Johns Hopkins Engineering

Computer Vision

2D Recognition Using SIFT (and other State-of-the-Art Descriptors)

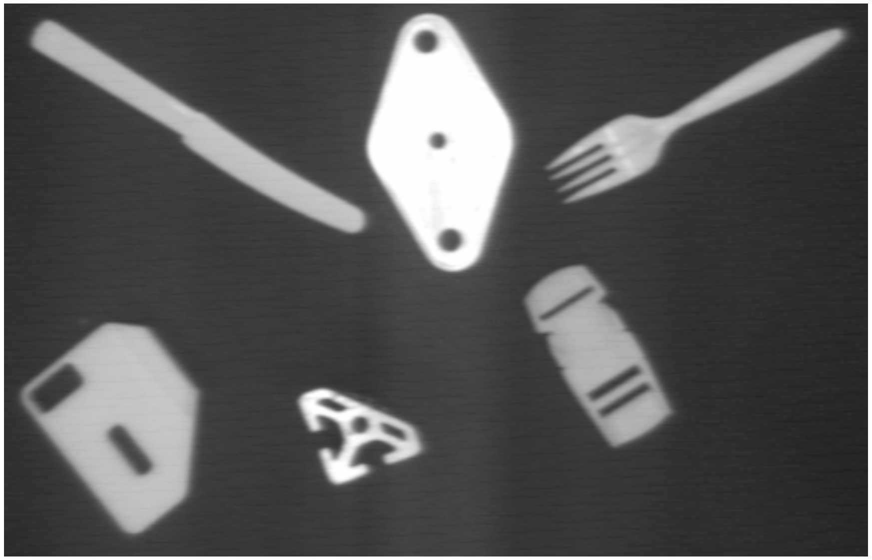
JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

2D Recognition Using SIFT

- Recognize 2D objects in real-world cluttered scenes using the Scale Invariant Feature Transform (SIFT).
- Topics:
  - Local Appearance and Interest Points
  - Blob Detection
  - Scale-Space

A Little Quiz (1)

How would you recognize the following types of objects?



Objects on an assembly line

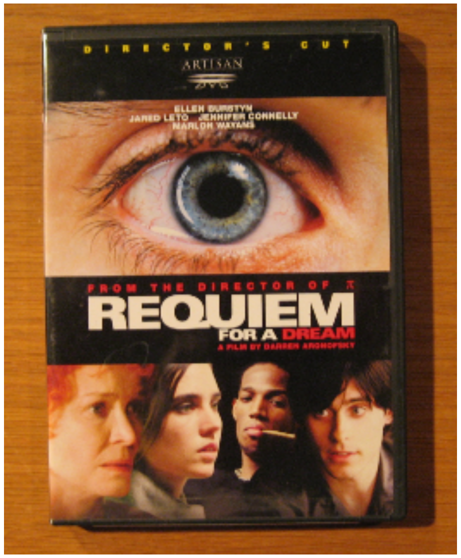
A Little Quiz (2)

How would you recognize the following types of objects?

License plates

A Little Quiz (3)

How would you recognize the following types of objects?



Template

Rich 2D Image

Match "Interesting Points or Features"

Raw Images are Hard to Match

Different Size, Orientation, Lighting, Brightness, …

Removing Sources of Variation

- Matching becomes easier if we can
- remove variations like size and orientation.

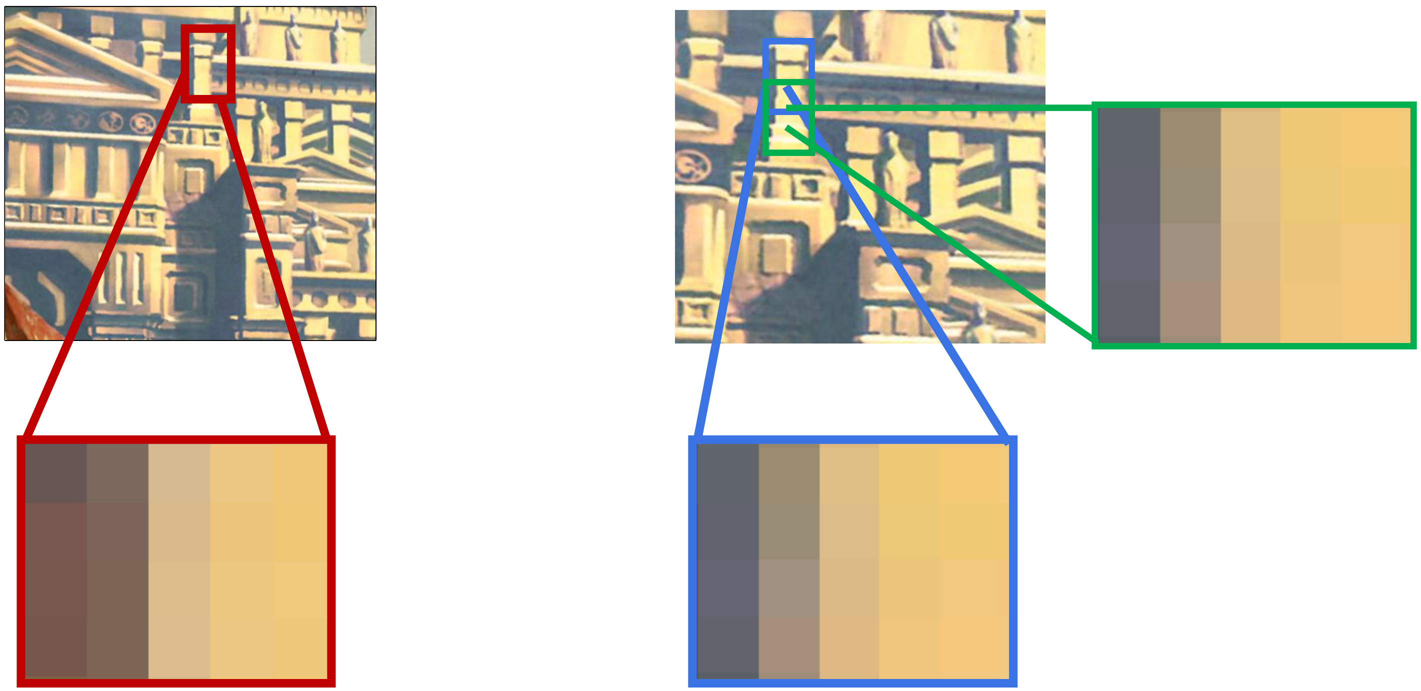
Some Patches are not "Interesting"

What is an Interesting Point/Feature?

