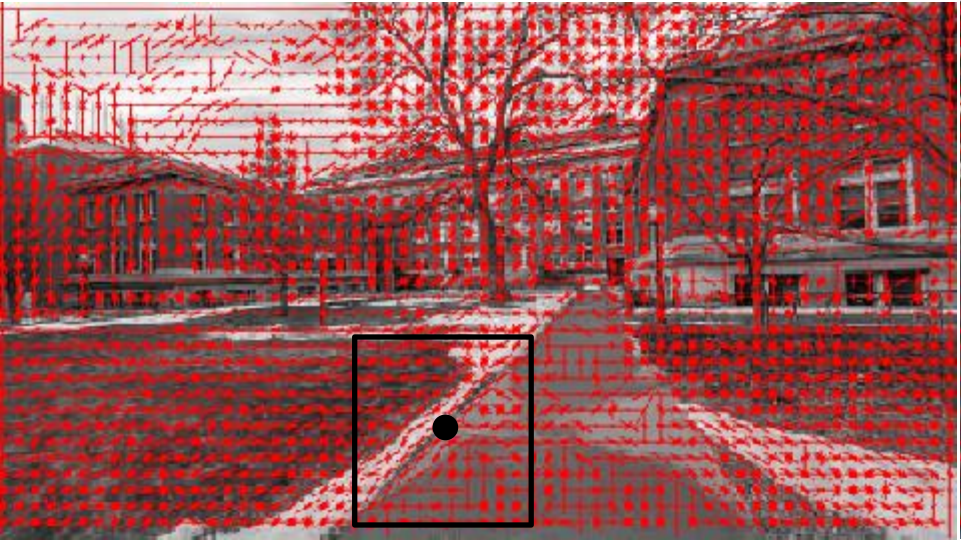
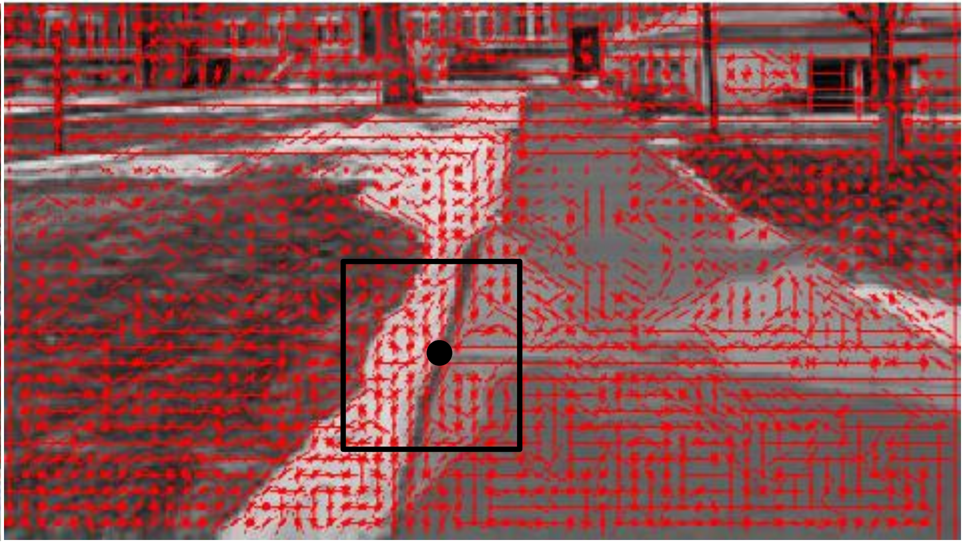
Scale-invariant Feature Transform

Hyun Soo Park

FEATURE MATCHING

RECALL: HOG





corr [ , ] = 0.15

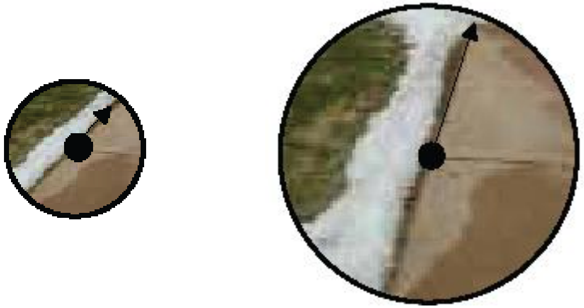
What's wrong?

Challenges of Feature Matching

- Scale
- Orientation

FEATURE NORMALIZATION

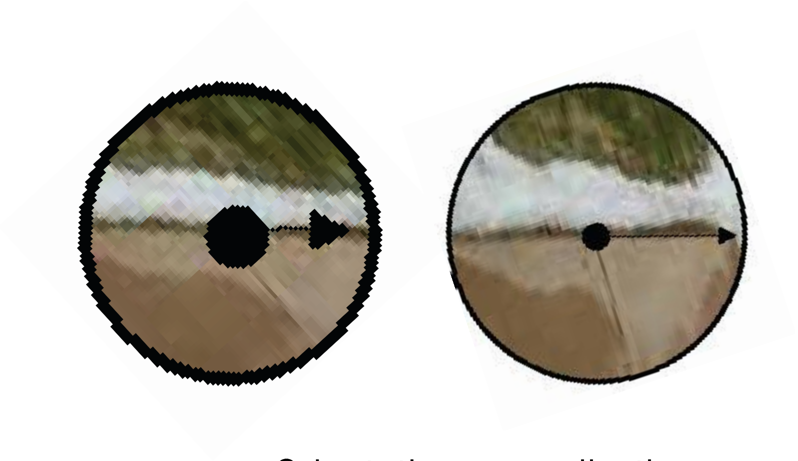
FEATURE NORMALIZATION



FEATURE NORMALIZATION

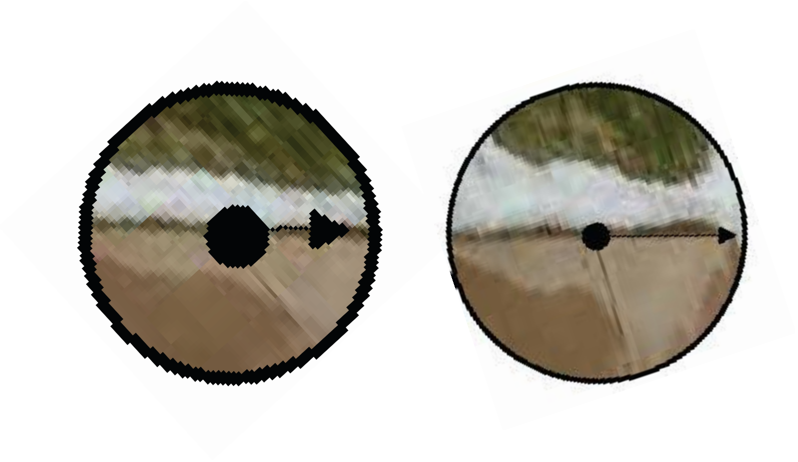
Scale normalization

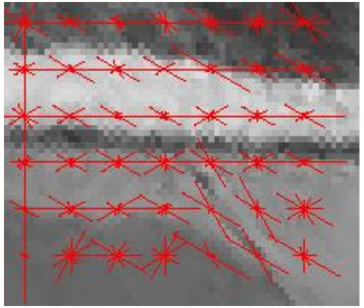
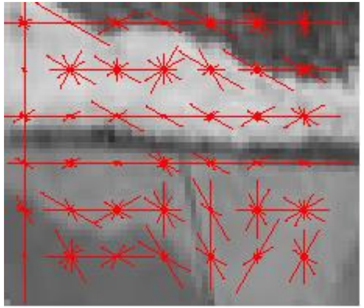
FEATURE NORMALIZATION



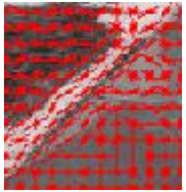
Orientation normalization

Scale-invariant Feature Transform (SIFT)



corr [ , ] = 0.91

cf)

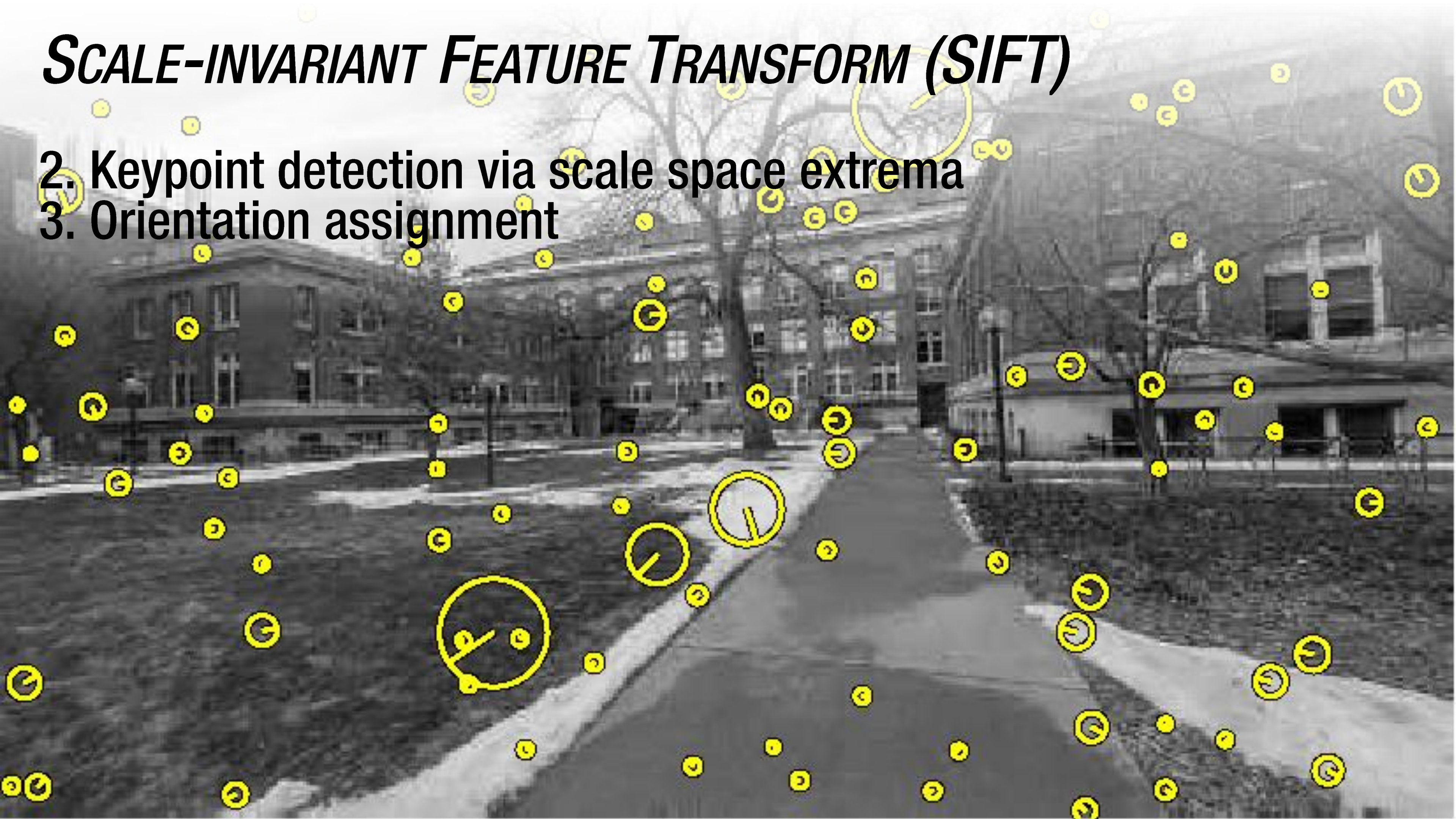
corr [ , ] = 0.15

Scale-invariant Feature Transform (SIFT)

