Local Features (contd.)

Readings: Mikolajczyk and Schmid;
F&P Ch 10

March 6, 2008

Motivation…

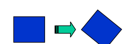
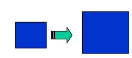
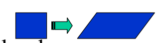
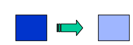
- **Feature points are used also for:**
  Image alignment (homography, fundamental matrix)
  3D reconstruction
  Motion tracking
  Object recognition
  Indexing and database retrieval
  Robot navigation
  … other

We want to:

detect *the same* interest points regardless of *image changes*

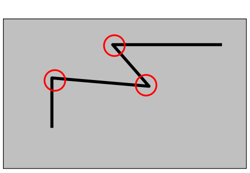
Darya Frolova, Denis Simakov
http://www.wisdom.weizmann.ac.il/~deniss/vision_spring04/files/InvariantFeatures.ppt

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 ($I \rightarrow a\,I + b$) 

Review: A Simple Example

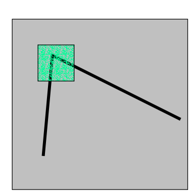
Harris corner detector



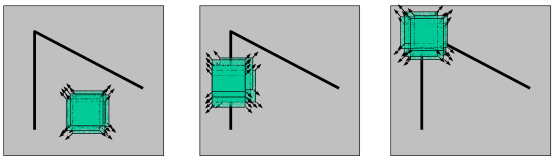
C.Harris, M.Stephens. "A Combined Corner and Edge Detector". 1988

The Basic Idea

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



Harris Detector: Basic Idea



"flat" region: no change in all directions

"edge": no change along the edge direction

"corner": significant change in all directions

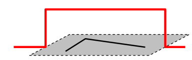
Harris Detector: Mathematics

Window-averaged change of intensity for the shift $[u,v]$:

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

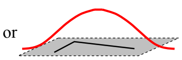
Window function — Shifted intensity — Intensity

Window function $w(x,y)$ =



1 in window, 0 outside

or



Gaussian

Harris Detector: Mathematics

Expanding E(u,v) in a 2nd order Taylor series expansion, we have,for small shifts $[u,v]$, a *bilinear* approximation:

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

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

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

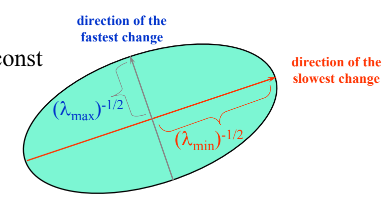
Harris Detector: Mathematics

Intensity change in shifting window: eigenvalue analysis

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

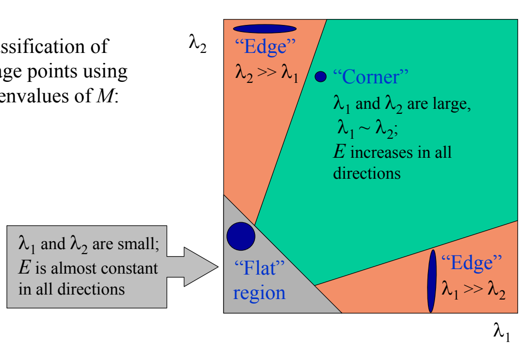
λ_1, λ_2 – eigenvalues of M

Ellipse $E(u,v)$ = const



Harris Detector: Mathematics

Classification of image points using eigenvalues of M:



Harris Detector: Mathematics

Measure of corner response:

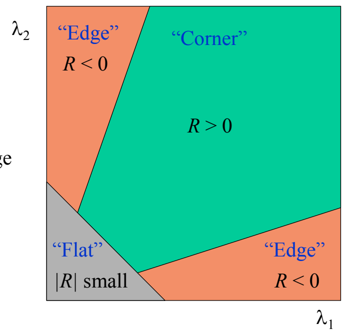
$$R = \det M - k\left(\operatorname{trace} M\right)^2$$

