SIFT

Scale-Invariant Feature Transform

Lowe, "Distinctive image features from scale-invariant keypoints", IJCV, 2004.

SIFT

Scale-Invariant Feature Transform

interest point detector + descriptor

Lowe, "Distinctive image features from scale-invariant keypoints", IJCV, 2004.

Object Recognition from Local Scale-Invariant Features

David G. Lowe

Computer Science Department
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada
lowe@cs.ubc.ca

Abstract

An object recognition system has been developed that uses a new class of local image features. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection.

translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. Previous approaches to local feature generation lacked invariance to scale and were more sensitive to projective distortion and illumination change. The SIFT features share a number of properties in common with the

Object Recognition from Local Scale-Invariant Features

David G. Lowe
Computer Science Department
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada
lowe@cs.ubc.ca

Abstract

An object recognition system has been developed that uses a new class of local image features. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection.

translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. Previous approaches to local feature generation lacked invariance to scale and were more sensitive to projective distortion and illumination change. The SIFT features share a number of properties in common

“I did submit papers on earlier versions of SIFT to both ICCV and CVPR (around 1997/98) and both were rejected. I then added more of a systems flavor and the paper was published at ICCV 1999, but just as a poster.”

— David Lowe

Object Recognition from Local Scale-Invariant Features

David G. Lowe

Computer Science Department
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada
lowe@cs.ubc.ca

Abstract

An object recognition system has been developed that uses a new class of local image features. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection.

translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. Previous approaches to local feature generation lacked invariance to scale and were more sensitive to projective distortion and illumination change. The SIFT features share a number of properties in common

Object Recognition from Local Scale-Invariant Features

David G. Lowe
Computer Science Department
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada
lowe@cs.ubc.ca

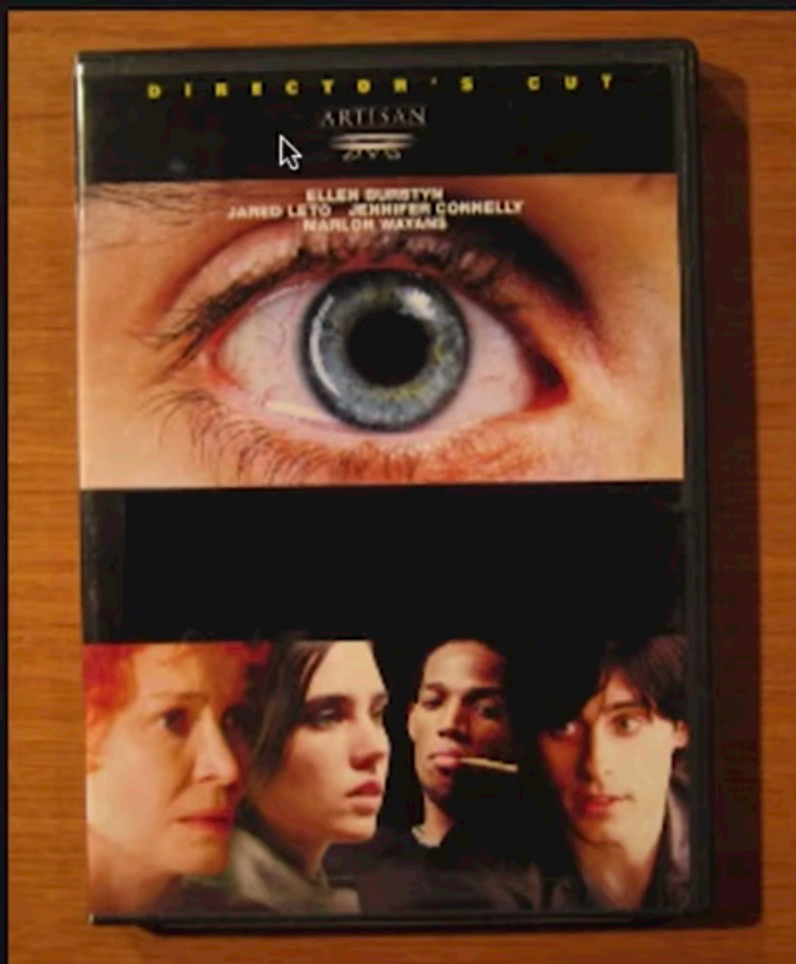
Abstract

An object recognition system has been developed that uses a new class of local image features. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection.

translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. Previous approaches to local feature generation lacked invariance to scale and were more sensitive to projective distortion and illumination change. The SIFT features share a number of

61+ Thousand

citations



Template

Rich 2D image

Raw Images are Hard to Match

Different size, orientation, lighting, brightness, etc.

What is an Interesting Point/Feature?

- Has rich image content (brightness variation, color variation, etc.) 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 insensitive to lighting changes

What stuff in an image matches with stuff in another?

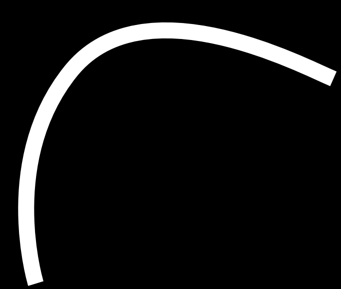

RU

Determine corresponding regions across images

Determine corresponding regions across images

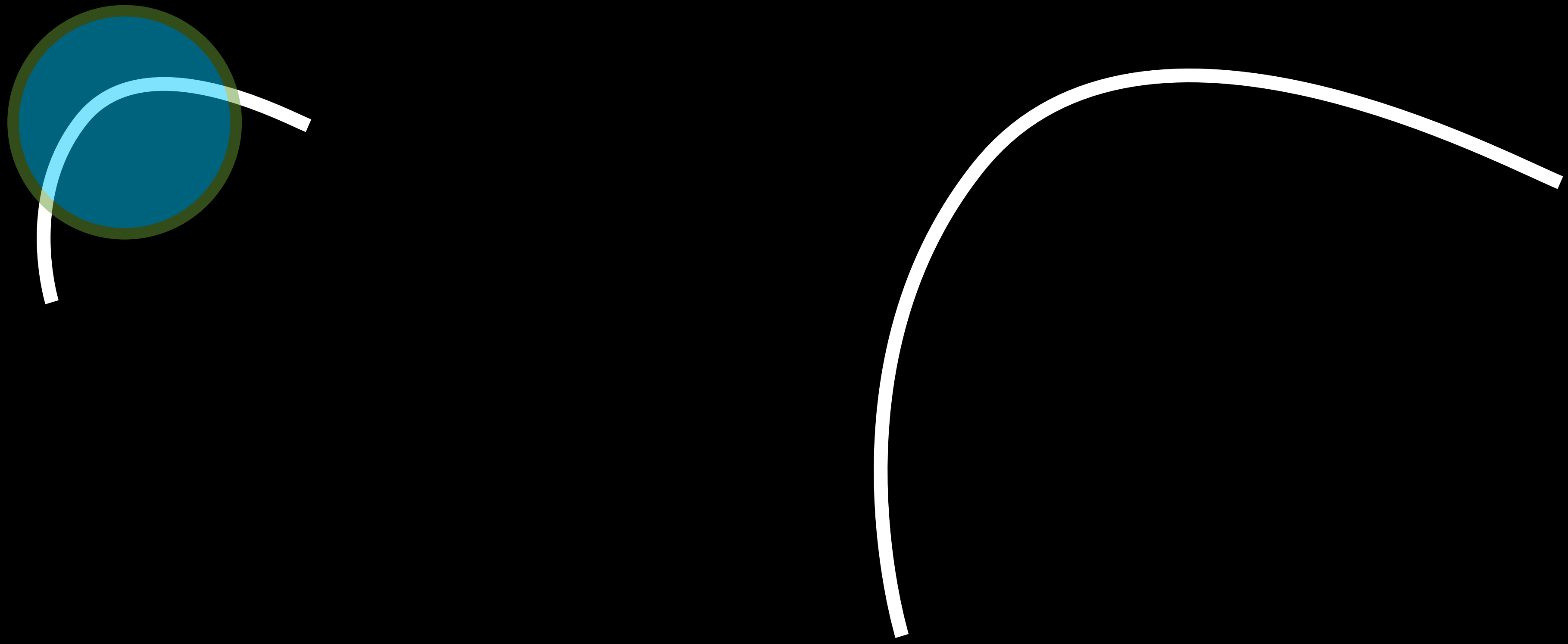
irrespective of image translation, scale and rotation

Scale Invariant Detection

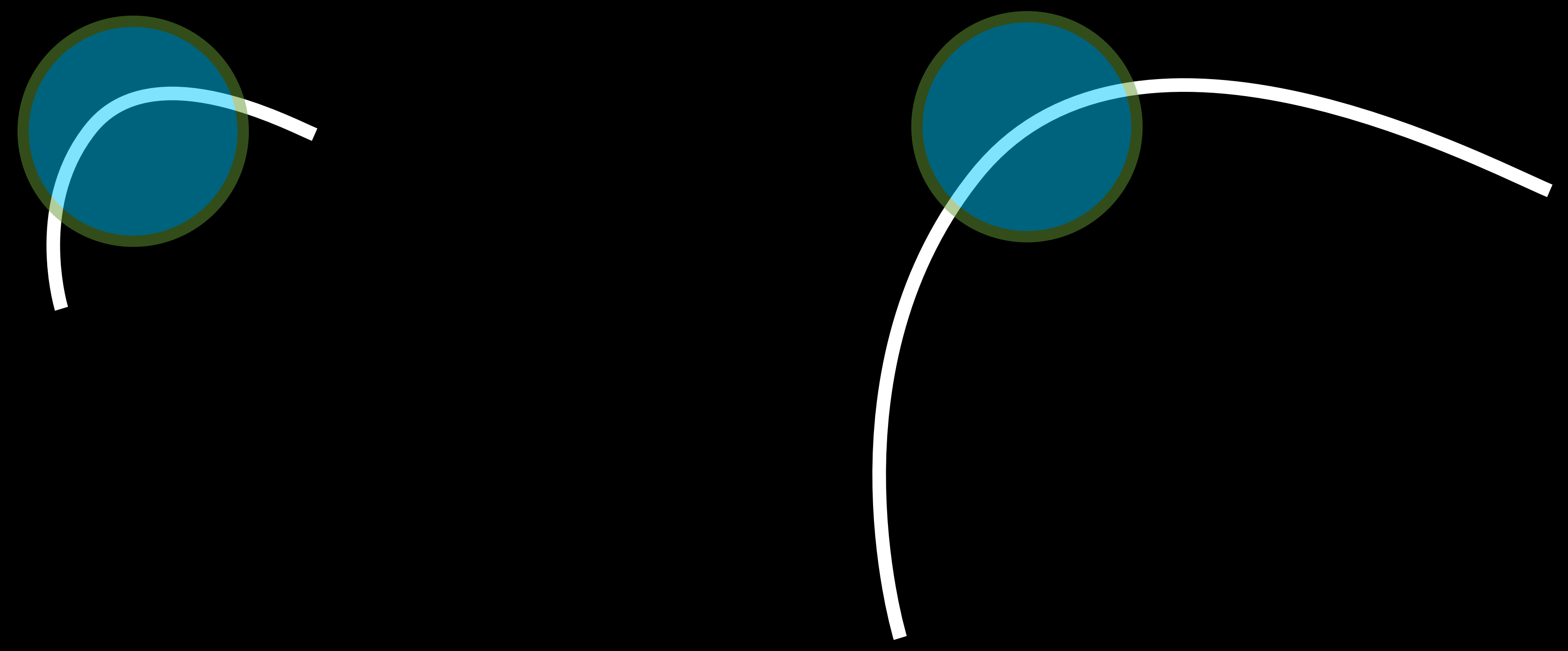


Scale Invariant Detection

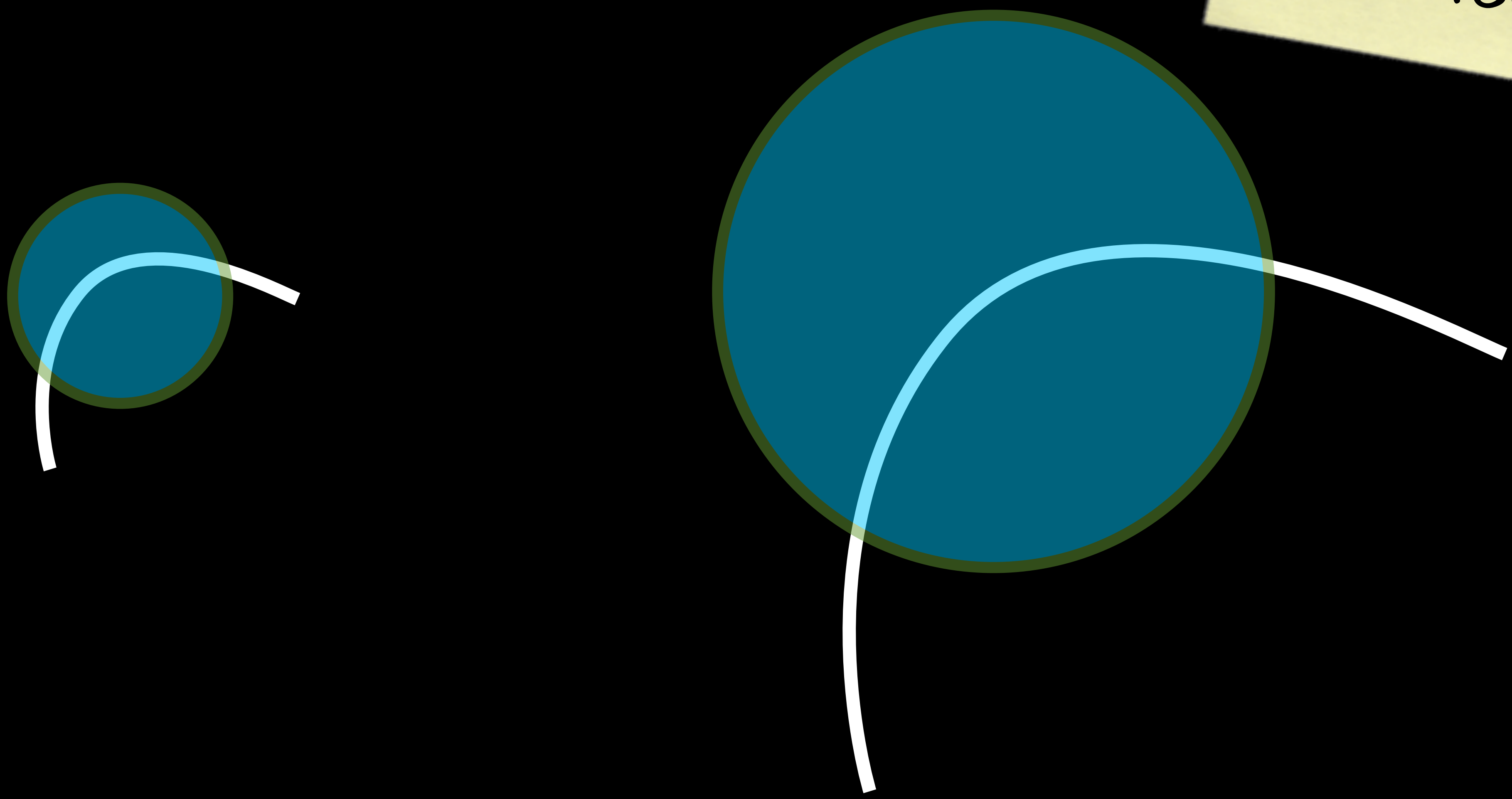
Scale Invariant Detection



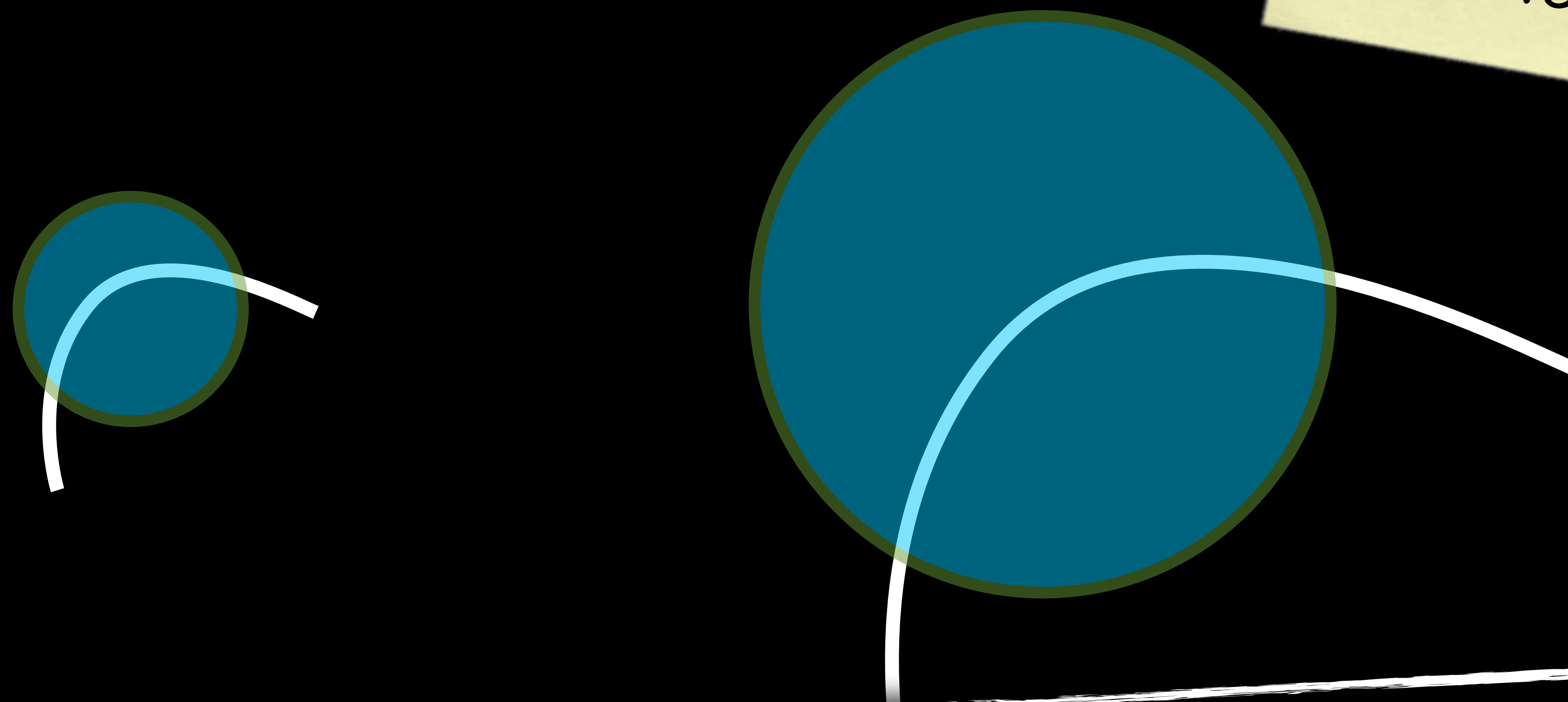
Scale Invariant Detection



Scale Invariant Detection

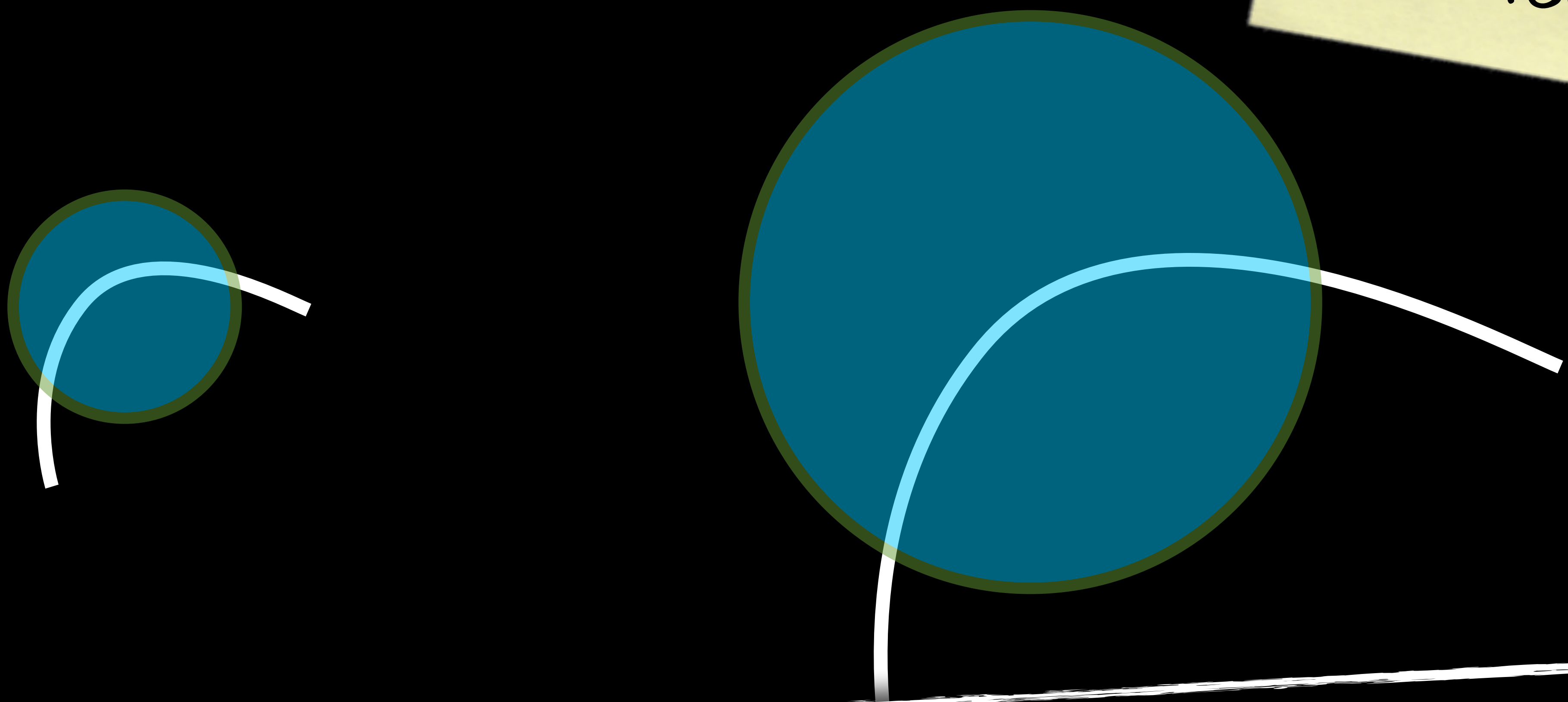


Scale Invariant Detection



How do we choose corresponding correct region sizes independently in each image?

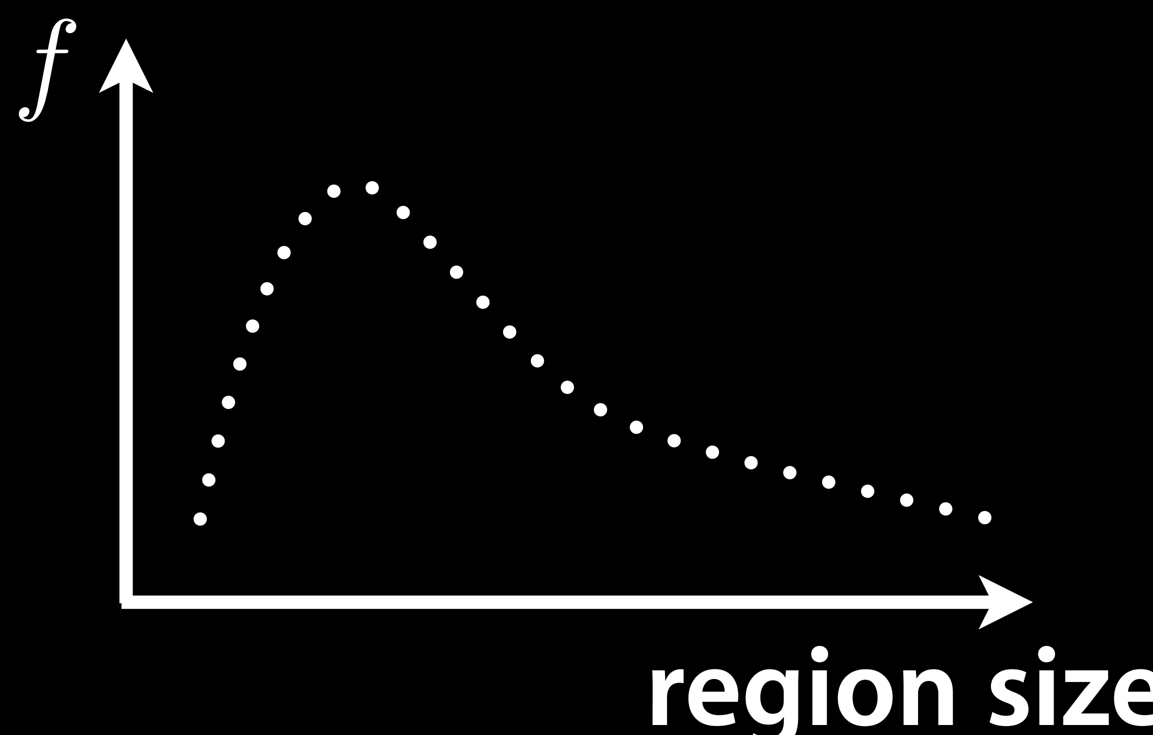
Scale Invariant Detection

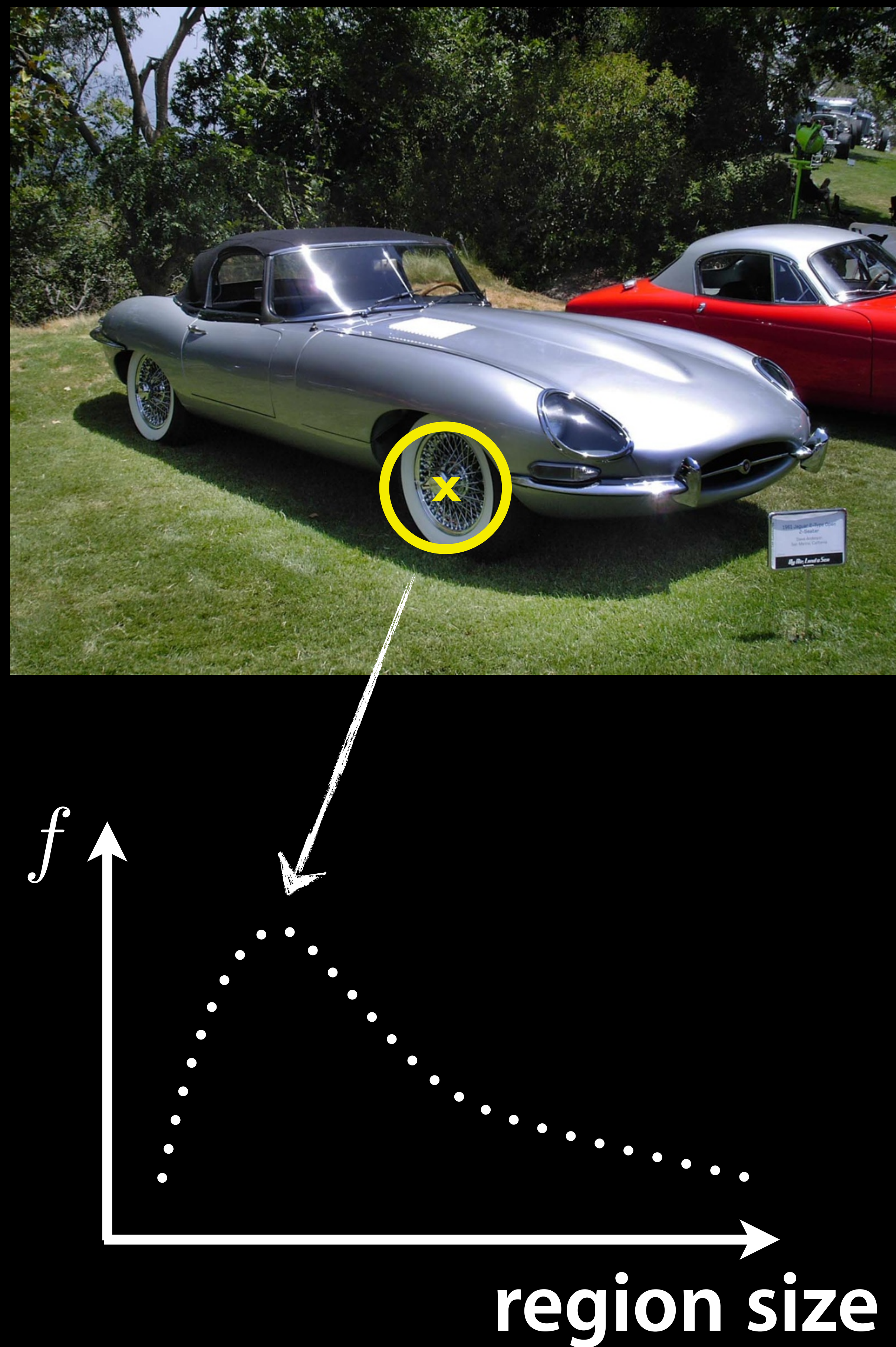


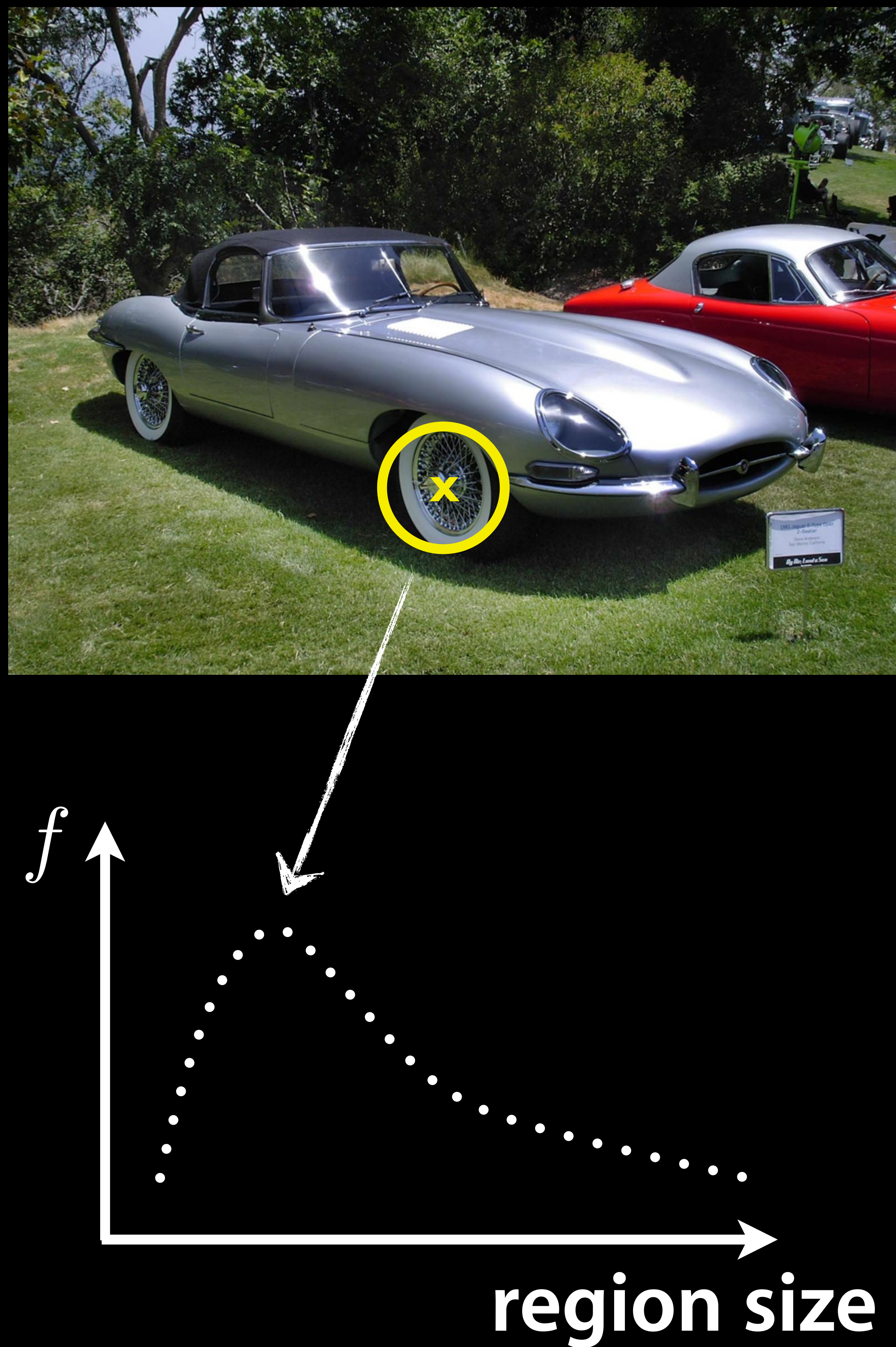
How do we choose corresponding correct region sizes **independently** in each image?

