6.891

Computer Vision and Applications

Prof. Trevor. Darrell

Lecture 7: Features and Geometry

- Affine invariant features
- Epipolar geometry
- Essential matrix

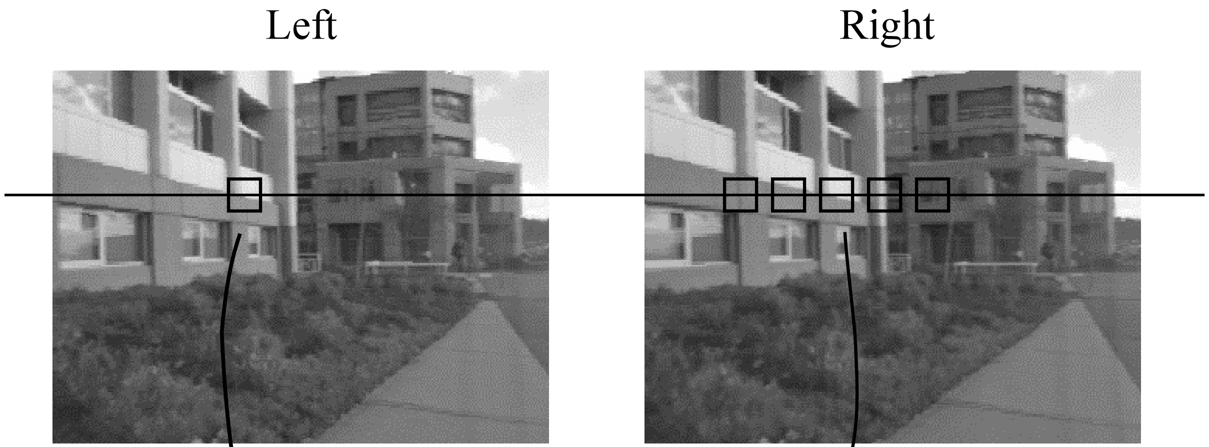
Readings: Mikolajczyk and Schmid; F&P Ch 10

Last time

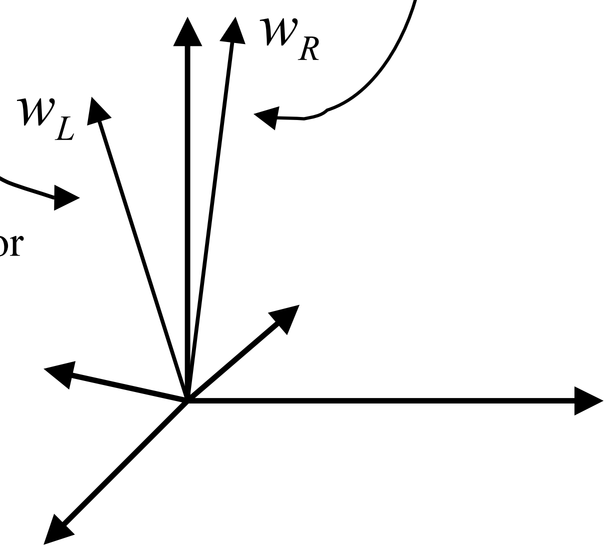
Interesting points, correspondence, affine patch tracking

Scale and rotation invariant descriptors [Lowe]

Images as Vectors



"Unwrap" image to form vector, using raster scan order



Each window is a vector in an m^2 dimensional vector space. Normalization makes them unit length.

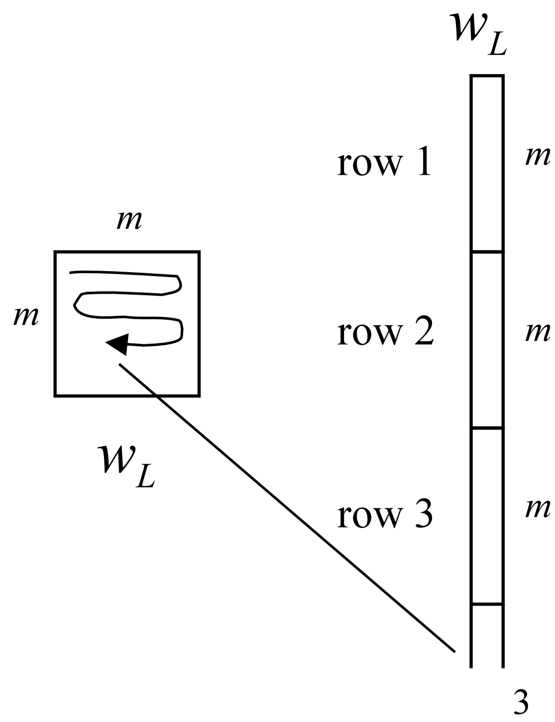
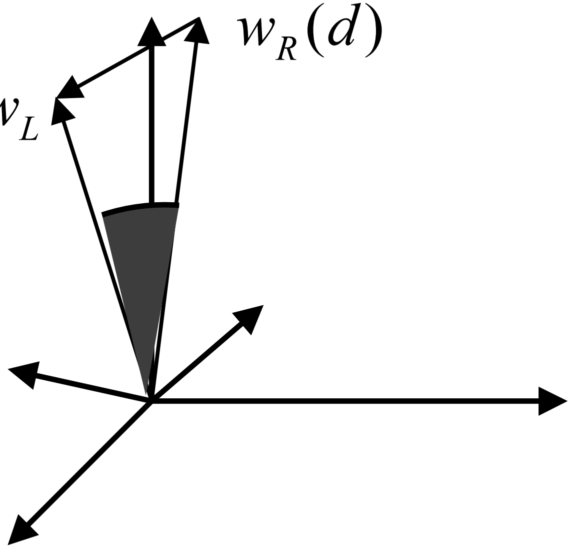


Image Metrics



(Normalized) Sum of Squared Differences

$$C_{\text{SSD}}(d) = \sum_{(u,v)\in W_m(x,y)} [\hat{I}_L(u,v) - \hat{I}_R(u-d,v)]^2$$

$$= \left\| w_L - w_R(d) \right\|^2$$

Normalized Correlation

$$C_{\text{NC}}(d) = \sum_{(u,v)\in W_m(x,y)} \hat{I}_L(u,v)\hat{I}_R(u-d,v)$$

$$= w_L \cdot w_R(d) = \cos\theta$$

$$\boxed{d^* = \arg\min_d \left\| w_L - w_R(d) \right\|^2 = \arg\max_d w_L \cdot w_R(d)}$$

Harris detector

Auto-correlation matrix

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

- Auto-correlation matrix
  - 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
- Interest point detection
  - threshold on the eigenvalues
  - local maximum for localization

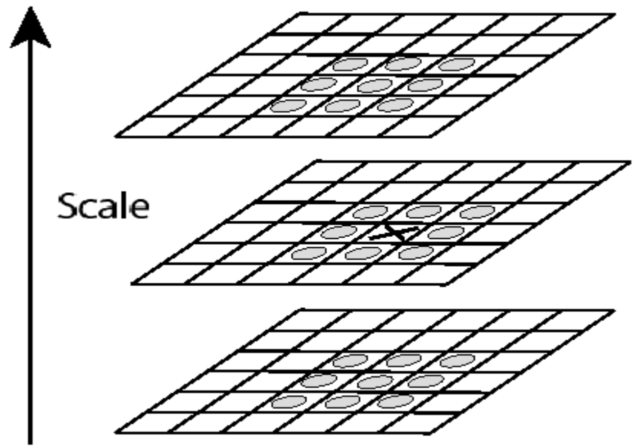
Key point localization

- 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:

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

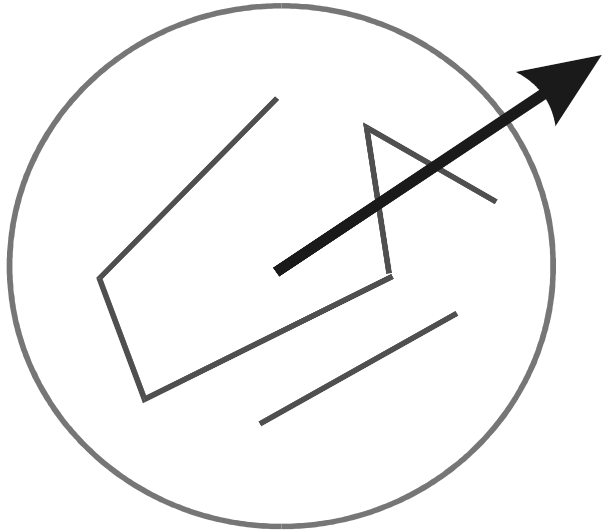
- Offset of extremum (use finite differences for derivatives):

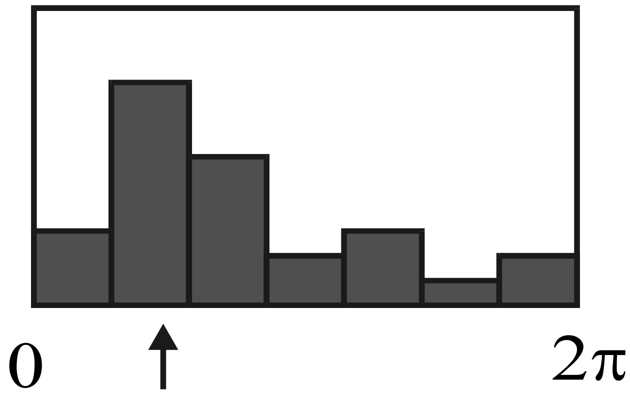
$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1} \frac{\partial D}{\partial \mathbf{x}}$$



Select canonical orientation

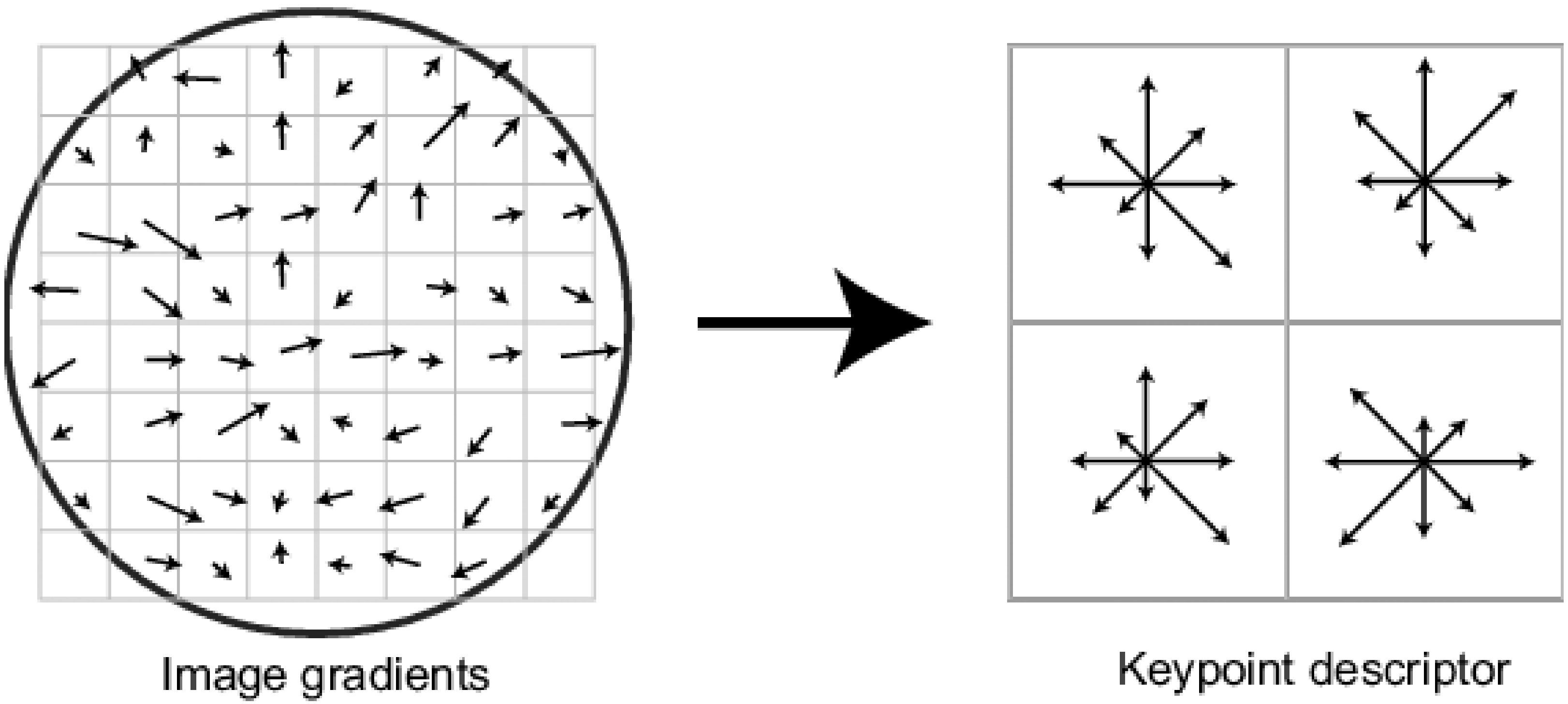
- Create histogram of local gradient directions computed at selected scale
- Assign canonical orientation at peak of smoothed histogram
- Each key specifies stable 2D coordinates (x, y, scale, orientation)





SIFT vector formation

- Thresholded image gradients are sampled over 16x16 array of locations in scale space
- Create array of orientation histograms
- 8 orientations x 4x4 histogram array = 128 dimensions



Today

Affine Invariant Interest points [Schmid]

Evaluation of interest points and descriptors [Schmid]

Epipolar geometry and the Essential Matrix

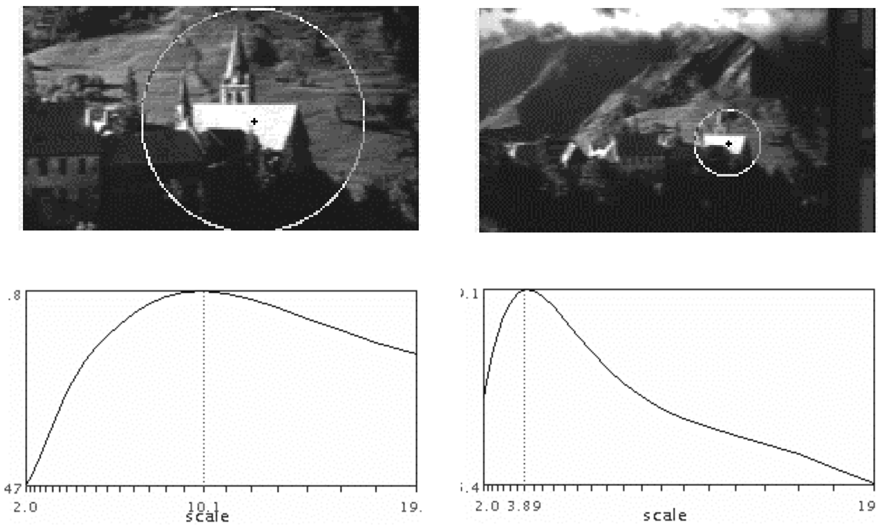
Affine invariance of interest points

Cordelia Schmid

CVPR’03 Tutorial

Scale invariant Harris points

- Multi-scale extraction of Harris interest points
- Selection of points at characteristic scale in scale space



