Matching Local Self-Similarities across Images and Videos

Eli Shechtman & Michal Irani
The Weizmann Institute of Science
Israel

Thomas Tiel Groenestege
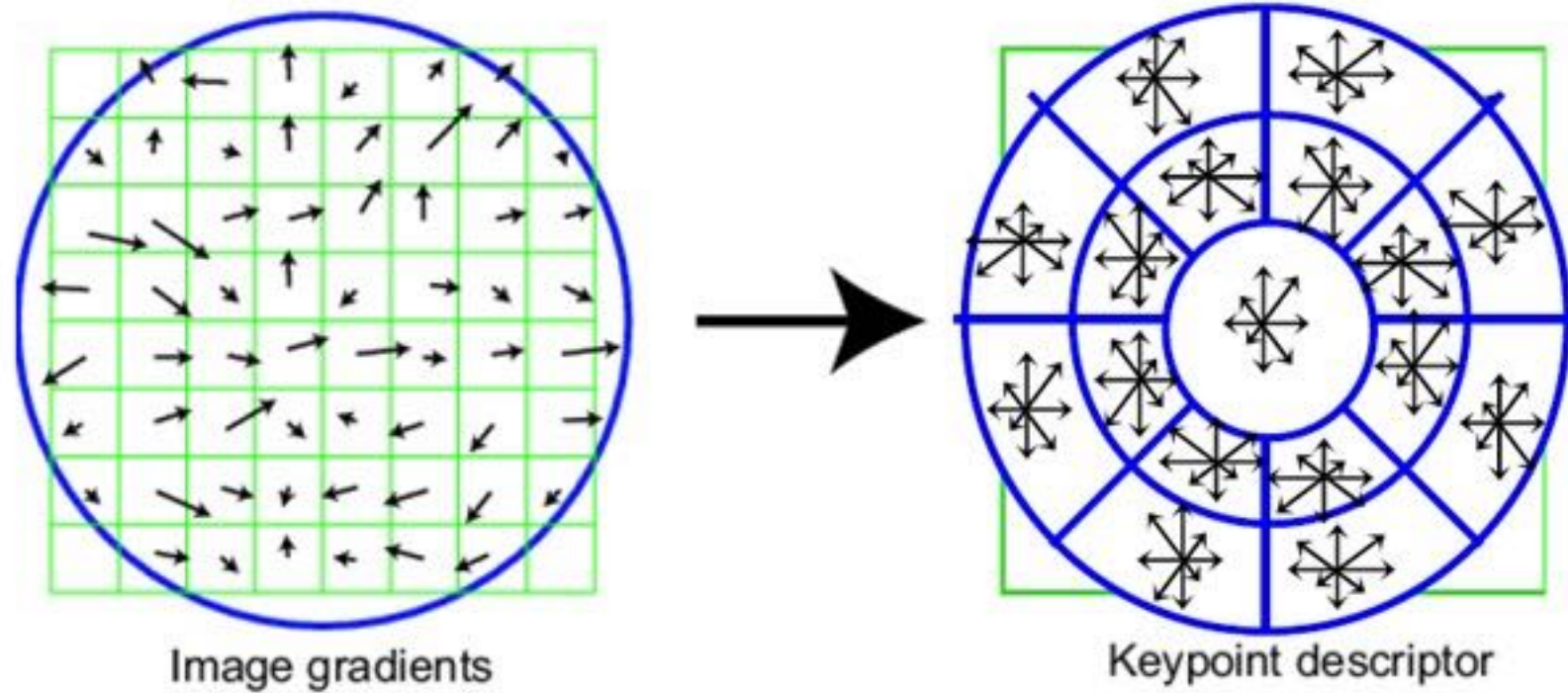Jorn Engelbart

Matching Local Self-Similarities

Matching Local Self-Similarities

Matching Local Self-Similarities

GLOH

Image gradients
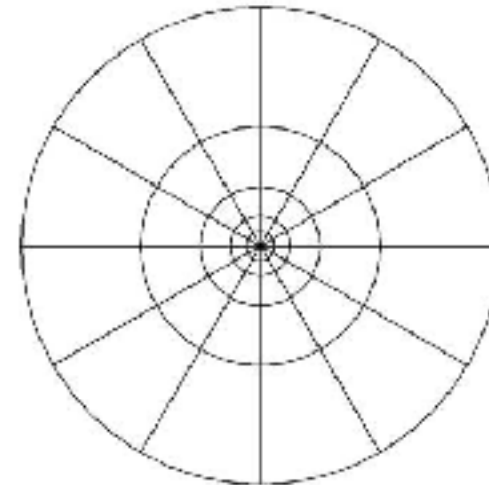