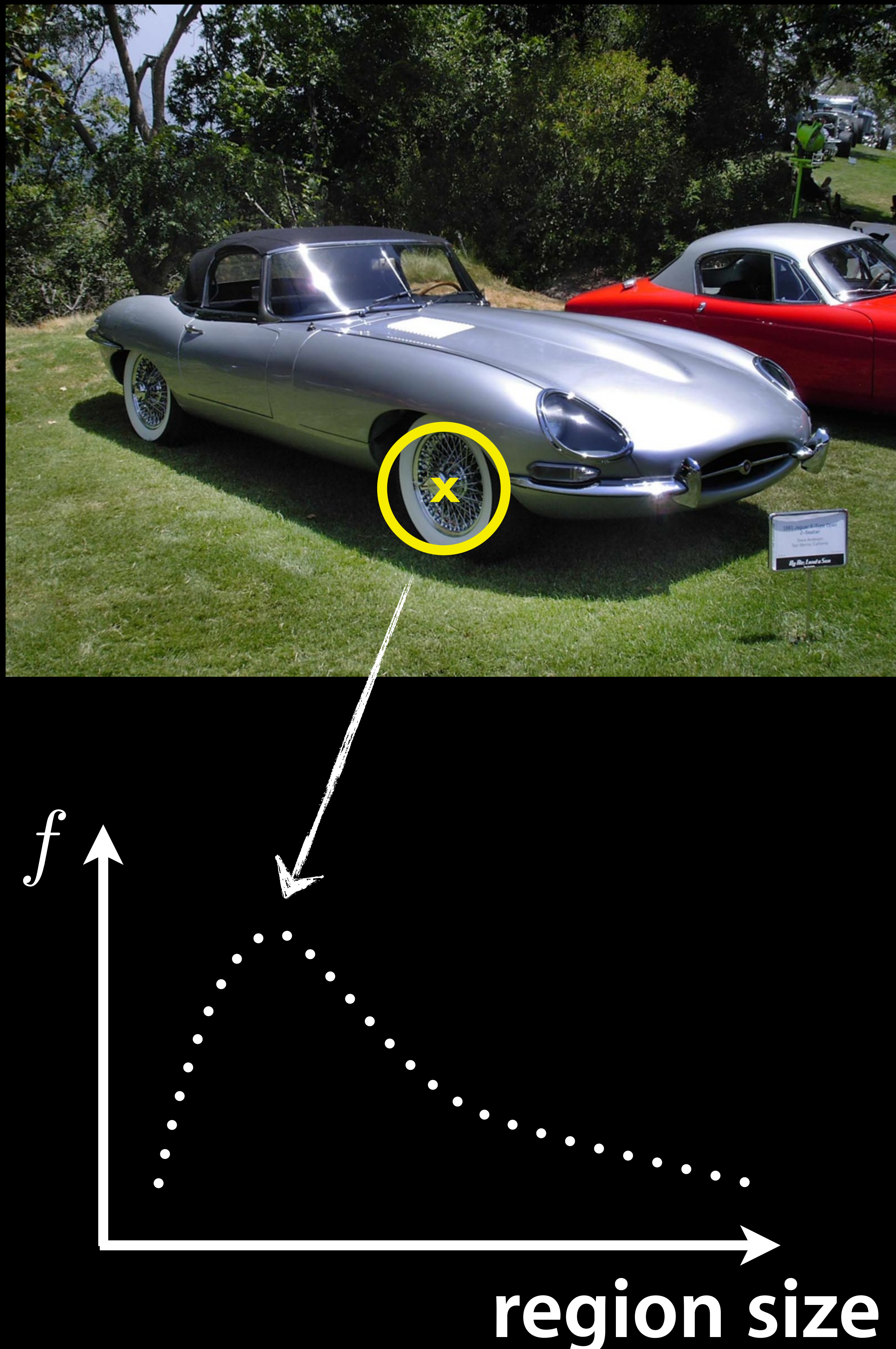
Design a **scale invariant function** on the region