- Has **rich image content** (color variations, gradient variations, …) within the local window
- Has well defined **representation** (**signature**) for matching/comparing with other points
- Has a well-defined **position** in the image
- Should be **invariant to image rotation** and **scaling**
- Should be relatively **invariant to lighting** changes

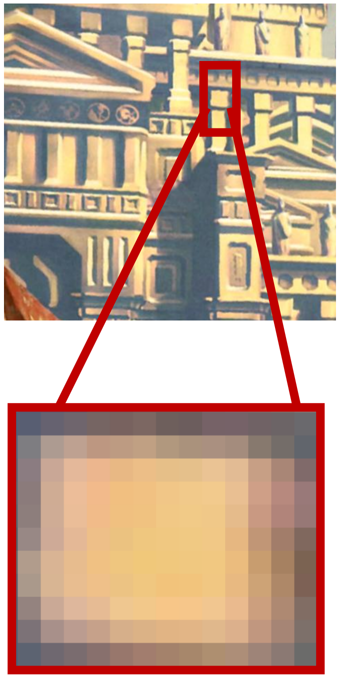
Are Lines/Edges Interesting?

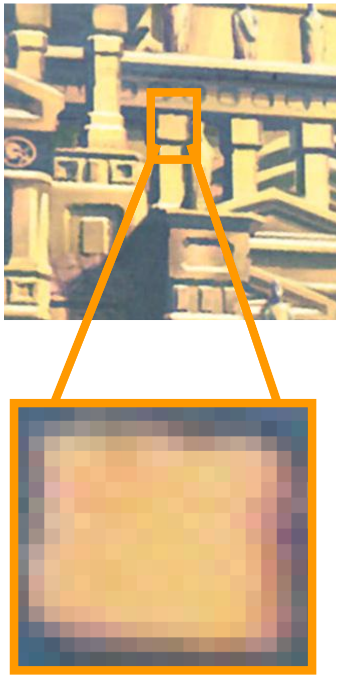
Cannot "**Localize**" an Edge



Are Blobs Interesting?

Yes! Blobs have **fixed position** and definite **size.**





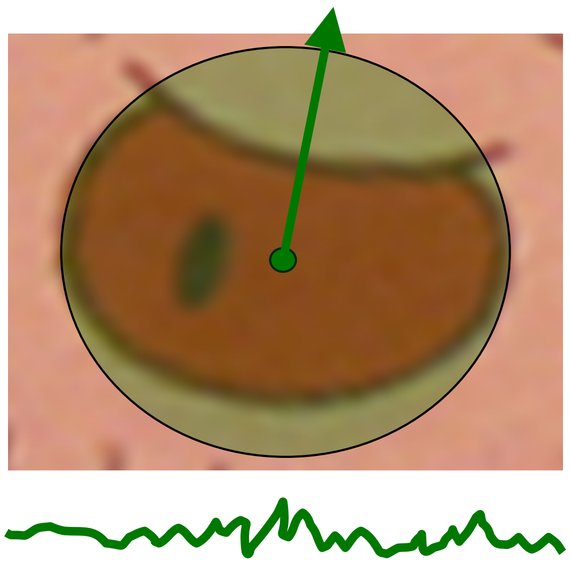
Blobs as Interest Points

We will use Blob-like Features for 2D recognition.

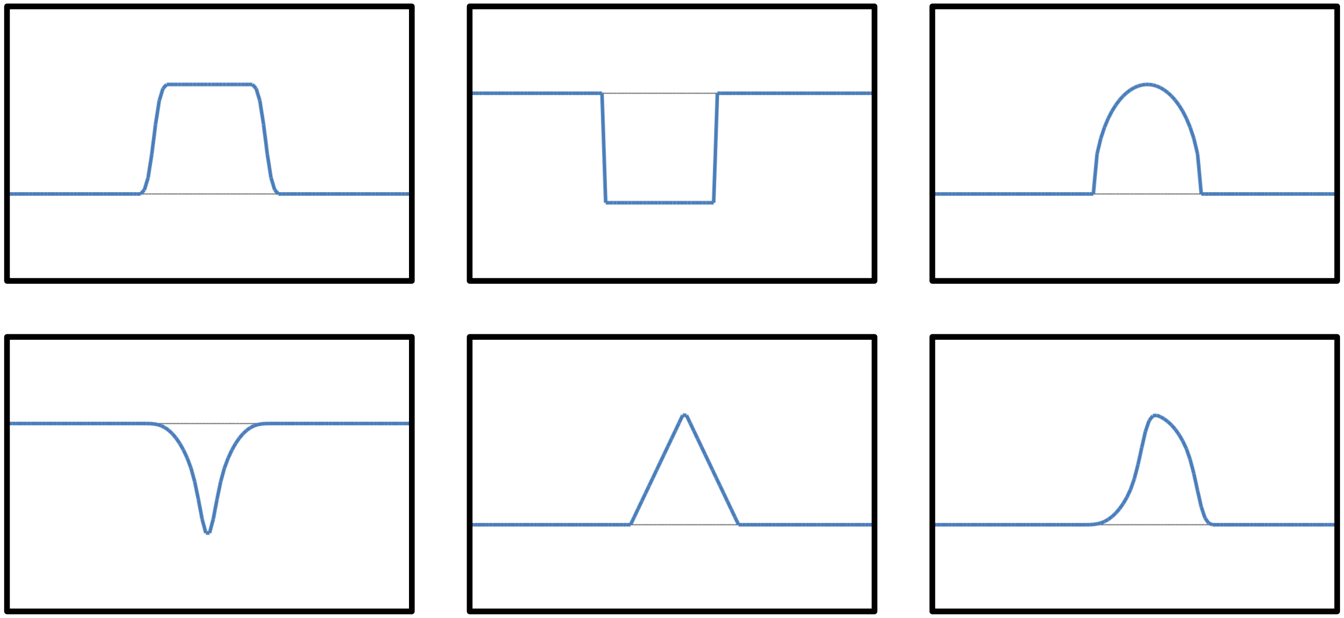
We need to:

Locate a blob

- Determine its **size**
- Determine its **orientation**
- Formulate a **description** or signature that is independent of size and orientation



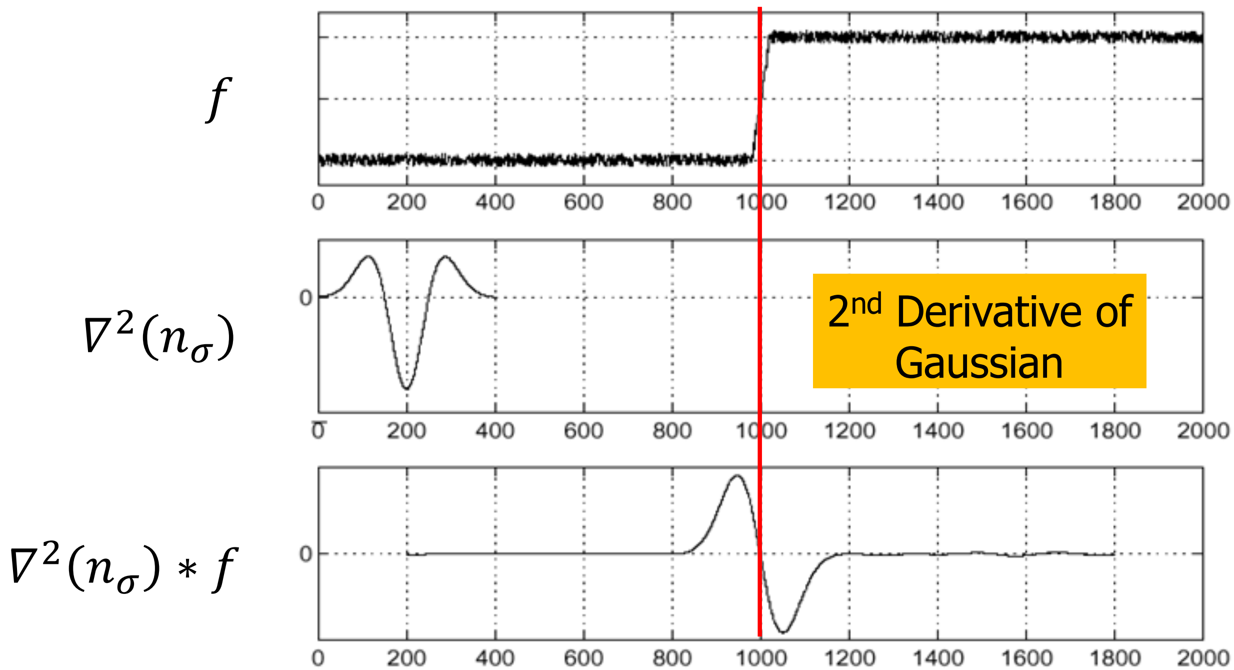
1D Blobs



