COMP 590/776: Computer Vision

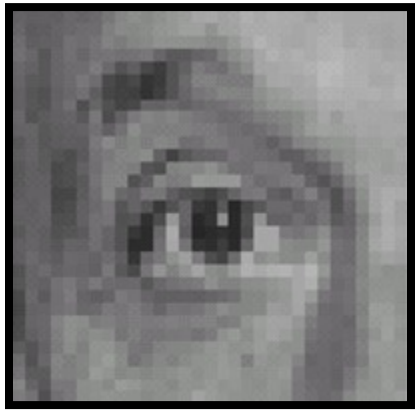
Lecture 7: Features 1

Instructor: Soumyadip (Roni) Sengupta

ULA: Andrea Dunn, William Li, Liujie Zheng

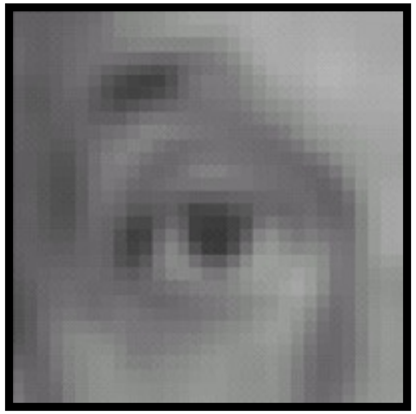
Course Website:
Scan Me!

Recap



Original

Linear Filtering:
Cross-correlation
& Convolution



Blur

Shifted left by 1 pixel

Sharpening

Convolution

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

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

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

Cross-correlation

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

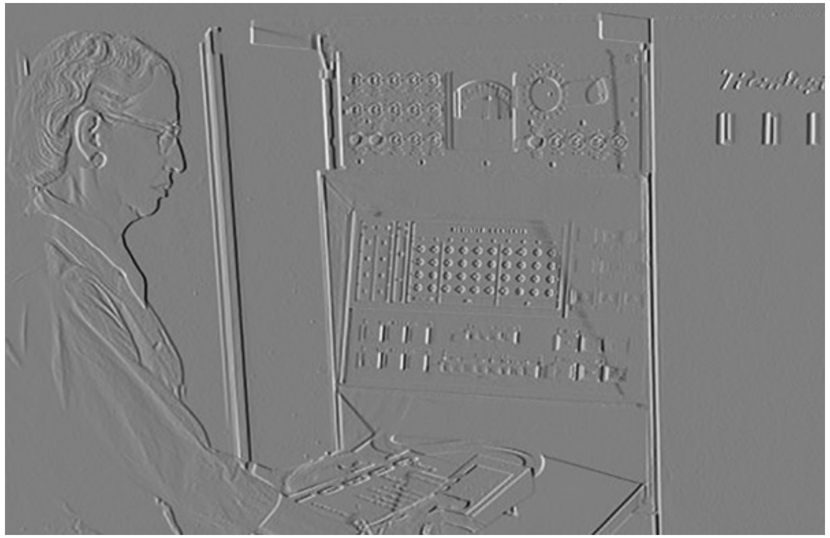
To accurately downsample a signal/image, # of samples >= 2*highest frequency in the signal. (Nyquist Rate!)

Gaussian pre-filtering

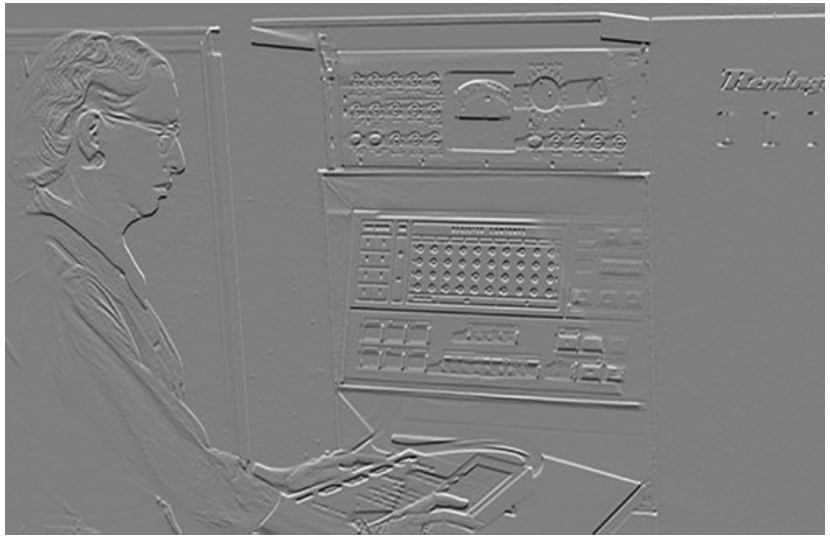
Partial Derivatives

Can be implemented as a convolution operation

-1	1



$$\frac{\partial f(x,y)}{\partial x}$$



$$\frac{\partial f(x,y)}{\partial y}$$

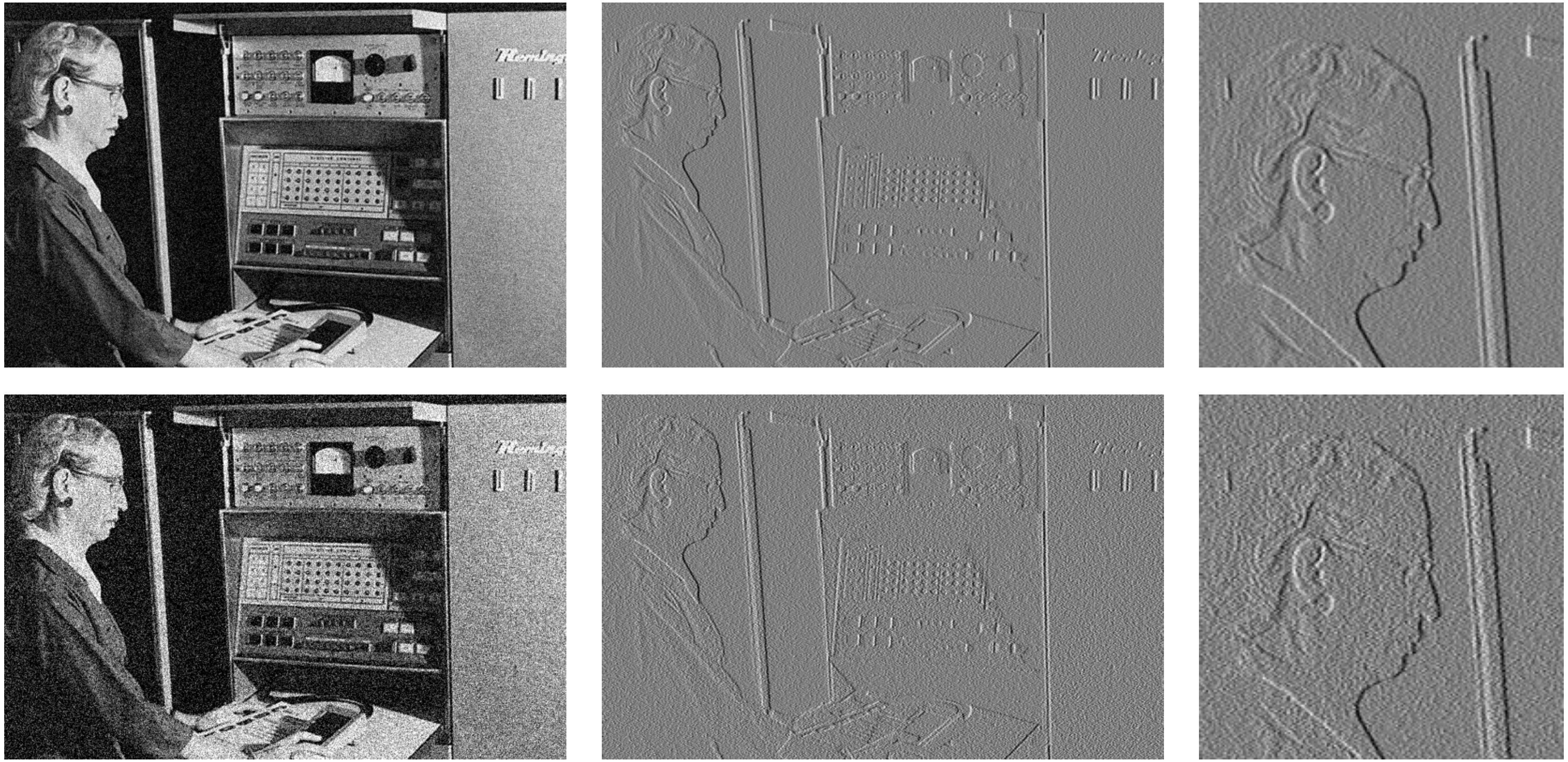
-1
1

or

1
-1

Noise in 2D

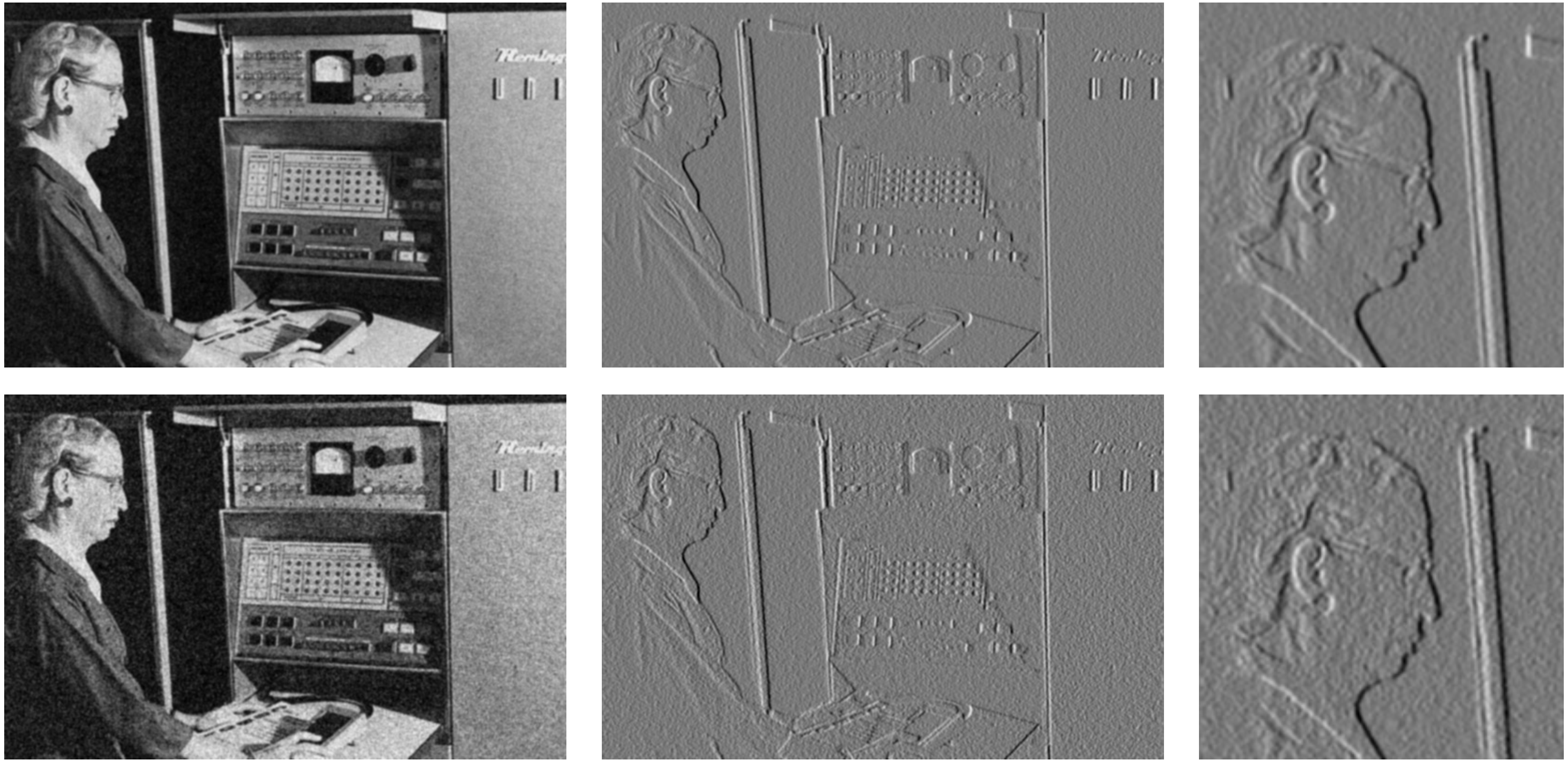
Noisy Input | Ix via [-1,0,1] | Zoom



Source: D. Fouhey

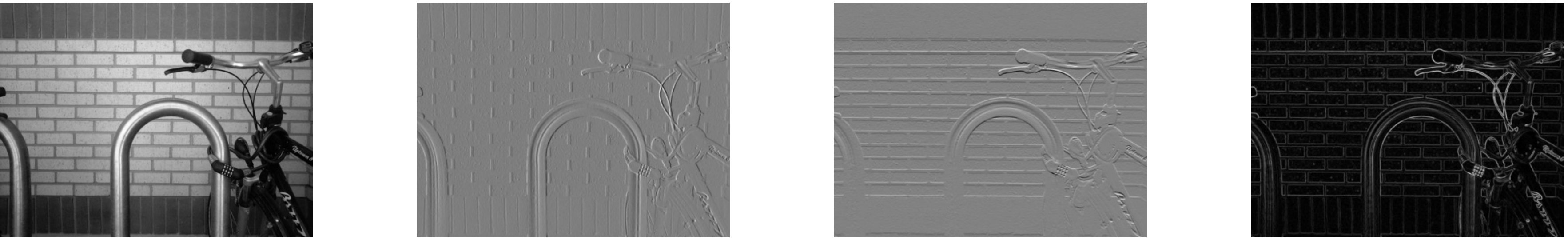
Noise + Smoothing

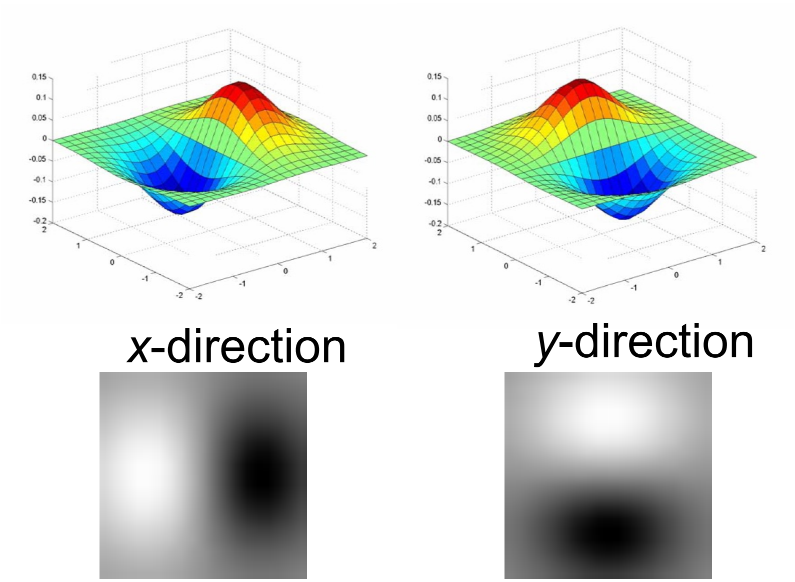
Smoothed Input | Ix via [-1,0,1] | Zoom



Source: D. Fouhey

Sobel operator: Approximation of derivative of gaussian





Source: Wikipedia

Canny edge detector

1. Filter image with derivative of Gaussian

2. Find magnitude and orientation of gradient.

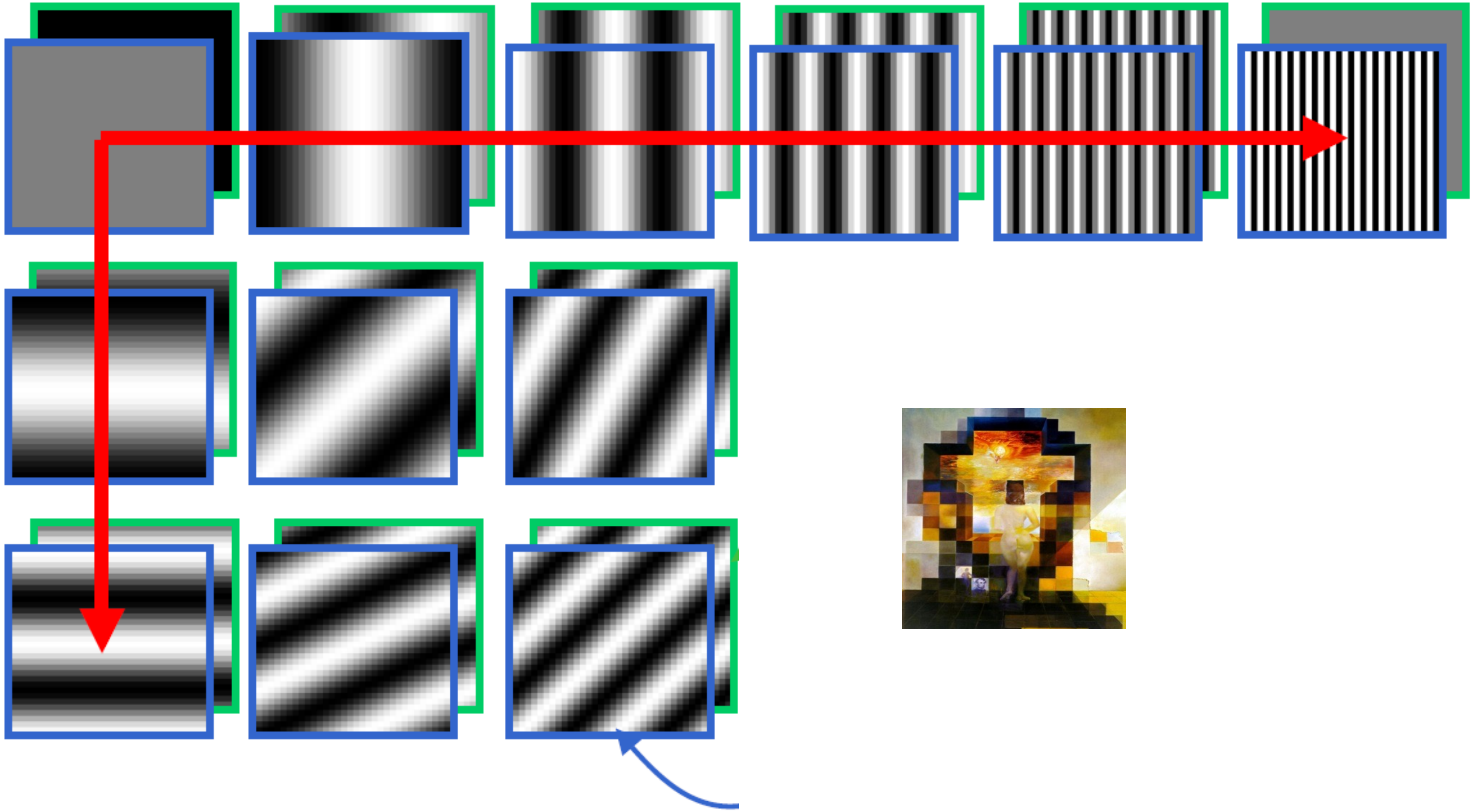
3. Non-maximum suppression

4. Linking and thresholding (hysteresis):
   - Define two thresholds: low and high
   - Use the high threshold to start edge curves and the low threshold to continue them

Source: D. Lowe, L. Fei-Fei, J. Redmon

Fourier Transform

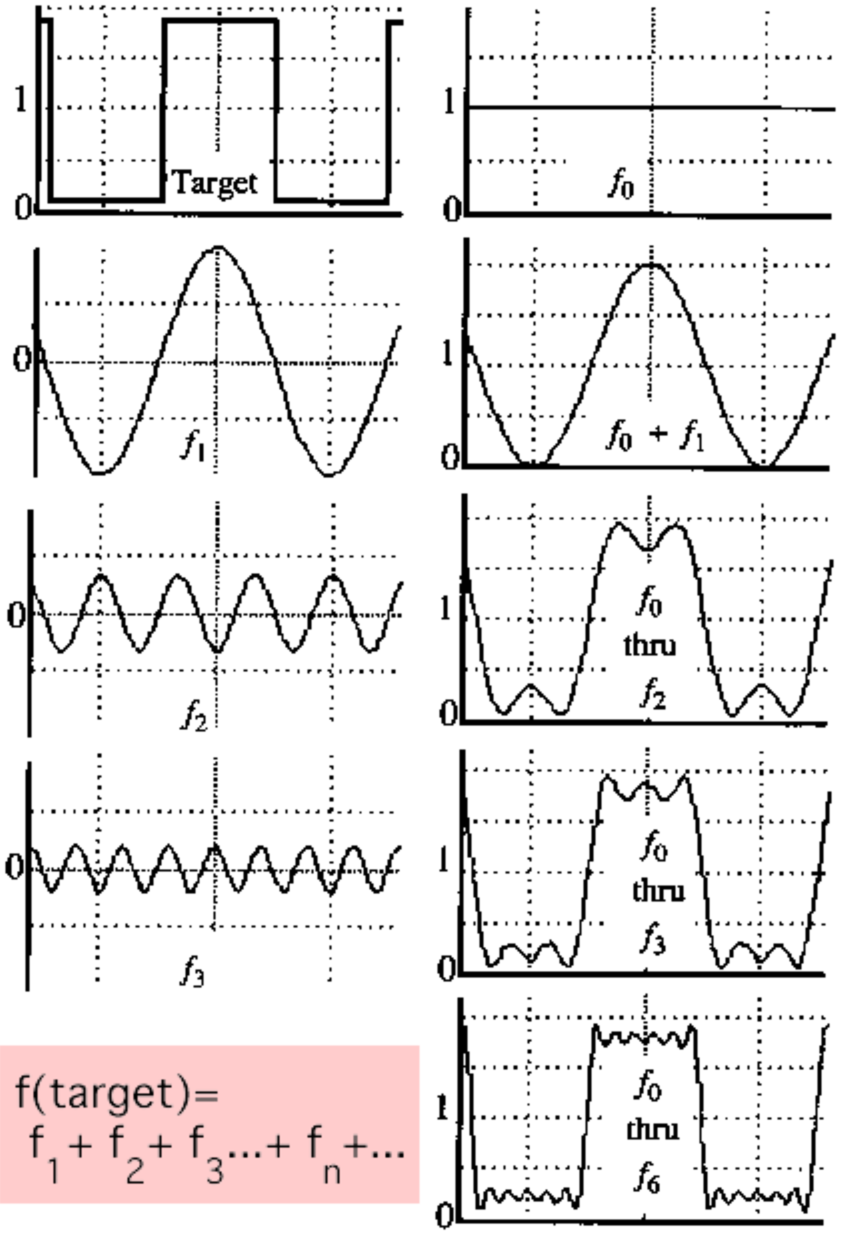
Teases away fast vs. slow changes in the image.