f
region size

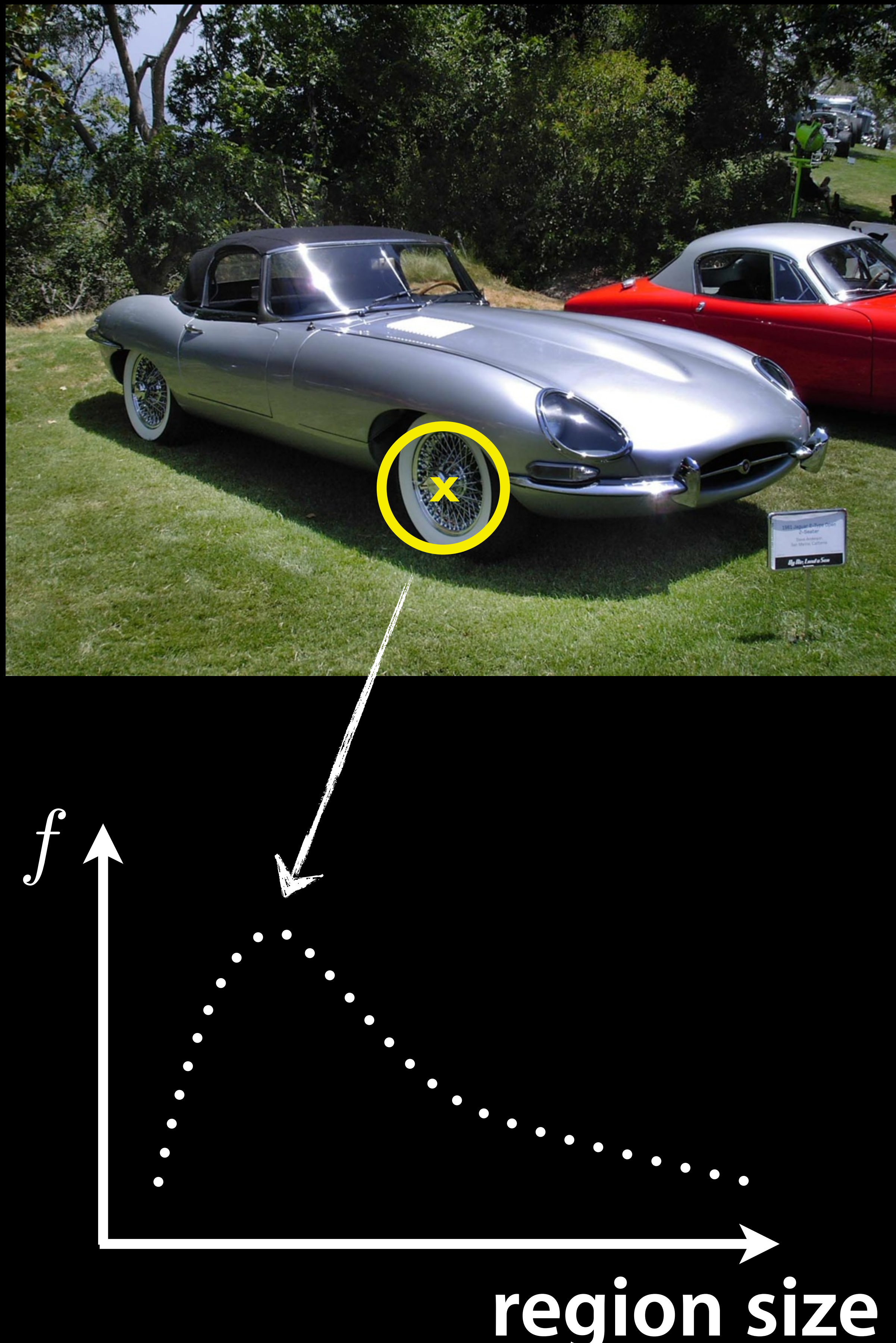

f
region size

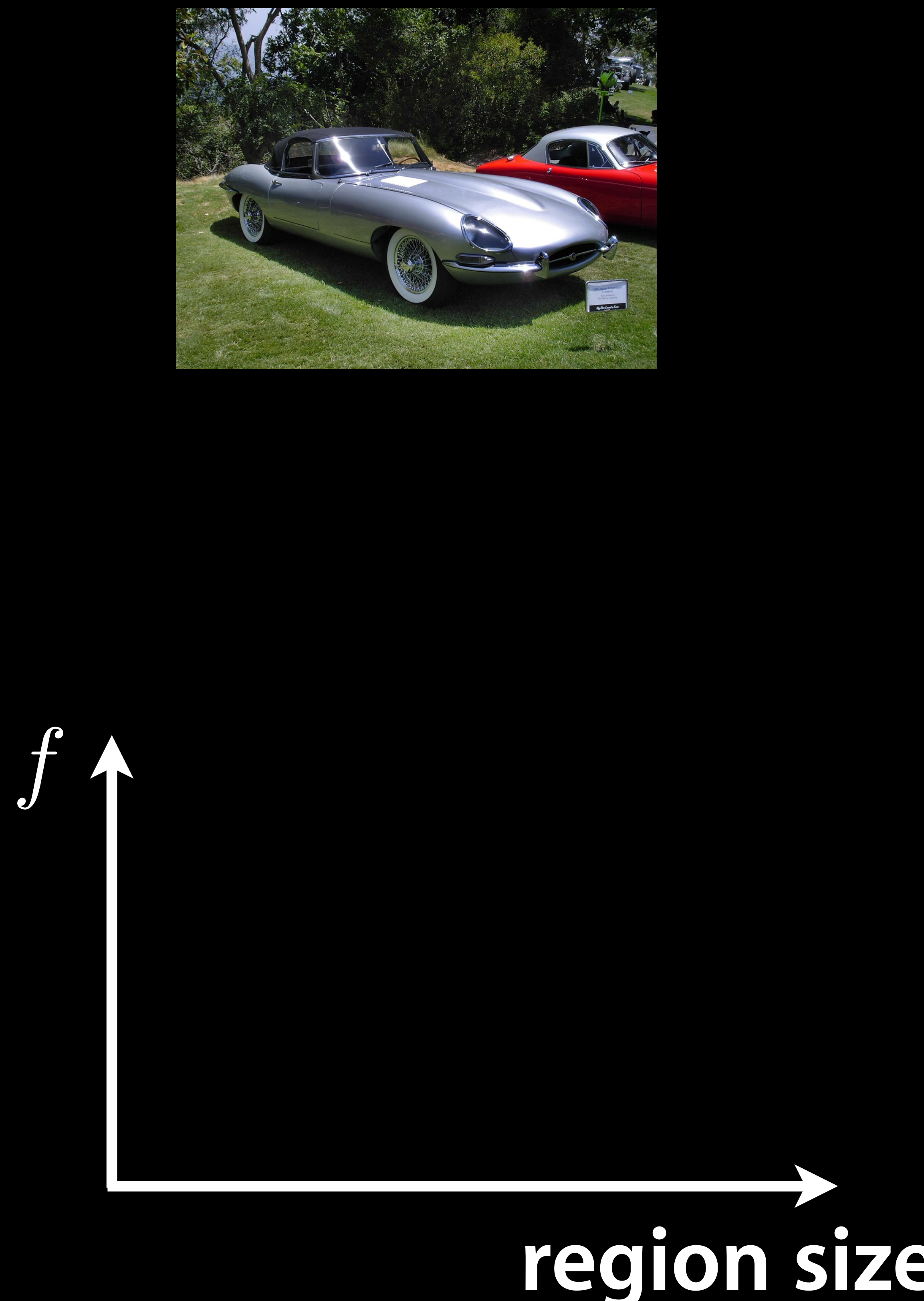

x
f
region size

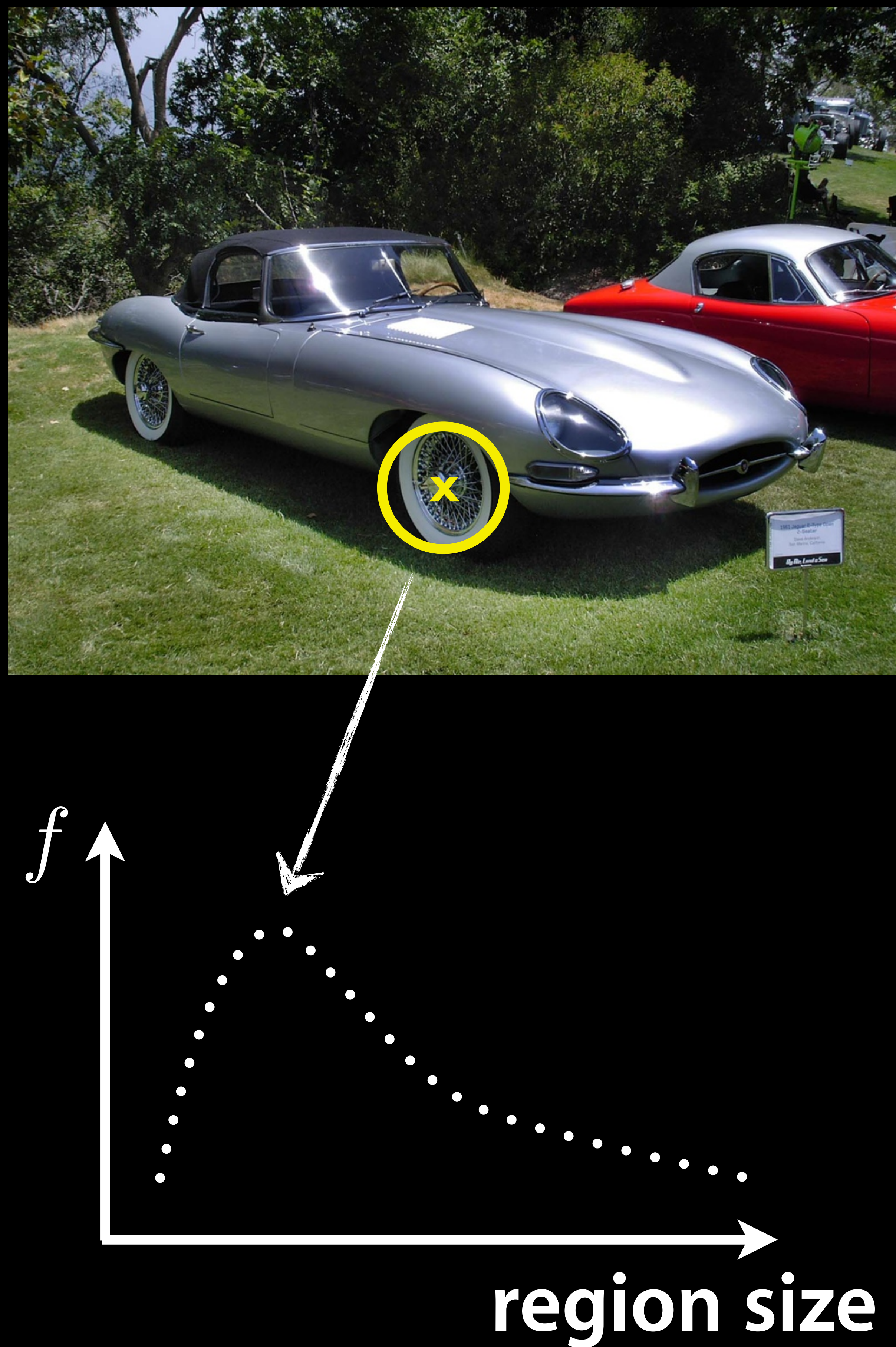

f
region size

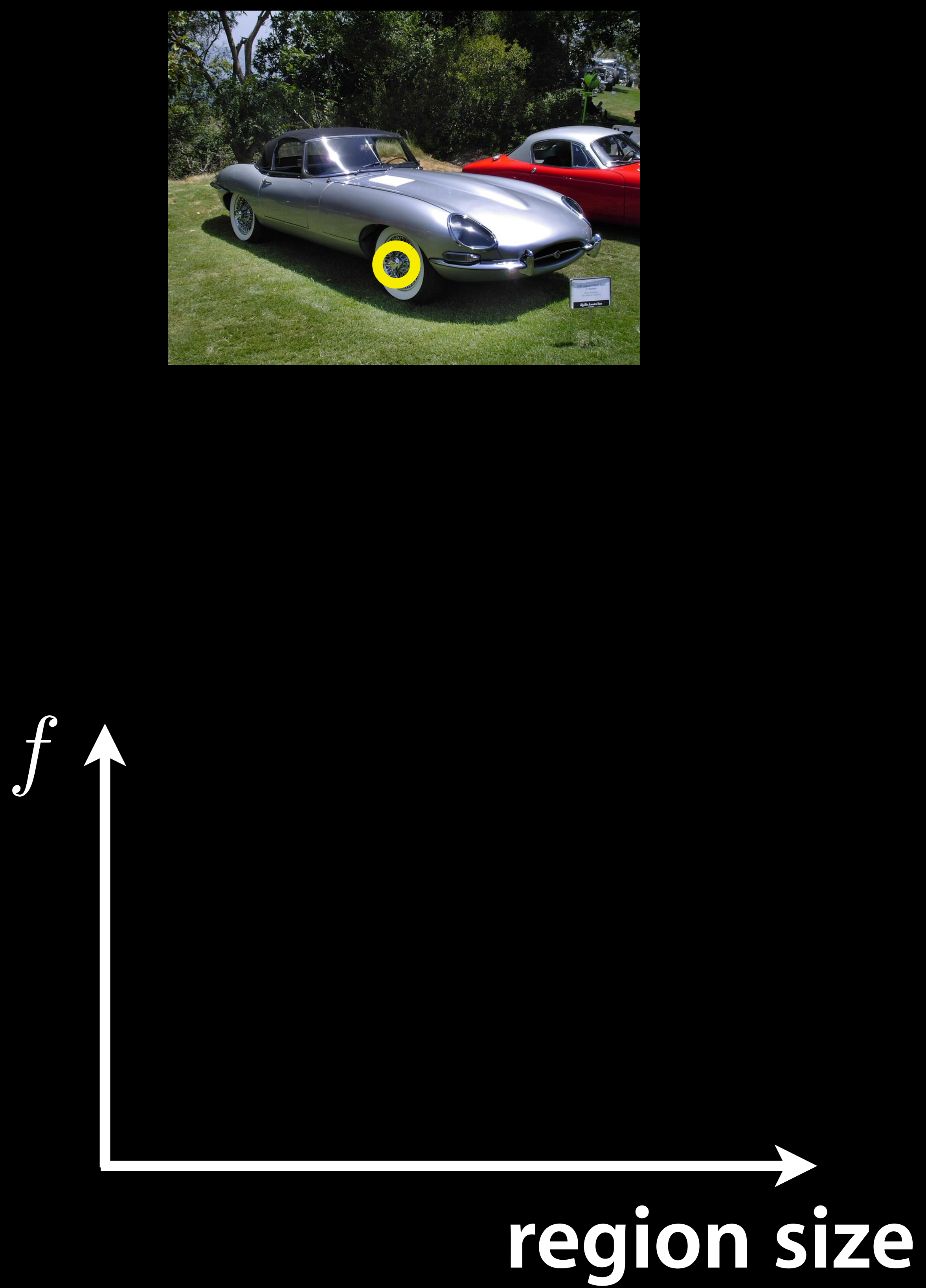


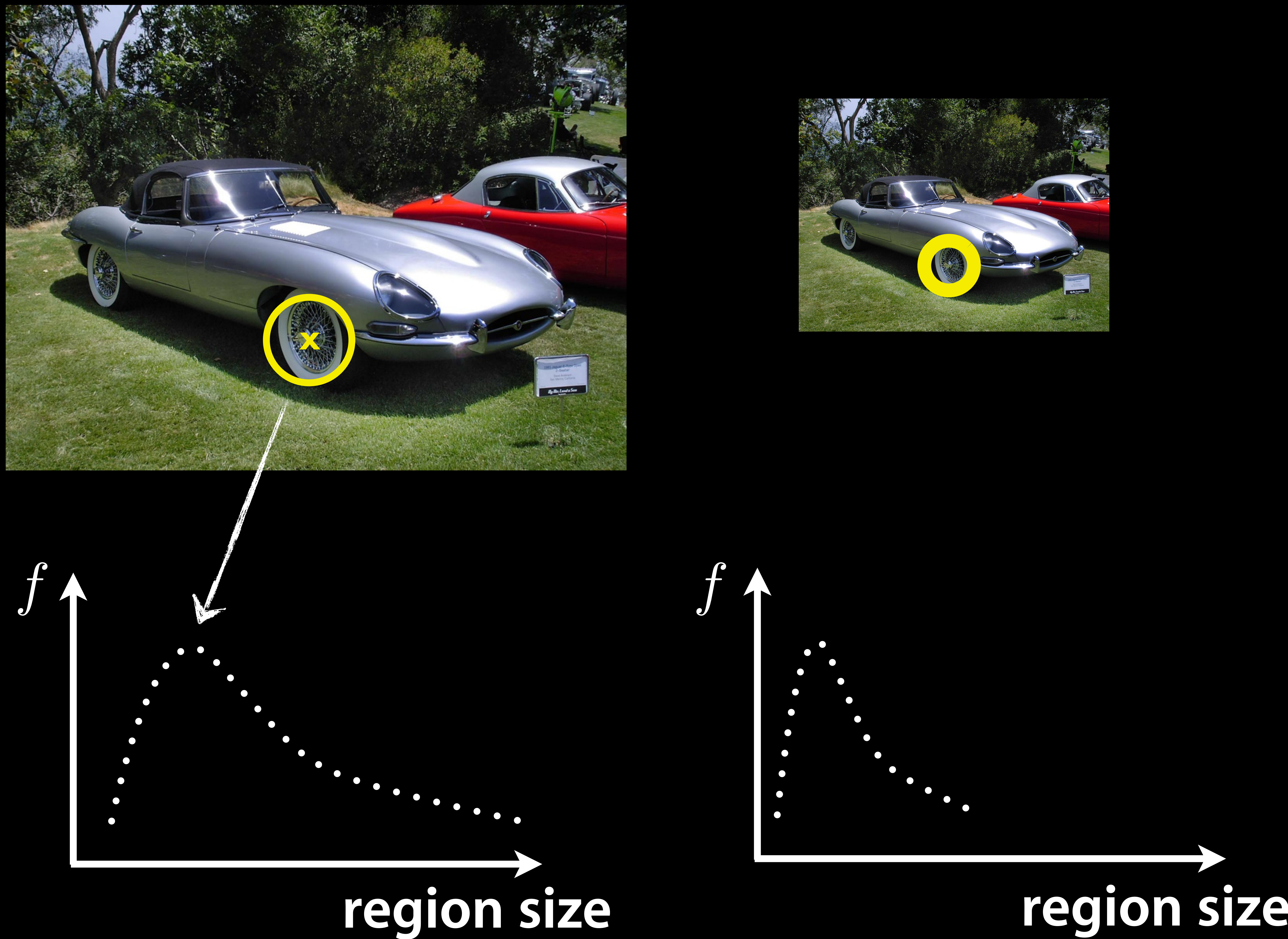
x
f
region size

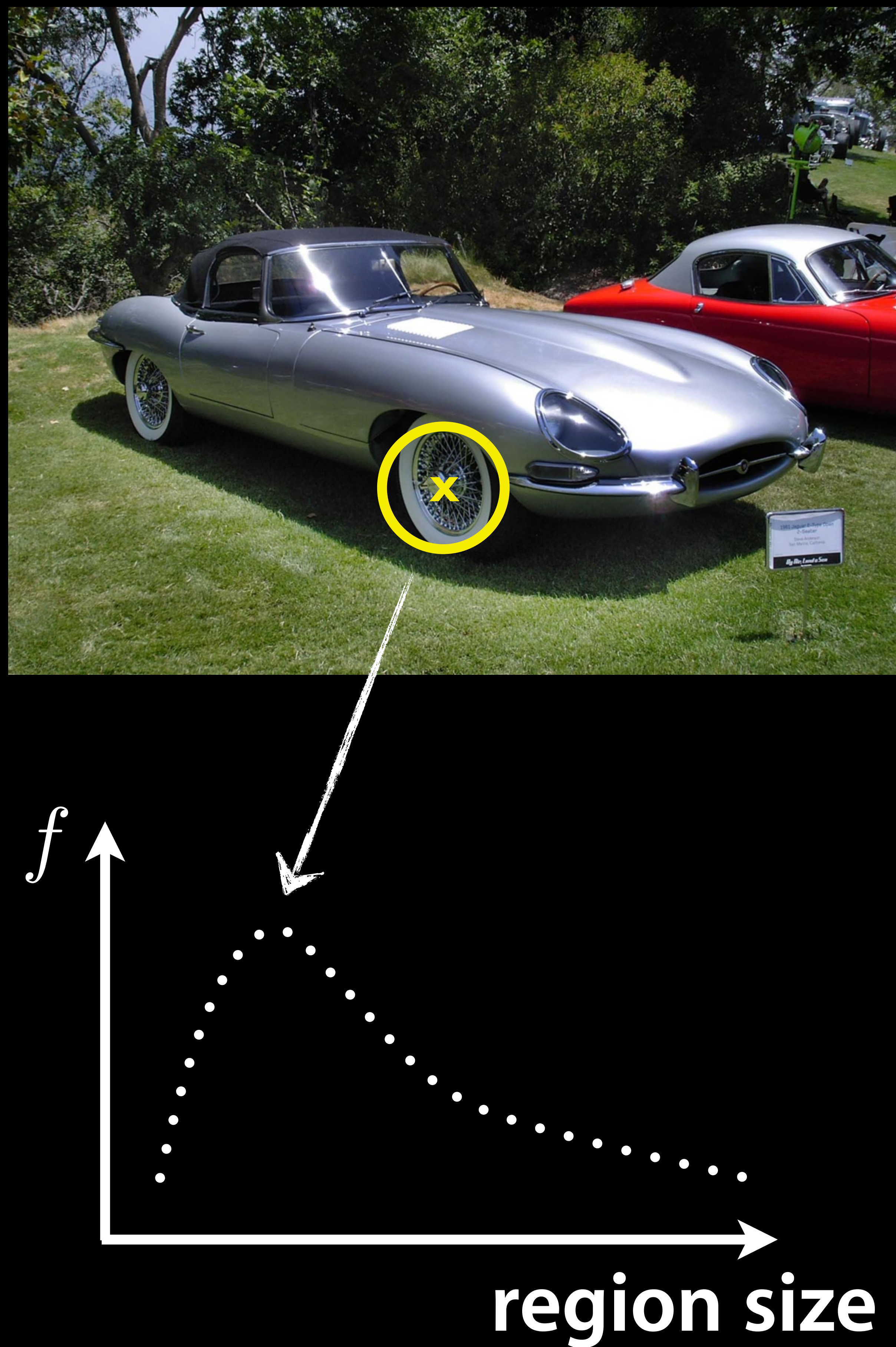

x
f
region size

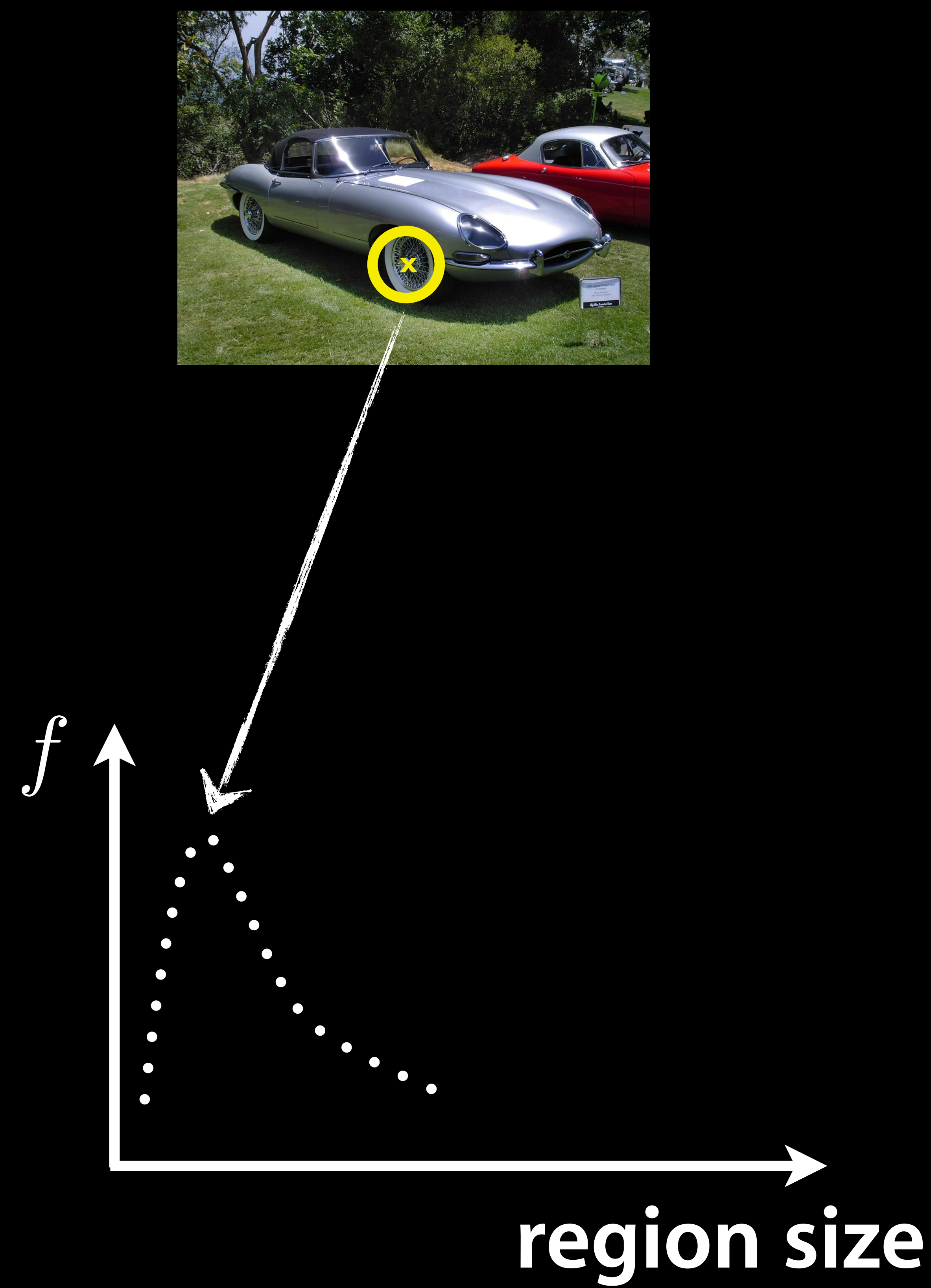

f
region size


f
region size


f
region size


x
f
region size
f
region size


f
region size


f
region size

Scale Invariant Description

image 1

image 2

Scale Invariant Description

image 1

image 2

Resample image about keypoint to a canonical size

Scale Invariant Description

image 1

image 2

Resample image about keypoint to a canonical size

Scale Invariant Description

image 1

image 2

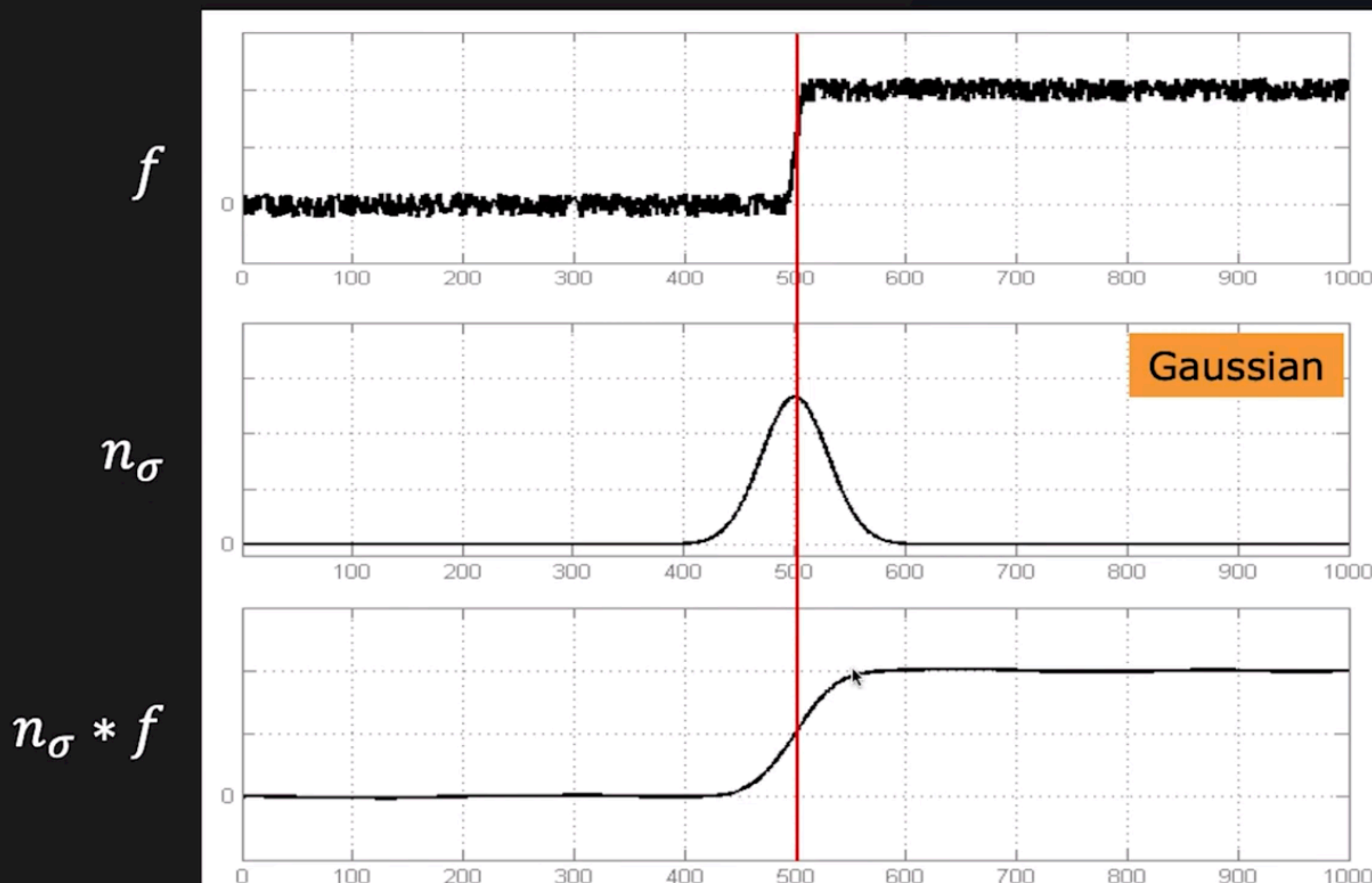
Resample image about keypoint to a canonical size

Which feature detector?

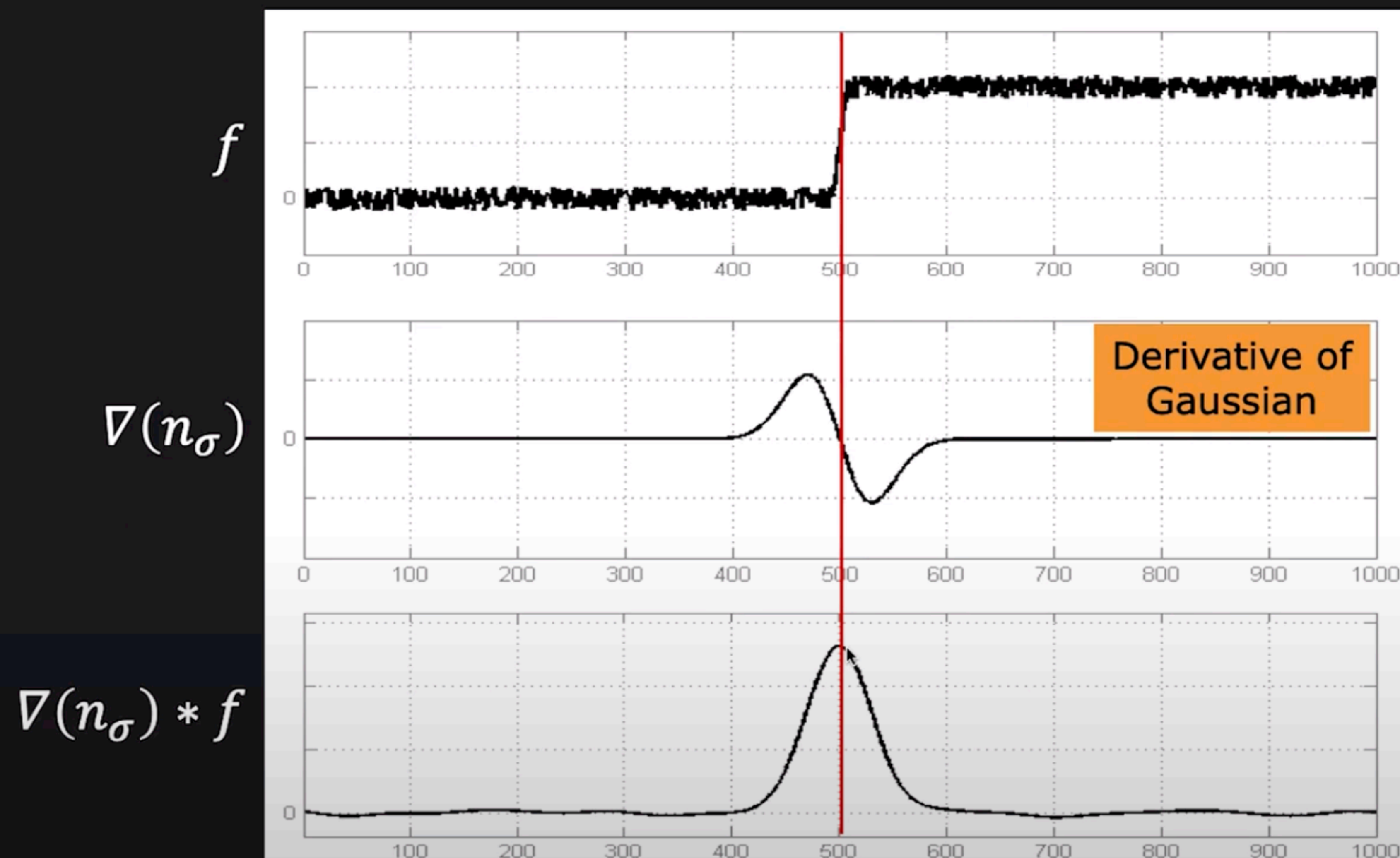
Which feature detector?

$$f = \text{Kernel} * \text{Image}$$

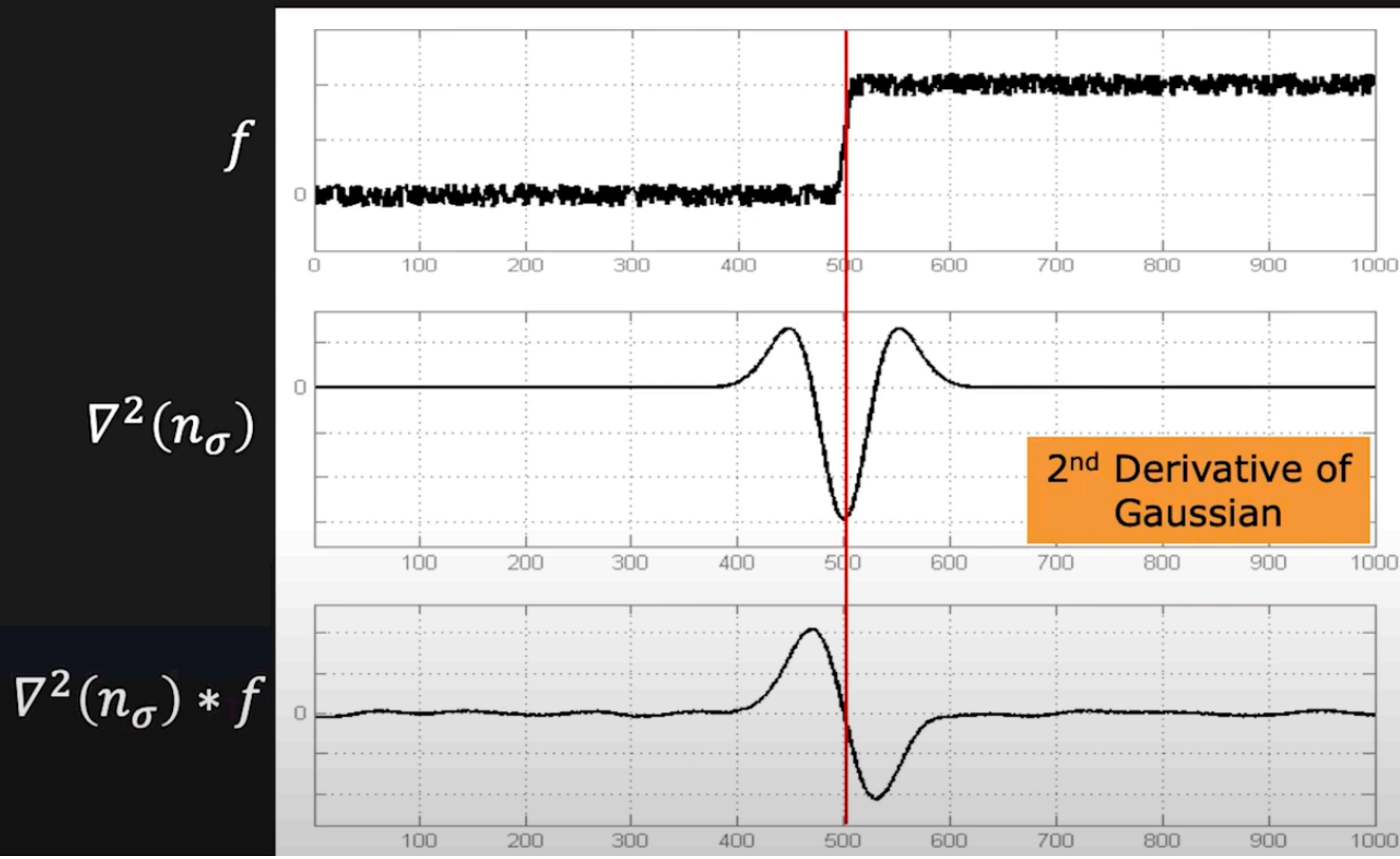
Review: Gaussian Filter



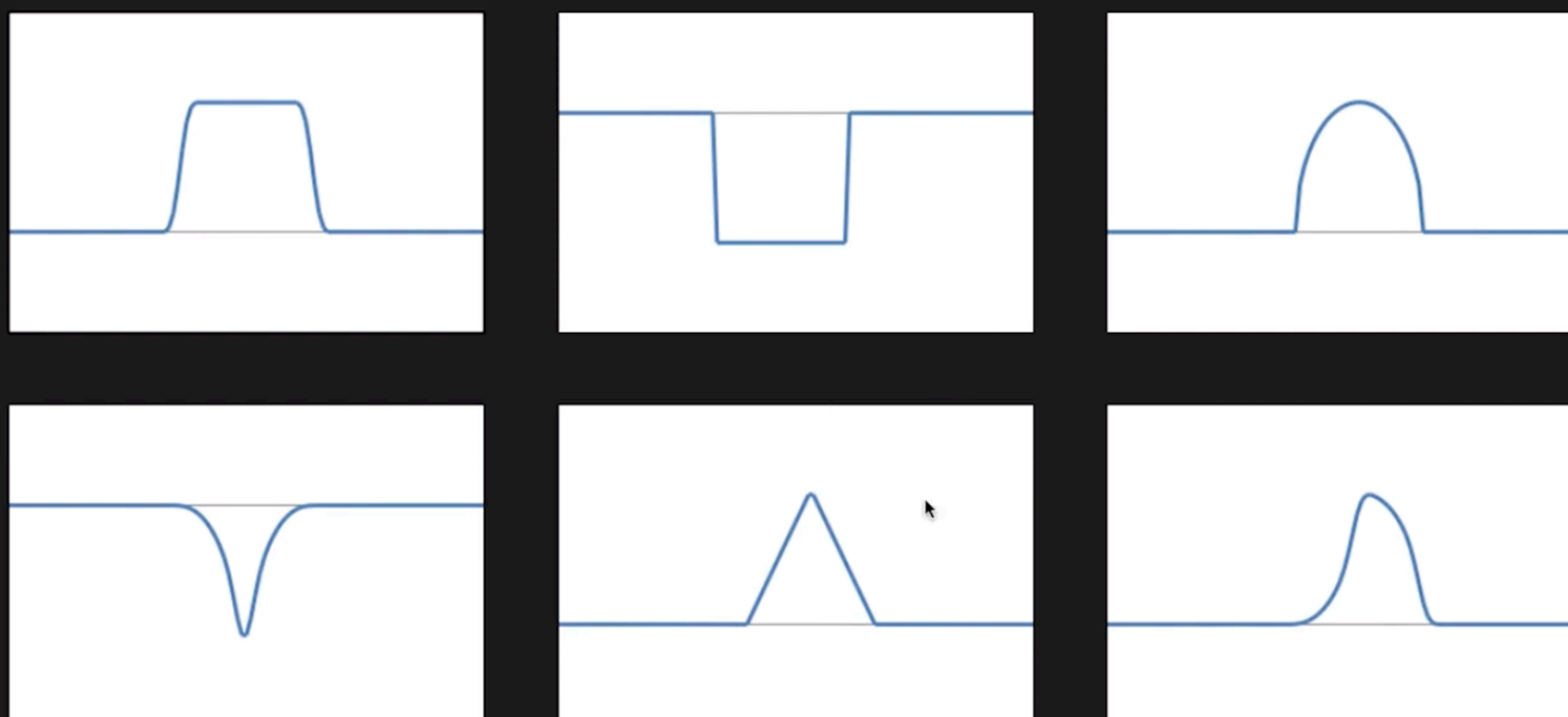
Review: Derivative of Gaussian



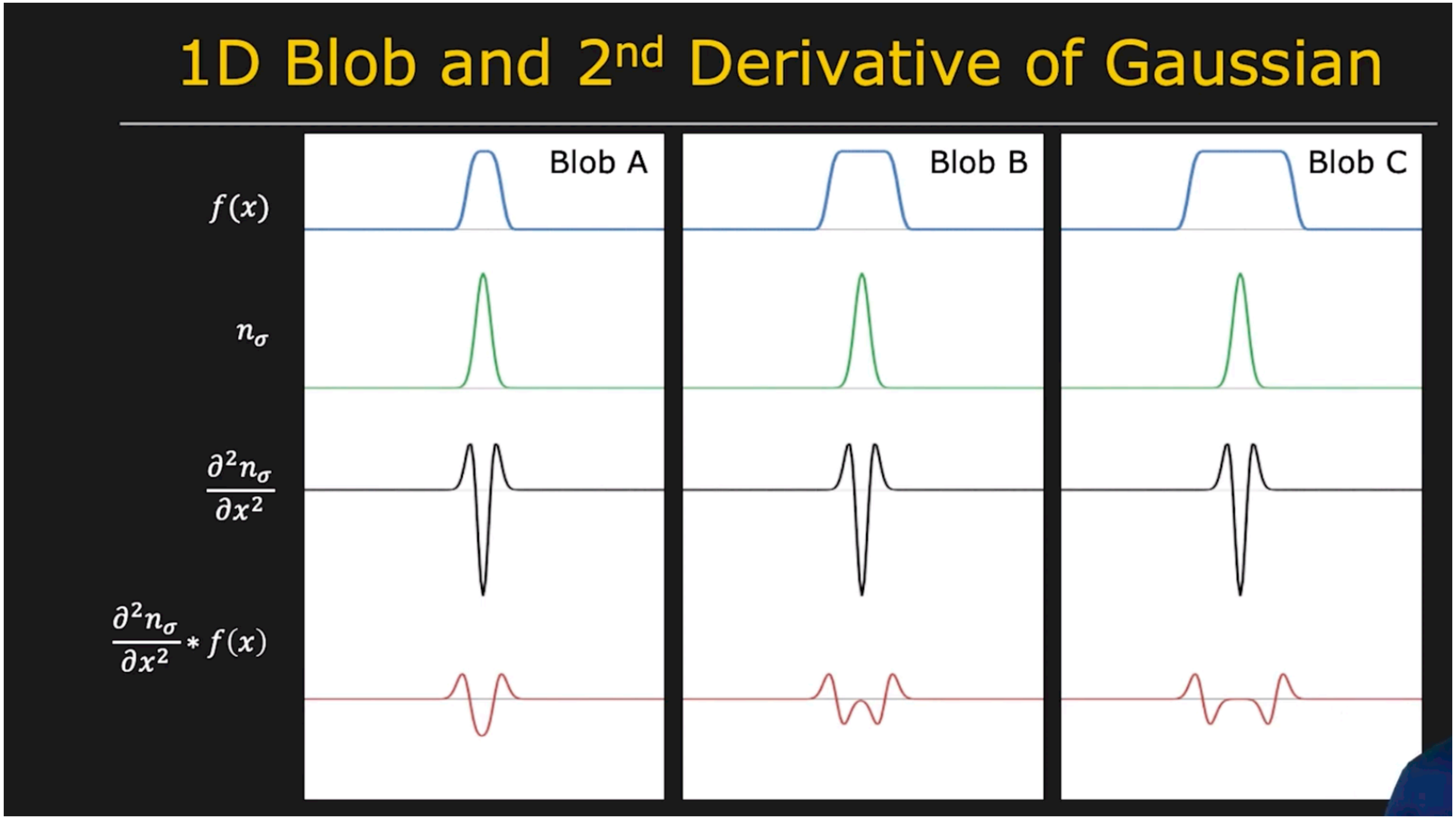
Review: 2nd Derivative of Gaussian



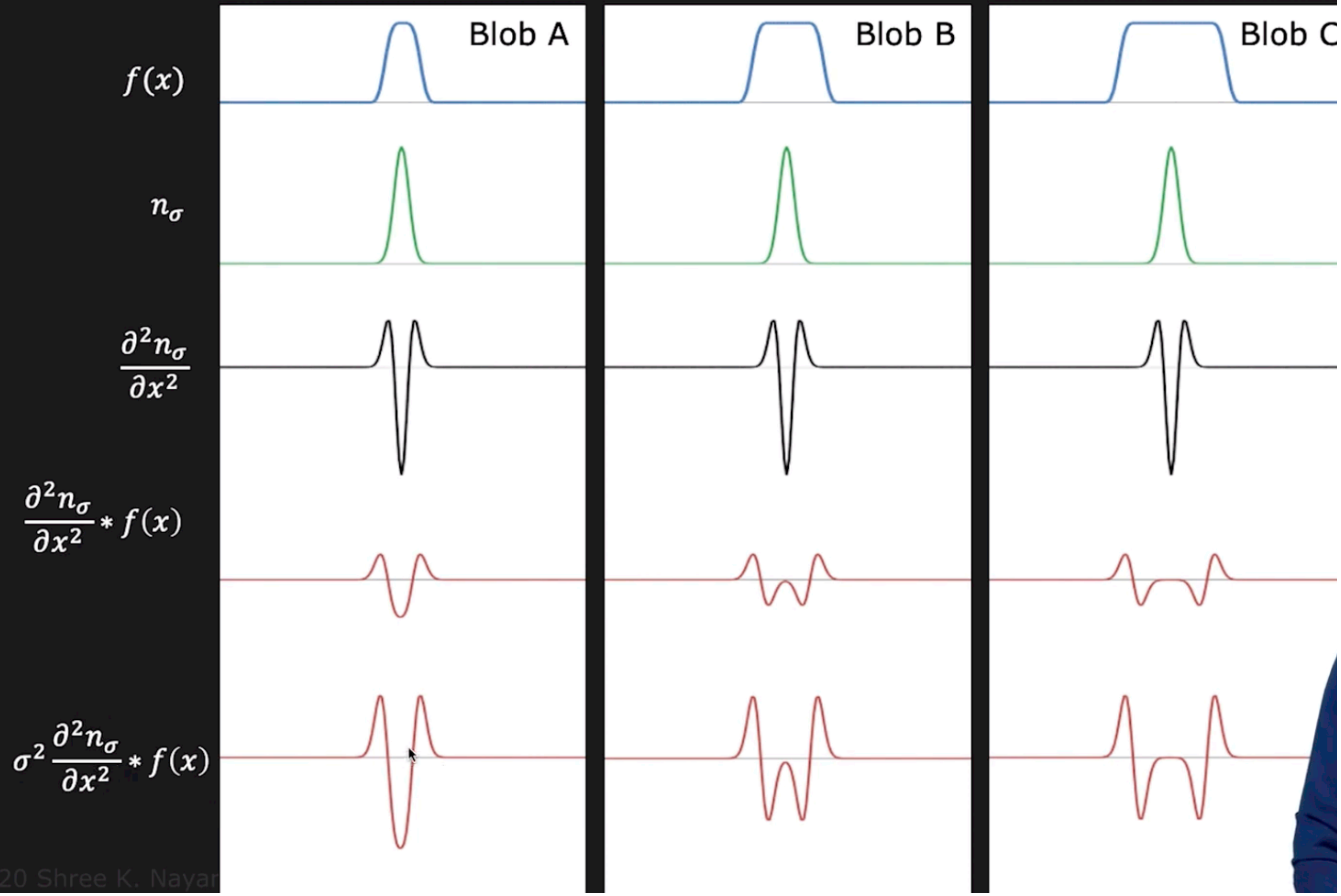
1D Blobs



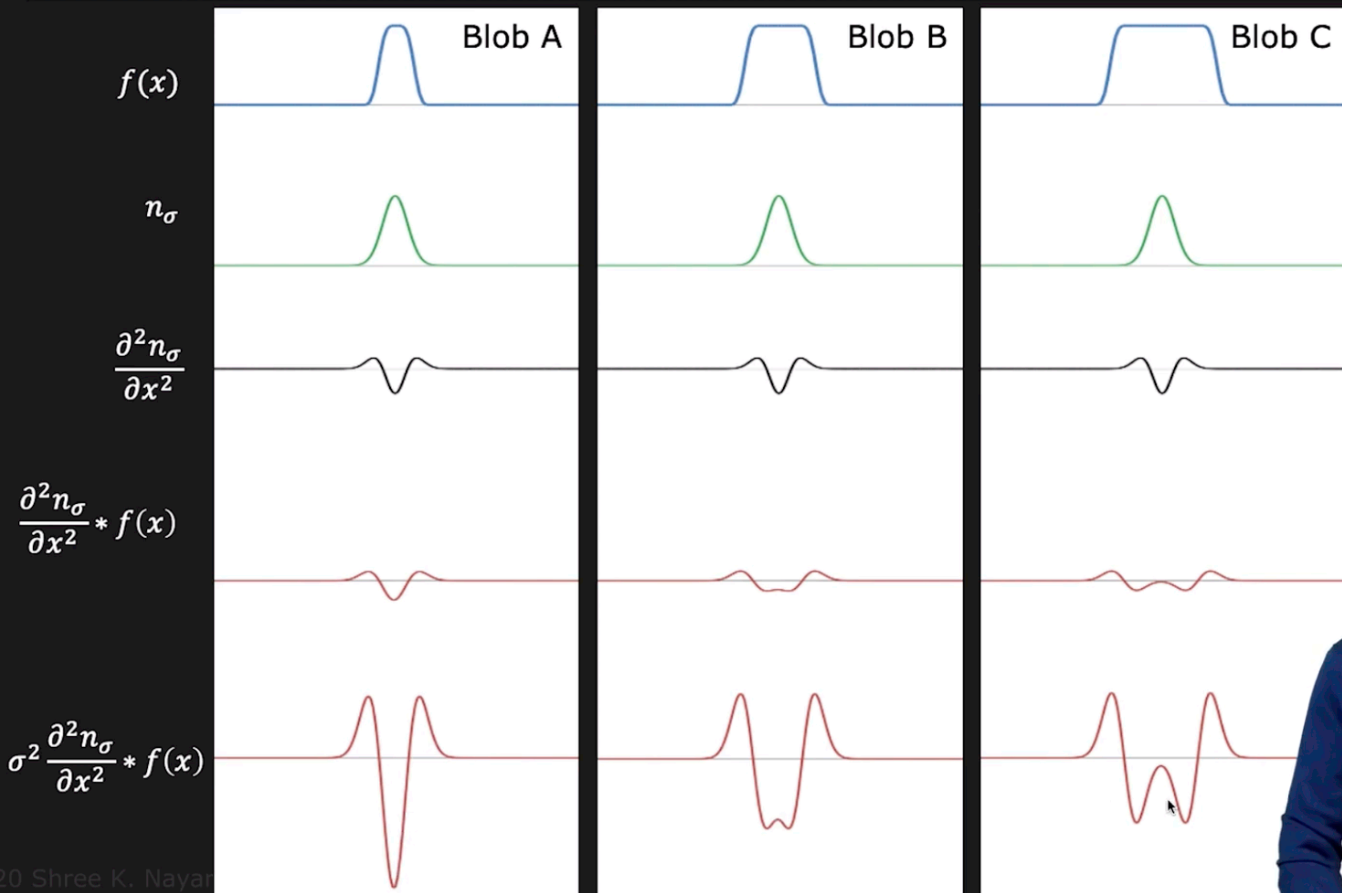
Examples of 1D Blob-like structures


1D Blob and 2nd Derivative of Gaussian
Blob A
Blob B
Blob C
$f(x)$
n_σ
$\frac{\partial^2 n_\sigma}{\partial x^2}$
$\frac{\partial^2 n_\sigma}{\partial x^2} * f(x)$

