ComVis1

ComVis1U3

- Interest points detectors
- Descriptors

ComVis1

ComVis1U3

- **Interest points detectors**
- Descriptors

From the talk:

Matching with Invariant Features

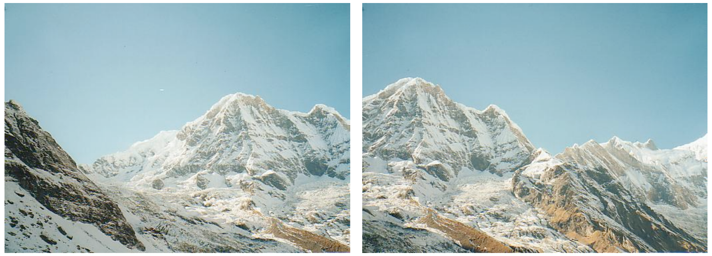
Darya Frolova, Denis Simakov
The Weizmann Institute of Science
March 2004

Example: Build a Panorama

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

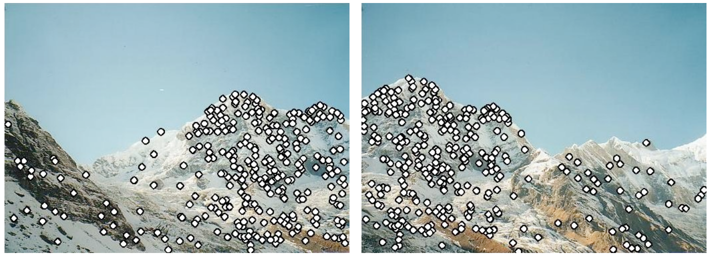
How do we build panorama?

- We need to match (align) images



Matching with Features

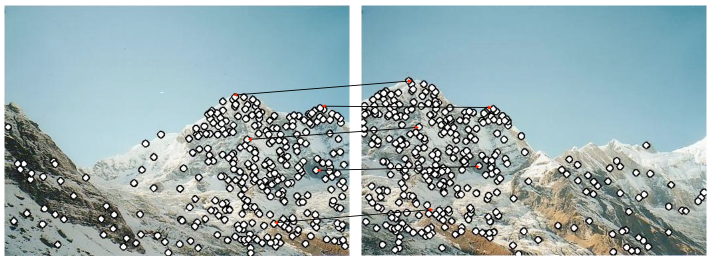
•Detect feature points in both images



Matching with Features

•Detect feature points in both images

•Find corresponding pairs

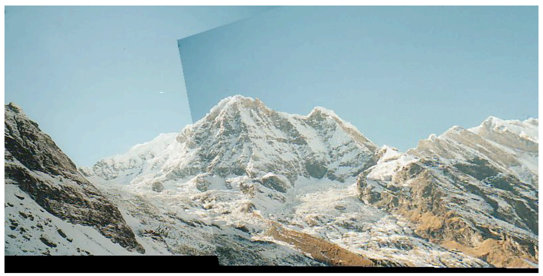


Matching with Features

•Detect feature points in both images

•Find corresponding pairs

•Use these pairs to align images



Matching with Features

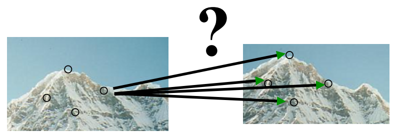
- Problem 1:
  - Detect the *same* point *independently* in both images

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

More motivation…

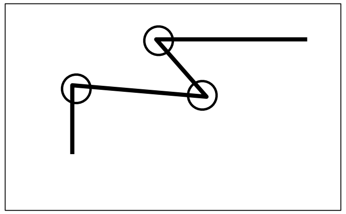
- Feature points are used also for:
  - Image alignment (homography, fundamental matrix)
  - 3D reconstruction
  - Motion tracking
  - Object recognition
  - Indexing and database retrieval
  - Robot navigation
  - … other

Contents

- Harris Corner Detector
  - Description
  - Analysis
- Detectors
  - Rotation invariant
  - Scale invariant
  - Affine invariant
- Descriptors
  - Rotation invariant
  - Scale invariant
  - Affine invariant

An introductory example:

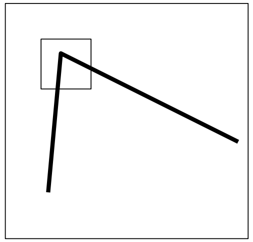
Harris corner detector



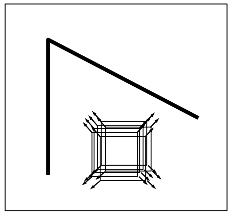
C.Harris, M.Stephens. “A Combined Corner and Edge Detector”. 1988

The Basic Idea

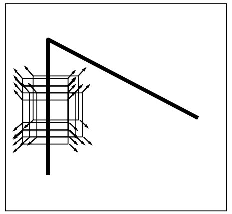
- We should easily recognize the point by looking through a small window
- Shifting a window in *any direction* should give *a large change* in intensity



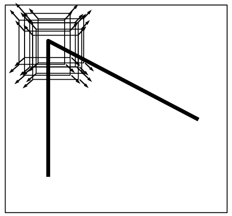
Harris Detector: Basic Idea



“flat” region:
no change in
all directions



“edge”:
no change along
the edge direction



“corner”:
significant change
in all directions

Contents

- Harris Corner Detector
  - Description
  - Analysis
- Detectors
  - Rotation invariant
  - Scale invariant
  - Affine invariant
- Descriptors
  - Rotation invariant
  - Scale invariant
  - Affine invariant

Harris Detector: Mathematics

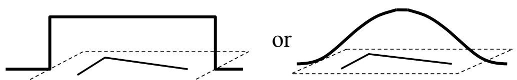
Change of intensity for the shift $[u,v]$:

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

Window function

Shifted intensity

Intensity

Window function $w(x,y)$ =  or

1 in window, 0 outside

Gaussian

Harris Detector: Mathematics

For small shifts $[u,v]$ we have 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:

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

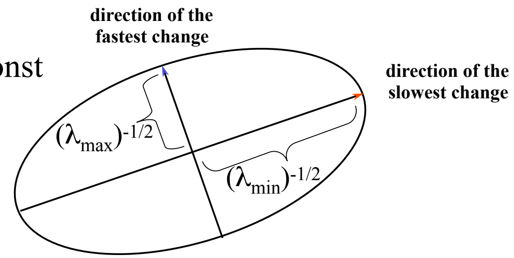
Harris Detector: Mathematics

Intensity change in shifting window: eigenvalue analysis

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

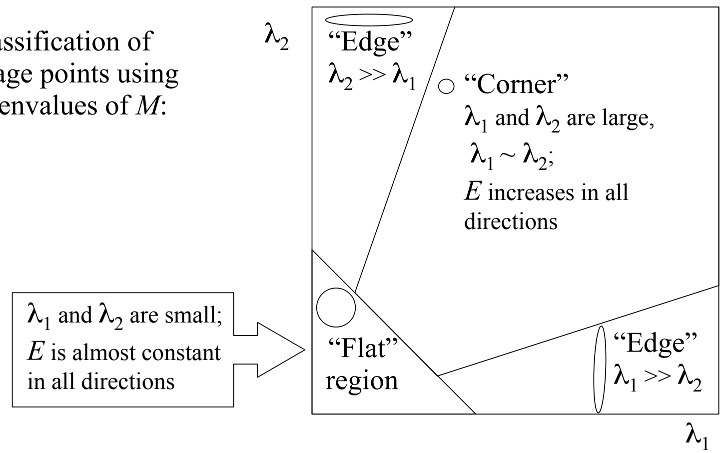
λ_1, λ_2 – eigenvalues of M

Ellipse $E(u,v)$ = const



Harris Detector: Mathematics

Classification of image points using eigenvalues of M:



Harris Detector: Mathematics

Measure of corner response:

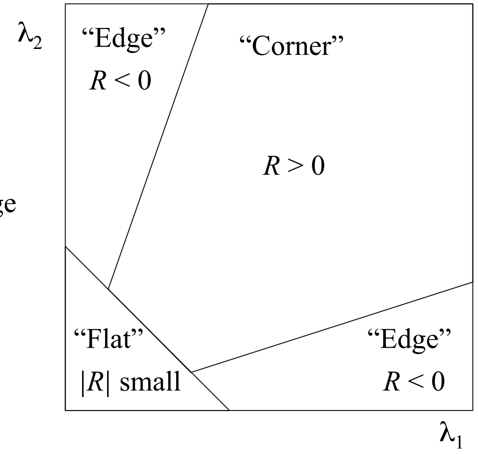
$$R = \det M - k\left(\operatorname{trace} M\right)^2$$

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

(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



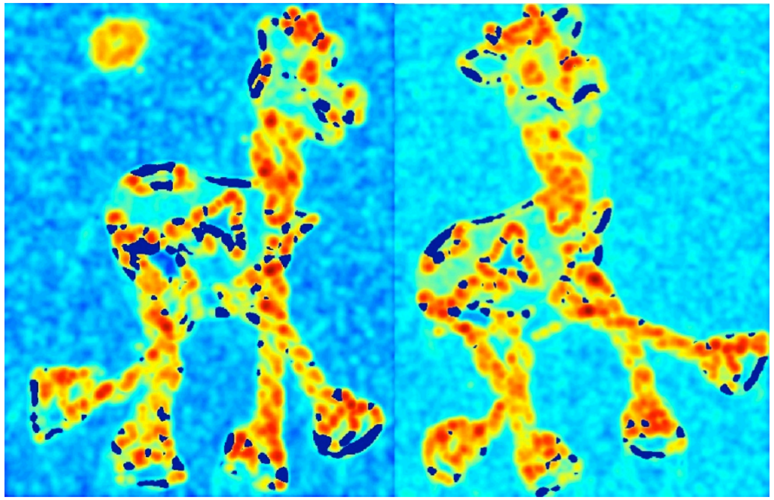
Harris Detector

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

Harris Detector: Workflow

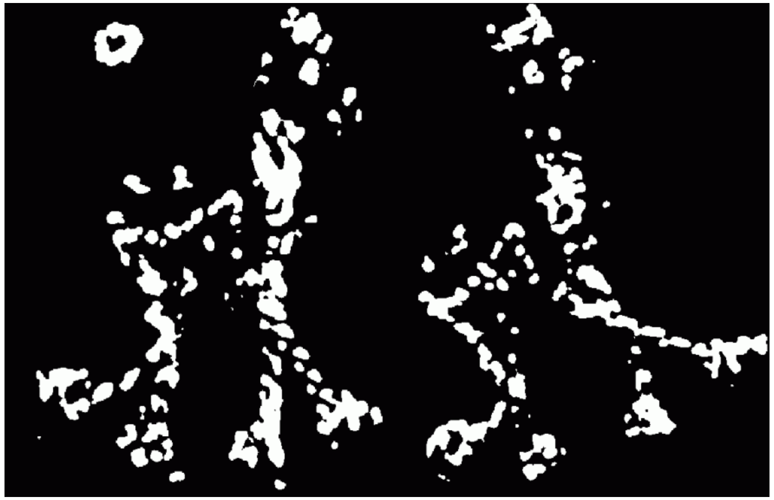
Harris Detector: Workflow

Compute corner response R



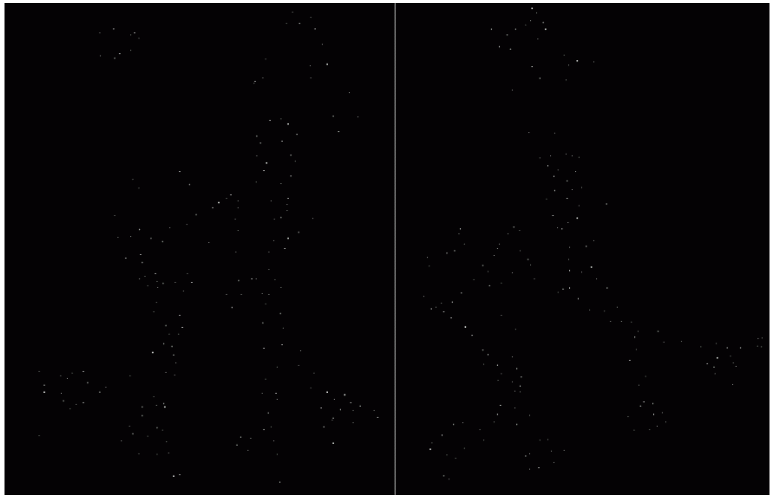
Harris Detector: Workflow

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*

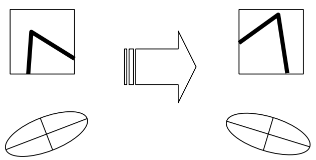
$$R = \lambda_1 \lambda_2 - k\left(\lambda_1 + \lambda_2\right)^2$$

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

Contents

Harris Detector: Some Properties

- Rotation invariance



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

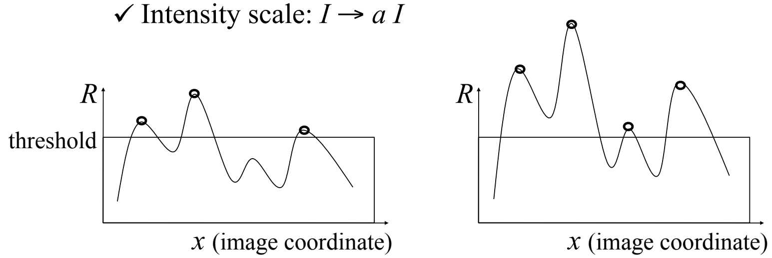
Corner response R is invariant to image rotation

Harris Detector: Some Properties

- Partial invariance to *affine intensity* change

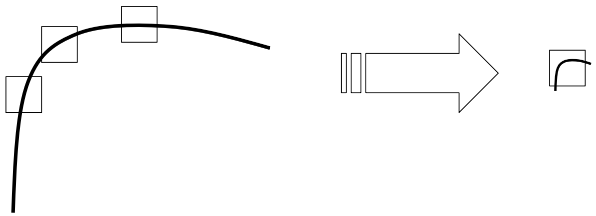
✓ Only derivatives are used => invariance to intensity shift $I \rightarrow I + b$

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



Harris Detector: Some Properties

- But: non-invariant to *image scale*!



All points will be classified as edges

Corner !

ComVis1

ComVis1U3

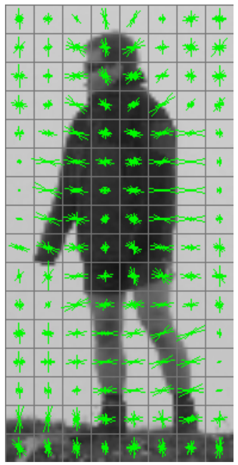
- Interest points detectors
- **Descriptors**

Image descriptors

- Descriptor: a high level representation of an image or video.
- Descriptor: a vector
- Useful for:
  - geometric based matching
  - image retrieval,
  - object recognition and categorization,
  - ....

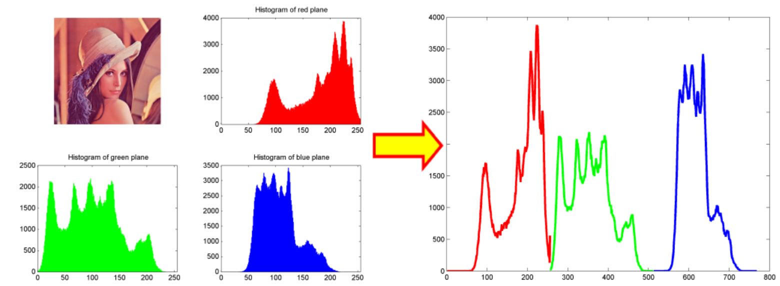
Introduction:

Dense and Sparse descriptors

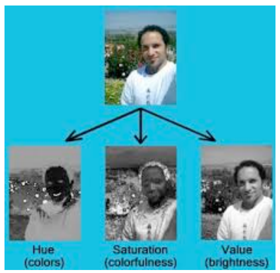


Some dense Descriptors

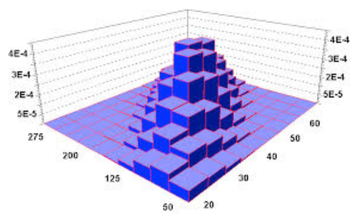
Color RGB histogram



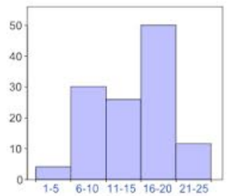
Color HS+V histogram



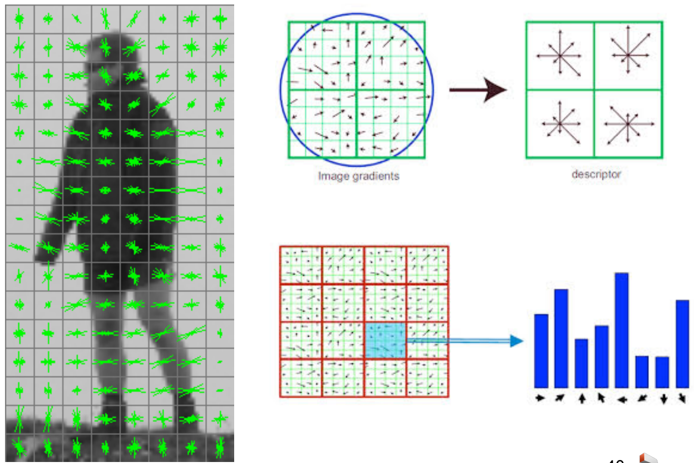
HS



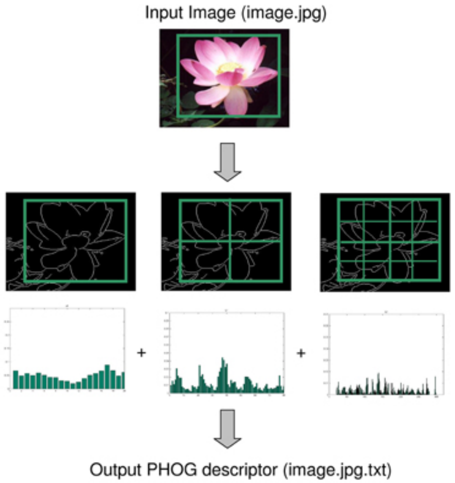
V



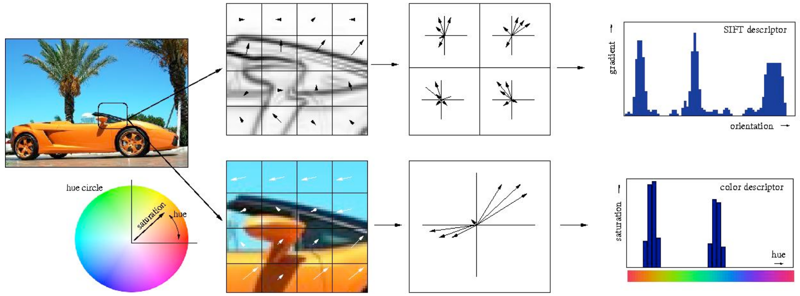
Histogram of oriented gradients (HOG)



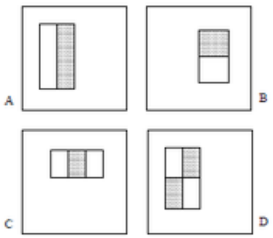
Histogram of oriented gradients (HOG)

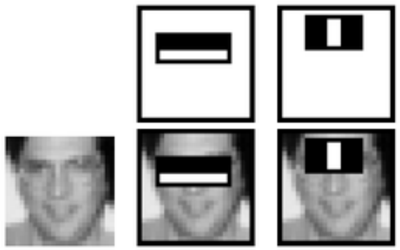


Combining histograms

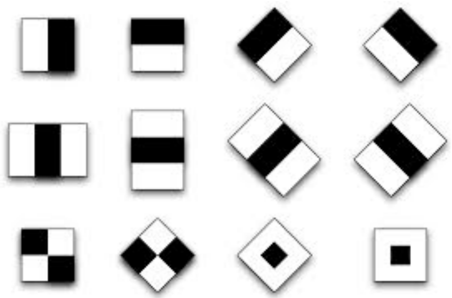


Haar Wavelets



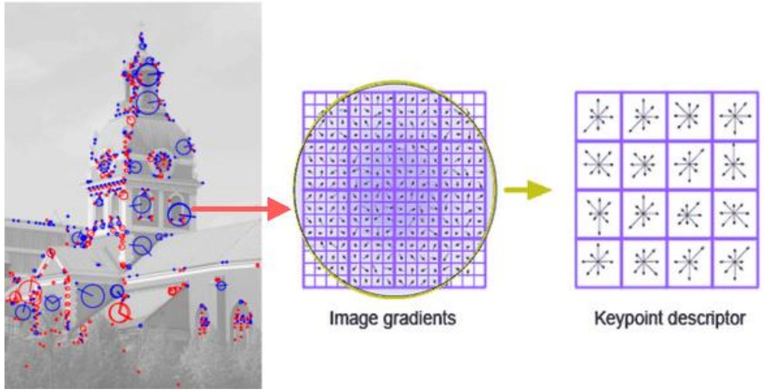


Haar Wavelets



Some Sparse Descriptors

SIFT

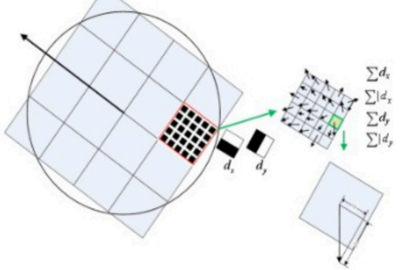


SURF

Speeded Up Robust Features SURF

- Approximate derivatives with Haar wavelets
- Exploit integral images

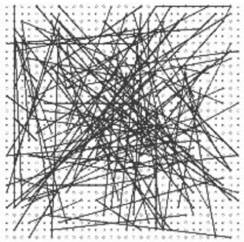
Citations:
4500 (2012)



BRIEF

Binary Robust Independent Elementary Features

- Random selection of pairs of intensity values.
- Fixed sampling pattern of 128, 256 or 512 pairs.
- Hamming distance to compare descriptors (XOR).



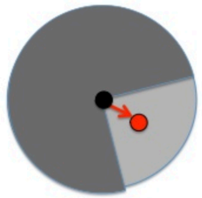
Citations:
149 (2012)

ORB

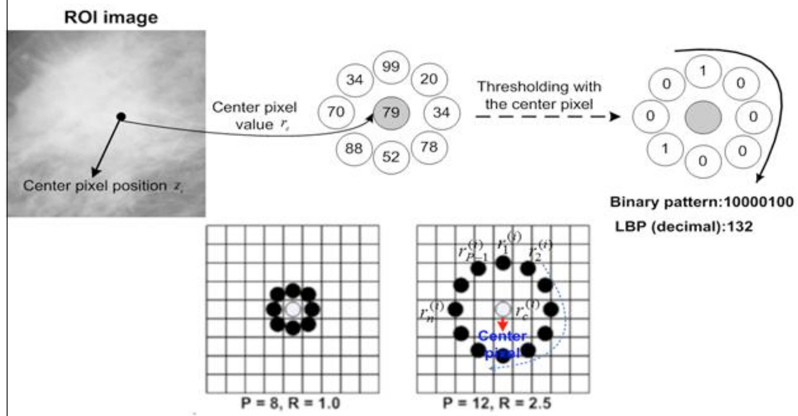
Oriented FAST and Rotated BRIEF

- Add rotation invariance to BRIEF
- Orientation assignment based on the intensity centroid

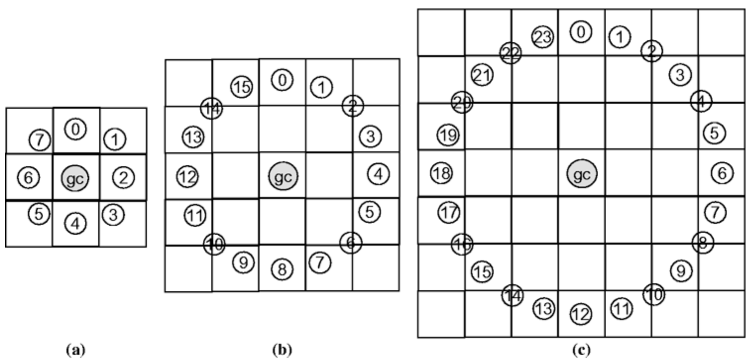
Citations:
43 (2012)



LBP: local binary pattern



LBP: local binary pattern

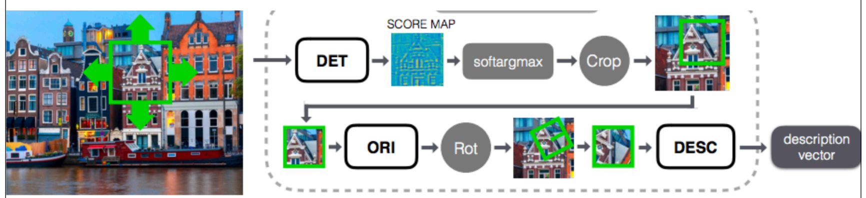


Notes: (a) P = 8, R = 1; (b) P = 16, R = 2; (c) P = 24, R = 3

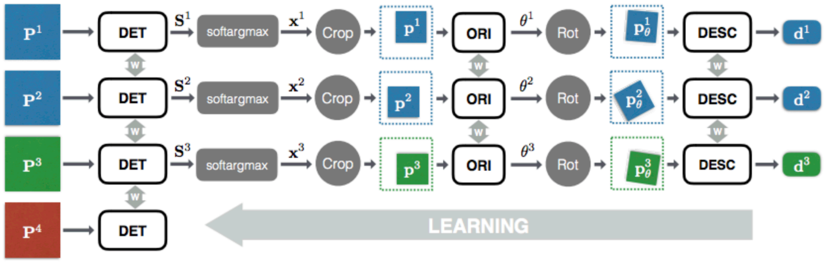
LIFT: Learned Invariant Feature Transform

Kwang Moo Yi*[,1], Eduard Trulls*[,1], Vincent Lepetit[2], Pascal Fua[1]

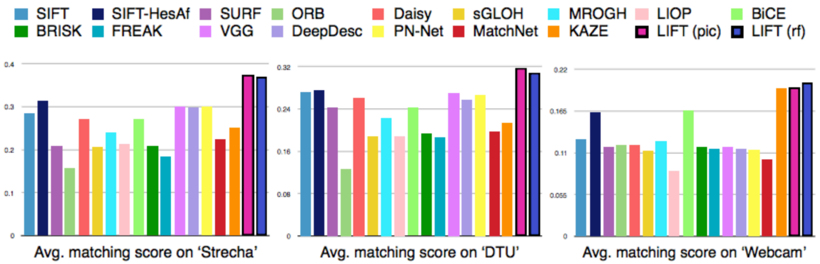
[1]Computer Vision Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL)
[2]Institute for Computer Graphics and Vision, Graz University of Technology
{kwang.yi, eduard.trulls, pascal.fua}@epfl.ch, lepetit@icg.tugraz.at



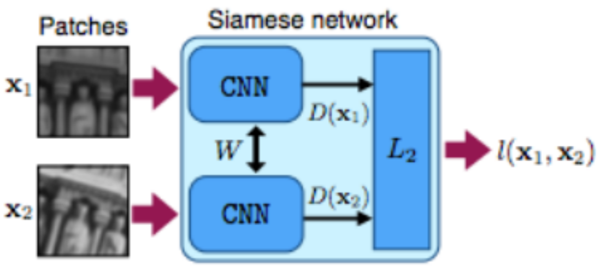
LIFT: Learned Invariant Feature Transform



LIFT: Learned Invariant Feature Transform



Deep Neural network descriptor for matching



Video based descriptors (HOG/HOF)

$$\mu = g(.;\sigma_i^2;\tau_i^2) * \begin{pmatrix} L_x^2 & L_xL_y & L_xL_t \\ L_xL_y & L_y^2 & L_yL_t \\ L_xL_t & L_yL_t & L_t^2 \end{pmatrix}$$