1. Input image

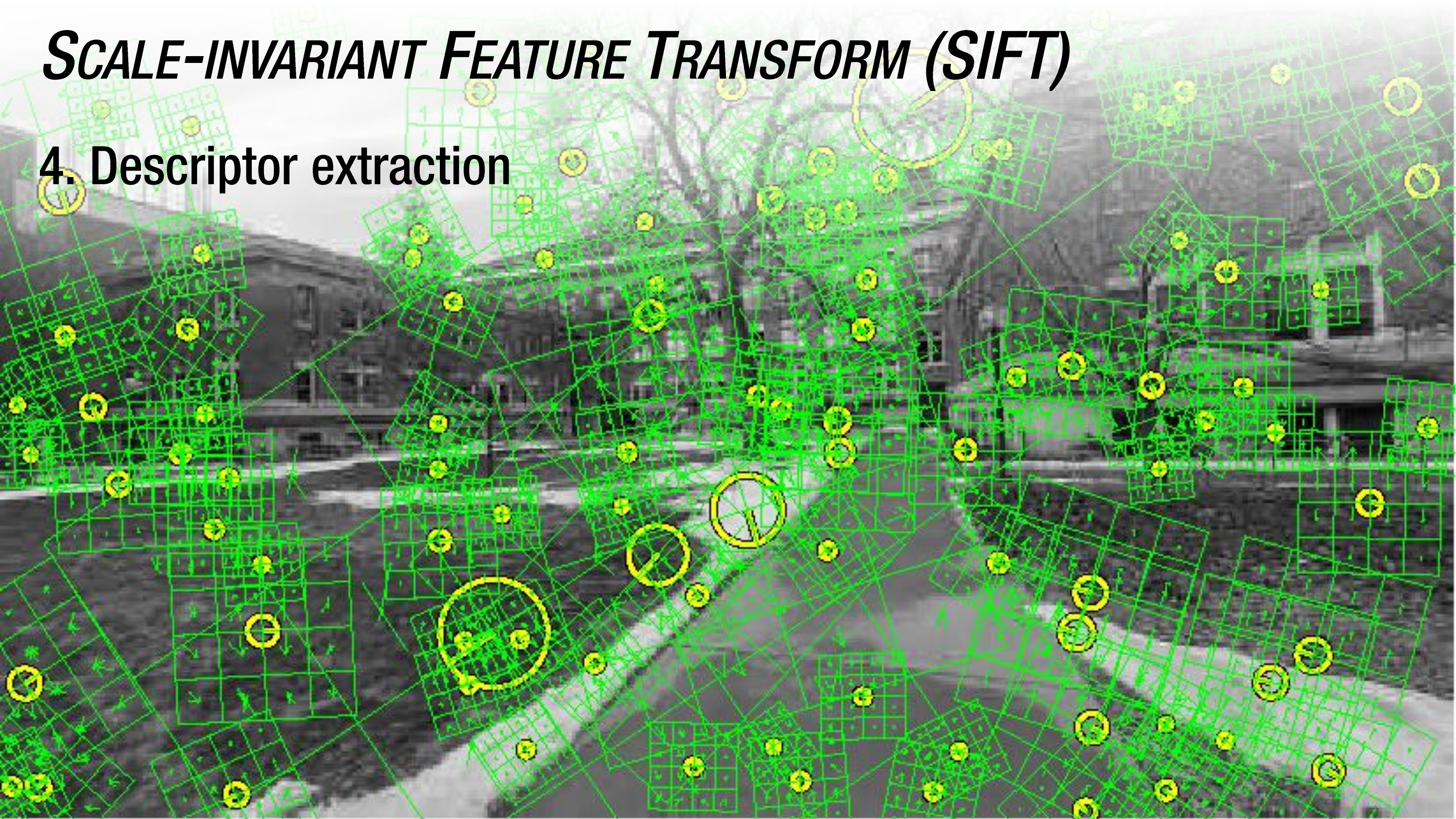
SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

2. Keypoint detection via scale space extrema
3. Orientation assignment



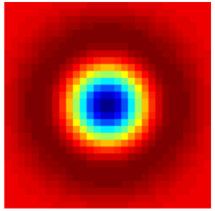
SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

4. Descriptor extraction



Keypoint Detection: Scale Space Extrema

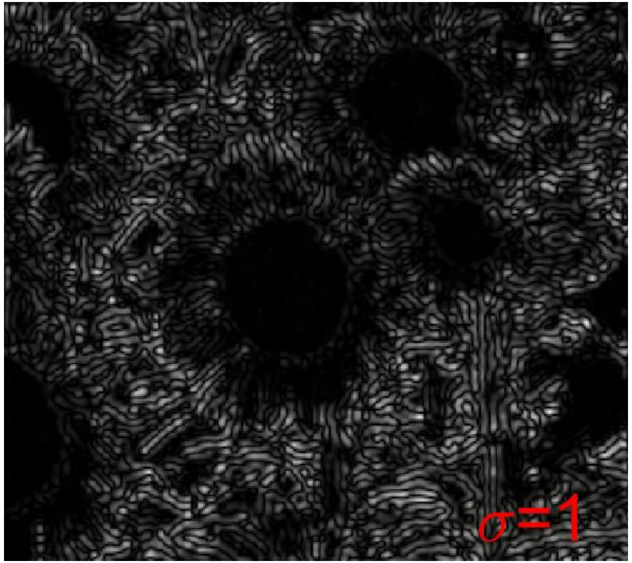
*



$$L = \sigma^2 \nabla^2 G$$

Laplacian of Gaussian (LoG)

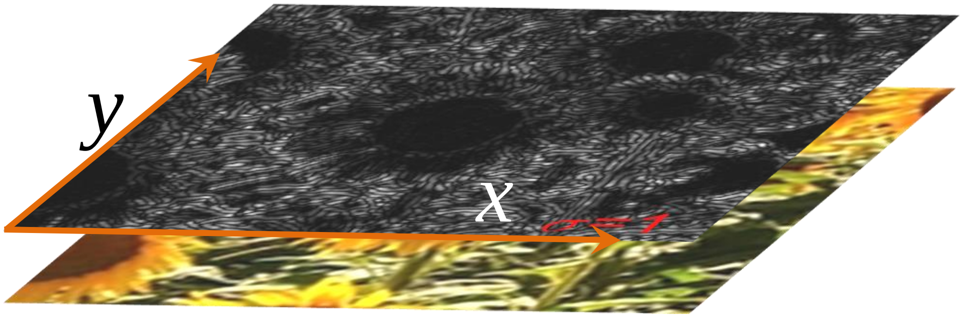
=



Keypoint Detection: Scale Space Extrema

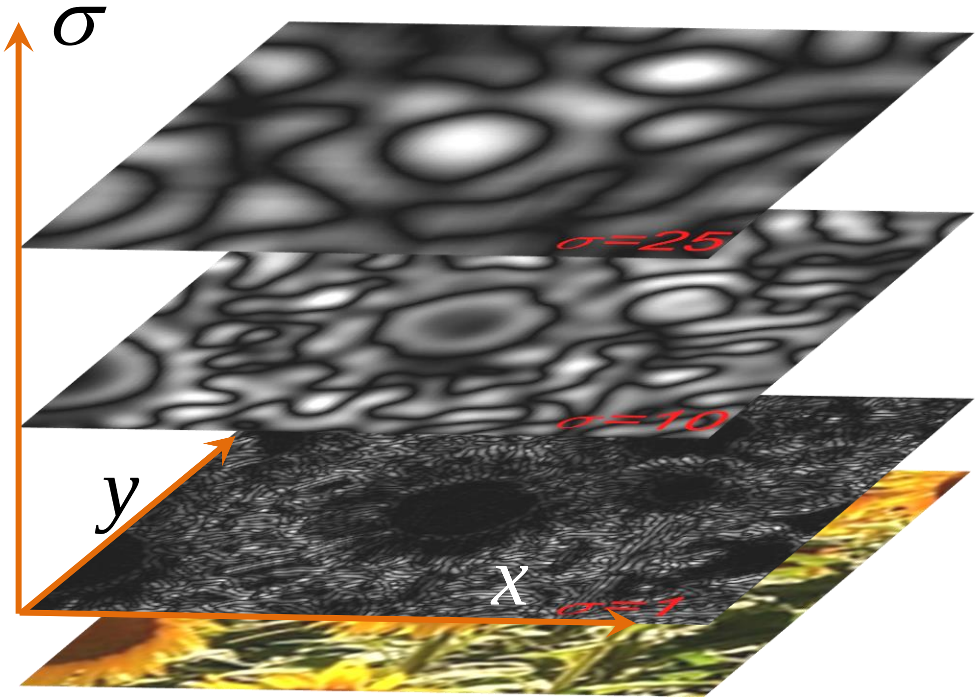
$I(x, y)$

Keypoint Detection: Scale Space Extrema



$$I(x, y) * L(\sigma)$$

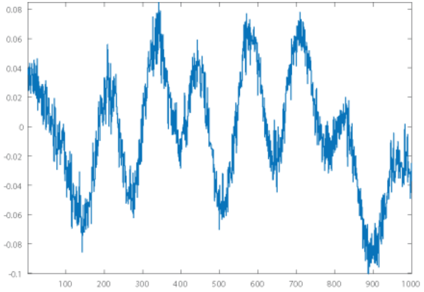
Keypoint Detection: Scale Space Extrema



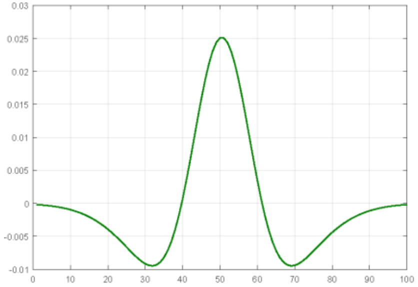
$$D(x, y, \sigma) = I(x, y) * L(\sigma)$$

