szeliski 2.1.5

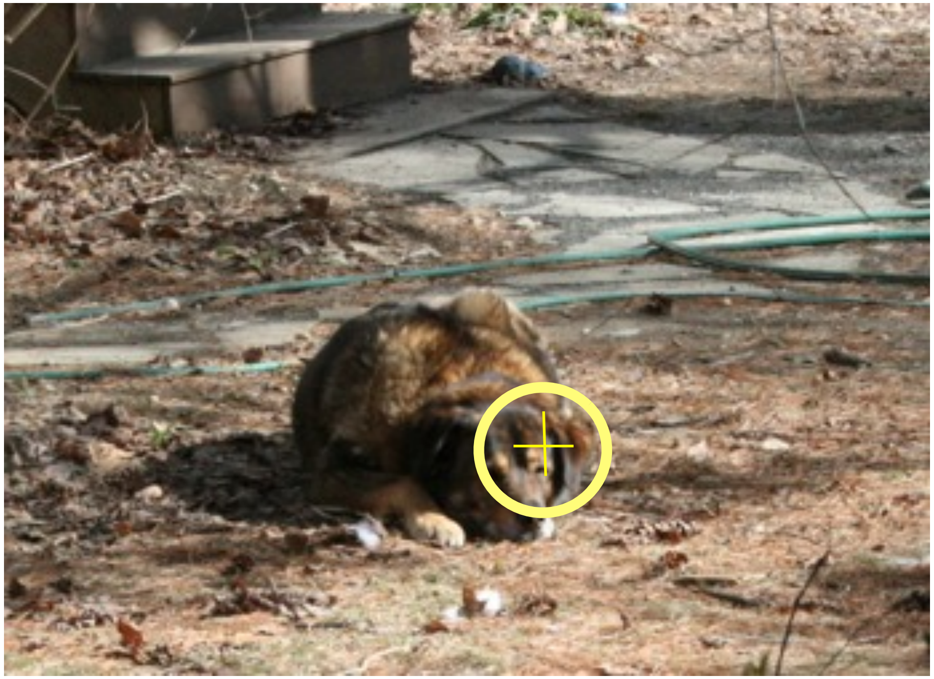
Image Features

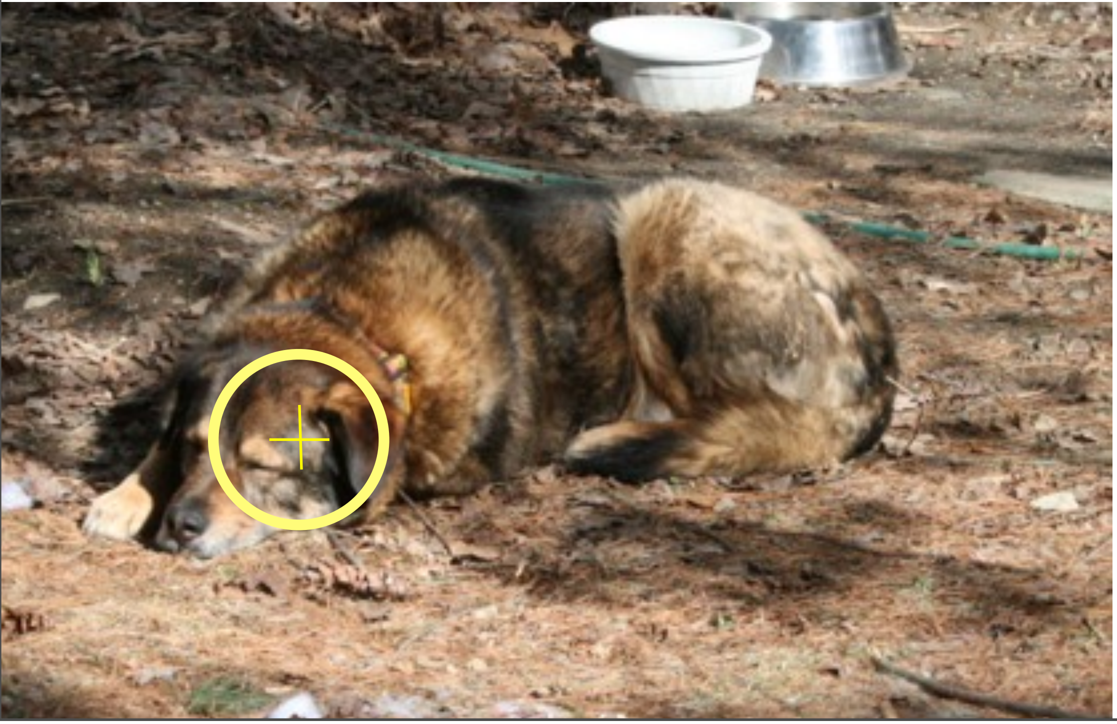
Computer Vision 6.869

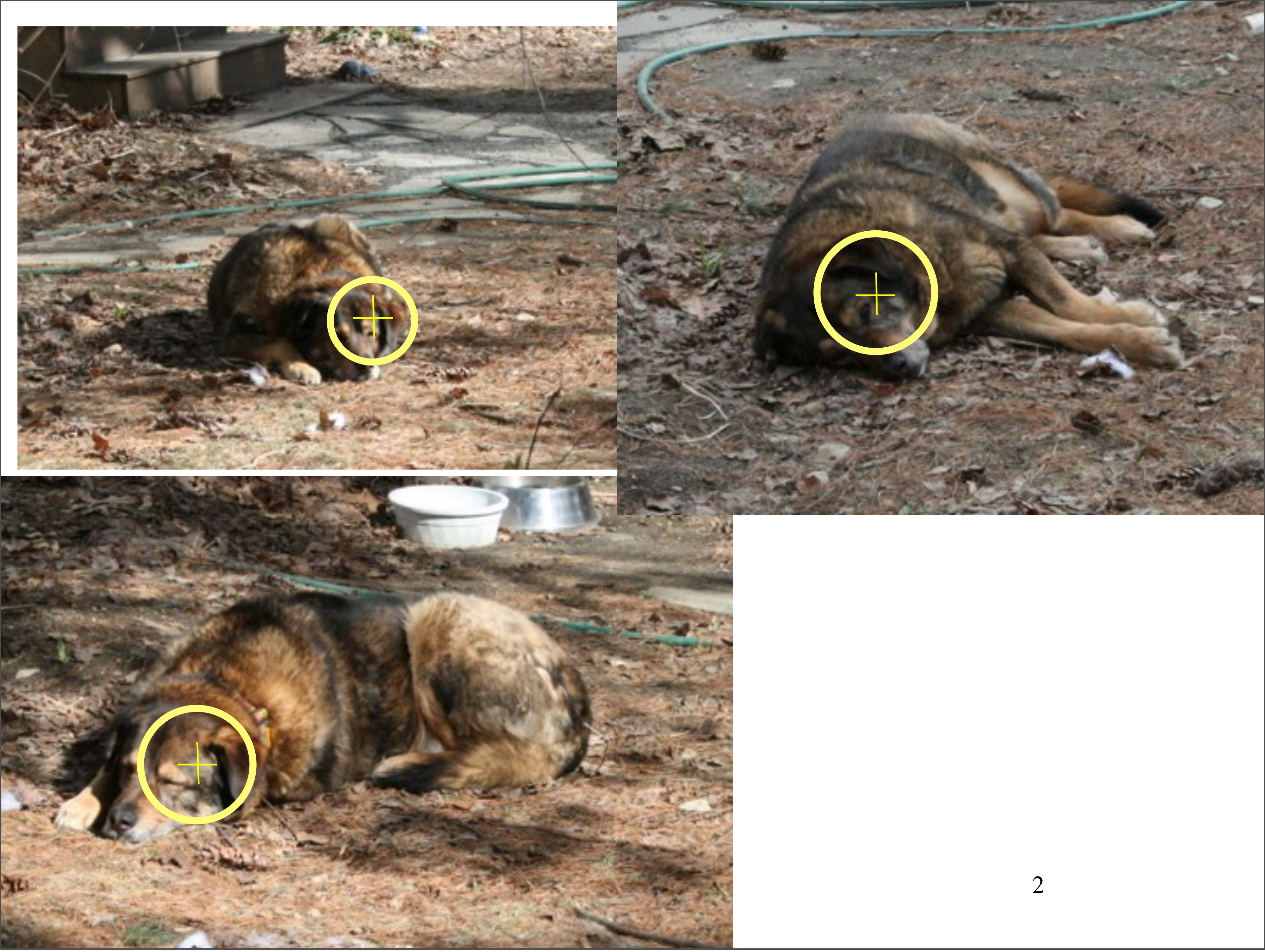
Bill Freeman

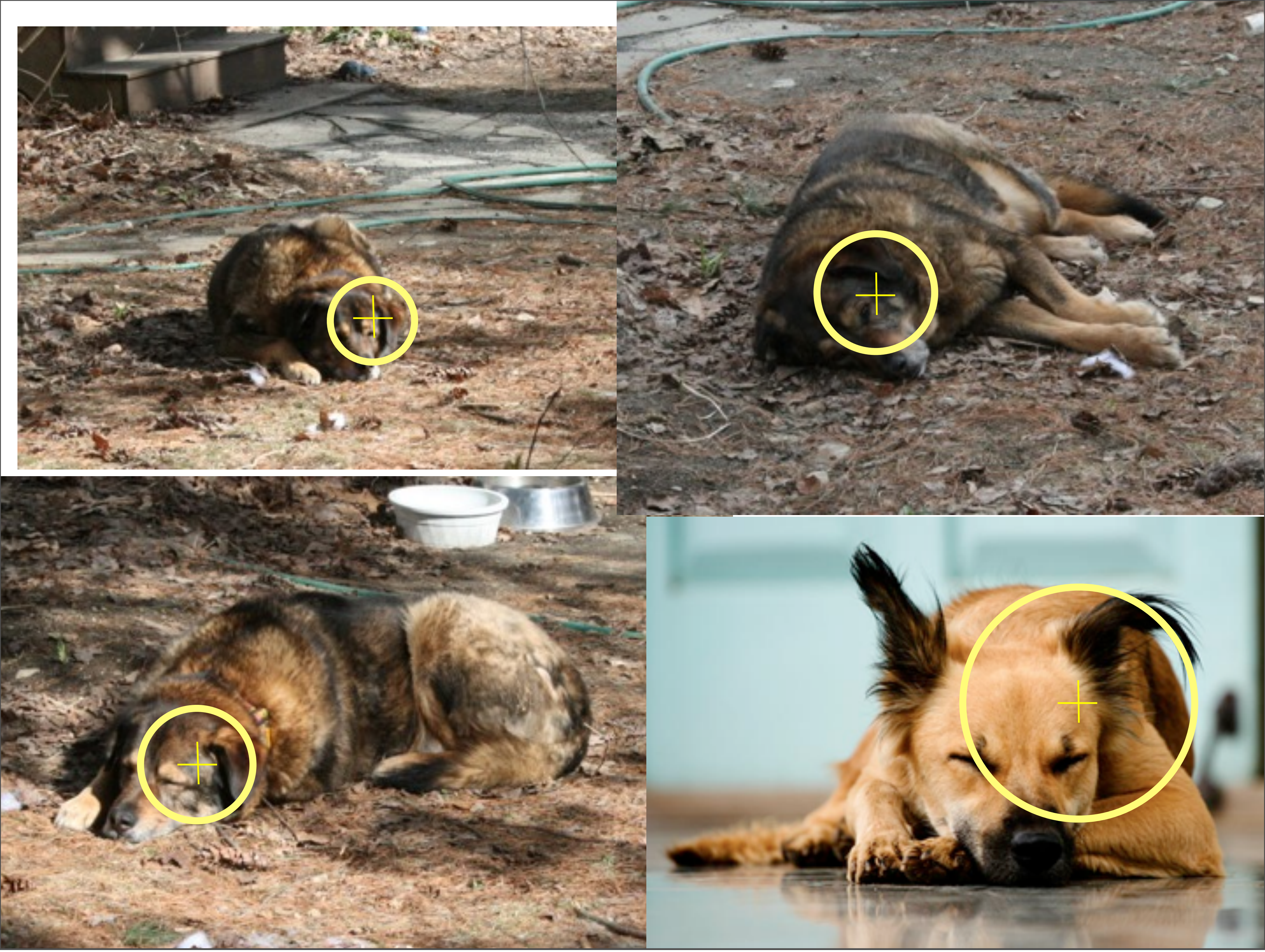
Image and shape descriptors:
Harris corner detectors and SIFT features.

With slides from
Darya Frolova, Denis Simakov, David Lowe







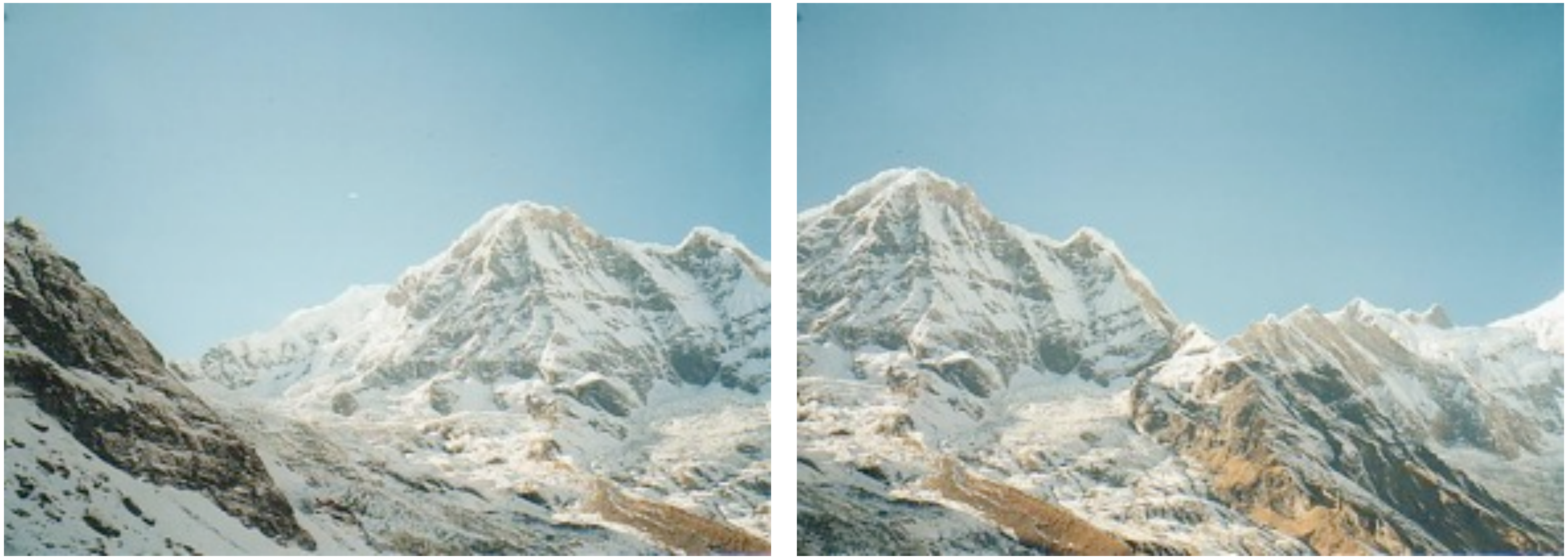


Uses for feature point detectors and descriptors in computer vision and graphics.

- Image alignment and building panoramas
- 3D reconstruction
- Motion tracking
- Object recognition
- Indexing and database retrieval
- Robot navigation
- … other

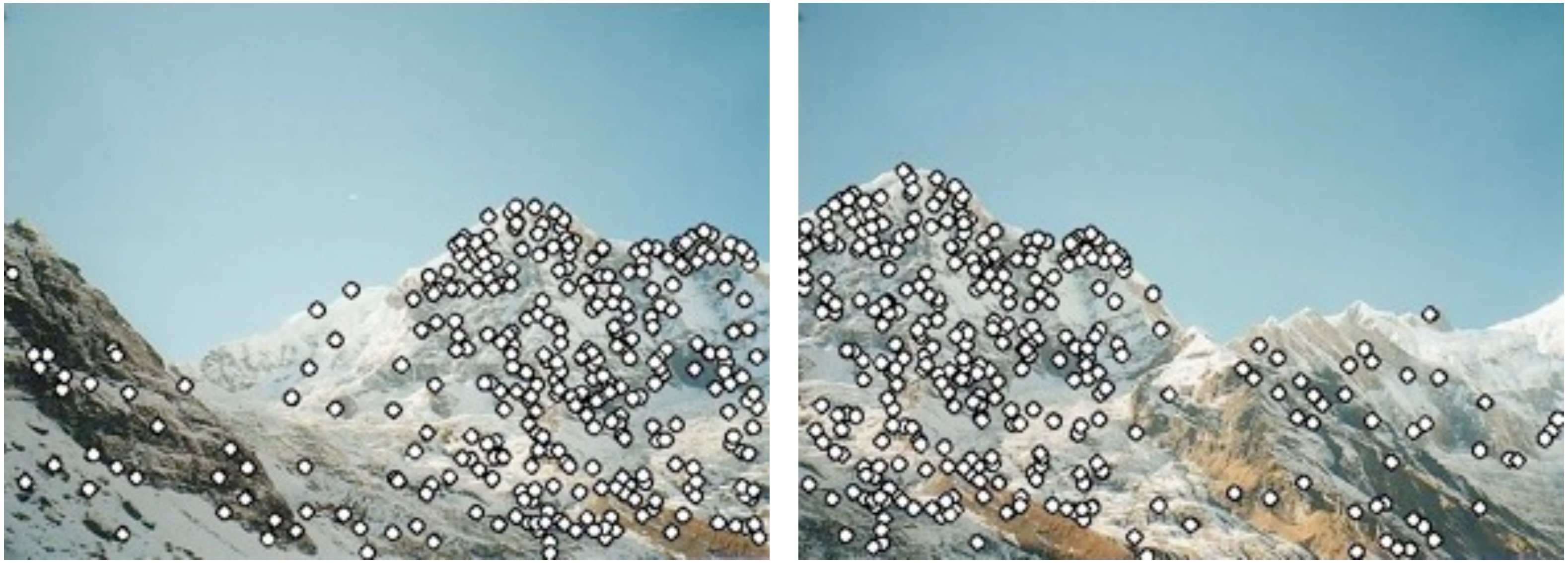
How do we build a panorama?

- We need to match (align) images



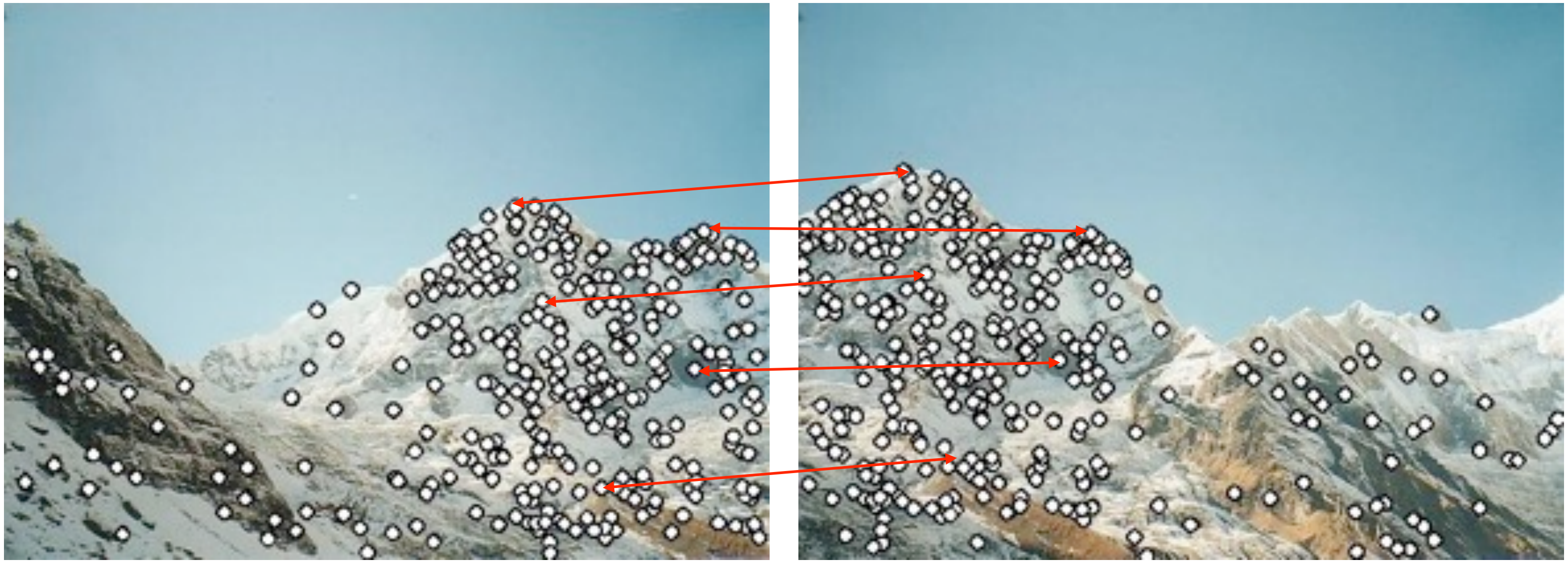
Matching with Features

- Detect feature points in both images



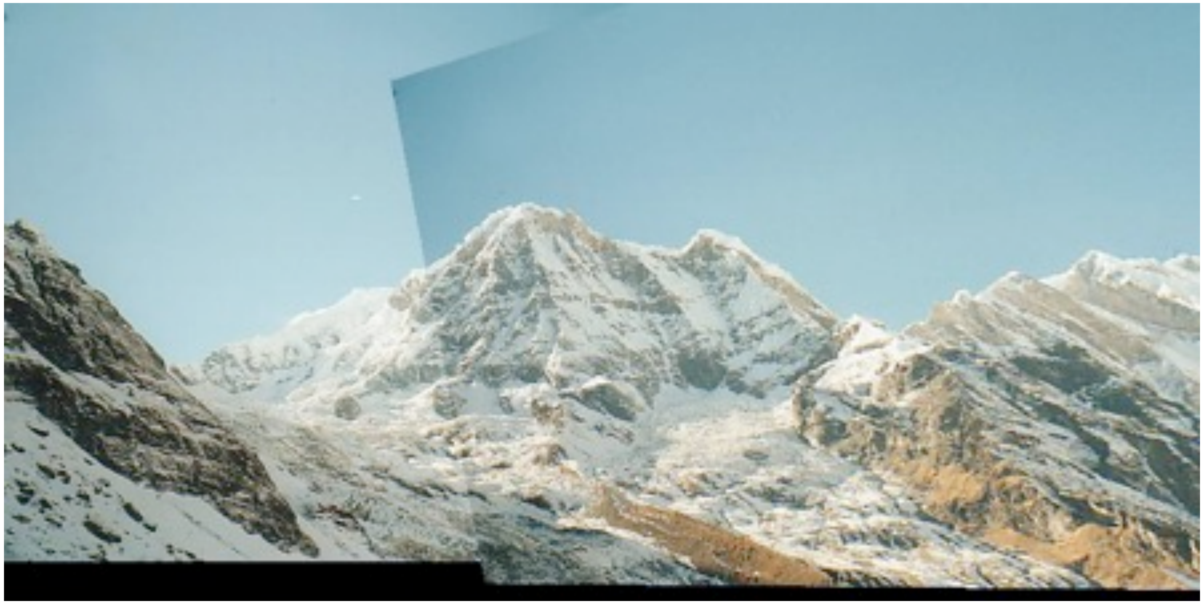
Matching with Features

- Detect feature points in both images
- Find corresponding pairs



Matching with Features

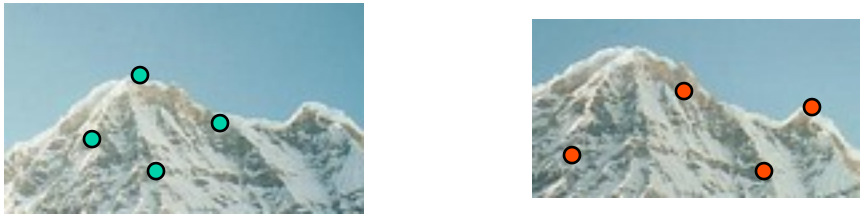
- •Detect feature points in both images
- •Find corresponding pairs
- •Use these matching pairs to align images - the required mapping is called a homography.



Matching with Features

- Problem 1:
  - Detect the *same* point *independently* in both images

counter-example:

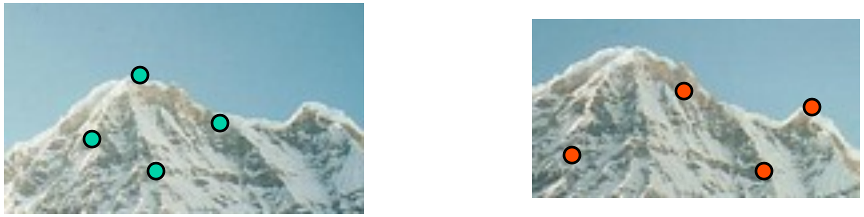


no chance to match!

Matching with Features

- Problem 1:
  - Detect the *same* point *independently* in both images

counter-example:

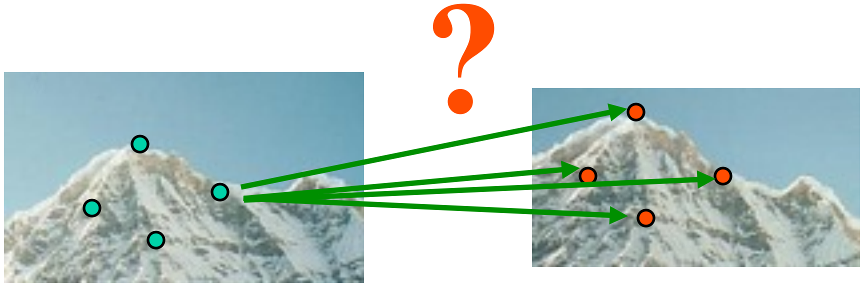


no chance to match!

We need a repeatable **detector**

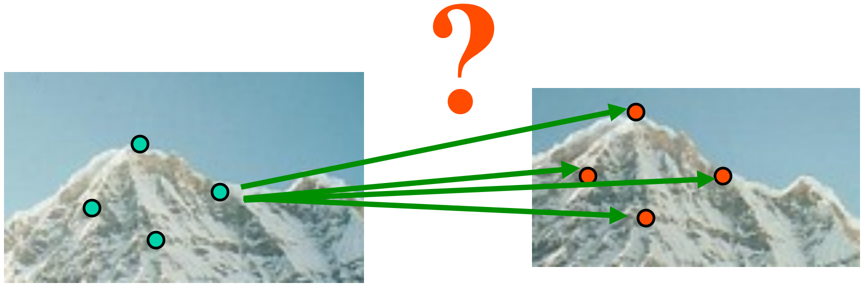
Matching with Features

- Problem 2:
  - For each point correctly recognize the corresponding one



Matching with Features

- Problem 2:
  - For each point correctly recognize the corresponding one

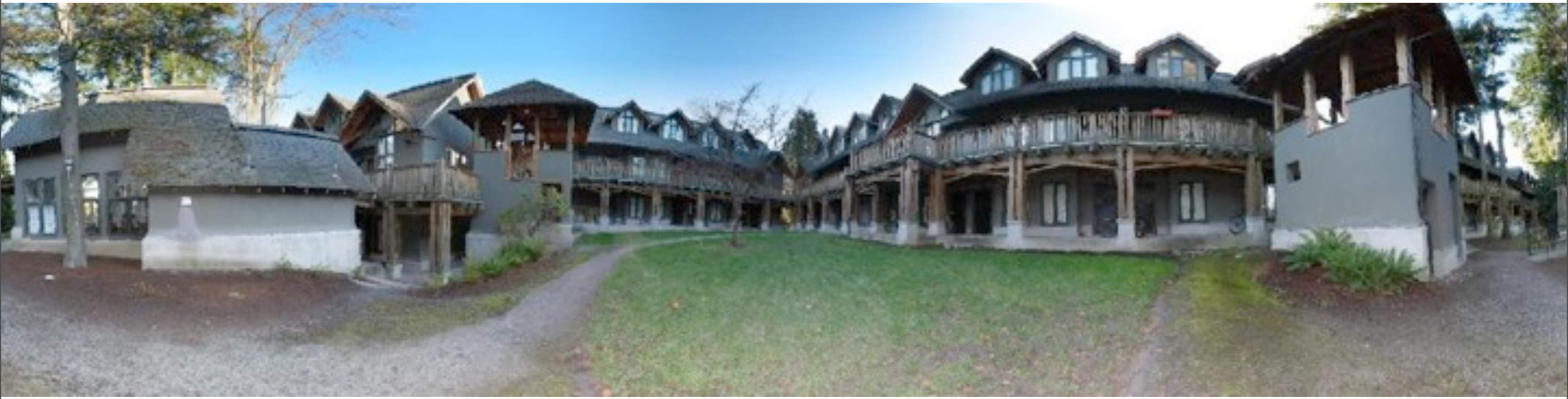


We need a reliable and distinctive **descriptor**

Building a Panorama

M. Brown and D. G. Lowe. Recognising Panoramas. ICCV 2003

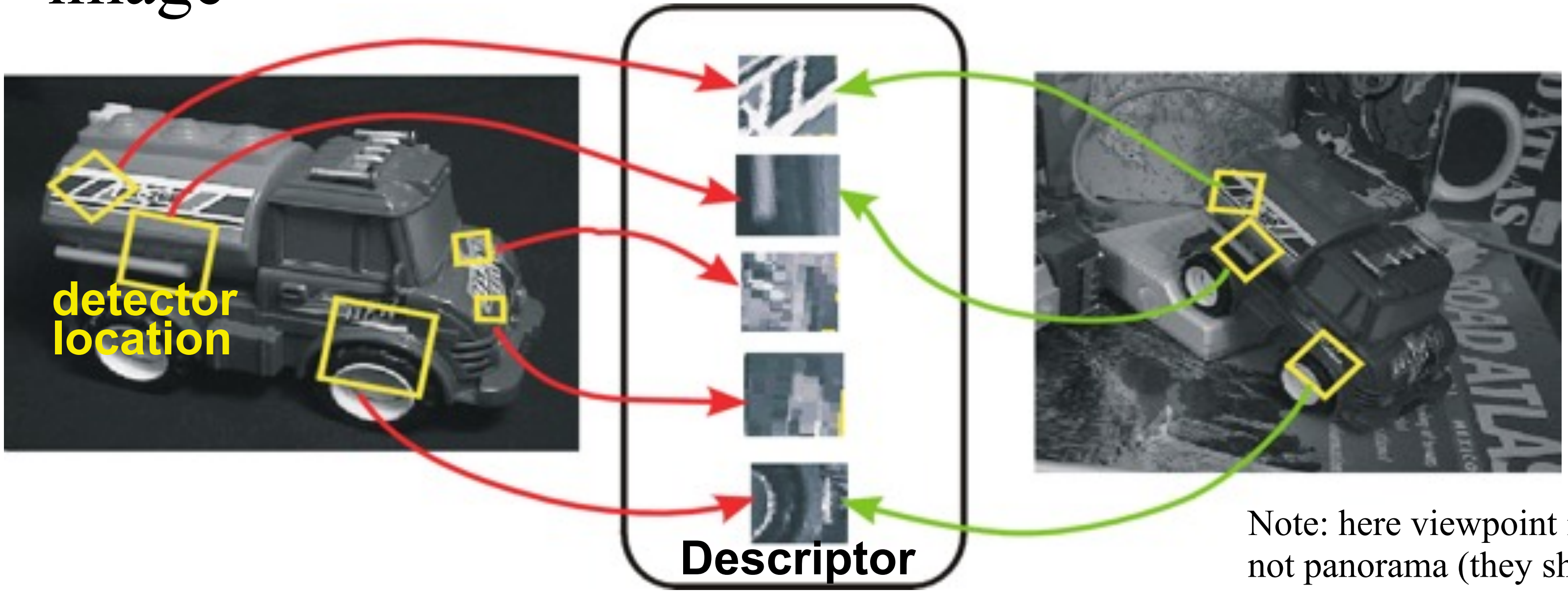
Building a Panorama



M. Brown and D. G. Lowe. Recognising Panoramas. ICCV 2003

Preview

- **Detector**: detect same scene points independently in both images
- **Descriptor**: encode local neighboring window
  - Note how scale & rotation of window are the same in both image (but computed independently)
- **Correspondence**: find most similar descriptor in other image



Note: here viewpoint is different, not panorama (they show off)

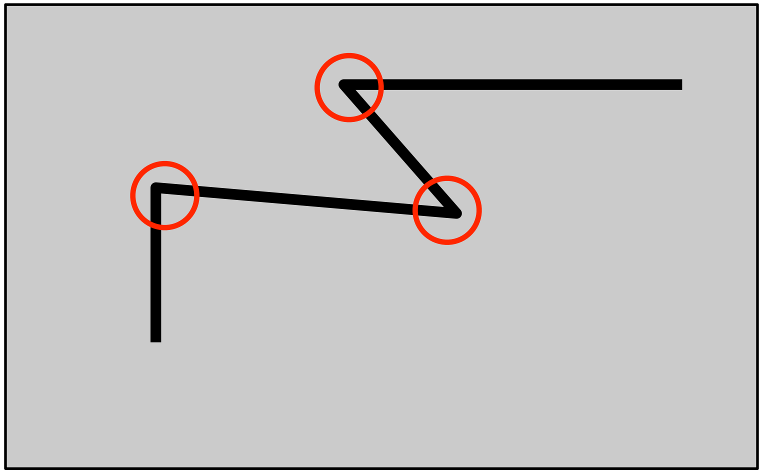
Outline

- Feature point <u>detection</u>
  - – Harris corner detector
  - – finding a characteristic scale
- Local image <u>description</u>
  - – SIFT features