Laplacian

Chacteristic scale :

- maximum in scale space
- scale invariant

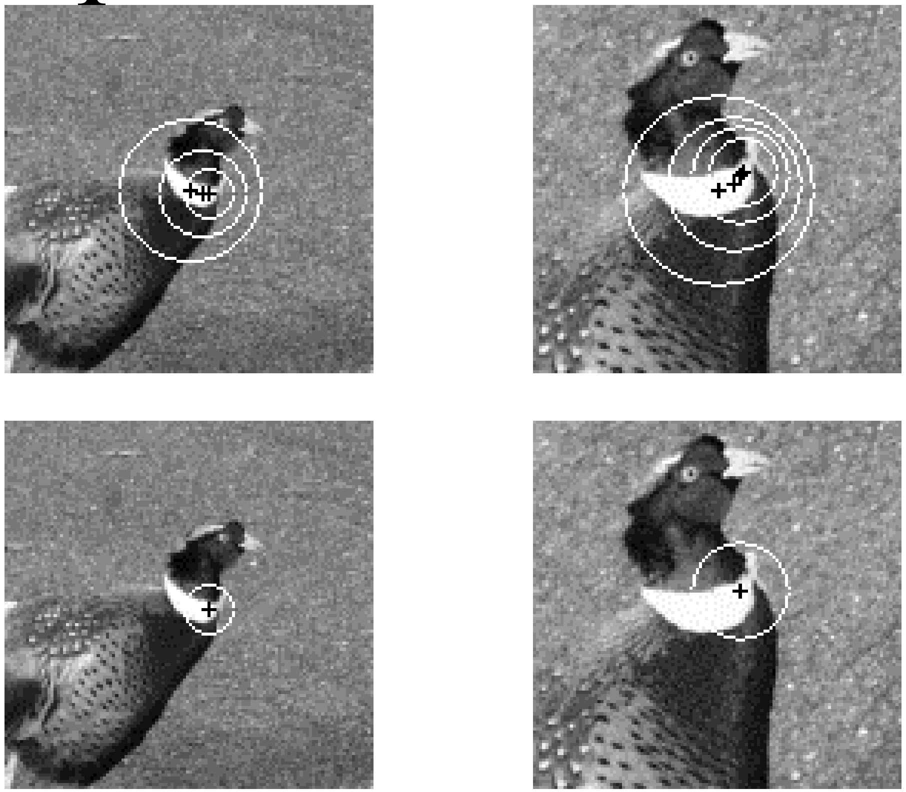
Scale invariant interest points

multi-scale Harris points

selection of points

at the characteristic scale

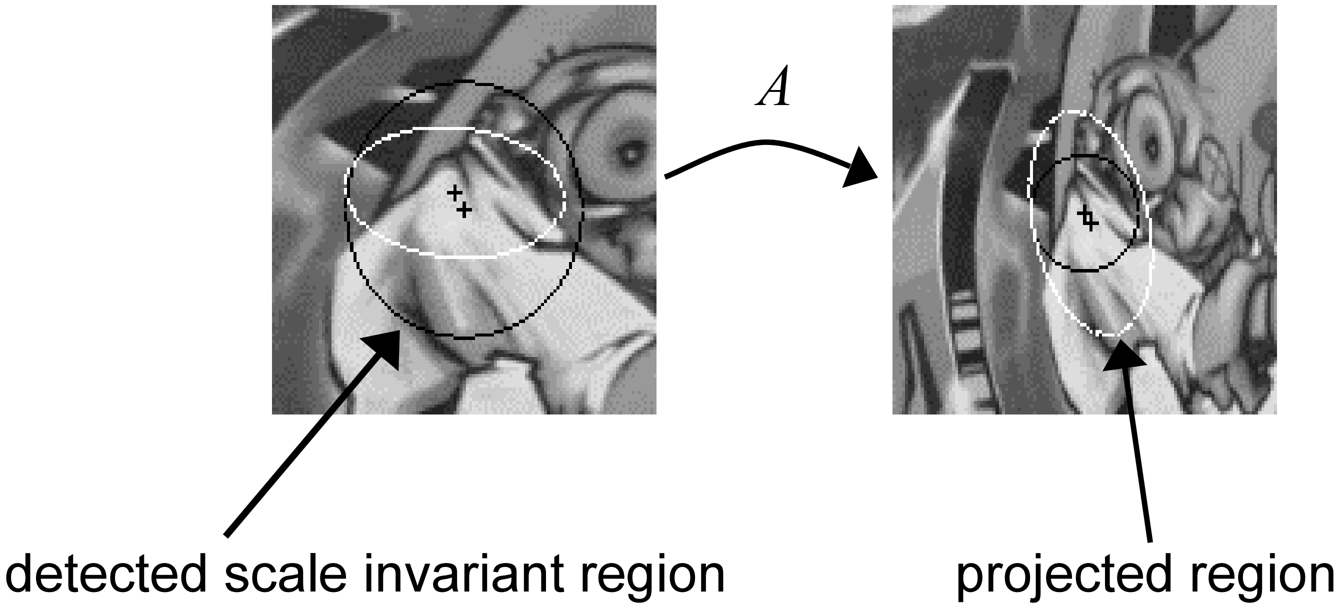
with Laplacian



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

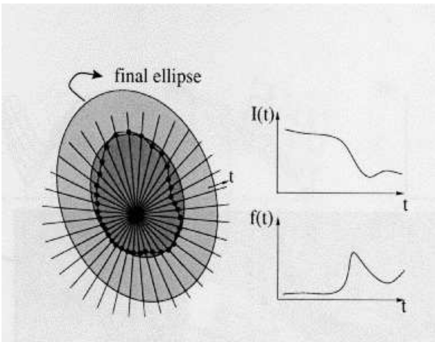
Viewpoint changes

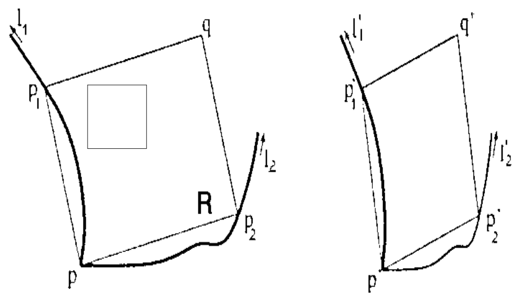
- Locally approximated by an affine transformation



State of the art

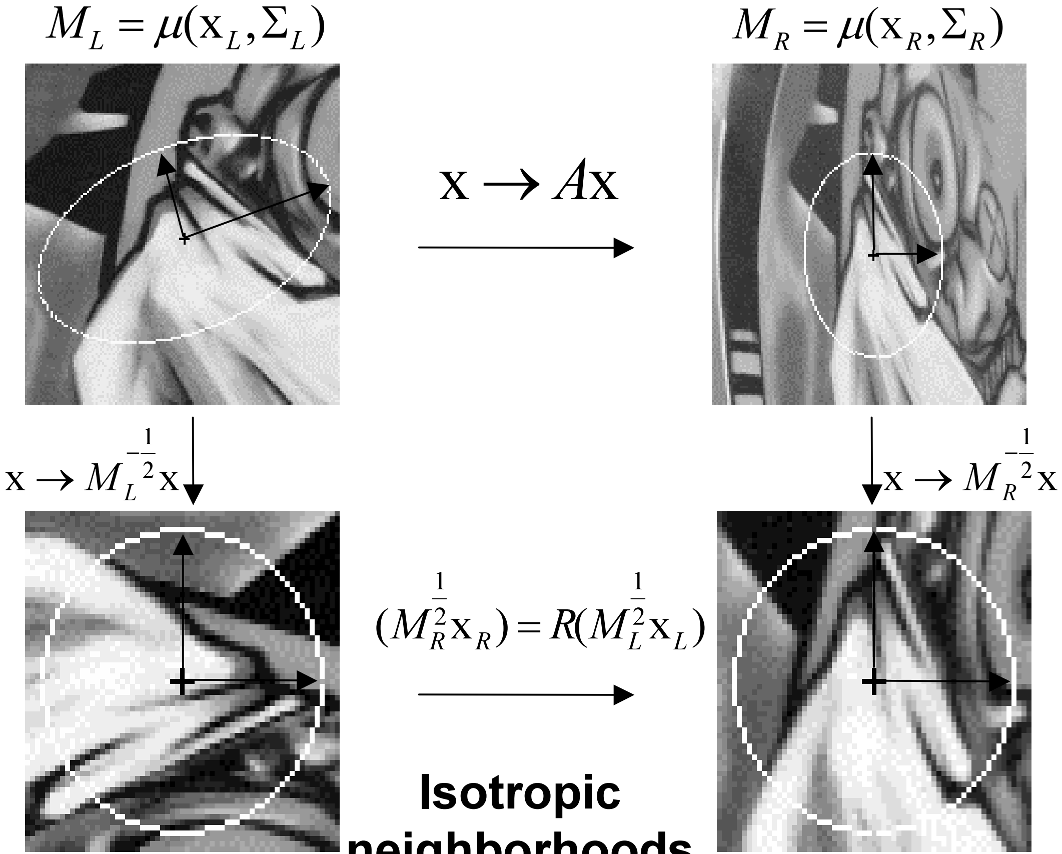
- Affine invariant regions (Tuytelaars et al.'00)
  - ellipses fitted to intensity maxima
  - parallelogram formed by interest points and edges





State of the art

- Theory for affine invariant neighborhood (Lindeberg'94)



Isotropic neighborhoods related by rotation

State of the art

- Localization & scale influence affine neighhorbood

  => affine invariant Harris points (Mikolajczyk & Schmid'02)

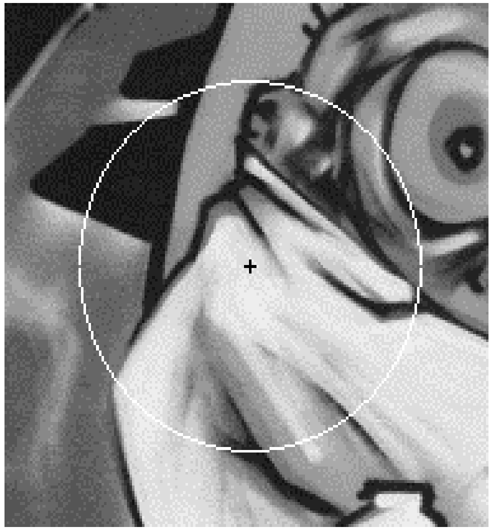
- Iterative estimation of these parameters
  1. localization – local maximum of the Harris measure
  2. scale – automatic scale selection with the Laplacian
  3. affine neighborhood – normalization with second moment matrix