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

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

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



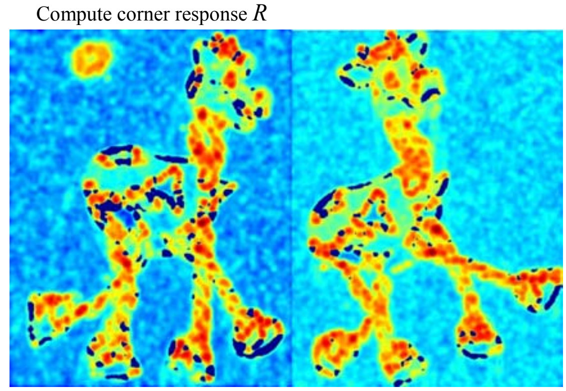
Harris Detector

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

Harris Detector: Workflow

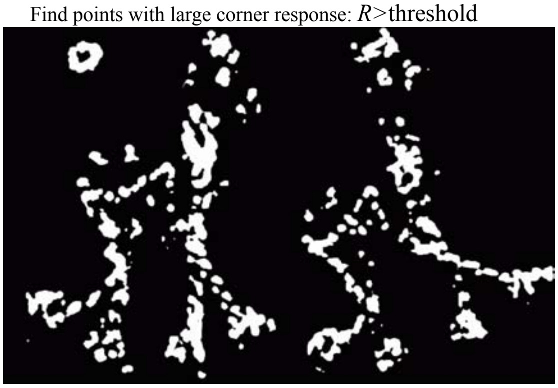
Harris Detector: Workflow

Compute corner response R



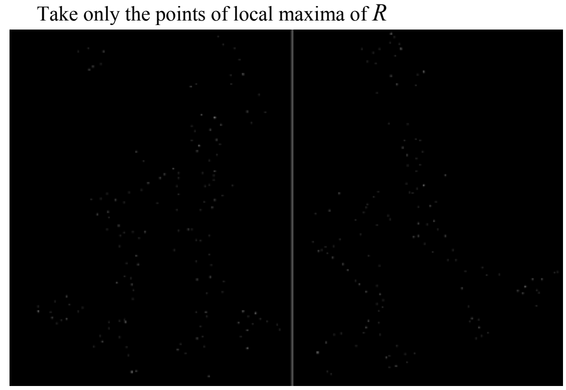
Harris Detector: Workflow

Find points with large corner response: R>threshold



Harris Detector: Workflow

Take only the points of local maxima of R



Harris Detector: Workflow

Harris Detector: Summary

- **Average intensity change in direction $[u,v]$ can be expressed as a bilinear form:**

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

- **Describe a point in terms of eigenvalues of M:** ***measure of corner response***

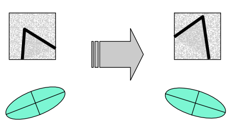
$$R = \lambda_1 \lambda_2 - k\left(\lambda_1 + \lambda_2\right)^2$$

- **A good (corner) point should have a *large intensity change* in *all directions*, i.e. R should be large positive**

Harris Detector: Some Properties

- **Rotation invariance**

Corner response R is invariant to image rotation



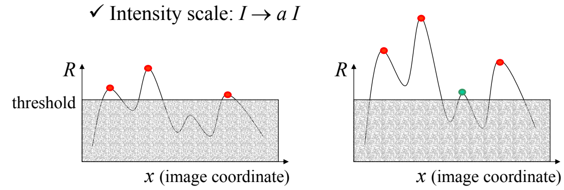
Ellipse rotates but its shape (i.e. eigenvalues) remains the same

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 scale: $I \rightarrow a\, I$

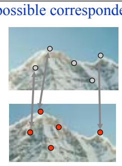


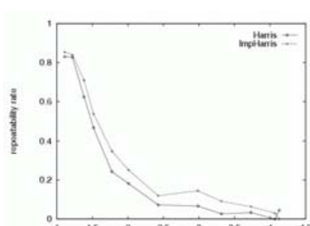
Harris Detector: Some Properties

- **Quality of Harris detector for different scale changes**

Repeatability rate:

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

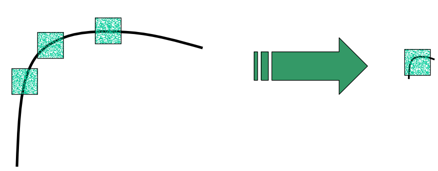




C.Schmid et.al. "Evaluation of Interest Point Detectors". IJCV 2000

Harris Detector: Some Properties

- **Not invariant to *image scale*!**

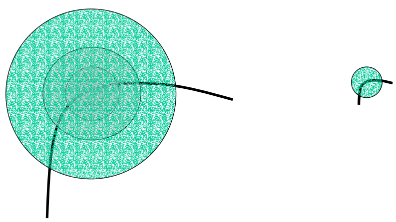


All points will be classified as edges

Corner !

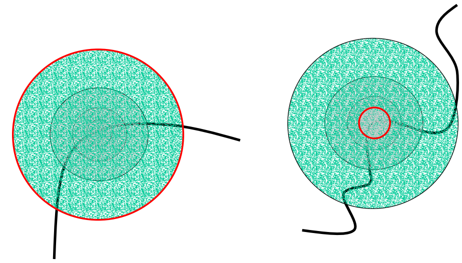
Scale Invariant Detection

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



Scale Invariant Detection

- **The problem: how do we choose corresponding circles** *independently* **in each image?**

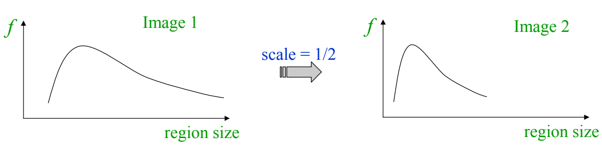


Scale Invariant Detection

- **Solution:**
  - Design a function on the region (circle), which is "scale invariant" (the same for corresponding regions, even if they are at different scales)

    Example: average intensity. For corresponding regions (even of different sizes) it will be the same.

  - For a point in one image, we can consider it as a function of region size (circle radius)



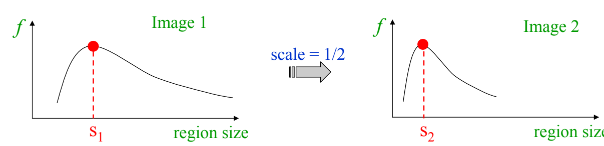
Scale Invariant Detection

- **Common approach:**

Take a local maximum of this function

Observation: region size, for which the maximum is achieved, should be *invariant* to image scale.

Important: this scale invariant region size is found in each image independently!



Scale Invariant Detection

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

Kernels:

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

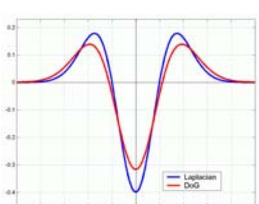
(Laplacian)

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

(Difference of Gaussians)

where Gaussian

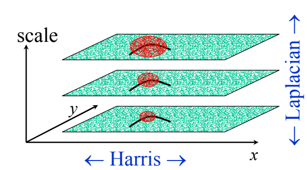
$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



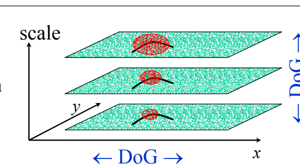
Note: both kernels are invariant to *scale* and *rotation*

Scale Invariant Detectors

- 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



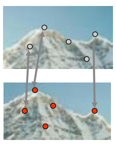
[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001
[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". Accepted to IJCV 2004

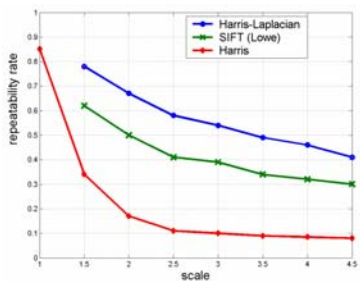
Scale Invariant Detectors

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

Repeatability rate:

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





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

Scale Invariant Detection: Summary

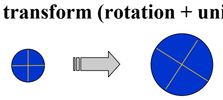
- **Given:** **two images of the same scene with a large *scale difference* between them**
- **Goal:** **find *the same* interest points *independently* in each image**
- **Solution:** **search for *maxima* of suitable functions in *scale* and in *space* (over the image)**