Darya Frolova, Denis Simakov The Weizmann Institute of Science

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

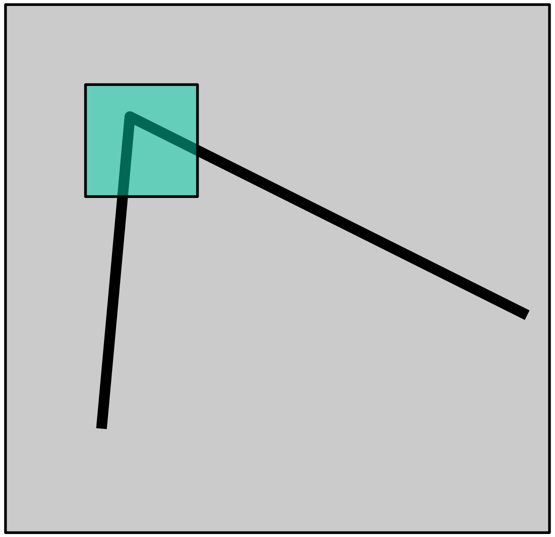
Harris corner detector



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

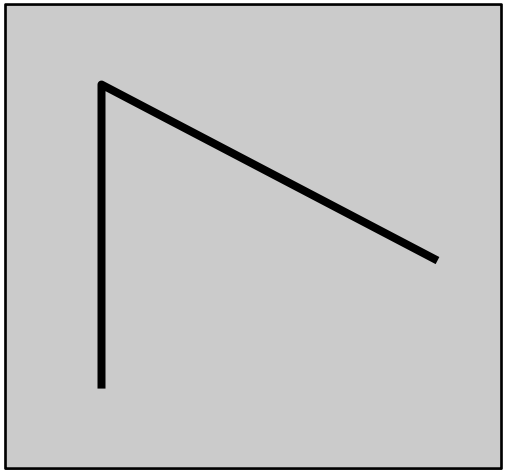
The Basic Idea

- We should easily localize the point by looking through a small window
- Shifting a window in *any direction* should give *a large change* in pixels intensities in window
  - – makes location precisely define



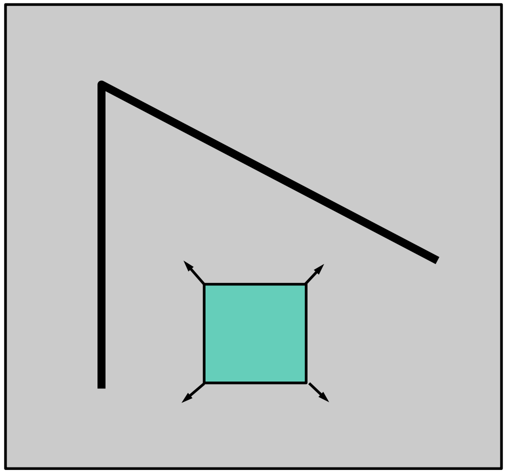
Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



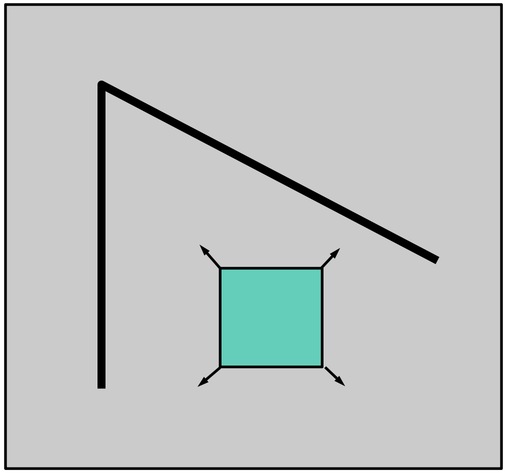
Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



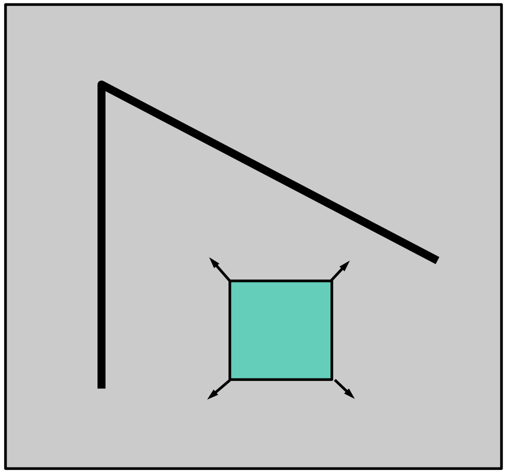
Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



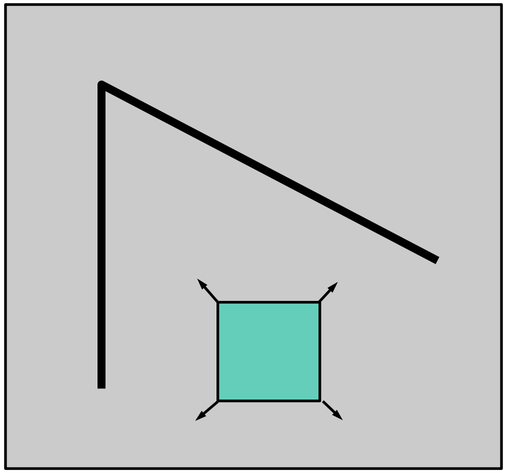
Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



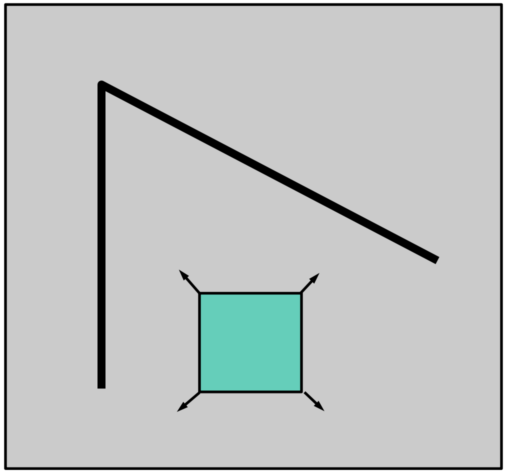
Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



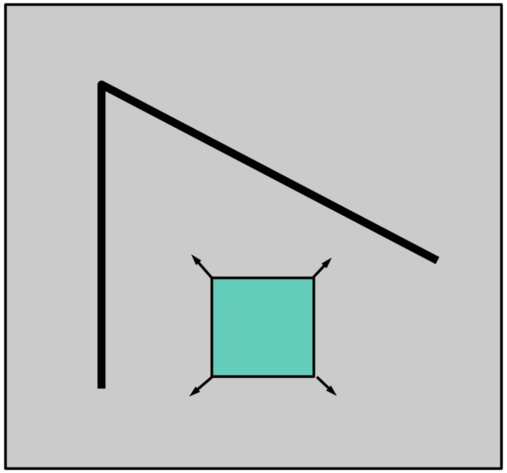
Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

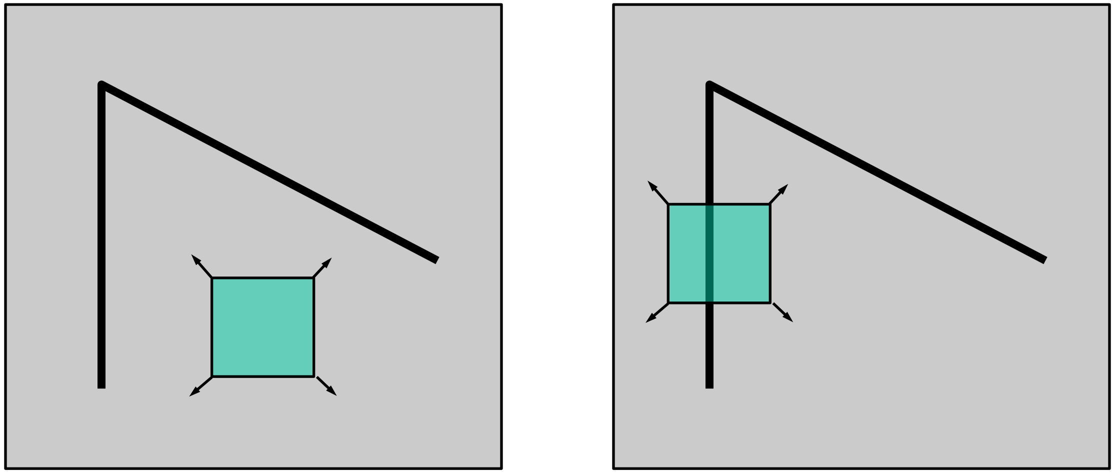
Corner Detector: Basic Idea



“flat” region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

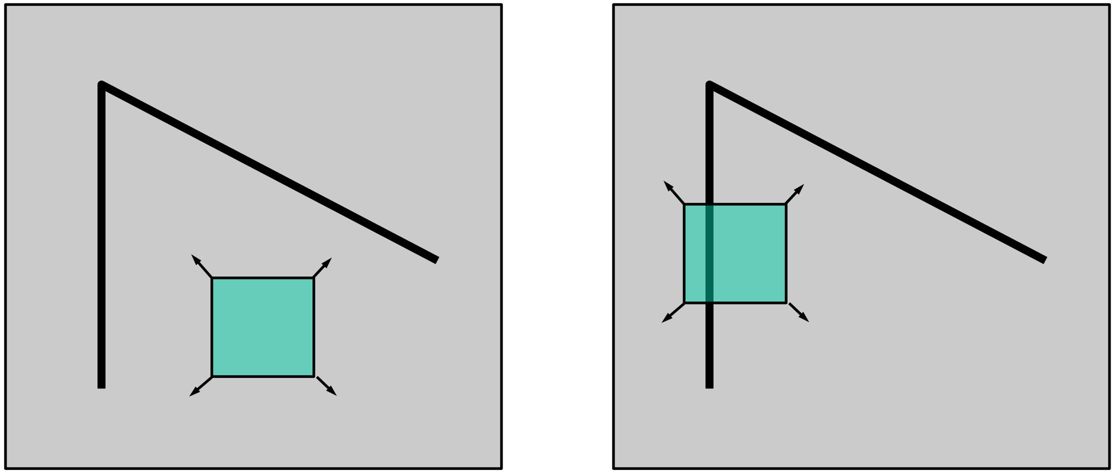
Corner Detector: Basic Idea



“flat” region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

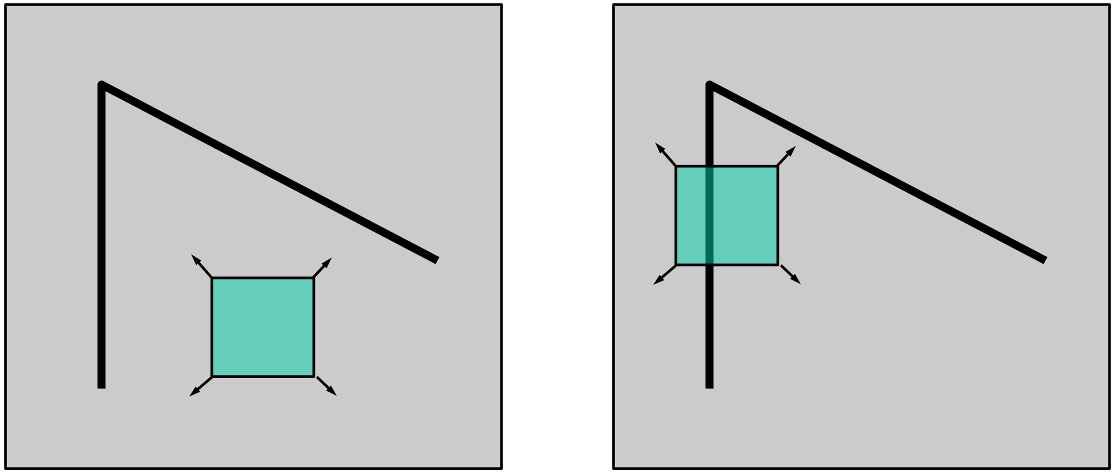
Corner Detector: Basic Idea



"flat" region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

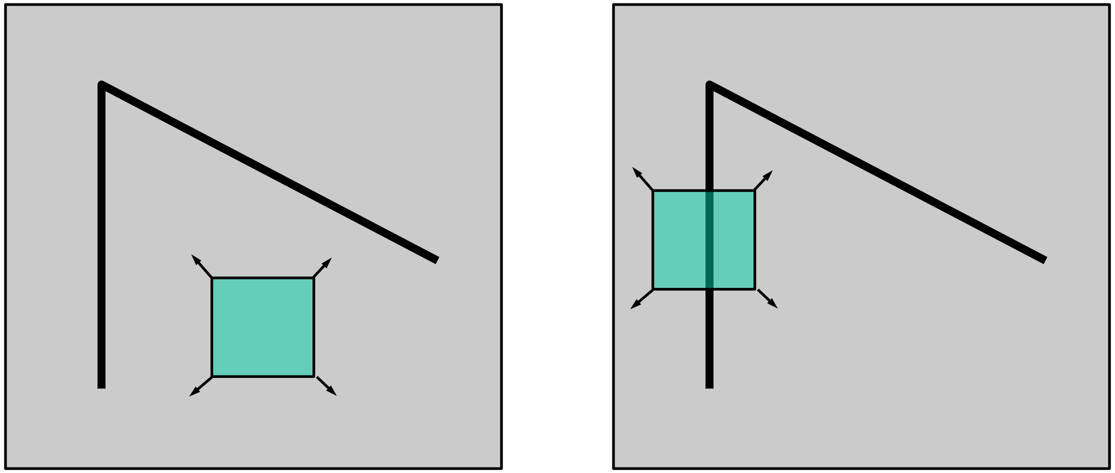
Corner Detector: Basic Idea



“flat” region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

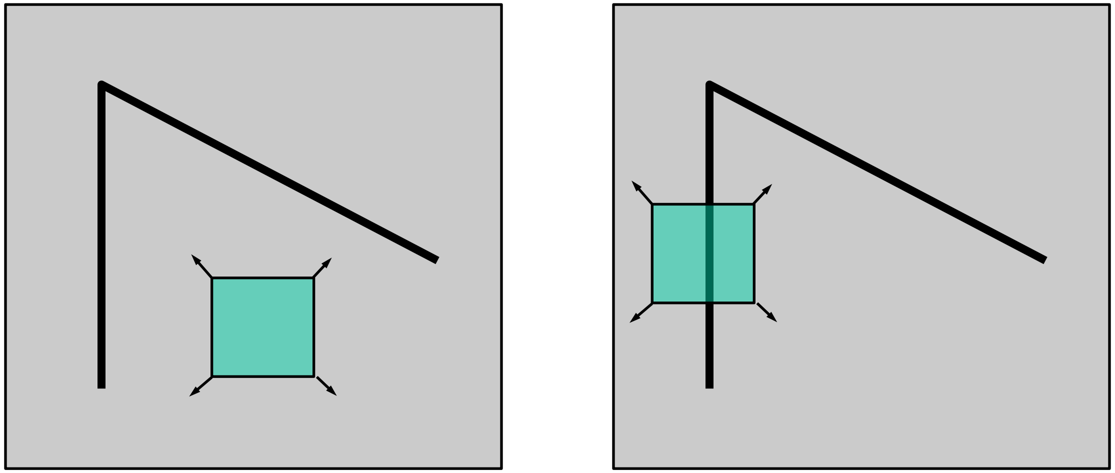
Corner Detector: Basic Idea



"flat" region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

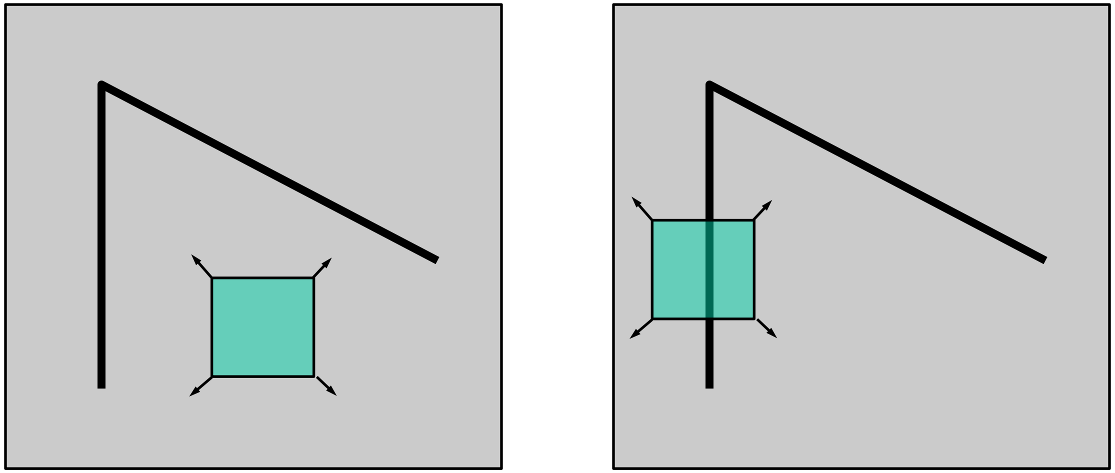
Corner Detector: Basic Idea



Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

“flat” region:
no change in
all directions

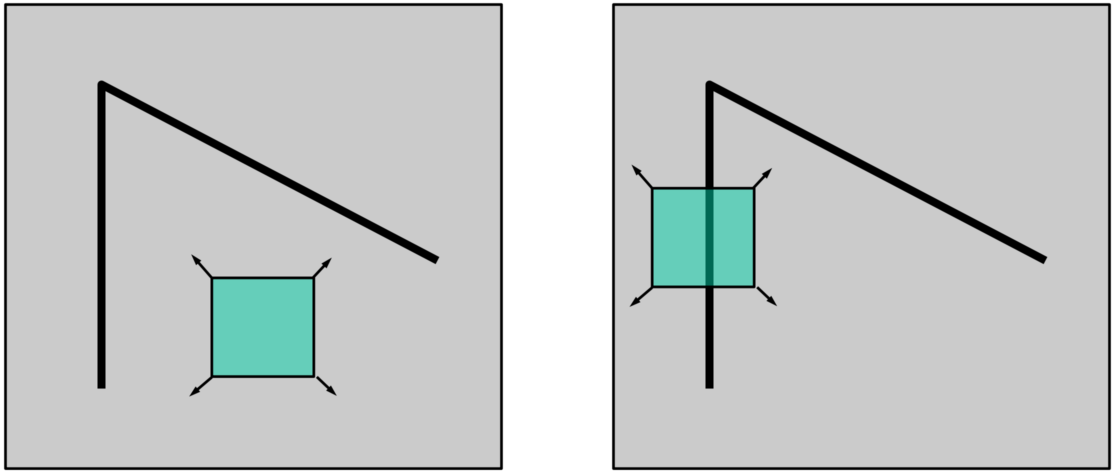
Corner Detector: Basic Idea



"flat" region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

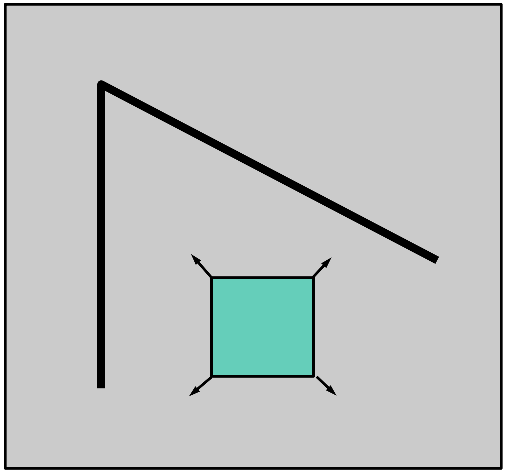
Corner Detector: Basic Idea



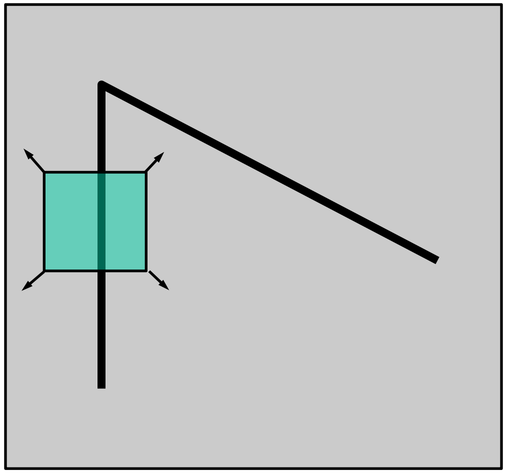
“flat” region:
no change in
all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



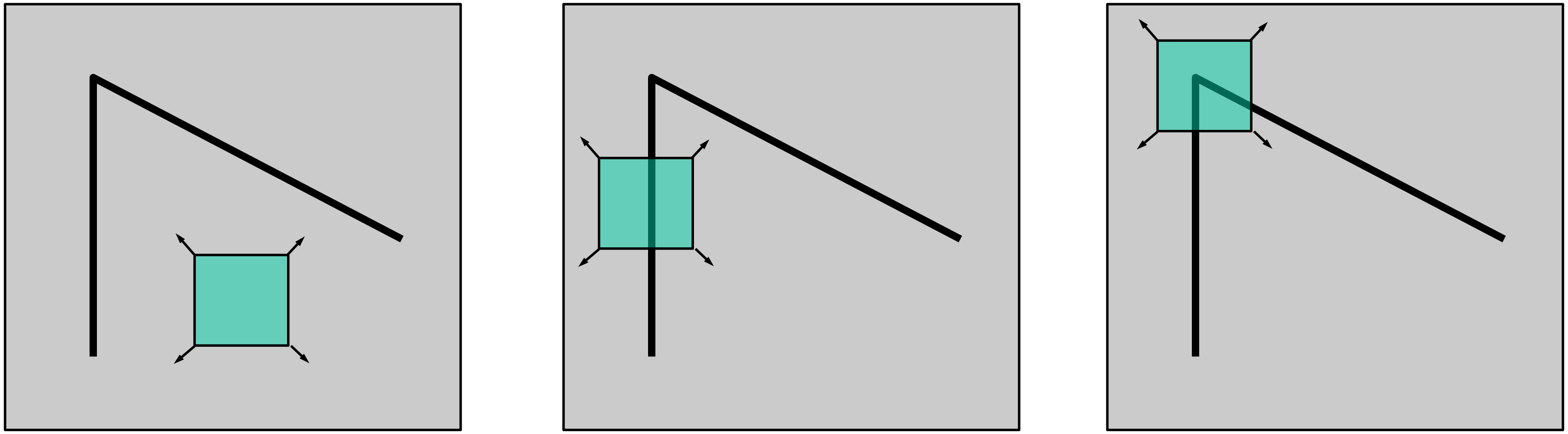
"flat" region:
no change in
all directions



"edge":
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

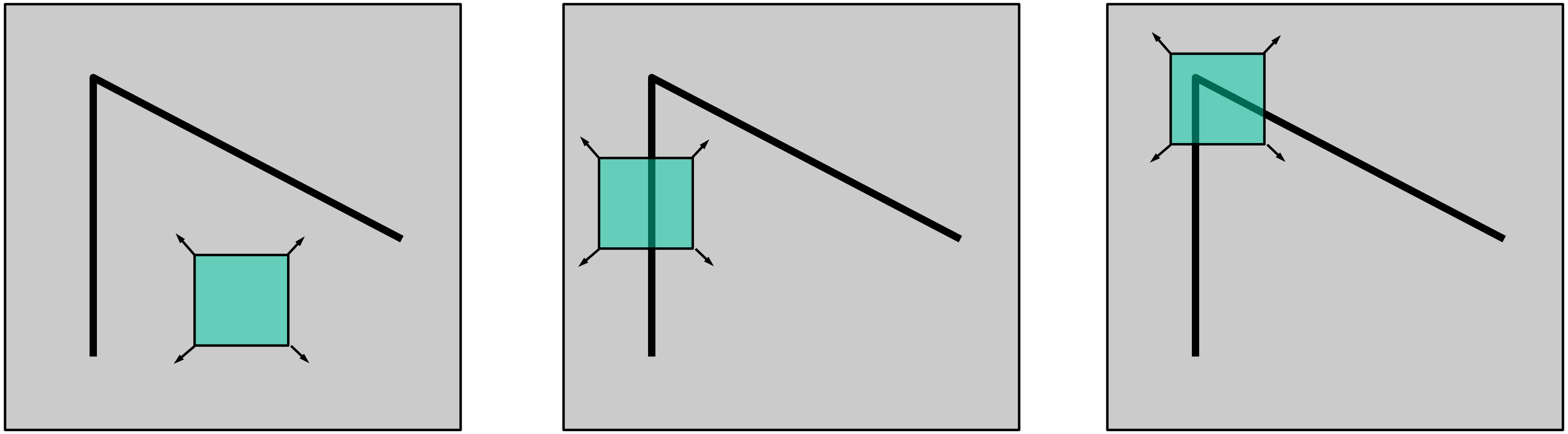


“flat” region:
no change in
all directions

“edge”:
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

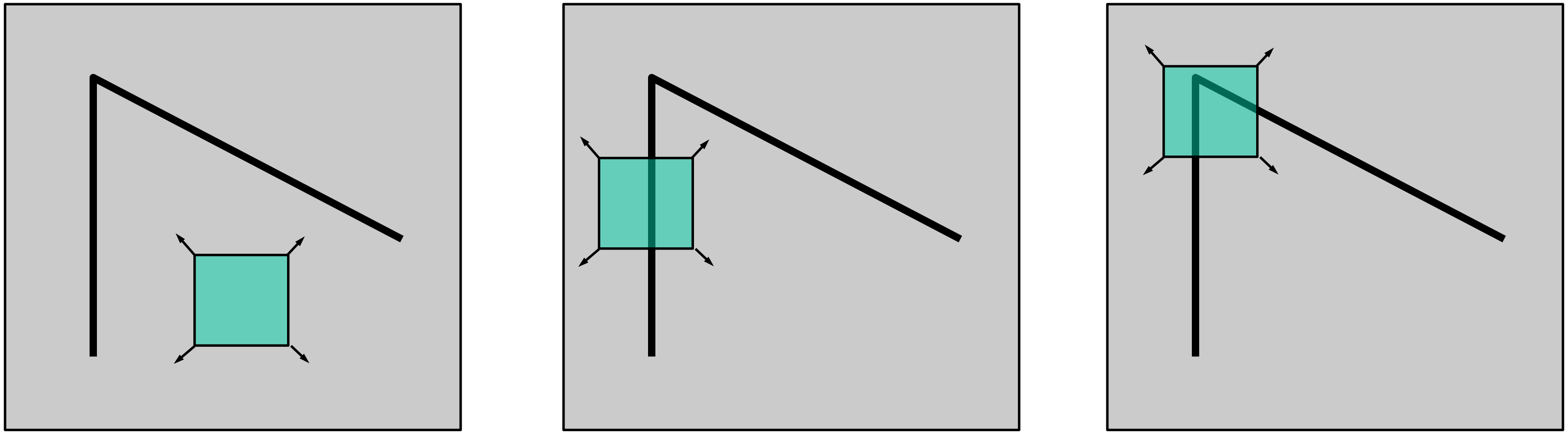


"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

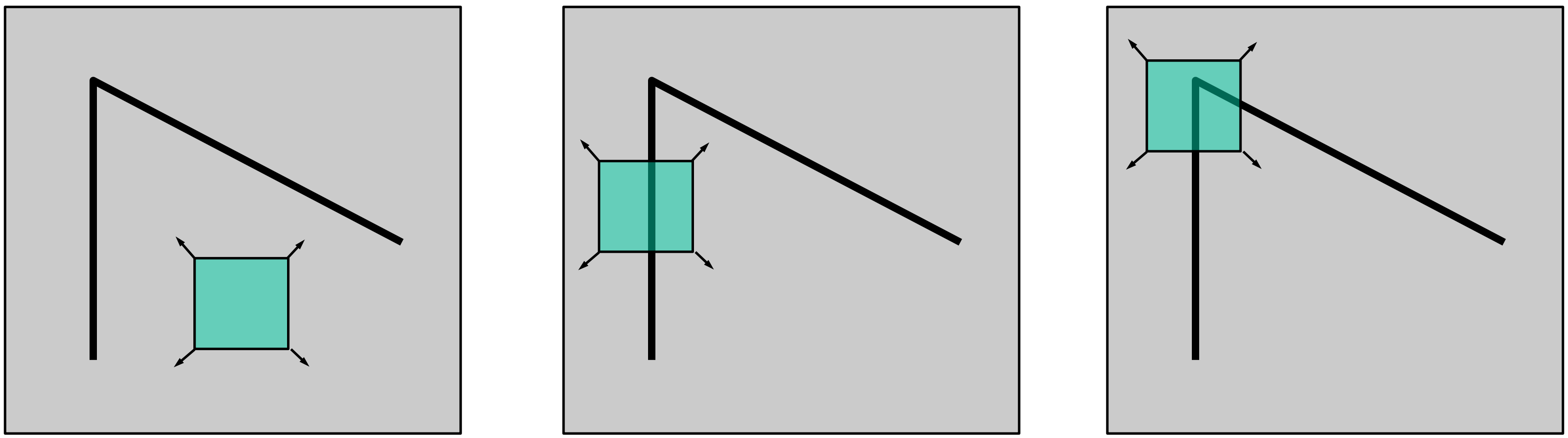


"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

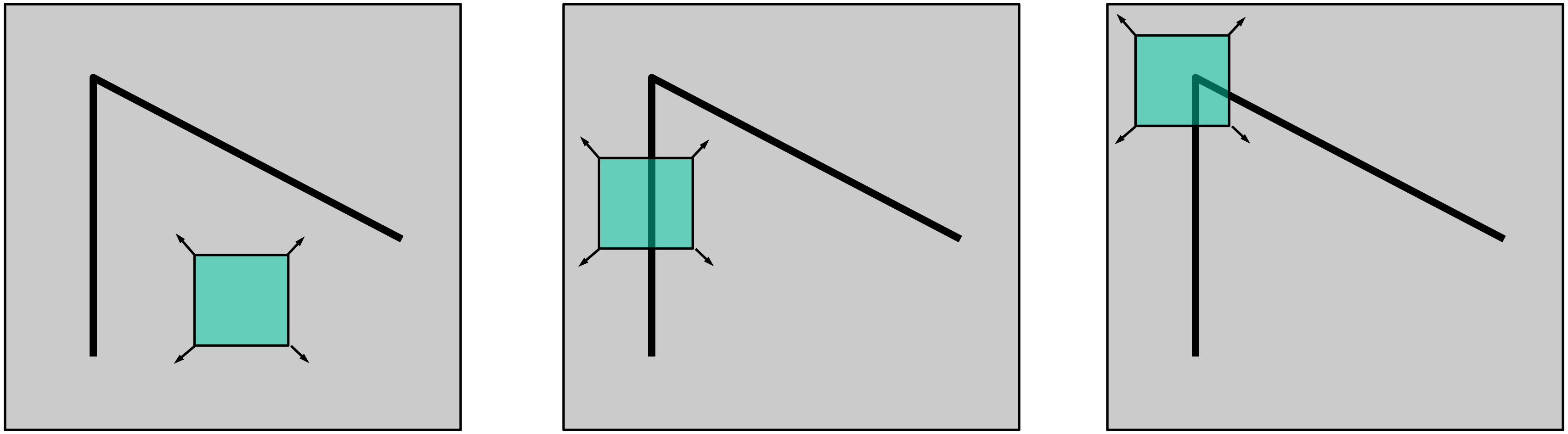


"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

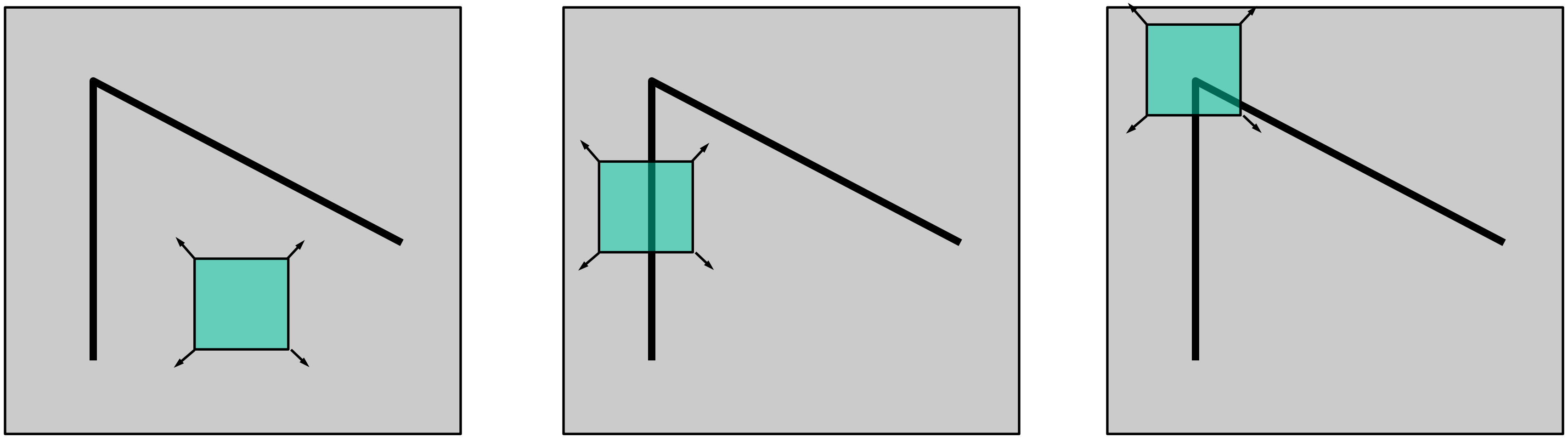


"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

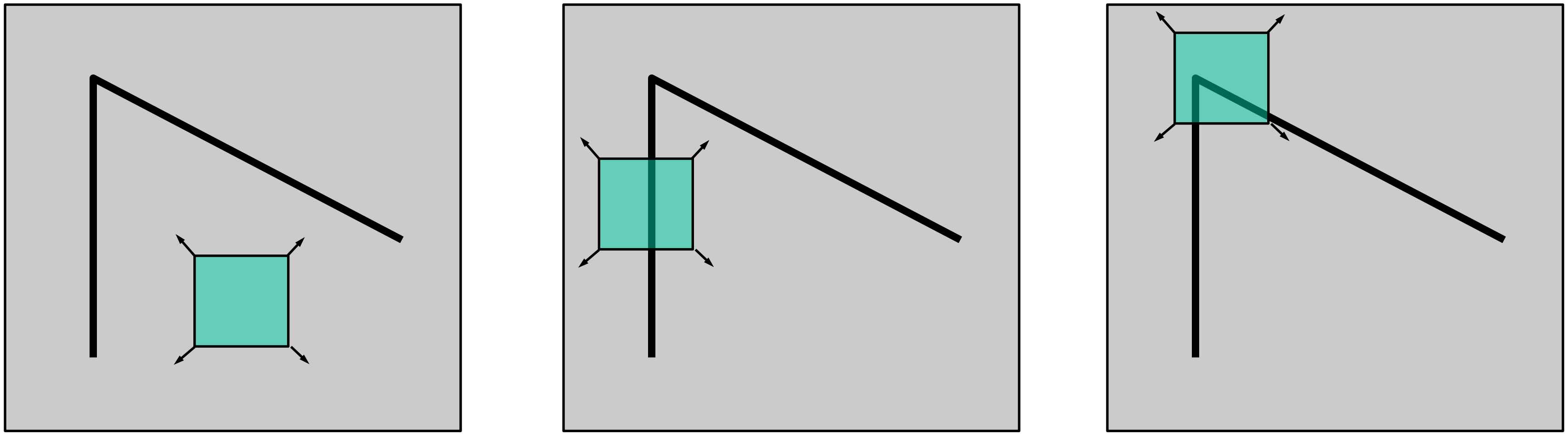


“flat” region:
no change in
all directions

“edge”:
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea

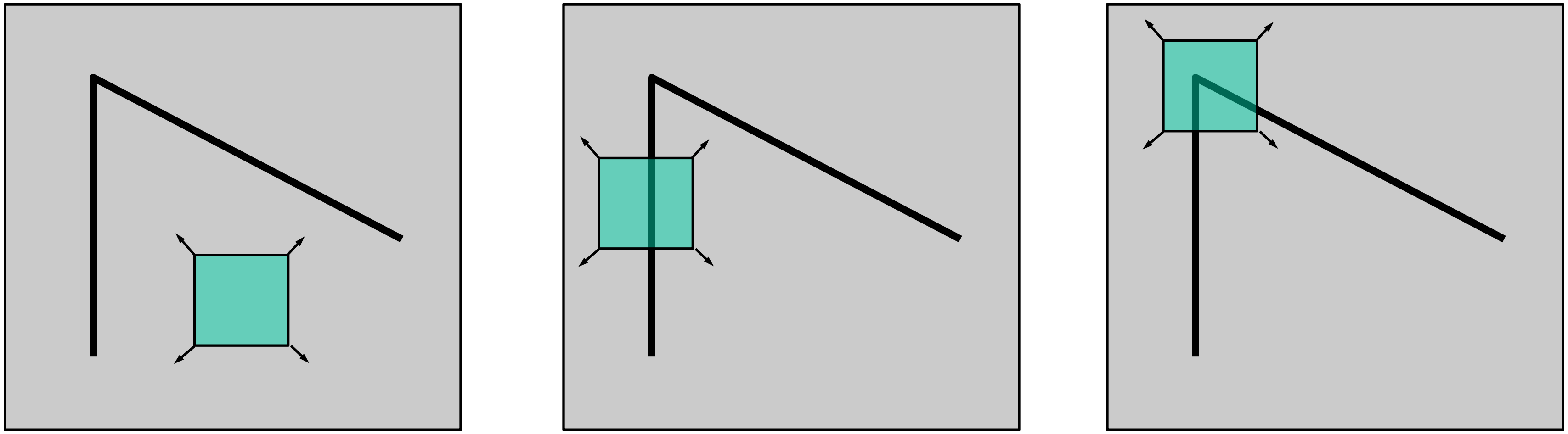


"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Corner Detector: Basic Idea