Keypoint descriptor

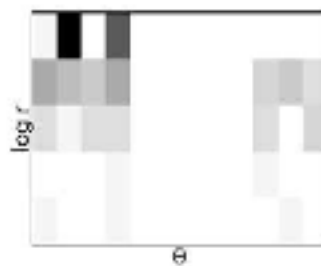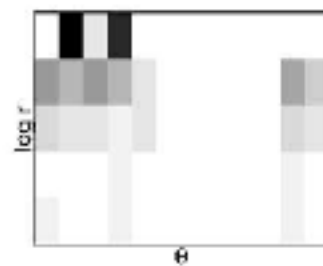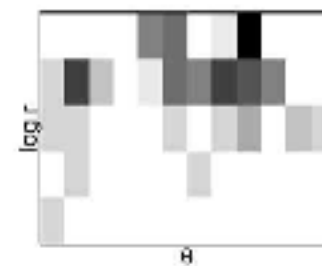# Shape Context

(a)   (b)   (c)

(d)   (e)   (f)   (g)

Template matching

- Template image F to target image G

- Not the same size

- Similarity not in appearance, but structure
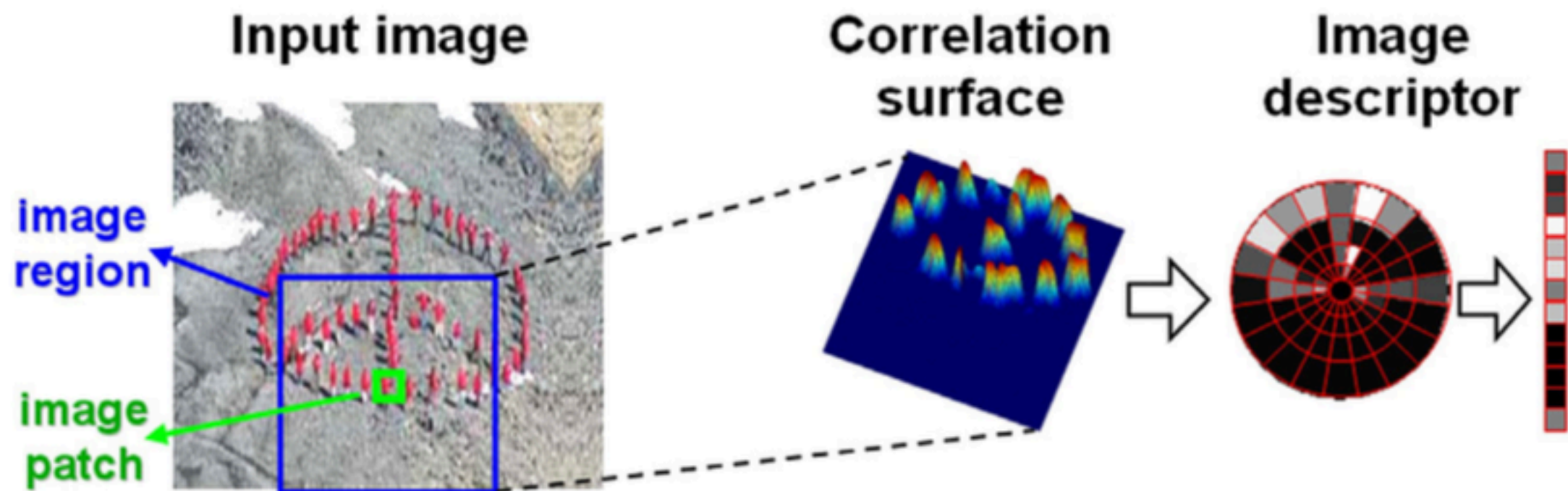
- Local self-similarity

