Instance-level recognition:

Local invariant features

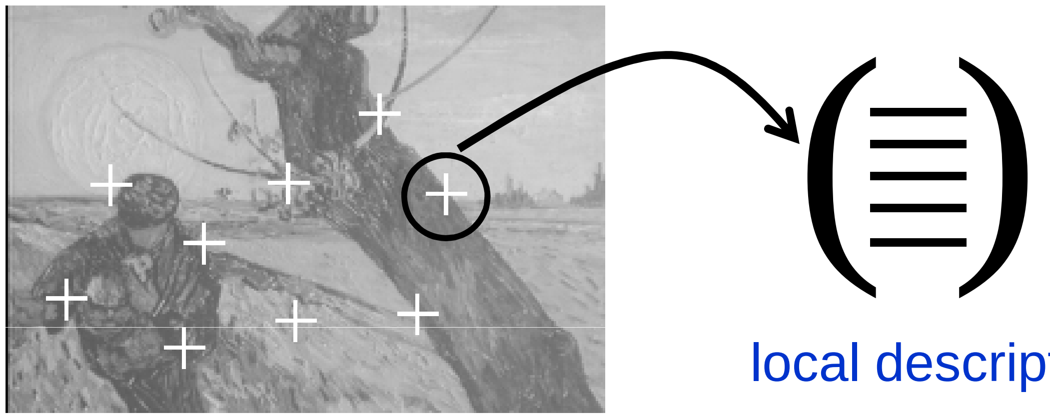
Cordelia Schmid

INRIA, Grenoble

Overview

- **Introduction to local features**
- Harris interest points + SSD, ZNCC, SIFT
- Scale & affine invariant interest point detectors
- Evaluation and comparison of different detectors
- Region descriptors and their performance

Local features



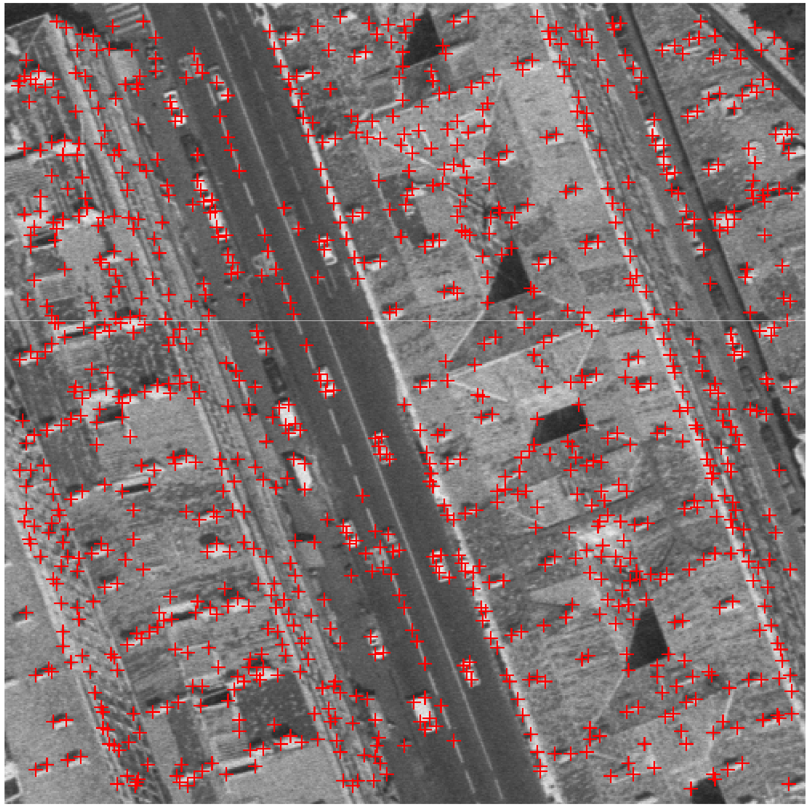
Several / many local descriptors per image

Robust to occlusion/clutter + no object segmentation required

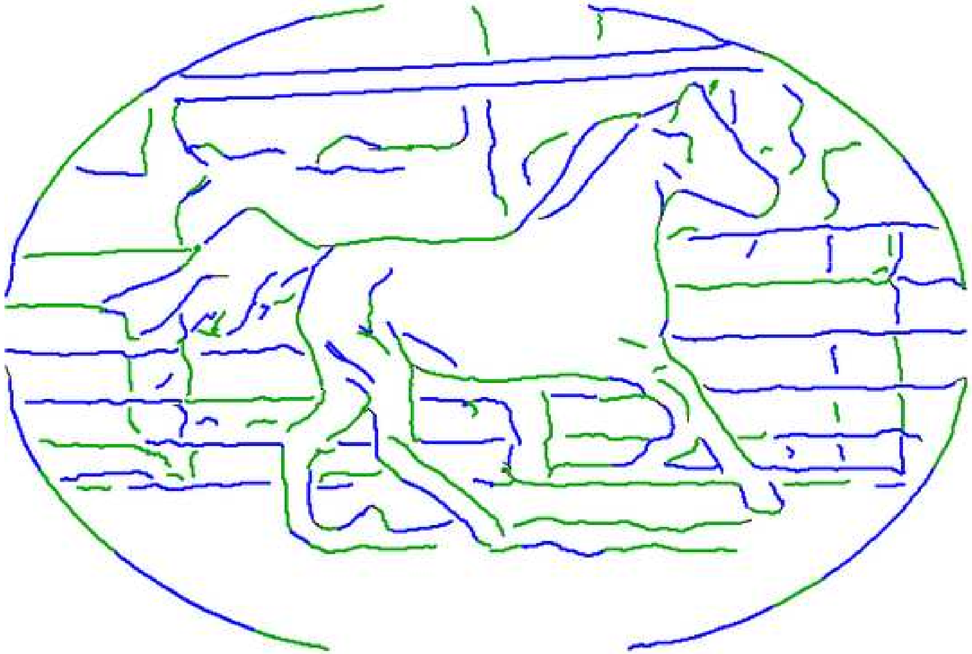
Photometric : distinctive

Invariant : to image transformations + illumination changes

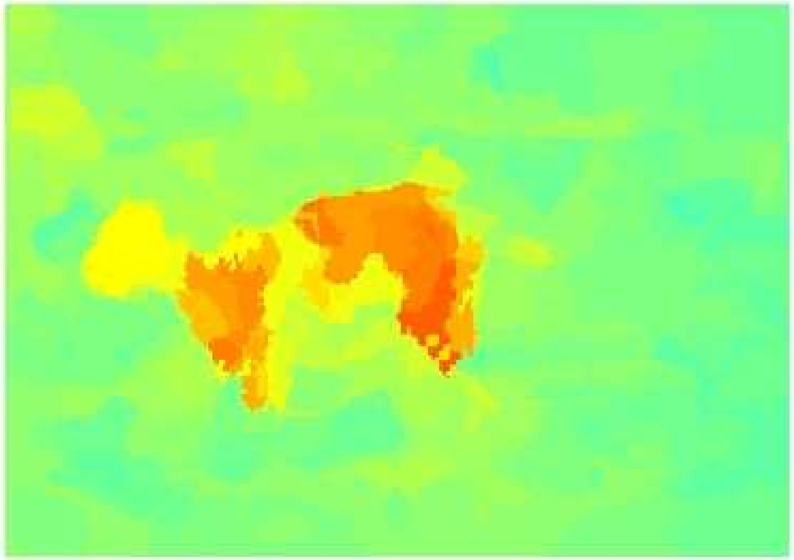
Local features: interest points



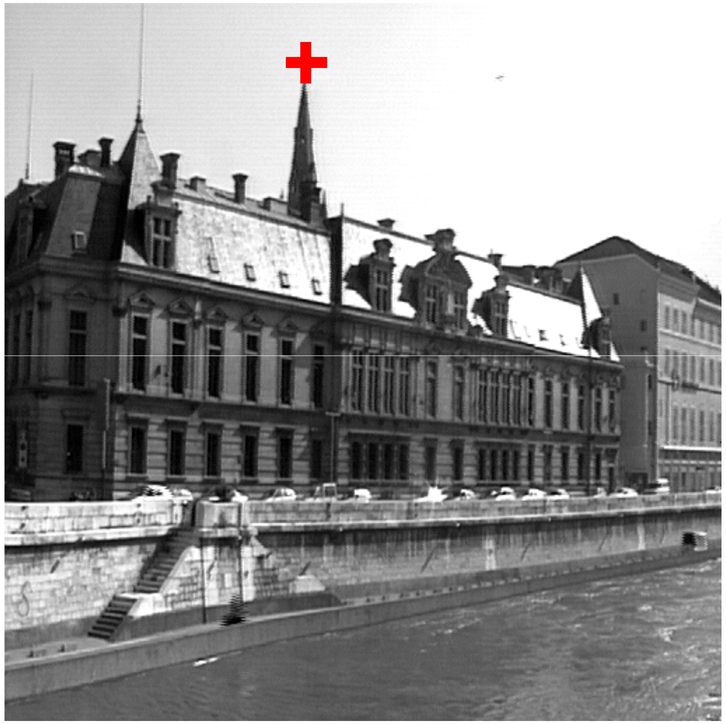
Local features: Contours/segments



Local features: segmentation

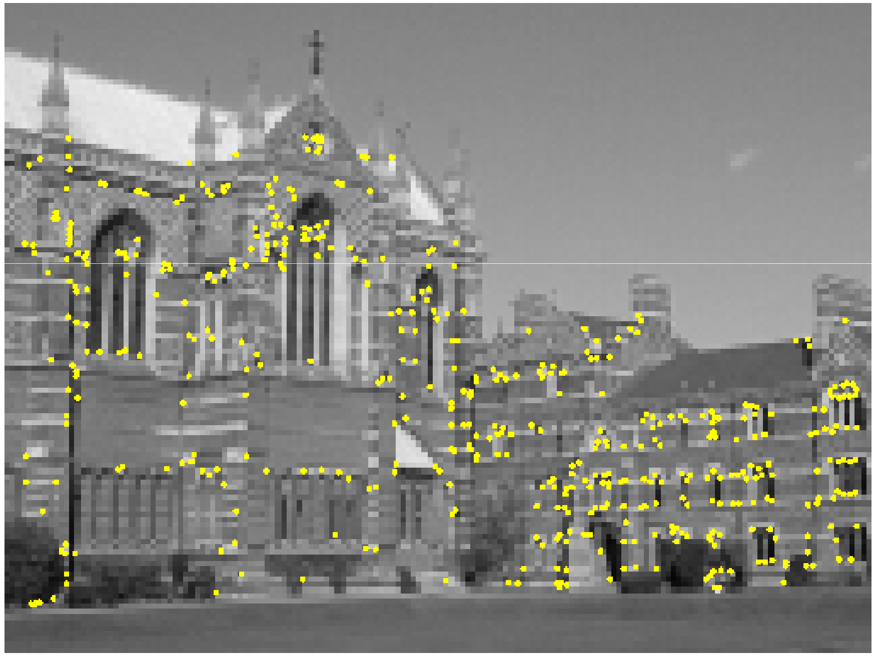


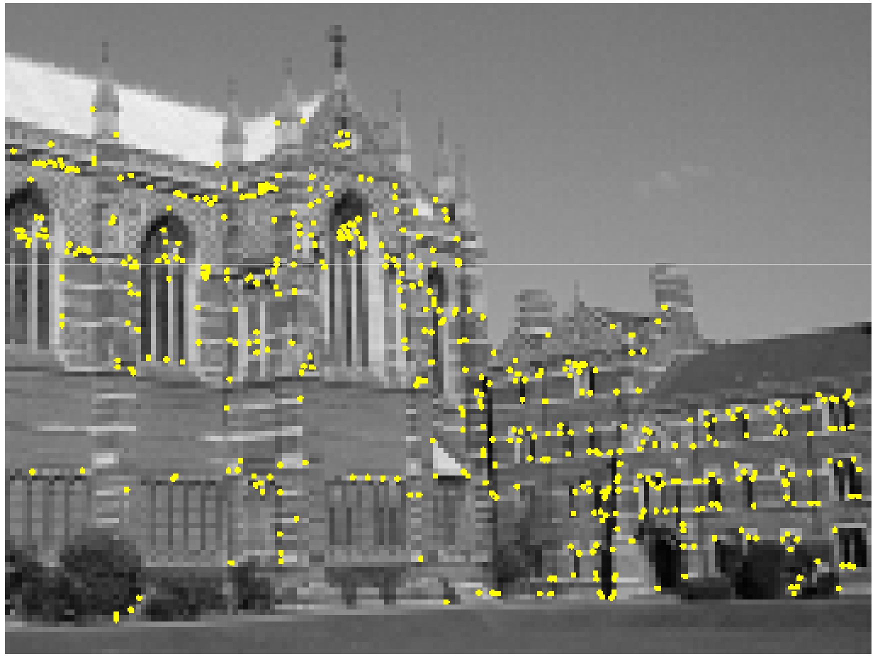
Application: Matching



Find corresponding locations in the image

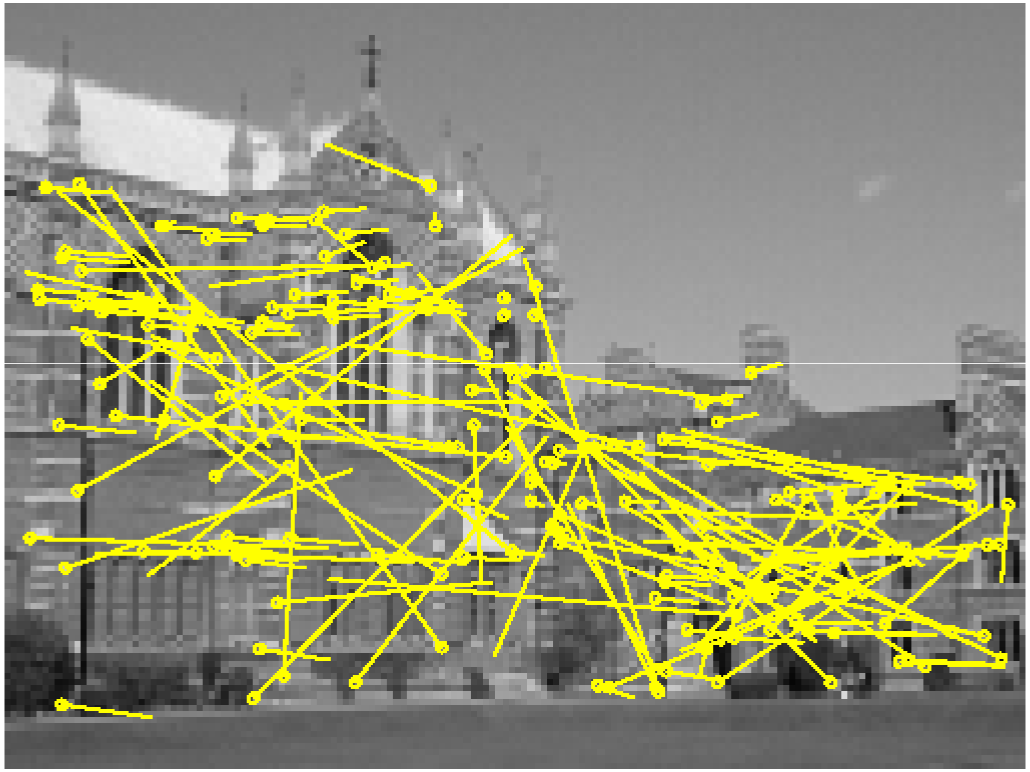
Illustration – Matching





Interest points extracted with Harris detector (~ 500 points)

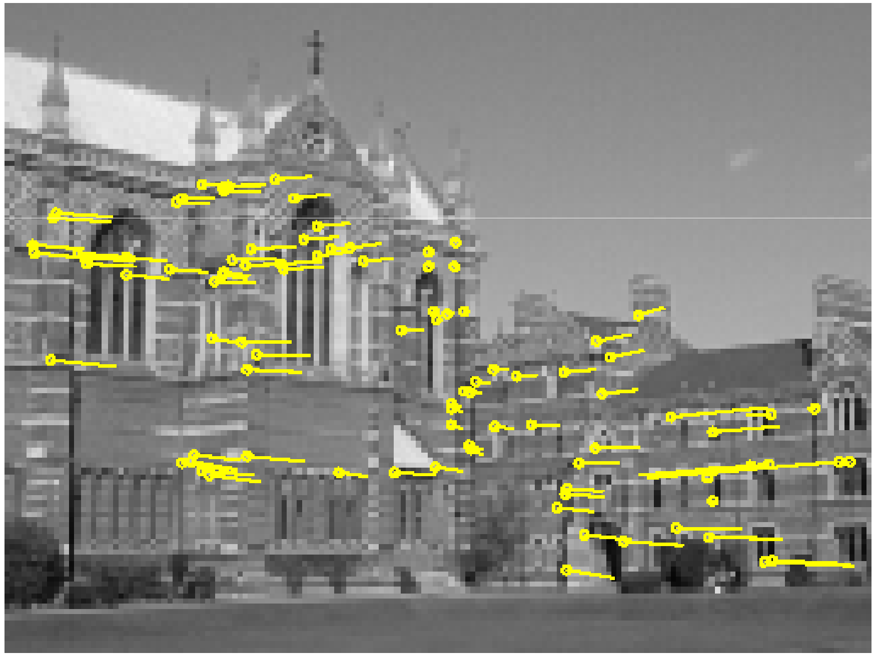
Illustration – Matching



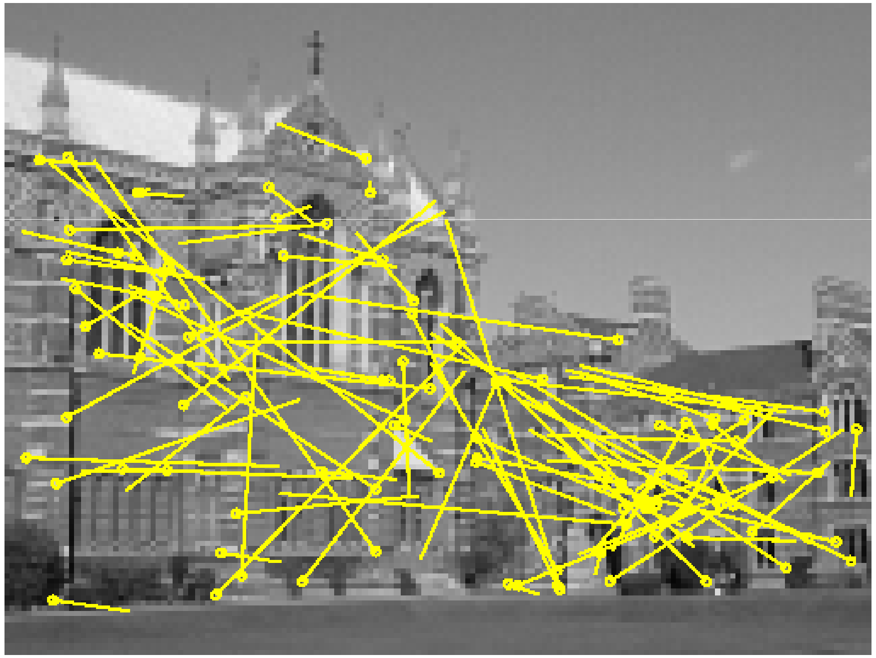
Interest points matched based on cross-correlation (188 pairs)

