Feature descriptors

Lihi Zelnik-Manor, Computer Vision

Today

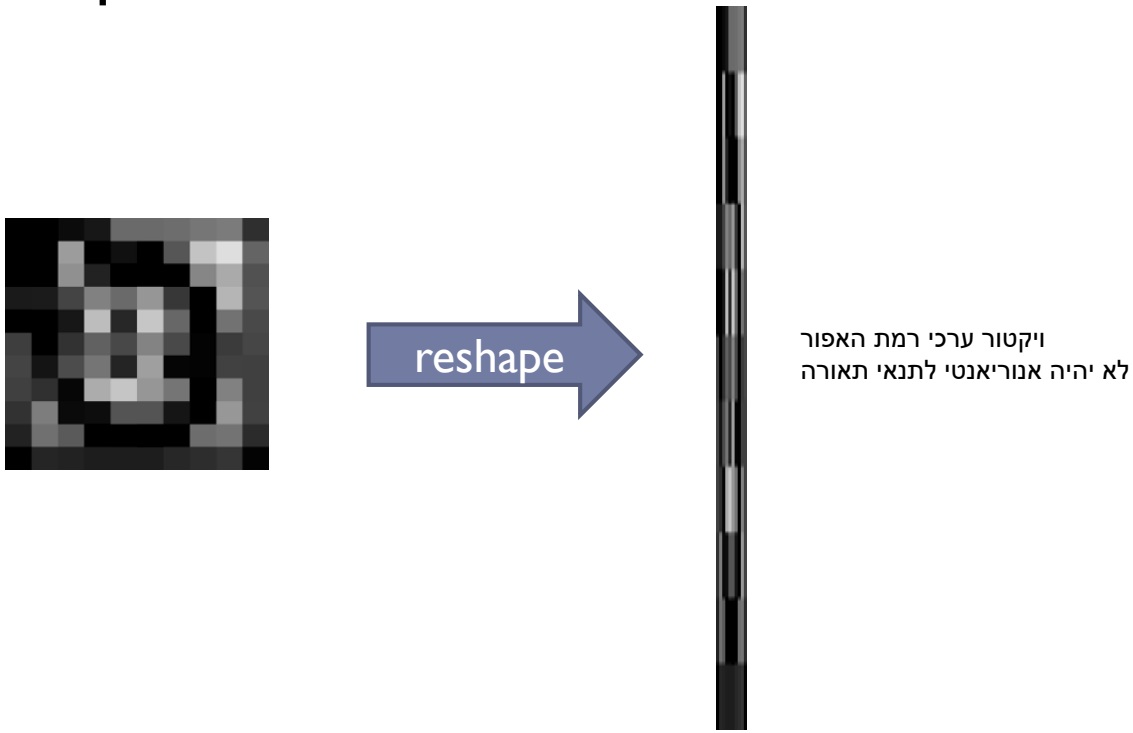
- Local descriptors
  - Selecting invariant regions
  - Feature descriptors:
    - SIFT and others

Today

- Local descriptors
  - Selecting invariant regions
  - Feature descriptors:
    - SIFT and others

The naïve descriptor – intensities vector

- The Simplest descriptor is a vector of the intensities within the patch.



What is this going to be invariant to?

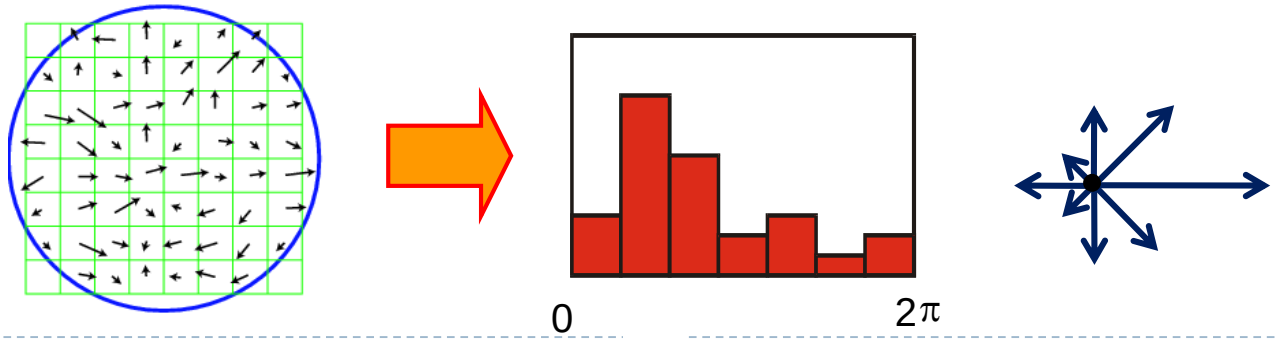
The naïve descriptor – intensities vector

- Disadvantage of the intensities vector
  1. Changes significantly with illumination
  2. Changes significantly with small shifts in position

Another naïve descriptor

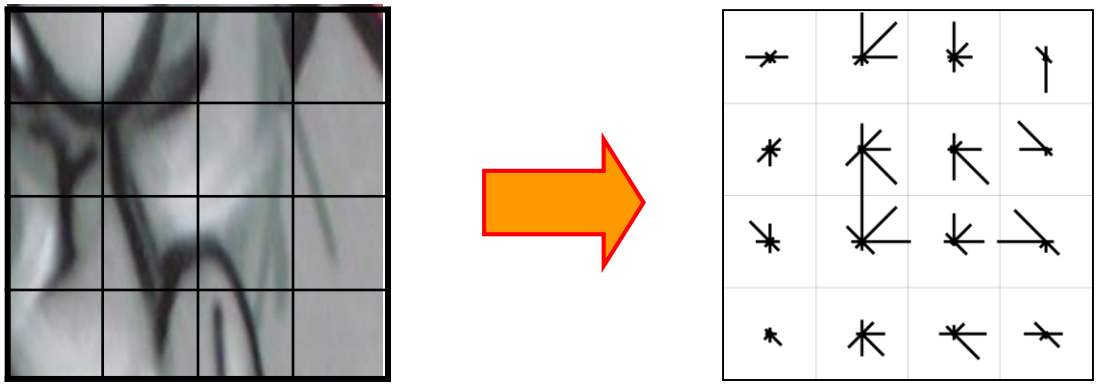
- Disadvantage of the intensities vector
  1. Changes significantly with illumination
  2. Changes significantly with small shifts in position

- Solutions
  1. Use gradients instead of intensities
  2. Histograms



A good feature descriptor: SIFT

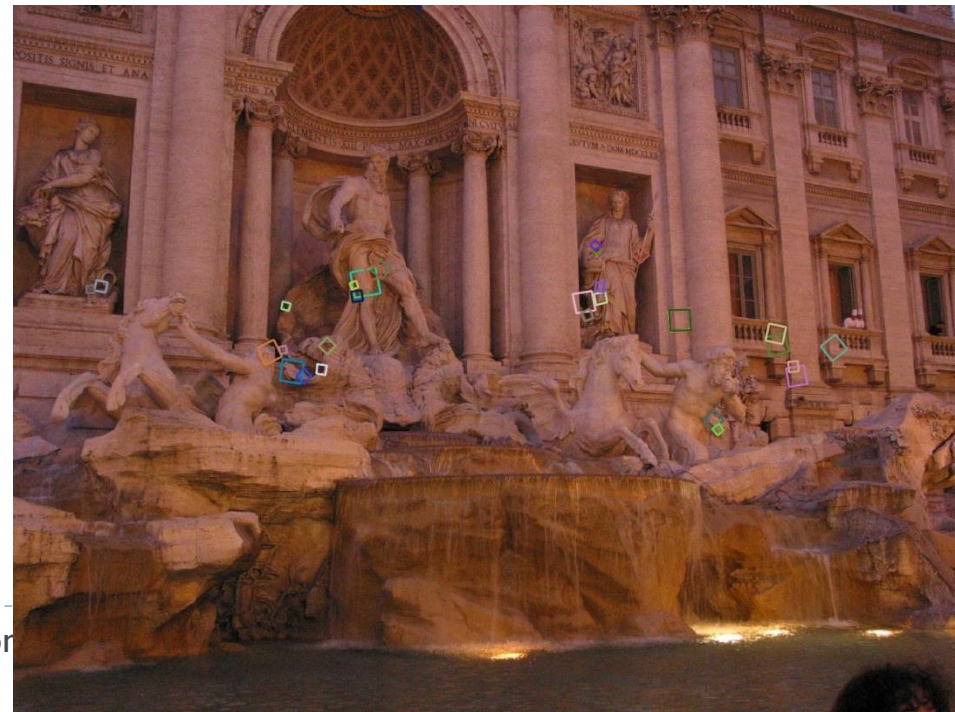
- **S**cale **I**nvariant **F**eature **T**ransform
- Descriptor computation:
  - Divide patch into 4x4 sub-patches: 16 cells
  - Compute histogram of gradient orientations (8 reference angles) for all pixels inside each sub-patch
  - Resulting descriptor: 4x4x8 = 128 dimensions



David G. Lowe. "Distinctive image features from scale-invariant keypoints." *IJCV*'2004.

SIFT overview

- Extraordinarily robust matching technique
  - Can handle changes in viewpoint up to about 60 degree out of plane rotation
  - Can handle significant changes in illumination
    - Sometimes even day vs. night (below)
  - Fast and efficient—can run in real time
  - Lots of code available
    - http://people.csail.mit.edu/albert/ladypack/wiki/index.php/Known_implementations_of_SIFT



Working with SIFT descriptors

- One image yields: 16 * 8
  - n 128-dimensional descriptors: each one is a histogram of the gradient orientations within a patch
    - [n x 128 matrix]
  - n scale parameters specifying the size of each patch
    - [n x 1 vector]
  - n orientation parameters specifying the angle of the patch
    - [n x 1 vector]
  - n 2d points giving positions of the patches
    - [n x 2 matrix]

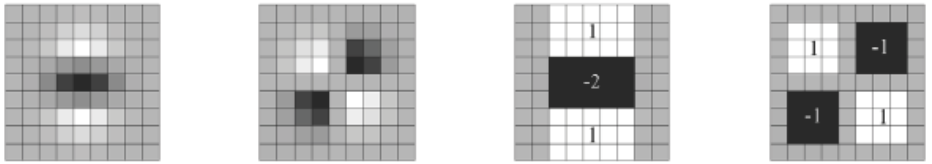
SURF descriptor

- Fast approximation of SIFT
  - Efficient computation by 2D box filters & integral images ➔ 6 times faster than SIFT
  - Equivalent quality for object identification

GPU implementation available

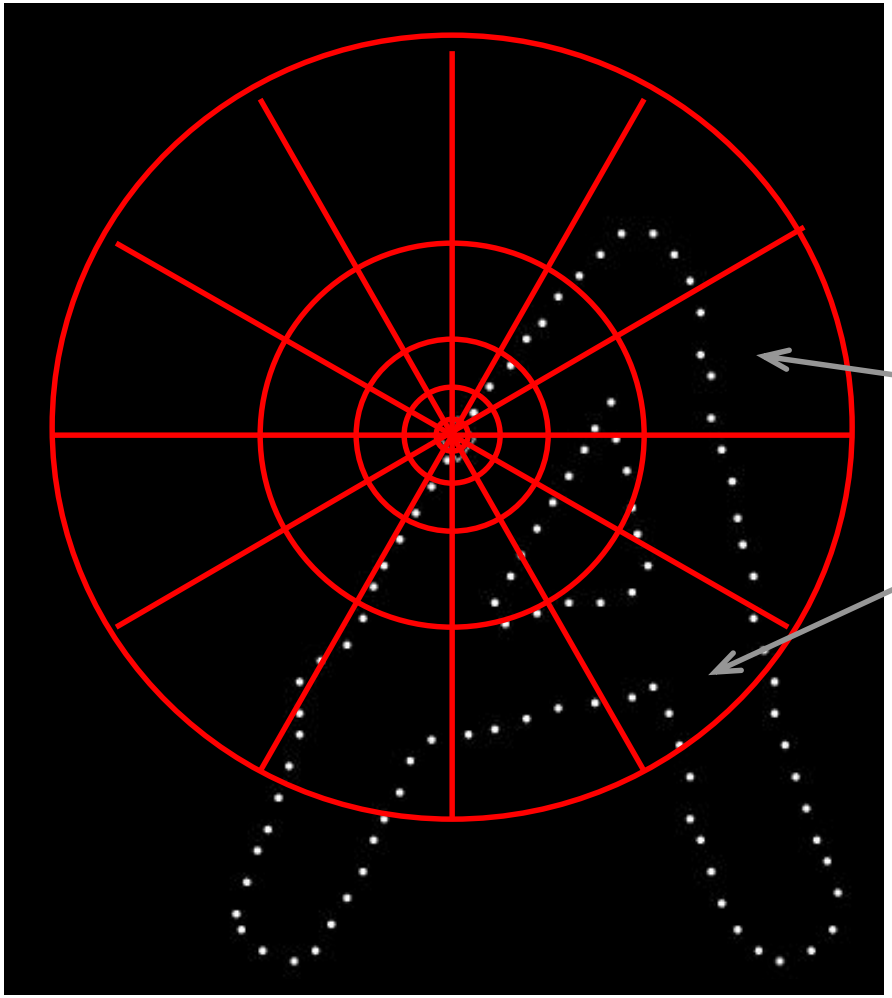
Feature extraction @ 200Hz (detector + descriptor, 640×480 img)

http://www.vision.ee.ethz.ch/~surf



[Bay, ECCV'06], [Cornelis, CVGPU'08]

Local Descriptors: Shape Context



Count the number of points inside each bin, e.g.:

Count = 4

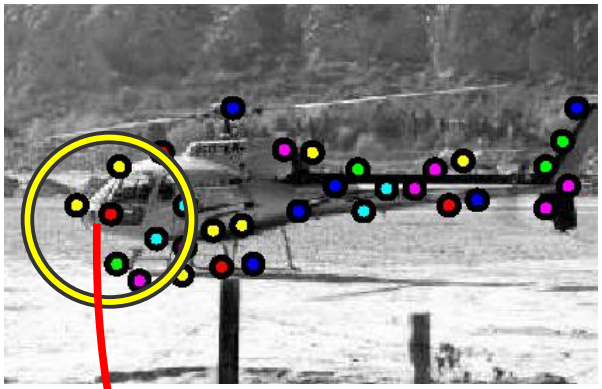
⋮

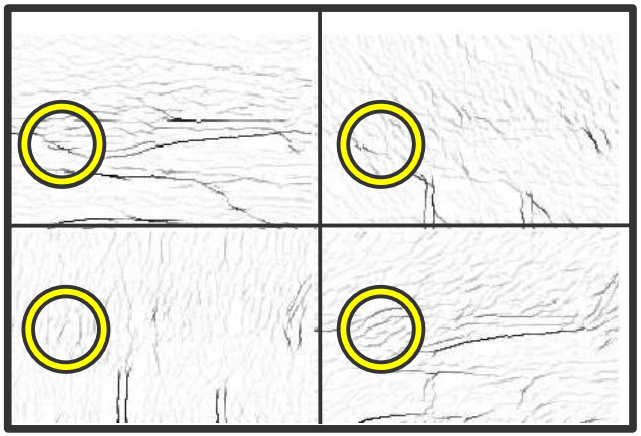
Count = 10

Log-polar binning: more precision for nearby points, more flexibility for farther points.

- Belongie & Malik, ICCV 2001

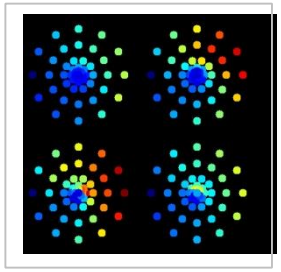
Local Descriptors: Geometric Blur



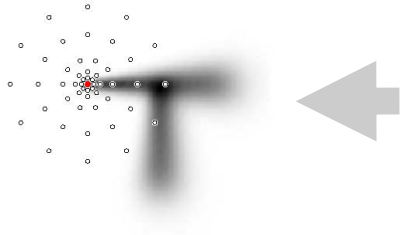


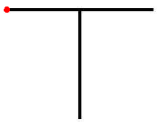
Compute edges at four orientations

Extract a patch in each channel



Example descriptor





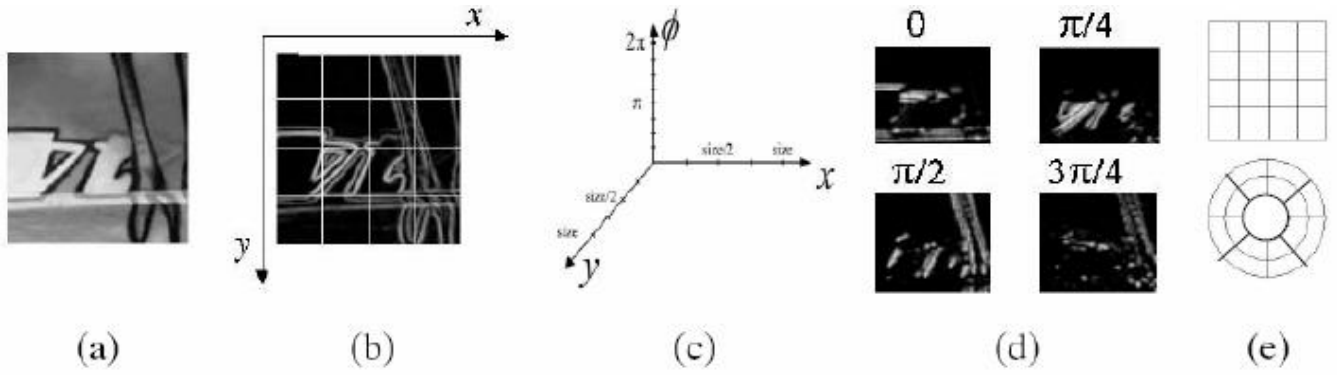
(Idealized signal)

Apply spatially varying blur and sub-sample

- Berg & Malik, CVPR 2001

GLOH

- Gradient Location and Orientation Histogram
  - Very similar to SIFT
  - Log-polar location grid
    - 3 bins in radial direction
    - 8 bins in angular direction
    - Gradient orientation quantized to 16 bins
  - Total dimension
    - (2x8+1)*16=272 bins ➔ PCA for dimension reduction



More on feature detection/description

Publications

Region detectors

- *Harris-Affine & Hessian Affine*: K. Mikolajczyk and C. Schmid, Scale and Affine invariant interest point detectors. In IJCV 1(60):63-86, 2004. PDF
- *MSER*: J.Matas, O. Chum, M. Urban, and T. Pajdla, Robust wide baseline stereo from maximally stable extremal regions. In BMVC p. 384-393, 2002. PDF
- *IBR & EBR*: T.Tuytelaars and L. Van Gool, Matching widely separated views based onaffine invariant regions. In IJCV 1 (59):61-85, 2004. PDF
- *Salient regions*: T. Kadir, A. Zisserman, and M. Brady, An affine invariant salient region detector. In ECCV p. 404-416, 2004. PDF

Region descriptors

- *SIFT*: D. Lowe, Distinctive image features from scale invariant keypoints. In IJCV 2(60):91-110, 2004. PDF

Performance evaluation

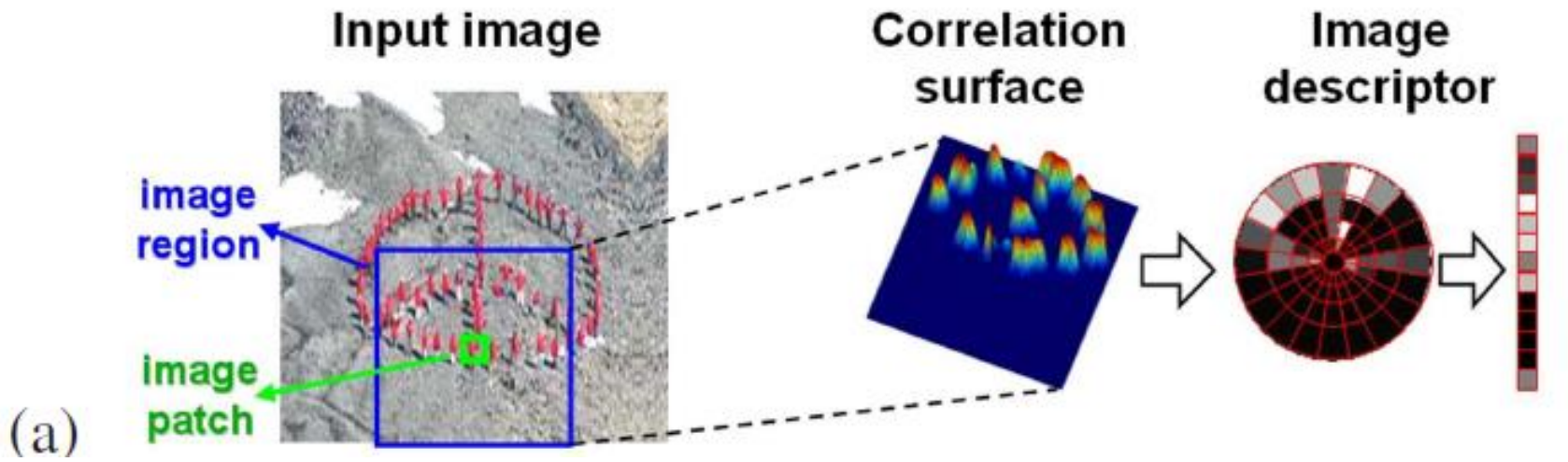
- K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir and L. Van Gool, A comparison of affine region detectors. Technical Report, accepted to IJCV. PDF
- K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors. Technical Report, accepted to PAMI. PDF

http://www.robots.ox.ac.uk/~vgg/research/affine/detectors.html#binaries

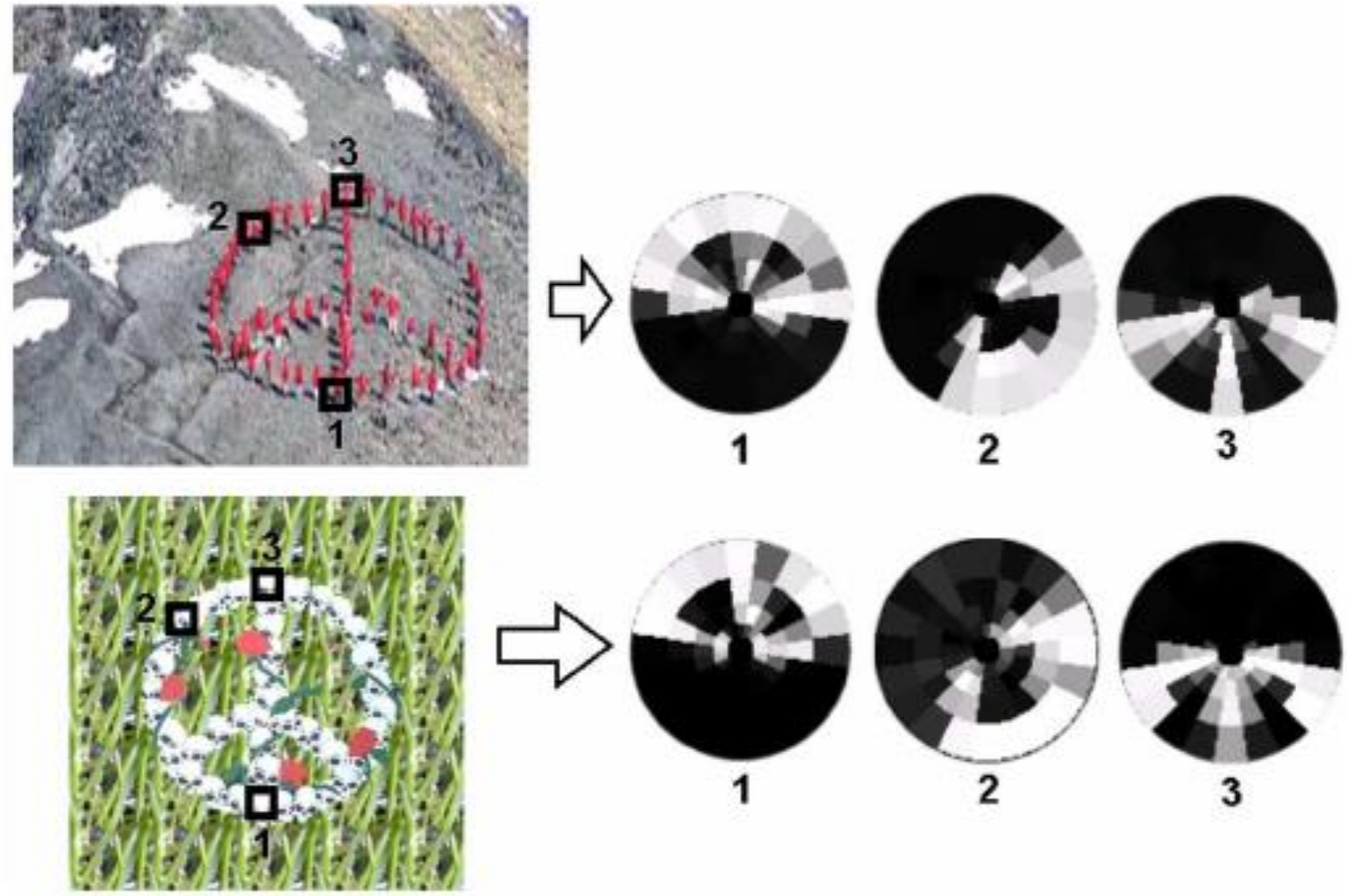
Self-similarity descriptor

- All the descriptors so far captured same appearance
- What can we do if the objects have the same shape but different appearance?

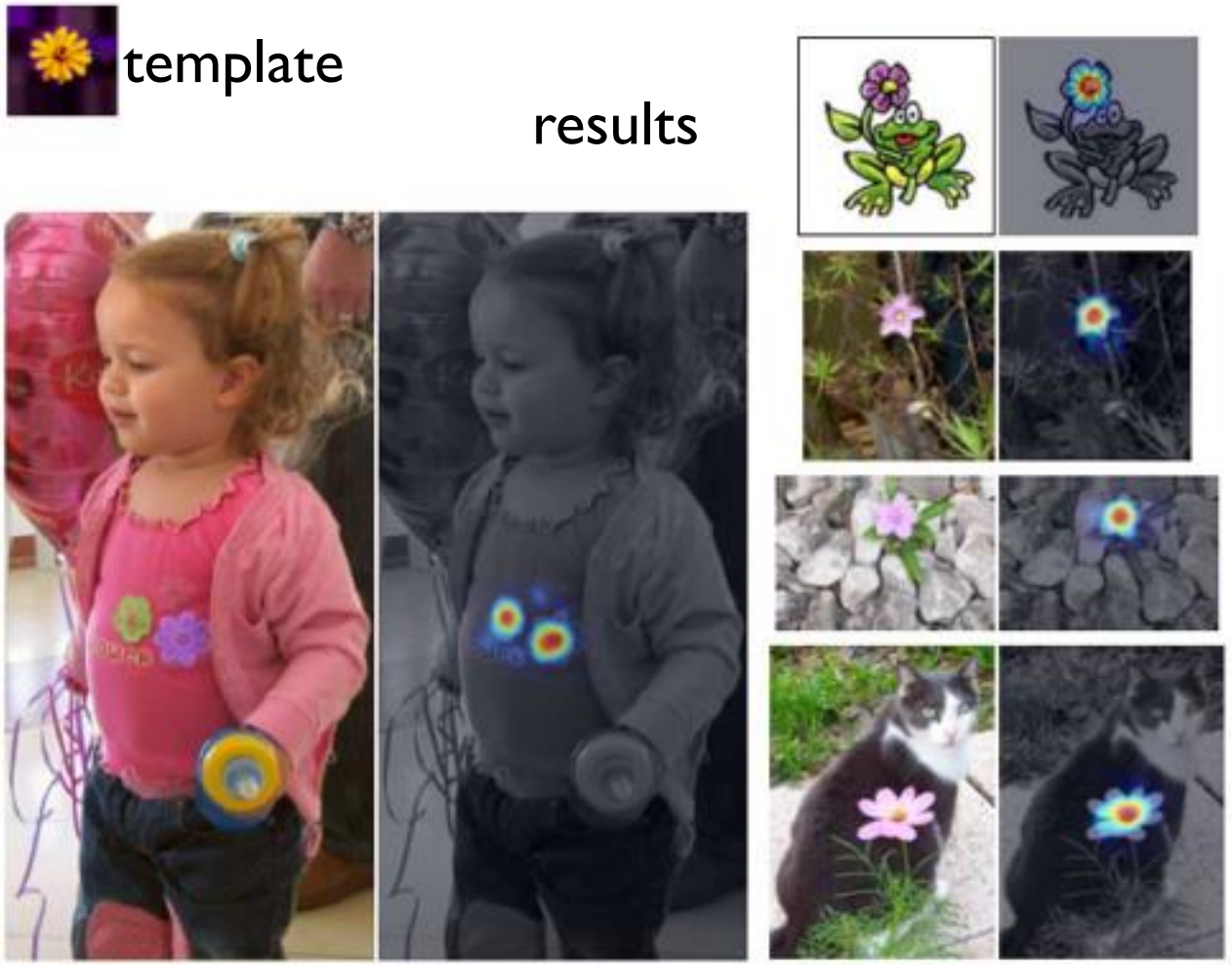
Self-similarity descriptor



Self-similarity descriptor



Self-similarity descriptor

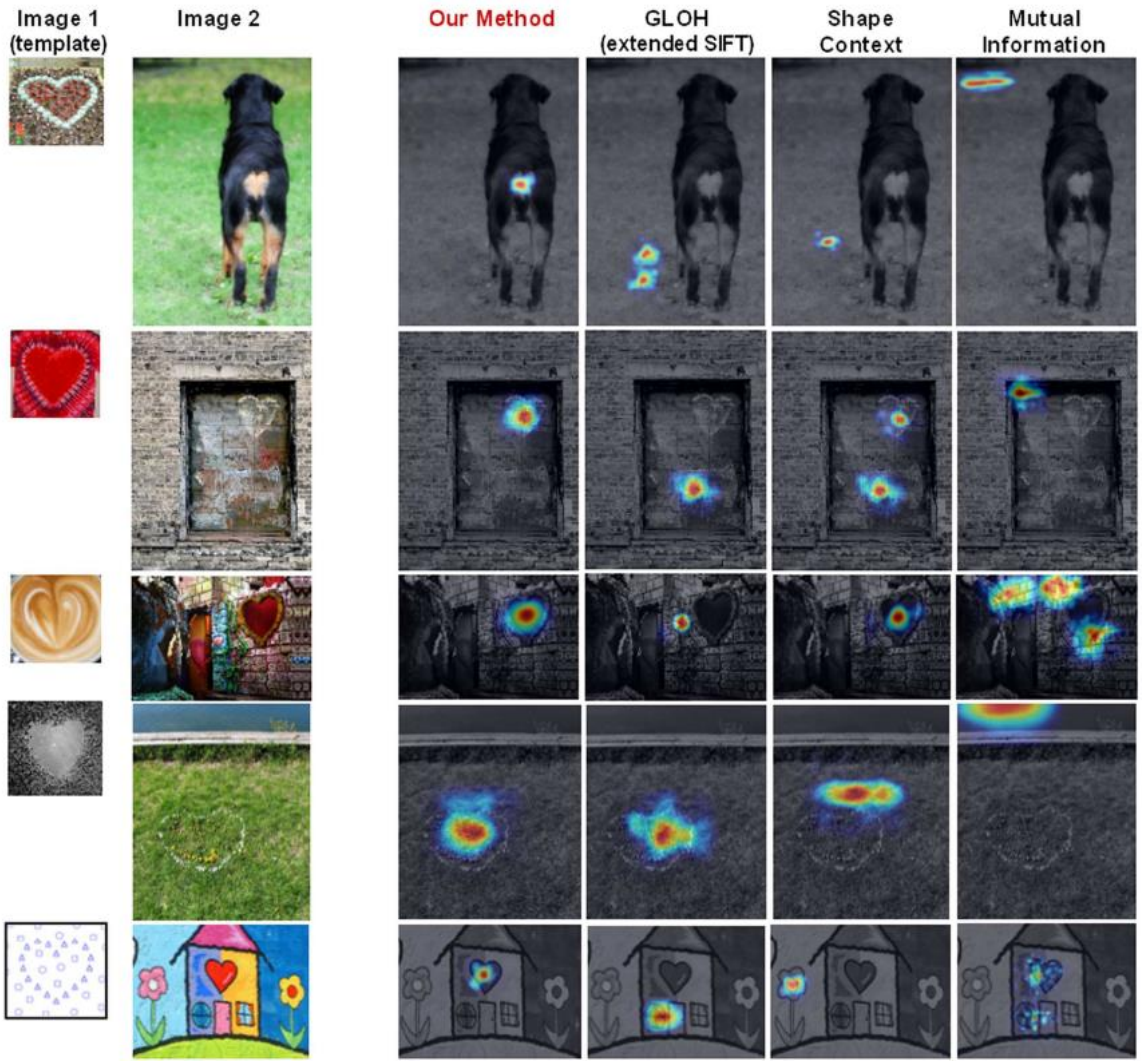


Self-similarity descriptor

template

results

Self-similarity descriptor



Advantages of local features

- Useful
  - It is critical to find distinctive and repeatable local regions for multi-view matching
- Complexity reduction
  - Selection of distinctive points reduces number of regions to process
- Compact description
  - Describe images, objects, parts without requiring segmentation;
- Robustness
  - To clutter & occlusion
  - Similar descriptors in spite of moderate view changes, noise, blur, etc.

End – Feature descriptors

Now you know how it works