Scale space response

RECALL: LAPLACIAN OF GAUSSIAN (LOG) ~ DOG

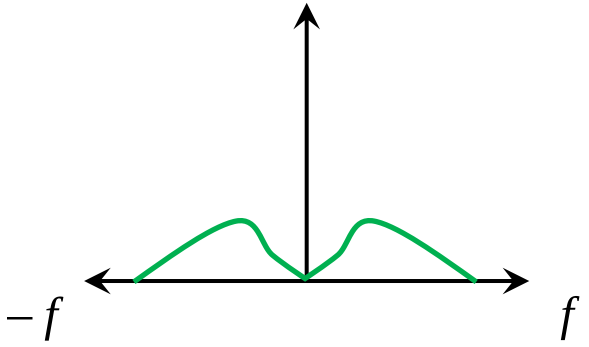


$*$



FT →

Inverse FT ←



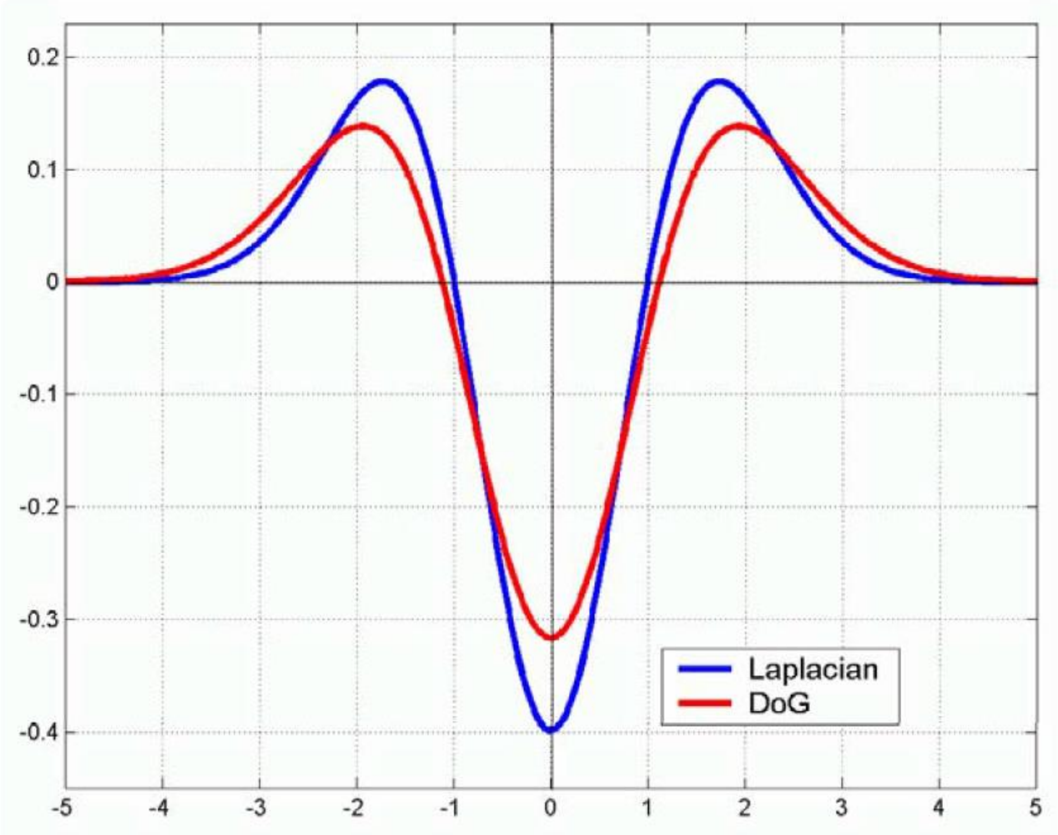
$x(t)$ $*$ $g(t;\sigma_1) - g(t;\sigma_2)$

$\underline{\approx \nabla \cdot \nabla g}$

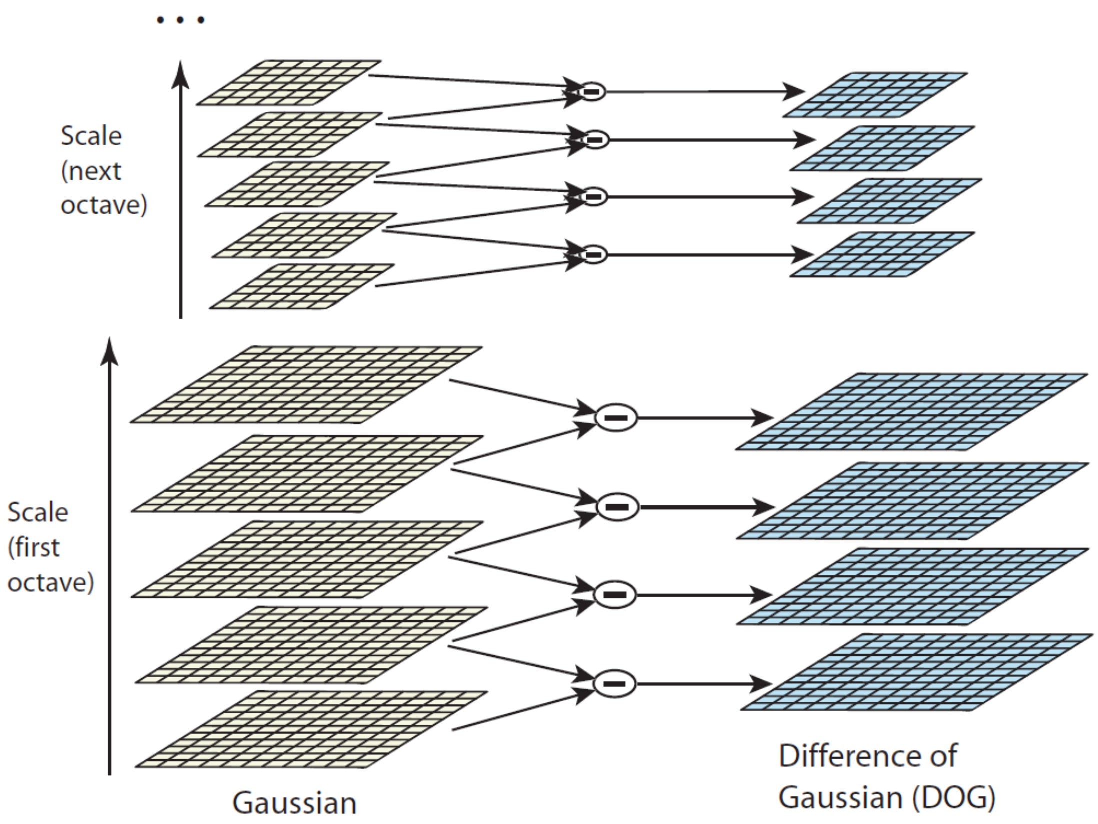
Laplacian of Gaussian

$X(f)\big(G(f;\sigma_1) - G(f;\sigma_2)\big)$

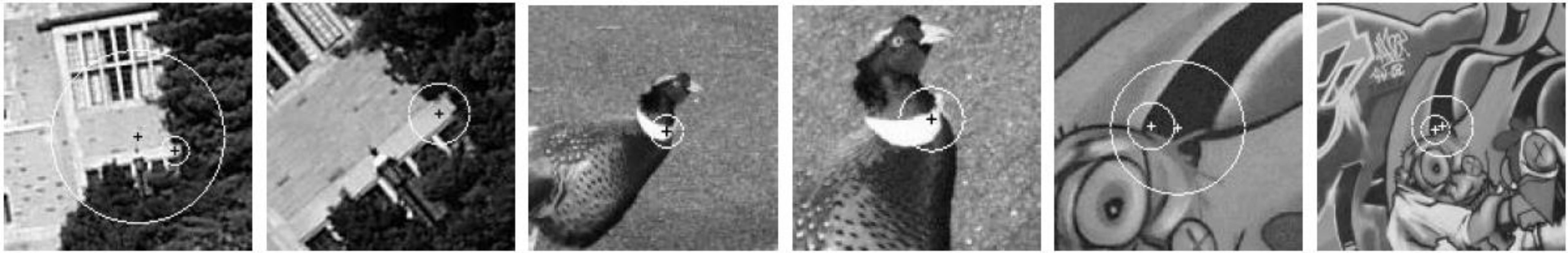
RECALL: LAPLACIAN OF GAUSSIAN (LOG) ~ DOG



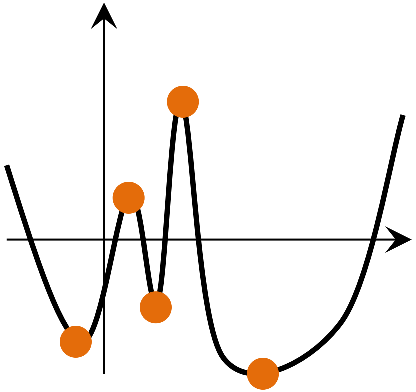
Laplacian ~ Difference of Gaussian



WHAT CHARACTERIZES KEYPOINT?