Illustration – Matching

Global constraint - Robust estimation of the fundamental matrix

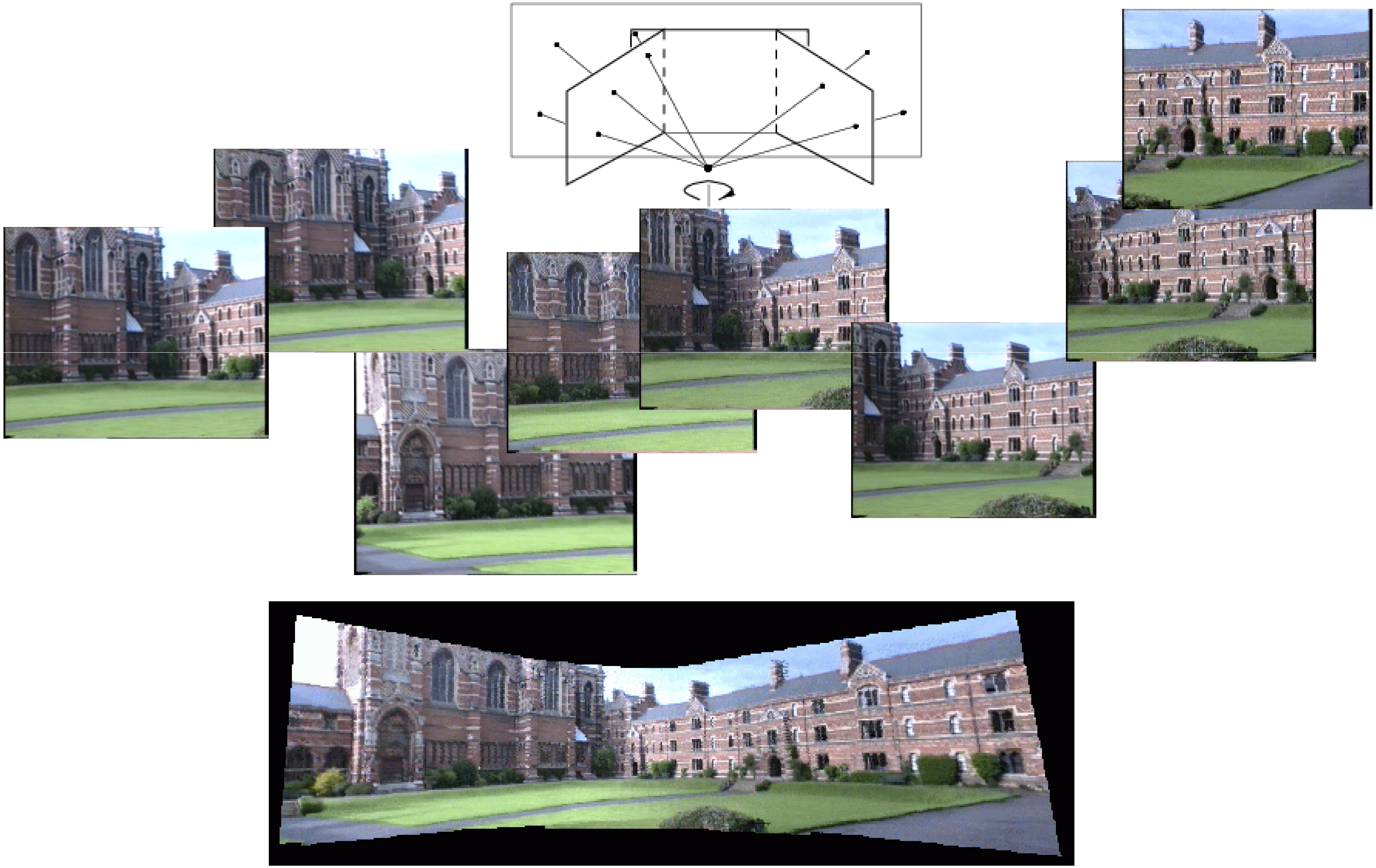


99 inliers



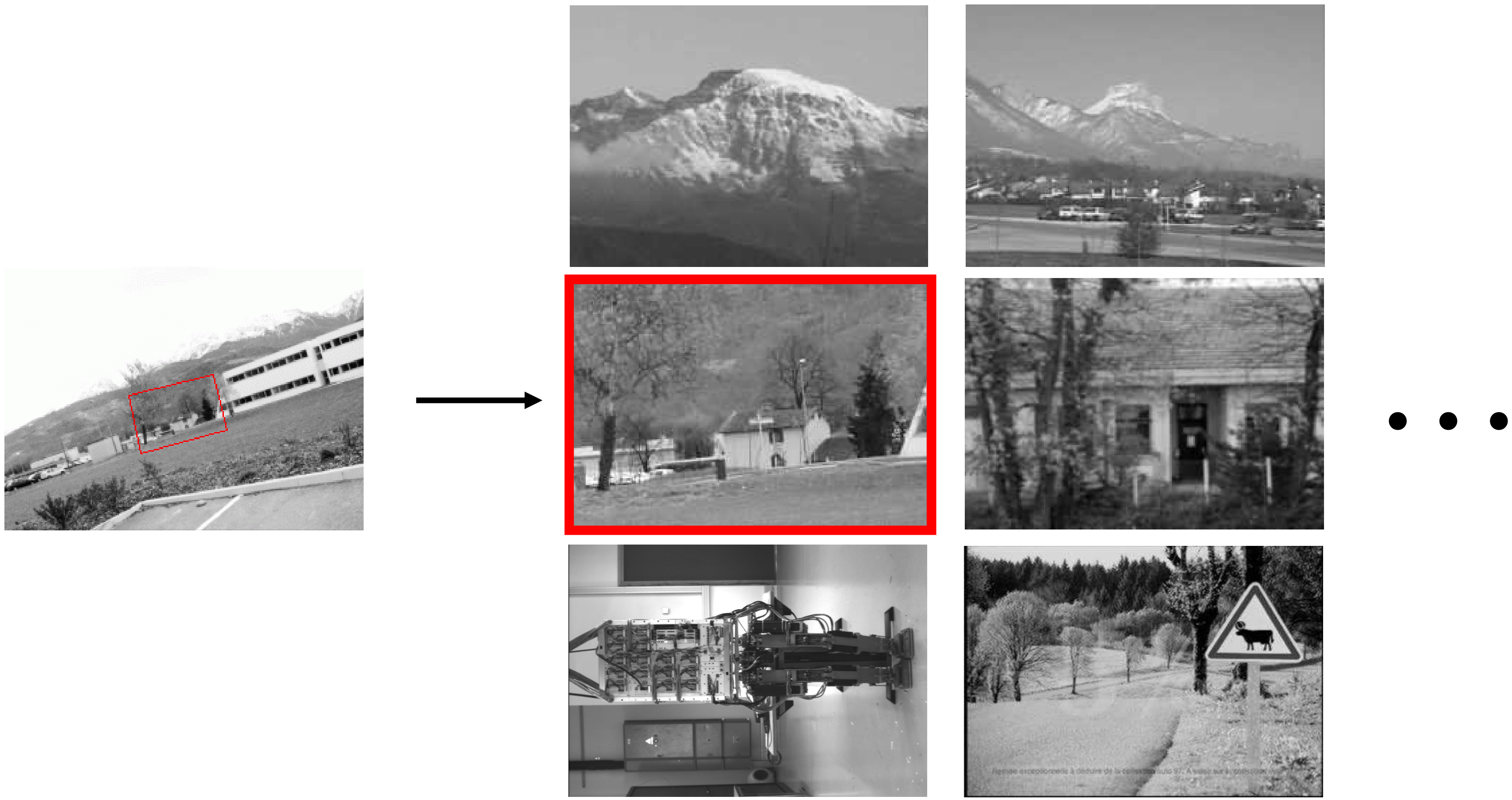
89 outliers

Application: Panorama stitching



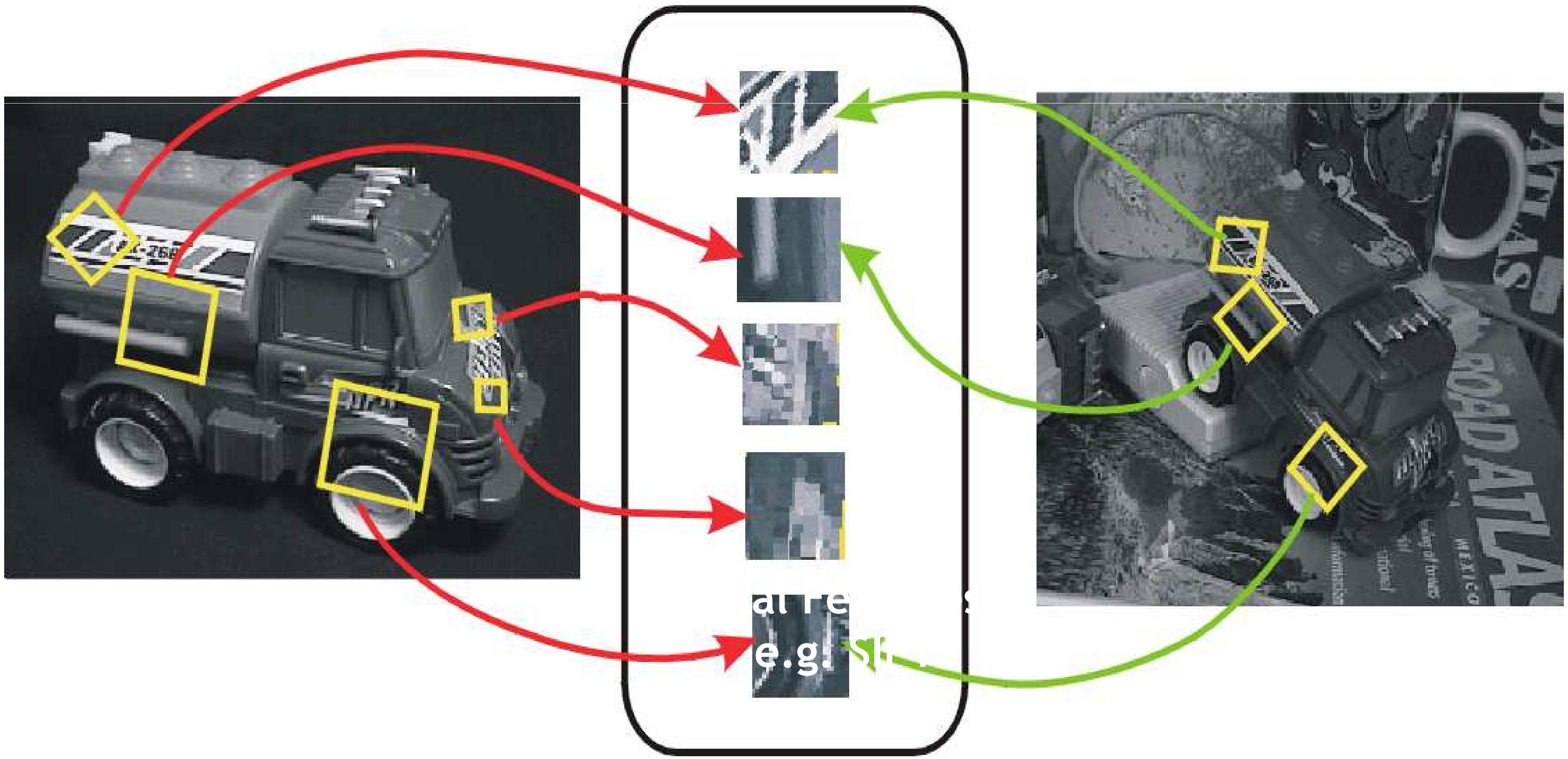
Application: Instance-level recognition

Search for particular objects and scenes in large databases



Instance-level recognition: Approach

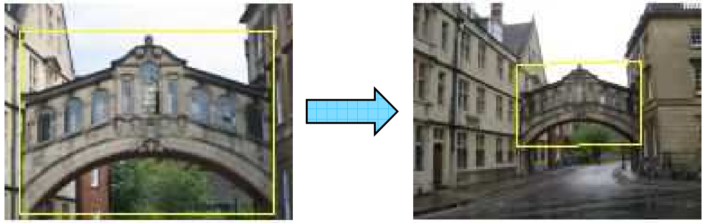
- Image content is transformed into local features invariant to geometric and photometric transformations
- Matching local invariant descriptors



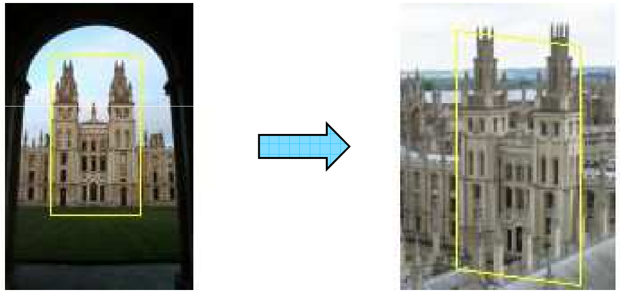
Difficulties

Finding the object despite possibly large changes in scale, viewpoint, lighting and partial occlusion

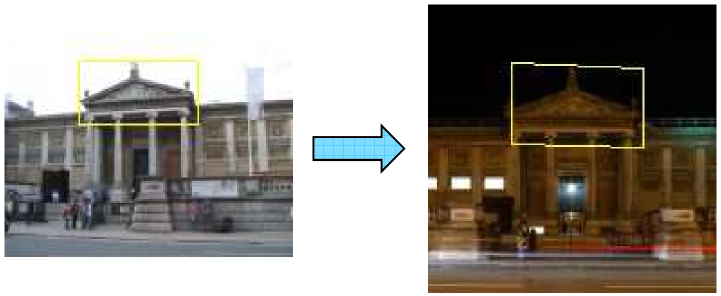
→ **requires invariant description**



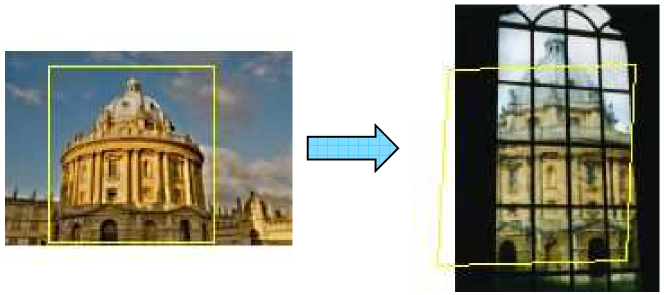
Scale



Viewpoint



Lighting



Occlusion

Difficulties

- Very large images collection → need for efficient indexing
  - Flickr has 2 billion photographs, more than 1 million added daily
  - Facebook has 15 billion images (~27 million added daily)
  - Large personal collections
  - Video collections, i.e., YouTube

Applications

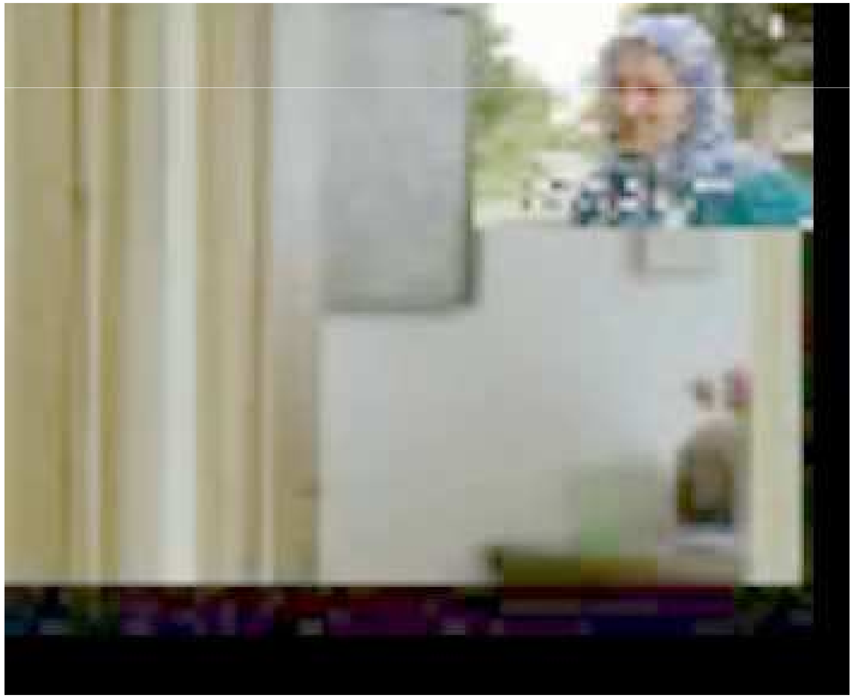
- Take a picture of a product or advertisement
  → find relevant information on the web

[Pixee – Milpix]

Applications

- Copy detection for images and videos

Query video



Search in 200h of video

Applications

- Sony Aibo – Robotics
  - Recognize docking station
  - Place recognition
  - Loop closure in SLAM

Local features - history

- Line segments [Lowe'87, Ayache'90]
- Interest points & cross correlation [Z. Zhang et al. 95]
- Rotation invariance with differential invariants [Schmid&Mohr'96]
- Scale & affine invariant detectors [Lindeberg'98, Lowe'99, Tuytelaars&VanGool'00, Mikolajczyk&Schmid'02, Matas et al.'02]
- Dense detectors and descriptors [Leung&Malik'99, Fei-Fei& Perona'05, Lazebnik et al.'06]
- Contour and region (segmentation) descriptors [Shotton et al.'05, Opelt et al.'06, Ferrari et al.'06, Leordeanu et al.'07]

Local features

1) Extraction of local features

- Contours/segments
- Interest points & regions
- Regions by segmentation
- Dense features, points on a regular grid

2) Description of local features

- Dependant on the feature type
- Contours/segments → angles, length ratios
- Interest points → greylevels, gradient histograms
- Regions (segmentation) → texture + color distributions

Line matching

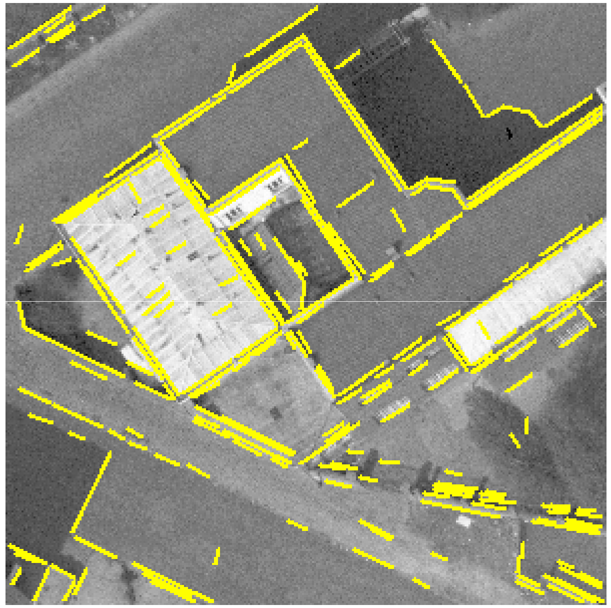
- Extraction de contours
  - Zero crossing of Laplacian
  - Local maxima of gradients
- Chain contour points (hysteresis)
- Extraction of line segments
- Description of segments
  - Mi-point, length, orientation, angle between pairs etc.