“flat” region:
no change in
all directions

“edge”:
no change along
the edge direction

“corner”:
significant change
in all directions

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Harris Detector: Mathematics

Window-averaged squared change of intensity induced by shifting the image data by $[u,v]$:

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

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Harris Detector: Mathematics

Window-averaged squared change of intensity induced by shifting the image data by [*u*,*v*]:

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

Intensity

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Harris Detector: Mathematics

Window-averaged squared change of intensity induced by shifting the image data by $[u,v]$:

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

Shifted intensity

Intensity

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Harris Detector: Mathematics

Window-averaged squared change of intensity induced by shifting the image data by $[u,v]$:

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

Window function

Shifted intensity

Intensity

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Harris Detector: Mathematics

Window-averaged squared change of intensity induced by shifting the image data by $[u,v]$:

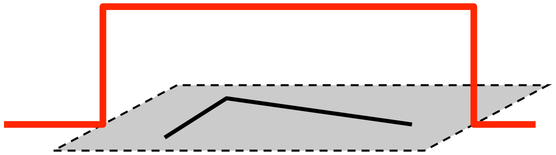
$$E(u,v)=\sum_{x,y} w(x,y)\left[I(x+u,y+v)-I(x,y)\right]^2$$

Window function

Shifted intensity

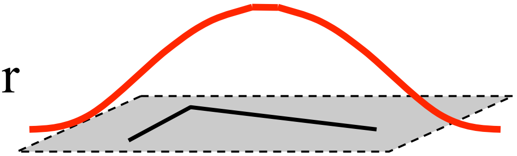
Intensity

Window function $w(x,y) =$



1 in window, 0 outside

or



Gaussian

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Taylor series approximation to shifted image gives quadratic form for error as function of image shifts.

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

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

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

Harris Detector: Mathematics

Expanding I(x,y) in a Taylor series expansion, we have, for small shifts $[u,v]$, a *quadratic* approximation to the error surface between a patch and itself, shifted by $[u,v]$:

$$E(u,v) \cong [u,v] \; M \begin{bmatrix} u \\ v \end{bmatrix}$$

where *M* is a 2×2 matrix computed from image derivatives:

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

M is often called structure tensor

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

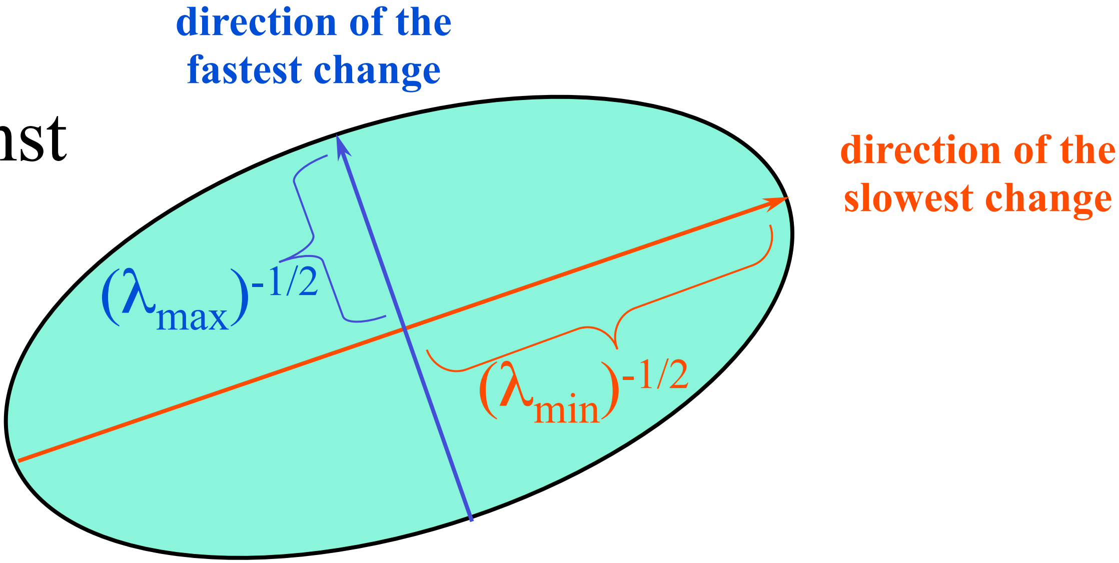
Harris Detector: Mathematics

Intensity change in shifting window: eigenvalue analysis

$$E(u,v) \cong [u,v] \; M \begin{bmatrix} u \\ v \end{bmatrix}$$

λ_1, λ_2 – eigenvalues of M

Ellipse $E(u,v)$ = const


direction of the fastest change
direction of the slowest change
$(\lambda_{max})^{-1/2}$
$(\lambda_{min})^{-1/2}$

Iso-contour of the squared error, $E(u,v)$

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Monday, March 28, 2011

Selecting Good Features

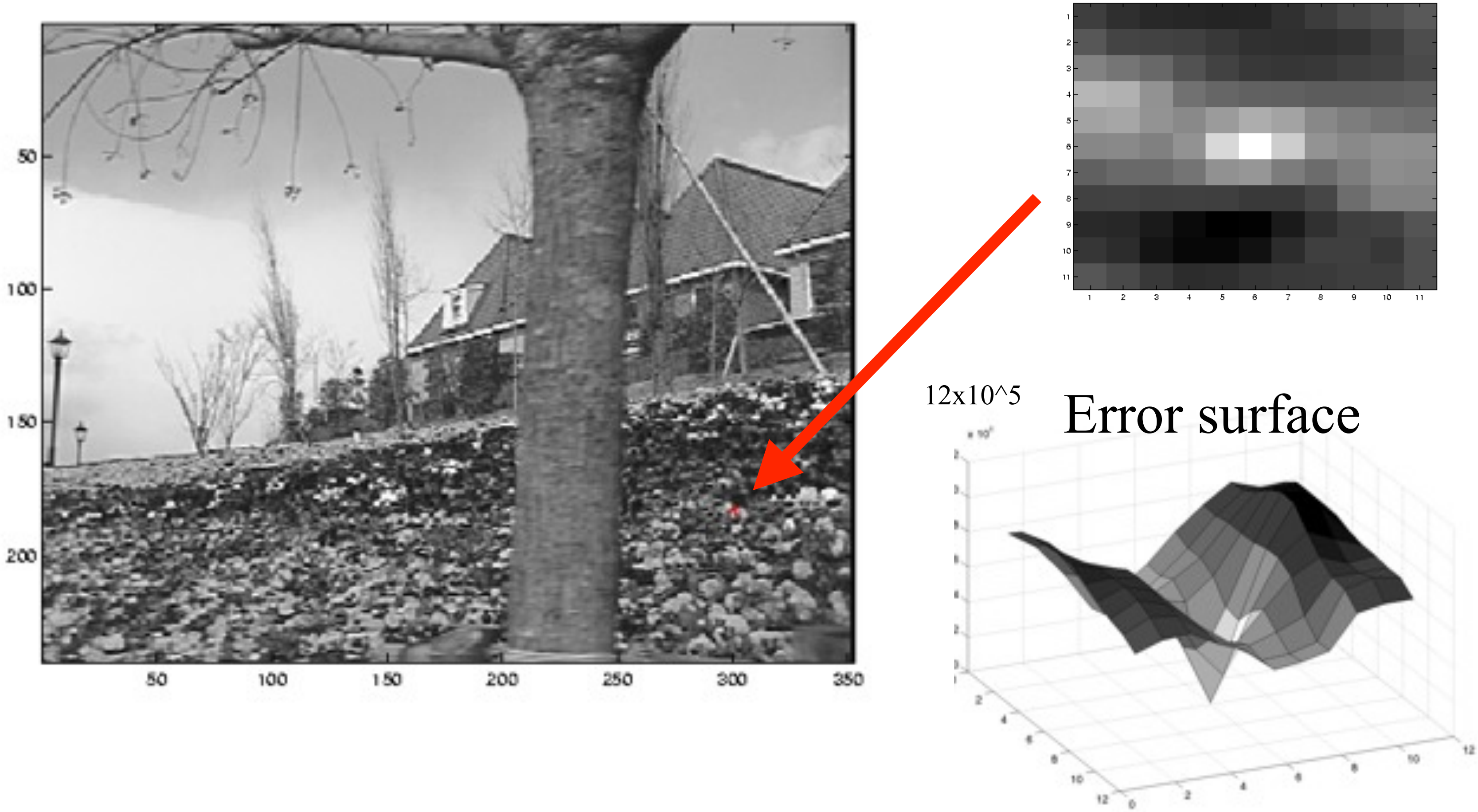


Image patch

12x10^5

Error surface

λ_1 and λ_2 are large

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Selecting Good Features

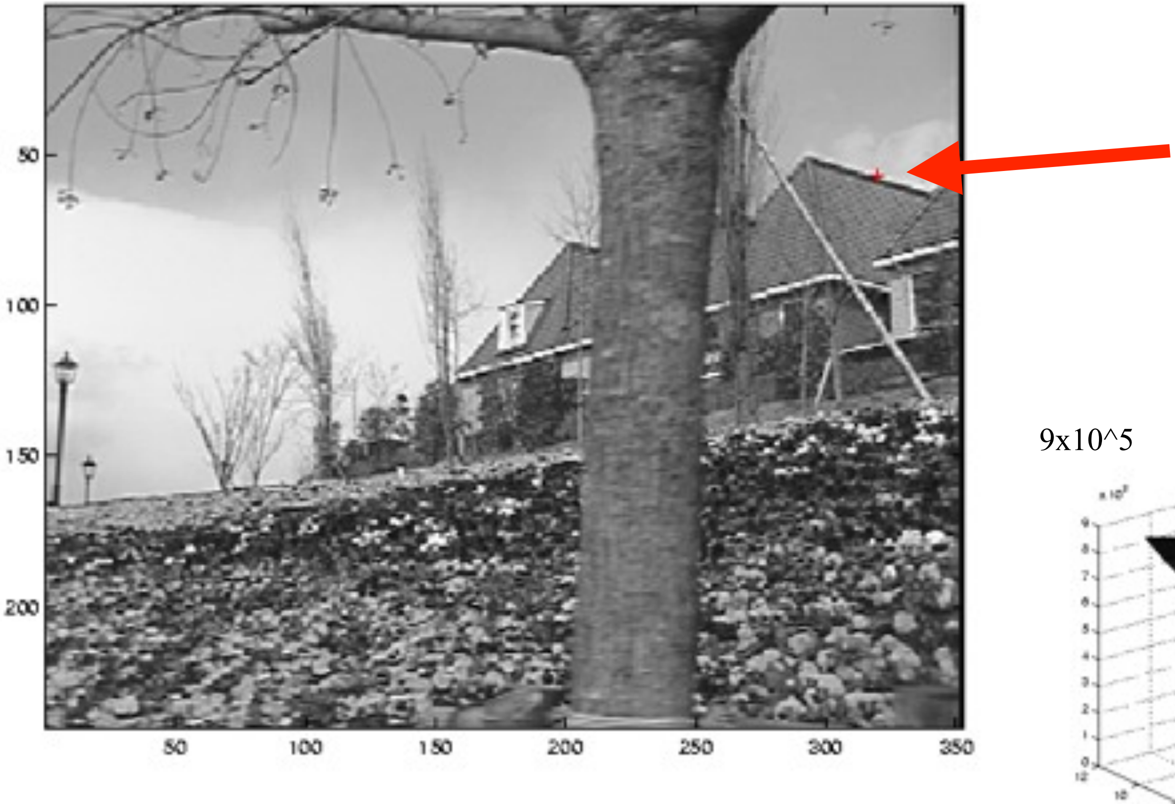
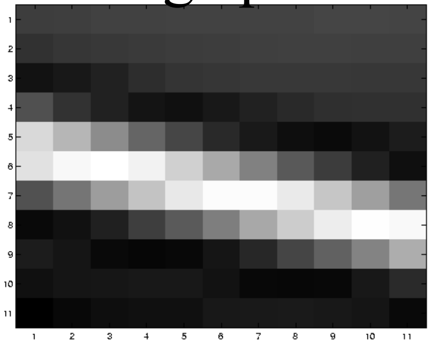
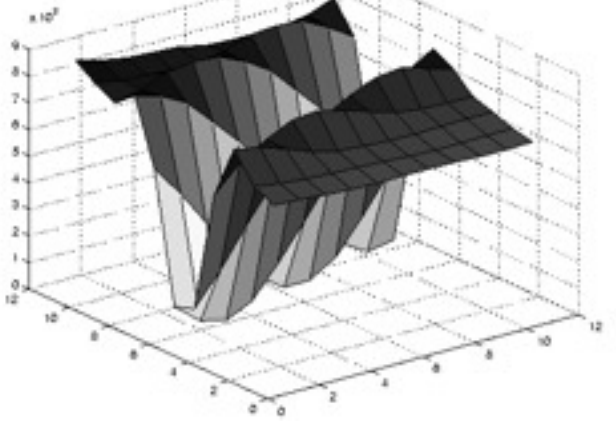


Image patch



9x10^5

Error surface



large λ_1, small λ_2

Darya Frolova, Denis Simakov The Weizmann Institute of Science
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Selecting Good Features

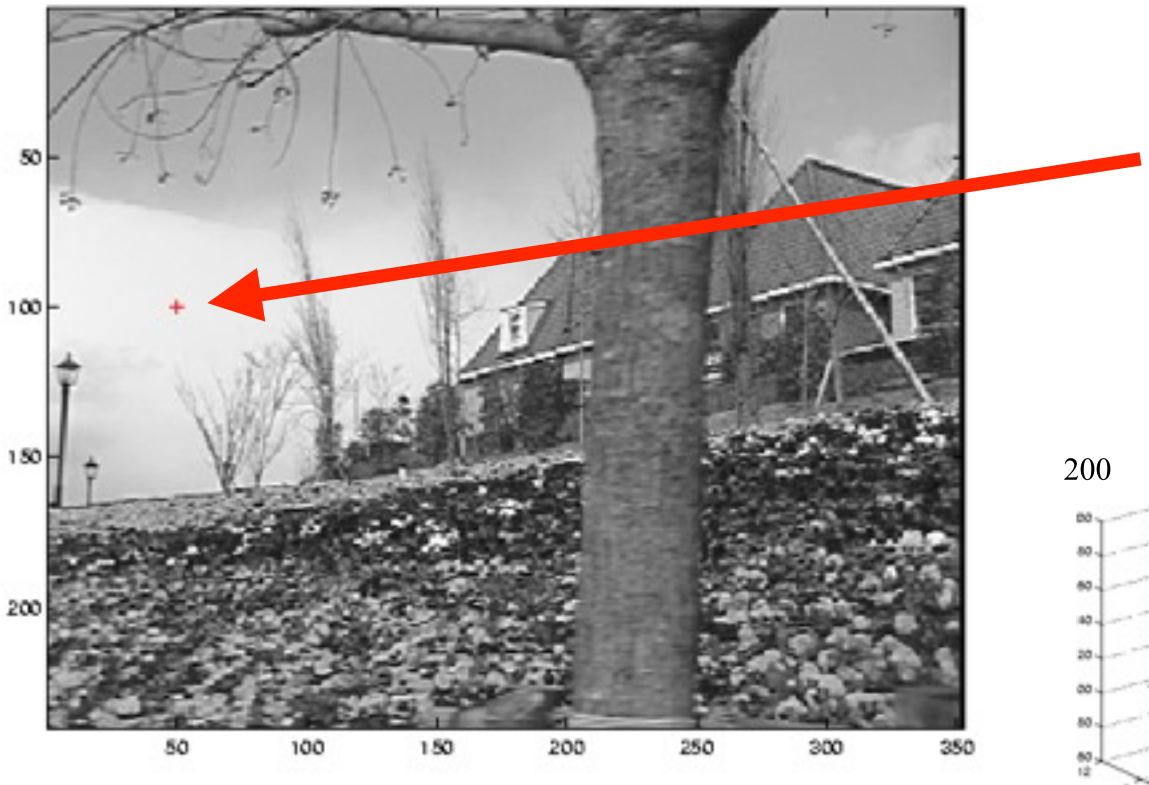
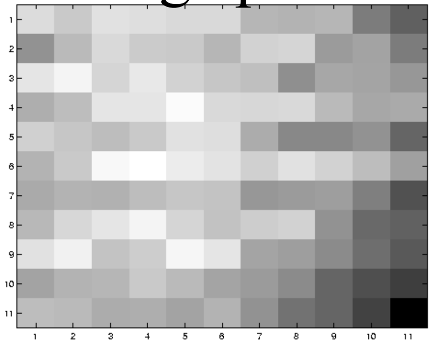
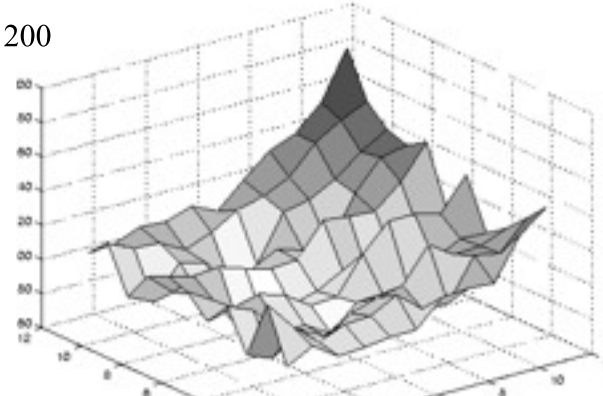


Image patch



(contrast auto-scaled)



Error surface

(vertical scale exaggerated relative to previous plots)

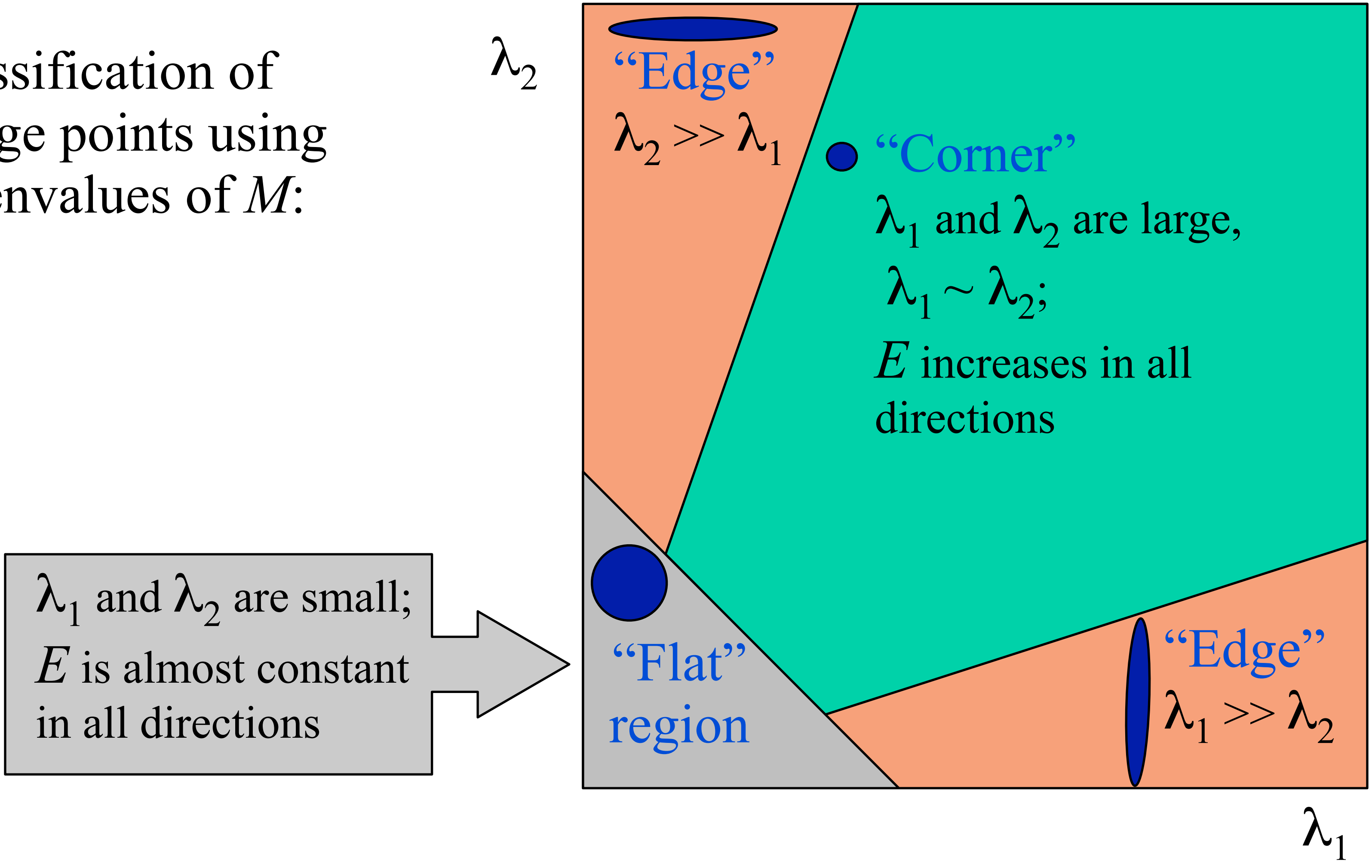
small λ_1, small λ_2

Darya Frolova, Denis Simakov The Weizmann Institute of Science

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Harris Detector: Mathematics

Classification of image points using eigenvalues of *M*:



λ_1 and λ_2 are small;
E is almost constant in all directions

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Mathematics

Measure of corner response:

$$R = \det M - k\left(\operatorname{trace} M\right)^2$$

$$\det M = \lambda_1 \lambda_2$$
$$\operatorname{trace} M = \lambda_1 + \lambda_2$$

(k – empirical constant, k = 0.04-0.06)

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Mathematics

Measure of corner response:

$$R = \det M - k\left(\operatorname{trace} M\right)^2$$

$$\det M = \lambda_1 \lambda_2$$
$$\operatorname{trace} M = \lambda_1 + \lambda_2$$

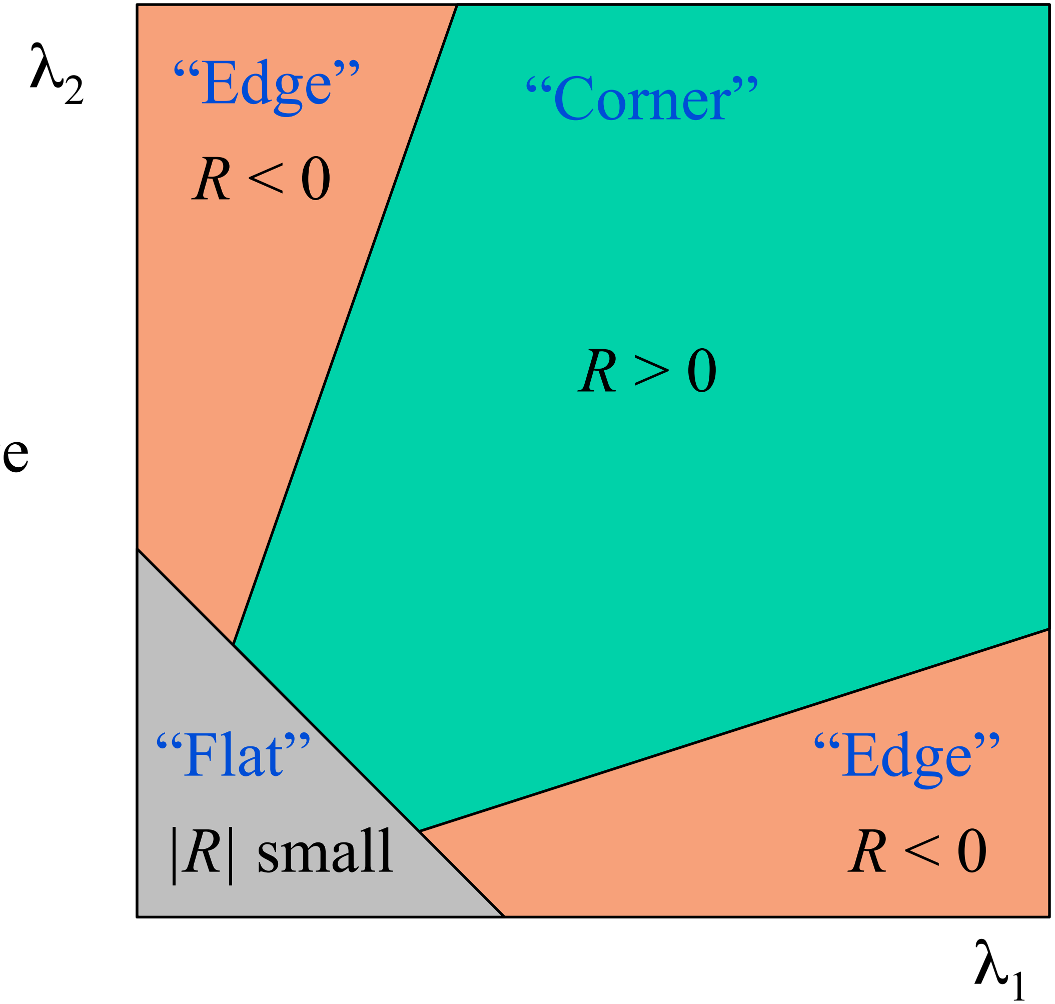
(k – empirical constant, k = 0.04-0.06)

(Shi-Tomasi variation: use min(λ_1,λ_2) instead of R)

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Mathematics

- R depends only on eigenvalues of M
- R is large for a corner
- R is negative with large magnitude for an edge
- $|R|$ is small for a flat region



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector

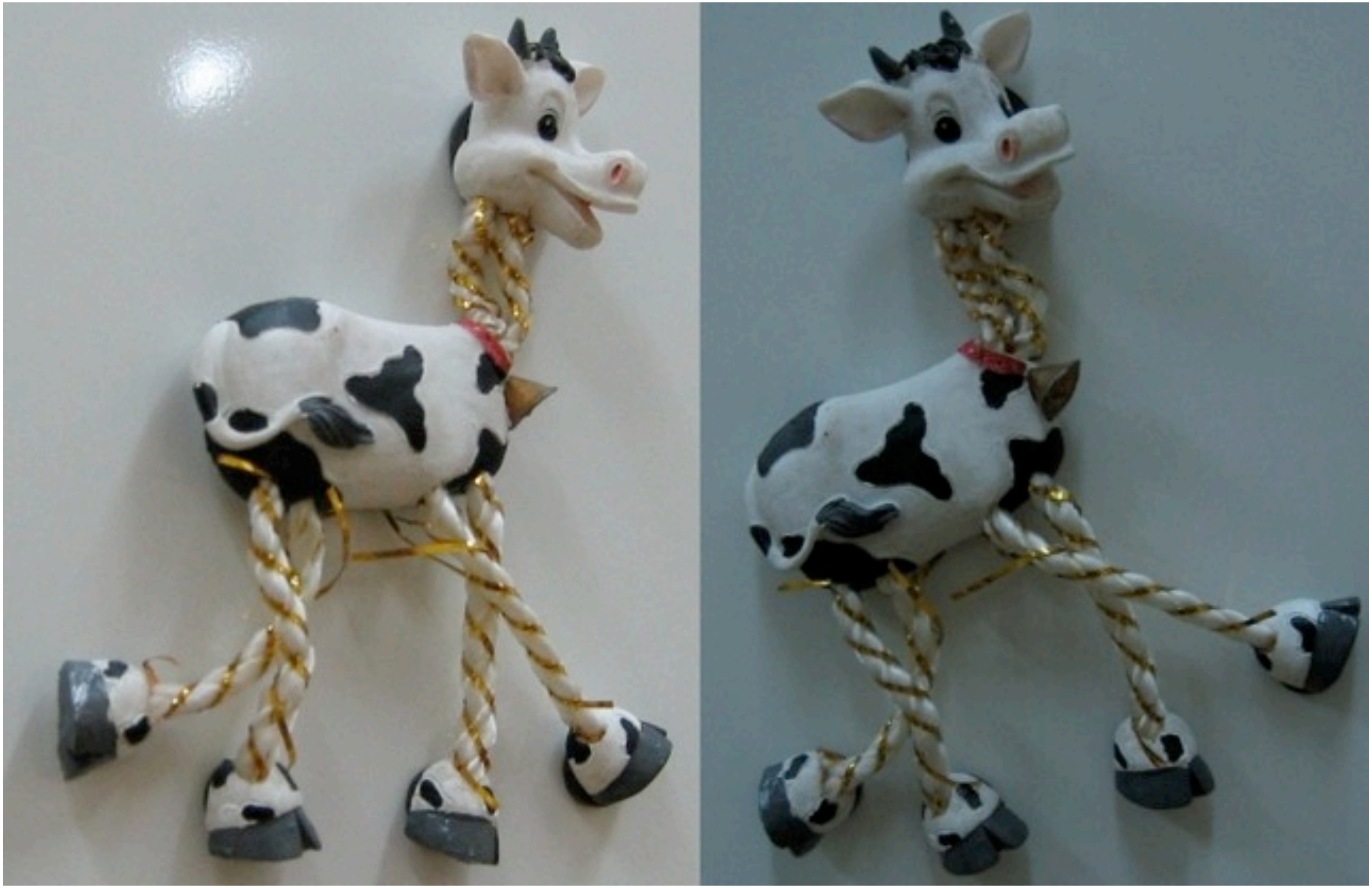
- The Algorithm:
  - Find points with large corner response function R ($R >$ threshold)
  - Take the points of *local maxima* of R

Harris corner detector algorithm

- Compute image gradients I_x I_y for all pixels
- For each pixel
  - Compute $$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$
    by looping over neighbors x,y
  - compute $$R = \det M - k\left(\operatorname{trace} M\right)^2$$
- Find points with large corner response function R ($R >$ threshold)
- Take the points of locally maximum R as the detected feature points (ie, pixels where R is bigger than for all the 4 or 8 neighbors).

