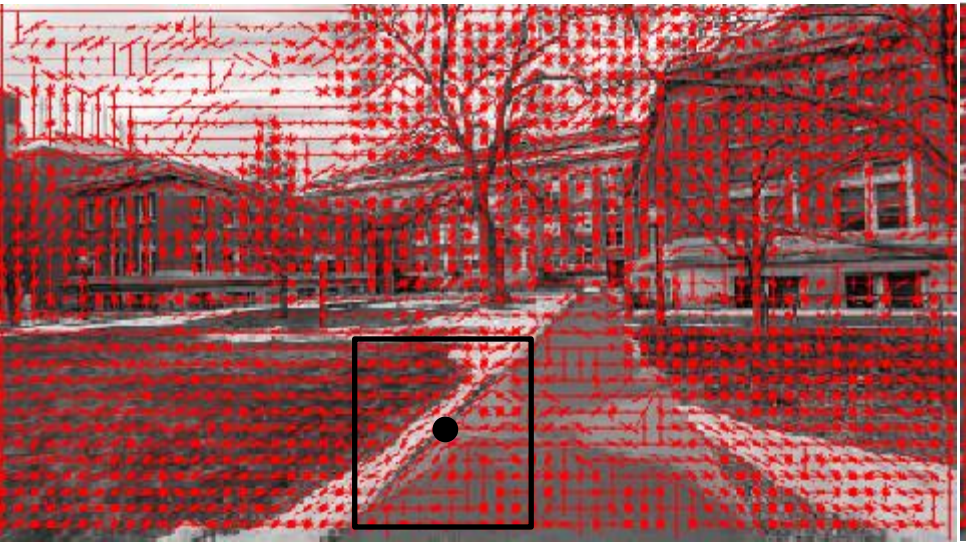
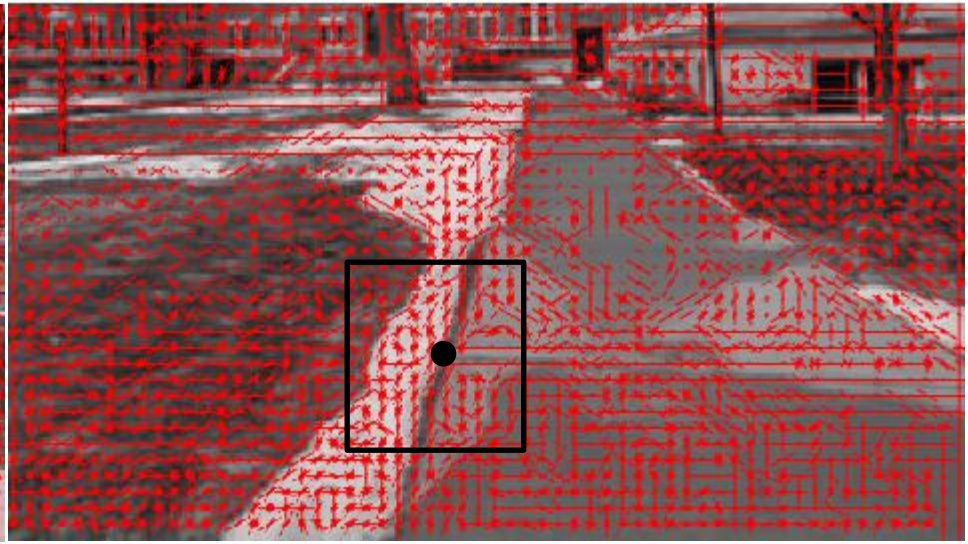
Scale-invariant Feature Transform

Hyun Soo Park

Feature Matching

RECALL: HOG





$$corr\left[\ \ ,\ \ \right] = 0.15$$

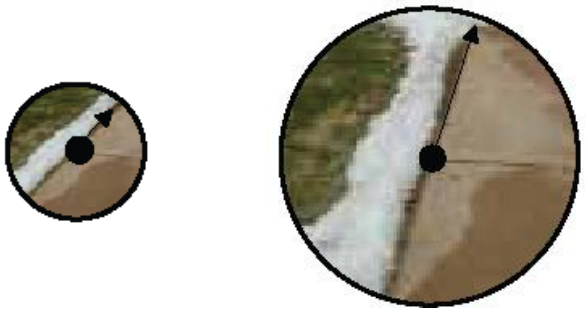
What's wrong?

Challenges of Feature Matching

- Scale
- Orientation

Feature Normalization

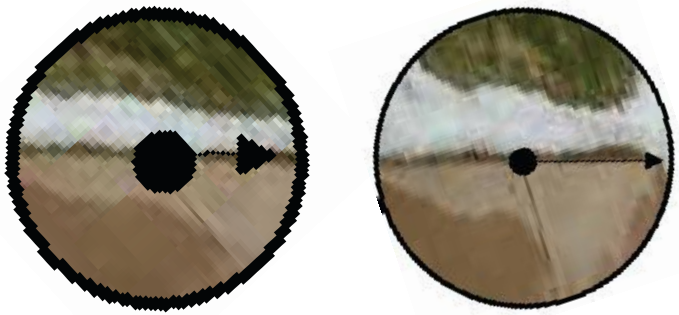
Feature Normalization



Feature Normalization

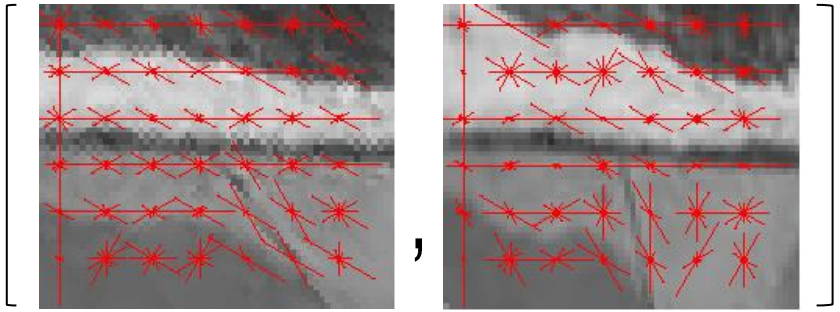
Scale normalization

Feature Normalization



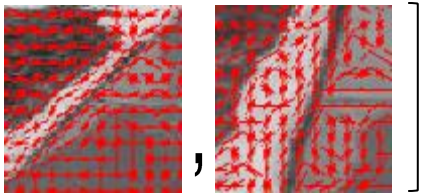
Orientation normalization

SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

corr 

= 0.91

cf)

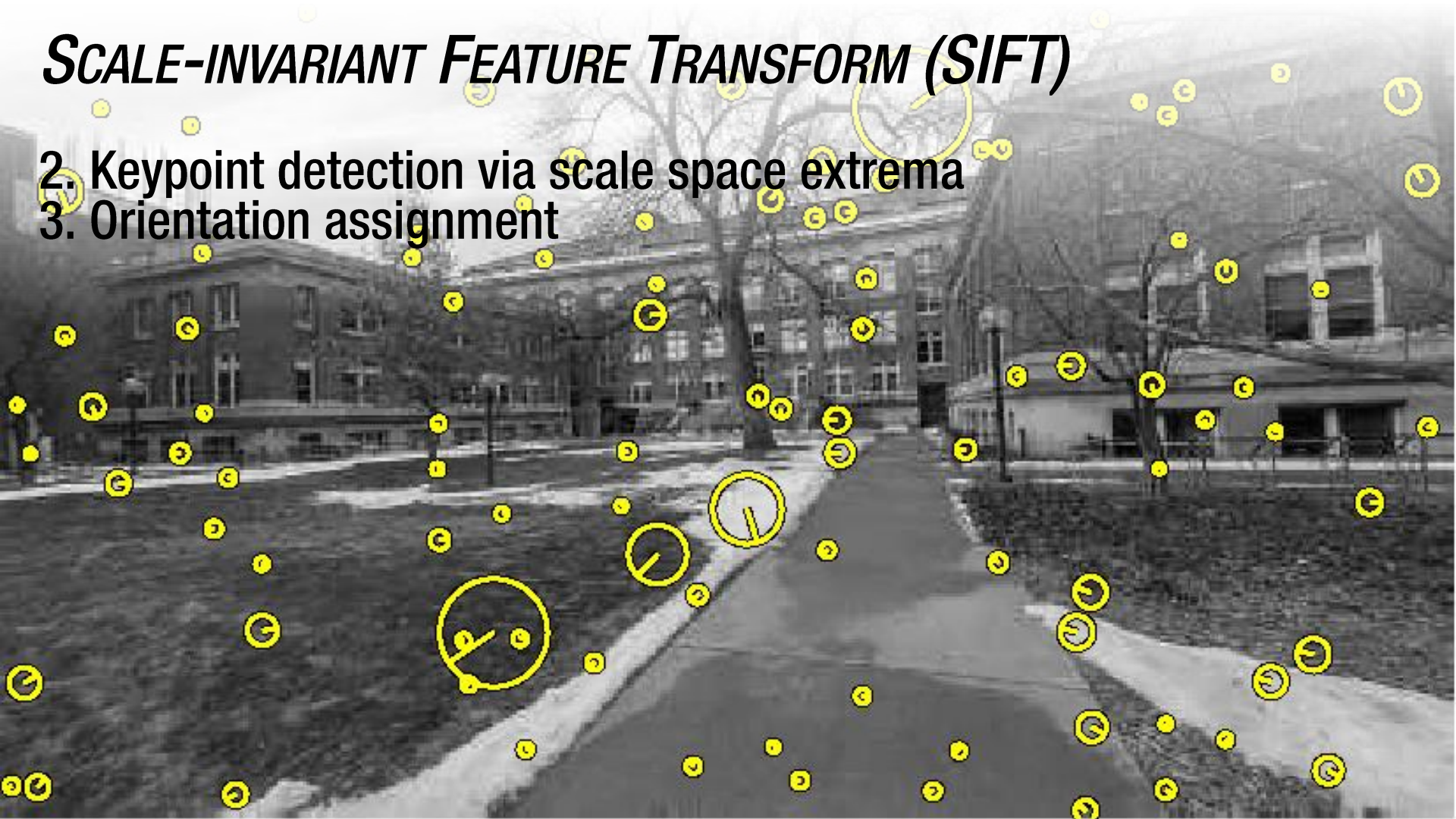
corr  = 0.15

Scale-invariant Feature Transform (SIFT)