Experimental results – line segments

images 600 x 600

Experimental results – line segments

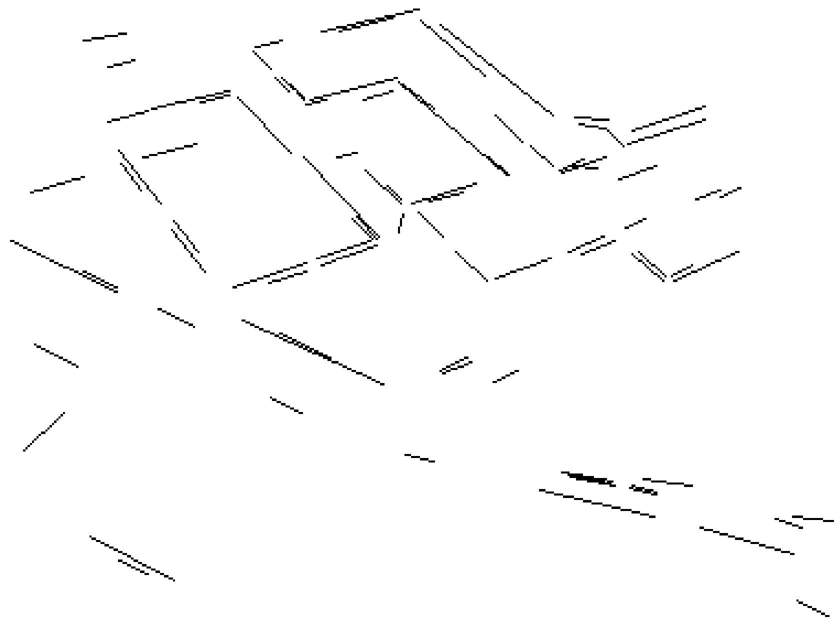


248 / 212 line segments extracted

Experimental results – line segments

89 matched line segments - 100% correct

Experimental results – line segments



3D reconstruction

Problems of line segments

- Often only partial extraction
  - Line segments broken into parts
  - Missing parts
- Information not very discriminative
  - 1D information
  - Similar for many segments
- Potential solutions
  - Pairs and triplets of segments
  - Interest points

Overview

- Introduction to local features
- **Harris interest points + SSD, ZNCC, SIFT**
- Scale & affine invariant interest point detectors
- Evaluation and comparison of different detectors
- Region descriptors and their performance

Harris detector [Harris & Stephens'88]

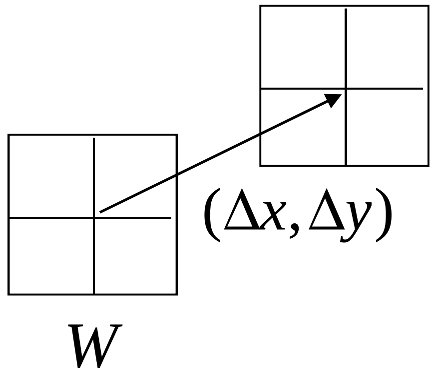
Based on the idea of auto-correlation

Important difference in all directions => interest point

Harris detector

Auto-correlation function for a point (x, y) and a shift $(\Delta x, \Delta y)$

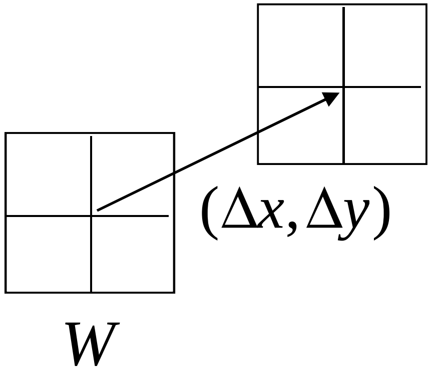
$$A(x, y) = \sum_{(x_k, y_k) \in W(x, y)} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2$$



Harris detector

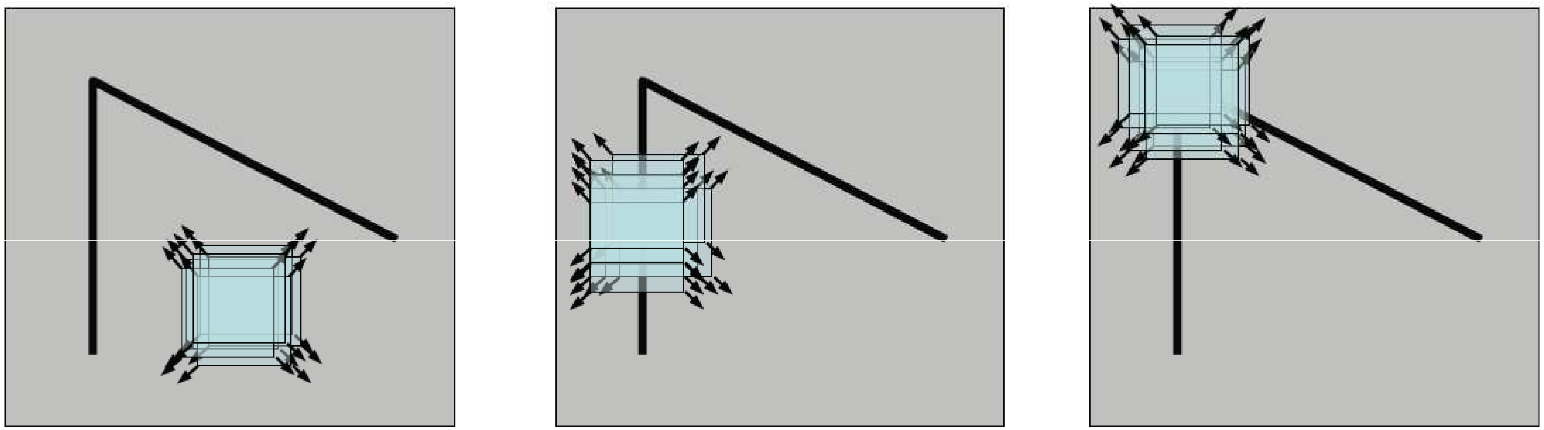
Auto-correlation function for a point (x, y) and a shift $(\Delta x, \Delta y)$

$$A(x, y) = \sum_{(x_k, y_k) \in W(x, y)} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2$$



$A(x, y)$
- small in all directions → uniform region
- large in one directions → contour
- large in all directions → interest point

Harris detector



"flat" region: no change in all directions

"edge": no change along the edge direction

"corner": significant change in all directions

Harris detector

Discret shifts are avoided based on the auto-correlation matrix

with first order approximation

$$I(x_k + \Delta x, y_k + \Delta y) = I(x_k, y_k) + (I_x(x_k, y_k) \quad I_y(x_k, y_k))\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}$$

$$A(x, y) = \sum_{(x_k, y_k) \in W(x, y)} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2$$

$$= \sum_{(x_k, y_k) \in W} \left(\left(I_x(x_k, y_k) \quad I_y(x_k, y_k)\right) \begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} \right)^2$$

Harris detector

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

Auto-correlation matrix

the sum can be smoothed with a Gaussian

$$= \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} G \otimes \begin{bmatrix} I_x{}^2 & I_x I_y \\ I_x I_y & I_y{}^2 \end{bmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

Harris detector

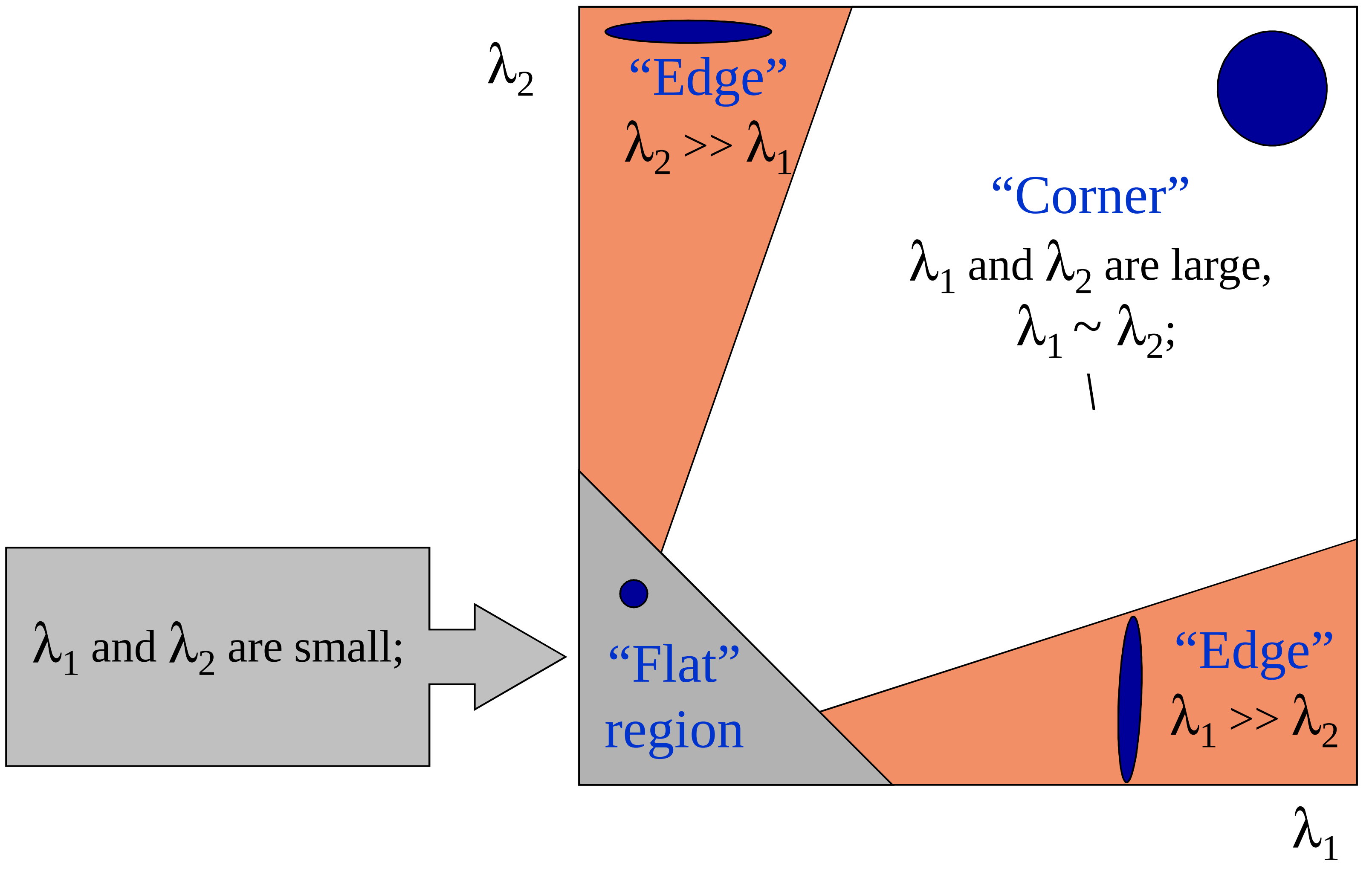
- Auto-correlation matrix

$$A(x,y) = G \otimes \begin{bmatrix} I_x^{\ 2} & I_x I_y \\ I_x I_y & I_y^{\ 2} \end{bmatrix}$$

  - captures the structure of the local neighborhood
  - measure based on eigenvalues of this matrix
    - 2 strong eigenvalues => interest point
    - 1 strong eigenvalue => contour
    - 0 eigenvalue => uniform region

Interpreting the eigenvalues

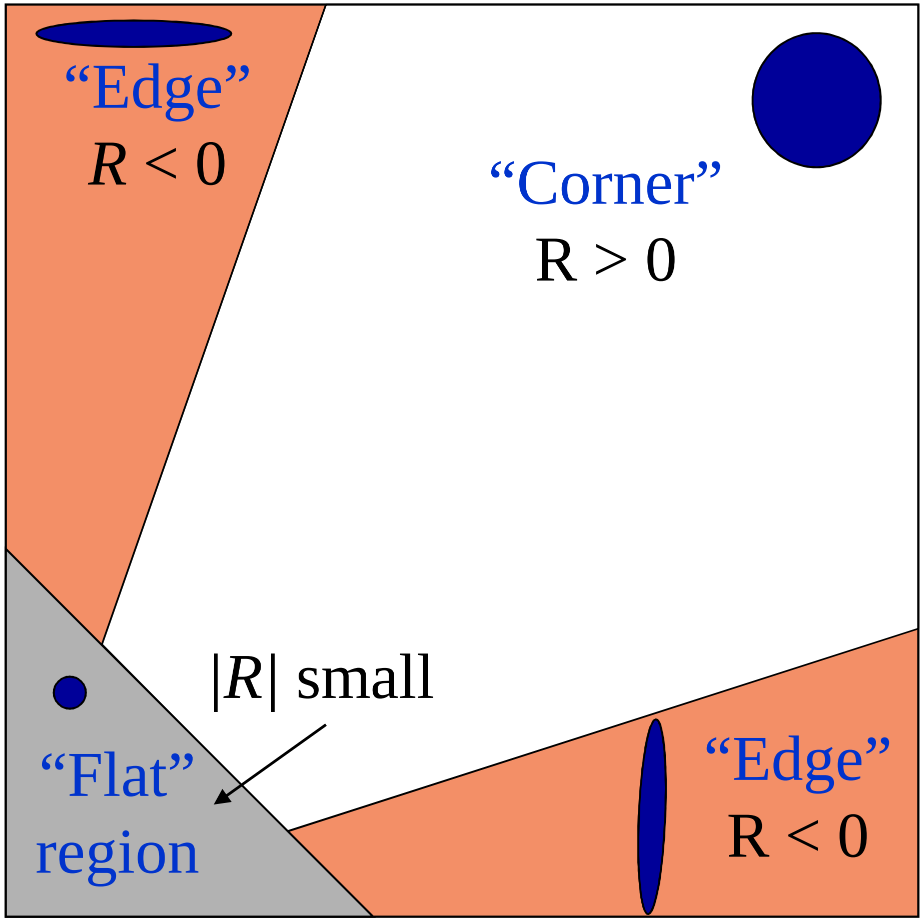
Classification of image points using eigenvalues of autocorrelation matrix:



Corner response function

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

α: constant (0.04 to 0.06)



Harris detector

- Cornerness function

$$f = \det(A) - k(trace(A))^2 = \lambda_1\lambda_2 - k(\lambda_1 + \lambda_2)^2$$

Reduces the effect of a strong contour

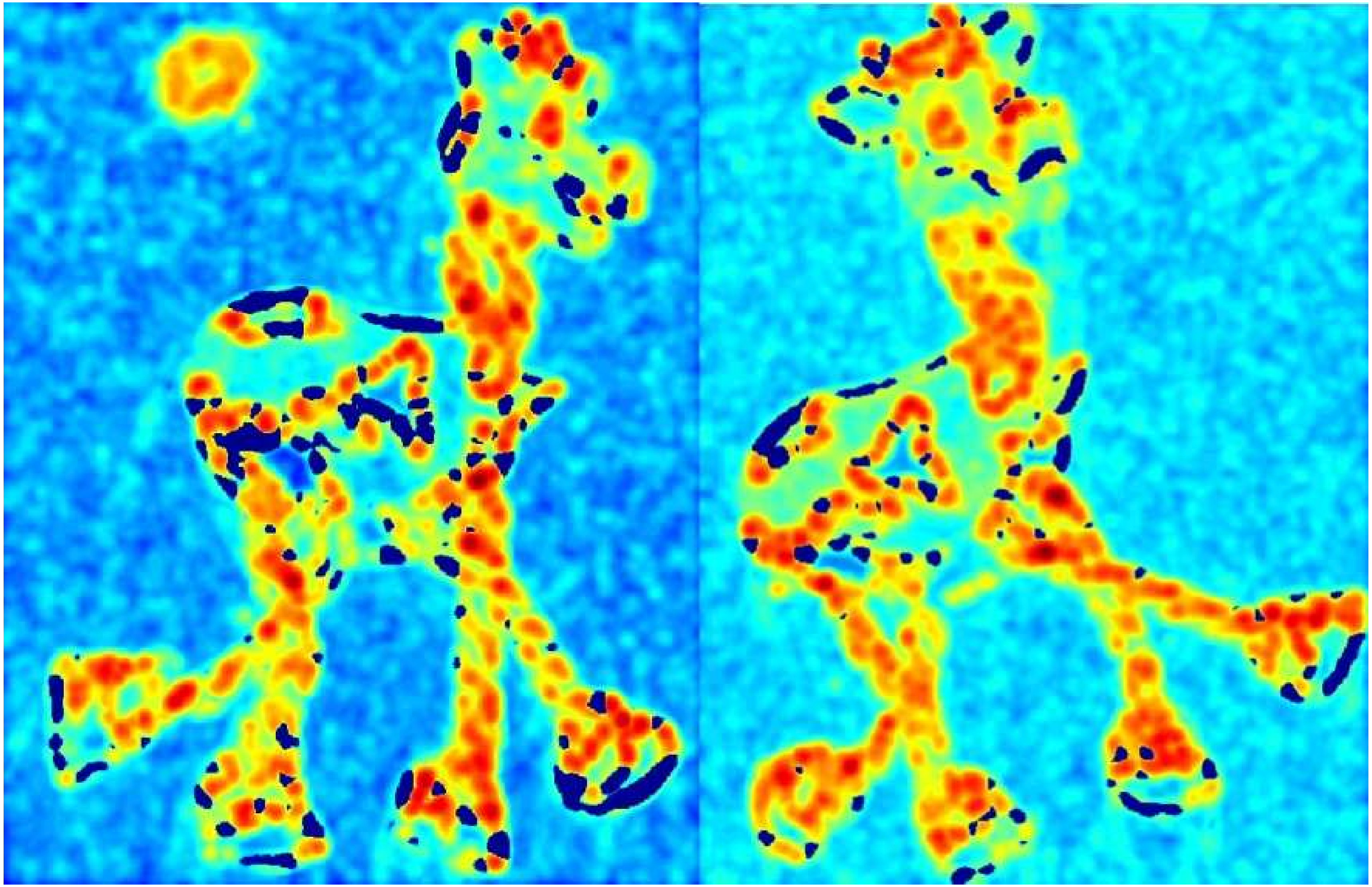
- Interest point detection
  - Treshold (absolut, relatif, number of corners)
  - Local maxima

$$f > thresh \ \wedge \ \forall x, y \in 8-neighbourhood \quad f(x,y) \geq f(x',y')$$