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

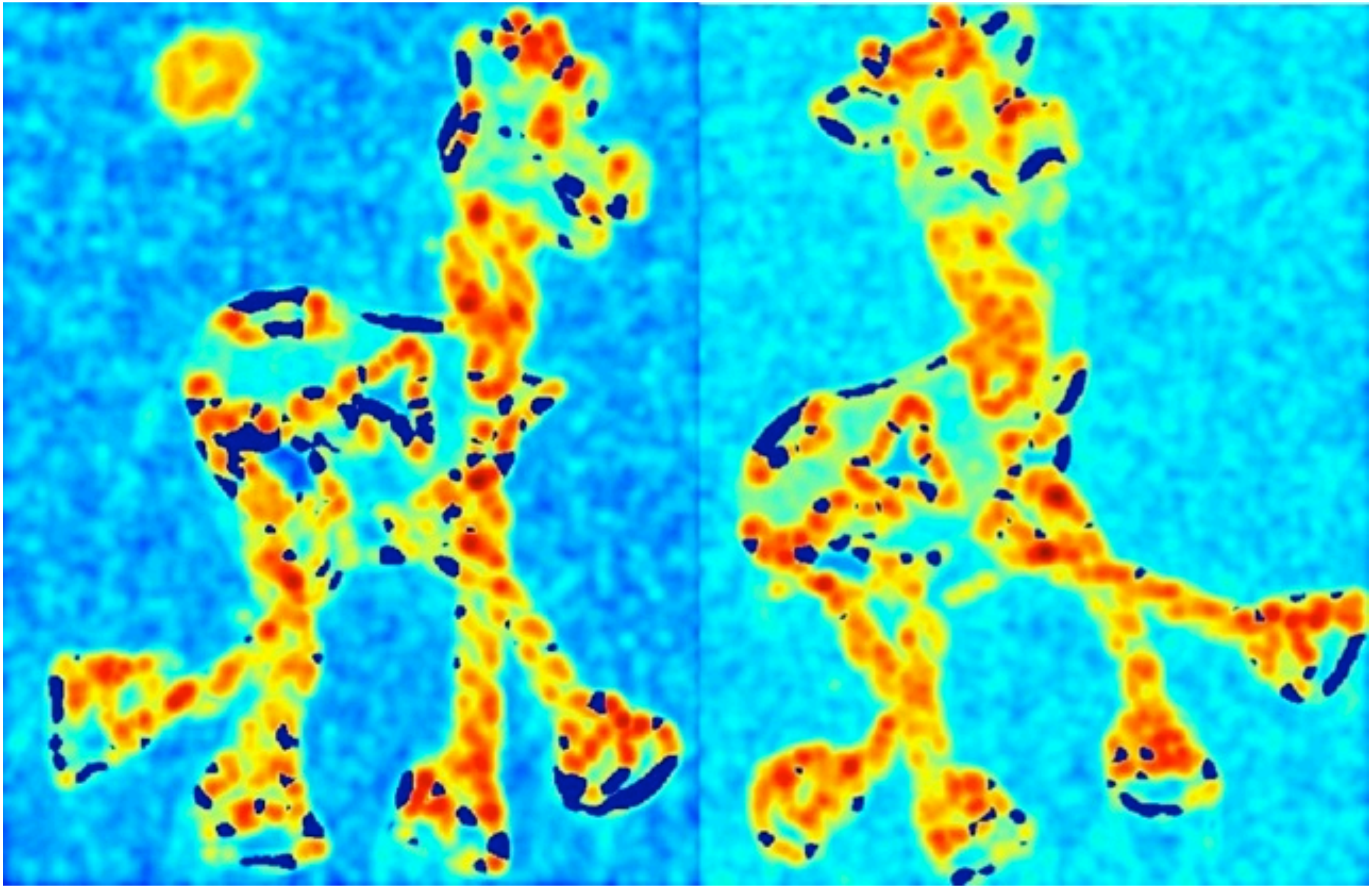
Harris Detector: Workflow



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Workflow

Compute corner response R

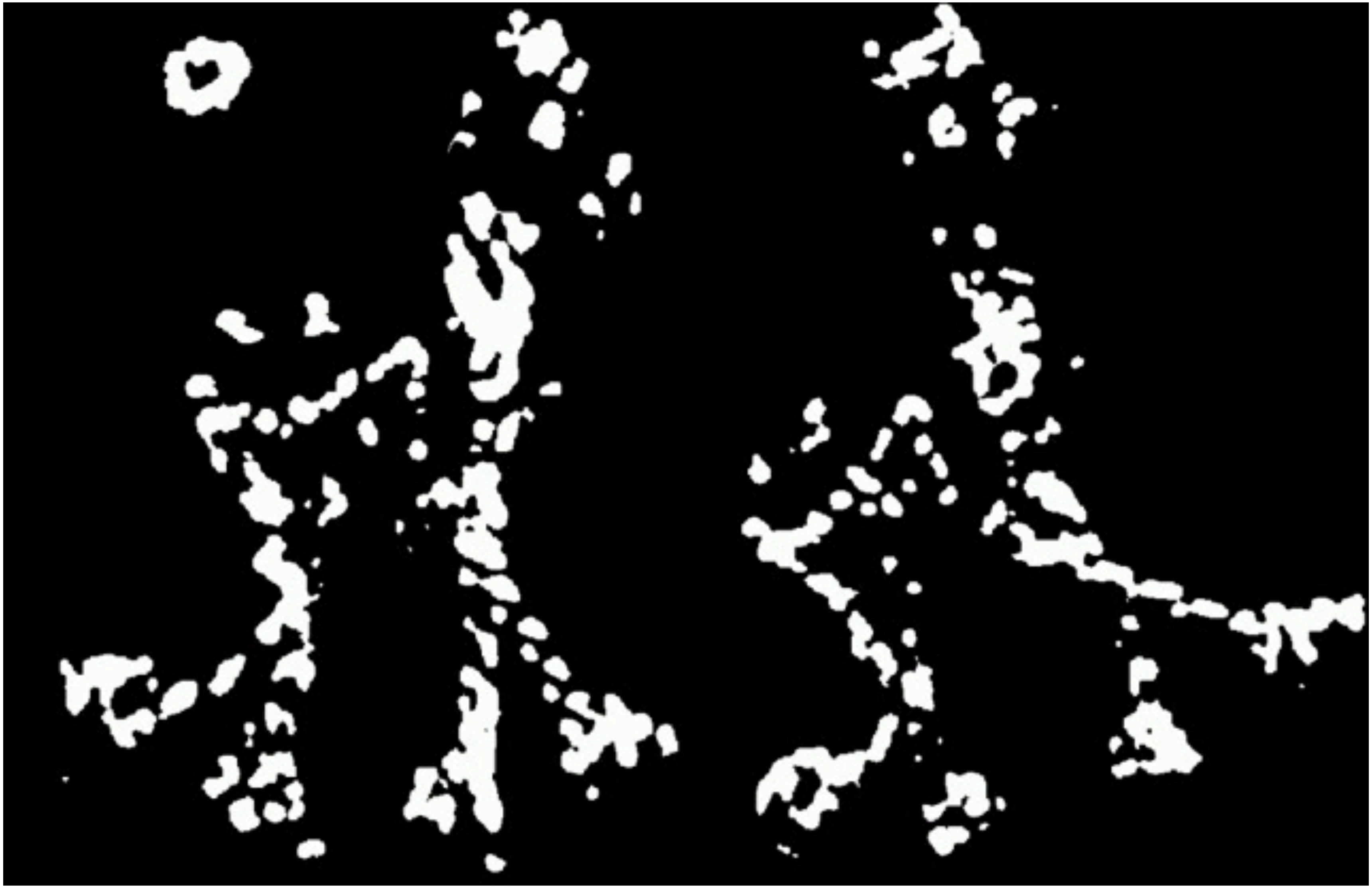


http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Workflow

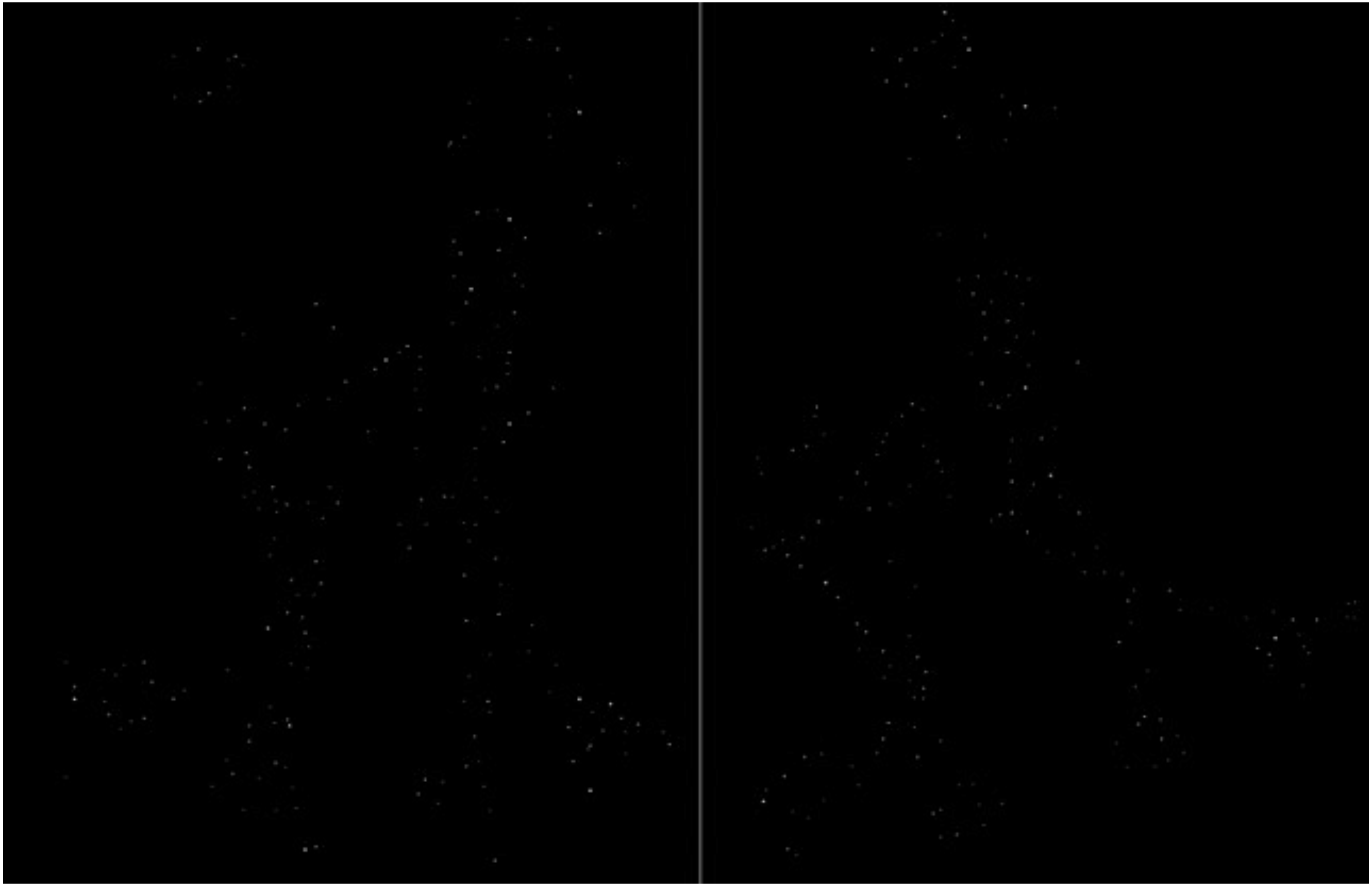
Find points with large corner response: R>threshold



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Workflow

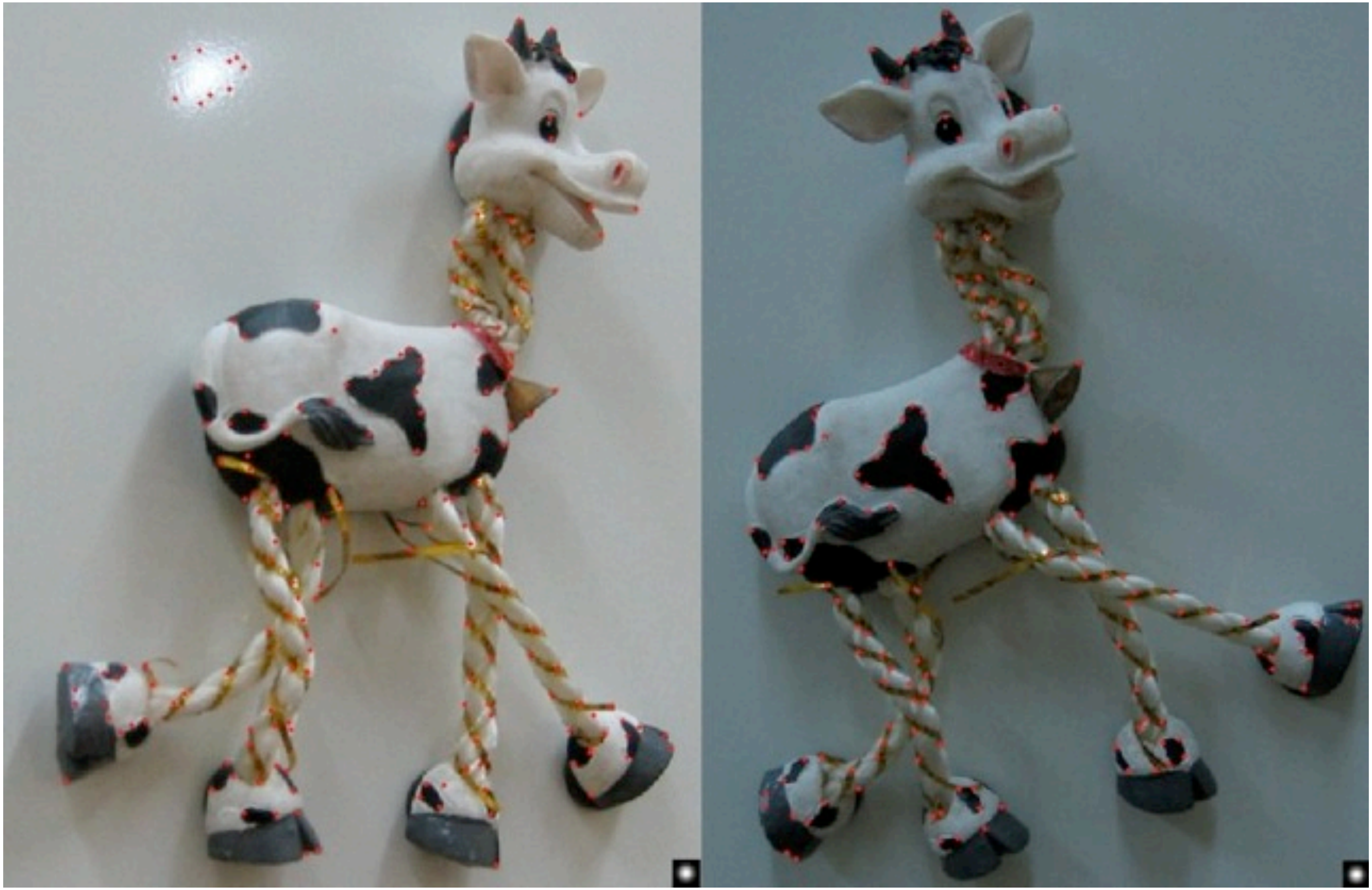
Take only the points of local maxima of R



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

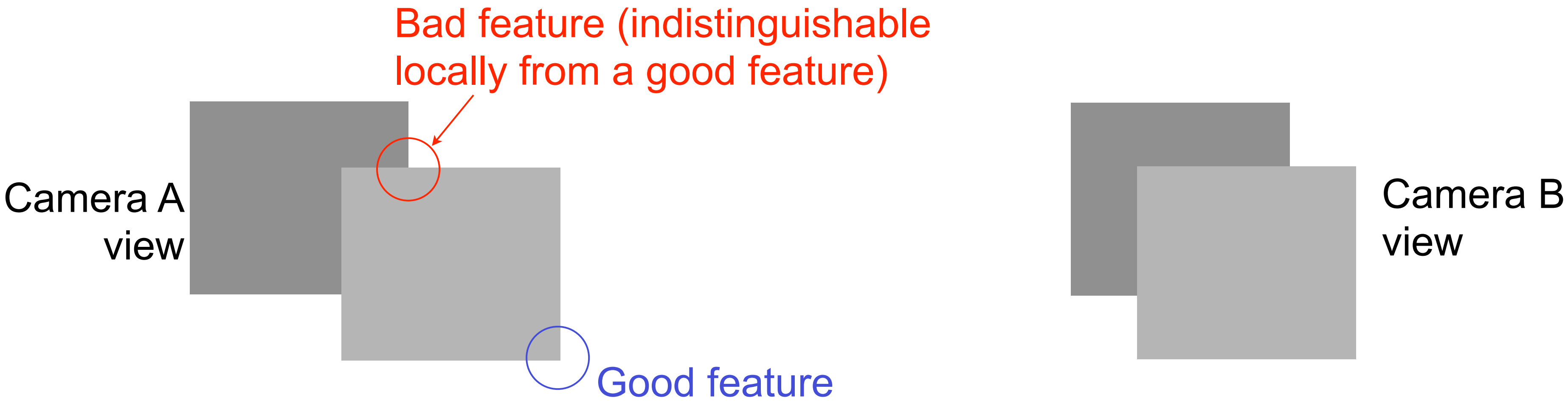
Harris Detector: Workflow



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Comments on Harris detector

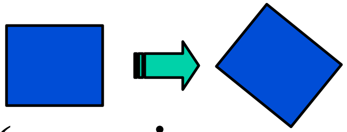
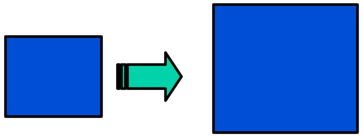
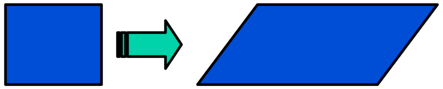
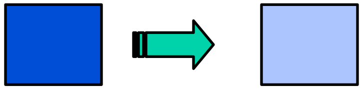
- We haven’t analyzed how consistent in position the local-max-of-R feature is. We’ll show some results on that later.
- If we reverse the image contrast, the local-max-of-R feature gives the exact same feature positions--do we want that?
- How might you address the problem below:



Analysis of Harris corner detector invariance properties

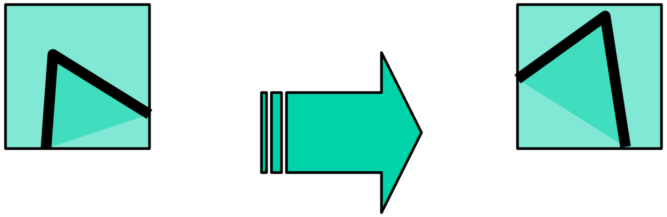
- Geometry
  - rotation
  - scale
- Photometry
  - intensity change

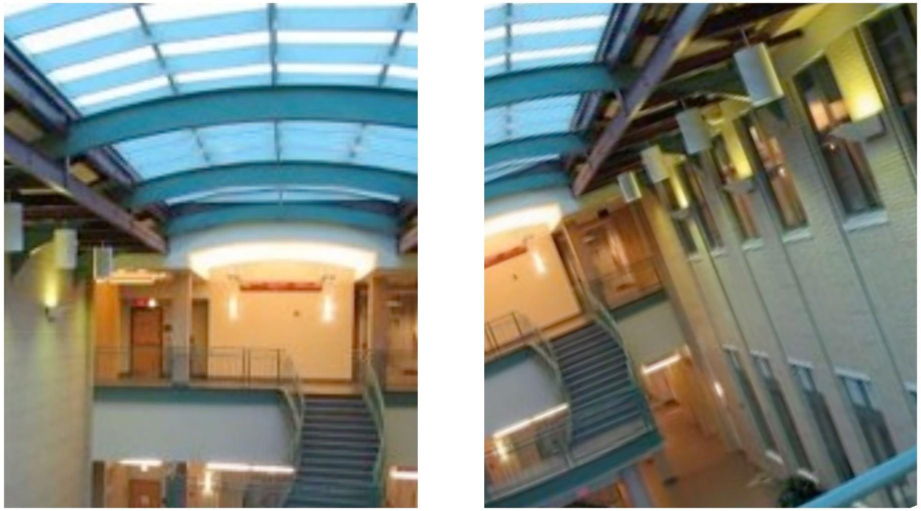
Models of Image Change

- Geometry
  - Rotation 
  - Similarity (rotation + uniform scale) 
  - Affine (scale dependent on direction) 
    valid for: orthographic camera, locally planar object
- Photometry
  - Affine intensity change ($I \rightarrow a\,I + b$) 

Harris Detector: Some Properties

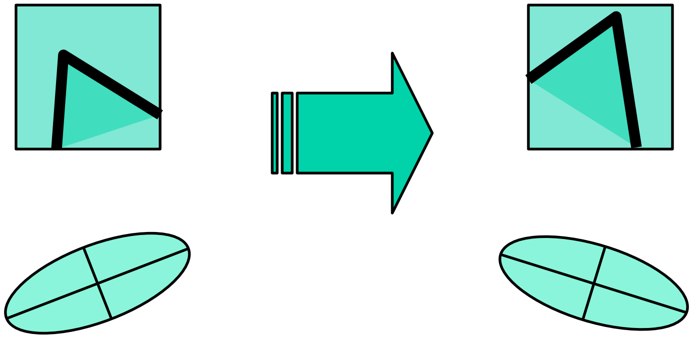
- Rotation invariance?





Harris Detector: Some Properties

- Rotation invariance



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

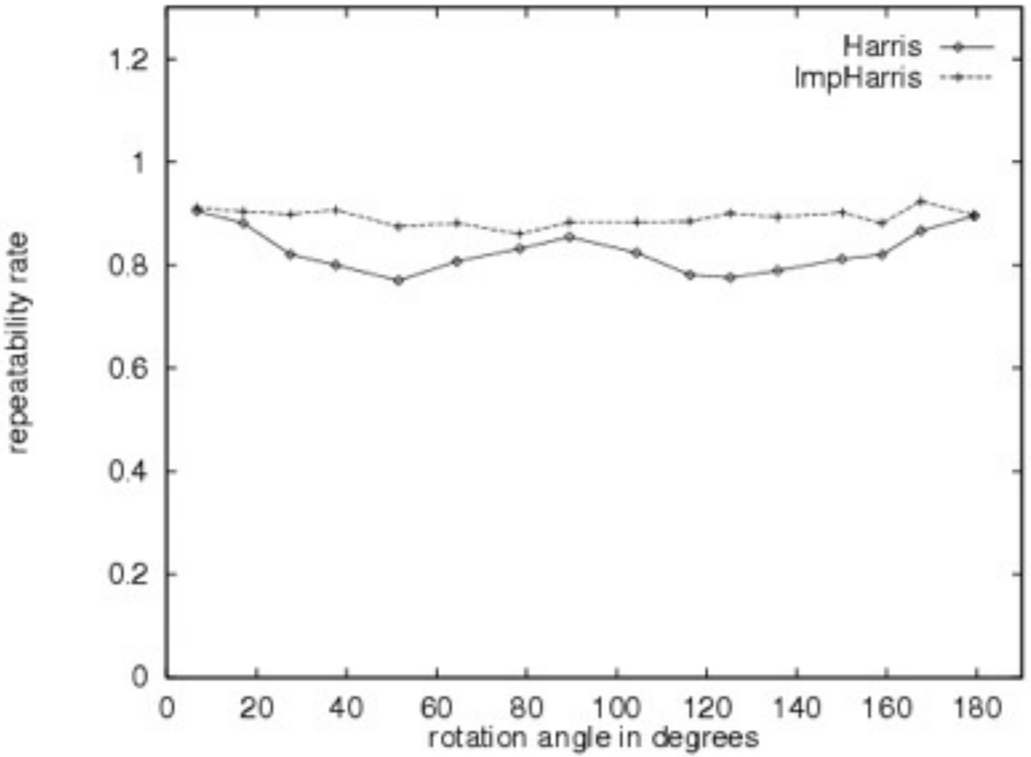
Corner response R is invariant to image rotation

Eigen analysis allows us to work in the canonical frame of the linear form.

Rotation Invariant Detection

ImpHarris: derivatives are computed more precisely by replacing the [-2 -1 0 1 2] mask with derivatives of a Gaussian (sigma = 1).

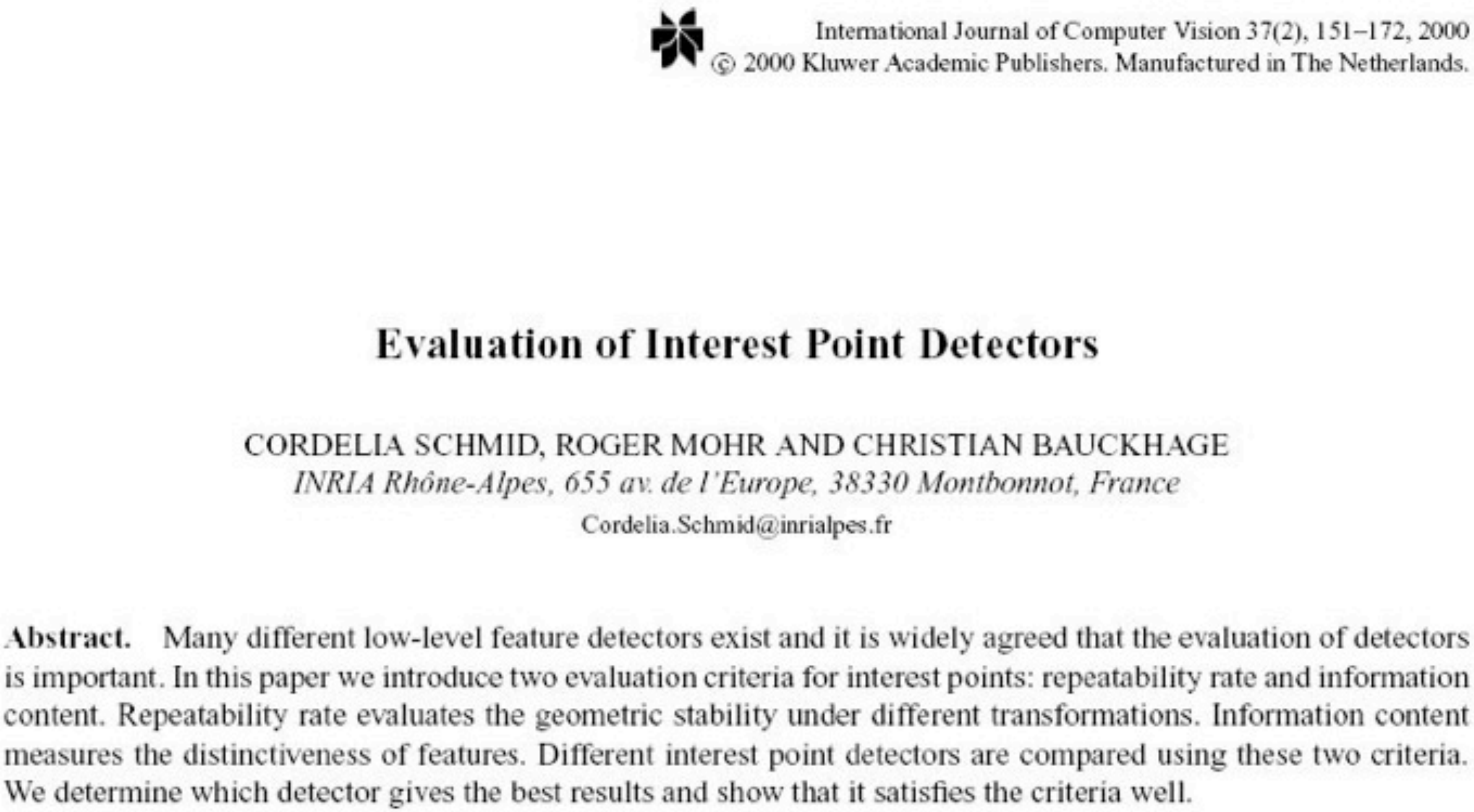
Harris Corner Detector



This gives us rotation invariant *detection*, but we'll need to do more to ensure a rotation invariant *descriptor*...

C.Schmid et.al. "Evaluation of Interest Point Detectors". IJCV 2000

Evaluation plots are from this paper



International Journal of Computer Vision 37(2), 151–172, 2000

Evaluation of Interest Point Detectors

CORDELIA SCHMID, ROGER MOHR AND CHRISTIAN BAUCKHAGE
INRIA Rhône-Alpes, 655 av. de l'Europe, 38330 Montbonnot, France
Cordelia.Schmid@inrialpes.fr

Abstract. Many different low-level feature detectors exist and it is widely agreed that the evaluation of detectors is important. In this paper we introduce two evaluation criteria for interest points: repeatability rate and information content. Repeatability rate evaluates the geometric stability under different transformations. Information content measures the distinctiveness of features. Different interest point detectors are compared using these two criteria. We determine which detector gives the best results and show that it satisfies the criteria well.

Harris Detector: Some Properties

- Invariance to image intensity change?

Harris Detector: Some Properties

- Partial invariance to additive and multiplicative intensity changes

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Some Properties

- Partial invariance to additive and multiplicative intensity changes

  ✓ Only derivatives are used => invariance to intensity shift $I \rightarrow I + b$

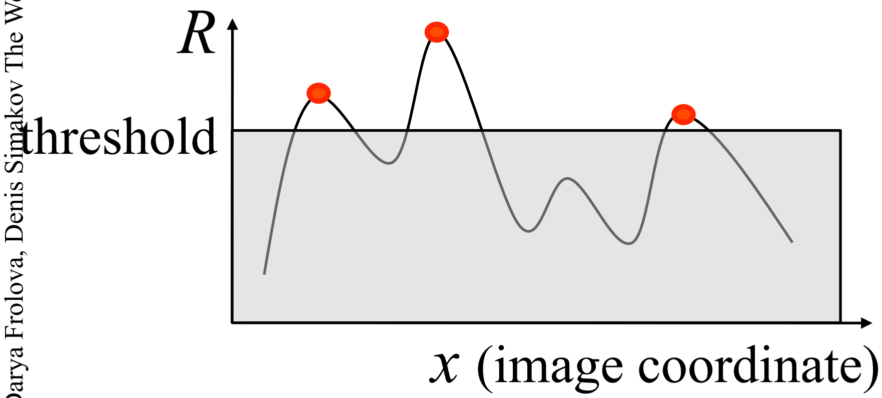
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

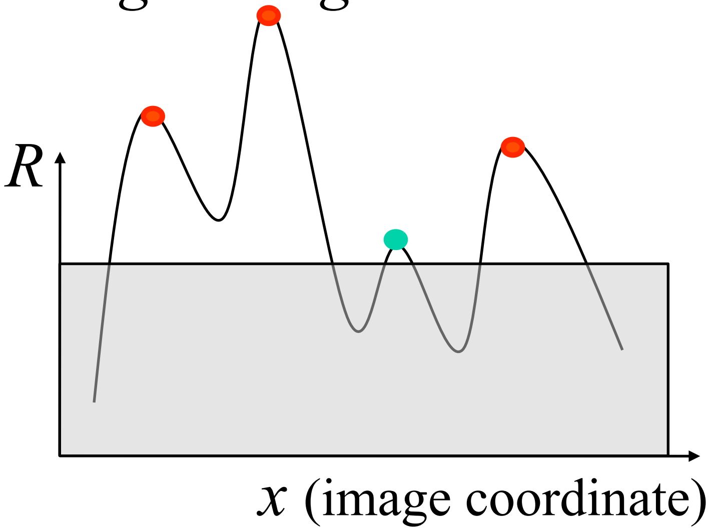
Harris Detector: Some Properties

- Partial invariance to additive and multiplicative intensity changes

✓ Only derivatives are used => invariance to intensity shift $I \rightarrow I + b$

✓ Intensity scaling: $I \rightarrow a\,I$ fine, *except for the threshold that's used to specify when R is large enough.*





http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Some Properties

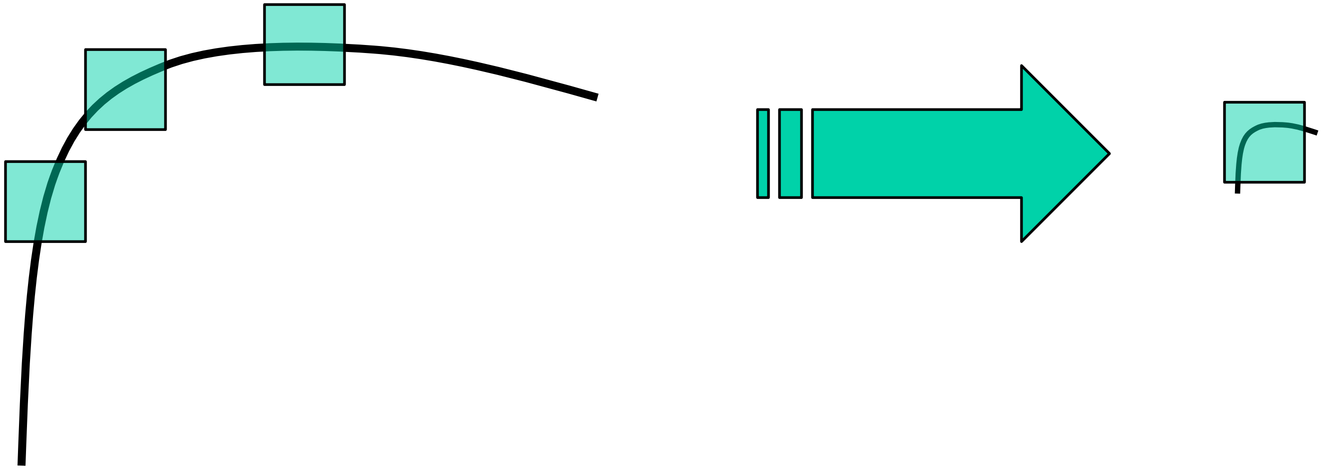
- Invariant to image scale?

image

zoomed image

Harris Detector: Some Properties

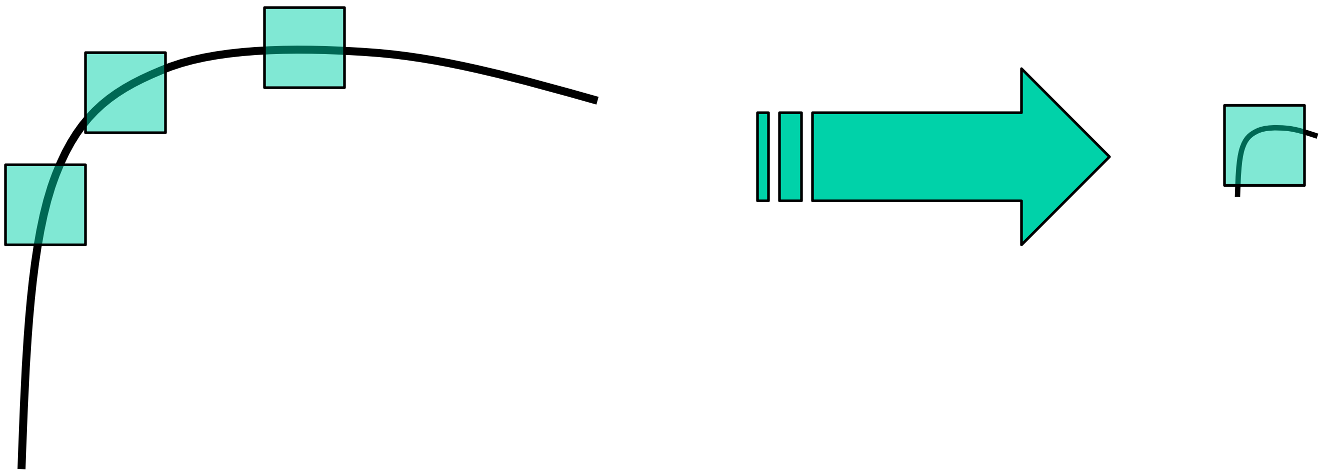
- Not invariant to *image scale*!



