ELEC 474 Machine Vision

1

FEATURE DETECTORS AND DESCRIPTORS

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Contents

- Local invariant features
  - Motivation
  - Requirements and invariances
- Keypoint/interest point detection
  - Harris corner detector
- Scale invariant region selection
  - Automatic scale selection
  - Difference of Gaussian Detector
- Scale Invariant Feature Transform (SIFT)

Motivation: Features for Recognition

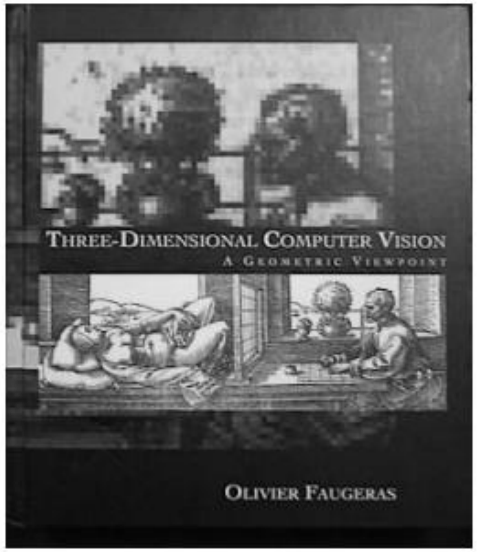


Image search: find the book in an image.

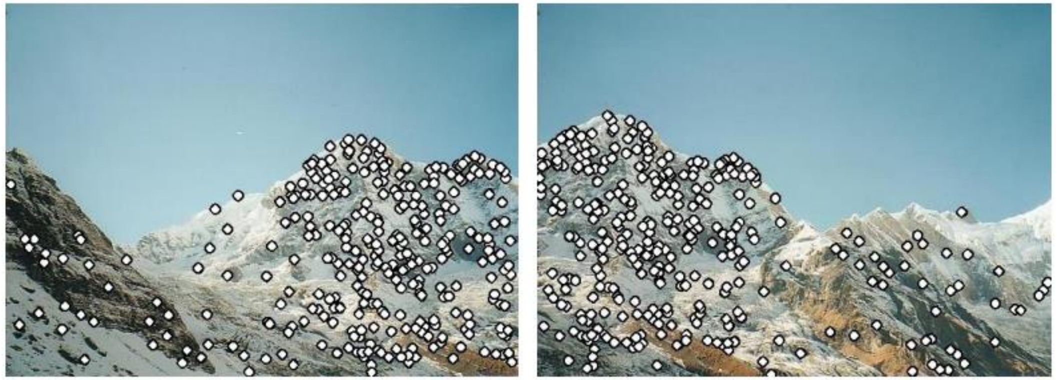
Motivation: Build a Panorama

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

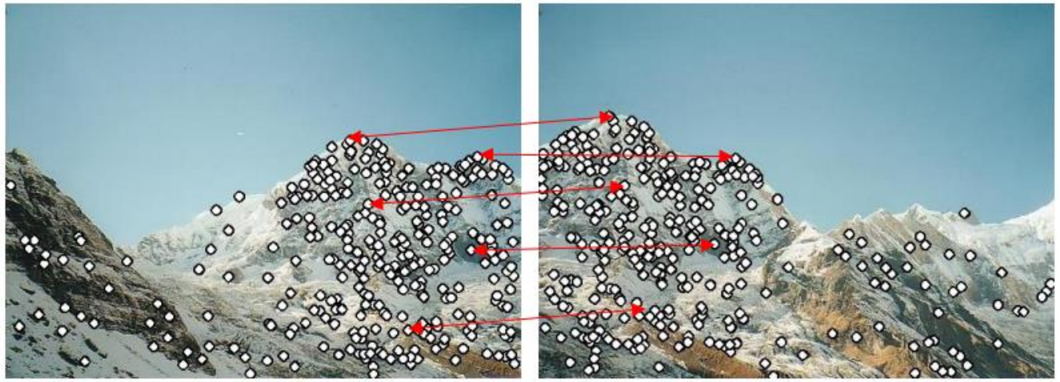
How do we build panorama?

- We need to match (align) images

1. Detect feature points in both images



2. Find corresponding pairs



3. Use these pairs to align images

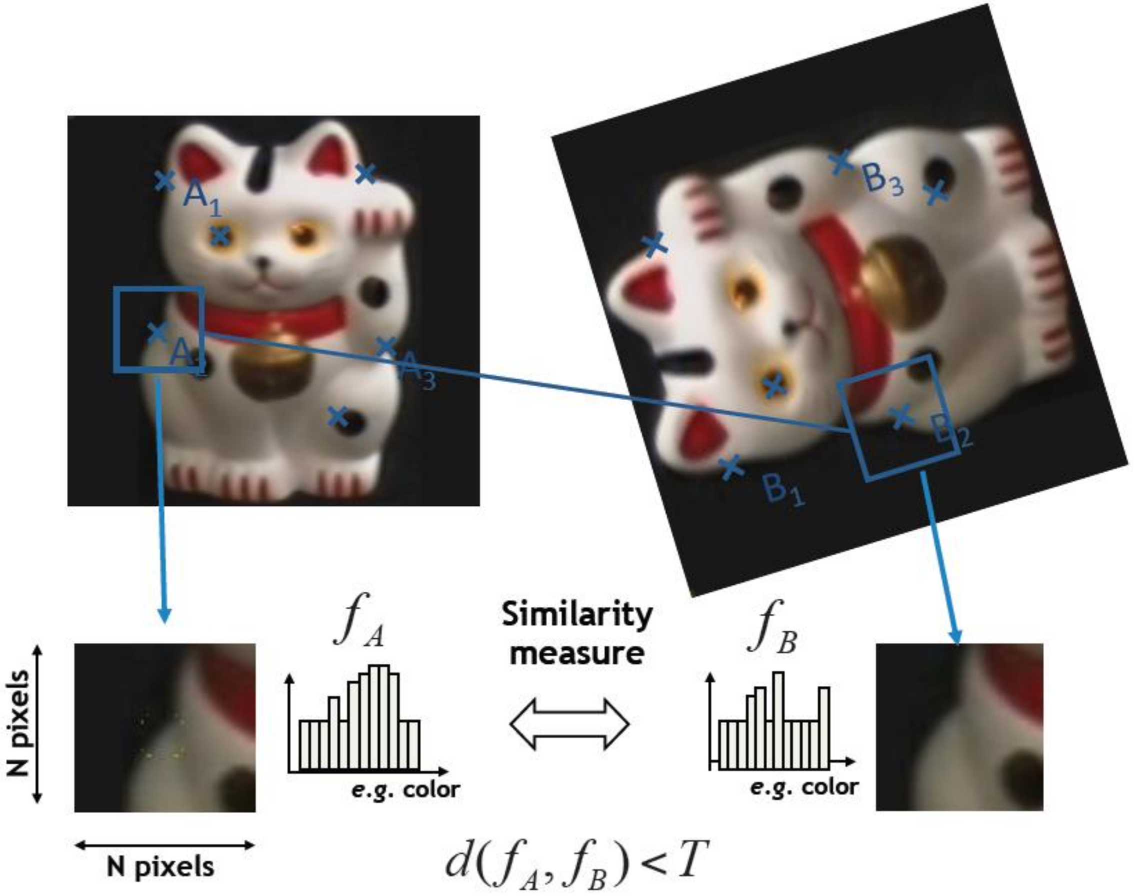
More Motivation

- Feature points are used also for:
  - Point matching for computing disparity
  - Image alignment (homography, fundamental matrix)
  - 3D reconstruction
  - Motion tracking
  - Motion based segmentation
  - Object recognition
  - Image retrieval and indexing
  - Robot navigation
  - … other

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Local Features

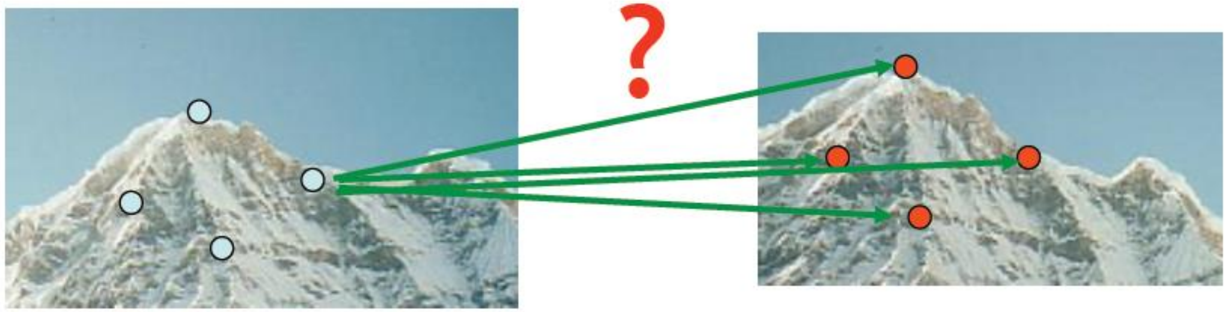
- Matching General Approach:
  1. Find a set of distinctive interest points
  2. Define a region around each interest point
  3. Extract and normalize the region content
  4. Compute a local descriptor from the normalized region
  5. Match local descriptors to establish correspondences
  6. Estimate transformation from matches



Local Features

- Challenges
  - Change in position, scale, and rotation
  - Change in viewpoint
  - Occlusion
  - Articulation, change in appearance
- Requirements
  - Detect the same interest points independently in both images
    - We need a **repeatable detector**
  - For each interest point correctly recognize the corresponding one
    - We need a reliable and **distinctive descriptor**