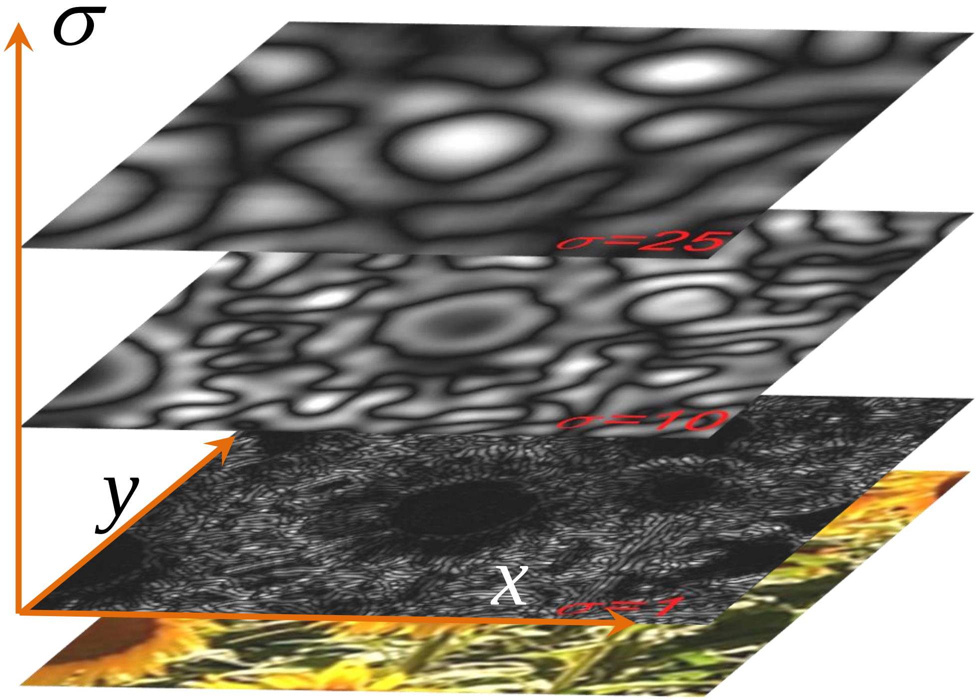


- Robust
- Repeatable
- Uniqueness



Local extrema (minima/maxma)

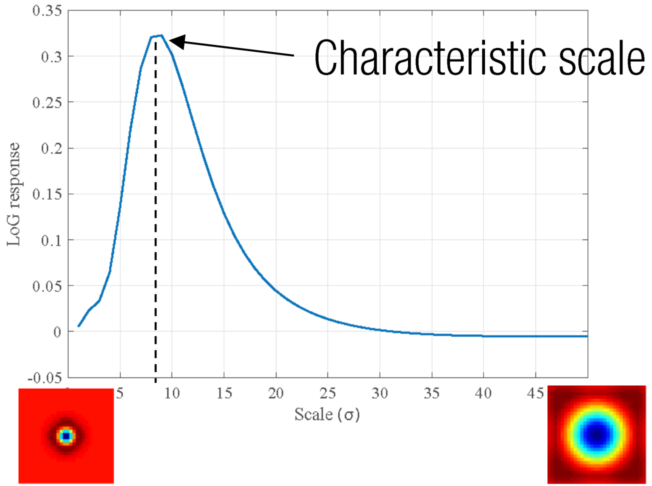
WHAT CHARACTERIZES KEYPOINT?



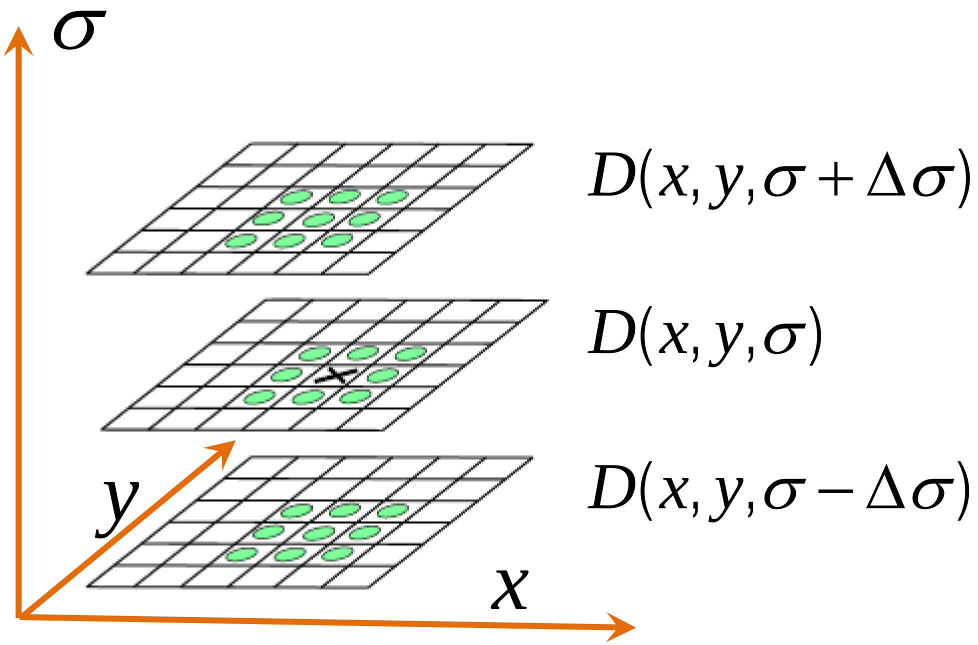
$$\underline{D(x,y,\sigma)} = I(x,y) * L(\sigma)$$

Scale space response

Local extrema (minima/maxma)



What Characterizes Keypoint?



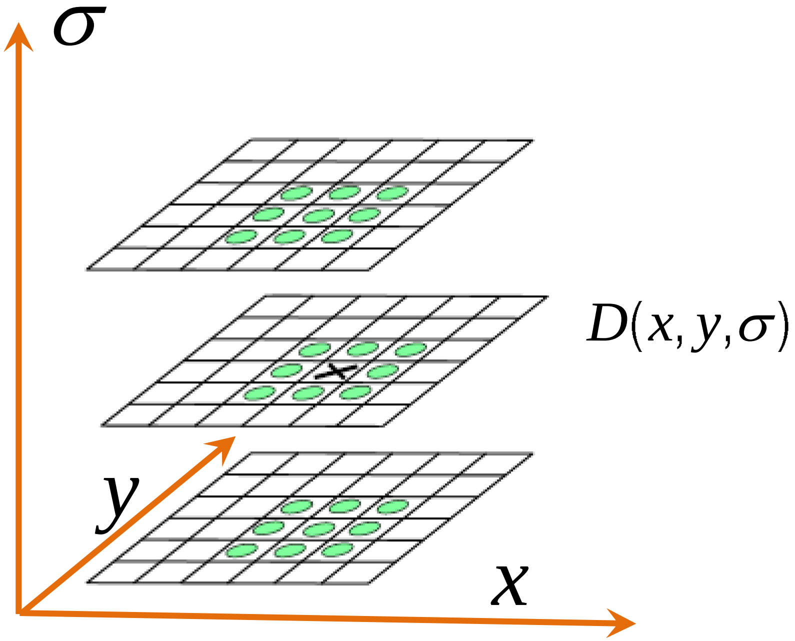
(x, y) is keypoint if

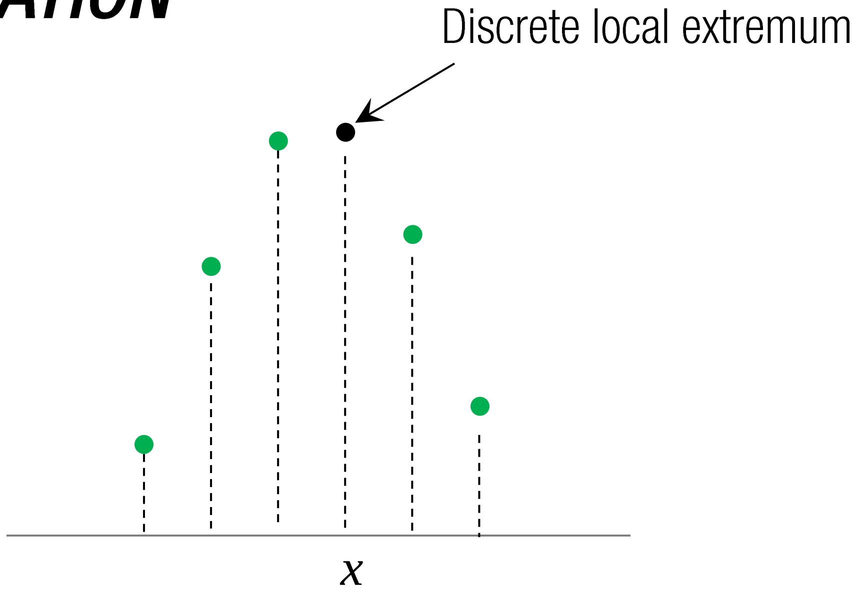
$$D(x,y,\sigma) > D(x \pm \Delta x, y \pm \Delta y, \sigma \pm \Delta\sigma)$$

or

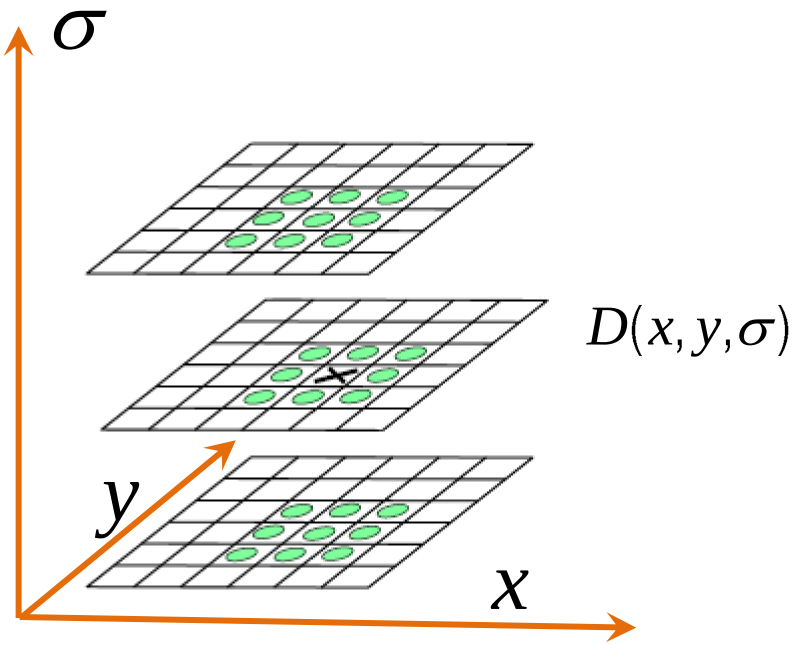
$$D(x,y,\sigma) < D(x \pm \Delta x, y \pm \Delta y, \sigma \pm \Delta\sigma)$$