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Some Properties

- Not invariant to *image scale*!



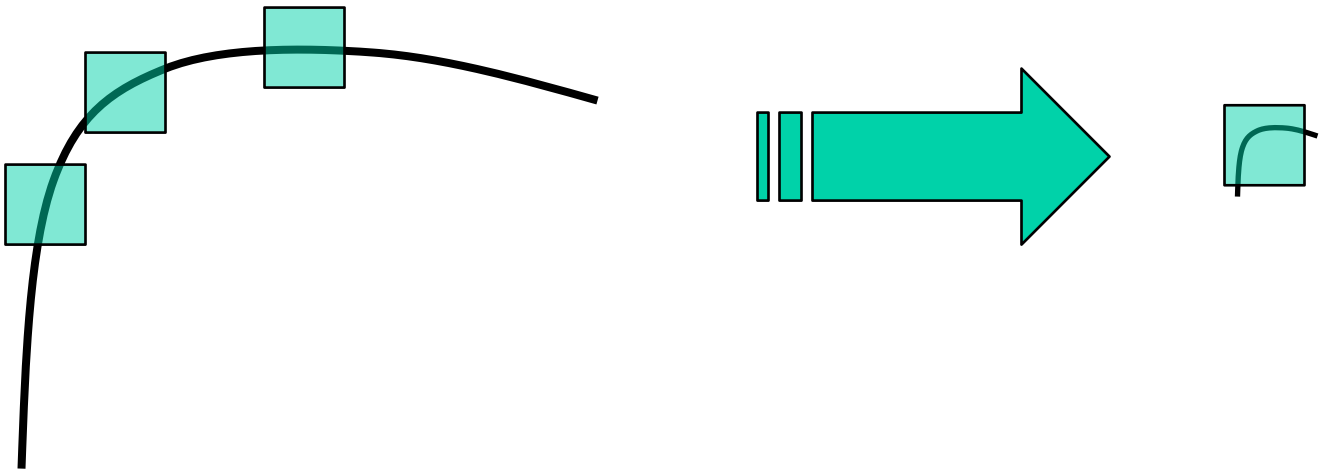
All points will be classified as edges

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Some Properties

- Not invariant to *image scale*!



All points will be classified as edges

Corner !

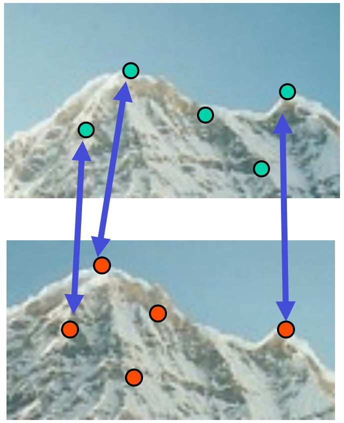
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Harris Detector: Some Properties

- Quality of Harris detector for different scale changes

Repeatability rate:

$$\frac{\text{\# correspondences}}{\text{\# possible correspondences}}$$



C.Schmid et.al. “Evaluation of Interest Point Detectors”. IJCV 2000

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

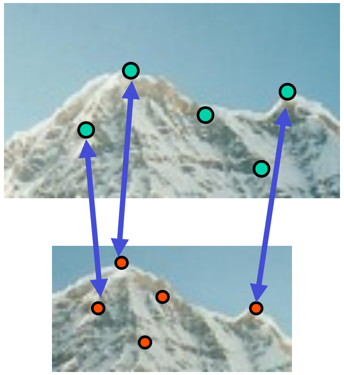
Darya Frolova, Denis Simakov The Weizmann Institute of Science

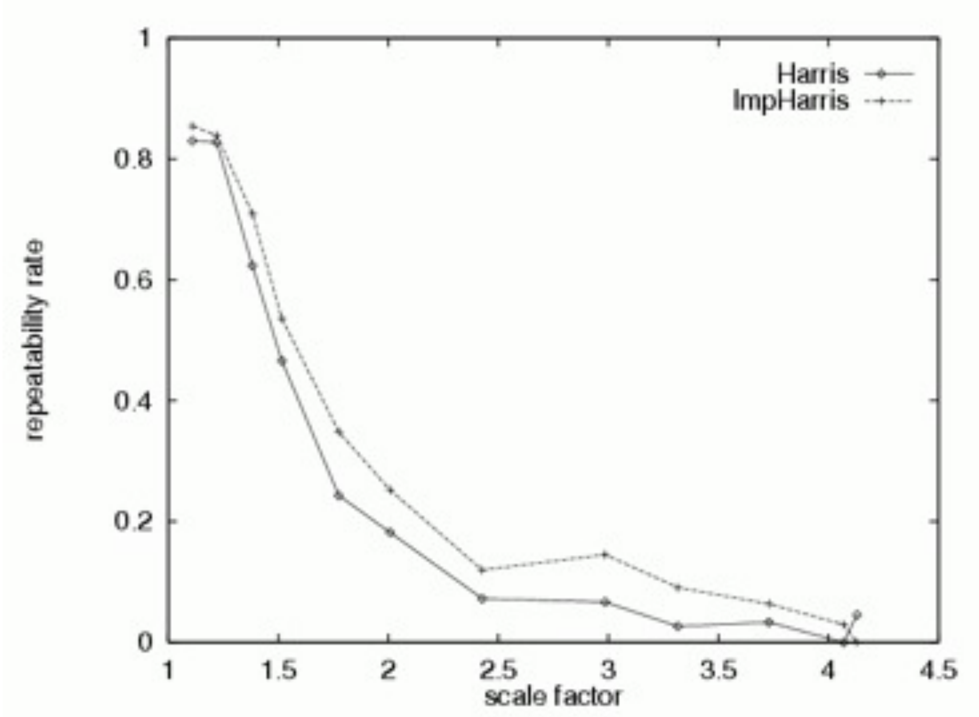
Harris Detector: Some Properties

- Quality of Harris detector for different scale changes

Repeatability rate:

$$\frac{\#\ \text{correspondences}}{\#\ \text{possible correspondences}}$$



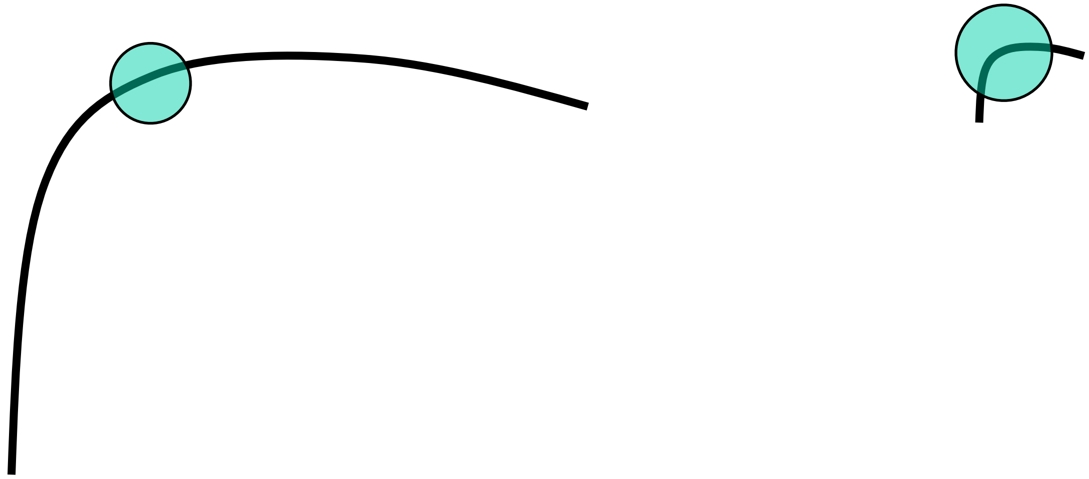


http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

C.Schmid et.al. “Evaluation of Interest Point Detectors”. IJCV 2000

Scale Invariant Detection

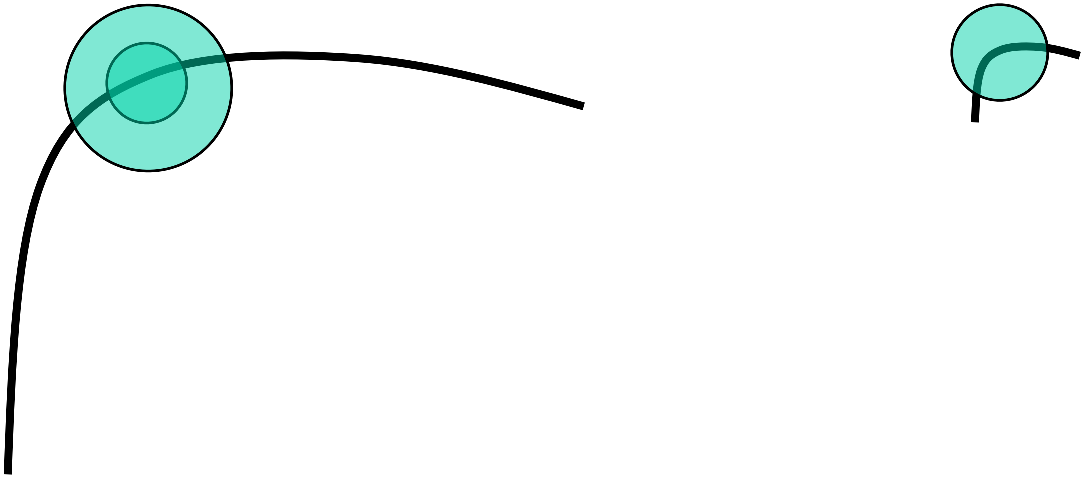
- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding sizes will look the same in both images



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Scale Invariant Detection

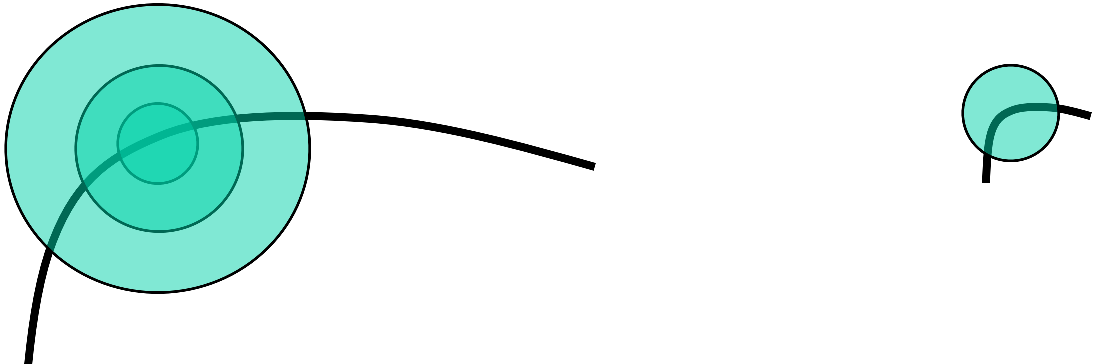
- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding sizes will look the same in both images



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Scale Invariant Detection

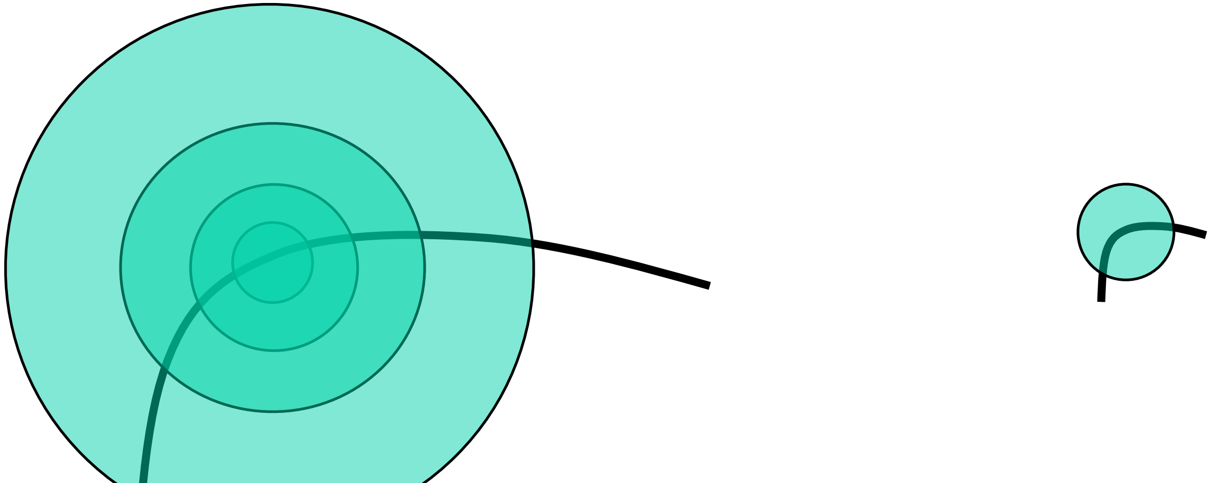
- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding sizes will look the same in both images



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Scale Invariant Detection

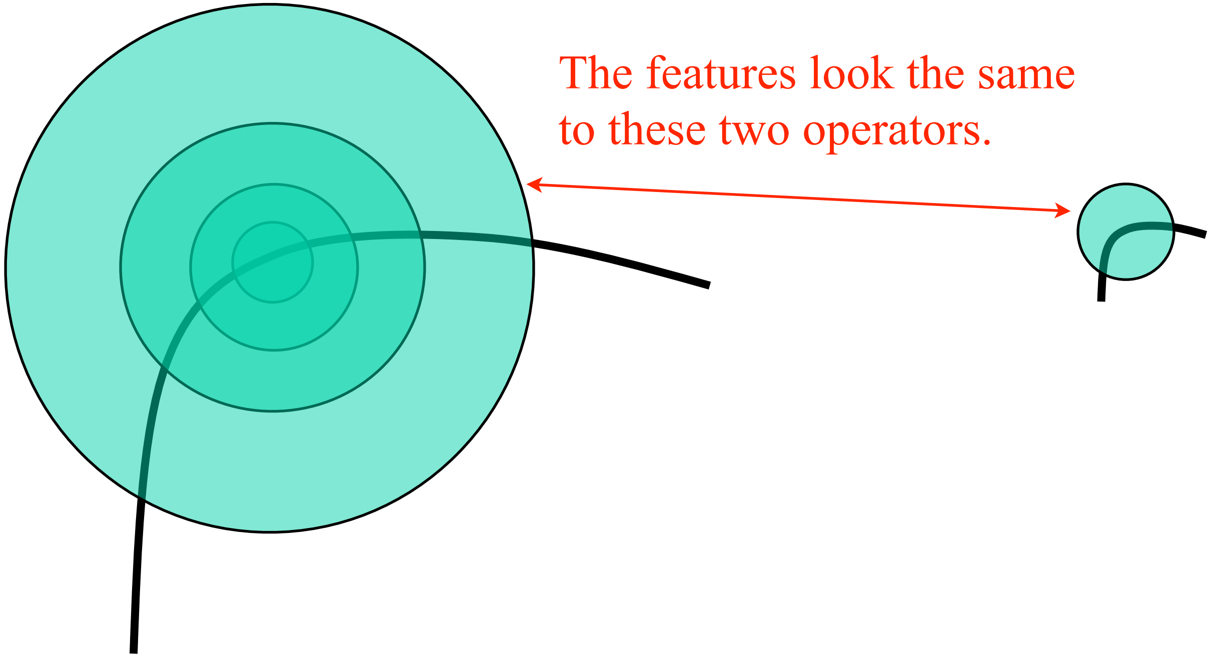
- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding sizes will look the same in both images



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

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



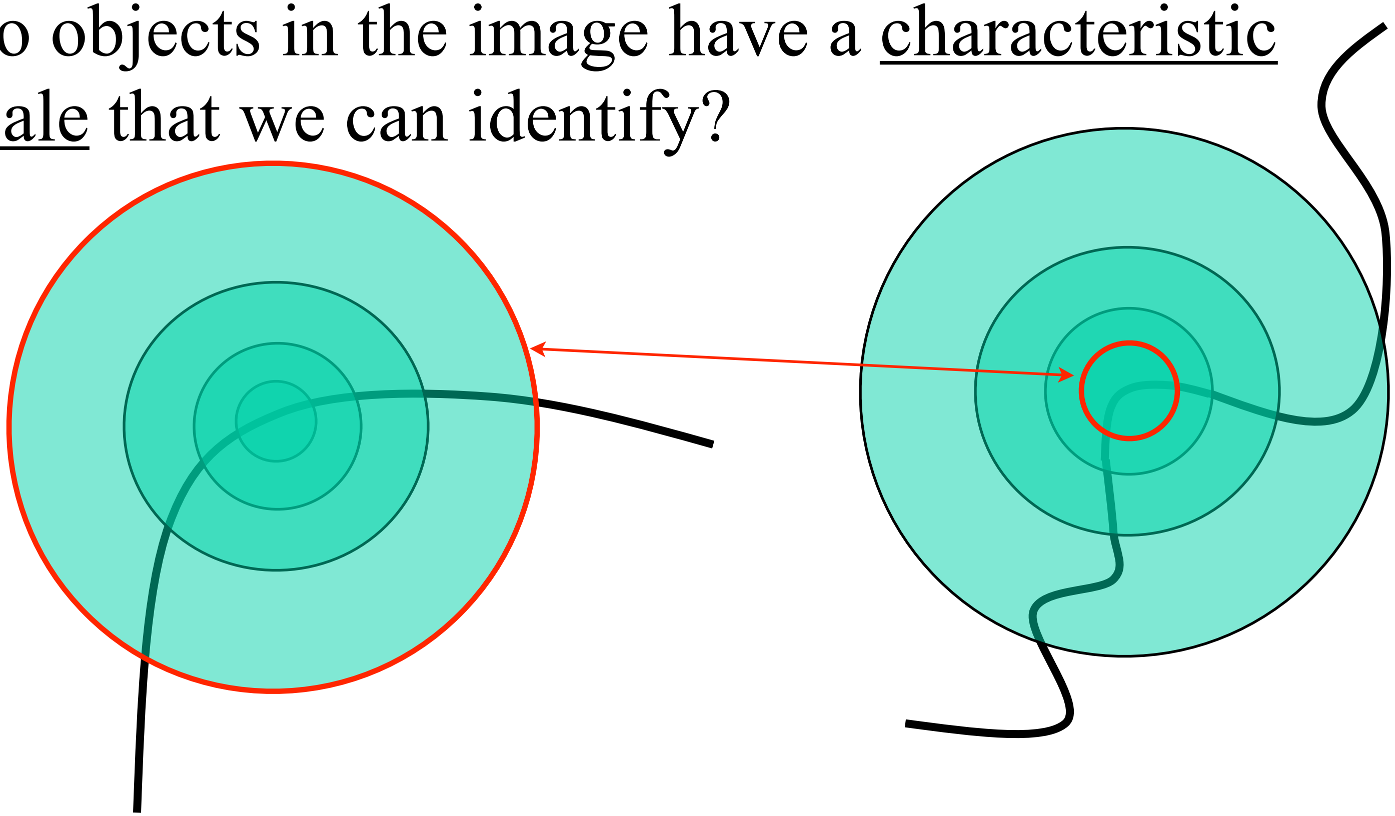
The features look the same to these two operators.

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Scale Invariant Detection

- The problem: how do we choose corresponding circles ***independently*** in each image?
- Do objects in the image have a characteristic scale that we can identify?



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Detection over scale

Requires a method to repeatably select points in location and scale:

- The only reasonable scale-space kernel is a Gaussian (Koenderink, 1984; Lindeberg, 1994)
- An efficient choice is to detect peaks in the difference of Gaussian pyramid (Burt & Adelson, 1983; Crowley & Parker, 1984 – but examining more scales)
- Difference-of-Gaussian with constant ratio of scales is a close approximation to Lindeberg’s scale-normalized Laplacian (can be shown from the heat diffusion equation)

Scale Invariant Detection

- Functions for determining scale $f = \text{Kernel} * \text{Image}$

Kernels:

$$L = \sigma^2 \left(G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma) \right)$$

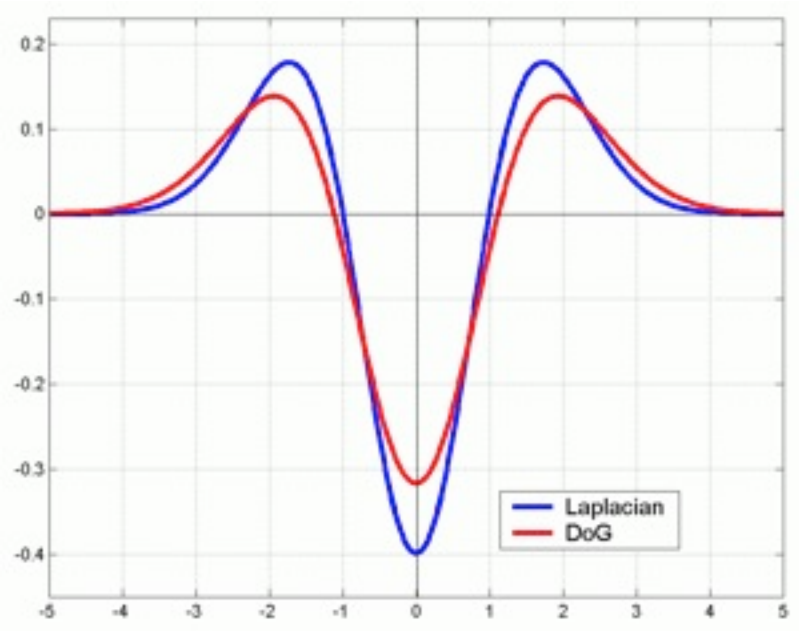
(Laplacian: 2nd derivative of Gaussian)

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

(Difference of Gaussians)

where Gaussian

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



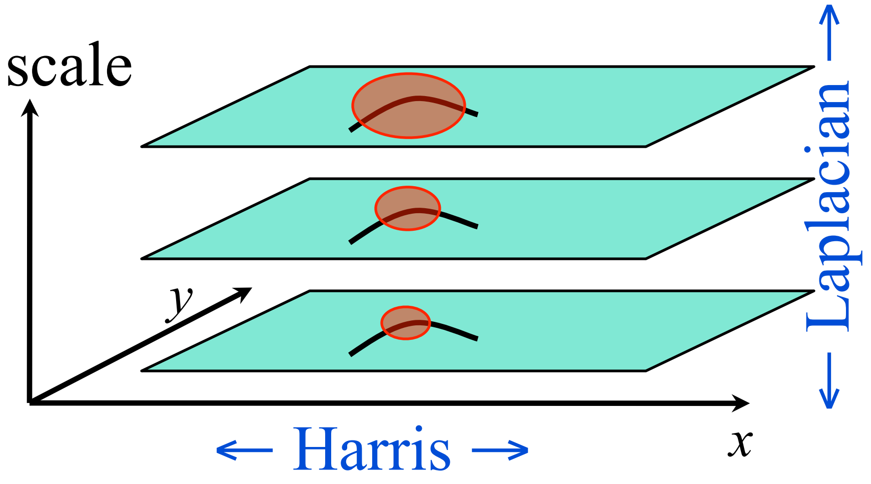
Note: both kernels are invariant to *scale* and *rotation*

http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Scale Invariant Detectors

- Harris-Laplacian[1]
  Find local maximum of:
  - Harris corner detector in space (image coordinates)
  - Laplacian in scale



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

Laplacian of Gaussian for selection of characteristic scale

http://www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/mikolajczyk_ijcv2004.pdf

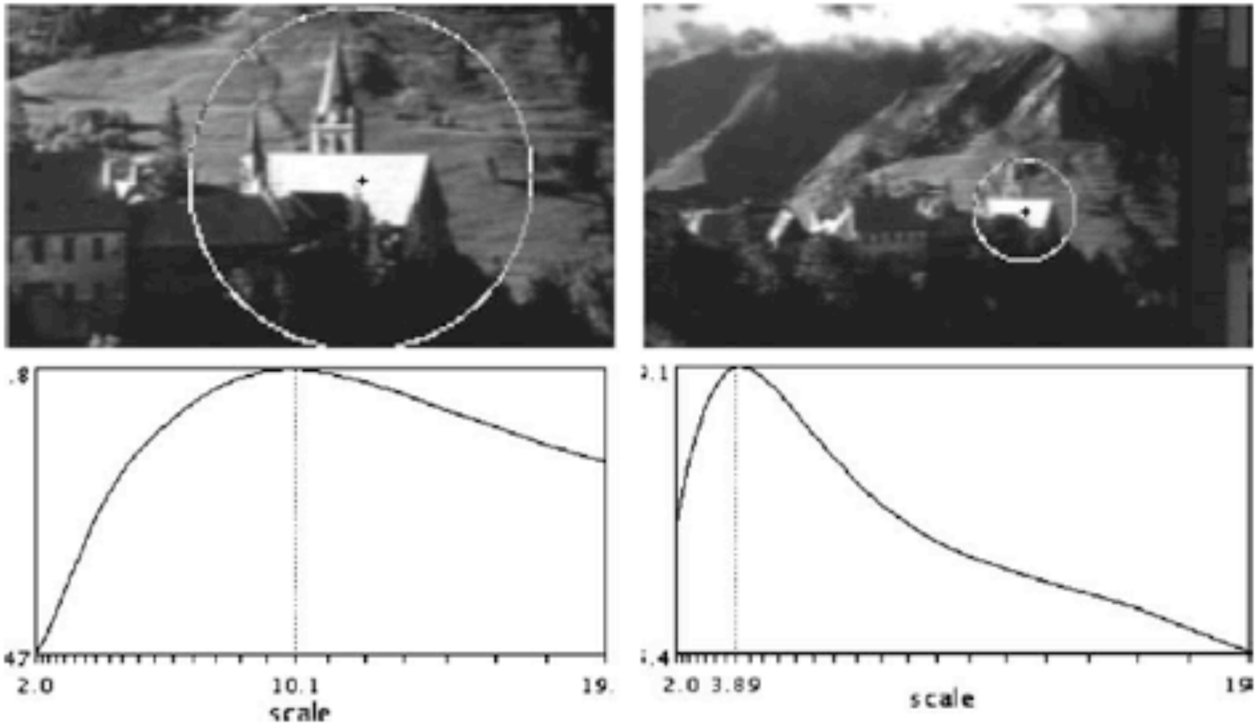
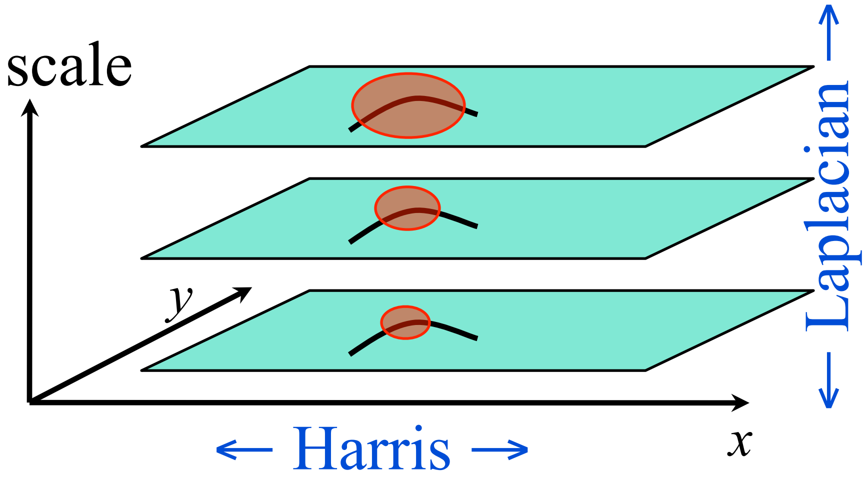


Figure 1. Example of characteristic scales. The top row shows two images taken with different focal lengths. The bottom row shows the response $F_{\text{norm}}(\mathbf{x}, \sigma_n)$ over scales where F_{norm} is the normalized LoG (cf. Eq. (3)). The characteristic scales are 10.1 and 3.89 for the left and right image, respectively. The ratio of scales corresponds to the scale factor (2.5) between the two images. The radius of displayed regions in the top row is equal to 3 times the characteristic scale.

Scale Invariant Detectors

- Harris-Laplacian[1]

  Find local maximum of:
  - Harris corner detector in space (image coordinates)
  - Laplacian in scale



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

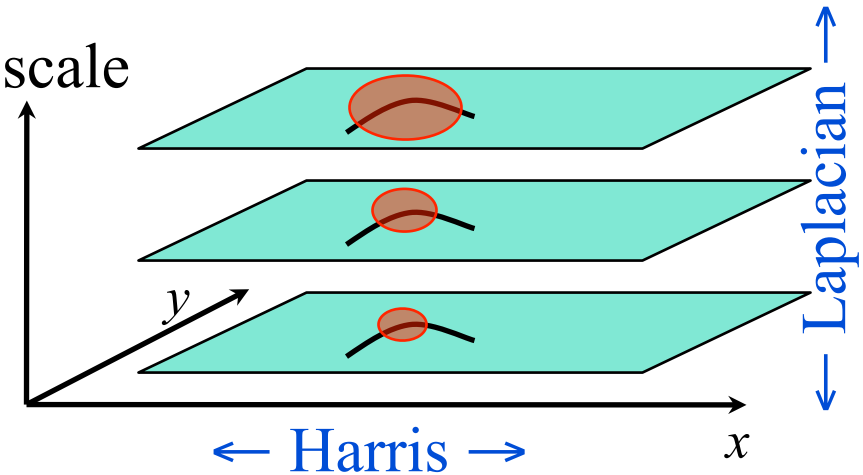
[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". Accepted to IJCV 2004

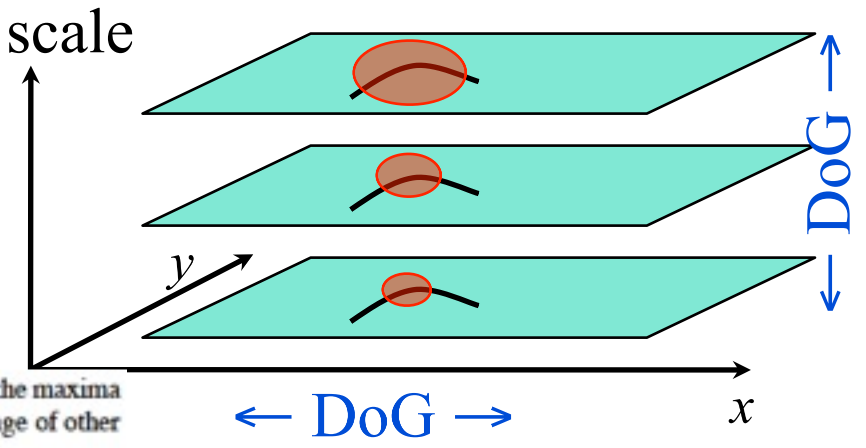
Scale Invariant Detectors

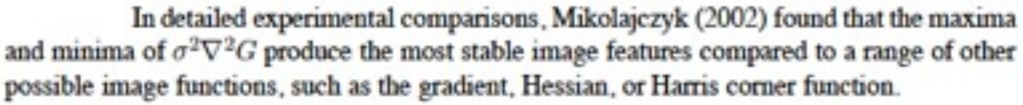
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

- Harris-Laplacian[1]
  Find local maximum of:
  - Harris corner detector in space (image coordinates)
  - Laplacian in scale



- SIFT (Lowe)[2]
  Find local maximum of:
  - Difference of Gaussians in space and scale



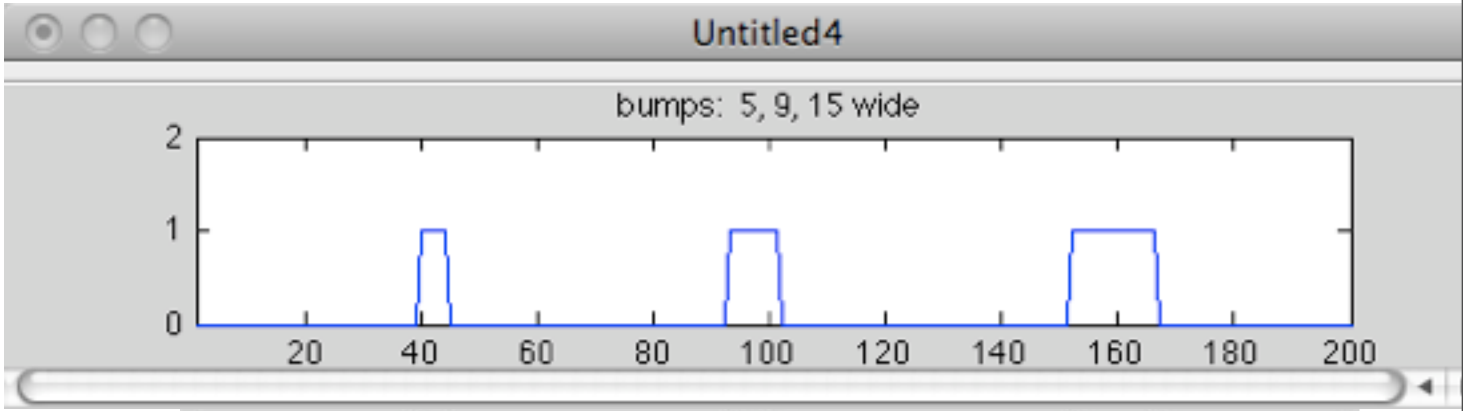

In detailed experimental comparisons, Mikolajczyk (2002) found that the maxima and minima of $\sigma^2\nabla^2 G$ produce the most stable image features compared to a range of other possible image functions, such as the gradient, Hessian, or Harris corner function.