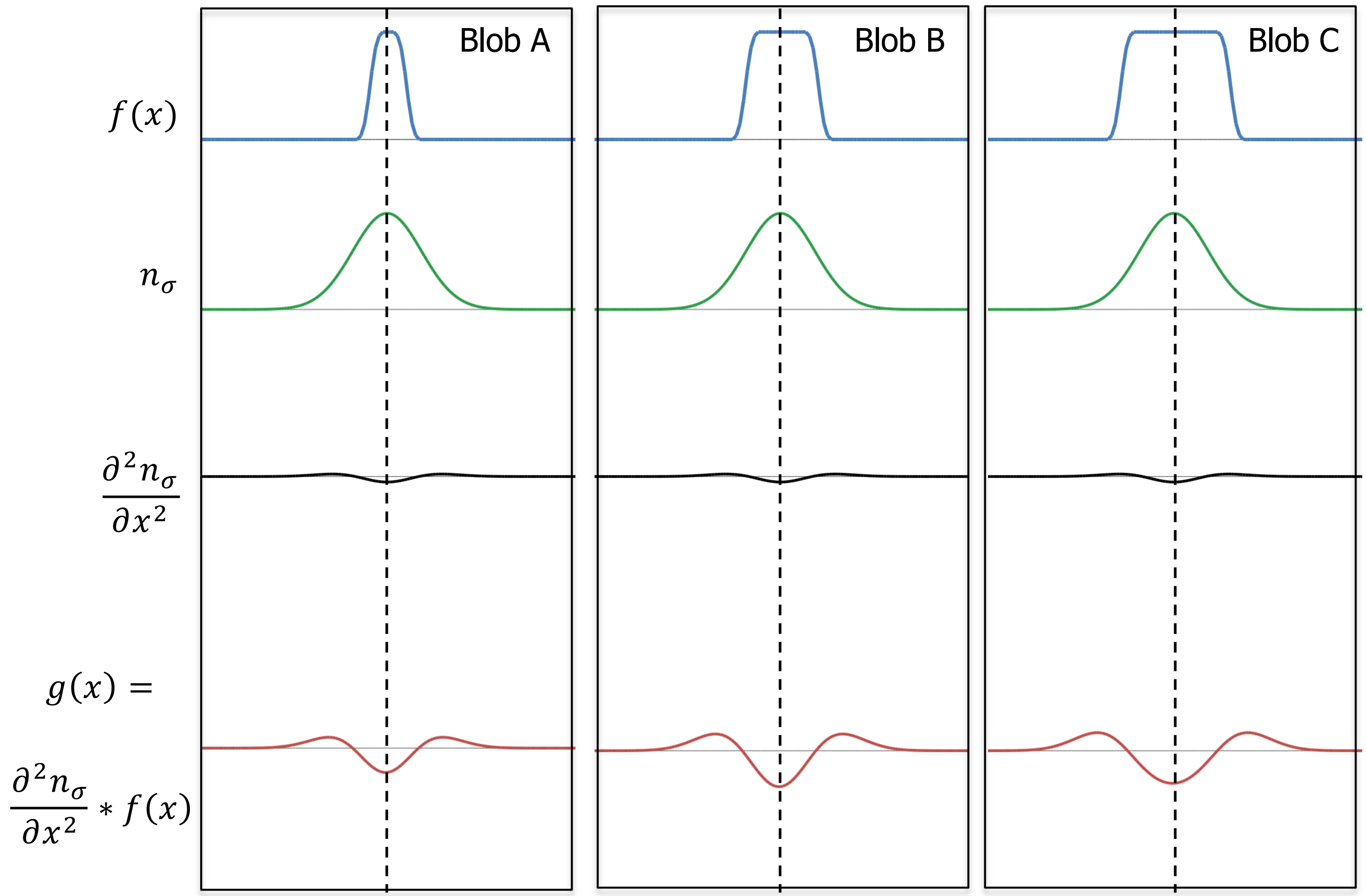
Examples of 1D Blob-like structures

Review: 2^{nd} Derivative of Gaussian

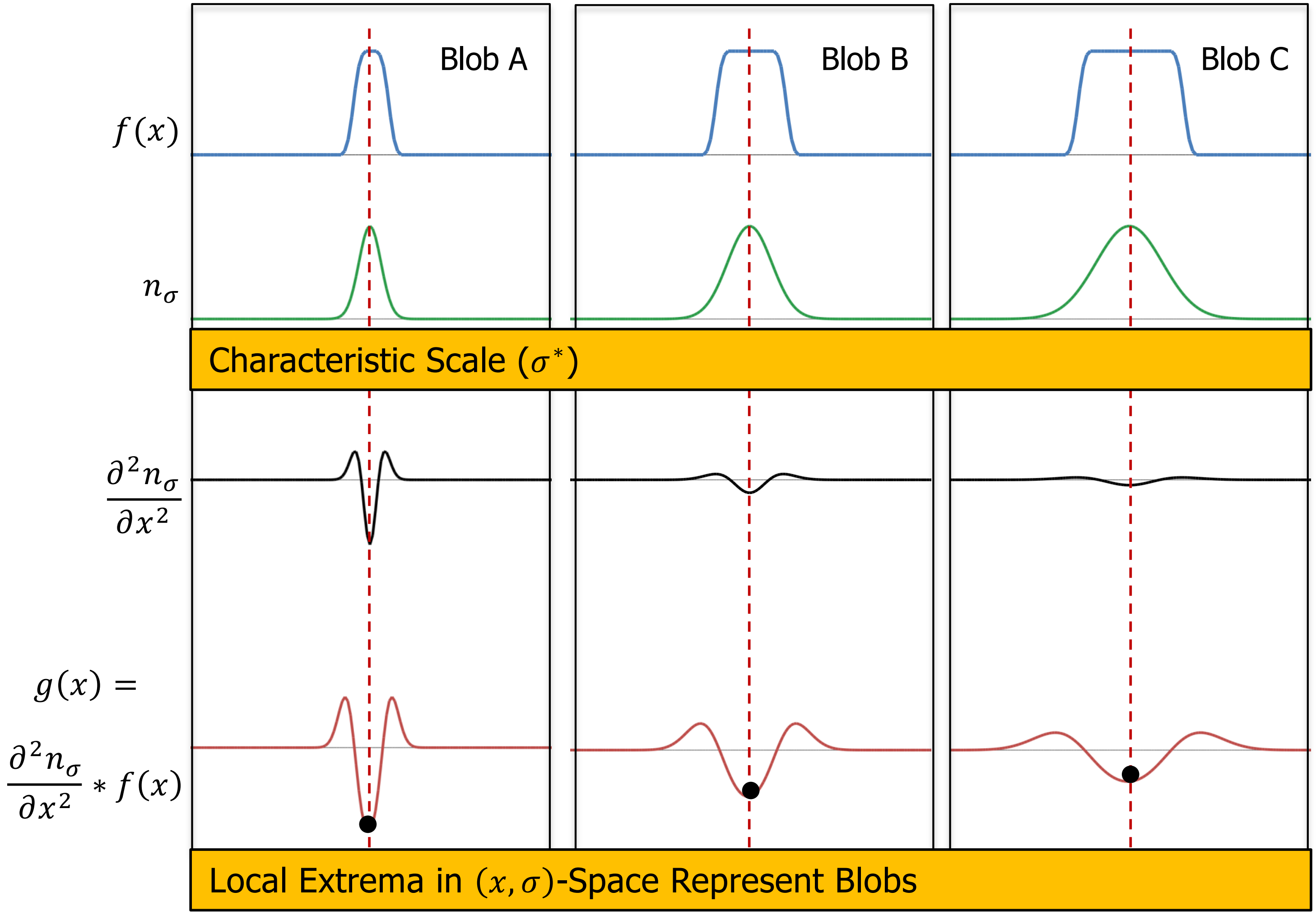
Zero Crossings in 2^{nd} Derivative of Gaussian denotes an Edge