Harris Detector: Steps

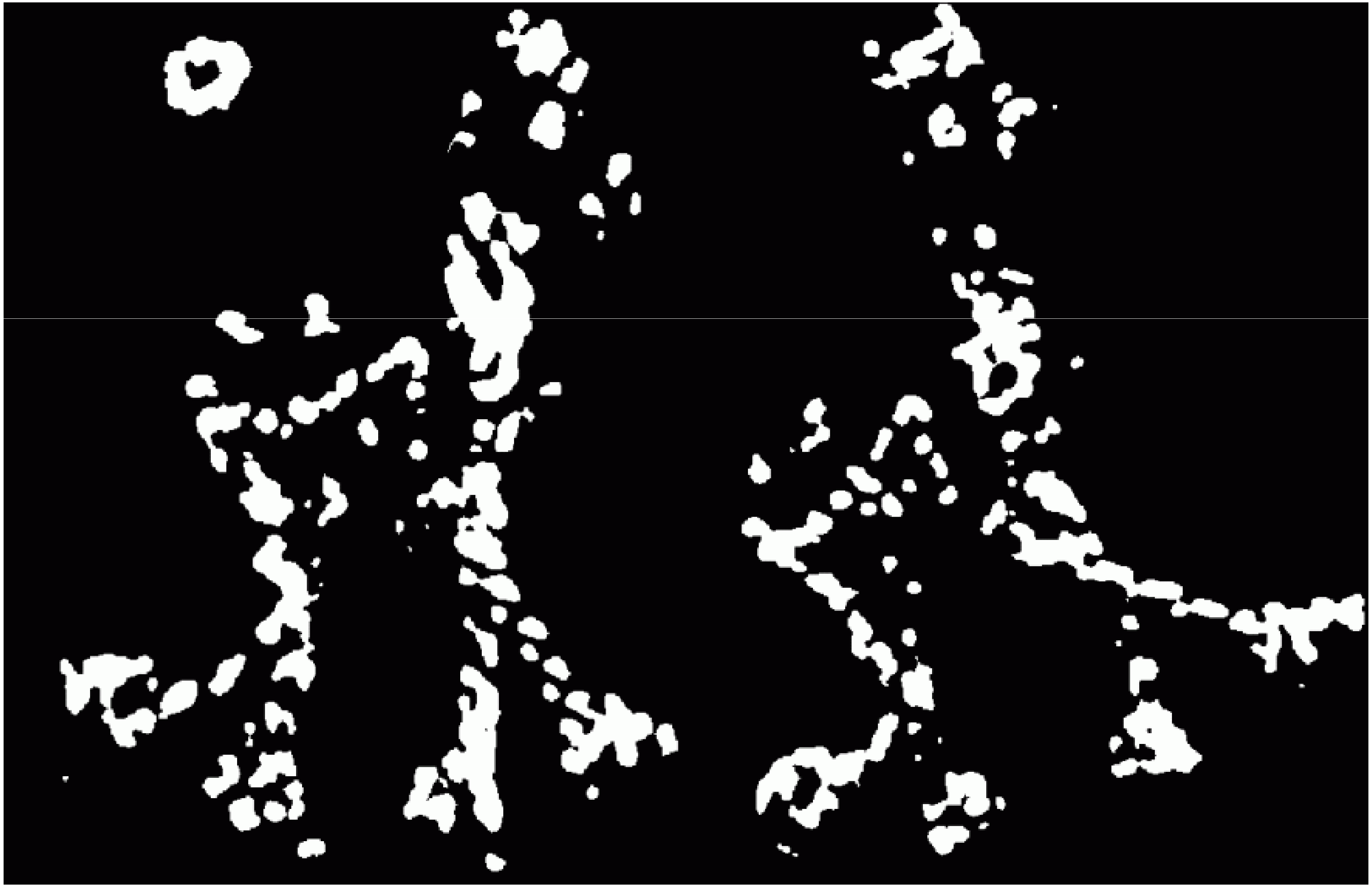
Harris Detector: Steps

Compute corner response R



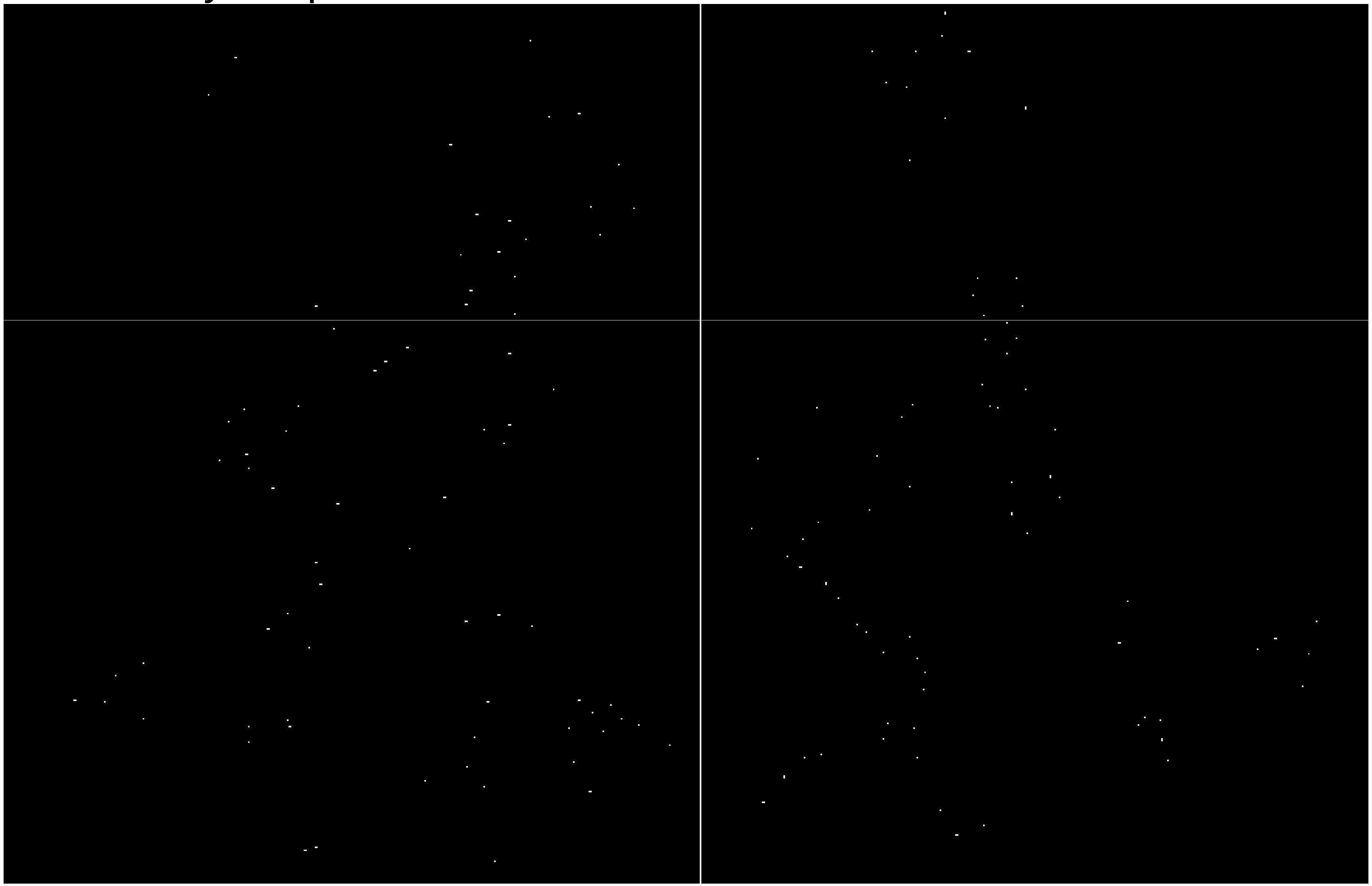
Harris Detector: Steps

Find points with large corner response: *R*>threshold



Harris Detector: Steps

Take only the points of local maxima of *R*



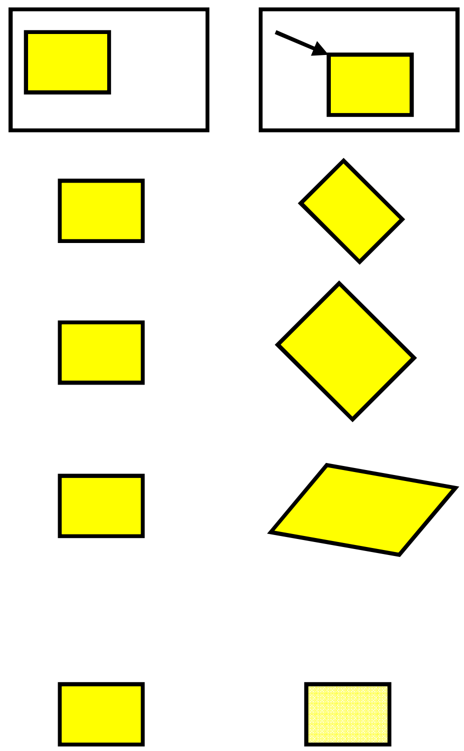
Harris Detector: Steps

Harris detector: Summary of steps

1. Compute Gaussian derivatives at each pixel
2. Compute second moment matrix A in a Gaussian window around each pixel
3. Compute corner response function R
4. Threshold R
5. Find local maxima of response function (non-maximum suppression)

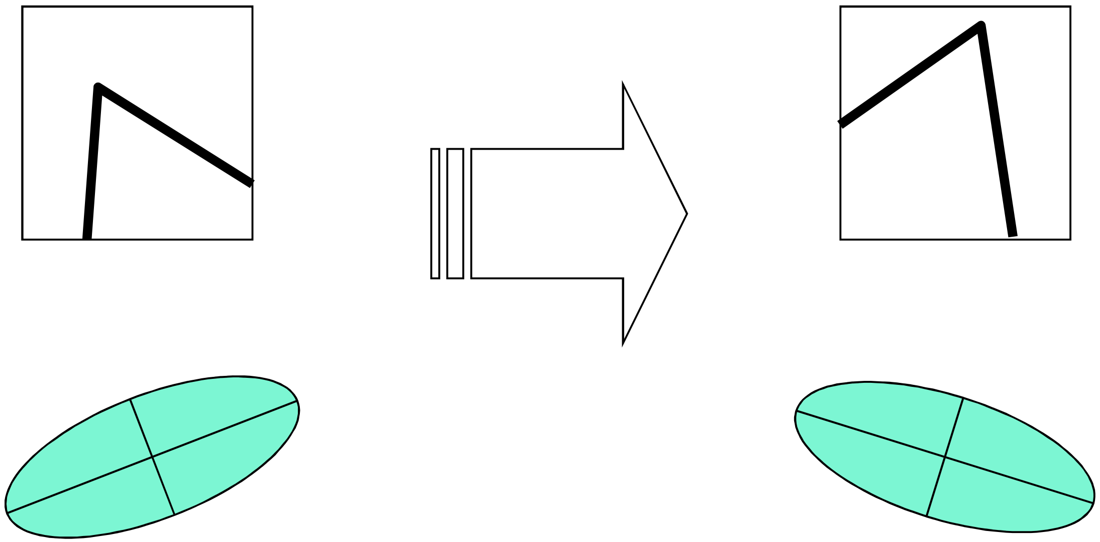
Harris - invariance to transformations

- Geometric transformations
  – translation
  – rotation
  – similitude (rotation + scale change)
  – affine (valide for local planar objects)
- Photometric transformations
  – Affine intensity changes (I $\rightarrow$ a I + b)



Harris Detector: Invariance Properties

- Rotation



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

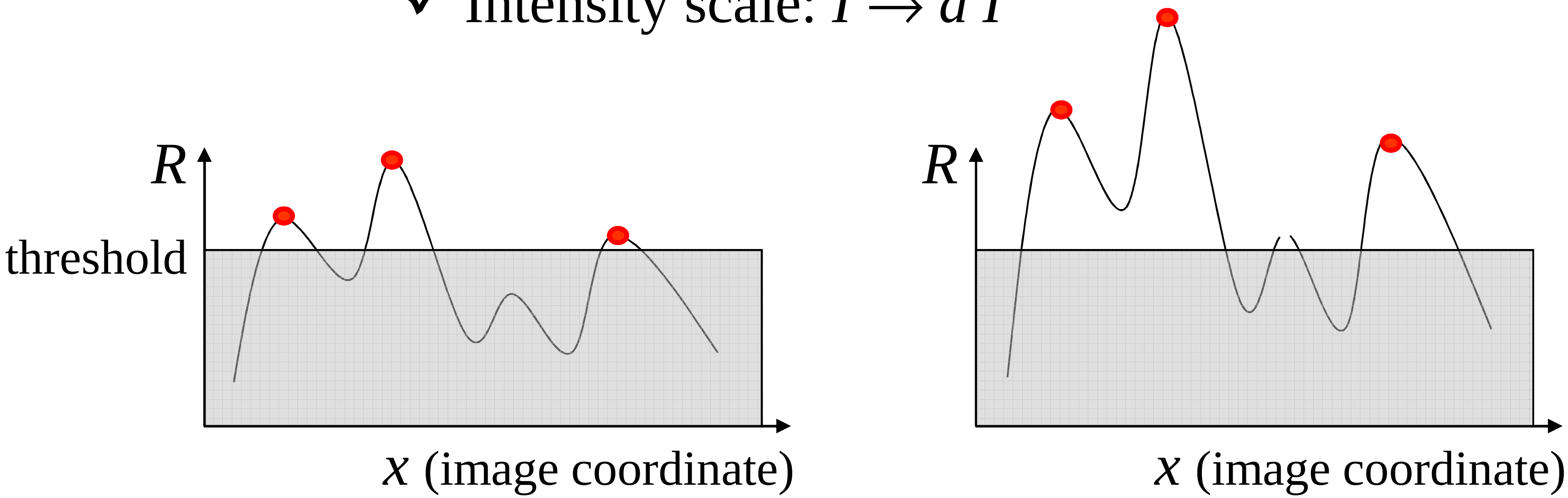
Corner response R is invariant to image rotation

Harris Detector: Invariance Properties

- Affine intensity change

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

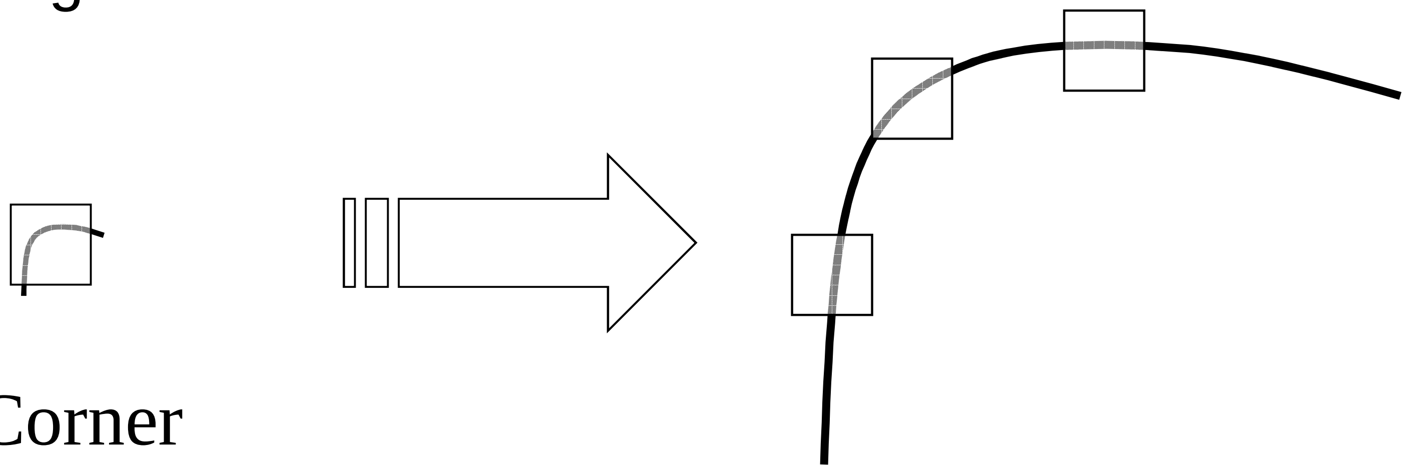
  ✓ Intensity scale: $I \rightarrow a\, I$



Partially invariant to affine intensity change, dependent on type of threshold

Harris Detector: Invariance Properties

- Scaling



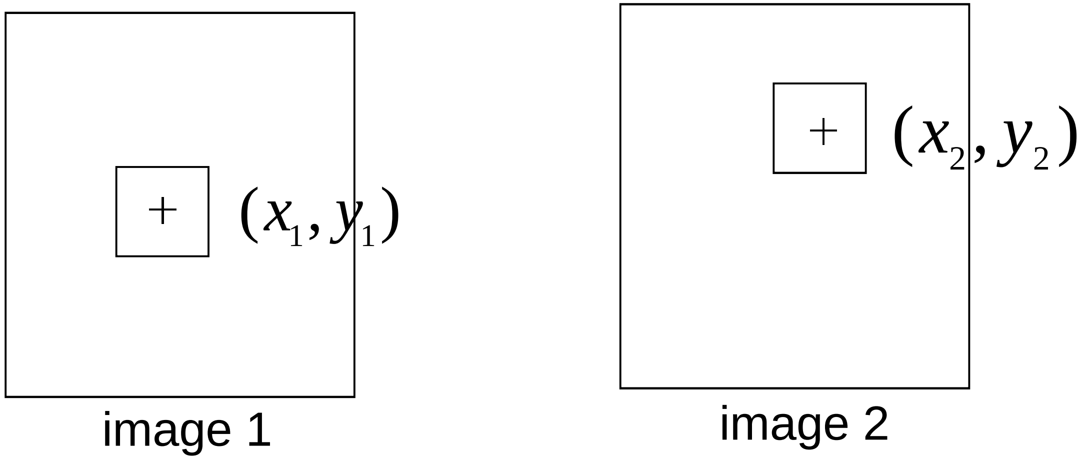
Corner

All points will be classified as edges

Not invariant to scaling

Comparison of patches - SSD

Comparison of the intensities in the neighborhood of two interest points



SSD : sum of square difference

$$\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\sum_{j=-N}^{N}\left(I_1(x_1+i, y_1+j)-I_2(x_2+i, y_2+j)\right)^2$$

Small difference values → similar patches

Comparison of patches

SSD : $$\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\sum_{j=-N}^{N}(I_1(x_1+i,y_1+j)-I_2(x_2+i,y_2+j))^2$$

Invariance to photometric transformations?

Intensity changes (I $\rightarrow$ I + b)

=> Normalizing with the mean of each patch

$$\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\sum_{j=-N}^{N}((I_1(x_1+i,y_1+j)-m_1)-(I_2(x_2+i,y_2+j)-m_2))^2$$

Intensity changes (I $\rightarrow$ aI + b)

=> Normalizing with the mean and standard deviation of each patch

$$\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\sum_{j=-N}^{N}\left(\frac{I_1(x_1+i,y_1+j)-m_1}{\sigma_1}-\frac{I_2(x_2+i,y_2+j)-m_2}{\sigma_2}\right)^2$$

Cross-correlation ZNCC

zero normalized SSD

$$\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\sum_{j=-N}^{N}\left(\frac{I_1(x_1+i,y_1+j)-m_1}{\sigma_1}-\frac{I_2(x_2+i,y_2+j)-m_2}{\sigma_2}\right)^2$$

⬇

ZNCC: zero normalized cross correlation

$$\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\sum_{j=-N}^{N}\left(\frac{I_1(x_1+i,y_1+j)-m_1}{\sigma_1}\right)\cdot\left(\frac{I_2(x_2+i,y_2+j)-m_2}{\sigma_2}\right)$$

ZNCC values between -1 and 1, 1 when identical patches

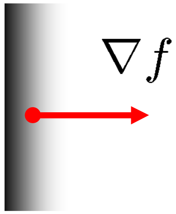
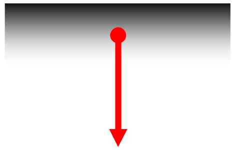
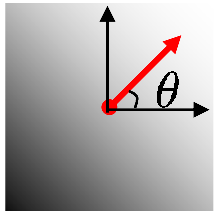
in practice threshold around 0.5

Local descriptors

- Greyvalue derivatives
- Differential invariants [Koenderink'87]
- SIFT descriptor [Lowe'99]

Greyvalue derivatives: Image gradient

- The gradient of an image: $$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]$$

- 
$\nabla f = \left[\frac{\partial f}{\partial x}, 0\right]$

$\nabla f = \left[0, \frac{\partial f}{\partial y}\right]$


$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]$

- The gradient points in the direction of most rapid increase in intensity

- The gradient direction is given by $$\theta = \tan^{-1}\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right)$$
  – how does this relate to the direction of the edge?