A sum of sines

- Our building block:
- $A\sin(\omega x + \phi)$

- Add enough of them to get any signal *f(x)* you want!

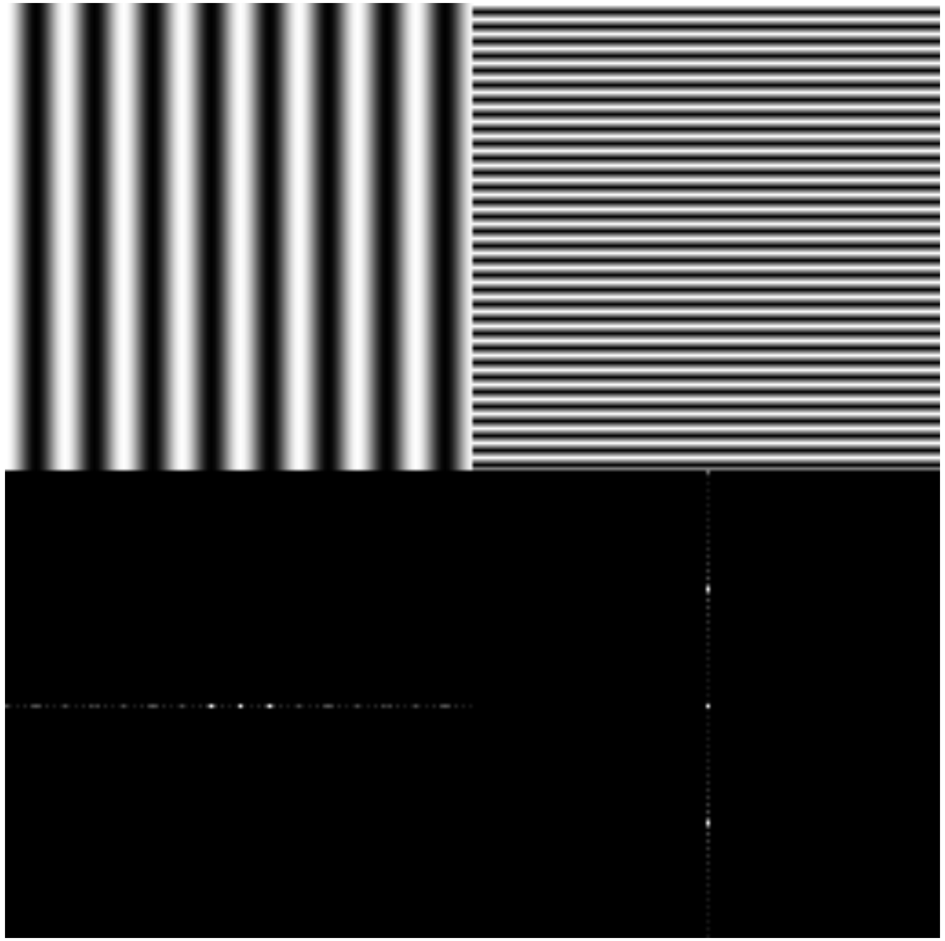


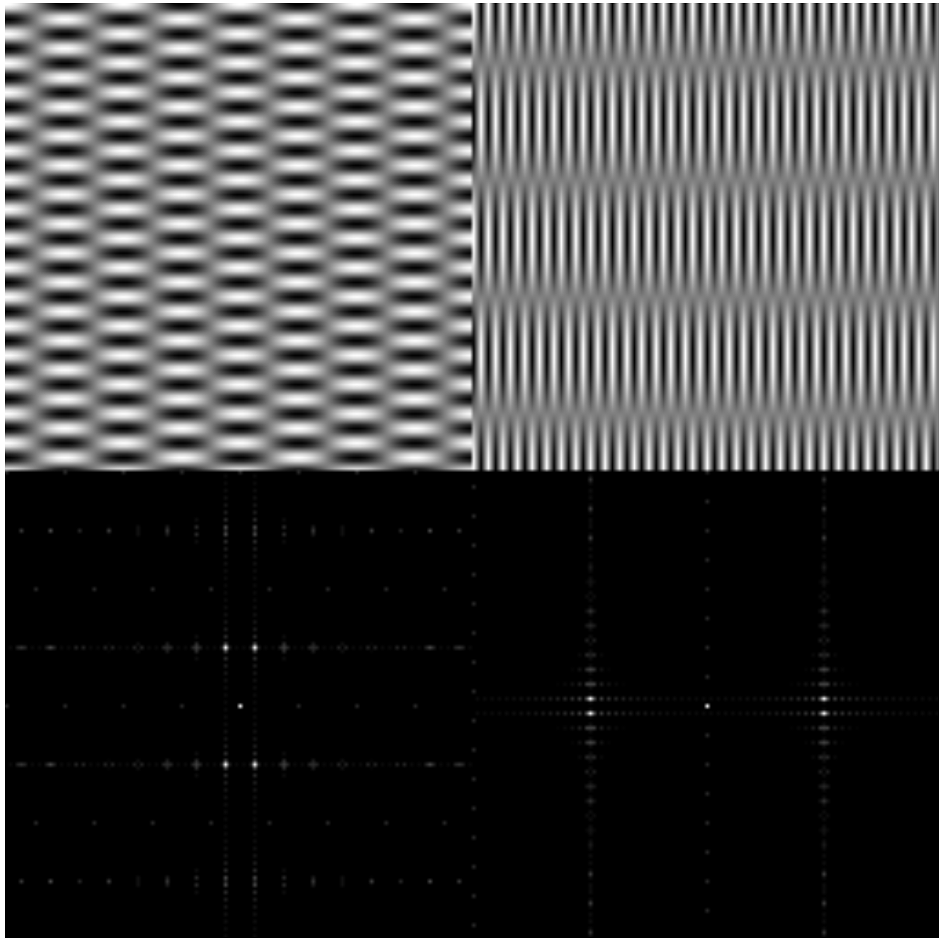
f(target)=
$f_1 + f_2 + f_3 ... + f_n + ...$

Amplitude Spectrum of DFT

Images

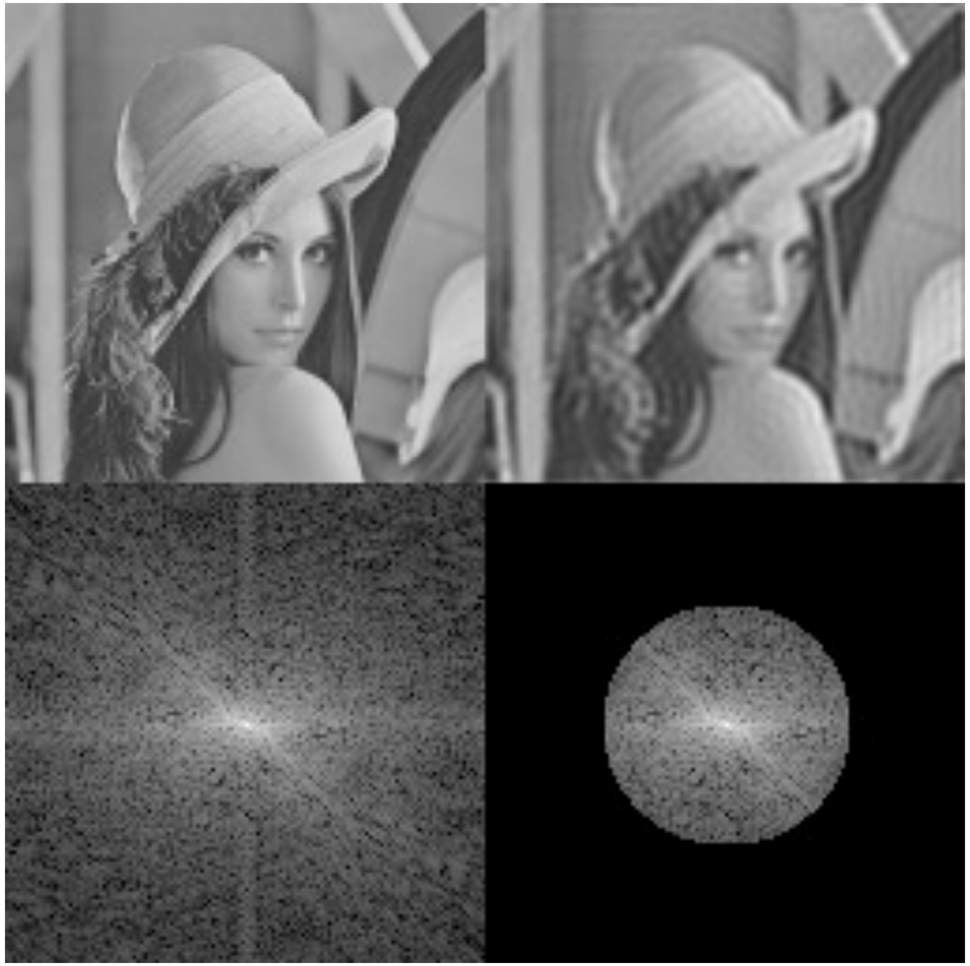
Amplitude spectrum



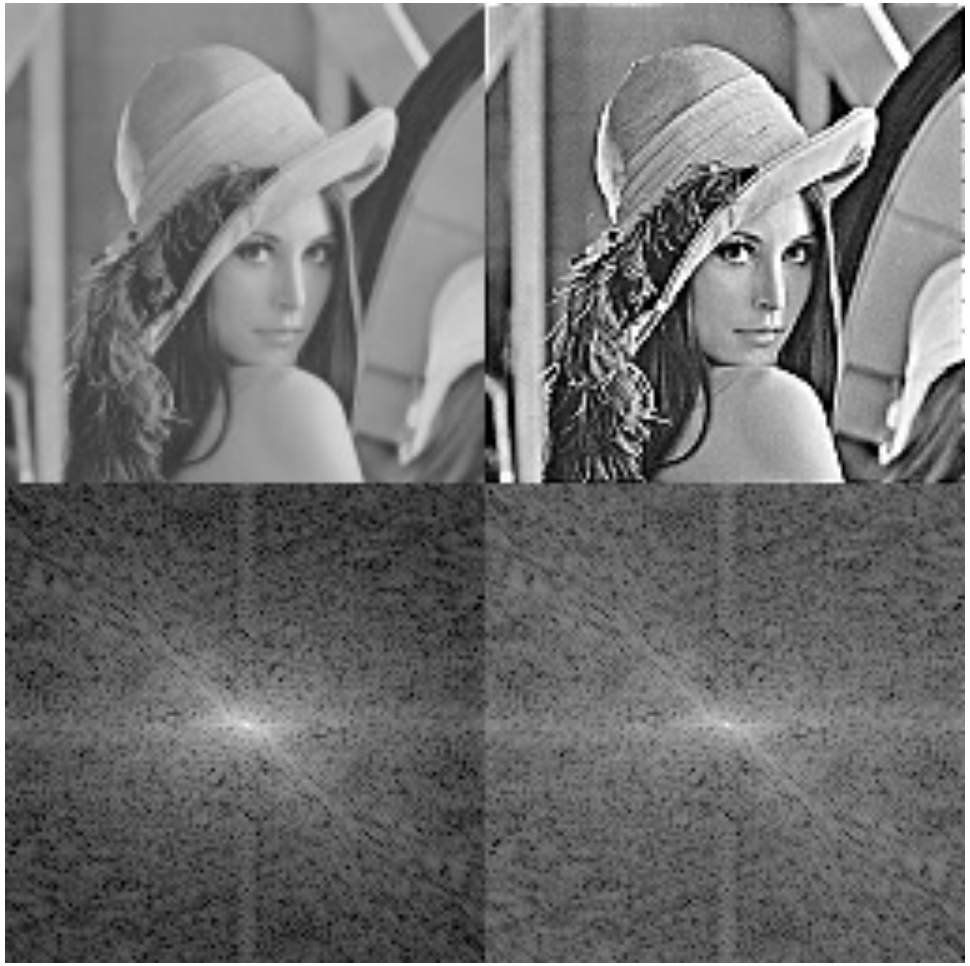


horizontal change (vertical stripes) in image -> frequency response in x-axis

Filtering & Frequency Domain



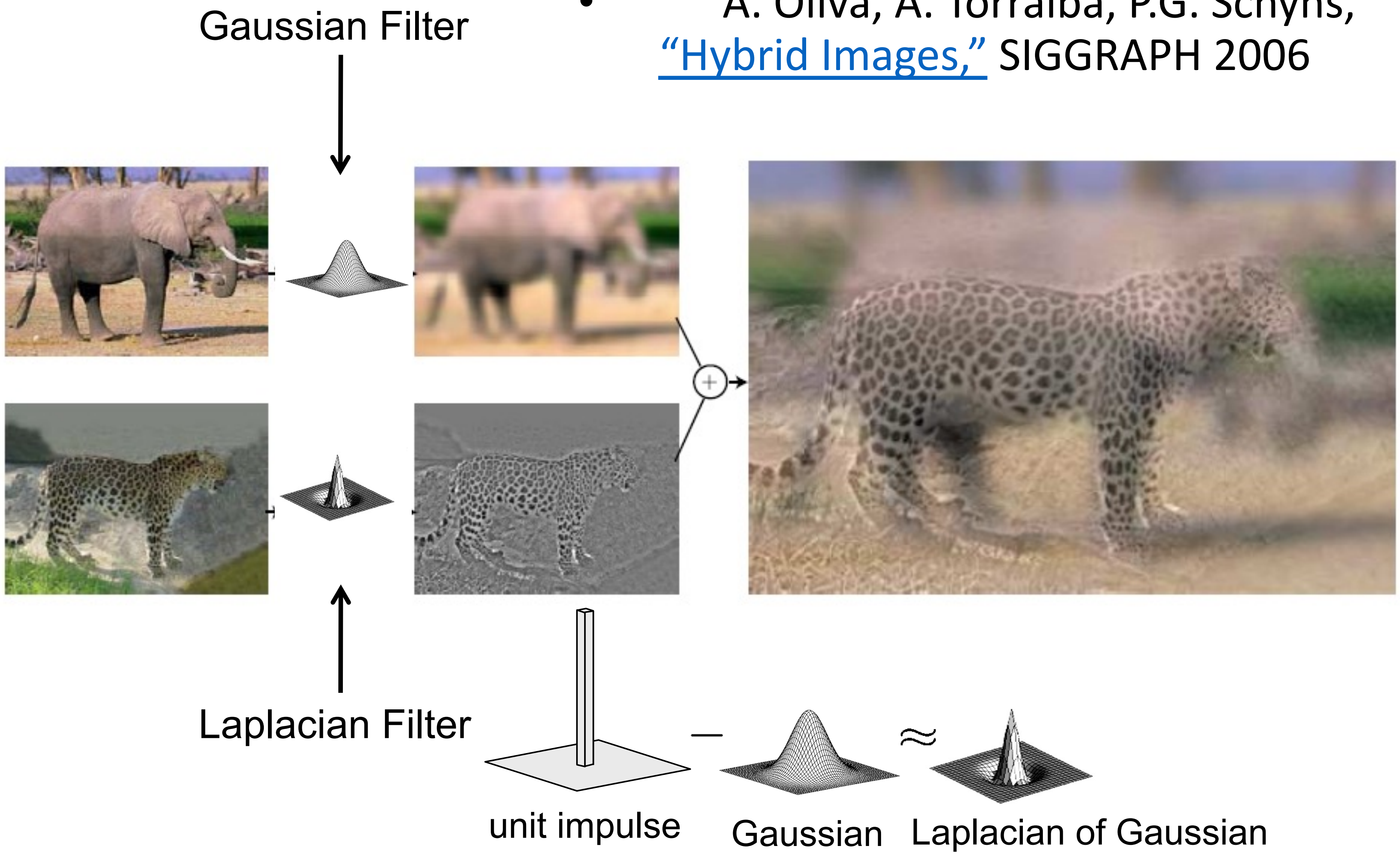
Low pass Filtering with Gaussian Filter



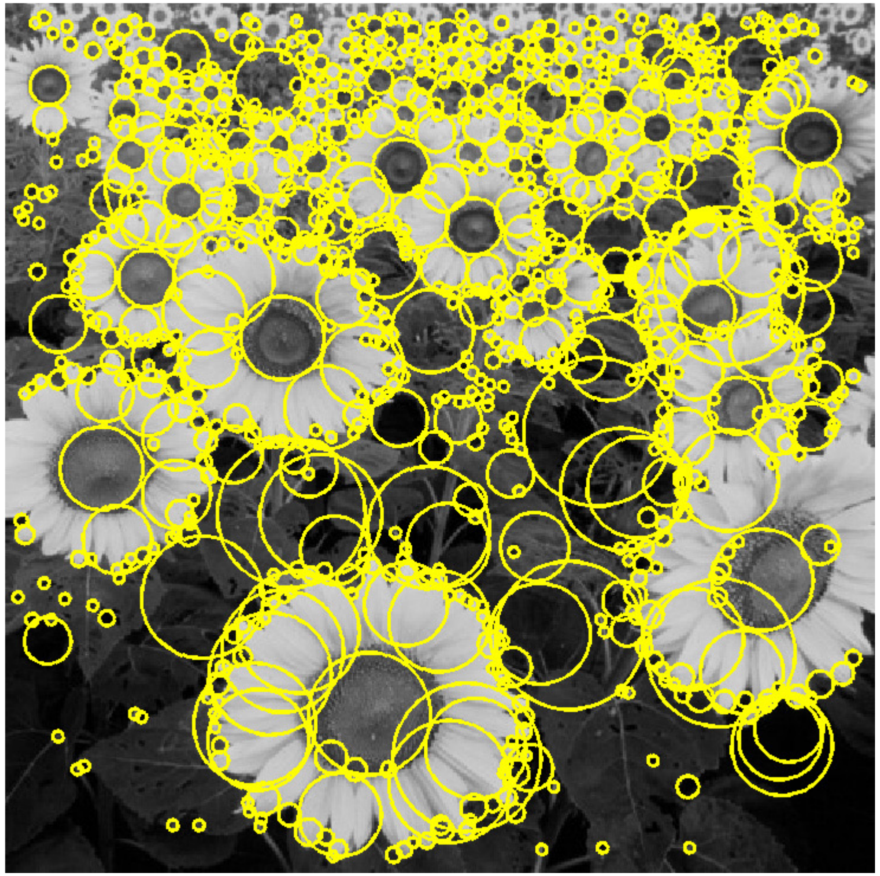
High pass Filtering with Laplacian Filter

Application: Hybrid Images

- A. Oliva, A. Torralba, P.G. Schyns, [“Hybrid Images,”](#) SIGGRAPH 2006



Today: Feature extraction—Corners and blobs



Motivation: Automatic panoramas

Credit: Matt Brown

Motivation: Automatic panoramas

GigaPan:
http://gigapan.com/

Also see Google Zoom Views:
https://www.google.com/culturalinstitute/beta/project/gigapixels

Steps of creating a Panorama
(For this & next week)

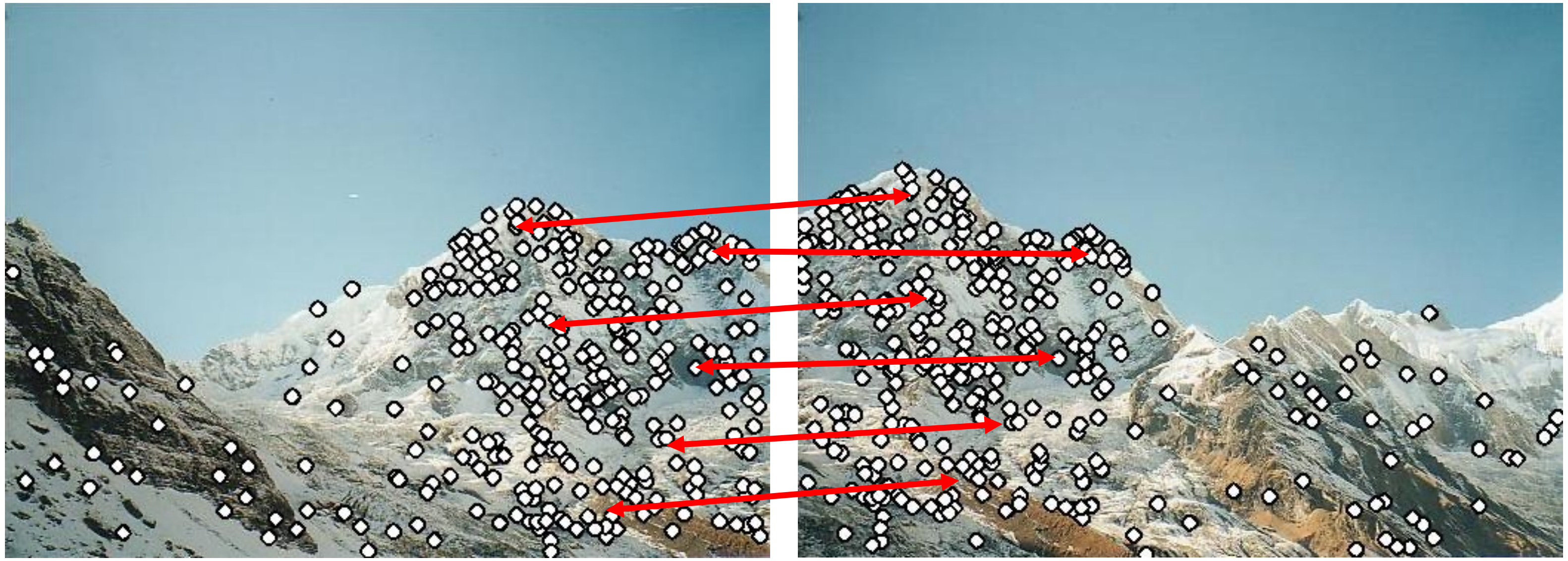
This is your next homework assignment!