Self-similarity descriptors
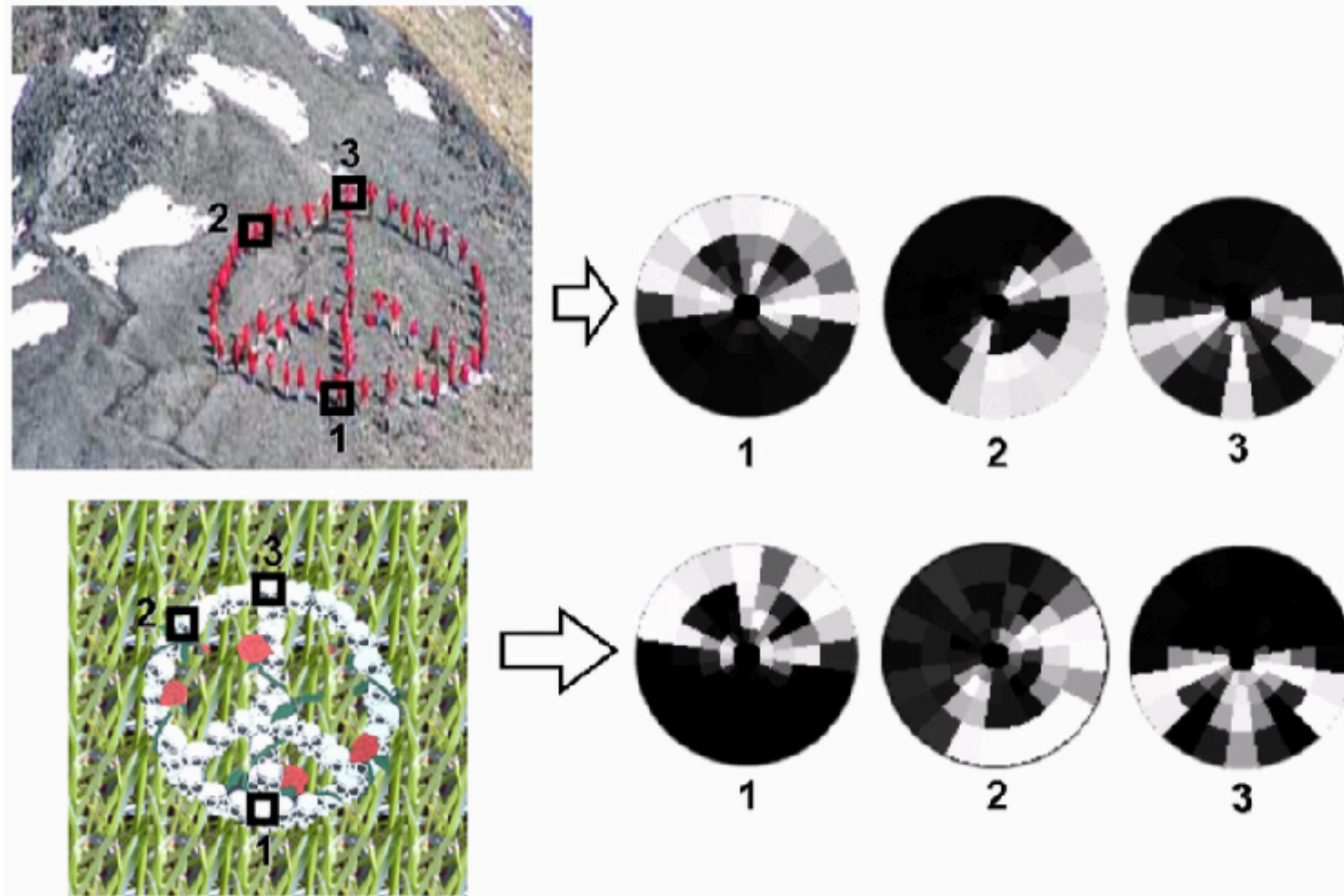
- Sum of Square Differences