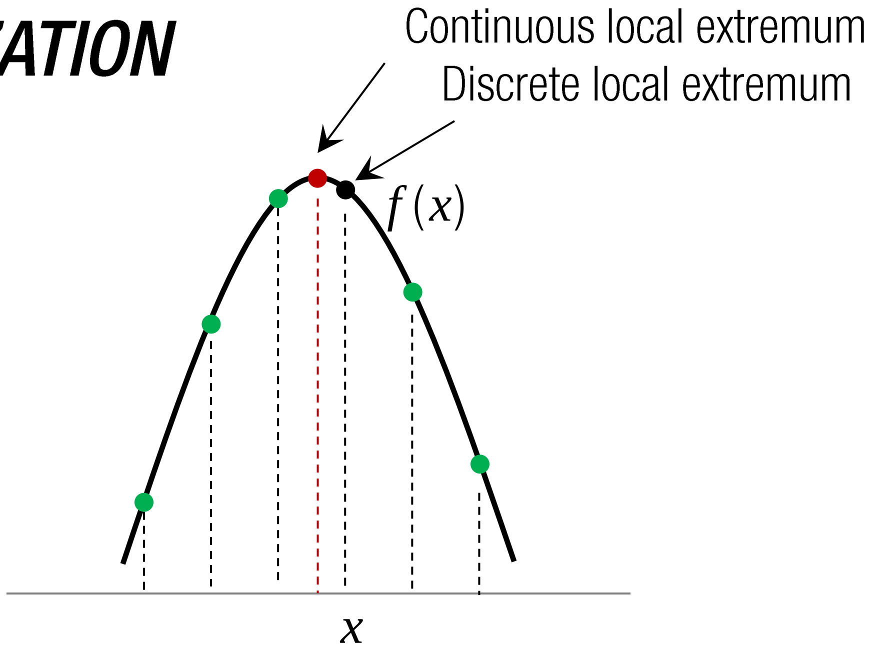
Subpixel Keypoint Localization



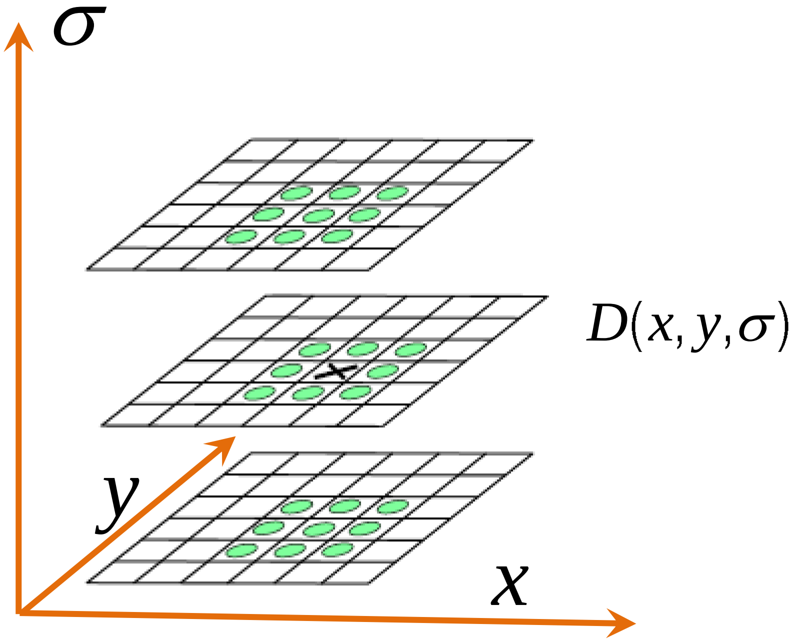


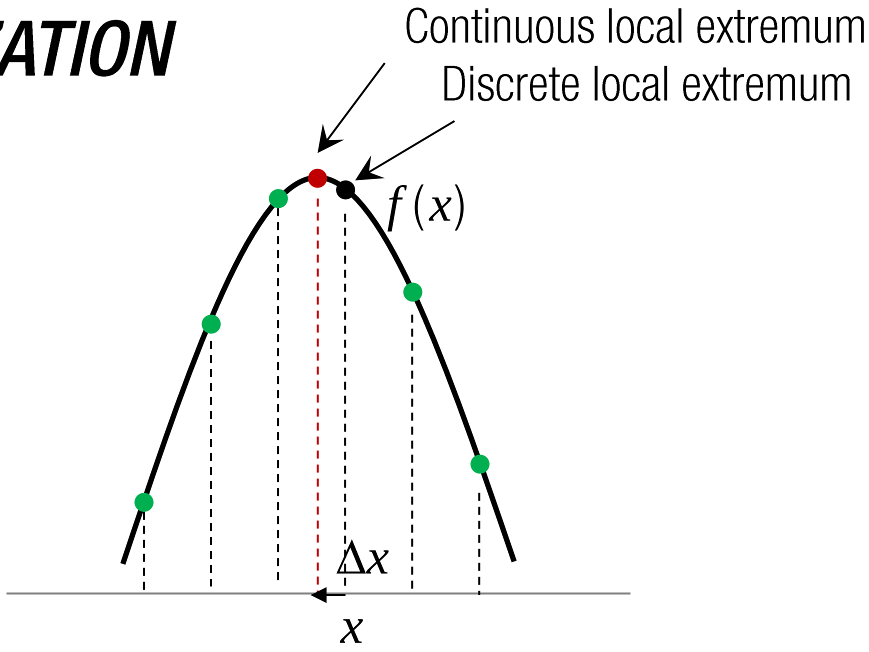
Subpixel Keypoint Localization





Subpixel Keypoint Localization



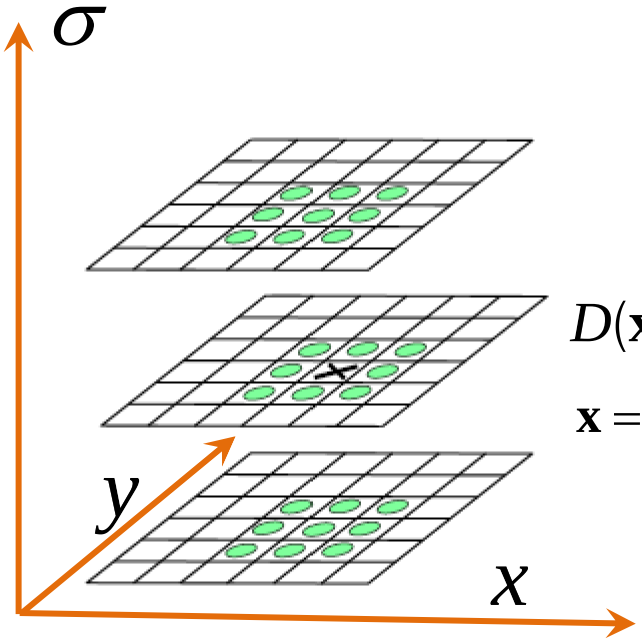


Newton's method

$$f(x+\Delta x) = f(x) + f'\Delta x + \frac{1}{2} f''\Delta x^2$$

$$\Delta x = -\frac{f'}{f''}$$

SUBPIXEL KEYPOINT LOCALIZATION



$D(\mathbf{x})$

$\mathbf{x} = (x, y, \sigma)$

Newton's method

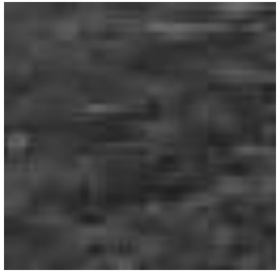
$$f(x+\Delta x) = f(x) + f'\Delta x + \frac{1}{2} f''\Delta x^2$$

$$\Delta x = -\frac{f'}{f''}$$

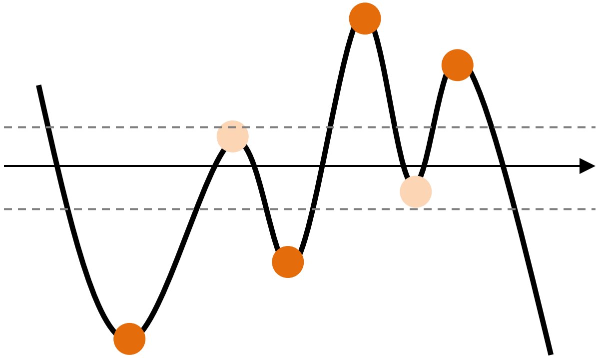
$$D(\mathbf{x}+\Delta\mathbf{x}) = D(\mathbf{x}) + \underbrace{\frac{\partial D}{\partial \mathbf{x}}}_{\text{Gradient}}\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathsf{T}} \underbrace{\frac{\partial^2 D}{\partial \mathbf{x}^2}}_{\text{Hessian}}\Delta\mathbf{x}$$

$$\Delta\mathbf{x} = -\left(\frac{\partial^2 D}{\partial \mathbf{x}^2}\right)^{-1} \frac{\partial D}{\partial \mathbf{x}}$$