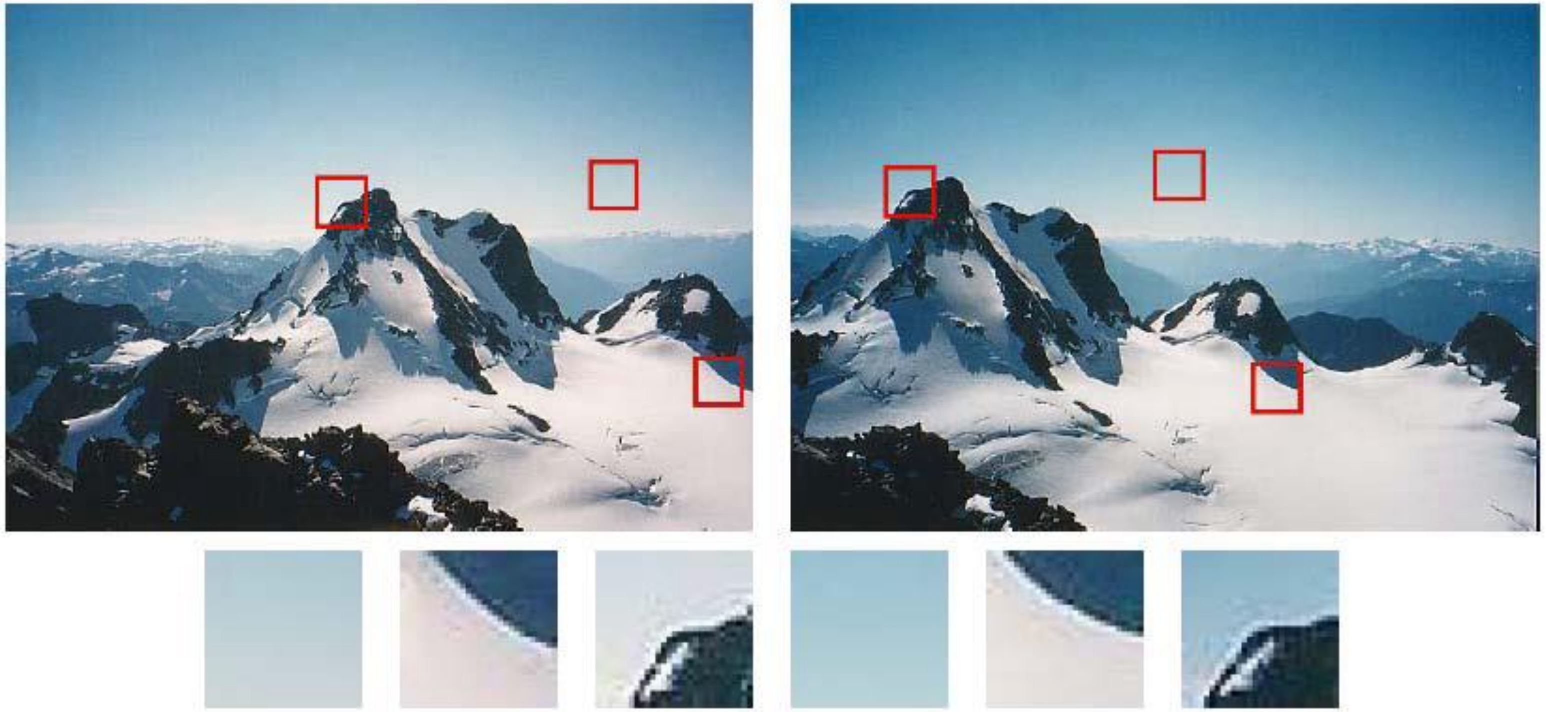
No chance of match!



Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

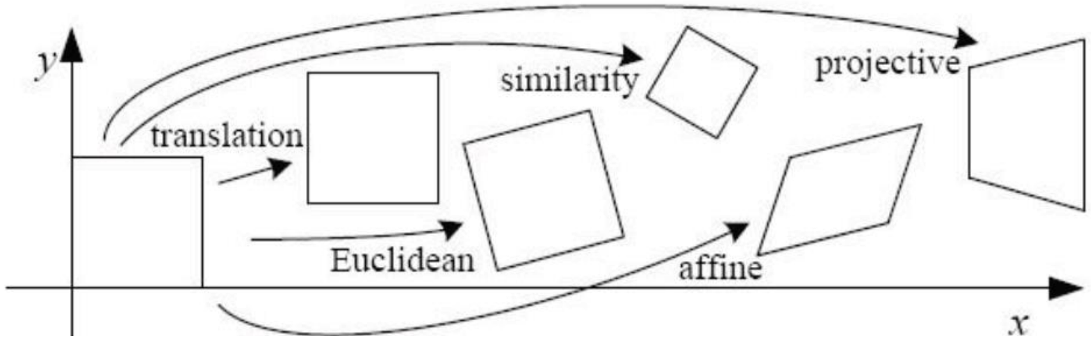
Local Features

- Some patches can be localized or matched with higher accuracy than others



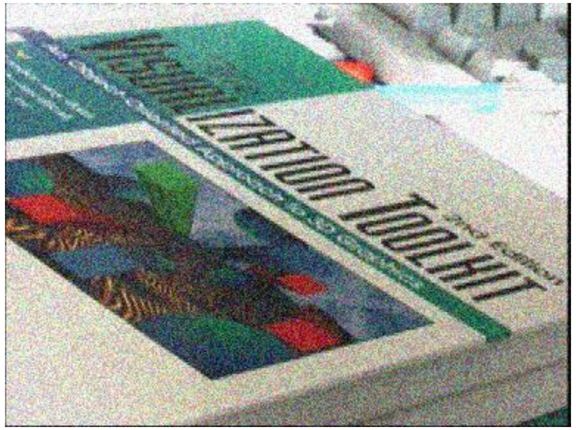
Invariance

Invariance to geometric transformations



Levels of geometric invariance

Invariance



Photometric Transformation

Detectors vs. Descriptors

Feature Detectors

- Also known as *"interest operators"*
- Identify pixels in the image which are worth processing further
- Are computationally very efficient
- Do not strongly describe the content in the pixel neighborhood
- Examples:
  - Harris Corner detector

Feature Descriptors

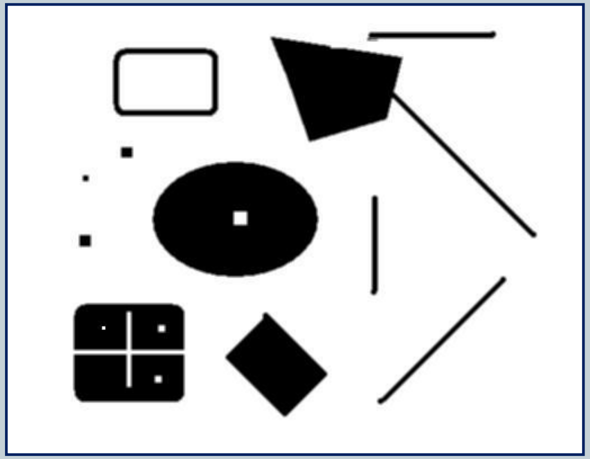
- Create a distinctive "code" that describes the neighborhood of a pixel
  - High dimensional
  - Invariant to noise, rotations, scale, etc.
- More computationally expensive than detectors
- Examples:
  - SIFT: Scale Invariant Feature Transform
  - SURF: Speeded Up Robust Features
  - FAST: Features from Accelerated Segment Test
  - BRIEF: Binary Robust Independent Elementary Features
  - ORB: Oriented FAST and Rotated BRIEF

Properties of the ideal feature

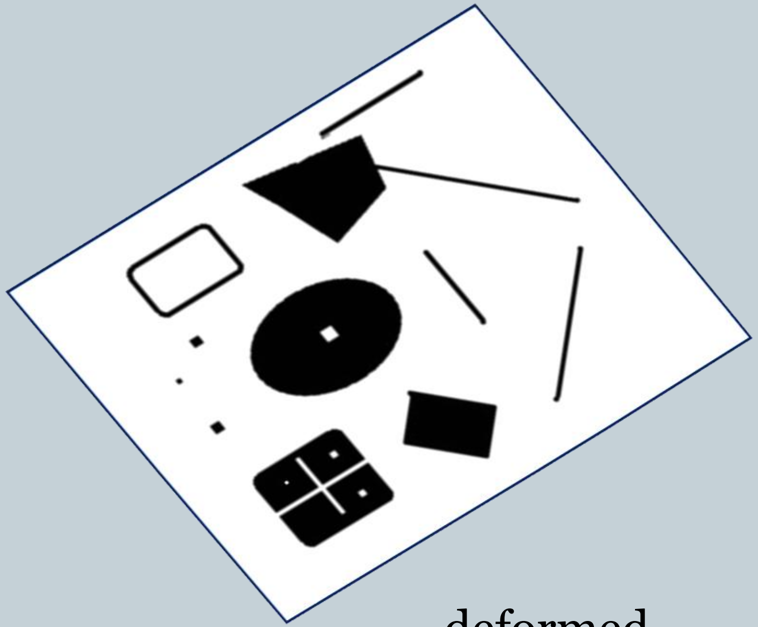
- **Local**
  - Defined by a neighborhood, and so robust to occlusion and clutter (no prior segmentation)
- **Invariant**
  - Remain unchanged under (some expected) changes in the image
- **Robust**
  - Noise, blur, discretization, compression, etc. do not have a big impact on the feature
- **Distinctive**
  - Individual features can be matched to a large database of objects
- **Quantity**
  - Many features can be generated for even small objects
- **Accurate**
  - Precise localization
- **Efficient**
  - Close to real-time performance