Why extract features?

- Motivation: panorama stitching
  - We have two images – how do we combine them?


Why extract features?

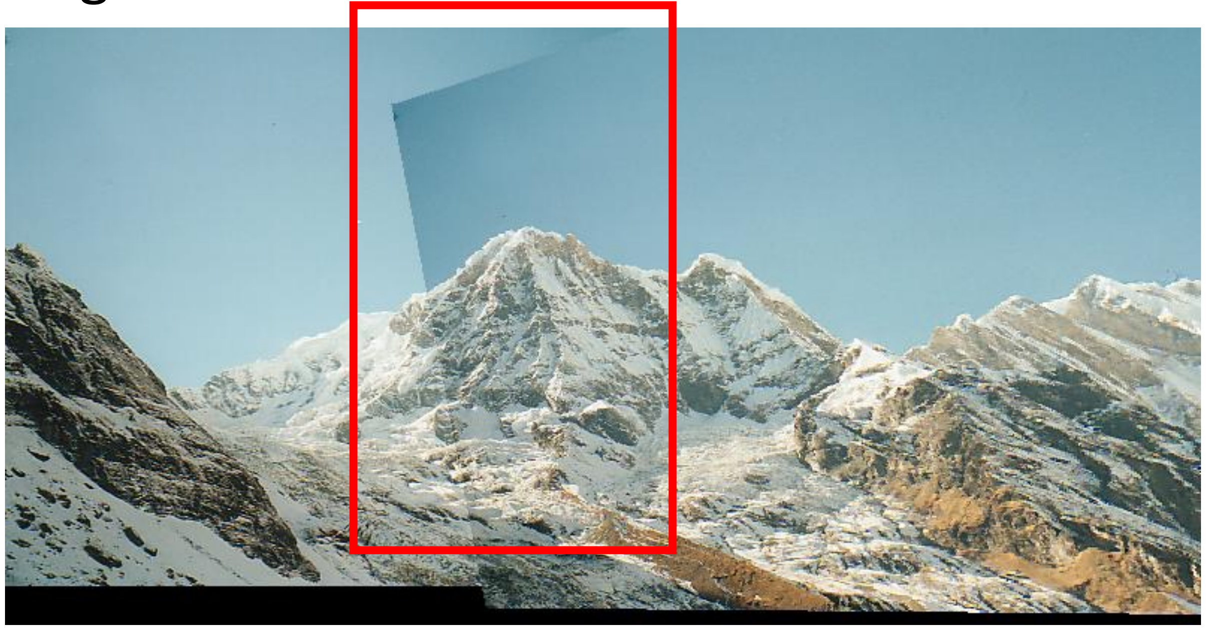
- Motivation: panorama stitching
  - We have two images – how do we combine them?



Step 1: extract features
Step 2: match features

Why extract features?

- Motivation: panorama stitching
  - We have two images – how do we combine them?



Step 1: extract features
Step 2: match features
Step 3: align images
Step 4: blending images

This Week

Next Week

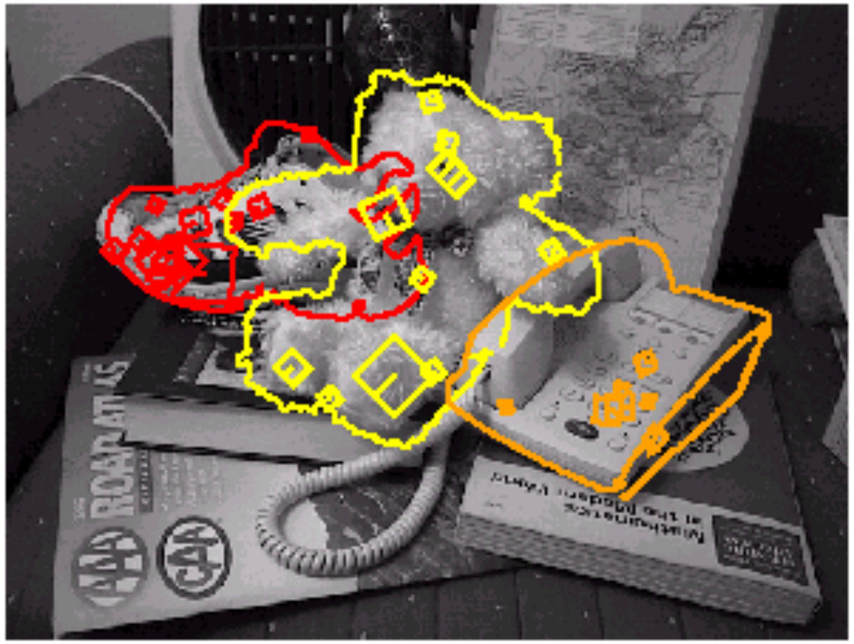
Content: Today’s class

- Why detect features?
- What is a good feature?
- Harris Corner Detector
- Properties of Harris Corner Detector
- Blob Detector

Content: Today's class

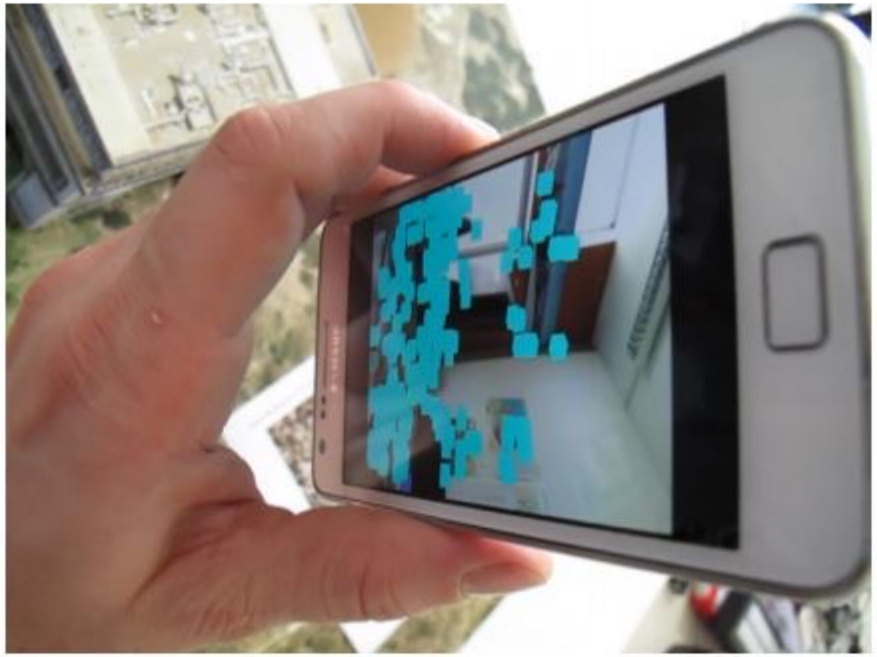
- Why detect features?
- What is a good feature?
- Harris Corner Detector
- Properties of Harris Corner Detector
- Blob Detector

Object recognition (David Lowe)



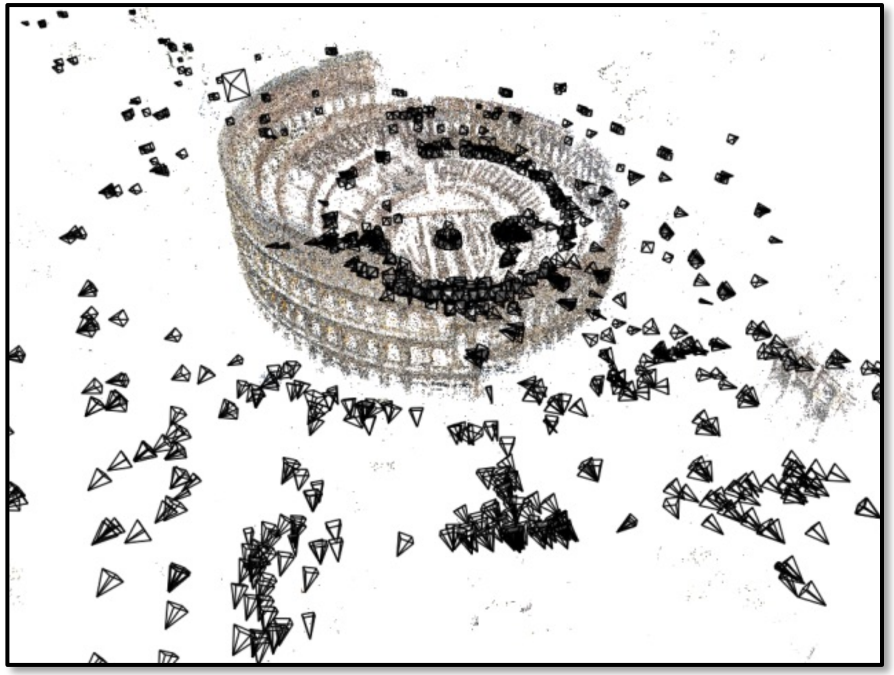
Application: Visual SLAM, Tracking in AR/VR

- (aka Simultaneous Localization and Mapping)



3D Reconstruction

Internet Photos ("Colosseum")



Reconstructed 3D cameras and points

Augmented Reality

Image matching

by Diva Sian

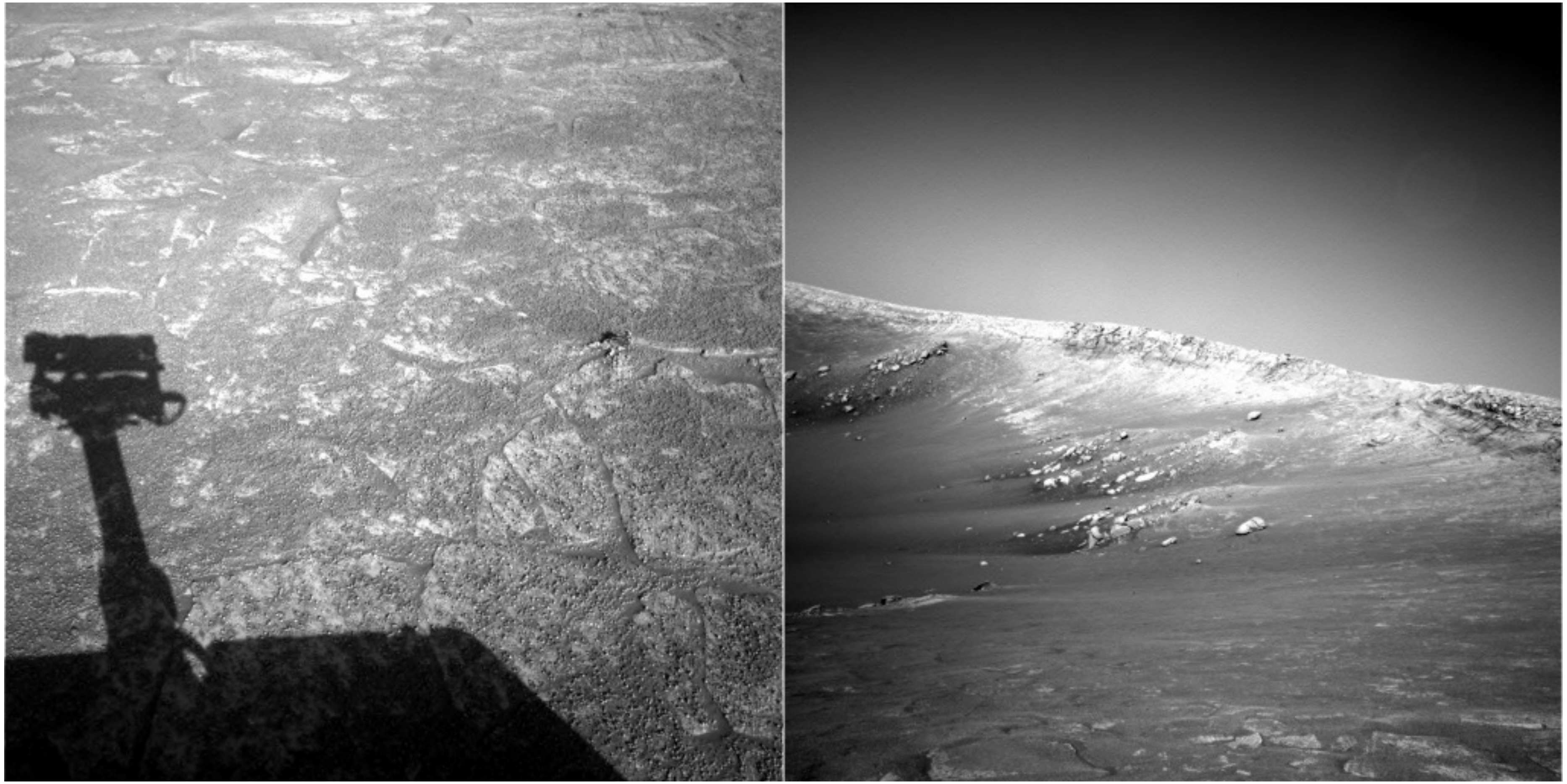
by swashford

Harder case

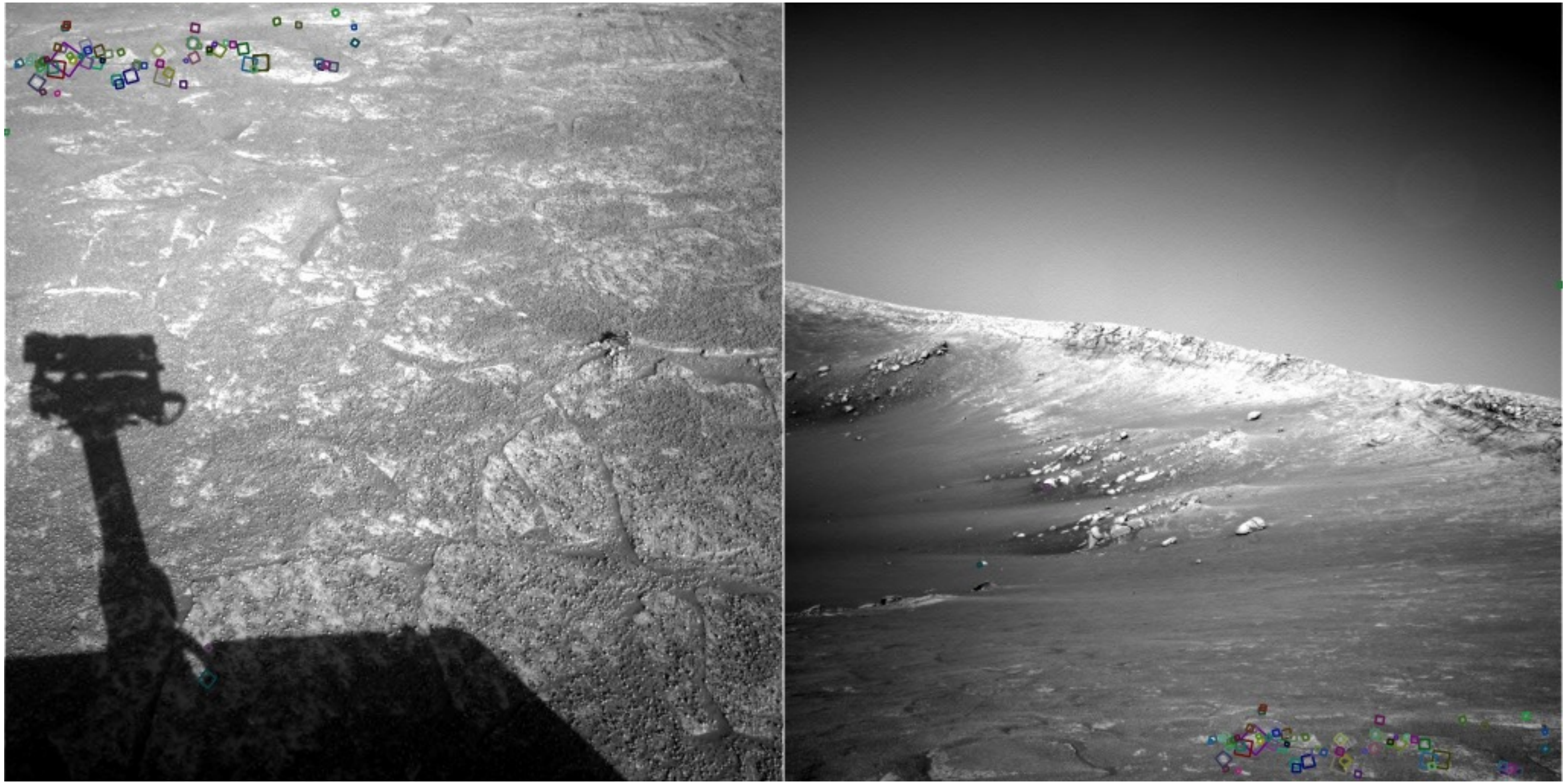
by Diva Sian

by scgbt

Harder still?



Answer below (look for tiny colored squares…)

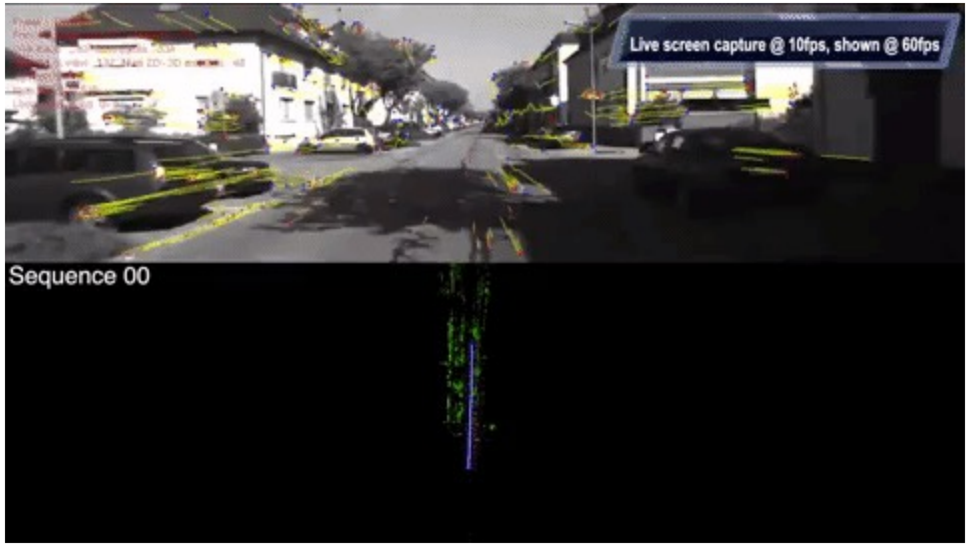


NASA Mars Rover images
with SIFT feature matches

More motivation…

Feature points are used for:

- Image alignment (e.g., mosaics)
- 3D reconstruction
- Motion tracking (e.g. for AR)
- Object recognition
- Image retrieval
- Robot/car navigation
- … other



Content: Today’s class

- Why detect features?
- What is a good feature?
- Harris Corner Detector
- Properties of Harris Corner Detector
- Blob Detector

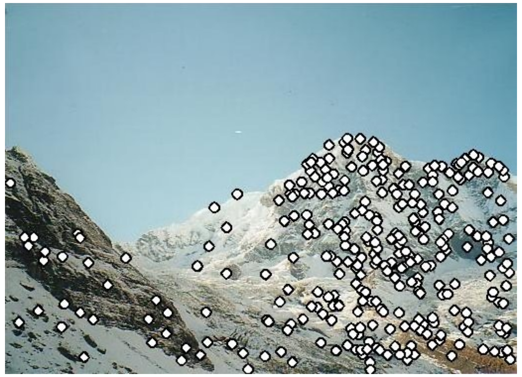
What makes a good feature?