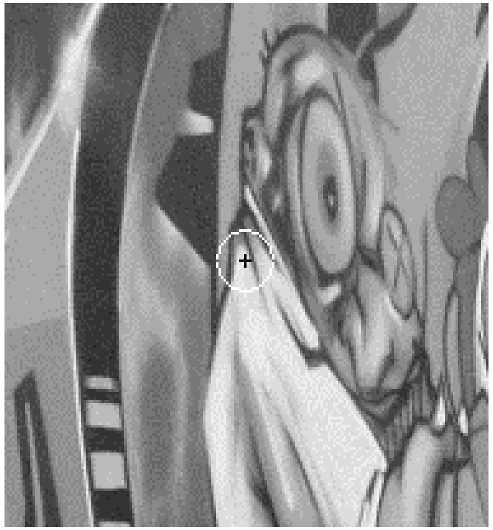
  Repeat estimation until convergence

Affine invariant Harris points

- Iterative estimation of localization, scale, neighborhood

Initial points

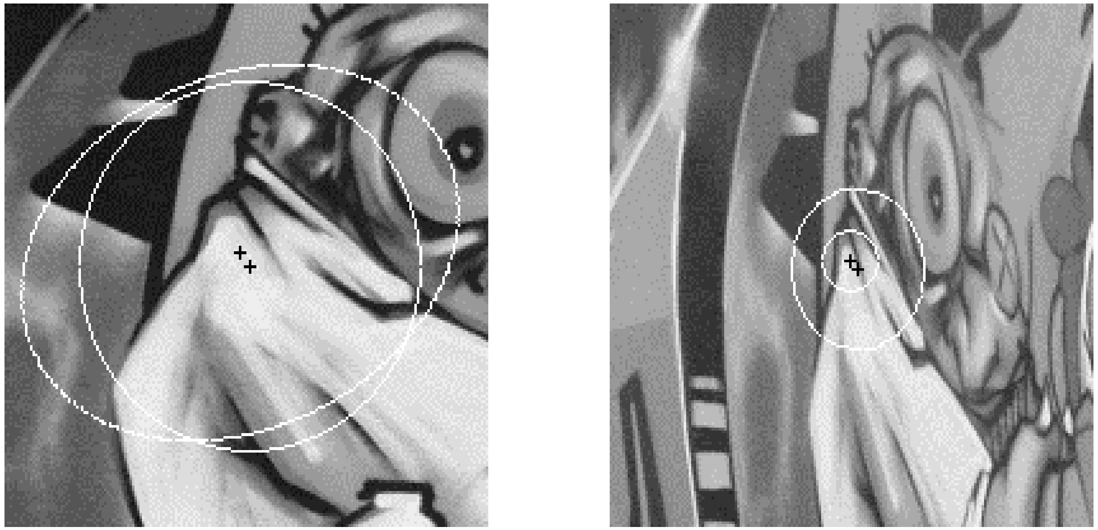




Affine invariant Harris points

- Iterative estimation of localization, scale, neighborhood

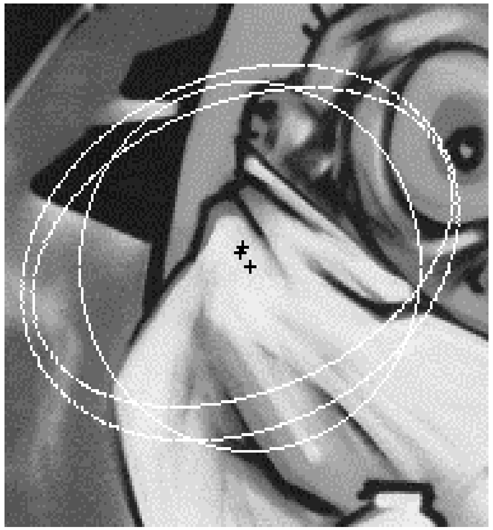
Iteration #1

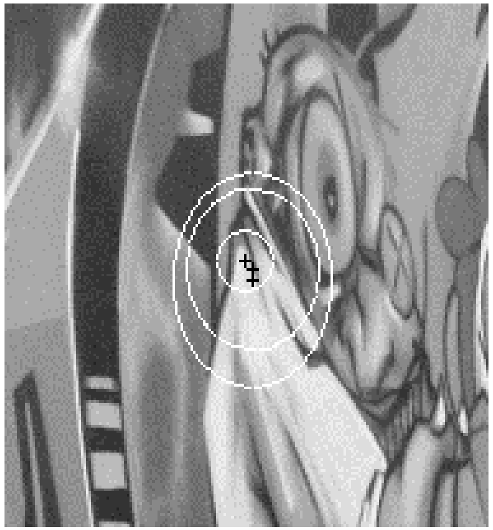


Affine invariant Harris points

- Iterative estimation of localization, scale, neighborhood

Iteration #2

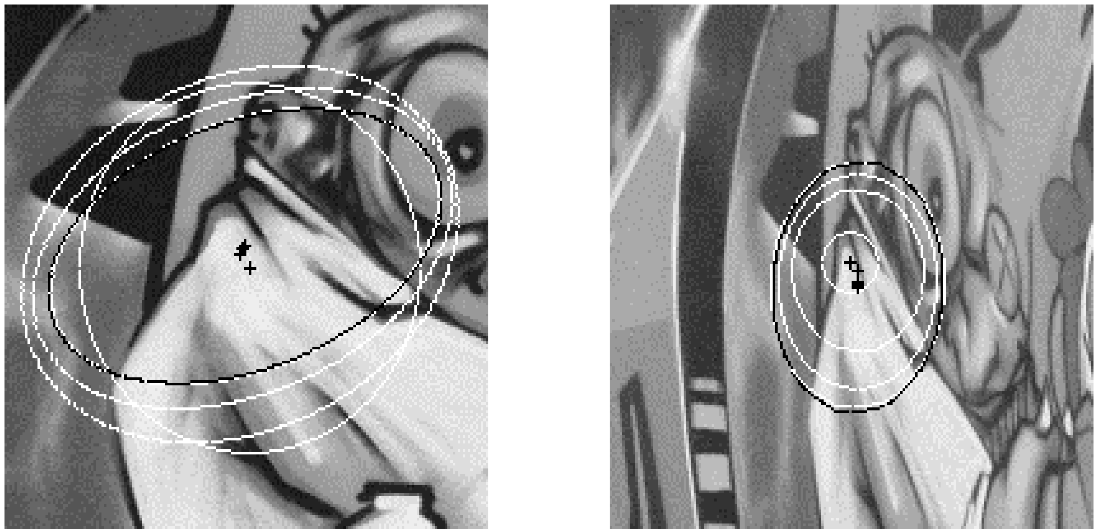




Affine invariant Harris points

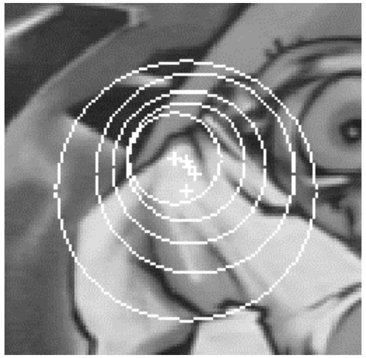
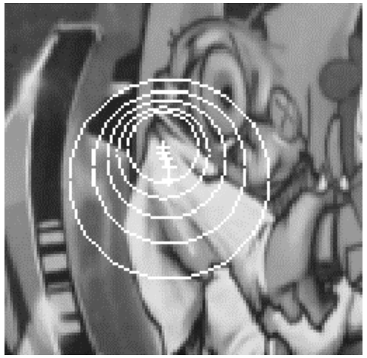
- Iterative estimation of localization, scale, neighborhood

Iteration #3, #4, ...

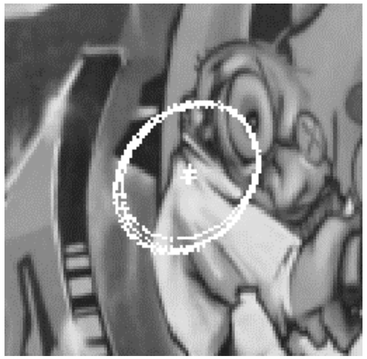
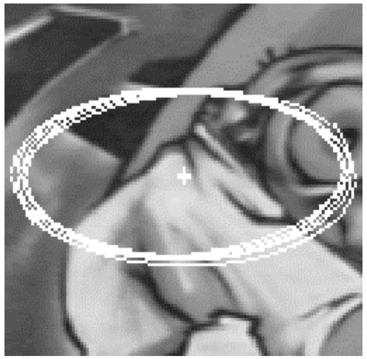


Affine invariant Harris points

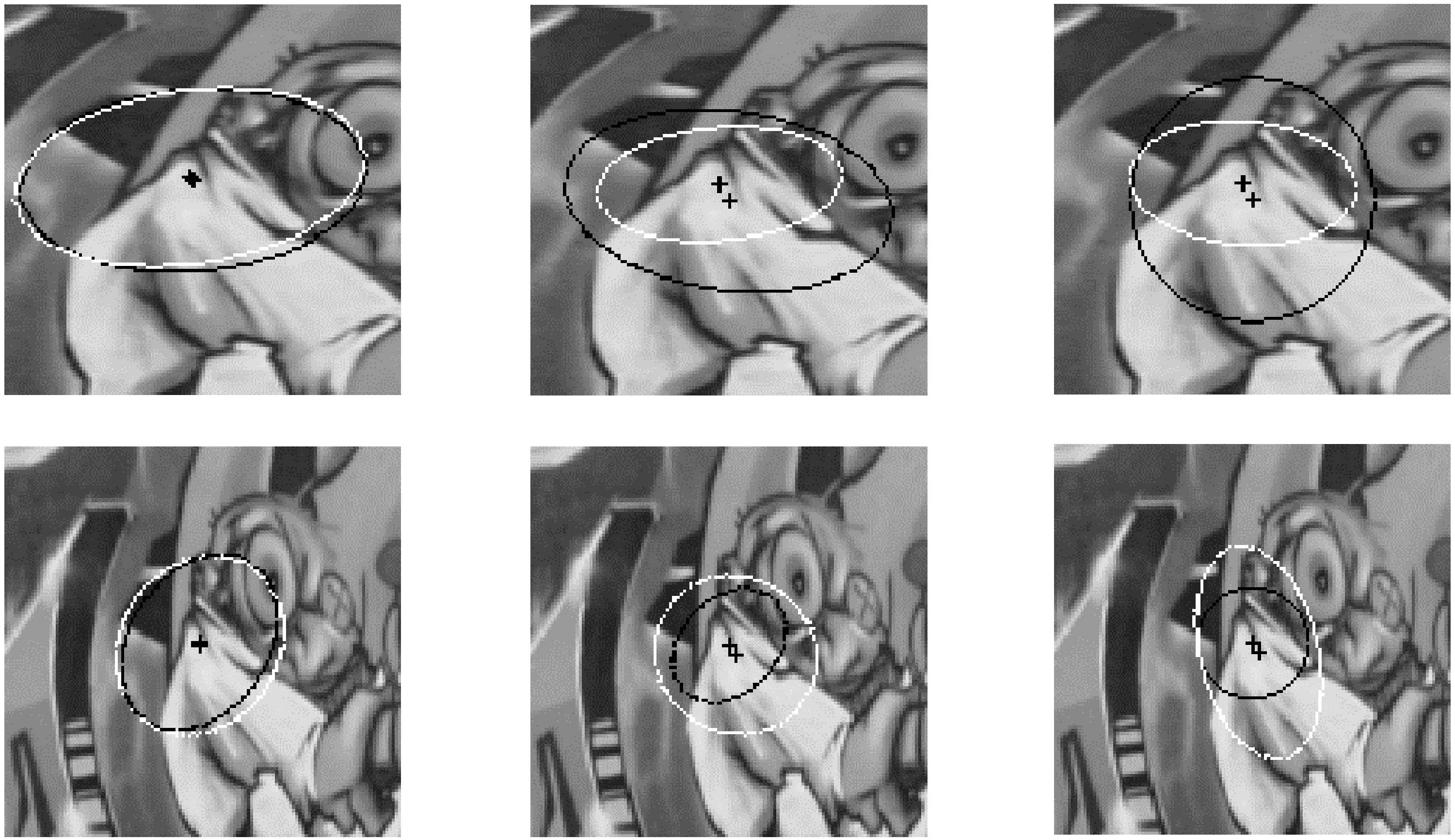
- Initialization with multi-scale interest points

- Iterative modification of location, scale and neighborhood

Affine invariant Harris points

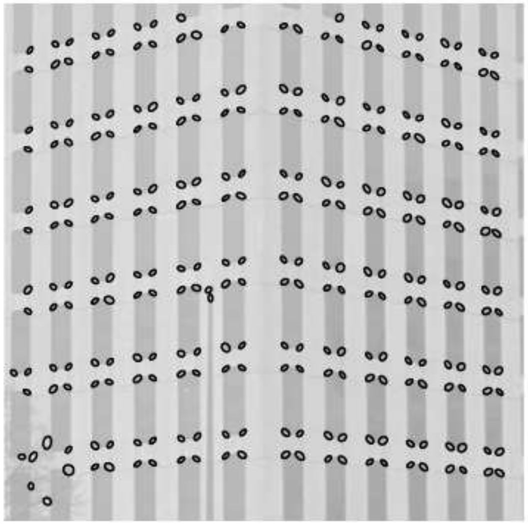
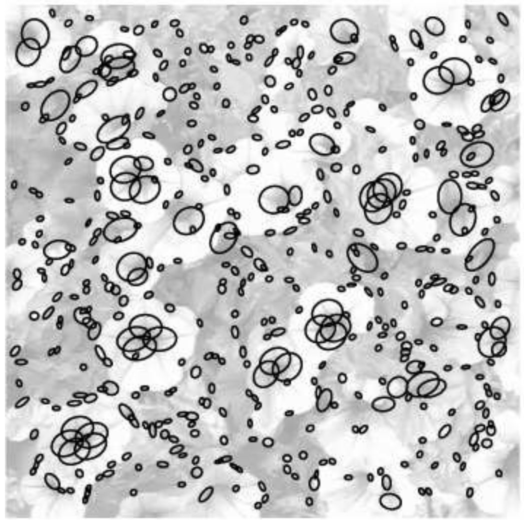


affine Harris

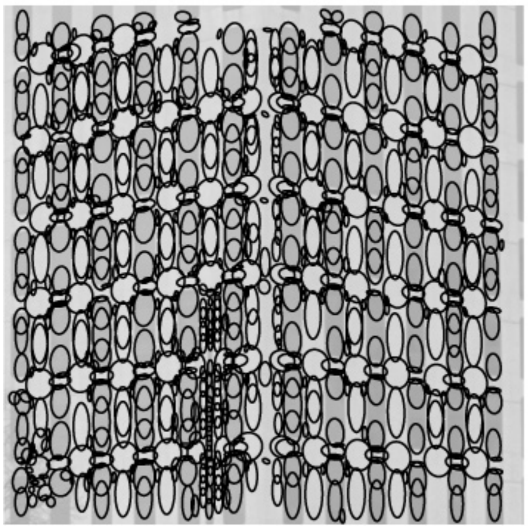
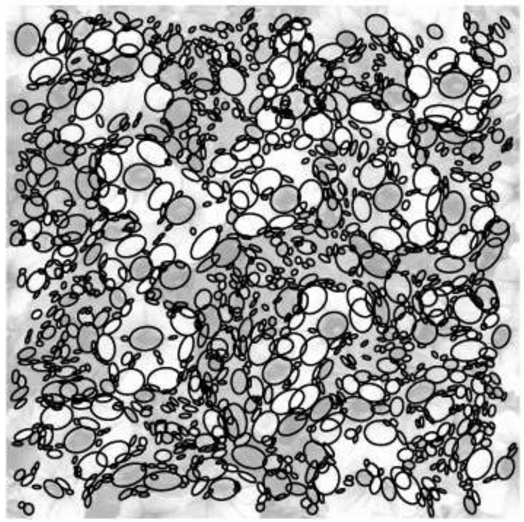
Harris-Laplace + affine regions