Harris Corner Detector

- Why corners as interest points?
  - Corners are *repeatable* and *distinctive*



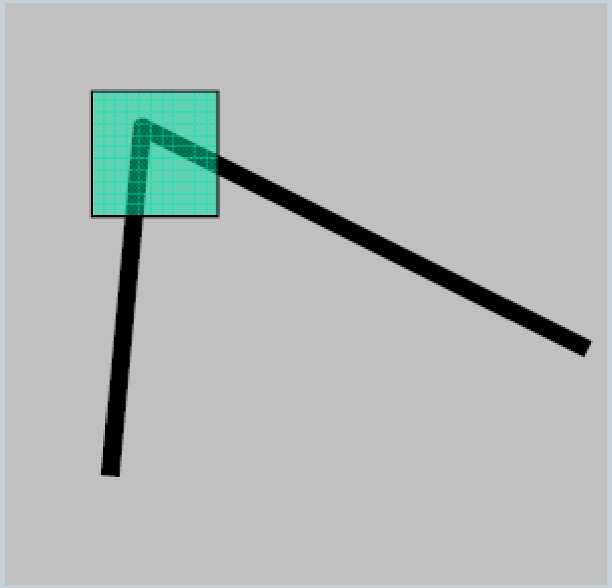
original



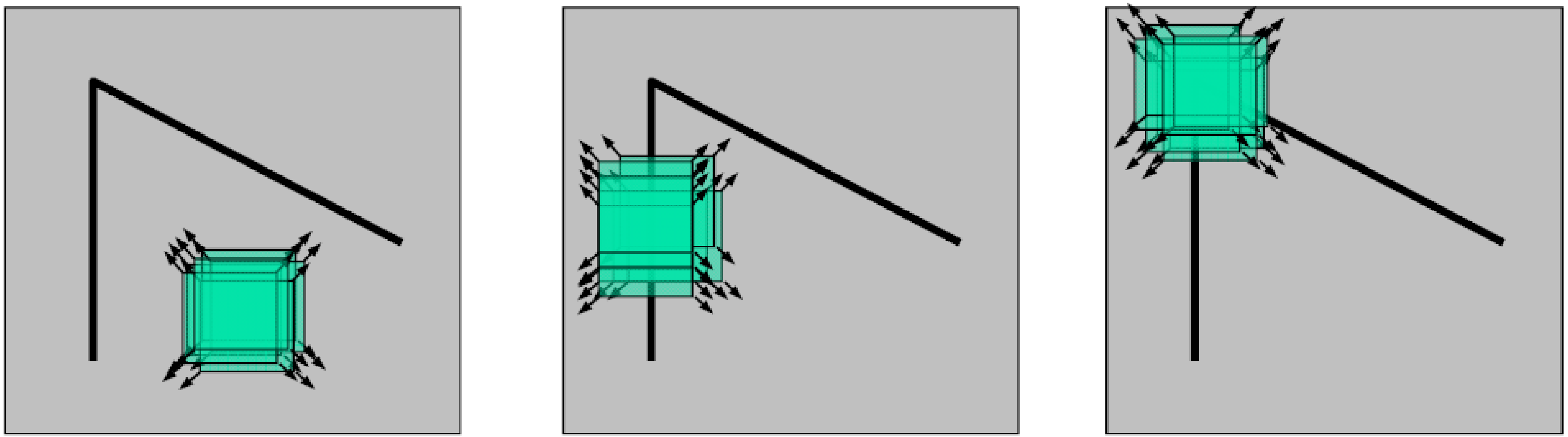
deformed

Harris Corner Detector

- Corner point can be recognized in a window
- Shifting a window in any direction should produce a large change in intensity



Basic Idea



"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

"corner":
significant change
in all directions

Harris Detector Formulation

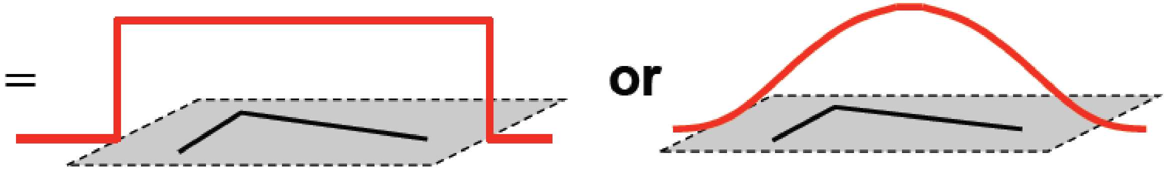
- Change of intensity for the shift [u,v]:

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

Window function

Shifted intensity

Intensity

Window function $w(x,y) =$  or

1 in window, 0 outside

Gaussian

Harris Detector Formulation

$$E(u,v)=\sum_{x,y}\underbrace{w(x,y)}_{\text{window function}}\,[\underbrace{I(x+u,y+v)}_{\text{shifted intensity}}-\underbrace{I(x,y)}_{\text{intensity}}]^2$$

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Harris Detector Formulation

$$E(u,v) = \sum_{x,y} \underbrace{w(x,y)}_{\text{window function}} [\underbrace{I(x+u, y+v)}_{\text{shifted intensity}} - \underbrace{I(x,y)}_{\text{intensity}}]^2$$

$$E(u,v) = \sum_{x,y} \underbrace{w(x,y)}_{\text{window function}} [\underbrace{I(x,y) + uI_x + vI_y}_{\text{shifted intensity}} - \underbrace{I(x,y)}_{\text{intensity}}]^2$$

Taylor Series

Harris Detector Formulation

$$E(u,v) = \sum_{x,y} \underbrace{w(x,y)}_{\text{window function}} \; [\underbrace{I(x+u, y+v)}_{\text{shifted intensity}} - \underbrace{I(x,y)}_{\text{intensity}}]^2$$

$$E(u,v) = \sum_{x,y} \underbrace{w(x,y)}_{\text{window function}} \; [\underbrace{I(x,y) + uI_x + vI_y}_{\text{shifted intensity}} - \underbrace{I(x,y)}_{\text{intensity}}]^2$$

Taylor Series

$$E(u,v) = \sum_{x,y} w(x,y)[\, uI_x + vI_y \,]^2$$

Harris Detector Formulation

$$E(u,v) = \sum_{x,y} w(x,y)[\,uI_x + vI_y\,]^2$$

$$E(u,v) = \sum_{x,y} w(x,y)\left[\begin{pmatrix} u & v \end{pmatrix}\begin{pmatrix} I_x \\ I_y \end{pmatrix}\right]^2$$

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Harris Detector Formulation

$$E(u,v)=\sum_{x,y} w(x,y)[\,uI_x+vI_y\,]^2$$

$$E(u,v)=\sum_{x,y} w(x,y)\left[\begin{pmatrix}u & v\end{pmatrix}\begin{pmatrix}I_x\\ I_y\end{pmatrix}\right]^2$$

$$E(u,v)=\sum_{x,y} w(x,y)\begin{pmatrix}u & v\end{pmatrix}\begin{pmatrix}I_x\\ I_y\end{pmatrix}\begin{pmatrix}I_x & I_y\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}$$

Harris Detector Formulation

$$E(u,v) = \sum_{x,y} w(x,y)[\,uI_x + vI_y\,]^2$$

$$E(u,v) = \sum_{x,y} w(x,y)\left[\begin{pmatrix} u & v \end{pmatrix}\begin{pmatrix} I_x \\ I_y \end{pmatrix}\right]^2$$

$$E(u,v) = \sum_{x,y} w(x,y)\begin{pmatrix} u & v \end{pmatrix}\begin{pmatrix} I_x \\ I_y \end{pmatrix}\begin{pmatrix} I_x & I_y \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

$$E(u,v) = \begin{pmatrix} u & v \end{pmatrix}\left[\sum_{x,y} w(x,y)\begin{pmatrix} I_x \\ I_y \end{pmatrix}\begin{pmatrix} I_x & I_y \end{pmatrix}\right]\begin{pmatrix} u \\ v \end{pmatrix}$$

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Harris Detector Formulation

$$E(u,v)=\sum_{x,y} w(x,y)[\,uI_x+vI_y\,]^2$$

$$E(u,v)=\sum_{x,y} w(x,y)\left[\begin{pmatrix}u & v\end{pmatrix}\begin{pmatrix}I_x\\ I_y\end{pmatrix}\right]^2$$

$$E(u,v)=\sum_{x,y} w(x,y)\begin{pmatrix}u & v\end{pmatrix}\begin{pmatrix}I_x\\ I_y\end{pmatrix}\begin{pmatrix}I_x & I_y\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}$$

$$E(u,v)=\begin{pmatrix}u & v\end{pmatrix}\left[\sum_{x,y} w(x,y)\begin{pmatrix}I_x\\ I_y\end{pmatrix}\begin{pmatrix}I_x & I_y\end{pmatrix}\right]\begin{pmatrix}u\\ v\end{pmatrix}$$

$$E(u,v)=\begin{pmatrix}u & v\end{pmatrix}M\begin{pmatrix}u\\ v\end{pmatrix}$$

Harris Detector Formulation

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Gradient with respect to x, times gradient with respect to y

Sum over image region – the area we are checking for corner

$$M = \begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} = \sum \begin{bmatrix} I_x \\ I_y \end{bmatrix} [I_x \ I_y]$$