Features = A set of salient keypoints (pixels) in an image

Local features: main components

1) **Detection**: Identify the interest points

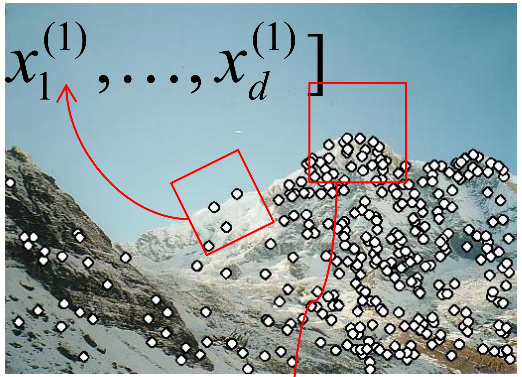
Today



2) **Description**: Extract vector feature descriptor surrounding each interest point

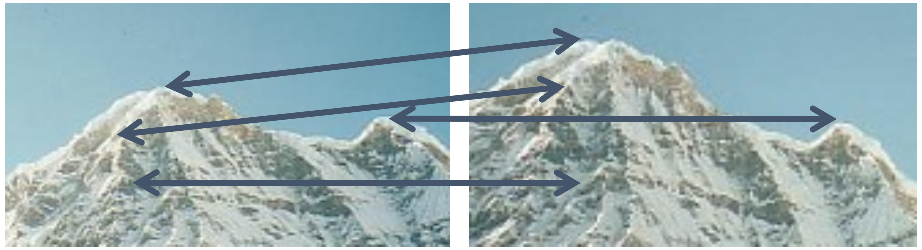
Next Class

$$\mathbf{x}_1 = [x_1^{(1)}, \ldots, x_d^{(1)}]$$



$$\mathbf{x}_2 = [x_1^{(2)}, \ldots, x_d^{(2)}]$$

3) **Matching**: Determine correspondence between descriptors in two views



Credit: Kristen Grauman

Advantages of local features

Locality

- features are local, so robust to occlusion and clutter

Quantity

- hundreds or thousands in a single image

Distinctiveness:

- can differentiate a large database of objects

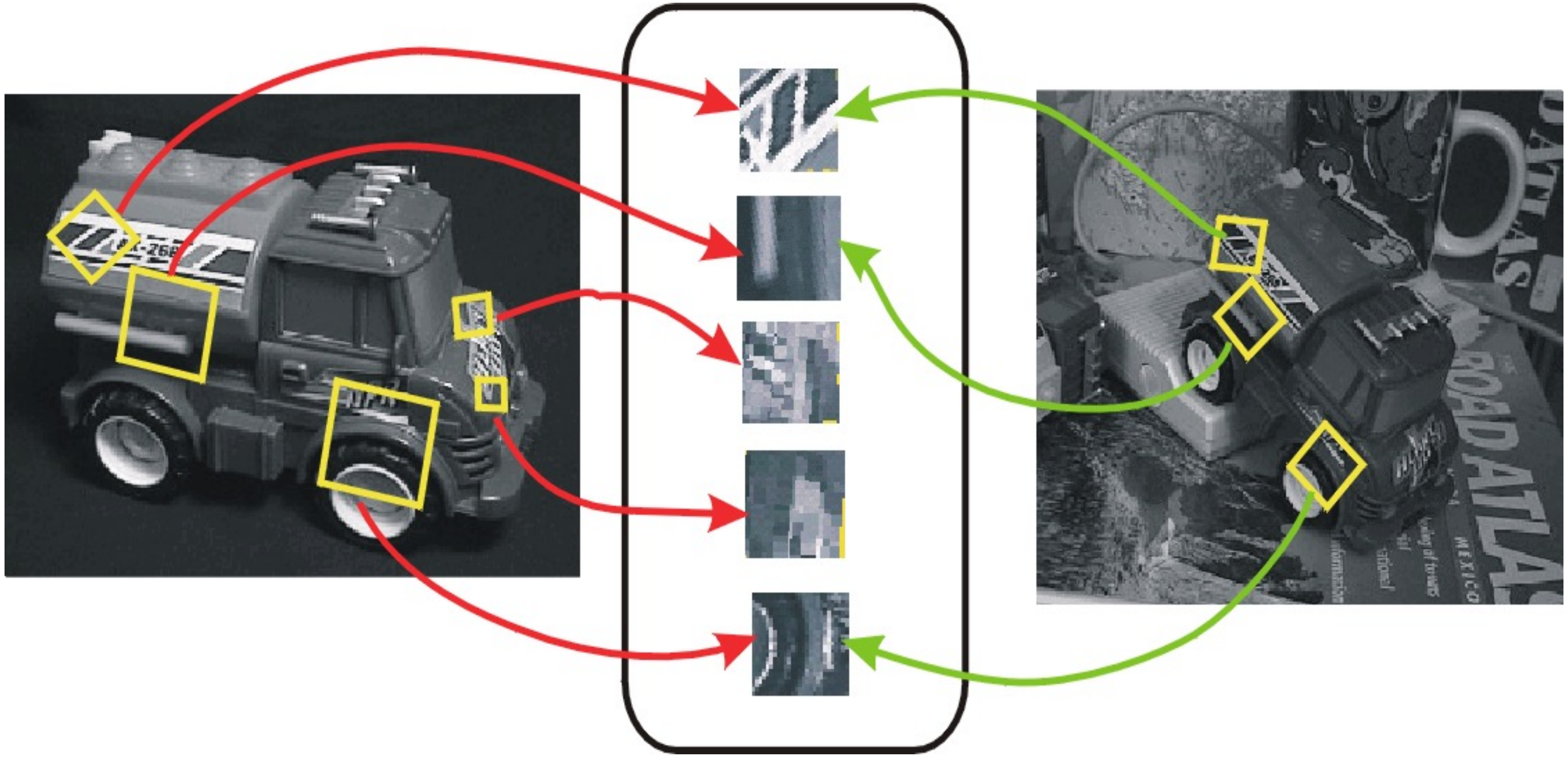
Efficiency

- real-time performance achievable

Invariant local features

Find features that are invariant to transformations

- geometric invariance: translation, rotation, scale
- photometric invariance: brightness, exposure, …



Feature Descriptors

Want uniqueness

Look for image regions that are unusual

- Lead to unambiguous matches in other images

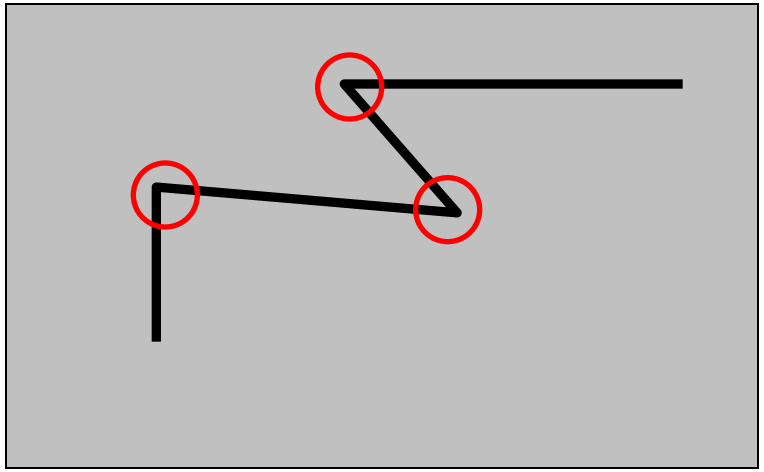
How to define “unusual”?

Content: Today’s class

- Why detect features?
- What is a good feature?
- Harris Corner Detector
- Properties of Harris Corner Detector
- Blob Detector

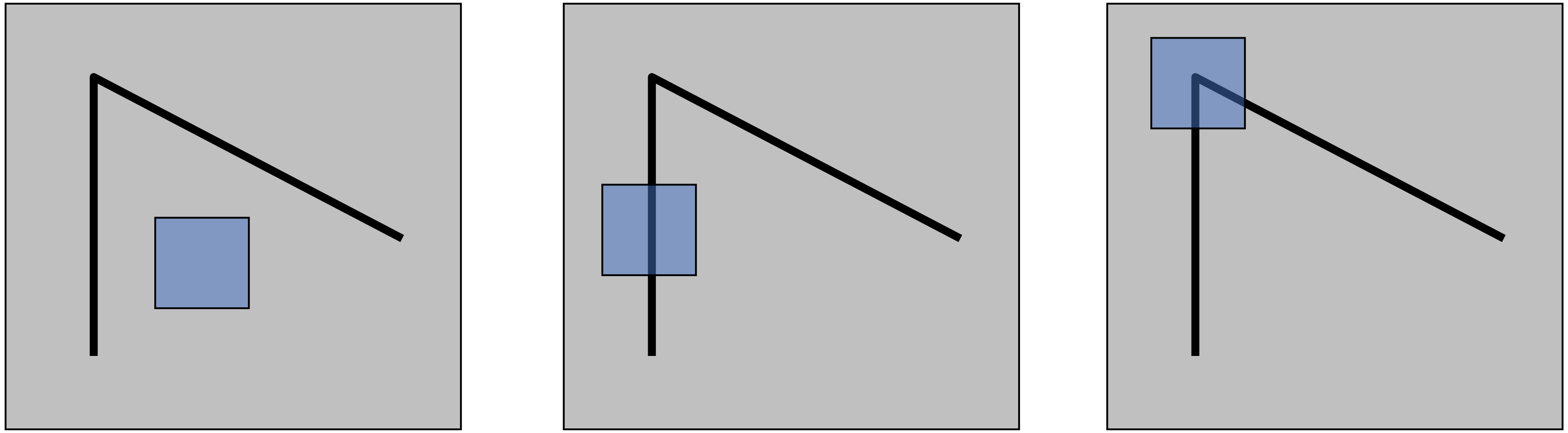
Harris corner detector

- C.Harris, M.Stephens. “A Combined Corner and Edge Detector”. 1988

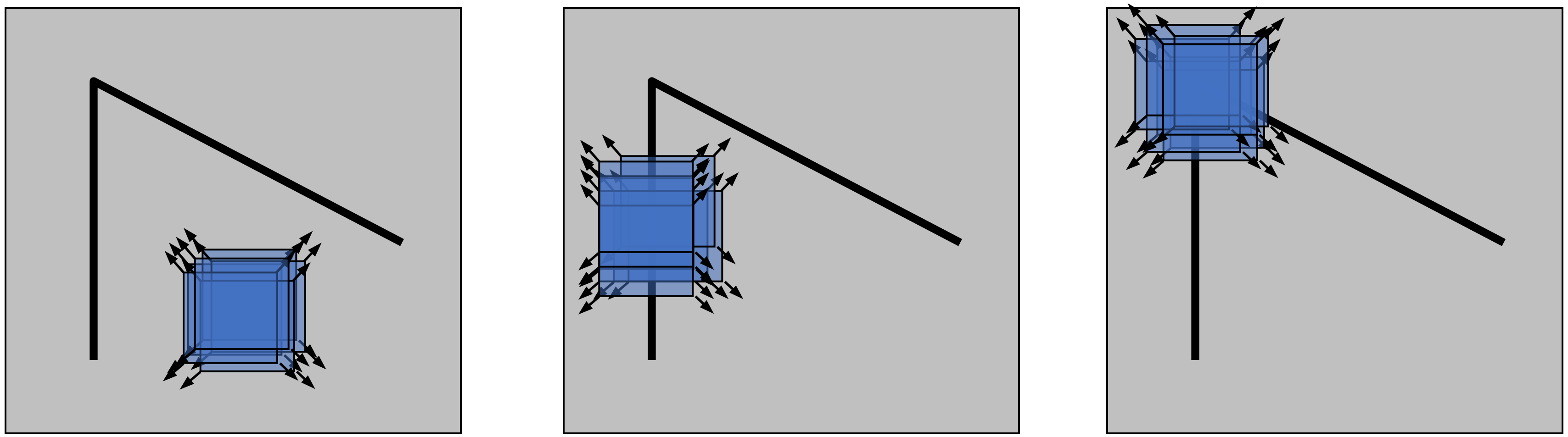


The Basic Idea

- We should easily recognize the point by looking through a small window
- Shifting a window in *any direction* should give *a large change* in intensity



Harris Detector: Basic Idea



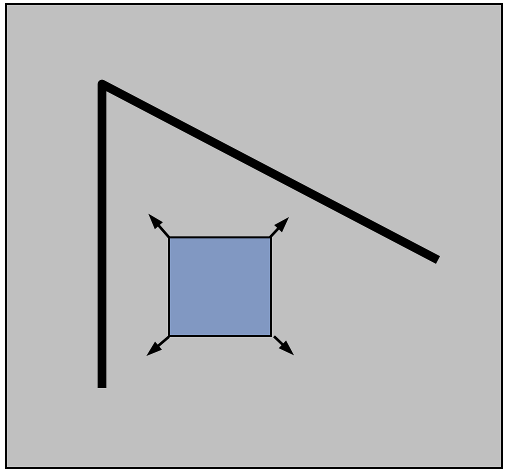
"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

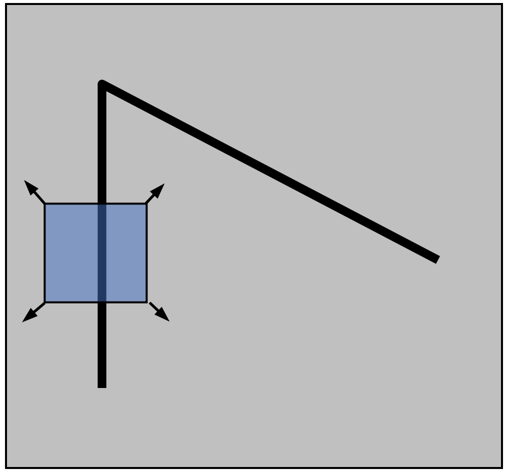
"corner":
significant change
in all directions

Local measures of uniqueness

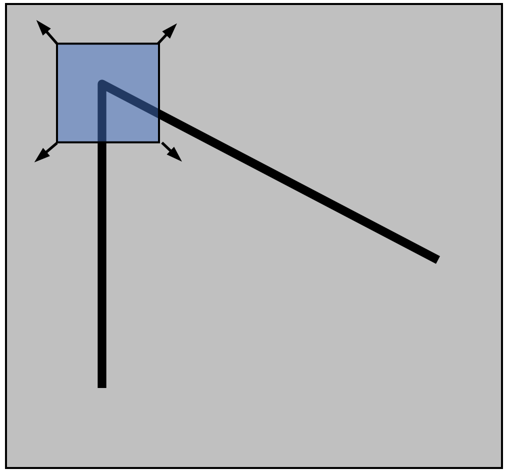
- How does the window change when you shift it?
- Shifting the window in any direction causes a big change



"flat" region:
no change in all directions



"edge":
no change along the edge direction



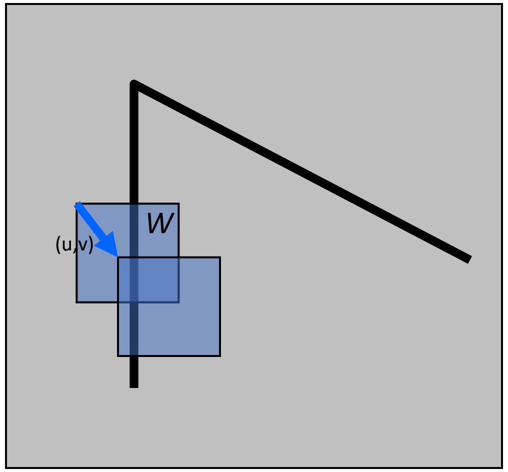
"corner":
significant change in all directions

Credit: S. Seitz, D. Frolova, D. Simakov

Harris corner detection: the math

Consider shifting the window *W* by (*u,v*)

- how do the pixels in W change?
- compare each pixel before and after by summing up the squared differences (SSD)
- this defines an SSD "error" *E*(*u,v*):



$$E(u,v) = \sum_{(x,y)\in W} \left[I(x+u,y+v) - I(x,y)\right]^2$$

- We are happy if this error is high. Corner has very high E(u,v) for all values of (u,v)
- Slow to compute exactly for each pixel and each offset (u,v)

Chris Harris and Mike Stephens (1988). "A Combined Corner and Edge Detector". *Alvey Vision Conference*.

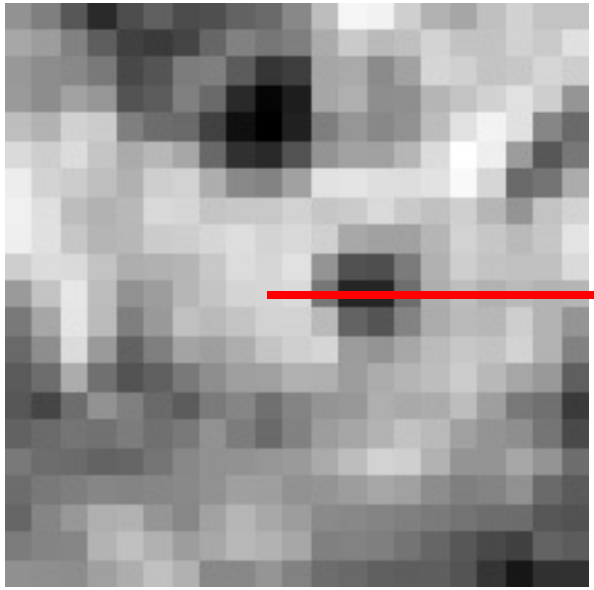
Corner Detection: Mathematics

Change in appearance of window *W* for the shift [*u*,*v*]:

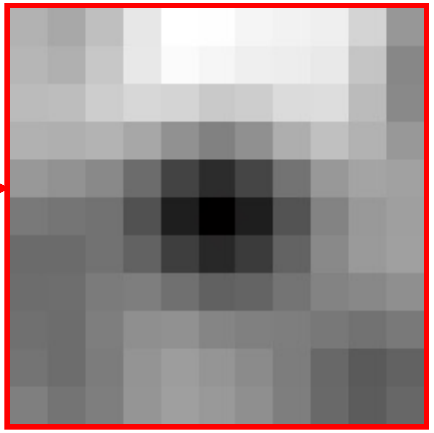
$$E(u,v) = \sum_{(x,y)\in W} [I(x+u, y+v) - I(x,y)]^2$$

We want to find out how this function behaves for small shifts

$I(x, y)$



$E(u, v)$



Compute E(u,v) for every pixel in the image.
Computationally inefficient

Small motion assumption

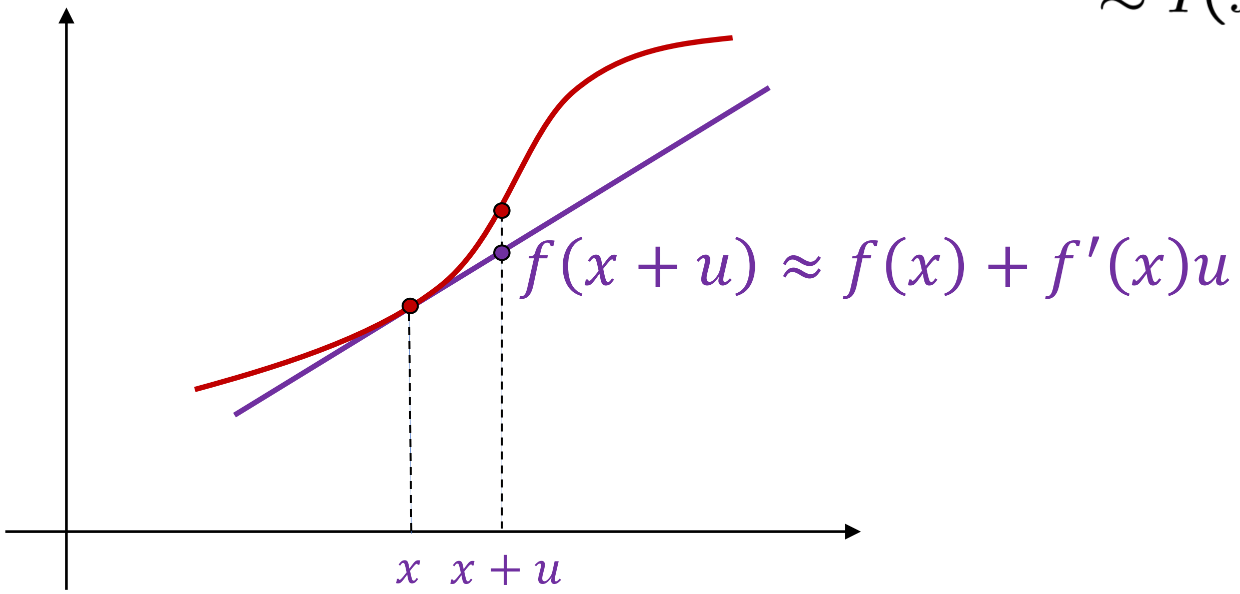
Taylor Series expansion of *I*:

$$I(x+u, y+v) = I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \text{higher order terms}$$

If the motion (*u*, *v*) is small, then first order approximation is good

$$\begin{aligned} I(x+u, y+v) &\approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v \\ &\approx I(x,y) + [I_x \;\; I_y] \begin{bmatrix} u \\ v \end{bmatrix} \end{aligned}$$

shorthand: $I_x = \frac{\partial I}{\partial x}$

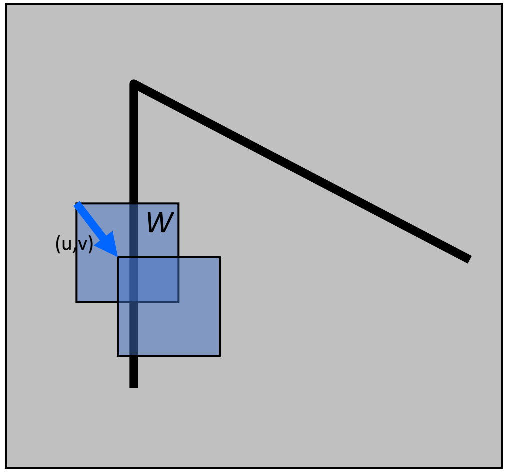


Plugging this into the formula on the previous slide…

Corner detection: the math

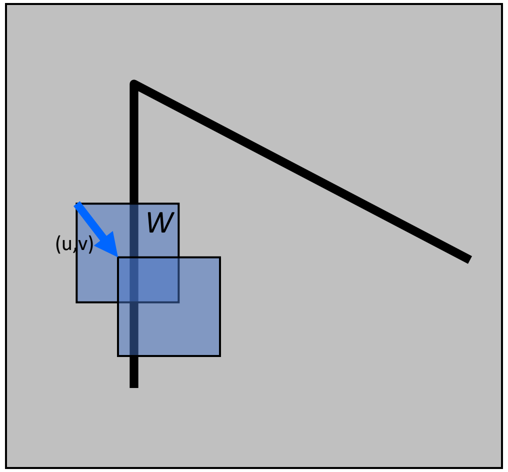
Consider shifting the window *W* by (*u,v*)

- define an SSD "error" *E(u,v)*:



$$
\begin{aligned}
E(u,v) &= \sum_{(x,y)\in W} \left[I(x+u,y+v) - I(x,y)\right]^2 \\
&\approx \sum_{(x,y)\in W} \left[I(x,y) + I_x u + I_y v - I(x,y)\right]^2 \\
&\approx \sum_{(x,y)\in W} \left[I_x u + I_y v\right]^2
\end{aligned}
$$

Corner detection: the math



Consider shifting the window *W* by (*u,v*)

- define an SSD "error" *E(u,v)*:

$$E(u,v) \approx \sum_{(x,y)\in W} [I_x u + I_y v]^2$$

$$\approx Au^2 + 2Buv + Cv^2$$

$$A = \sum_{(x,y)\in W} I_x^2 \qquad B = \sum_{(x,y)\in W} I_x I_y \qquad C = \sum_{(x,y)\in W} I_y^2$$

- Thus, *E*(*u,v*) is locally approximated as a quadratic error function

The second moment matrix

The surface $E(u,v)$ is locally approximated by a quadratic form.