1. Input image

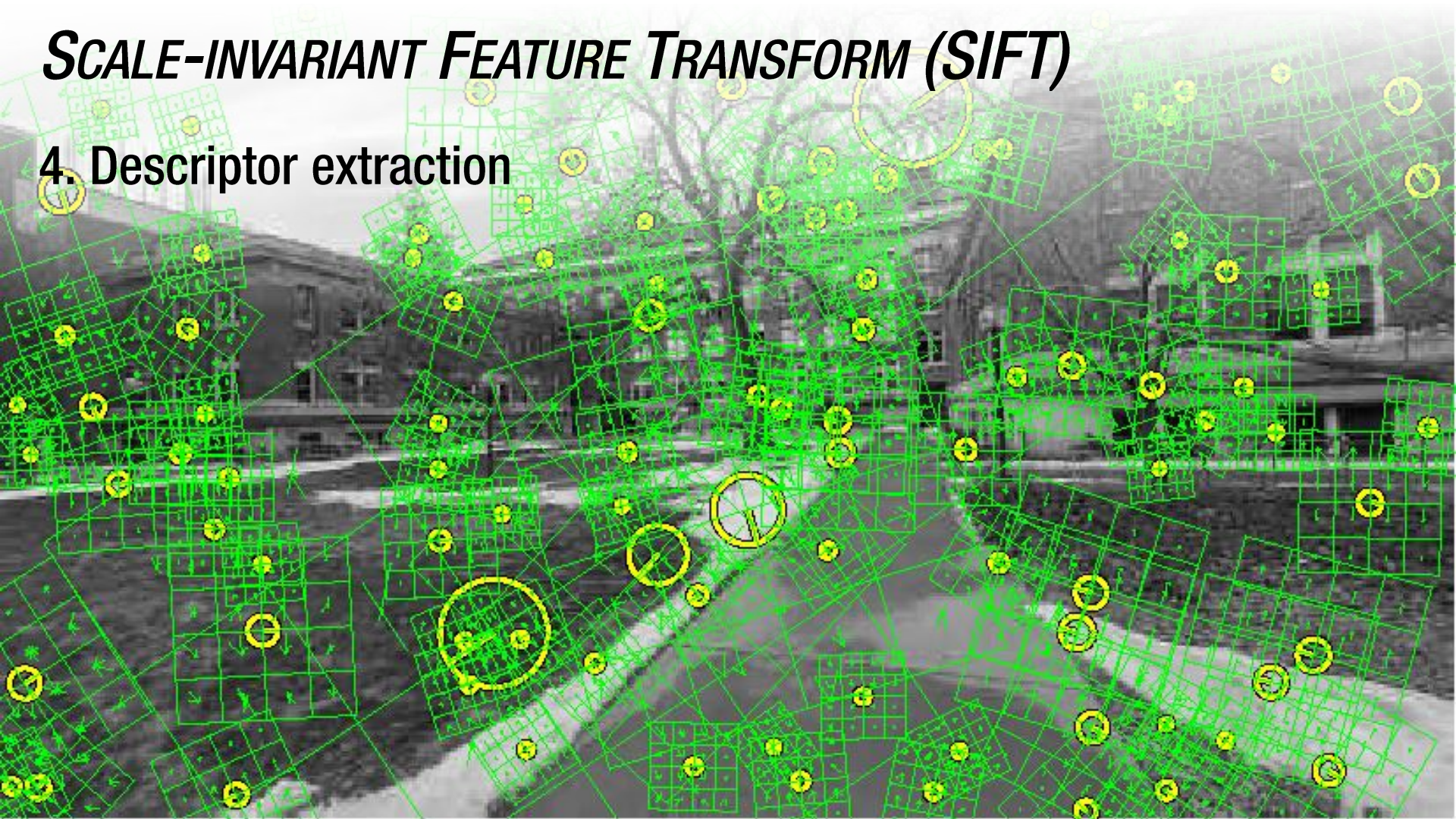
Scale-invariant Feature Transform (SIFT)

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

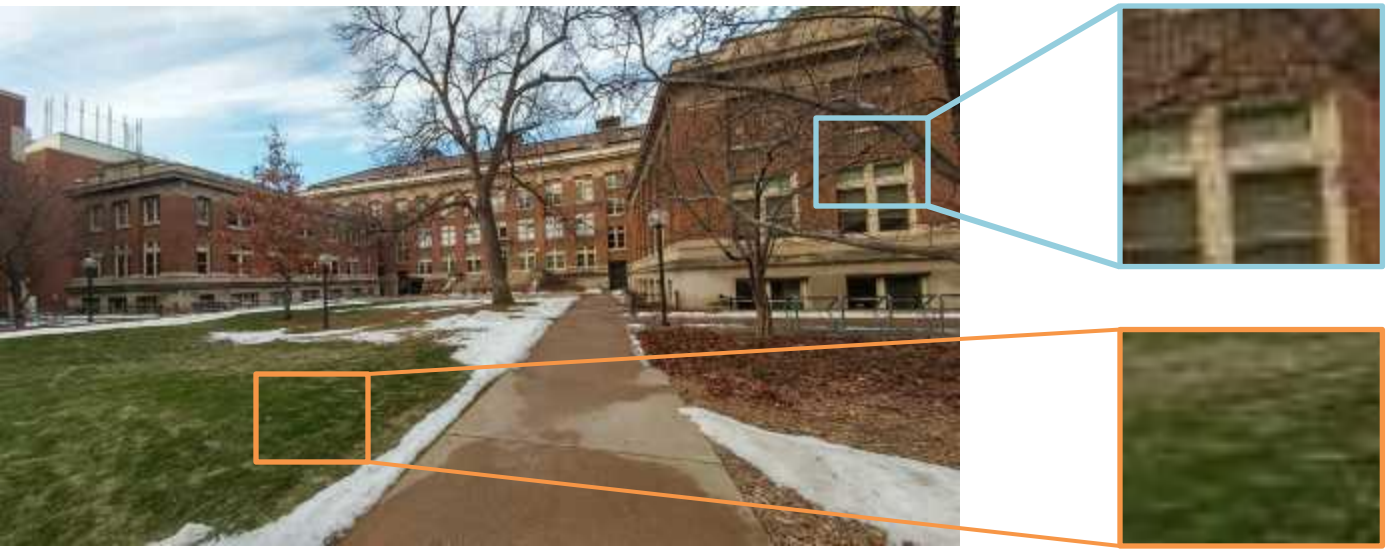


SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

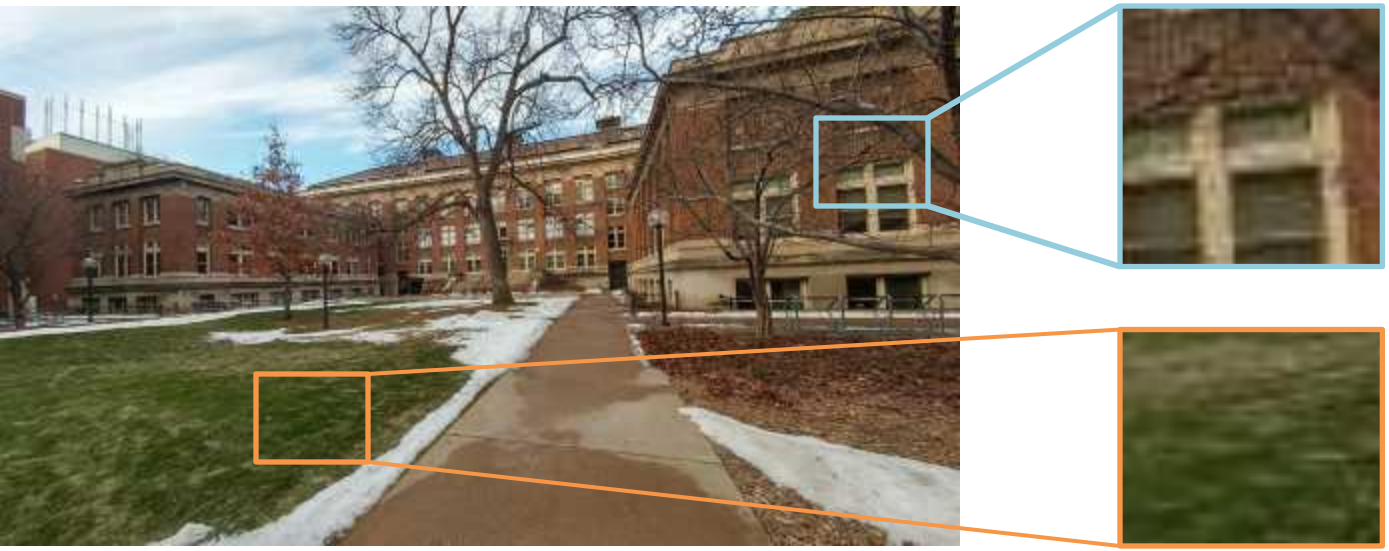
4. Descriptor extraction



What Characterizes Keypoint?