Methods:

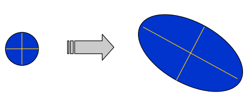
1. Harris-Laplacian [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. SIFT [Lowe]: maximize Difference of Gaussians over scale and space

Affine Invariant Detection

- **Above we considered:**
  Similarity transform (rotation + uniform scale)

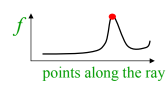


- Now we go on to:
  Affine transform (rotation + non-uniform scale)



Affine Invariant Detection

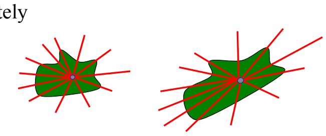
- **Take a local intensity extremum as initial point**
- **Go along every ray starting from this point and stop when extremum of function f is reached**



$$f(t) = \frac{|I(t) - I_0|}{\frac{1}{t}\int_o^t |I(t) - I_0|\,dt}$$

- We will obtain approximately corresponding regions

Remark: we search for scale in every direction



T.Tuytelaars, L.V.Gool. "Wide Baseline Stereo Matching Based on Local, Affinely Invariant Regions". BMVC 2000.

Affine Invariant Detection

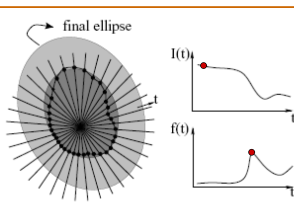
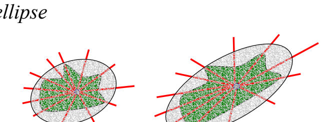


Figure 2: *The intensity along "rays" emanating from a local extremum are examined. The point on each ray for which a function f(t) reaches an extremum is selected. Linking these points together yields an affinely invariant region, to which an ellipse is fitted using moments.*

- all points corresponding to extremum of f(t) along rays originating from the same local extremum are linked to enclose an (affinely invariant) region (see figure 2).
- This often irregularly-shaped region is then replaced by an ellipse having the same shape moments up to the second order. This ellipse-fitting is affinely invariant as well.

Affine Invariant Detection

- **Algorithm summary (detection of affine invariant region):**
  - Start from a *local intensity extremum* point
  - Go in *every direction* until the point of extremum of some function f
  - Curve connecting the points is the region boundary
  - Compute *geometric moments* of orders up to 2 for this region
  - Replace the region with *ellipse*



T.Tuytelaars, L.V.Gool. "Wide Baseline Stereo Matching Based on Local, Affinely Invariant Regions". BMVC 2000.

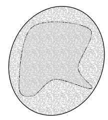
Affine Invariant Detection

- **The regions found may not exactly correspond, so we approximate them with ellipses**
- Geometric Moments:

$$m_{pq} = \int_{\mathbb{R}^2} x^p y^q f(x,y)\,dxdy$$

Fact: moments m_{pq} uniquely determine the function f

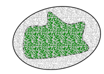
Taking f to be the characteristic function of a region (1 inside, 0 outside), moments of orders up to 2 allow to approximate the region by an ellipse



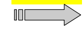
This ellipse will have the same moments of orders up to 2 as the original region

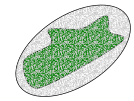
Affine Invariant Detection

- Covariance matrix of region points defines an ellipse:



$q = Ap$





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

$$\Sigma_1 = \langle pp^T \rangle_{\text{region 1}}$$

($p = [x, y]^T$ is relative to the center of mass)

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

$$\Sigma_2 = \langle qq^T \rangle_{\text{region 2}}$$

$$\Sigma_2 = A \Sigma_1 A^T$$

Ellipses, computed for corresponding regions, also correspond!

Affine Invariant Detection : Summary

- **Under affine transformation, we do not know in advance shapes of the corresponding regions**
- **Ellipse given by geometric covariance matrix of a region robustly approximates this region**
- **For corresponding regions ellipses also correspond.**

Methods:

1. Search for extremum along rays [Tuytelaars, Van Gool]:
2. Maximally Stable Extremal Regions [Matas et.al.]