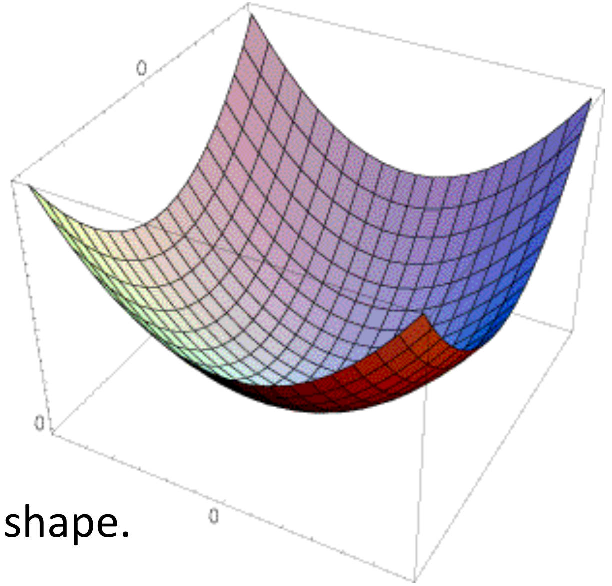
$$E(u,v) \approx Au^2 + 2Buv + Cv^2$$

$$\approx \begin{bmatrix} u & v \end{bmatrix} \underbrace{\begin{bmatrix} A & B \\ B & C \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$A = \sum_{(x,y)\in W} I_x^2$$

$$B = \sum_{(x,y)\in W} I_x I_y$$

$$C = \sum_{(x,y)\in W} I_y^2$$



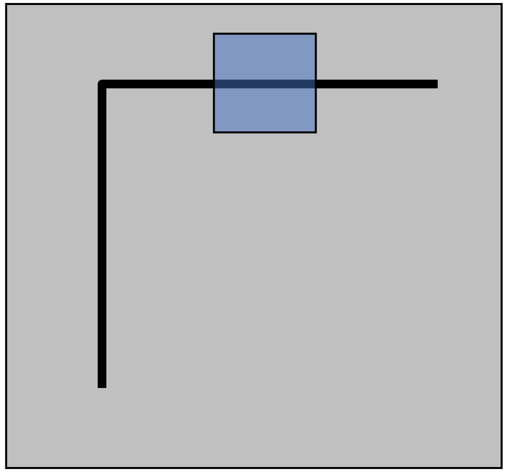
Let's try to understand its shape.

$$E(u,v) \approx \begin{bmatrix} u & v \end{bmatrix} \underbrace{\begin{bmatrix} A & B \\ B & C \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$A = \sum_{(x,y)\in W} I_x^2$$

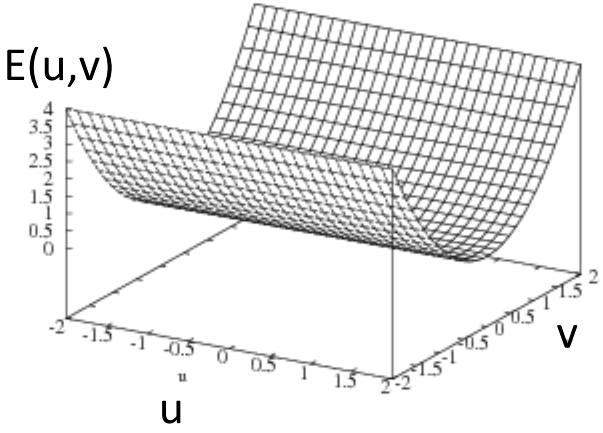
$$B = \sum_{(x,y)\in W} I_x I_y$$

$$C = \sum_{(x,y)\in W} I_y^2$$



Horizontal edge: $I_x = 0$

$$H = \begin{bmatrix} 0 & 0 \\ 0 & C \end{bmatrix}$$

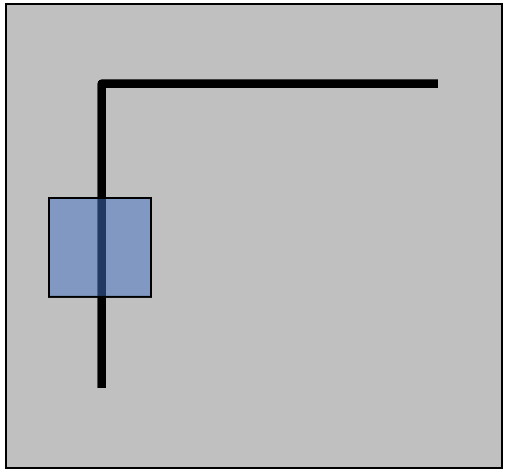


$$E(u,v) \approx \begin{bmatrix} u & v \end{bmatrix} \underbrace{\begin{bmatrix} A & B \\ B & C \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$A = \sum_{(x,y)\in W} I_x^2$$

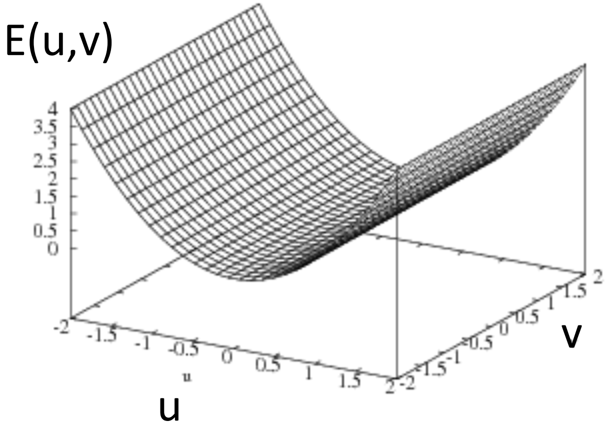
$$B = \sum_{(x,y)\in W} I_x I_y$$

$$C = \sum_{(x,y)\in W} I_y^2$$



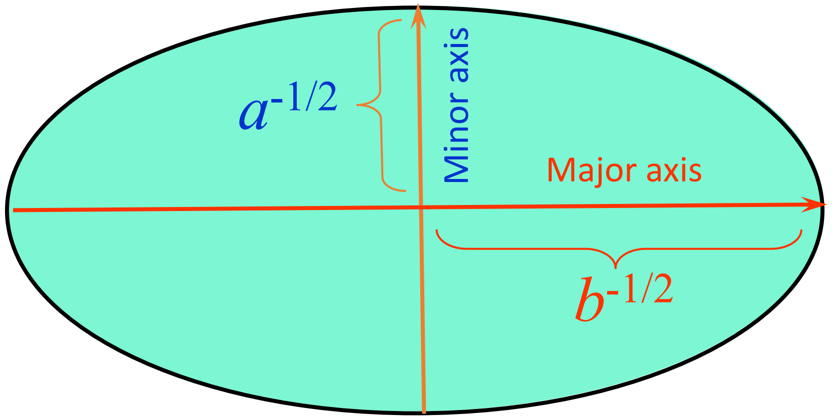
Vertical edge: $I_y = 0$

$$H = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$$



Interpreting the second moment matrix

- Consider the axis-aligned case (gradients are either horizontal or vertical):
- $(u \quad v)\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = 1$
- $au^2 + bv^2 = 1$
- $\frac{u^2}{\left(a^{-1/2}\right)^2} + \frac{v^2}{\left(b^{-1/2}\right)^2} = 1$

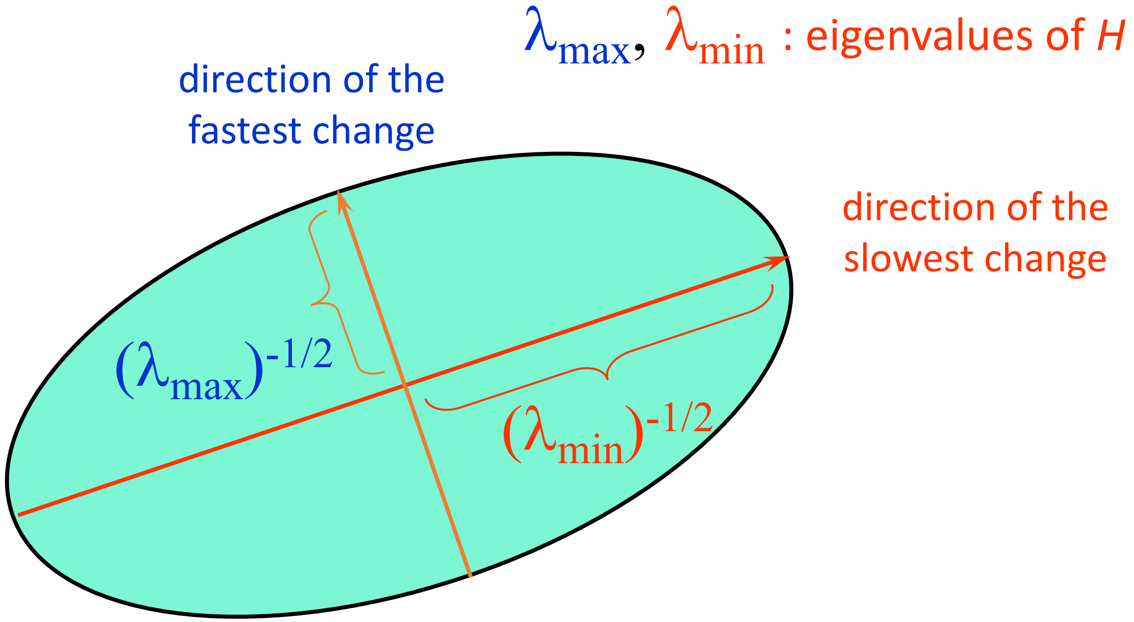


General case

We can visualize *H* as an ellipse with axis lengths determined by the *eigenvalues* of *H* and orientation determined by the *eigenvectors* of *H*

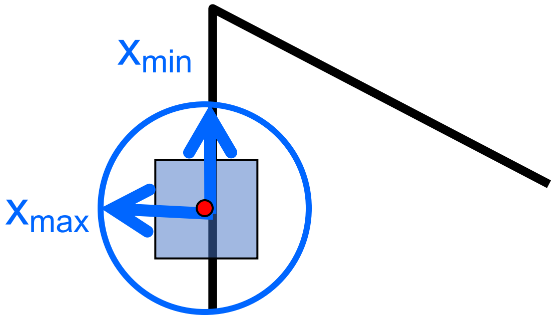
Ellipse equation:

$$[u \;\; v] \; H \begin{bmatrix} u \\ v \end{bmatrix} = \text{const}$$



Corner detection: the math

$$E(u,v) \approx \begin{bmatrix} u & v \end{bmatrix} \underbrace{\begin{bmatrix} A & B \\ B & C \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$



$$Hx_{\max} = \lambda_{\max} x_{\max}$$

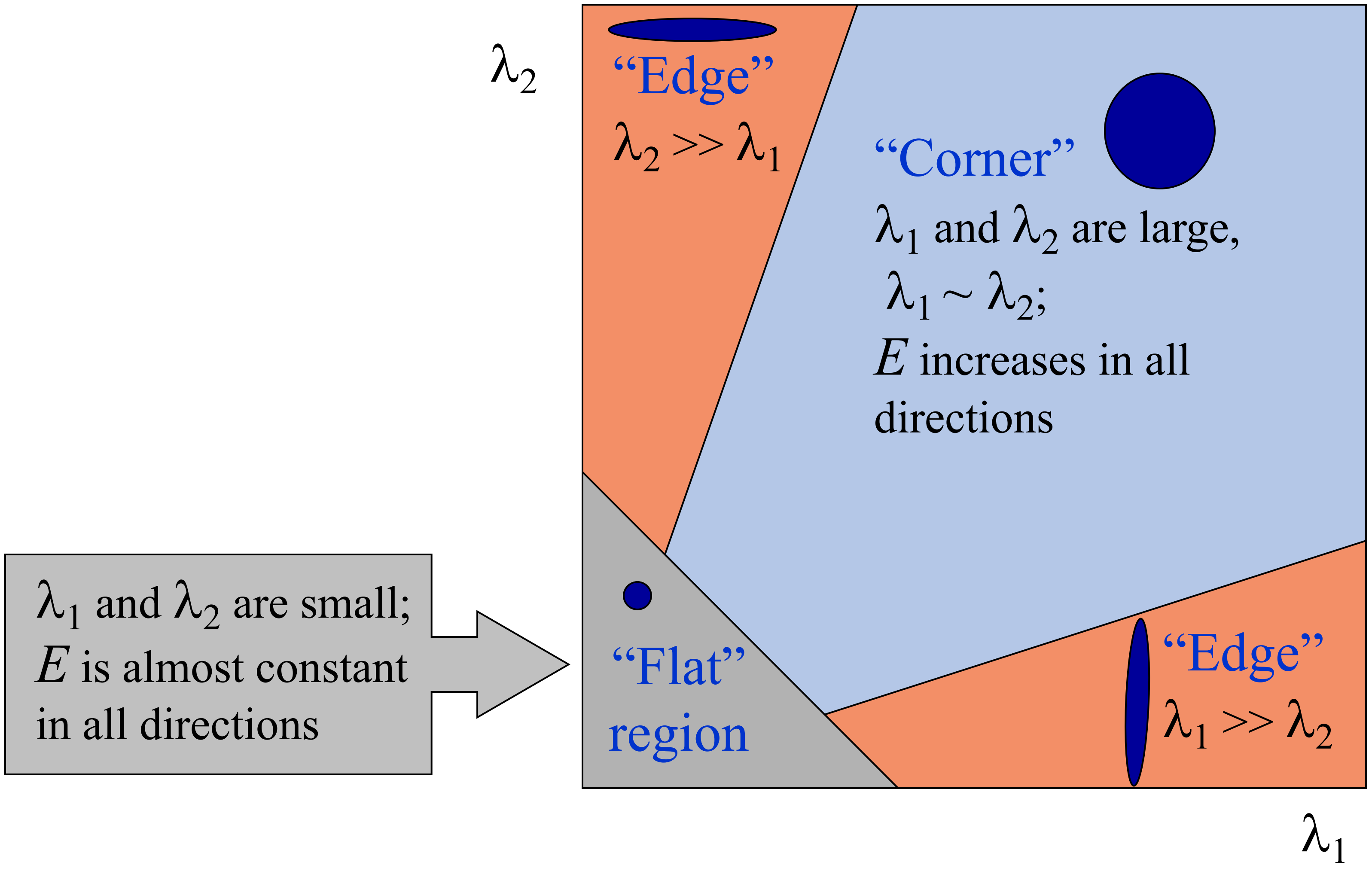
$$Hx_{\min} = \lambda_{\min} x_{\min}$$

Eigenvalues and eigenvectors of H

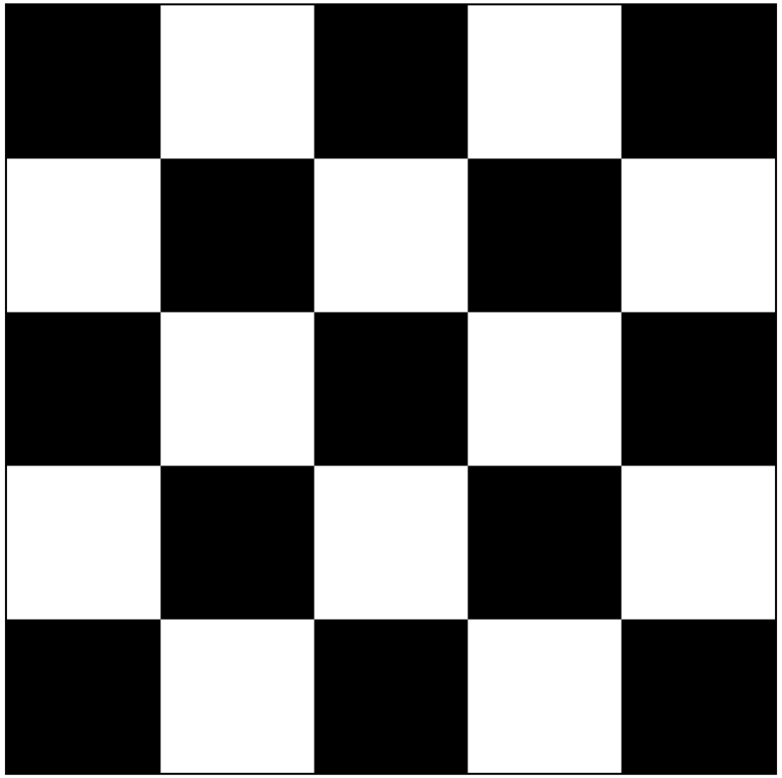
- Define shift directions with the smallest and largest change in error
- x_{max} = direction of largest increase in *E*
- λ_{max} = amount of increase in direction x_{max}
- x_{min} = direction of smallest increase in *E*
- λ_{min} = amount of increase in direction x_{min}

Interpreting the eigenvalues

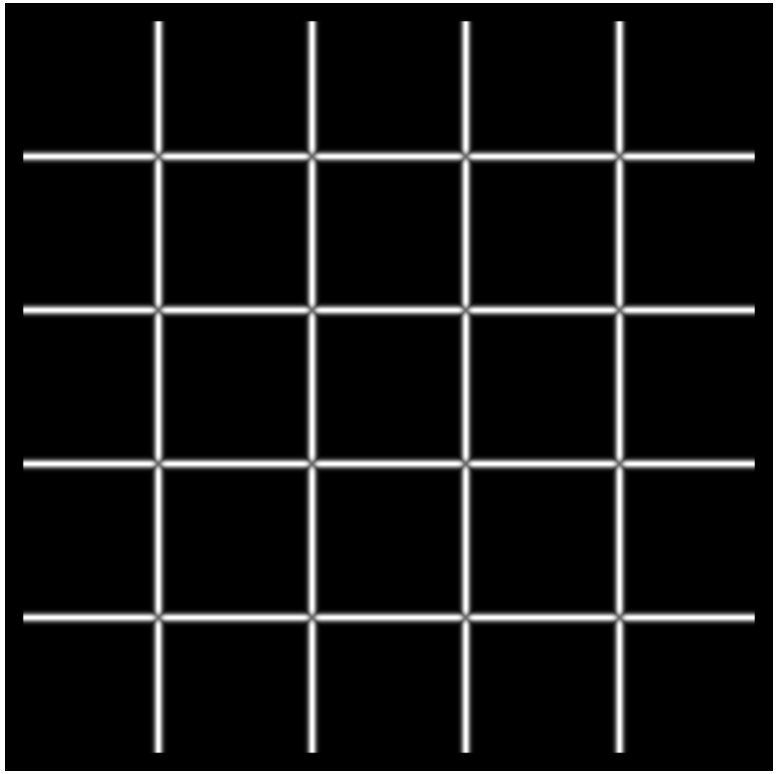
Classification of image points using eigenvalues of *M*:



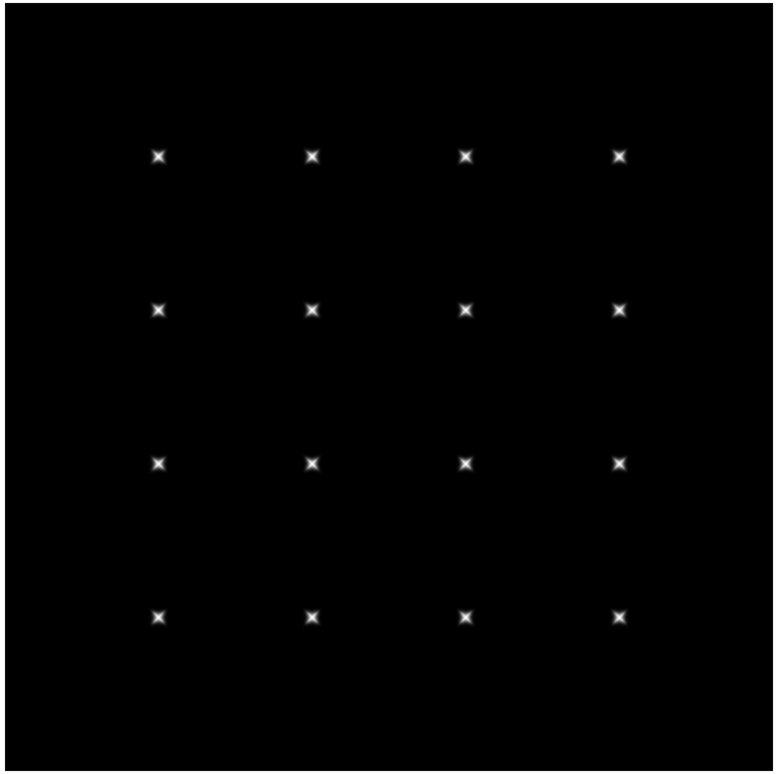
Visualizing Corner detection



I



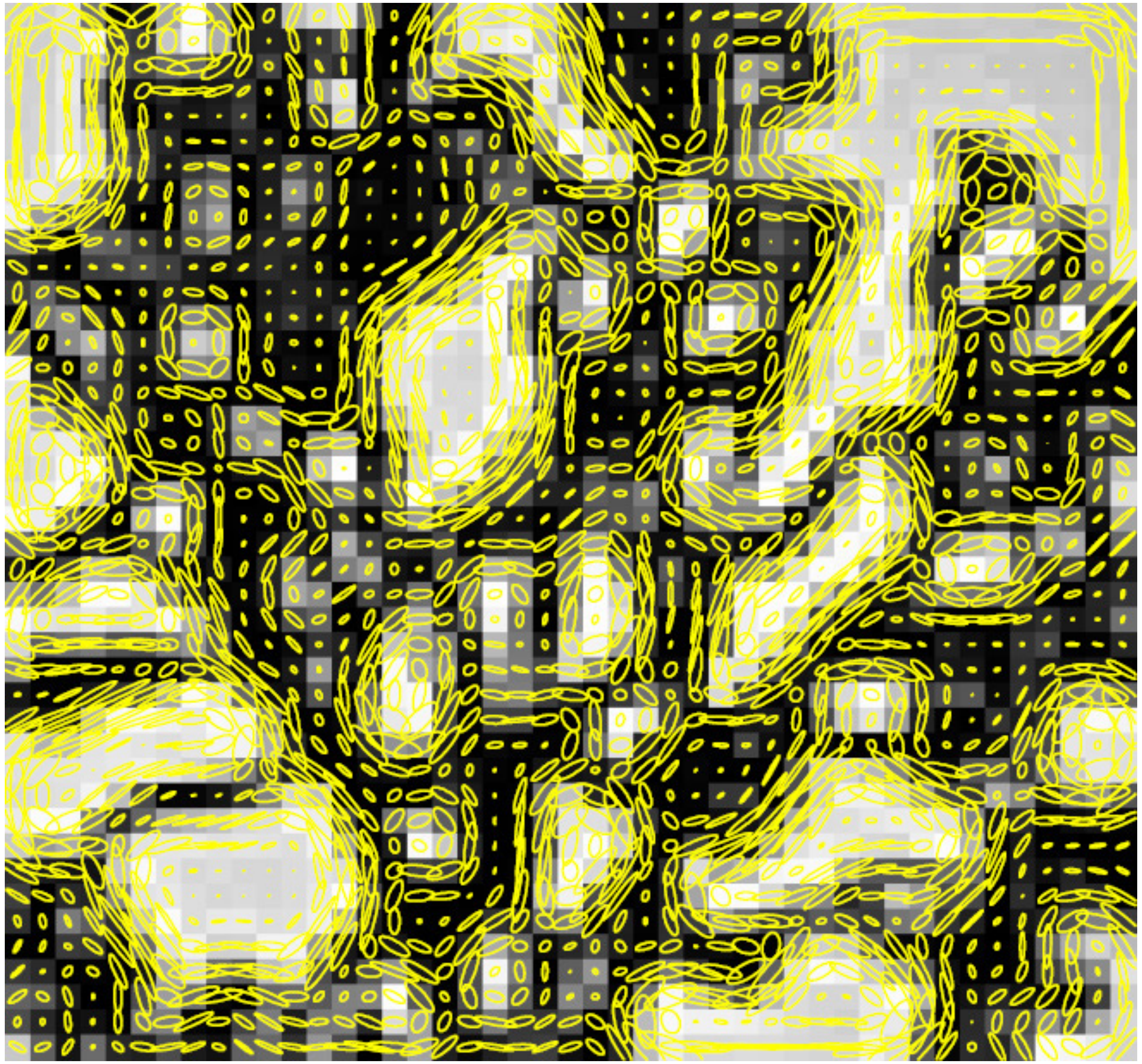
$\lambda_{\max}$



$\lambda_{\min}$

Visualization of second moment matrices

Visualization of second moment matrices



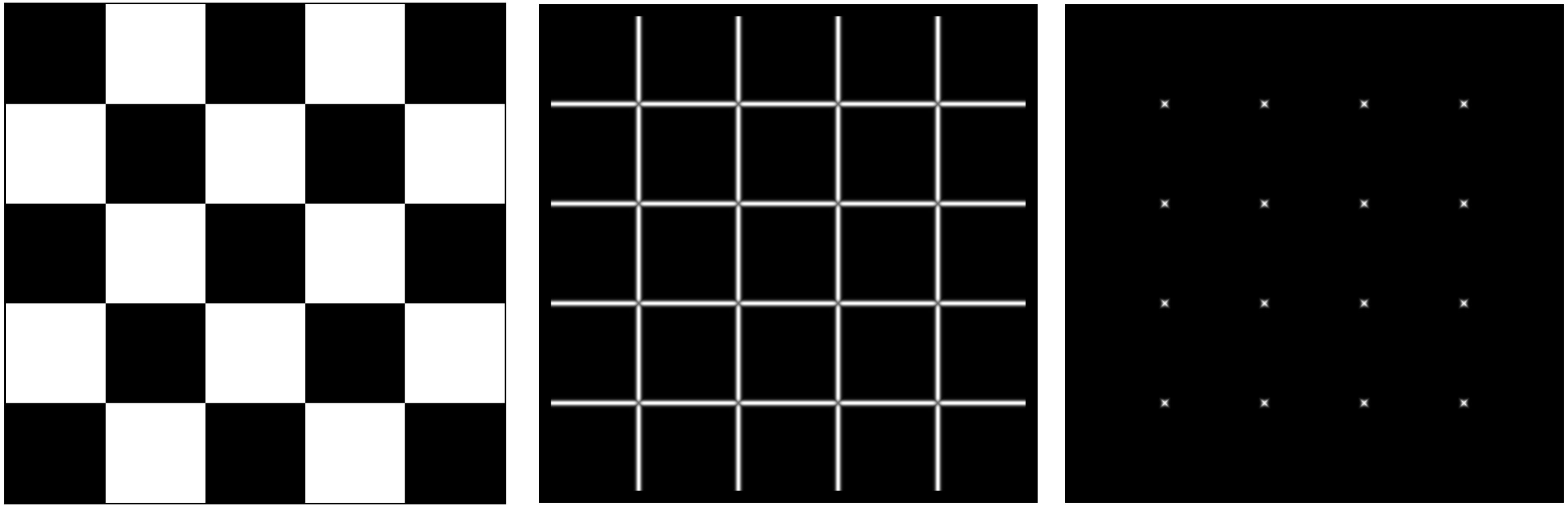
Note: axes are rescaled so ellipse areas are proportional to edge energy (i.e., bigger ellipses correspond to stronger edges)

Corner detection summary

Here's what you do:

- Compute the gradient at each point in the image
- For each pixel:
  - Create the *H* matrix from nearby gradient values
  - Compute the eigenvalues.
  - Find points with large response (λ_{min} > threshold)
- Choose those points where λ_{min} is a local maximum as features

$$H = \sum_{(x,y)\in W} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

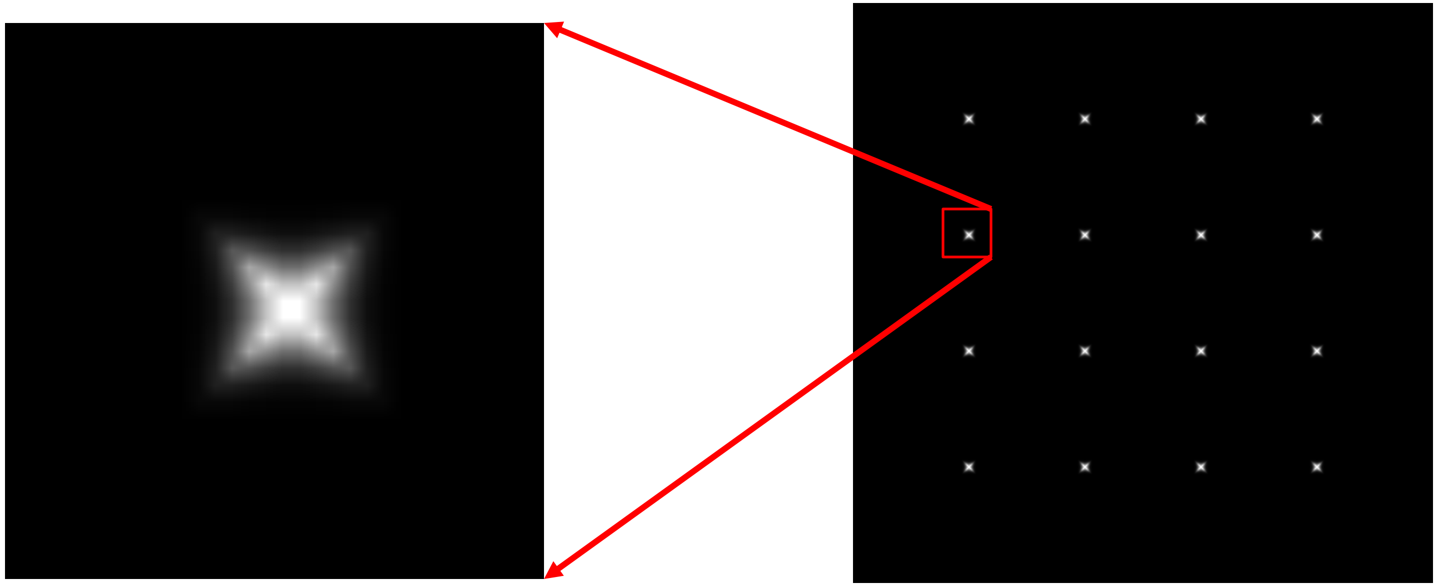


I $\qquad$ $\lambda_{\max}$ $\qquad$ $\lambda_{\min}$

Corner detection summary

Here's what you do:

- Compute the gradient at each point in the image
- For each pixel:
  - Create the *H* matrix from nearby gradient values
  - Compute the eigenvalues.
  - Find points with large response (λ_{min} > threshold)
- Choose those points where λ_{min} is a local maximum as features



$\lambda_{\min}$

The Harris operator

λ_{min} is a variant of the "Harris operator" for feature detection

$$f = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$

$$= \frac{determinant(H)}{trace(H)}$$

- The *trace* is the sum of the diagonals, i.e., $trace(H) = h_{11} + h_{22}$
- Very similar to λ_{min} but less expensive (no square root)
- Called the *Harris Corner Detector* or *Harris Operator*
- Lots of other detectors, this is one of the most popular

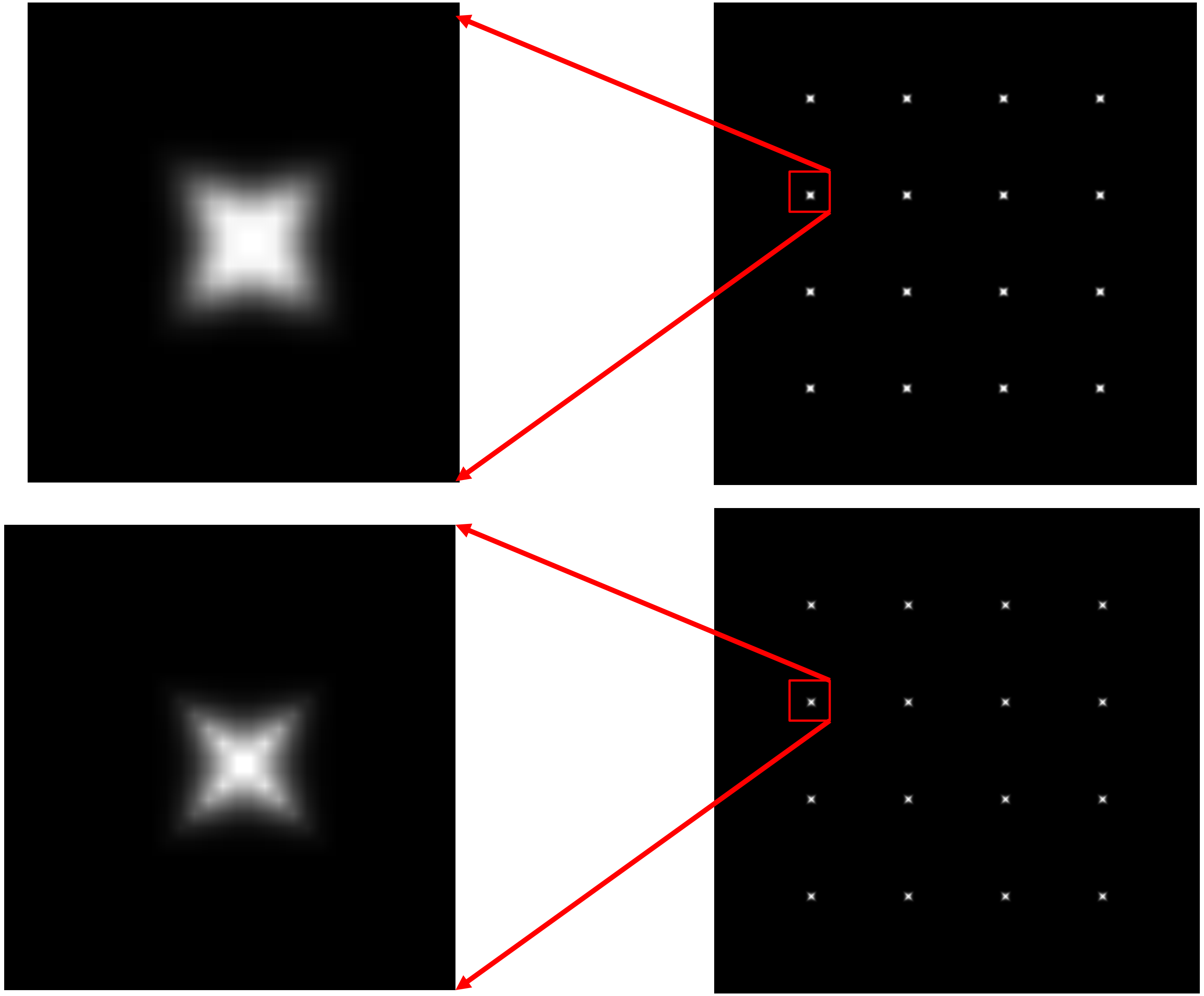
Alternate Version of Harris Detector $R = \lambda_1 \lambda_2 - k \cdot (\lambda_1 + \lambda_2)^2 = \det(M) - k \cdot \operatorname{tr}(M)^2$ $M = H$

Harris detector: Steps

1. Compute Gaussian derivatives at each pixel
2. Compute second moment matrix H in a Gaussian window around each pixel
3. Compute corner response function f or R
4. Threshold f or R
5. Find local maxima of response function (nonmaximum suppression)

C.Harris and M.Stephens. "A Combined Corner and Edge Detector." *Proceedings of the 4th Alvey Vision Conference*: pages 147—151, 1988.

The Harris operator

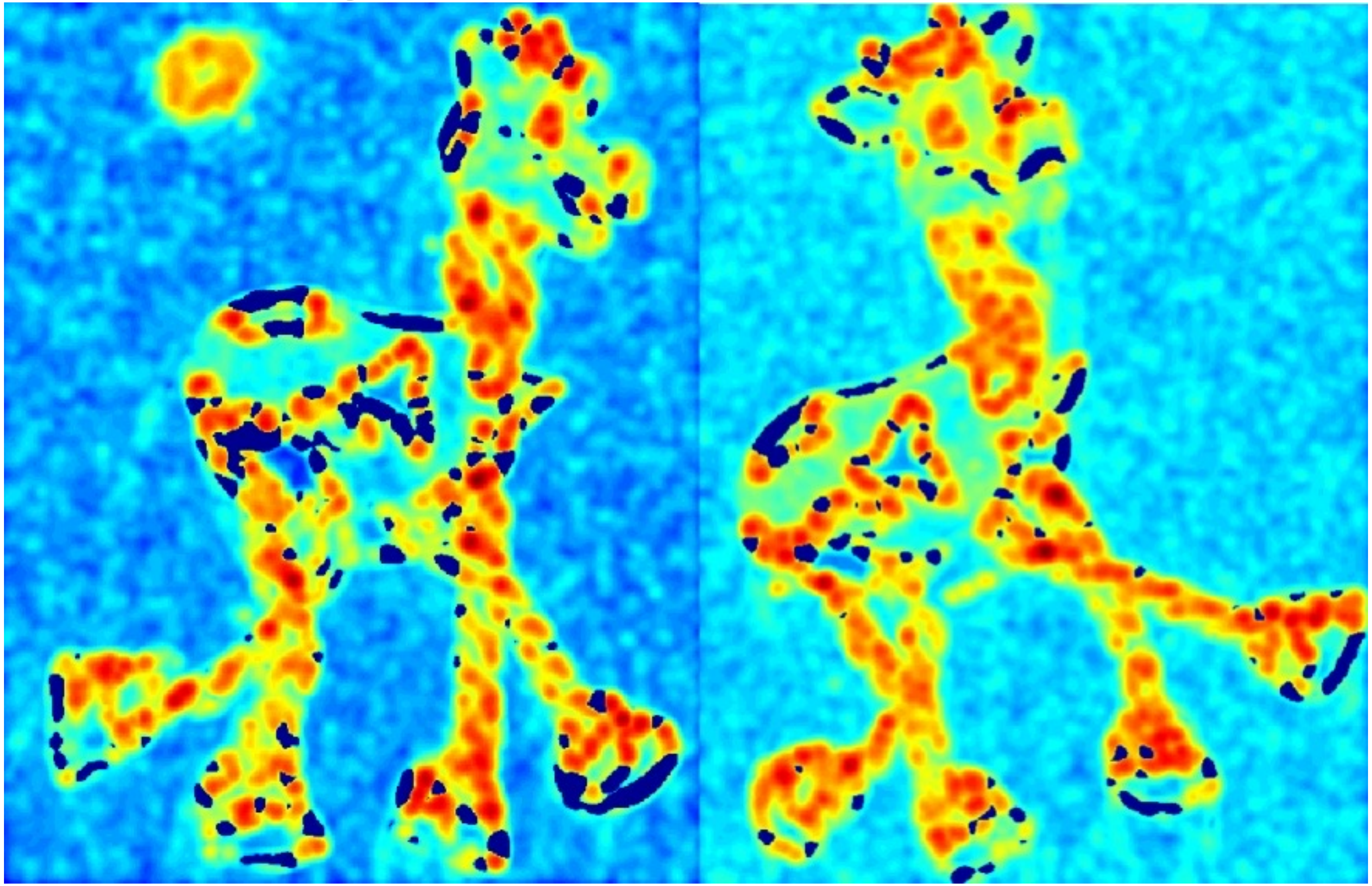


Harris operator

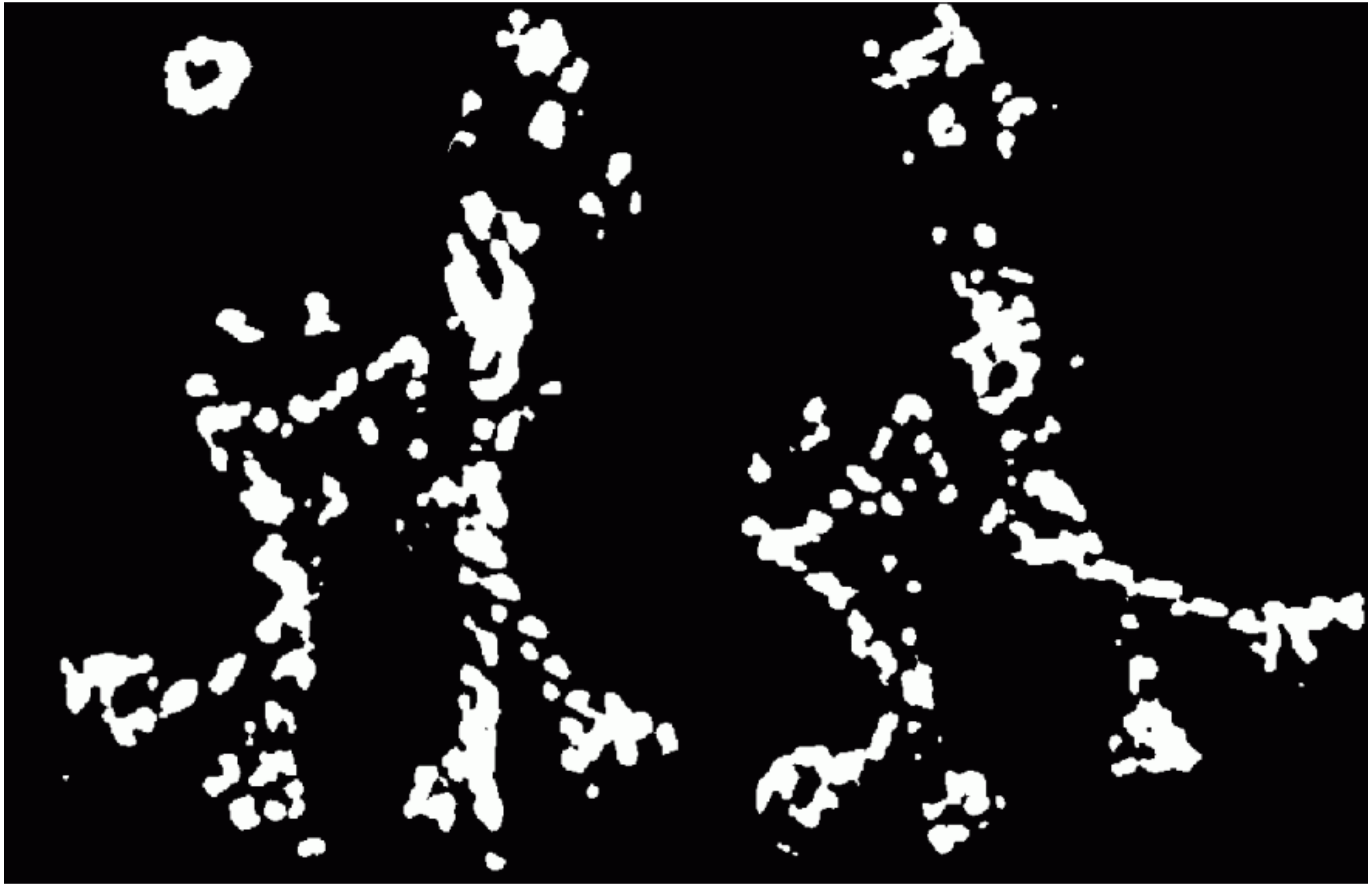
$\lambda_{\min}$

Harris detector example

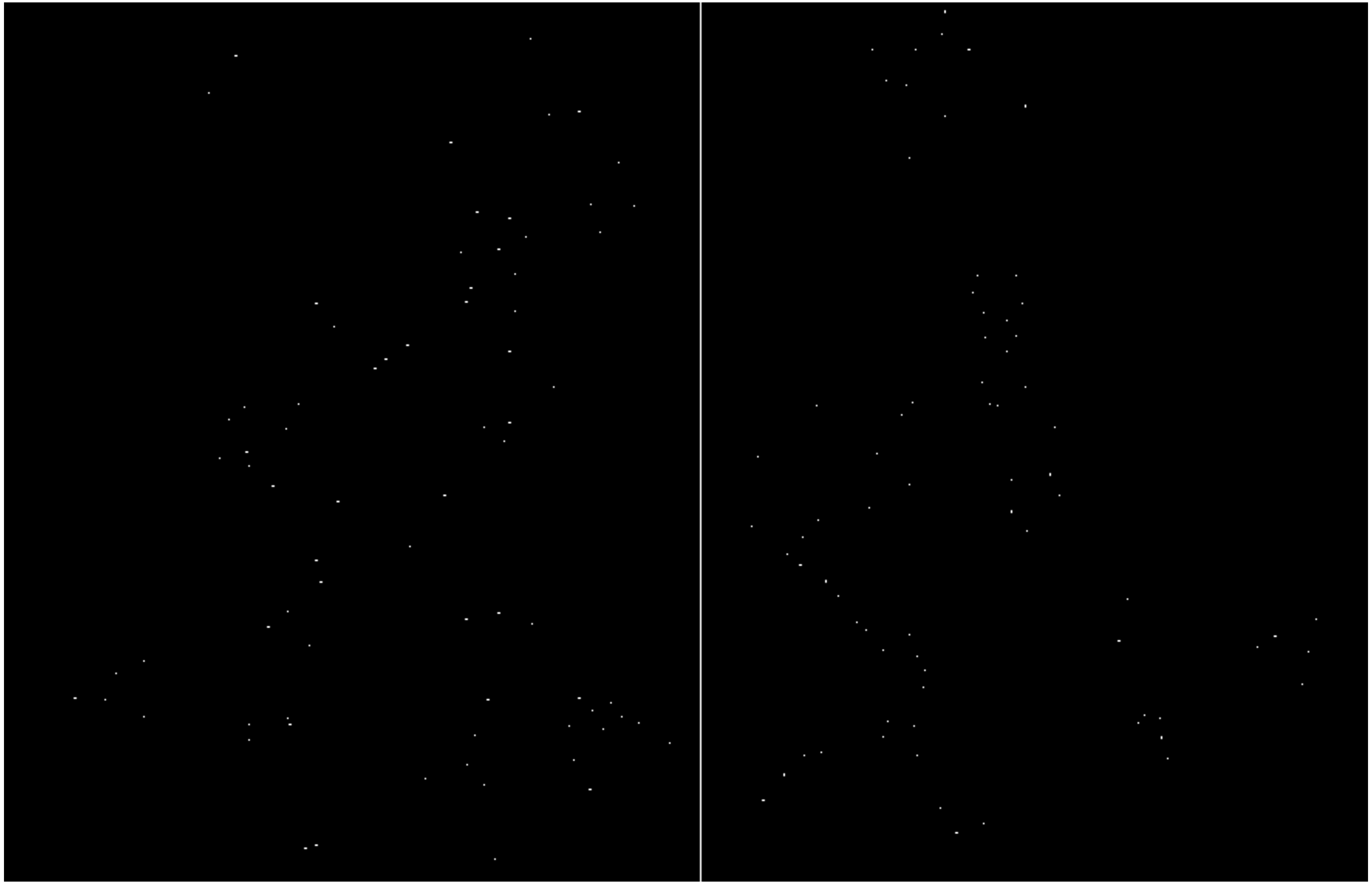
f value (red high, blue low)