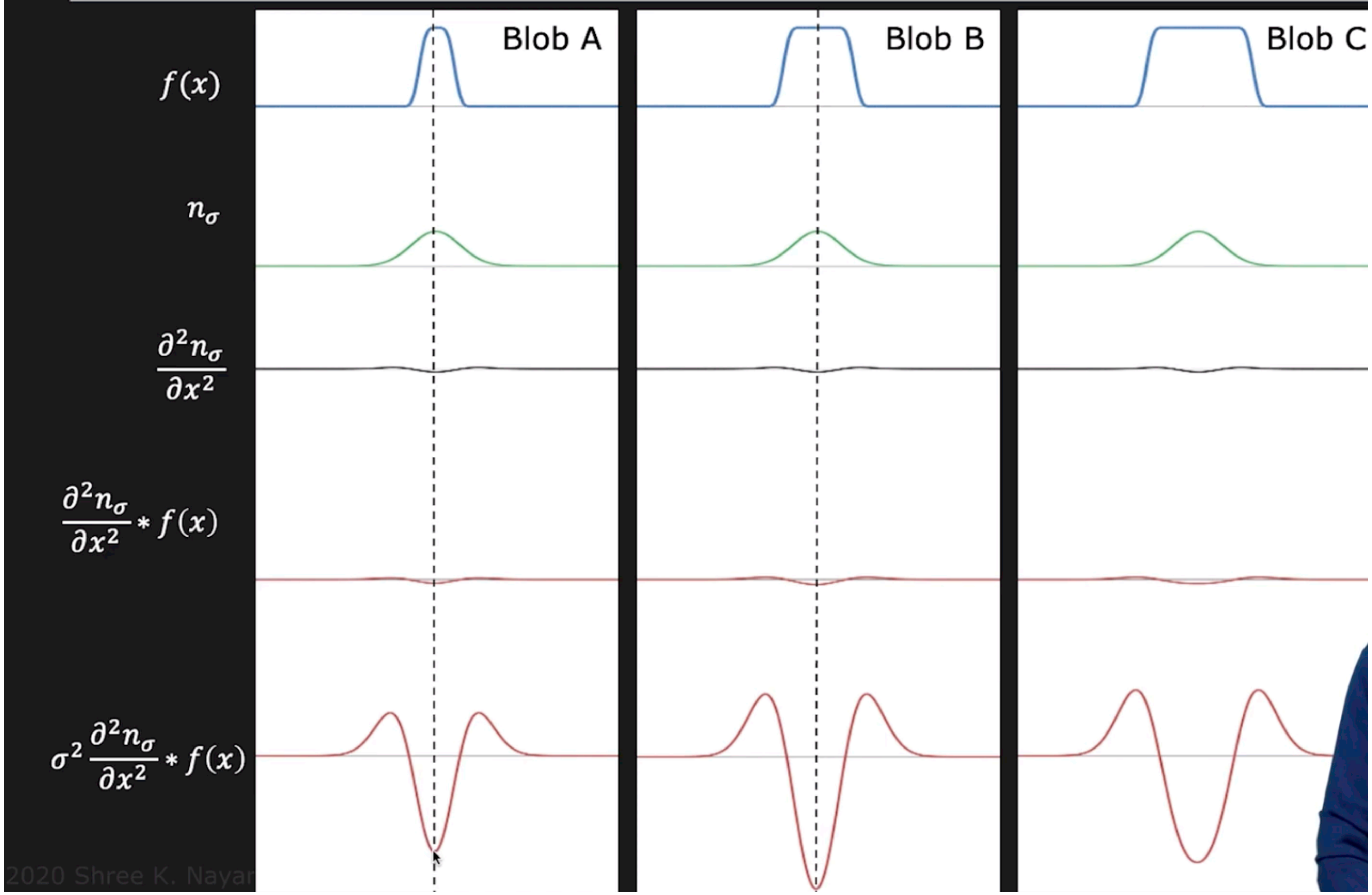
1D Blob and 2nd Derivative of Gaussian



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



1D Blob and 2nd Derivative of Gaussian



1D Blob Detection Summary

Given: 1D signal $f(x)$

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

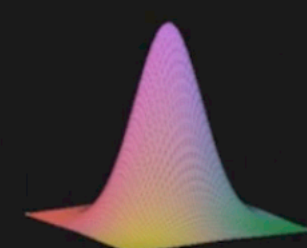
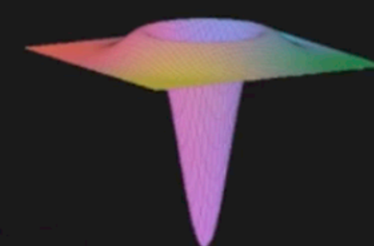
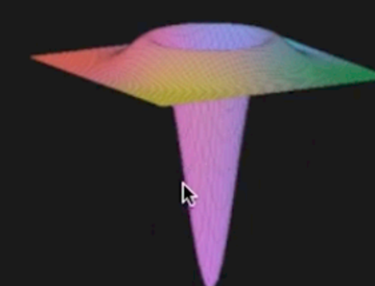
Find: $$\boxed{(x^*, \sigma^*) = \underset{(x,\sigma)}{\arg\max} \left| \sigma^2 \frac{\partial^2 n_\sigma}{\partial x^2} * f(x) \right|}$$

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	Gaussian	LoG	NLoG
$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$			
	n_σ	$\nabla^2 n_\sigma$	$\sigma^2 \nabla^2 n_\sigma$

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

2D Blob Detection Summary

Given an image $I(x, y)$

Convolve the image using NLoG at many scales σ

Find:

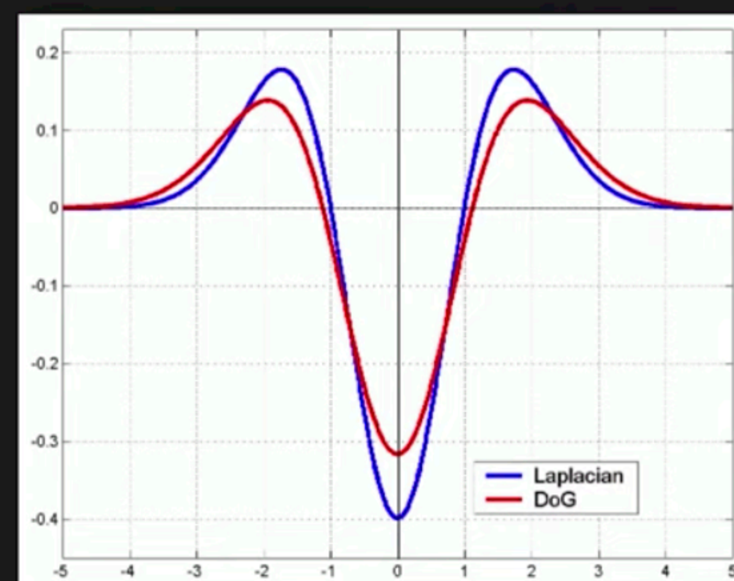
$$\boxed{(x^*, y^*, \sigma^*) = \underset{(x,y,\sigma)}{\arg\max} \left|\sigma^2 \nabla^2 n_\sigma * I(x, y)\right|}$$

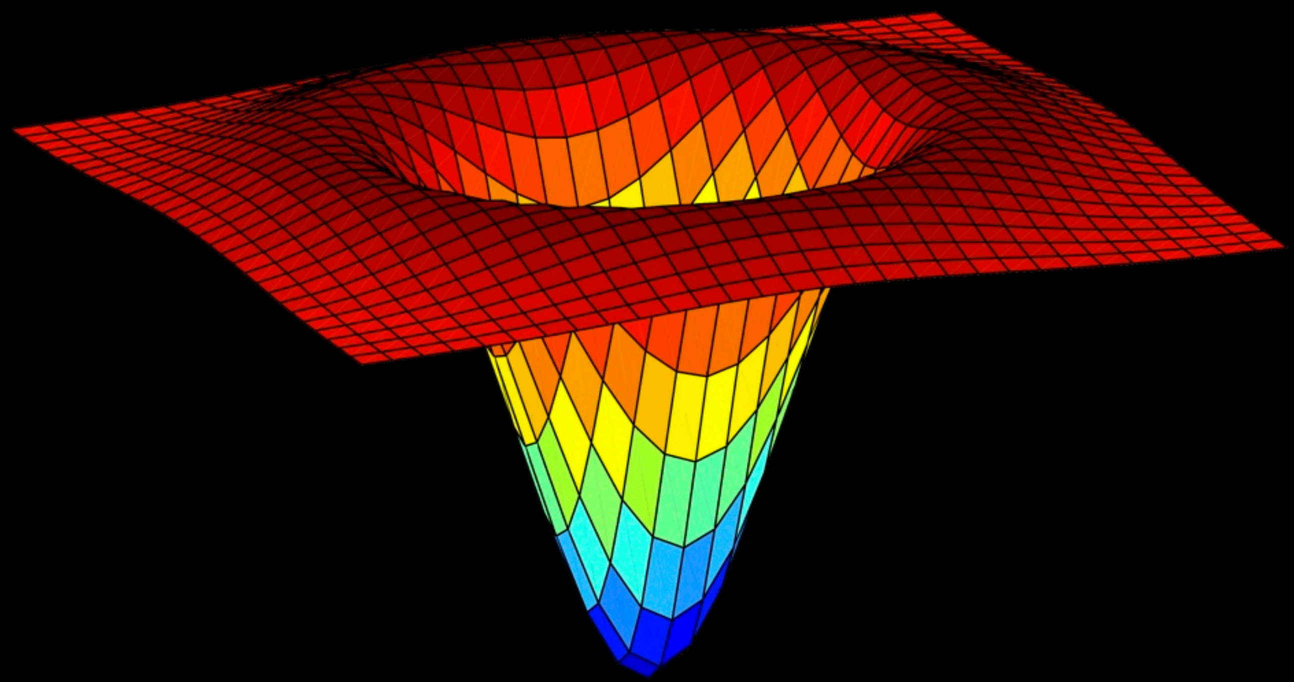
(x^*, y^*): Position of the blob

σ^*: Size of the blob

Fast NLoG Approximation: DoG

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

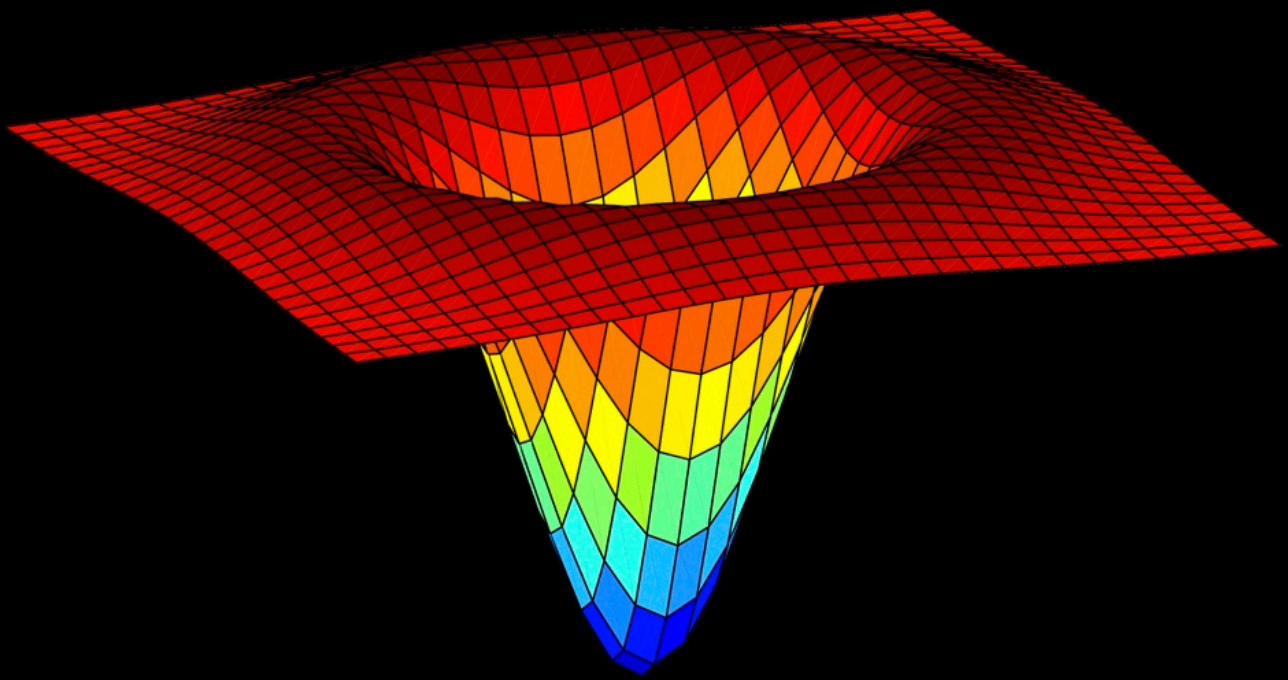




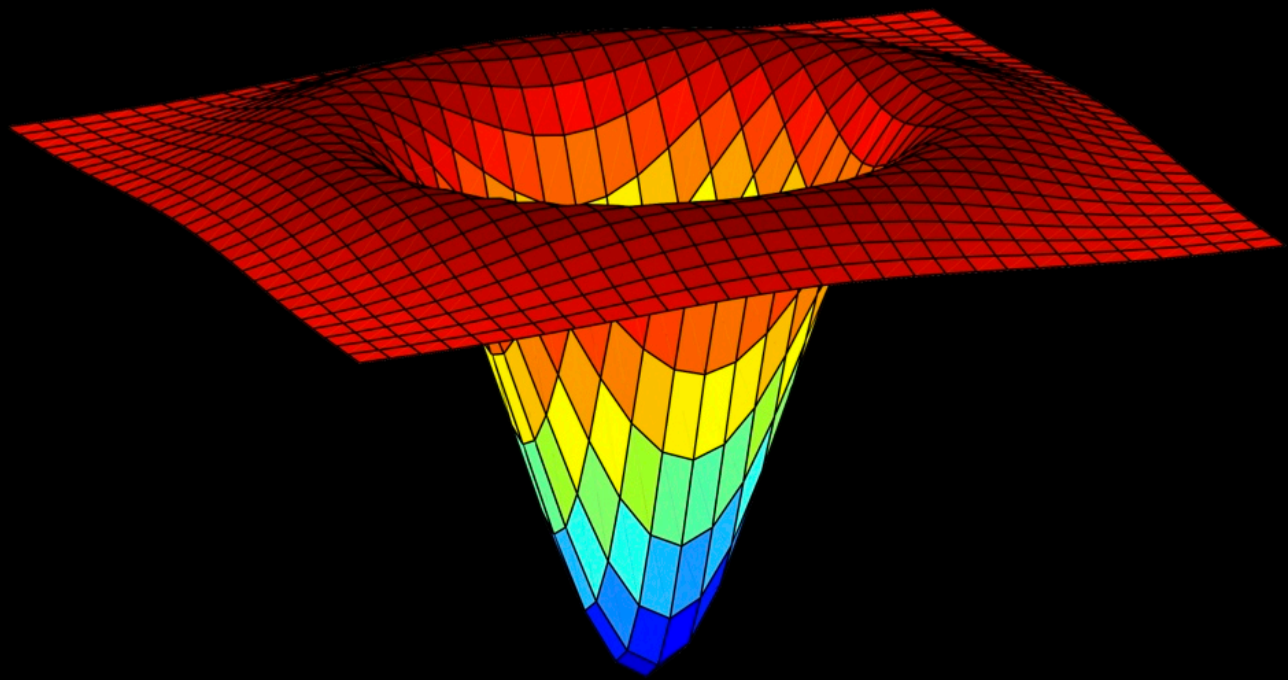
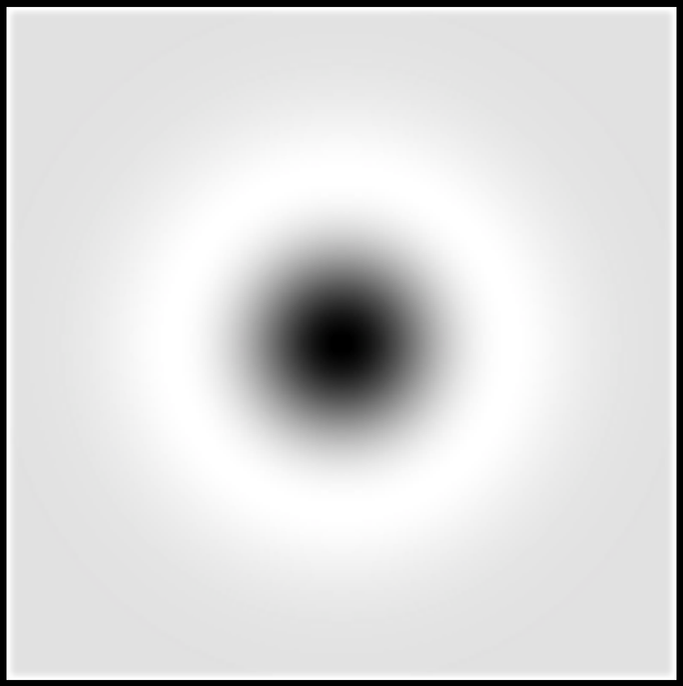
Laplacian of Gaussian (LoG)

Mexican hat operator

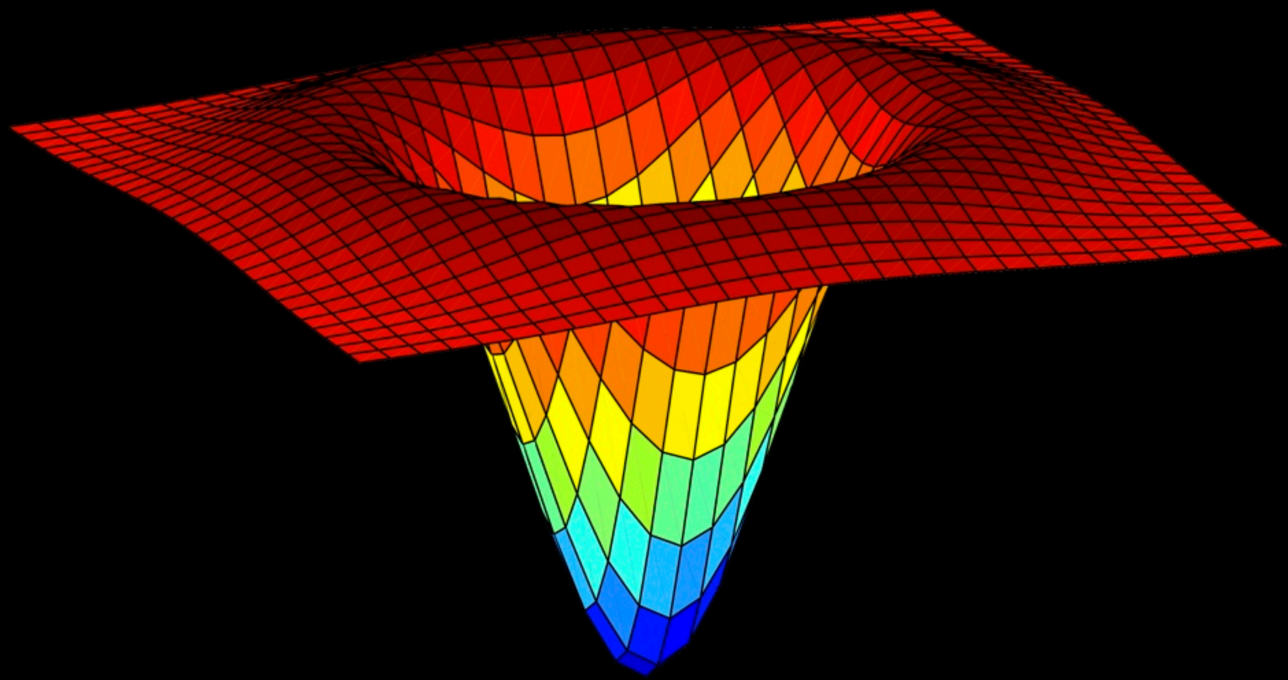
Mexican hat operator

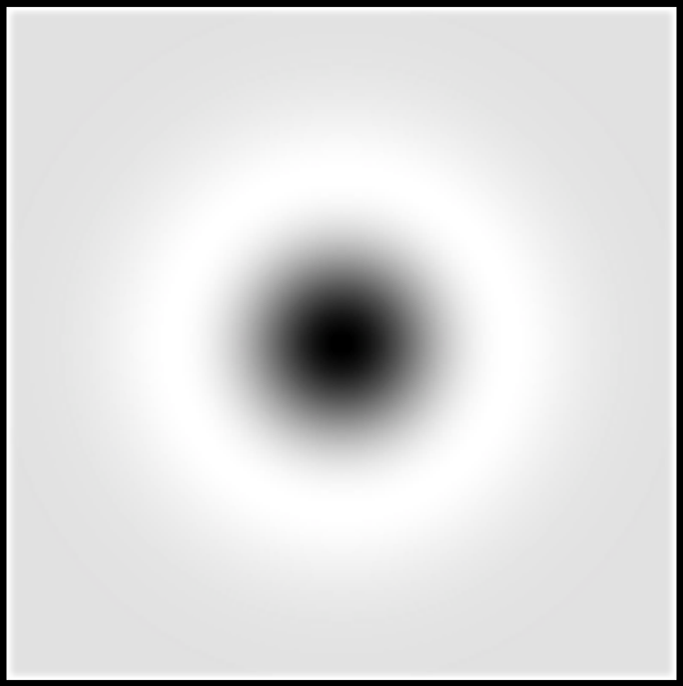


Laplacian of Gaussian (LoG)

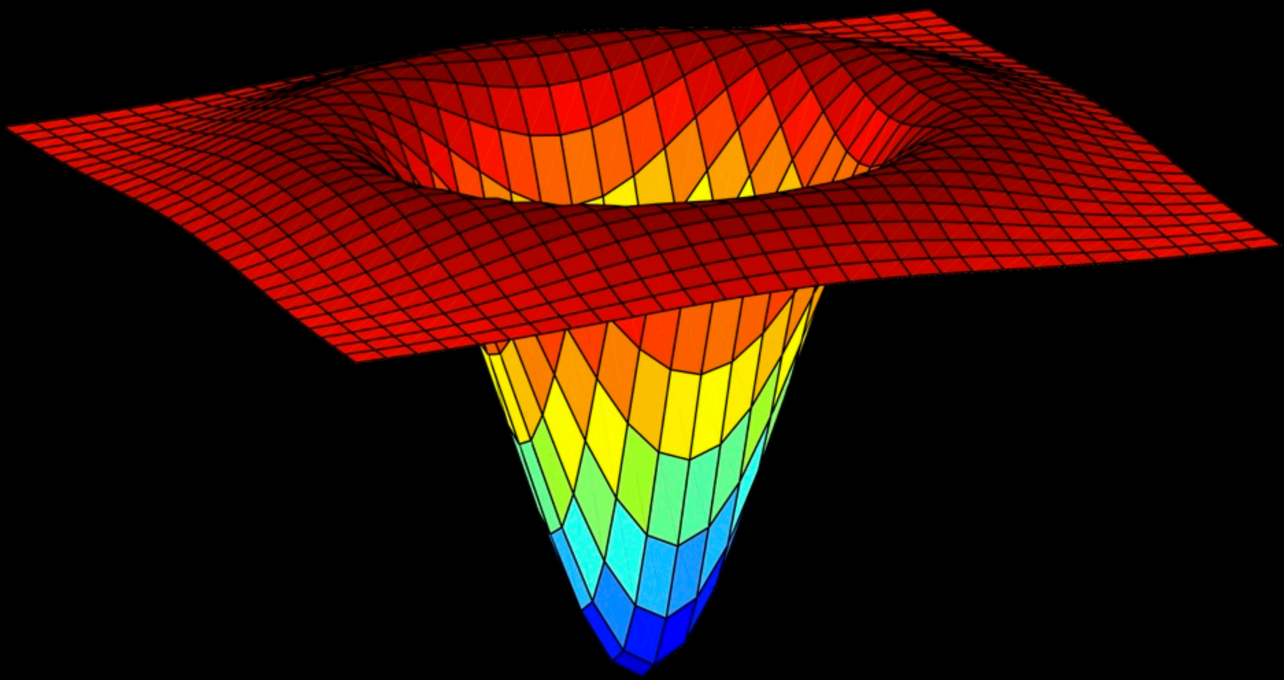
Laplacian of Gaussian (LoG)

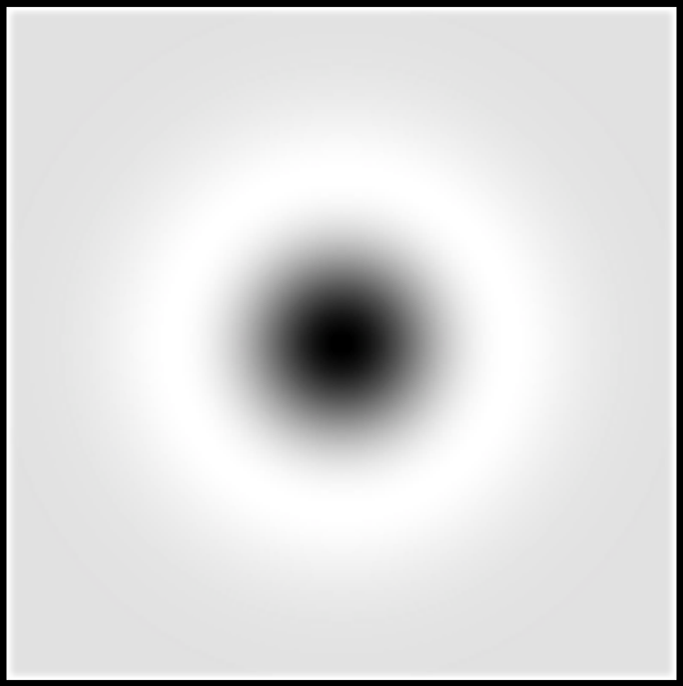




Laplacian of Gaussian (LoG)

What pattern does this filter detect?

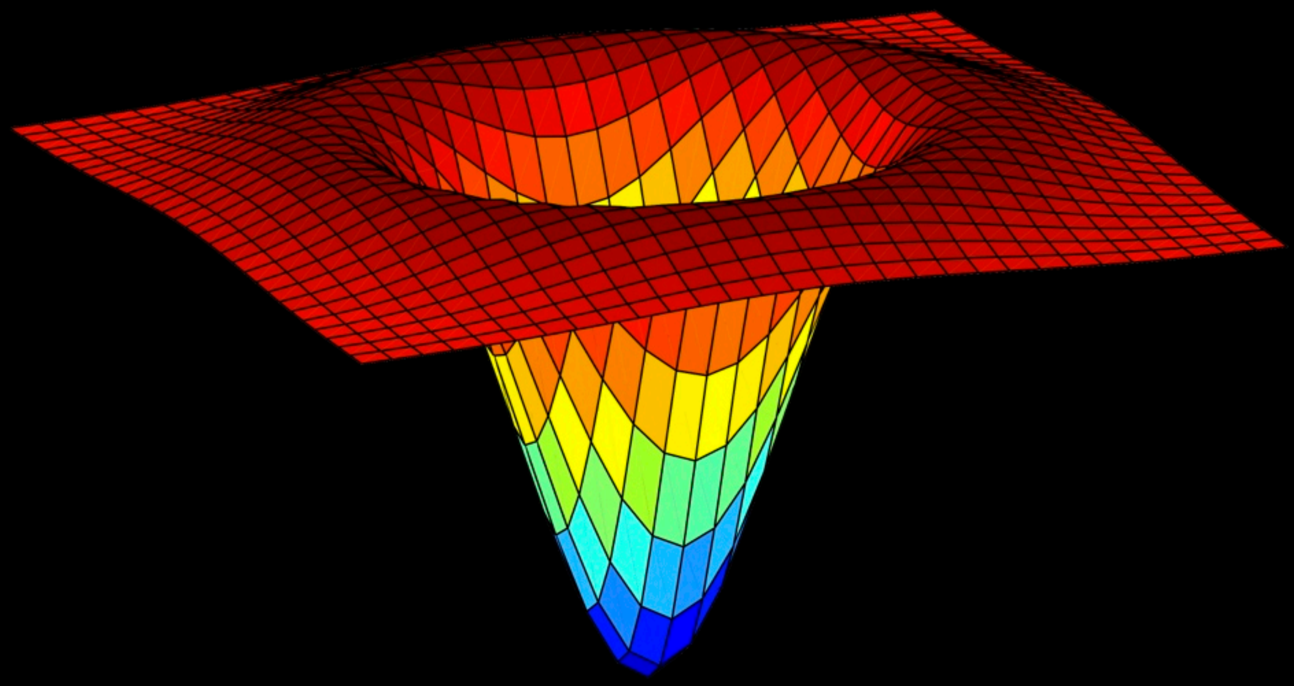




Laplacian of Gaussian (LoG)

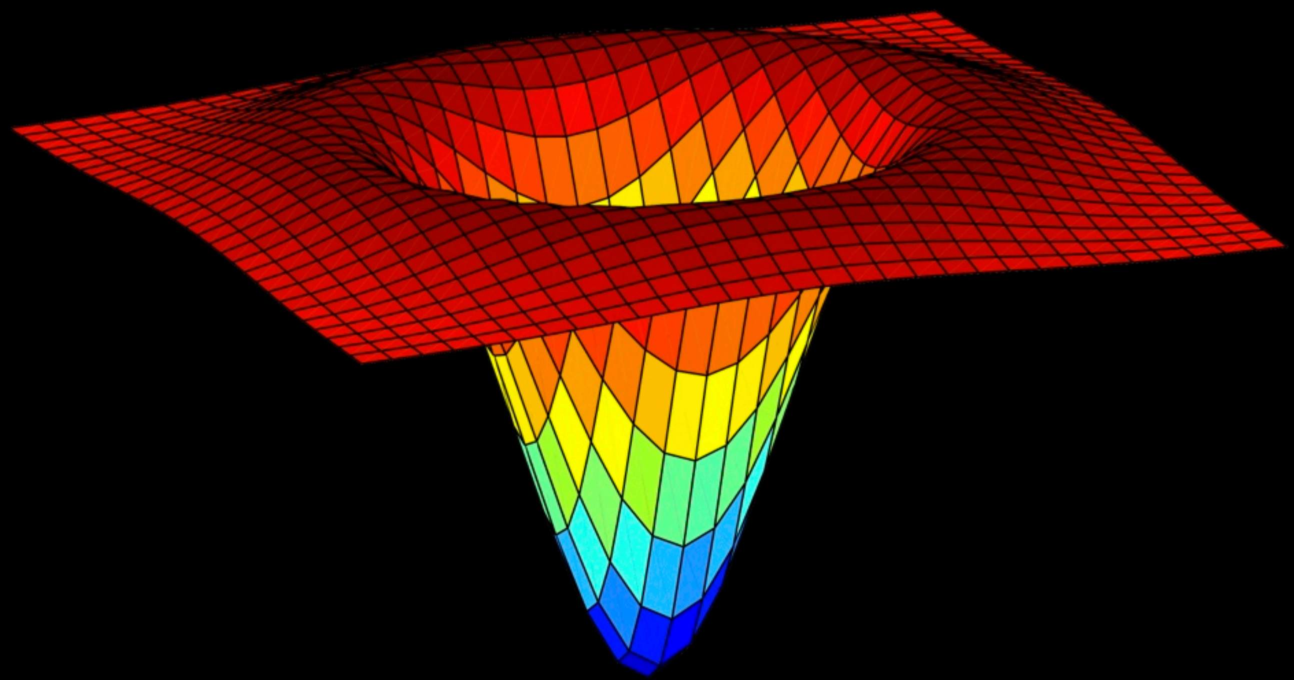
What pattern does this filter detect?