- The *edge strength* is given by the gradient magnitude

$$||\nabla f|| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

Differentiation and convolution

- Recall, for 2D function, f(x,y): $$\frac{\partial f}{\partial x} = \lim_{\varepsilon \to 0}\left(\frac{f(x+\varepsilon, y)}{\varepsilon} - \frac{f(x,y)}{\varepsilon}\right)$$

- We could approximate this as $$\frac{\partial f}{\partial x} \approx \frac{f(x_{n+1}, y) - f(x_n, y)}{\Delta x}$$

- Convolution with the filter

-1	**1**

Finite difference filters

- Other approximations of derivative filters exist:

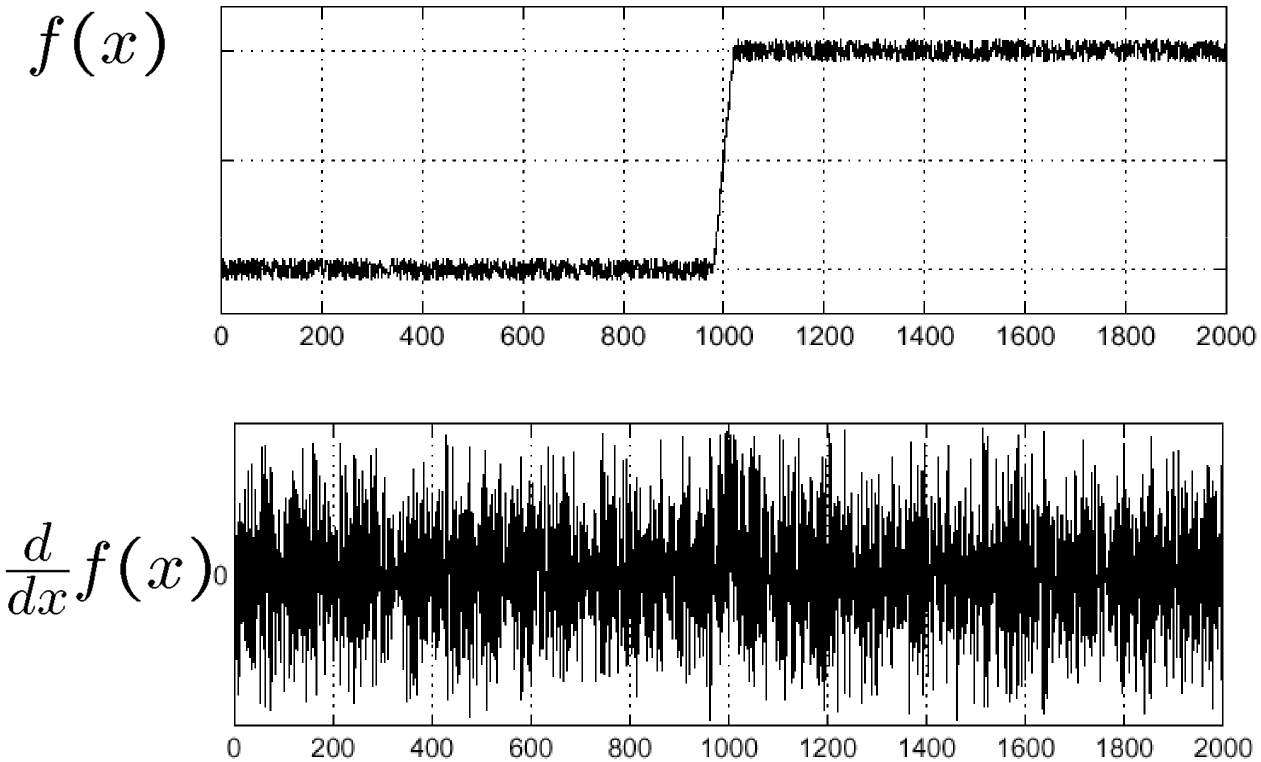
Prewitt: $$M_x = \begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad ; \quad M_y = \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -1 & -1 \\ \hline \end{array}$$

Sobel: $$M_x = \begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad ; \quad M_y = \begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

Roberts: $$M_x = \begin{array}{|c|c|}\hline 0 & 1 \\ \hline -1 & 0 \\ \hline \end{array} \quad ; \quad M_y = \begin{array}{|c|c|}\hline 1 & 0 \\ \hline 0 & -1 \\ \hline \end{array}$$

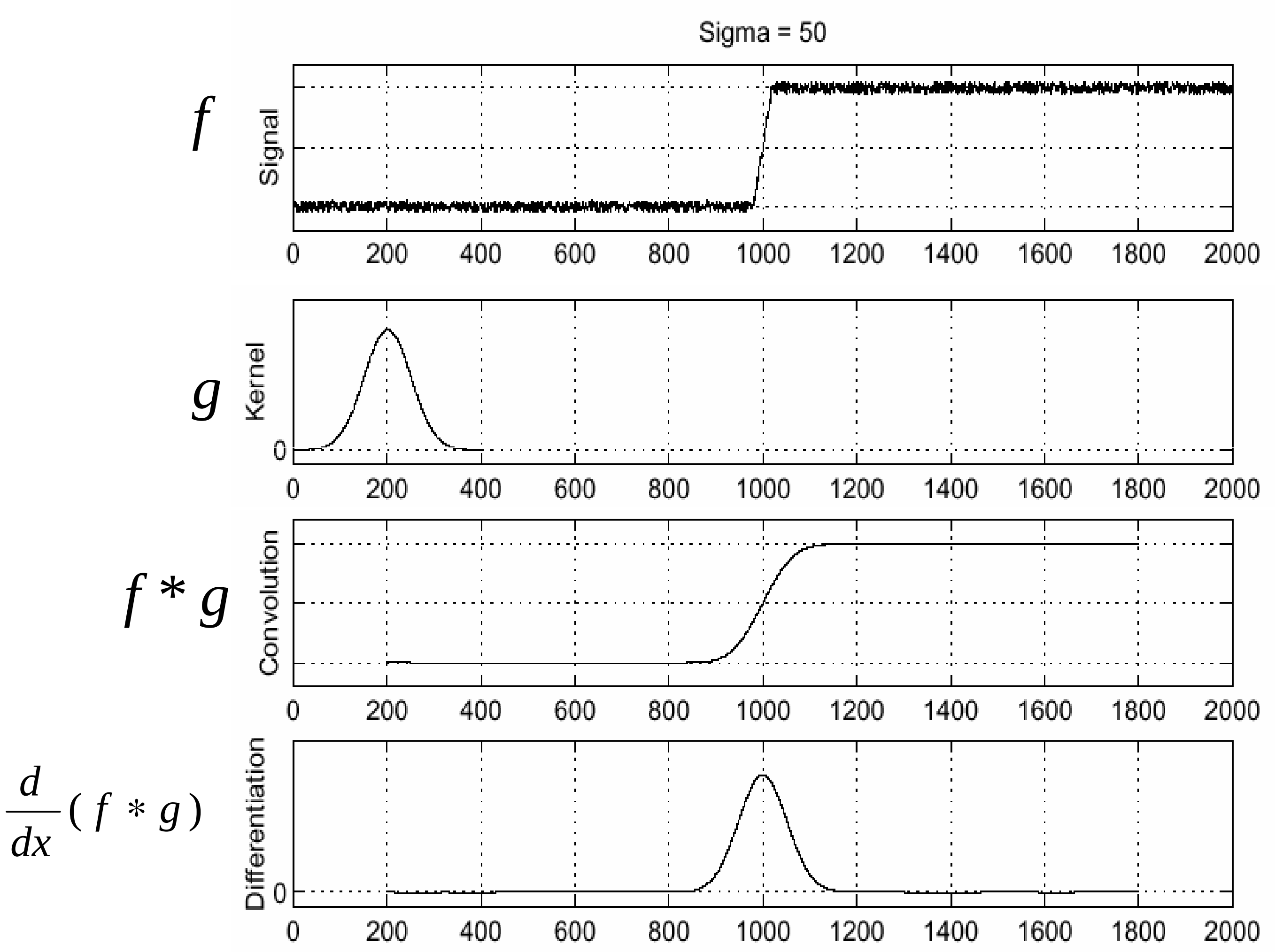
Effects of noise

- Consider a single row or column of the image
  - Plotting intensity as a function of position gives a signal



- Where is the edge?

Solution: smooth first



- To find edges, look for peaks in $\frac{d}{dx}(f * g)$

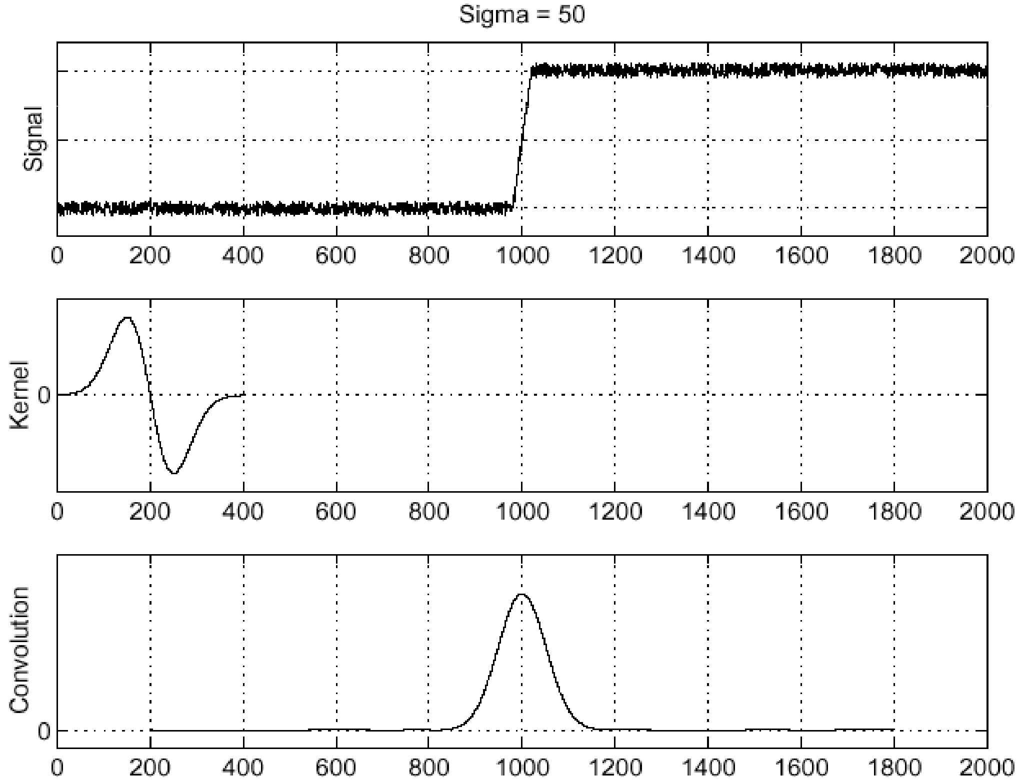
Derivative theorem of convolution

- Differentiation is convolution, and convolution is associative: $\frac{d}{dx}(f * g) = f * \frac{d}{dx} g$
- This saves us one operation:

f

$\frac{d}{dx} g$

$f * \frac{d}{dx} g$



Local descriptors

- Greyvalue derivatives
  - Convolution with Gaussian derivatives

$$\mathbf{v}(x,y) = \begin{pmatrix} I(x,y) * G(\sigma) \\ I(x,y) * G_x(\sigma) \\ I(x,y) * G_y(\sigma) \\ I(x,y) * G_{xx}(\sigma) \\ I(x,y) * G_{xy}(\sigma) \\ I(x,y) * G_{yy}(\sigma) \\ \vdots \end{pmatrix}$$

$$I(x,y) * G(\sigma) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x', y', \sigma) I(x - x', y - y') dx' dy'$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

Local descriptors

Notation for greyvalue derivatives [Koenderink'87]

$$\mathbf{v}(x,y) = \begin{pmatrix} I(x,y) * G(\sigma) \\ I(x,y) * G_x(\sigma) \\ I(x,y) * G_y(\sigma) \\ I(x,y) * G_{xx}(\sigma) \\ I(x,y) * G_{xy}(\sigma) \\ I(x,y) * G_{yy}(\sigma) \\ \vdots \end{pmatrix} = \begin{pmatrix} L(x,y) \\ L_x(x,y) \\ L_y(x,y) \\ L_{xx}(x,y) \\ L_{xy}(x,y) \\ L_{yy}(x,y) \\ \vdots \end{pmatrix}$$

Invariance?

Local descriptors – rotation invariance

Invariance to image rotation : differential invariants [Koen87]

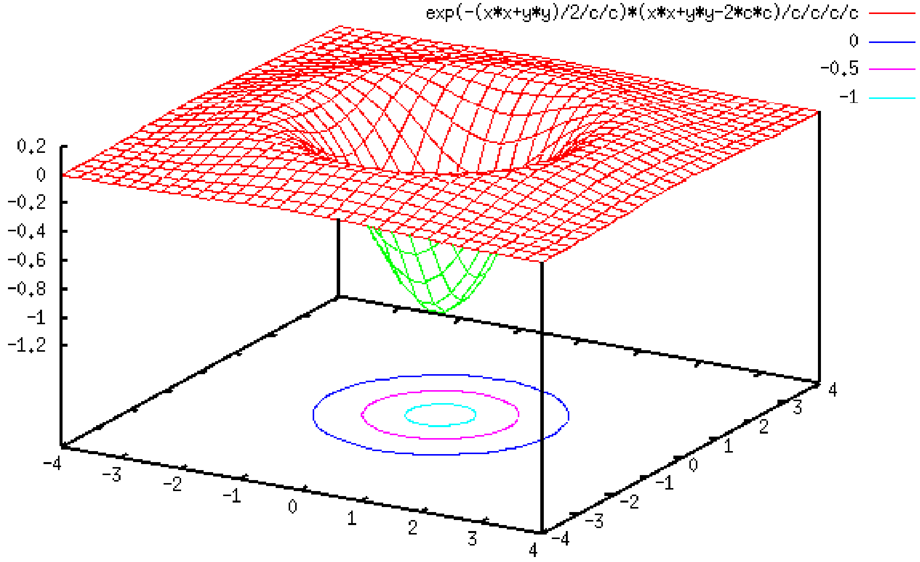
gradient magnitude → $L_x L_x + L_y L_y$

Laplacian → $L_{xx} + L_{yy}$

$$
\begin{bmatrix}
L \\
L_x L_x + L_y L_y \\
L_{xx} L_x L_x + 2 L_{xy} L_x L_y + L_{yy} L_{yy} \\
L_{xx} + L_{yy} \\
L_{xx} L_{xx} + 2 L_{xy} L_{xy} + L_{yy} L_{yy} \\
\cdots \\
\cdots \\
\cdots \\
\cdots
\end{bmatrix}
$$

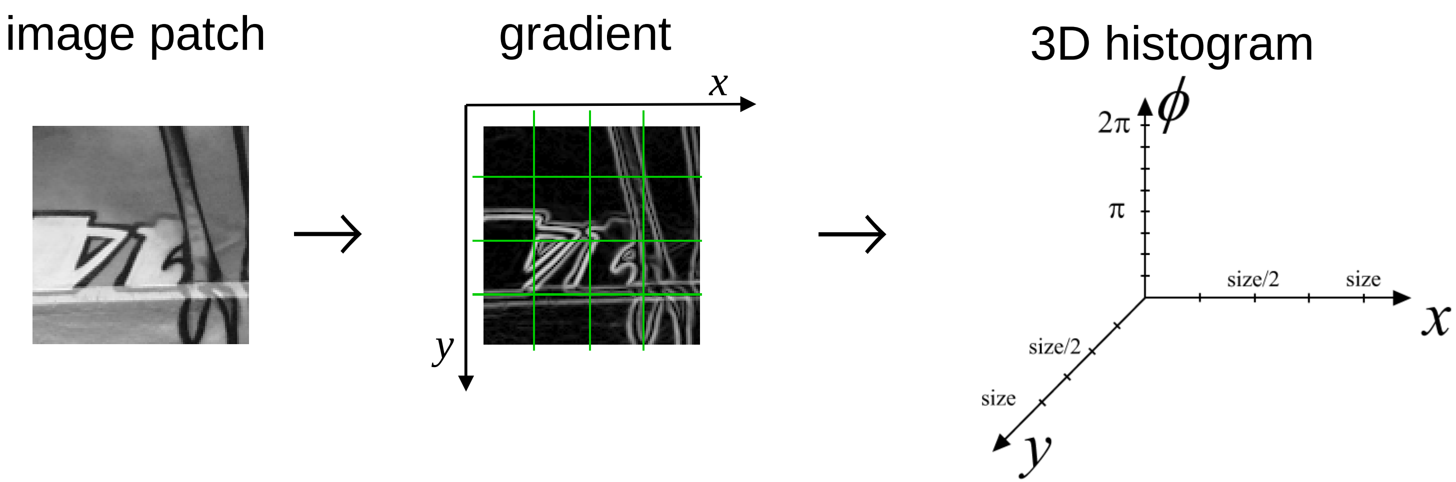
Laplacian of Gaussian (LOG)

$$LOG = G_{xx}(\sigma) + G_{yy}(\sigma)$$



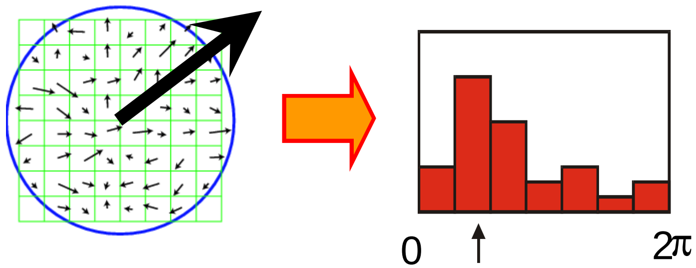
SIFT descriptor [Lowe'99]

- Approach
  - 8 orientations of the gradient
  - 4x4 spatial grid
  - Dimension 128
  - soft-assignment to spatial bins
  - normalization of the descriptor to norm one
  - comparison with Euclidean distance

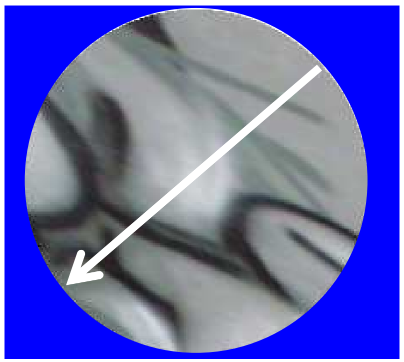


Local descriptors - rotation invariance

- Estimation of the dominant orientation
  - extract gradient orientation
  - histogram over gradient orientation
  - peak in this histogram



- Rotate patch in dominant direction



Local descriptors – illumination change

- Robustness to illumination changes

in case of an affine transformation $I_1(\mathbf{x}) = aI_2(\mathbf{x}) + b$

Local descriptors – illumination change

- Robustness to illumination changes

in case of an affine transformation $I_1(\mathbf{x}) = aI_2(\mathbf{x}) + b$

- Normalization of derivatives with gradient magnitude

$$(L_{xx} + L_{yy}) \Big/ \sqrt{L_x L_x + L_y L_y}$$

Local descriptors – illumination change

- Robustness to illumination changes

in case of an affine transformation $I_1(\mathbf{x}) = aI_2(\mathbf{x}) + b$

- Normalization of derivatives with gradient magnitude

$$(L_{xx} + L_{yy}) \Big/ \sqrt{L_x L_x + L_y L_y}$$

- Normalization of the image patch with mean and variance

Invariance to scale changes

- Scale change between two images
- Scale factor s can be eliminated
- Support region for calculation!!
  - In case of a convolution with Gaussian derivatives defined by σ

$$I(x,y)*G(\sigma)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}G(x',y',\sigma)I(x-x',y-y')dx'dy'$$

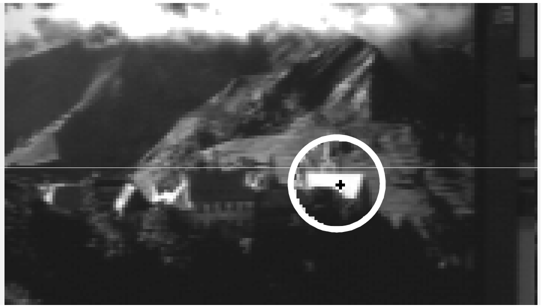
$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}\exp(-\frac{x^2+y^2}{2\sigma^2})$$

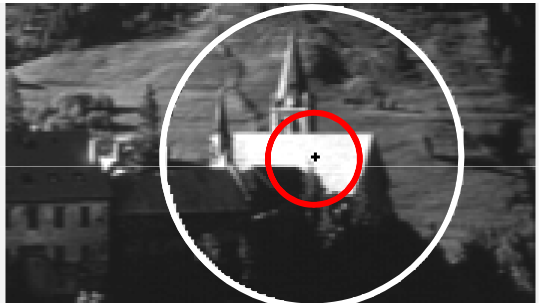
Overview

- Introduction to local features
- Harris interest points + SSD, ZNCC, SIFT
- **Scale & affine invariant interest point detectors**
- Evaluation and comparison of different detectors
- Region descriptors and their performance

Scale invariance - motivation

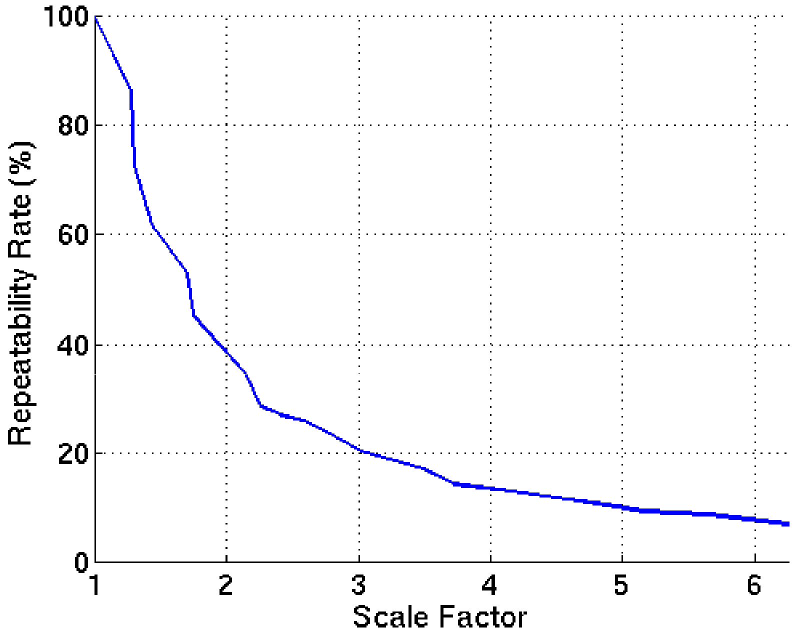
- Description regions have to be adapted to scale changes