Threshold (f > value)



Find local maxima of f (non-max suppression)



Harris features (in red)

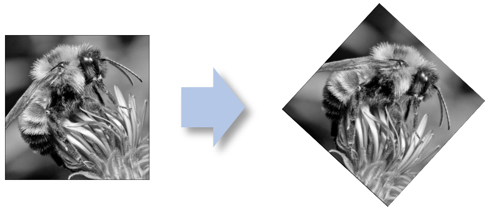
Content: Today’s class

- Why detect features?
- What is a good feature?
- Harris Corner Detector
- Properties of Harris Corner Detector
- Blob Detector

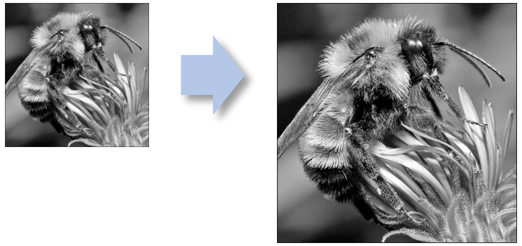
Image transformations

- Geometric

Rotation

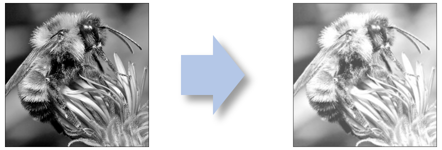


Scale



- Photometric

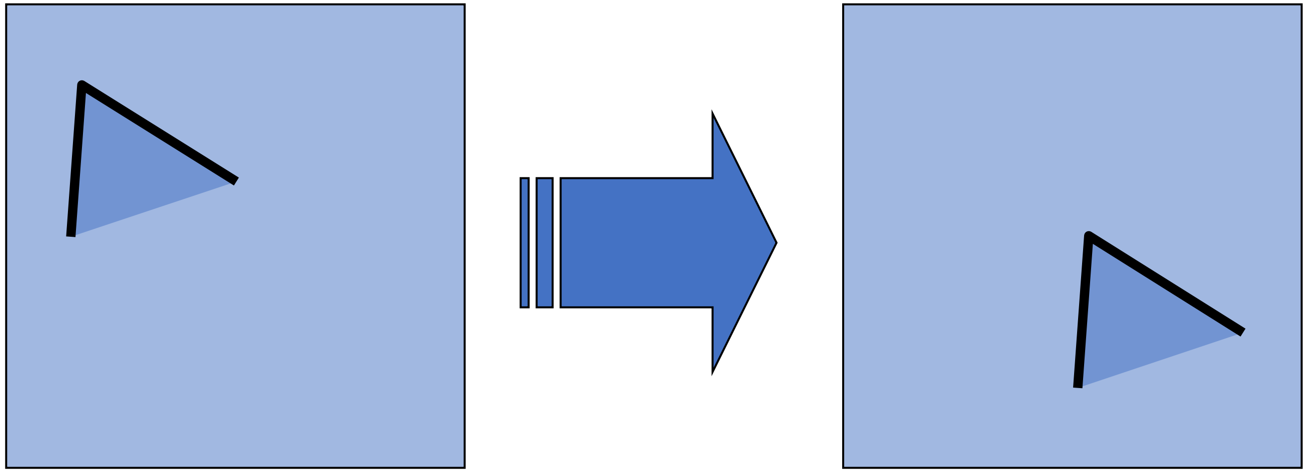
Intensity change



Invariance and equivariance

- We want corner locations to be *invariant* to photometric transformations and *equivariant* to geometric transformations
  - **Invariance:** image is transformed and corner locations do not change
  - **Equivariance:** if we have two transformed versions of the same image, features should be detected in corresponding locations
  - (Sometimes “invariant” and “equivariant” are both referred to as “invariant”)
  - (Sometimes “equivariant” is called “covariant”)

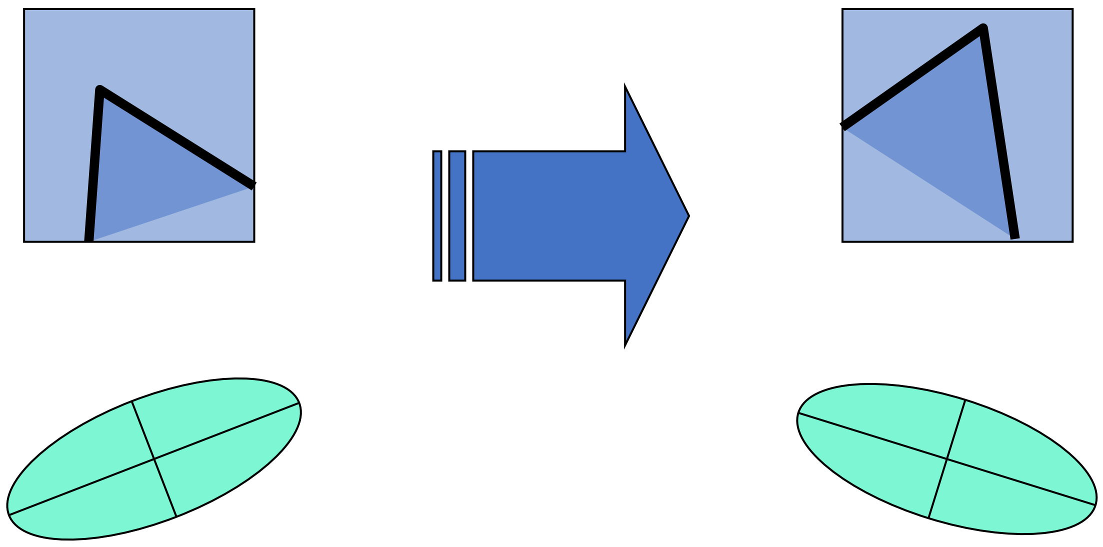
Harris detector invariance properties: image translation



- Derivatives and window function are equivariant

Corner location is equivariant w.r.t. translation

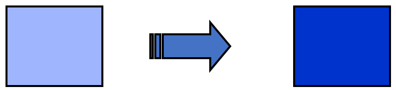
Harris detector invariance properties: image rotation



Second moment ellipse rotates but its shape (i.e. eigenvalues) remains the same

Corner location is equivariant w.r.t. image rotation

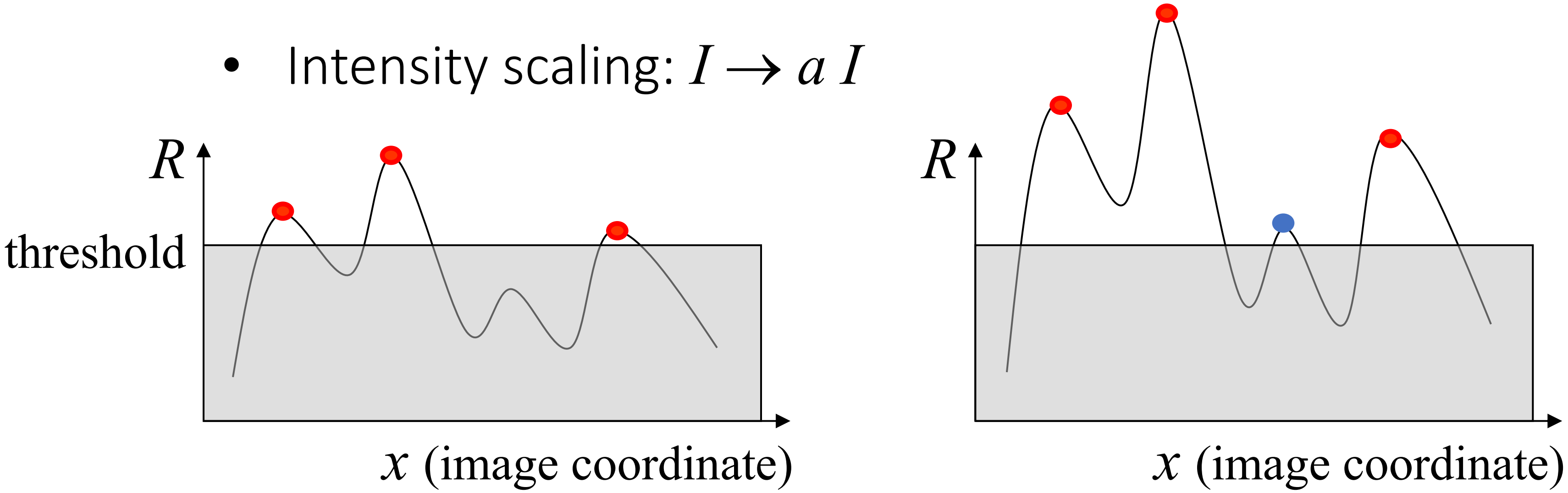
Harris detector invariance properties: Affine intensity change



$I \rightarrow a\,I + b$

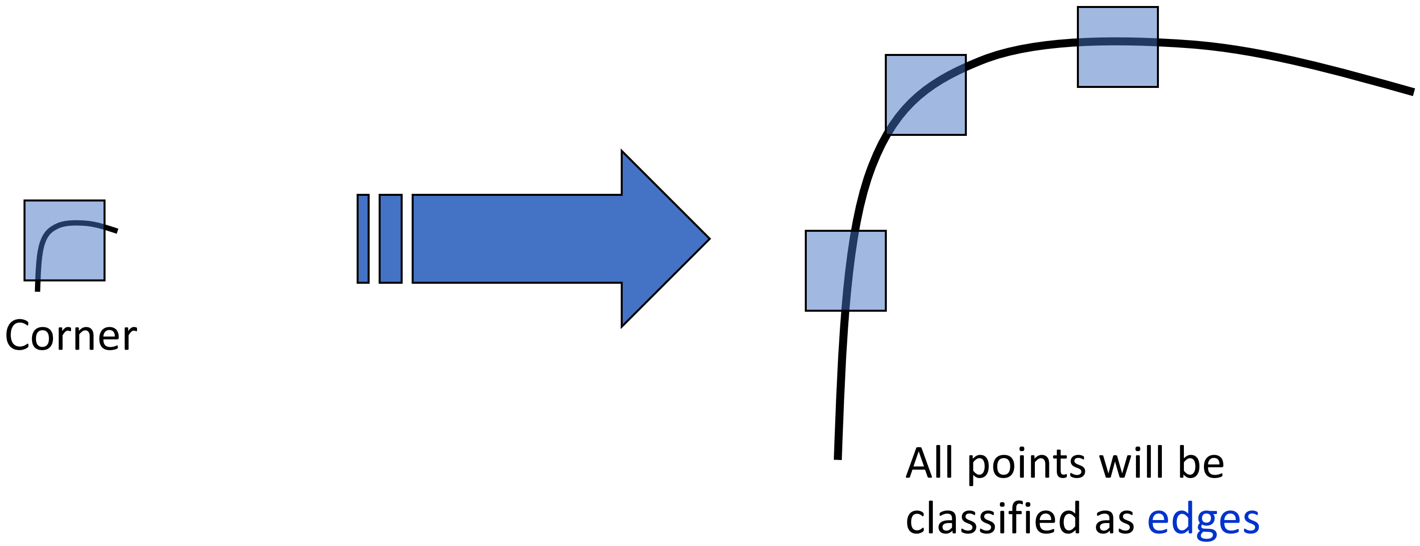
- Only derivatives are used → invariance to intensity shift $I \rightarrow I + b$

- Intensity scaling: $I \rightarrow a\,I$



Partially invariant to affine intensity change

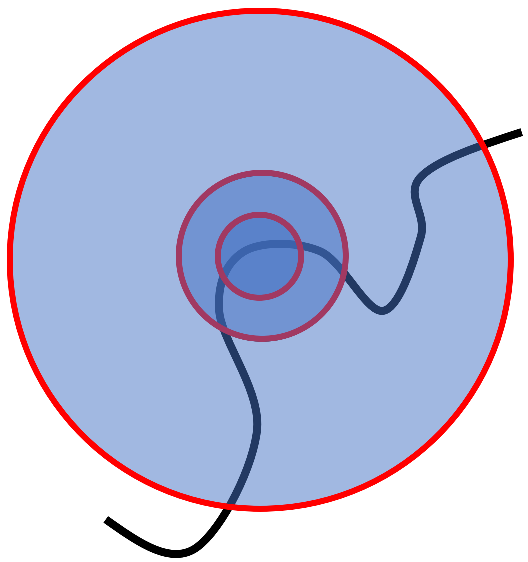
Harris detector invariance properties: scaling



Neither invariant nor equivariant to scaling

Scale invariant detection

Suppose you're looking for corners

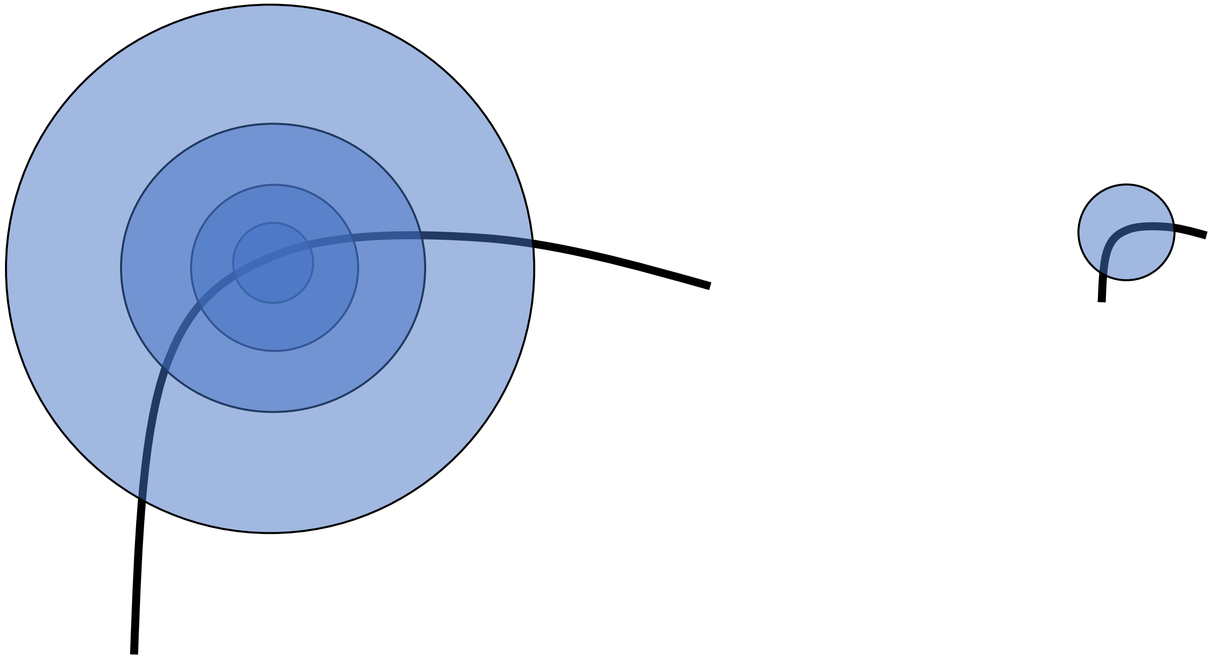


Key idea: find scale that gives local maximum of *f*

- in both position and scale
- One definition of *f*: the Harris operator

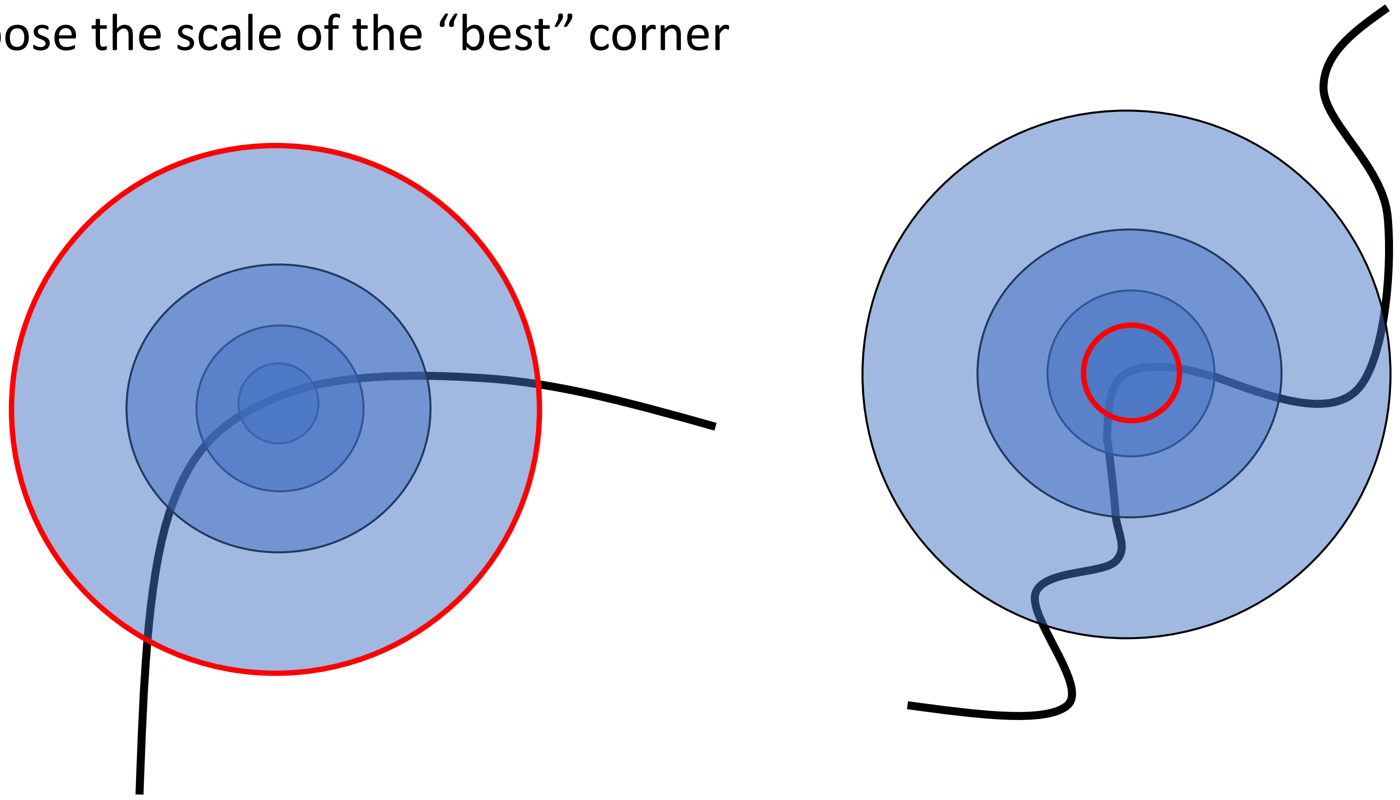
Scale Invariant Detection

- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding sizes will look the same in both images



Scale Invariant Detection

- The problem: how do we choose corresponding circles ***independently*** in each image?
- Choose the scale of the "best" corner