blob



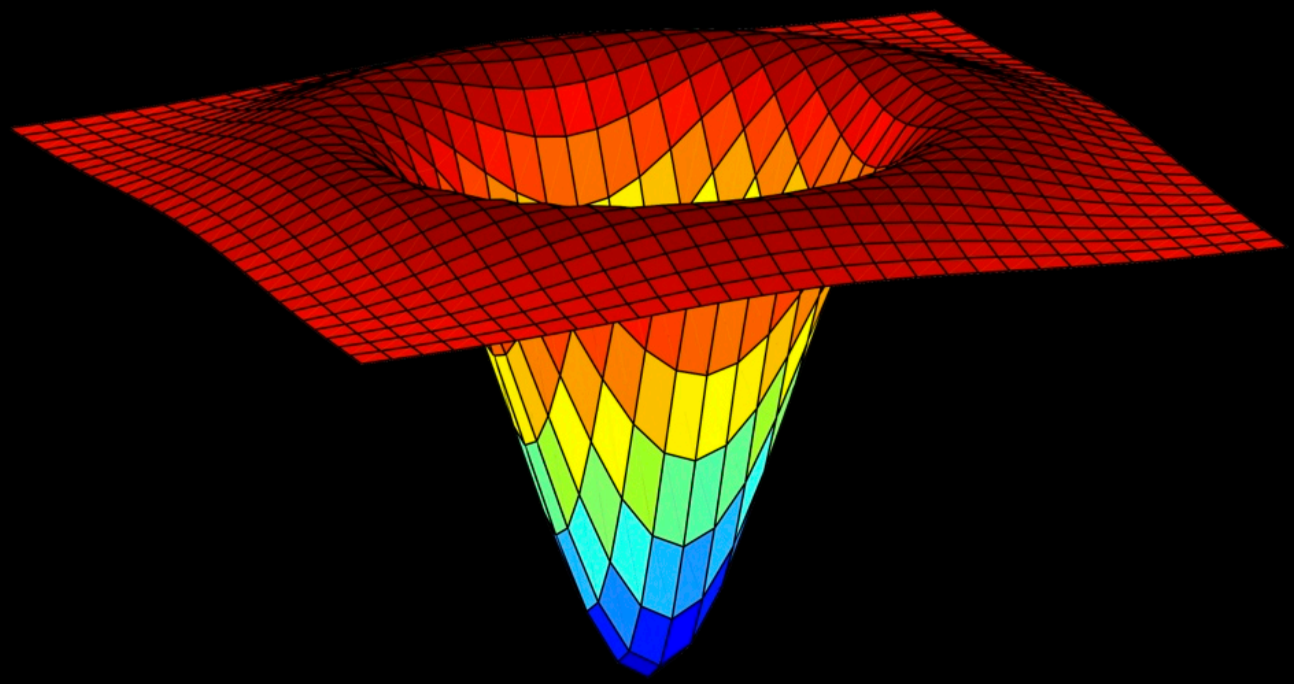
Laplacian of Gaussian (LoG)

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



Laplacian of Gaussian (LoG)

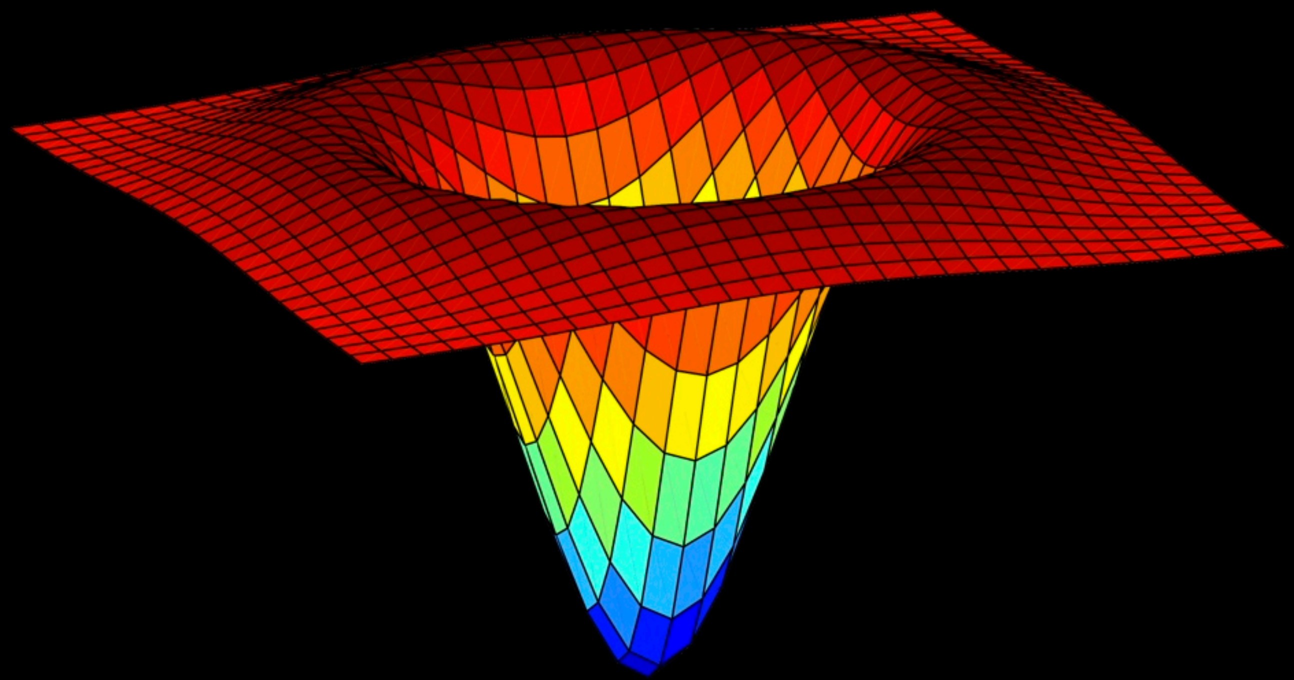
$$L(x, y; \sigma) = \sigma^2 (G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$



Laplacian of Gaussian (LoG)

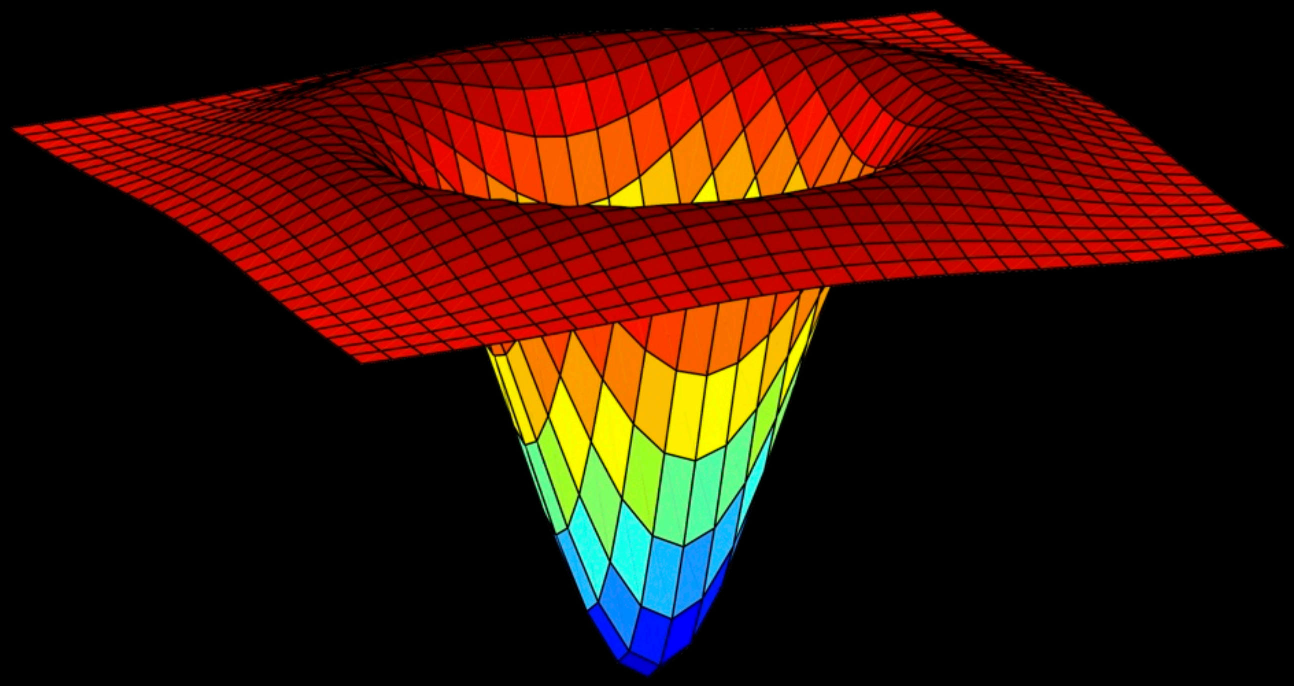
$$L(x, y; \sigma) = \sigma^2 (G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$

$$f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$$



Laplacian of Gaussian (LoG)

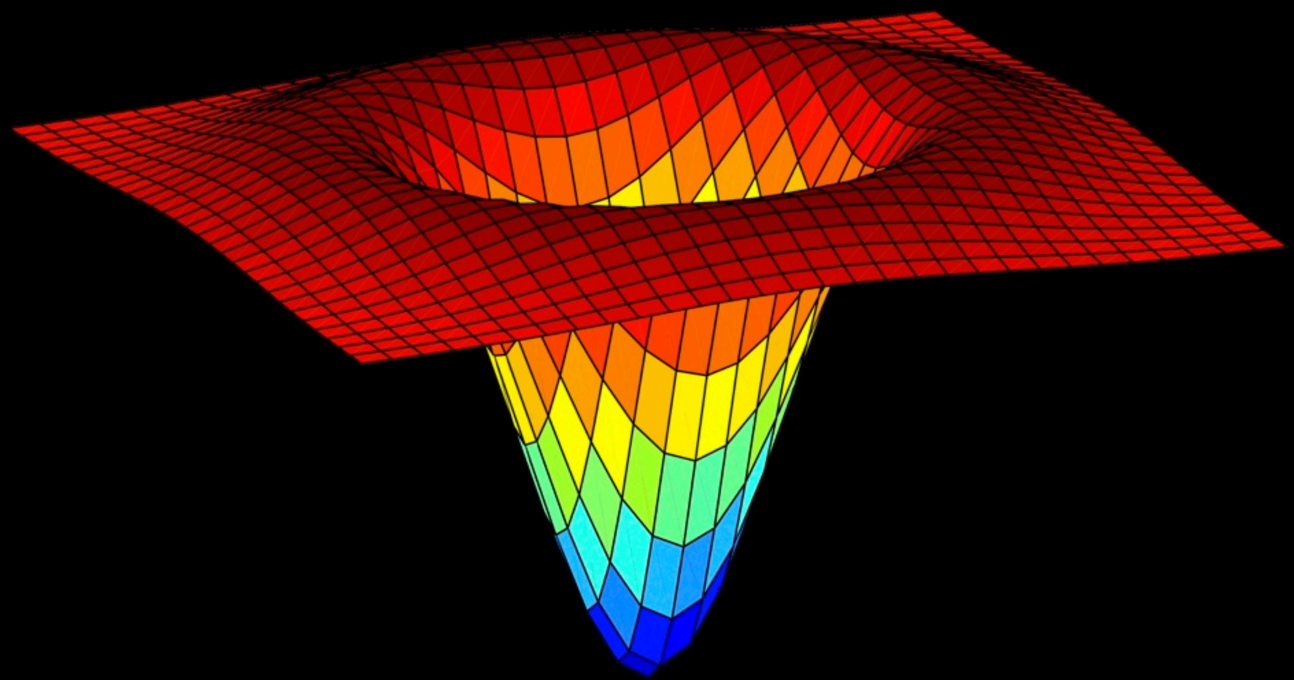
$$L(x, y; \sigma) = \sigma^2 (G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$



Laplacian of Gaussian (LoG)

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

$$\approx G(x, y; k\sigma) - G(x, y; \sigma)$$

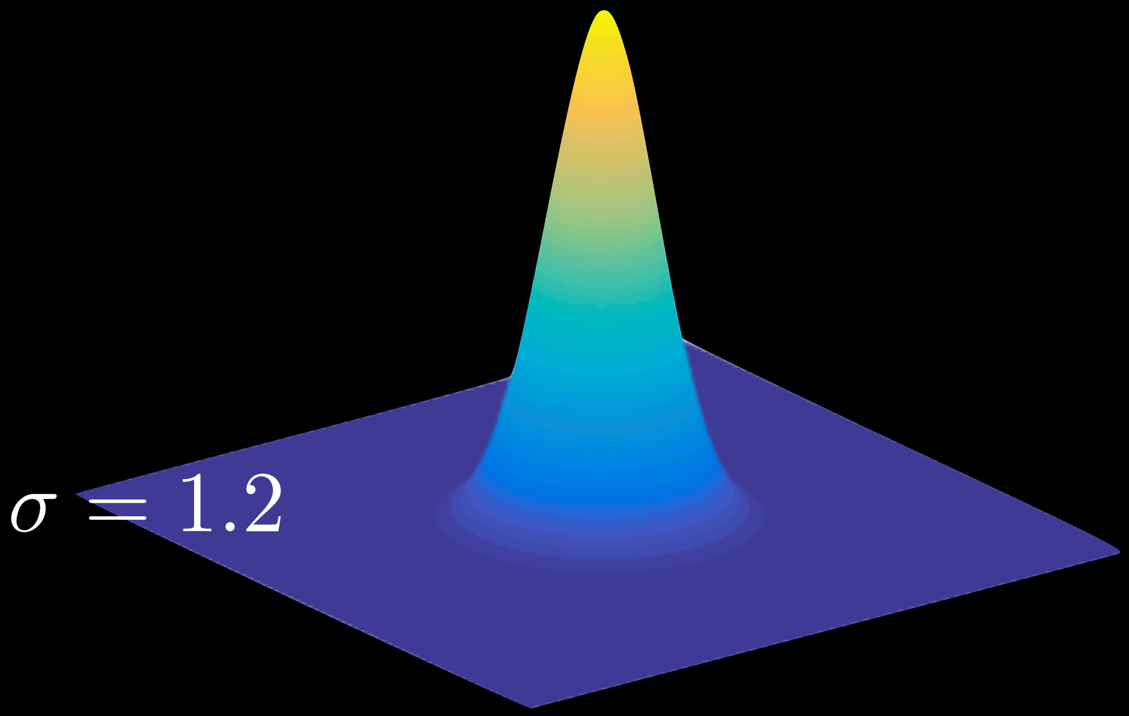


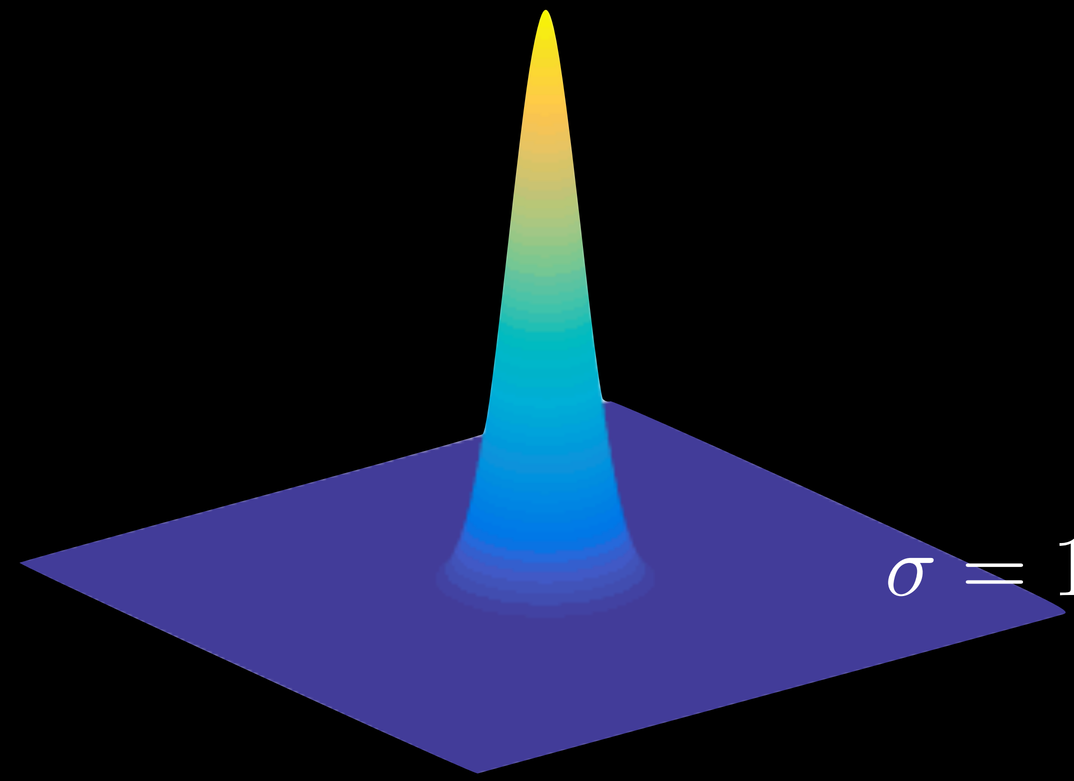
Laplacian of Gaussian (LoG)

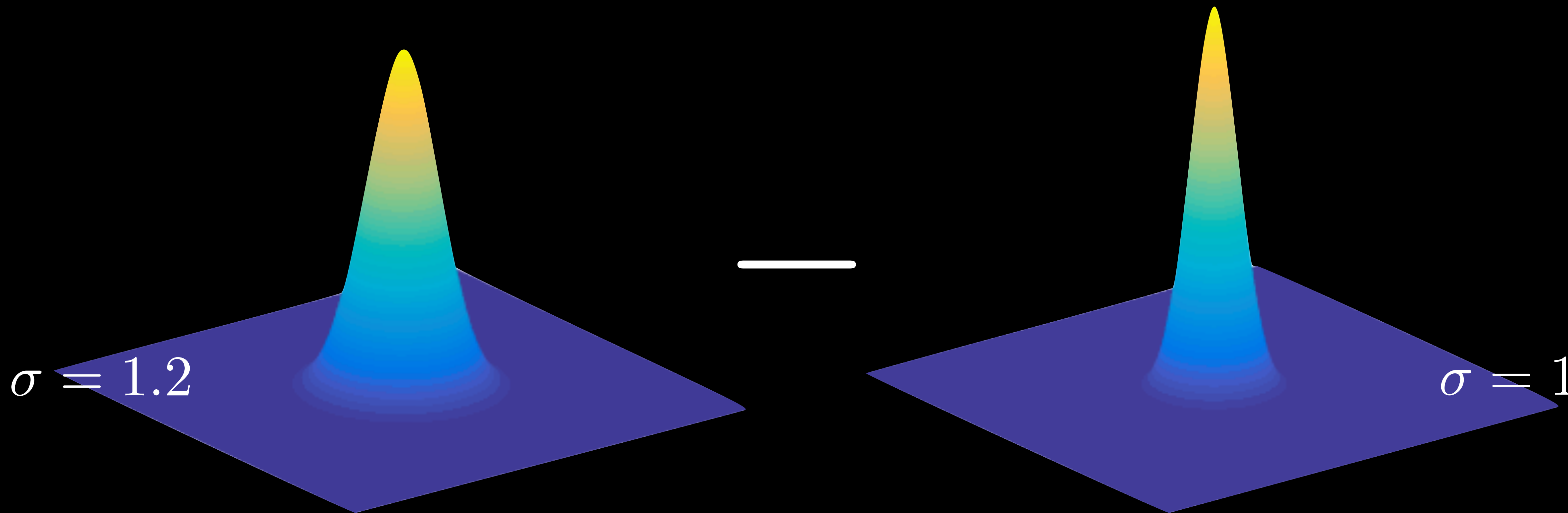
$$L(x, y; \sigma) = \sigma^2 (G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$

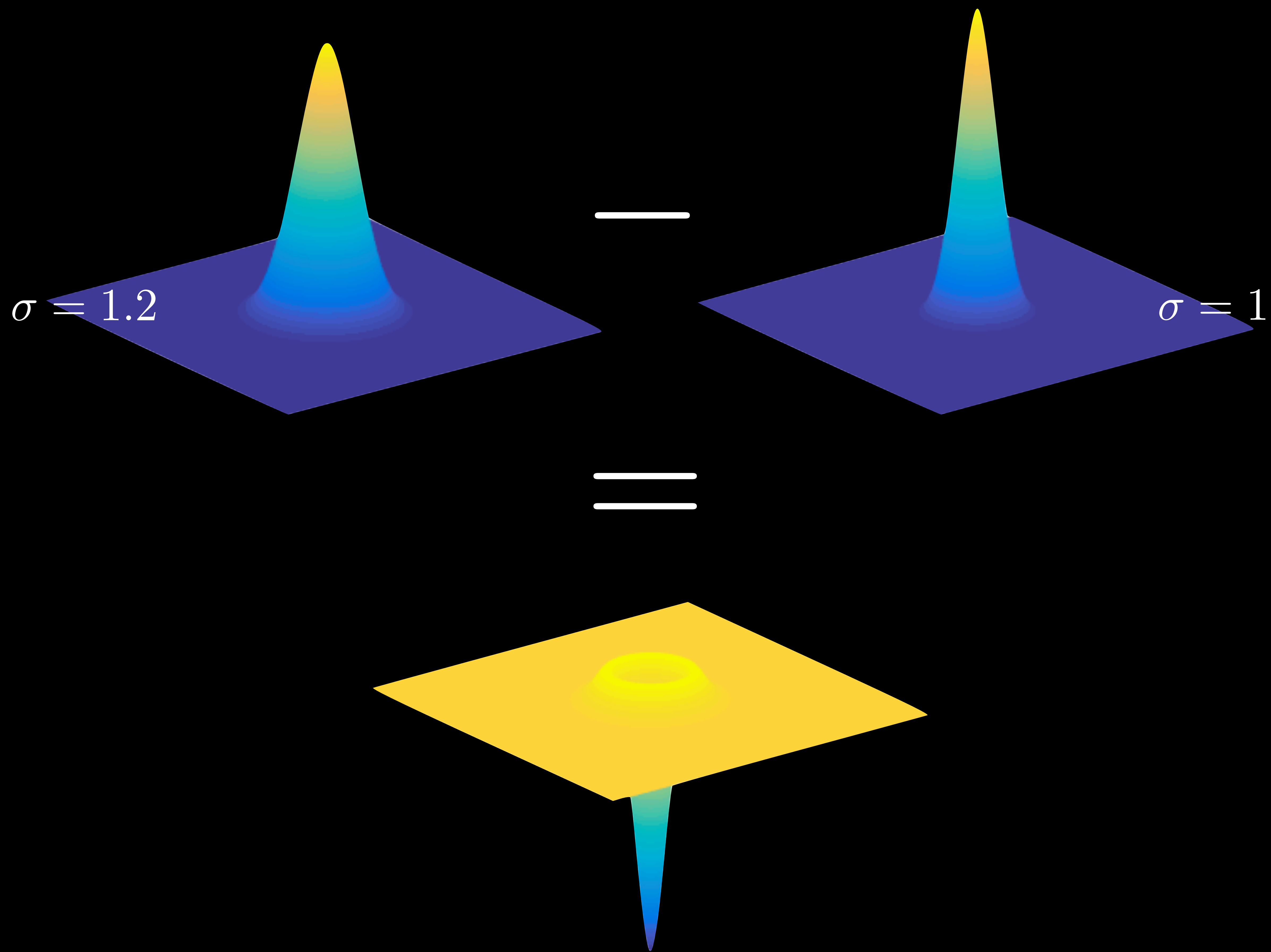
$$\approx G(x, y; k\sigma) - G(x, y; \sigma)$$

Difference of Gaussians (DoG)


$\sigma = 1.2$


$\sigma = 1$


$\sigma = 1.2$
$-$
$\sigma = 1$

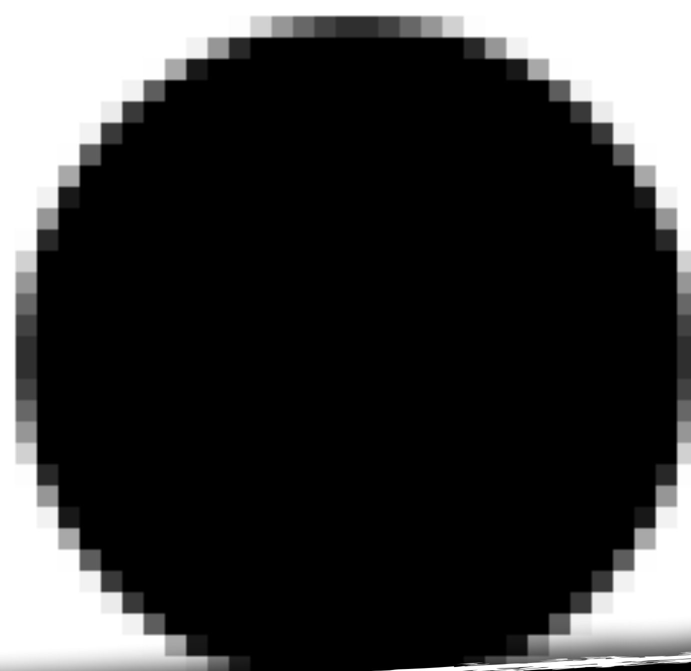

$\sigma = 1.2$
$\sigma = 1$

Blob Detection

Blob Detection

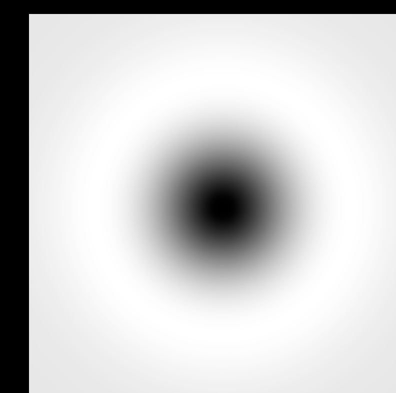
Convolve image with “blob” filter at multiple scales and find extrema

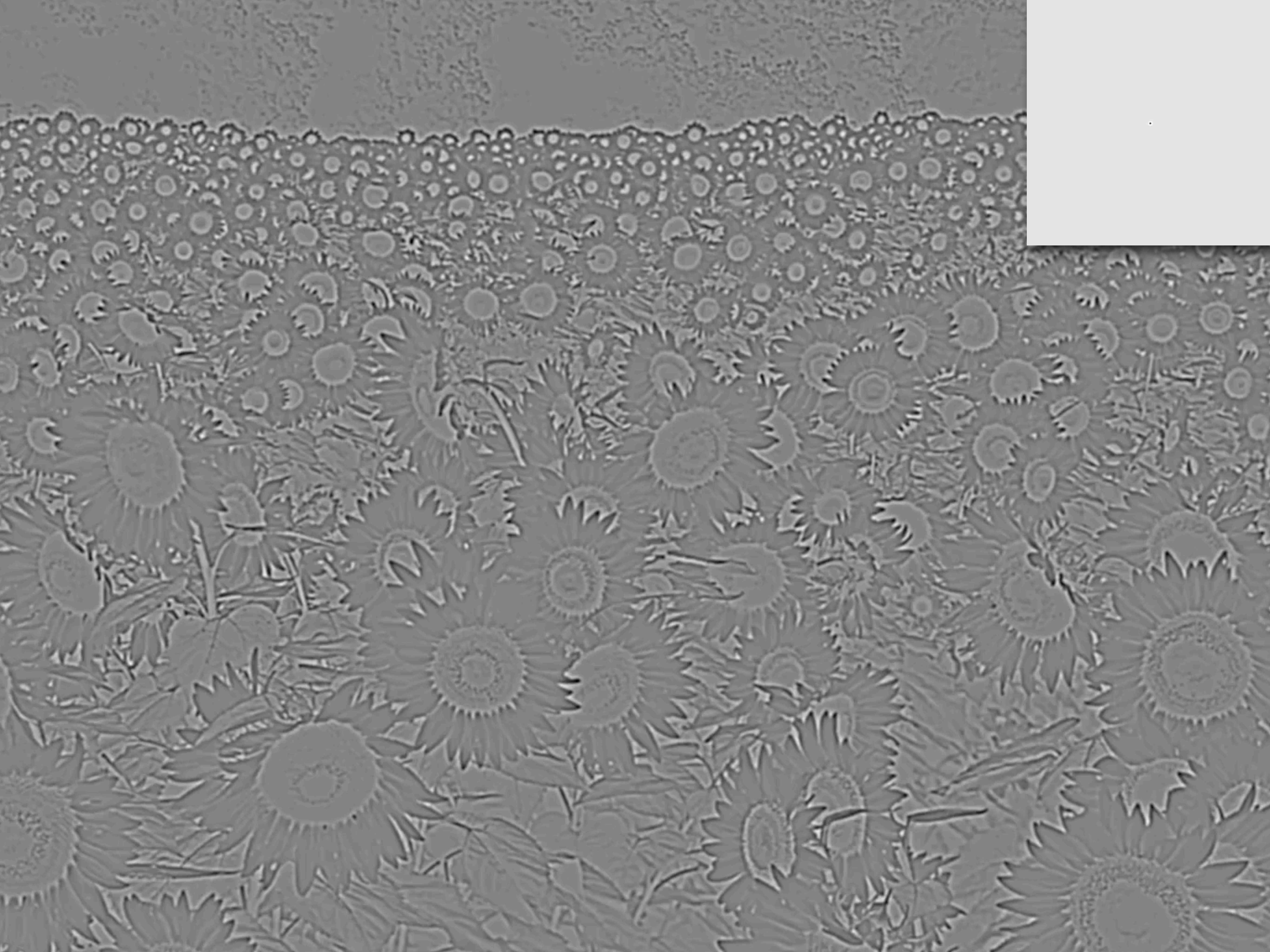
Laplacian extreme response is at $\sigma = \mathrm{radius}/\sqrt{2}$