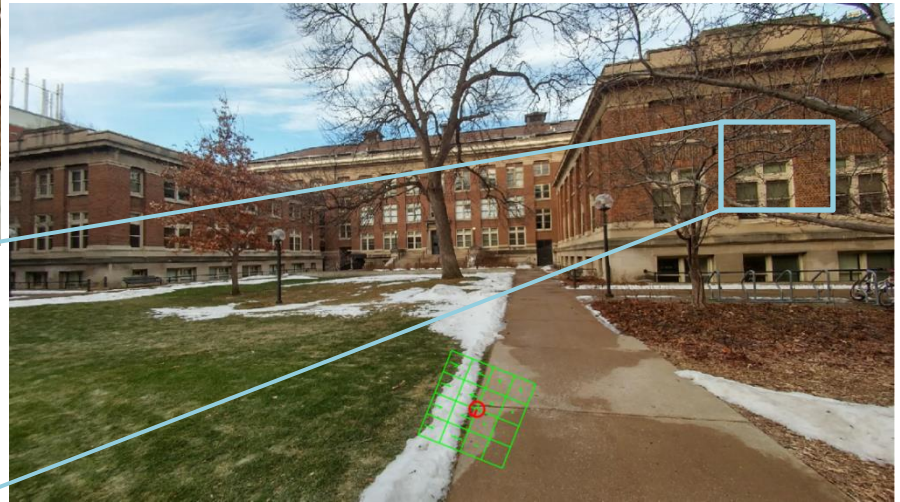
What Characterizes Keypoint?



- Distinctiveness

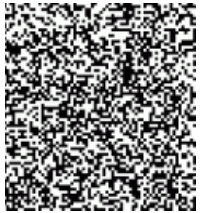
What Characterizes Keypoint?

- Distinctiveness
- Repeatability



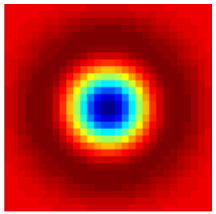
What Characterizes Keypoint?

- Distinctiveness
- Repeatability



KEYPOINT DETECTION: SCALE SPACE EXTREMA

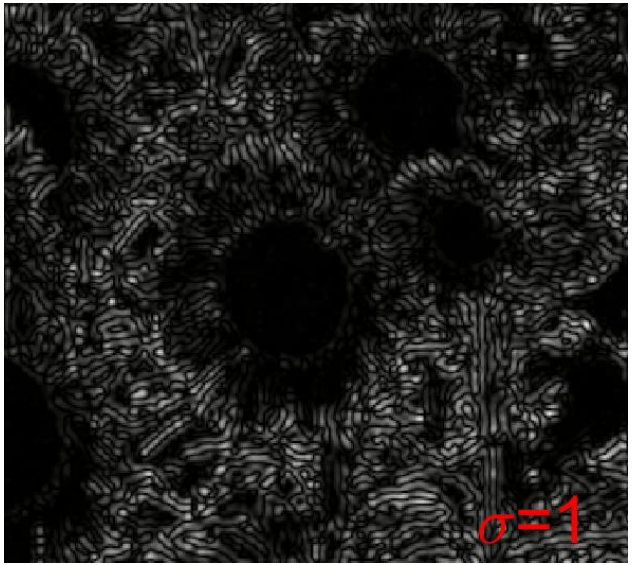
*



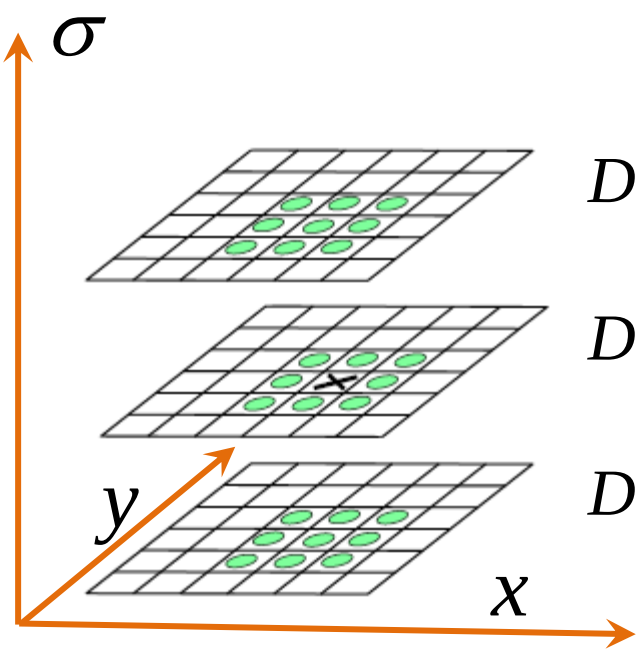
$$L = \sigma^2 \nabla^2 G$$

Laplacian of Gaussian (LoG)

=



What Characterizes Keypoint?



$D(x, y, \sigma + \Delta\sigma)$

$D(x, y, \sigma)$

$D(x, y, \sigma - \Delta\sigma)$

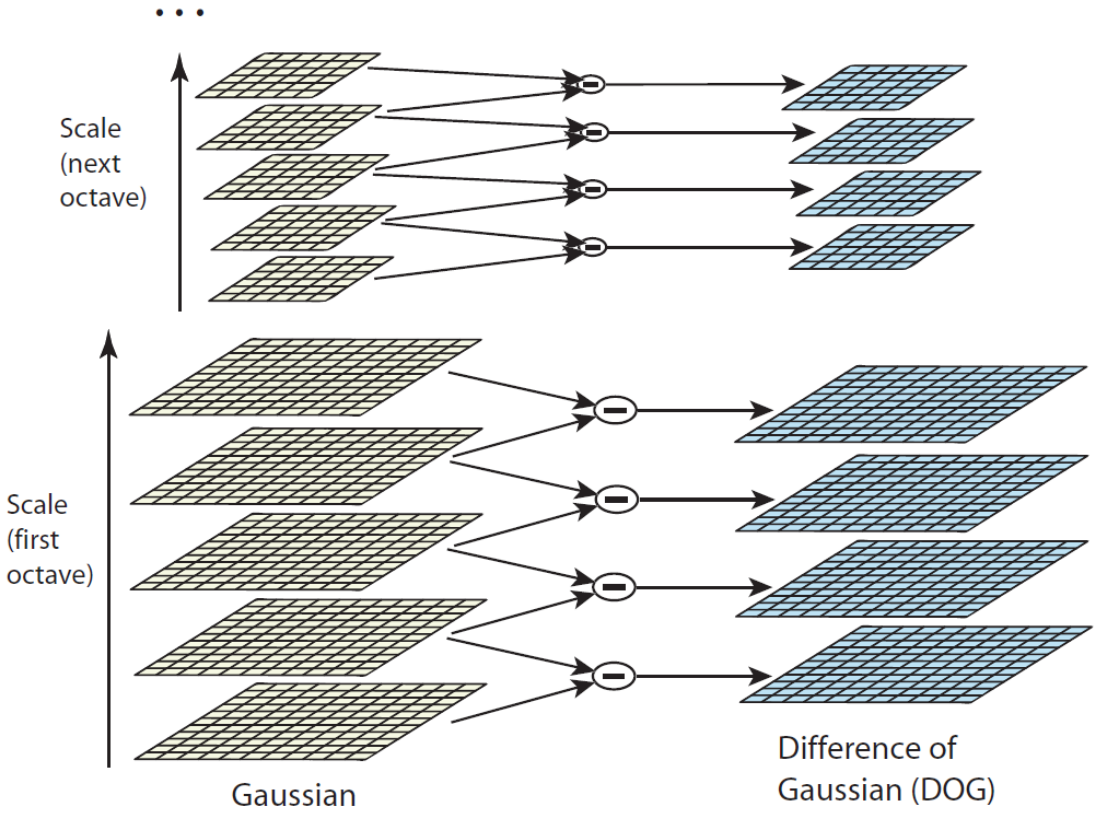
(x, y) is keypoint if

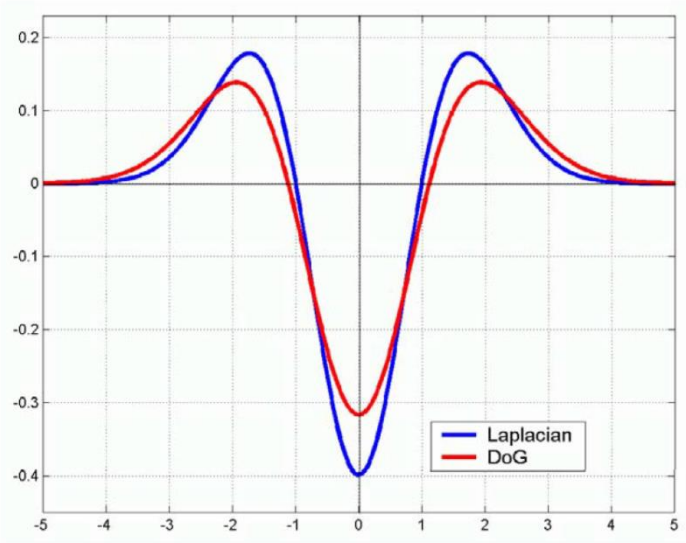
$$D(x, y, \sigma) > D(x \pm \Delta x, y \pm \Delta y, \sigma \pm \Delta\sigma)$$

or

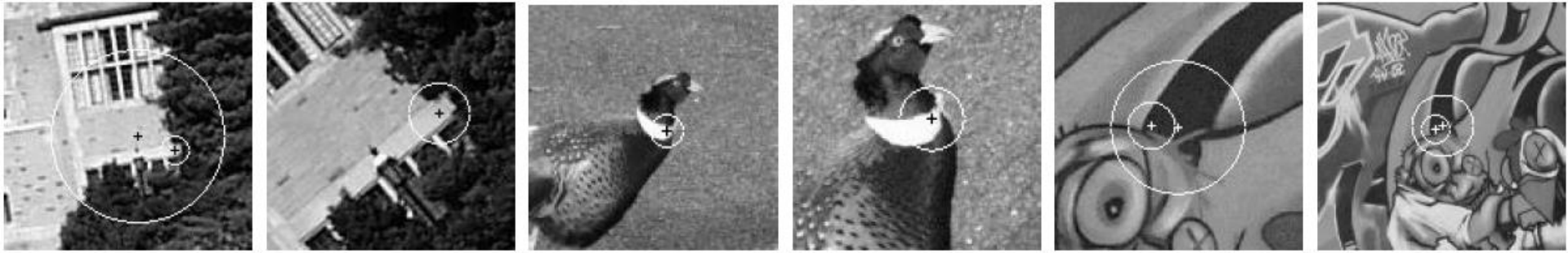
$$D(x, y, \sigma) < D(x \pm \Delta x, y \pm \Delta y, \sigma \pm \Delta\sigma)$$