[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

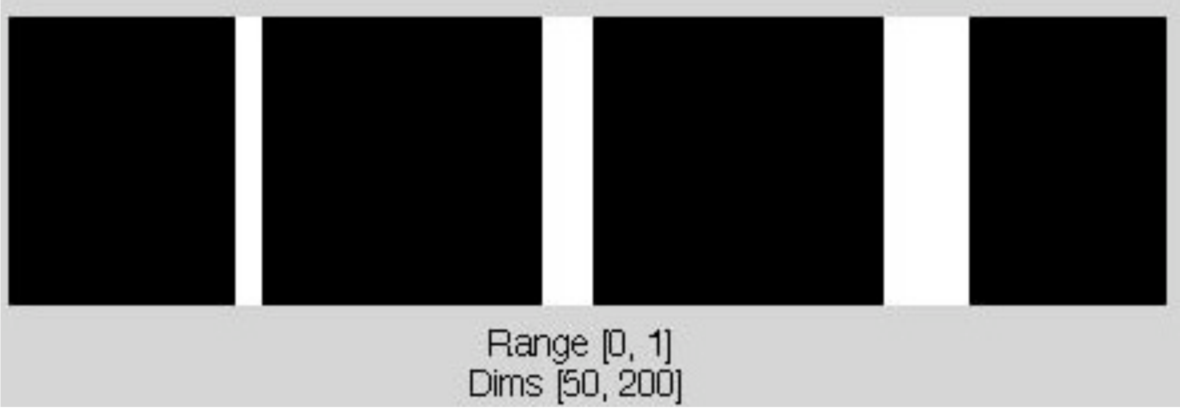
[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". Accepted to IJCV 2004

Scale-space example: 3 bumps of different widths.

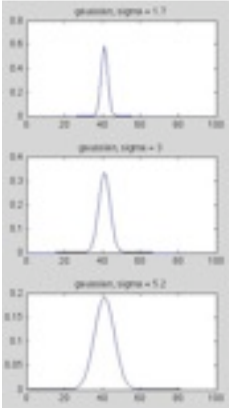
1-d bumps



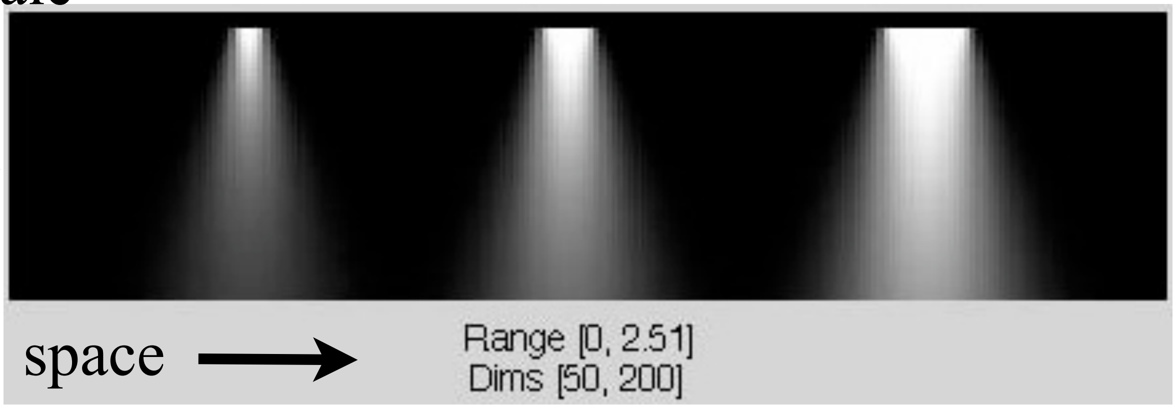
display as an image


Range [0, 1]
Dims [50, 200]

blur with Gaussians of increasing width

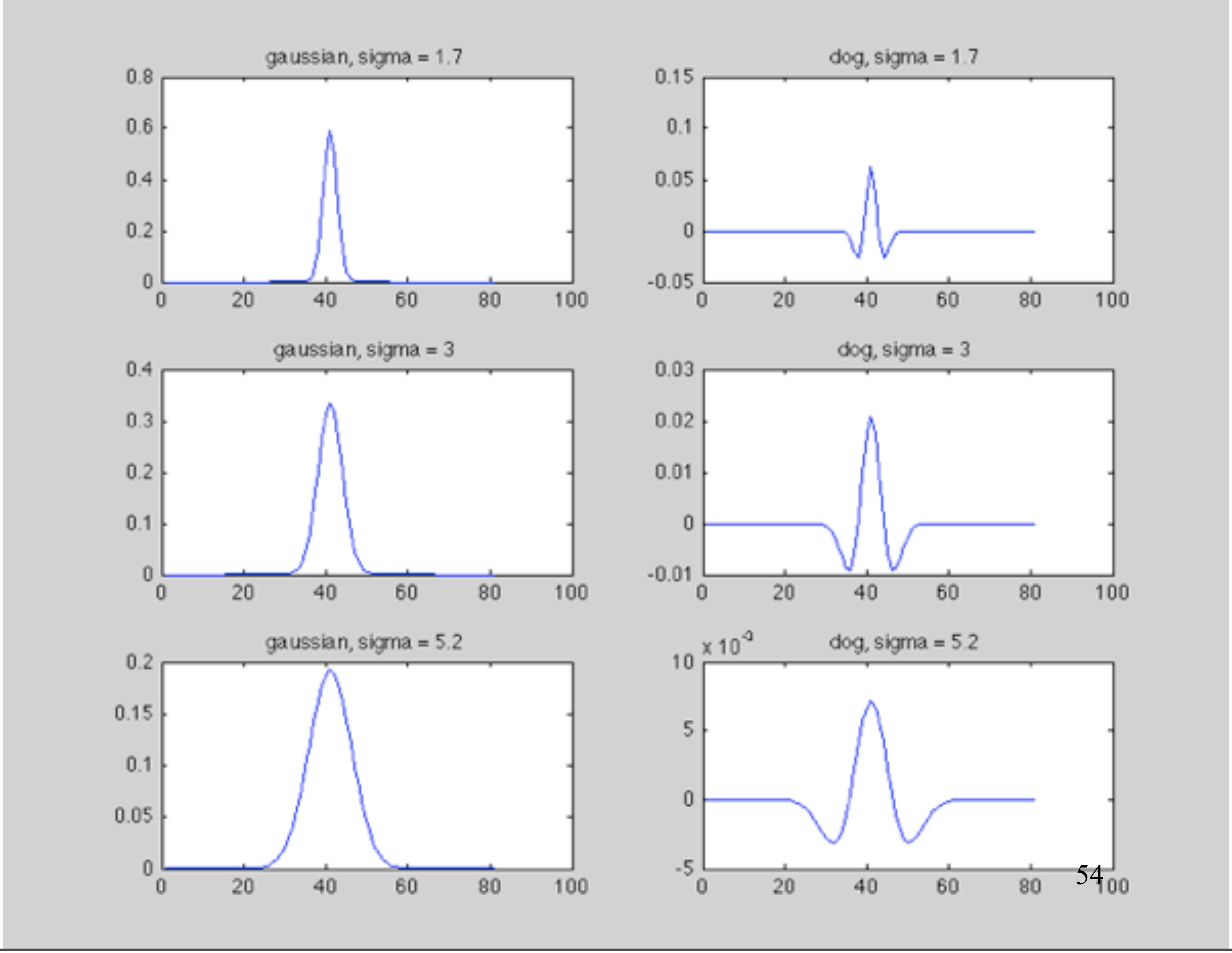


scale ↓

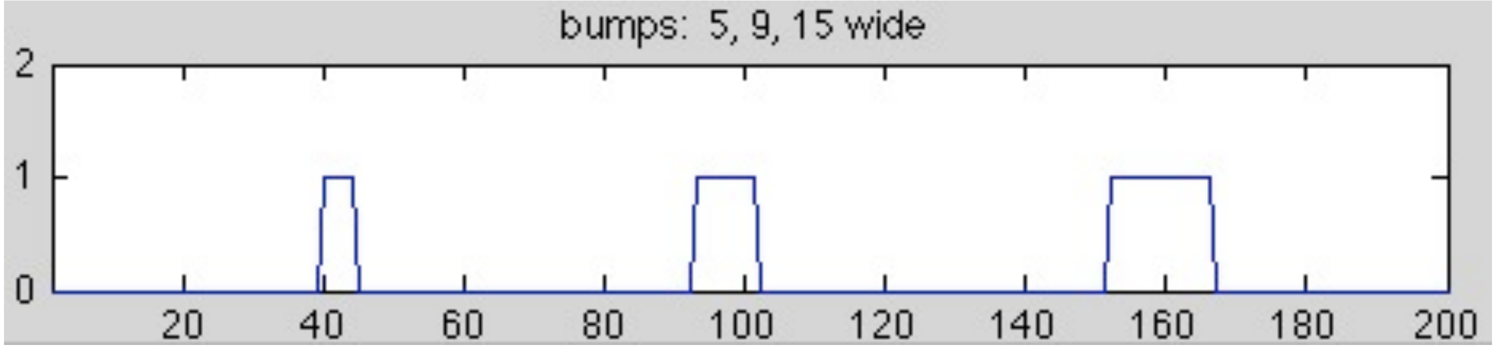

space ⟶

Range [0, 2.51]
Dims [50, 200]

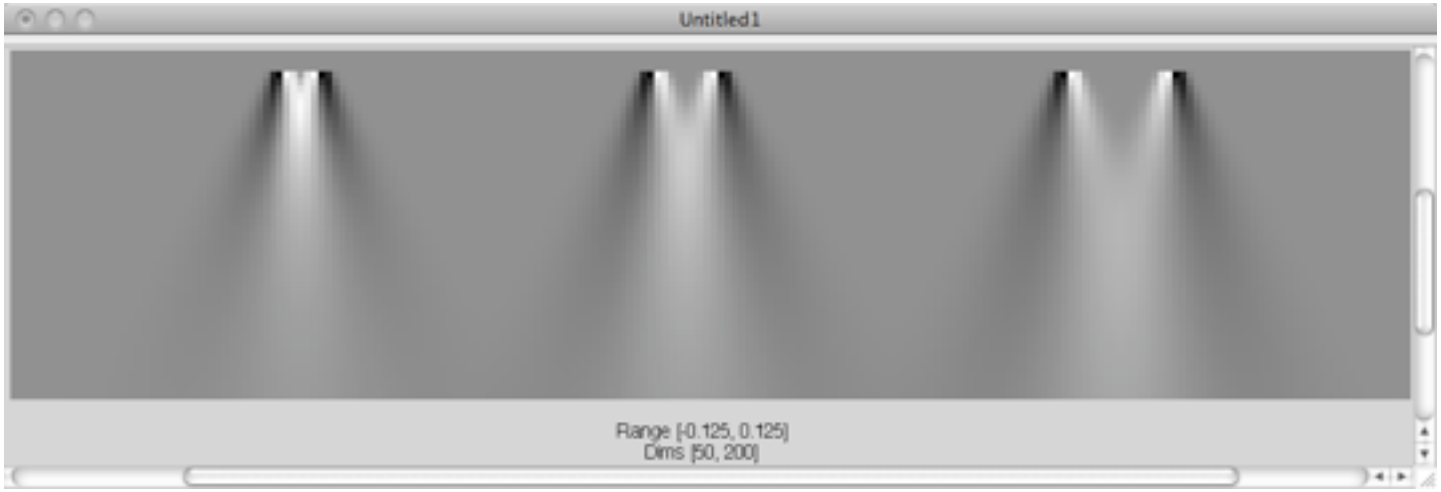
Gaussian and difference-of-Gaussian filters



The bumps, filtered by difference-of-Gaussian filters

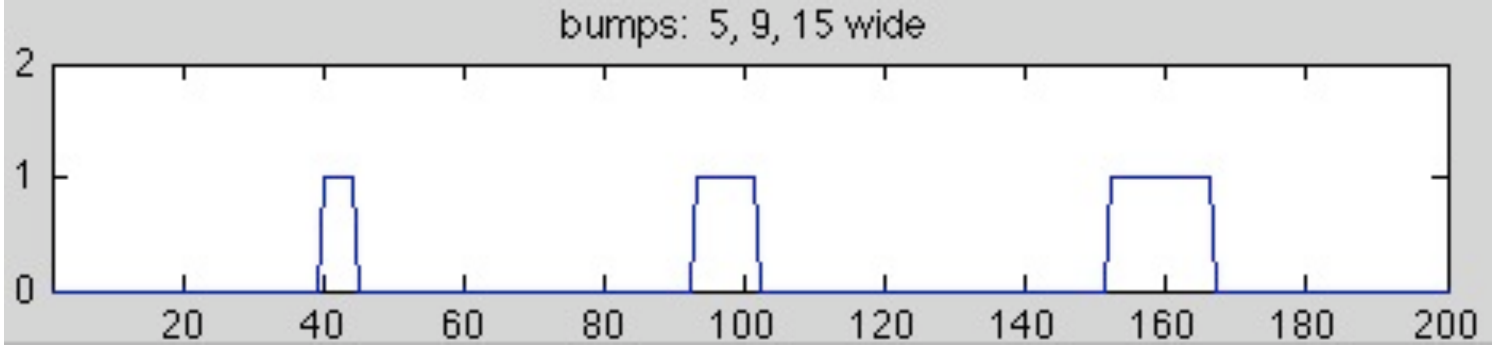


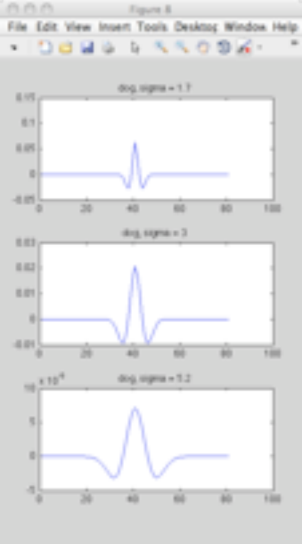
scale

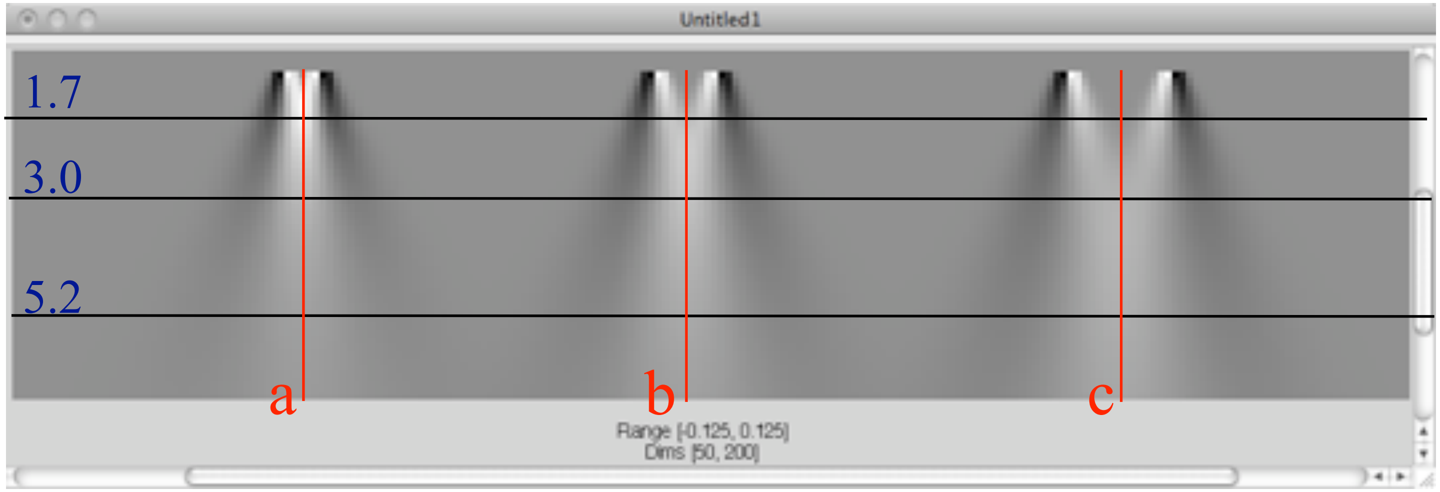


space

The bumps, filtered by difference-of-Gaussian filters



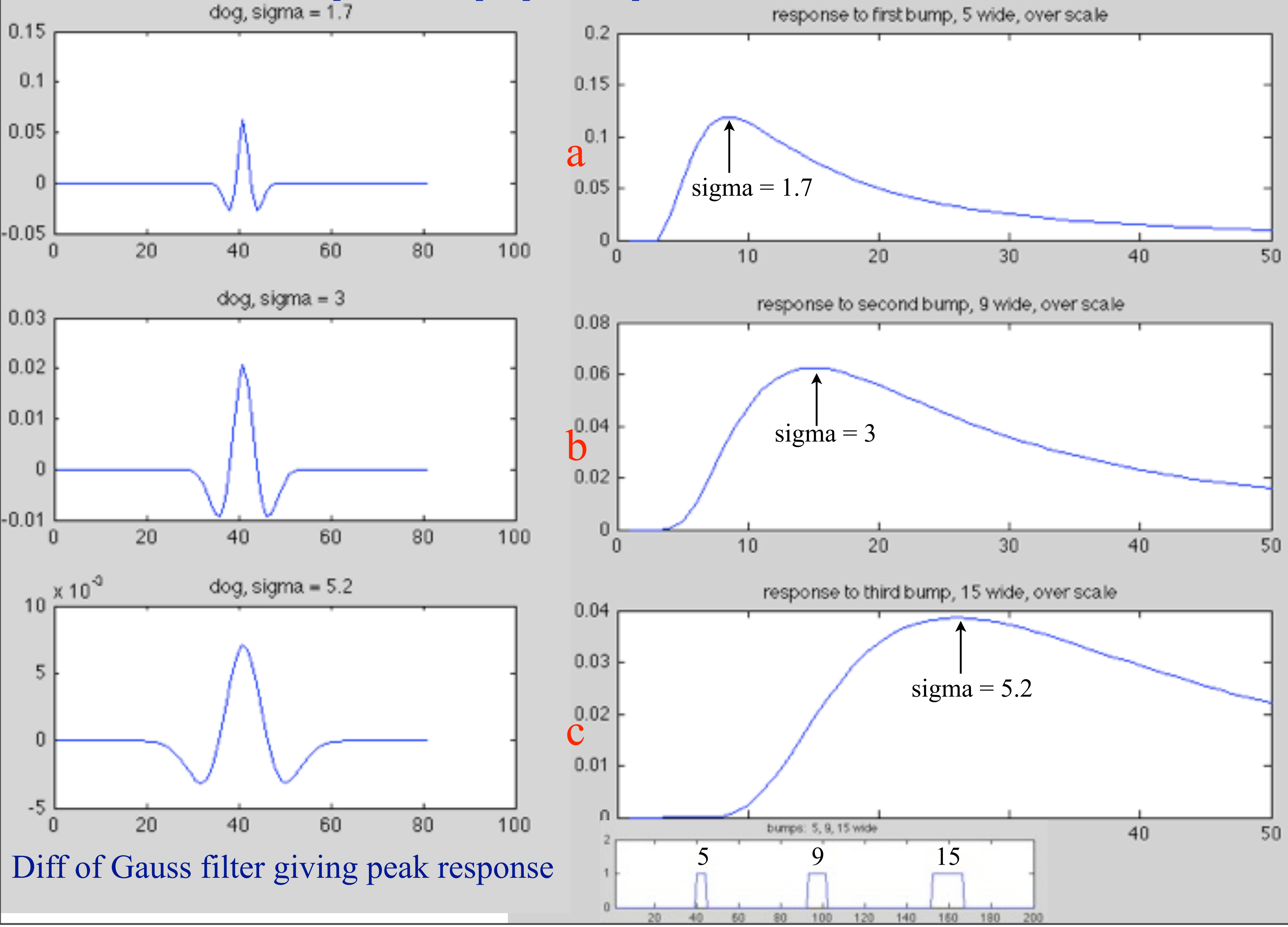




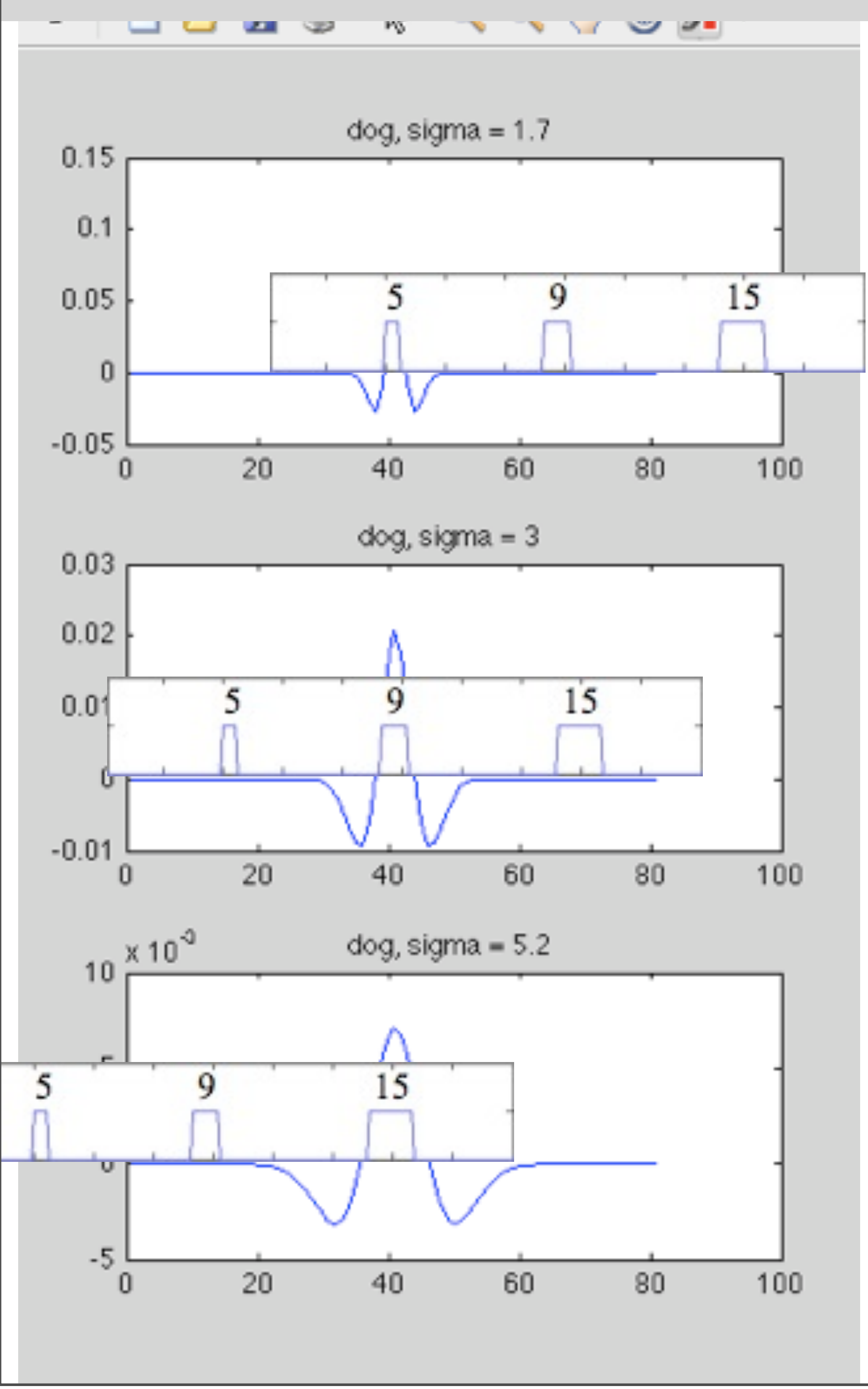
cross-sections along red lines plotted next slide

Scales of peak responses are proportional to bump width (the characteristic scale of each bump):

[1.7, 3, 5.2] ./ [5, 9, 15] = 0.3400   0.3333   0.3467



Diff of Gauss filter giving peak response



Scales of peak responses are proportional to bump width (the characteristic scale of each bump): [1.7, 3, 5.2] ./ [5, 9, 15] = 0.3400 0.3333   0.3467

Note that the max response filters each has the same relationship to the bump that it favors (the zero crossings of the filter are about at the bump edges). So the scale space analysis correctly picks out the “characteristic scale” for each of the bumps.

More generally, this happens for the features of the images we analyze.

Scale Invariant Detectors

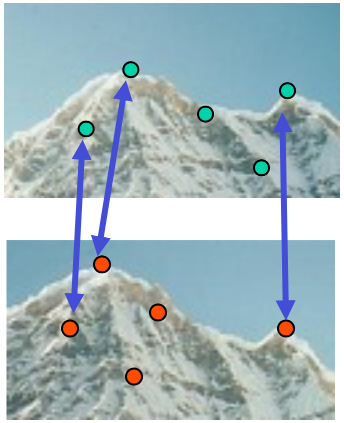
- Experimental evaluation of detectors w.r.t. scale change

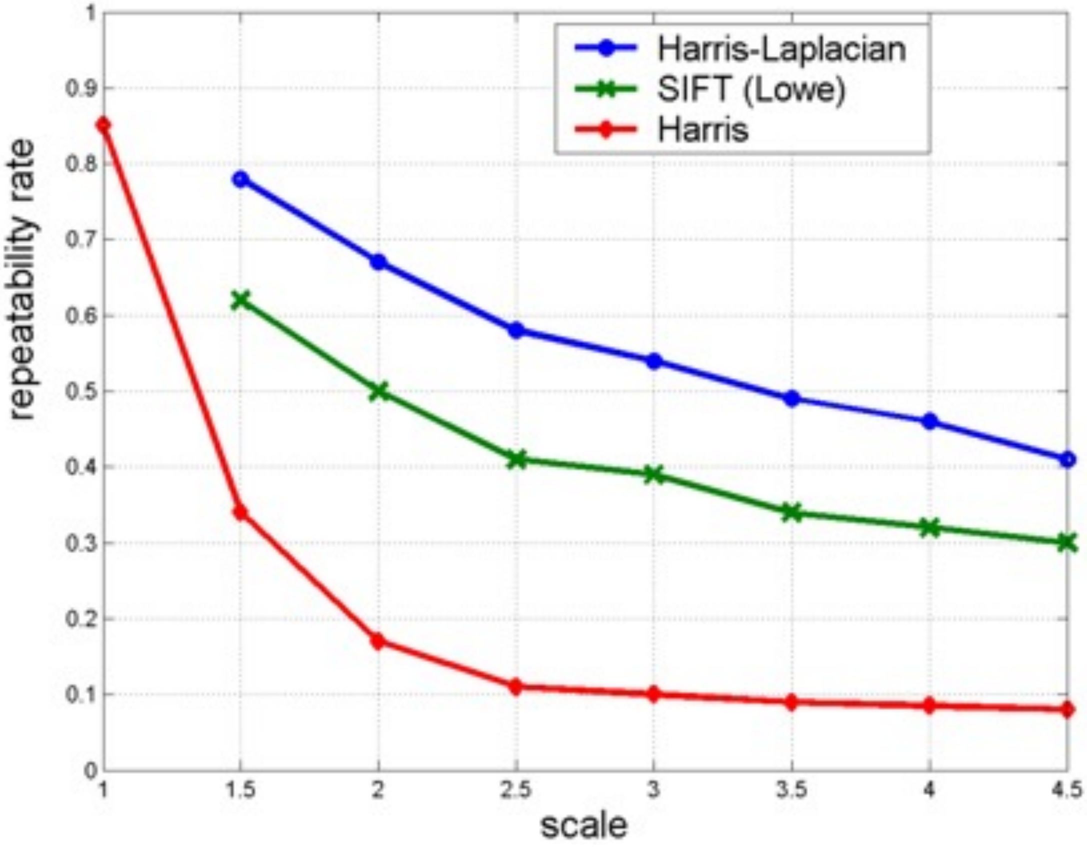
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

Darya Frolova, Denis Simakov The Weizmann Institute of Science

Repeatability rate:

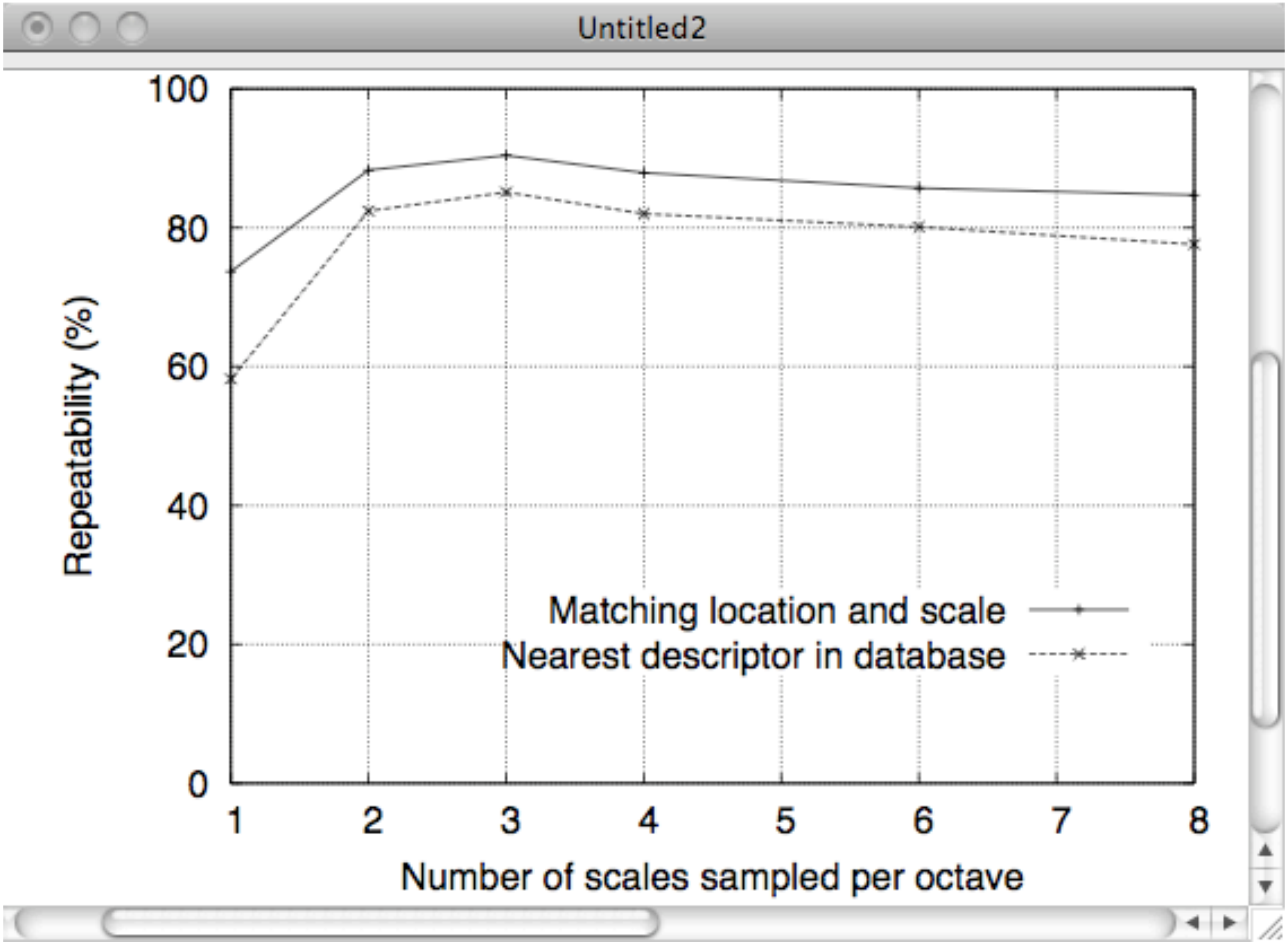
$$\frac{\text{\# correspondences}}{\text{\# possible correspondences}}$$





K.Mikolajczyk, C.Schmid. “Indexing Based on Scale Invariant Interest Points”. ICCV 2001

Repeatability vs number of scales sampled per octave



David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2 (2004), pp. 91-110

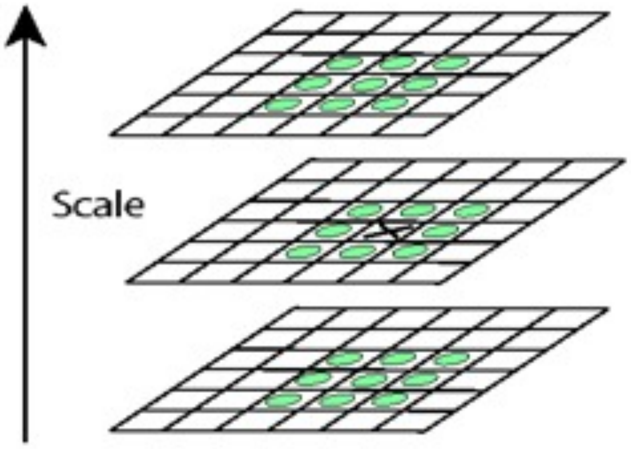
Some details of key point localization over scale and space

- Detect maxima and minima of difference-of-Gaussian in scale space
- Fit a quadratic to surrounding values for sub-pixel and sub-scale interpolation (Brown & Lowe, 2002)
- Taylor expansion around point:

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^T\frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x}$$

- Offset of extremum (use finite differences for derivatives):

$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1}\frac{\partial D}{\partial \mathbf{x}}$$



http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt
Darya Frolova, Denis Simakov The Weizmann Institute of Science

Scale and Rotation Invariant Detection: Summary

Scale and Rotation Invariant Detection: Summary

- Given: two images of the same scene with a large *scale difference and/or rotation* between them

Scale and Rotation Invariant Detection: Summary

- Given: two images of the same scene with a large *scale difference and/or rotation* between them
- Goal: find *the same* interest points *independently* in each image

Scale and Rotation Invariant Detection: Summary

- Given: two images of the same scene with a large *scale difference and/or rotation* between them
- Goal: find *the same* interest points *independently* in each image
- Solution: search for *maxima* of suitable functions in *scale* and in *space* (over the image). Also, find characteristic *orientation*.