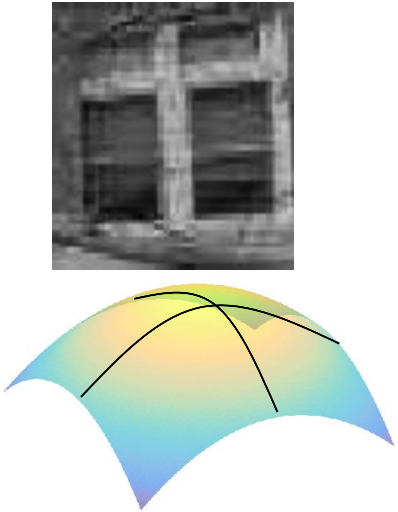
Thresholding: Low Contrast



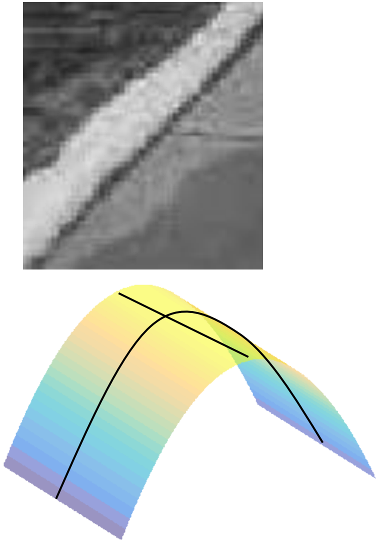
$$|D(\mathbf{x})| < 0.03$$



THRESHOLDING: EDGE



$\lambda_1 \approx \lambda_2$

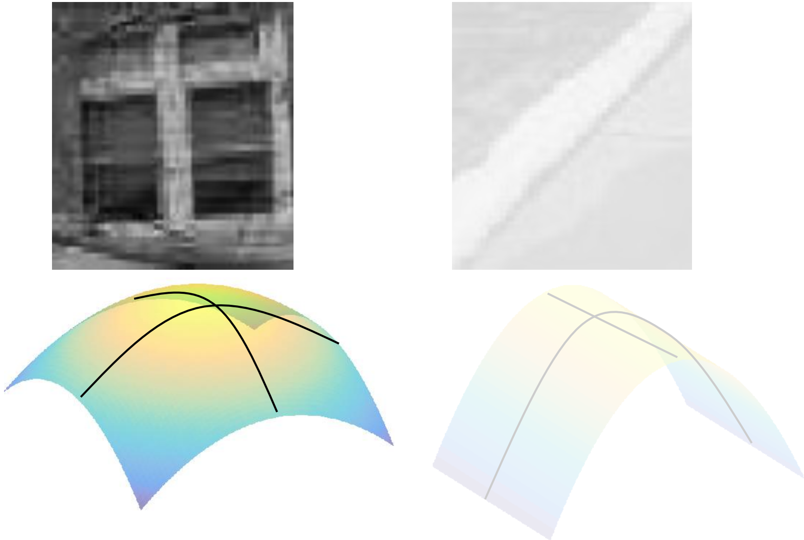


$\lambda_1 \gg \lambda_2$

Principal curvatures are eigen values of Hessian matrix:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix}$$

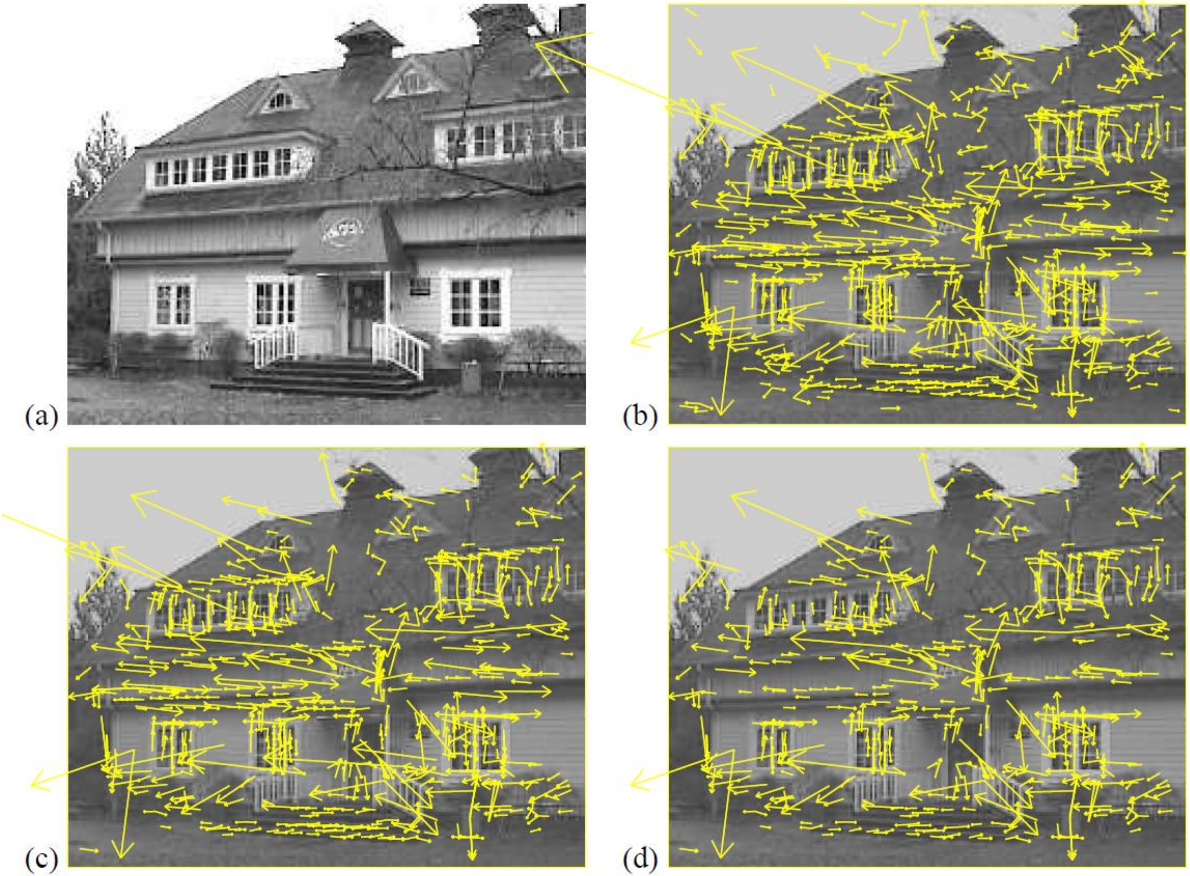
THRESHOLDING: EDGE



$\lambda_1 \approx \lambda_2$ $\lambda_1 \gg \lambda_2$

Principal curvatures are eigen values of Hessian matrix:

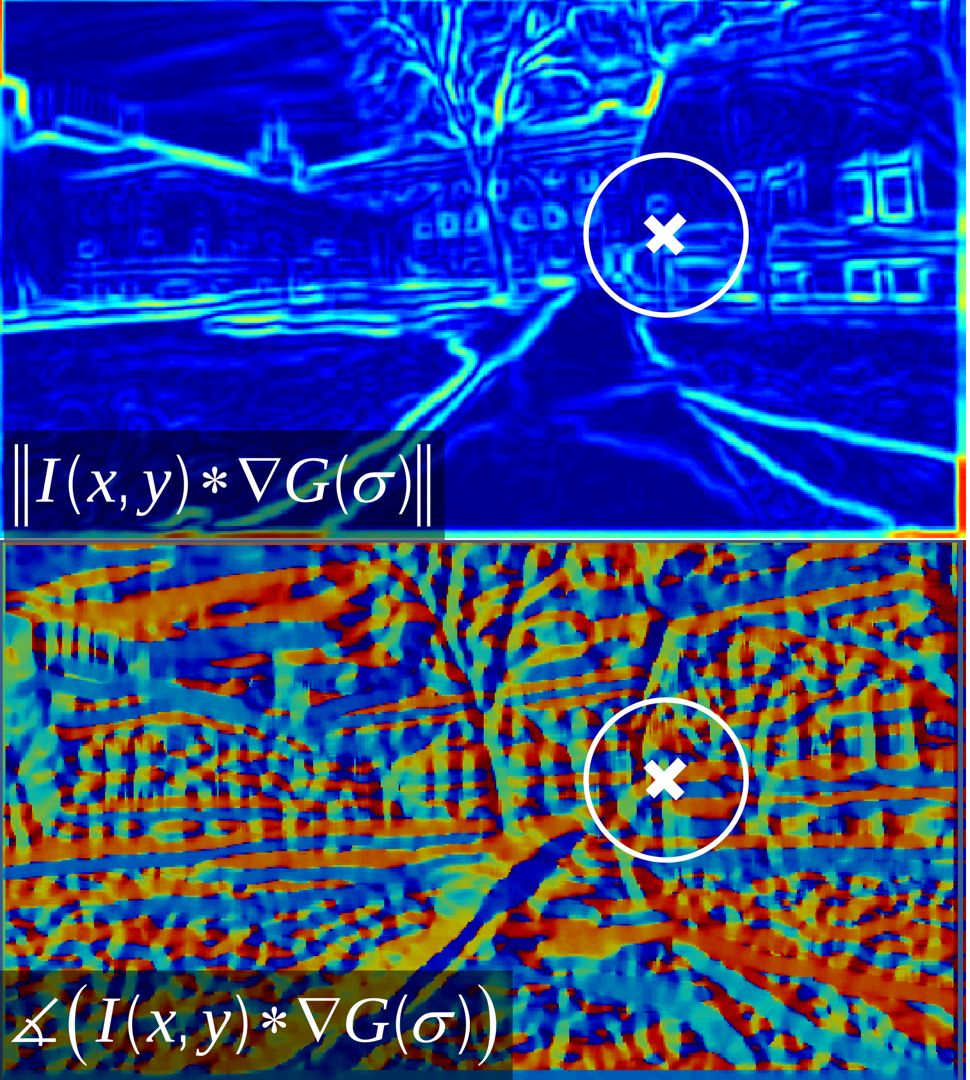
$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix}$$



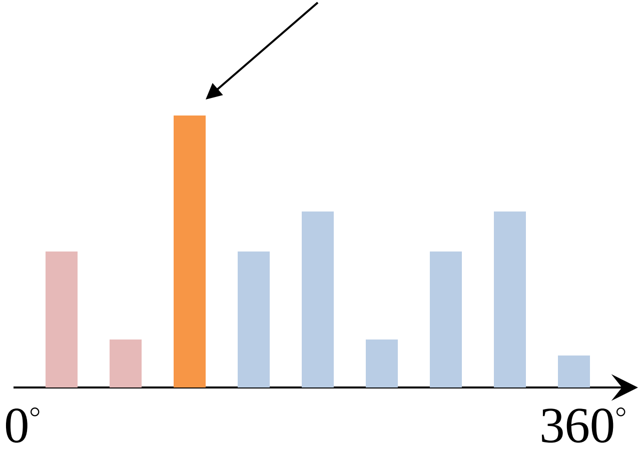
(a) (b) (c) (d)

Figure 5: This figure shows the stages of keypoint selection. (a) The 233x189 pixel original image. (b) The initial 832 keypoints locations at maxima and minima of the difference-of-Gaussian function. Keypoints are displayed as vectors indicating scale, orientation, and location. (c) After applying a threshold on minimum contrast, 729 keypoints remain. (d) The final 536 keypoints that remain following an additional threshold on ratio of principal curvatures.