Harris-Laplace

Affine invariant neighborhhood

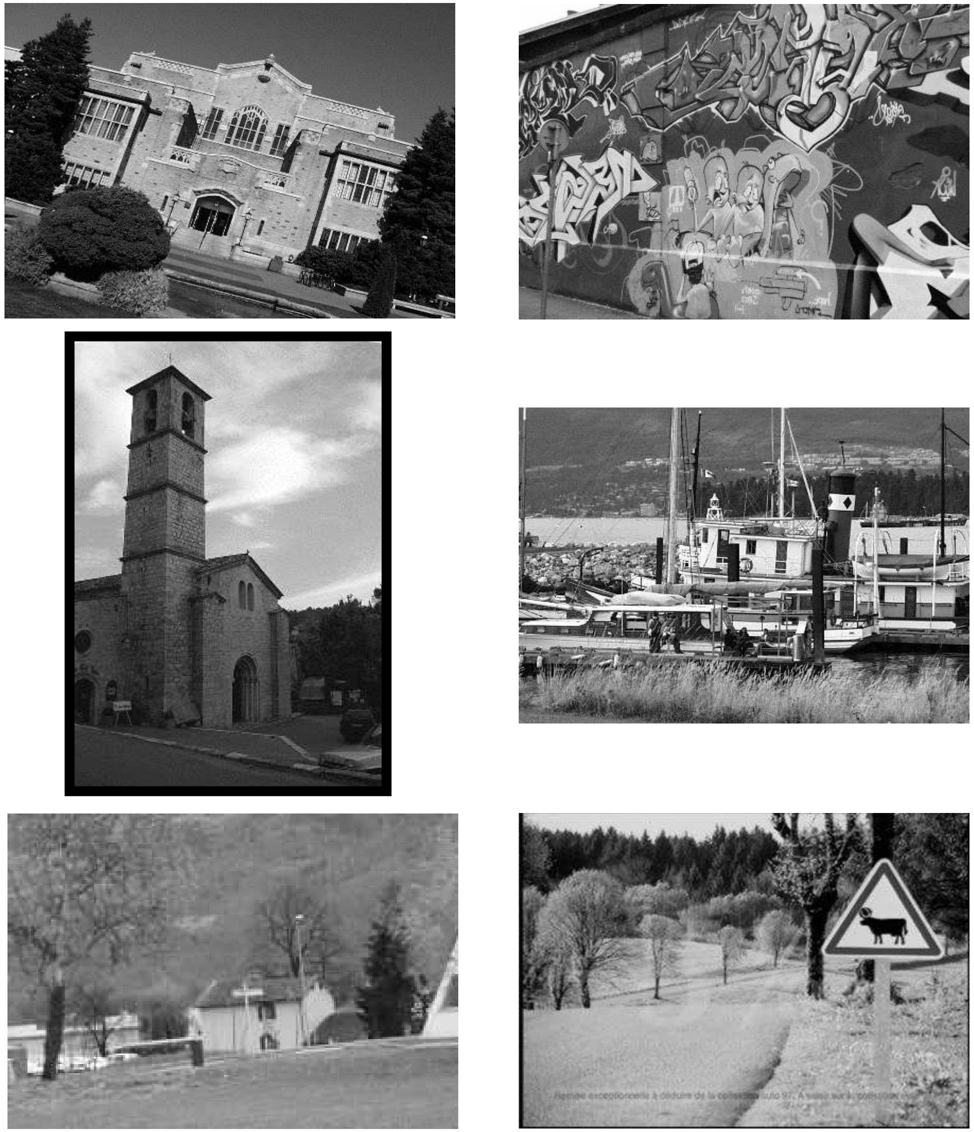
affine Harris
detector

affine Laplace
detector

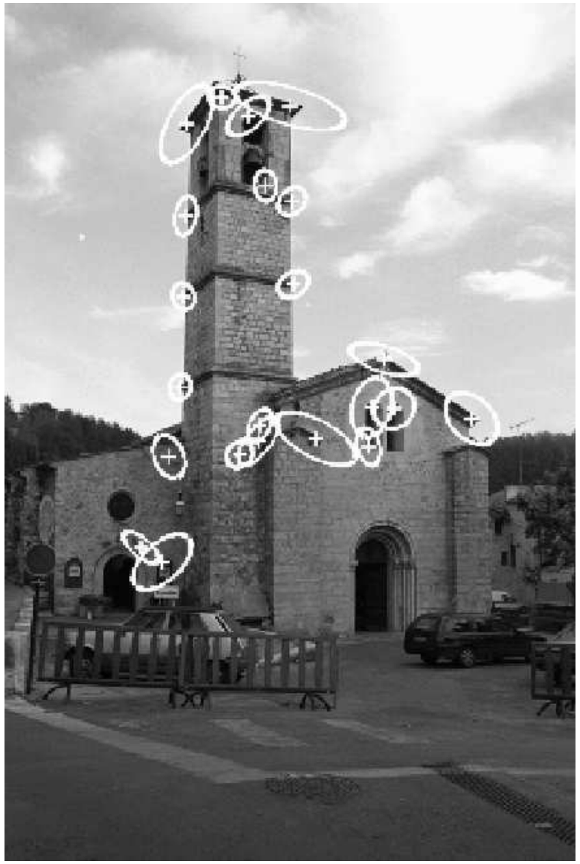
Image retrieval

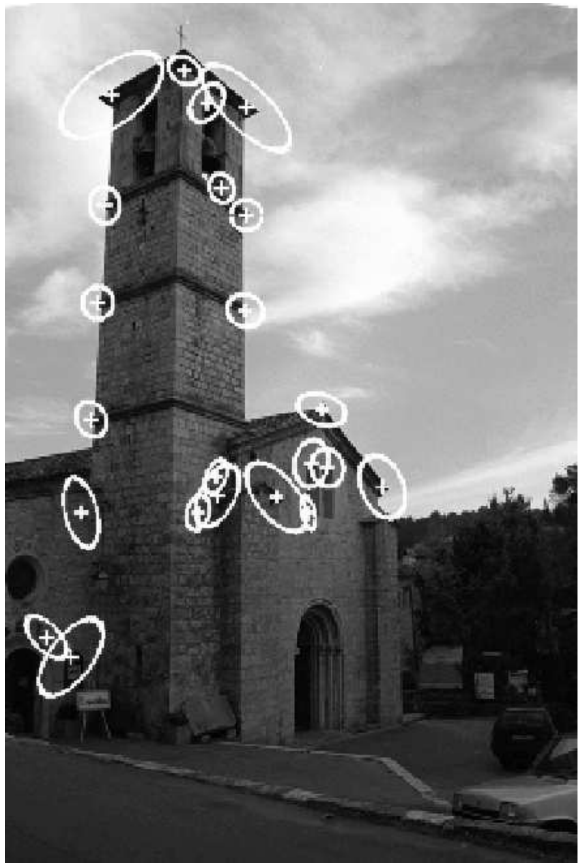
change in viewing angle



• • •
> 5000
images

Matches





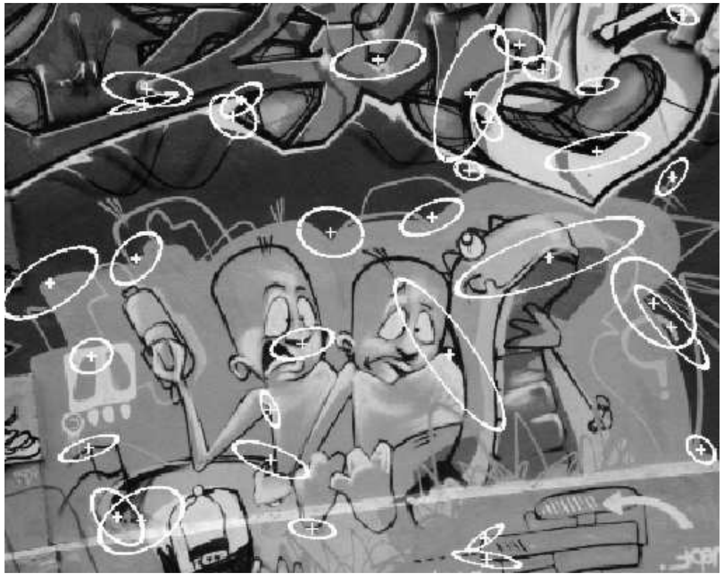
22 correct matches

Image retrieval

change in viewing angle
+ scale change

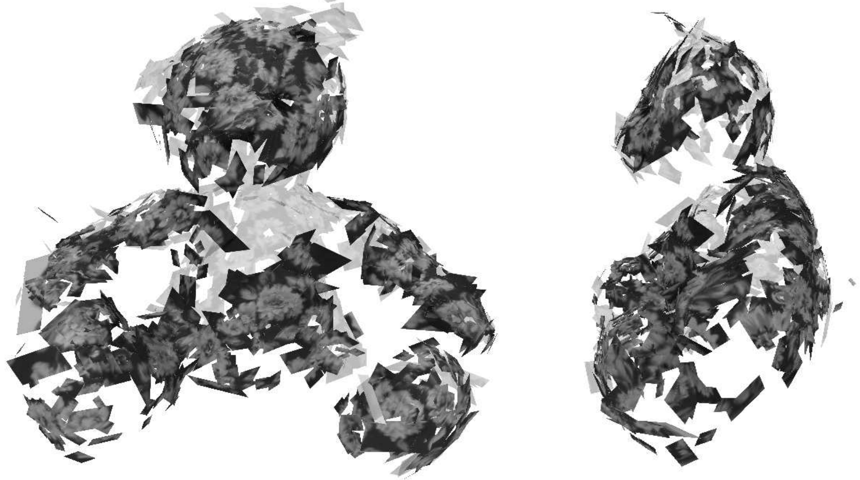
• • •
> 5000
images

Matches

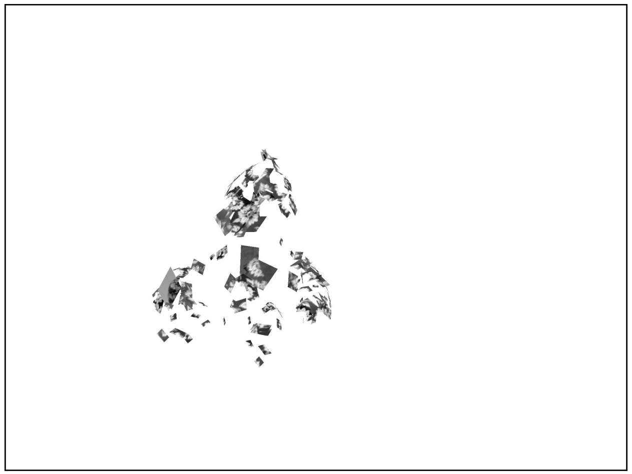


33 correct matches

3D Recognition



3D Recognition



3D object modeling and recognition using affine-invariant patches and multi-view spatial constraints,

F. Rothganger, S. Lazebnik, C. Schmid, J. Ponce,

CVPR 2003

Evaluation of interest points and descriptors

Cordelia Schmid

CVPR'03 Tutorial

Introduction

- Quantitative evaluation of interest point detectors
  - points / regions at the same relative location

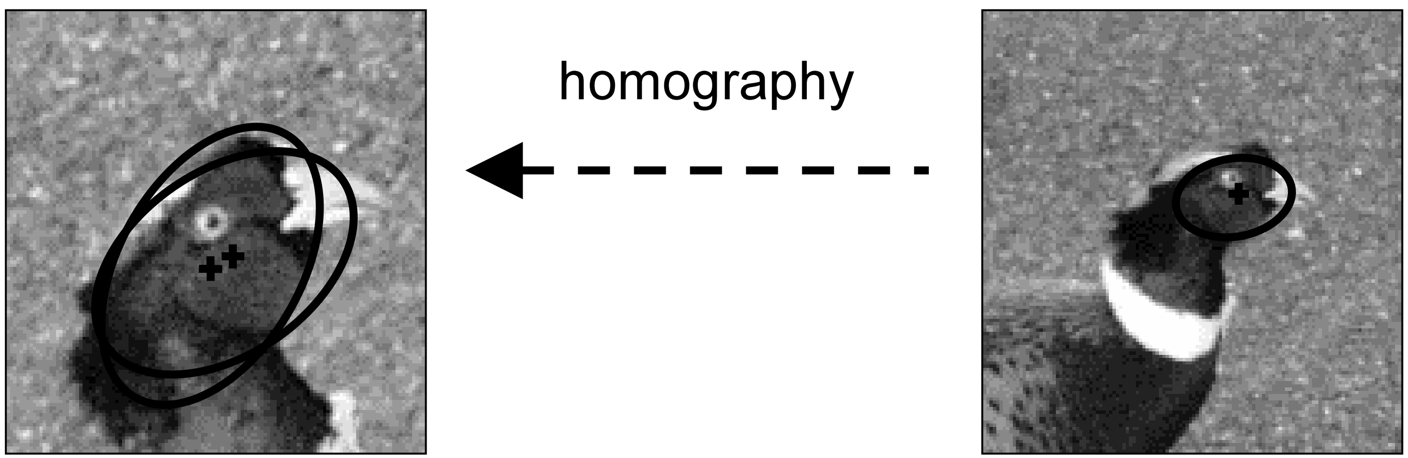
  => repeatability rate

- Quantitative evaluation of descriptors
  - distinctiveness

  => detection rate with respect to false positives

Quantitative evaluation of detectors

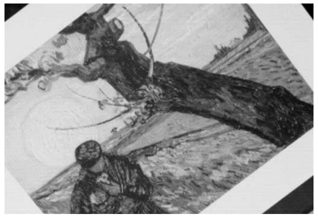
- Repeatability rate : percentage of corresponding points

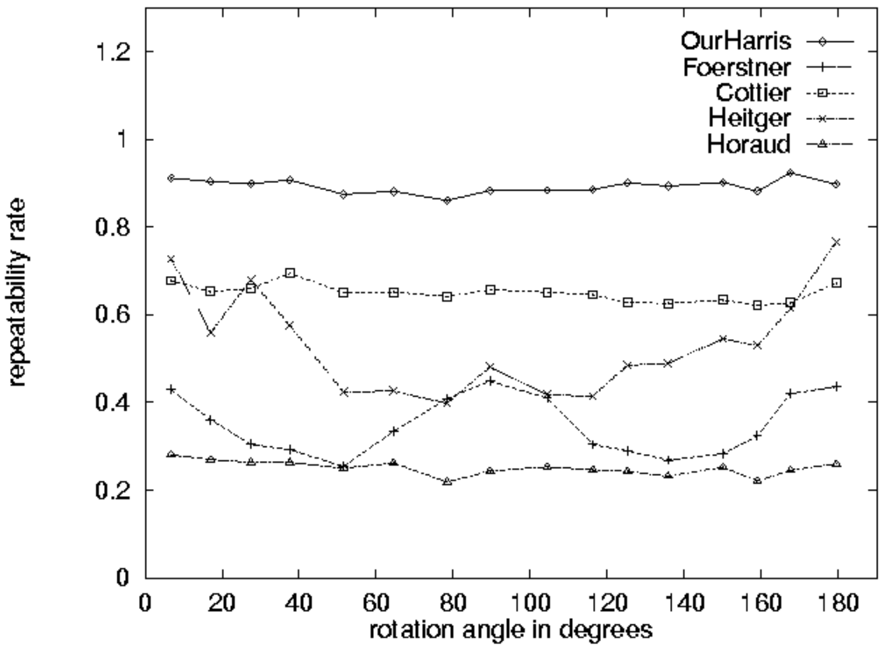


- Two points are corresponding if
  1. The location error is less than 1.5 pixel
  2. The intersection error is less than 20%