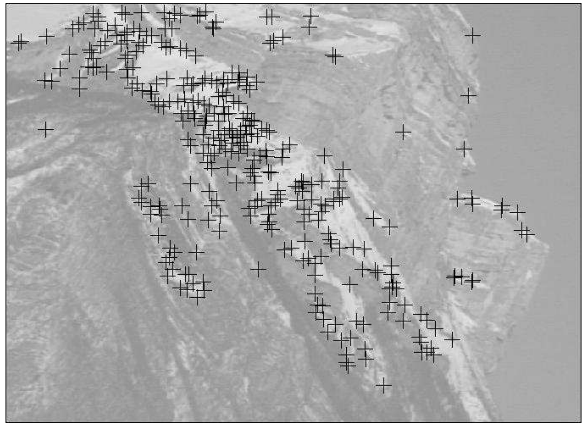


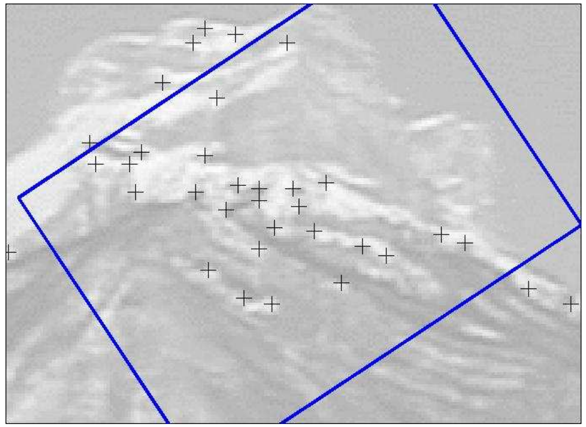


- Interest points have to be repeatable for scale changes

Harris detector + scale changes







Repeatability rate

$$R(\varepsilon) = \frac{|\{(\mathbf{a}_i, \mathbf{b}_i) \mid dist(H(\mathbf{a}_i), \mathbf{b}_i) < \varepsilon\}|}{\max(|\mathbf{a}_i|, |\mathbf{b}_i|)}$$

Scale adaptation

Scale change between two images

$$I_1\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = I_2\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = I_2\begin{pmatrix} sx_1 \\ sy_1 \end{pmatrix}$$

Scale adapted derivative calculation

Scale adaptation

Scale change between two images

$$I_1\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = I_2\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = I_2\begin{pmatrix} sx_1 \\ sy_1 \end{pmatrix}$$

Scale adapted derivative calculation

$$I_1\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \otimes G_{i_1 \ldots i_n}(\sigma) = \boldsymbol{s}^{\boldsymbol{n}} I_2\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \otimes G_{i_1 \ldots i_n}(\boldsymbol{s\sigma})$$

Scale adaptation

$$G(\tilde{\sigma}) \otimes \begin{bmatrix} L_x^2(\sigma) & L_x L_y(\sigma) \\ L_x L_y(\sigma) & L_y^2(\sigma) \end{bmatrix}$$

where $L_i(\sigma)$ are the derivatives with Gaussian convolution

Scale adaptation

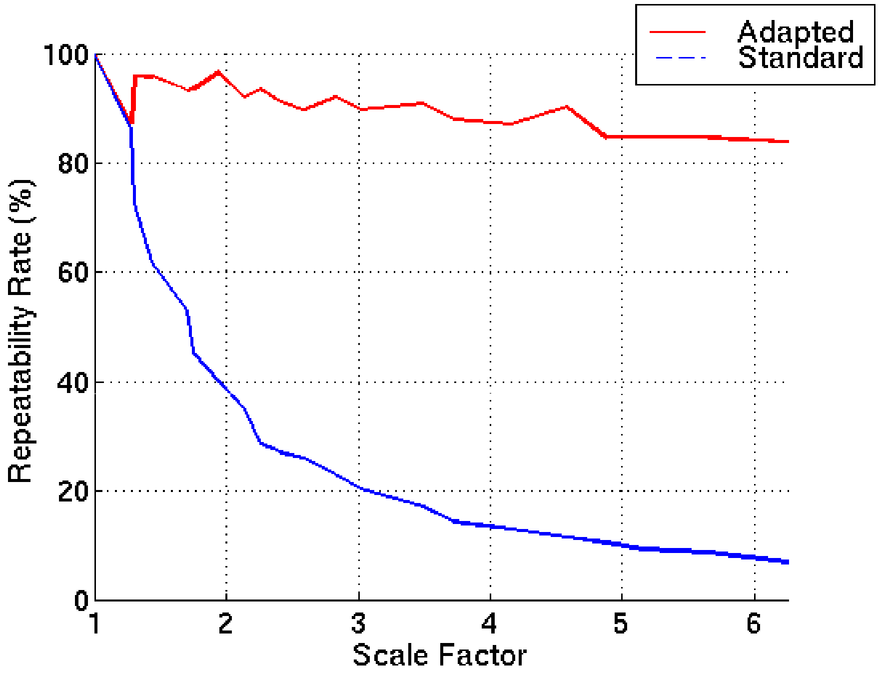
$$G(\tilde{\sigma}) \otimes \begin{bmatrix} L_x^2(\sigma) & L_x L_y(\sigma) \\ L_x L_y(\sigma) & L_y^2(\sigma) \end{bmatrix}$$

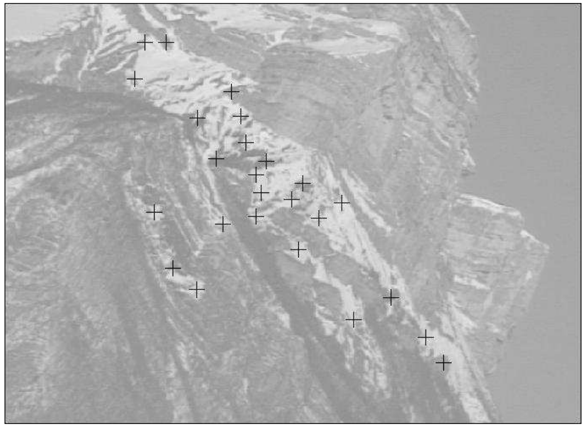
where $L_i(\sigma)$ are the derivatives with Gaussian convolution

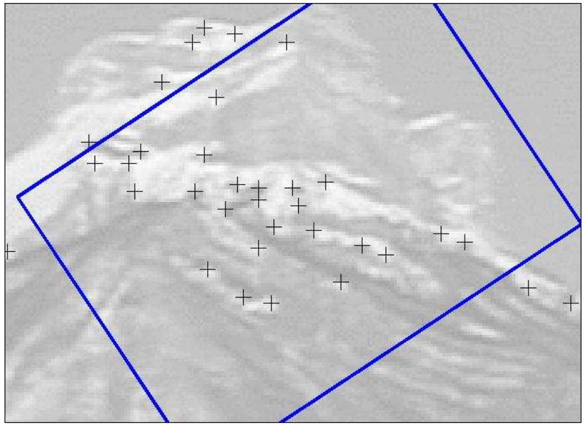
Scale adapted auto-correlation matrix

$$s^2 G(s\tilde{\sigma}) \otimes \begin{bmatrix} L_x^2(s\sigma) & L_x L_y(s\sigma) \\ L_x L_y(s\sigma) & L_y^2(s\sigma) \end{bmatrix}$$