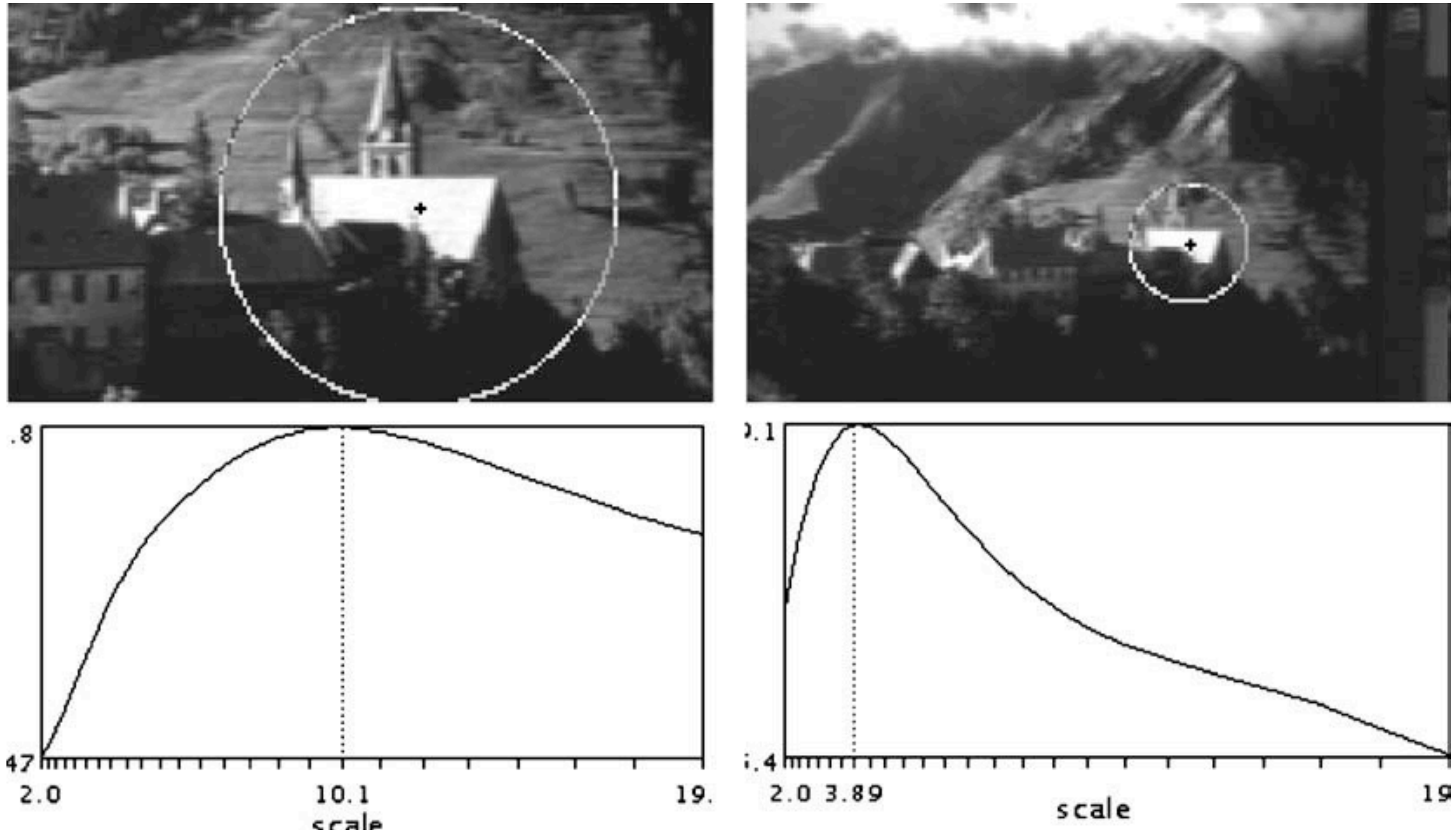


Keypoint detection with scale selection

- We want to extract keypoints with *characteristic scales* that are *equivariant* (or *covariant*) w.r.t. to scaling of the image

Harris Corner Response

Scale-invariant response function

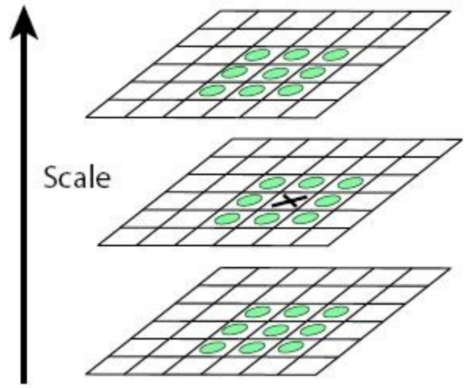


K. Mikolajczyk and C. Schmid. [Indexing based on scale invariant interest points](). ICCV 2001

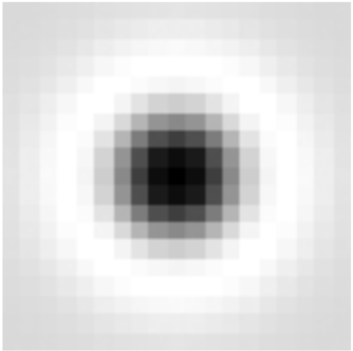
T. Lindeberg, [Feature detection with automatic scale selection](), *IJCV* 30(2), pp. 77-116, 1998

Keypoint detection with scale selection

- We want to extract keypoints with *characteristic scales* that are *equivariant* (or *covariant*) w.r.t. to scaling of the image
- Approach: compute a *scale-invariant* response function over neighborhoods centered at each location (x, y) and a range of scales (σ), find *scale-space locations* (x, y, σ) where this function reaches a local maximum
- A particularly convenient response function is given by the *scale-normalized Laplacian of Gaussian (LoG) filter*:



"scale space"

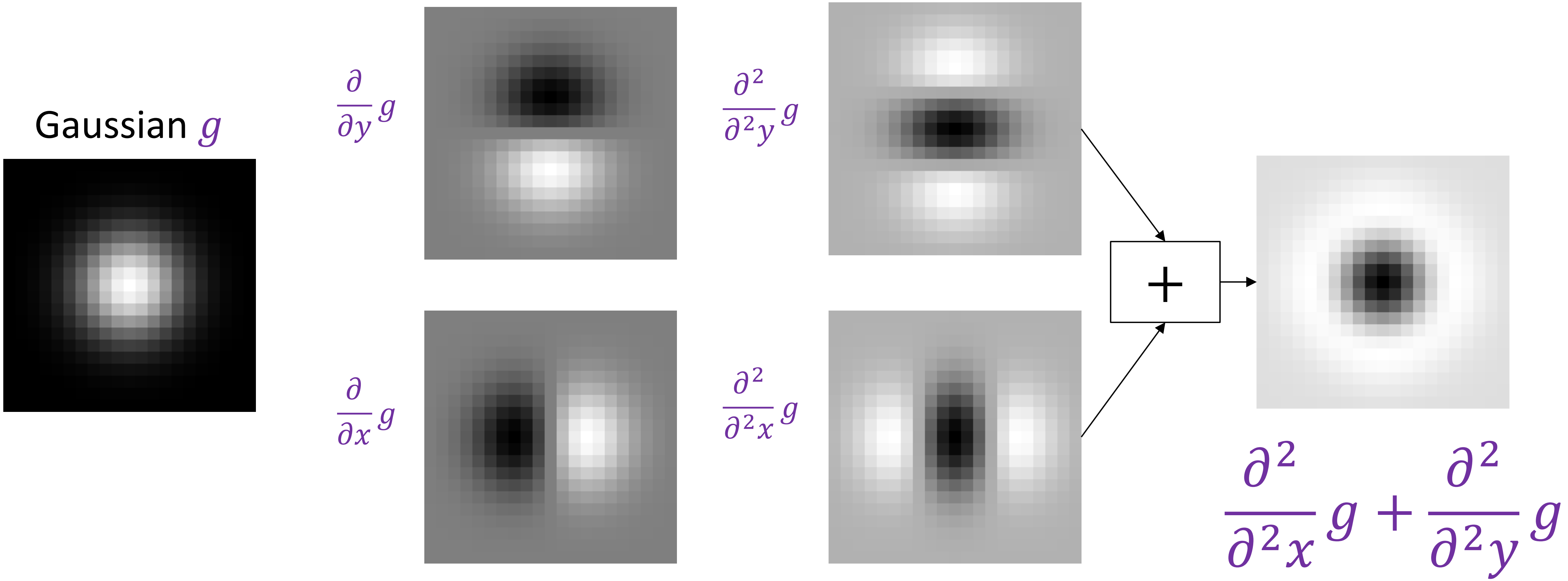


$$\nabla^2_{\text{norm}} = \sigma^2 \left(\frac{\partial^2}{\partial x^2} g + \frac{\partial^2}{\partial^2 y} g \right)$$

Content: Today’s class

- Why detect features?
- What is a good feature?
- Harris Corner Detector
- Properties of Harris Corner Detector
- Blob Detector

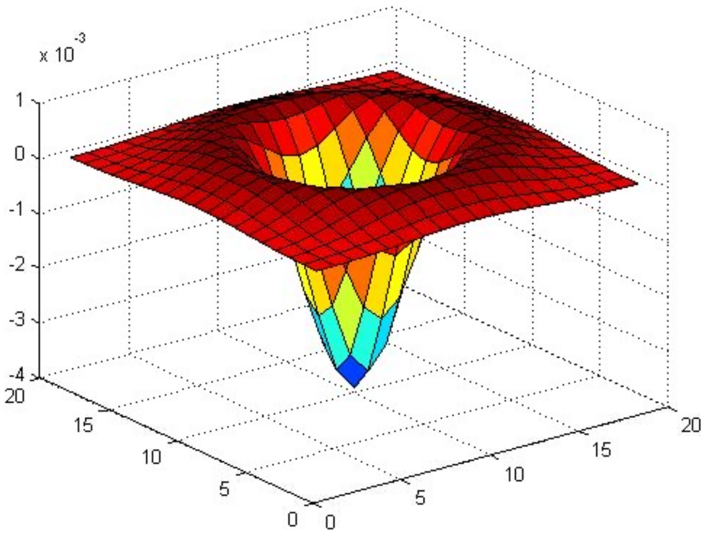
Laplacian of Gaussian

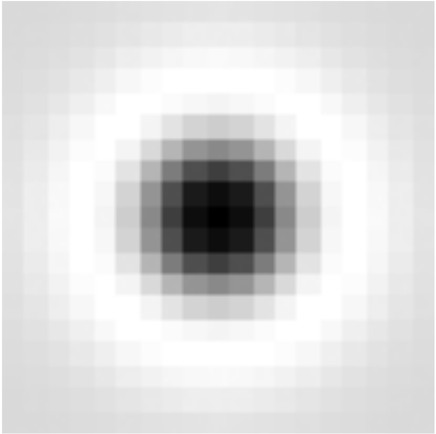


Source: [J. Johnson and D. Fouhey]

Scale-normalized Laplacian

- You need to multiply the LoG by σ^2 to make responses comparable across scales

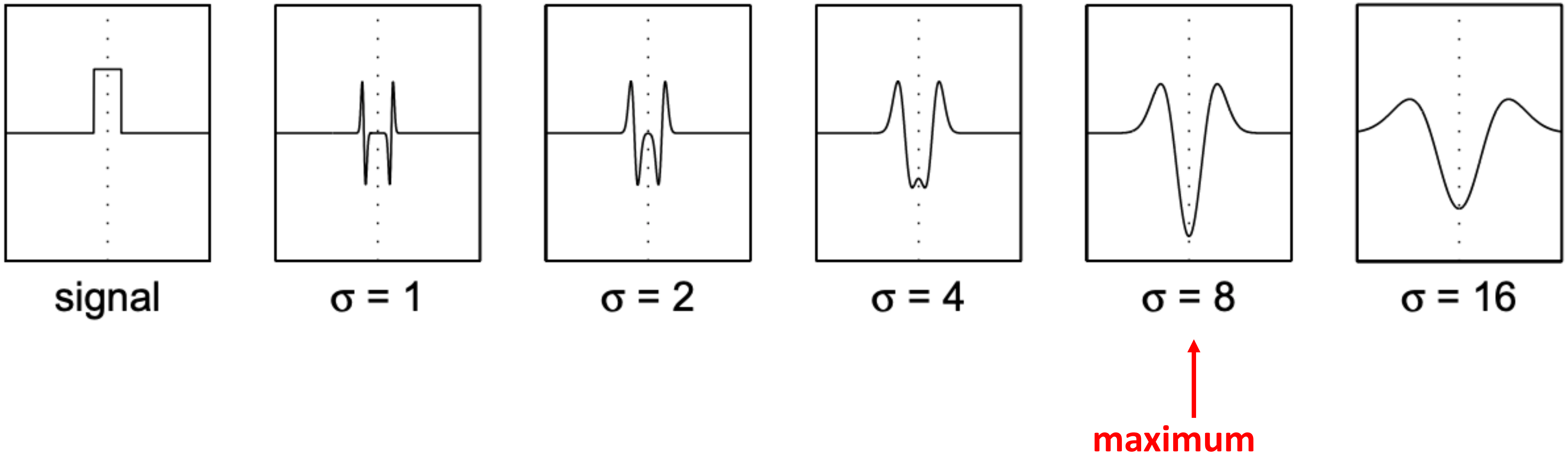




$$\nabla^2_{\text{norm}} = \sigma^2 \left(\frac{\partial^2}{\partial x^2} g + \frac{\partial^2}{\partial^2 y} g \right)$$

Scale selection: Characteristic Scale

- We can find the *characteristic scale* of the blob by convolving it with *scale-normalized* Laplacians at several scales (σ) and looking for the maximum response

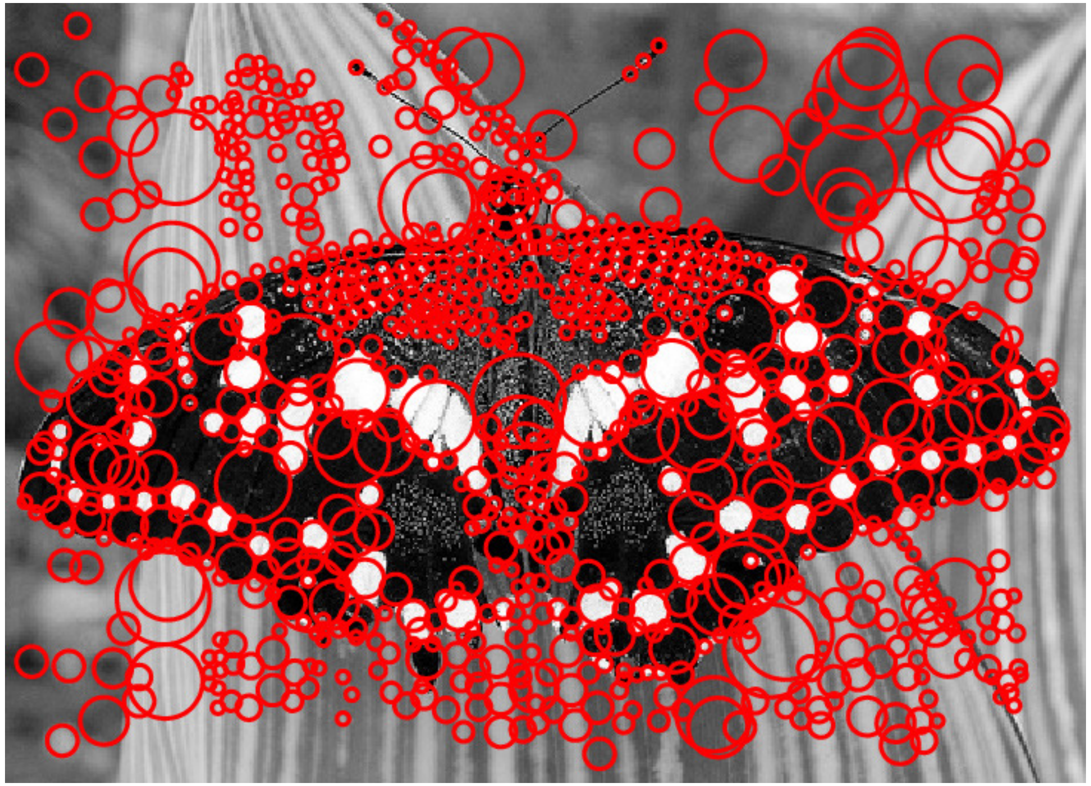


Scale-space blob detector: Example

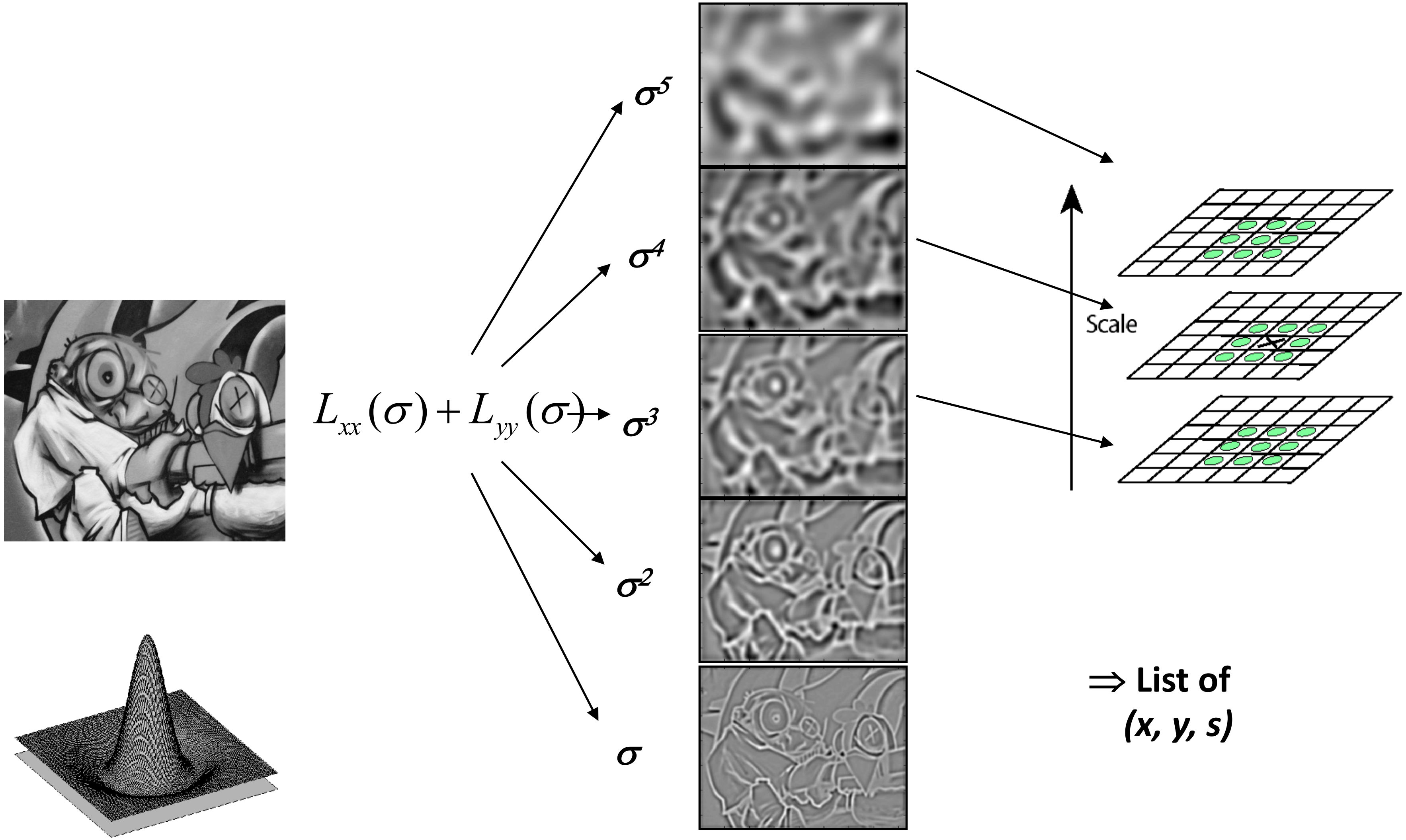
Scale-space blob detector: Example

sigma = 11.9912

Scale-space blob detector: Example



Find local maxima in 3D position-scale space

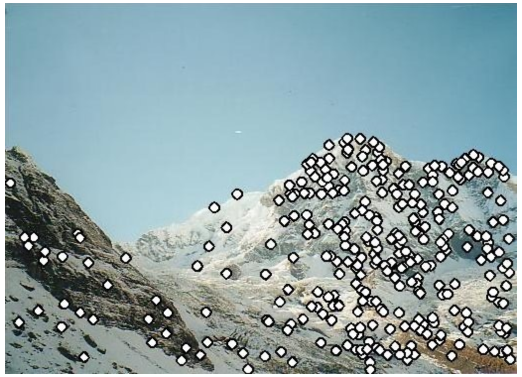


K. Grauman, B. Leibe

Local features: main components

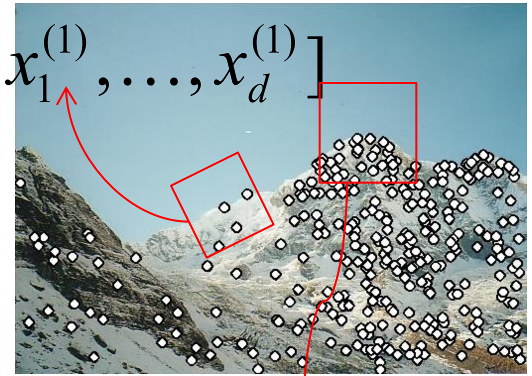
This Class

1) **Detection**: Identify the interest points



2) **Description**: Extract vector feature descriptor surrounding each interest point.

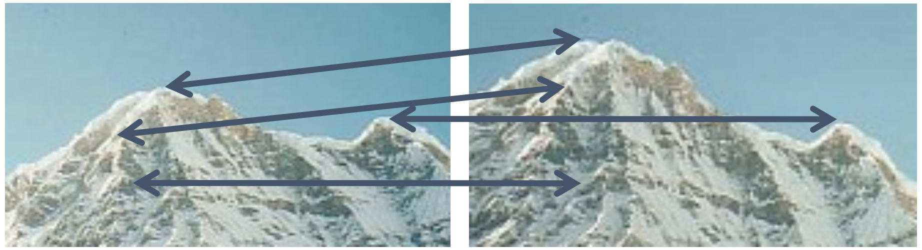
$$\mathbf{x}_1 = [x_1^{(1)}, \ldots, x_d^{(1)}]$$



$$\mathbf{x}_2 = [x_1^{(2)}, \ldots, x_d^{(2)}]$$

Next Class:
We will learn about what is SIFT feature! The most famous feature in Computer Vision!!

3) **Matching**: Determine correspondence between descriptors in two views



Kristen Grauman

Slide Credits

- CS5670, Introduction to Computer Vision**, Cornell Tech, by Noah Snavely.**
- CS 194-26/294-26: Intro to Computer Vision and Computational Photography**, UC Berkeley, by Alyosha Efros.**
- Fall 2022 CS 543/ECE 549: Computer Vision, UIUC, by **Svetlana Lazebnik.**