-1	0	1
-2	0	2
-1	0	1

I_x

-1	-2	-1
0	0	0
1	2	1

I_y

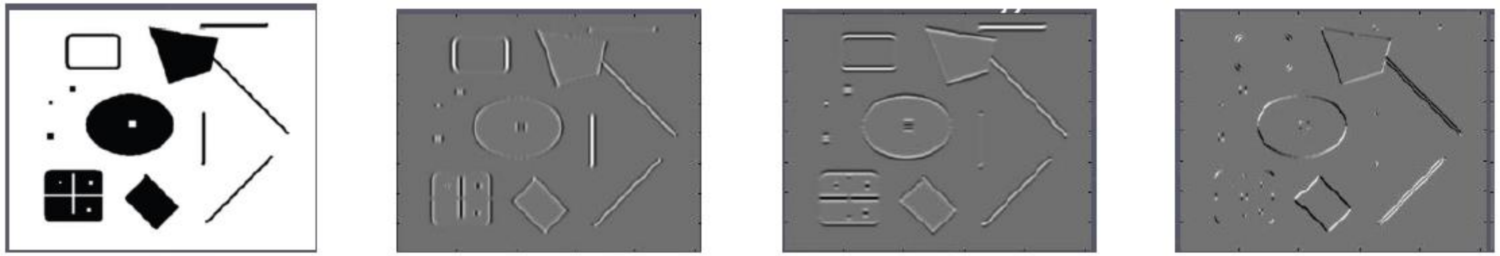


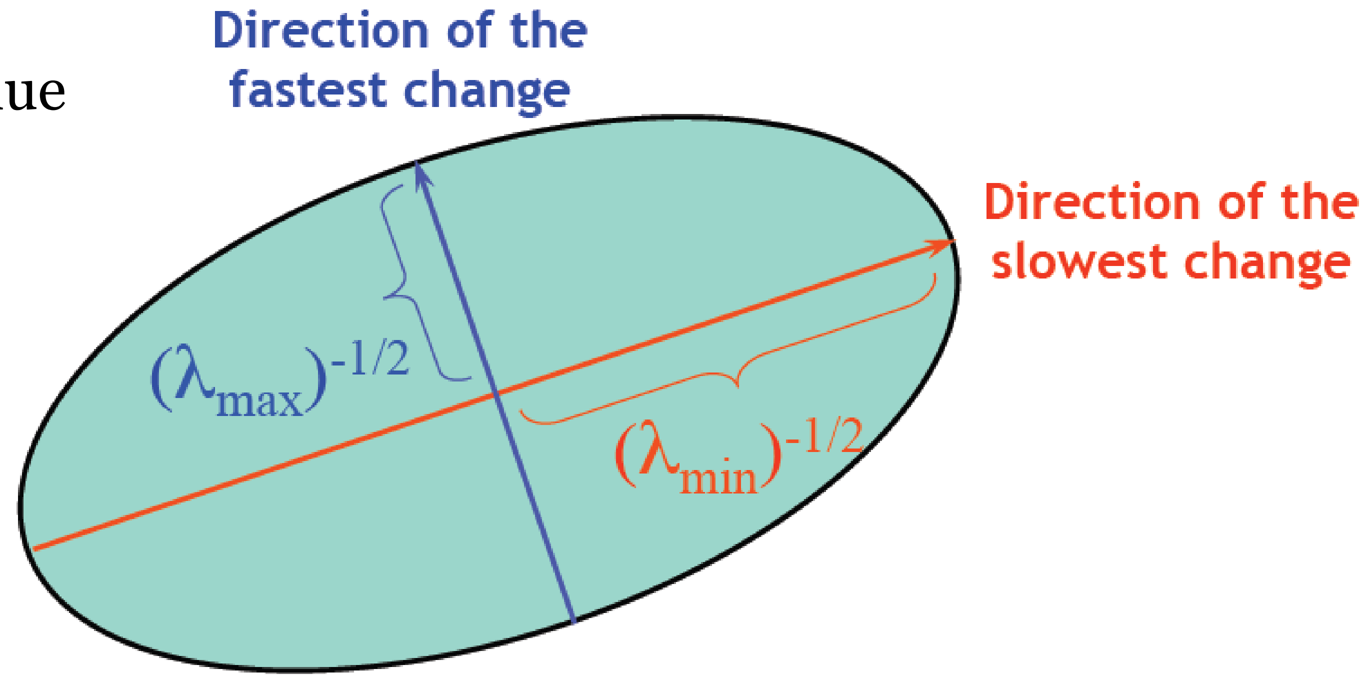
Image I $\quad I_x \quad I_y \quad I_x I_y$

Harris Detector Formulation

- M is symmetric
  - So can decompose M using eigenvalue decomposition:

$$M = A^{-1} \Lambda \, A$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$



- Can visualize M as an ellipse
  - axis lengths determined by the eigenvalues Λ
  - orientation determined by A.

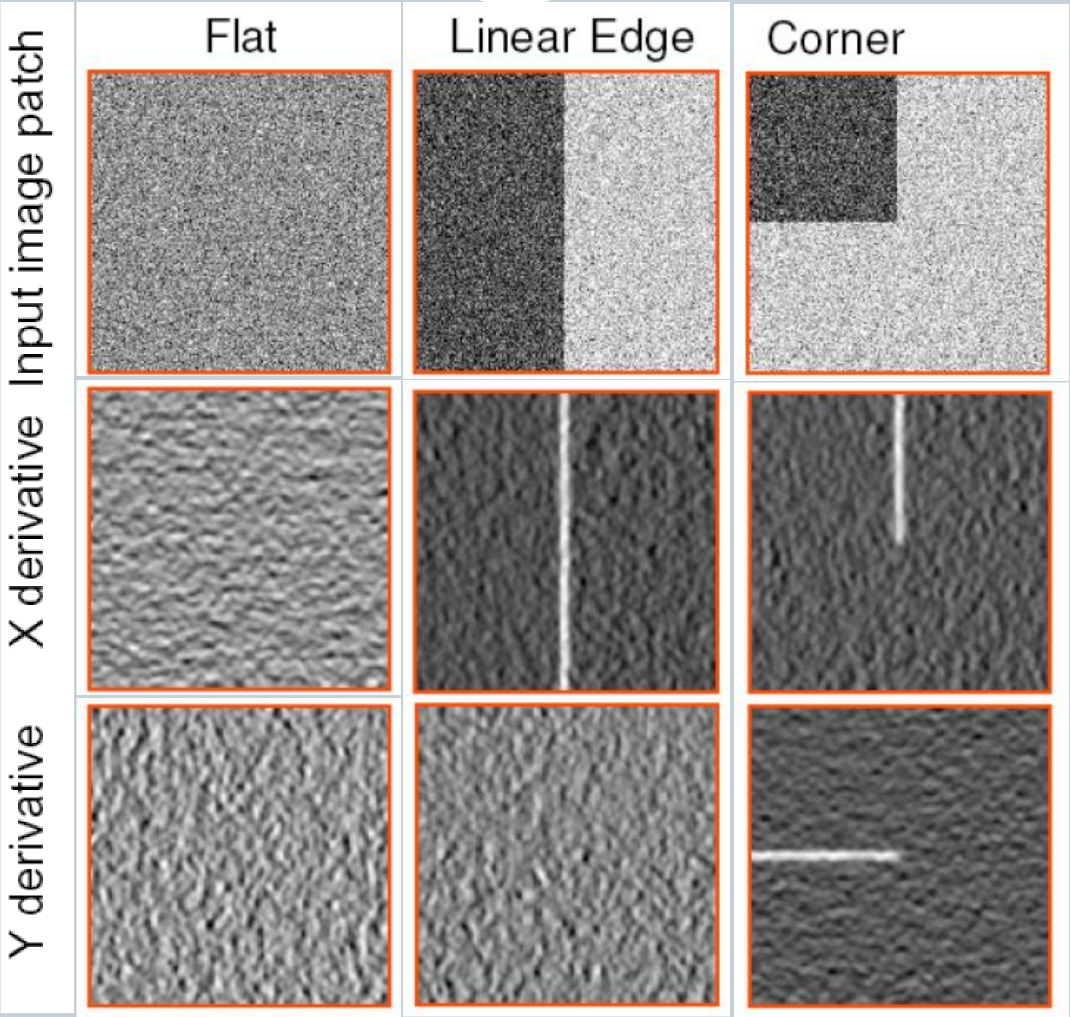
Intuitive Way to Understand Harris

- Treat gradient vectors as a set of (dx,dy) points with a center of mass defined as being at (0,0).

- Fit an ellipse to that set of points via scatter matrix

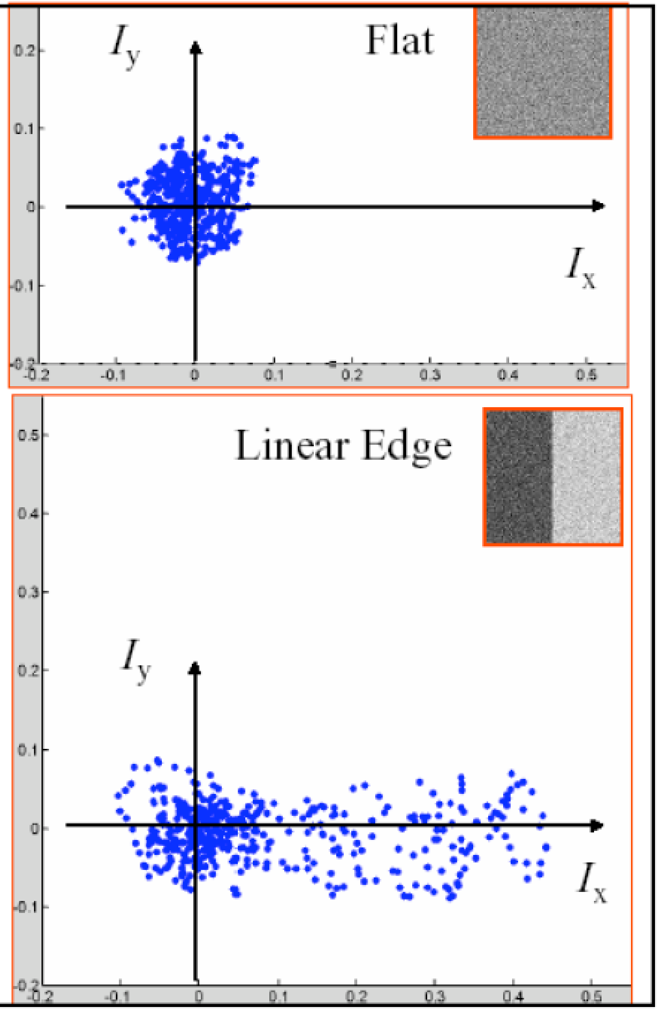
- Analyze ellipse parameters for varying cases…