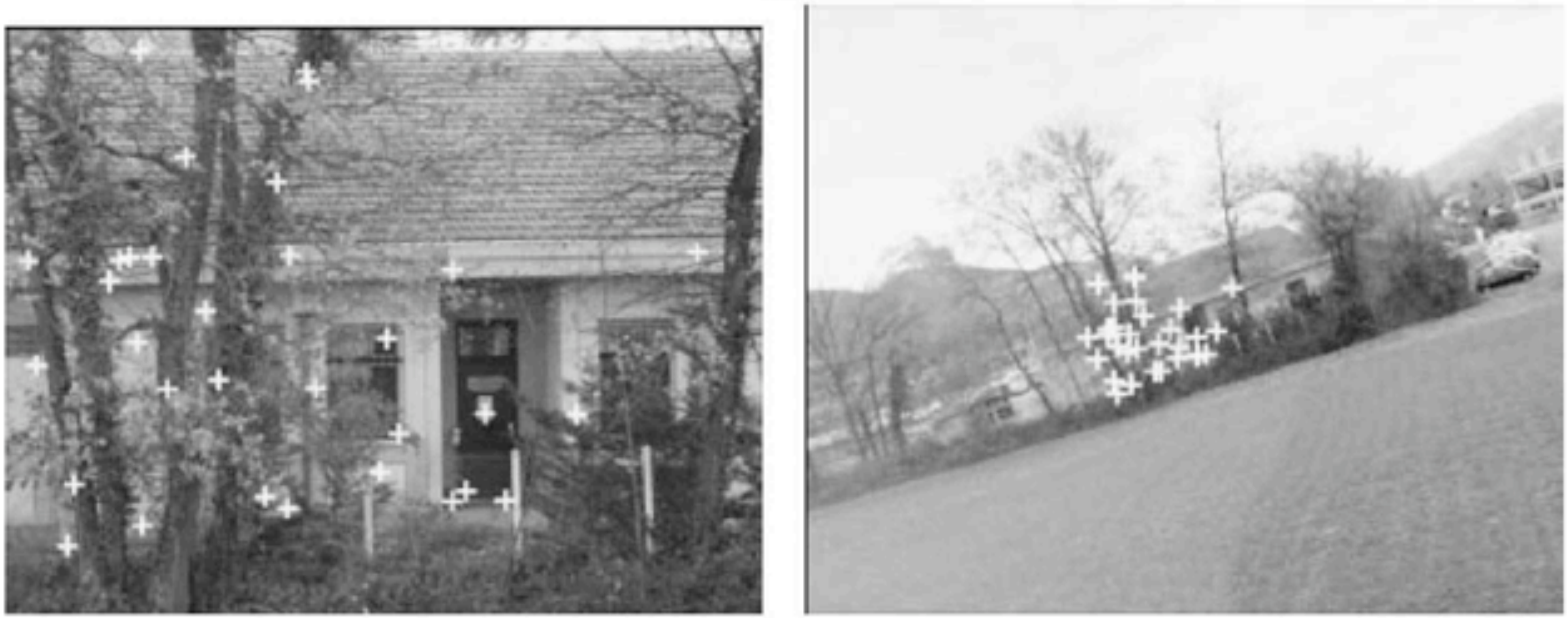
Scale and Rotation Invariant Detection: Summary

- Given: two images of the same scene with a large *scale difference and/or rotation* between them
- Goal: find *the same* interest points *independently* in each image
- Solution: search for *maxima* of suitable functions in *scale* and in *space* (over the image). Also, find characteristic *orientation*.

Methods:

1. Harris-Laplacian [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. SIFT [Lowe]: maximize Difference of Gaussians over scale and space

Example of keypoint detection



(c)

Figure 12. Robust matching: Harris-Laplace detects 190 and 213 points in the left and right images, respectively (a). 58 points are initially matched (b). There are 32 inliers to the estimated homography (c), all of which are correct. The estimated scale factor is 4.9 and the estimated rotation angle is 19 degrees.

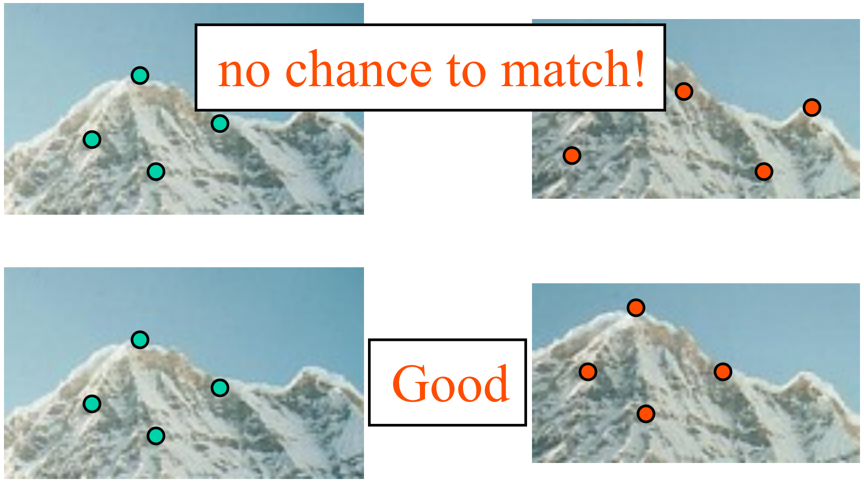
http://www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/mikolajczyk_ijcv2004.pdf

Outline

- Feature point detection
  - Harris corner detector
  - finding a characteristic scale
- Local image description
  - SIFT features

Recall: Matching with Features

- Problem 1:
  - Detect the *same* point *independently* in both images

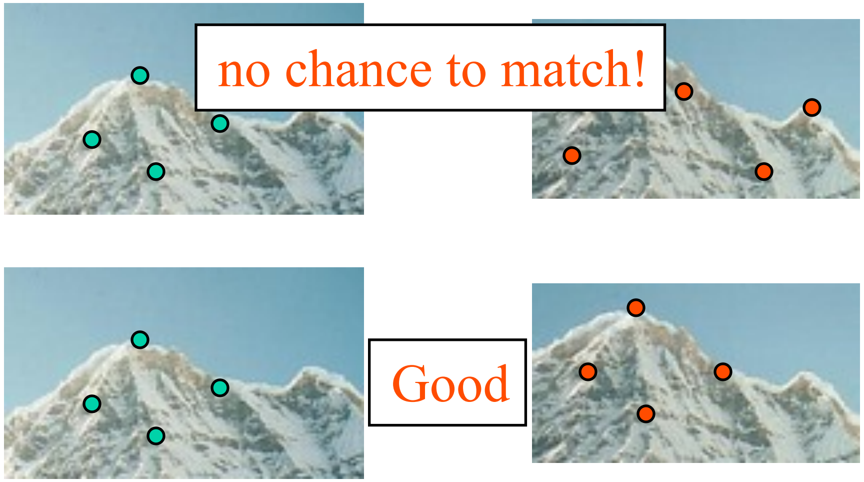


no chance to match!

Good

Recall: Matching with Features

- Problem 1:
  - Detect the *same* point *independently* in both images



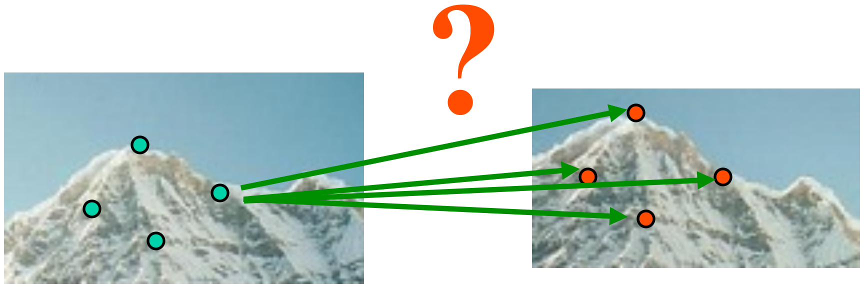
no chance to match!

Good

We need a repeatable detector

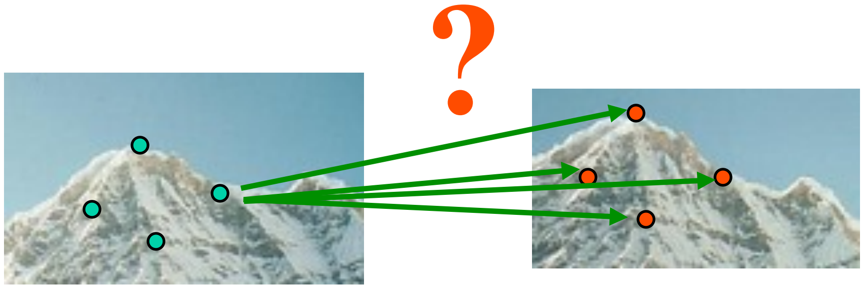
Recall: Matching with Features

- Problem 2:
  - For each point correctly recognize the corresponding one



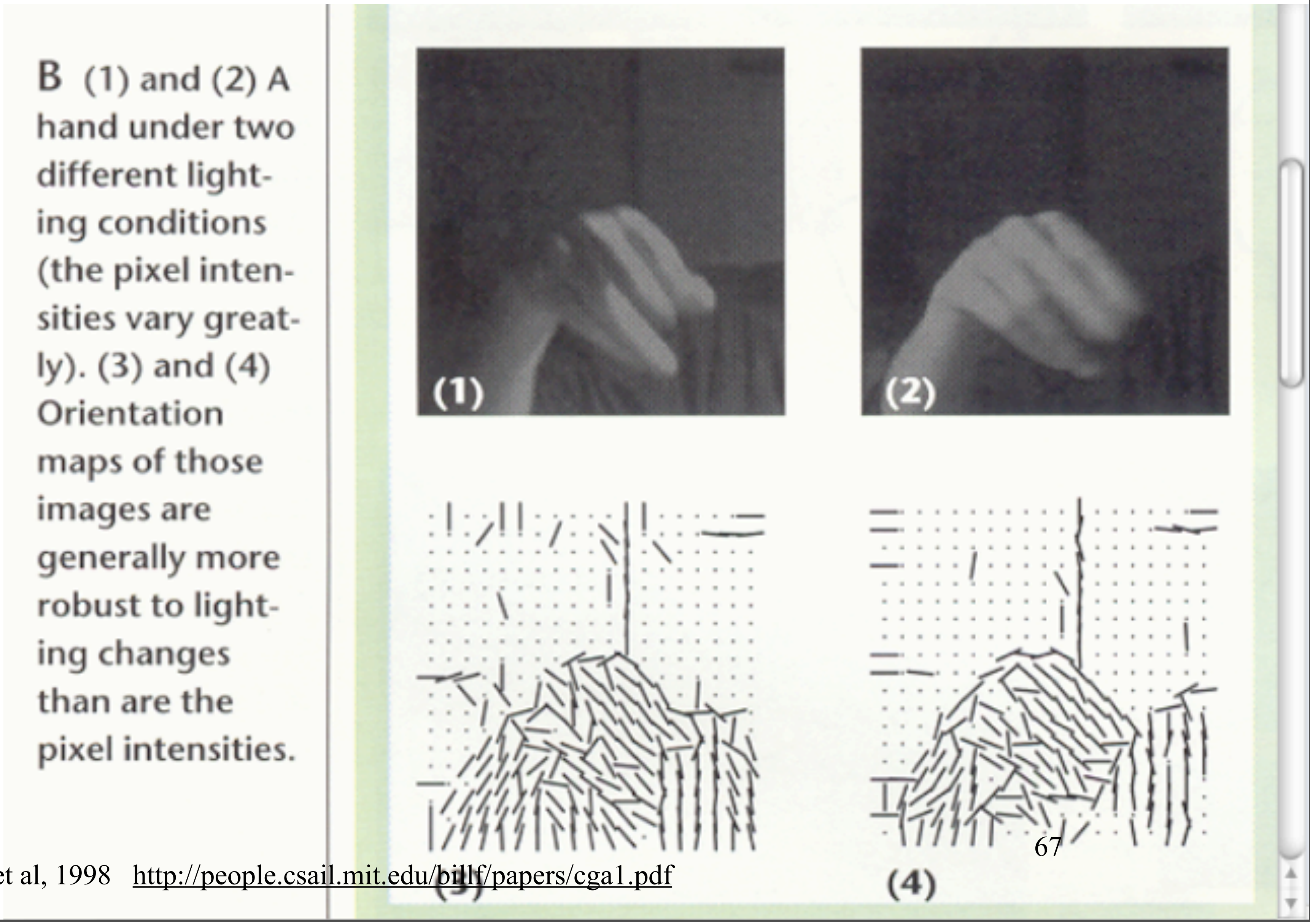
Recall: Matching with Features

- Problem 2:
  - For each point correctly recognize the corresponding one



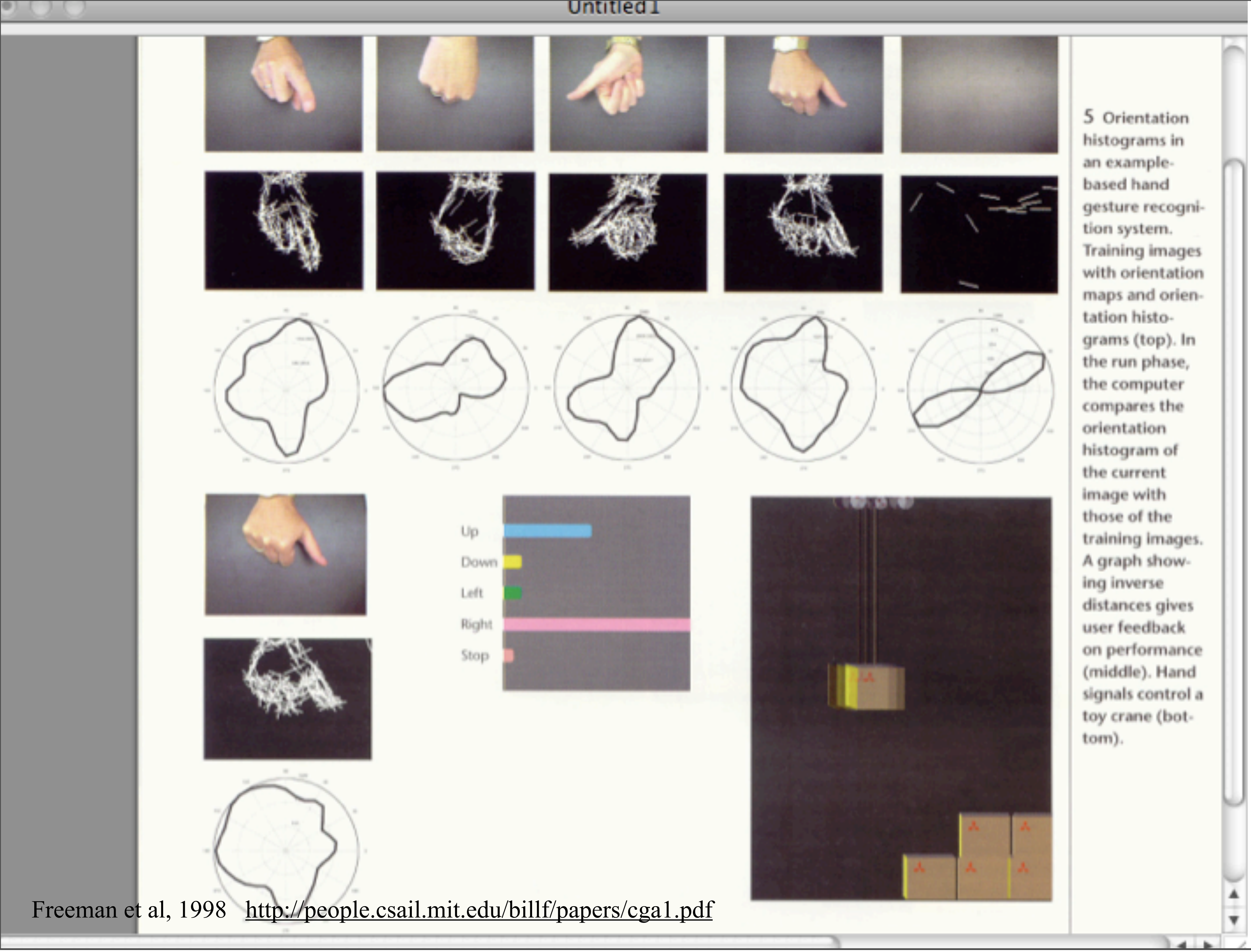
We need a reliable and distinctive descriptor

Motivation for using gradients: gradient orientations somewhat invariant to lighting variations



B (1) and (2) A hand under two different lighting conditions (the pixel intensities vary greatly). (3) and (4) Orientation maps of those images are generally more robust to lighting changes than are the pixel intensities.

Freeman et al, 1998 http://people.csail.mit.edu/billf/papers/cga1.pdf



5 Orientation histograms in an example-based hand gesture recognition system. Training images with orientation maps and orientation histograms (top). In the run phase, the computer compares the orientation histogram of the current image with those of the training images. A graph showing inverse distances gives user feedback on performance (middle). Hand signals control a toy crane (bottom).

Freeman et al, 1998 http://people.csail.mit.edu/billf/papers/cga1.pdf

CVPR 2003 Tutorial

Recognition and Matching Based on Local Invariant Features

David Lowe

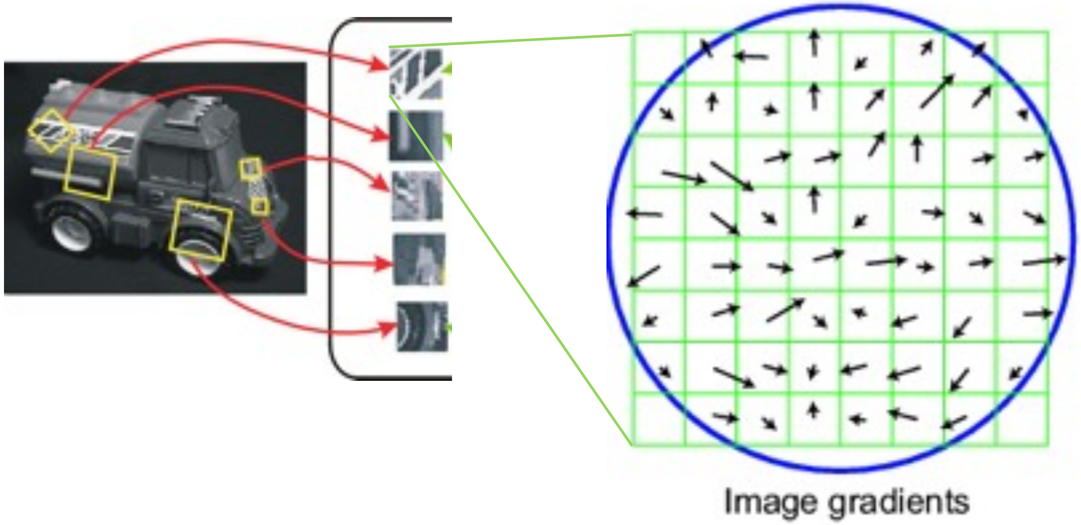
Computer Science Department

University of British Columbia

http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf

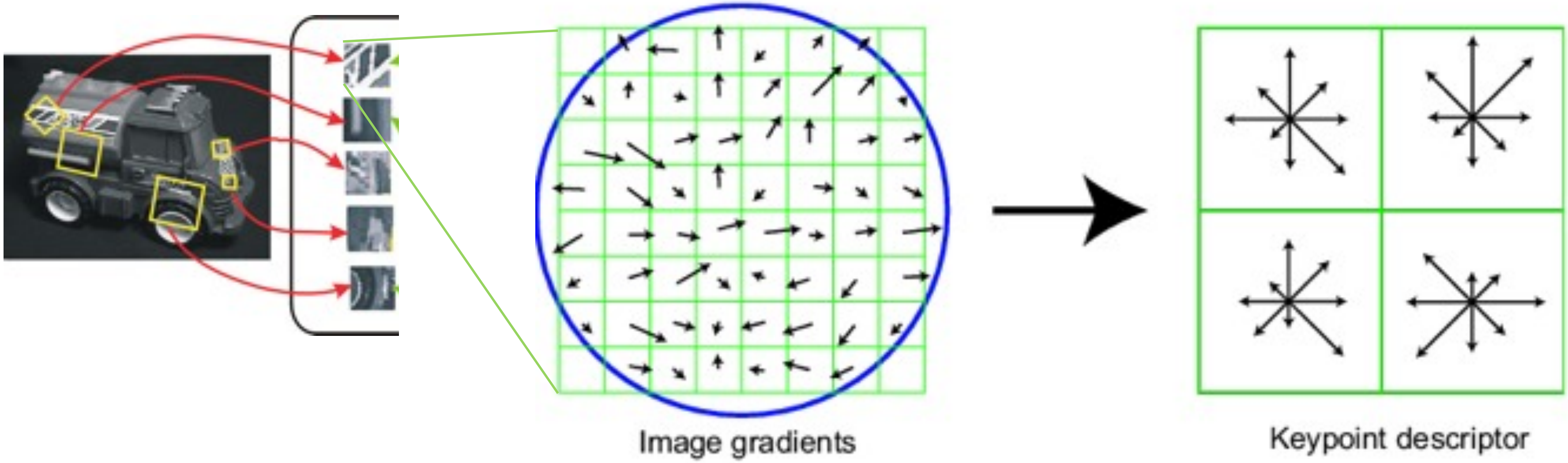
SIFT vector formation

- Computed on rotated and scaled version of window according to computed orientation & scale
  - resample the window
- Based on gradients weighted by a Gaussian of variance half the window (for smooth falloff)



SIFT vector formation

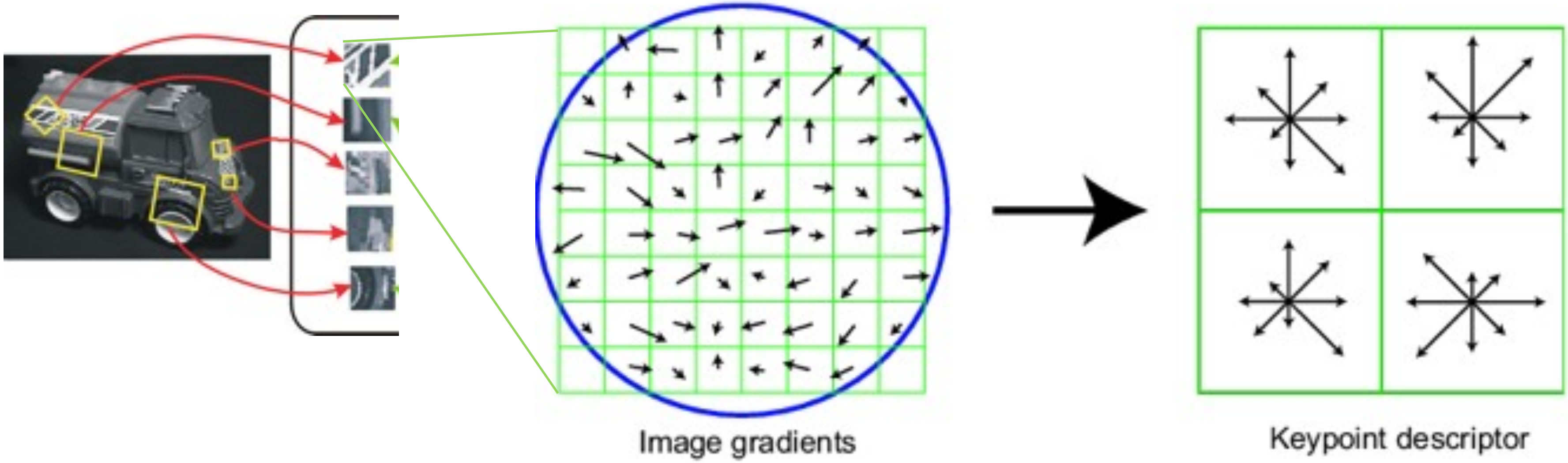
- 4x4 array of gradient orientation histograms
  - not really histogram, weighted by magnitude
- 8 orientations x 4x4 array = 128 dimensions
- Motivation: some sensitivity to spatial layout, but not too much.



showing only 2x2 here but is 4x4

Reduce effect of illumination

- 128-dim vector normalized to 1
- Threshold gradient magnitudes to avoid excessive influence of high gradients
  - after normalization, clamp gradients >0.2
  - renormalize



Tuning and evaluating the SIFT descriptors

Database images were subjected to rotation, scaling, affine stretch, brightness and contrast changes, and added noise. Feature point detectors and descriptors were compared before and after the distortions, and evaluated for:

- Sensitivity to number of histogram orientations and subregions.
- Stability to noise.
- Stability to affine change.
- Feature distinctiveness

Sensitivity to number of histogram orientations and subregions, n.

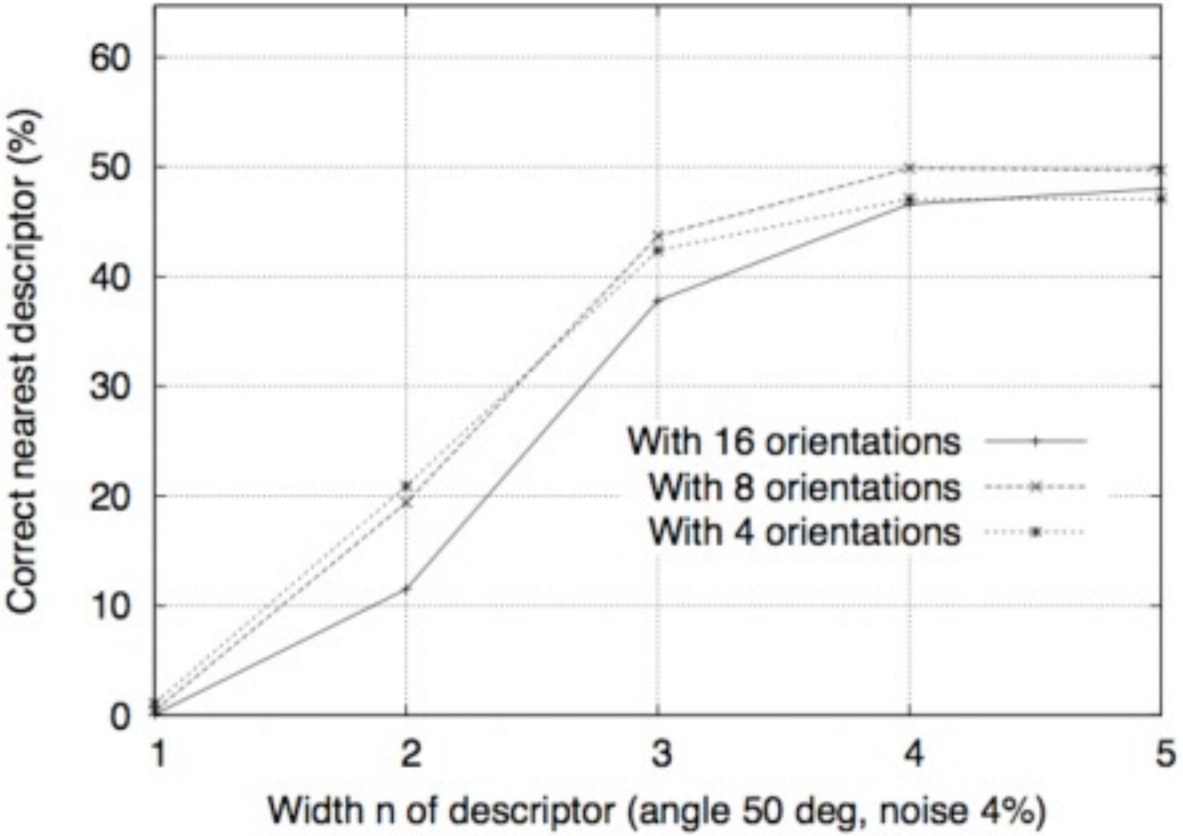
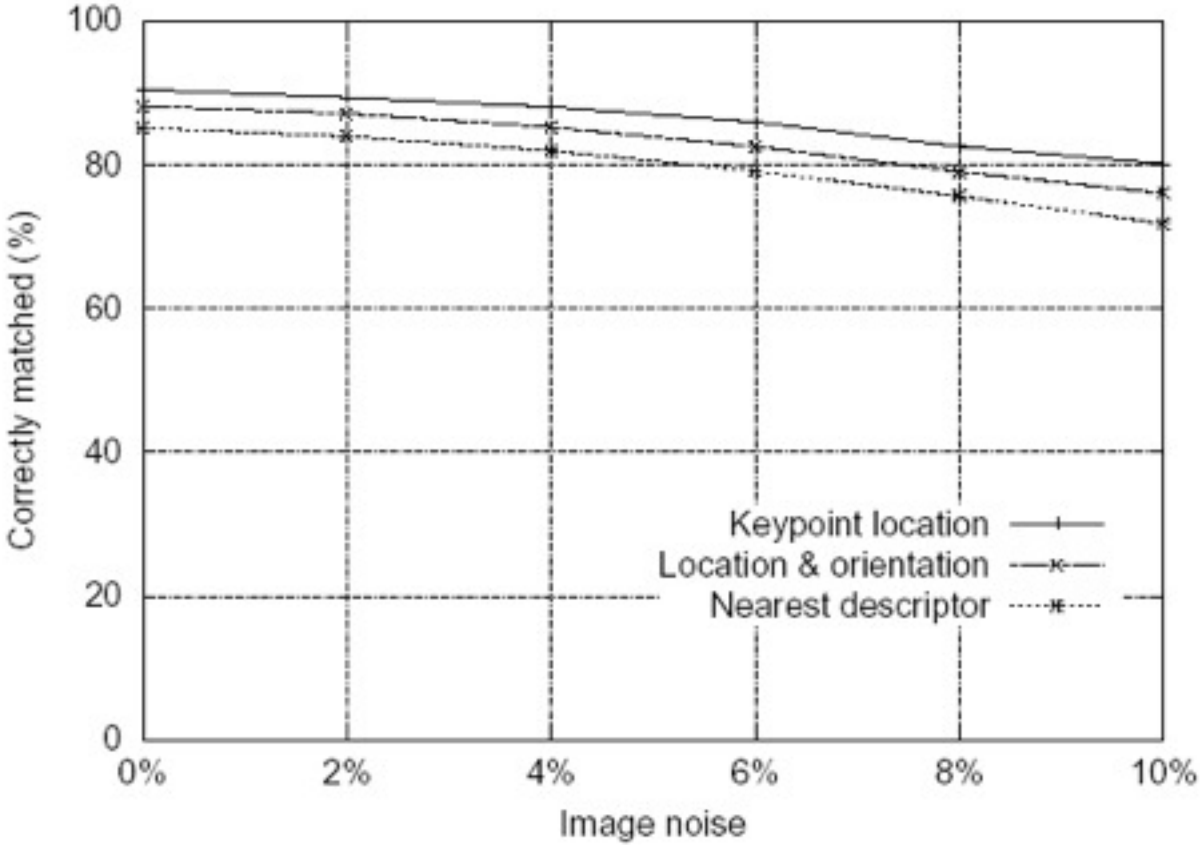


Figure 8: This graph shows the percent of keypoints giving the correct match to a database of 40,000 keypoints as a function of width of the $n \times n$ keypoint descriptor and the number of orientations in each histogram. The graph is computed for images with affine viewpoint change of 50 degrees and addition of 4% noise.