Laplacian extreme response is at $\sigma = \text{radius}/\sqrt{2}$

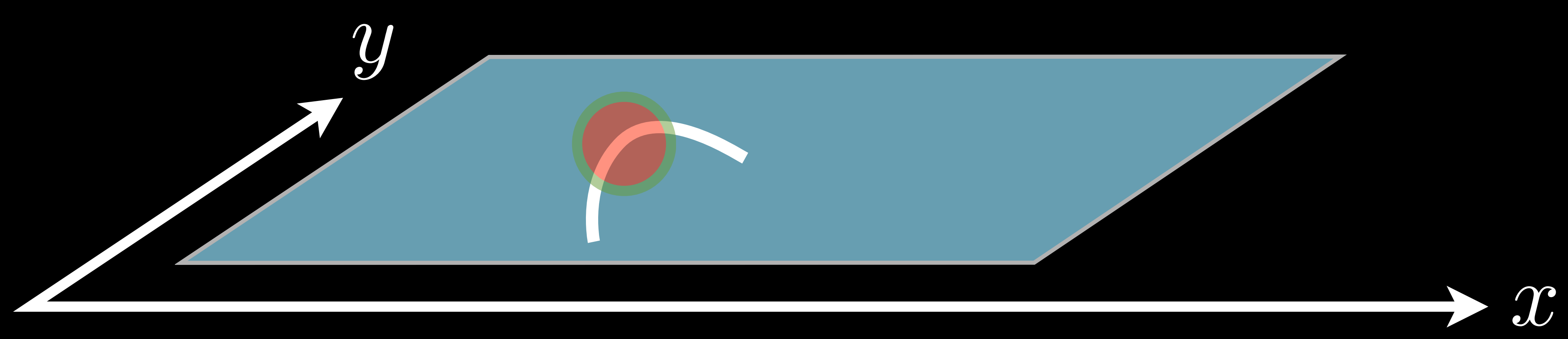
*



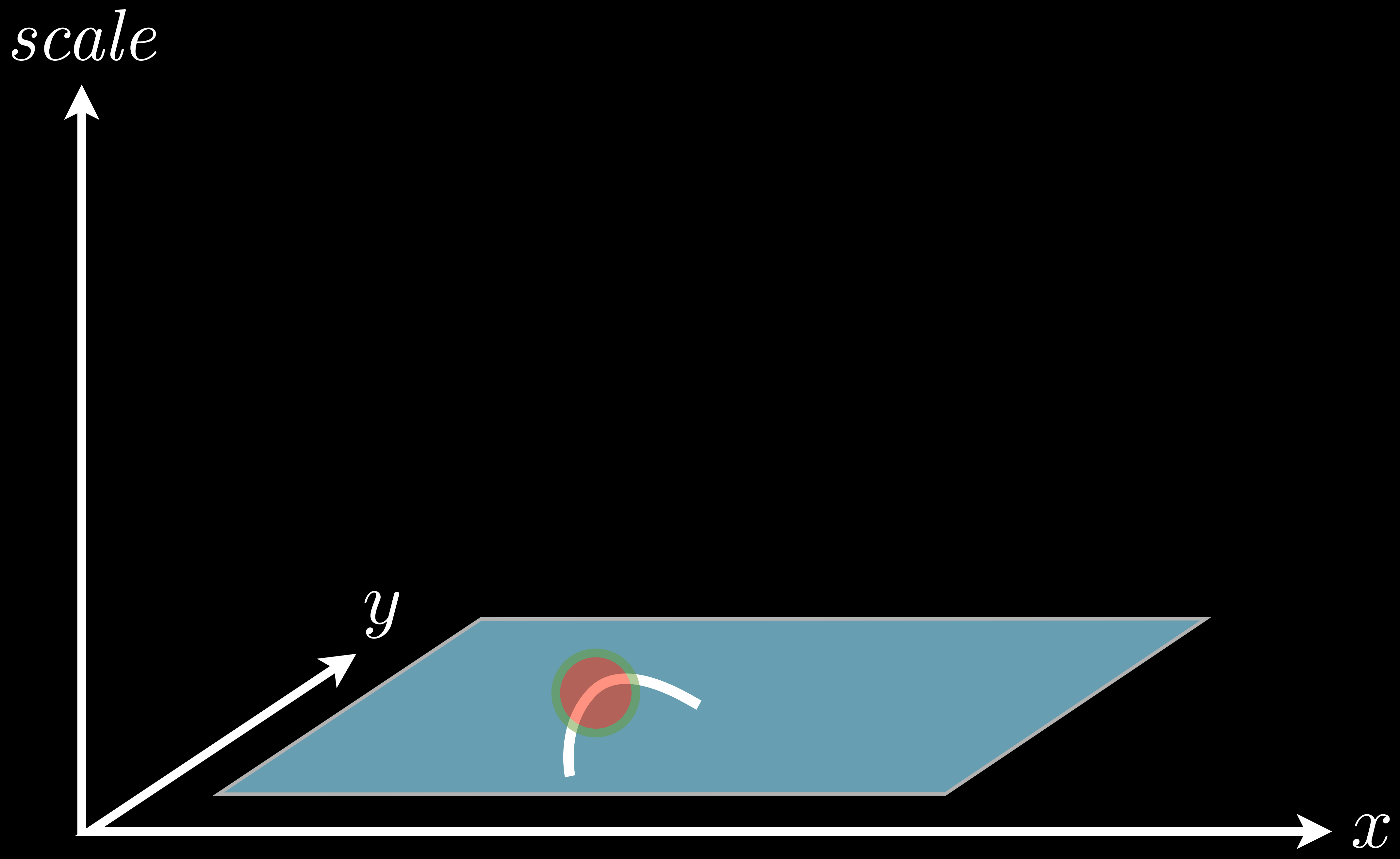



y
x

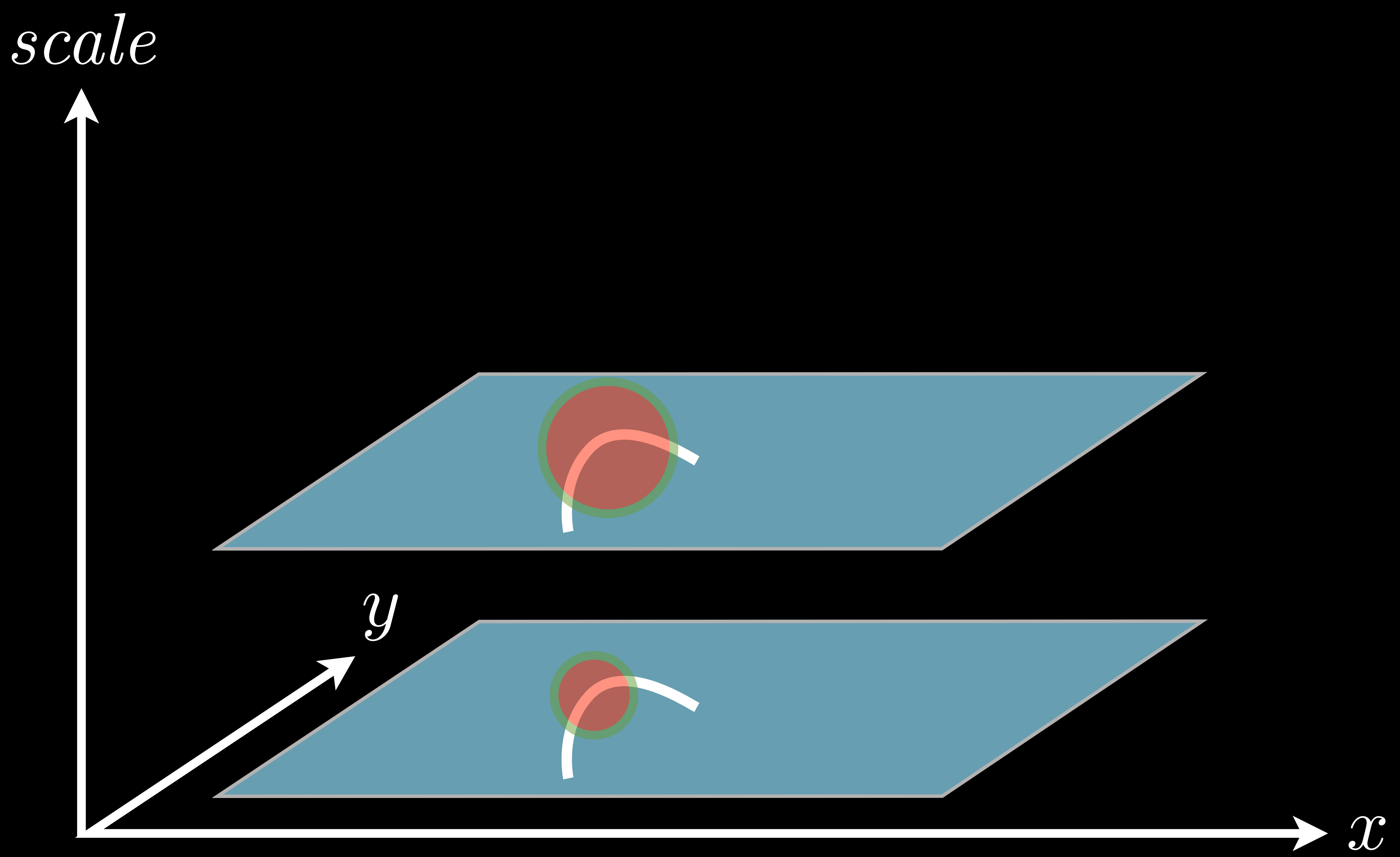
Convolve DoG across space at several scales



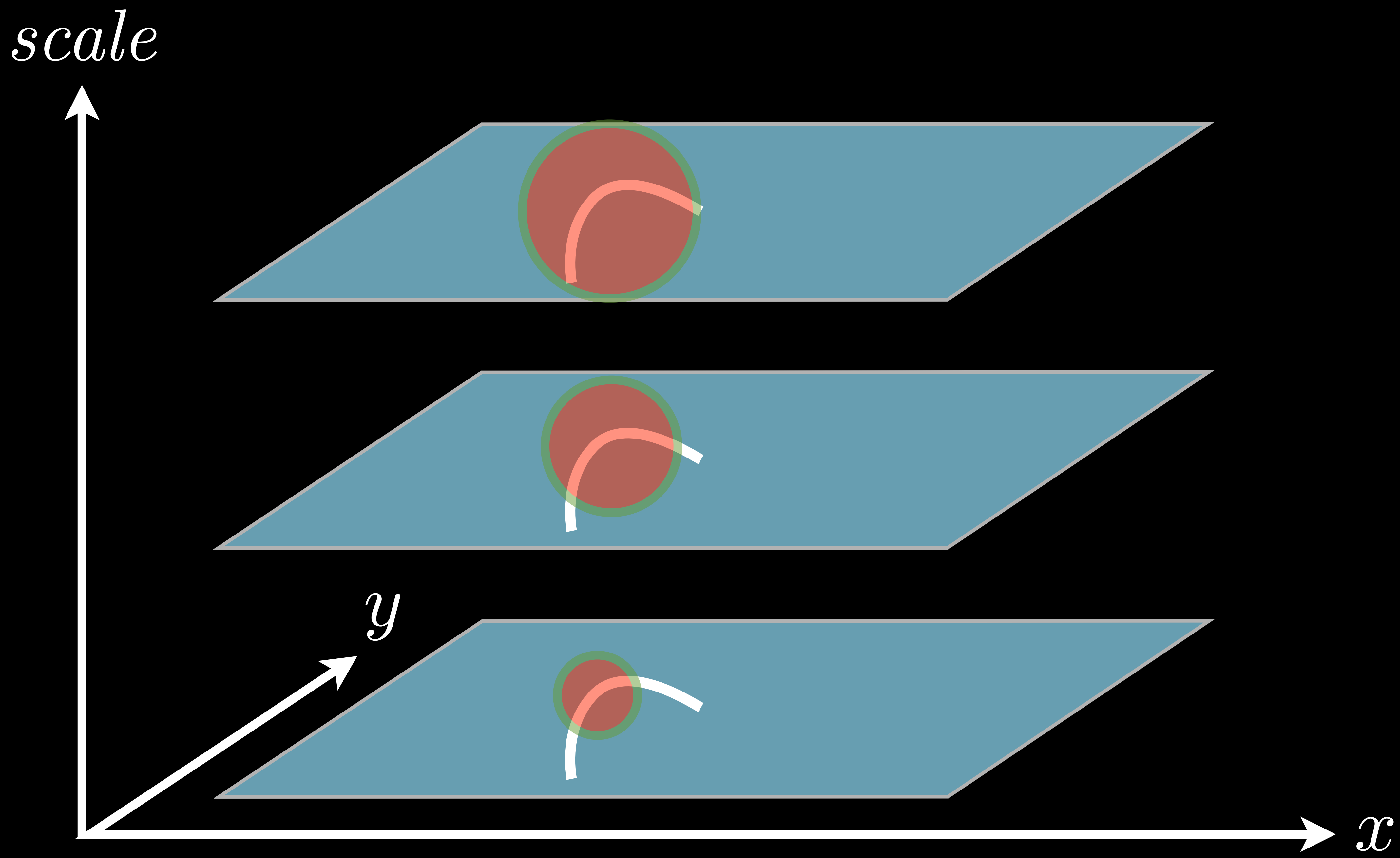
Convolve DoG across space at several scales



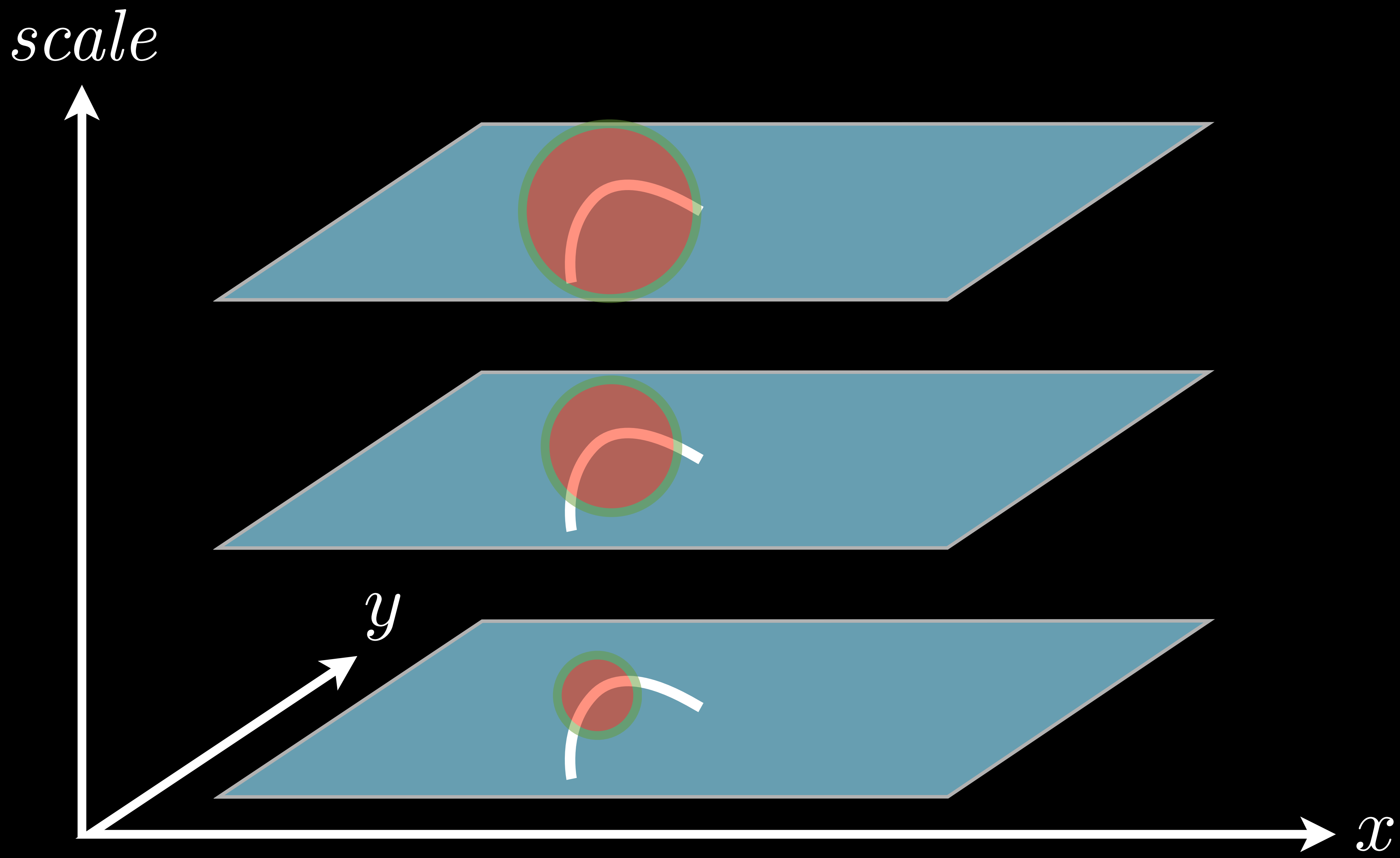
Convolve DoG across space at several scales



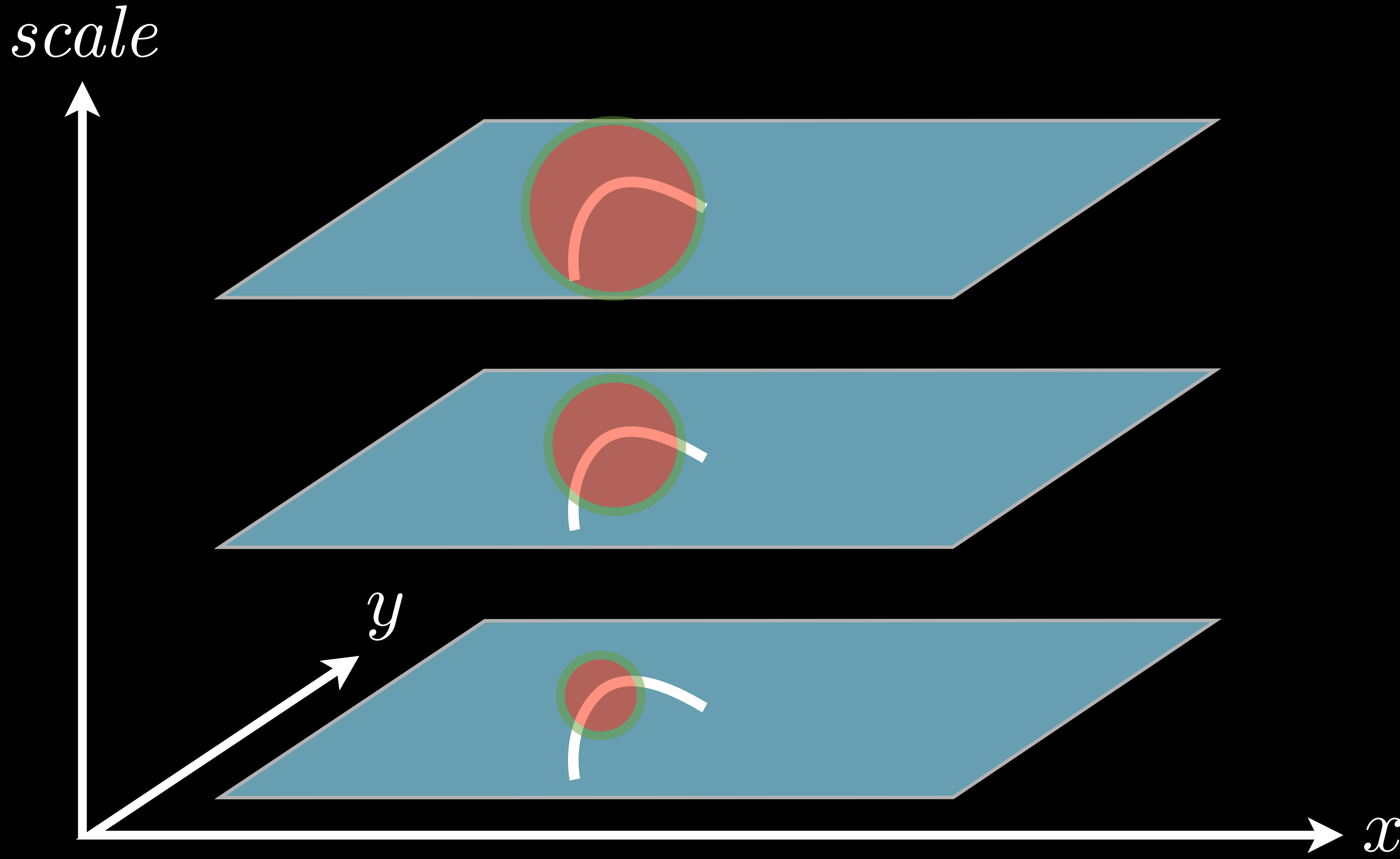
Convolve DoG across space at several scales