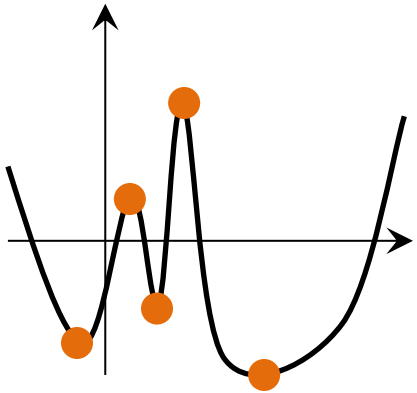
LAPLACIAN ~ DIFFERENCE OF GAUSSIAN





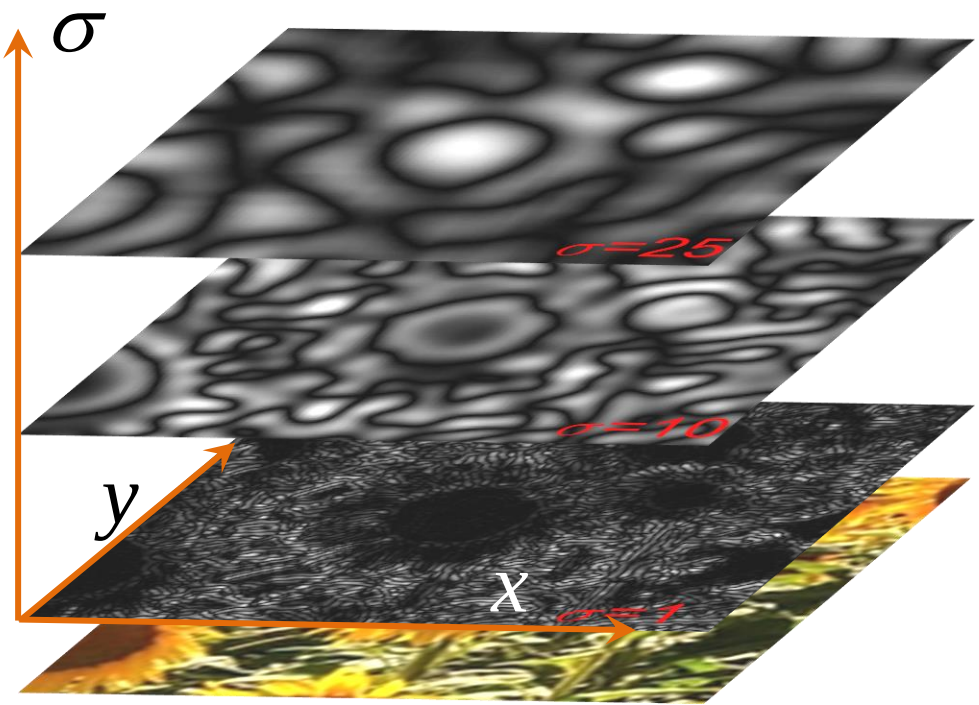
What Characterizes Keypoint?





Local extrema (minima/maxma)

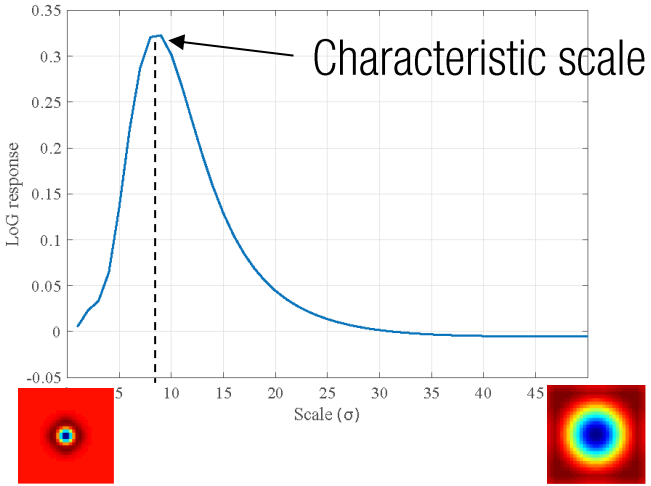
What Characterizes Keypoint?



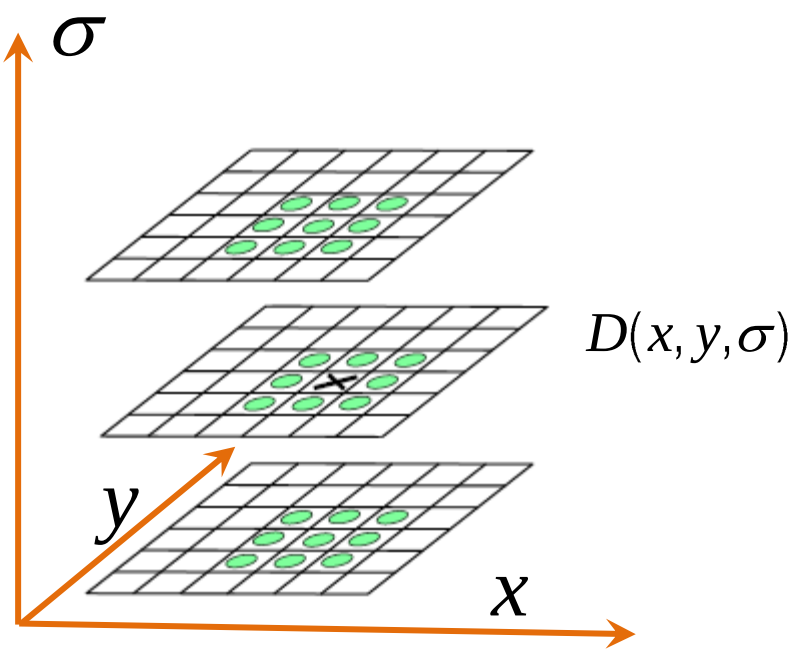
$$D(x, y, \sigma) = I(x, y) * L(\sigma)$$

Scale space response

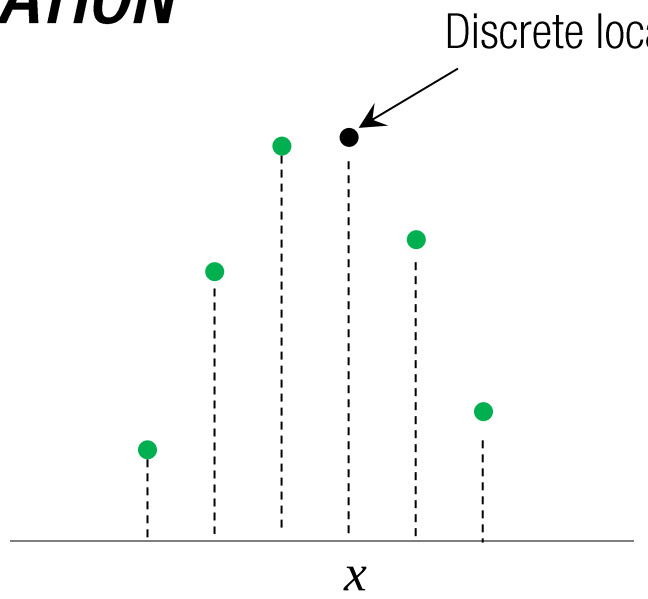
Local extrema (minima/maxma)