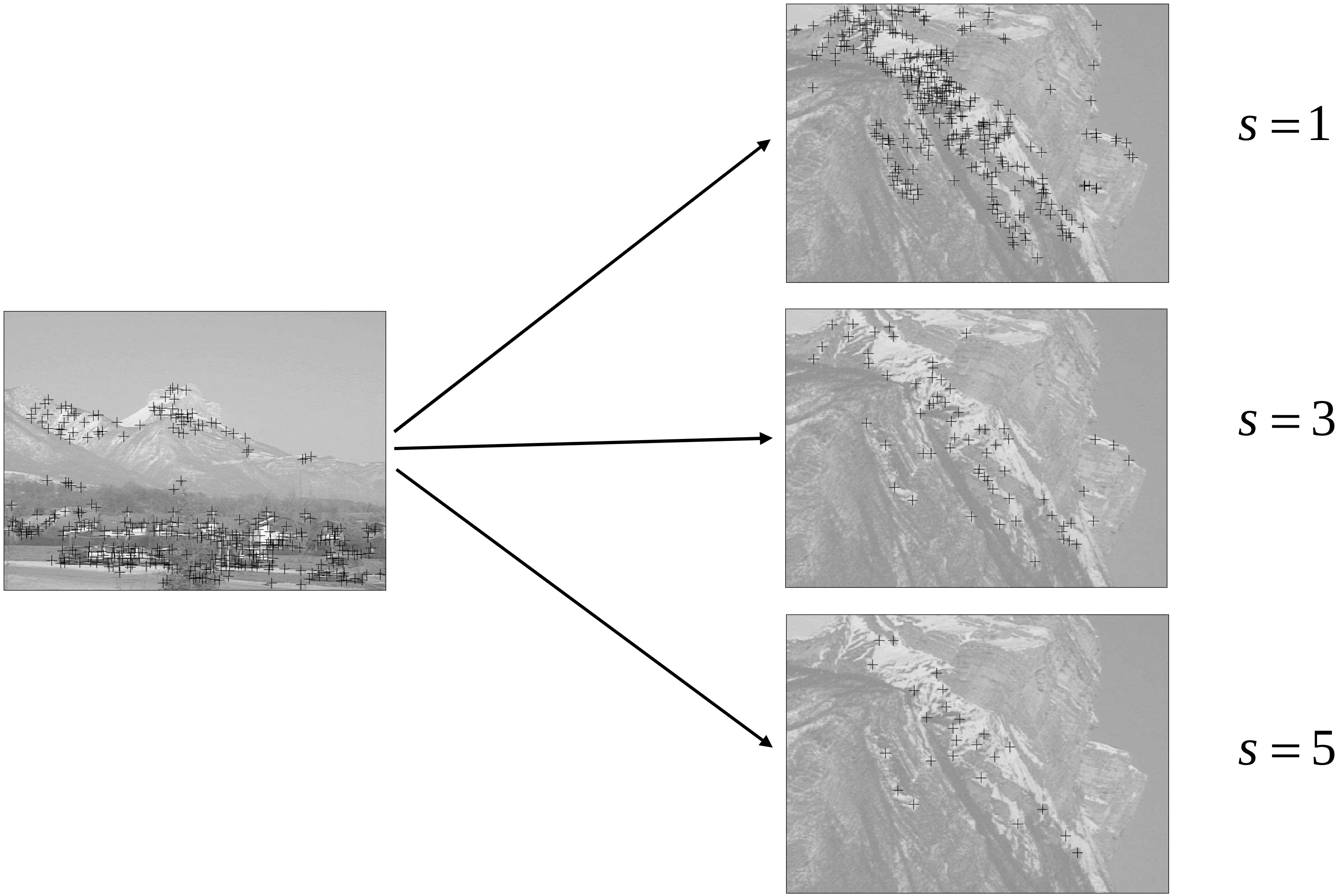
Harris detector – adaptation to scale



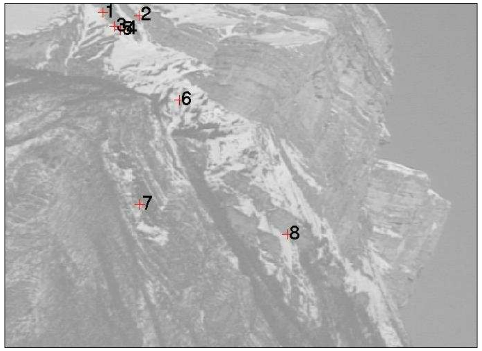




Multi-scale matching algorithm

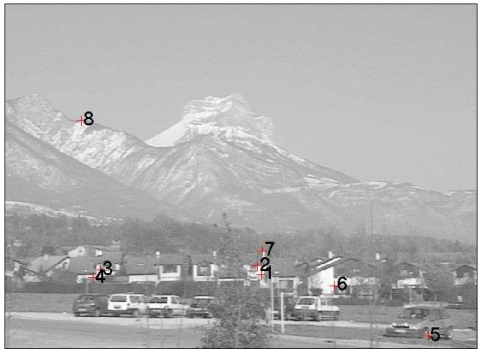


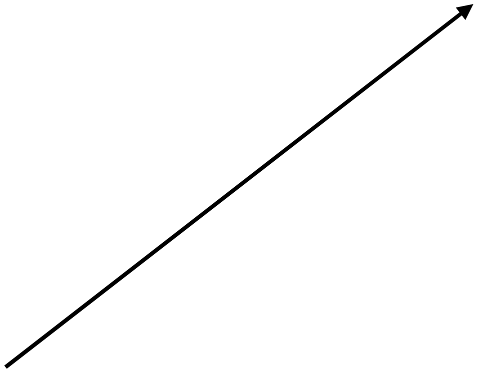
Multi-scale matching algorithm



$s = 1$

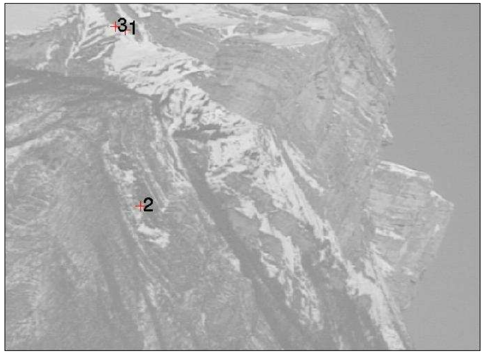
8 matches





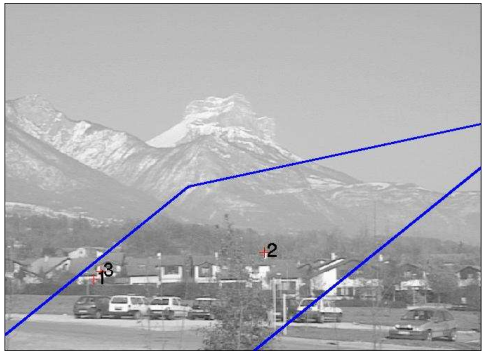
Multi-scale matching algorithm

Robust estimation of a global affine transformation

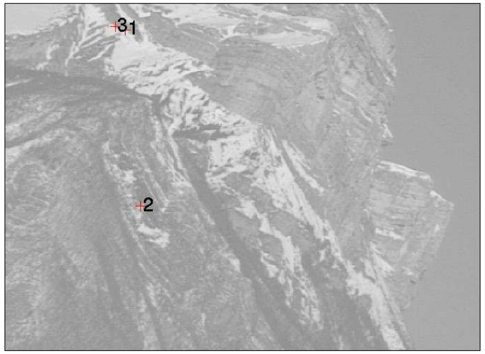


$s = 1$

3 matches

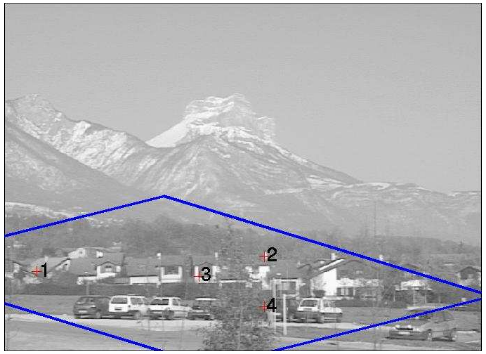


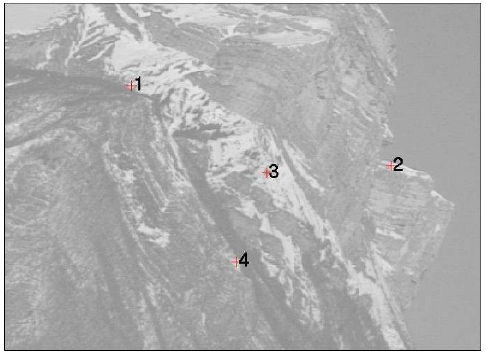
Multi-scale matching algorithm



$s = 1$

3 matches

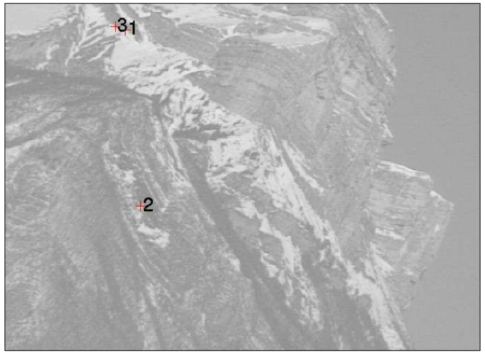




$s = 3$

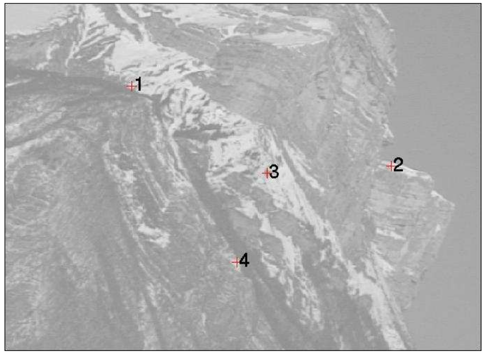
4 matches

Multi-scale matching algorithm



$s = 1$

3 matches

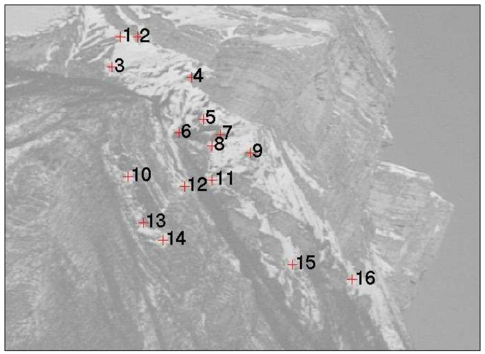


$s = 3$

4 matches

highest number of matches

correct scale



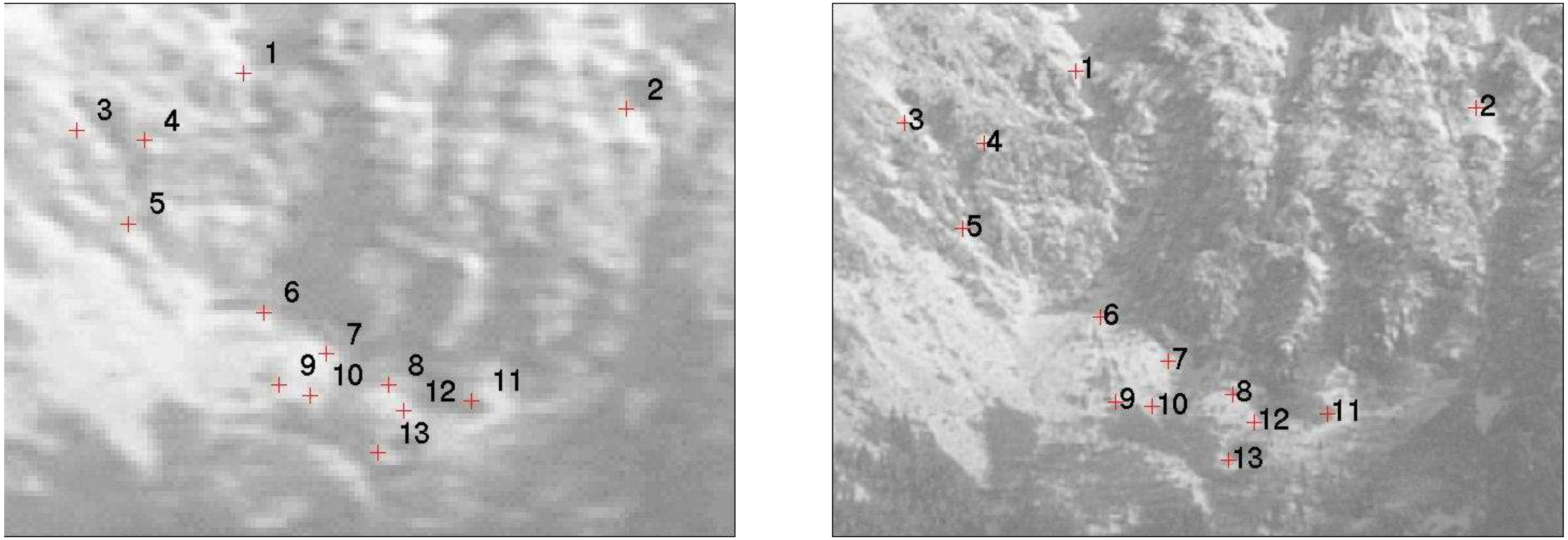
$s = 5$

16 matches

Matching results

Scale change of 5.7

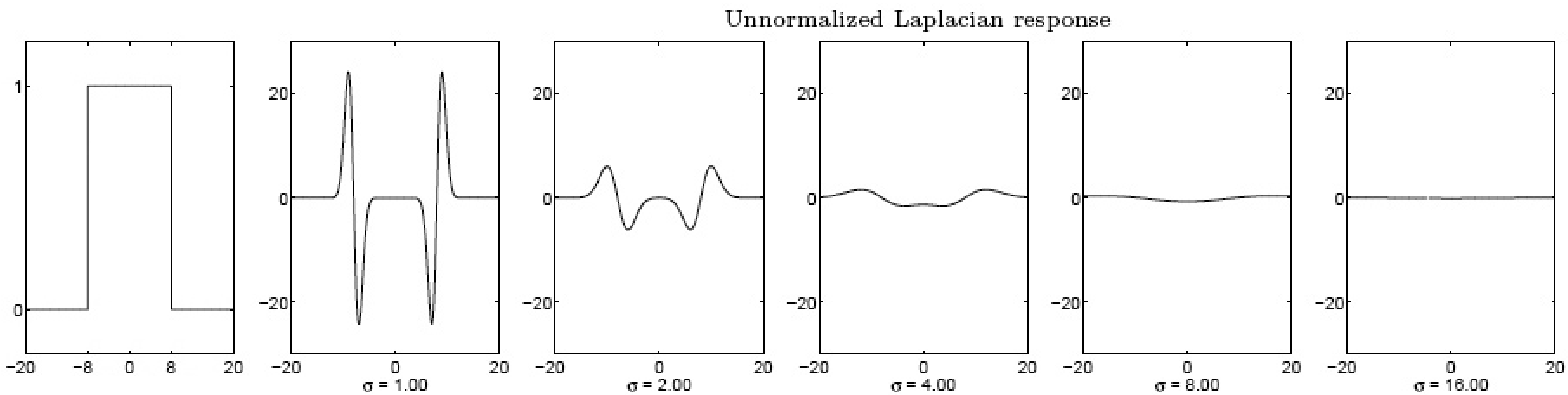
Matching results



100% correct matches (13 matches)

Scale selection

- We want to find the characteristic scale of the blob by convolving it with Laplacians at several scales and looking for the maximum response
- However, Laplacian response decays as scale increases:



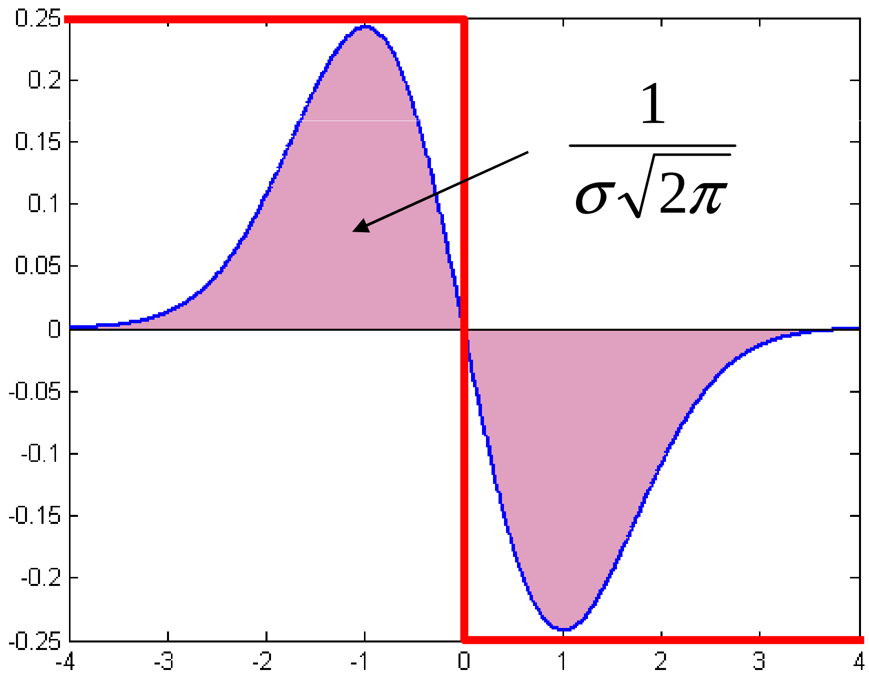
original signal (radius=8)

increasing σ →

Why does this happen?

Scale normalization

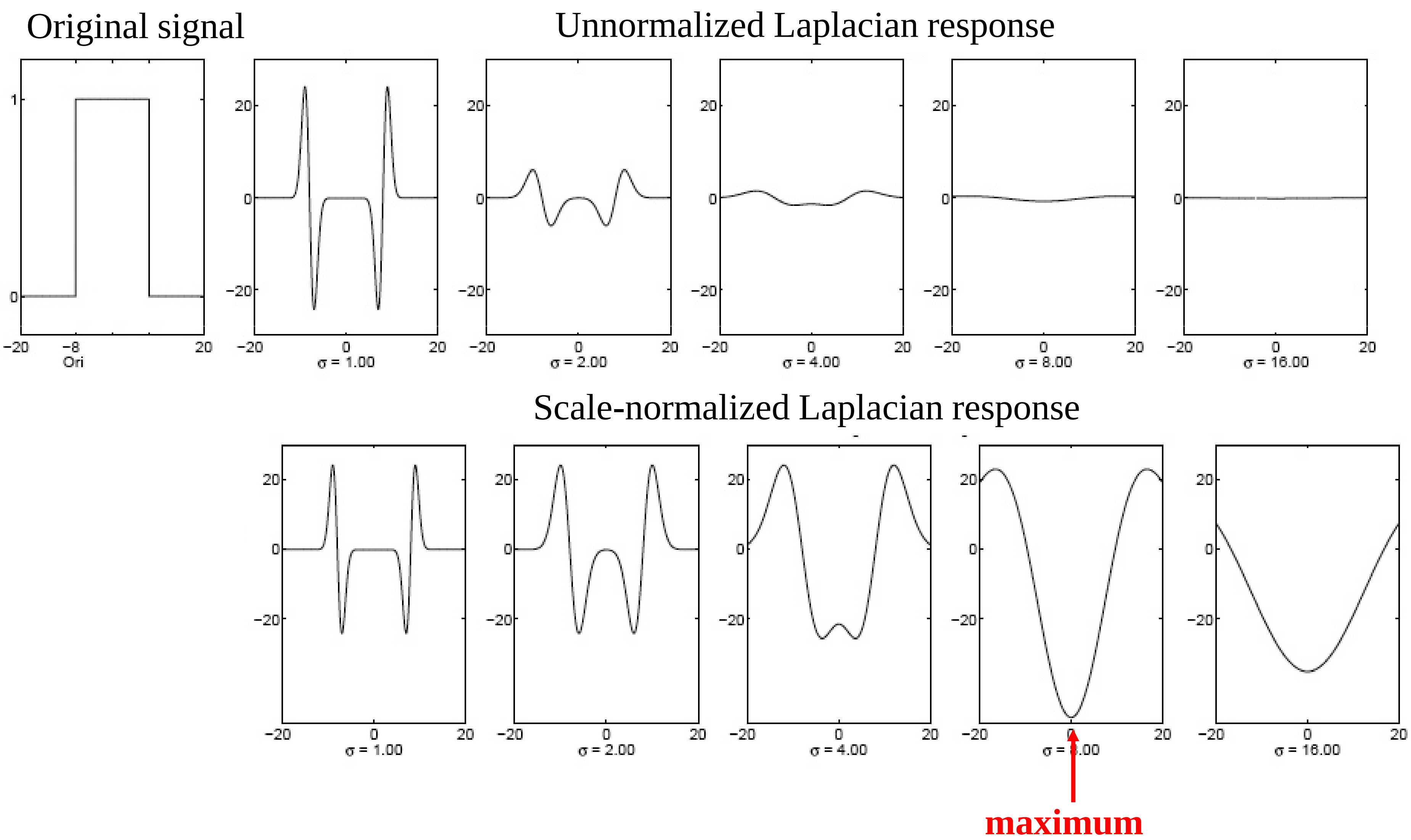
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases



Scale normalization

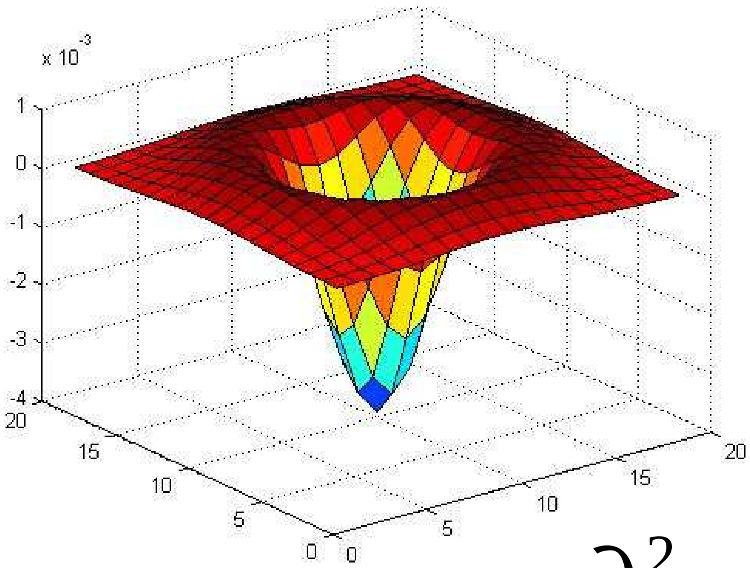
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases
- To keep response the same (scale-invariant), must multiply Gaussian derivative by σ
- Laplacian is the second Gaussian derivative, so it must be multiplied by σ^2

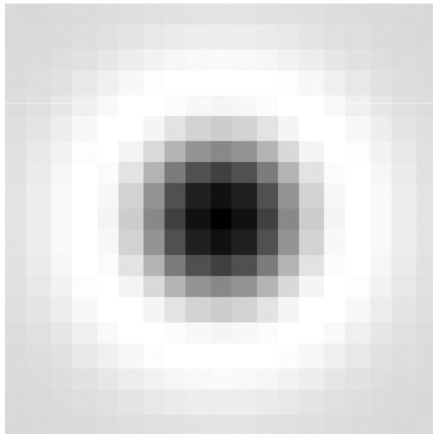
Effect of scale normalization



Blob detection in 2D

- Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D

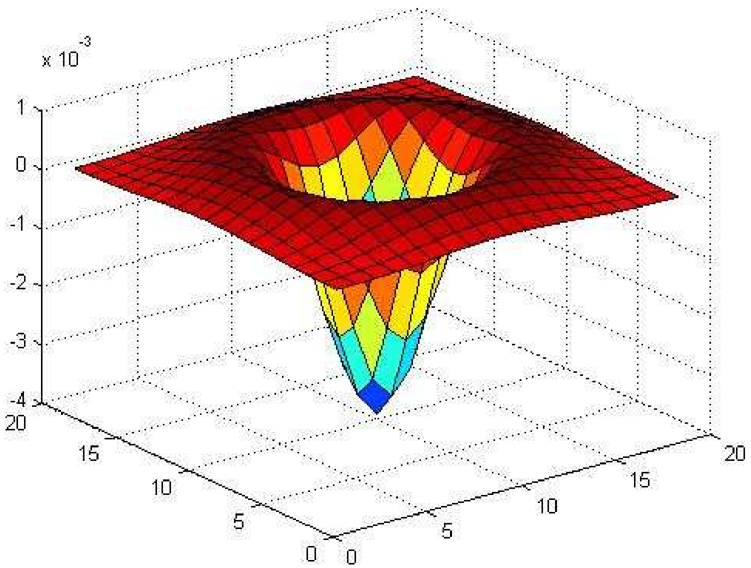


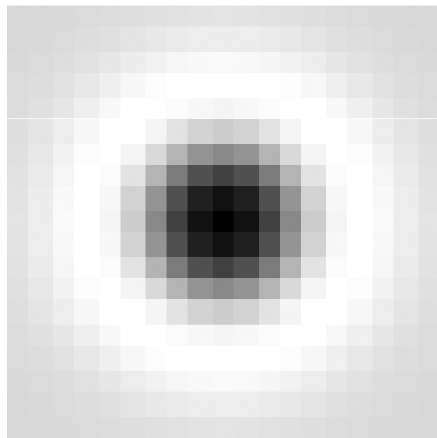


$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

Blob detection in 2D

- Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D





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

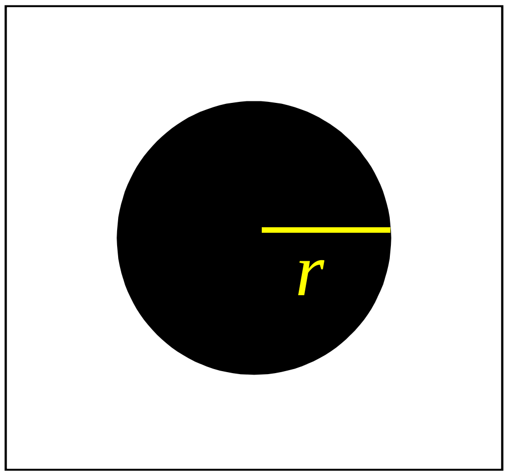
Scale selection

- The 2D Laplacian is given by

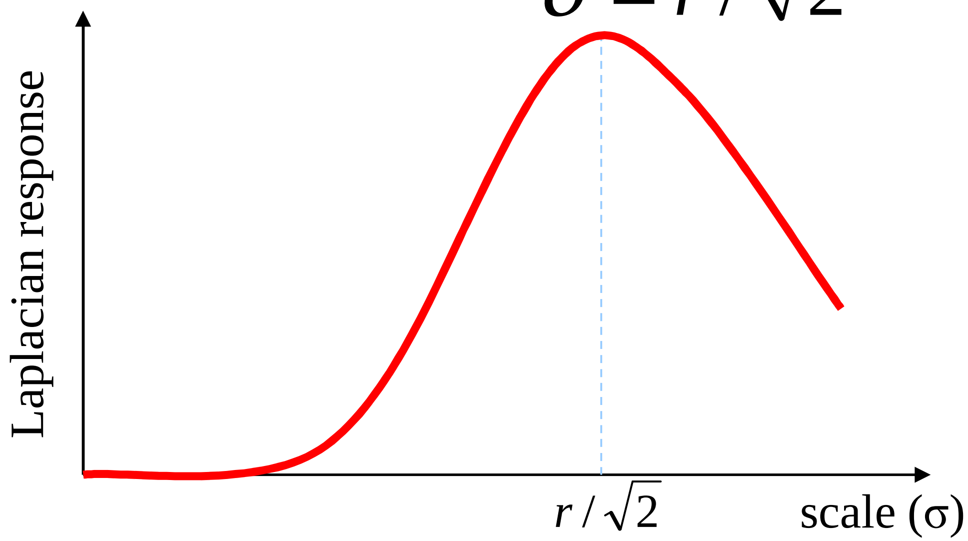
$$(x^2 + y^2 - 2\sigma^2)\, e^{-(x^2+y^2)/2\sigma^2} \quad \text{(up to scale)}$$

- For a binary circle of radius r, the Laplacian achieves a maximum at

$$\sigma = r / \sqrt{2}$$

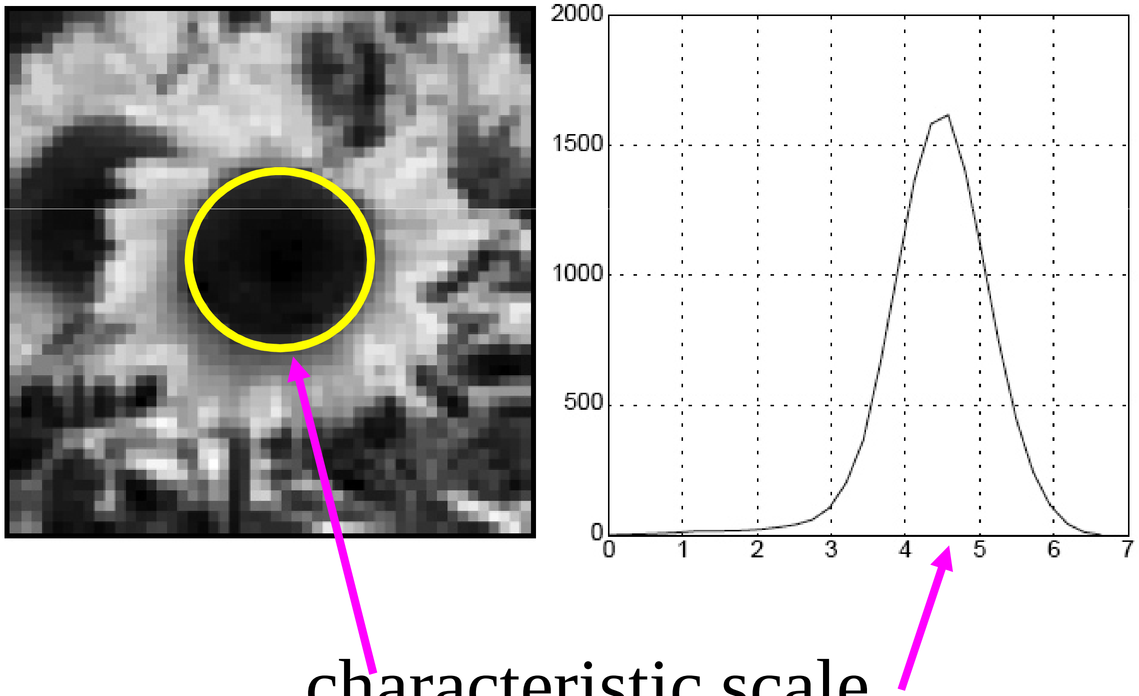


image



Characteristic scale

- We define the characteristic scale as the scale that produces peak of Laplacian response



T. Lindeberg (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision* **30** (2): pp 77--116.

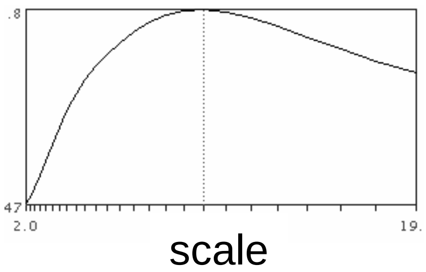
Scale selection

- For a point compute a value (gradient, Laplacian etc.) at several scales
- Normalization of the values with the scale factor

  e.g. Laplacian $| s^2 (L_{xx} + L_{yy}) |$

- Select scale s^* at the maximum $\rightarrow$ characteristic scale

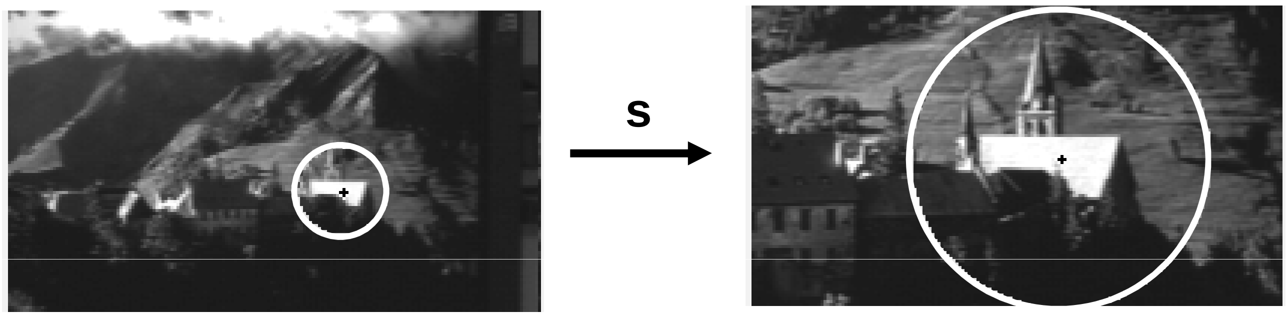
$| s^2 (L_{xx} + L_{yy}) |$

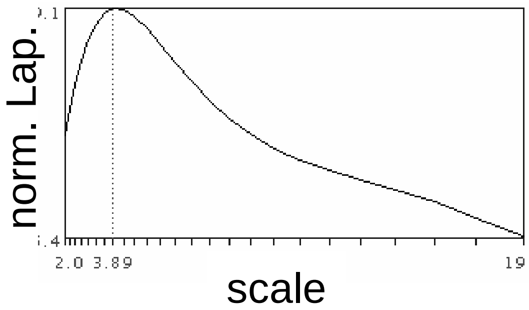


- Exp. results show that the Laplacian gives best results

Scale selection

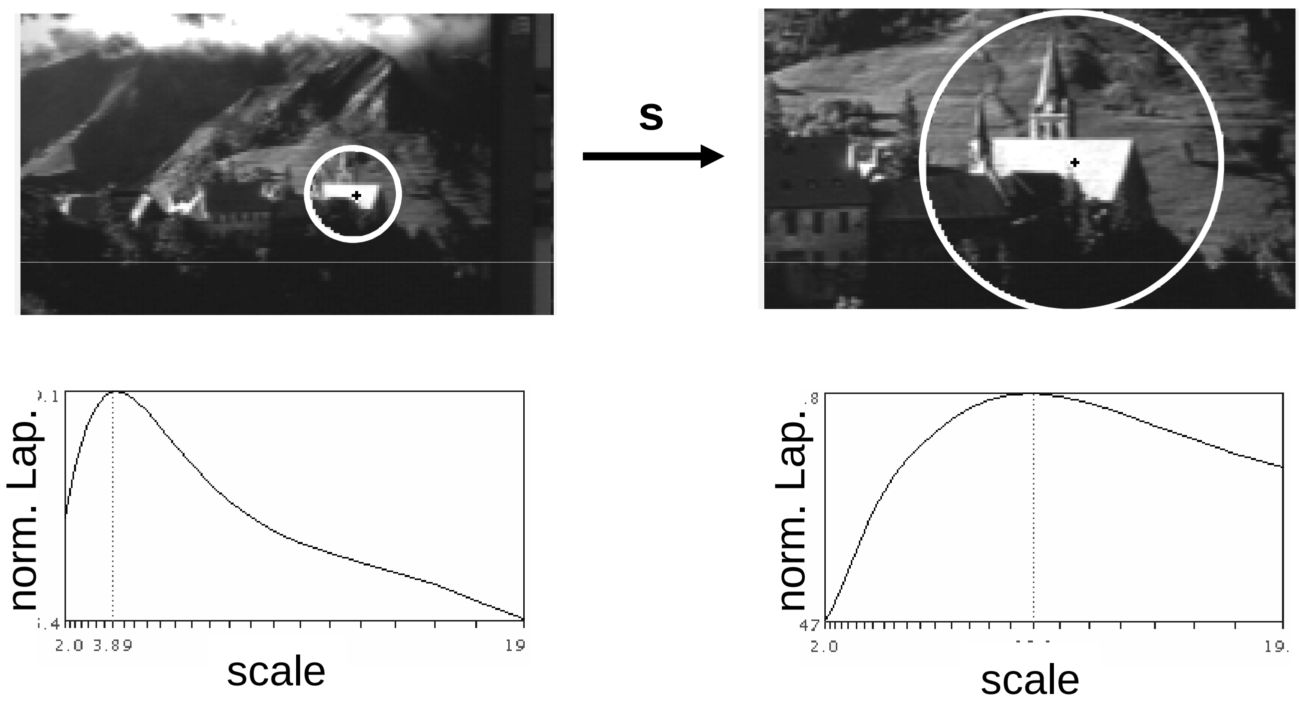
- Scale invariance of the characteristic scale





Scale selection

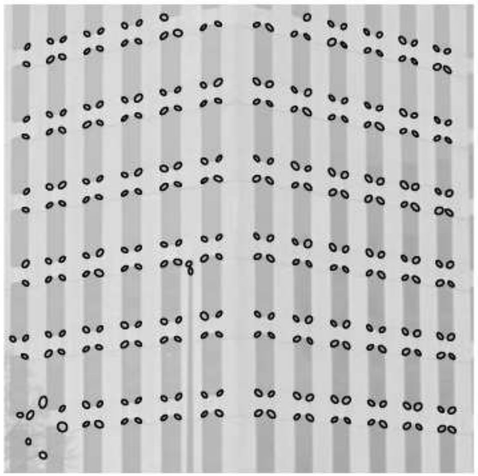
- Scale invariance of the characteristic scale



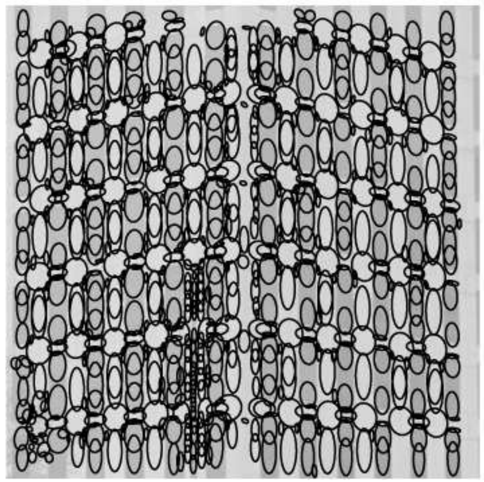
- Relation between characteristic scales $s \cdot s_1^* = s_2^*$

Scale-invariant detectors

- Harris-Laplace (Mikolajczyk & Schmid'01)
- Laplacian detector (Lindeberg'98)
- Difference of Gaussian (Lowe'99)



Harris-Laplace



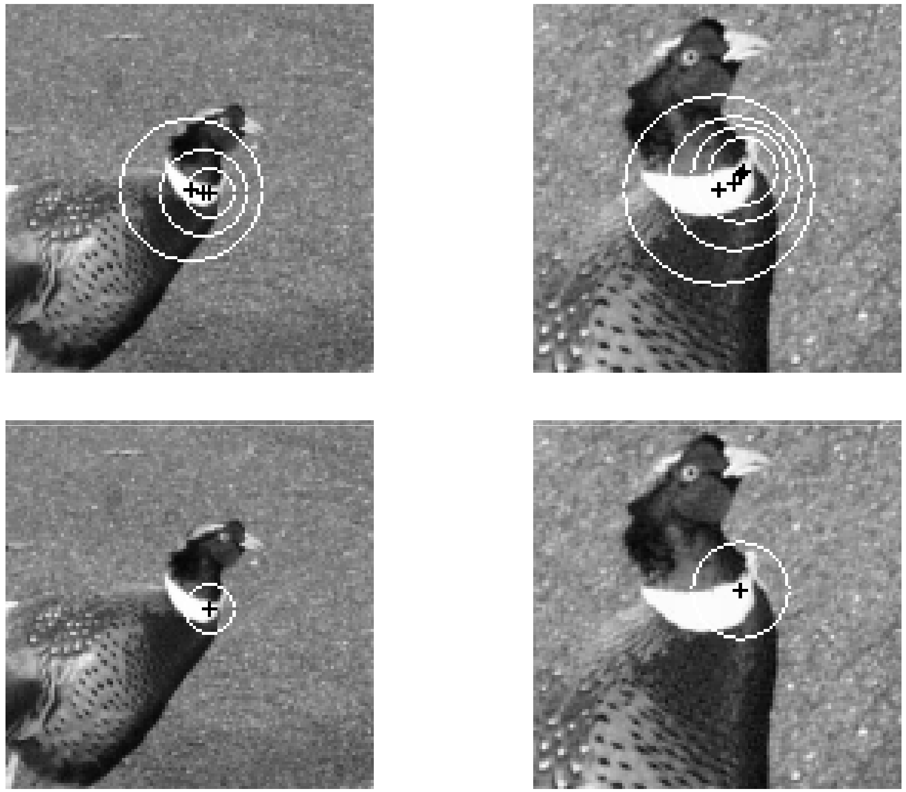
Laplacian

Harris-Laplace

multi-scale Harris points

selection of points at

maximum of Laplacian

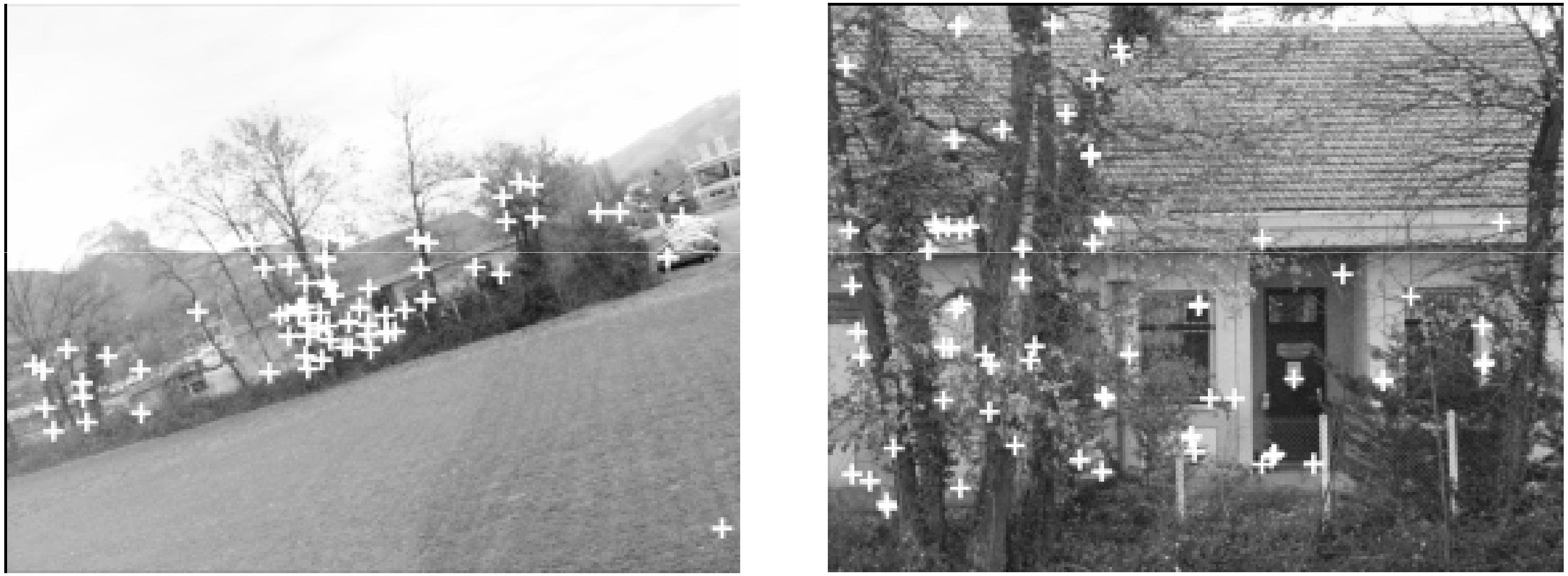


➡ invariant points + associated regions [Mikolajczyk & Schmid'01]

Matching results

213 / 190 detected interest points

Matching results



58 points are initially matched

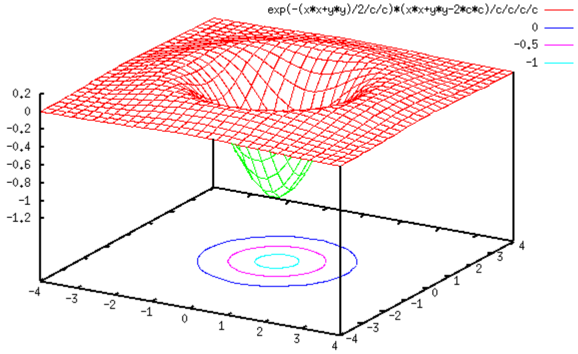
Matching results

32 points are matched after verification – all correct

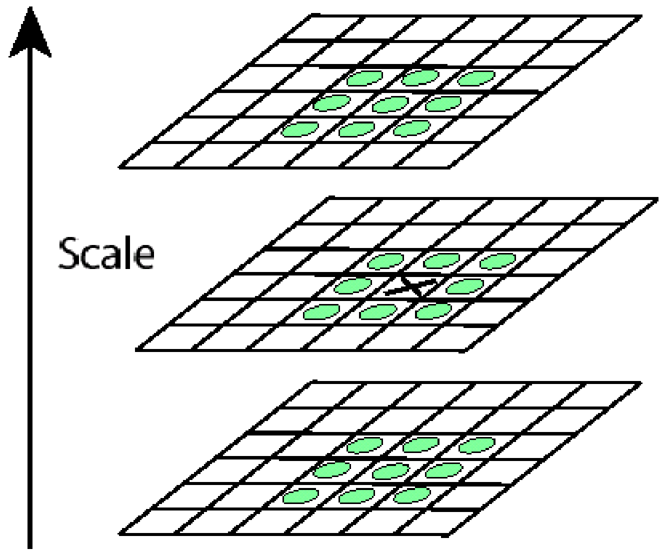
LOG detector

Convolve image with scale-normalized Laplacian at several scales



$$LOG = s^2 (G_{xx}(\sigma) + G_{yy}(\sigma))$$

Detection of maxima and minima of Laplacian in scale space



Hessian detector

Hessian matrix $$H(x) = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix}$$

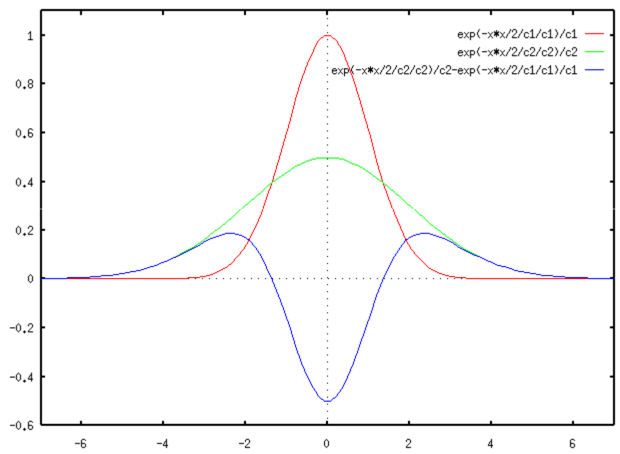
Determinant of Hessian matrix $$DET = L_{xx} L_{yy} - L_{xy}{}^2$$

Penalizes/eliminates long structures

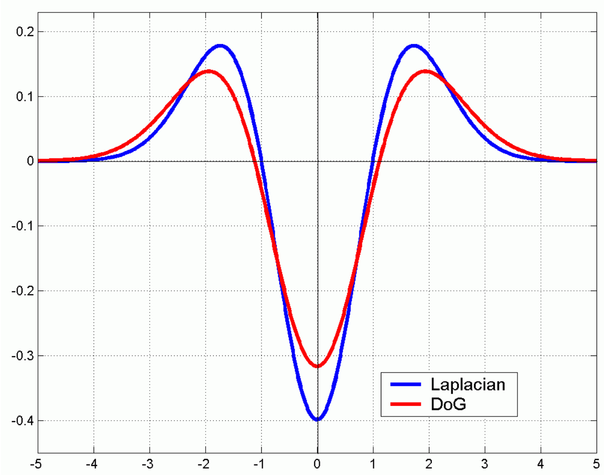
- with small derivative in a single direction

Efficient implementation

- Difference of Gaussian (DOG) approximates the Laplacian $DOG = G(k\sigma) - G(\sigma)$

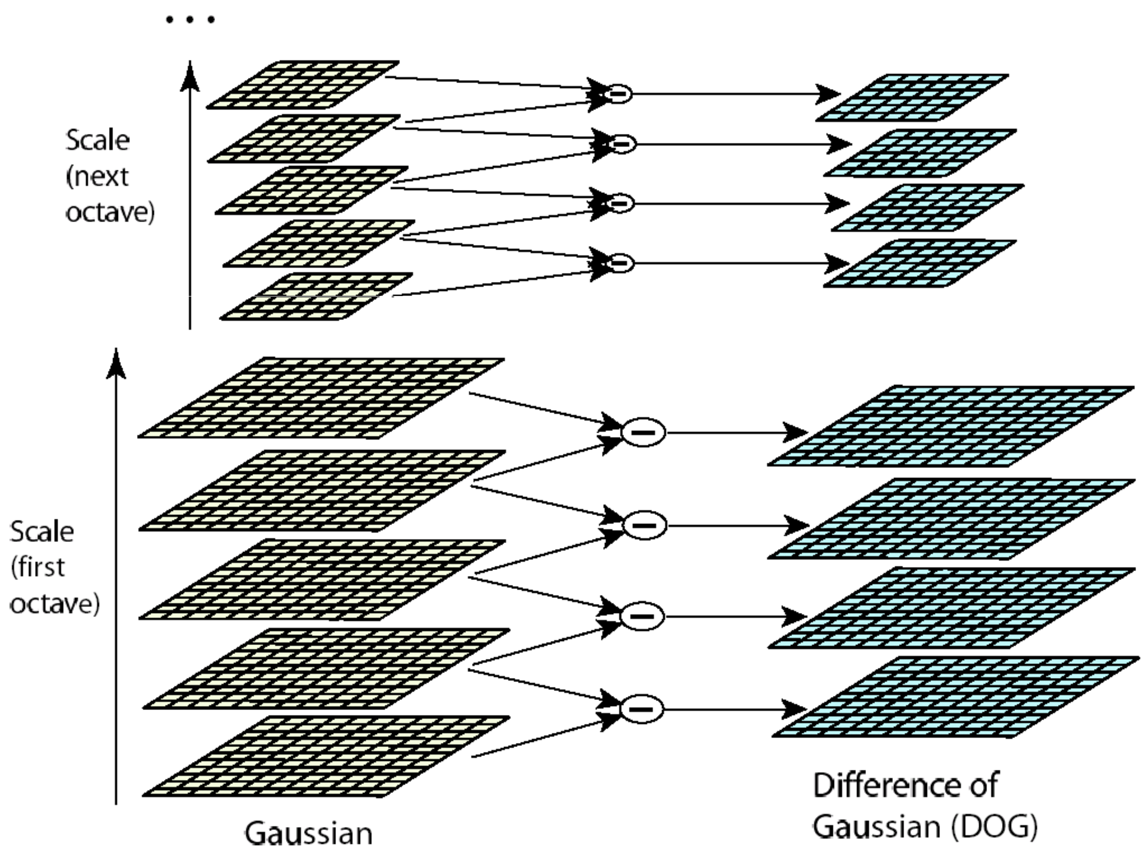


- Error due to the approximation



DOG detector

- Fast computation, scale space processed one octave at a time



David G. Lowe. "Distinctive image features from scale-invariant keypoints."I*JCV* 60 (2).