Convolve DoG across space at several scales

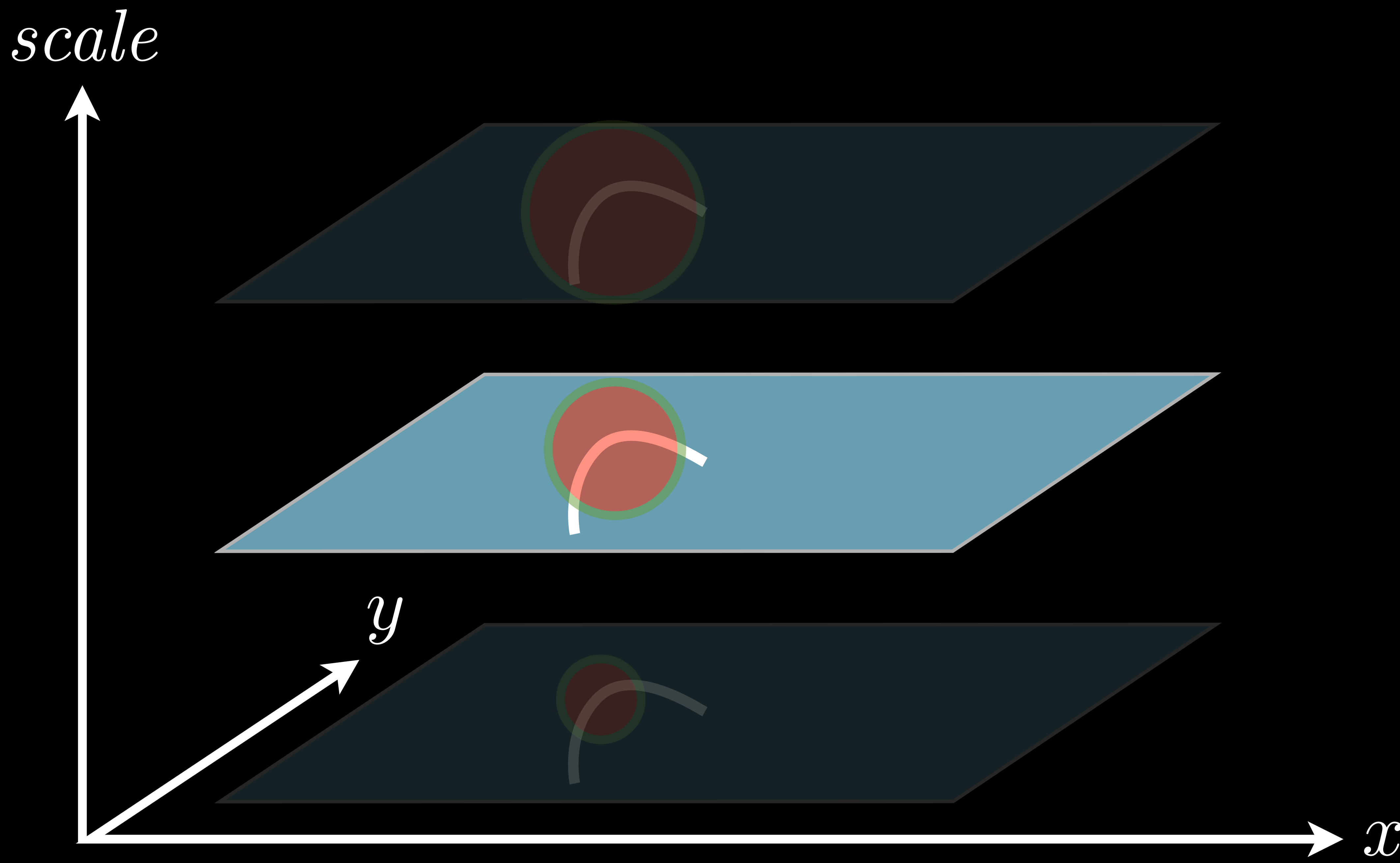


Find local maxima of DoG in space



Find local maxima of DoG in space

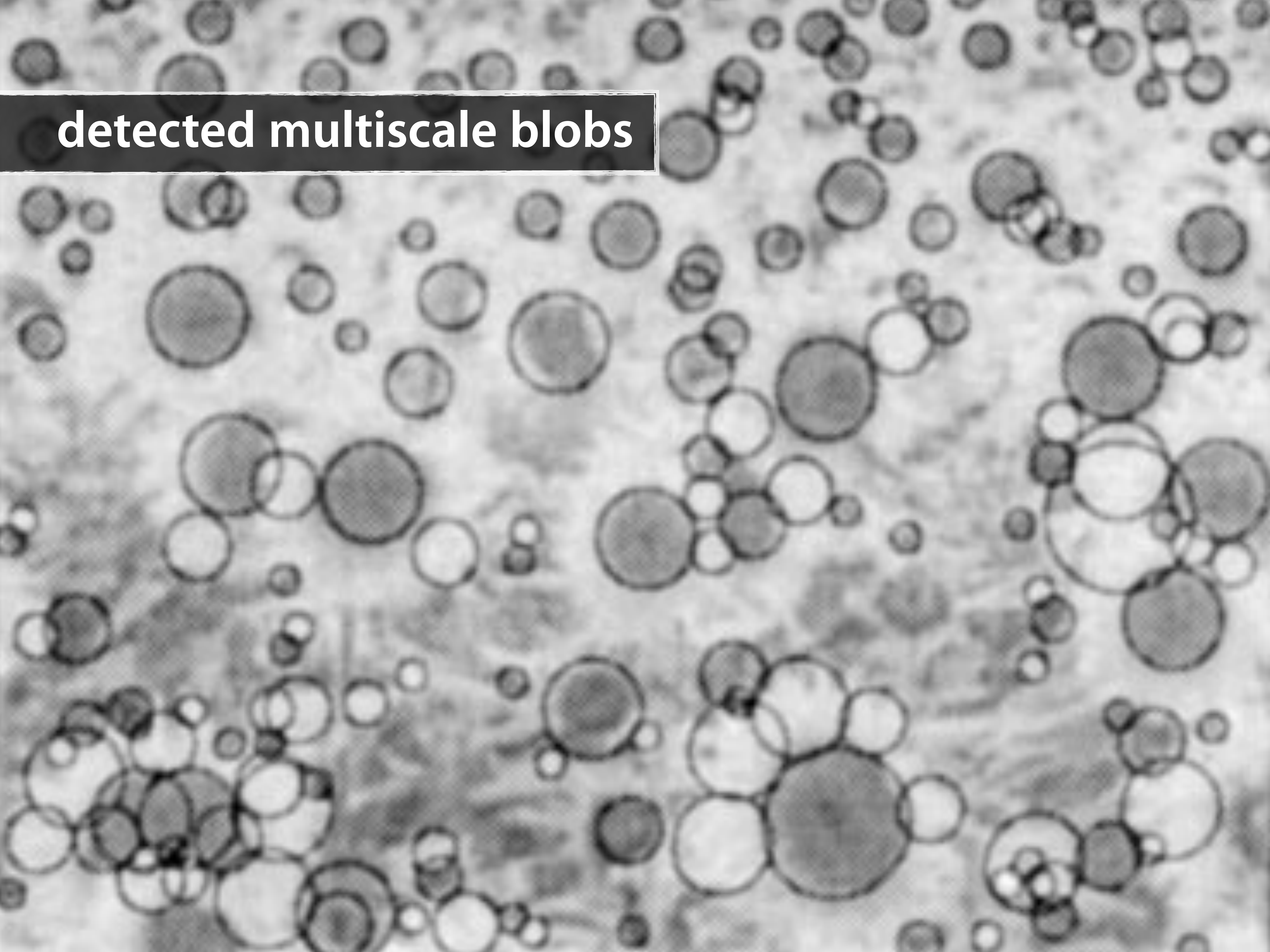
and scale

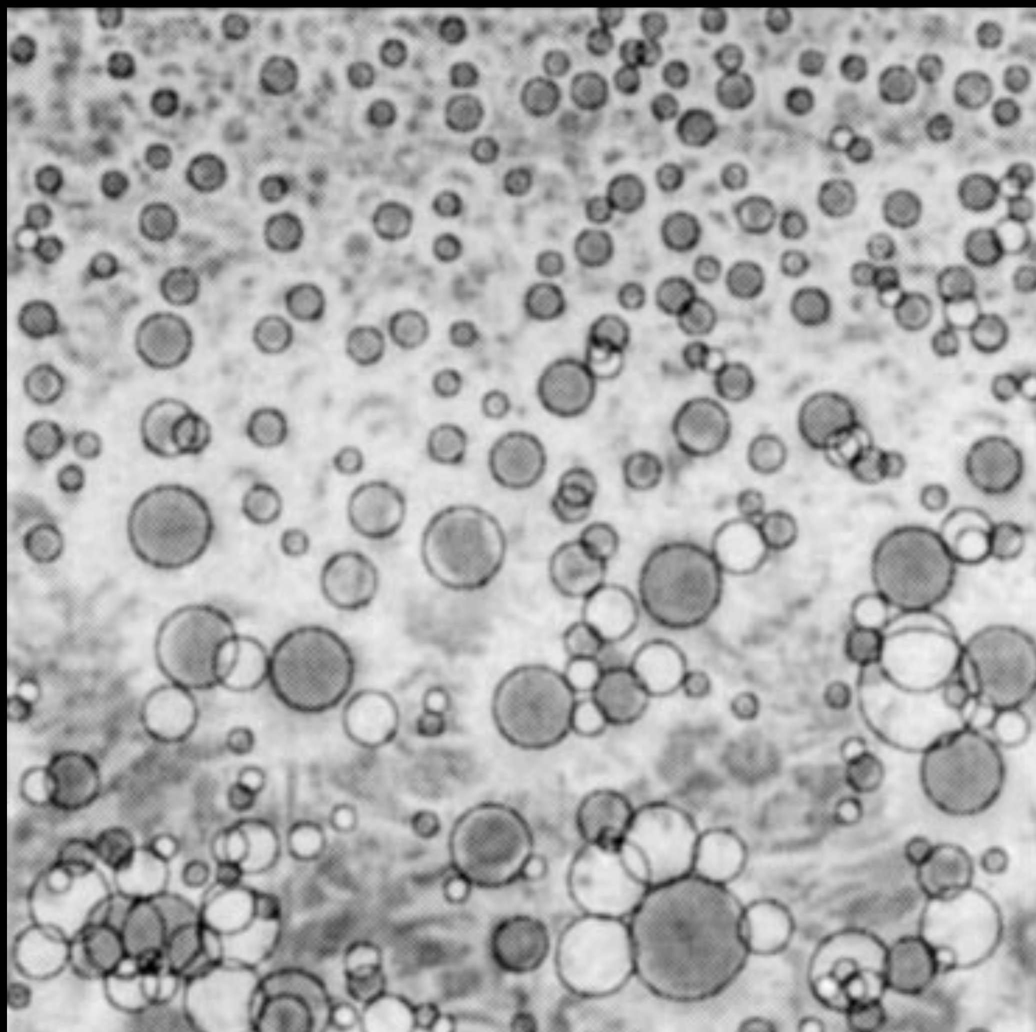


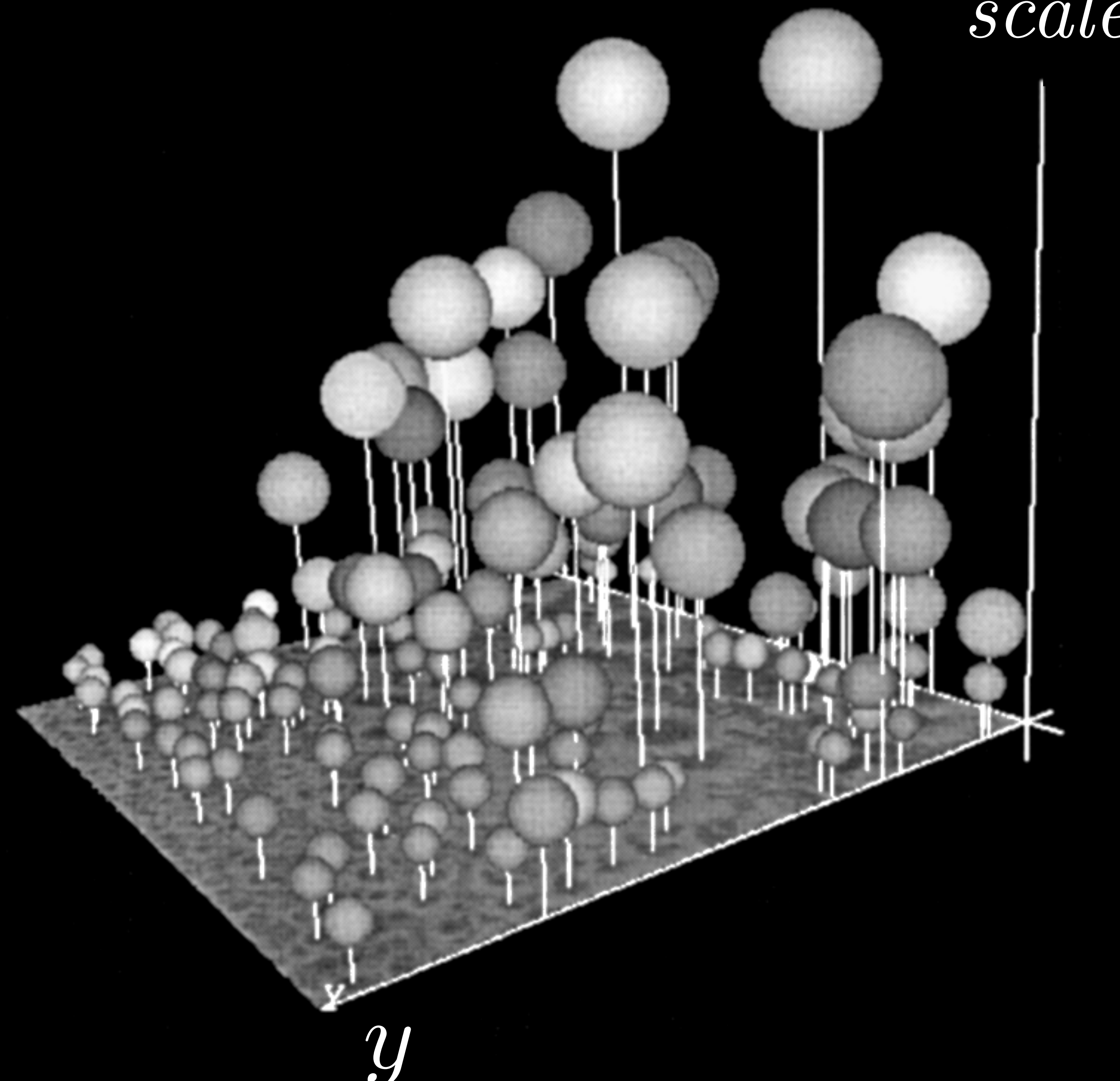
Find local maxima of DoG in space

and scale


input image


detected multiscale blobs

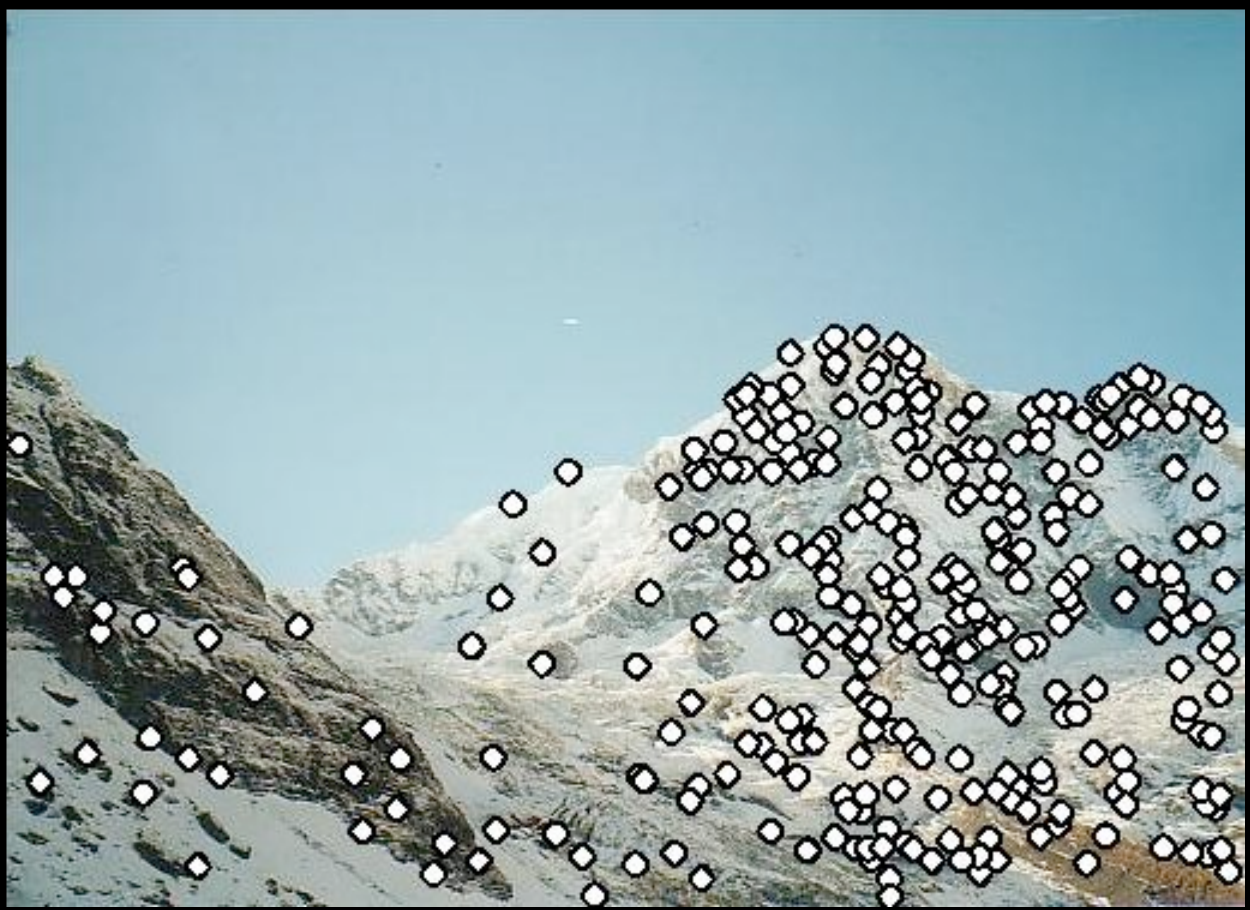
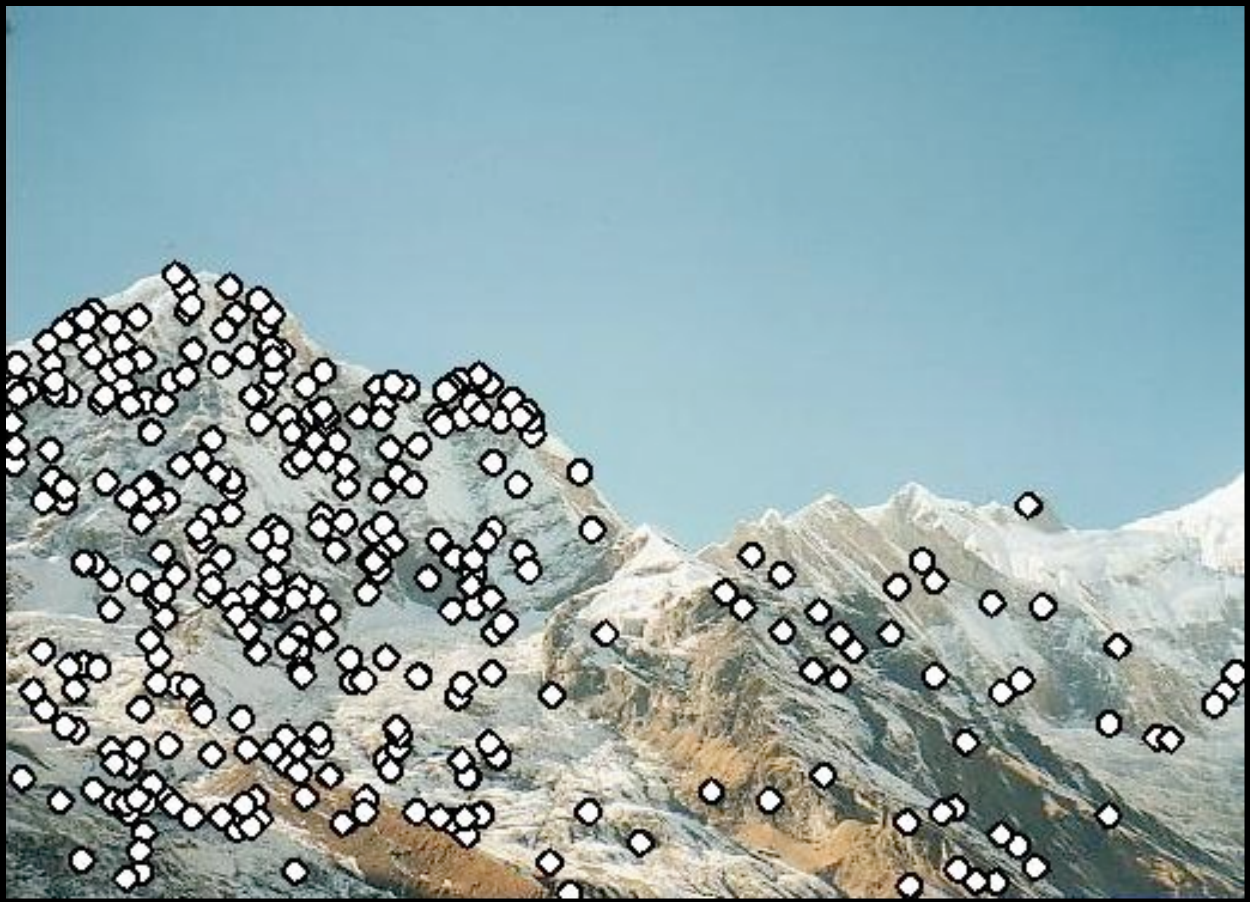


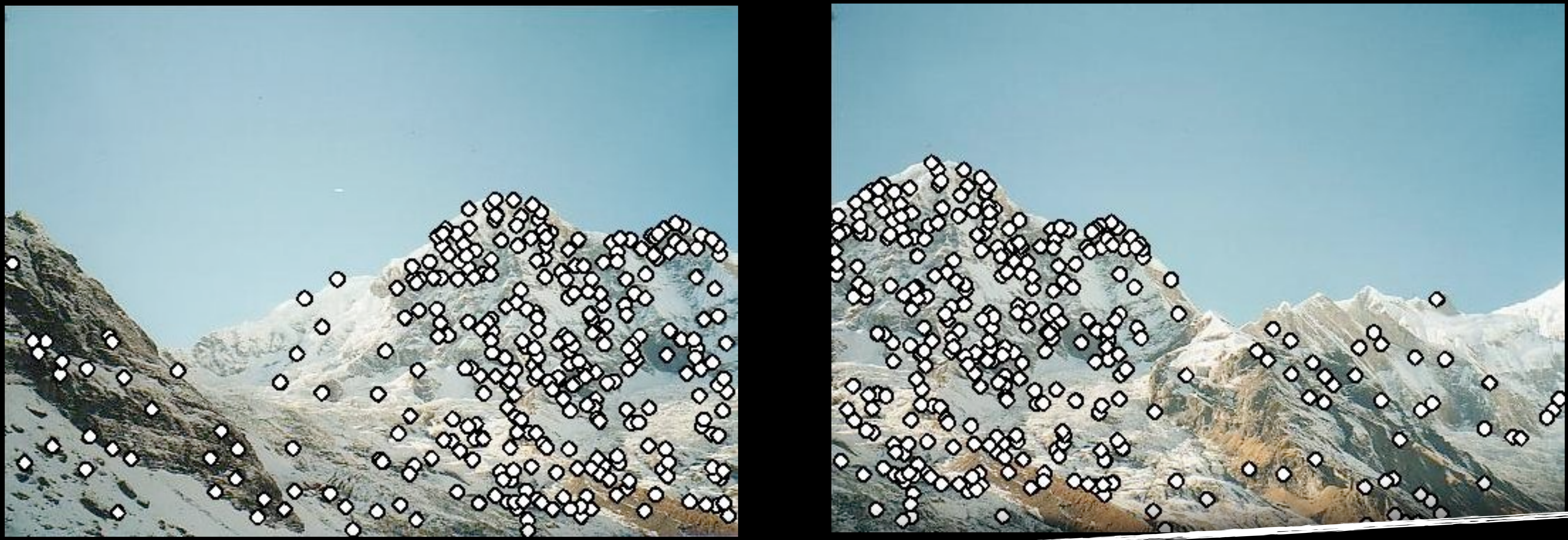


Lindeberg, Feature Detection with Automatic Scale Selection, IJCV, 1998

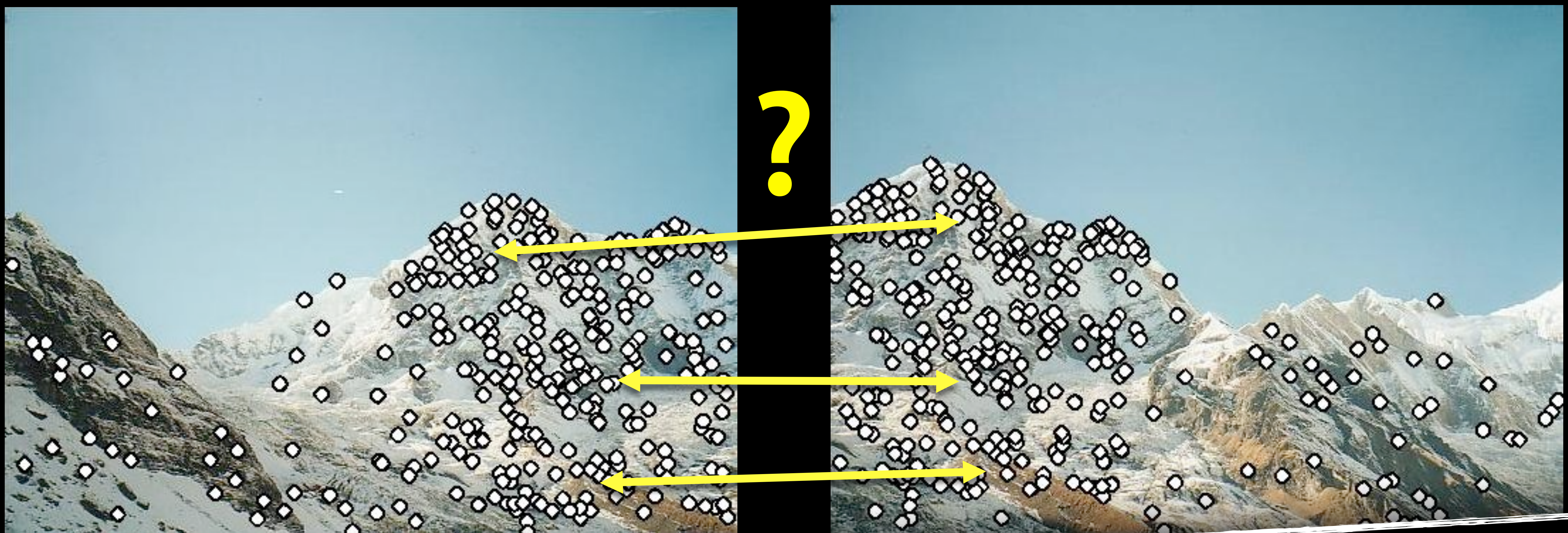
SIFT

Feature Descriptor



How do we match detected features across images?

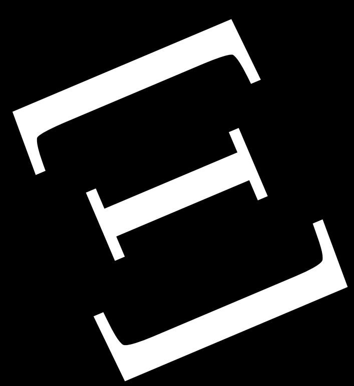


How do we match detected features across images?

Rotation Invariant
Description

Step 1: Find dominant orientation of the patch

Step 1: Find dominant orientation of the patch

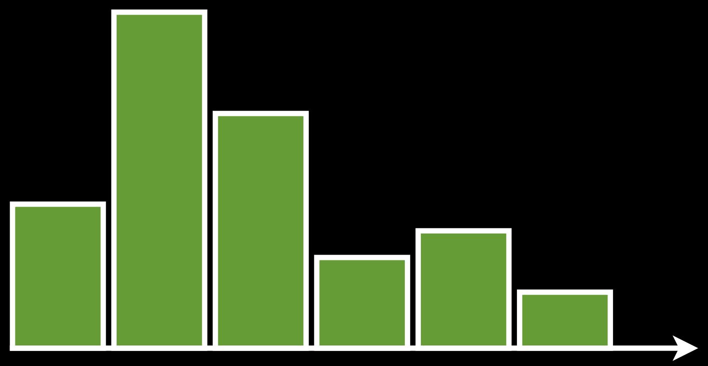


Rotation Invariant Description

Step 1: Find dominant orientation of the patch

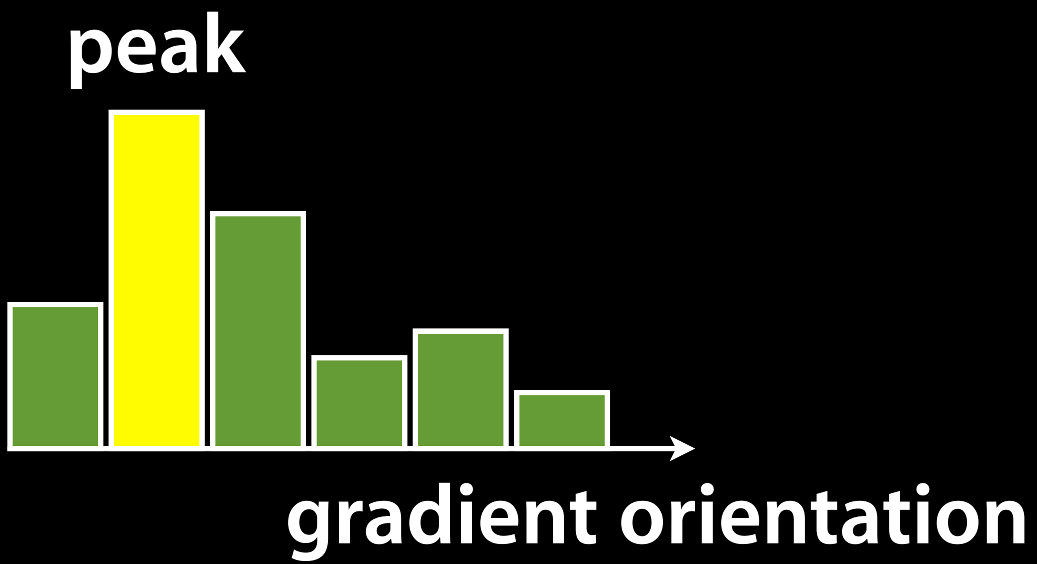
Rotation Invariant Description

Step 1: Find dominant orientation of the patch



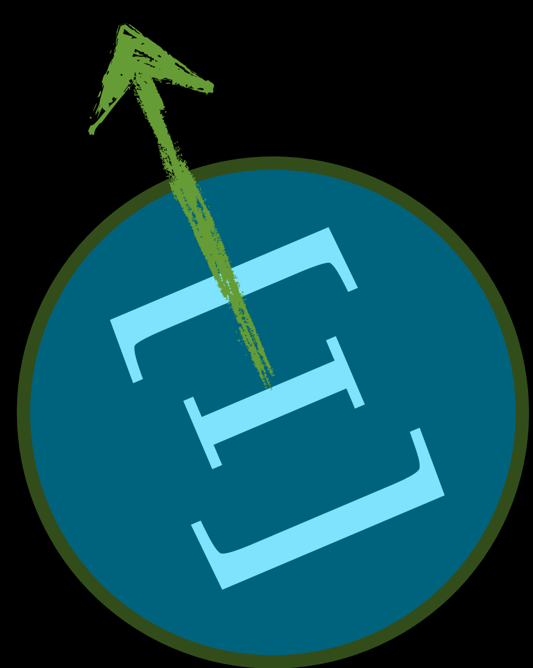
Rotation Invariant Description

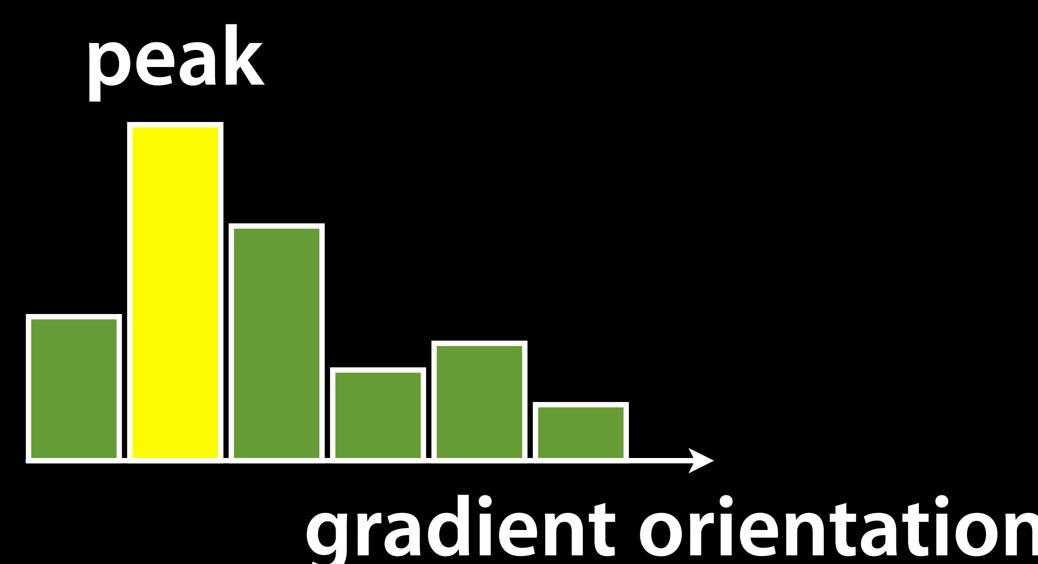
Step 1: Find dominant orientation of the patch



Rotation Invariant Description

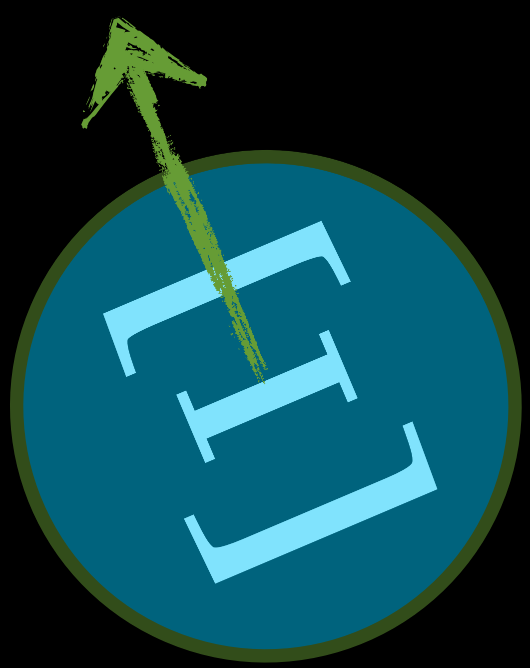
Step 1: Find dominant orientation of the patch