- $S_q(x,y) = \exp\left(-\dfrac{SSD_q(x,y)}{\max\left(var_{noise}, var_{auto}(q)\right)}\right)$

Self-similarity descriptors

- 20 angles for 4 intervals = 80 bins
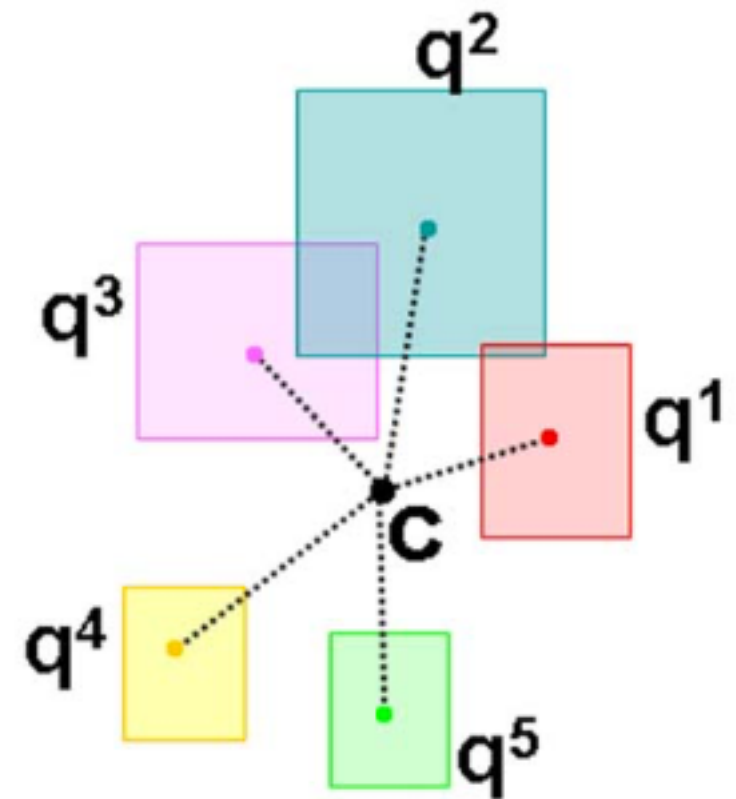
Self-similarity descriptors

Self-similarity descriptors

- Very local

- Takes maximal values for bins

- Accounts for small affine transformations

- Patch over pixel

Matching ensembles of descriptors

- Creating "ensembles"