Orientation Assignment

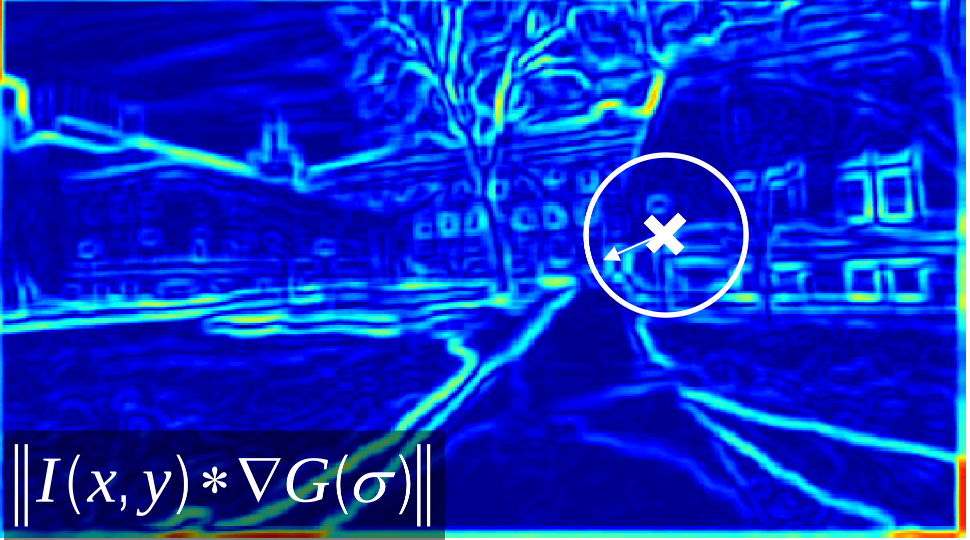


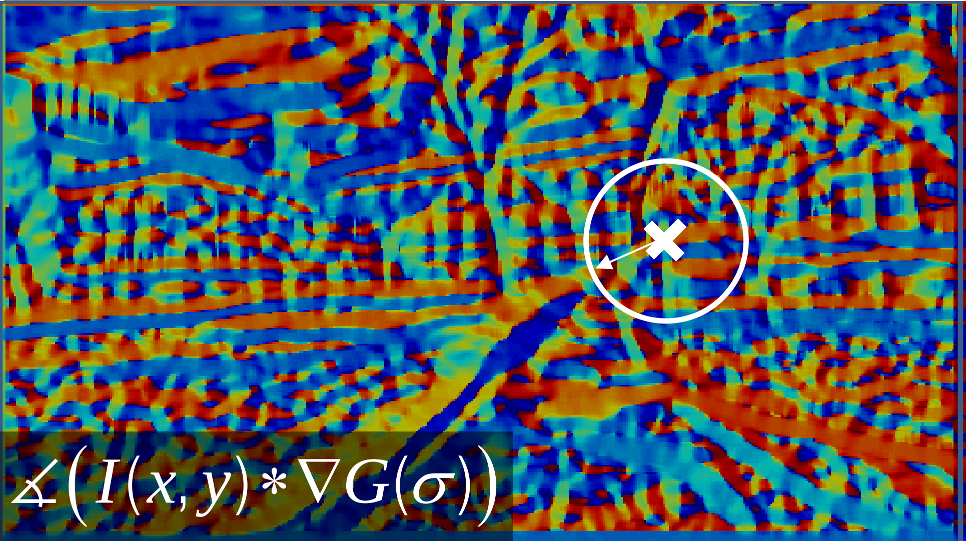
Peak orientation ~ dominant orientation

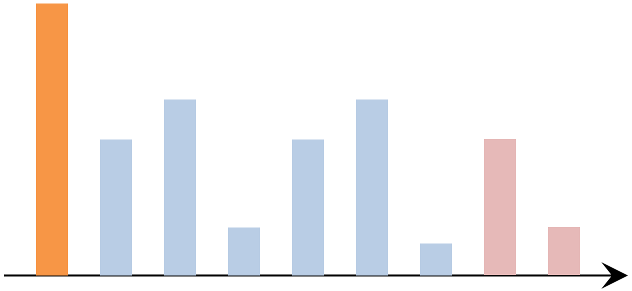


Histogram of orientation
weighted by gradient magnitude and Gaussian

Orientation Normalization

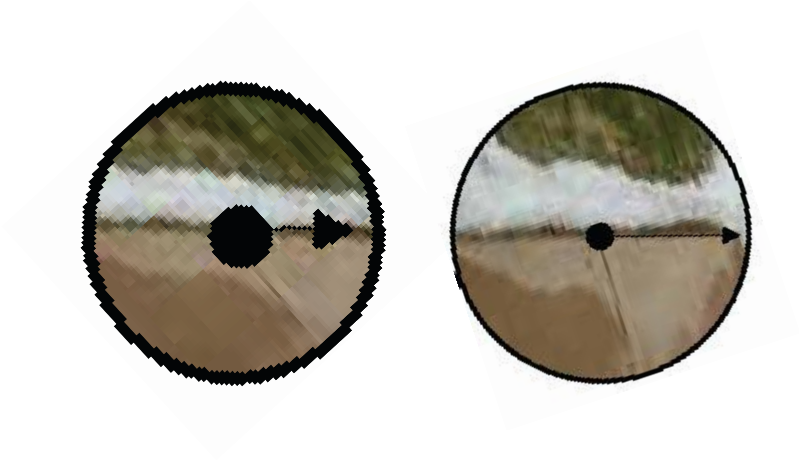


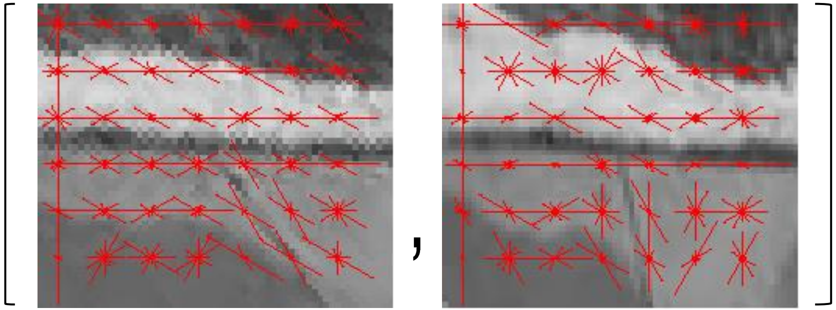




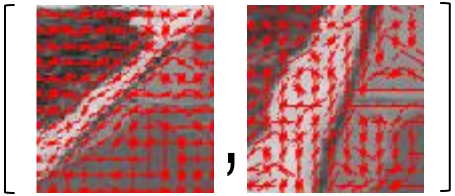
Histogram of orientation
weighted by gradient magnitude and Gaussian

Scale-invariant Feature Transform (SIFT)



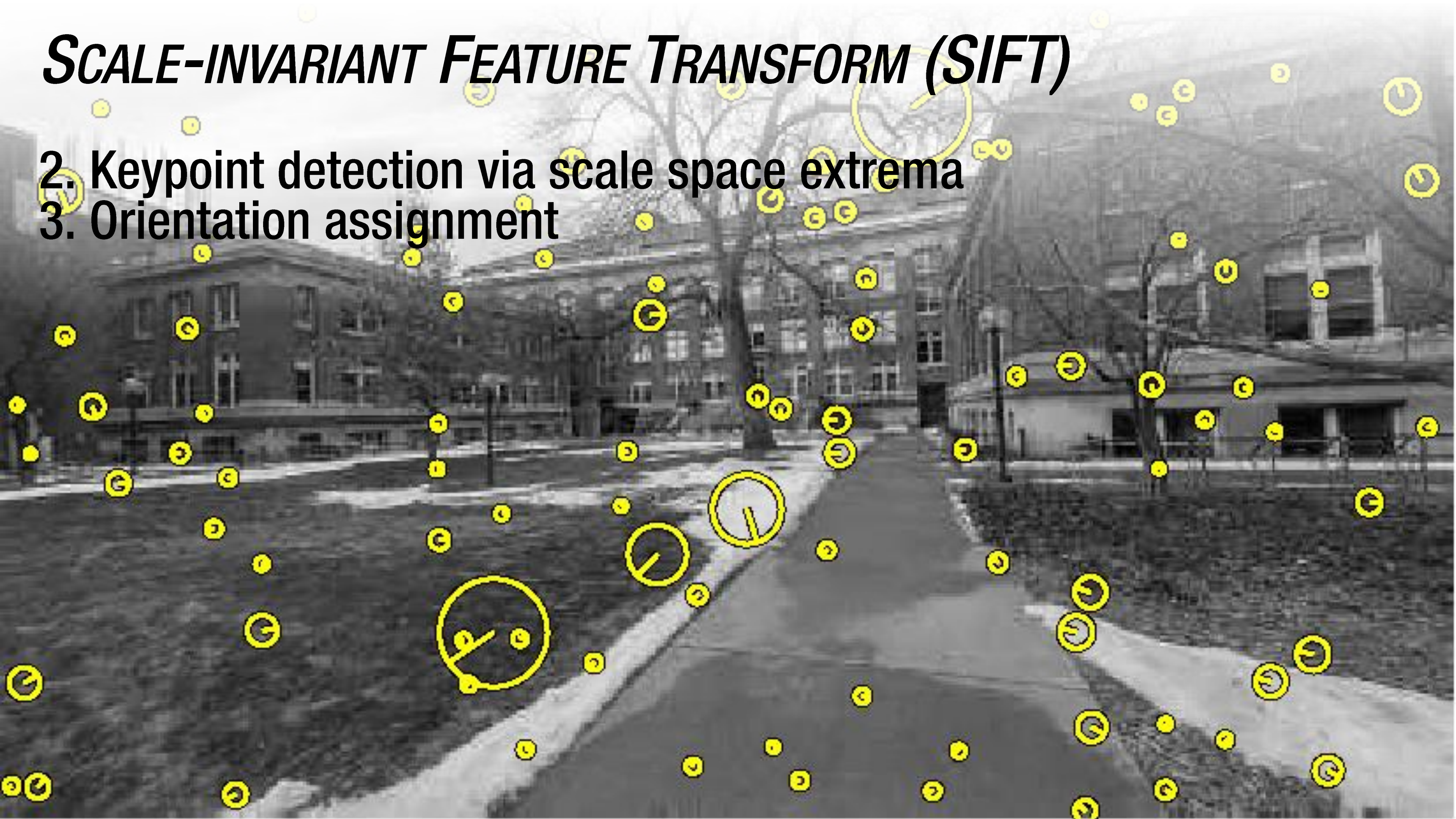
corr  = 0.91

cf)