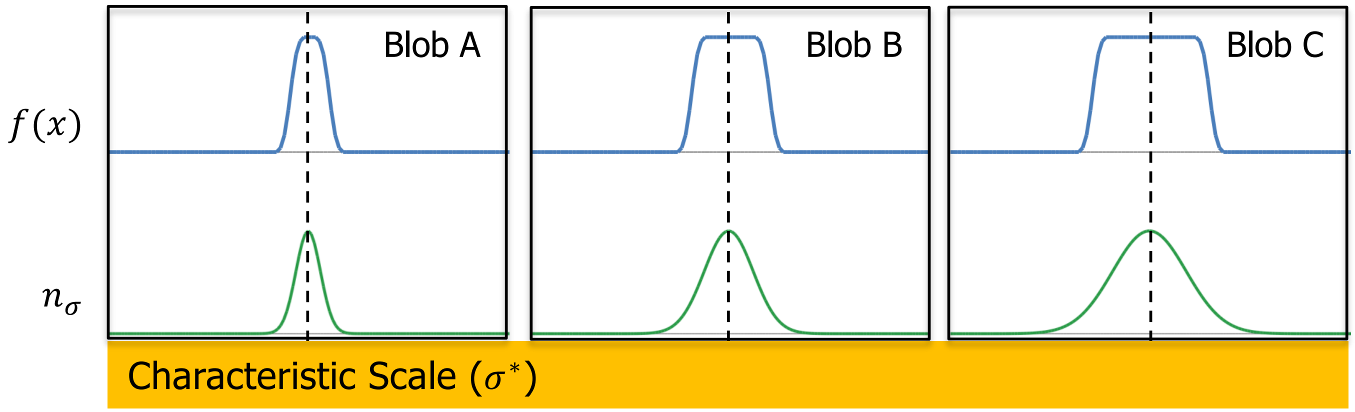
1D Blob and 2nd Derivative of Gaussian



1D Blob and 2^{nd} Derivative of Gaussian



Characteristic Scale and Blob Size



$\sigma_A{}^* = \sigma_1$ $\sigma_B{}^* = 2\sigma_1$ $\sigma_C{}^* = 3\sigma_1$

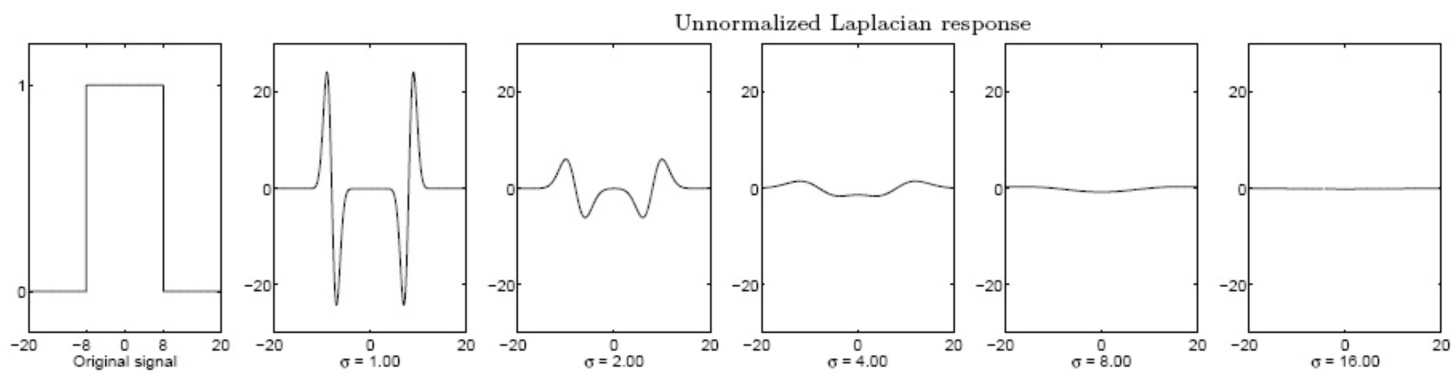
Characteristic Scale: The σ at which 2nd Derivative attains its extreme value.

$\boxed{\text{Characteristic Scale} \propto \text{Size of Blob}}$

$$\Rightarrow \quad \frac{\text{Size of Blob A}}{\text{Size of Blob B}} = \frac{\sigma_A{}^*}{\sigma_B{}^*} \; ; \quad \frac{\text{Size of Blob B}}{\text{Size of Blob C}} = \frac{\sigma_B{}^*}{\sigma_C{}^*}$$

Scale Selection

- We want to find the characteristic scale of the blob by convolving it with Laplacians at several scales and looking for the maximum response
- However, Laplacian response decays as scale increases:



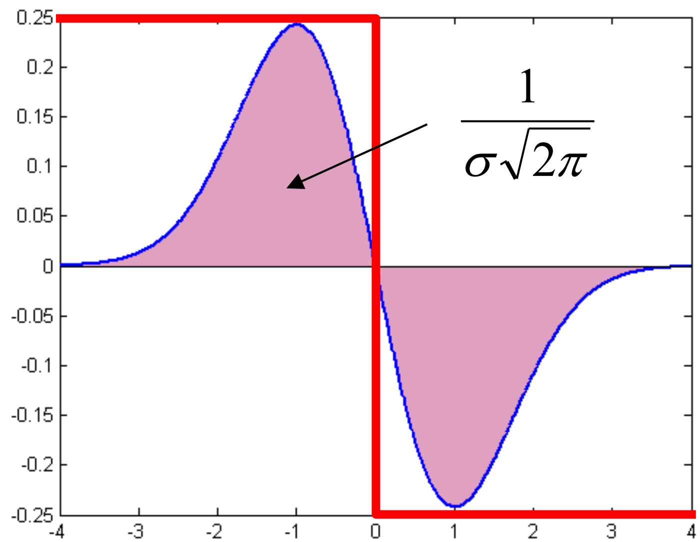
original signal
(radius=8)

increasing σ

Why does this happen?

