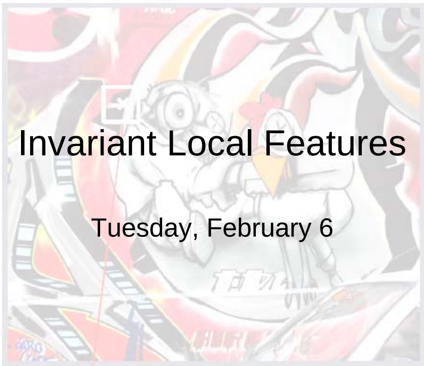
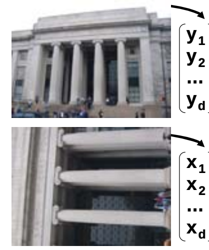


Invariant Local Features

Tuesday, February 6

Invariant local features

Subset of local feature types designed to be invariant to
- Scale
- Translation
- Rotation
- Affine transformations
- Illumination



1) Detect distinctive interest points
2) Extract invariant descriptors

[Mikolajczyk & Schmid, Matas et al., Tuytelaars & Van Gool, Lowe, Kadir et al.,…]

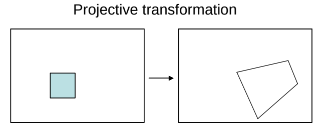
(Good) invariant local features

- Reliably detected
- Distinctive
- Robust to noise, blur, etc.
- Description normalized properly

Classes of transformations

- **Euclidean/rigid**: Translation + rotation
- **Similarity**: Translation + rotation + uniform scale
- **Affine**: Similarity + shear
- **Projective**: Affine + projective warps

Projective transformation

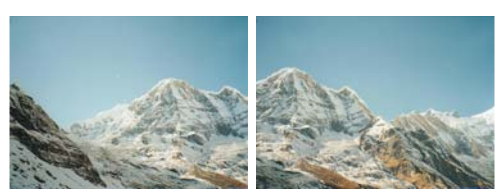


Case study: panorama stitching

[Brown, Szeliski, and Winder, CVPR 2005]

How do we build panorama?

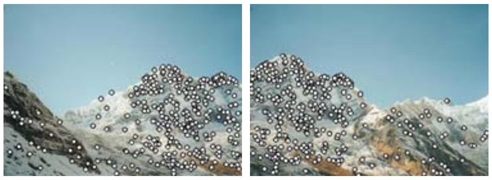
- We need to match (align) images



[These slides are from Darya Frolova and Denis Simakov]

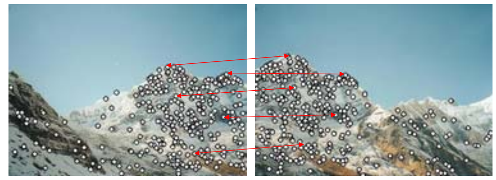
Matching with Features

- Detect feature points in both images



Matching with Features

- Detect feature points in both images
- Find corresponding pairs

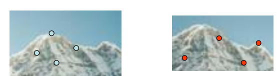


Matching with Features

- Detect feature points in both images
- Find corresponding pairs
- Use these pairs to align images

Matching with Features

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

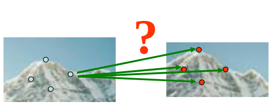


no chance to match!

We need a repeatable detector

Matching with Features

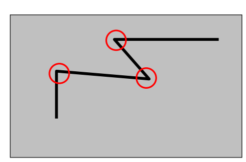
- Problem 2:
  - For each point correctly recognize the corresponding one



We need a reliable and distinctive descriptor

Interest operators: an introductory example

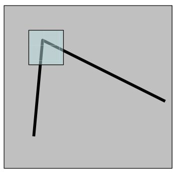
Harris corner detector



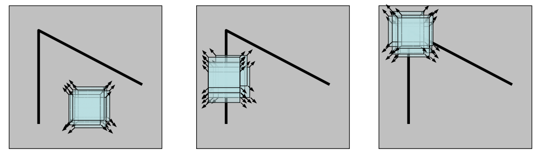
C.Harris, M.Stephens. “A Combined Corner and Edge Detector”. 1988

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: Basic Idea

Corner: significant change in all directions.

[Figure from C. Schmid]

Harris Detector: Mathematics

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

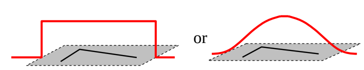
$$E(u,v) = \sum_{x,y} w(x,y)\left[I(x+u, y+v) - I(x,y)\right]^2$$

Window function

Shifted intensity

Intensity

Window function $w(x,y) =$



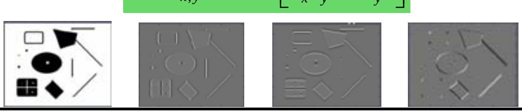
Harris Detector: Mathematics

A bilinear approximation for average intensity change for small shifts in direction [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:

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

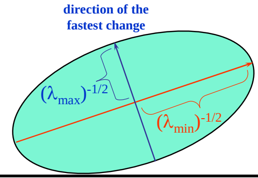


Harris Detector: Mathematics

Intensity change in shifting window: eigenvalues tell us how intensity changes in different directions

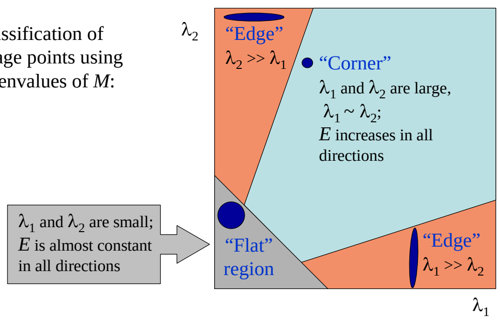
$$E(u,v) \cong [u,v] \; M \begin{bmatrix} u \\ v \end{bmatrix}$$

λ_1, λ_2 – eigenvalues of M



Harris Detector: Mathematics

Classification of image points using eigenvalues of M:



Harris Detector: Mathematics

Measure of corner response:

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

$$\det M = \lambda_1 \lambda_2$$

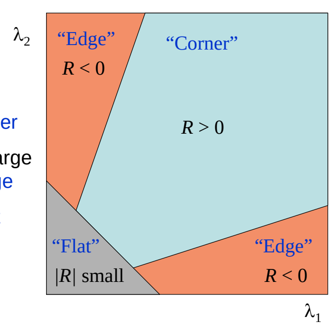
$$\operatorname{trace} M = \lambda_1 + \lambda_2$$

Avoid computing eigenvalues themselves.

(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



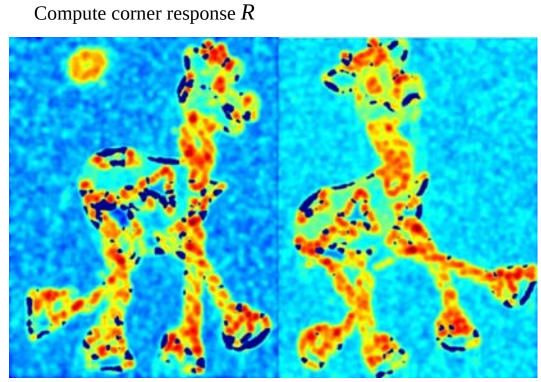
Harris Detector

- 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

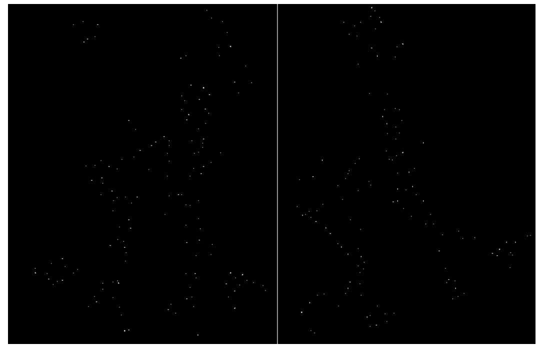


Harris Detector: Workflow

Find points with large corner response: R>threshold

Harris Detector: Workflow

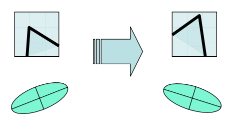
Take only the points of local maxima of R



Harris Detector: Workflow

Harris Detector: Some Properties

- Rotation invariance

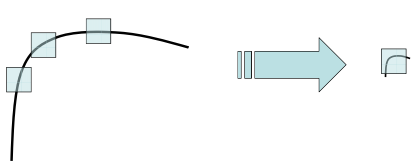


Ellipse rotates but its shape (i.e. eigenvalues) remains the same

Corner response R is invariant to image rotation

Harris Detector: Some Properties

- Not invariant to *image scale*!

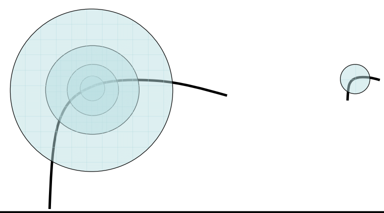


Scale Invariant Detection

[Images from T. Tuytelaars]

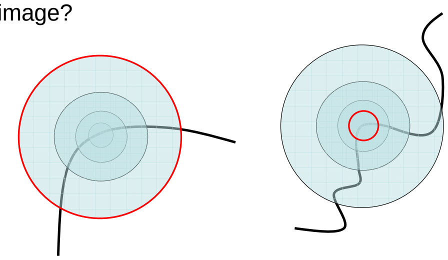
Scale Invariant Detection

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



Scale Invariant Detection

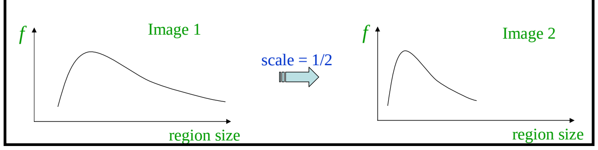
- The problem: how do we choose corresponding circles ***independently*** in each image?



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)



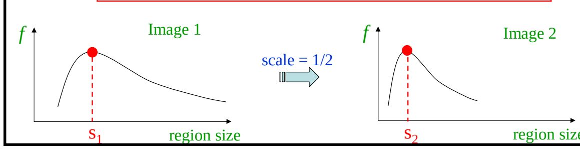
Scale Invariant Detection

- Common approach:

  Take a local maximum of this function

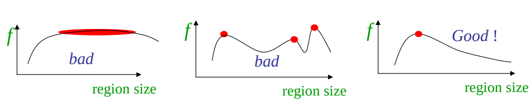
  Observation: region size, for which the maximum is achieved, should be *invariant* to image scale.

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



Scale Invariant Detection

- A "good" function for scale detection: has one stable sharp peak



- For usual images: a good function would be a one which responds to contrast (sharp local intensity change)

Scale Invariant Detection

- Functions for determining scale

$$f = \text{Kernel} * \text{Image}$$

Kernels:

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

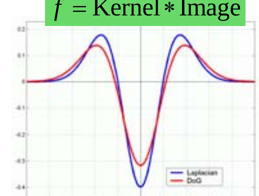
(Laplacian)

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

(Difference of Gaussians)

where Gaussian

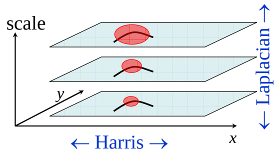
$$G(x,y,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



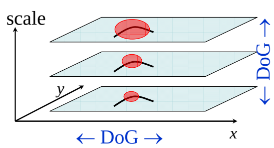
Note: both kernels are invariant to *scale* and *rotation*

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



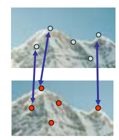
[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001
[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". IJCV 2004

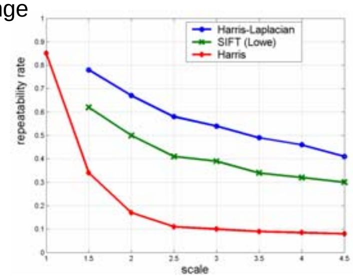
Scale Invariant Detectors

- 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

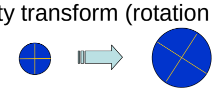
- Given: two images of the same scene with a large *scale difference* between them
- Goal: find *the same* interest points *independently* in each image
- Solution: search for *maxima* of suitable functions in *scale* and in *space* (over the image)

Methods:

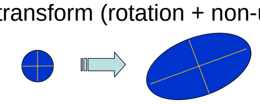
1. Harris-Laplacian [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. SIFT [Lowe]: maximize Difference of Gaussians over scale and space

Affine Invariant Detection

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



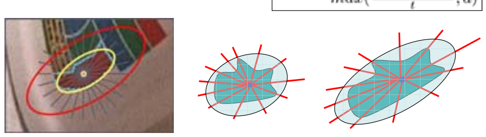
- Now we go on to:
  Affine transform (rotation + non-uniform scale)



Affine Invariant Detection

- Intensity-based regions (IBR):
  - Start from a local intensity extrema
  - Consider intensity profile along rays
  - Select maximum of f(t) along each ray
  - Connect local maxima
  - Fit an ellipse

$$f(t) = \frac{|I(t) - I_0|}{max\left(\frac{\int_0^t |I(t) - I_0| dt}{t}, d\right)}$$



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

Affine Invariant Detection

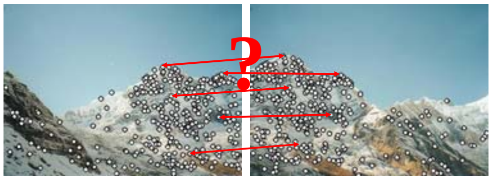
- Maximally Stable Extremal Regions (MSER)
  - *Threshold* image intensities: $I > I_0$
  - Extract *connected components* ("Extremal Regions")
  - Seek extremal regions that remain "Maximally Stable" under range of thresholds

Matas et al. Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. BMVC 2002.

Point Descriptors

- We know how to detect points
- Next question:

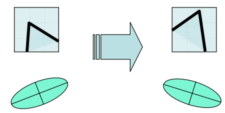
How to *describe* them for matching?



Point descriptor should be:
1. Invariant
2. Distinctive

Rotation Invariant Descriptors

- Harris corner response measure: depends only on the eigenvalues of the matrix *M*

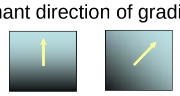


C.Harris, M.Stephens. "A Combined Corner and Edge Detector". 1988

Rotation Invariant Descriptors

- Find local orientation

Dominant direction of gradient

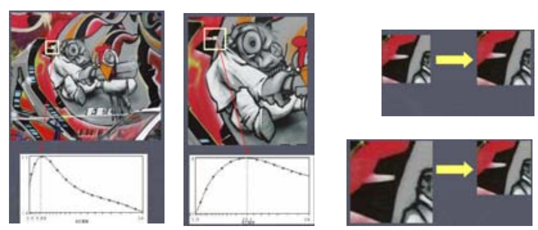


- Rotate description relative to dominant orientation

[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 Descriptors

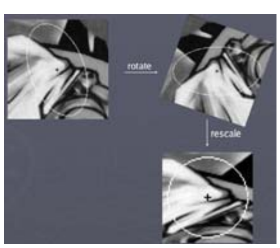
- Use the scale determined by detector to compute descriptor in a normalized frame



[Images from T. Tuytelaars]

Affine Invariant Descriptors

- Compute rotation invariant descriptor in the affine normalized frame (deskew)



[Image from T. Tuytelaars]

Applications

- Wide baseline stereo
- Motion tracking
- Panoramas
- Mobile robot navigation
- 3D reconstruction
- Recognition
  - Specific objects
  - Textures
  - Categories
- …

Wide baseline stereo

[Image from T. Tuytelaars ECCV 2006 tutorial]

Panorama stitching

Brown, Szeliski, and Winder, 2005

Recognition of specific objects, scenes

Schmid and Mohr 1997

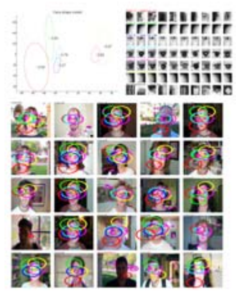
Sivic and Zisserman, 2003

Rothganger et al. 2003

Lowe 2002

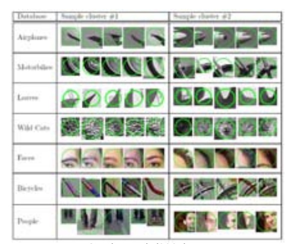
Recognition of categories

Constellation model



Weber et al. (2000)
Fergus et al. (2003)

Bags of words



Csurka et al. (2004)
Dorko & Schmid (2005)
Sivic et al. (2005)
Lazebnik et al. (2006), …

[Slide from Lazebnik, Sicily 2006]

Comparative evaluations

Testing various detector and descriptor options for relative *repeatability* and *distinctiveness*

Planar objects / flat scenes:
Mikolajczyk & Schmid (2004)

3D objects:
Moreels & Perona (2005)

[Images from Lazebnik, Sicily 2006]

Issues

- For *specific-level* recognition – scaling the search?
  - Complexity
  - Distinctiveness
- For *category-level* recognition – are features most appropriate?
  - Sparse
  - Strict appearance description
  - Texture vs. shape
- Expense of detecting interest points