Interest Point Detectors and Descriptors

IPAM - UCLA

July 24, 2013

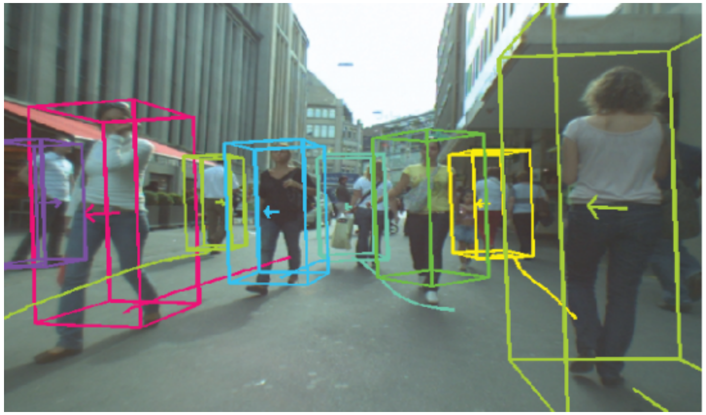
Iasonas Kokkinos
Center for Visual Computing
Ecole Centrale Paris / INRIA Saclay

Image processing

Image to Image

Computer Vision

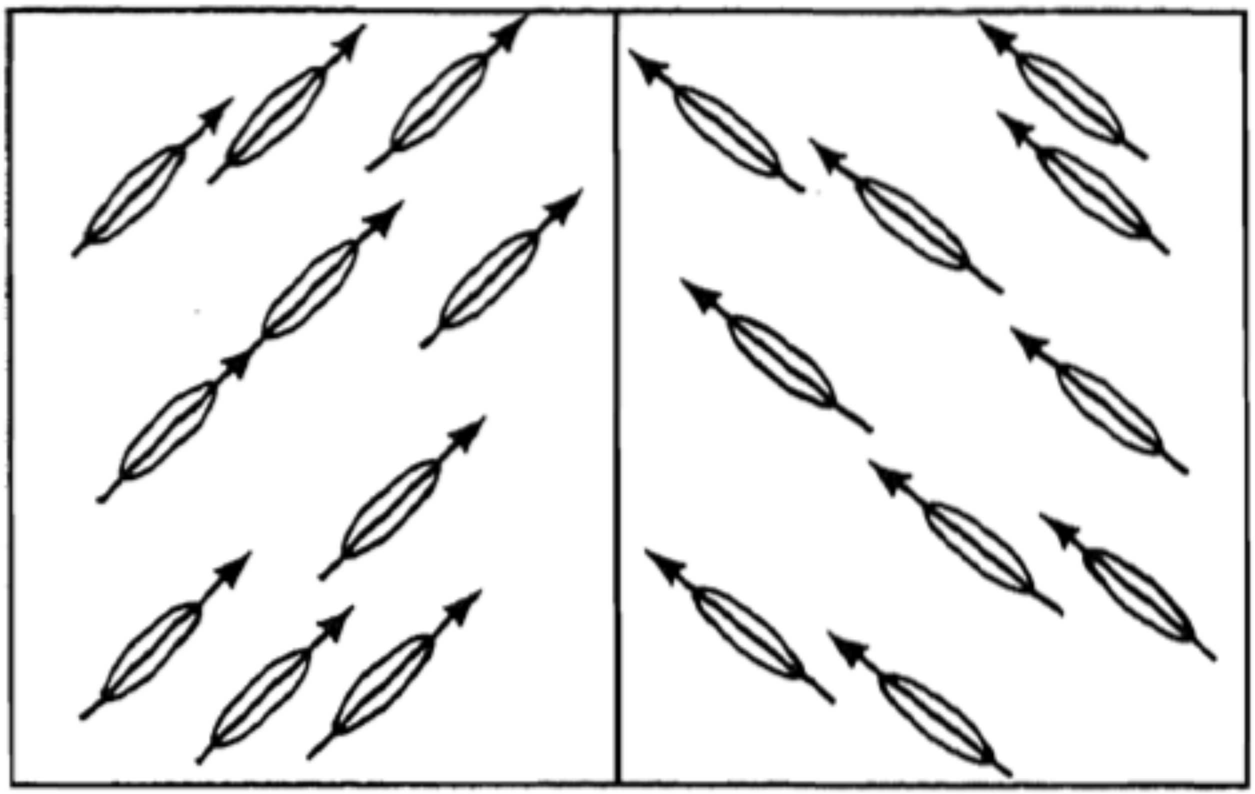
Image to Symbols



Low-level symbolic representations

- Primal sketch

Level 2
boundary



D. Marr, 'Vision', 1982

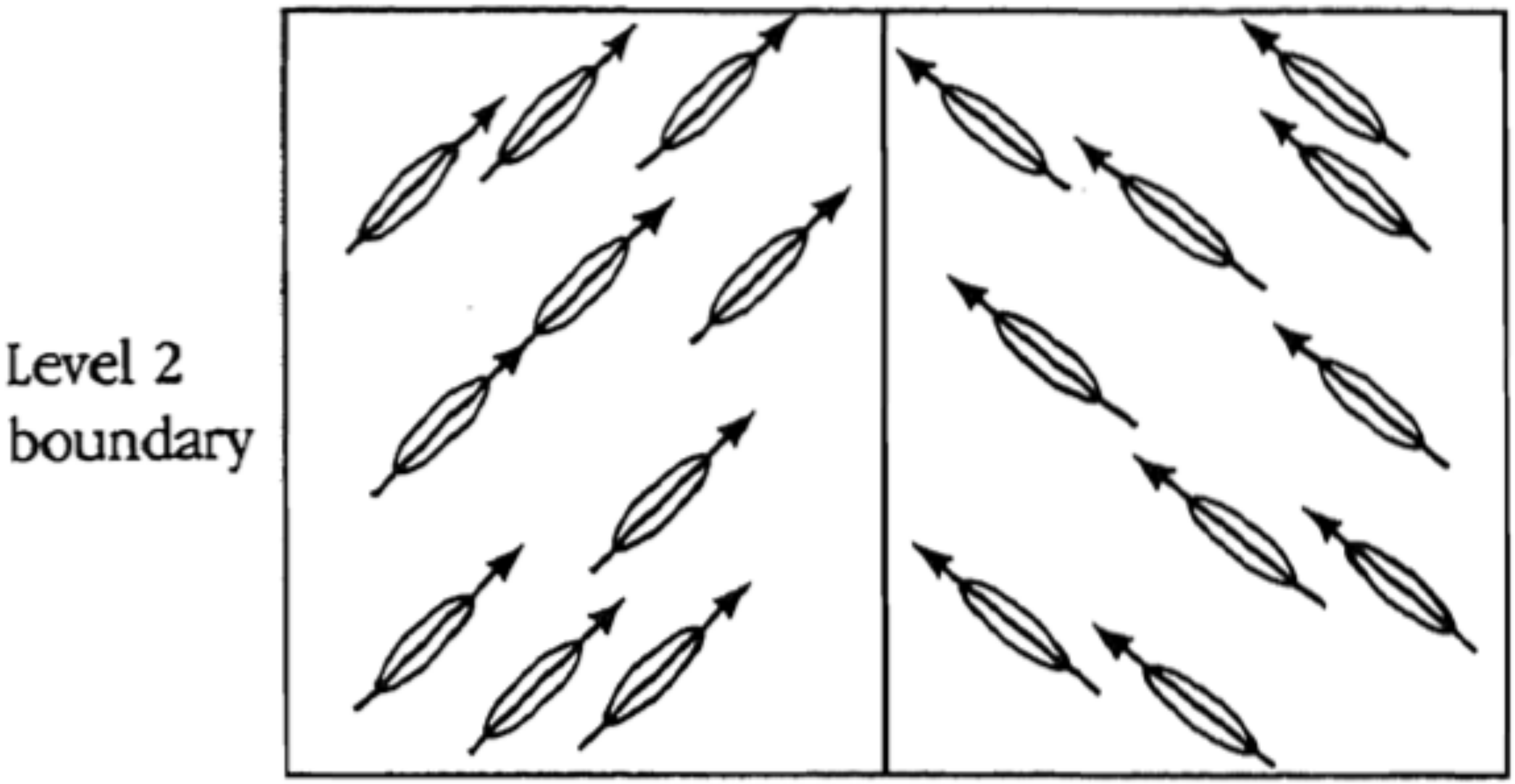
Low-level symbolic representations

- Primal sketch

D. Marr, ‘Vision’, 1982

Low-level symbolic representations

- Primal sketch



D. Marr, 'Vision', 1982

Goals for a low-level representation

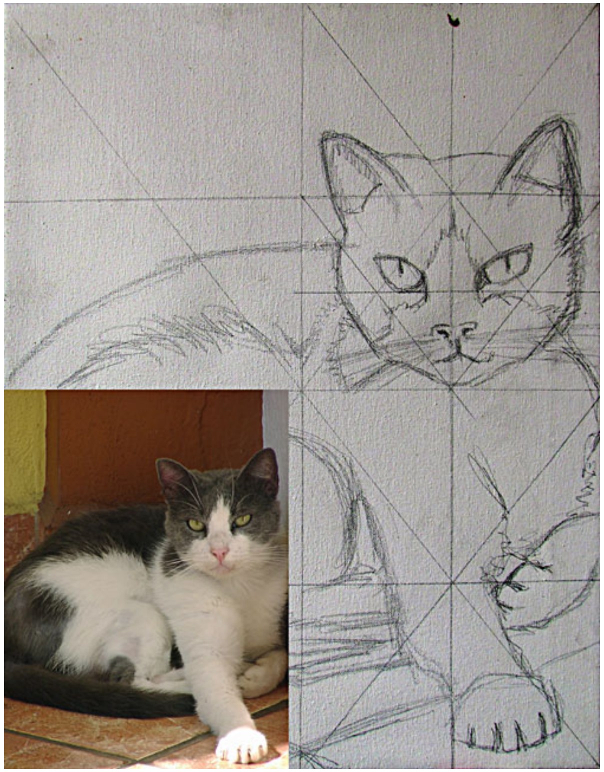
- Compact
  - Reduce the number of processed image locations

    0000000111111000001111111111000

    = 0000000111111000001111111111000

    = 7x0 5x1 5x0 7x1 3x0

- Sufficient
  - No need to look back into the image



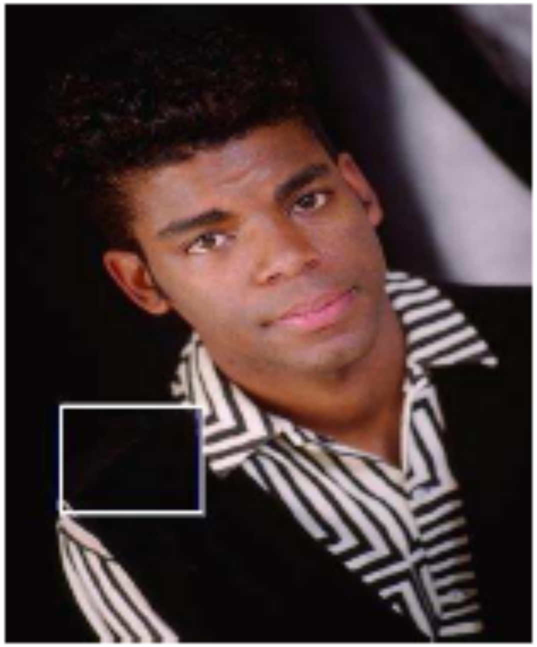
Boundaries

Object/Surface Boundaries

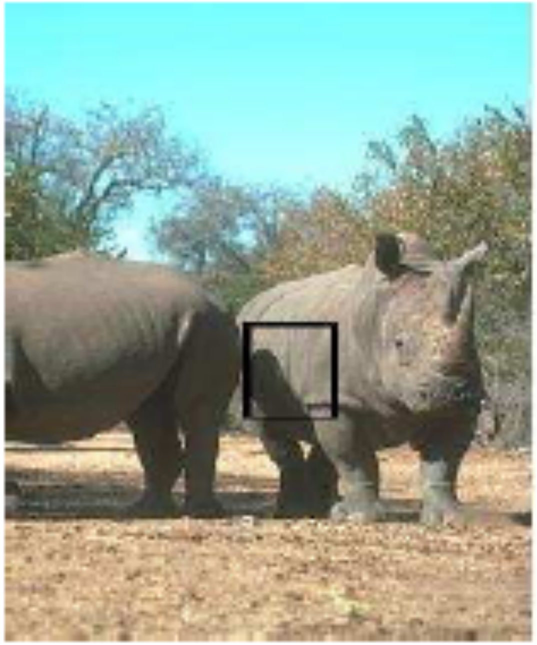
A biref anlaogy wtih txet

- Waht mttares is waht hpapens on wrod baounrdies
- Mocpera iwht htsi sesm
  (compare with this mess)
- Evidence of our visual system employing boundary detection

Signal-level challenges



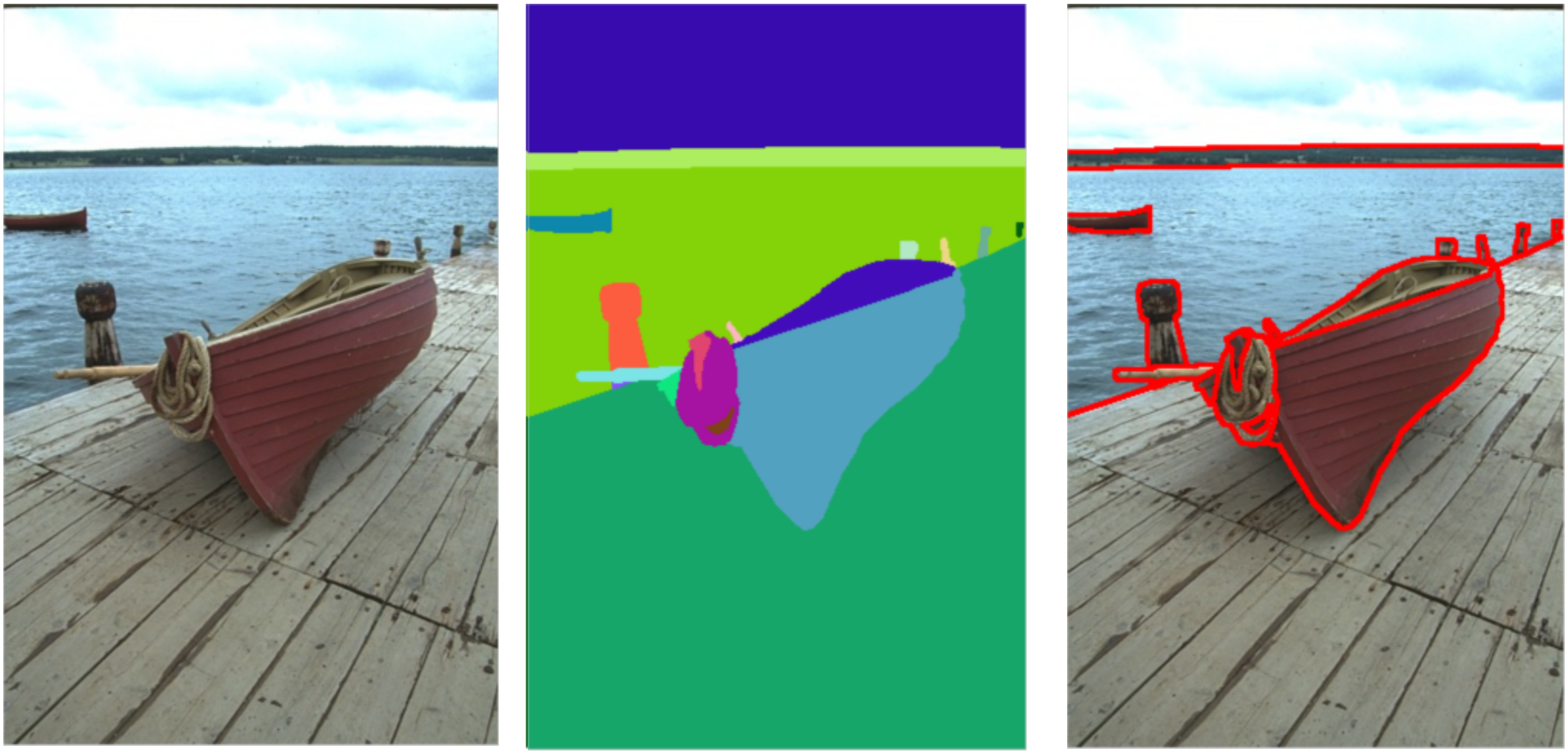
Poor contrast



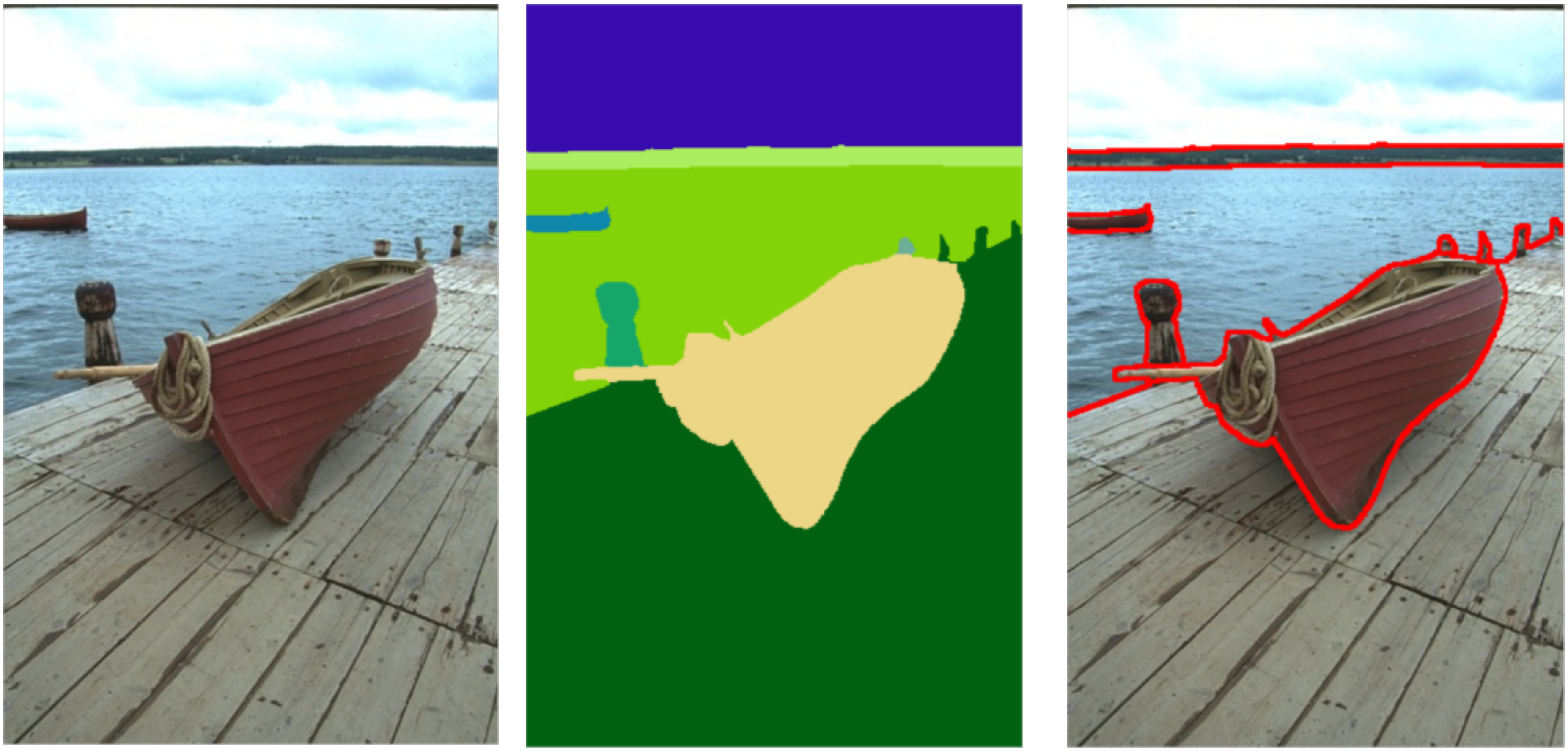
Shadows

Texture

Fundamental challenges: can humans do it?



Fundamental challenges: can humans do it?

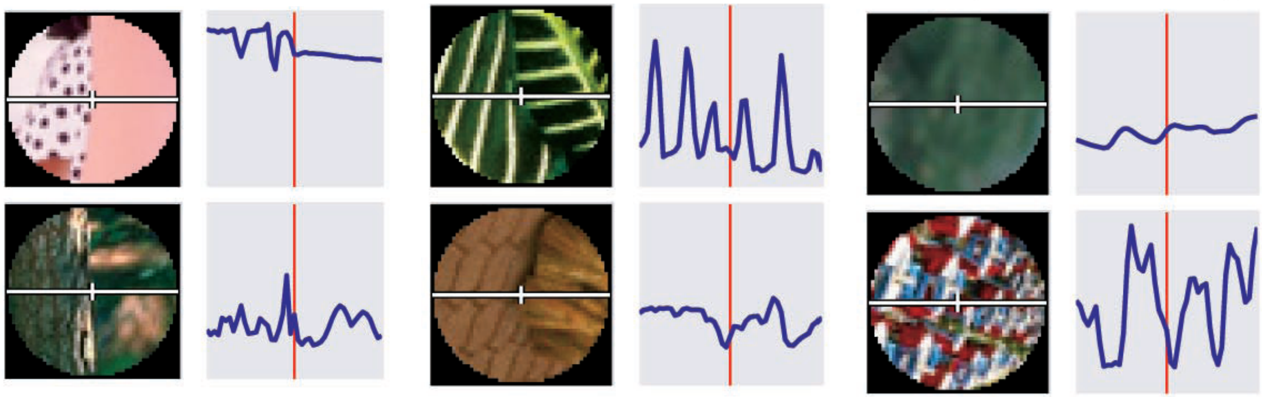


Fundamental challenges: can humans do it?

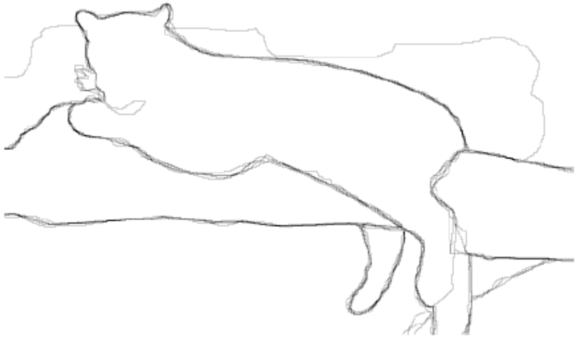


Learning-based approaches

Boundary or non-boundary?



Use human-annotated segmentations

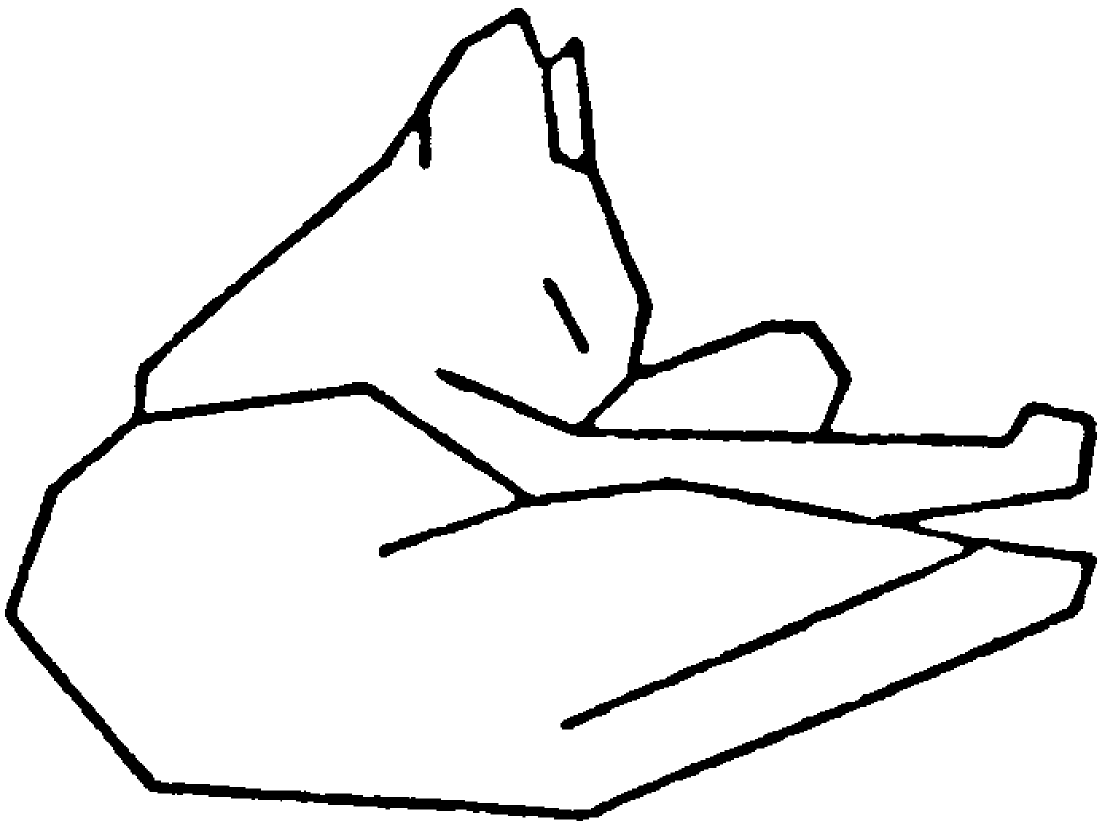


Use any visual cue as input to the decision function.
Use decision trees/logisitic regression/boosting/… and *learn* to combine the individual inputs.

S. Konishi, A.Yuille, J. Coughlan, S.C. Zhu, “Statistical Edge Detection: Learning and Evaluating Edge Cues”, IEEE PAMI, 2003

D. Martin, C. Fowlkes, J. Malik. "Learning to Detect Natural Image Boundaries Using Local Brightness, Color and Texture Cues", IEEE PAMI, 2004

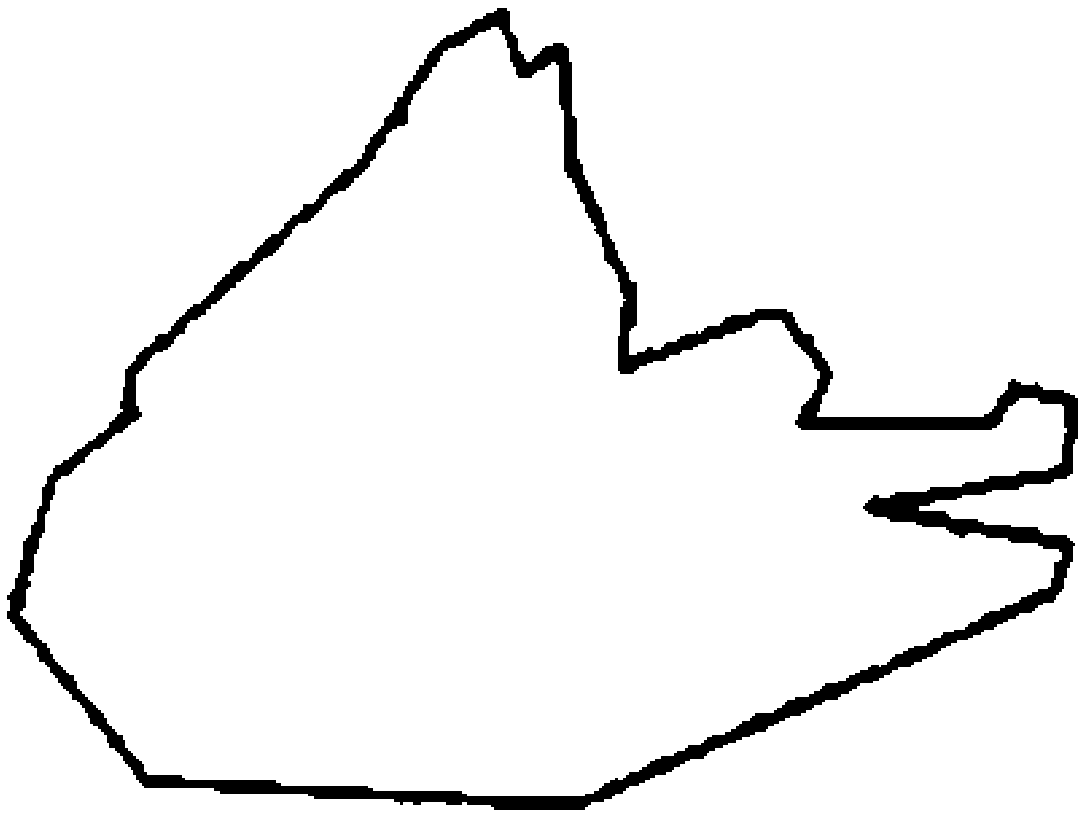
Boundaries or edges?



F. Attneave, "Some informational aspects of visual perception," *Psychological Review,* vol. 61, pp. 183–193, 1954

S. C. Zhu and A. Yuille, "FORMS: A Flexible Object Recognition and Modeling System," *IJCV,* 96

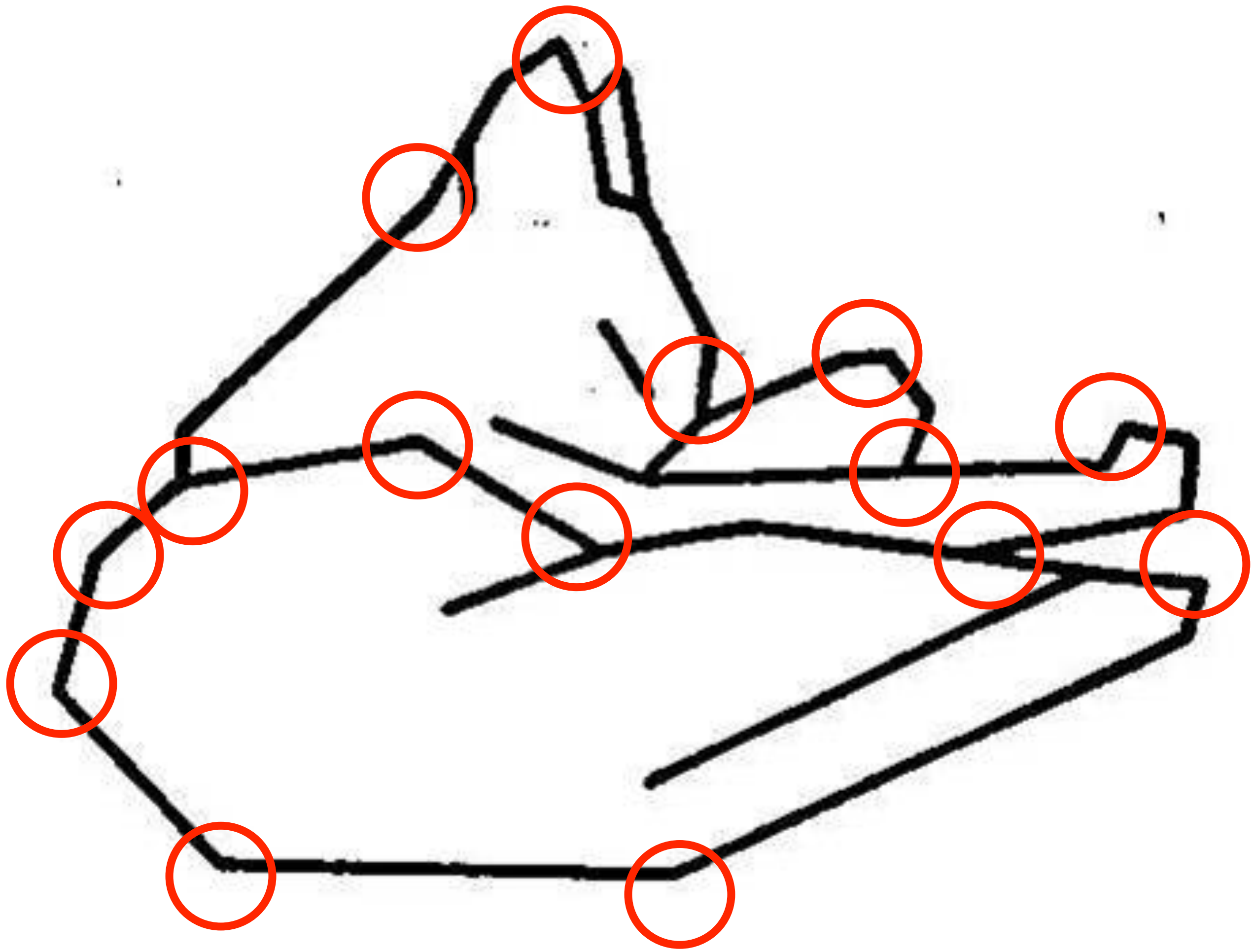
Boundaries or edges?