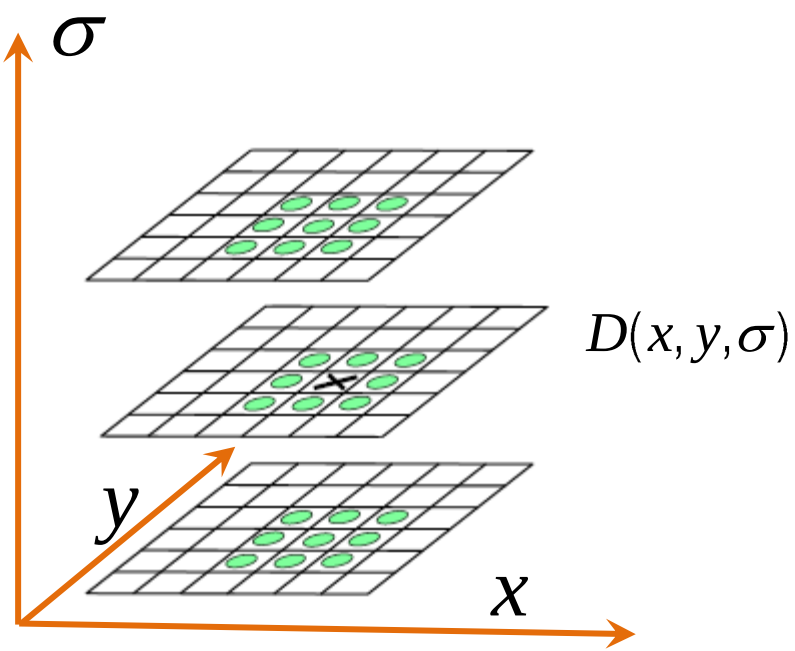
SUBPIXEL KEYPOINT LOCALIZATION



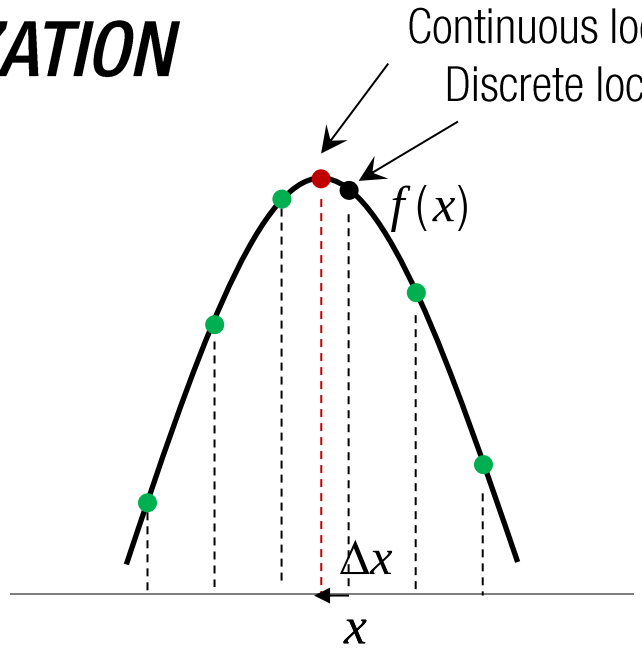
$D(x, y, \sigma)$



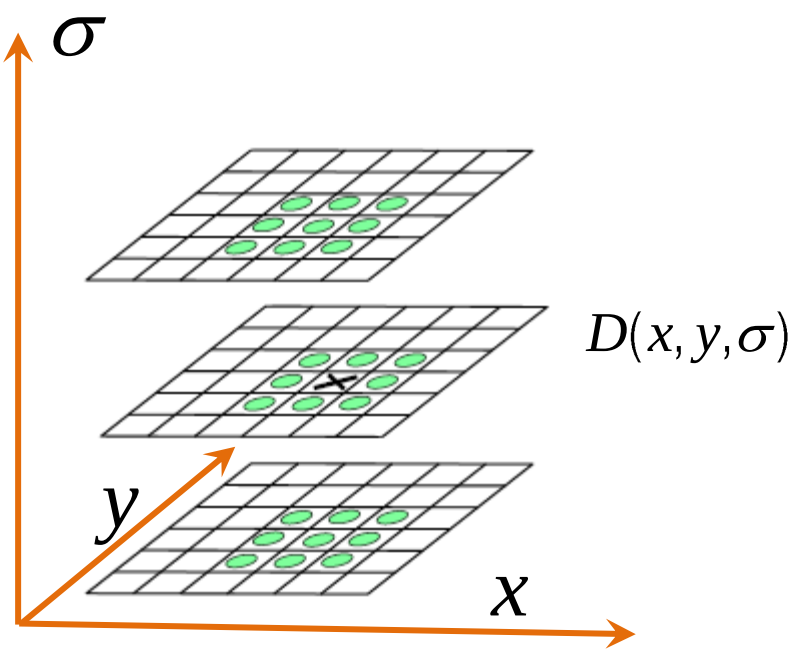
SUBPIXEL KEYPOINT LOCALIZATION



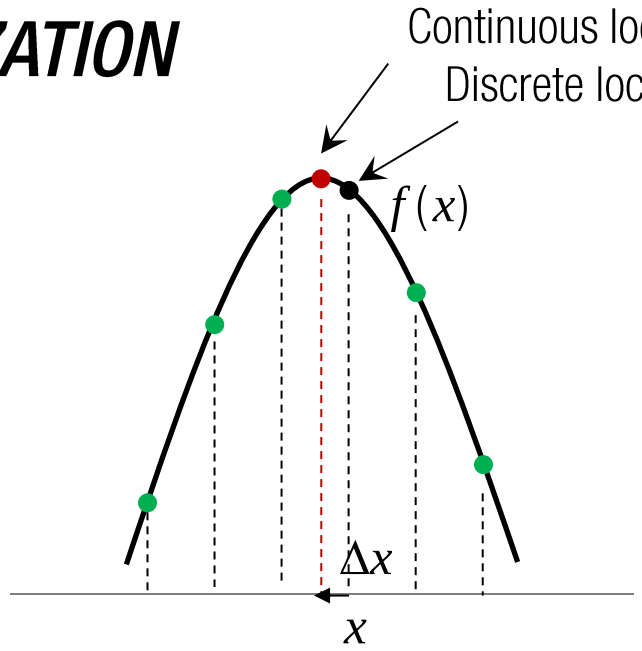
$D(x, y, \sigma)$



Subpixel Keypoint Localization



$D(x, y, \sigma)$

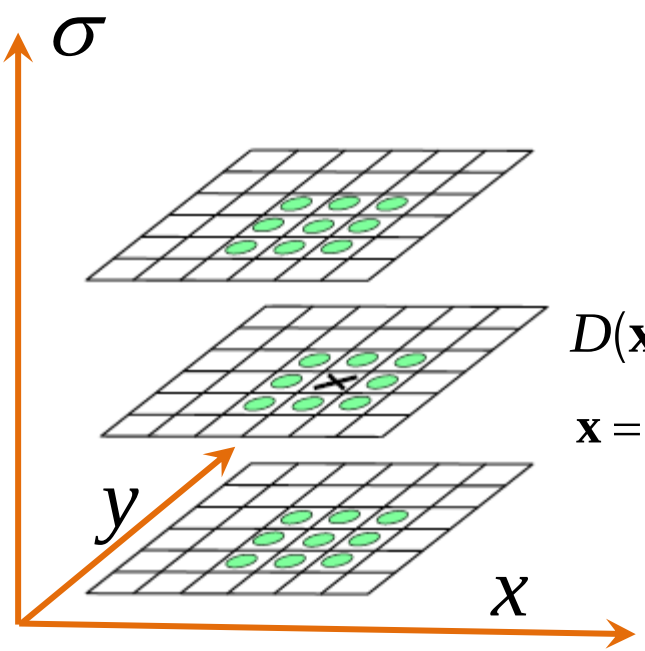


Newton's method

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

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

SUBPIXEL KEYPOINT LOCALIZATION



$D(\mathbf{x})$

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

Newton's method

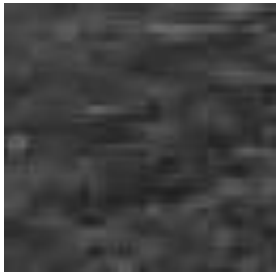
$$f(x + \Delta x) = f(x) + f'\Delta x + \frac{1}{2} f''\Delta x^2$$

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

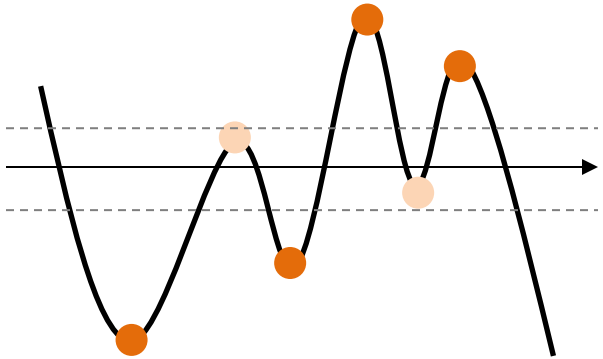
$$D(\mathbf{x} + \Delta\mathbf{x}) = D(\mathbf{x}) + \underbrace{\frac{\partial D}{\partial \mathbf{x}}}_{\text{Gradient}}\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathsf{T}}\underbrace{\frac{\partial^2 D}{\partial \mathbf{x}^2}}_{\text{Hessian}}\Delta\mathbf{x}$$

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

THRESHOLDING: LOW CONTRAST

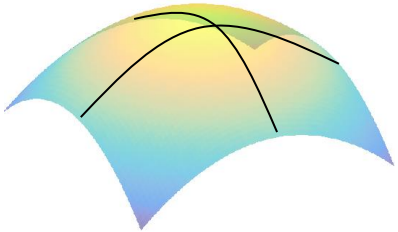


$$\left|D(\mathbf{x})\right| < 0.03$$

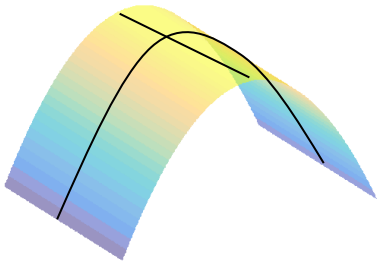


THRESHOLDING: EDGE

THRESHOLDING: EDGE



$\lambda_1 \approx \lambda_2$

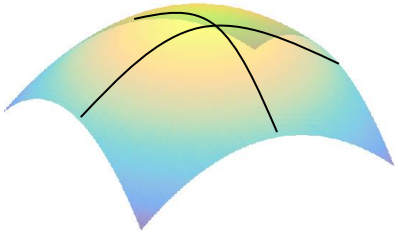


$\lambda_1 \gg \lambda_2$

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:

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

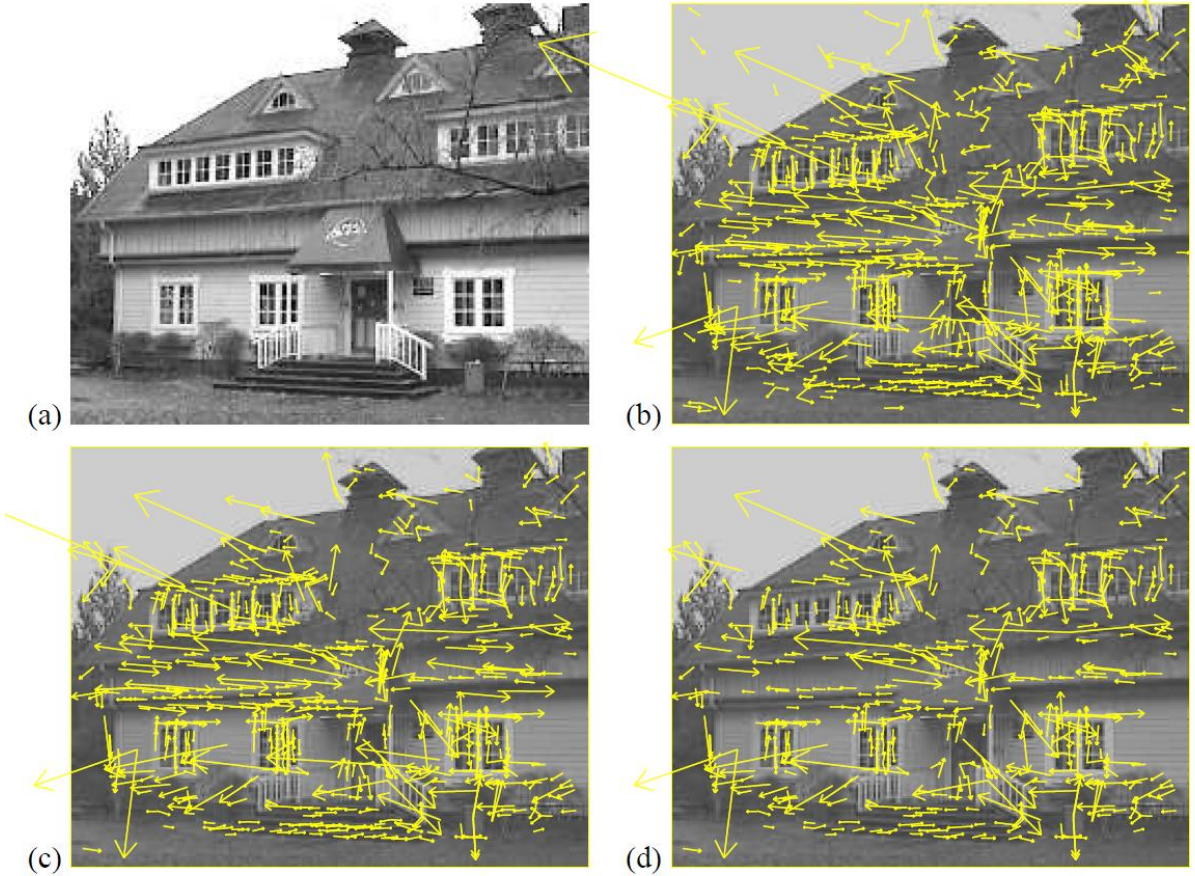
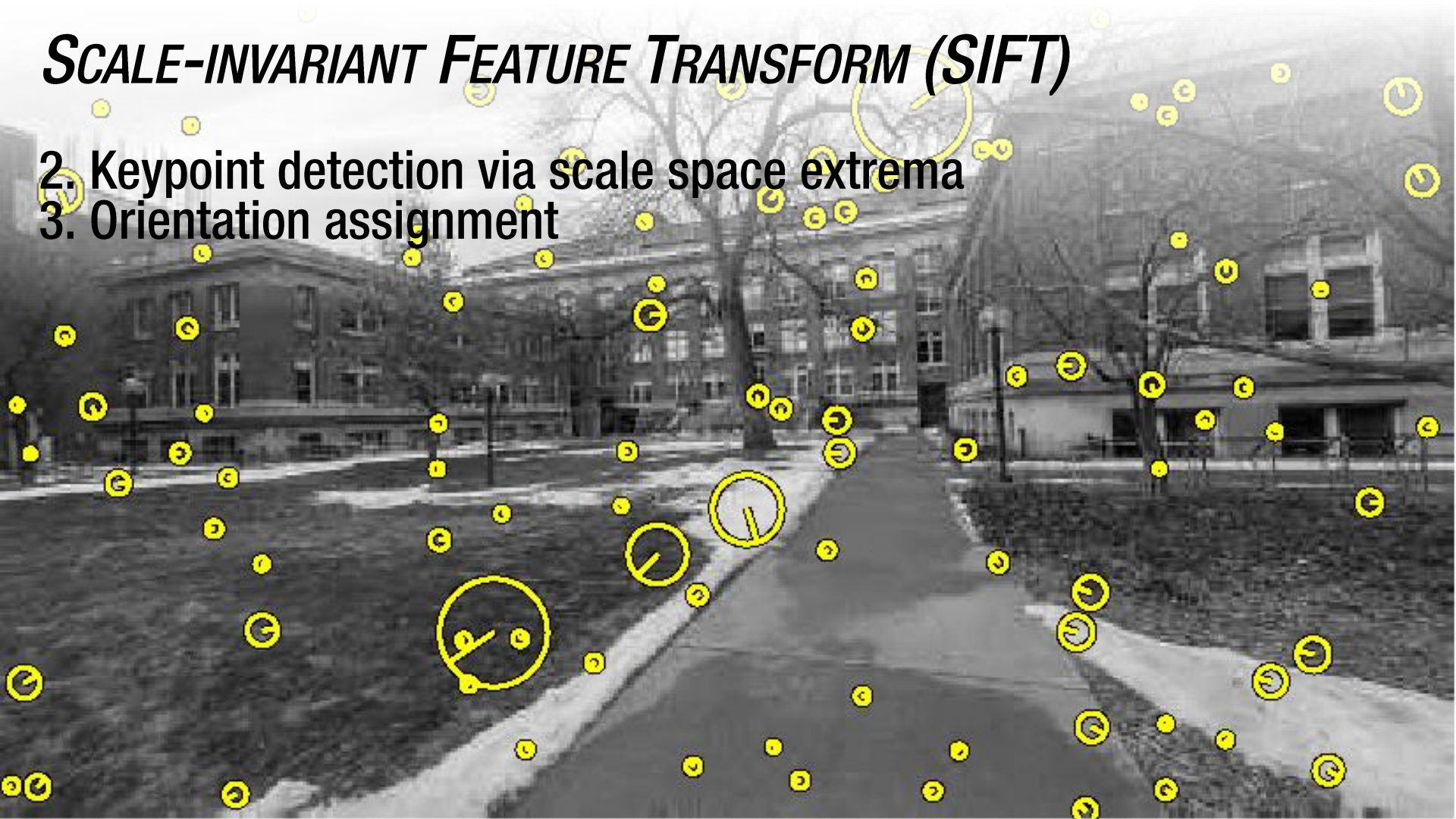


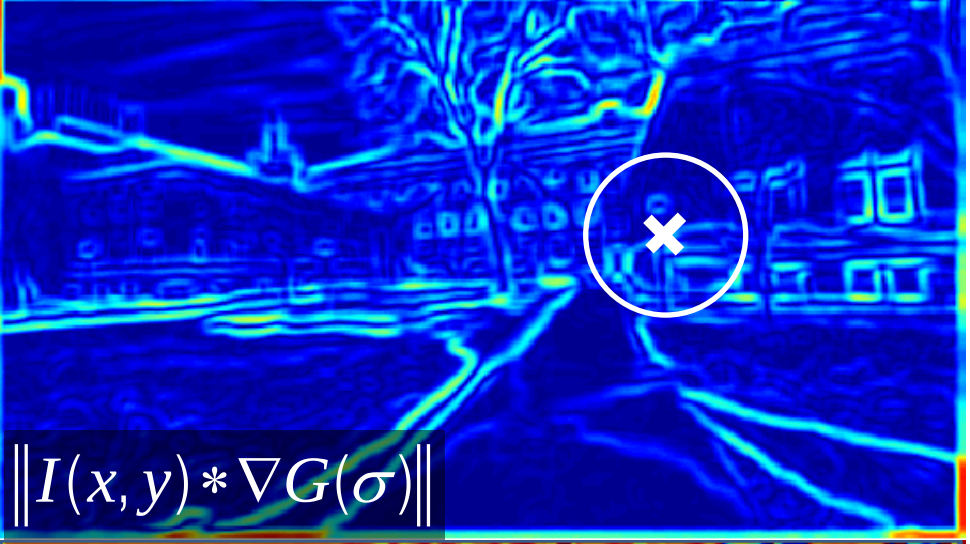
Figure 5: This figure shows the stages of keypoint selection. (a) The 233x189 pixel original image. (b) The initial 832 keypoints locations at maxima and minima of the difference-of-Gaussian function. Keypoints are displayed as vectors indicating scale, orientation, and location. (c) After applying a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