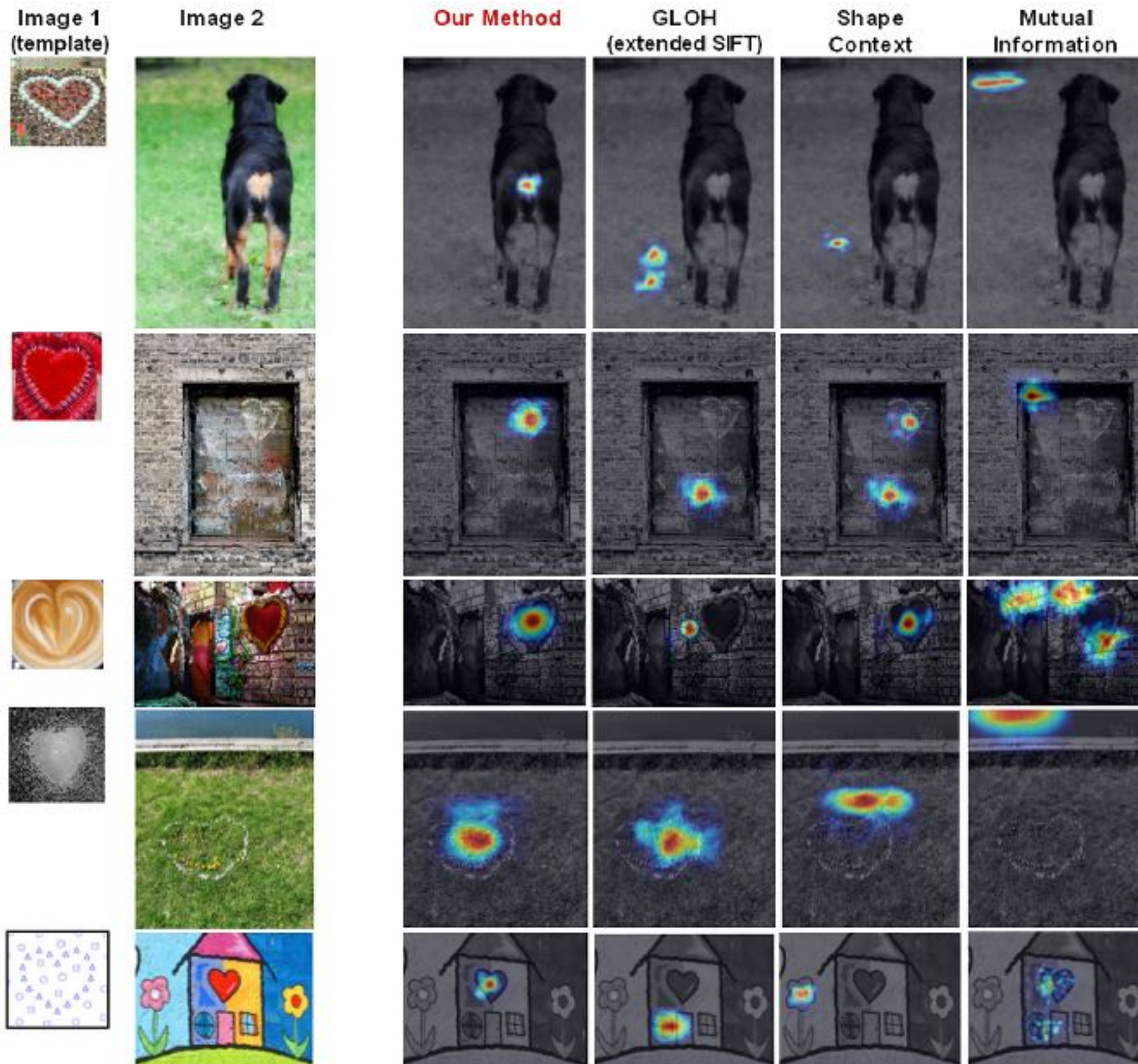- Relative positions of descriptors

- Filter out non-information

Matching ensembles of descriptors

- Similarity with sigmoid of L1 distance

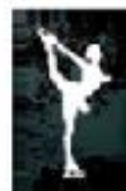- Likelihood map of the template

- Gaussian image pyramid, normalisation

Results

- 60 challenging image pairs