F. Attneave, “Some informational aspects of visual perception,” *Psychological Review,* vol. 61, pp. 183–193, 1954

S. C. Zhu and A. Yuille, “FORMS: A Flexible Object Recognition and Modeling System,” *IJCV,* 96

Contours or points?

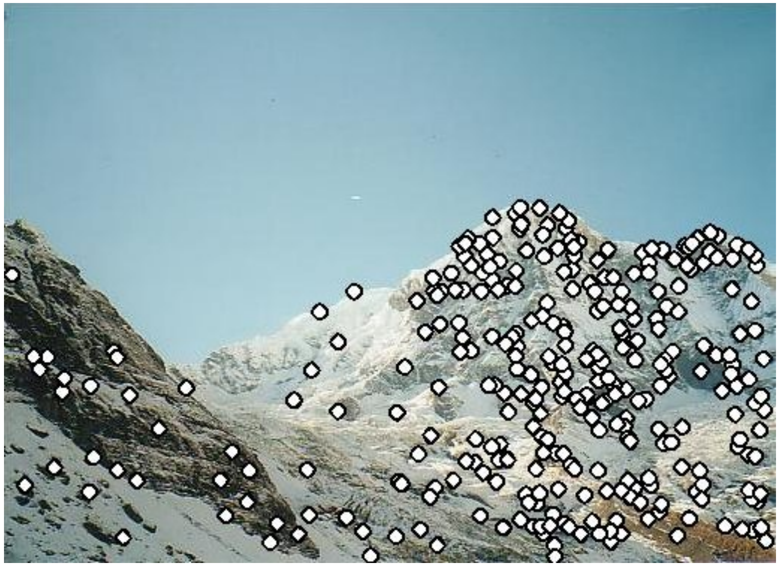
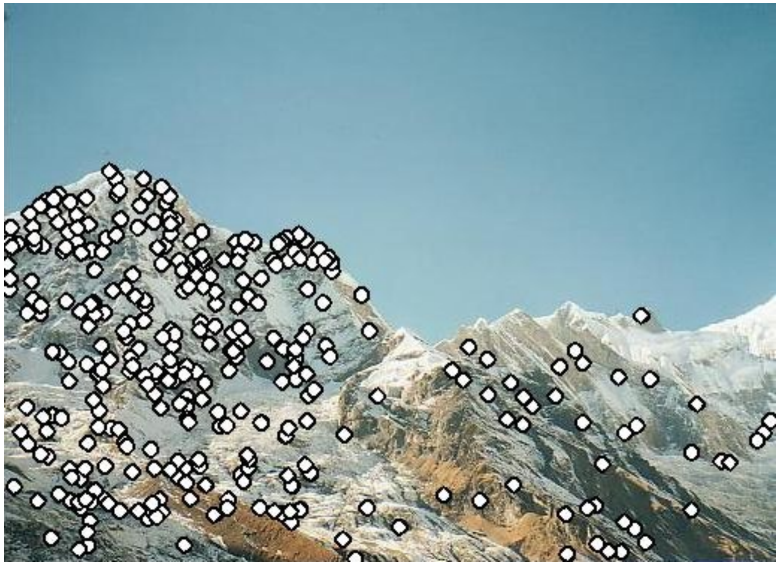


F. Attneave, “Some informational aspects of visual perception,” *Psychological Review,* vol. 61, pp. 183–193, 1954

Application: Image Stitching

Slide credit: Darya Frolova, Denis Simakov

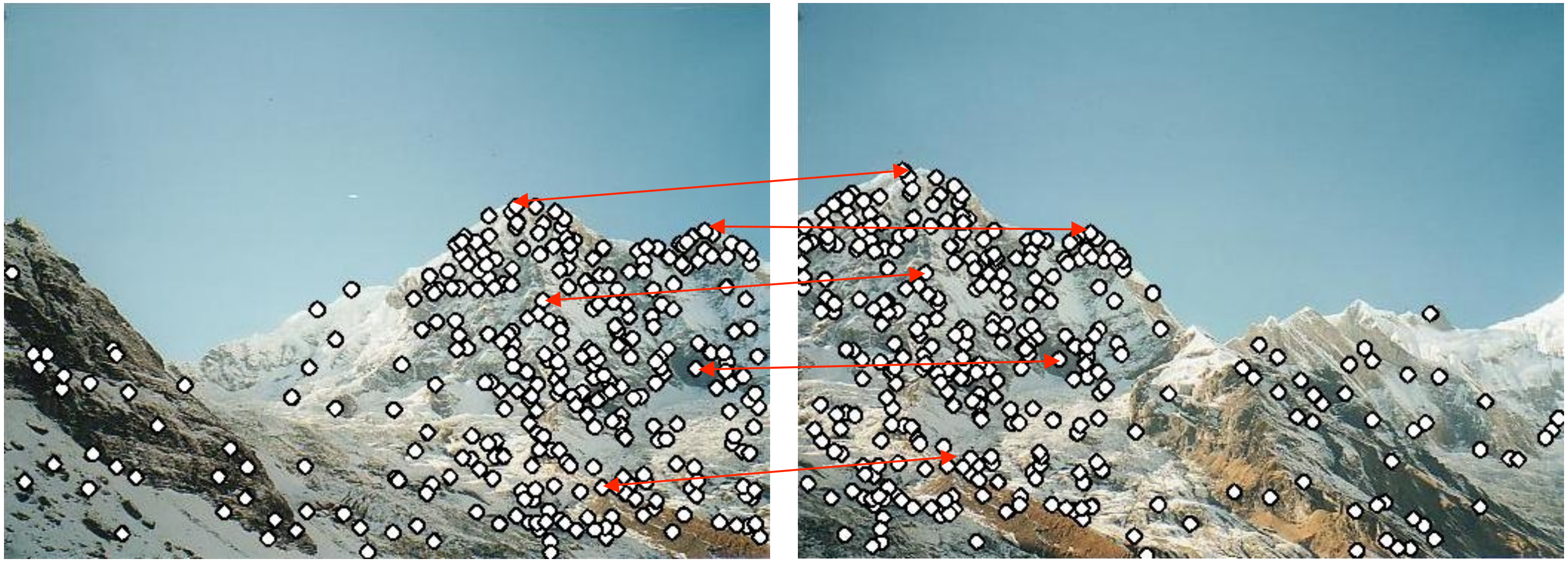
Application: Image Stitching

- Procedure:
  - Detect feature points in both images

Slide credit: Darya Frolova, Denis Simakov

Application: Image Stitching



- Procedure:
  - Detect feature points in both images
  - Find corresponding pairs

Slide credit: Darya Frolova, Denis Simakov

Application: Image Stitching

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

Slide credit: Darya Frolova, Denis Simakov

Common Requirements

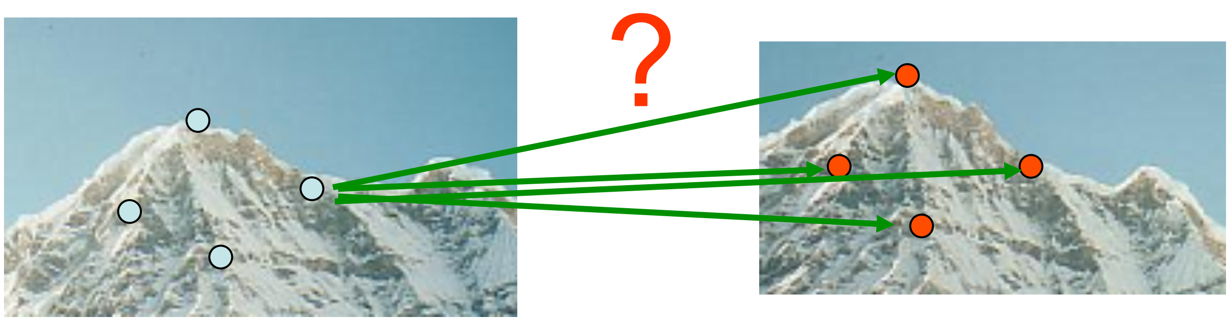
- Problem 1:
  - Detect the same point *independently* in both images

No chance to match!

Slide credit: Darya Frolova, Denis Simakov

Common Requirements

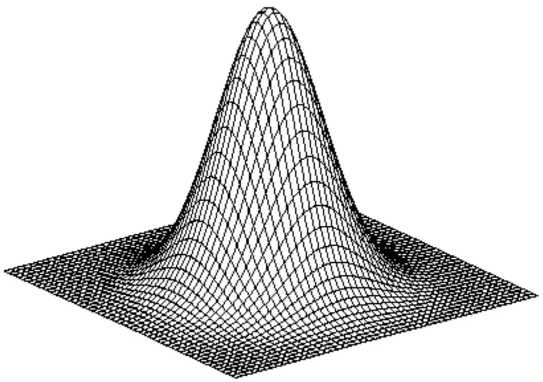
- Problem 1:
  - Detect the same point *independently* in both images
- Problem 2:
  - For each point correctly recognize the corresponding one



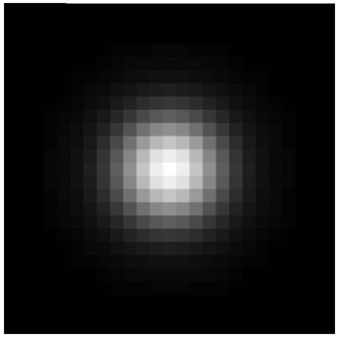
Slide credit: Darya Frolova, Denis Simakov

Laplacian-of-Gaussian

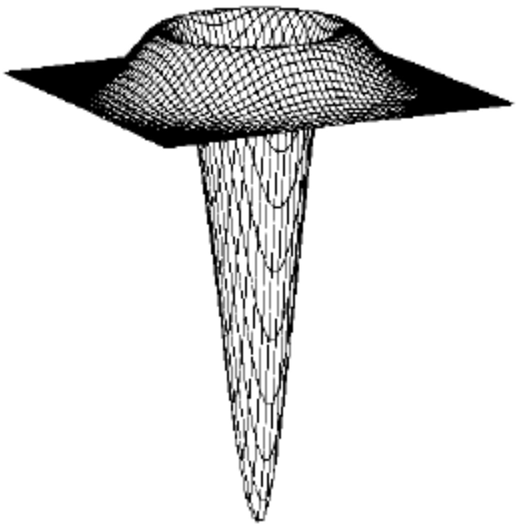
Gaussian



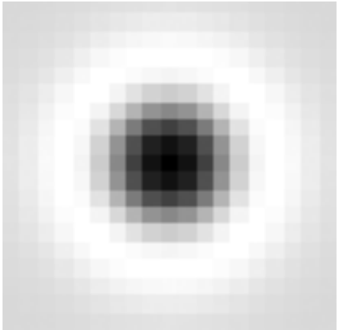
$$g_\sigma(x, y)$$



Laplacian of Gaussian

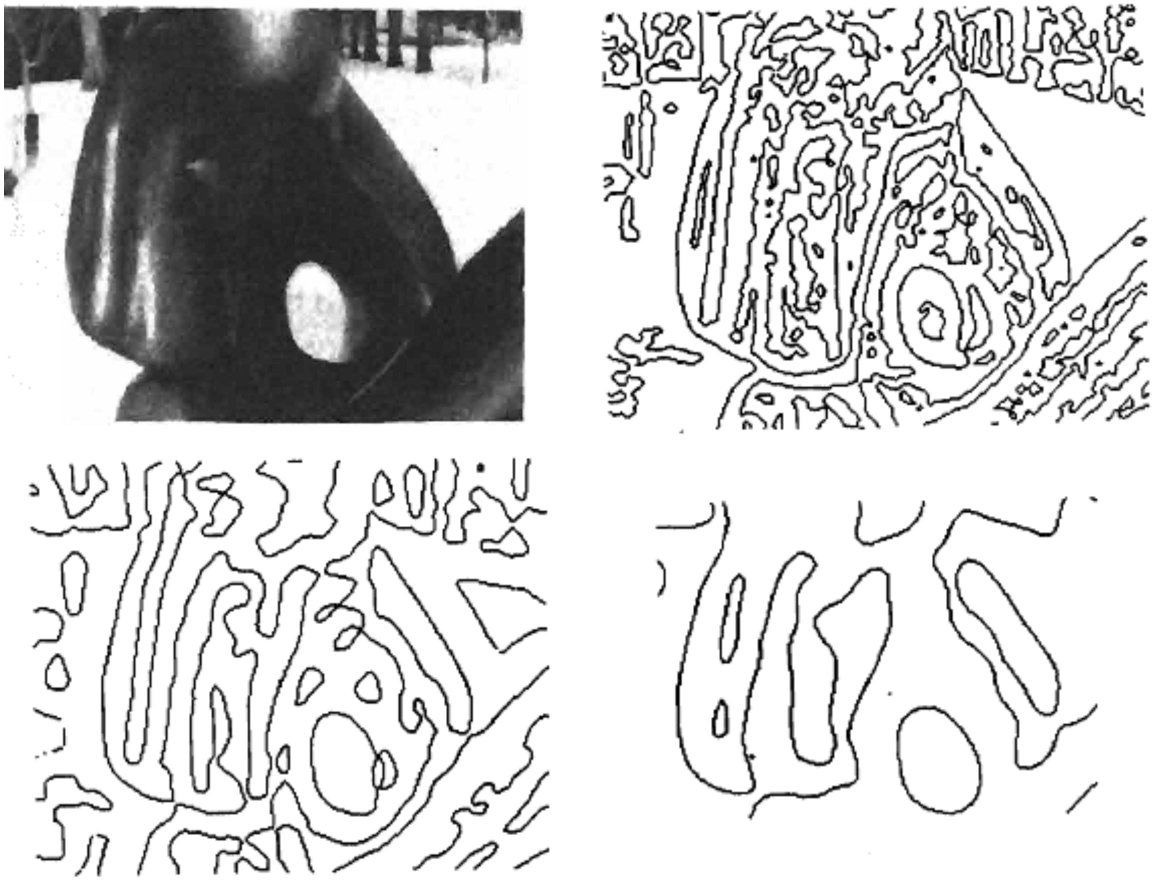


$$\nabla^2 g_\sigma(x, y) = \frac{\partial^2 g_\sigma(x, y)}{\partial x^2} + \frac{\partial^2 g_\sigma(x, y)}{\partial y^2}$$



Early edge detection research

- Zero-crossings of LoG operator at increasing scales

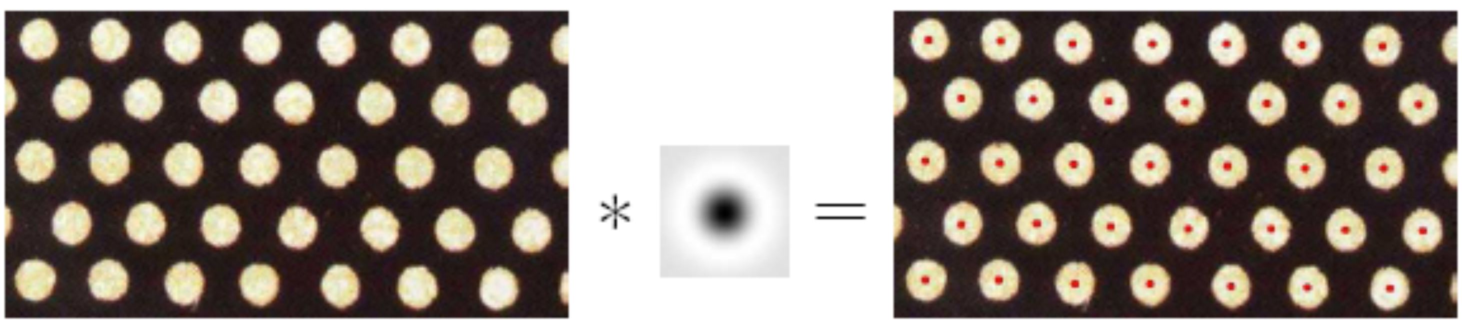


- Different take: go for the maxima/minima

Finding blobs

Filtering= inner product between image patch and filter: template matching

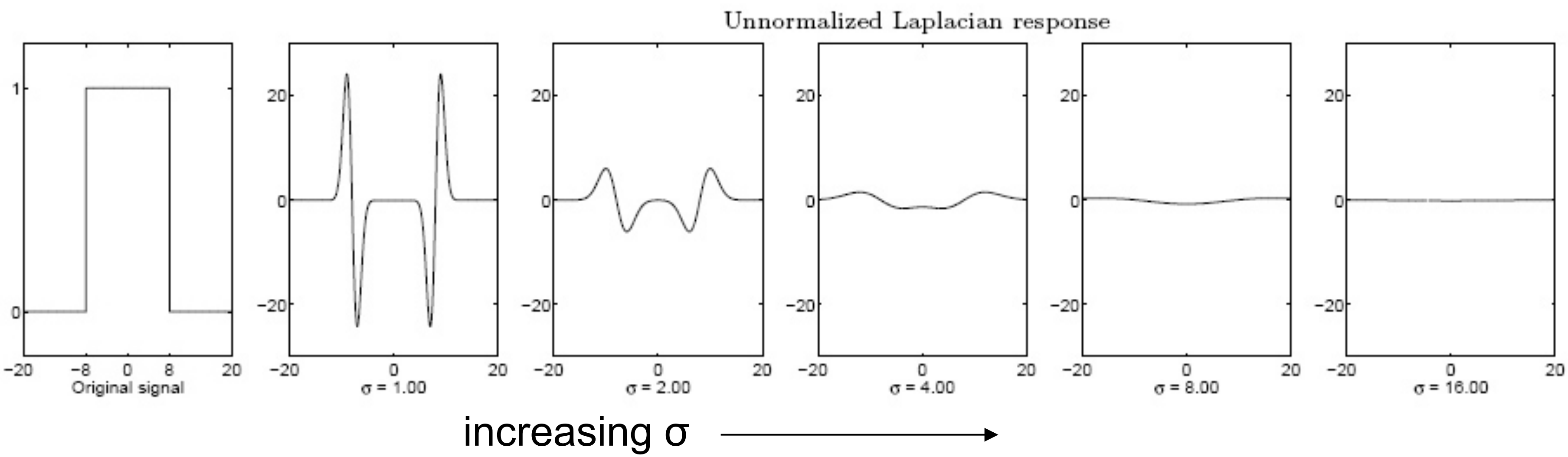
$$
\begin{aligned}
|I-f|^2 &= \langle I-f, I-f\rangle \\
&= \langle I, I\rangle + \langle f, f\rangle - 2\langle f, I\rangle \\
&= C - 2\langle I, f\rangle
\end{aligned}
$$



Polka Dots $*$ filter $=$ Detected Blobs

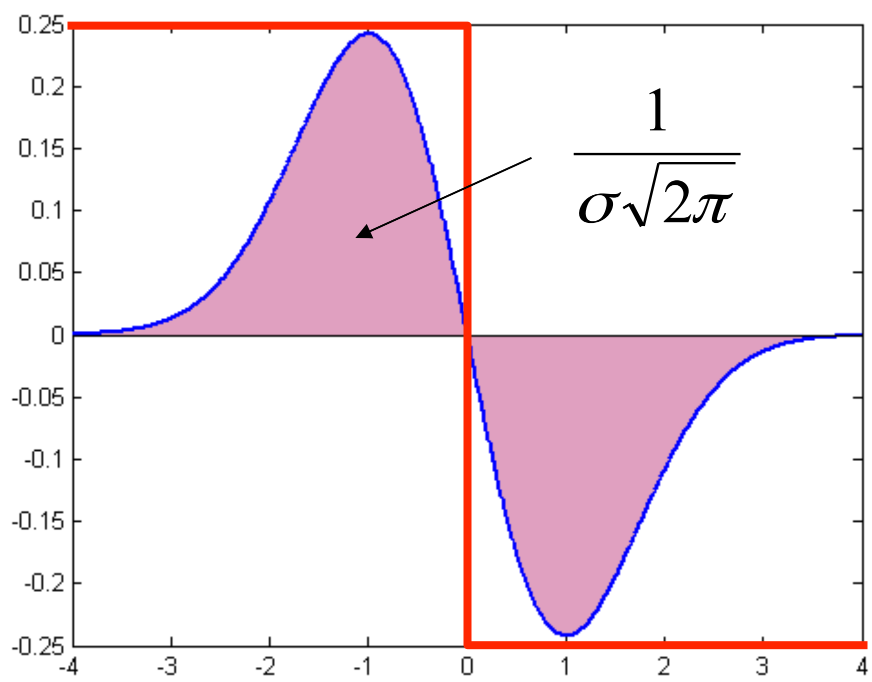
Scale selection

- First idea: convolve with Laplacians at several scales and find maximum in scale
- Observation: Laplacian decays as scale increases:



Scale normalization

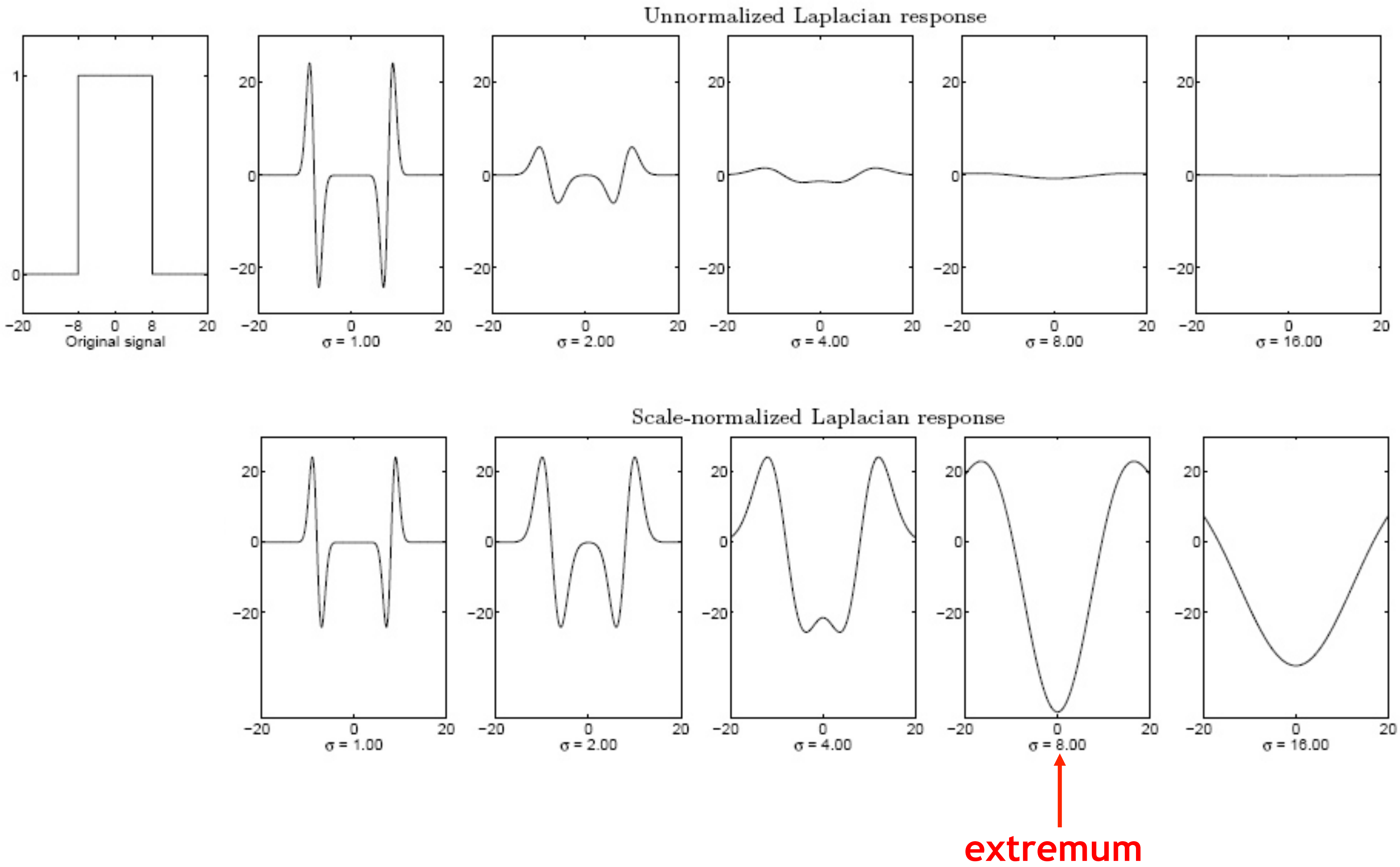
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases



Scale normalization

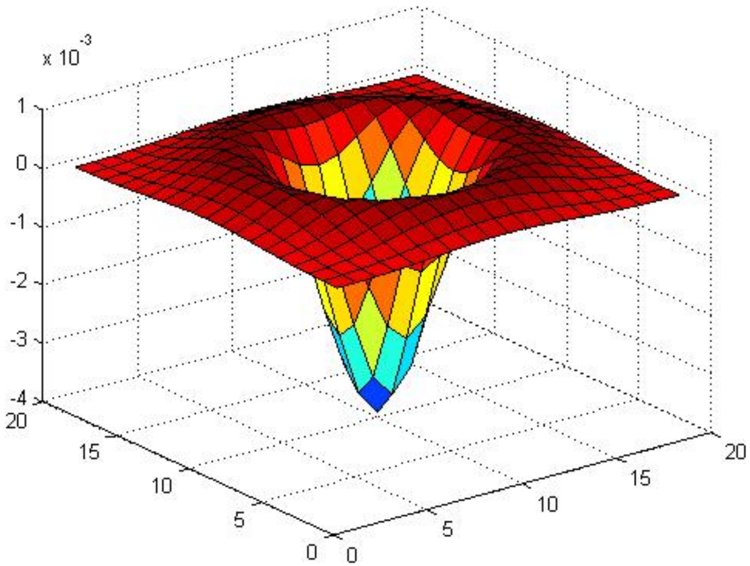
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases
- To keep response the same (scale-invariant), must multiply Gaussian derivative by σ
- Laplacian is the second Gaussian derivative, so it must be multiplied by σ^2

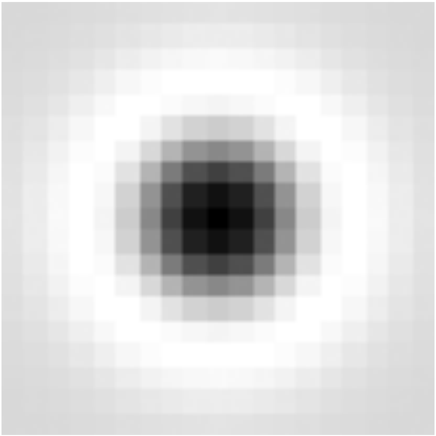
Effect of scale normalization



Blob detection in 2D

Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D

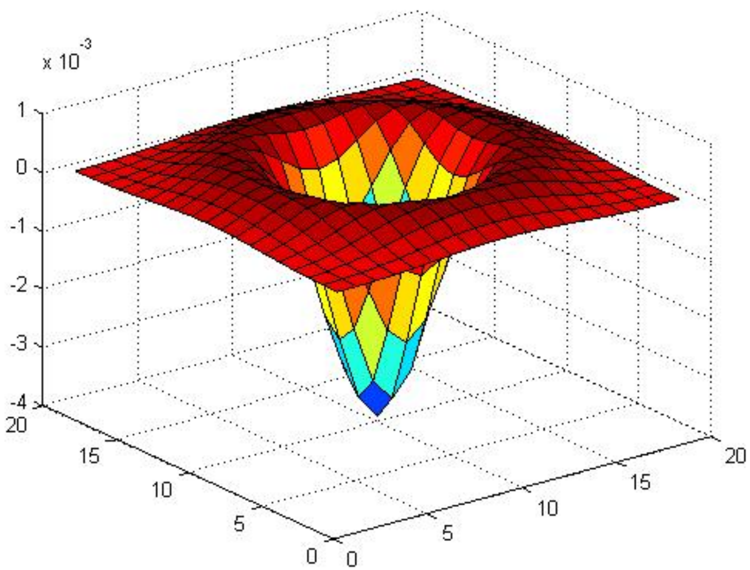


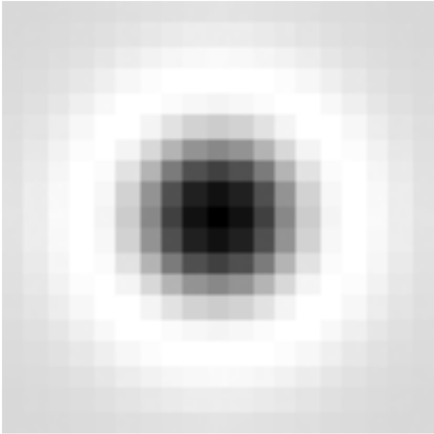


$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

Blob detection in 2D

Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D



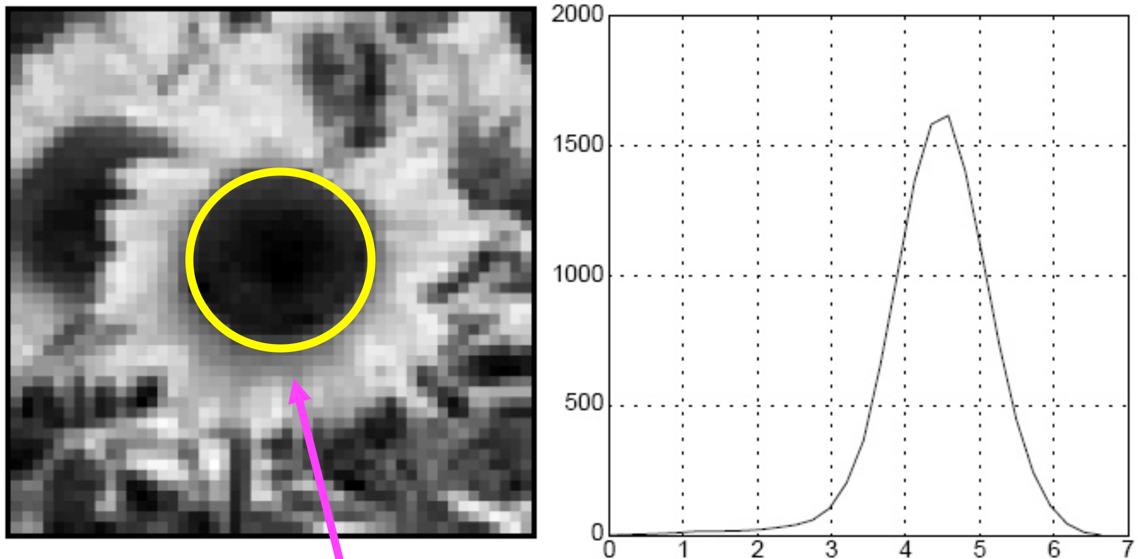


Scale-normalized:

$$\nabla^2_{\text{norm}} g = \sigma^2 \left(\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} \right)$$