Scale Normalization

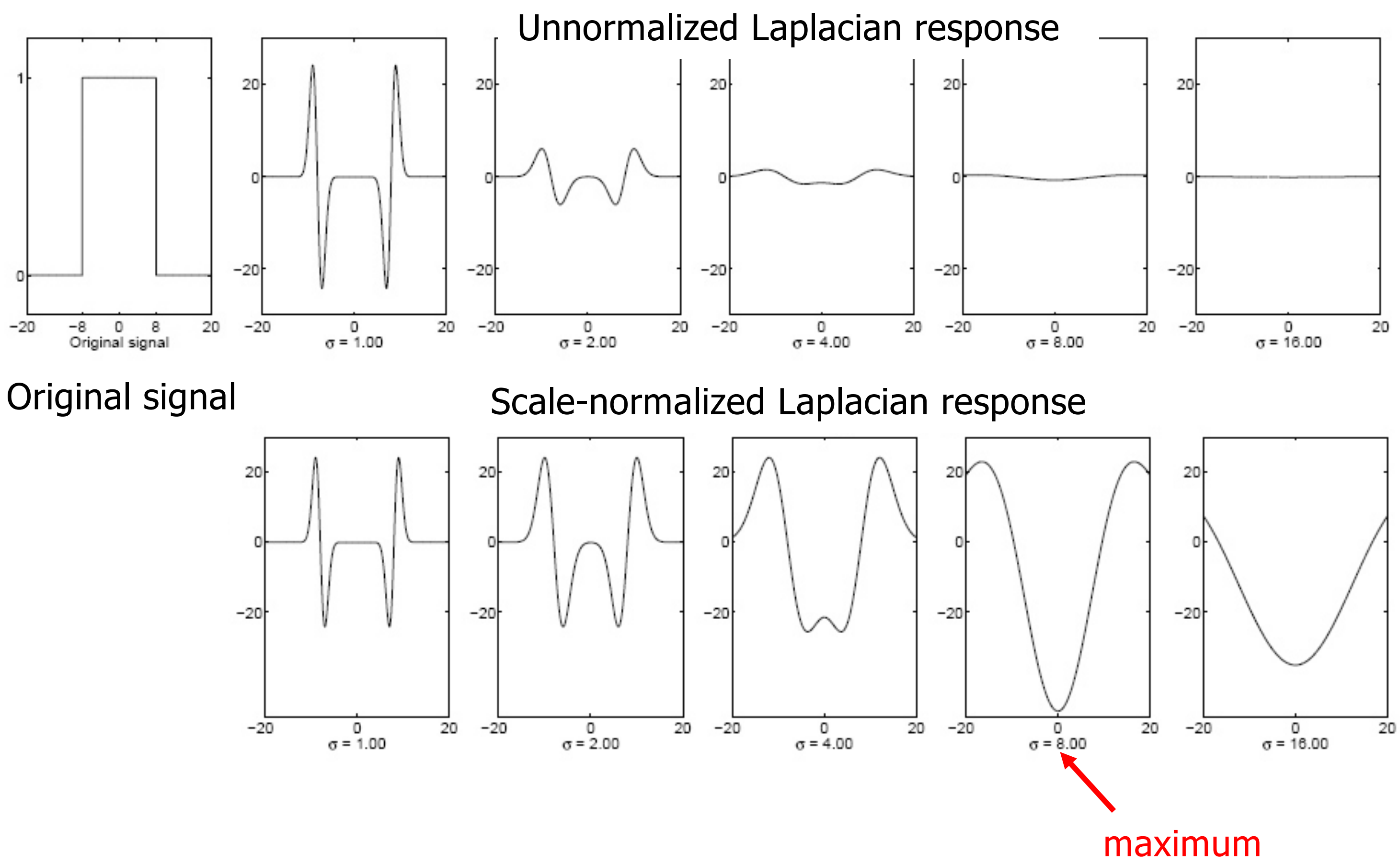
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases



Scale Normalization (cont.)

- 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



1D Blob Detection Summary

Given a 1D signal $f(x)$.

Compute $\frac{\partial^2 n_\sigma}{\partial x^2} * f(x)$ at many scales $(\sigma_0, \sigma_1, \sigma_2, \ldots, \sigma_k)$.

Find:
$$\boxed{(x^*, \sigma^*) = \underset{(x,\sigma)}{\arg\max} \left| \frac{\partial^2 n_\sigma}{\partial x^2} * f(x) \right|}$$

x^*: Blob Position

σ^*: Characteristic Scale (Blob Size)

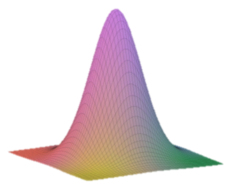
2D Blob Detector

Normalized Laplacian of Gaussian (NLoG) is used as the 2D equivalent for Blob Detection.

Laplacian

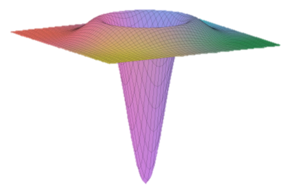
$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

Gaussian



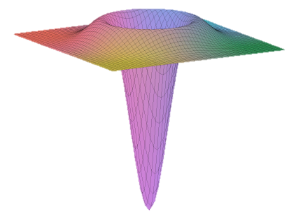
n_σ

LoG



$\nabla^2 n_\sigma$

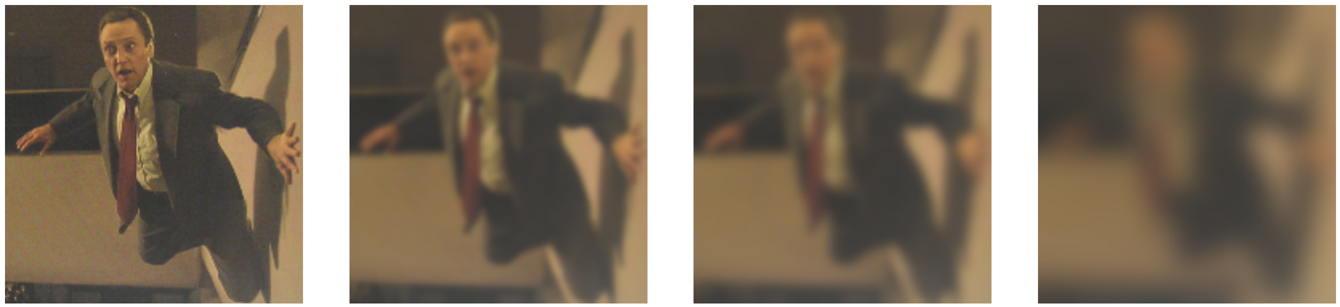
NLoG



$\sigma^2 \nabla^2 n_\sigma$

Location of Blobs given by **Local Extrema** after applying Normalized Laplacian of Gaussian at many scales.