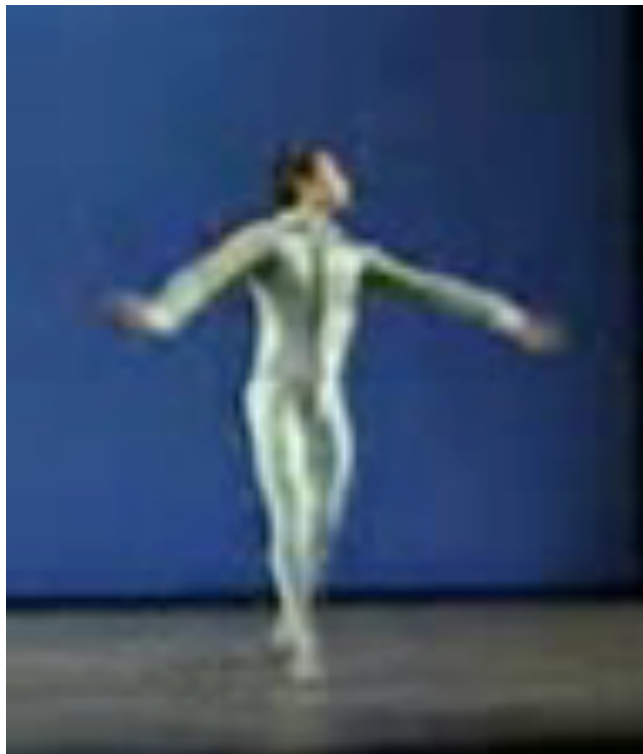- Other methods failed in majority of cases
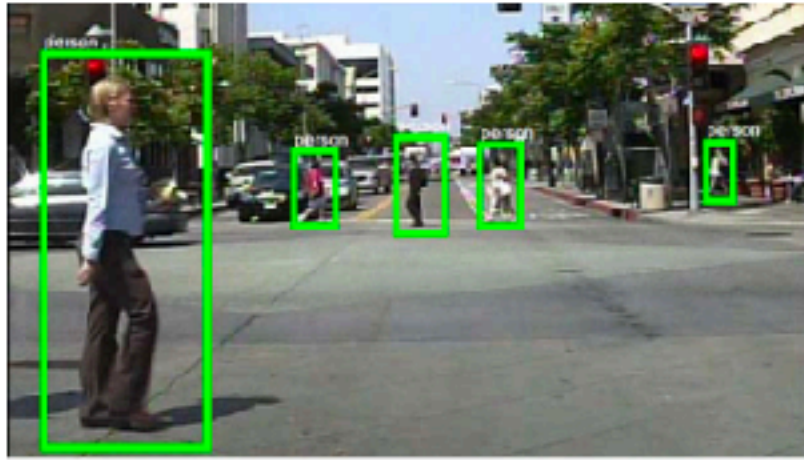
- LSS 86% correct

Other methods

| Image 1 (template) | Image 2 | | Our Method | GLOH (extended SIFT) | Shape Context | Mutual Information |

Pose sketches

Video

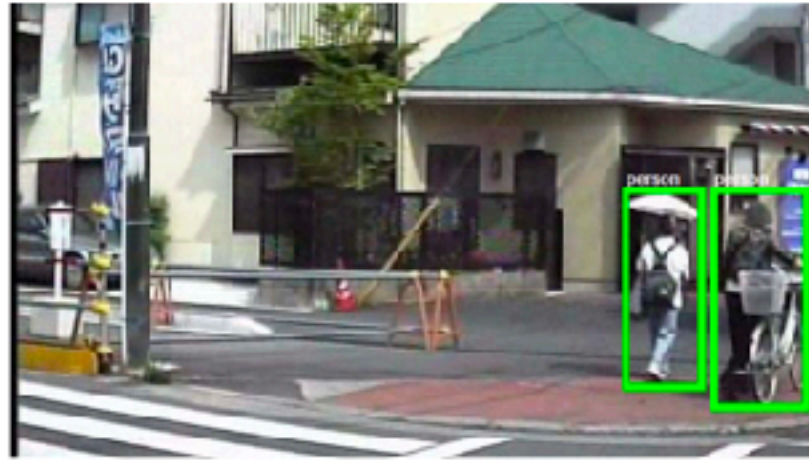Input video

[Shechtman-Irani CVPR'05]