Scale selection

- Characteristic scale: peak of normalized Laplacian response

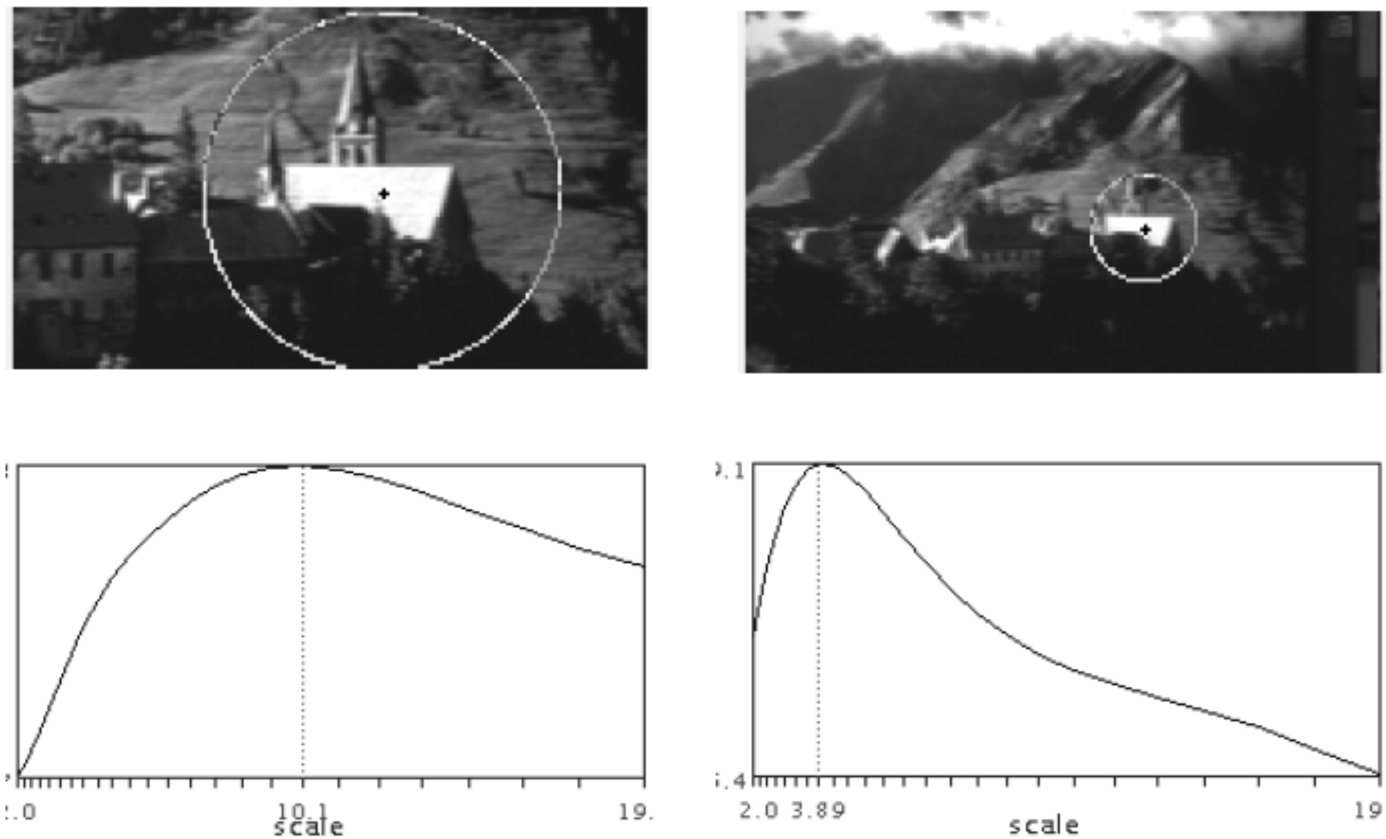


characteristic scale

Tony Lindeberg: Feature Detection with Automatic Scale Selection. International Journal of Computer Vision 30(2): 79-116 (1998)

Tony Lindeberg: Edge Detection and Ridge Detection with Automatic Scale Selection. International Journal of Computer Vision 30(2): 117-156 (1998)

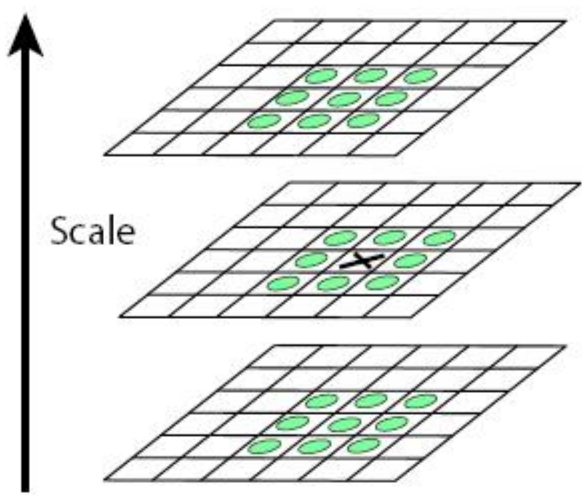
Scale invariance using scale selection



Laplacian

Scale-space blob detector

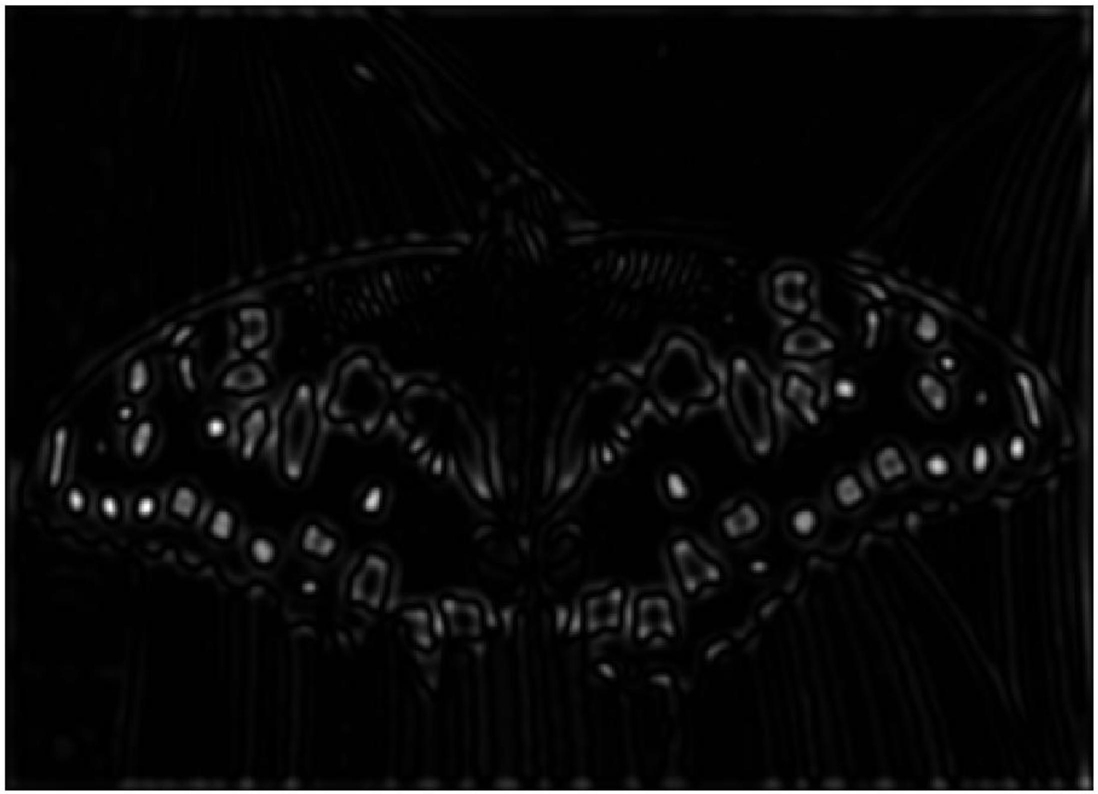
1. Convolve image with scale-normalized Laplacian at several scales
2. Find maxima of squared Laplacian response in scale-space



Scale-space blob detector: Example

Slide credit: S. Lazebnik

Scale-space blob detector: Example



sigma = 3.1296

Slide credit: S. Lazebnik

Scale-space blob detector: Example

sigma = 4.8972

Slide credit: S. Lazebnik

Scale-space blob detector: Example

sigma = 7.6631

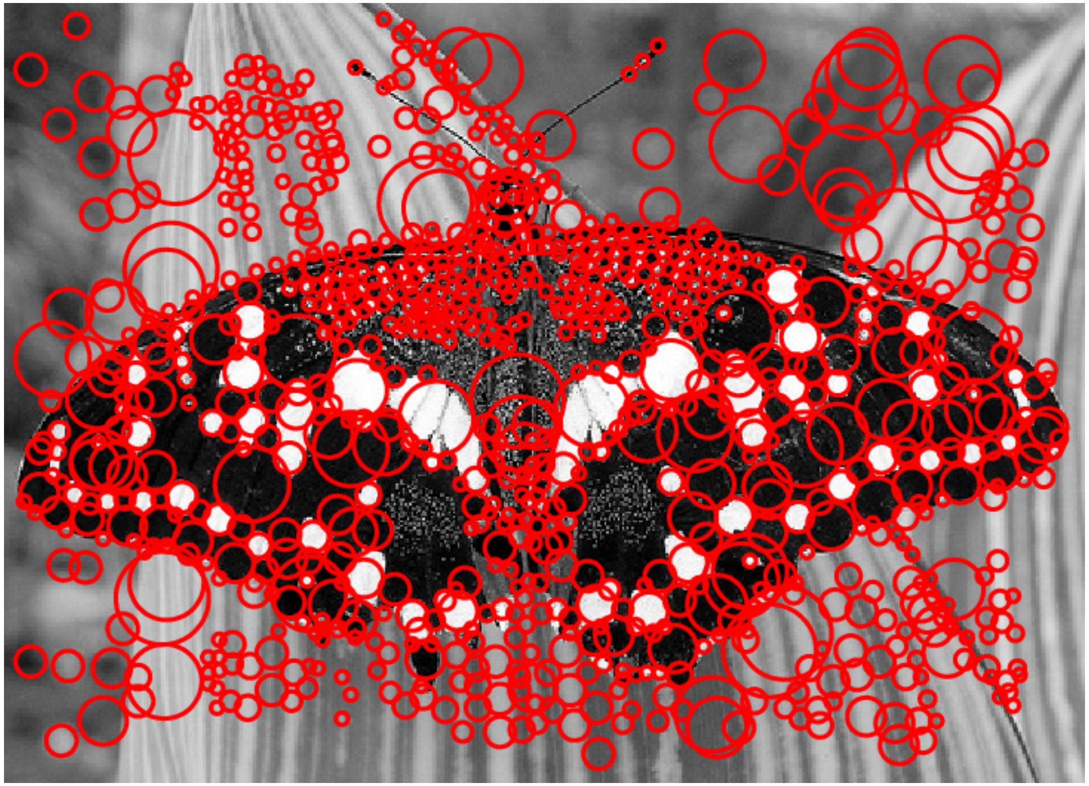
Slide credit: S. Lazebnik

Scale-space blob detector: Example

sigma = 11.9912

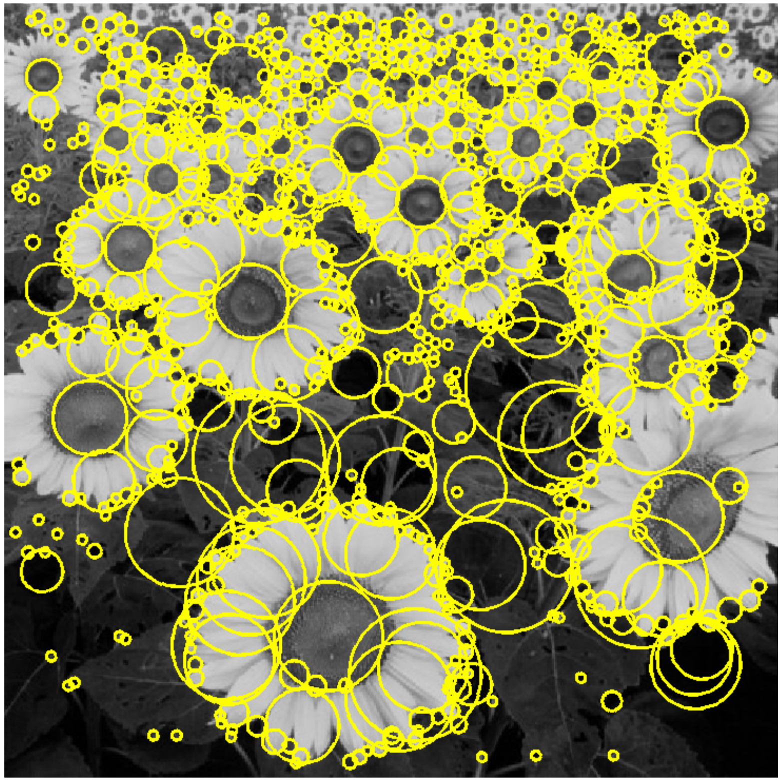
Slide credit: S. Lazebnik

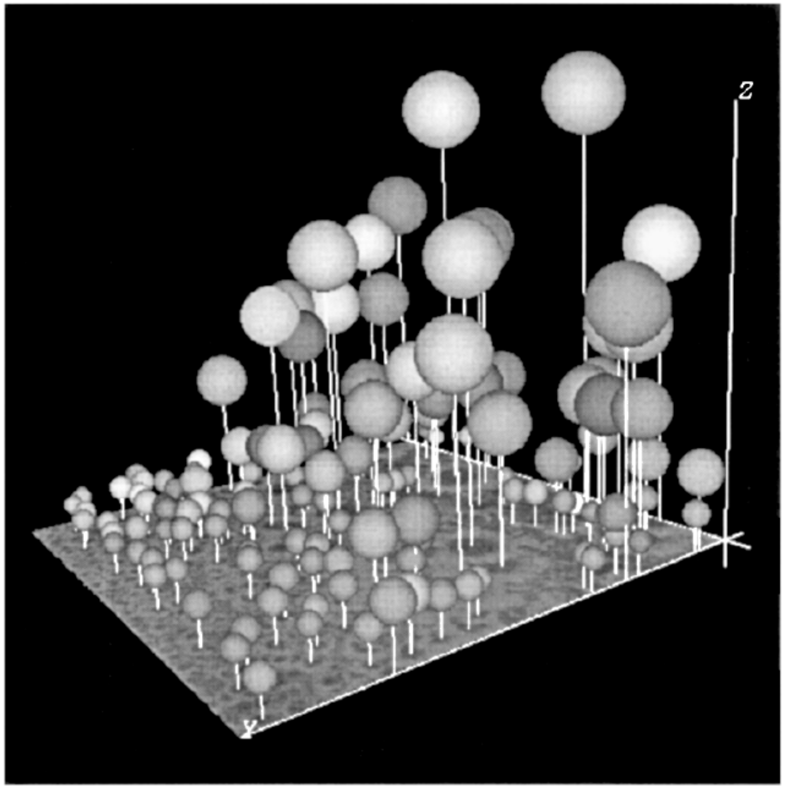
Scale-space blob detector: Example



Tony Lindeberg: Feature Detection with Automatic Scale Selection. International Journal of Computer Vision 30(2): 79-116 (1998)

Blob coordinates: (x,y,scale)





Tony Lindeberg: Feature Detection with Automatic Scale Selection. International Journal of Computer Vision 30(2): 79-116 (1998)

Laplacian of Gaussian ~= Difference of Gaussian

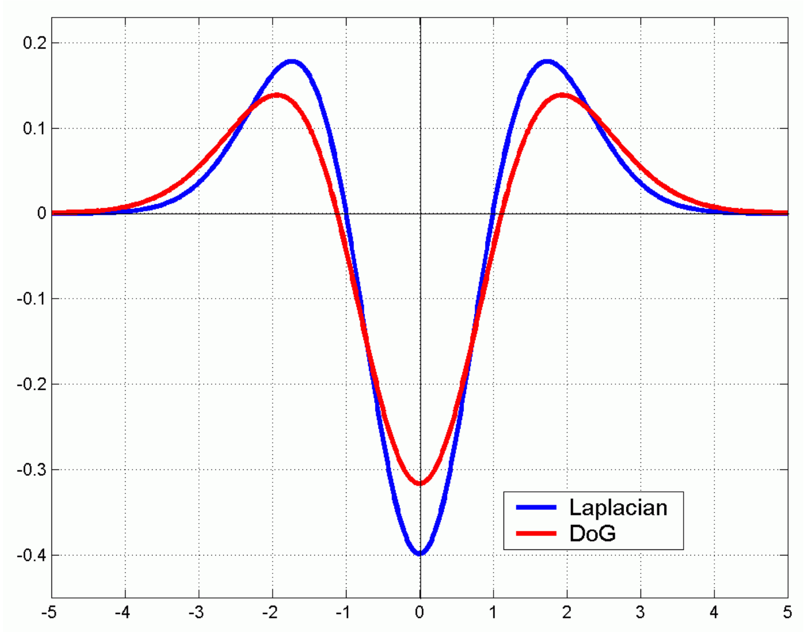
- We can efficiently approximate the Laplacian with a difference of Gaussians:

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

(Laplacian)

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

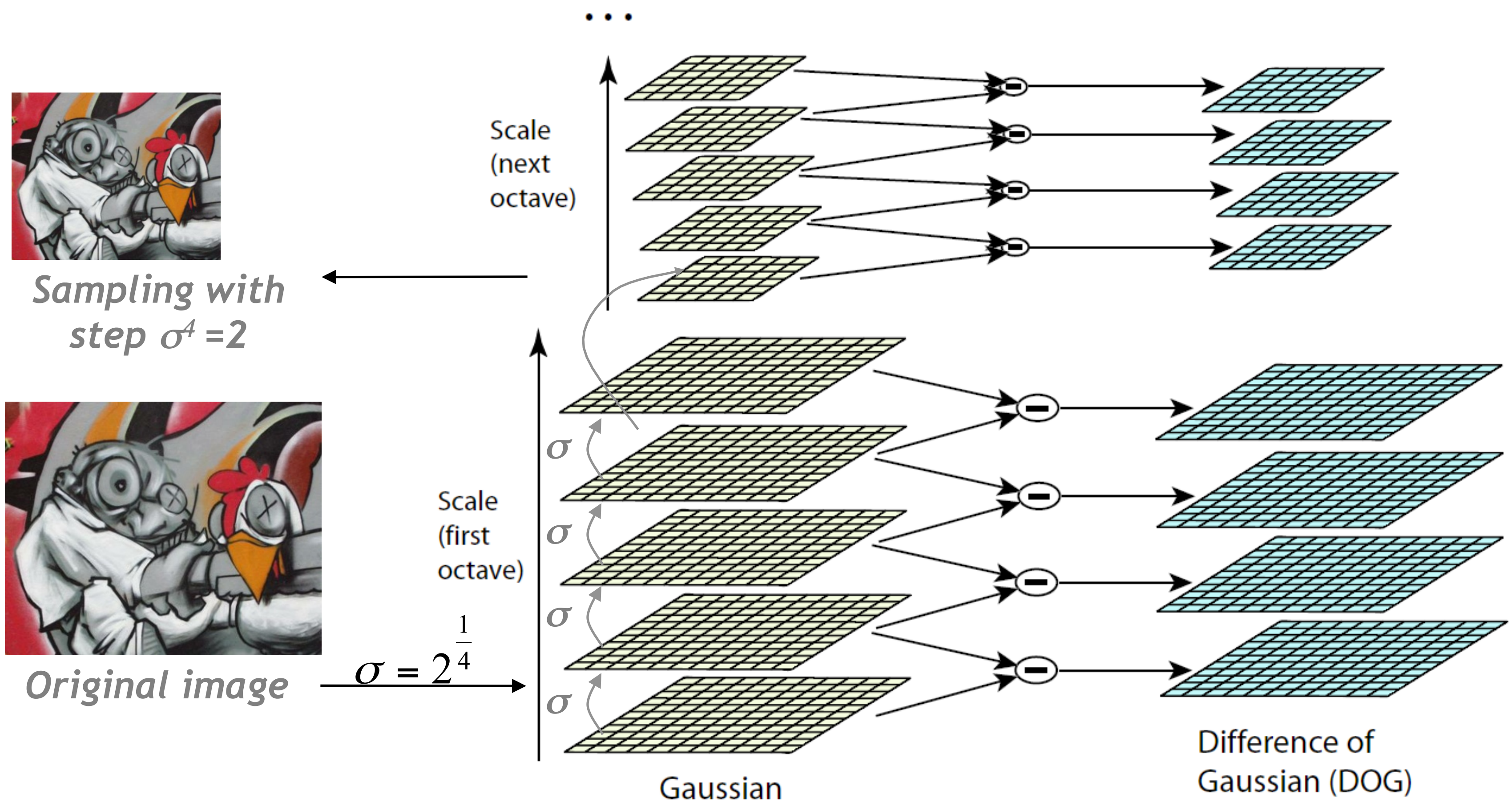
(Difference of Gaussians)



David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60(2): 91-110 (2004)

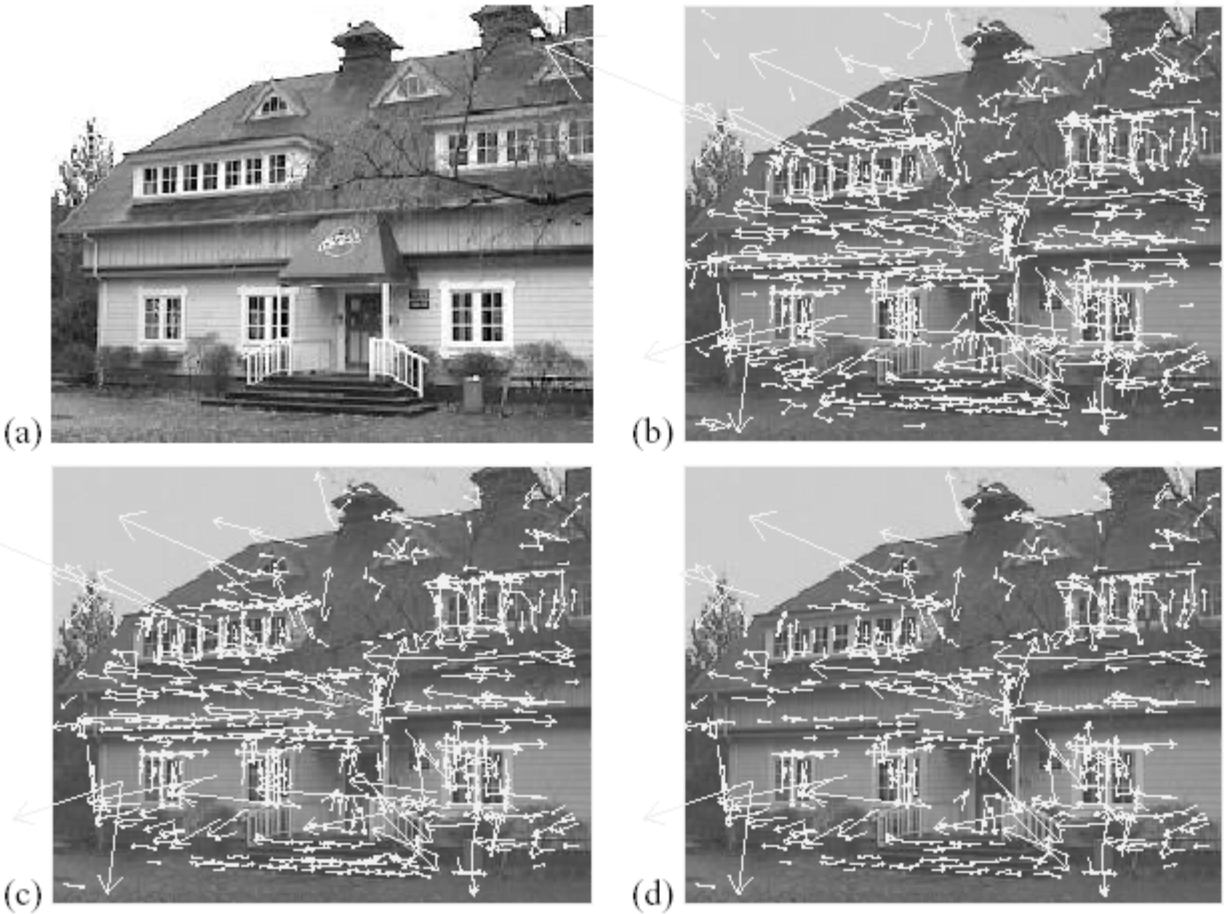
Efficient Computation (SIFT)

- Computation in Gaussian scale pyramid



David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60(2): 91-110 (2004)

Keypoind Detection (SIFT)



(a) 233x189 image

(b) 832 DoG extrema

(c) 729 left after peak value threshold

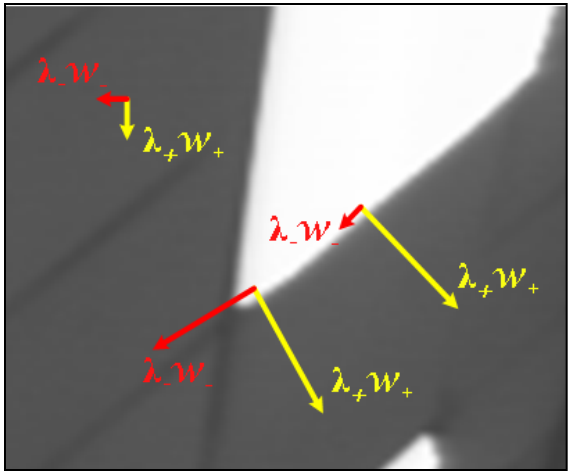
(d) 536 left after testing ratio of principle curvatures (removing edge responses)

David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60(2): 91-110 (2004)

Second Moment Matrix

$$J = G_\rho * \left[(\nabla G_\sigma * u)^T (\nabla G_\sigma * u) \right]$$

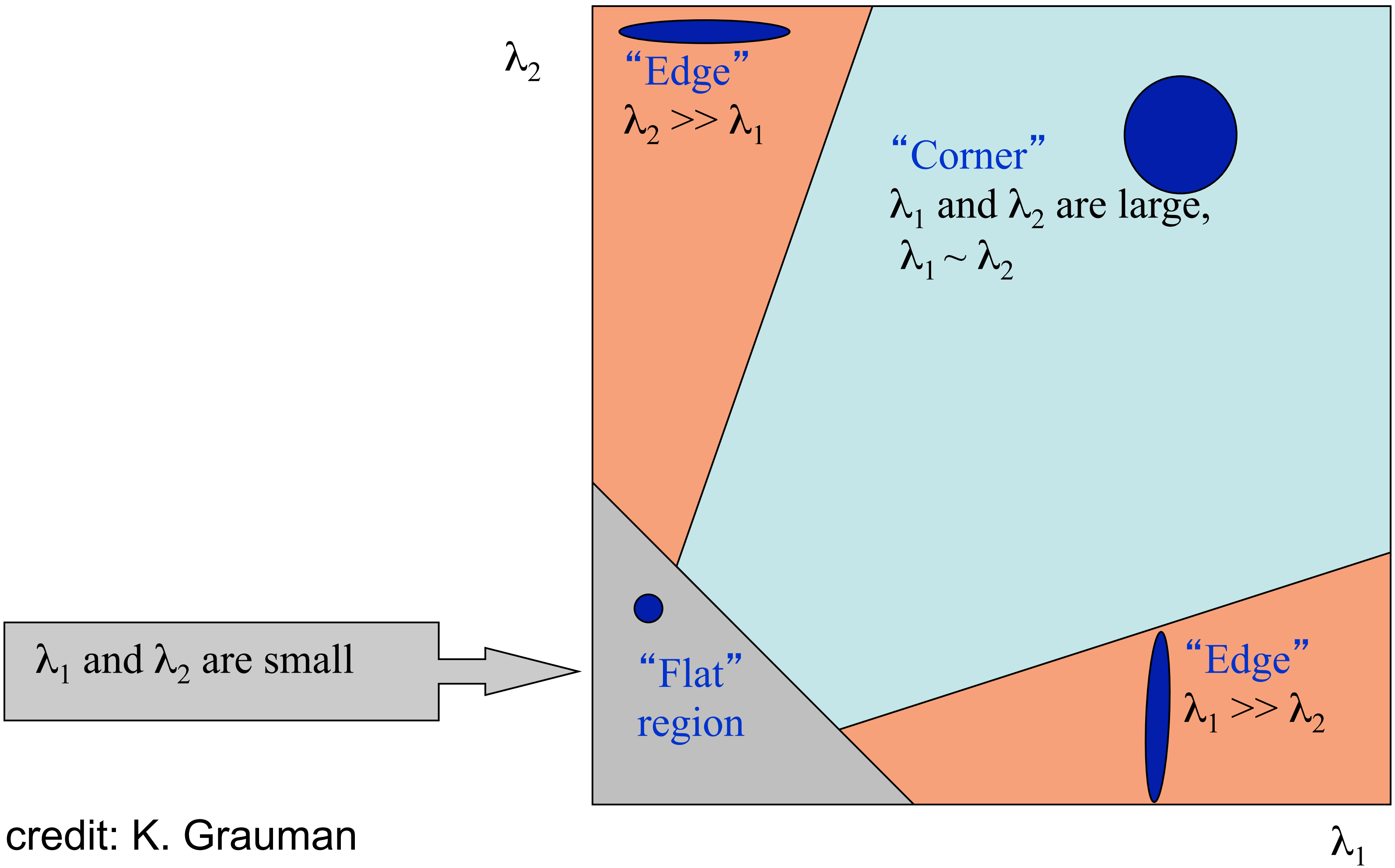
- Eigenvectors w_+, w_- : directions of maximal and minimal variation of u
- Eigenvalues: amounts of minimal and maximal variation u



C.Harris and M.Stephens. "A Combined Corner and Edge Detector." *Proceedings of the 4th Alvey Vision Conference*, 1988.

Interpreting the eigenvalues

Classification of image points using eigenvalues of the second moment matrix:



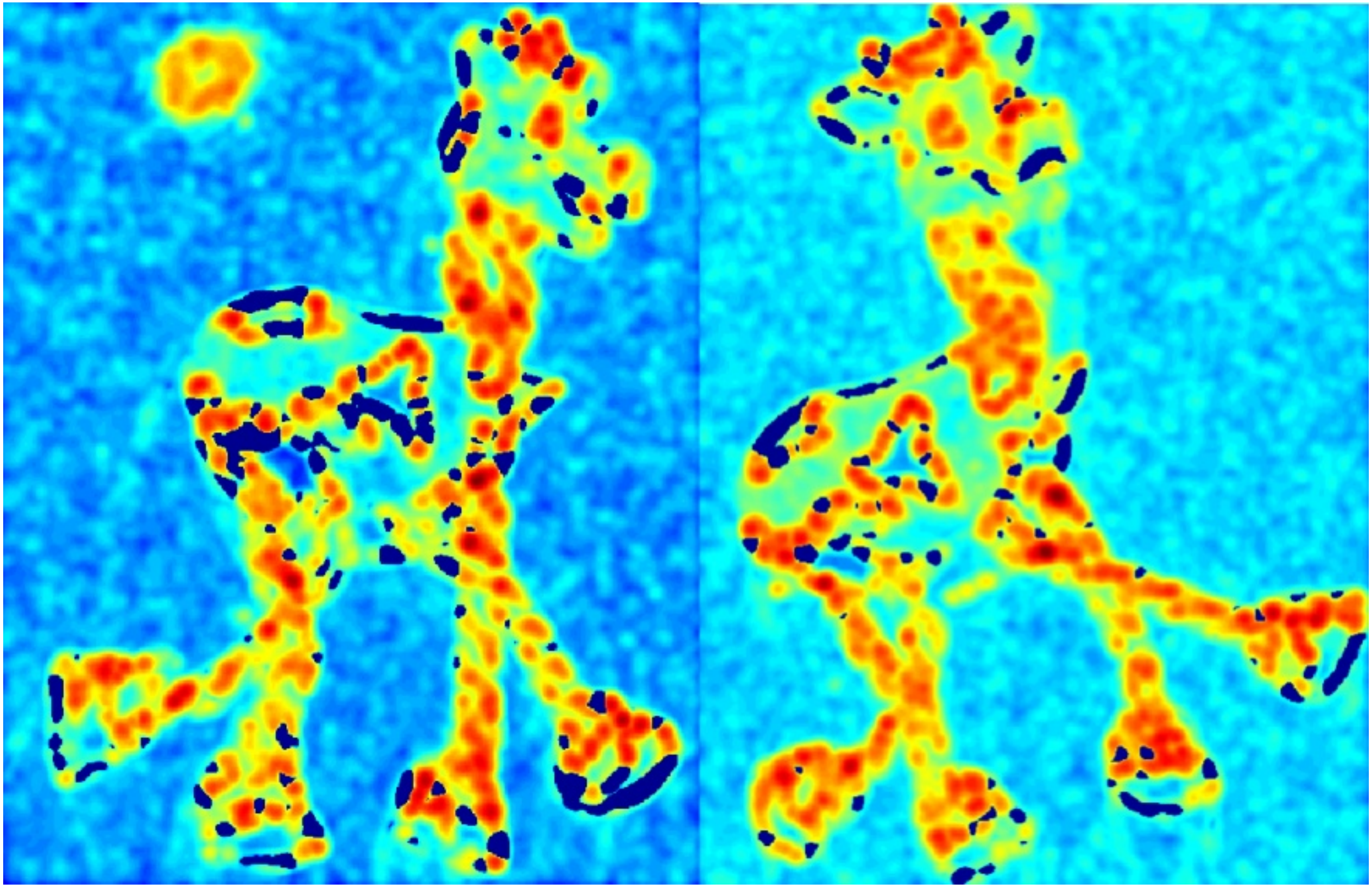
Slide credit: K. Grauman

Harris Detector: Steps

Slide credit: S. Lazebnik

Harris Detector: Steps

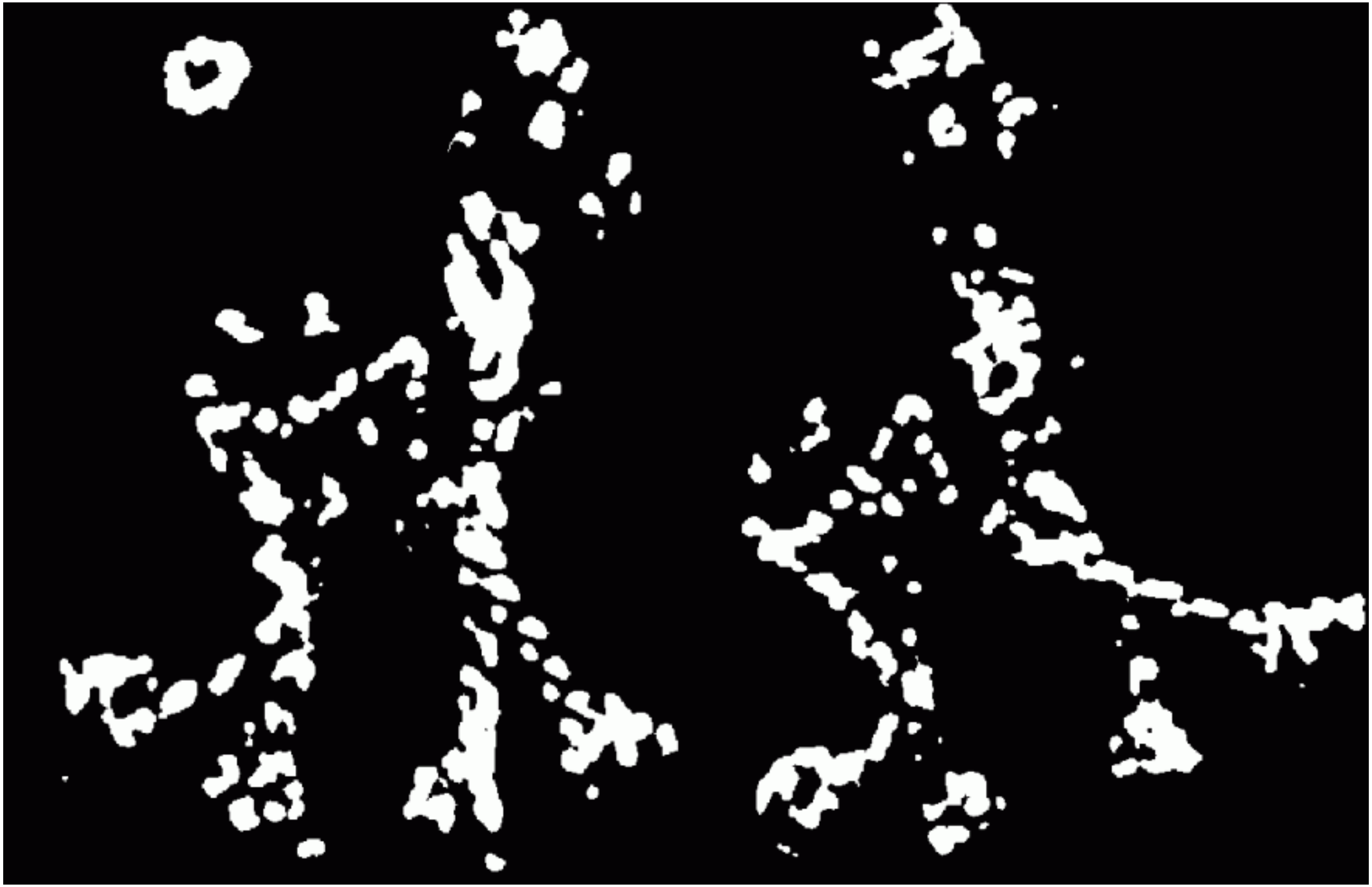
Compute corner response R



Slide credit: S. Lazebnik

Harris Detector: Steps

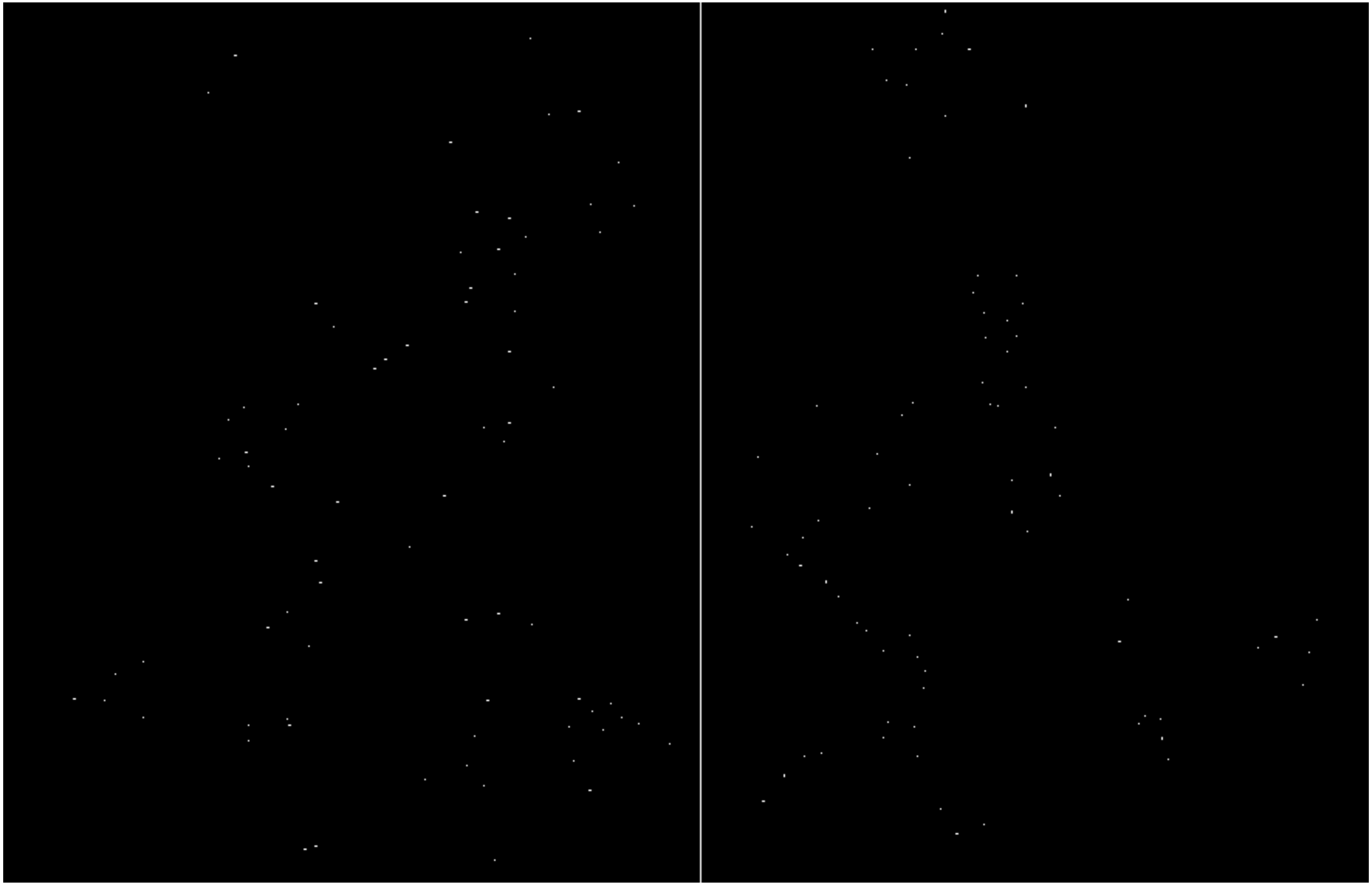
Find points with large corner response: R>threshold



Slide credit: S. Lazebnik

Harris Detector: Steps

Take only the points of local maxima of *R*



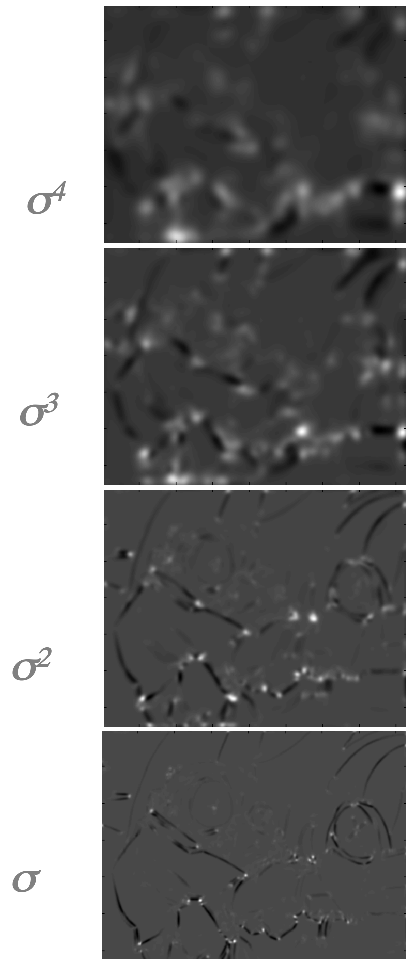
Slide credit: S. Lazebnik

Harris Detector: Steps

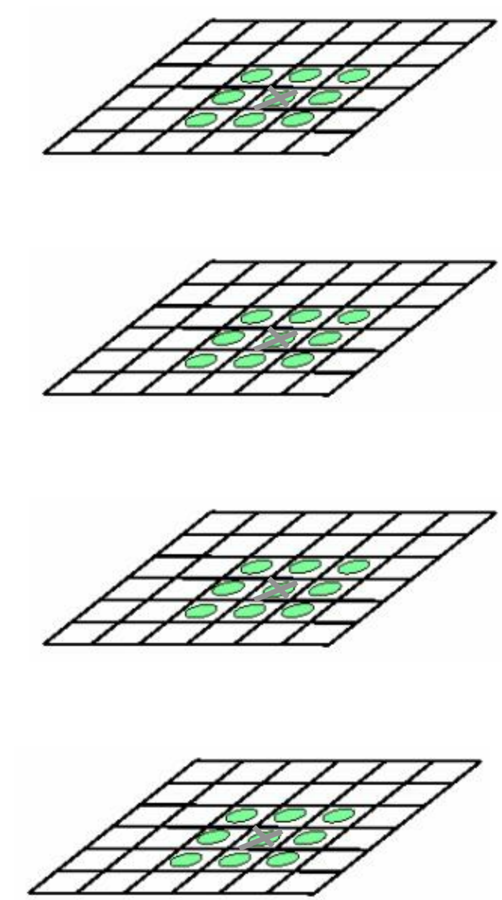
Slide credit: S. Lazebnik

Harris-Laplace [Mikolajczyk ‘01]

1. Initialization: Multiscale Harris corner detection



Computing Harris function



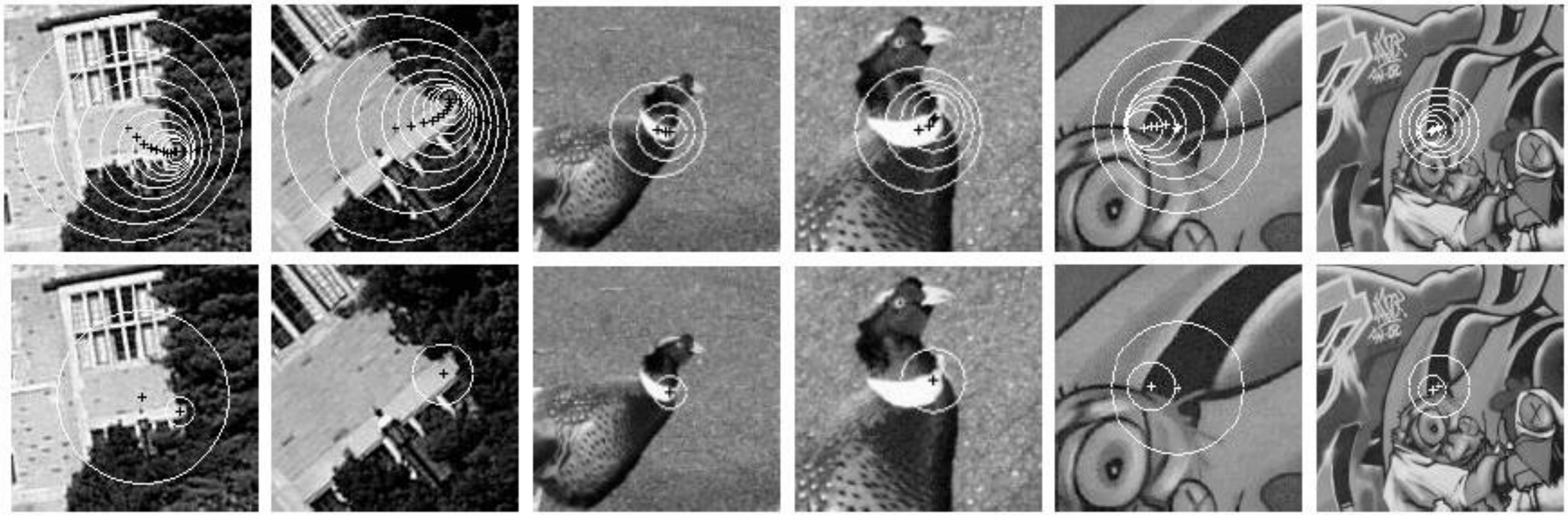
Detecting local maxima

Slide adapted from Krystian Mikolajczyk

Harris-Laplace [Mikolajczyk ‘01]

1. Initialization: Multiscale Harris corner detection
2. Scale selection based on Laplacian
   (same procedure with Hessian $\Rightarrow$ Hessian-Laplace)

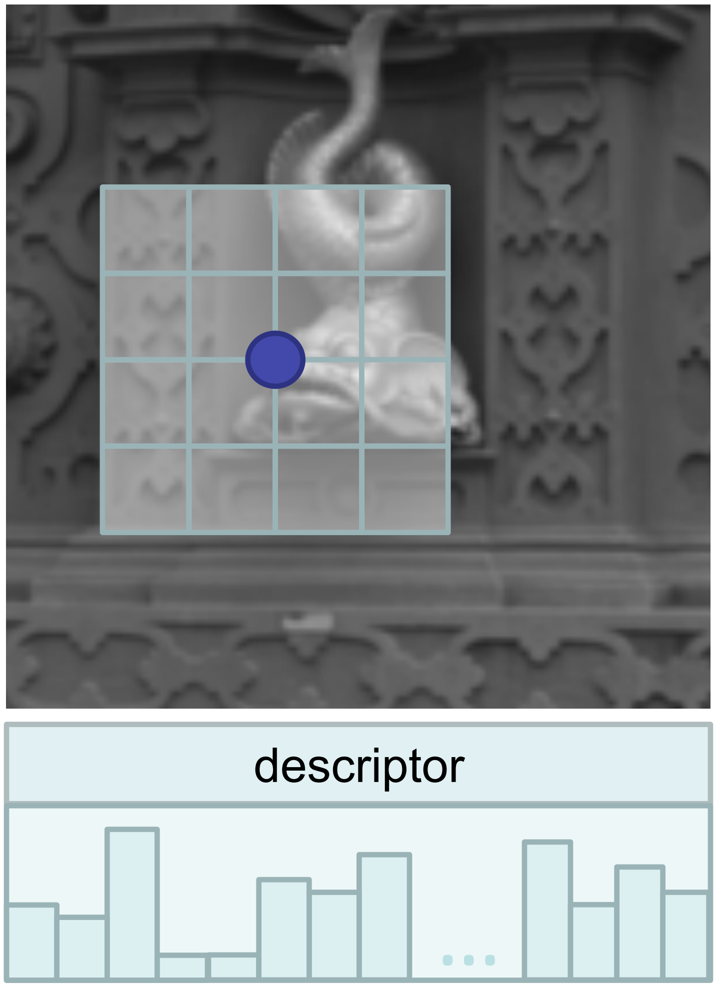
Harris points

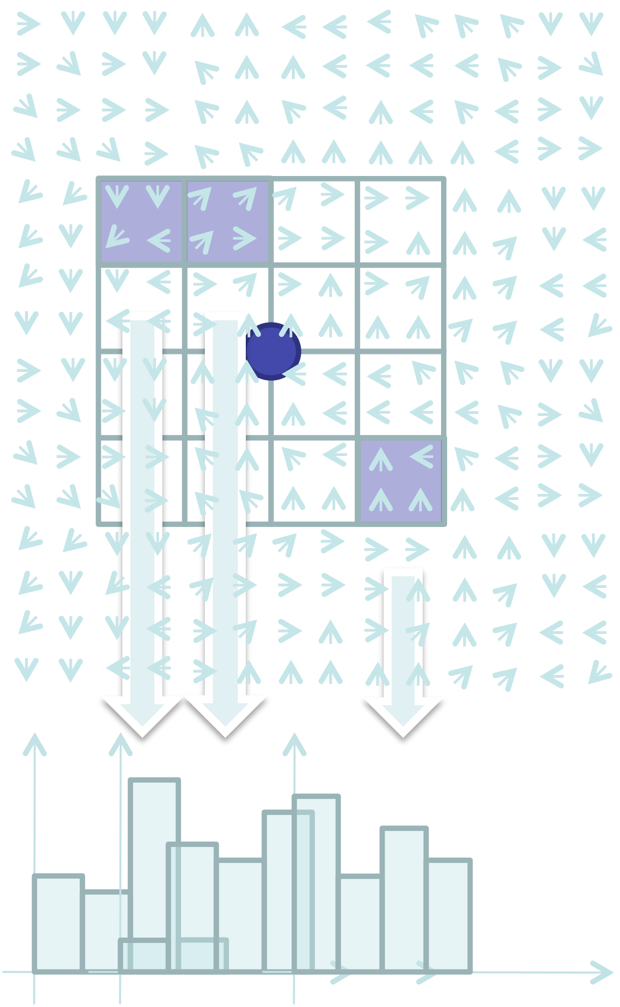


Harris-Laplace points

Slide adapted from Krystian Mikolajczyk

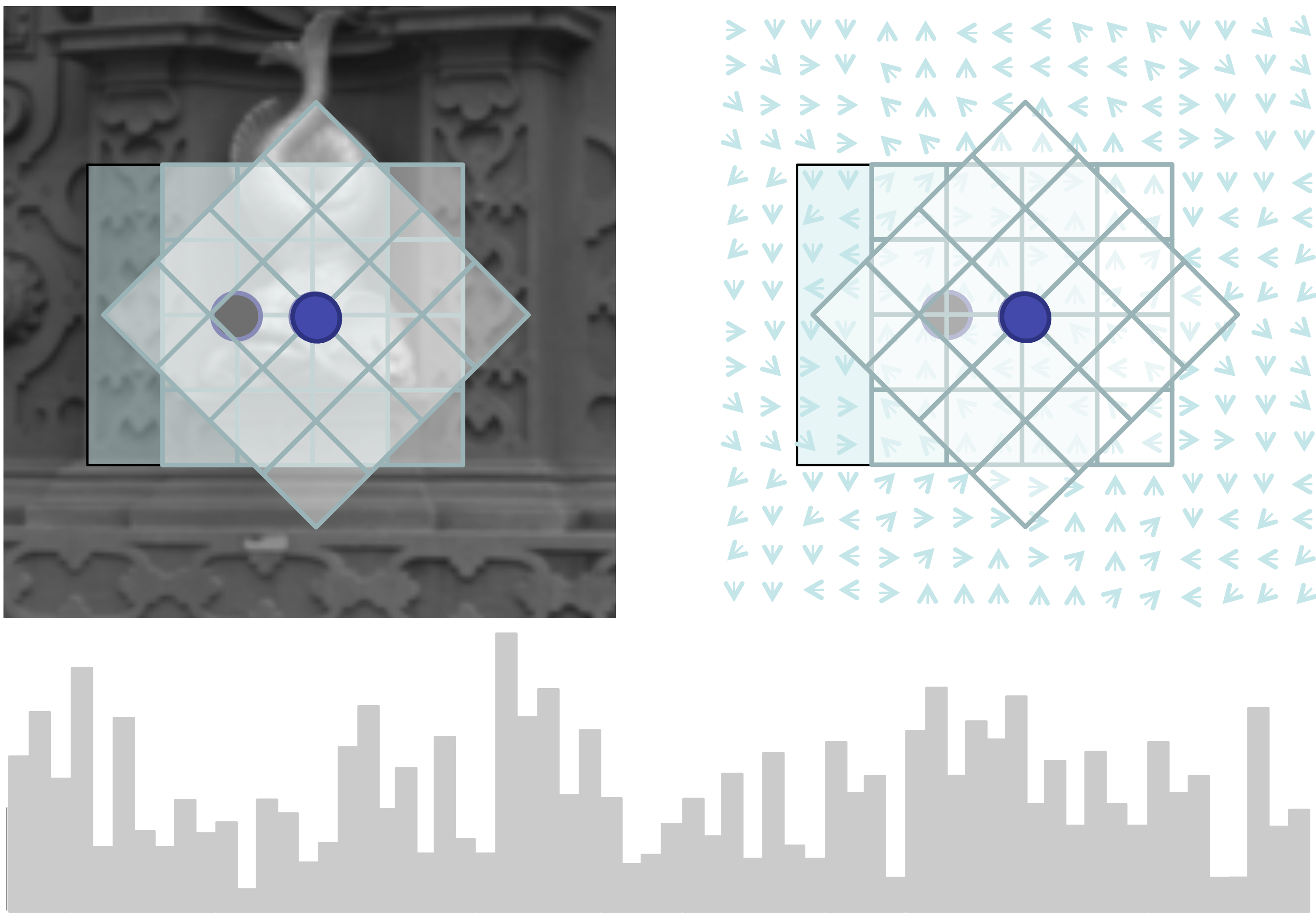
SIFT computation





Slide credit: E. Tola

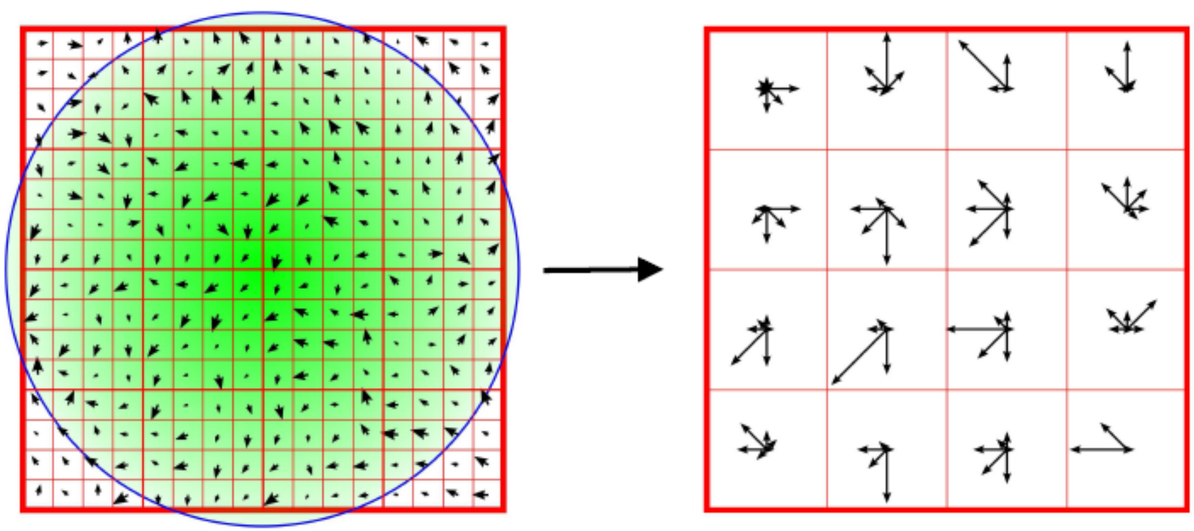
SIFT computation



Slide credit: E. Tola

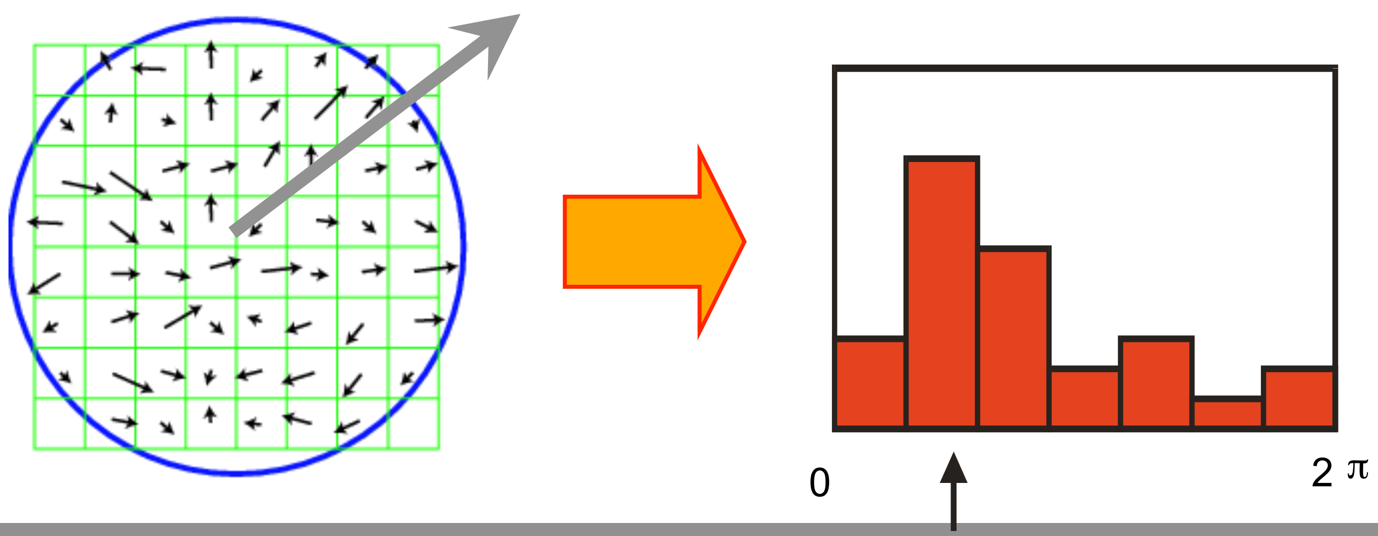
Orientation Histogram

- 4x4 spatial bins (16 bins total)
- 8-bin orientation histogram per bin
- 8 x 16 = 128 dimensions total
- Normalized to unit norm



SIFT descriptor

- Image patch descriptor
- Location and characteristic scale: Blob/Corner detector
- Find orientation from orientation histogram



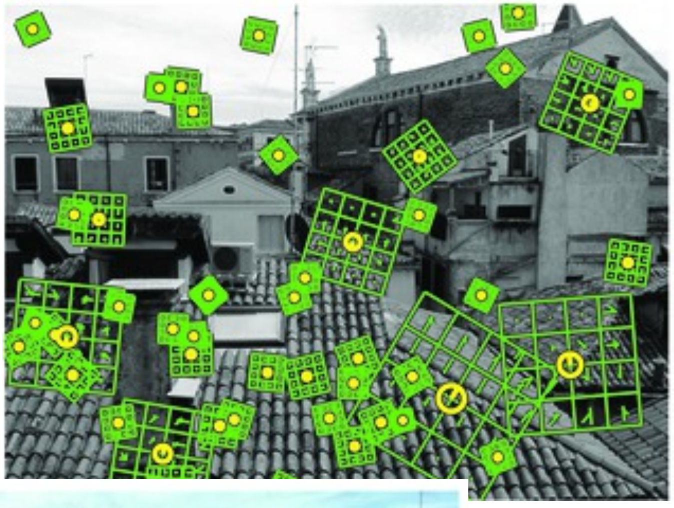
SIFT invariances

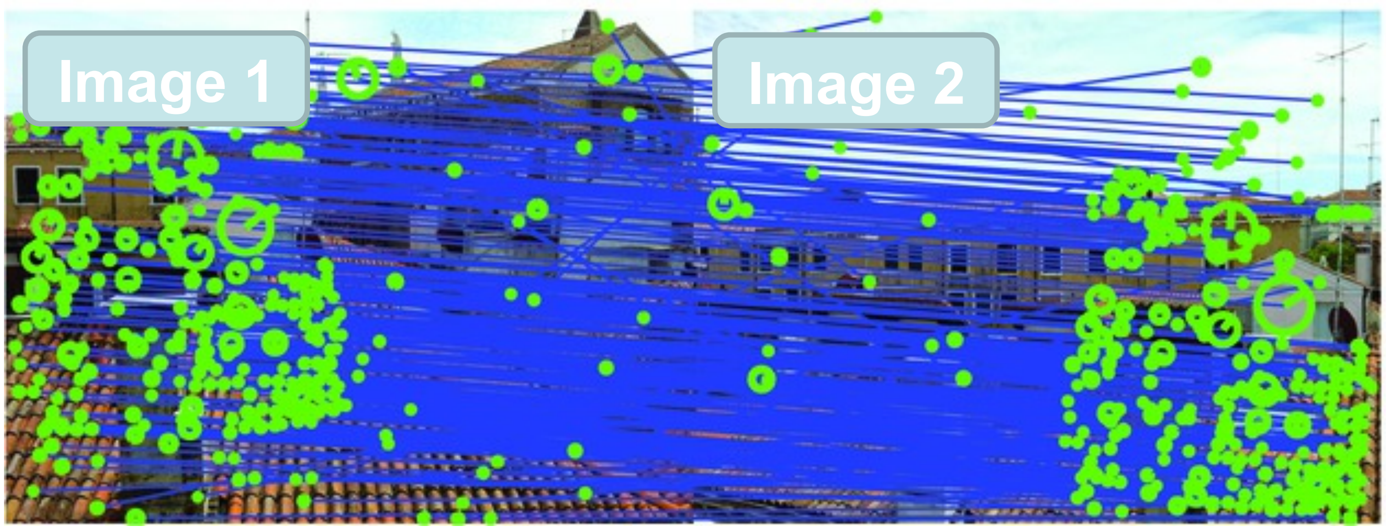
- Spatial binning: tolerance to small shifts in location
- Orientation normalization
- Photometric normalization by making all vectors unit norm
- Orientation histogram: robustness to small local deformations

Application: Image Matching

Assumption: images undergo global deformations with a few degrees-of-freedom (e.g. scaling, rotation)

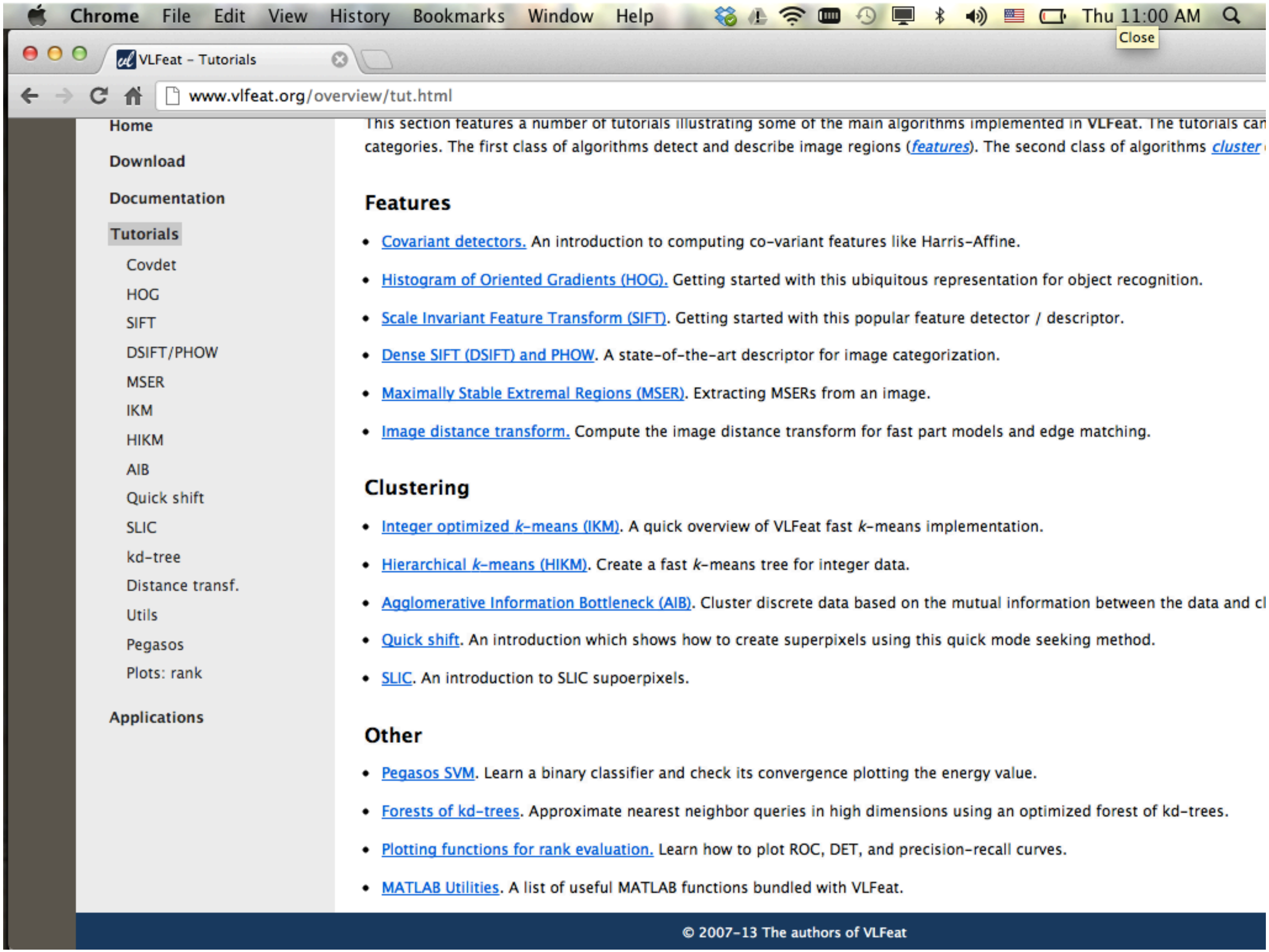
Correspondences of a few points suffice (found e.g. with SIFT)





Richard Szeliski's talk, tomorrow

Open source implementation: www.vlfeat.org



Further reading (literature ‘seeds’)

- **Compact Codes & Large-scale Retrieval**
  - J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. ICCV, 2003.
  - Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree. CVPR. (2006)
  - M. Perdoch, O. Chum, and J. Matas. Efficient representation of local geometry for large scale object retrieval. In Proc. CVPR, 2009
  - H. Jegou, M. Douze, C. Schmid, and P. Perez. Aggregating local descriptors into a compact image representation. CVPR, 10
  - A. Babenko and V. Lempitsky, The Inverted Multi-Index, CVPR 12
  - R. Arandjelović, A. Zisserman, All about VLAD, CVPR 2013
- **Fast/Compact Descriptors**
  - M. Calonder, V. Lepetit, C. Strecha, and P. Fua, BRIEF: Binary Robust Independent Elementary Features, (ECCV), 2010.
  - T. Trzcinski, M. Christoudias, P. Fua, and V. Lepetit, Boosting Binary Keypoint Descriptors. (CVPR), 2013.
  - SURF, FAST, ORB, FREAK,…

Further reading (literature 'seeds')

- **Feature encoding**
  - Improving the fisher kernel for large-scale image classification, F. Perronnin, J. Sánchez, and T. Mensink. In Proc. ECCV, 2010.
  - The devil is in the details: an evaluation of recent feature encoding methods, K. Chatfield, V. Lempitsky, A. Vedaldi, and A. Zisserman, BMVC, 2011
  - Sparse Kernel Approximations for Efficient Classification and Detection, A. Vedaldi and A. Zisserman, in Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2012
- **Descriptor Learning**
  - Simon A. J. Winder, Matthew Brown: Learning Local Image Descriptors. CVPR 2007
  - S. Winder, G. Hua, and M. Brown. Picking the best daisy. In Proc. CVPR, 2009.
  - Descriptor Learning for Efficient Retrieval, J. Philbin, M. Isard, J. Sivic, A. Zisserman,ECCV 10
  - K. Simonyan, A. Vedaldi, and A. Zisserman. Descriptor learning using convex optimisation. In Proc. ECCV, 2012

Dense descriptors

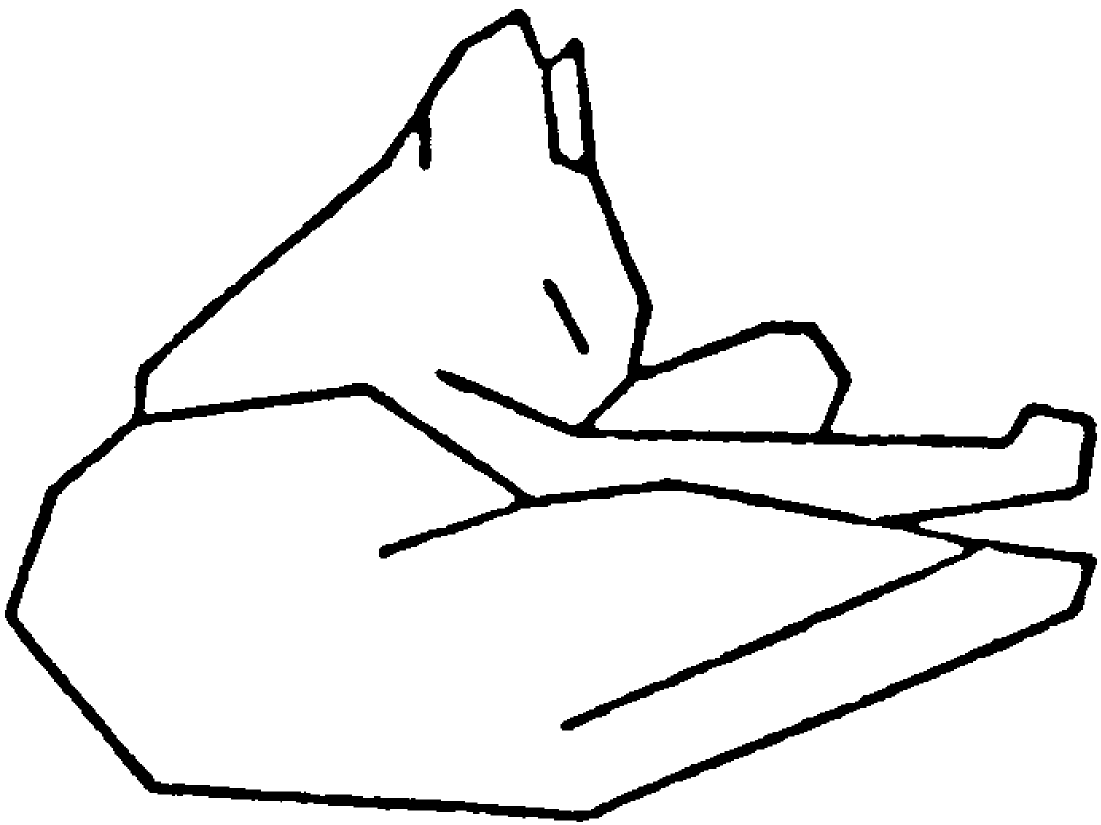
- Interest point detection revisited

D. Lowe, Perceptual Organization and Visual Recognition, 1985

F. Jurie, B. Triggs, Sampling strategies for bag-of-work classification, 2005

Dense descriptors

- Interest point detection revisited: there is nothing special about corners

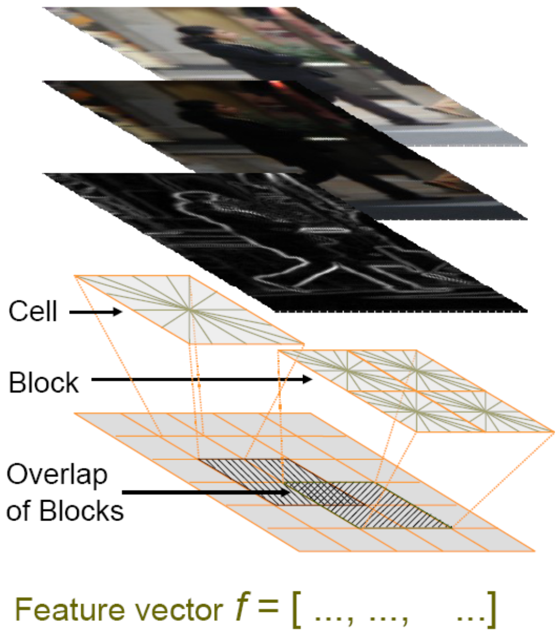


D. Lowe, Perceptual Organization and Visual Recognition, 1985

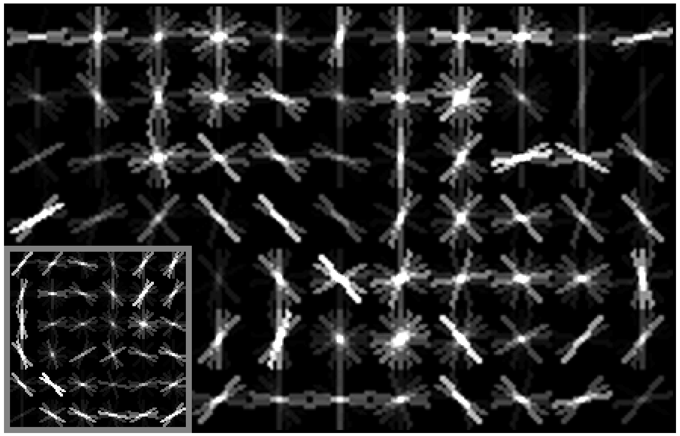
F. Jurie, B. Triggs, Sampling strategies for bag-of-work classification, 2005

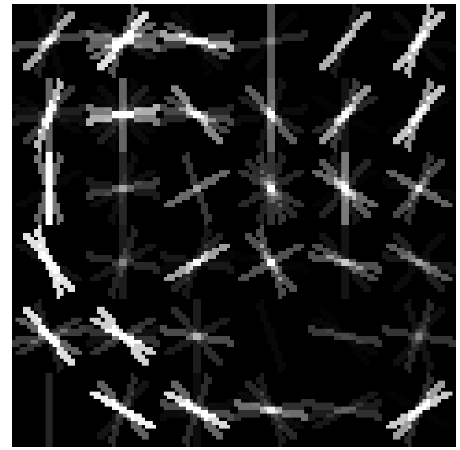
Histogram of Orientated Gradients (HOG) descriptor

- Dalal and Triggs, ICCV 2005
  - Like SIFT descriptor, but for arbitrary box aspect ratio, and computed over all image locations and scales
  - Highly accurate detection using linear classifier

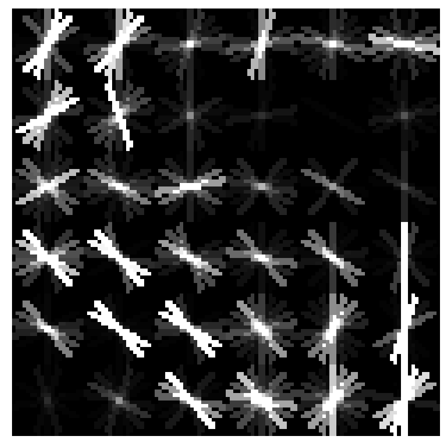


Part score computation





$$\mathbf{w}[y]$$

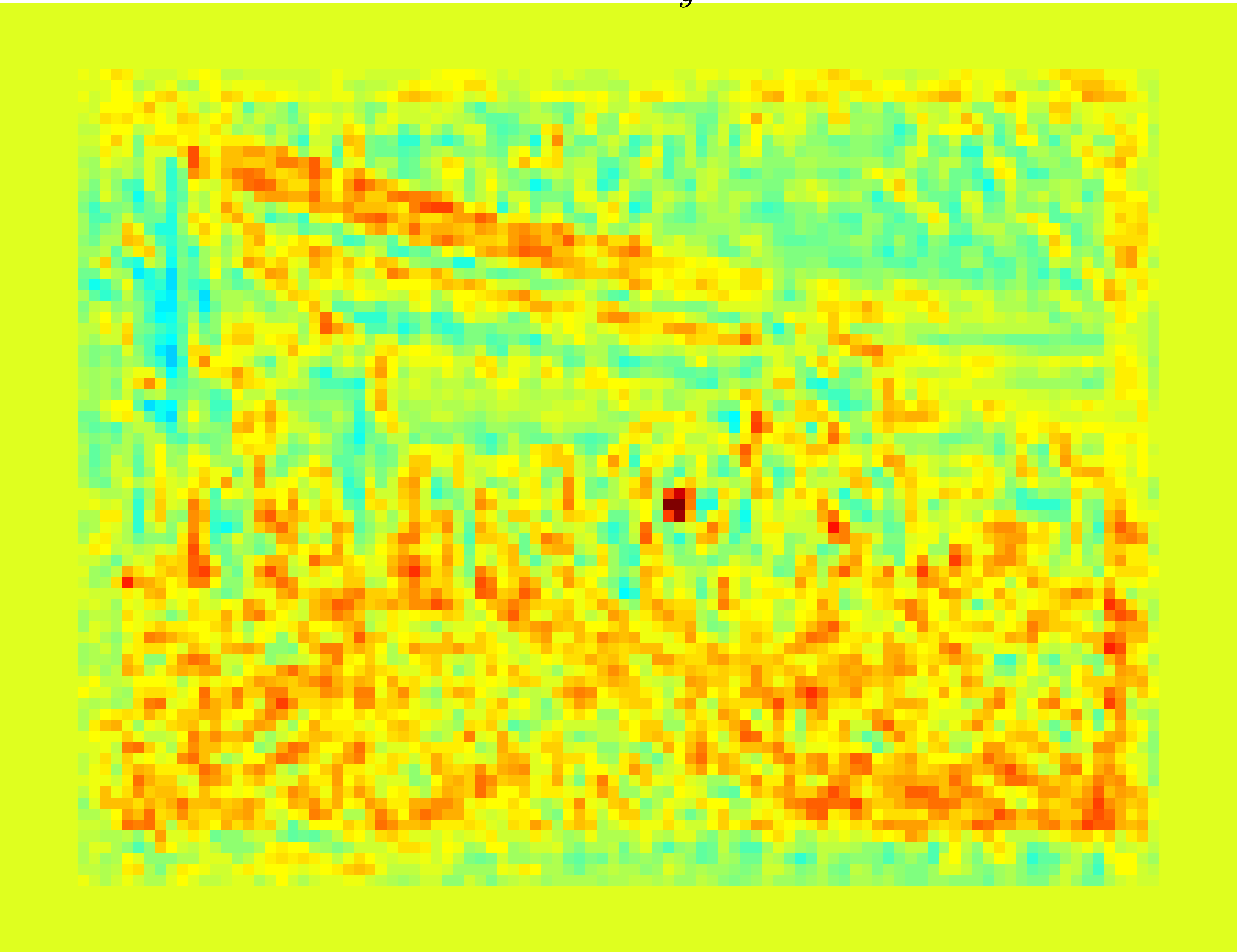


$$\mathbf{h}[x+y]$$

$$s[x] = \sum_{y} \langle \mathbf{h}[x+y], \mathbf{w}[y] \rangle$$

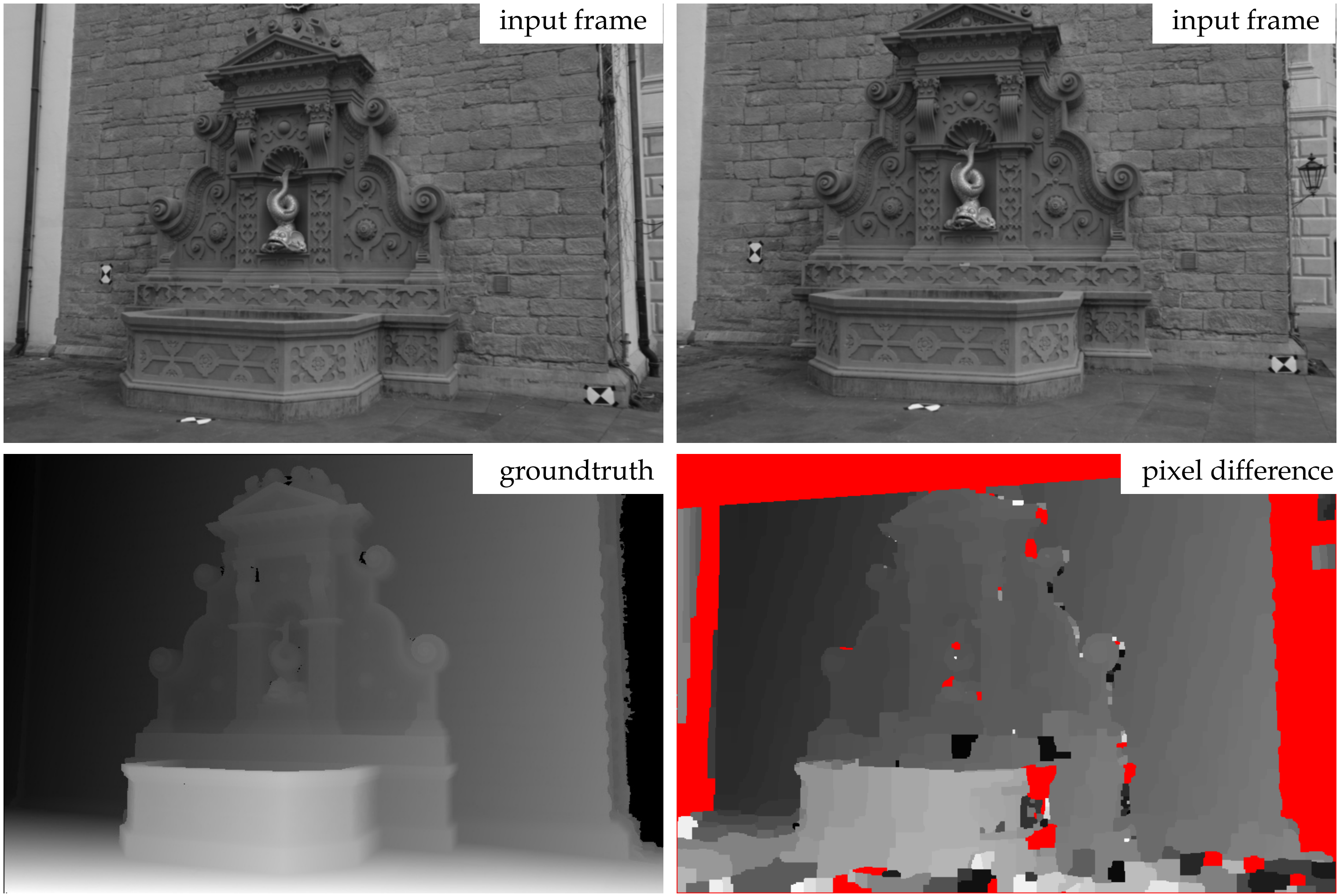
Part score

$$\mathbf{h}[x] \qquad s[x] = \sum_y \langle \mathbf{h}[x+y], \mathbf{w}[y] \rangle$$



Dense descriptors: motivation

Narrow baseline : Pixel Difference + Graph Cuts

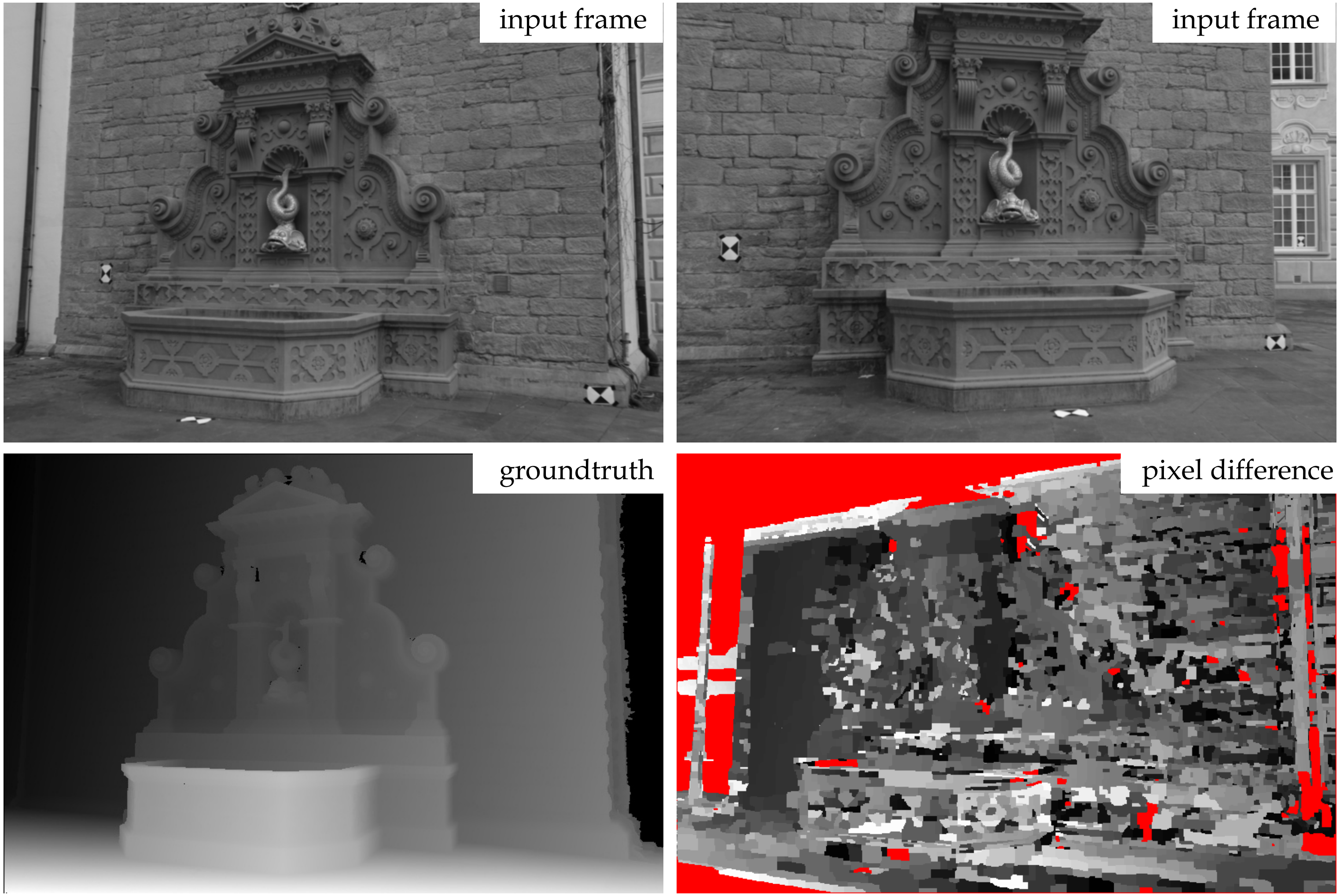


DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

Dense descriptors: motivat

USE A DESCRIPTOR

Wide baseline : Pixel Difference + Graph Cuts

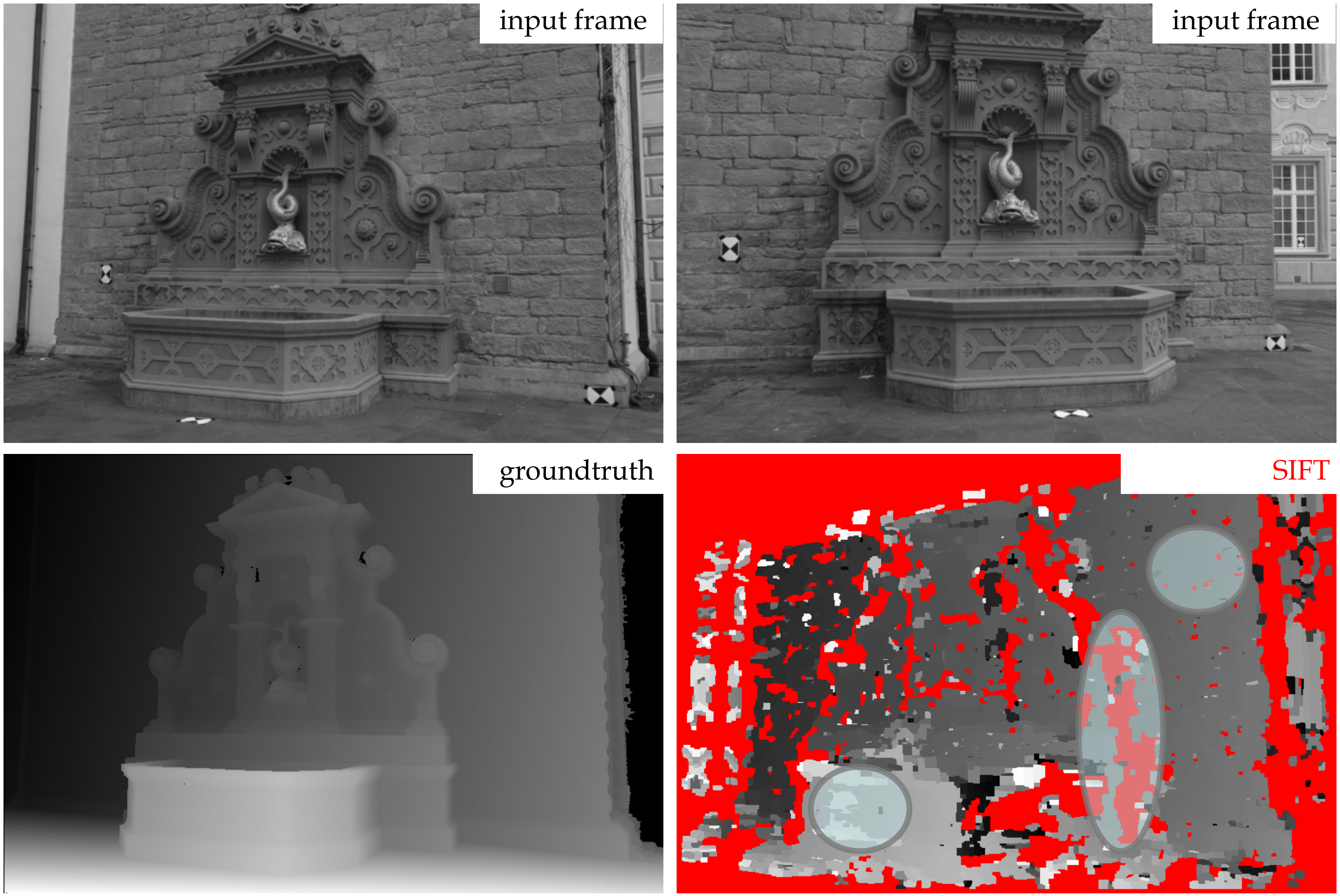


DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

Dense descriptors

250 Seconds

Wide baseline : SIFT Descriptor*+ Graph Cuts

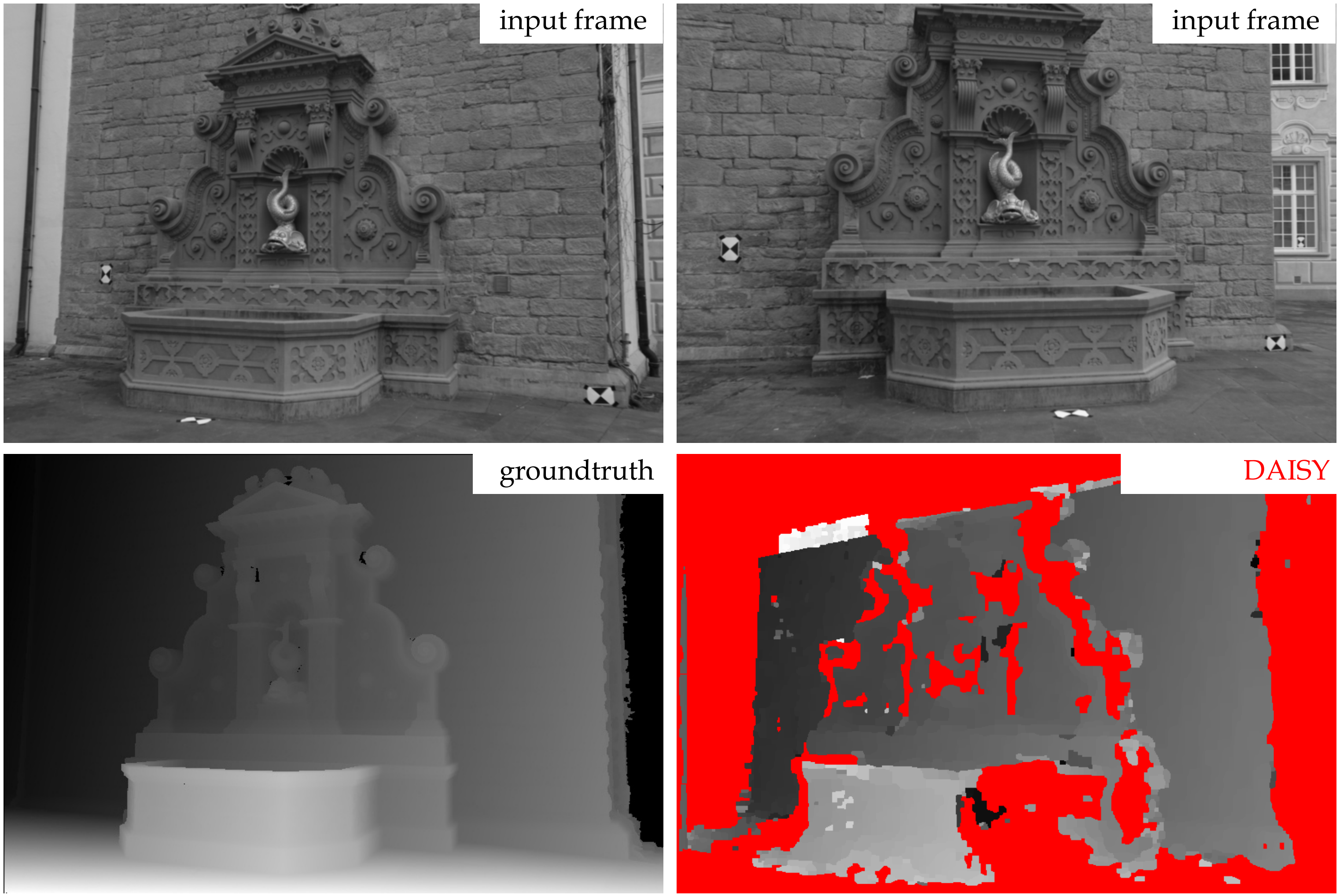


DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

Fast dense descriptors

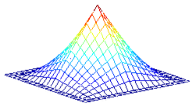
5 Seconds

Wide baseline : DAISY Descriptor+ Graph Cuts

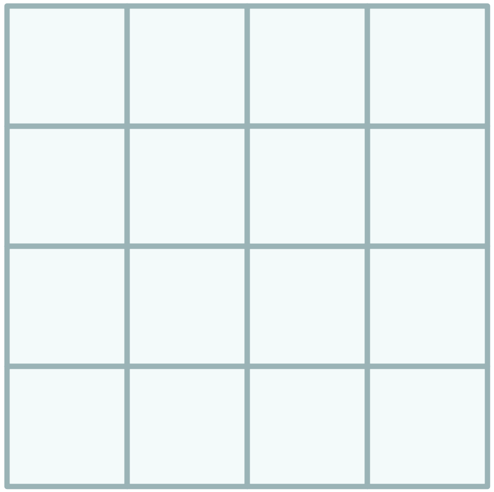


DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

SIFT-> DAISY



SIFT

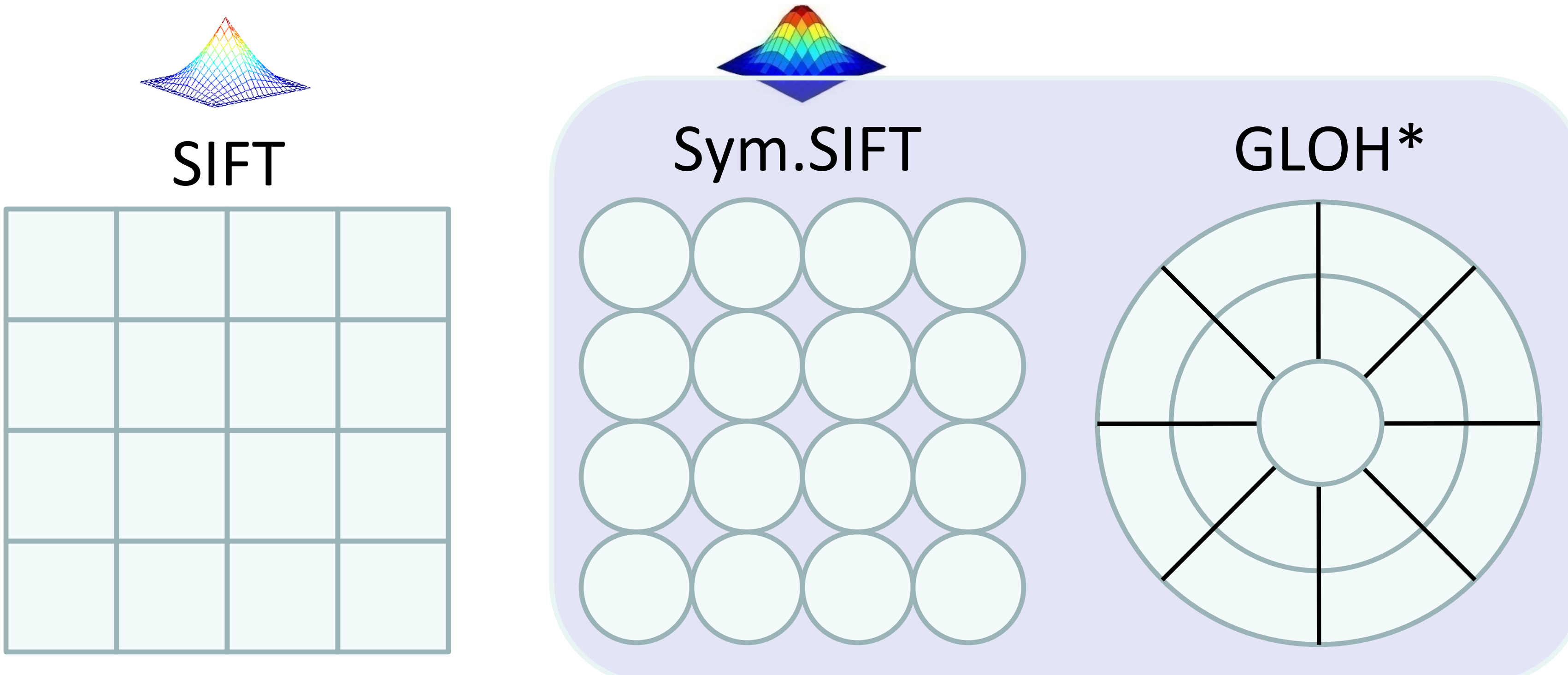


+ Good Performance

- Not suitable for dense computation

DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

SIFT-> DAISY



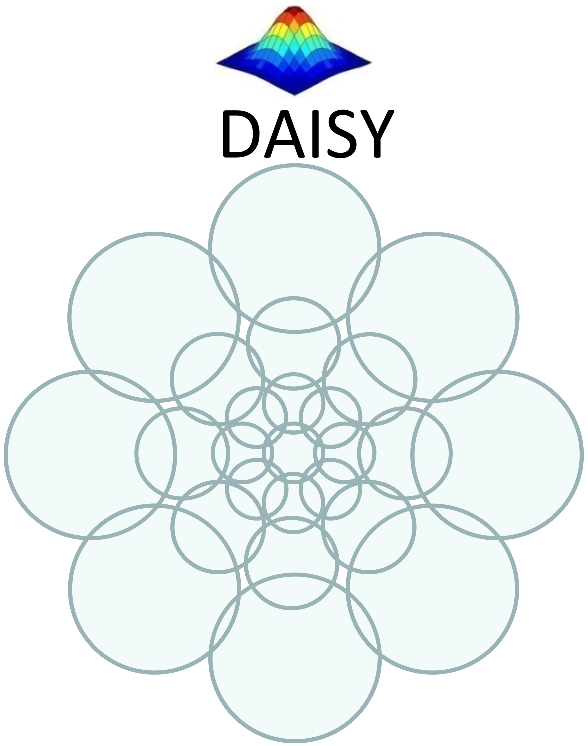
+ Good Performance
- Not suitable for dense computation

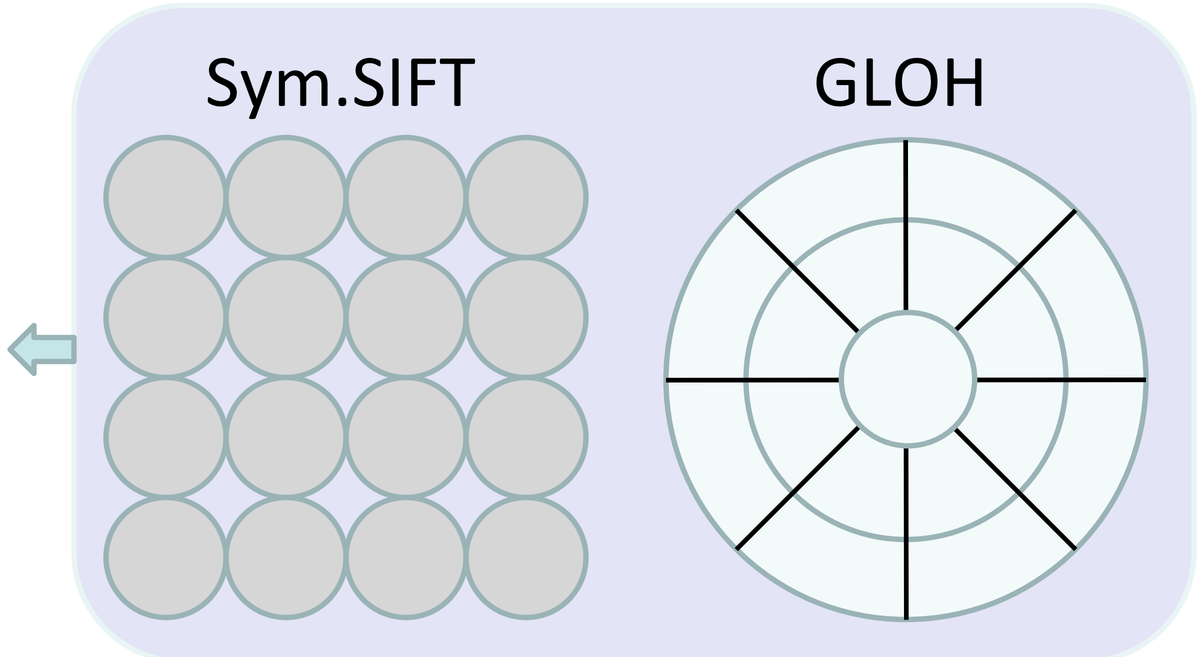
+ Gaussian Kernels : Suitable for Dense Computation

+ Good Performance
+ Better Localization
- Not suitable for dense computation

* K. Mikolajczyk and C. Schmid. A Performance Evaluation of Local Descriptors. PAMI'04.

SIFT-> DAISY





Suitable for dense computation
+ Improved performance:*
+ Precise localization
+ Rotational Robustness

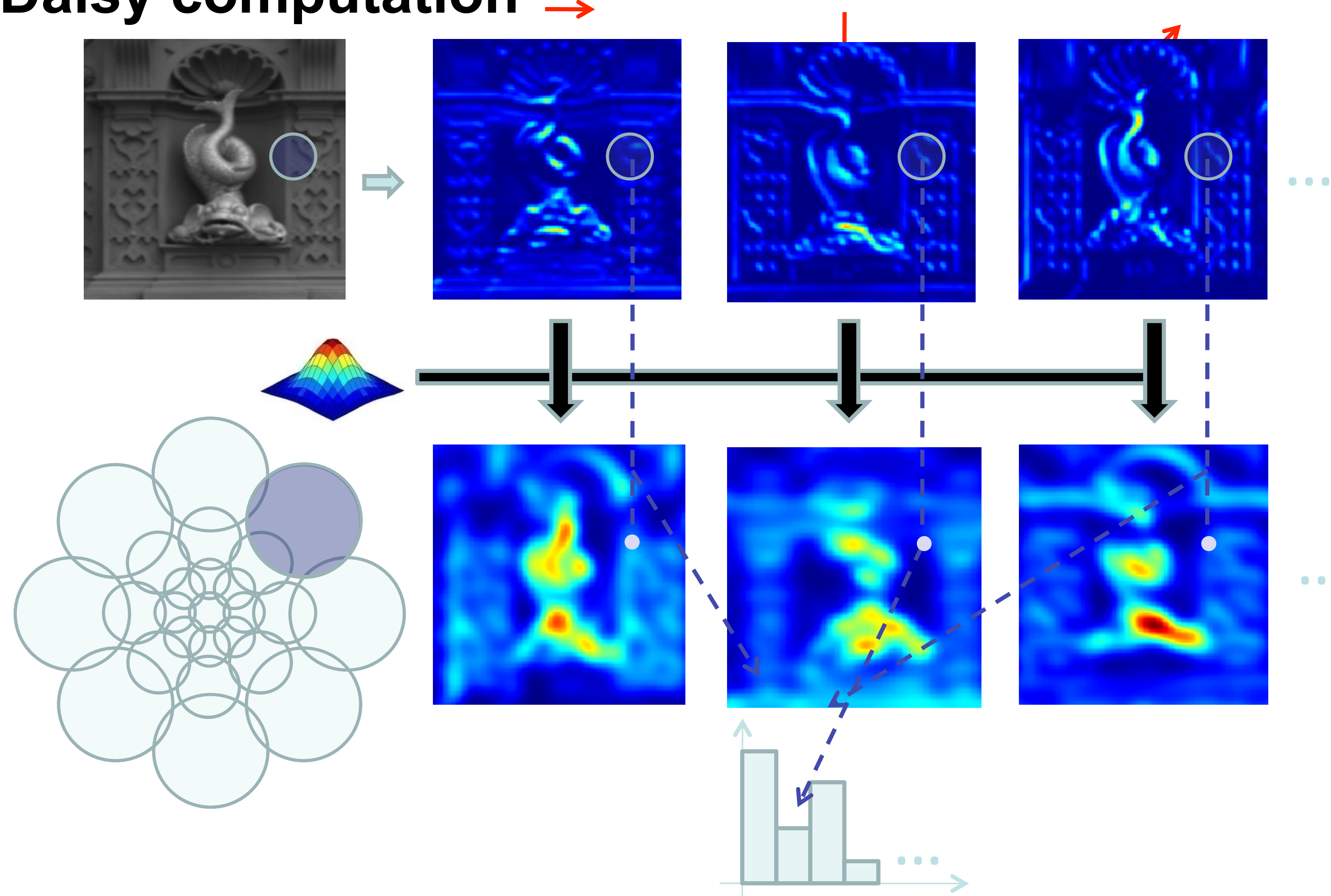
+ Suitable for Dense Computation

+ Good Performance
+ Better Localization

- Not suitable for dense computation

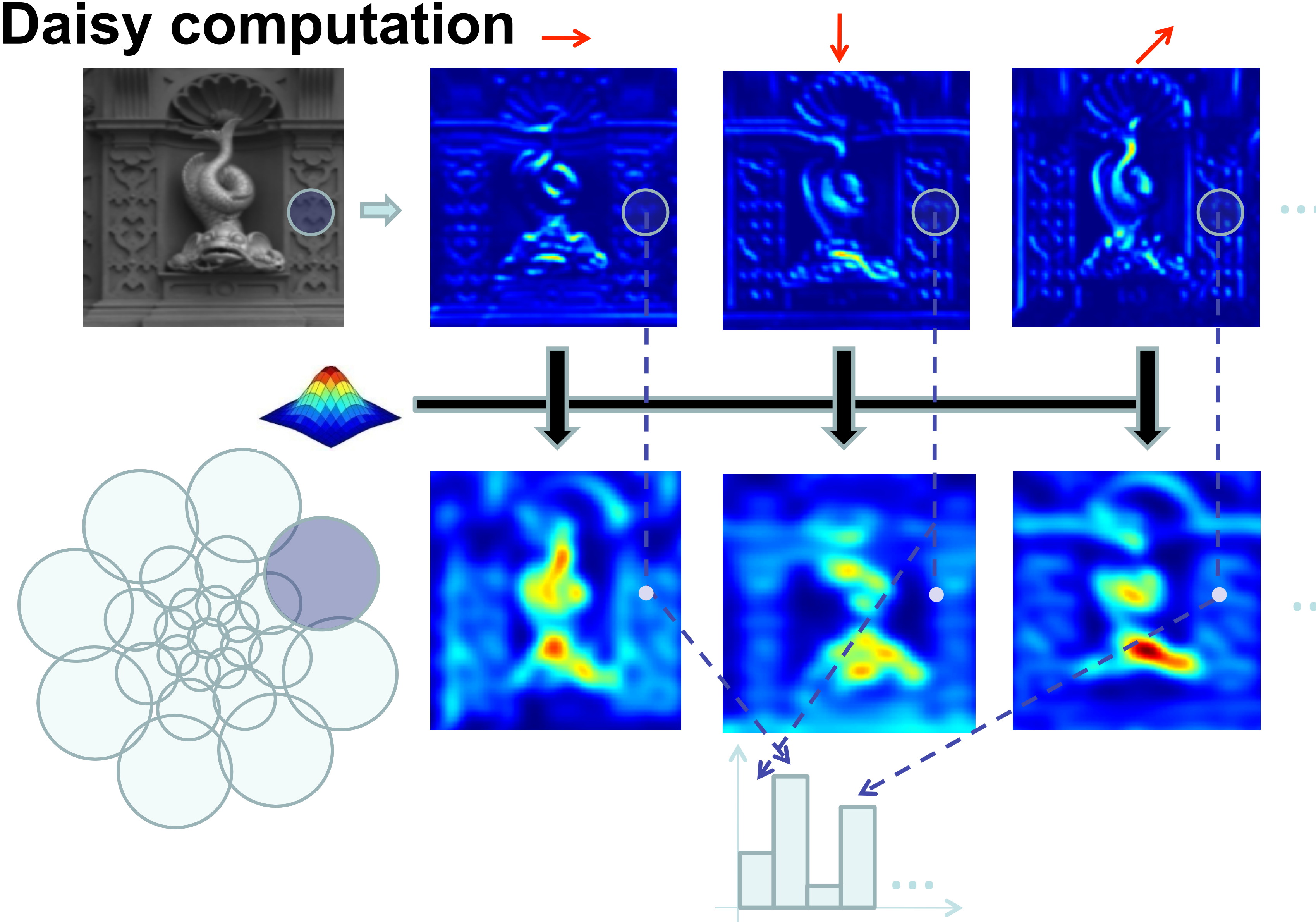
* S. Winder and M. Brown. *Learning Local Image Descriptors* in CVPR'07

Daisy computation



DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

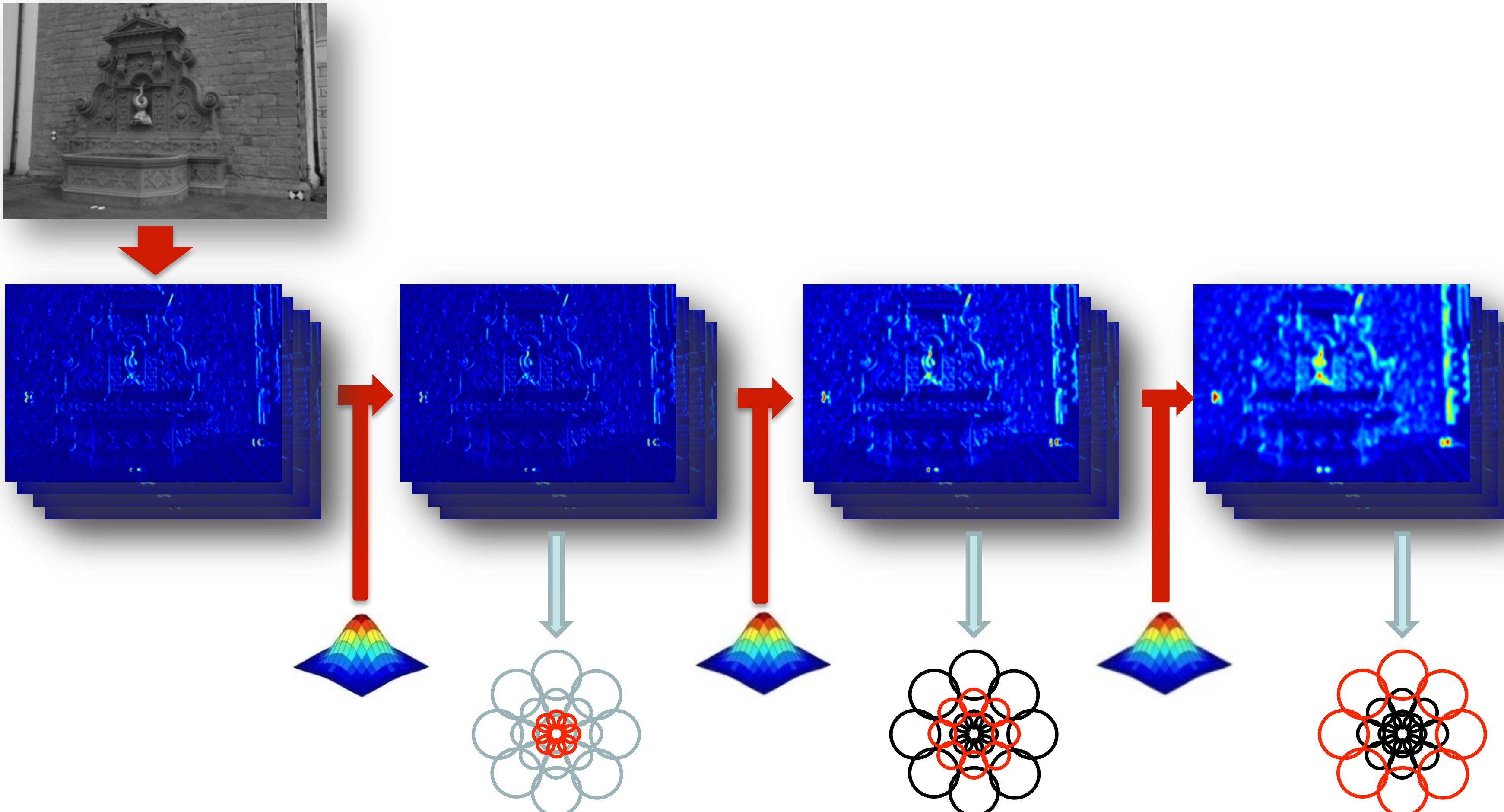
Daisy computation



DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

Daisy computation

DAISY :	5s
SIFT :	250s



- Rotating the descriptor only involves reordering the histograms.
- The computation mostly involves 1D convolutions, which is fast.

DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo, E. Tola, V. Lepetit, P. Fua, PAMI, 10

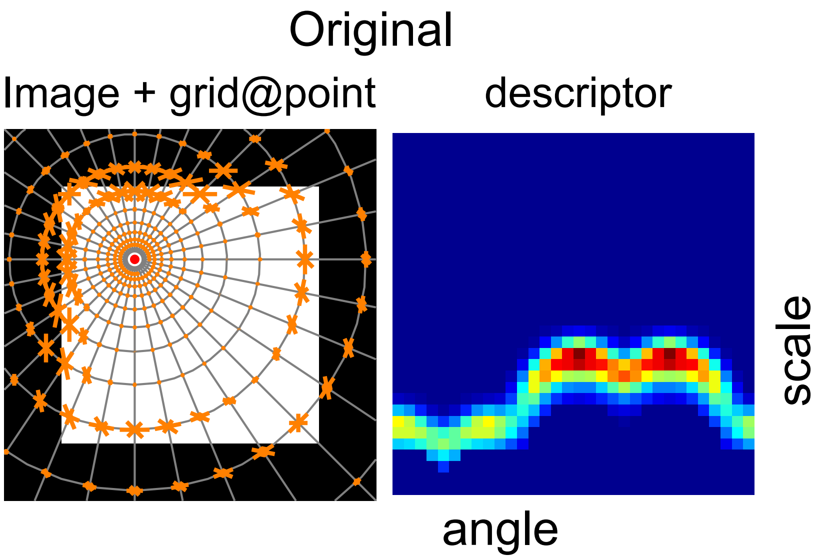
Scale- and rotation- invariance through Fourier Transforms

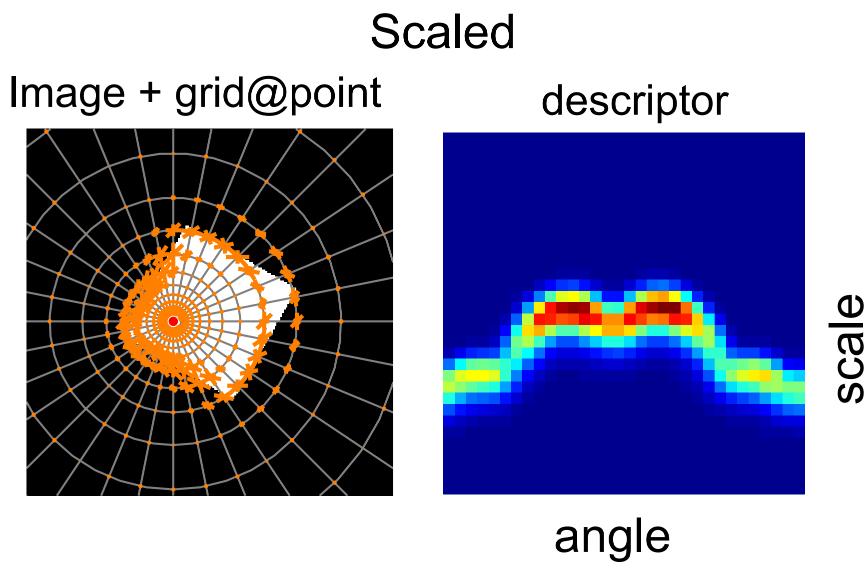
Lecture 1: Modulation property: $f(x) \leftrightarrow F(\omega) \quad f(x)e^{j\omega_c x} \leftrightarrow F(\omega - \omega_c)$

Fact 1: Fourier shifting property: $f[i - n_i, j - n_j] \xrightarrow{\mathcal{F}} F \exp\left(-j\left(n_i \frac{2\pi}{N} + n_j \frac{2\pi}{K}\right)\right) \xrightarrow{|\cdot|} |F|$

Signal translation does not affect the signal's Fourier Transform Magnitude

Fact 2: the log-polar sampling *turns image scaling and rotation to translation*:



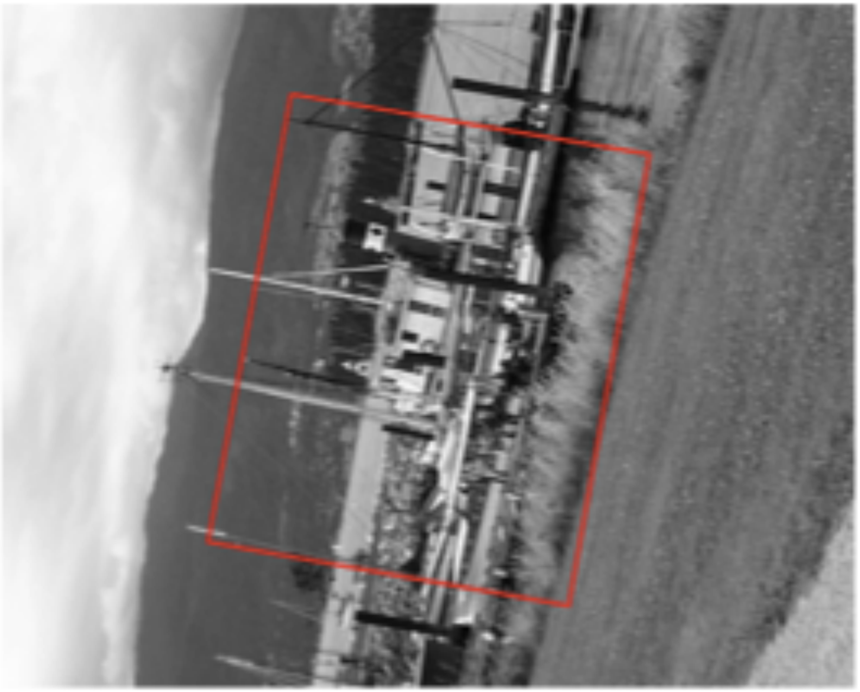


Fact 1+2: the *Fourier Transform Modulus of log-polar descriptors is invariant*

I. Kokkinos and A. Yuille, Scale Invariance without Scale Selection, CVPR, 2008.
D. Casasent and D. Psaltis, Rotation and scale-invariant optical correlation, Applied Optics, 1976

Dense Scale-Invariant Descriptors

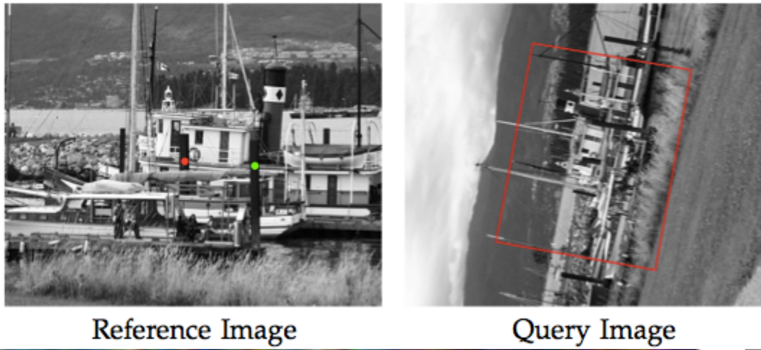
Reference Image

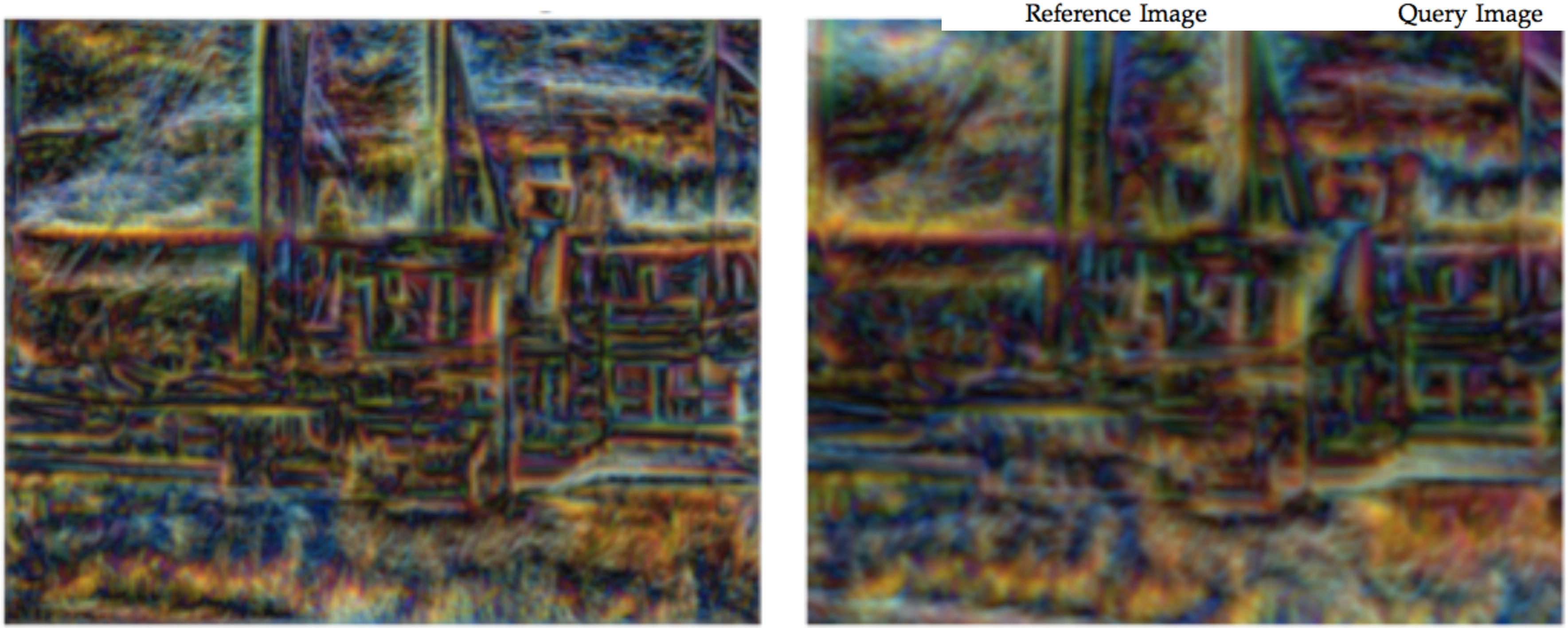


Query Image

I. Kokkinos and A. Yuille, Scale Invariance without Scale Selection, CVPR, 2008.
D. Casasent and D. Psaltis, Rotation and scale-invariant optical correlation, Applied Optics, 1976

Dense Scale-Invariant Descriptors

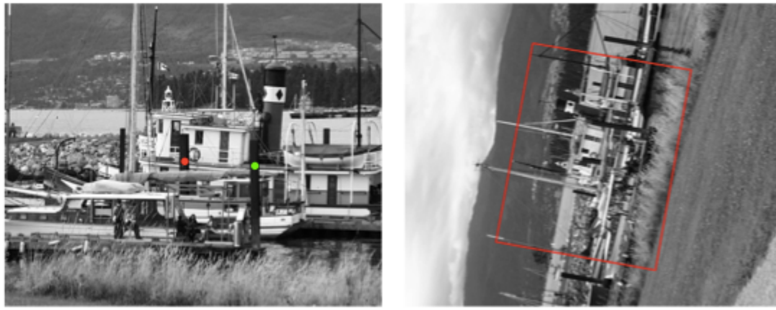




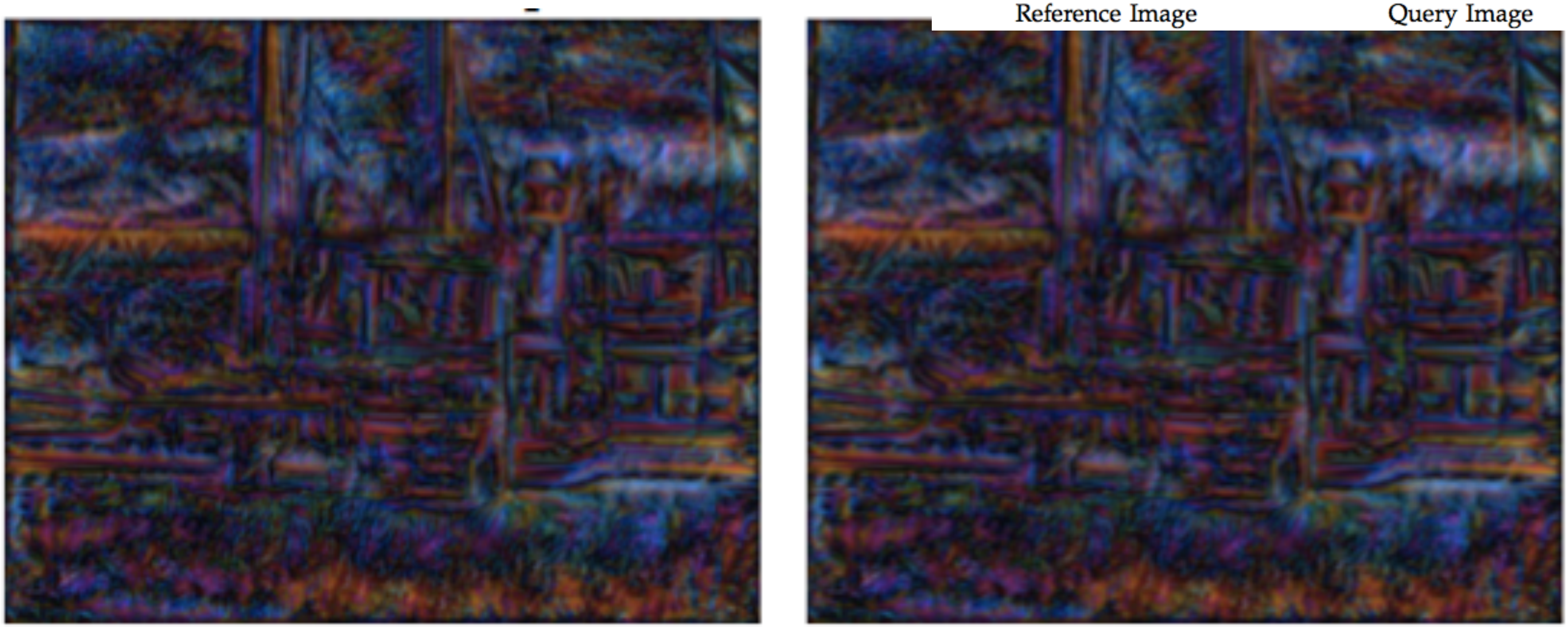
Dense Descriptors, $\omega_n = 1, \omega_k = 1, d \in \{1, 2, 3\}$

I. Kokkinos and A. Yuille, Scale Invariance without Scale Selection, CVPR, 2008.
D. Casasent and D. Psaltis, Rotation and scale-invariant optical correlation, Applied Optics, 1976

Dense Scale-Invariant Descriptors



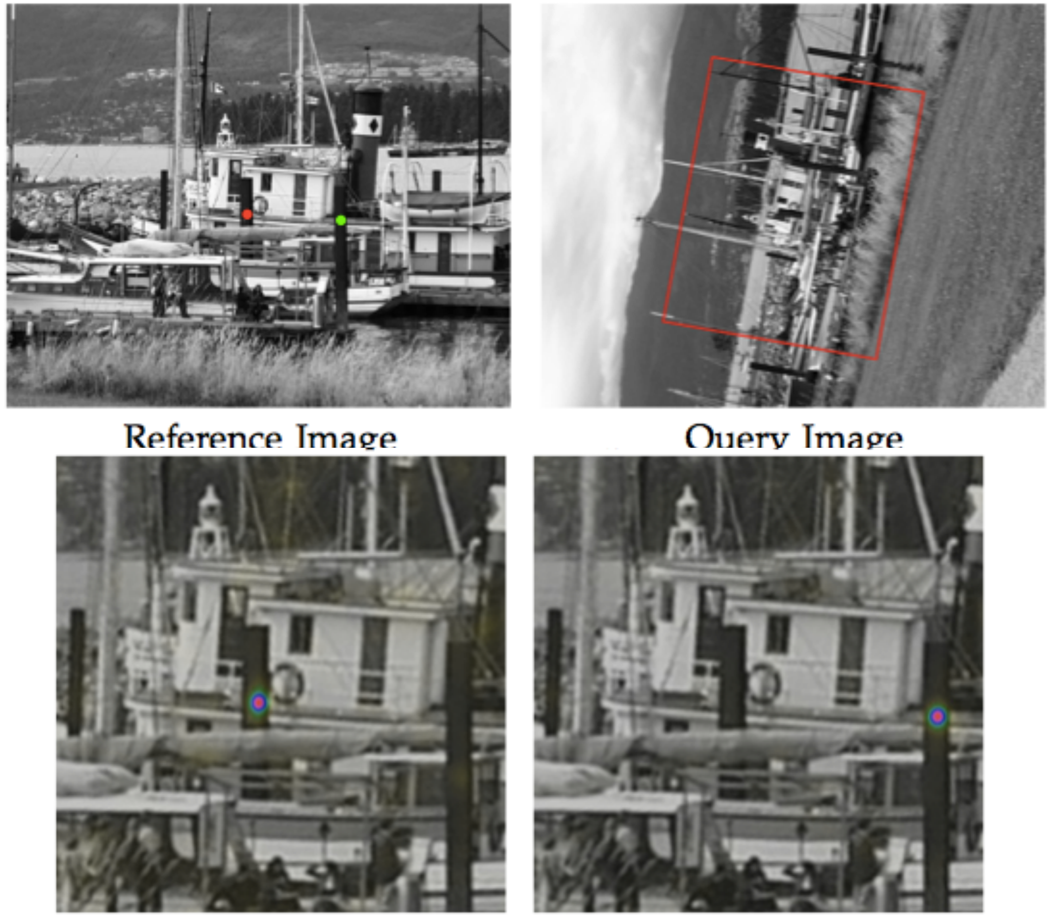
Reference Image Query Image



Dense Descriptors, $\omega_n = 1, \omega_k = 2, d \in \{1, 2, 3\}$

I. Kokkinos and A. Yuille, Scale Invariance without Scale Selection, CVPR, 2008.
D. Casasent and D. Psaltis, Rotation and scale-invariant optical correlation, Applied Optics, 1976

Dense Scale-Invariant Descriptors

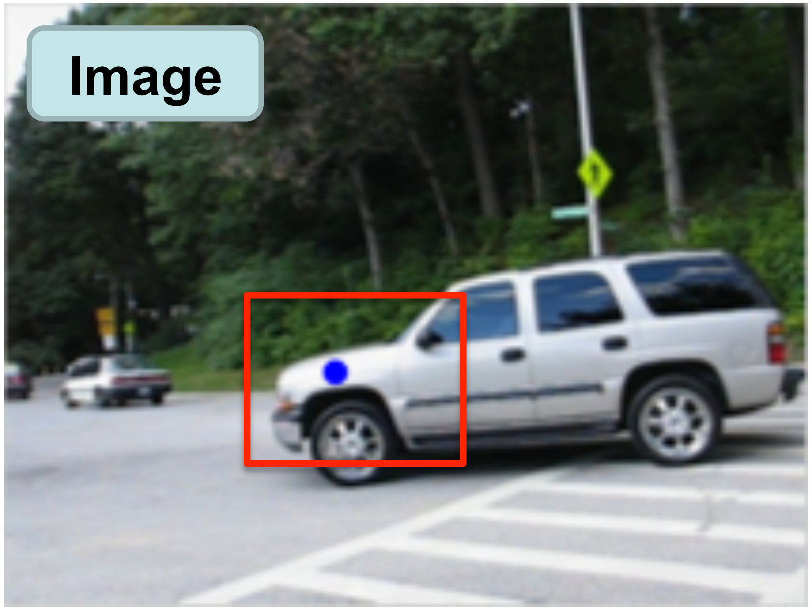


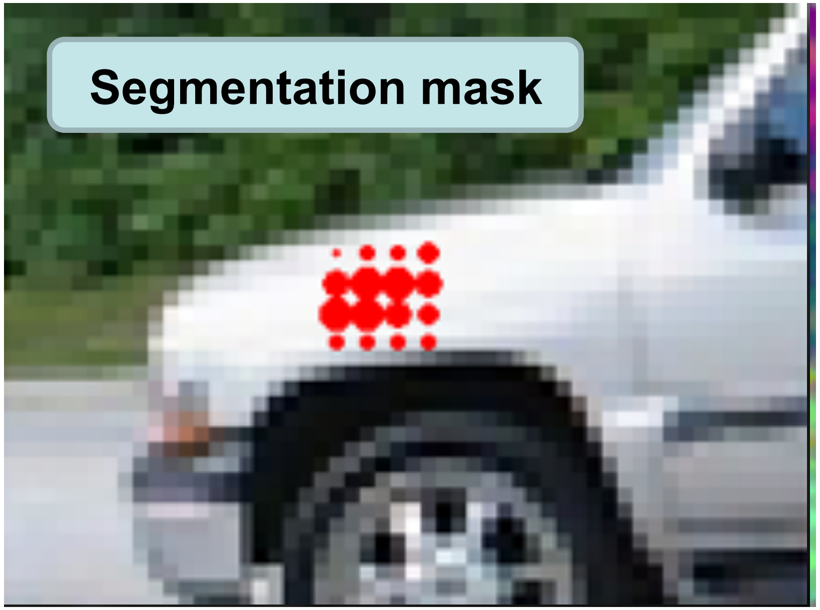
Query similarity to red (left) and green (right) reference points.

I. Kokkinos, M. Bronstein and A. Yuille, Dense Scale-Invariant Descriptors for Images and Surfaces, Technical report 2012
I. Kokkinos and A. Yuille, Scale Invariance without Scale Selection, CVPR, 2008.
D. Casasent and D. Psaltis, Rotation and scale-invariant optical correlation, Applied Optics, 1976

Segmentation-aware descriptors

- We improve **descriptors** by suppressing the **background** with **soft segmentation**





RGB representation (actually 8 layers)

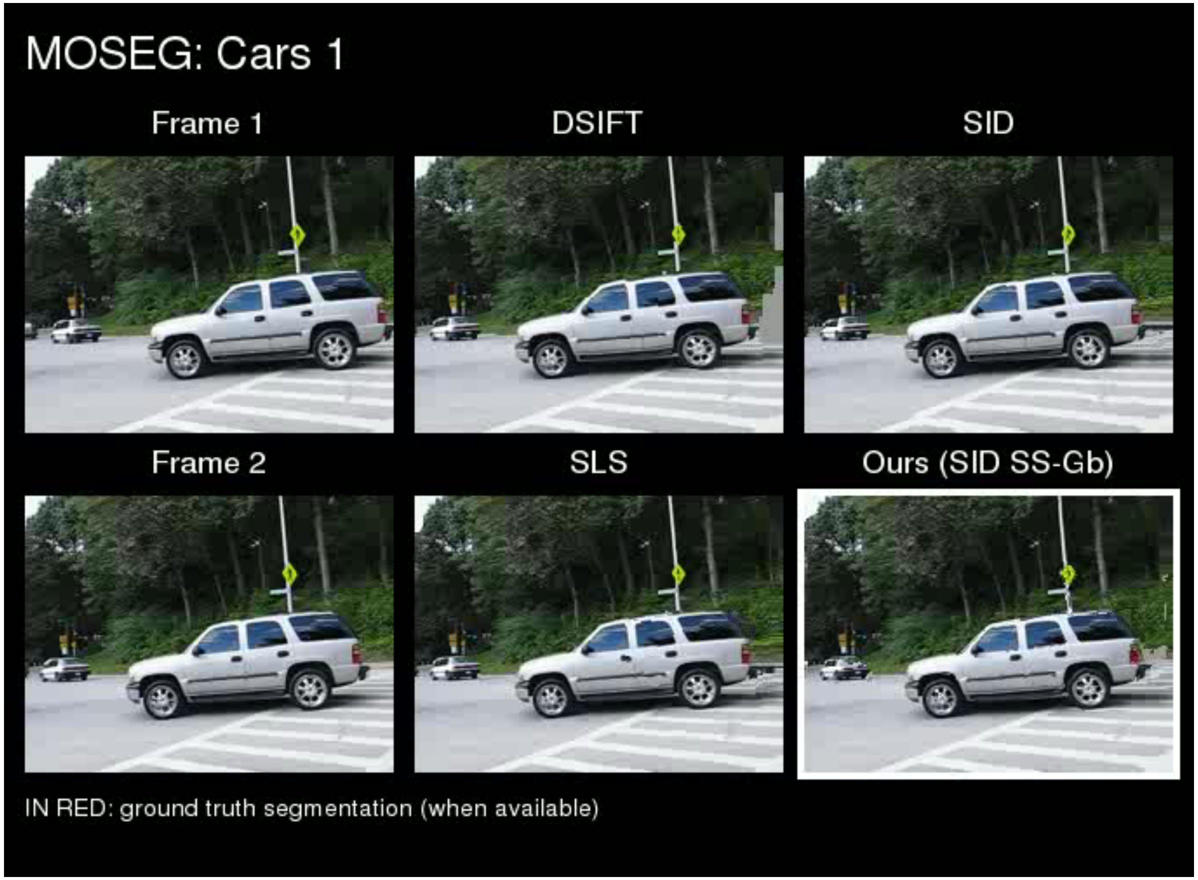
Segmentation masks: $\mathbf{w}^{[i]} = \exp\left(-\lambda \cdot d(\mathbf{x}, \mathbf{G}^{[i]}(\mathbf{x}))\right)$

- Simple, generic and fast.
- Scale- and rotation- invariant, if need be.

E. Trulls et al. Dense segmentation-aware descriptors, (CVPR), 2013.

Application-1: large displacement optical flow

SIFT-flow for optical flow estimation
JHU/MOSEG benchmark for evaluation

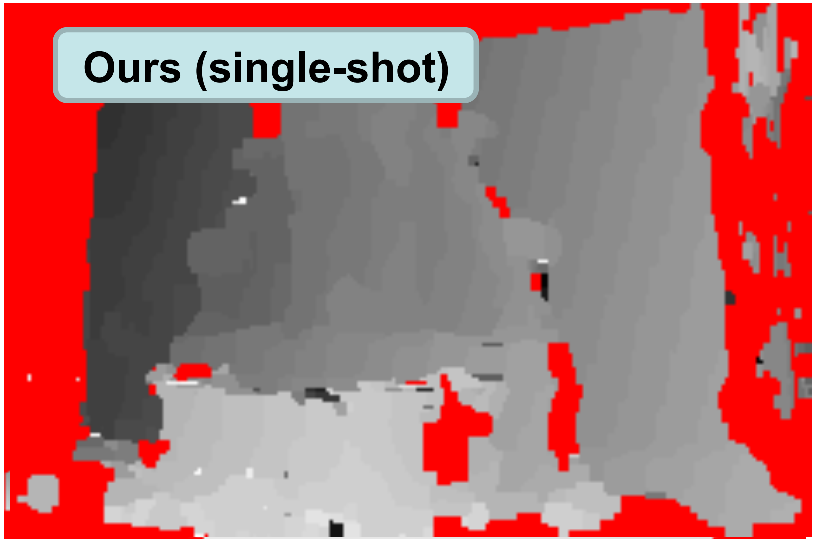


C. Liu, J. Yuen, and A. Torralba. Sift flow: dense correspon- dence across difference scenes. T.PAMI, 33(5), 2011.
T. Brox and J. Malik. Object segmentation by long term analysis of point trajectories. In ECCV, 2010

E. Trulls et al. Dense segmentation-aware descriptors, (CVPR), 2013.

Application 2: Wide-baseline stereo

Dense wide-baseline stereo reconstruction with Graph Cuts

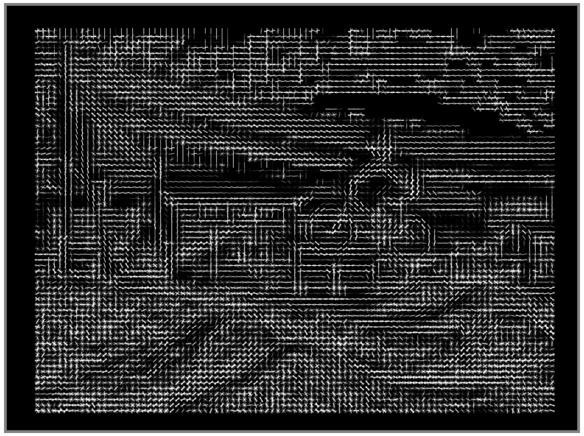


Code

- http://vision.mas.ecp.fr/Personnel/iasonas/descriptors.html
- http://www.iri.upc.edu/people/etrulls/#code

Features for 3D data

- Intensity data: Histogram of Gradients, SIFT etc.



- Depth/surface data?

Kinect

Kinect Fusion

Dense invariant surface descriptors

- Recent advances in surface analysis
- Combination with computer vision techniques

Michael Bronstein

Alex Bronstein

Roee Litman

M. Bronstein, I. Kokkinos, Scale-Invariant Heat Kernel Signatures, CVPR 2010

I. Kokkinos, M. Bronstein, R. Litman, A. Bronstein, Intrinsic Shape Context, CVPR 2012

Slides from:

http://www.cs.technion.ac.il/~mbron/teaching.html
http://tosca.cs.technion.ac.il/book/course_milano08.html

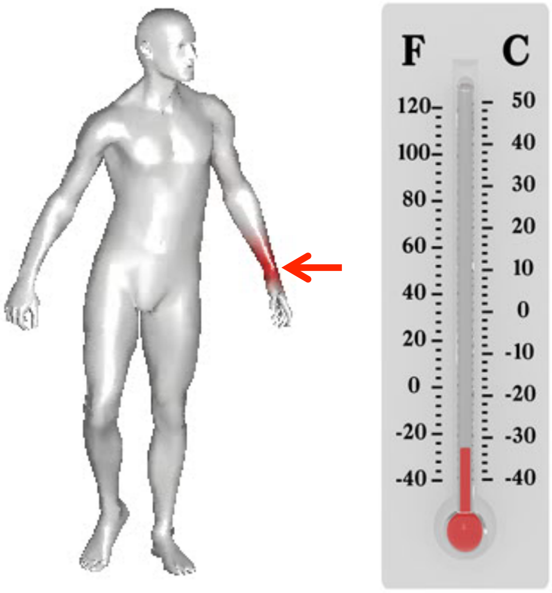
Diffusion geometry

Heat equation $\left(\Delta_S + \frac{\partial}{\partial t}\right) u(t,x) = 0$

where

Δ_S - Laplace-Beltrami operator

u - heat distribution



Fundamental solution: solution for initial conditions $u(x,0) = \delta(x-y)$

Represented in the Laplace-Beltrami eigenbasis as

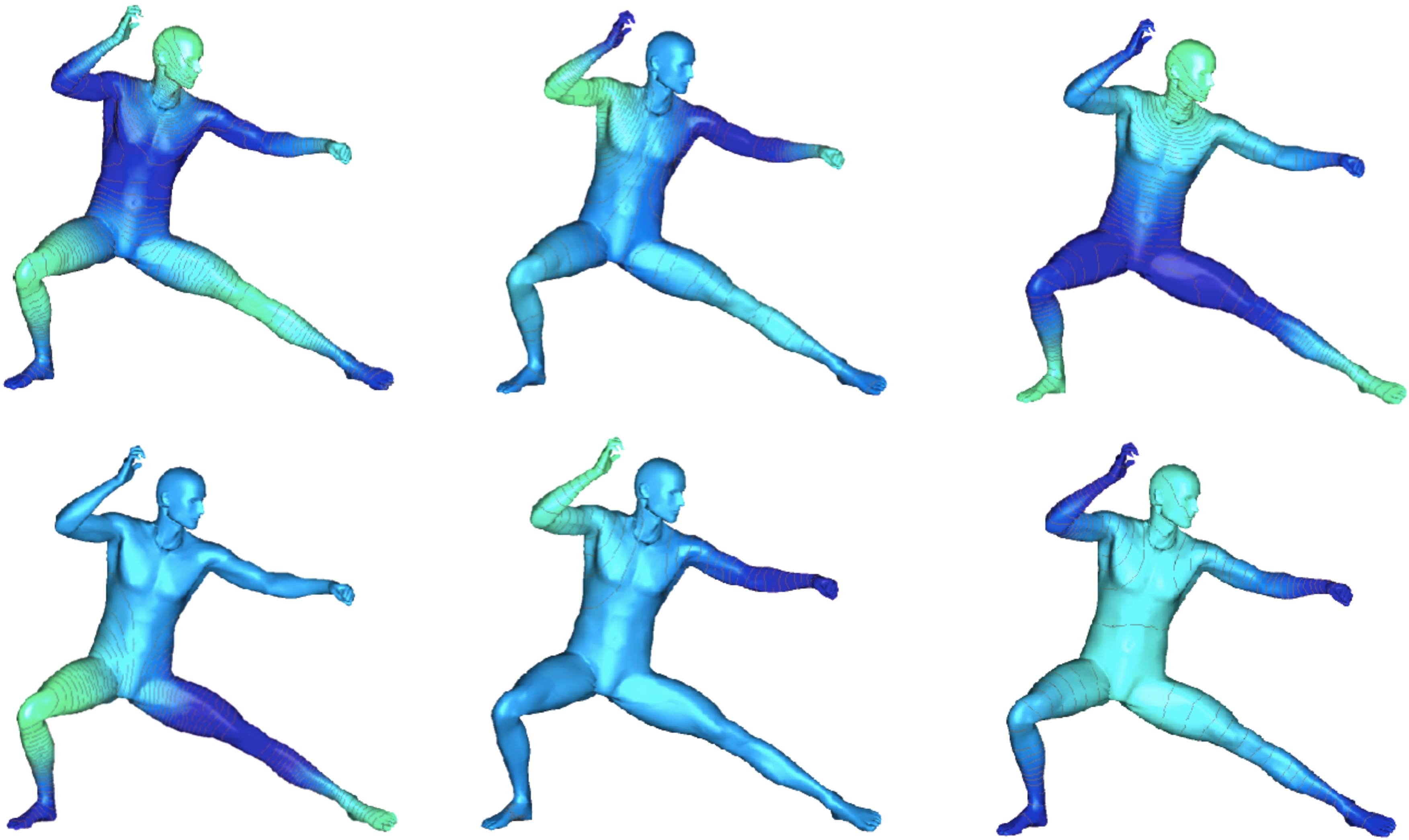
$$k_t(x,y) = \sum_{i=0}^{\infty} e^{-t\lambda_i} \phi_i(x)\phi_i(y)$$

Coefficients decay fast: approximation by **truncated sum**

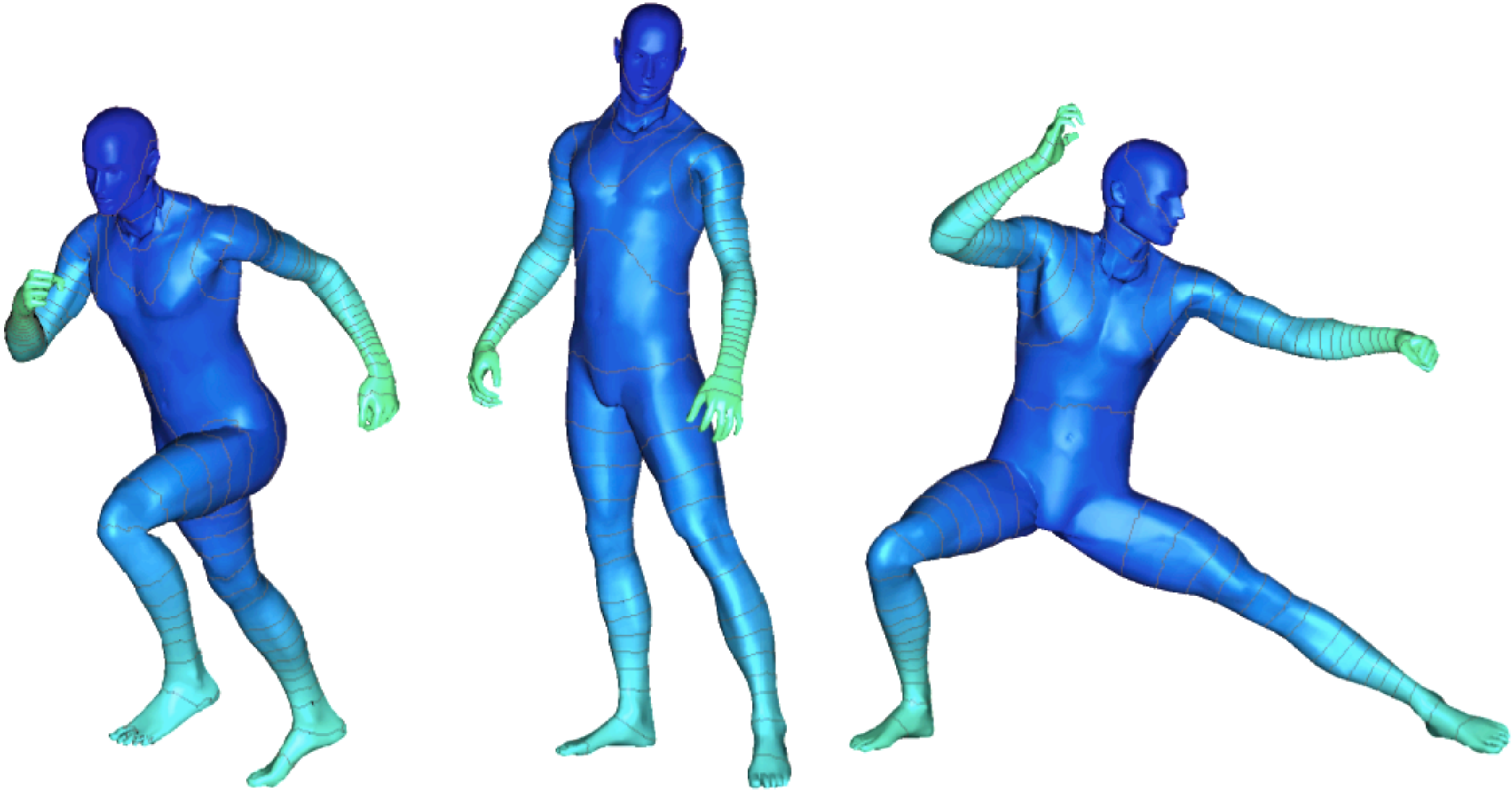
$$k_t(x,y) \approx \sum_{i=0}^{N} e^{-t\lambda_i} \phi_i(x)\phi_i(y)$$

Chladni's patterns
http://www.youtube.com/watch?v=cT30XOfd1yI

Laplace-Beltrami eigenfunctions



Laplace-Beltrami eigenfunctions: invariance

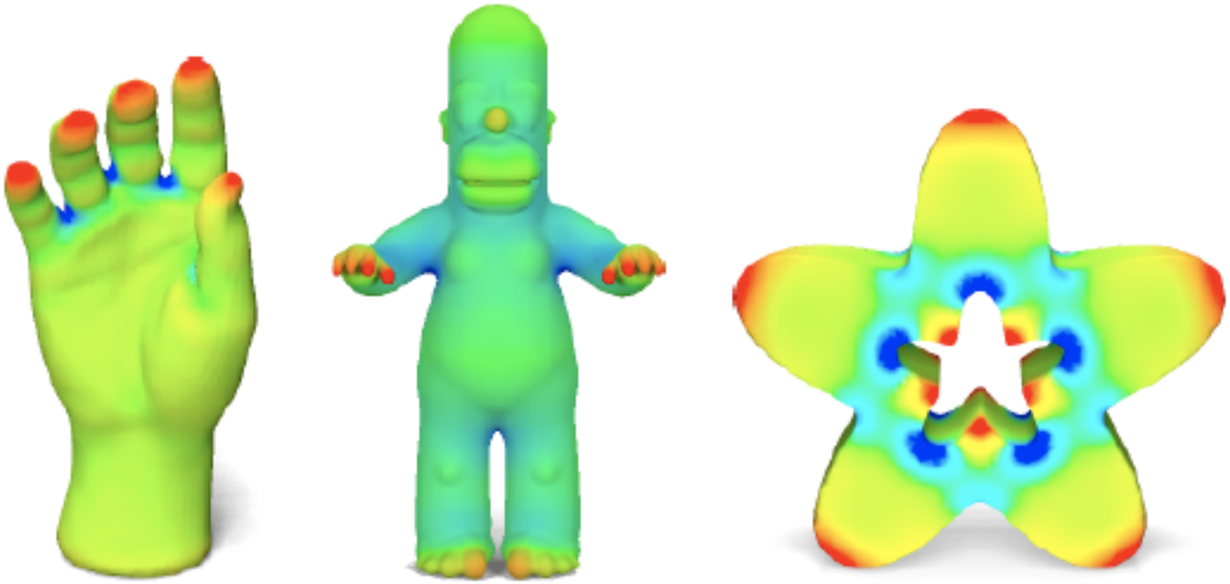


From the heat kernel to heat kernel signature

Multiscale local shape descriptor

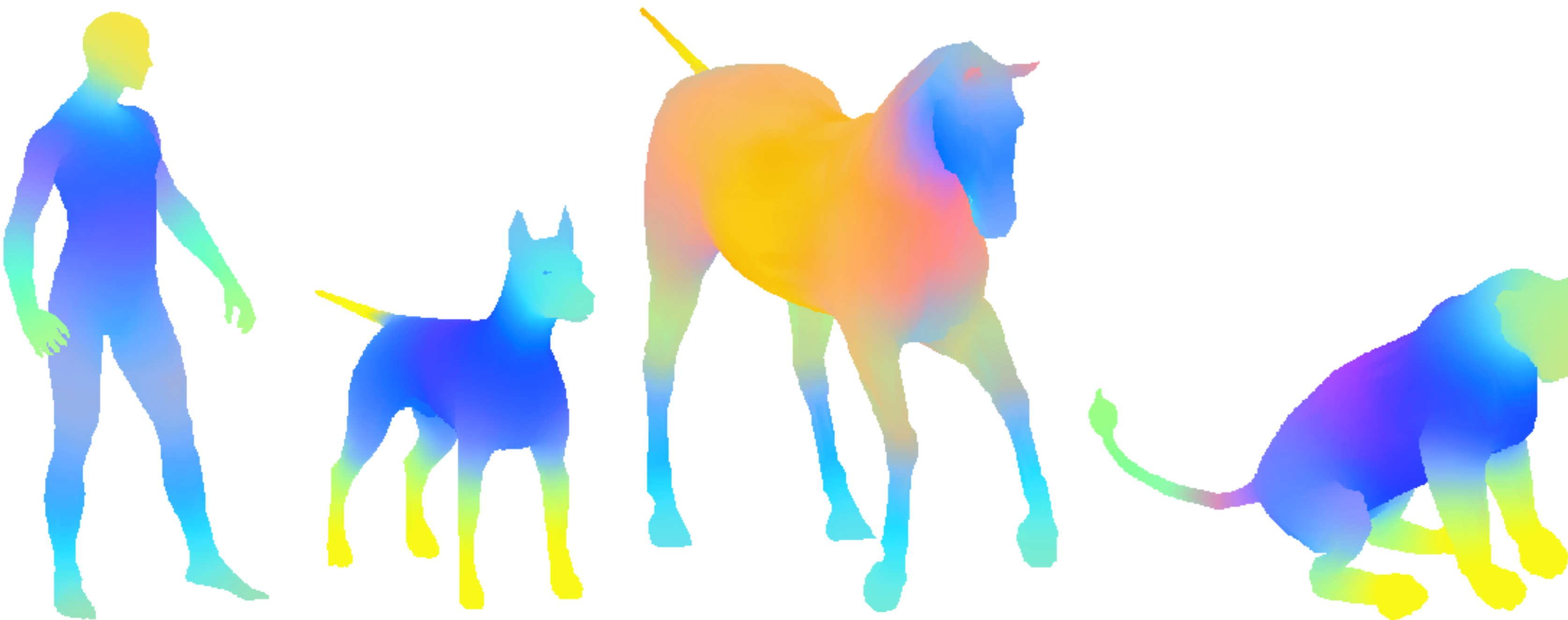
$$p(x) = (K_{t_1}(x,x), ..., K_{t_n}(x,x))$$

Interpretation: multiscale gaussian curvature (time= scale)



J. Sun, M. Ovsjanikov, L. Guibas, *SGP* 2009

Heat kernel signature



Heat kernel signatures represented in RGB space

J. Sun, M. Ovsjanikov, L. Guibas, *SGP* 2009

Heat kernel descriptors

Invariant to isometric deformations

Localized sensitivity to topological noise

J. Sun, M. Ovsjanikov, L. Guibas, *SGP* 2009

Scale invariance?

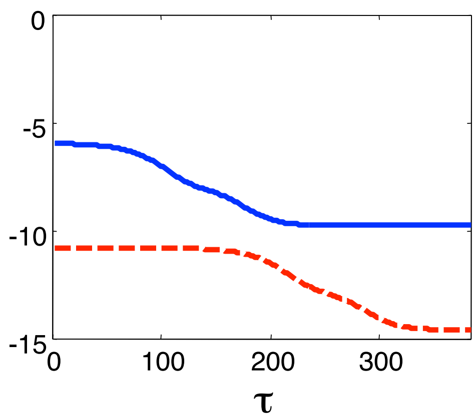
Original shape	Scaled by $1/\alpha$
λ, ϕ	$\alpha^2\lambda, \alpha\phi$
HKS= $K_t(x,x)$	HKS= $\alpha^2 K_{\alpha^2 t}(x,x)$

Not scale invariant!

Scale-invariant heat kernel signature

Log scale-space

$$K_{\alpha^\tau} \rightarrow \alpha^2 K_{\alpha^{\tau+2}}$$

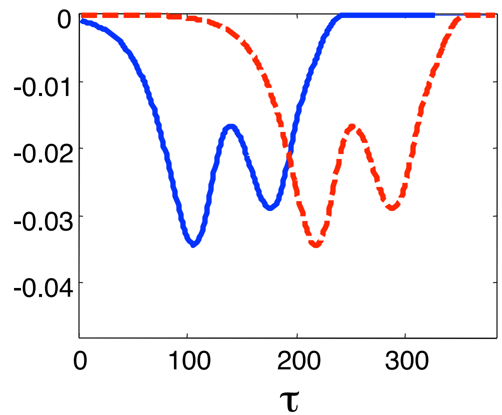


Scaling = shift and multiplicative constant in HKS

log + d/dτ

$$\log \alpha^2 + \log K_{\alpha^{\tau+2}}$$

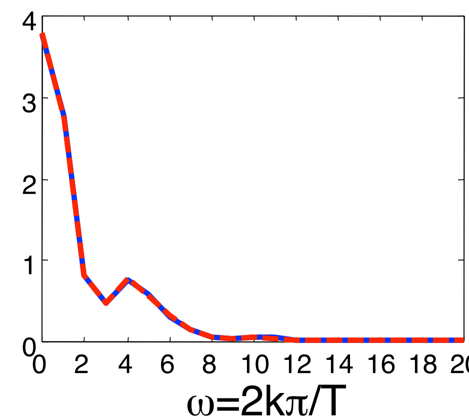
$$d/d\tau \log K_{\alpha^{\tau+2}}$$



Undo scaling

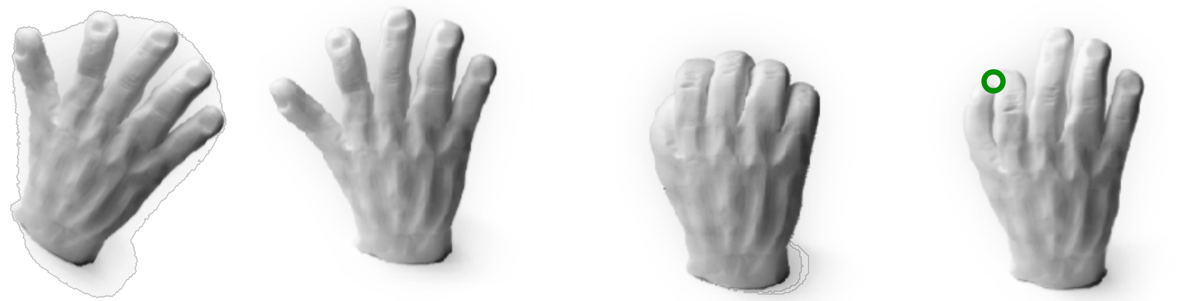
Fourier transform magnitude

$$\mathcal{F} d/d\tau \log K_{\alpha^{\tau+2}} = e^{2i\omega\pi} \mathcal{F} d/d\tau \log K_{\alpha^\tau}$$



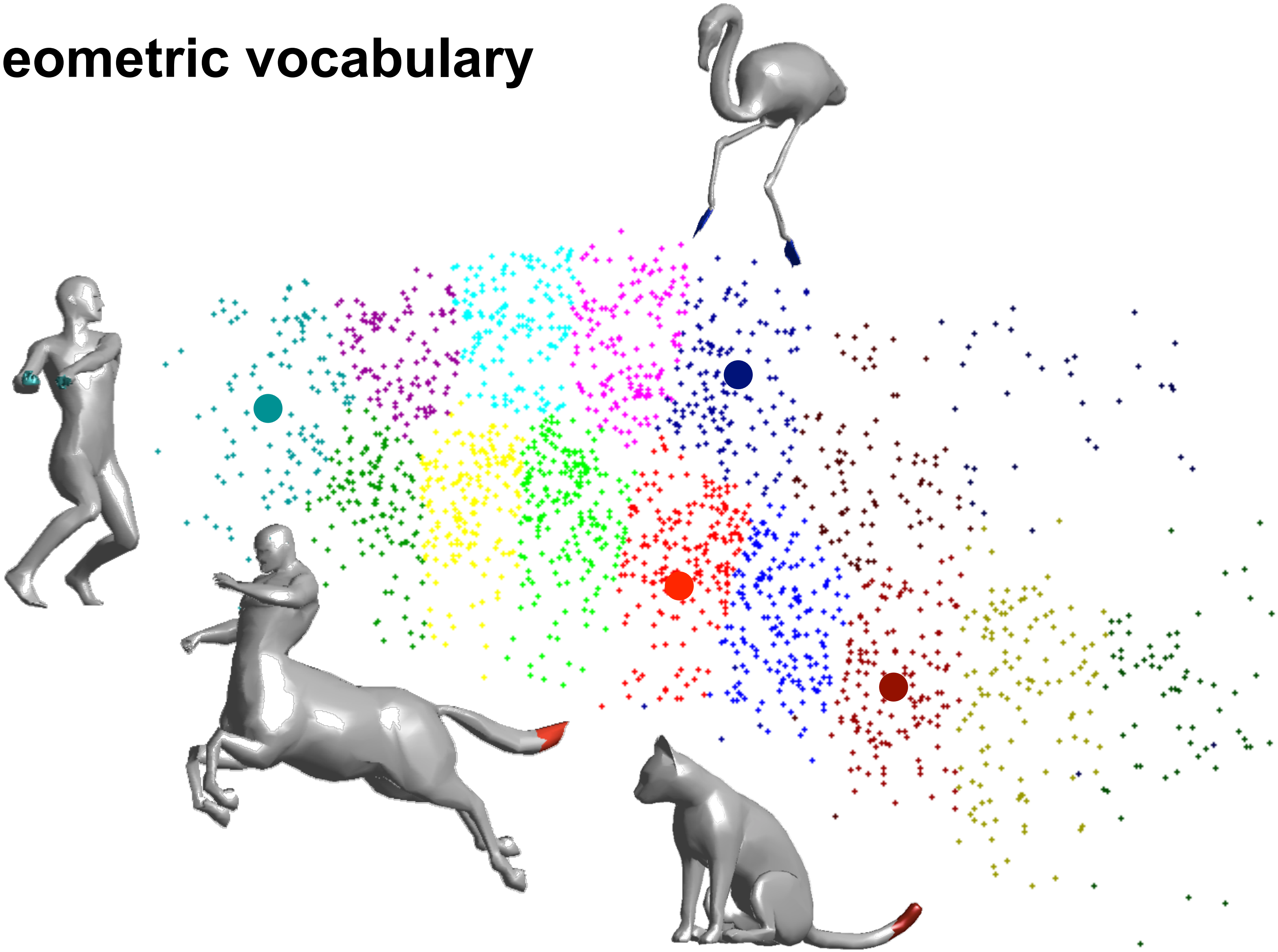
Undo shift

Invariance



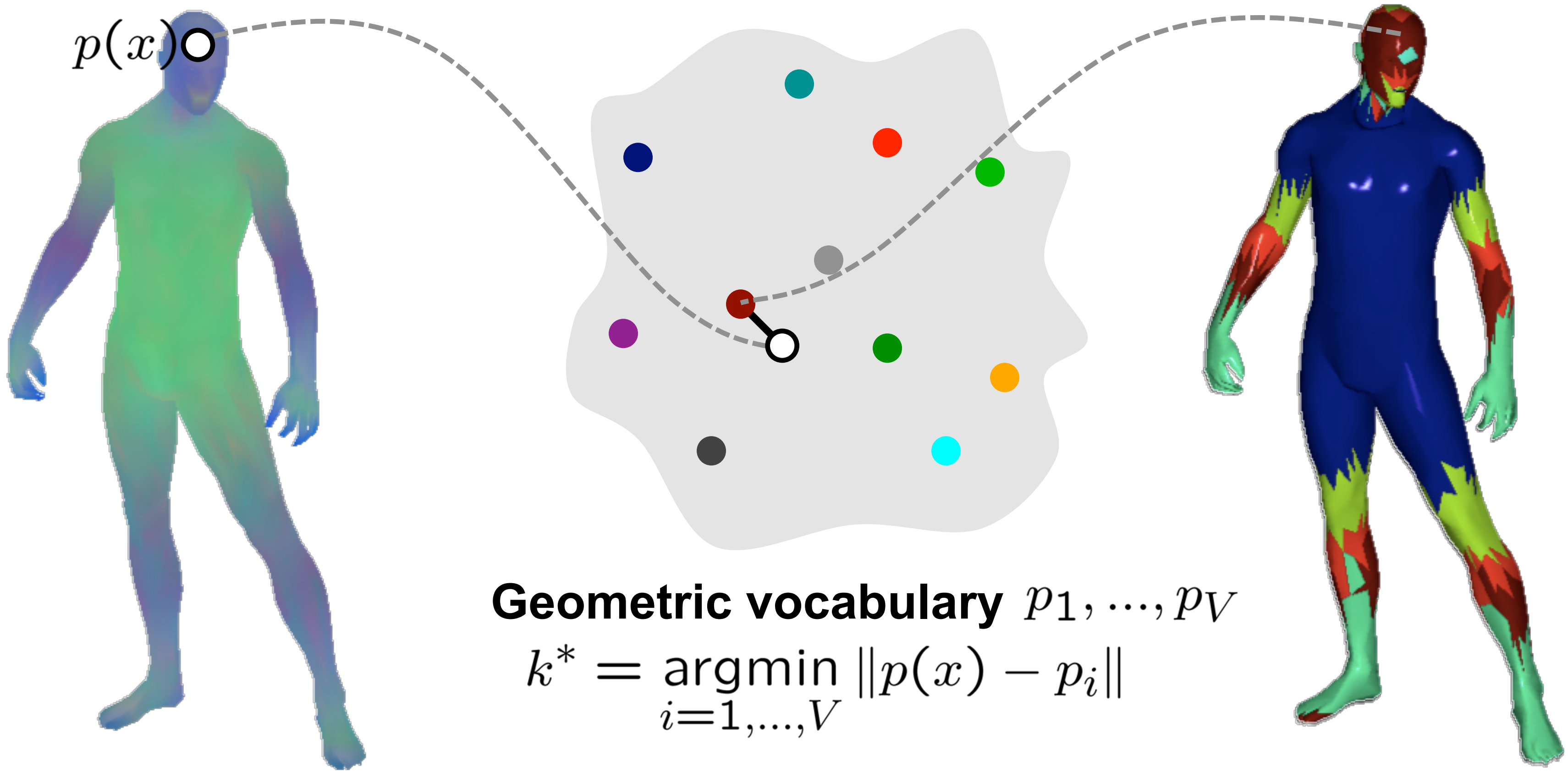
	Rigid	Scale	Inelastic	Topology
Heat kernel signature (HKS)	✓	✗	✓	✓
Scale-invariant HKS (SI-HKS)	✓	✓	✓	✓

Geometric vocabulary



M. Ovsjanikov, M Bronstein A. Bronstein, L. Guibas, 2009

Bags of features



Geometric vocabulary $p_1, \dots, p_V$

$$k^* = \operatorname*{argmin}_{i=1,\dots,V} \|p(x) - p_i\|$$

Nearest neighbor in the descriptor space

$$\theta_k(x) = \begin{cases} 1 & k = k^* \\ 0 & \text{else} \end{cases}$$

M. Ovsjanikov, M Bronstein A. Bronstein, L. Guibas, 2009

Bags of features

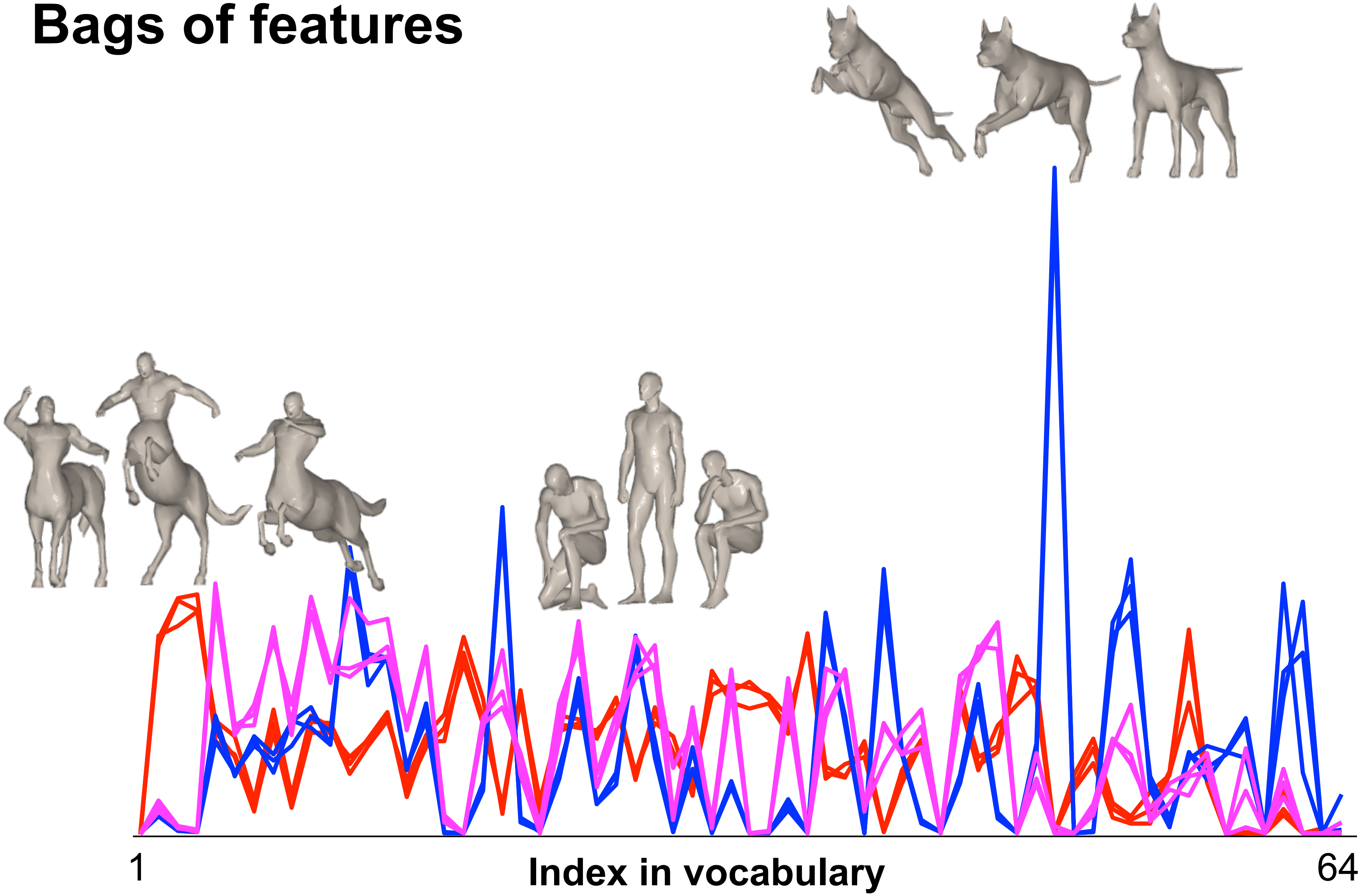
Statistics of different geometric words over the entire shape

$$f(X) = \int_X \theta(x)dx$$

Shape distance = distance between bags of features

$$d_{\text{BoF}}(X,Y) = \|f(X) - f(Y)\|$$

Bags of features



M. Ovsjanikov, M Bronstein A. Bronstein, L. Guibas, 2009

Transform.	Strength 1	≤2	≤3	≤4	≤5
Isometry	100.00	100.00	100.00	100.00	100.00
Topology	100.00	98.08	97.44	96.79	96.41
Holes	100.00	100.00	97.44	95.19	90.13
Micro holes	100.00	100.00	100.00	100.00	100.00
Scale	0.98	40.68	43.31	33.72	27.42
Local scale	100.00	100.00	98.72	89.38	80.22
Sampling	100.00	100.00	100.00	100.00	99.23
Noise	100.00	100.00	100.00	100.00	100.00
Shot noise	100.00	100.00	100.00	100.00	100.00
Partial	7.54	5.70	4.51	3.58	2.95
Mixed	53.13	55.86	47.77	37.54	30.34
Average	94.94	93.12	90.84	87.82	85.00

ShapeGoogle with HKS descriptor (mAP %)

	Strength				
Transform.	**1**	$\leq$**2**	$\leq$**3**	$\leq$**4**	$\leq$**5**
Isometry	100.00	100.00	100.00	100.00	100.00
Topology	96.15	96.15	94.87	93.27	92.69
Holes	100.00	100.00	100.00	94.71	89.97
Micro holes	100.00	100.00	100.00	100.00	100.00
Scale	91.03	95.51	97.01	97.76	98.21
Local scale	100.00	100.00	97.44	89.38	82.08
Sampling	100.00	100.00	100.00	100.00	97.69
Noise	100.00	100.00	100.00	100.00	100.00
Shot noise	100.00	100.00	100.00	100.00	100.00
Partial	17.43	10.31	9.57	8.06	6.61
Mixed	56.47	57.44	63.59	67.47	65.07
Average	97.05	95.16	94.03	92.54	90.79

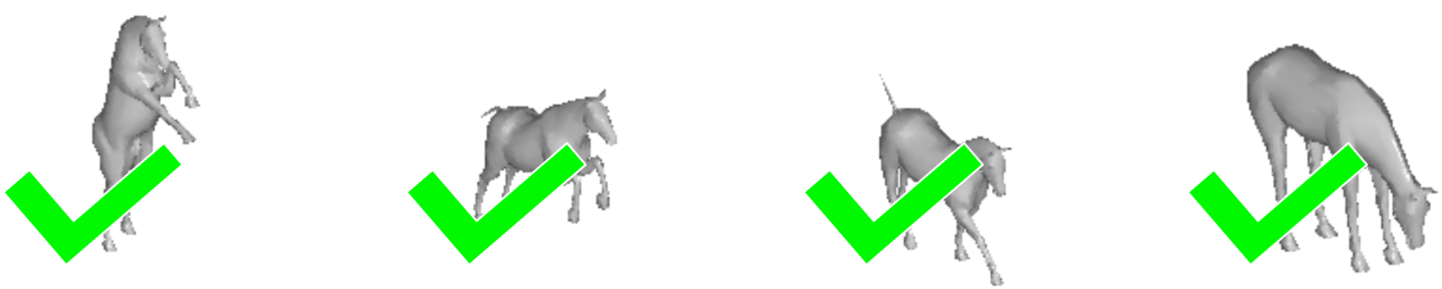
ShapeGoogle with SI-HKS descriptor (mAP %)

M. Bronstein, I.K., CVPR 2010

Scale-invariant retrieval

Scale 1.3

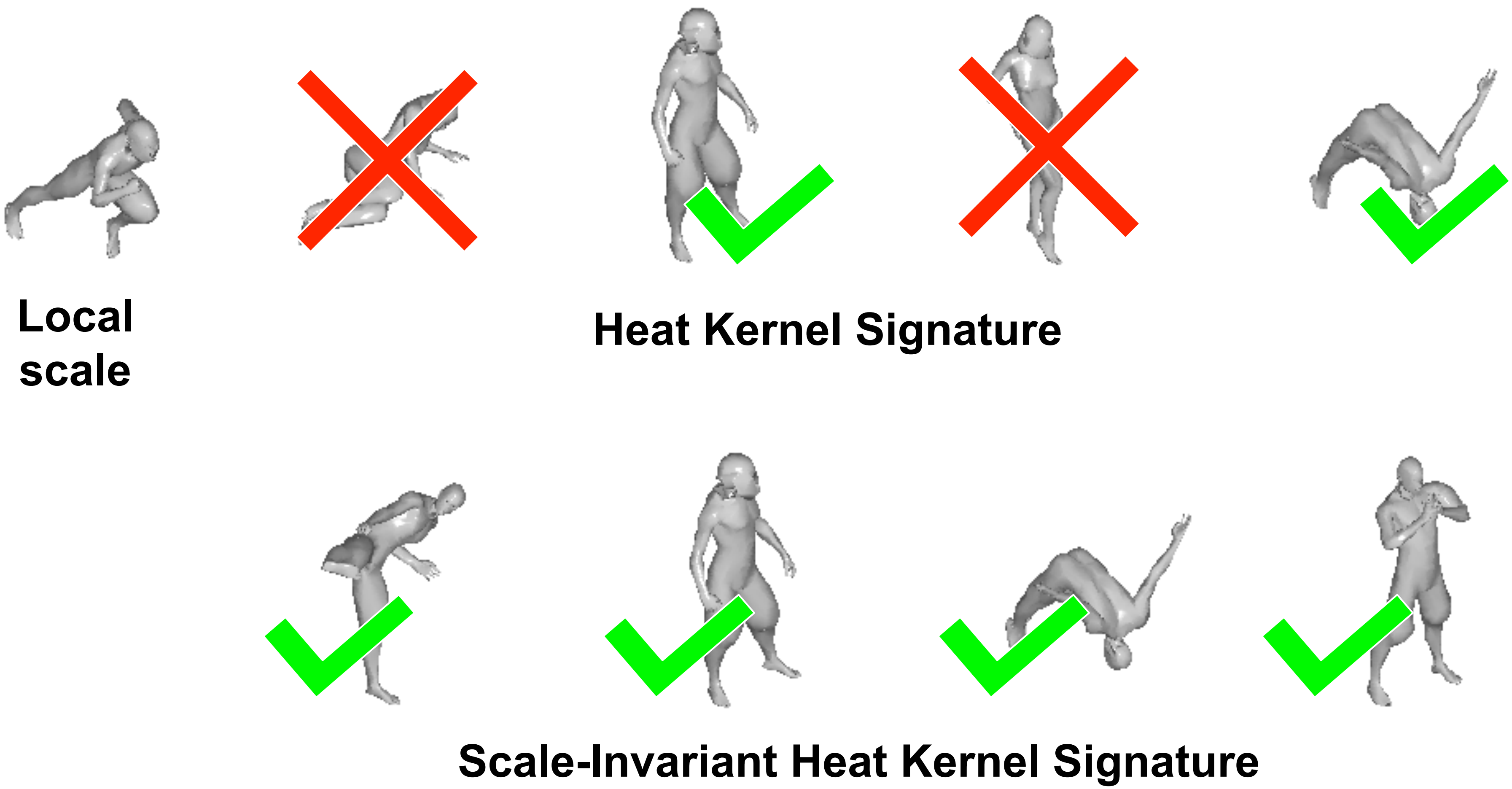
Heat Kernel Signature



Scale-Invariant Heat Kernel Signature

M. Bronstein, I.K., 2010

Scale-invariant retrieval



M. Bronstein, I.K., 2010

Intrinsic Shape Contexts

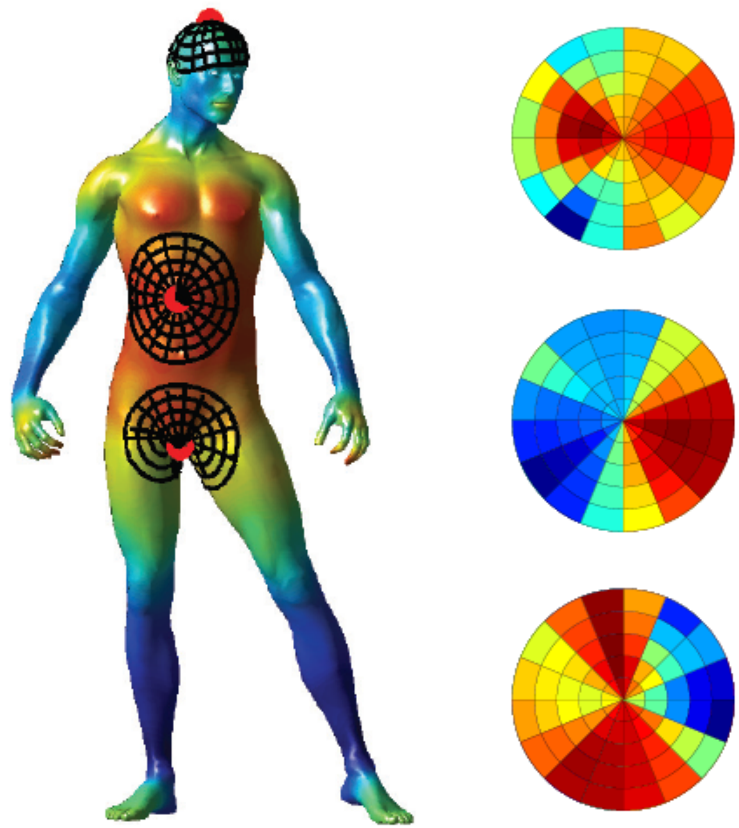
Goal: use context information for descriptor construction

Naïve: work at larger feature scale (more smoothing)

Better: stack together neighbors (meta-descriptor)

Surface processing: local surface charting

Intrinsic: invariant to isometric deformations



IK, M. Bronstein, R. Littman, A. Bronstein, Intrinsic Shape Contexts , CVPR 2012

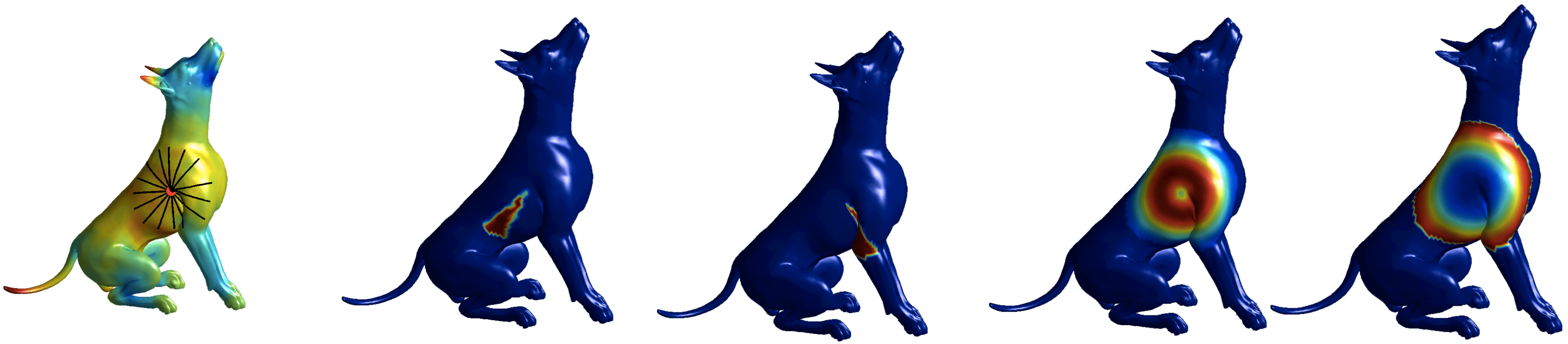
Intrinsic Shape Context Construction

Uniformly sample directions on the tangent plane of a point

Shoot and track geodesics outwards from the point

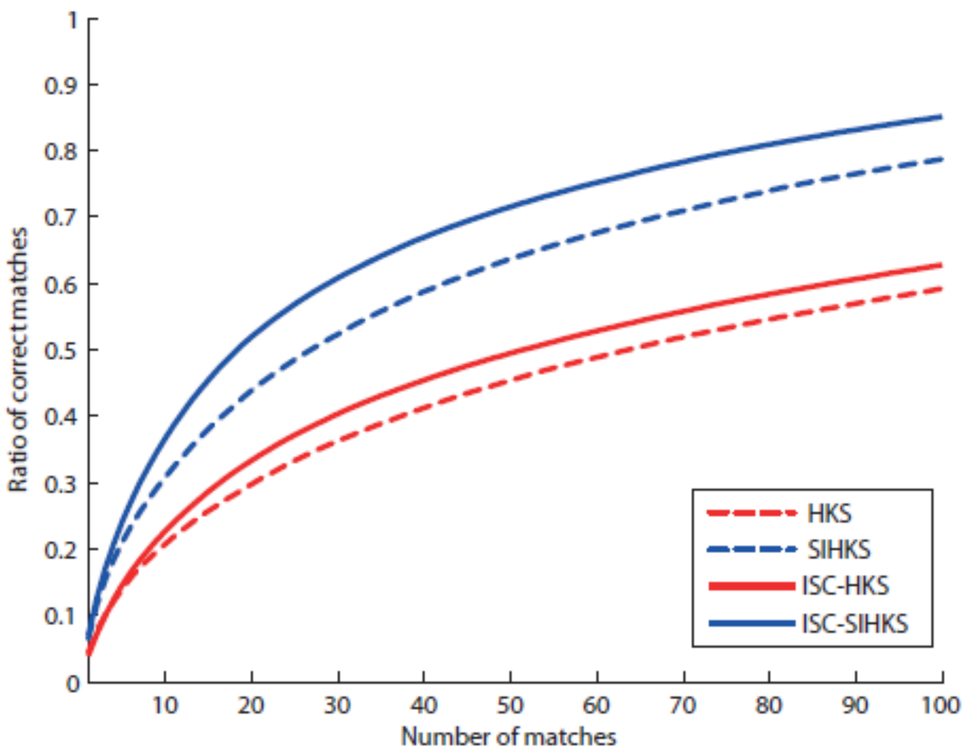
Construct soft angular membership functions based on distance from geodesics

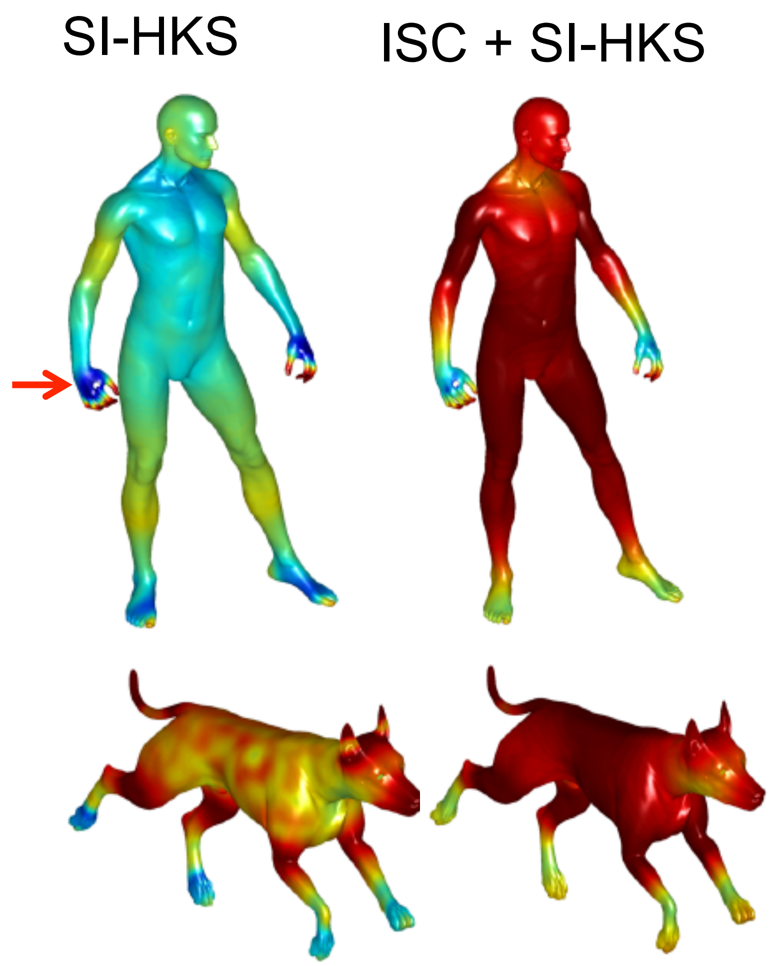
Construct soft radial membership functions based on distance from point



IK, M. Bronstein, R. Littman, A. Bronstein, Intrinsic Shape Contexts, CVPR 2012

Intrinsic Shape Contexts: higher discrimination





IK, M. Bronstein, R. Littman, A. Bronstein, ISC, CVPR 2012