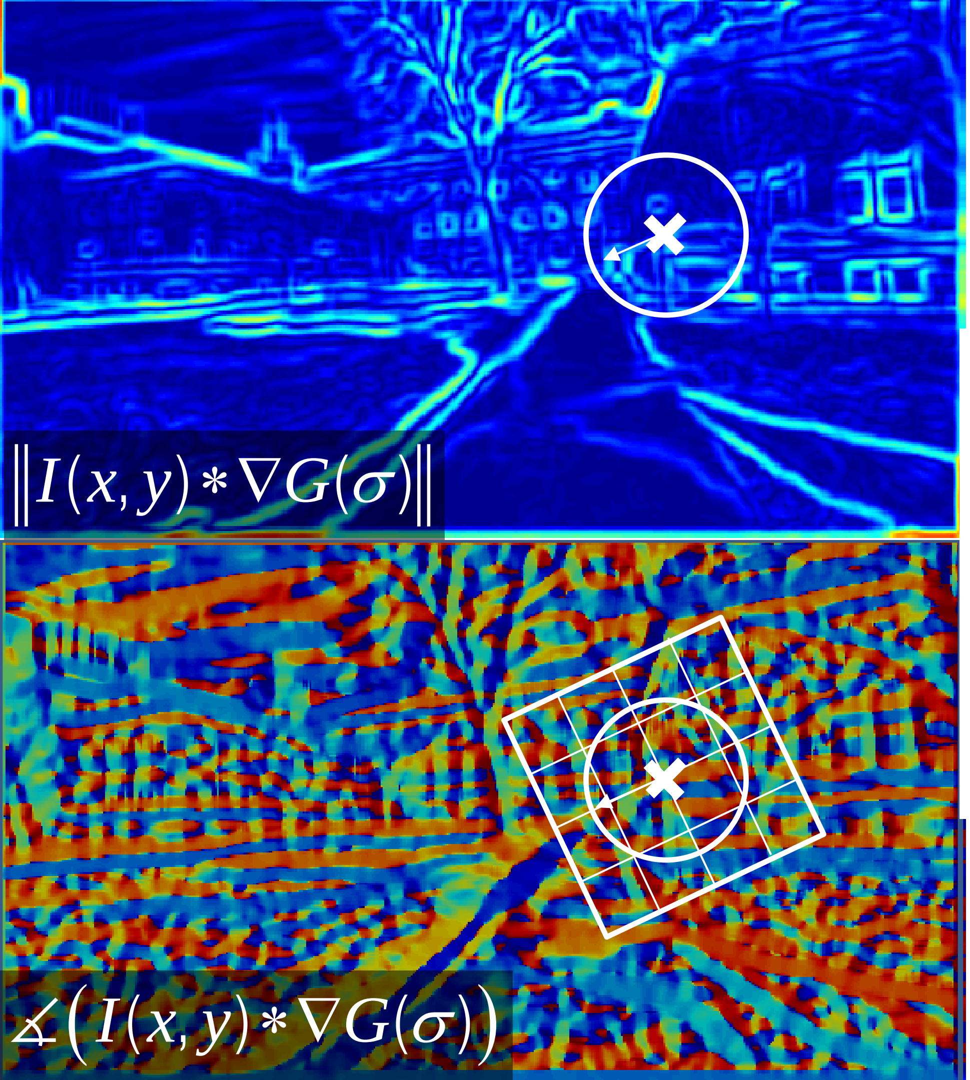
corr  = 0.15

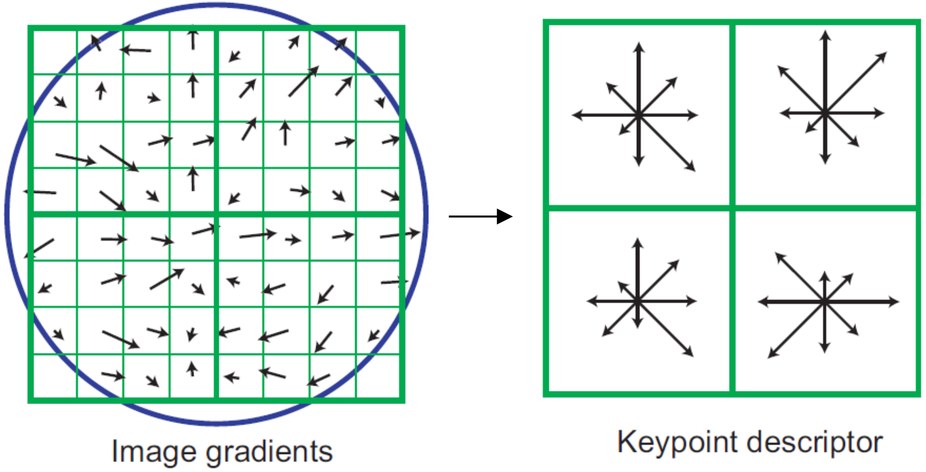
SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

2. Keypoint detection via scale space extrema
3. Orientation assignment



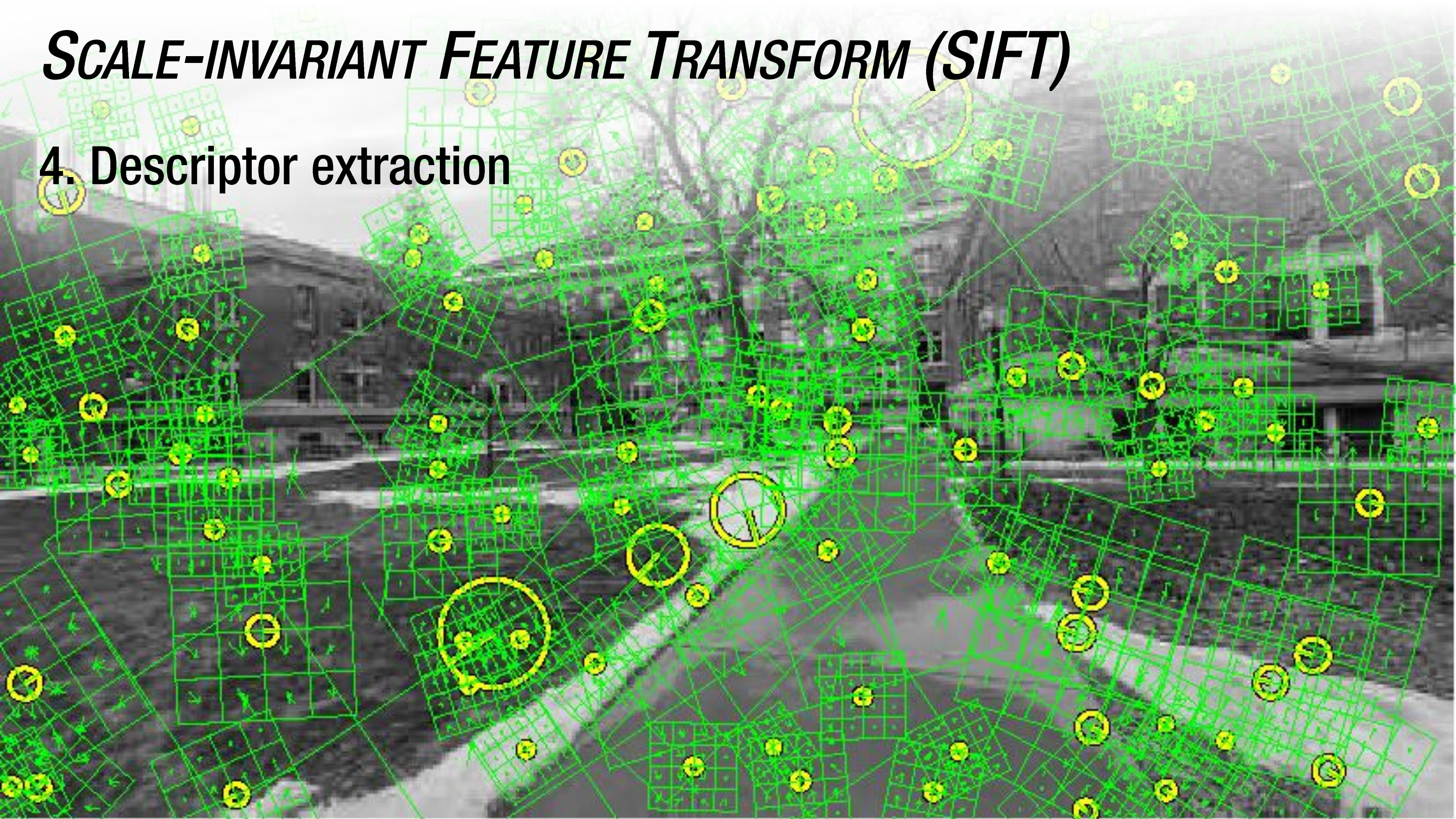
Descriptor Computation





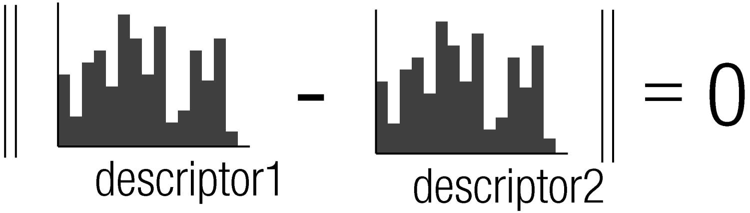
SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

4. Descriptor extraction




Local visual descriptor


Local visual descriptor


-
= 0
descriptor1
descriptor2