Comparison of different detectors

repeatability - image rotation

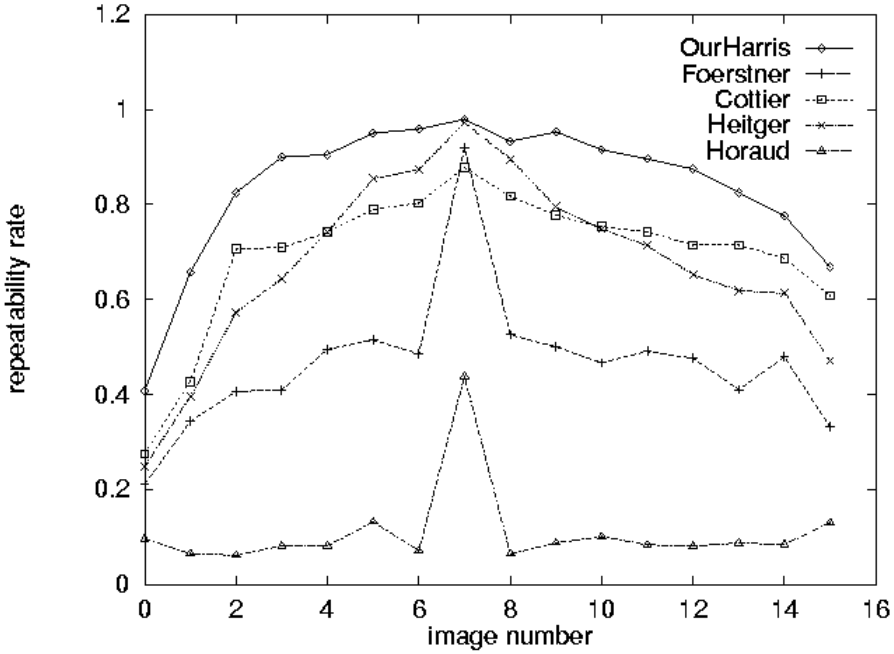




[Comparing and Evaluating Interest Points, Schmid, Mohr & Bauckhage, ICCV 98]

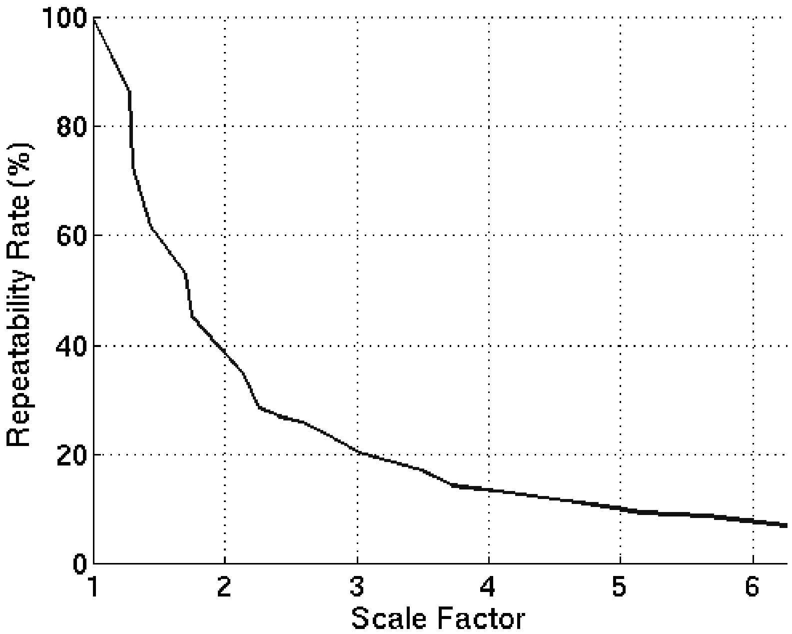
Comparison of different detectors

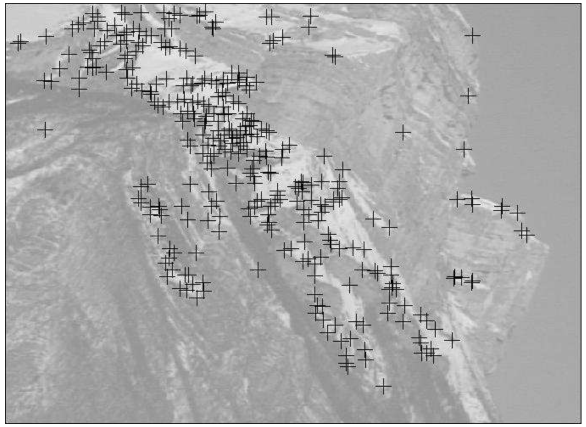
repeatability – perspective transformation

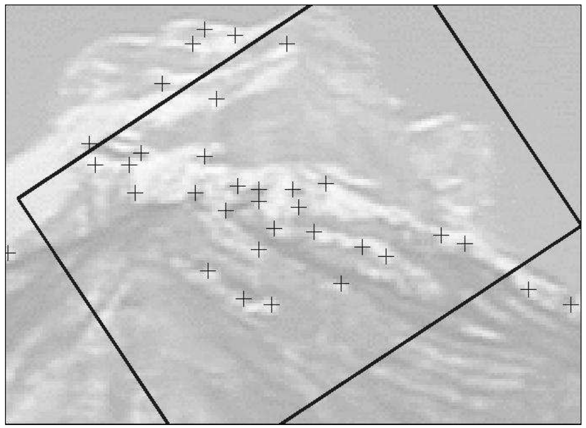


[Comparing and Evaluating Interest Points, Schmid, Mohr & Bauckhage, ICCV 98]

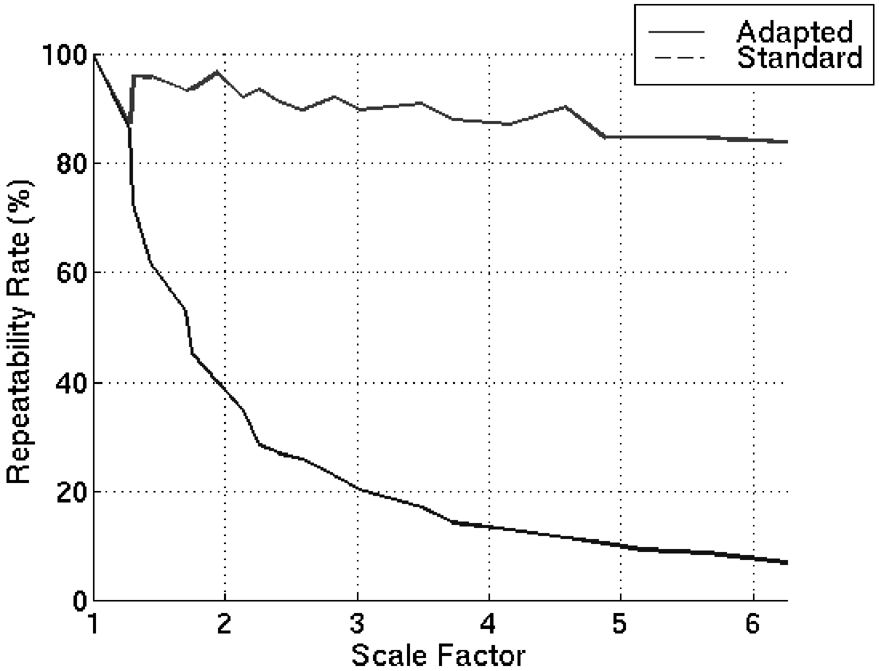
Harris detector + scale changes

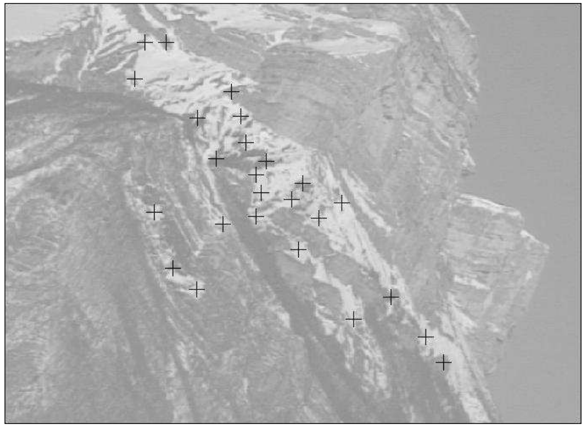


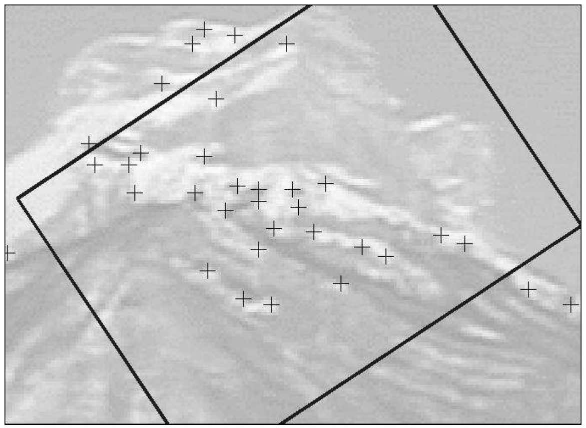




Harris detector – adaptation to scale

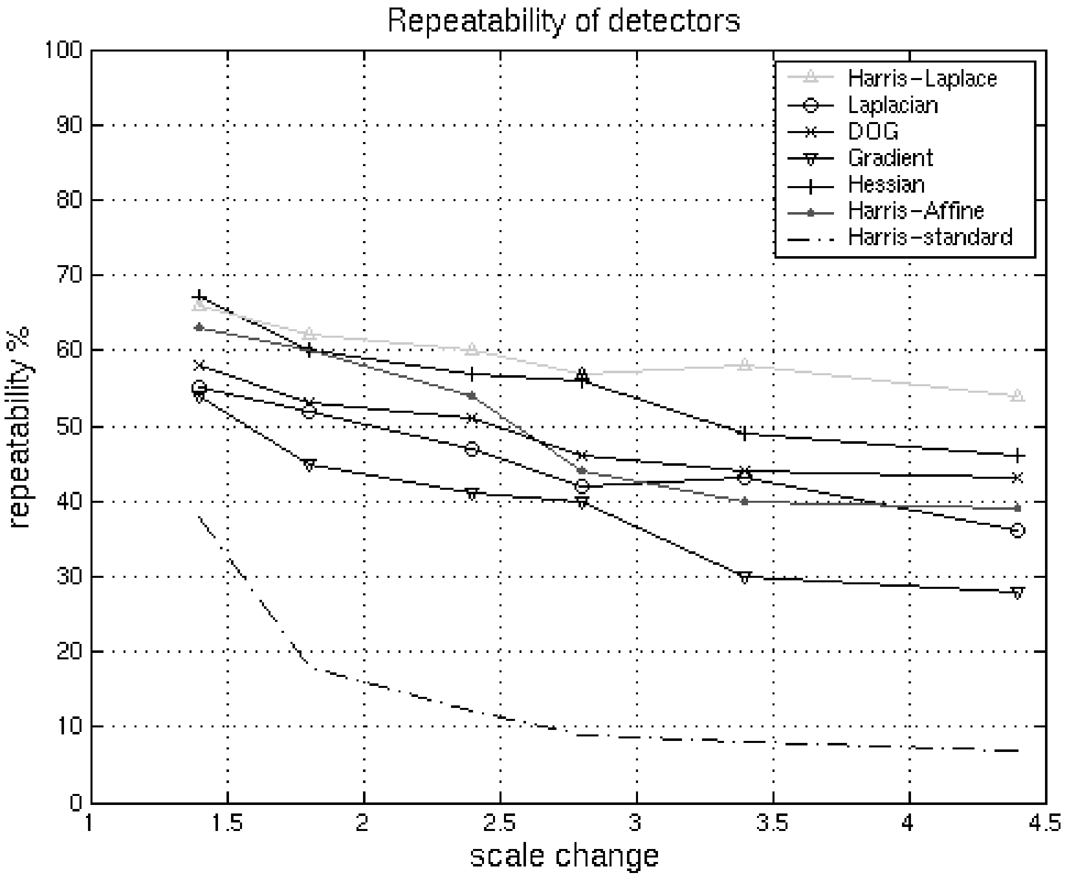






Evaluation of scale invariant detectors

repeatability – scale changes



Evaluation of affine invariant detectors

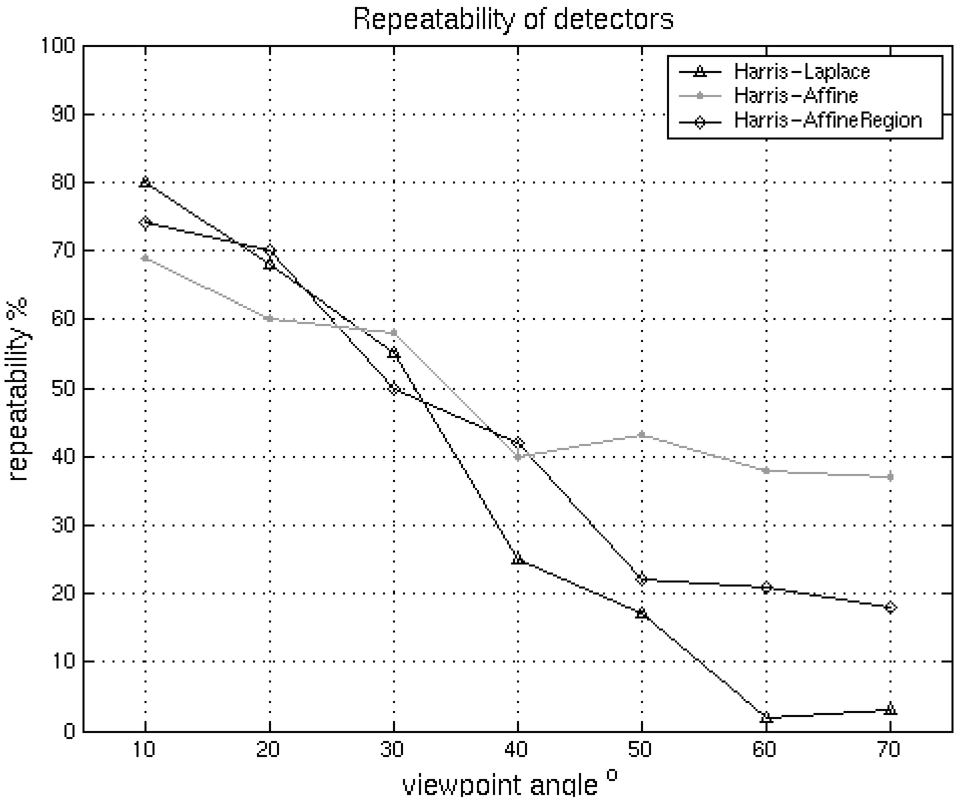
repeatability – perspective transformation