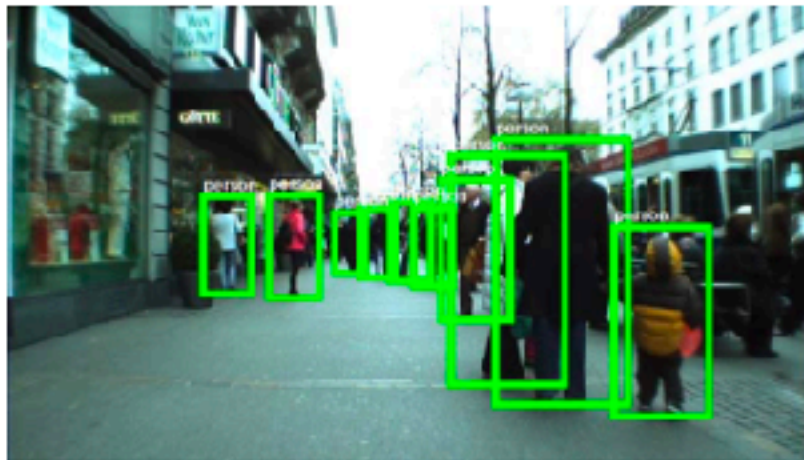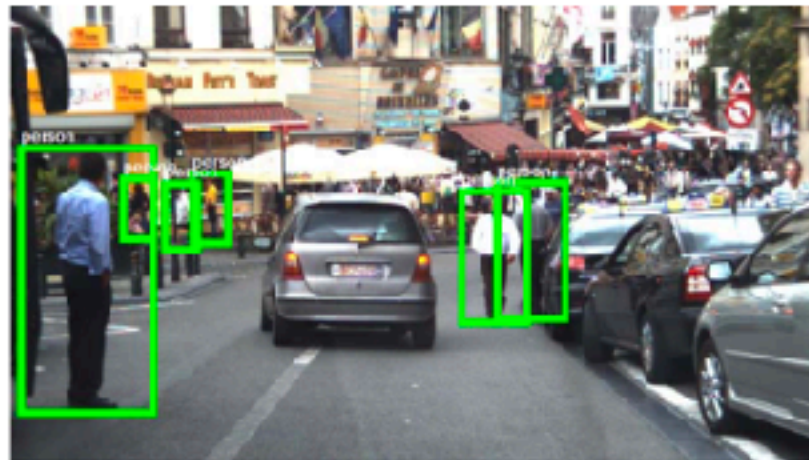[Shechtman-Irani CVPR'07]

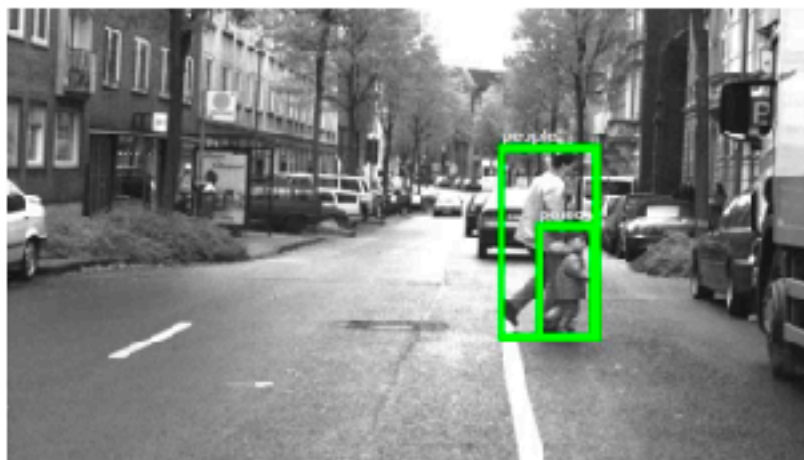# Object classification

(a) Caltech [3]

(b) Caltech-Japan [3]

(c) ETH [4]