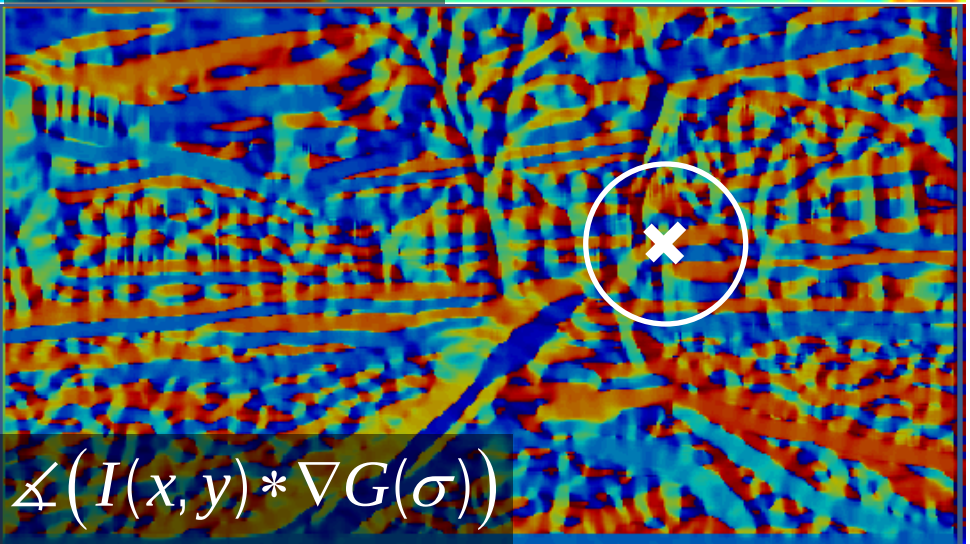
Scale-invariant Feature Transform (SIFT)

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

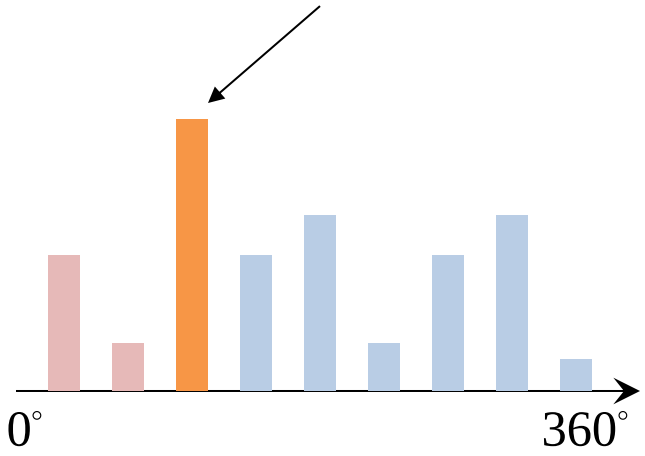


ORIENTATION ASSIGNMENT



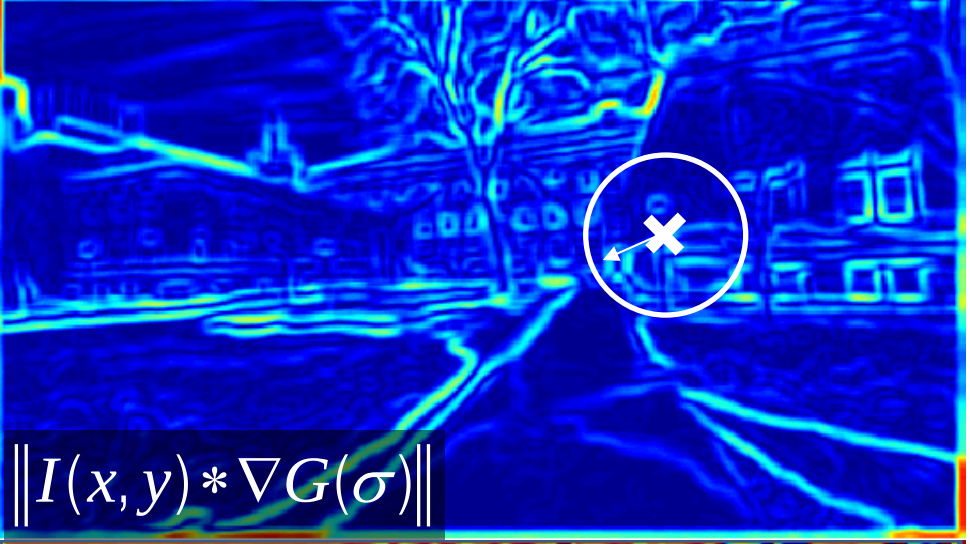


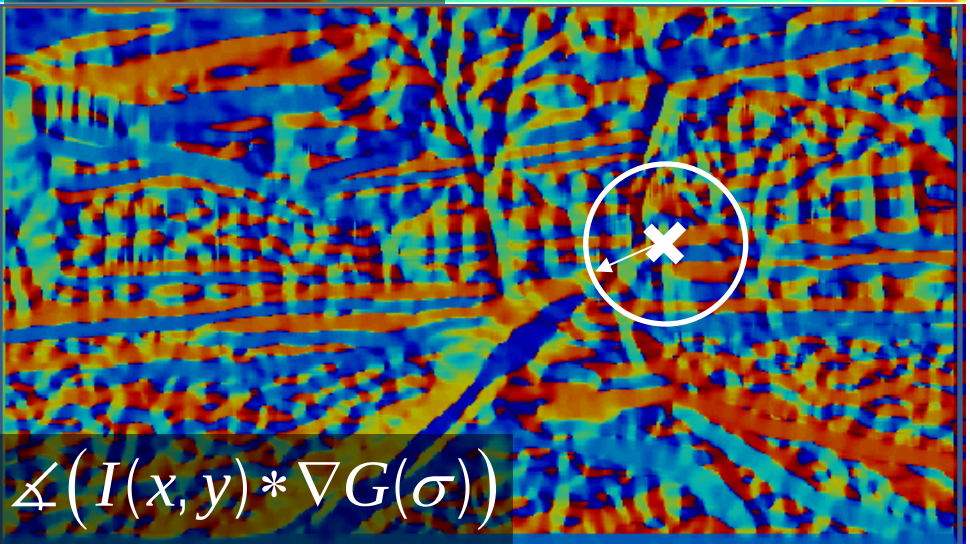
Peak orientation ~ dominant orientation

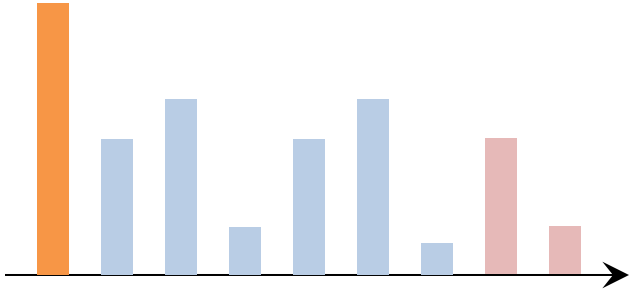


Histogram of orientation
weighted by gradient magnitude and Gaussian

Orientation Normalization

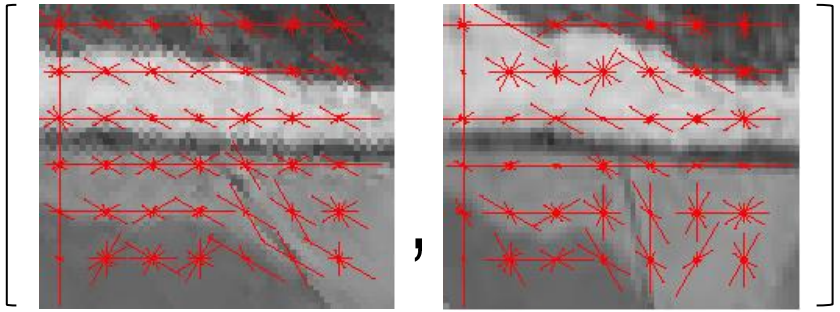






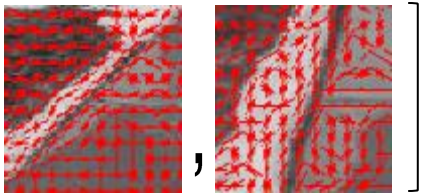
Histogram of orientation
weighted by gradient magnitude and Gaussian

SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

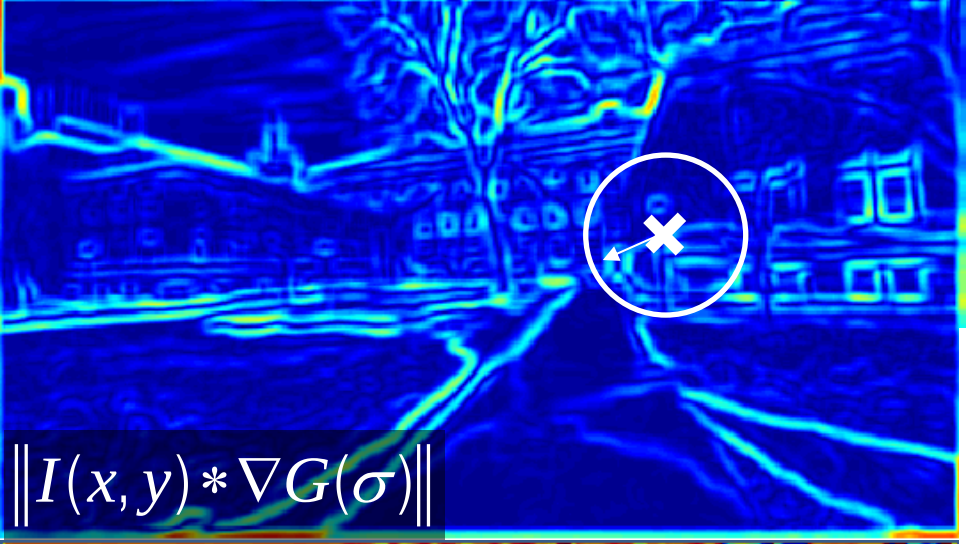
corr 

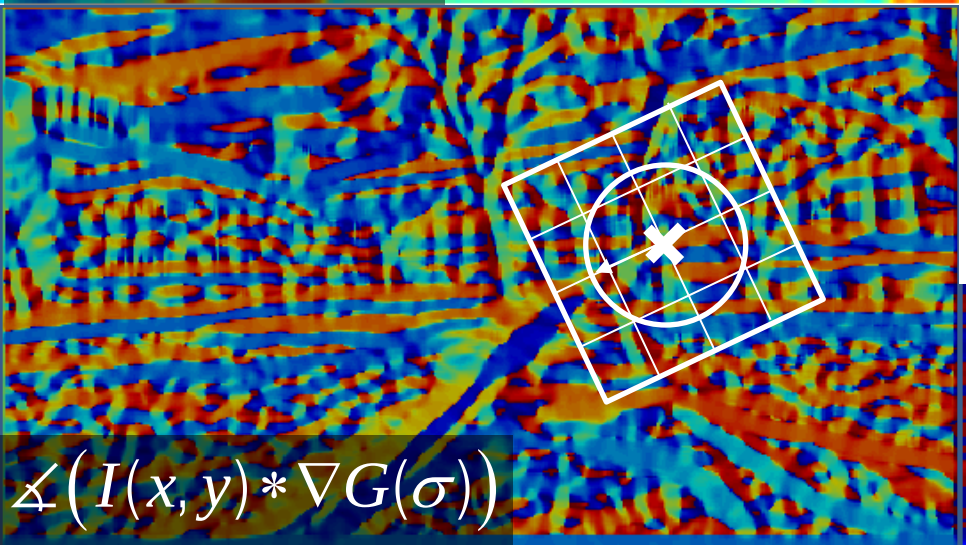
= 0.91

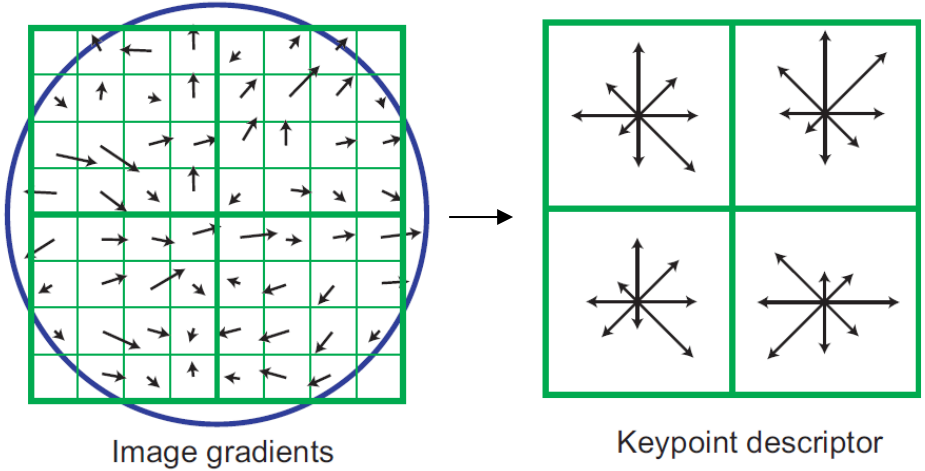
cf)

corr  = 0.15

Descriptor Computation

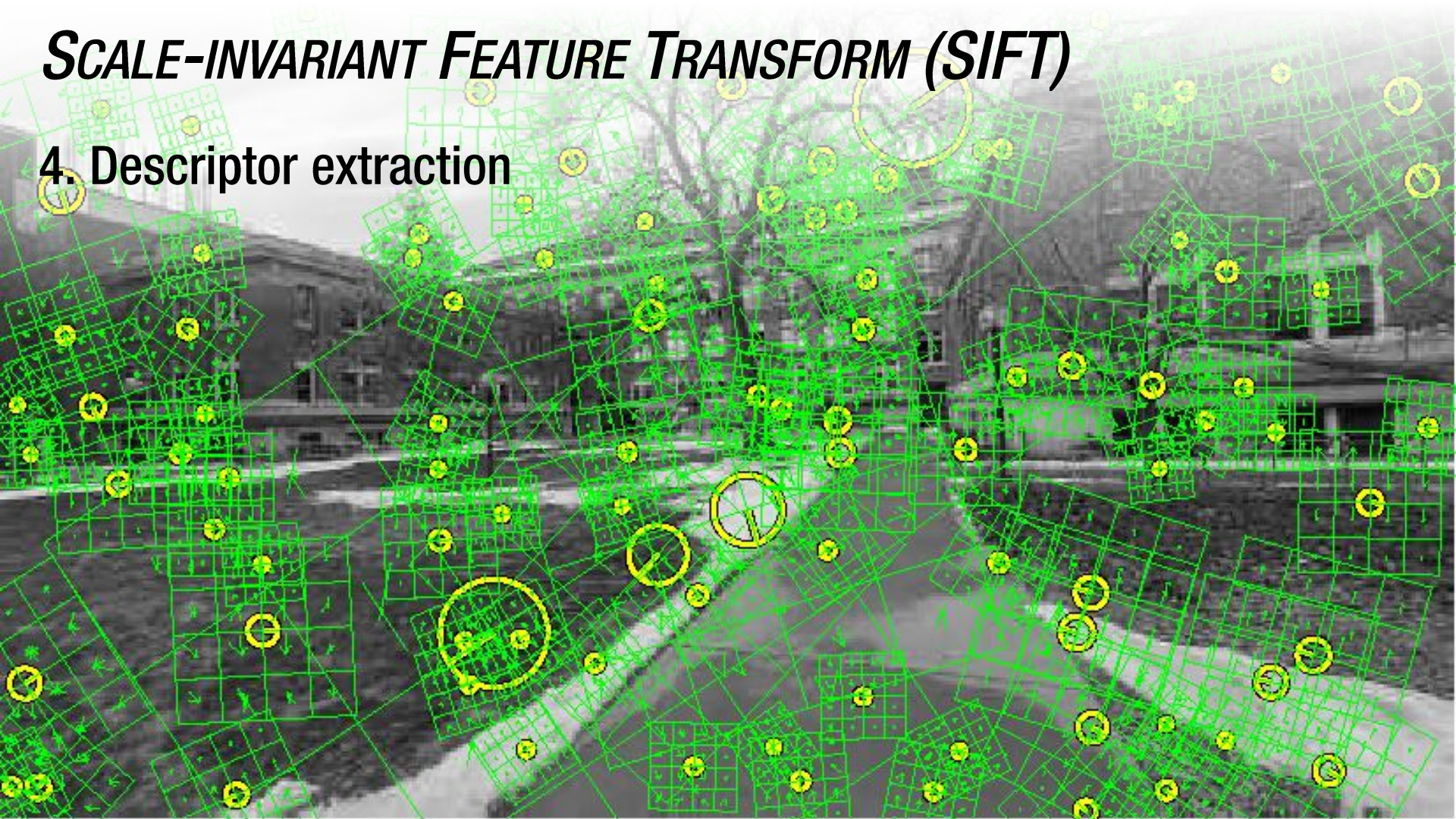




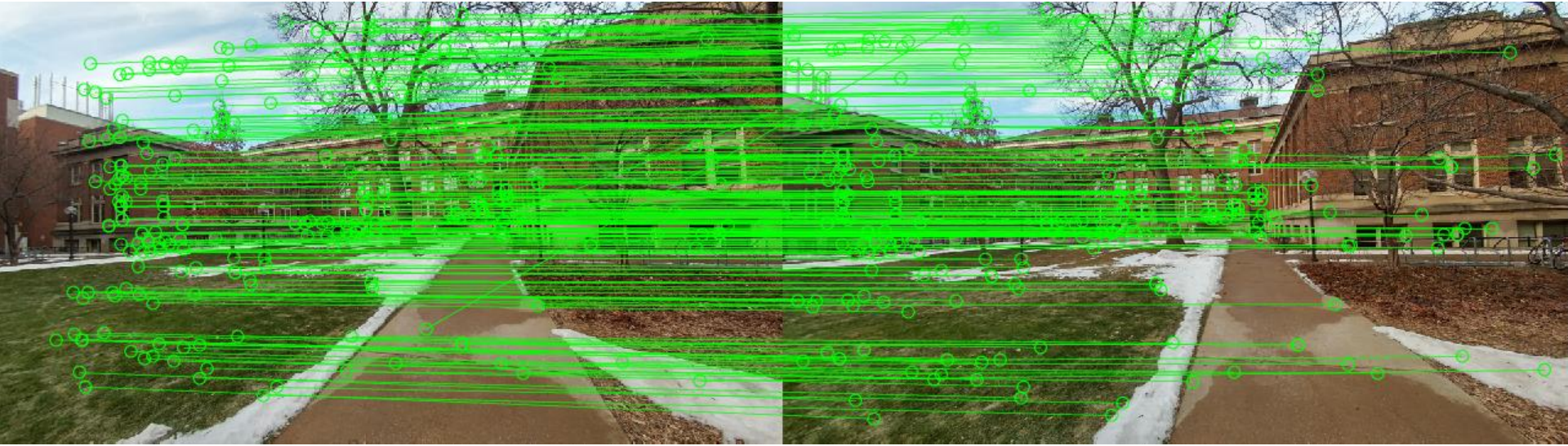


SCALE-INVARIANT FEATURE TRANSFORM (SIFT)

4. Descriptor extraction



Local visual descriptor

Local visual descriptor

$$\left\| \text{descriptor1} - \text{descriptor2} \right\| = 0$$