Scale-Space



…

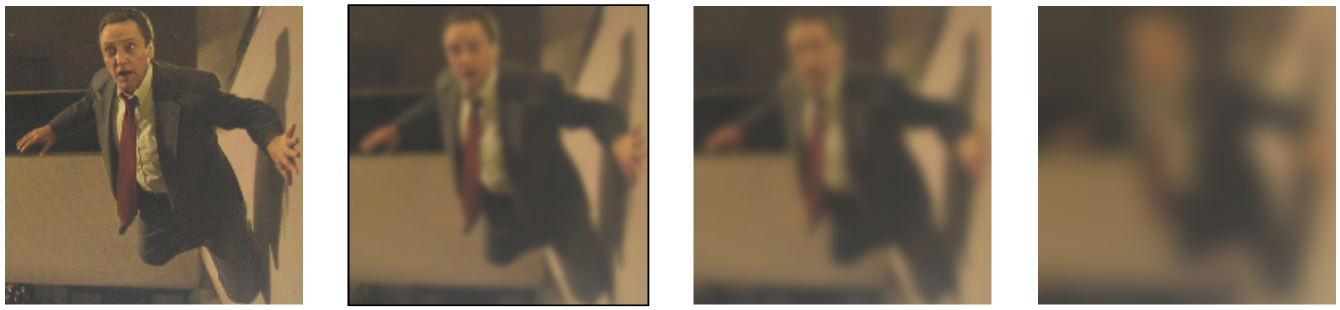
$S(x, y, \sigma_0)$ $S(x, y, \sigma_1)$ $S(x, y, \sigma_2)$ $S(x, y, \sigma_3)$

Increasing σ, Higher Scale, Lower Resolution

Scale Space: Stack created by filtering an image with Gaussians of different sigma(σ)

$$\boxed{S(x, y, \sigma) = n(x, y, \sigma) * I(x, y)}$$

Creating Scale-Space



$S(x, y, \sigma_0)$ $S(x, y, \sigma_1)$ $S(x, y, \sigma_2)$ $S(x, y, \sigma_3)$ …

Increasing σ, Higher Scale, Lower Resolution

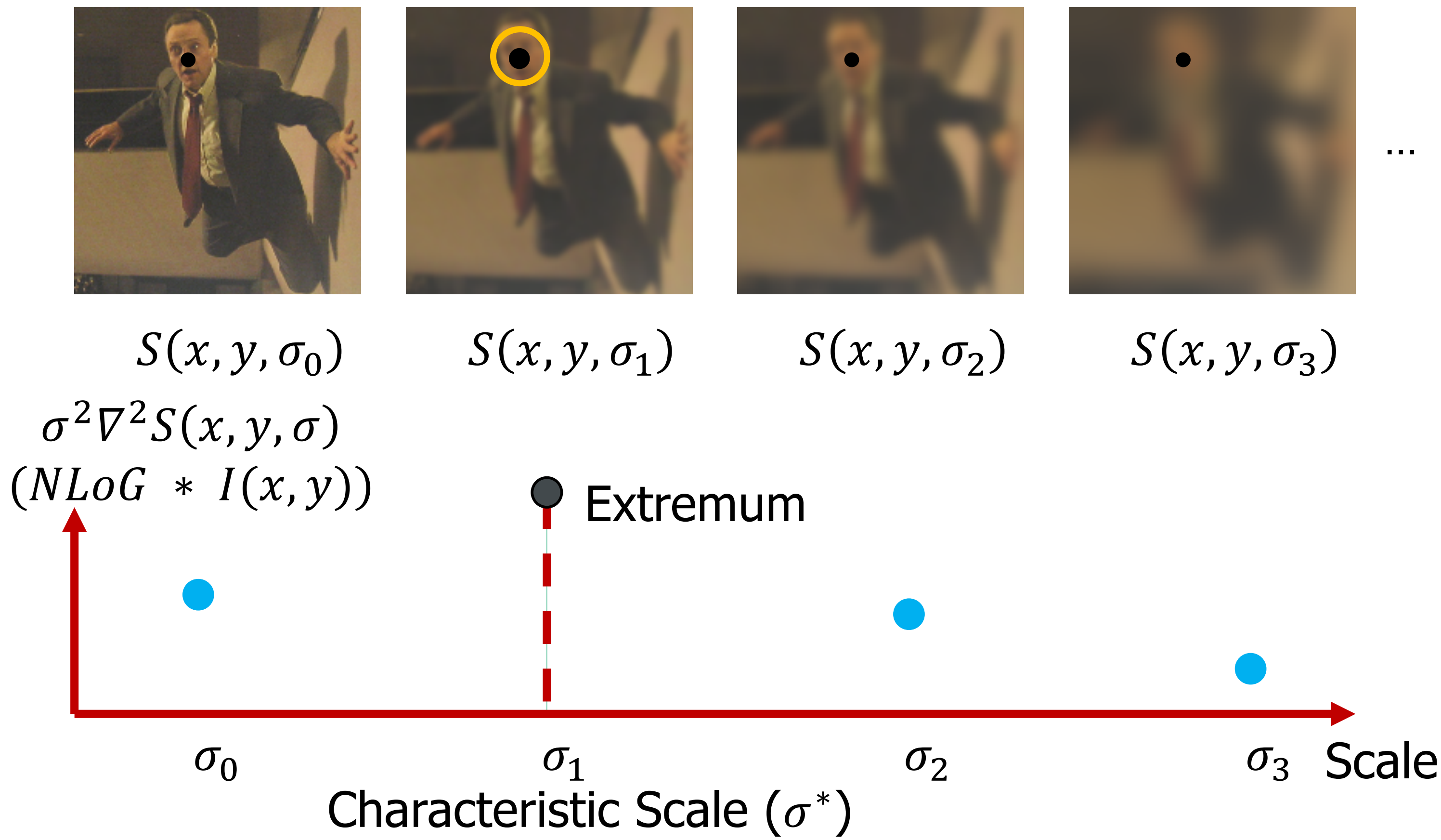
Selecting sigma's to generate the scale-space:

$$\sigma_k = \sigma_0 s^k \quad k = 0,1,2,3, \ldots$$

where, s: Constant multiplier

σ_0: Initial Scale

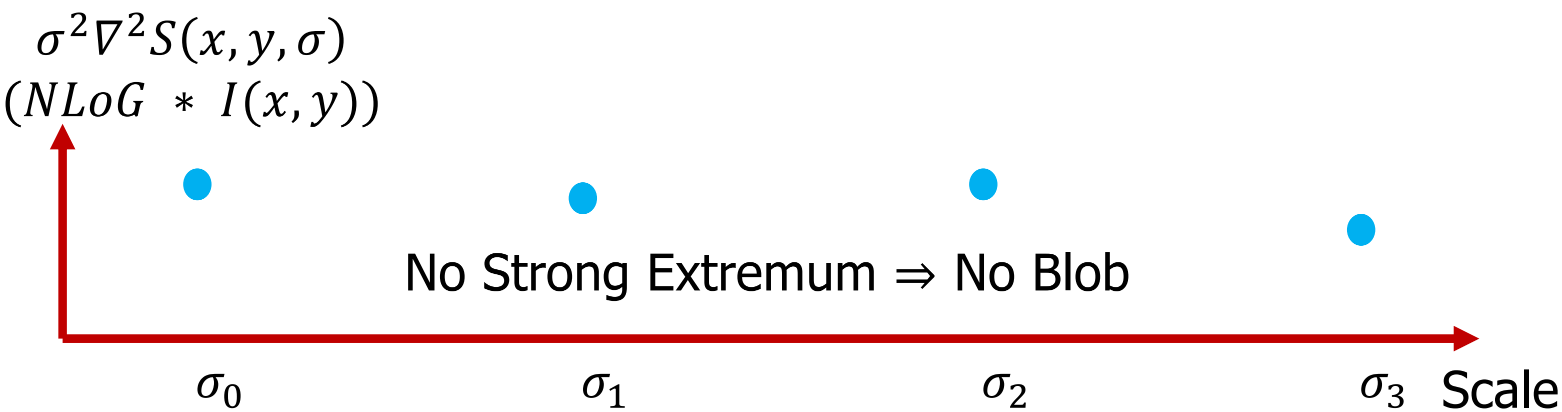
Blob Detection using Local Extrema



[Lindeberg 1994]

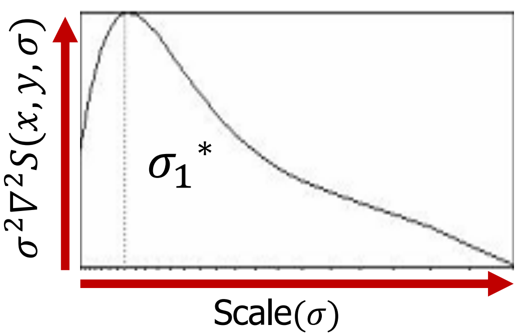
Blob Detection using Local Extrema (cont.)

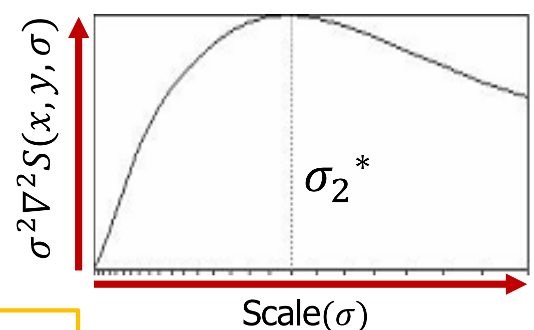
$S(x, y, \sigma_0)$ $S(x, y, \sigma_1)$ $S(x, y, \sigma_2)$ $S(x, y, \sigma_3)$ …



[Lindeberg 1994]

Comparison of Characteristic Scales





$\frac{{\sigma_1}^*}{{\sigma_2}^*}$:Ratio of Blob Sizes

[Mikolajczyk 2002]

2D Blob Detection Summary

Given an image $\mathrm{I}(x, y)$.

Convolve the image using NLoG at many scales σ.

Find:

$$(x^*, y^*, \sigma^*) = \underset{(x,y,\sigma)}{\arg\max} \, |\sigma^2 \nabla^2 n_\sigma * I(x, y)|$$

or

$$(x^*, y^*, \sigma^*) = \underset{(x,y,\sigma)}{\arg\max} \, |\sigma^2 \nabla^2 S(x, y, \sigma)|$$

(x^*, y^*): Position of the blob

σ^*: Size of the blob

Summary

- **Blobs a feature anchors**: A way to identify robust feature locations
- Essential concepts in this lecture:
  - Rotation, scale, and photometric invariance
  - The idea of using Laplacian of Gaussian to detect blobs
  - Scale space of Laplacians

Johns Hopkins Engineering

Computer Vision

2D Recognition Using SIFT (and other State-of-the-Art Descriptors)

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING