2D Recognition Using SIFT

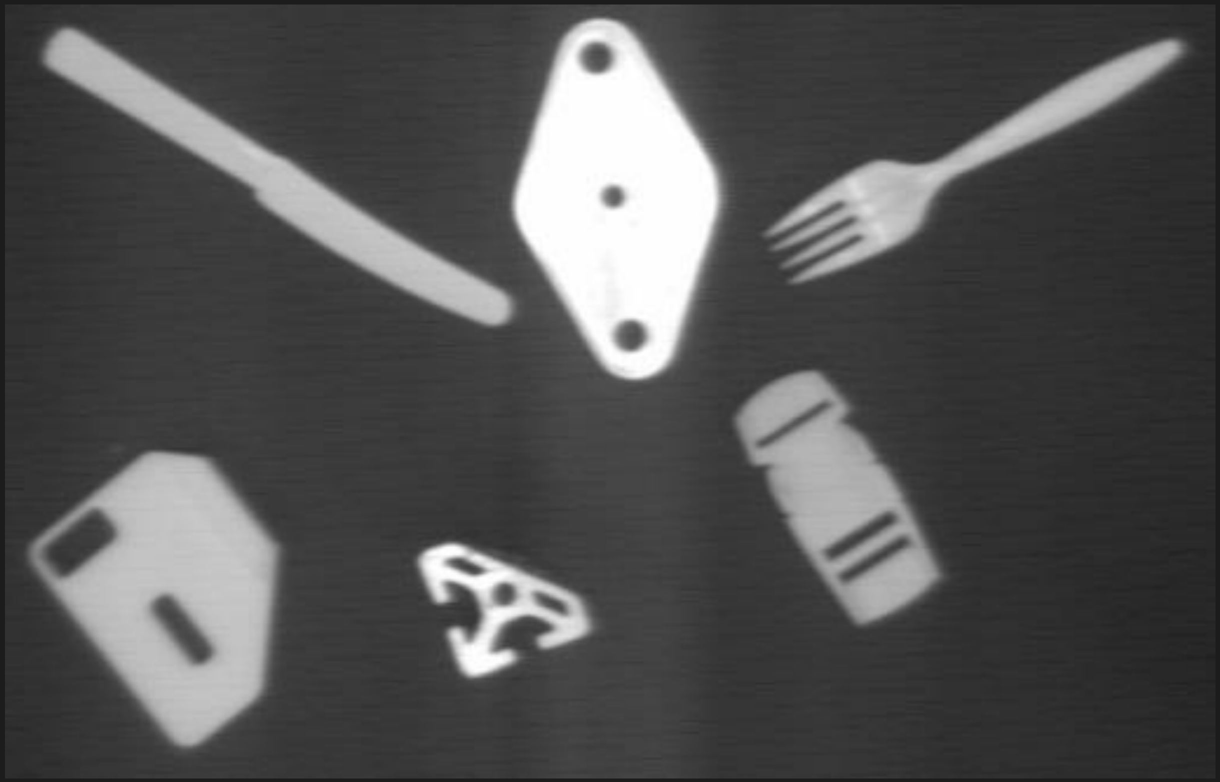
Introduction to Computational Photography:

EECS 395/495

Northwestern University

A Little Quiz

How would you recognize the following types of objects?



Objects on an assembly line

A Little Quiz

How would you recognize the following types of objects?

License plates

A Little Quiz

How would you recognize the following types of objects?

Template

Rich 2D Image

Match "Interesting Points or Features"

2D Recognition Using SIFT

Recognize 2D objects in real-world cluttered scenes using the Scale Invariant Feature Transform (SIFT).

Topics:

(1) Local Appearance and Interest Points

(2) Blob Detection

(3) Scale-Space

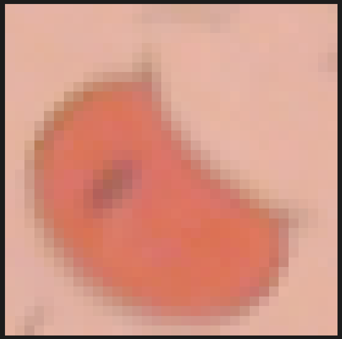
(4) The SIFT interest point detector

(5) Matching and Results

Raw Images are Hard to Match


Different Size, Orientation, Lighting, Brightness, …

Removing Sources of Variation



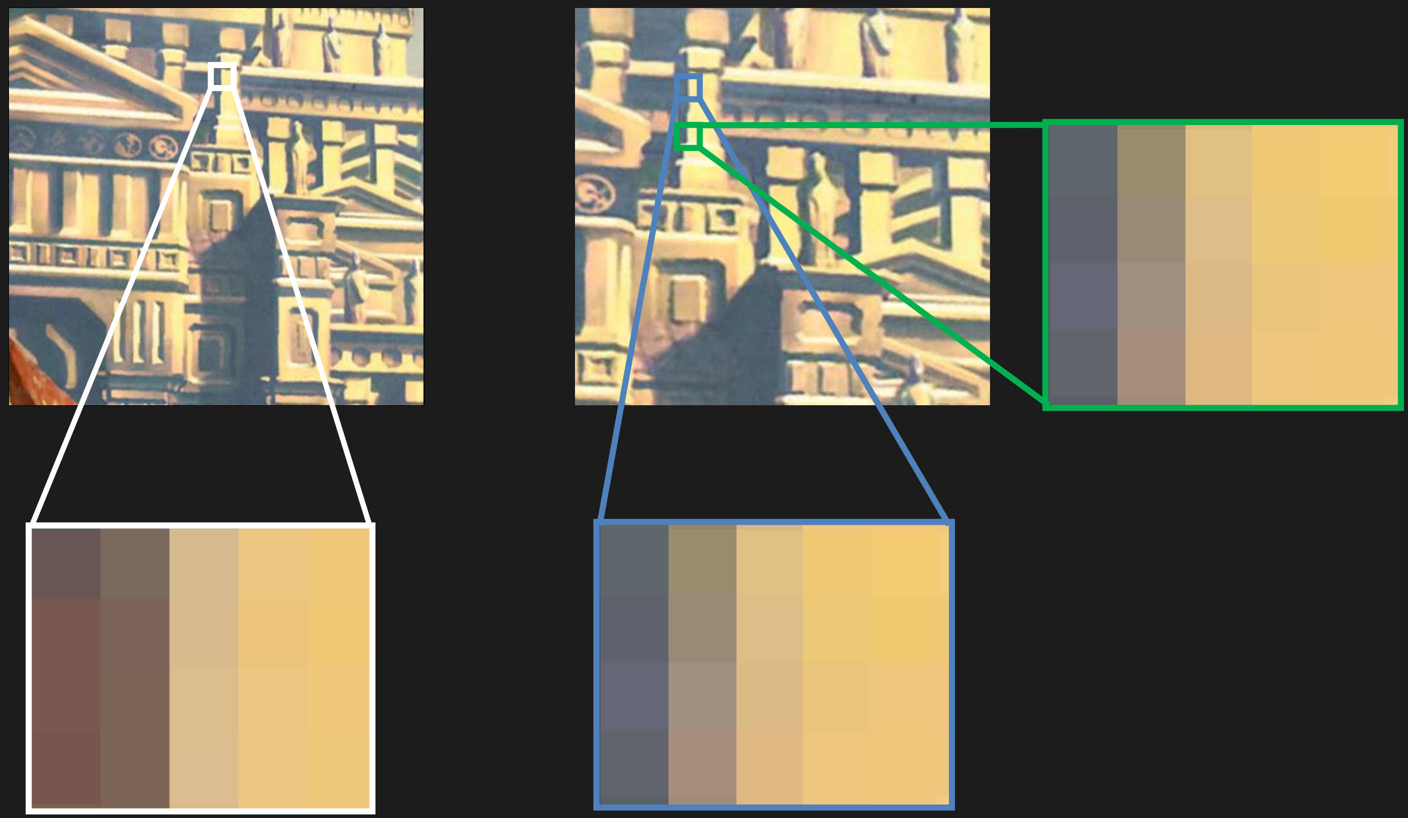
Matching becomes easier if we can
remove variations like size and orientation.

Some Patches are not "Interesting"

What is an Interesting Point/Feature?

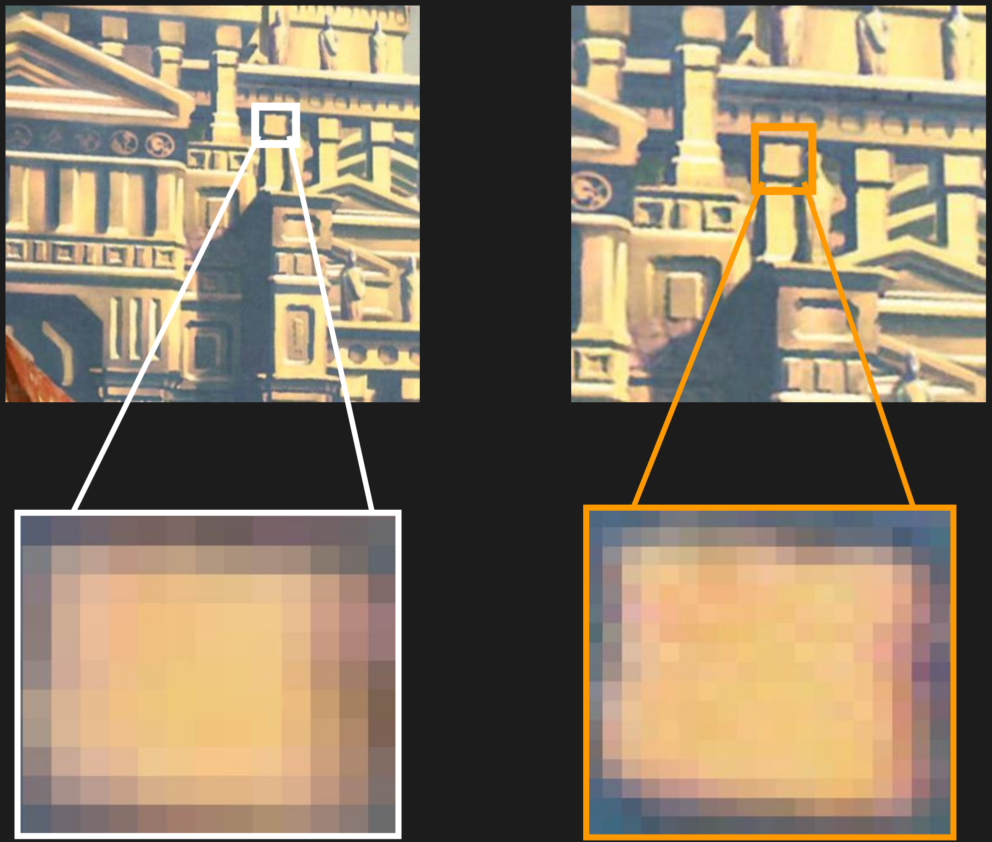
- Has rich image content (color variations, gradient variations, …) within the local window
- Has well defined representation (signature) for matching/comparing with other points
- Has a well defined position in the image
- Should be invariable to image rotation and scaling
- Should be relatively invariable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

Are Lines/Edges Interesting?



Cannot "Localize" an Edge

Are Blobs Interesting?



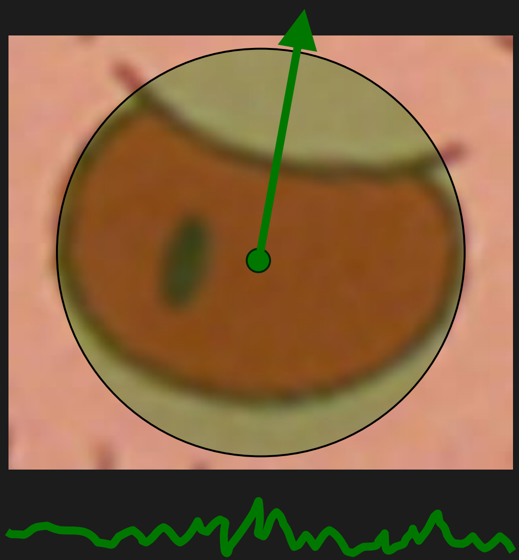
Yes! Blobs have fixed position and definite size.

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

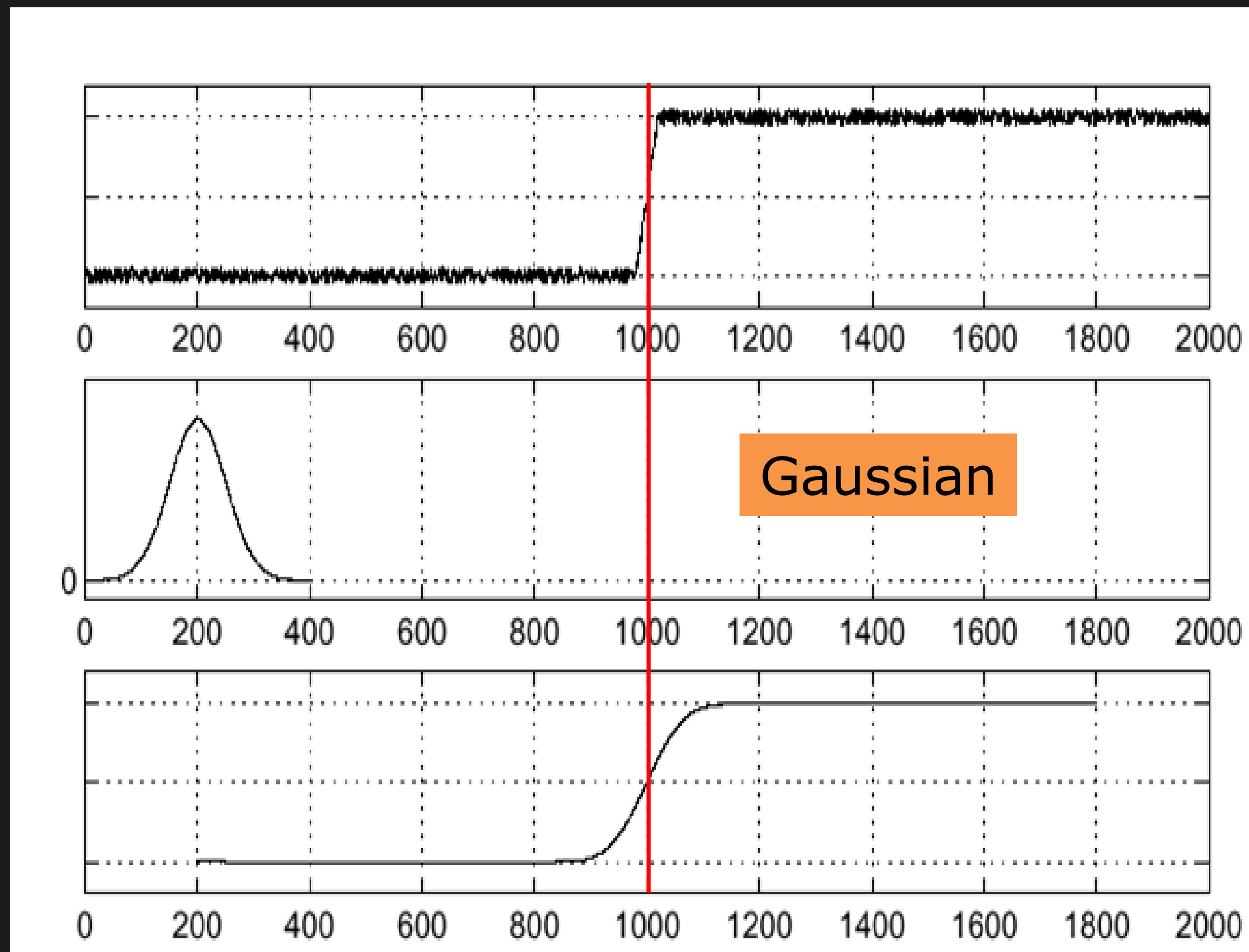


Review: Gaussian Filter

f

n_σ

$n_\sigma * f$



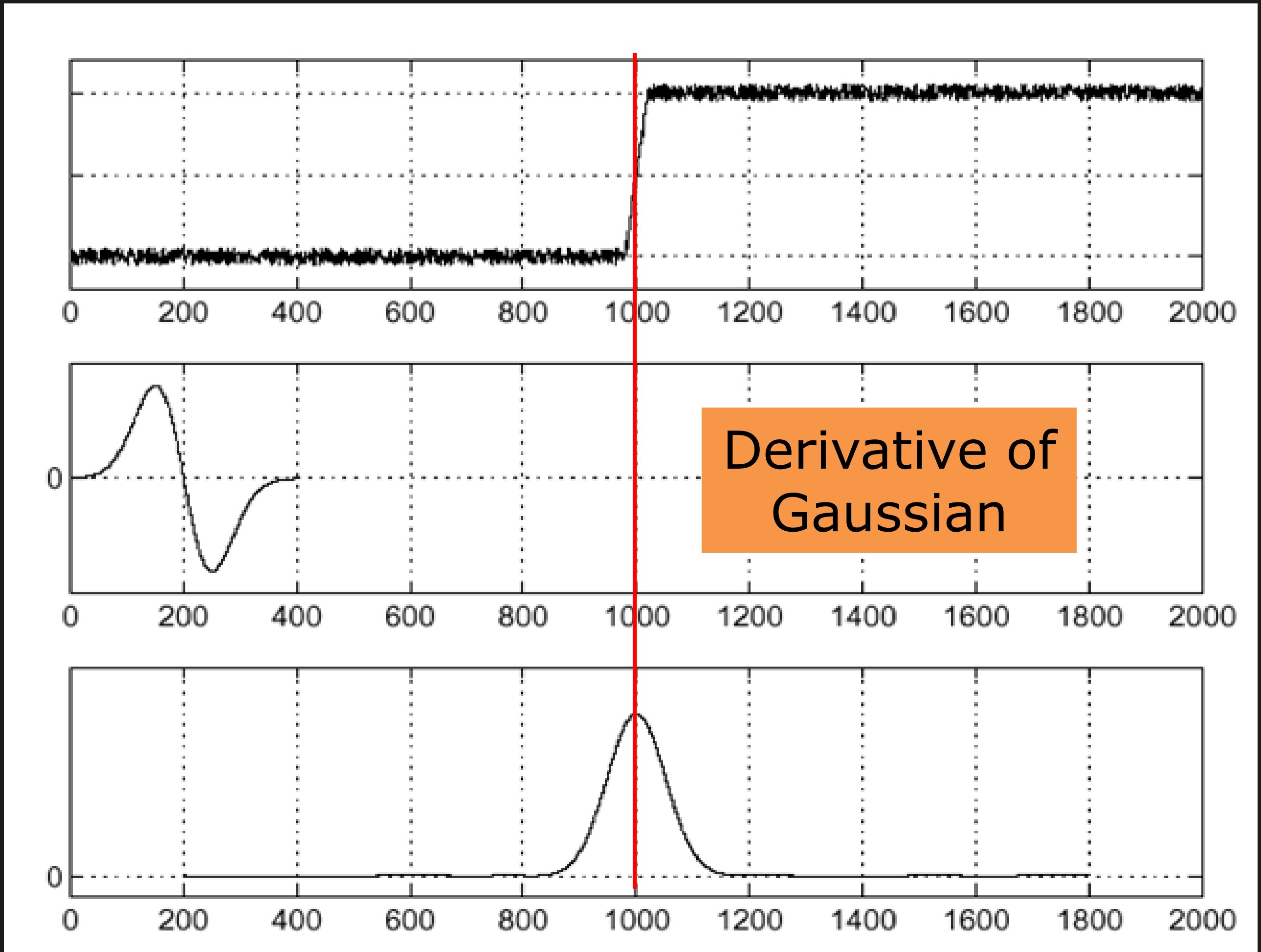
Gaussian Filter is used for removing noise by smoothing

Review: Derivative of Gaussian

f

$\nabla(n_\sigma)$

$\nabla(n_\sigma) * f$



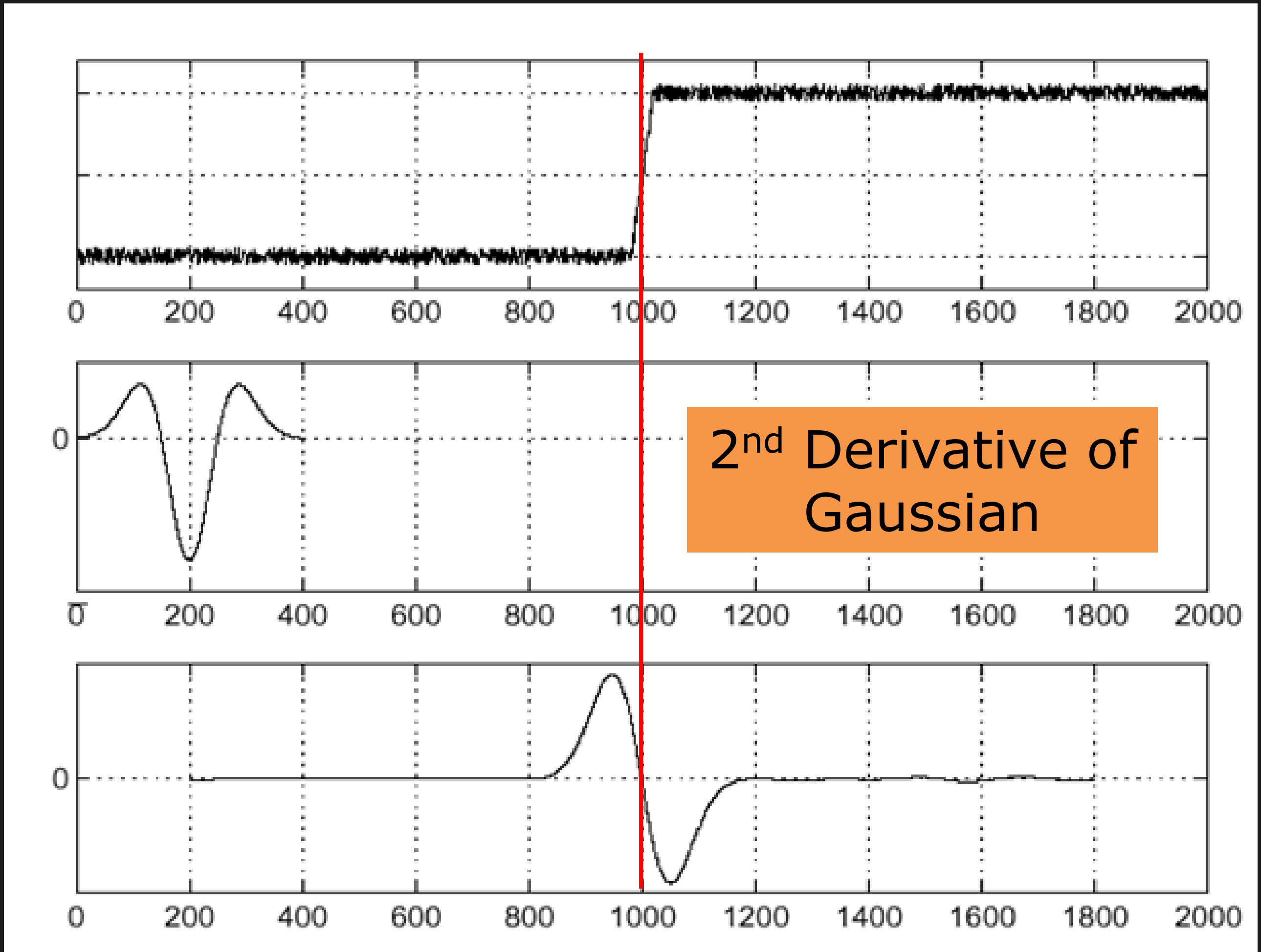
Extremum of Derivative of Gaussian denotes an Edge

Review: 2nd Derivative of Gaussian

f

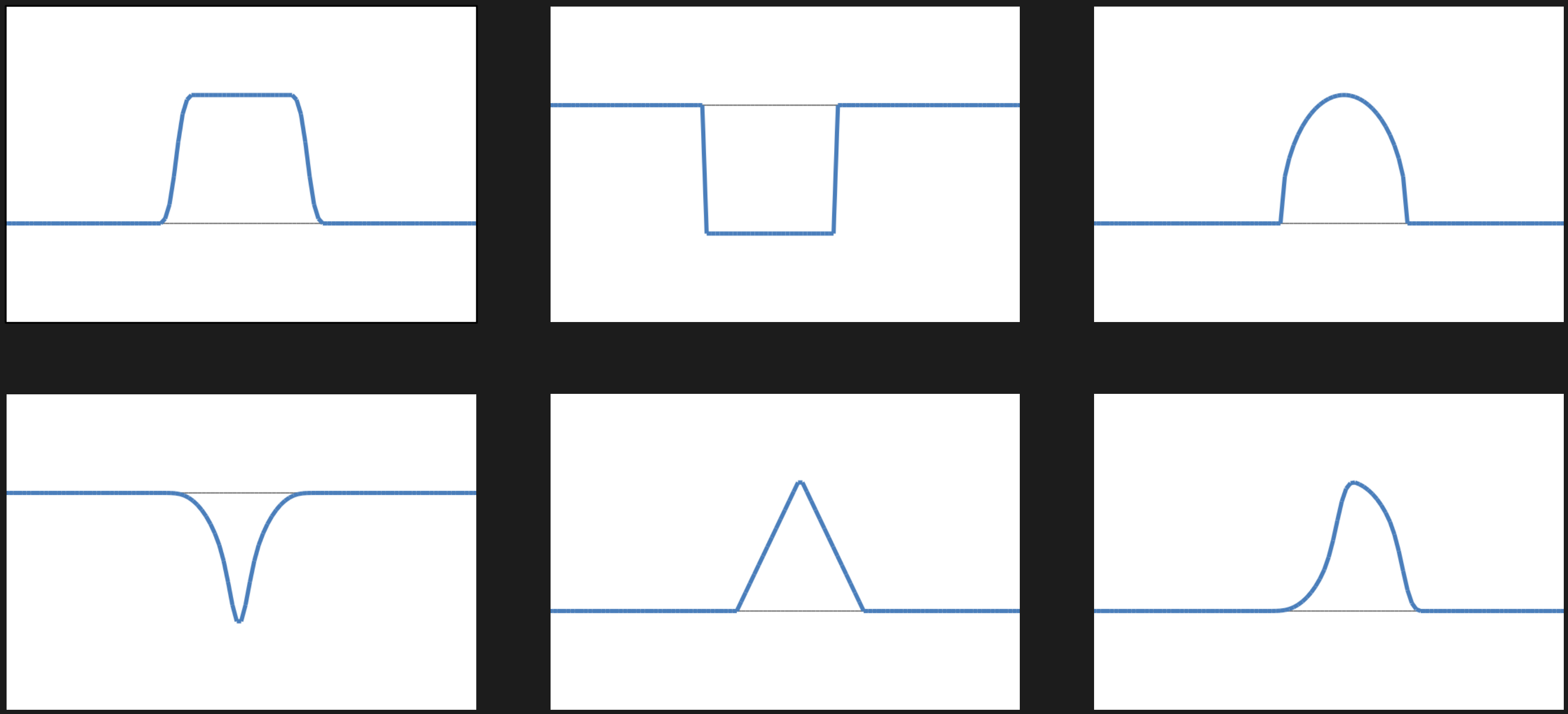
$\nabla^2(n_\sigma)$

$\nabla^2(n_\sigma) * f$



Zero Crossings in 2nd Derivative of Gaussian denotes an Edge

1D Blobs



Examples of 1D Blob-like structures

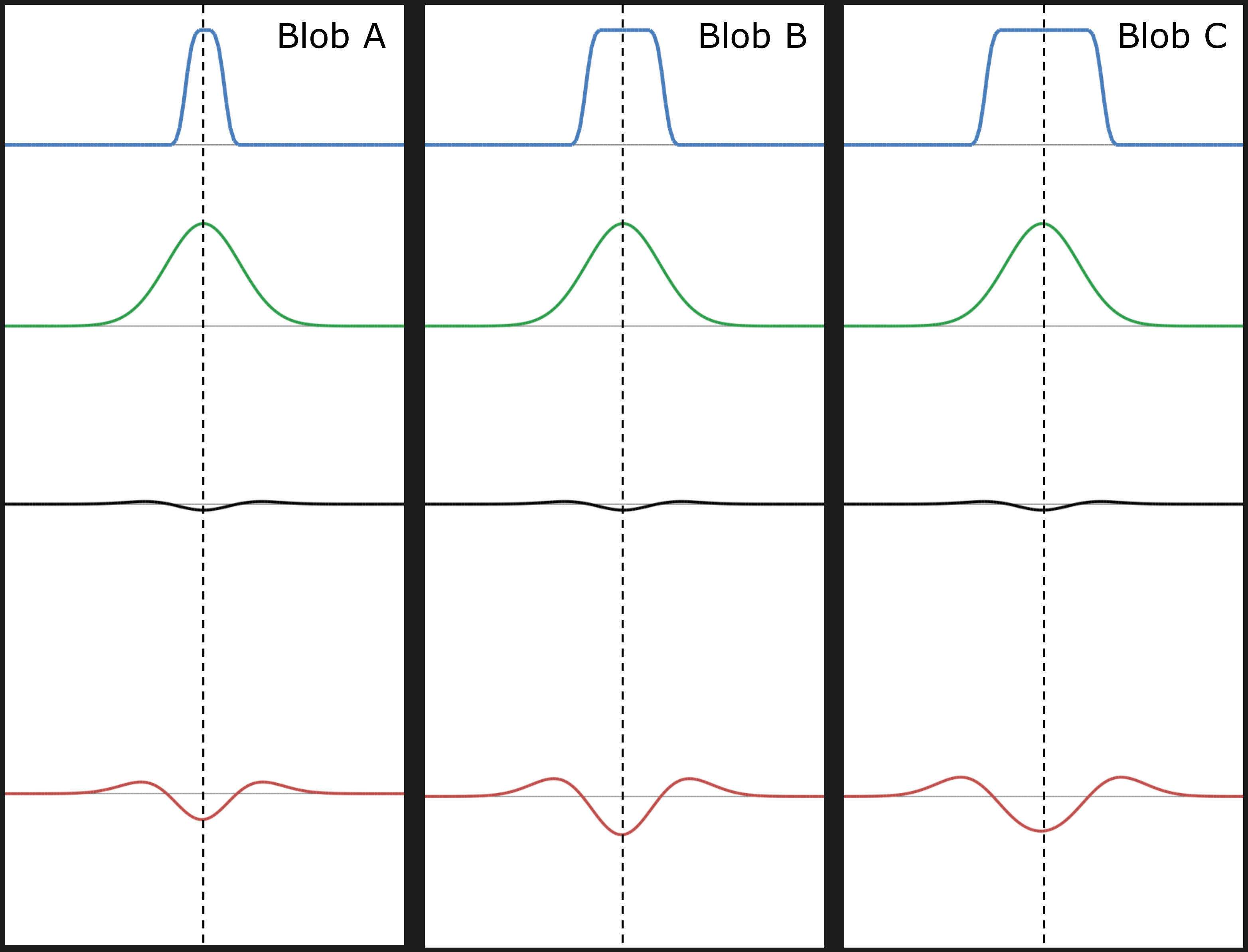
1D Blob and 2nd Derivative of Gaussian

$f(x)$

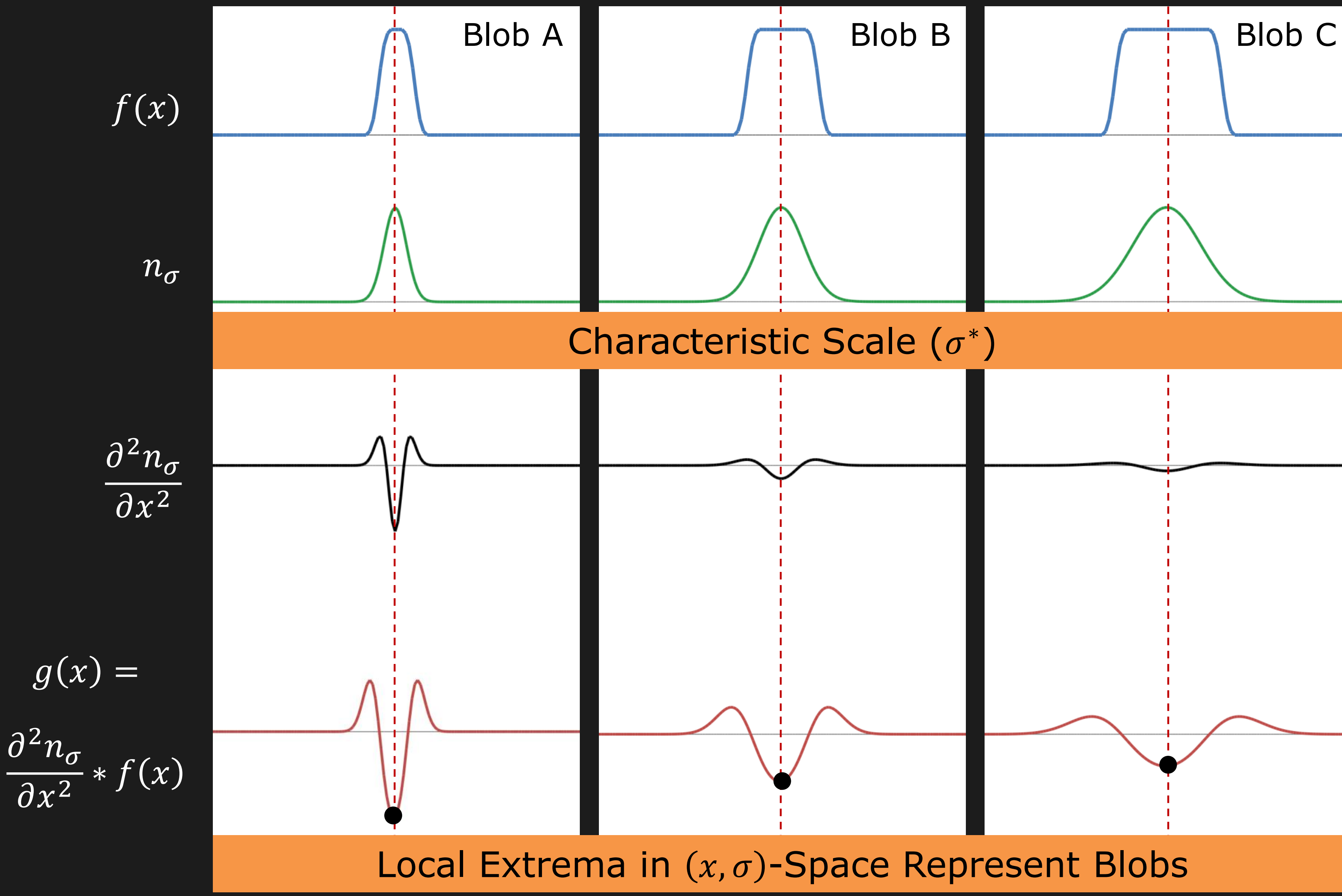
n_σ

$\frac{\partial^2 n_\sigma}{\partial x^2}$

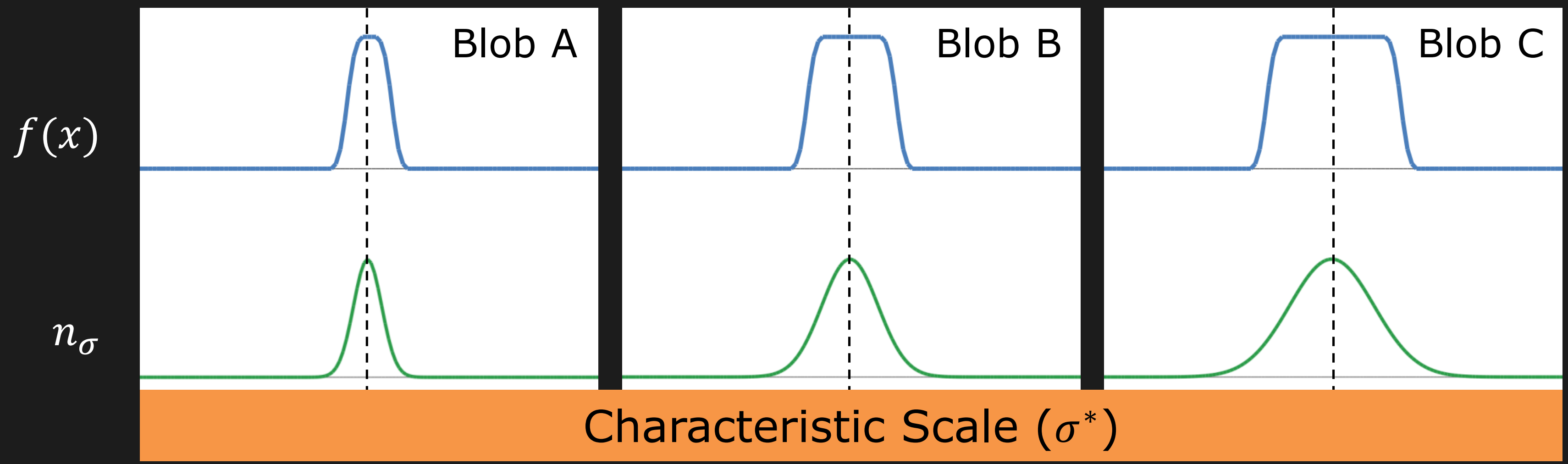
$g(x) = \frac{\partial^2 n_\sigma}{\partial x^2} * f(x)$



1D Blob and 2nd 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.

Characteristic Scale ∝ Size of Blob

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

1D Blob Detection Summary

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

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

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)

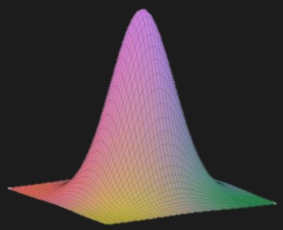
2D Blob Detector

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

Laplacian

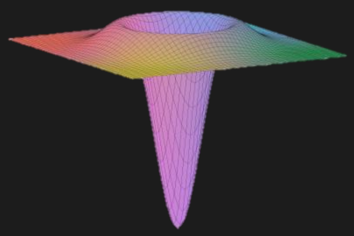
$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

Gaussian



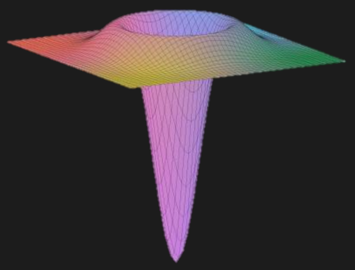
n_σ

LoG



$\nabla^2 n_\sigma$

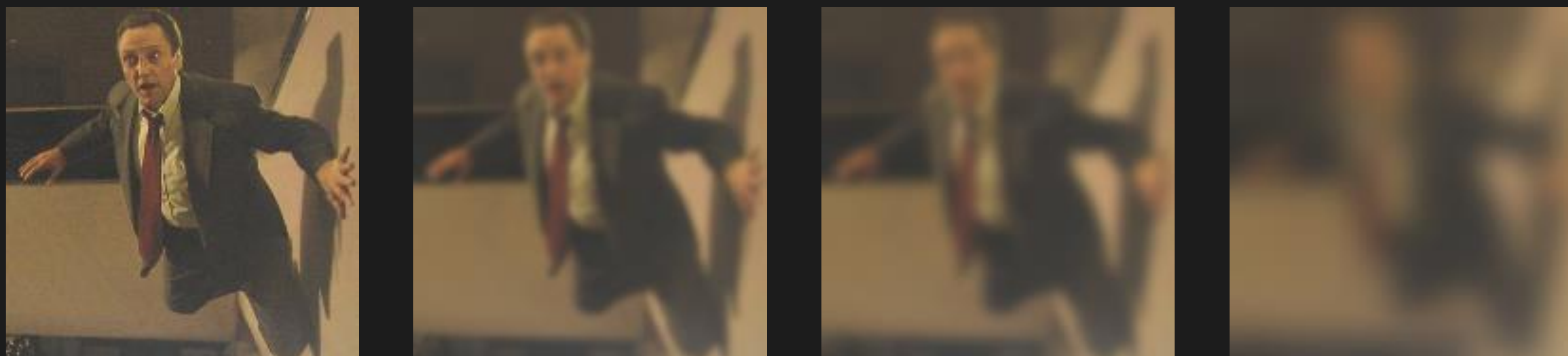
NLoG



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

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

Scale-Space



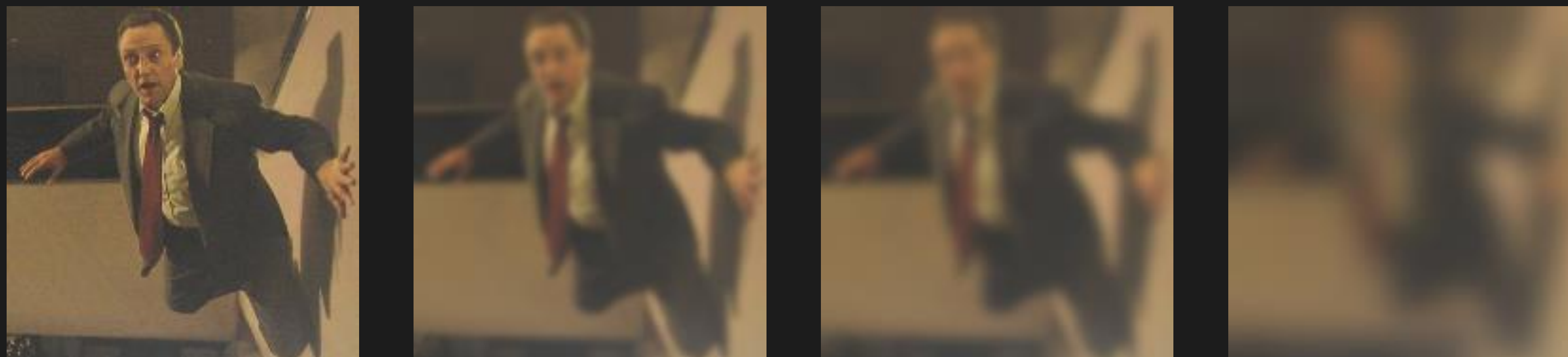
$S(x, y, \sigma_0)$ $S(x, y, \sigma_1)$ $S(x, y, \sigma_2)$ $S(x, y, \sigma_3)$ …

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

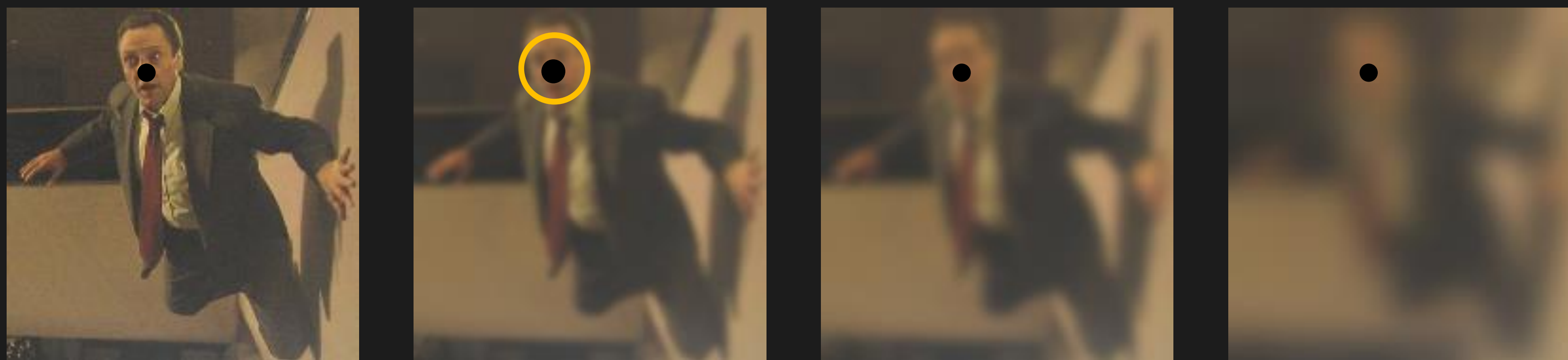
Selecting sigma's to generate the scale-space:

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

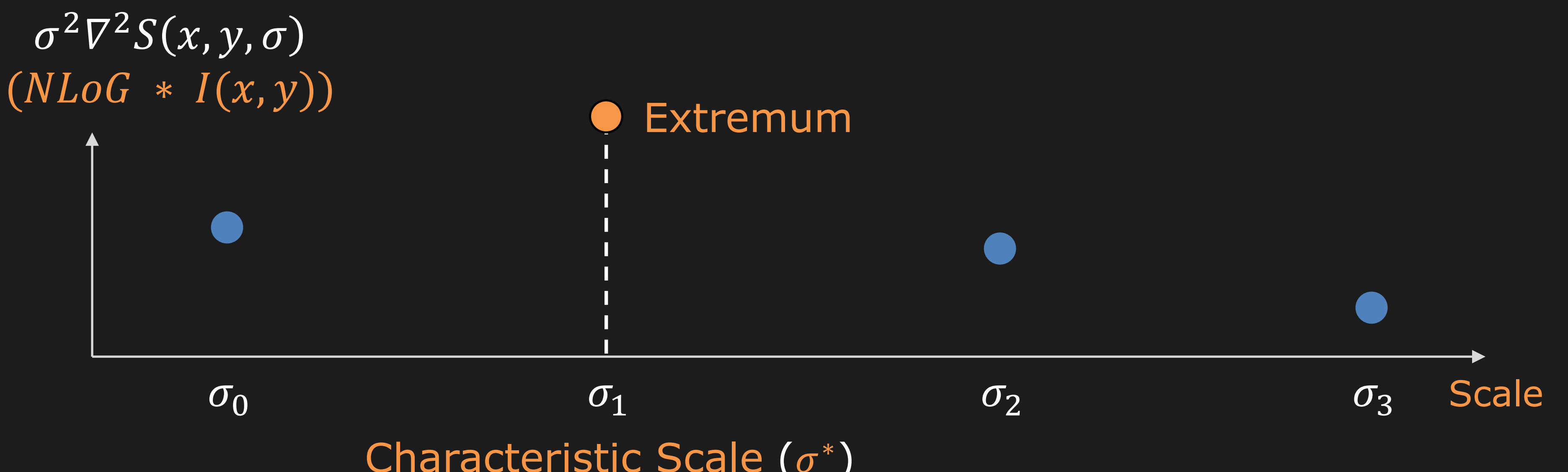
where, s: Constant multiplier

σ_0: Initial Scale

Blob Detection using Local Extrema

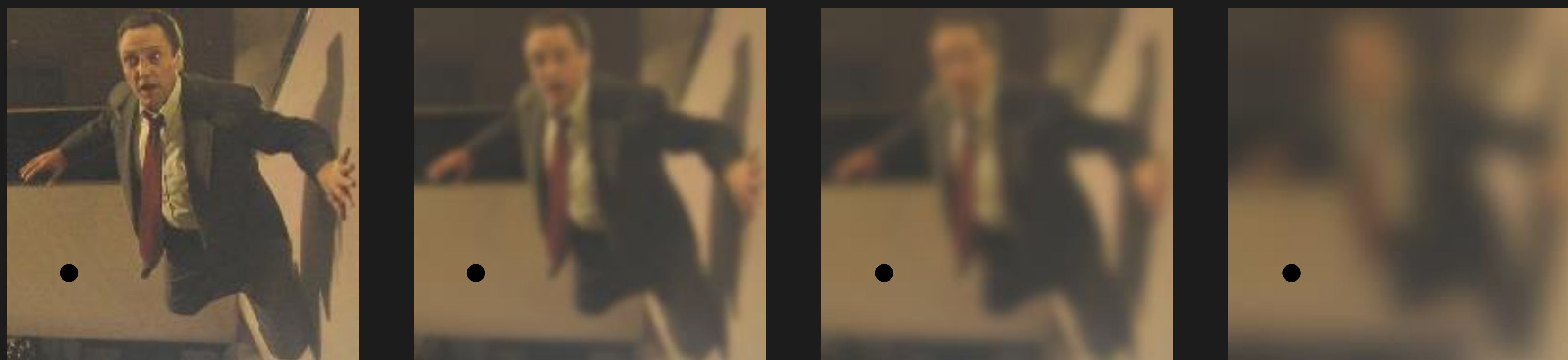


$S(x, y, \sigma_0)$ $S(x, y, \sigma_1)$ $S(x, y, \sigma_2)$ $S(x, y, \sigma_3)$ ...



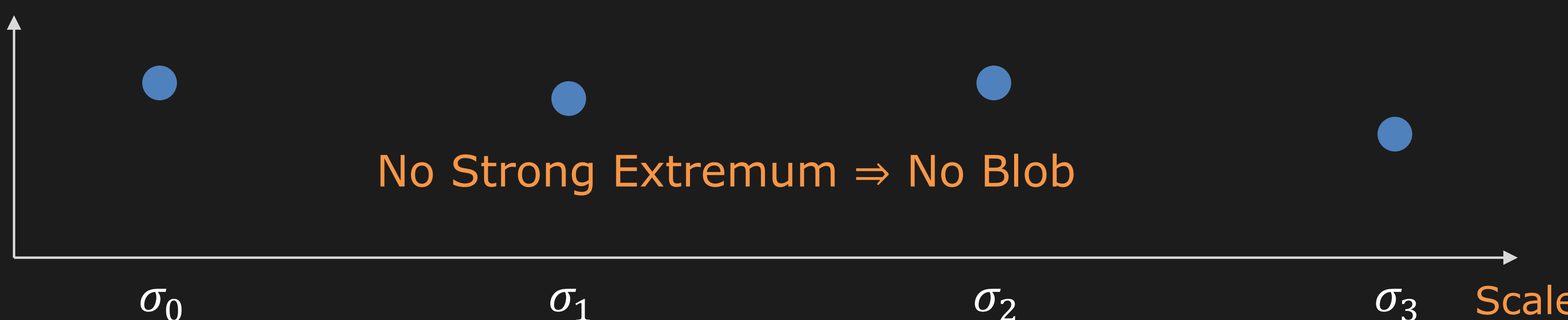
[Lindeberg 1994]

Blob Detection using Local Extrema



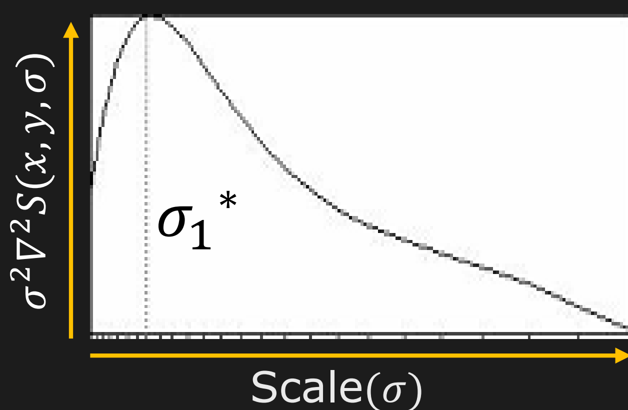
$S(x, y, \sigma_0)$ $S(x, y, \sigma_1)$ $S(x, y, \sigma_2)$ $S(x, y, \sigma_3)$ ...

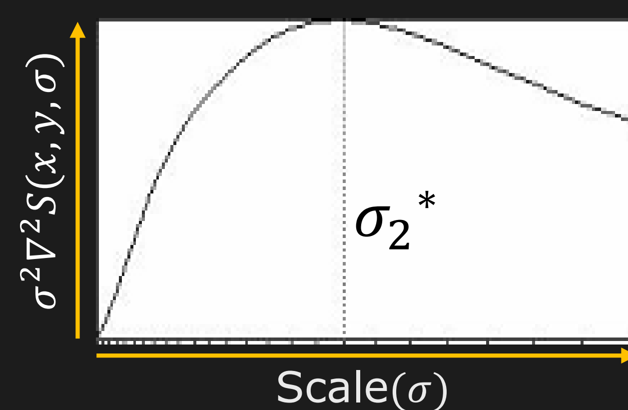
$\sigma^2 \nabla^2 S(x, y, \sigma)$
$(NLoG * I(x, y))$



[Lindeberg 1994]

Comparison of Characteristic Scales





$\dfrac{\sigma_1{}^*}{\sigma_2{}^*}$: 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:

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

or

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

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

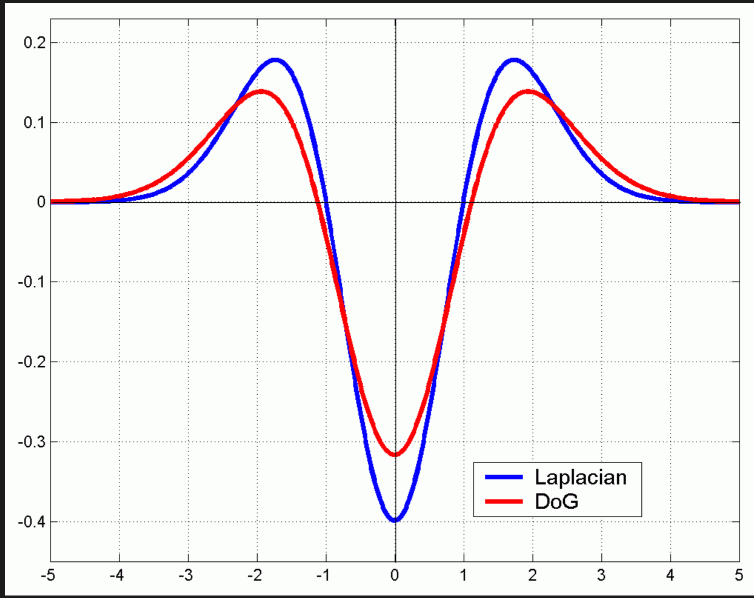
σ^*: Size of the blob

The SIFT Detector

An Efficient Implementation of Blob Detector

Uses Difference of Gaussian (DoG) as an approximation of NLoG

Difference of Gaussian(DoG) $= (n_{s\sigma} - n_{\sigma}) \approx (s-1)\sigma^2 \nabla^2 n_{\sigma}$



$\boxed{\text{DoG} \approx \text{NLoG}}$

Extracting SIFT Interest Points

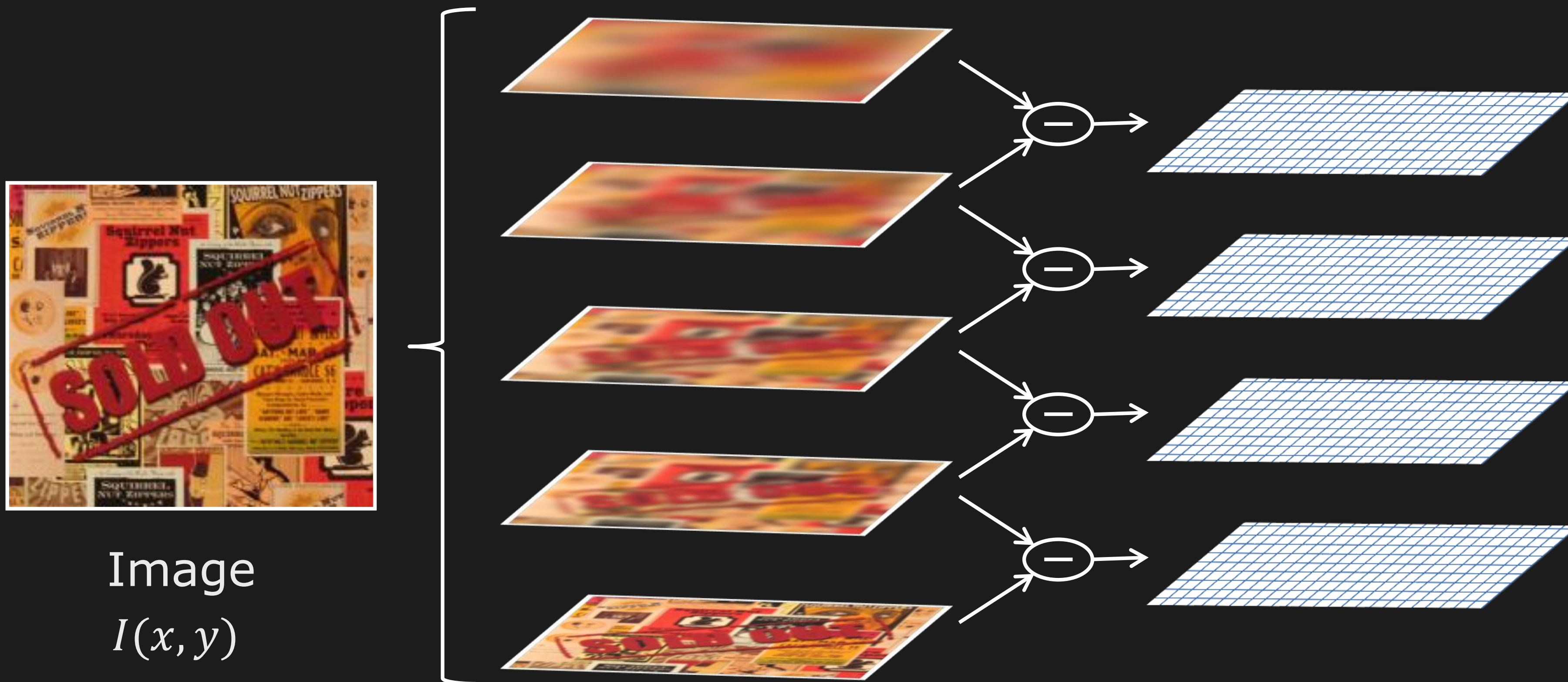
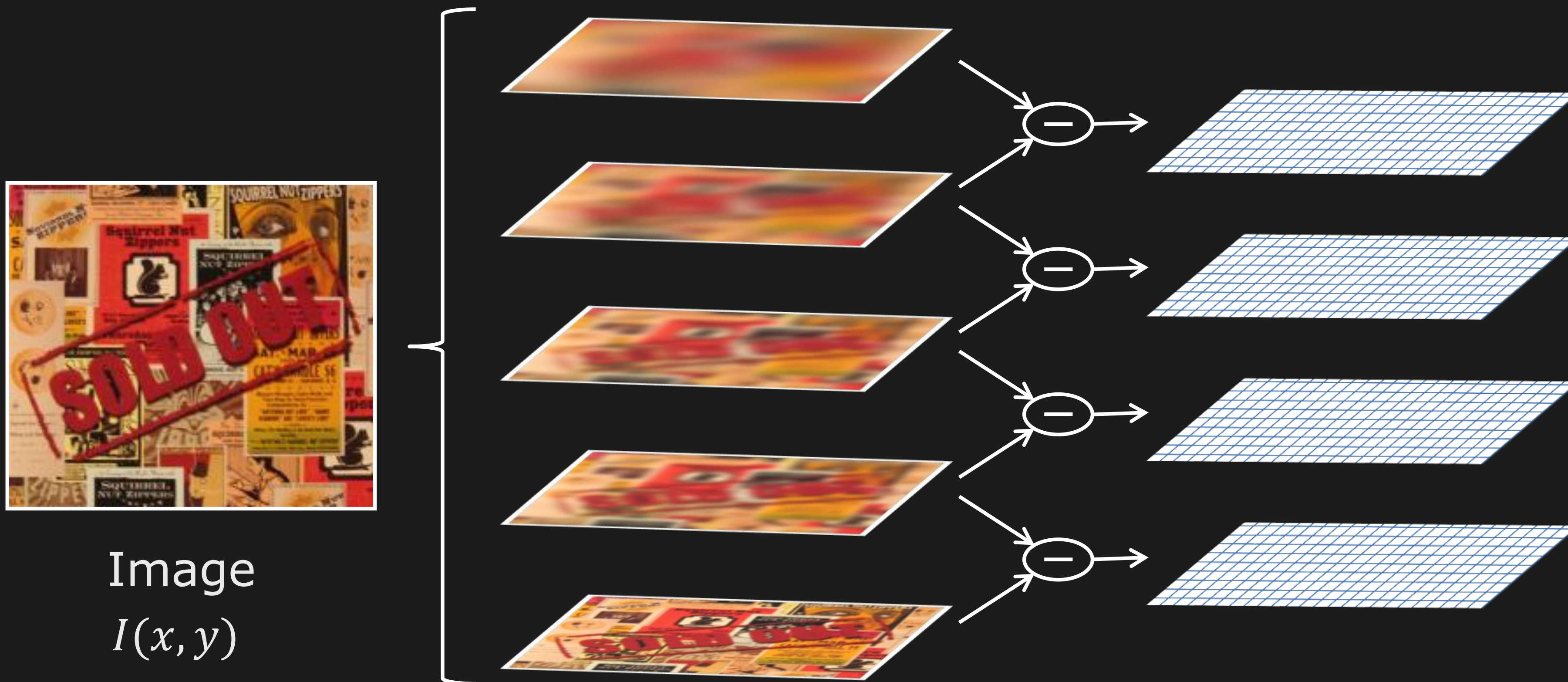


Image
$I(x, y)$

Gaussian
Scale-Space
$S(x, y, \sigma)$

Difference of
Gaussians (DoG)
$\approx \sigma^2 \nabla^2 S(x, y, \sigma)$

Extracting SIFT Interest Points



Image

$I(x, y)$

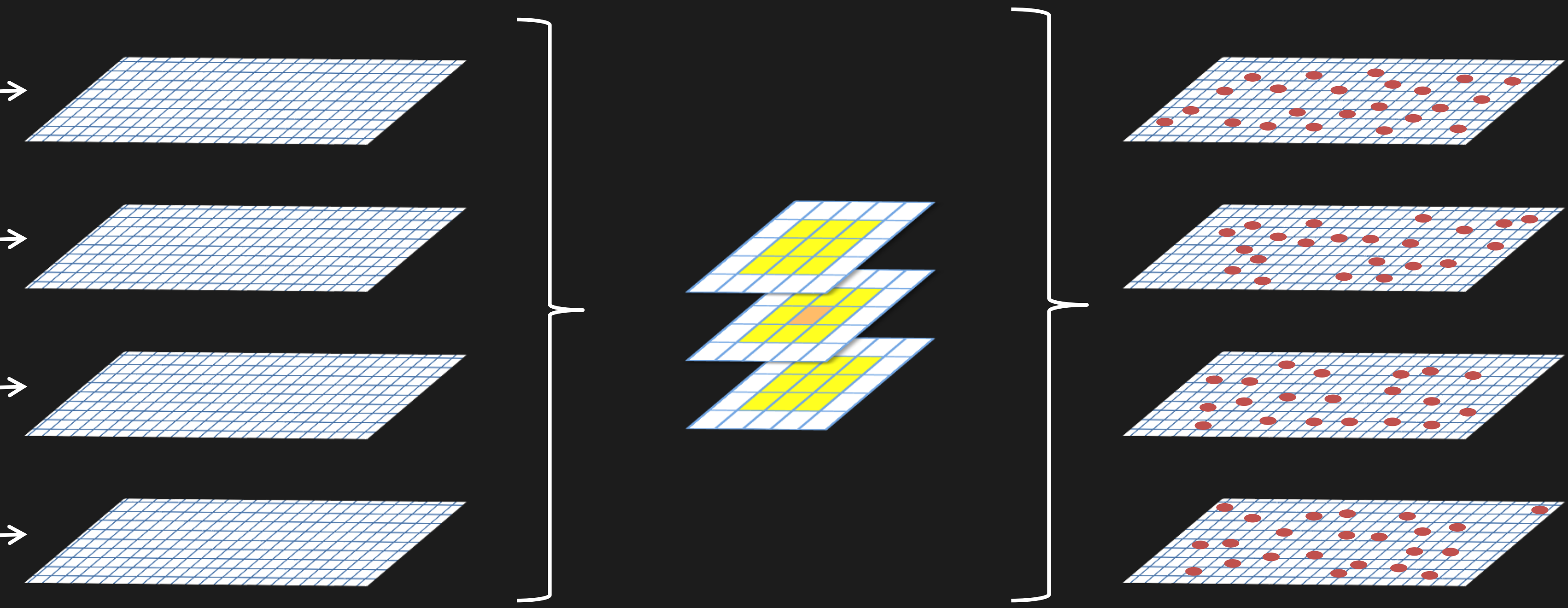
Gaussian Scale-Space

$S(x, y, \sigma)$

Difference of Gaussians (DoG)

$\sim\sigma^2 \nabla^2 S(x, y, \sigma)$

Extracting SIFT Interest Points



Difference of Gaussians (DoG)

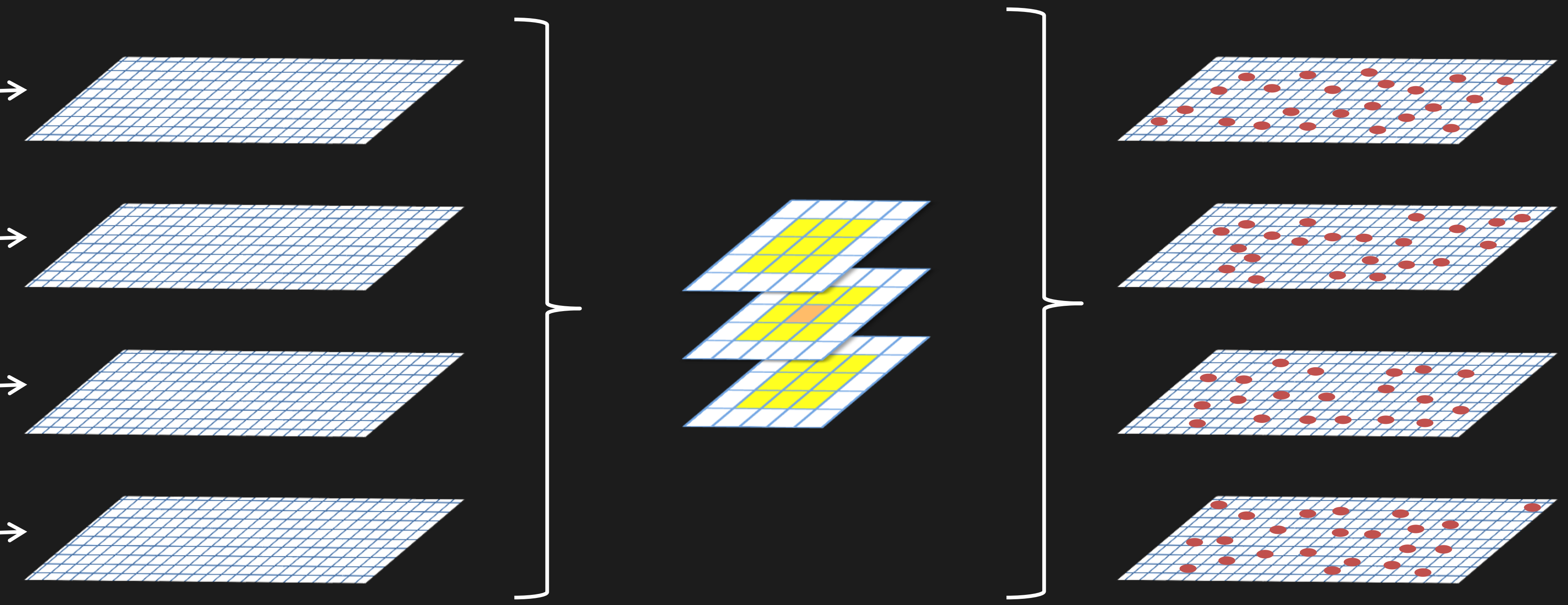
$\sim\sigma^2\nabla^2 S(x, y, \sigma)$

Find Extrema in every 3x3x3 grid

Interest Point Candidates

(includes weak extrema, bad contrast, …)

Extracting SIFT Interest Points



Difference of Gaussians (DoG)

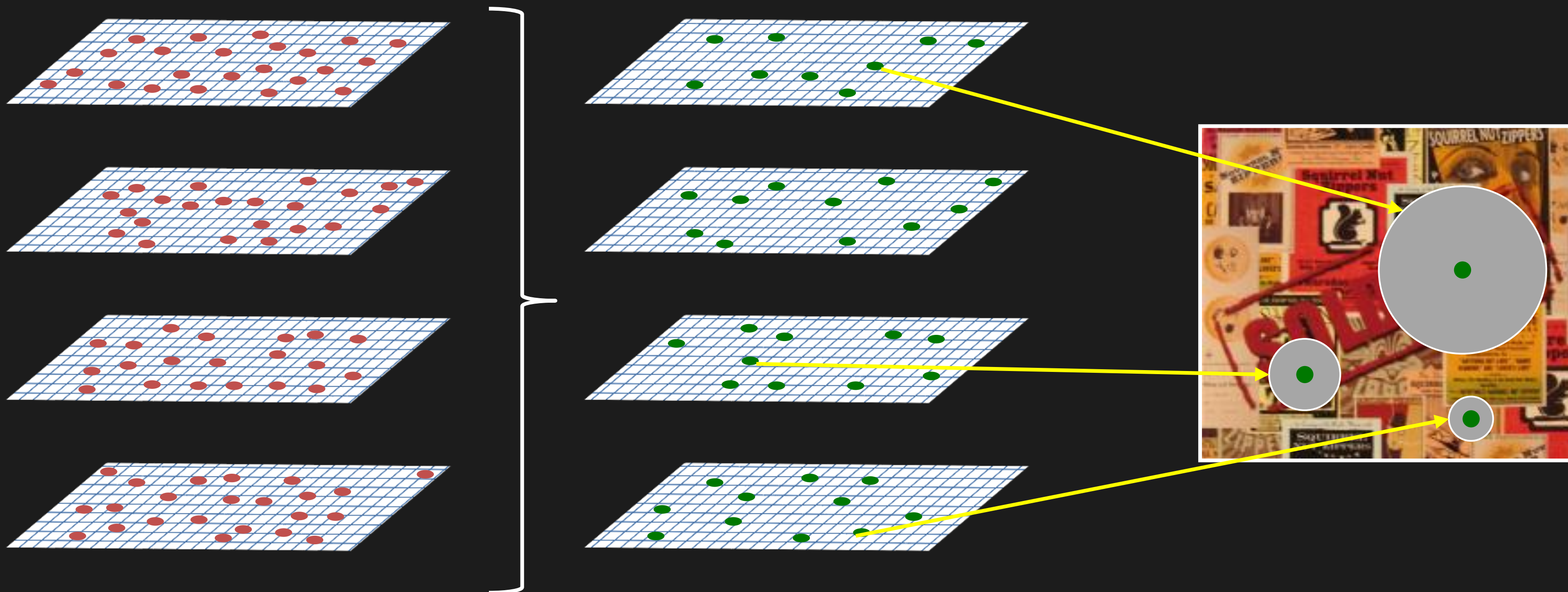
$\sim\sigma^2\nabla^2 S(x, y, \sigma)$

Find Extrema in every 3x3x3 grid

Interest Point Candidates

(includes weak extrema, bad contrast, …)

Extracting SIFT Interest Points



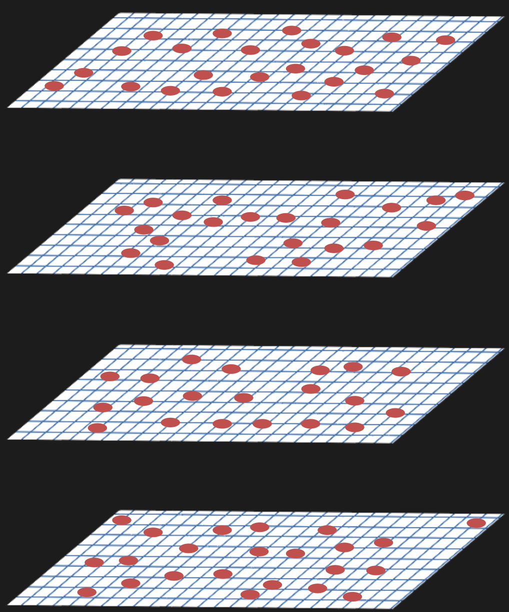
Interest Point Candidates

(includes weak extrema, bad contrast, …)

SIFT Interest Points

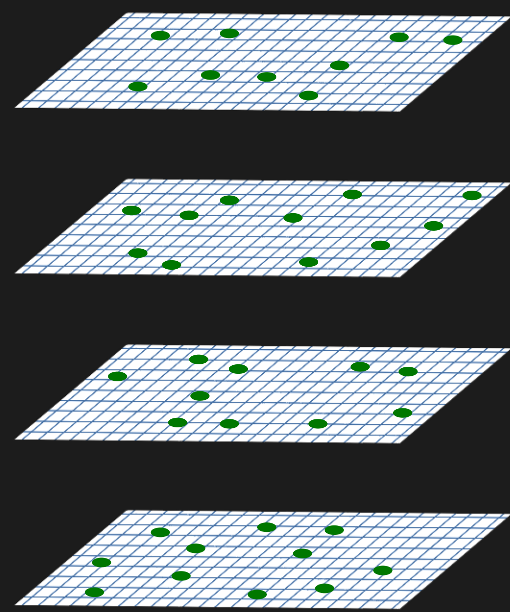
(after removing weak points)

Extracting SIFT Interest Points



Interest Point Candidates

(includes weak extrema, bad contrast, …)

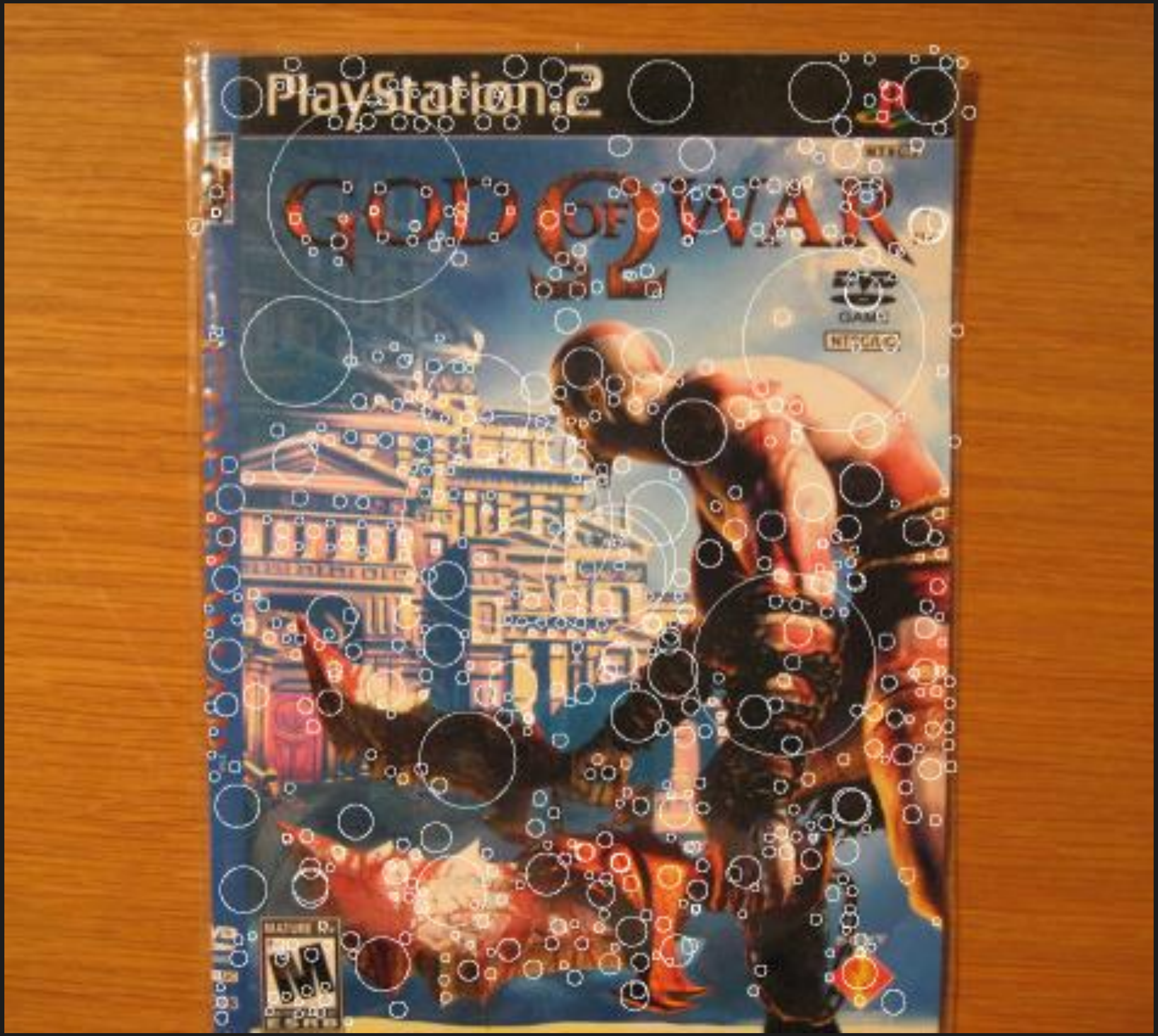


SIFT Interest Points

(after removing weak points)

Interest Point Depiction

SIFT Detection Examples

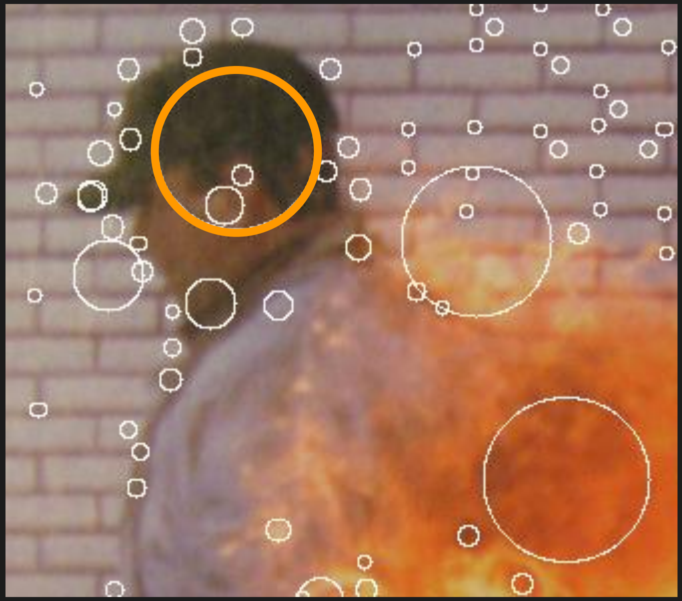


SIFT Detection Examples

SIFT Detection Examples

SIFT Scale Invariance

Use the characteristic scales to match sizes



Computing the Principal Orientation

Use the histogram of gradient directions

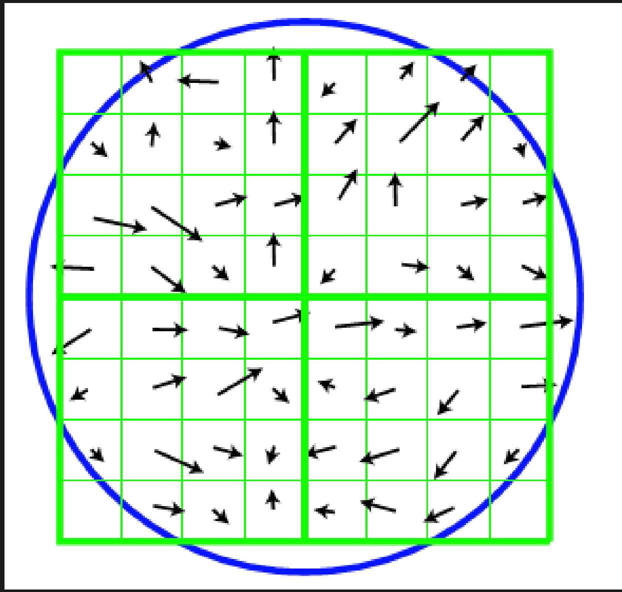
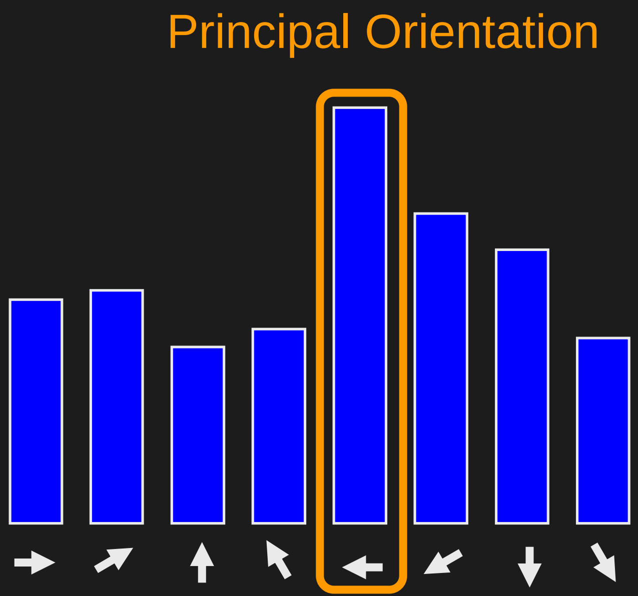


Image gradient directions

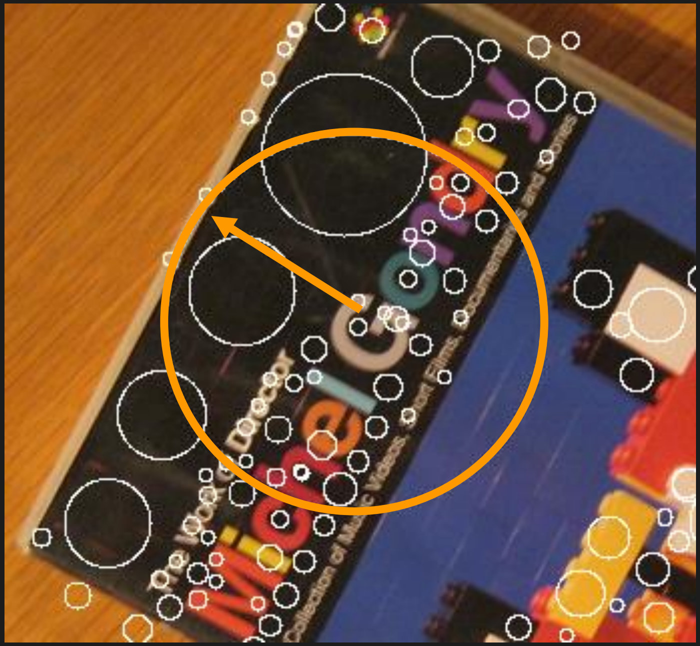
$$\theta = \tan^{-1}\left(\frac{\partial I}{\partial y} / \frac{\partial I}{\partial x}\right)$$



Choose the most prominent gradient direction

SIFT Rotation Invariance

Use the principal orientations to match rotation

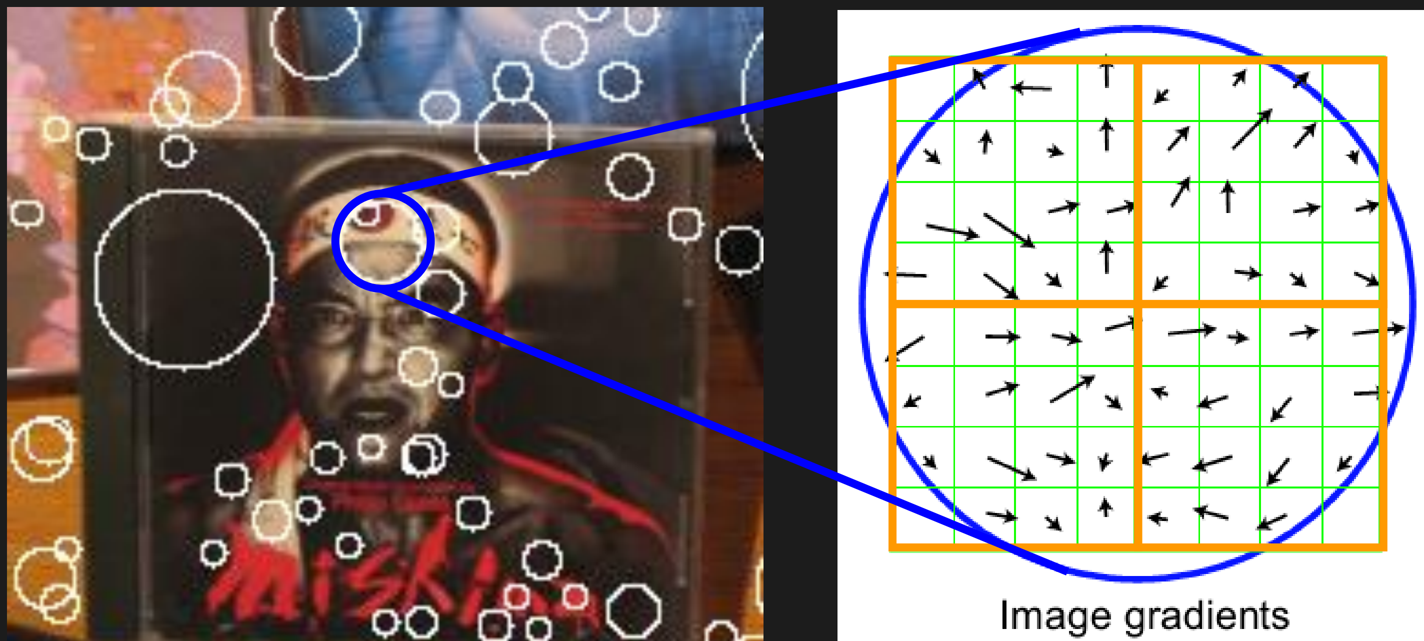


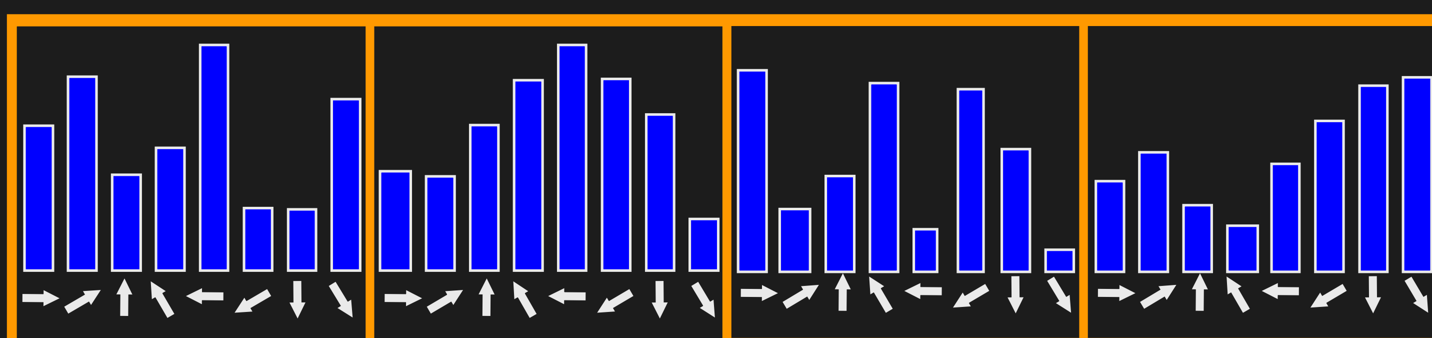
The SIFT Descriptor

"Describe" points so they can be compared

Computing the SIFT Descriptor

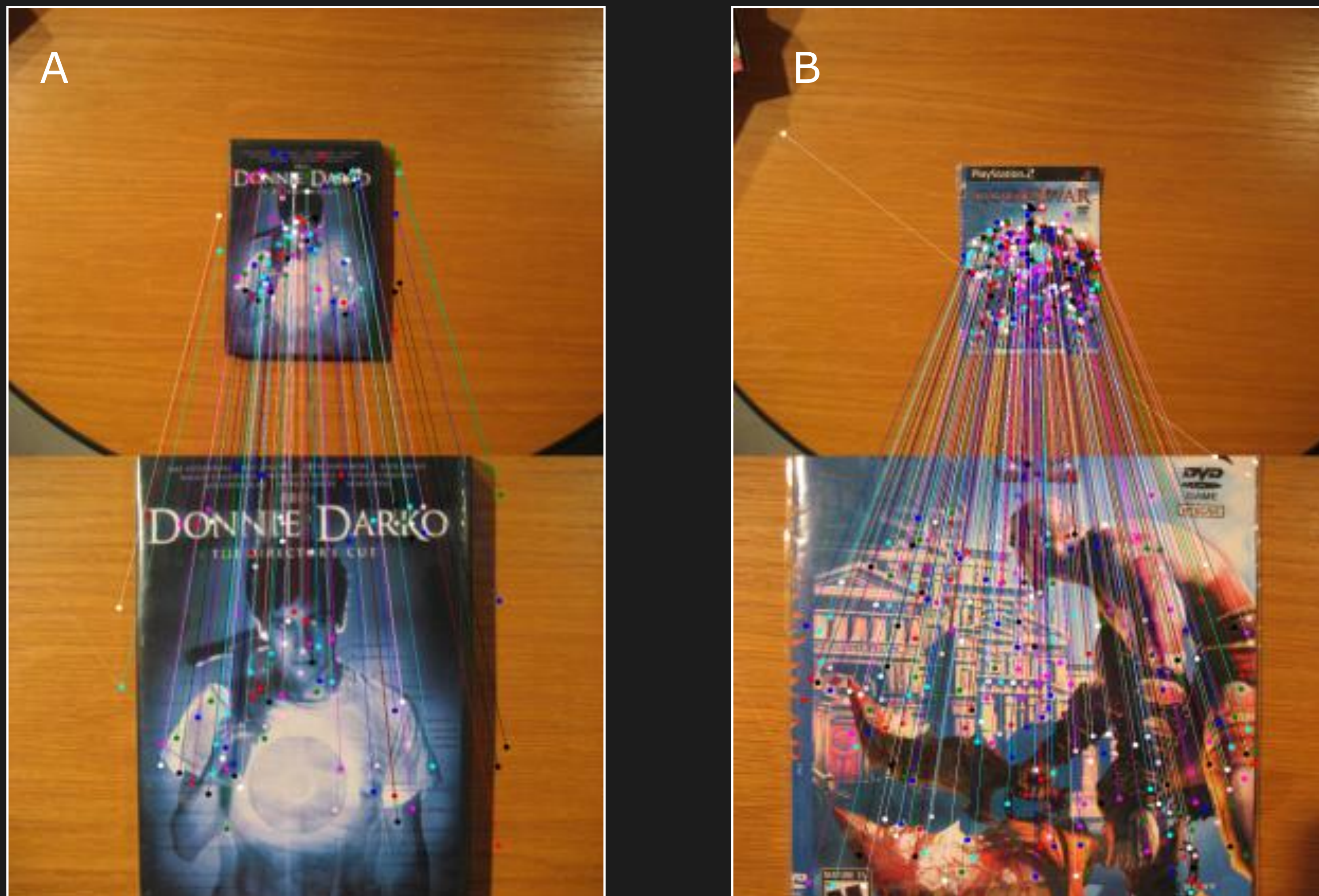
Histograms of gradient directions over spatial regions





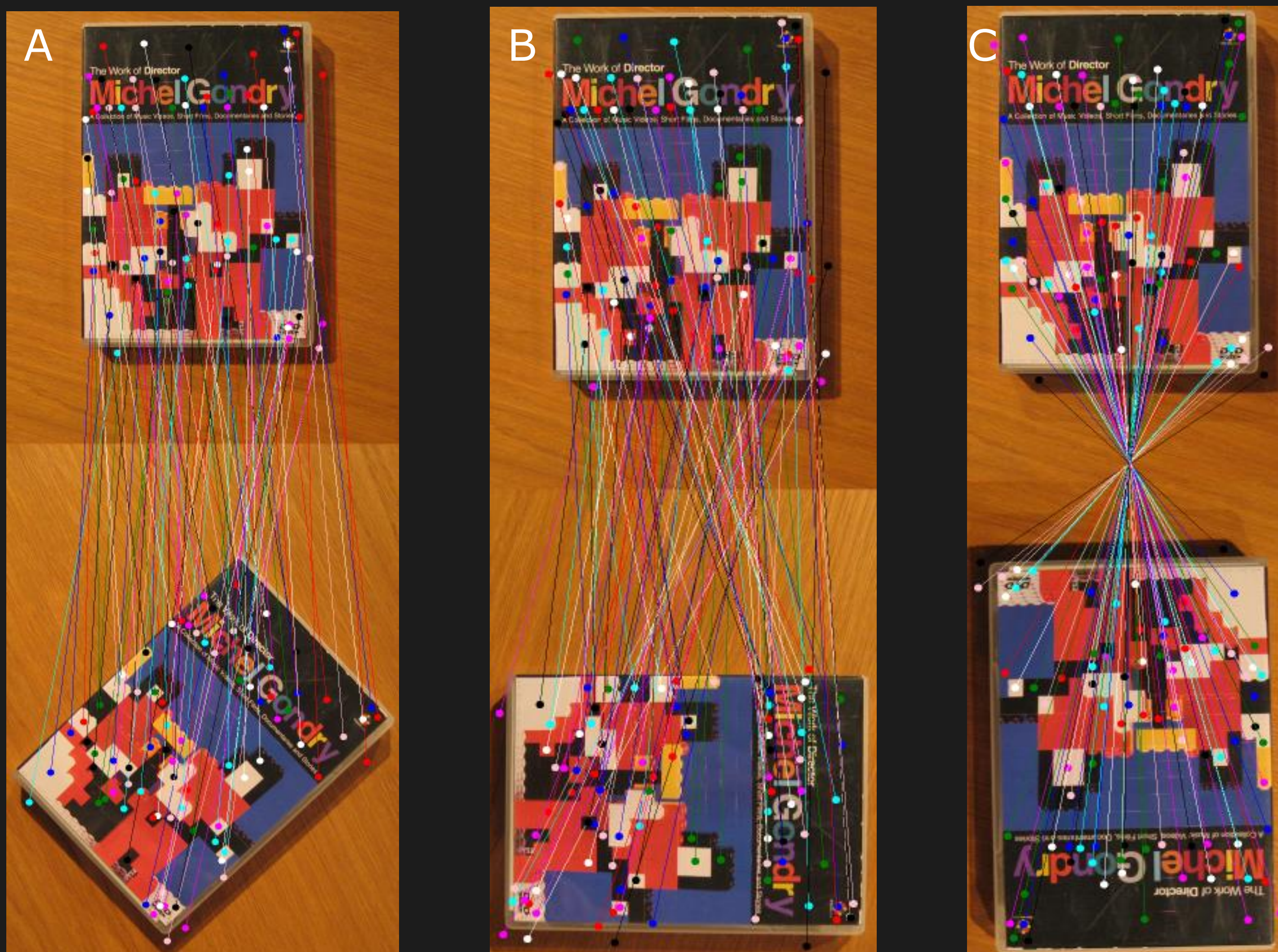
Invariant to Scale, Lighting, Brightness

SIFT Results: Scale Invariance



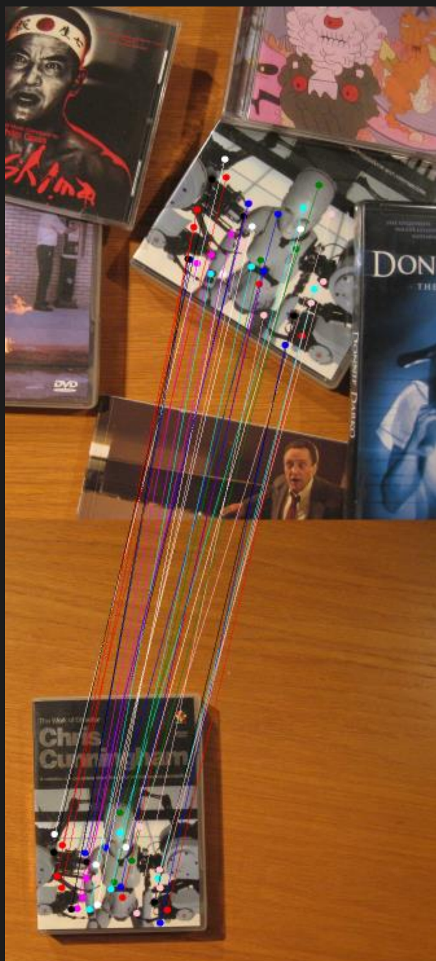
SIFT detects corresponding features in images at different resolutions

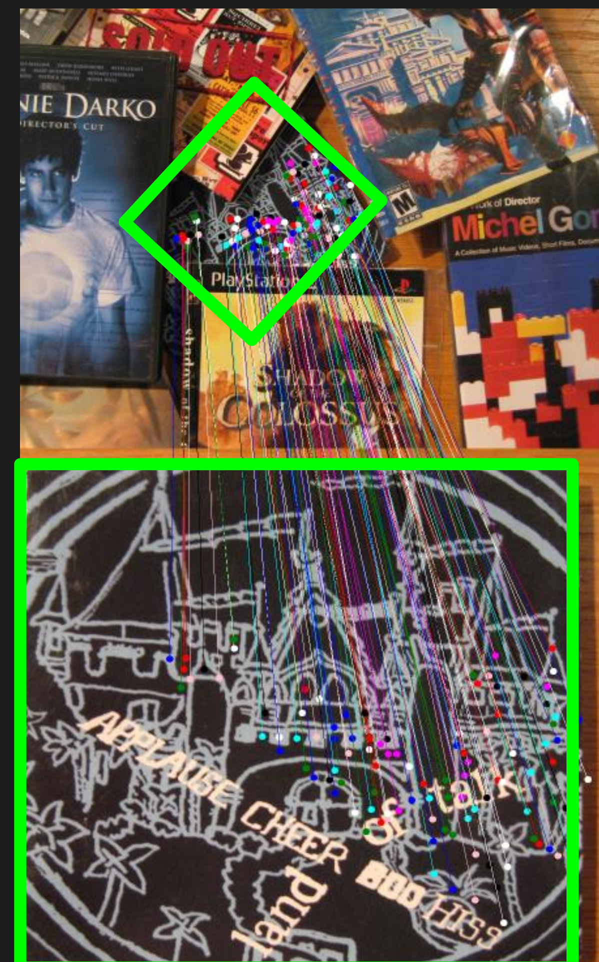
SIFT Results: Rotation Invariance



SIFT detects corresponding features in rotated images

SIFT Robustness to Clutter

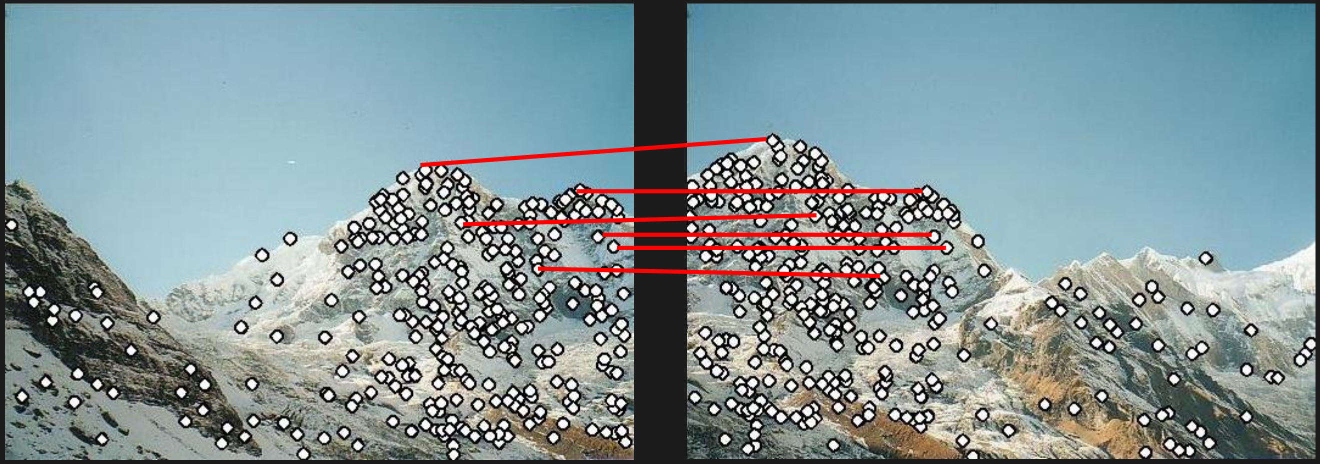




Panorama Stitching using SIFT

Image 1

Image 2



Match SIFT Interest Points

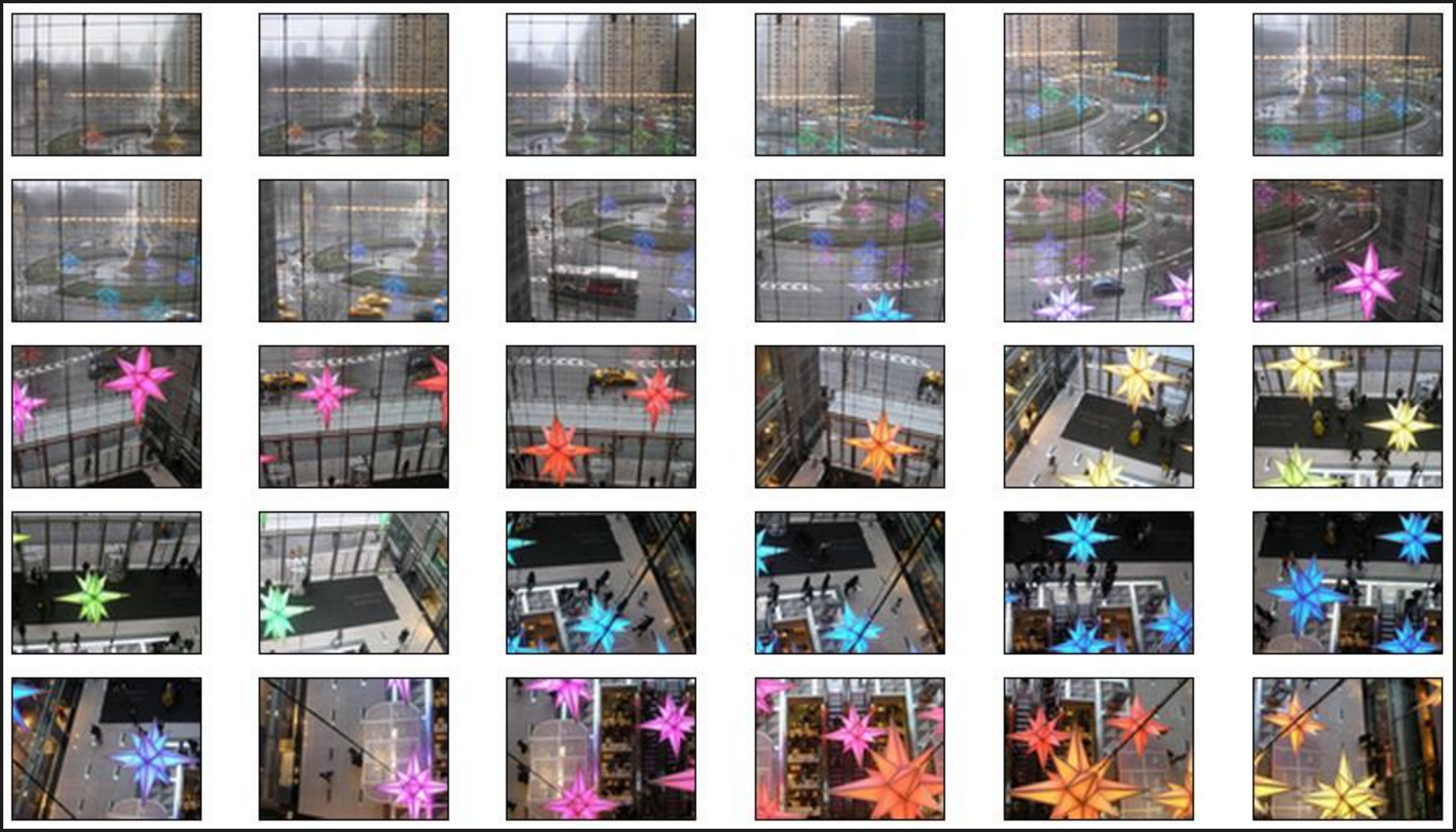
[Autopano]

Panorama Stitching using SIFT

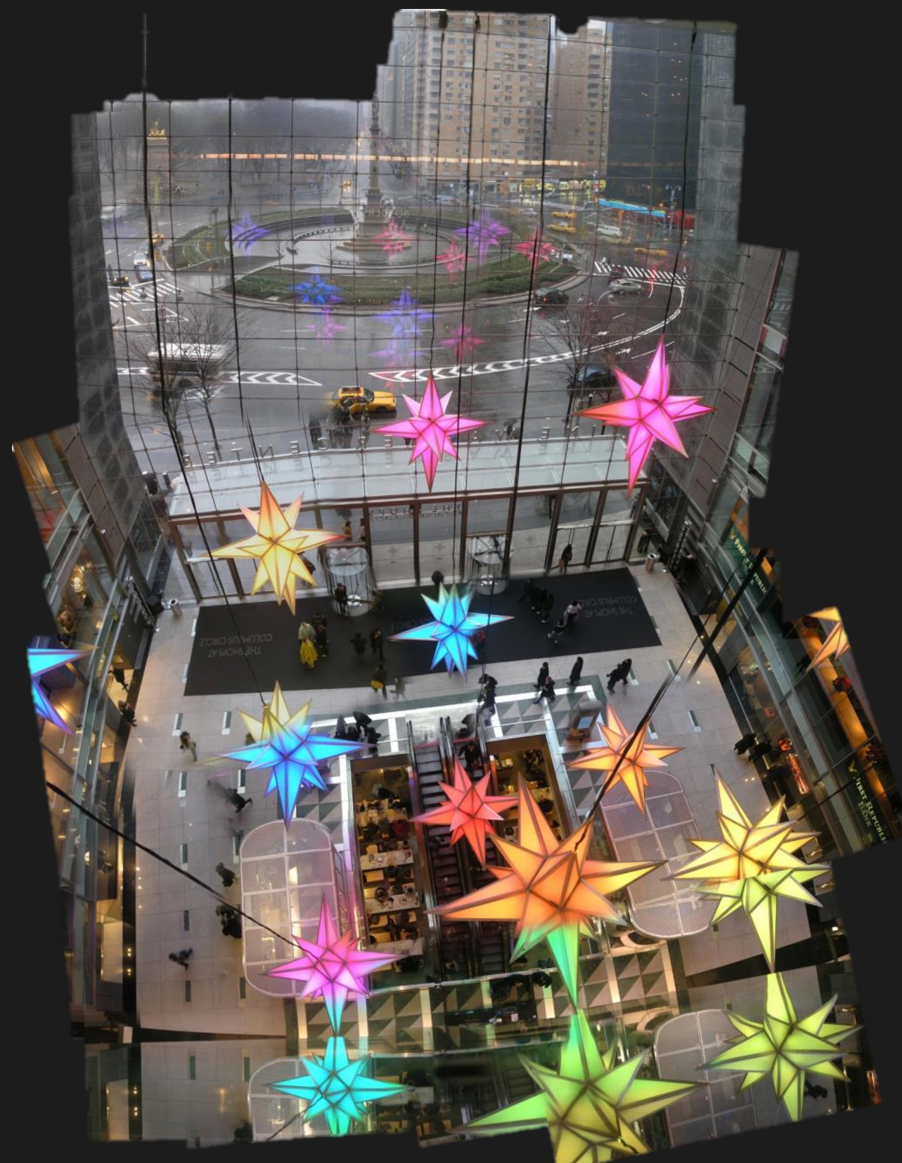
Transform/Warp one or both images so that corresponding SIFT points in images are aligned.

[Autopano]

Auto Collage using SIFT

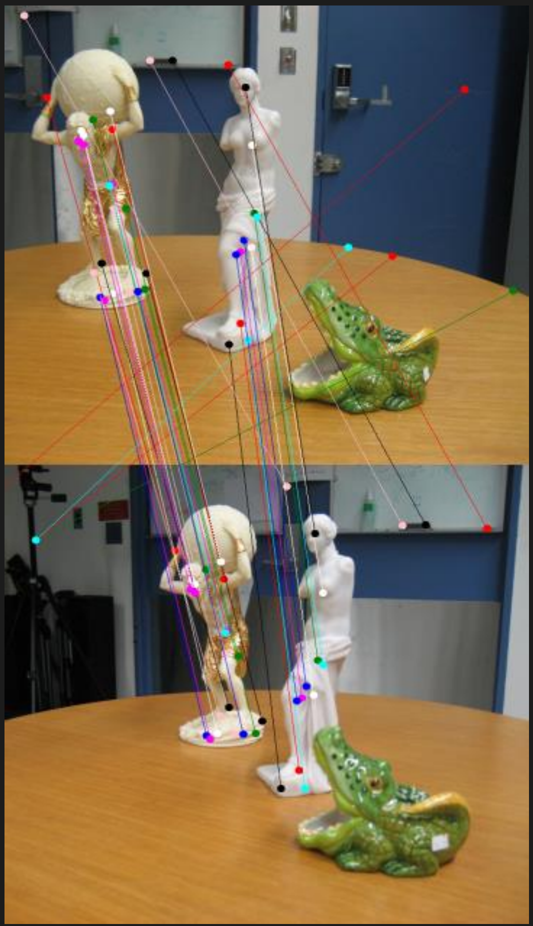


Auto Collage using SIFT



SIFT for 3D Objects?

No Change in Viewpoint



30° Change in Viewpoint

90° Change in Viewpoint

References

[Autopano] Software to make panaromas using SIFT. http://user.cs.tu-berlin.de/~nowozin/autopano-sift/

[Brown and Lowe 2002] M. Brown and D. Lowe. "Invariant Features from Interest Point Groups". *BMVC*, 2002.

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

[Lowe 2004] D. Lowe."Distinctive Image Features from Scale-Invariant Keypoints". *IJCV*, 2004.

[Lindeberg 1994] T. Lindeberg. "Scale-Space Theory: A Basic Tool for Analysing Structures at Different Scales." *J. of Applied Statistics,* 1994.

[Matas 2002] J. Matas, O. Chum, M. Urban, and T. Pajdla. "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. *BMVC*, 2002.

[Mikolajczyk 2002] K. Mikolajczyk. "Detection of Local Features Invariant to Affine Transformations." *Ph.D. Thesis*, 2002.

References

[Mikolajczyk 2004] K. Mikolajczyk and C. Schmid. "Scale and Affine Invariant Interest Point Detectors." *IJCV,* 2004.

[Mikolajczyk 2005] K. Mikolajczyk and C. Schmid. "A Performance Evaluation of Local Descriptors." *PAMI,* 2005.

[SIFT] SIFT Binaries. http://www.cs.ubc.ca/~lowe/keypoints/

[Witkin 1983] A. Witkin. "Scale-Space Filtering". *IJCAI*, 1983.