s add it back:

+ α



=

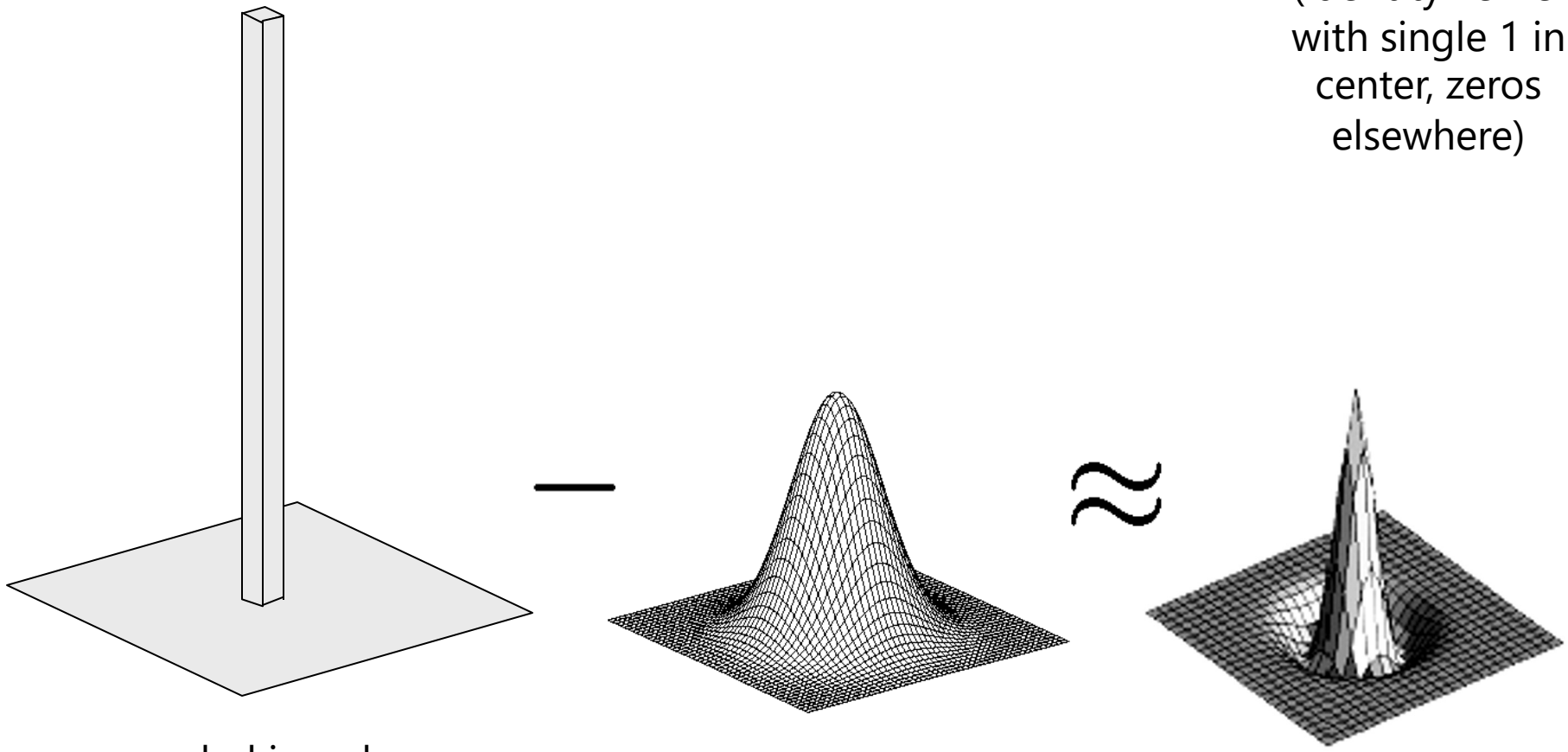
Photo credit: https://www.flickr.com/photos/geezaweezer/16089096376/

Sharpen filter

$$F + \alpha\,(F - \underbrace{F * H}_{\text{blurred image}}) =$$

image

unit impulse (identity kernel with single 1 in center, zeros elsewhere)

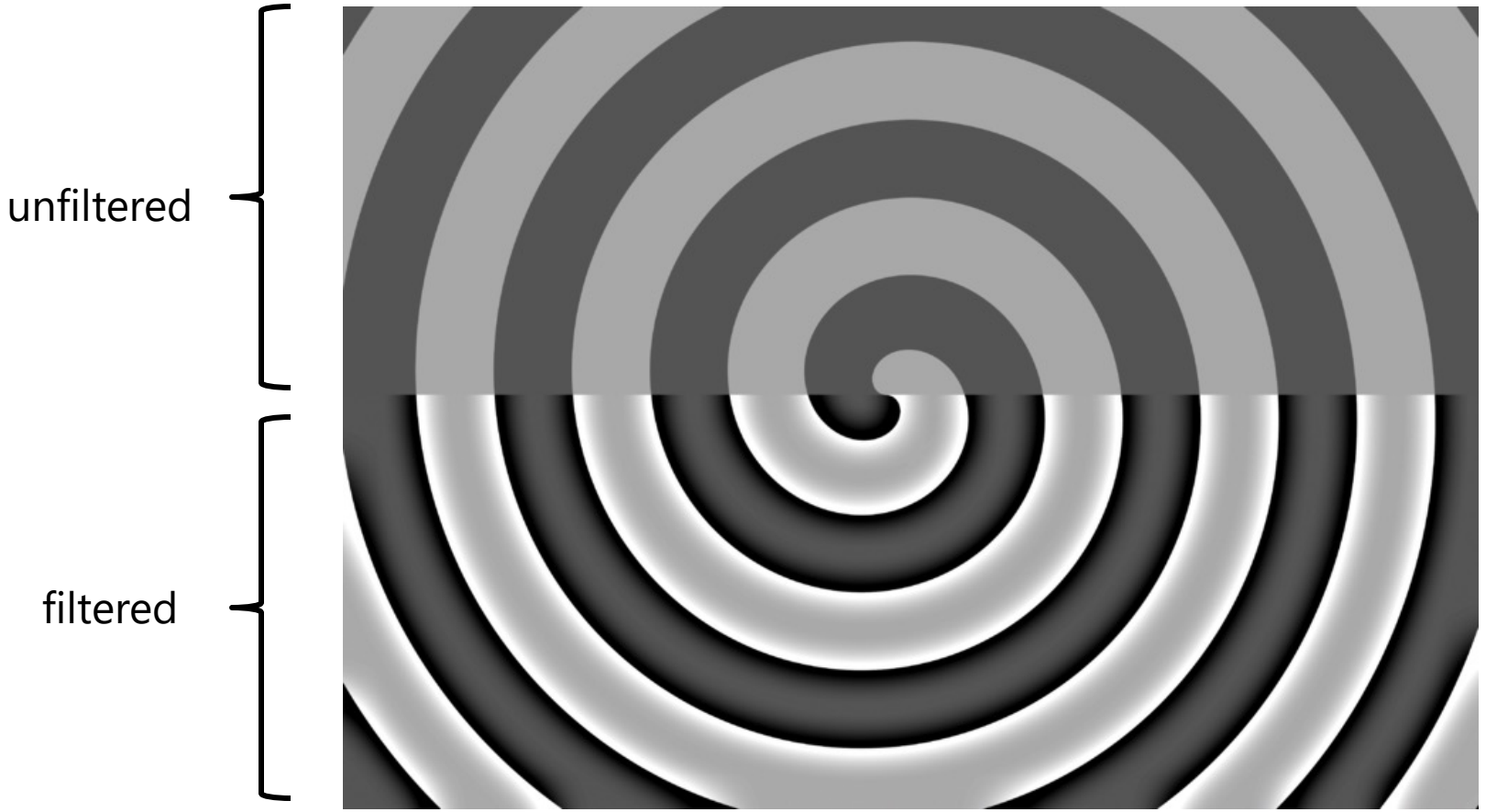


scaled impulse

Gaussian

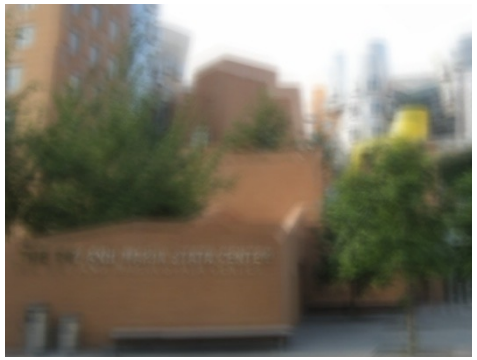
Sharpen filter

Sharpen filter



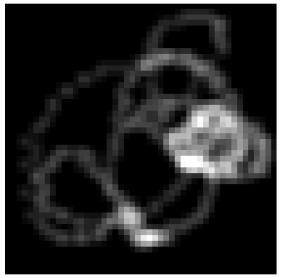
“Optical” convolution

Camera shake



=

*



Source: Fergus, *et al.* “Removing Camera Shake from a Single Photograph”, SIGGRAPH 2006

Bokeh: Blur in out-of-focus regions of an image.

Source: https://www.diyphotography.net/diy_create_your_own_bokeh/

Filters: Thresholding

$$g(m,n) = \begin{cases} 255, & f(m,n) > A \\ 0 & otherwise \end{cases}$$

Linear filters

- Can thresholding be implemented with a linear filter?

slido

Can thresholding be implemented with a linear filter?

ⓘ Start presenting to display the poll results on this slide.

Questions?