0

40

60

70



Quantitative evaluation of descriptors

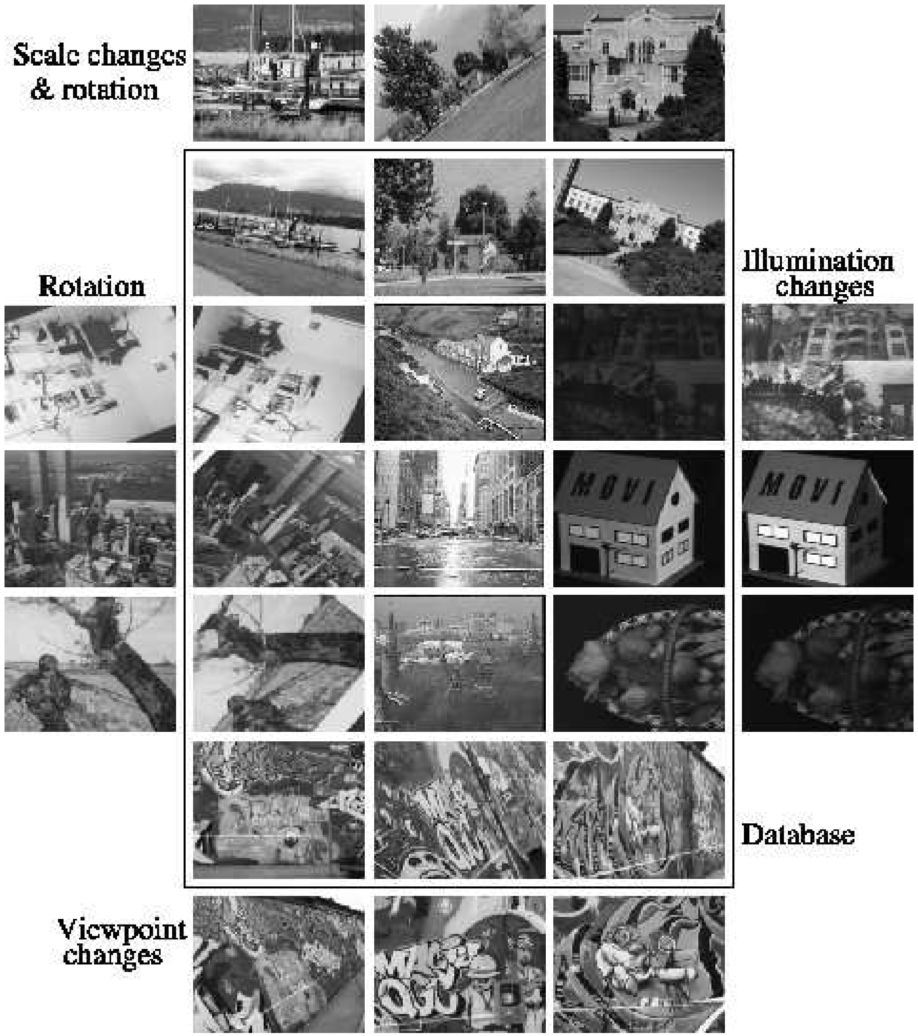
- Evaluation of different local features
  - SIFT, steerable filters, differential invariants, moment invariants, cross-correlation

- Measure : distinctiveness
  - receiver operating characteristics of
    detection rate with respect to false positives

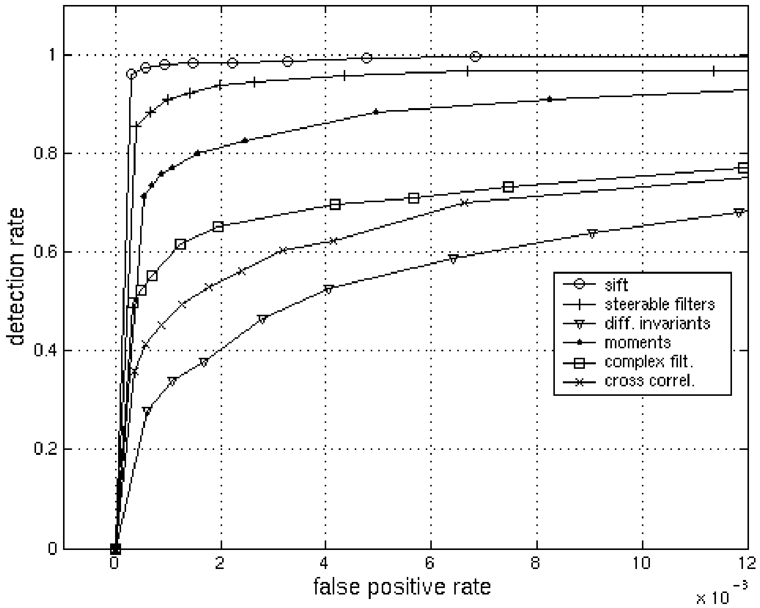
  - detection rate = correct matches / possible matches
  - false positives = false matches / (database points * query points)

[A performance evaluation of local descriptors, Mikolajczyk & Schmid, CVPR'03]

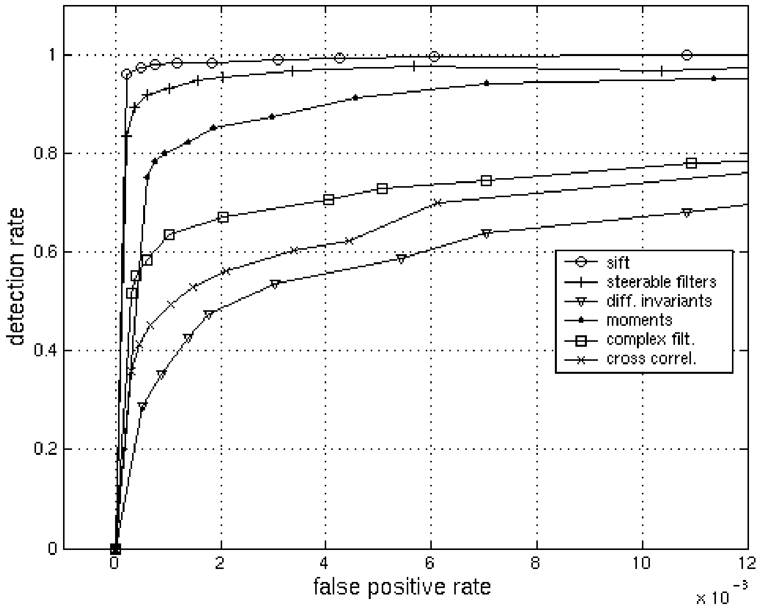
Experimental evaluation



Scale change (factor 2.5)

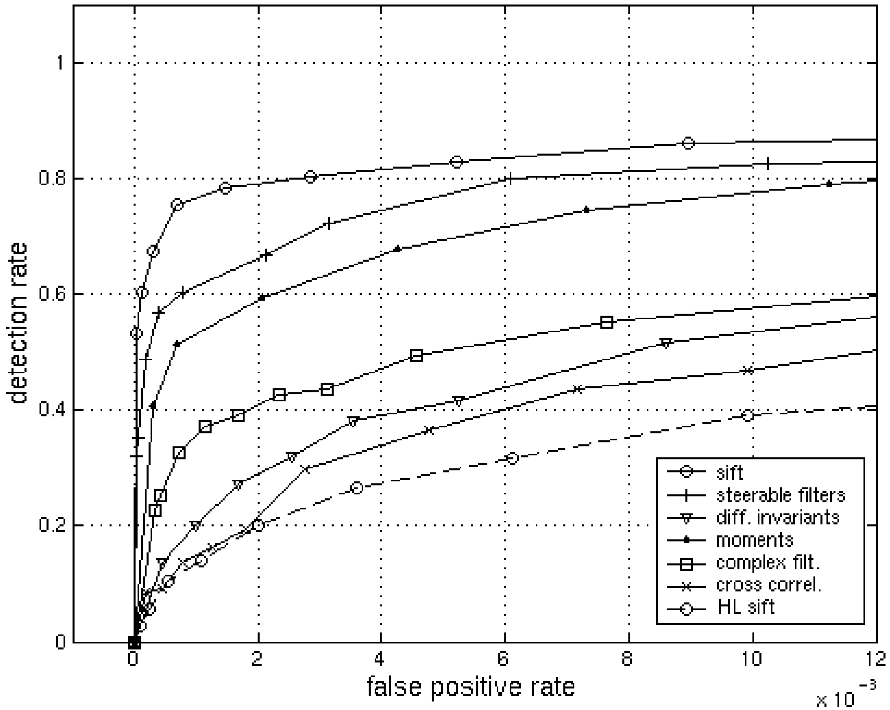


Harris-Laplace



DoG

Viewpoint change (60 degrees)



Harris-Affine (Harris-Laplace)

Descriptors - conclusion

- SIFT + steerable perform best
- Performance of the descriptor independent of the detector
- Errors due to imprecision in region estimation, localization

Today

Affine Invariant Interest points [Schmid]

Evaluation of interest points and descriptors [Schmid]

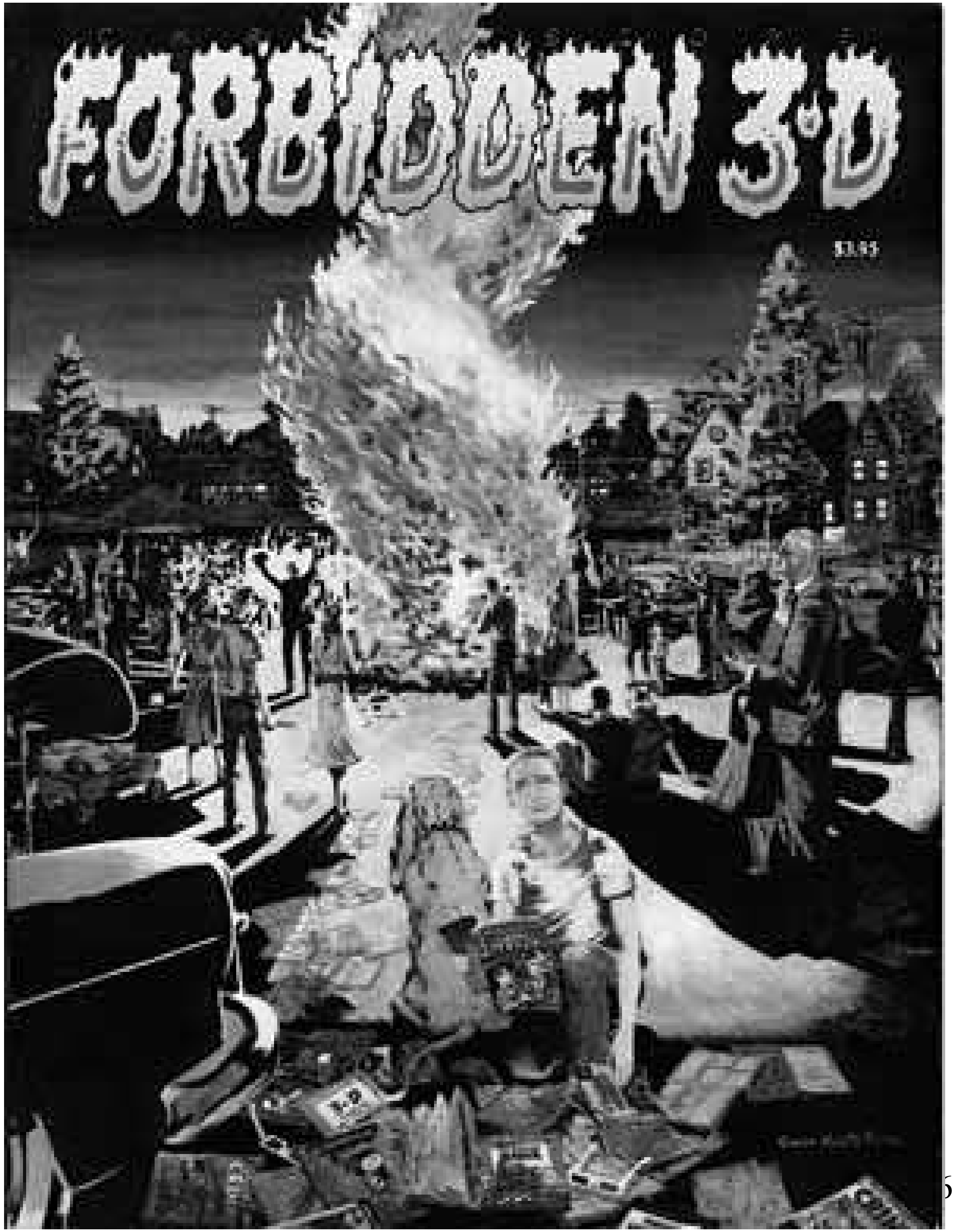
Epipolar geometry and the Essential Matrix

Multi-view geometry and 3-D

We have 2 eyes, yet we see 3-D!

Using multiple views allows inference of hidden dimension.

3-D: The hidden dimension…



Multiple views to the rescue!

How to see in 3-D

(Using geometry…)

- Find features
- Triangulate & reconstruct depth

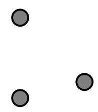
Multi-view geometry

Relate

Multi-view geometry

Relate

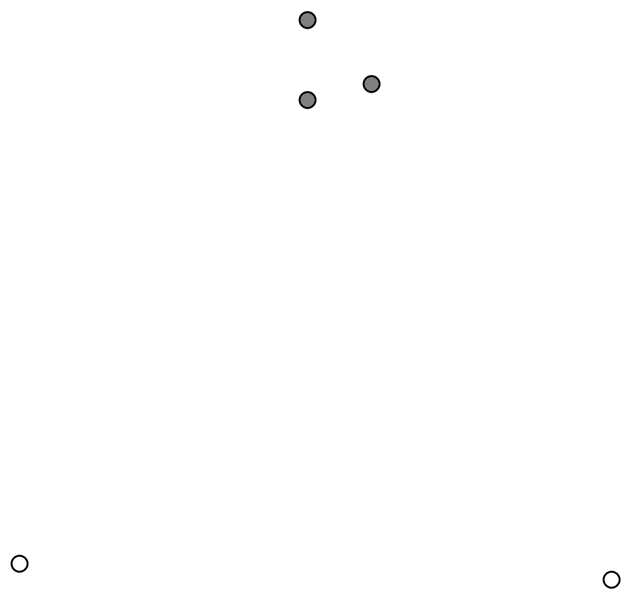
- 3-D points



Multi-view geometry

Relate

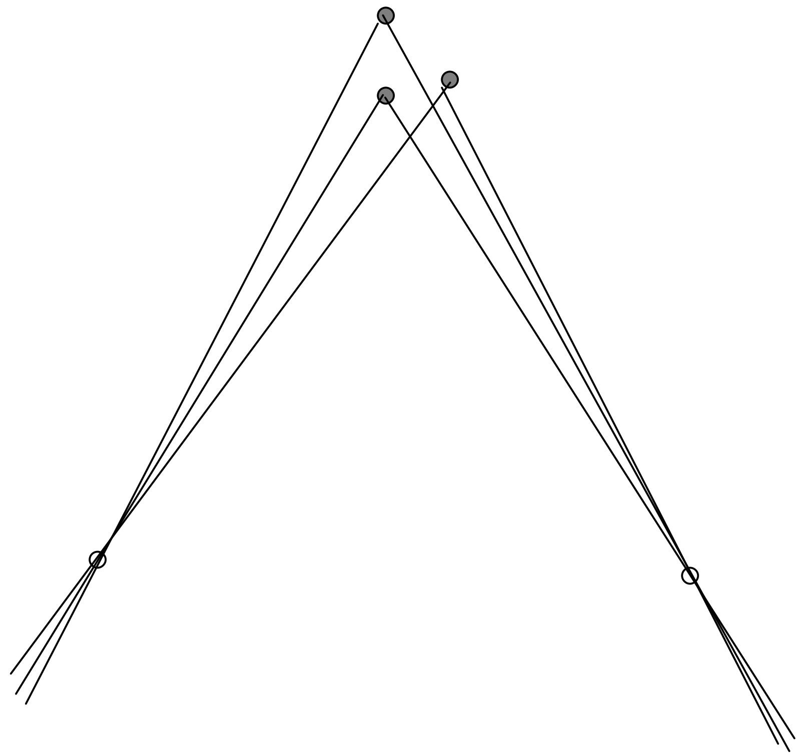
- 3-D points
- Camera centers



Multi-view geometry

Relate

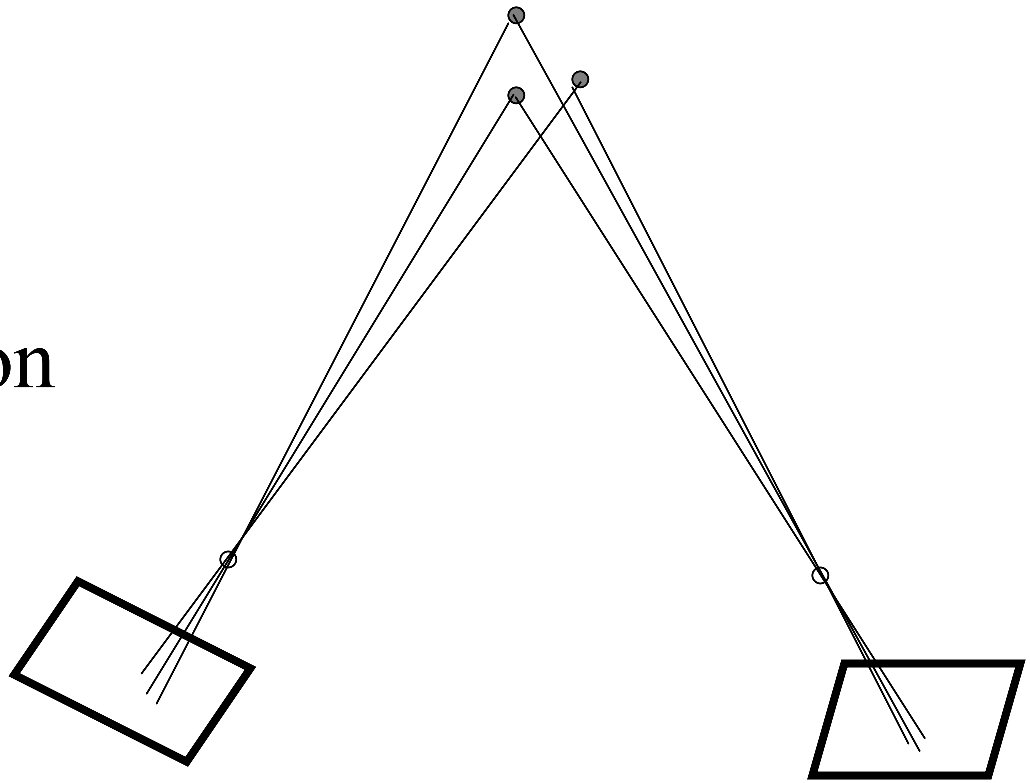
- 3-D points
- Camera centers
- Camera orientation



Multi-view geometry

Relate

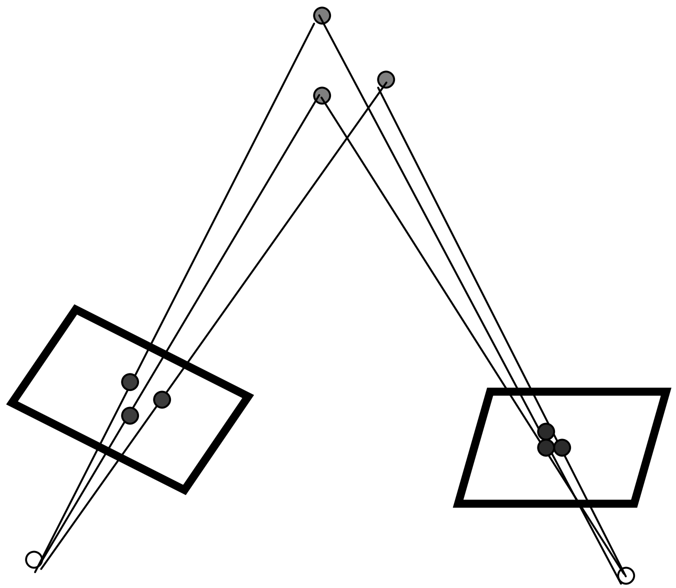
- 3-D points
- Camera centers
- Camera orientation
- Camera intrinsics



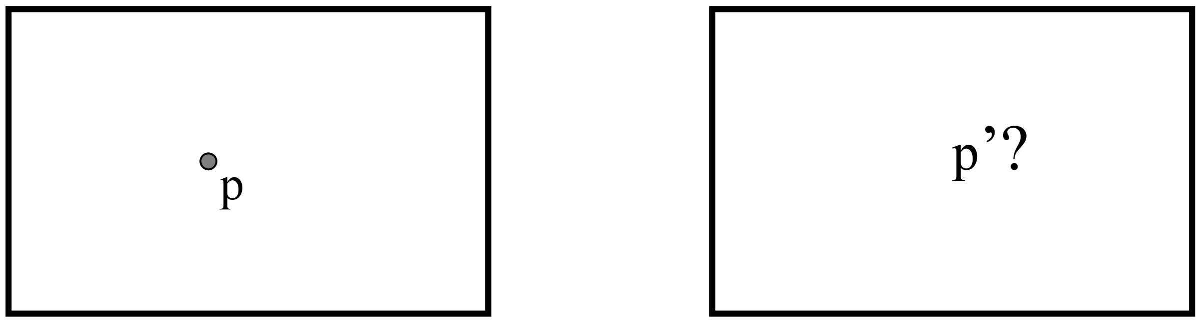
Multi-view geometry

Relate

- 3-D points
- Camera centers
- Camera orientation
- Camera intrinsics



Stereo constraints

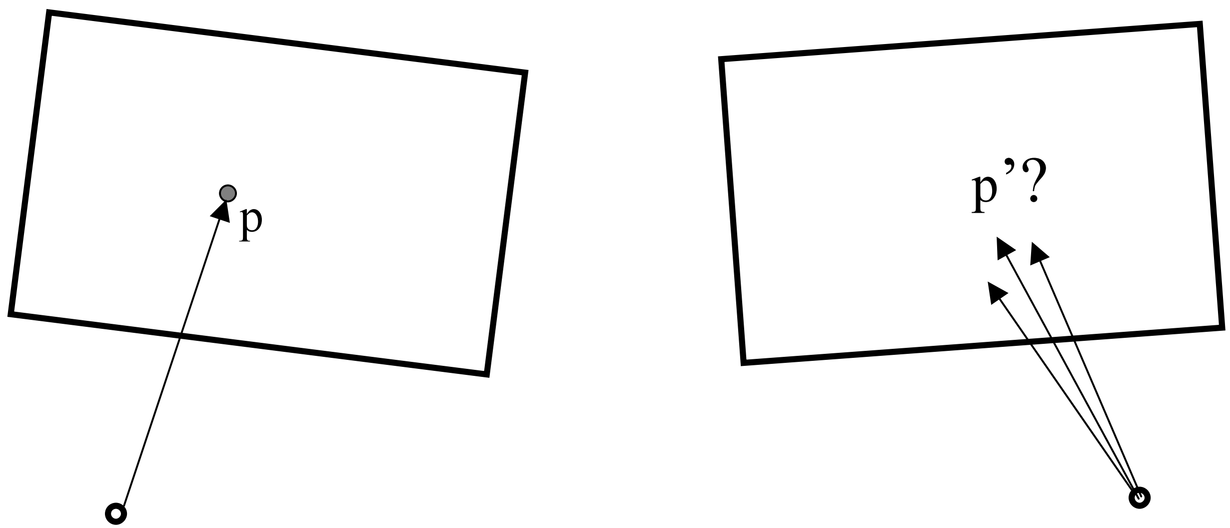


Given p in left image, where can corresponding point p' be?

Could be anywhere! Might not be same scene!

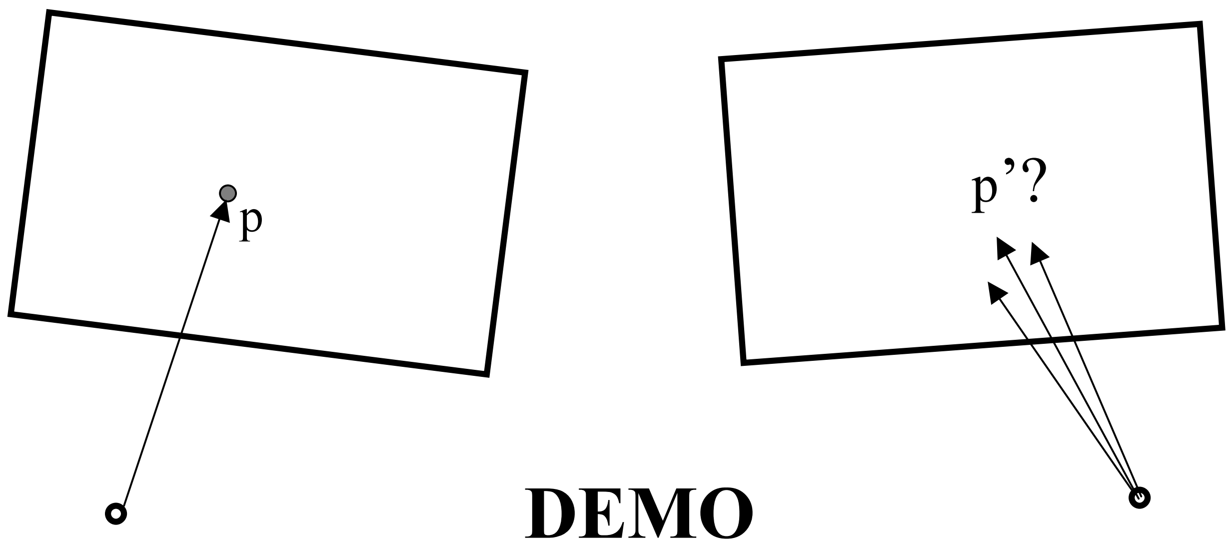
Stereo constraints

Given p in left image, where can p’ be?



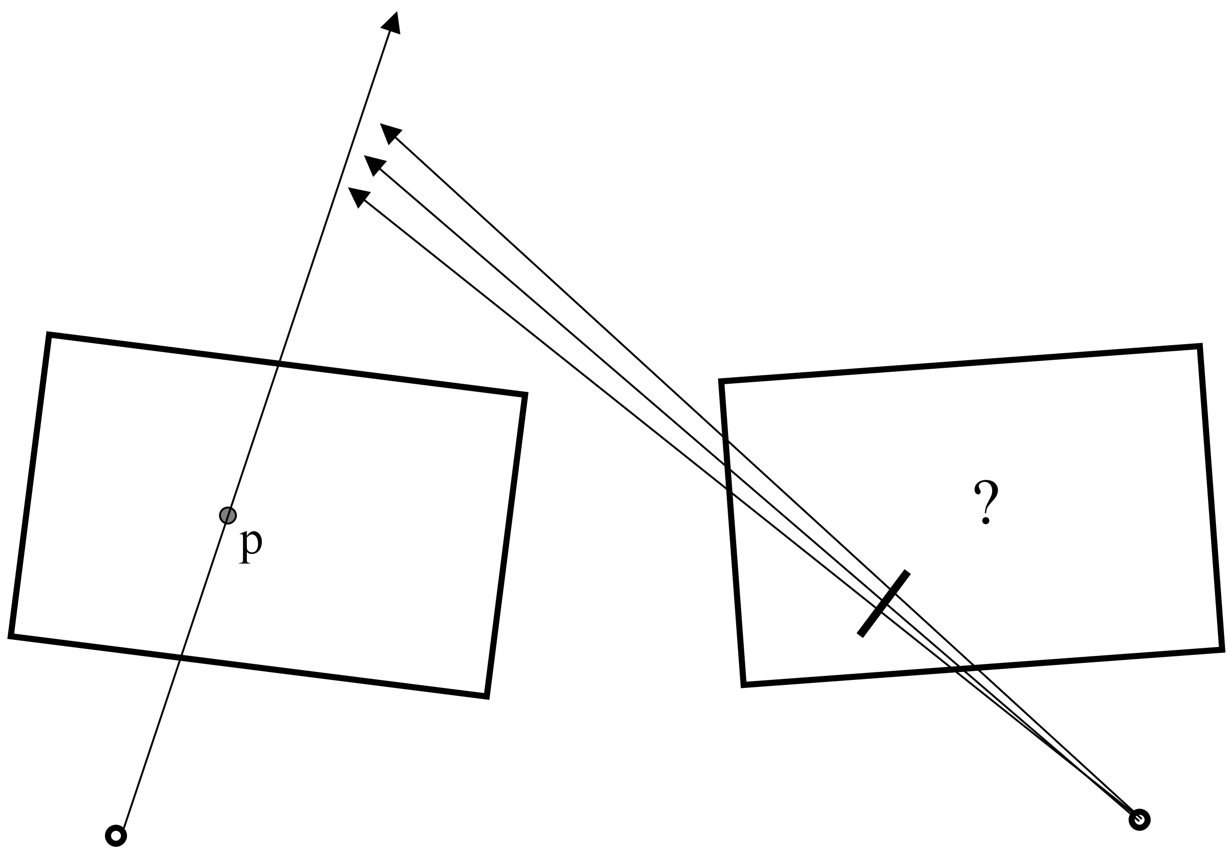
Stereo constraints

Given p in left image, where can p' be?



DEMO

Epipolar line



Epipolar constraint

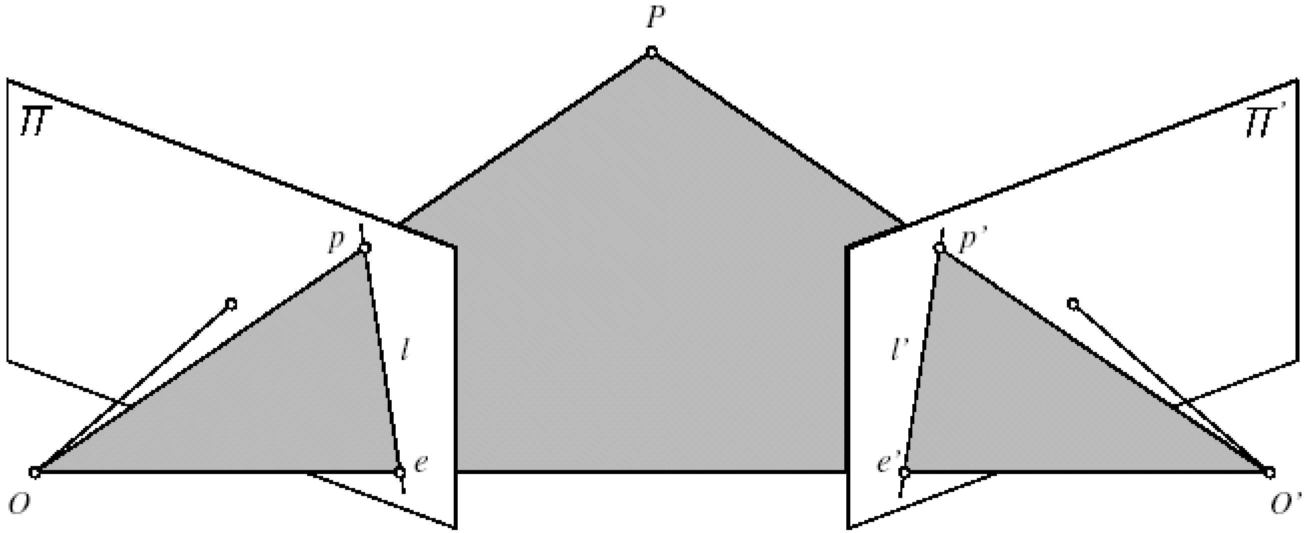


FIGURE 11.1: Epipolar geometry: the point P, the optical centers O and O' of the two cameras, and the two images p and p' of P all lie in the same plane.

All epipolar lines contain epipole, the image of other camera center.

From geometry to algebra…

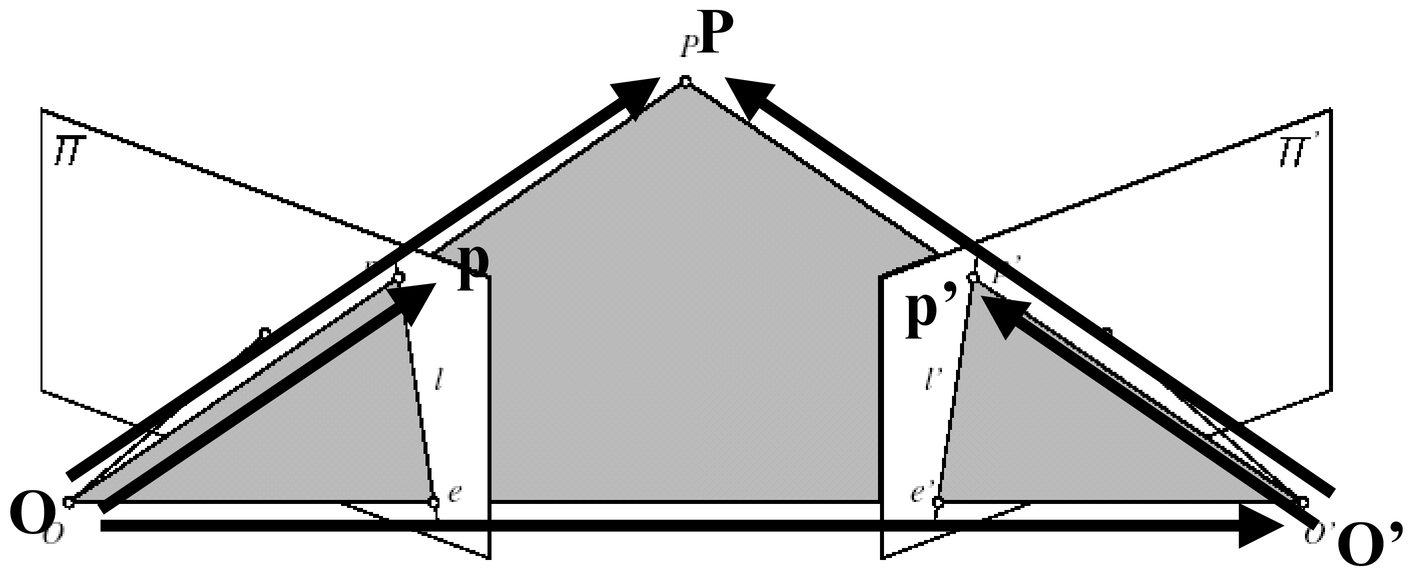
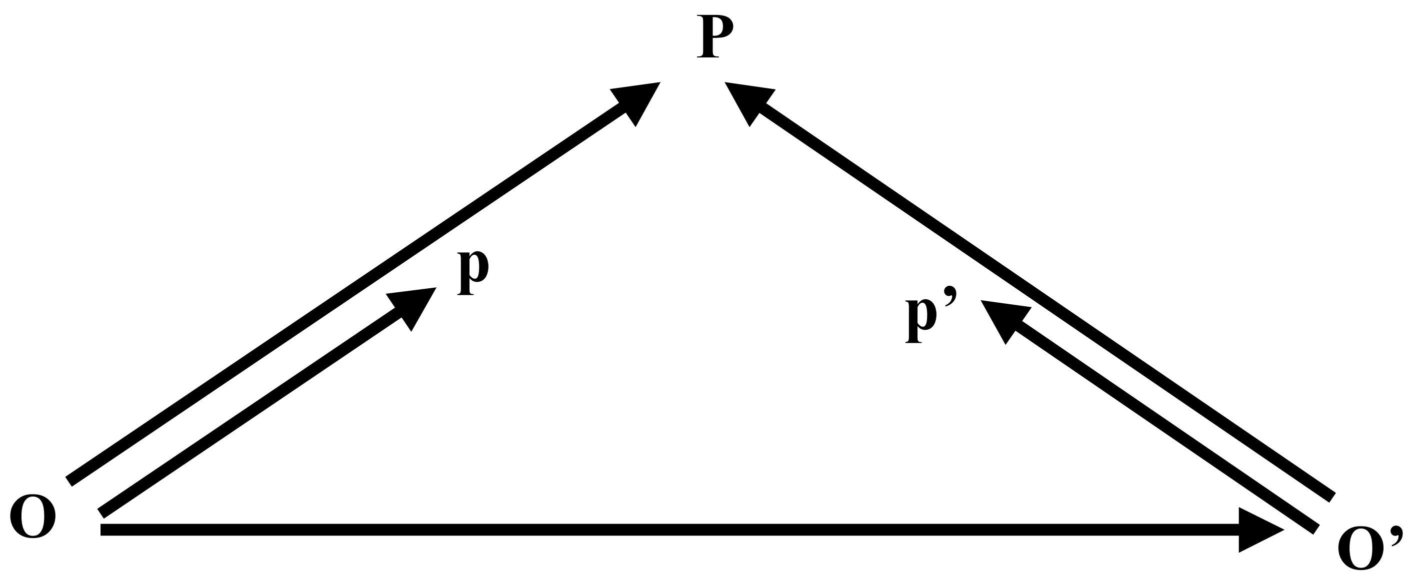


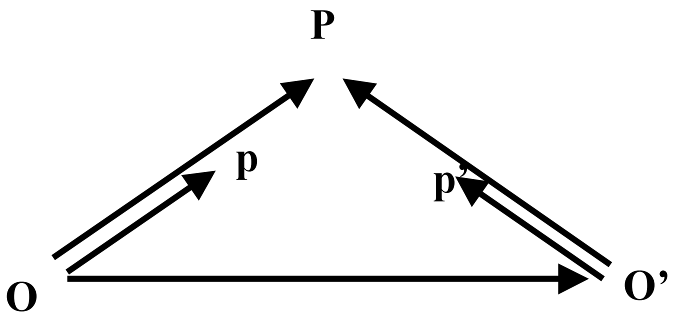
FIGURE 11.1: Epipolar geometry: the point P, the optical centers O and O' of the two cameras, and the two images p and p' of P all lie in the same plane.

From geometry to algebra…

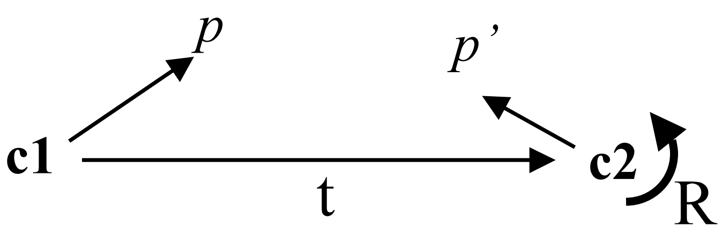


The epipolar constraint: these vectors are coplanar:

$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$



$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$



p,p' are image coordinates of P in c1 and c2…

c2 is related to c1 by rotation R and translation t

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0$$

Matrix form

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0$$

Linear constraint, should be able to express as matrix equation…

Review: matrix form of cross-product

The vector cross product also acts on two vectors and returns a third vector. Geometrically, this new vector is constructed such that its projection onto either of the two input vectors is zero.

$$\vec{a} \times \vec{b} = \begin{bmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{bmatrix}$$

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \vec{c}$$

$$\vec{a} \cdot \vec{c} = 0$$
$$\vec{b} \cdot \vec{c} = 0$$

Review: matrix form of cross-product

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \vec{c}$$

$$\vec{a} \cdot \vec{c} = 0$$

$$\vec{b} \cdot \vec{c} = 0$$

$$[a_x] = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

$$\boxed{\vec{a} \times \vec{b} = [a_x]\vec{b}}$$

Matrix form

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0$$

$$\vec{a} \times \vec{b} = [a_x]\vec{b}$$

$$p^T [t_x] \Re p' = 0$$

$$\varepsilon = [t_x] \Re$$

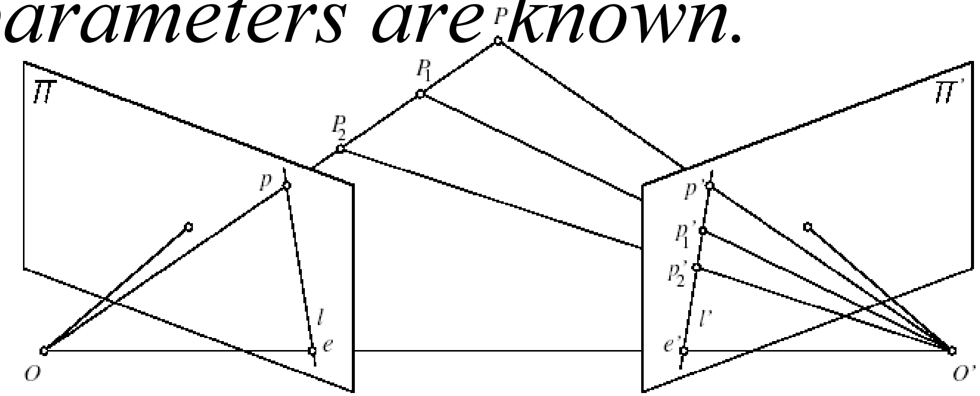
$$\boxed{\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0}$$

The Essential Matrix

Matrix that relates image of point in one camera to a second camera, given translation and rotation.

Assumes intrinsic parameters are known.

$$\varepsilon = [t_x]\Re$$

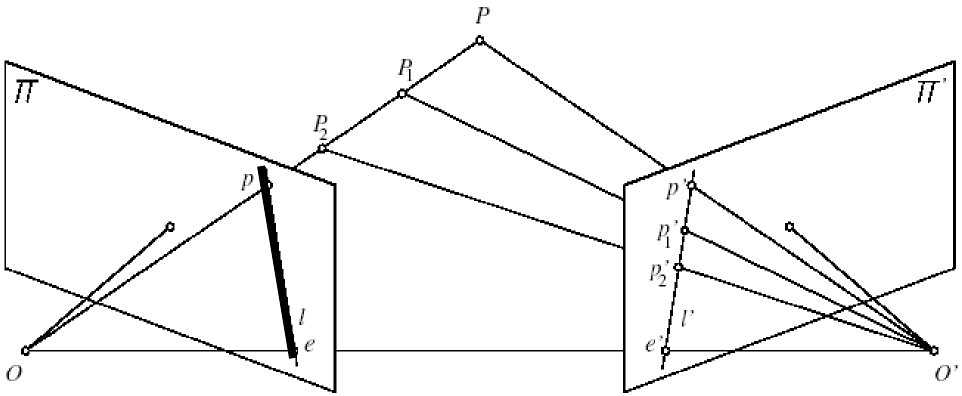


$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0$$

$$\vec{a} \times \vec{b} = [a_x]\vec{b}$$

The Essential Matrix

$\mathcal{E}p'$ is the epipolar line corresponding to p' in the left camera.



$\mathcal{E}p'$

$$au + bv + c = 0$$

$$p = (u, v, 1)^T$$

$$l = (a, b, c)^T$$

$$l \cdot p = 0$$

$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0$$

$$\mathcal{E}p' \cdot p = 0$$

Today

Affine Invariant Interest points [Schmid]

Evaluation of interest points and descriptors [Schmid]

Epipolar geometry and the Essential Matrix