Example: Cases and 2D Derivatives



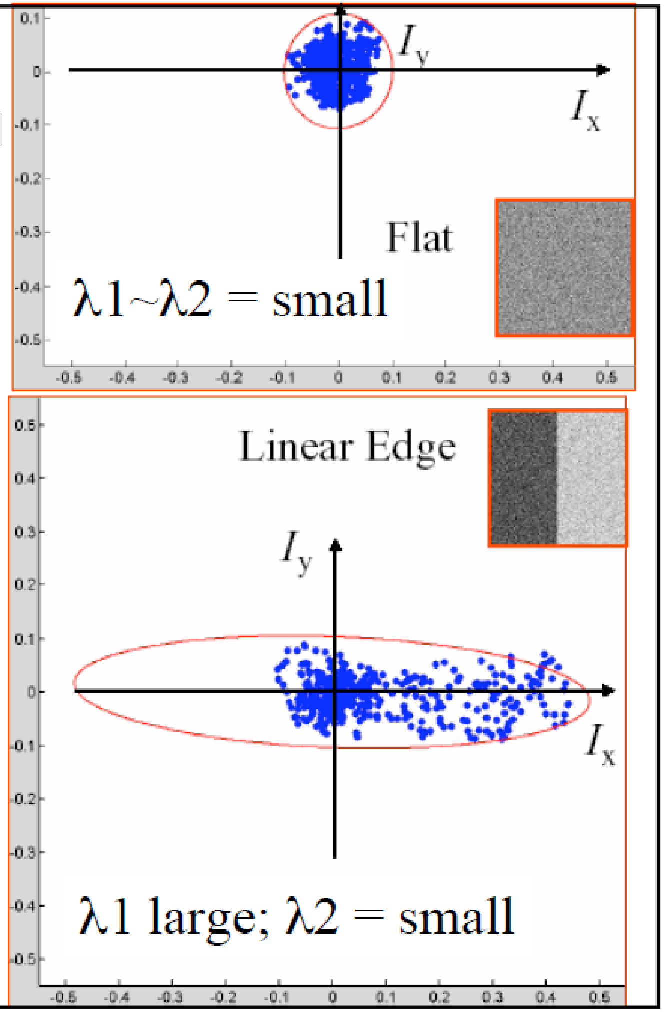
Plotting Derivatives as 2D Points

The distribution of the x and y derivatives is very different for all three types of patches



Fitting Ellipse to each Set of Points

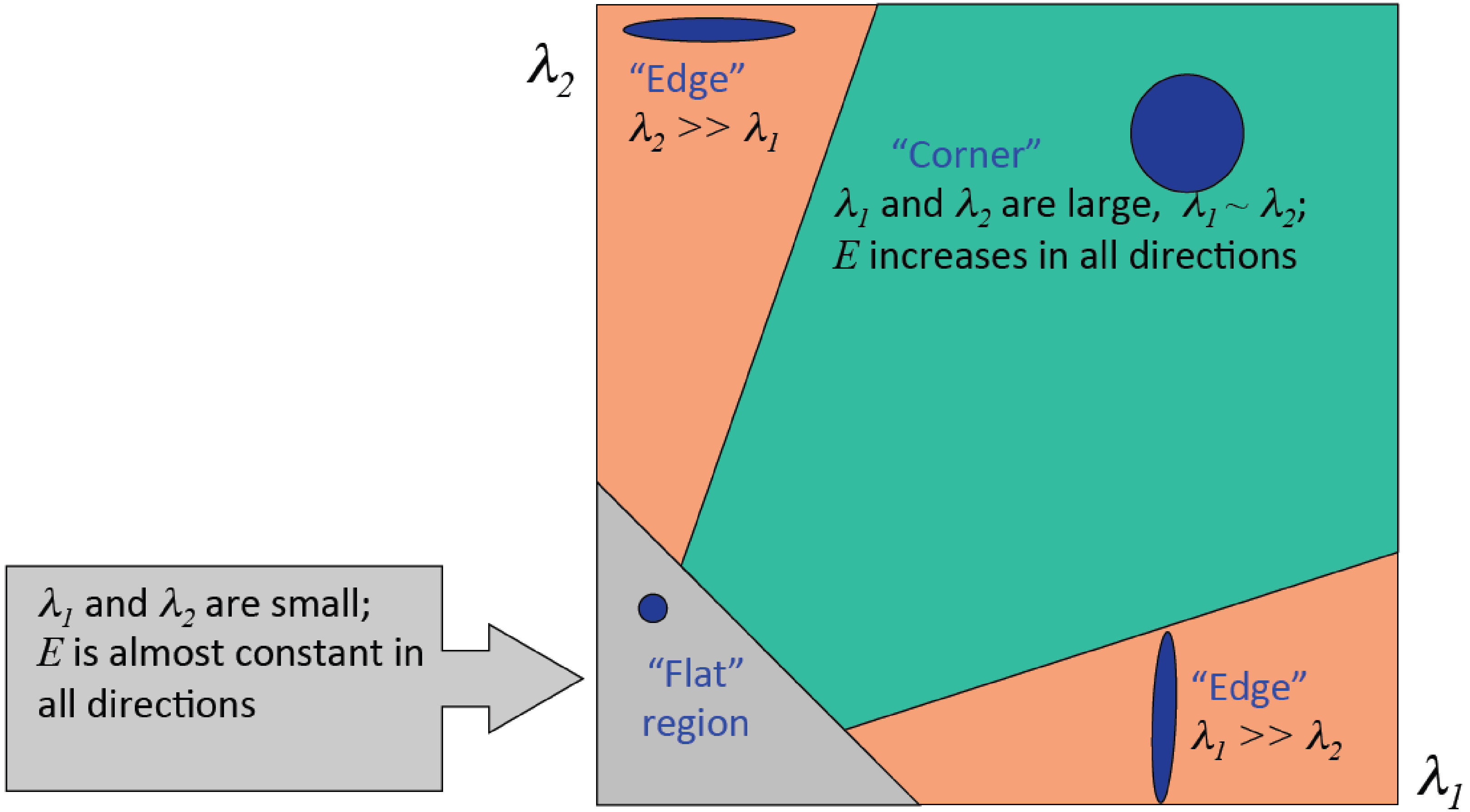
The distribution of x and y derivatives can be characterized by the shape and size of the principal component ellipse



Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Harris Detector Formulation

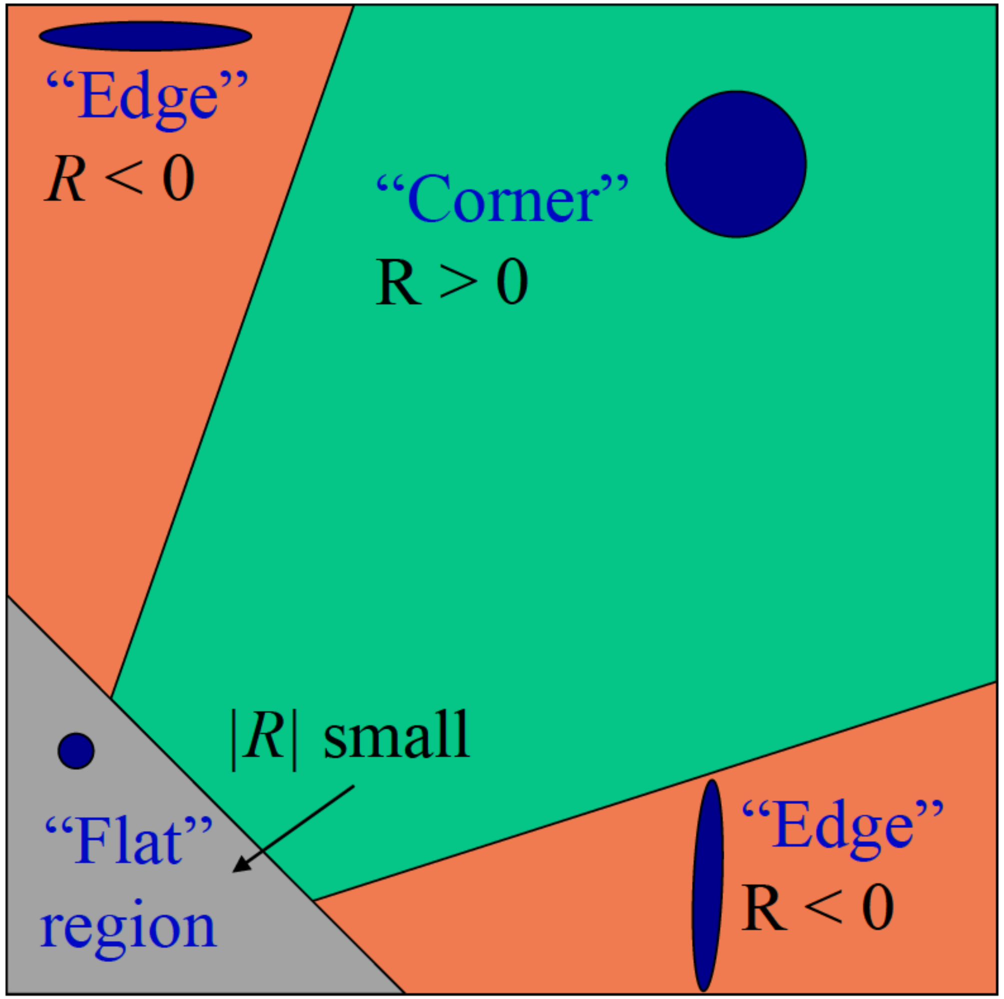
- Classification of image points using eigenvalues of M



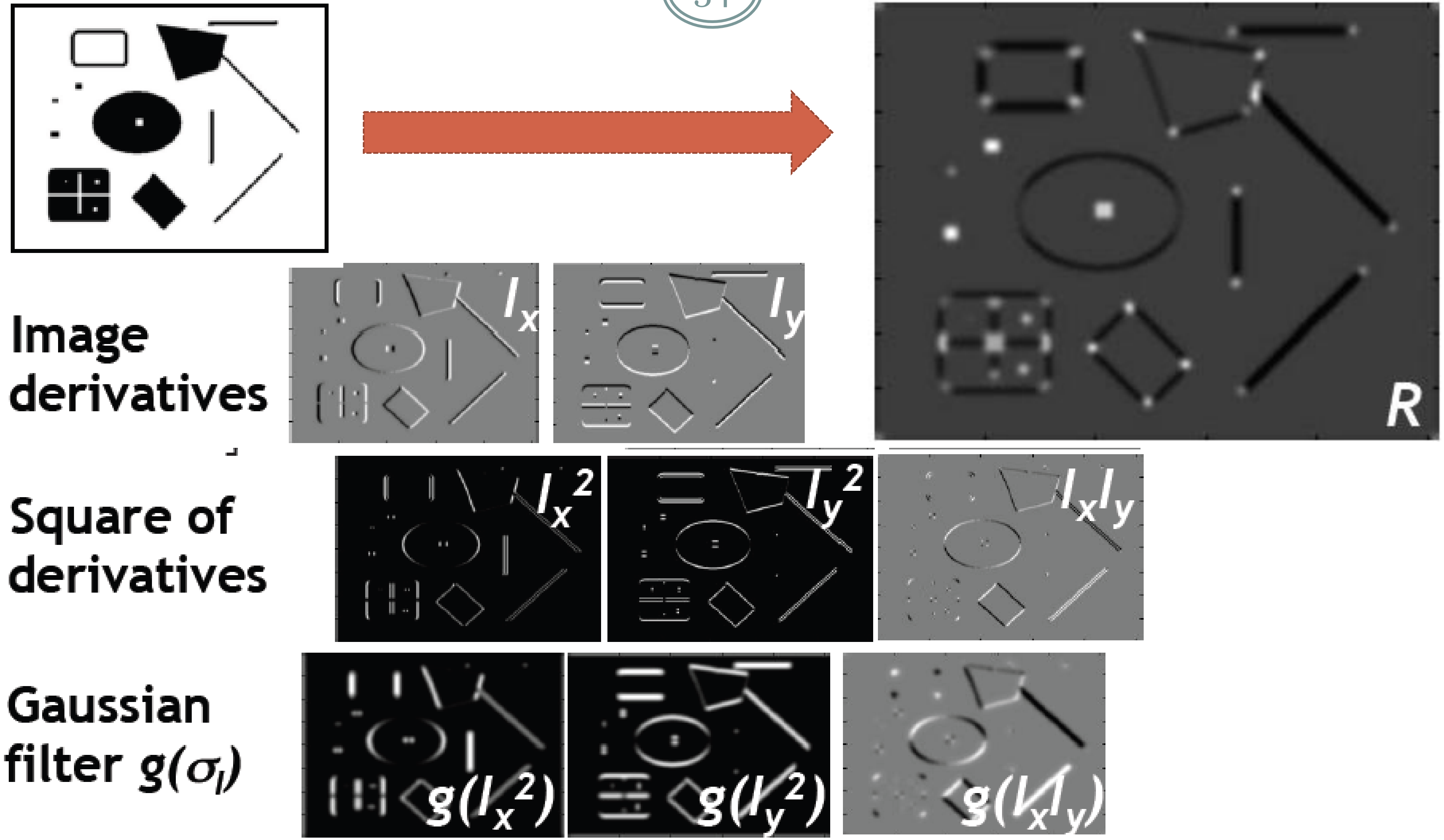
Harris Detector Formulation

$$R = \det(M) - \alpha\ \mathrm{trace}(M)^2$$
$$= \lambda_1\lambda_2 - \alpha\ (\lambda_1 + \lambda_2)^2$$

- Fast approximation
  - Avoid computing the eigenvalues
  - α: constant (0.04 to 0.06)

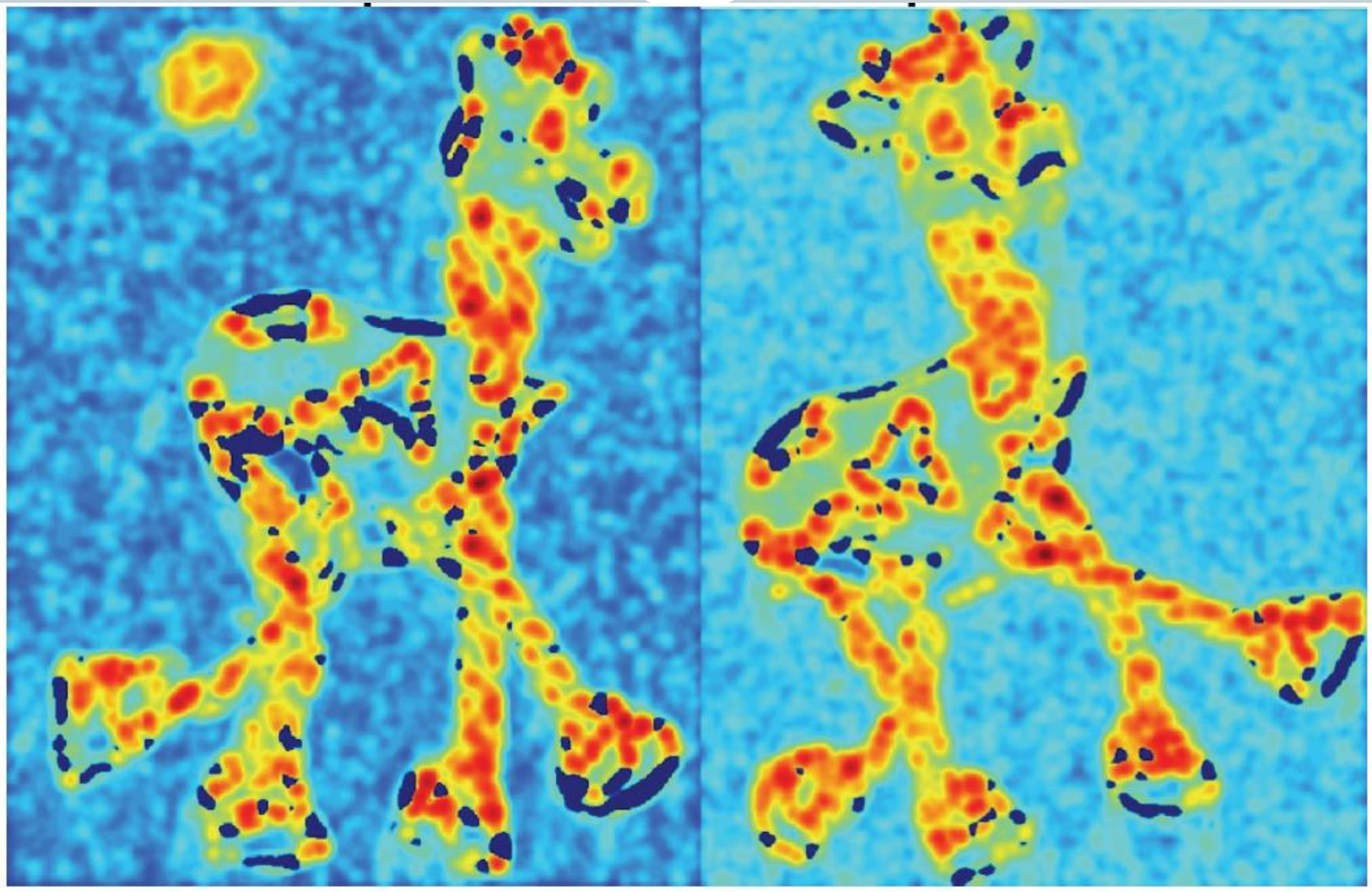


Summary: Harris Detector

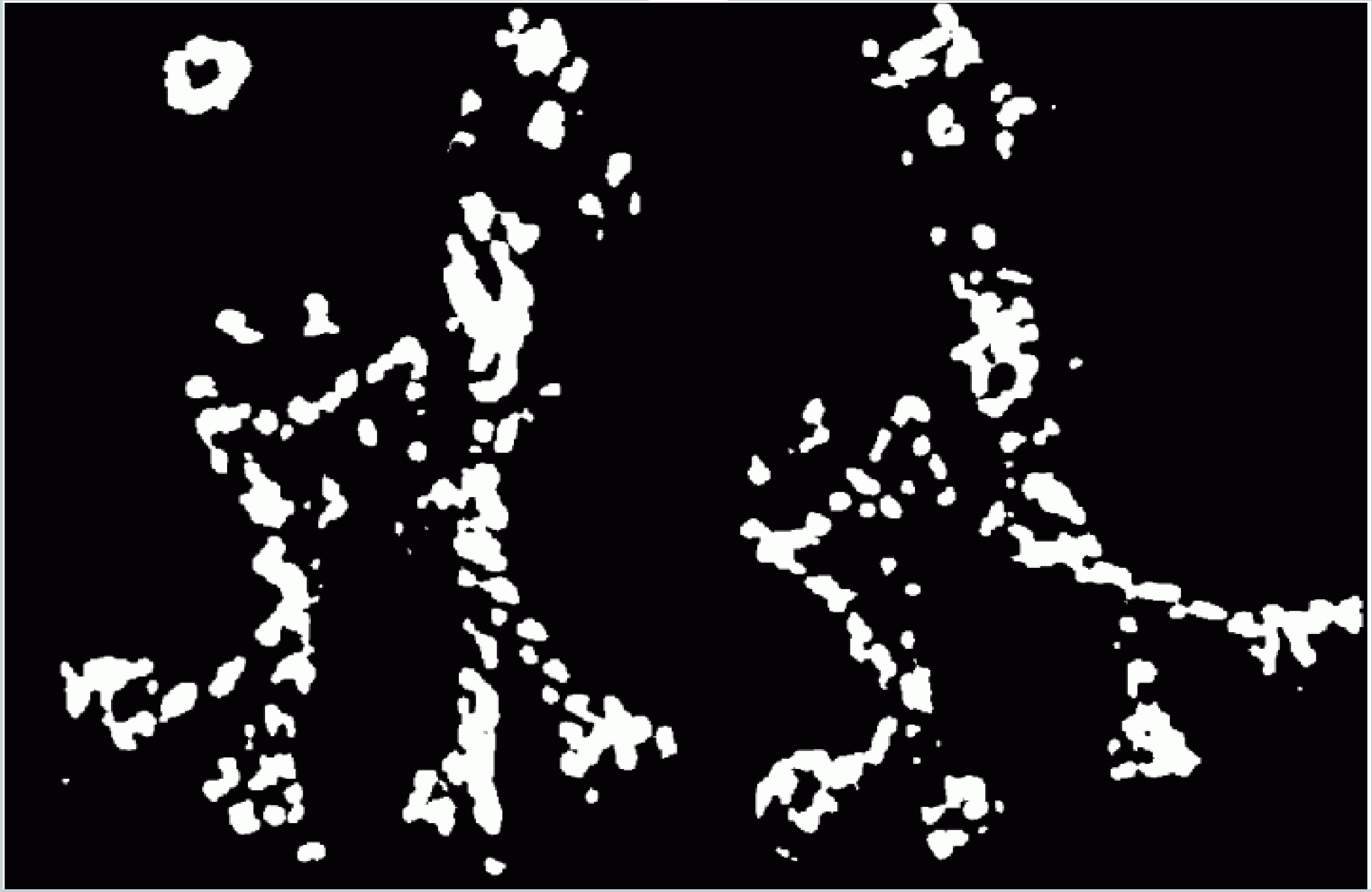


Harris Detector: Workflow

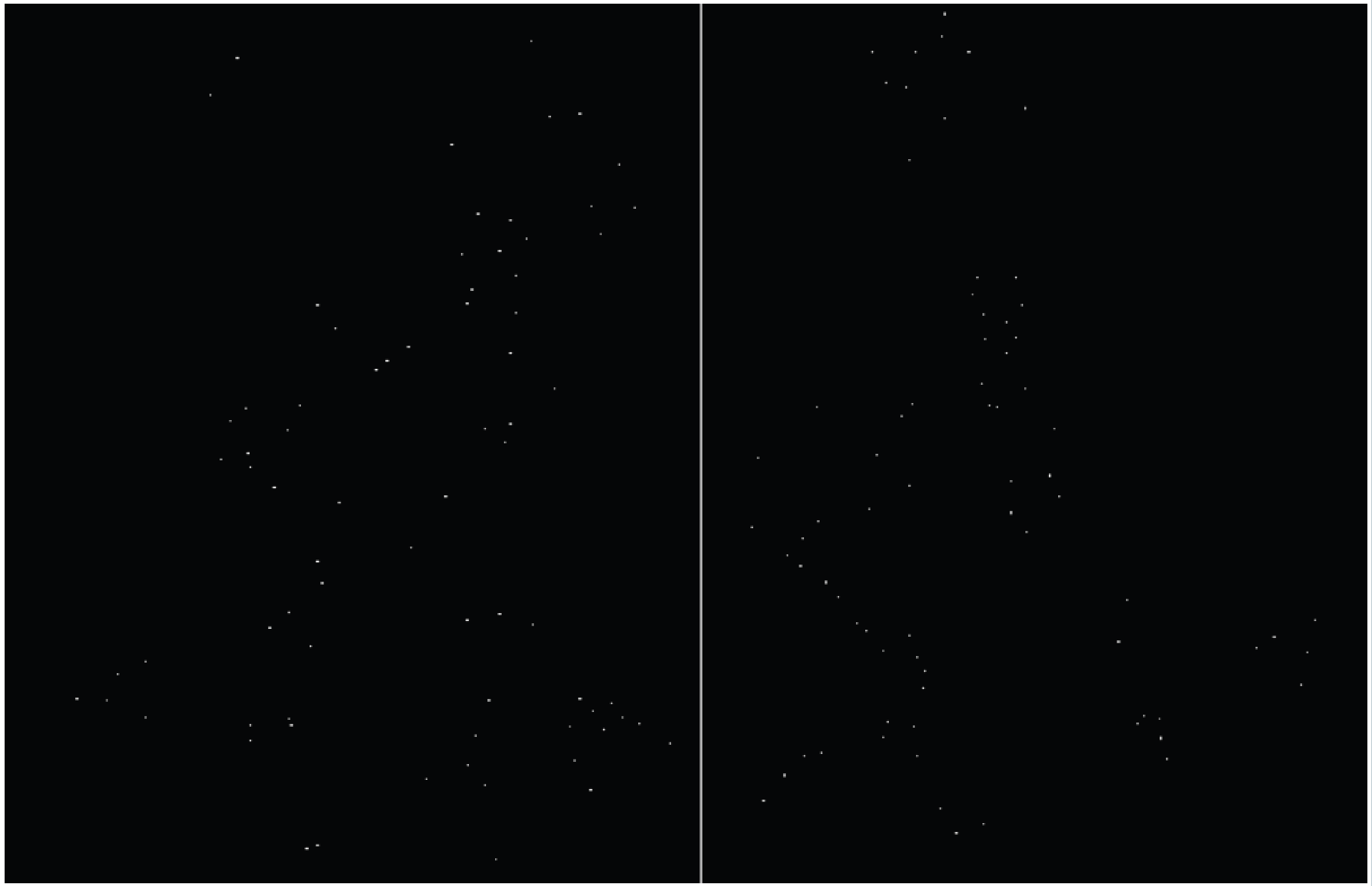
Compute corner responses R



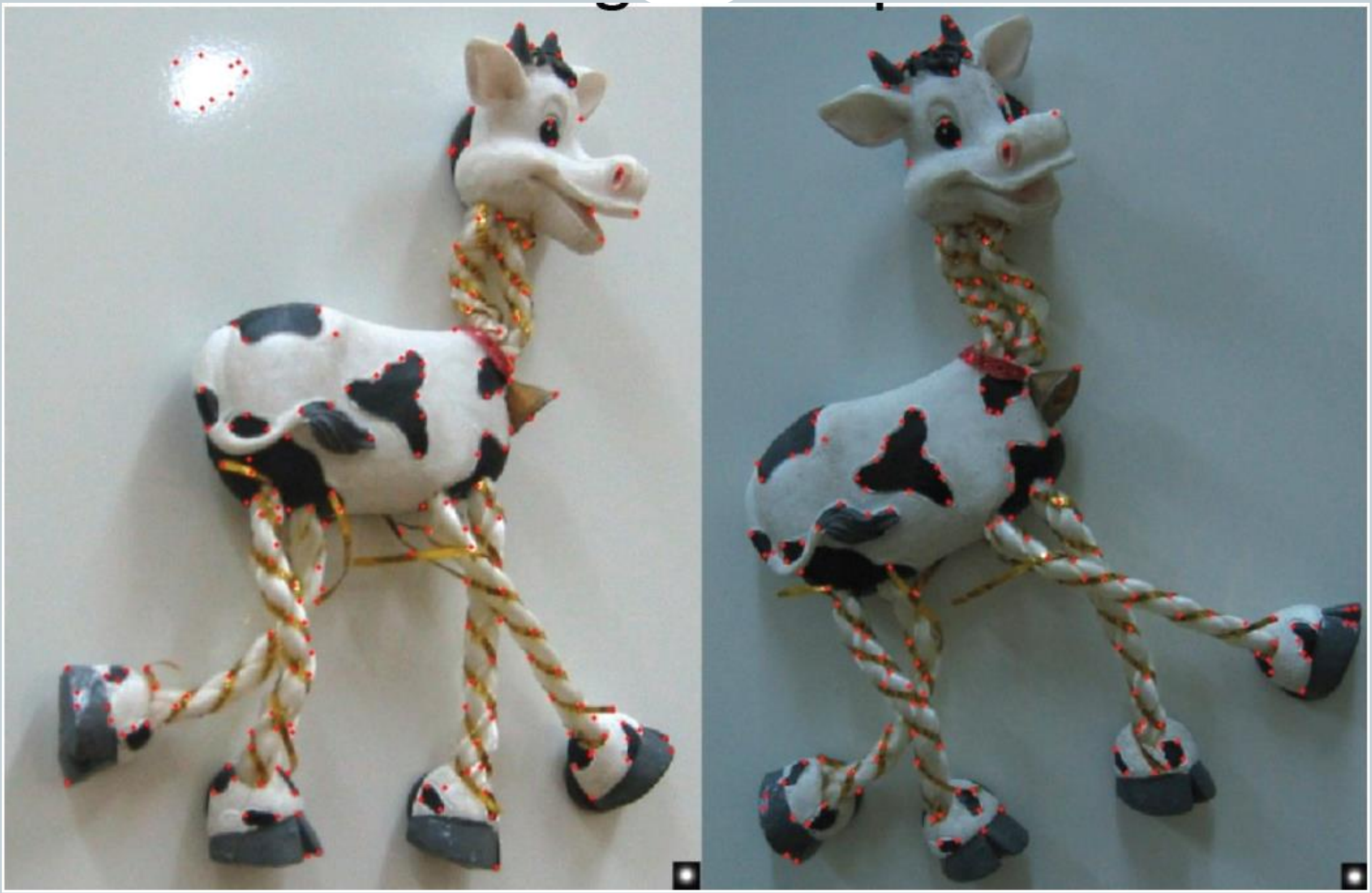
Select points with large corner response (R > t1)



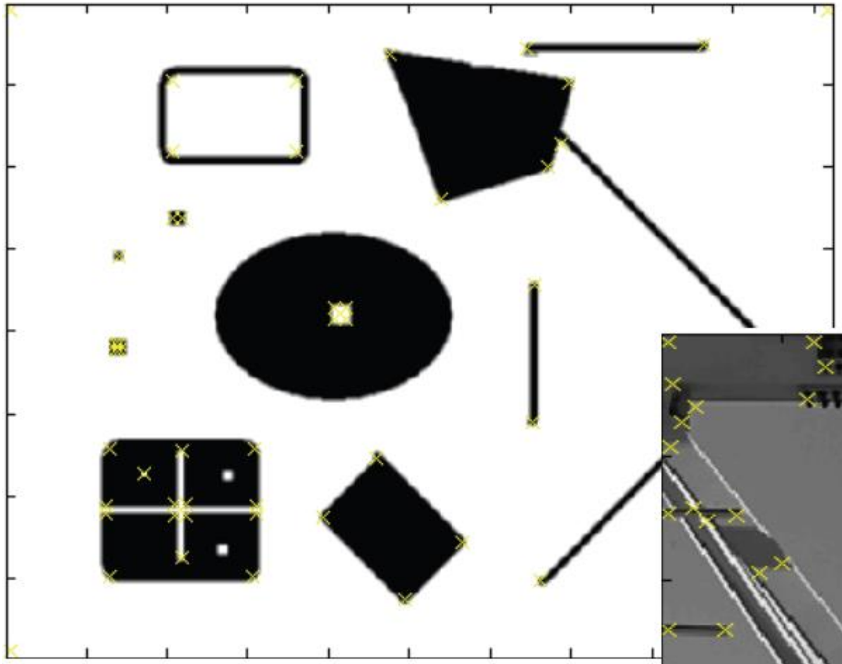
Take only the local maxima of R



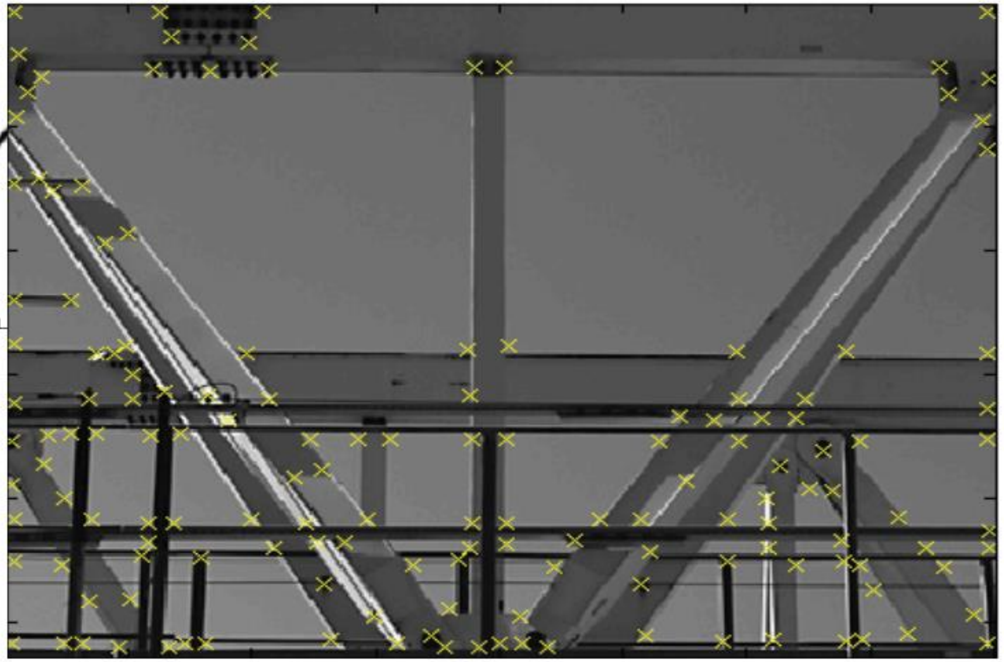
Resulting Harris corner points



More results …

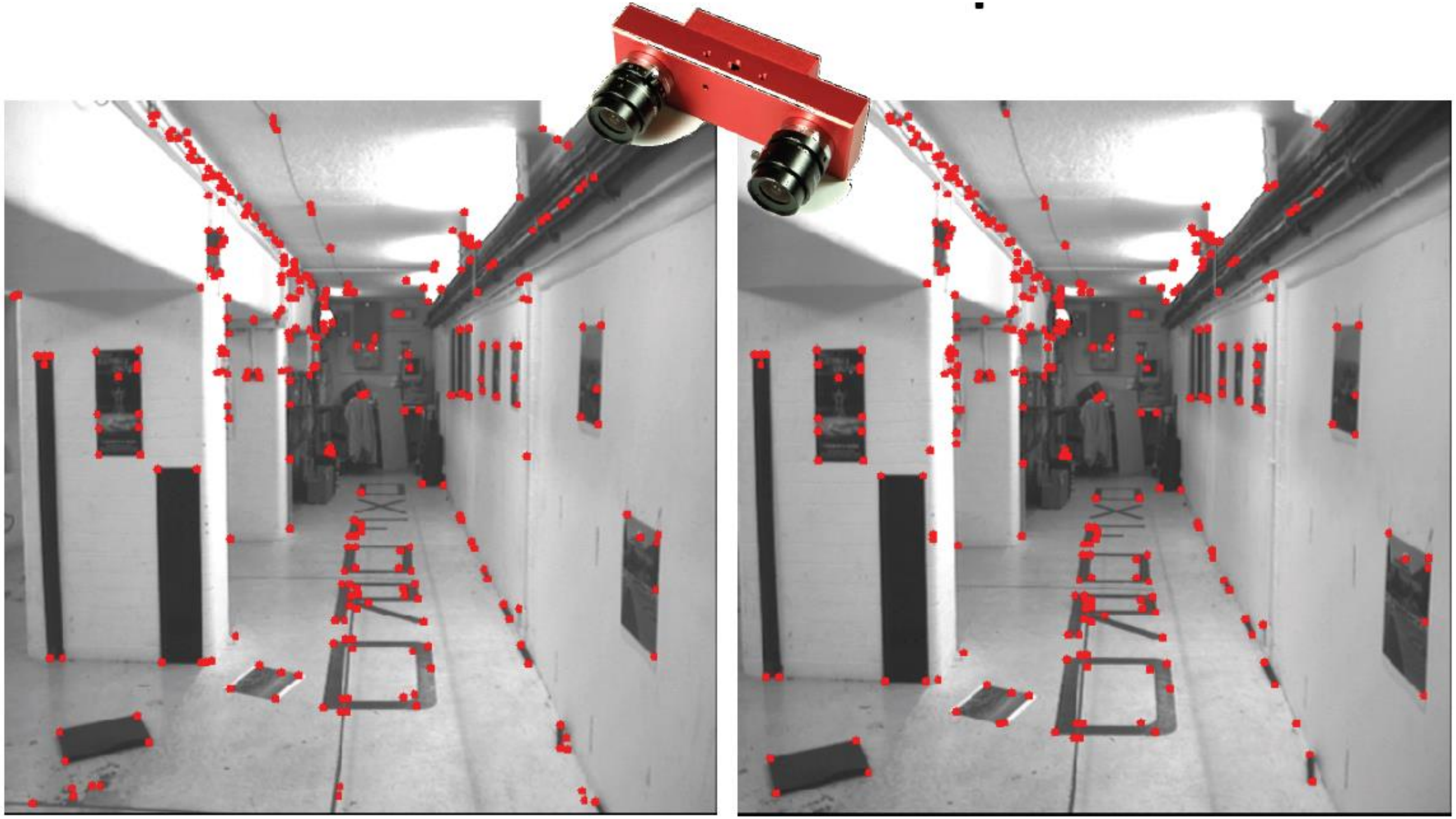


Effect: A very precise corner detector.



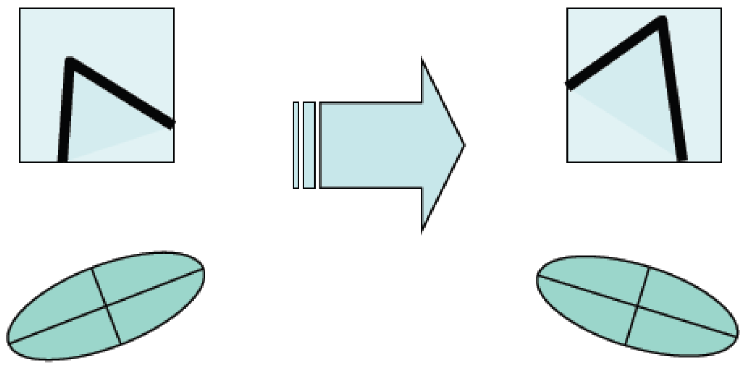
More results ...

Results are well suited for finding stereo correspondences



Harris Detector: Properties

- Translation invariance?
- Rotation invariance?

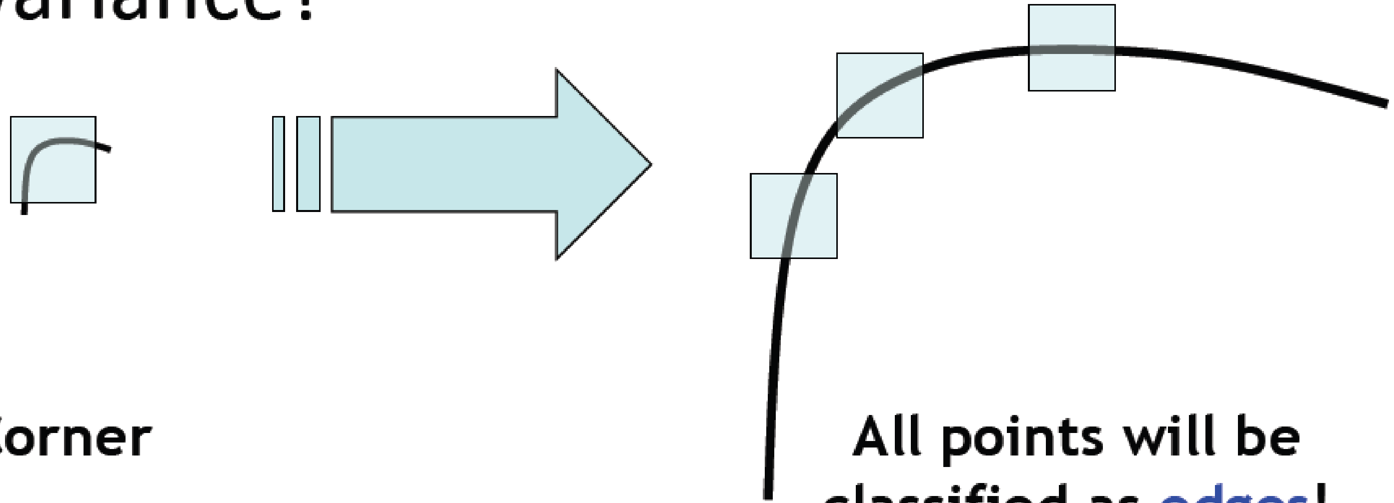


Ellipse rotates but its shape (i.e. eigenvalues) remains the same

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Harris Detector: Properties

- Scale invariance?



Corner

All points will be classified as edges!

Harris Detector: Properties

- Harris Corner detector is
  - Translation invariant?
    - YES ☺
  - Rotation invariant?
    - YES ☺
  - Scale invariant
    - NO ☹

We Learnt

- Local invariant features
  - Motivation
  - Requirements and invariances
- Keypoint/interest point detection
  - Harris corner detector
- Scale invariant region selection
  - Automatic scale selection
  - Difference of Gaussian Detector
- Scale Invariant Feature Transform (SIFT)

Mirza Tahir Ahmed, Zia Ud Din, Michael Greenspan

Resources & Reference

- C.Harris and M.Stephens. "A Combined Corner and Edge Detector.“ Proceedings of the 4th Alvey Vision Conference, 1988

- Slides source credits:
  - Fei-Fei Li
  - David Lowe
  - Robert Collins
  - Mubarak Shah
  - Chang Shu