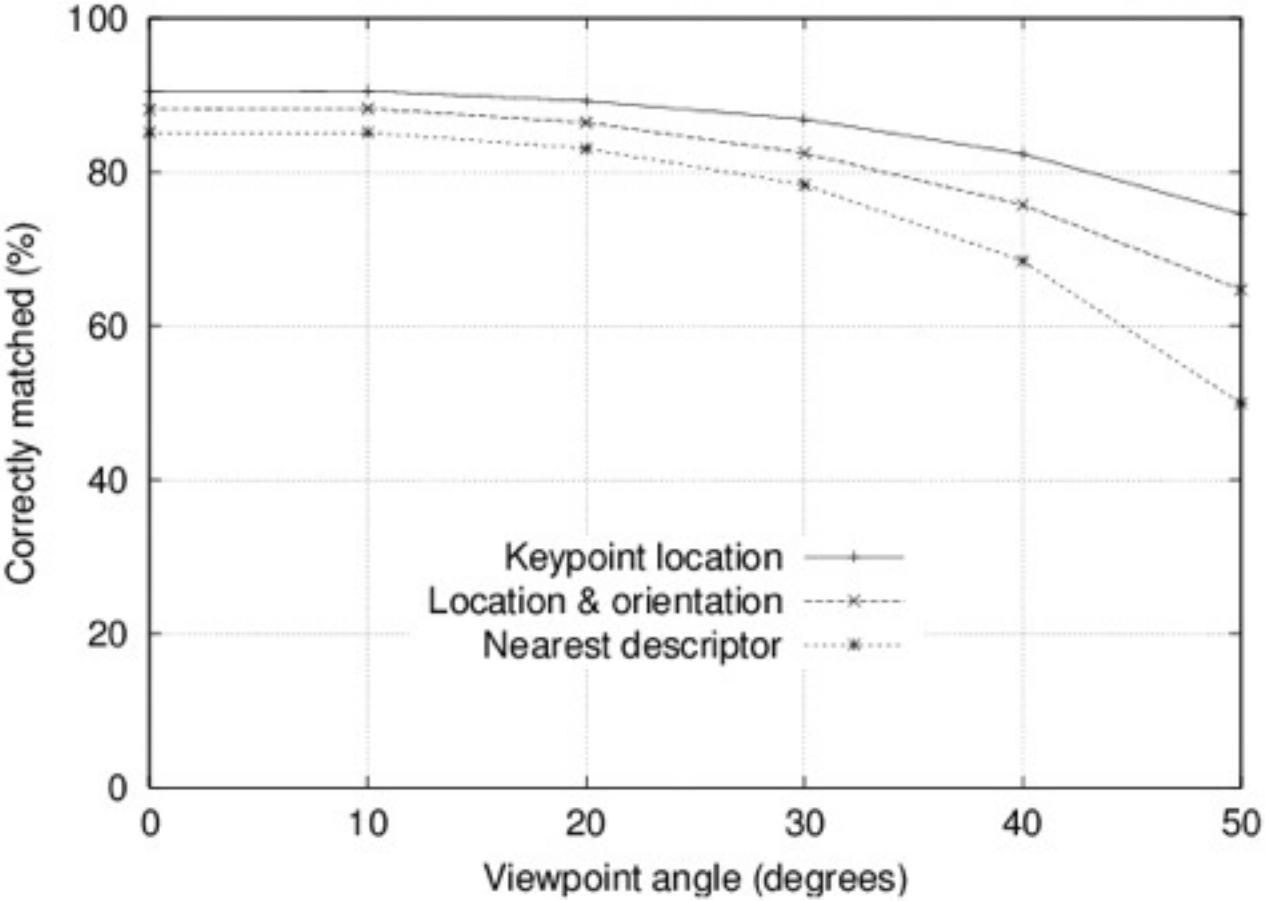
Feature stability to noise

- Match features after random change in image scale & orientation, with differing levels of image noise
- Find nearest neighbor in database of 30,000 features

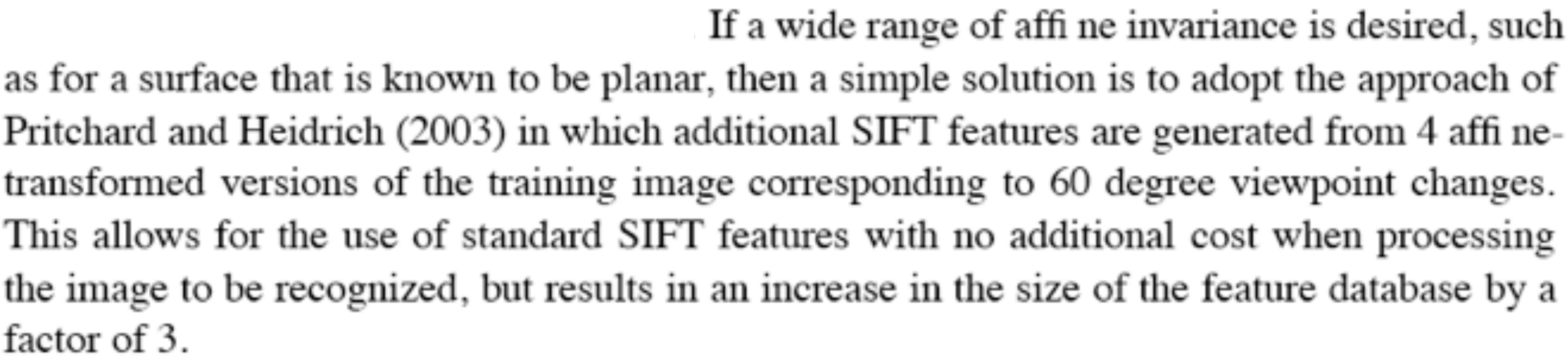


Feature stability to affine change

- Match features after random change in image scale & orientation, with 2% image noise, and affine distortion
- Find nearest neighbor in database of 30,000 features

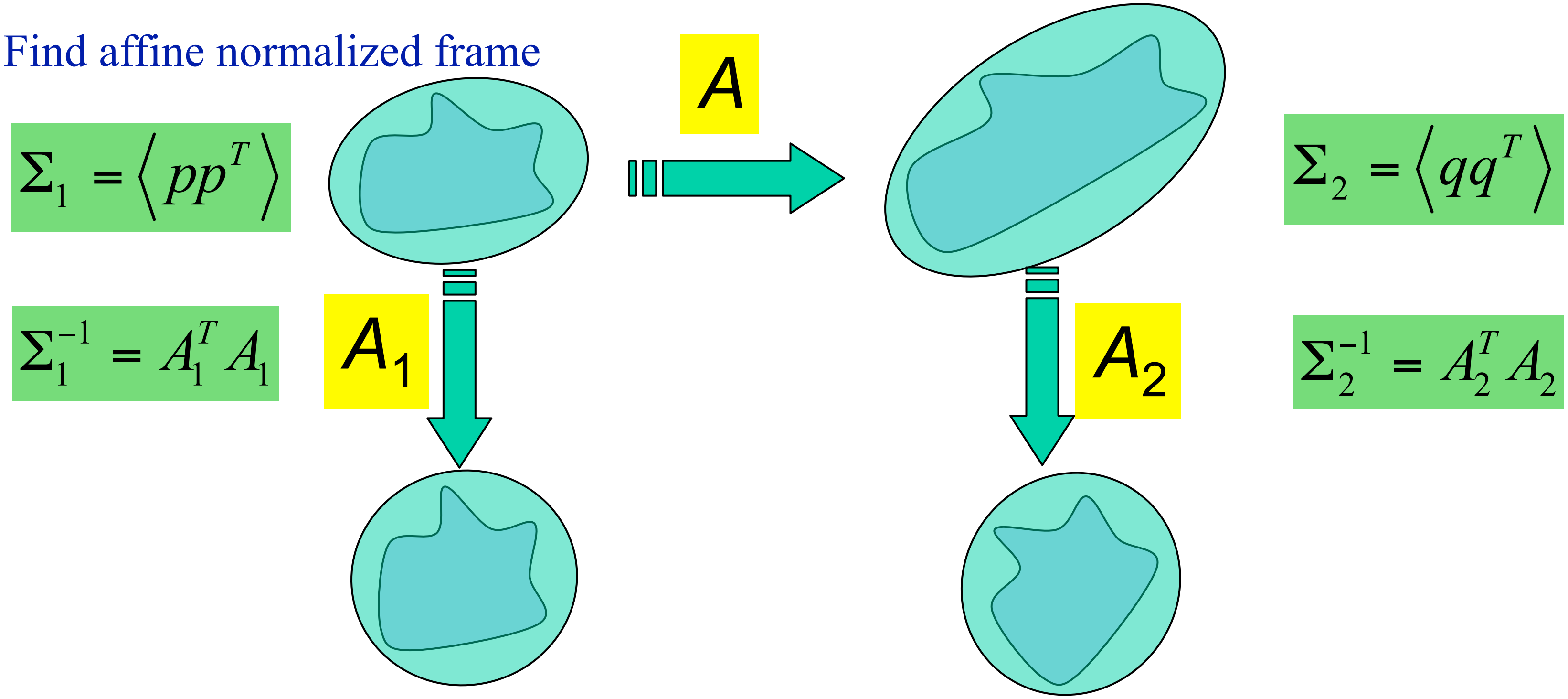


Affine Invariant Descriptors



If a wide range of affine invariance is desired, such as for a surface that is known to be planar, then a simple solution is to adopt the approach of Pritchard and Heidrich (2003) in which additional SIFT features are generated from 4 affine-transformed versions of the training image corresponding to 60 degree viewpoint changes. This allows for the use of standard SIFT features with no additional cost when processing the image to be recognized, but results in an increase in the size of the feature database by a factor of 3.

Find affine normalized frame



$\Sigma_1 = \langle pp^T \rangle$

$\Sigma_1^{-1} = A_1^T A_1$

A

A_1

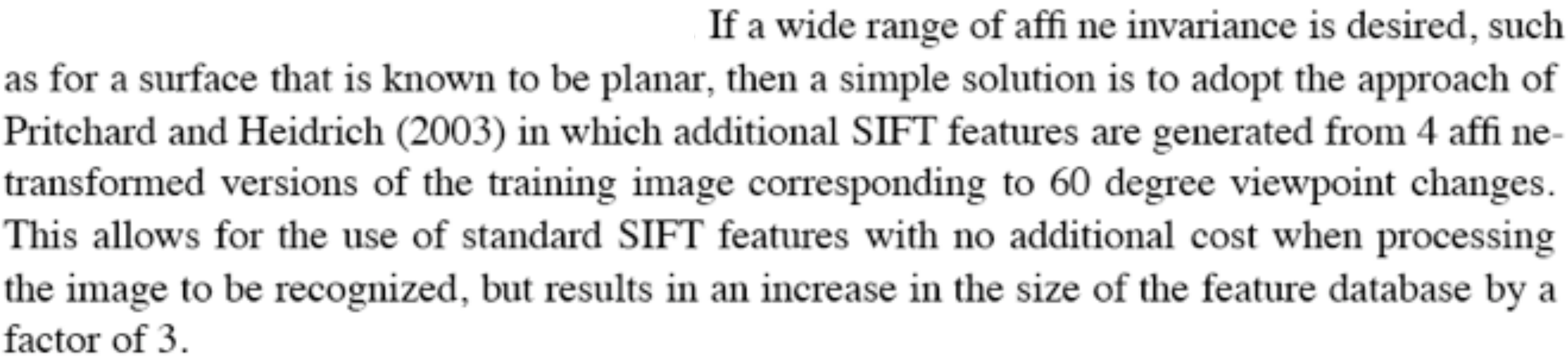
A_2

$\Sigma_2 = \langle qq^T \rangle$

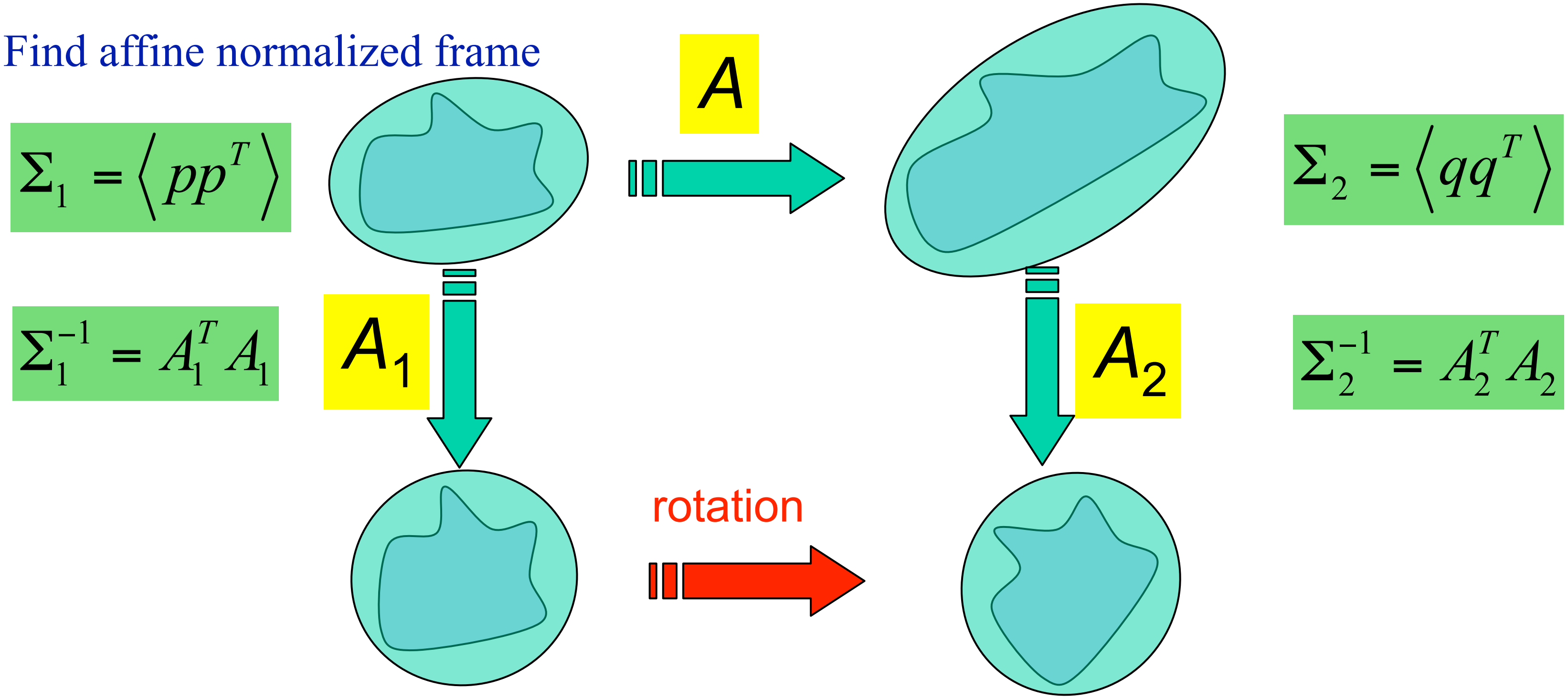
$\Sigma_2^{-1} = A_2^T A_2$

J.Matas et.al. “Rotational Invariants for Wide-baseline Stereo”. Research Report of CMP, 2003

Affine Invariant Descriptors

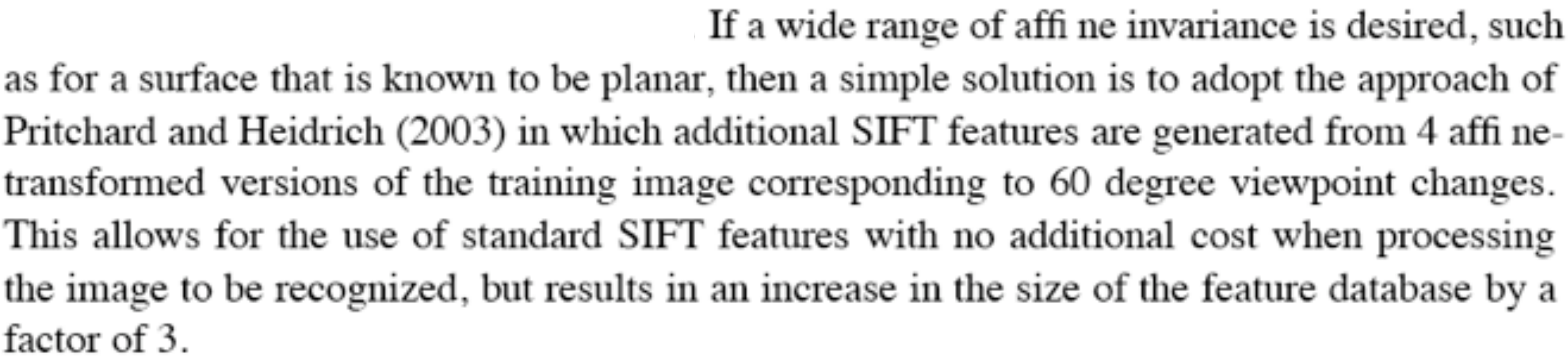


If a wide range of affine invariance is desired, such as for a surface that is known to be planar, then a simple solution is to adopt the approach of Pritchard and Heidrich (2003) in which additional SIFT features are generated from 4 affine-transformed versions of the training image corresponding to 60 degree viewpoint changes. This allows for the use of standard SIFT features with no additional cost when processing the image to be recognized, but results in an increase in the size of the feature database by a factor of 3.


Find affine normalized frame
$\Sigma_1 = \langle pp^T \rangle$
$\Sigma_1^{-1} = A_1^T A_1$
A
A_1
A_2
$\Sigma_2 = \langle qq^T \rangle$
$\Sigma_2^{-1} = A_2^T A_2$
rotation

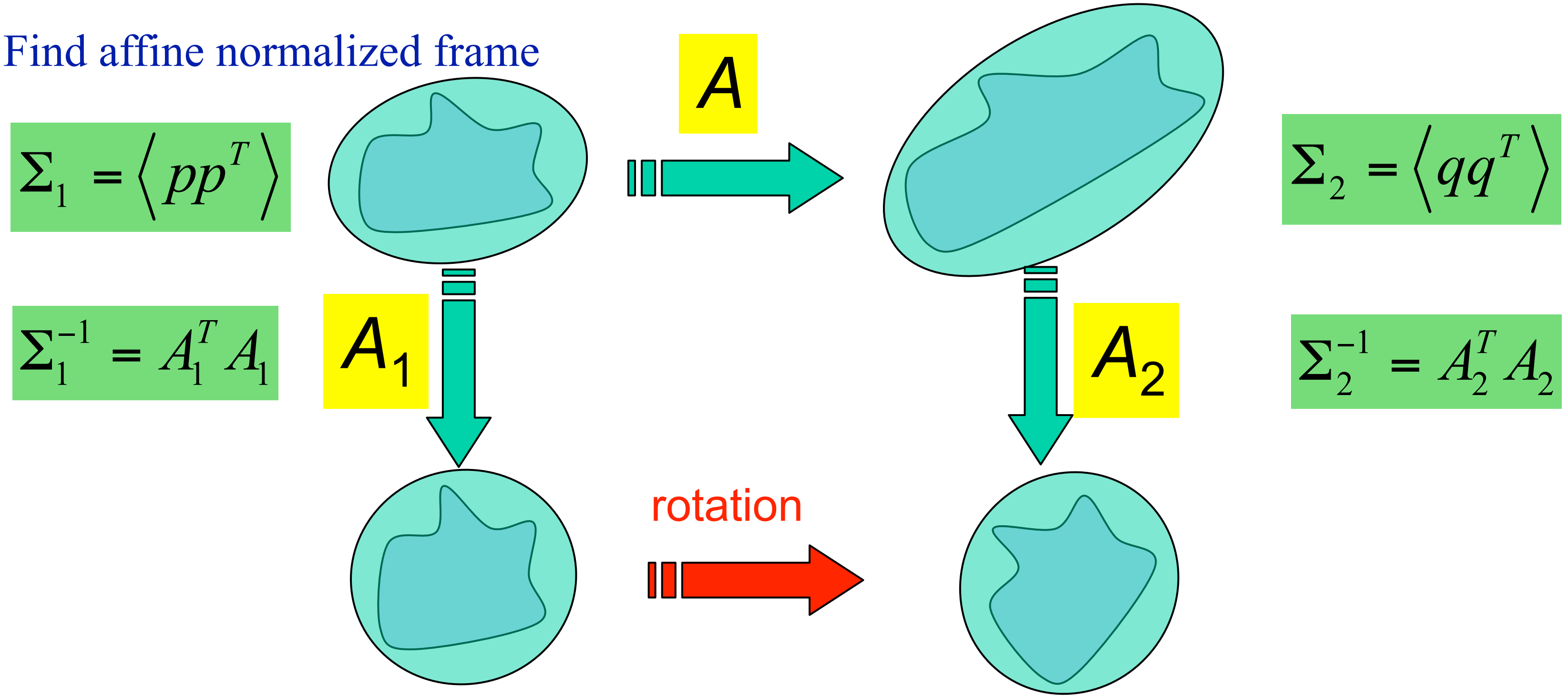
J.Matas et.al. “Rotational Invariants for Wide-baseline Stereo”. Research Report of CMP, 2003

Affine Invariant Descriptors



If a wide range of affine invariance is desired, such as for a surface that is known to be planar, then a simple solution is to adopt the approach of Pritchard and Heidrich (2003) in which additional SIFT features are generated from 4 affine-transformed versions of the training image corresponding to 60 degree viewpoint changes. This allows for the use of standard SIFT features with no additional cost when processing the image to be recognized, but results in an increase in the size of the feature database by a factor of 3.

Find affine normalized frame



$\Sigma_1 = \langle pp^T \rangle$

$\Sigma_1^{-1} = A_1^T A_1$

$\Sigma_2 = \langle qq^T \rangle$

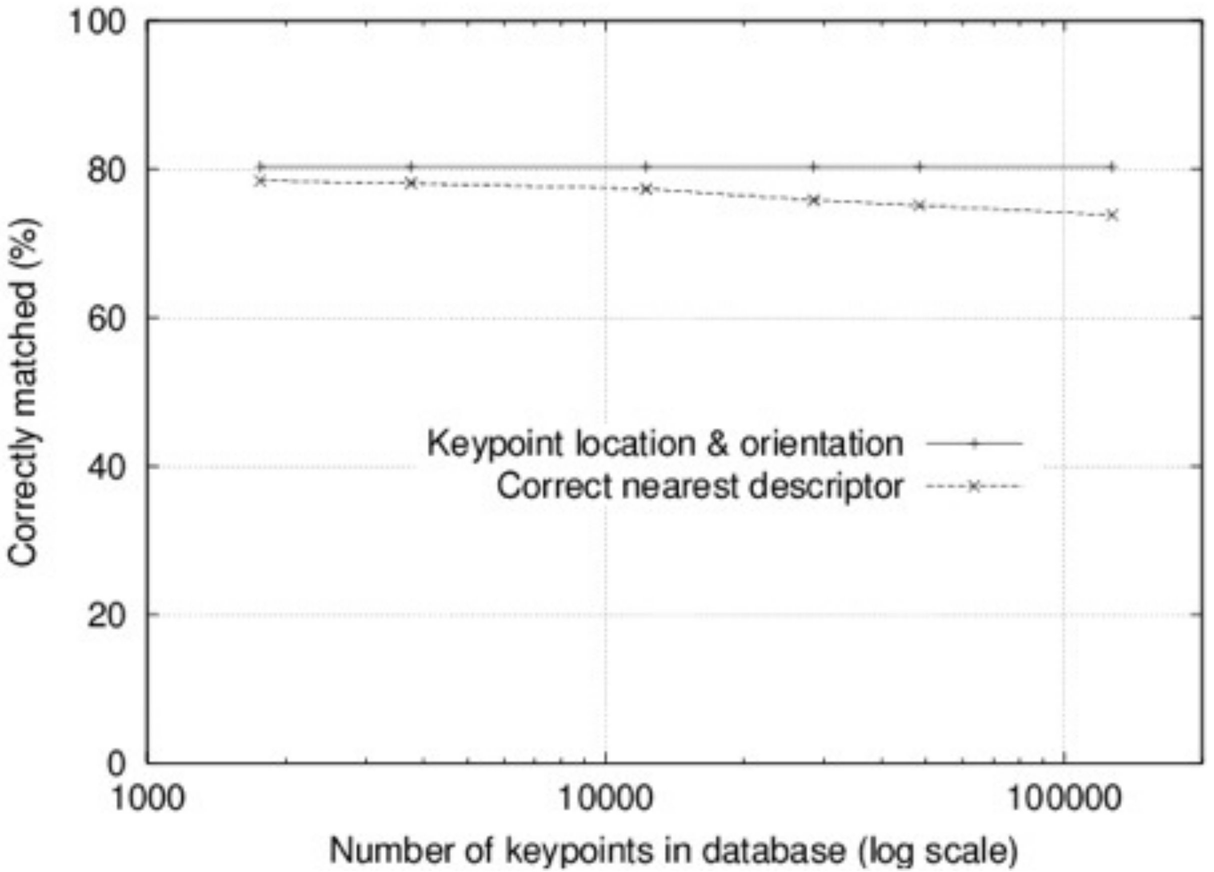
$\Sigma_2^{-1} = A_2^T A_2$

Compute rotational invariant descriptor in this normalized frame

J.Matas et.al. “Rotational Invariants for Wide-baseline Stereo”. Research Report of CMP, 2003

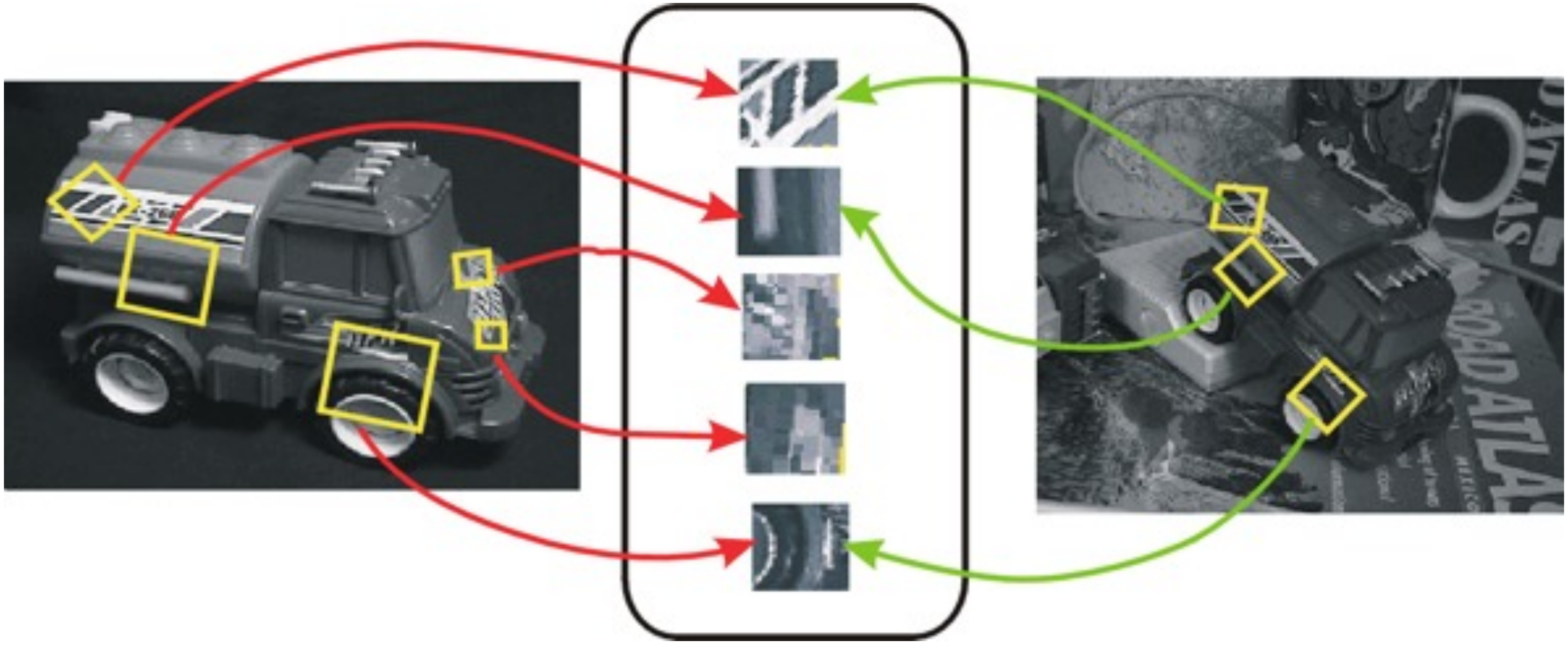
Distinctiveness of features

- Vary size of database of features, with 30 degree affine change, 2% image noise
- Measure % correct for single nearest neighbor match



Application of invariant local features to object (instance) recognition.

Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters



SIFT Features

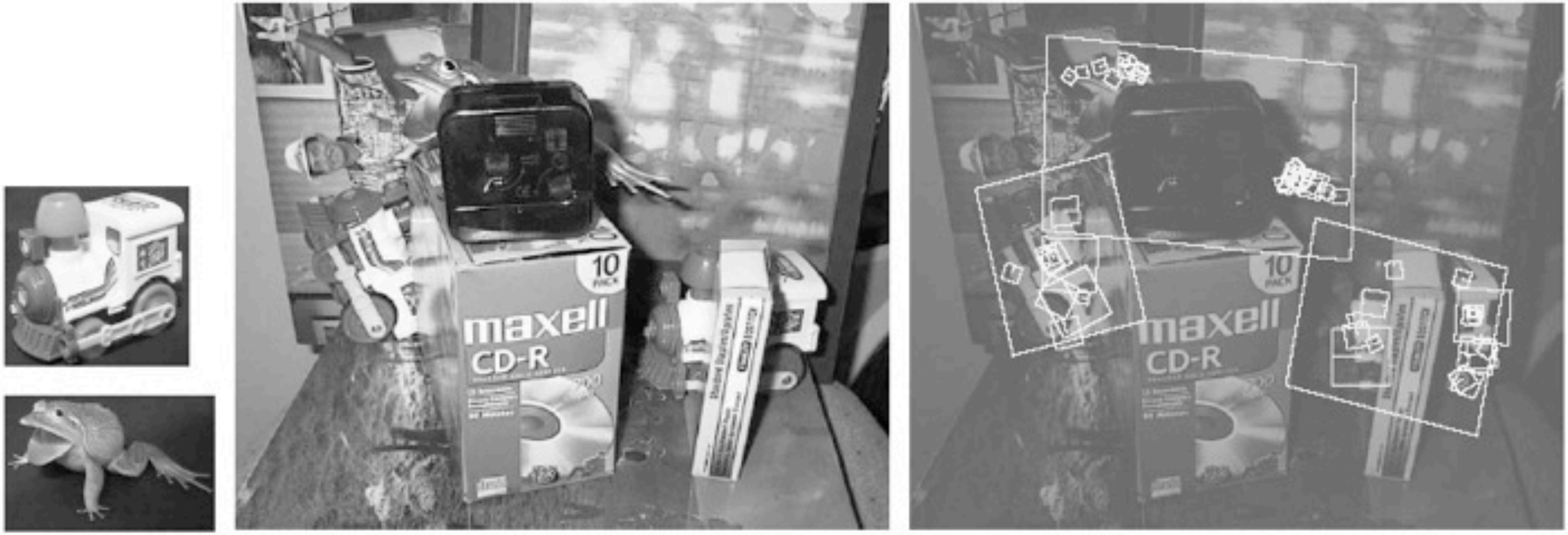


Figure 12: The training images for two objects are shown on the left. These can be recognized in a cluttered image with extensive occlusion, shown in the middle. The results of recognition are shown on the right. A parallelogram is drawn around each recognized object showing the boundaries of the original training image under the affine transformation solved for during recognition. Smaller squares indicate the keypoints that were used for recognition.

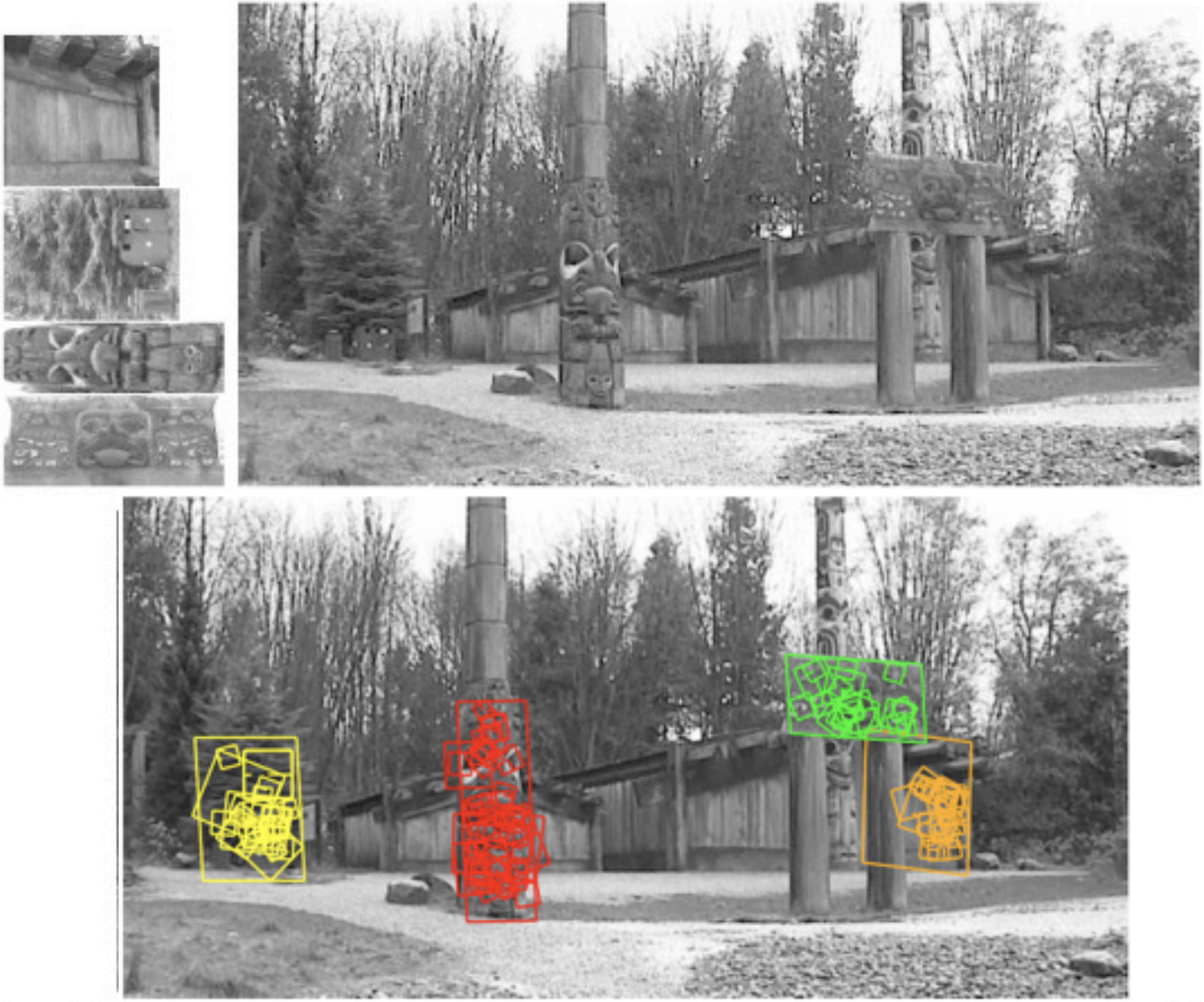


Figure 13: This example shows location recognition within a complex scene. The training images for locations are shown at the upper left and the 640x315 pixel test image taken from a different viewpoint is on the upper right. The recognized regions are shown on the lower image, with keypoints shown as squares and an outer parallelogram showing the boundaries of the training images under the affine transform used for recognition.

SIFT features impact

SIFT feature paper citations:

Distinctive image features from scale-invariant keypoints

DG Lowe - International journal of computer vision, 2004 - Springer
International Journal of Computer Vision 60(2), 91–110, 2004 cc 2004 Kluwer Academic Publishers. Computer Science Department, University of British Columbia ...

Cited by 6537

A good SIFT features tutorial:

http://www.cs.toronto.edu/~jepson/csc2503/tutSIFT04.pdf

By Estrada, Jepson, and Fleet.

The original SIFT paper:
http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf

Now we have

- Well-localized feature points
- Distinctive descriptor

- Now we need to
  - match pairs of feature points in different images
  - Robustly compute homographies
    (in the presence of errors/outliers)