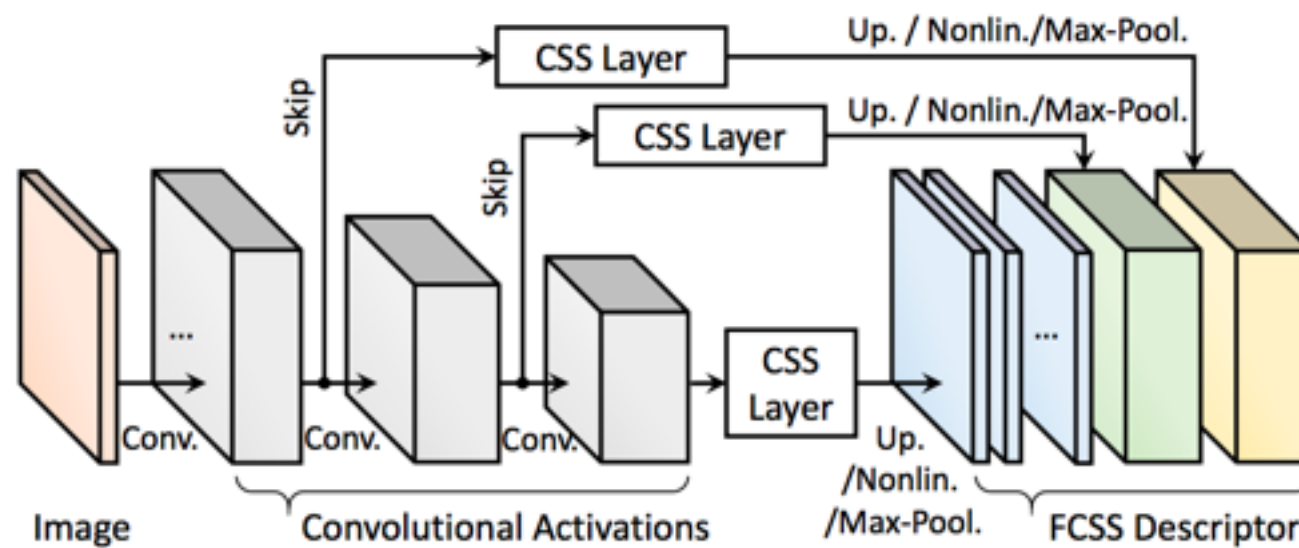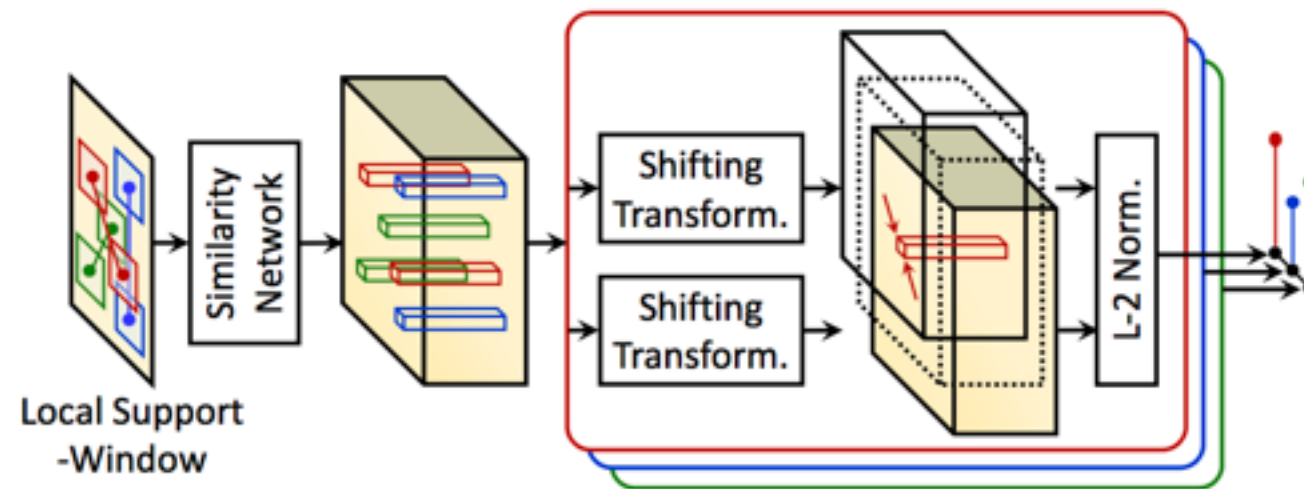(d) TUD-Brussels [5]

(e) Daimler [6]

(f) INRIA [7]

Cross-modal matching

Deep learning

Discussion

- No extensive evaluation on good dataset

Questions?