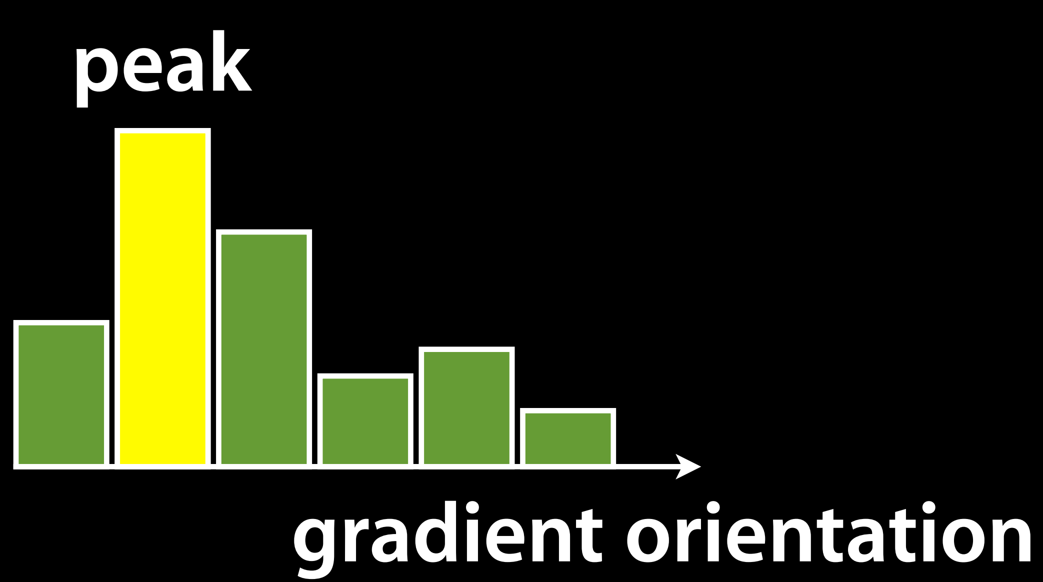




Rotation Invariant Description

Step 1: Find dominant orientation of the patch

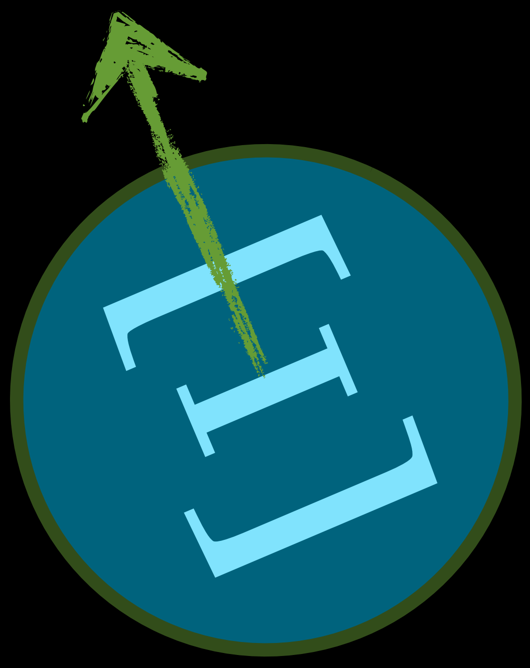


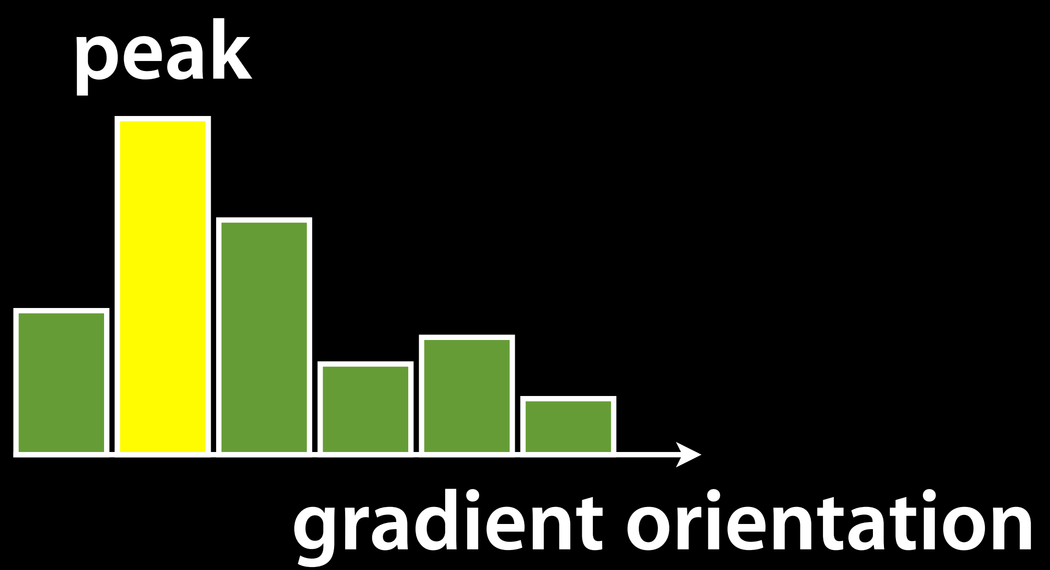


Rotation Invariant Description

Step 1: Find dominant orientation of the patch

Step 2: Rotate the patch to point along x-axis

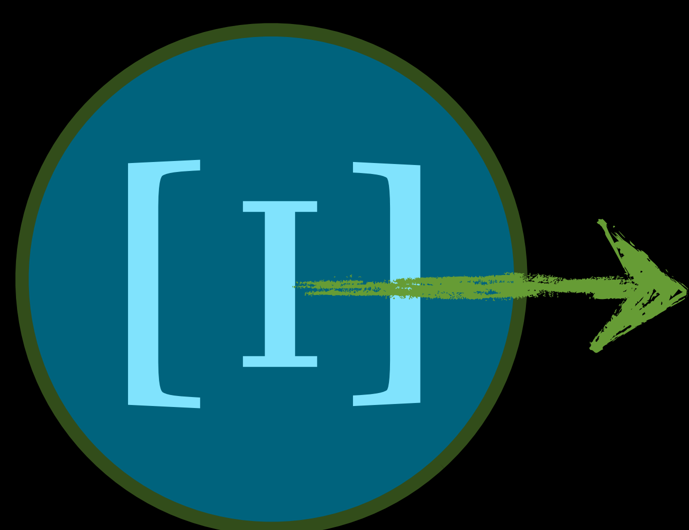


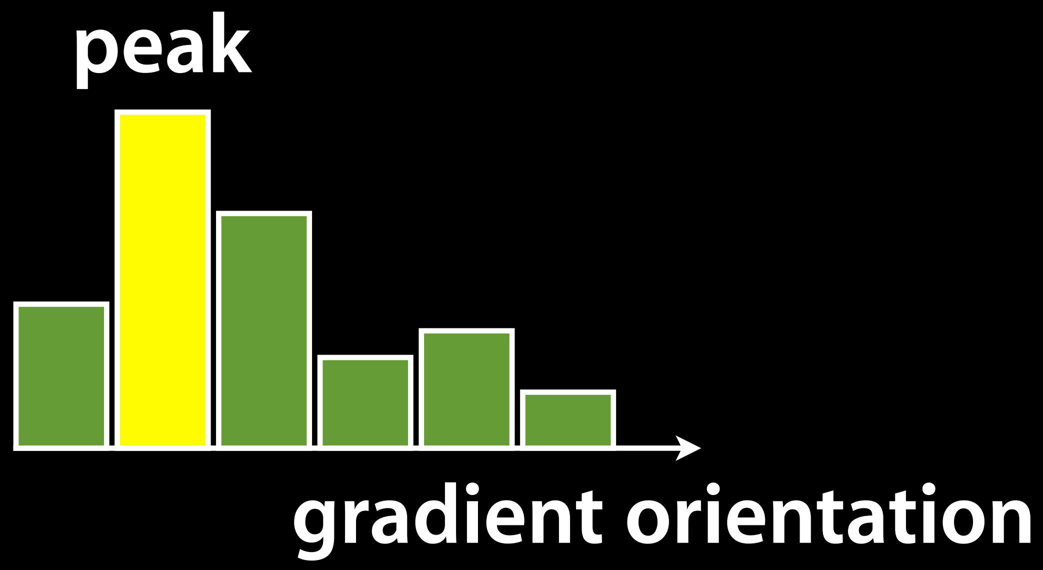


Rotation Invariant Description

Step 1: Find dominant orientation of the patch

Step 2: Rotate the patch to point along x-axis





SIFT Descriptor

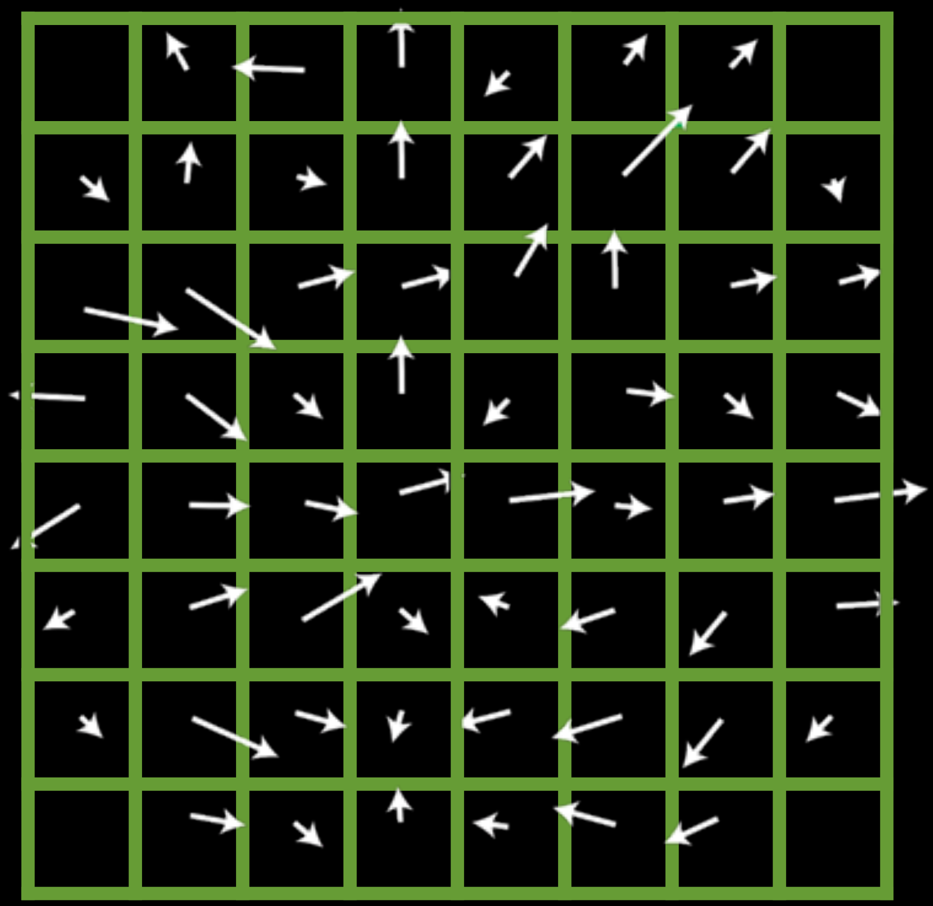


image gradients

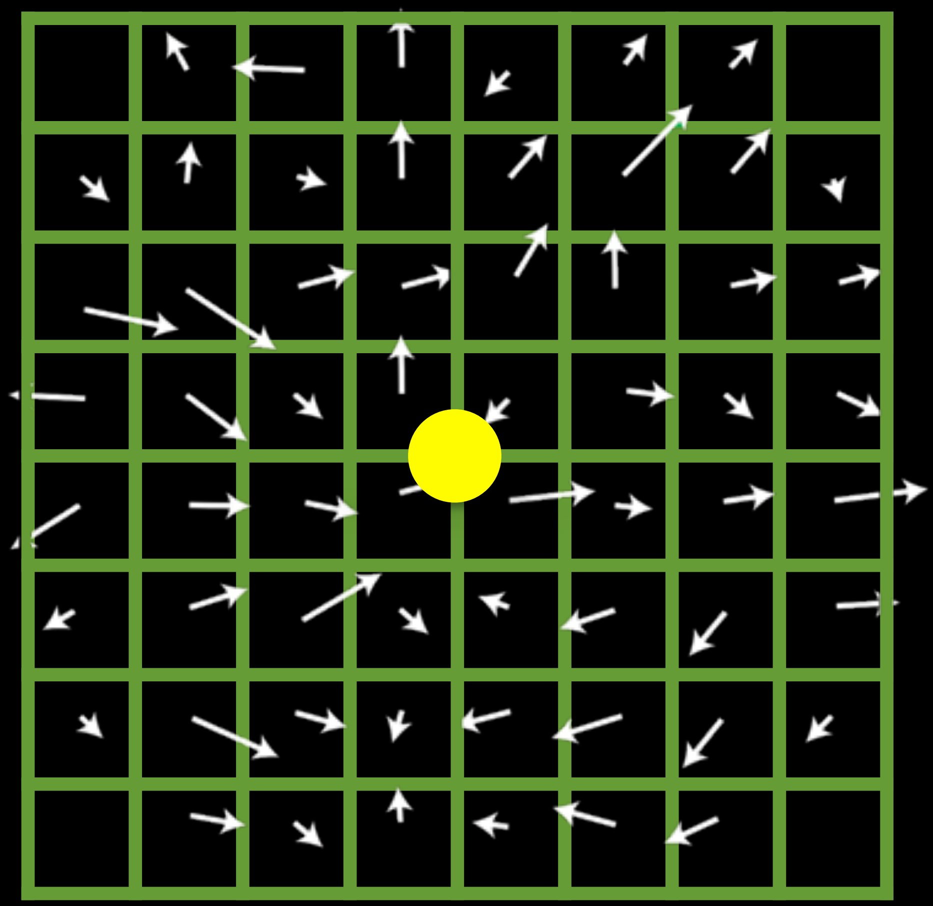


image gradients

SIFT Descriptor

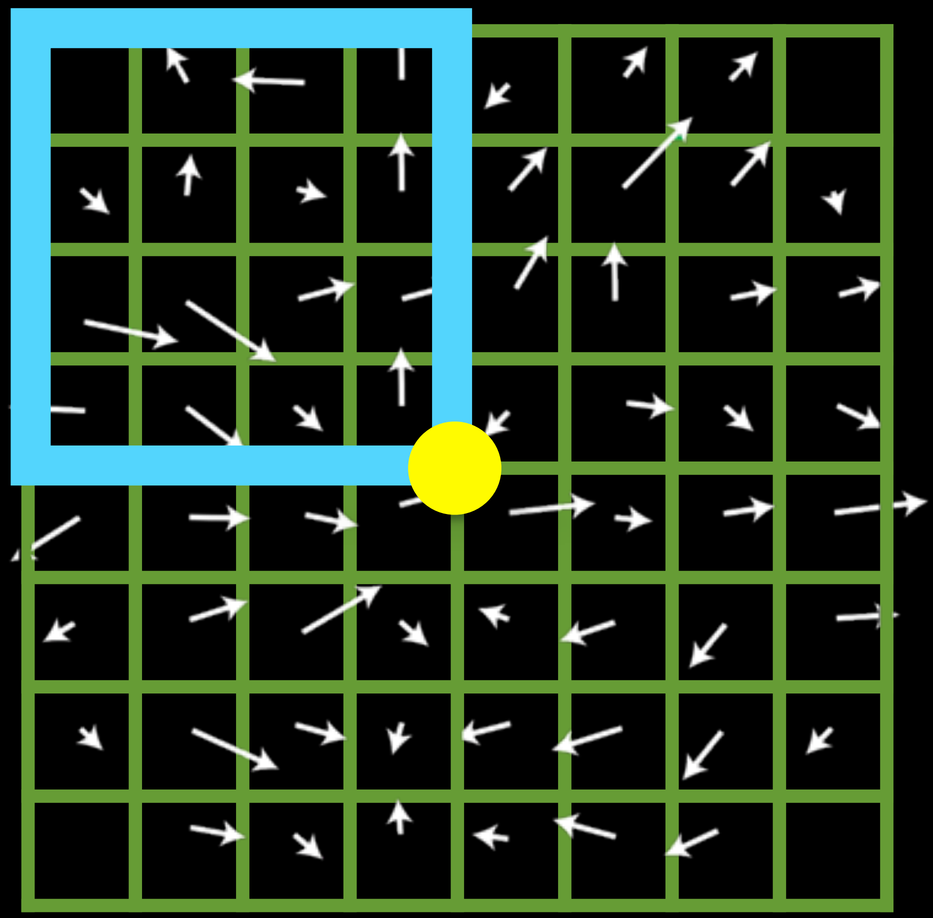


image gradients

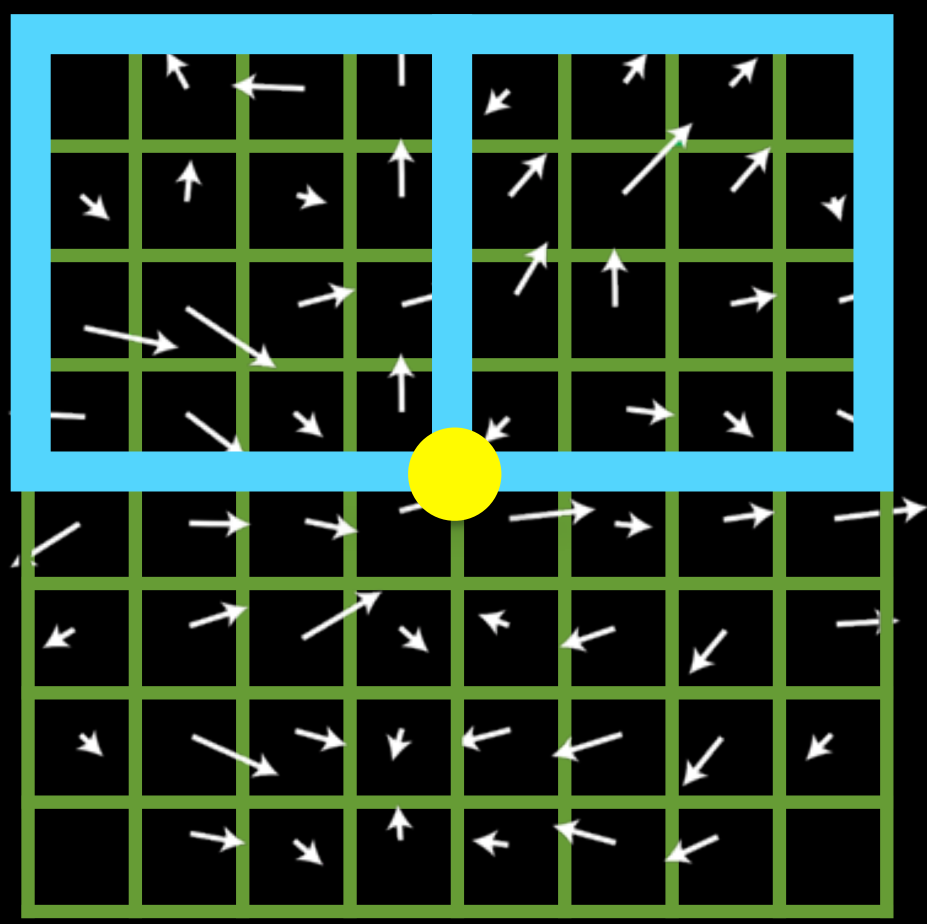


image gradients

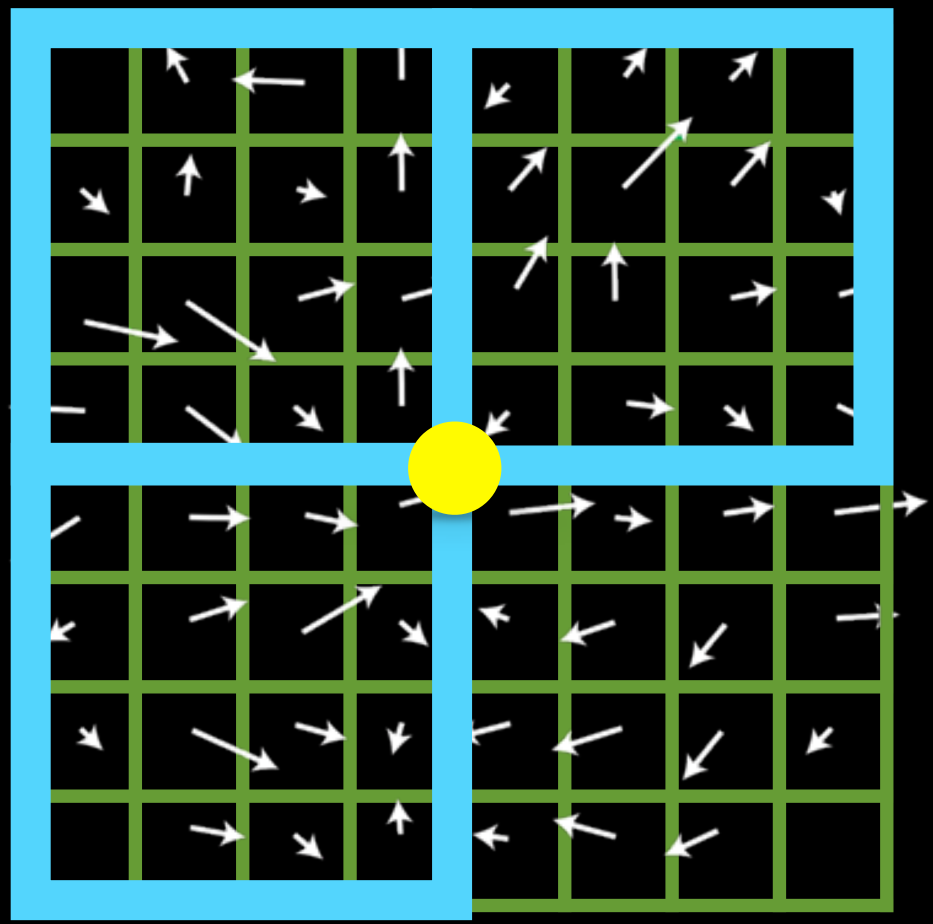


image gradients

SIFT Descriptor

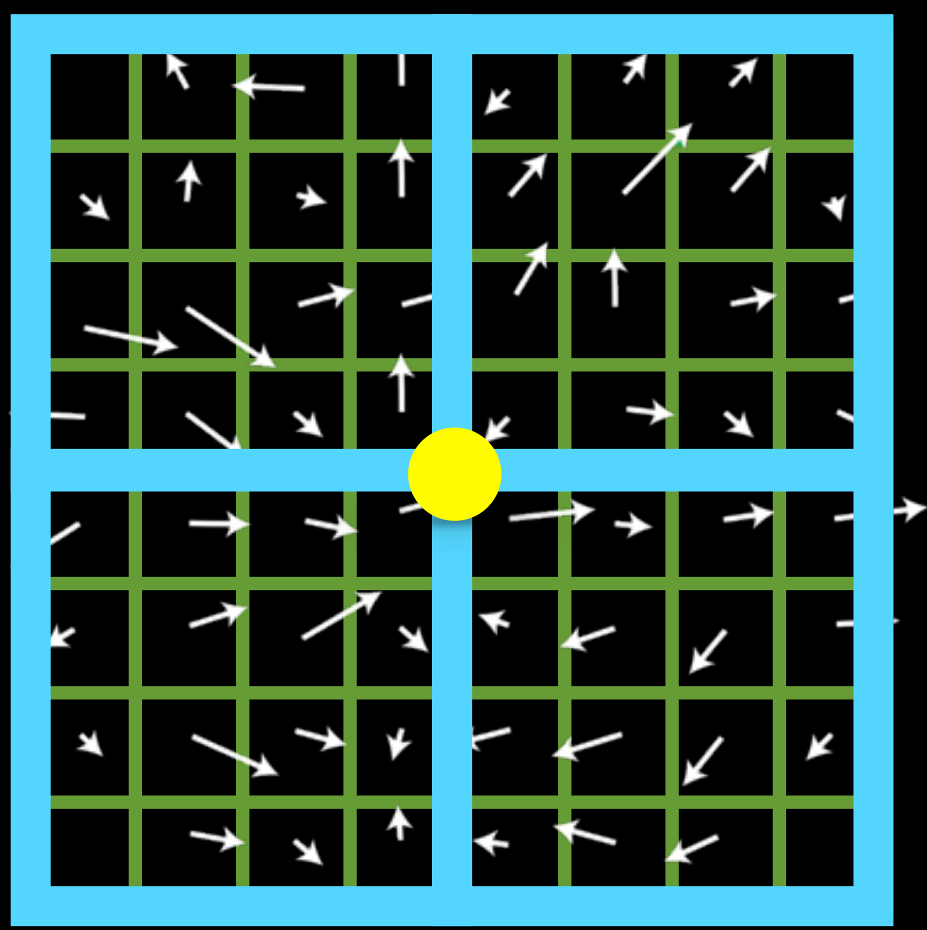


image gradients

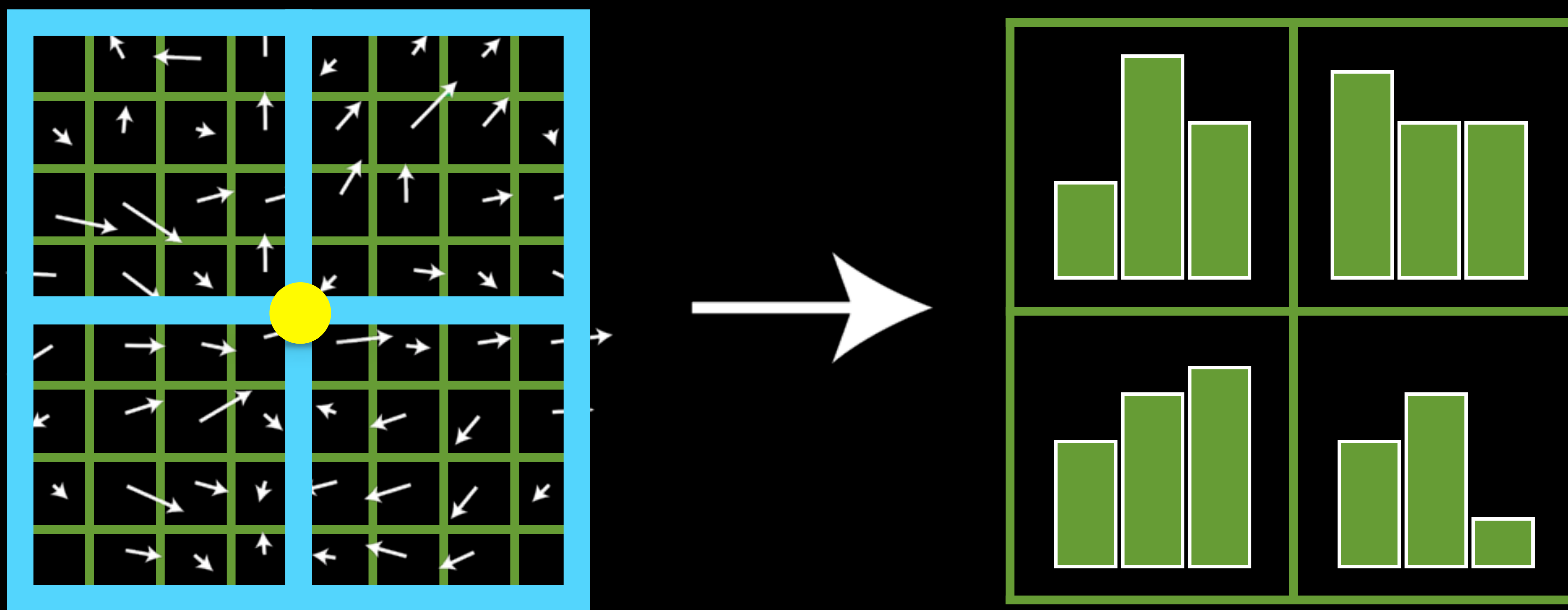
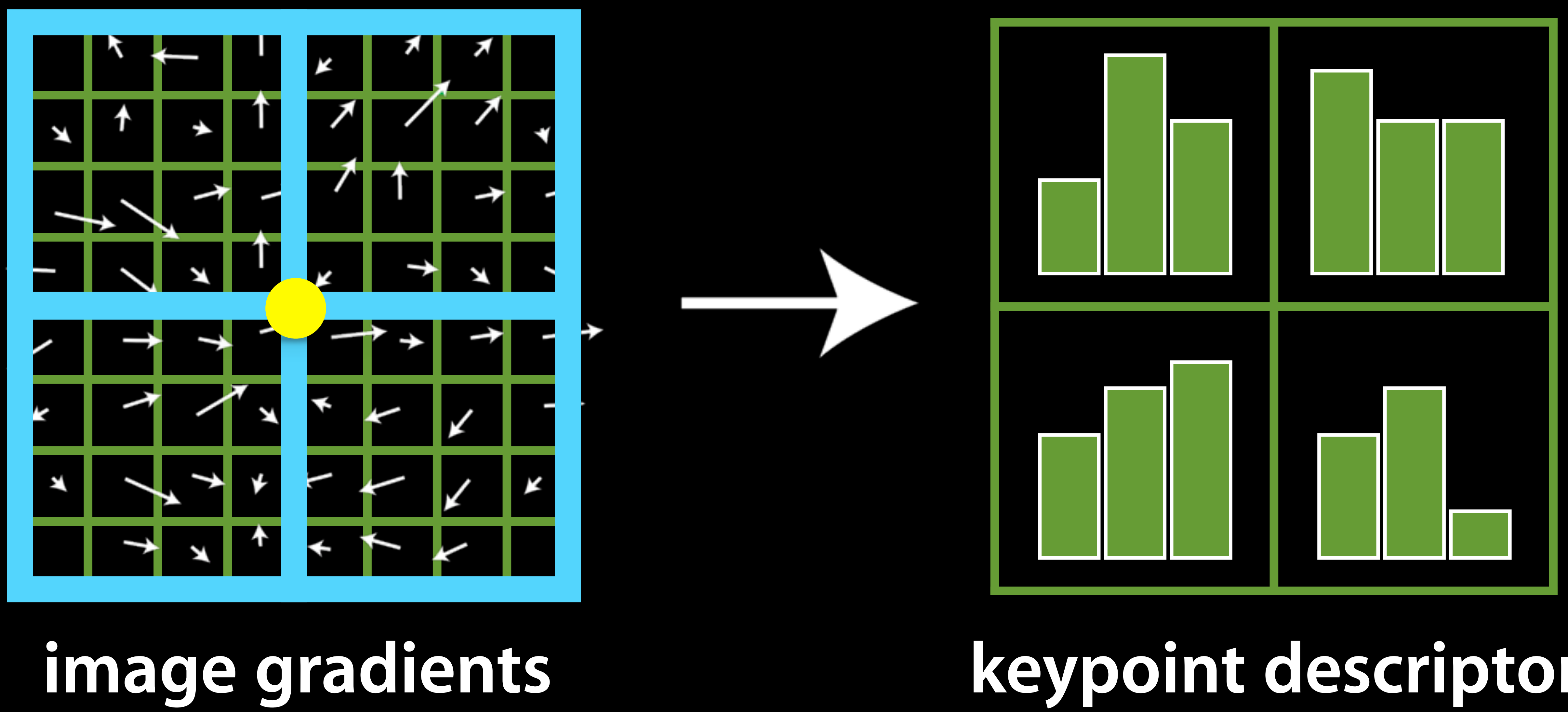
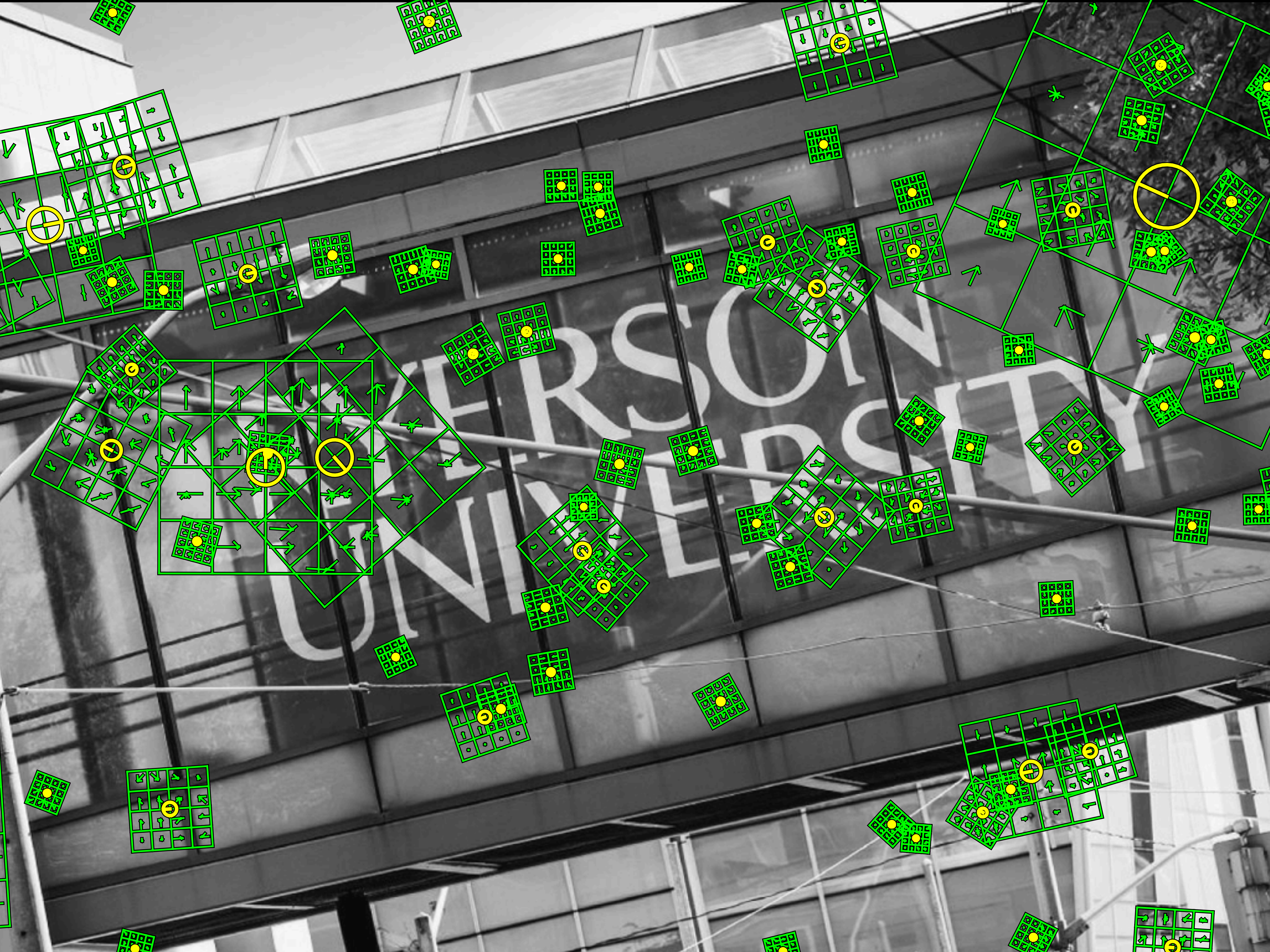


image gradients **keypoint descriptor**



16 cells x 8 bins = 128 dimensional descriptor


RYERSON
UNIVERSITY


ERSON
UNIVERSITY


RAC
NO

SIFT Properties

SIFT Properties

Robust to image rotation, scale and intensity change

SIFT Properties

Robust to image rotation, scale and intensity change

Robust to moderate out of plane rotation

SIFT Properties

Robust to image rotation, scale and intensity change

Robust to moderate out of plane rotation

Fast and efficient

too long; didn’t listen

tl;dl

SIFT
tl;dl

SIFT
tl;dl

1. **Keypoint detection**

1. Keypoint detection

Search across image locations and scales for feature response extrema

1. Keypoint detection
   Search across image locations and scales for feature response extrema

2. Keypoint orientation assignment

1. **Keypoint detection**
   Search across image locations and scales for feature response extrema

2. **Keypoint orientation assignment**
   Determine best orientation(s) for each keypoint

SIFT
tl;dr

1. **Keypoint detection**

    Search across image locations and scales for feature response extrema

2. **Keypoint orientation assignment**

    Determine best orientation(s) for each keypoint

3. **Keypoint description**

1. Keypoint detection

Search across image locations and scales for feature response extrema

2. Keypoint orientation assignment

Determine best orientation(s) for each keypoint

3. Keypoint description

Describe keypoint region at selected scale and rotation with image gradients

1. Keypoint detection

Search across image locations and scales for feature response extrema

2. Keypoint orientation assignment

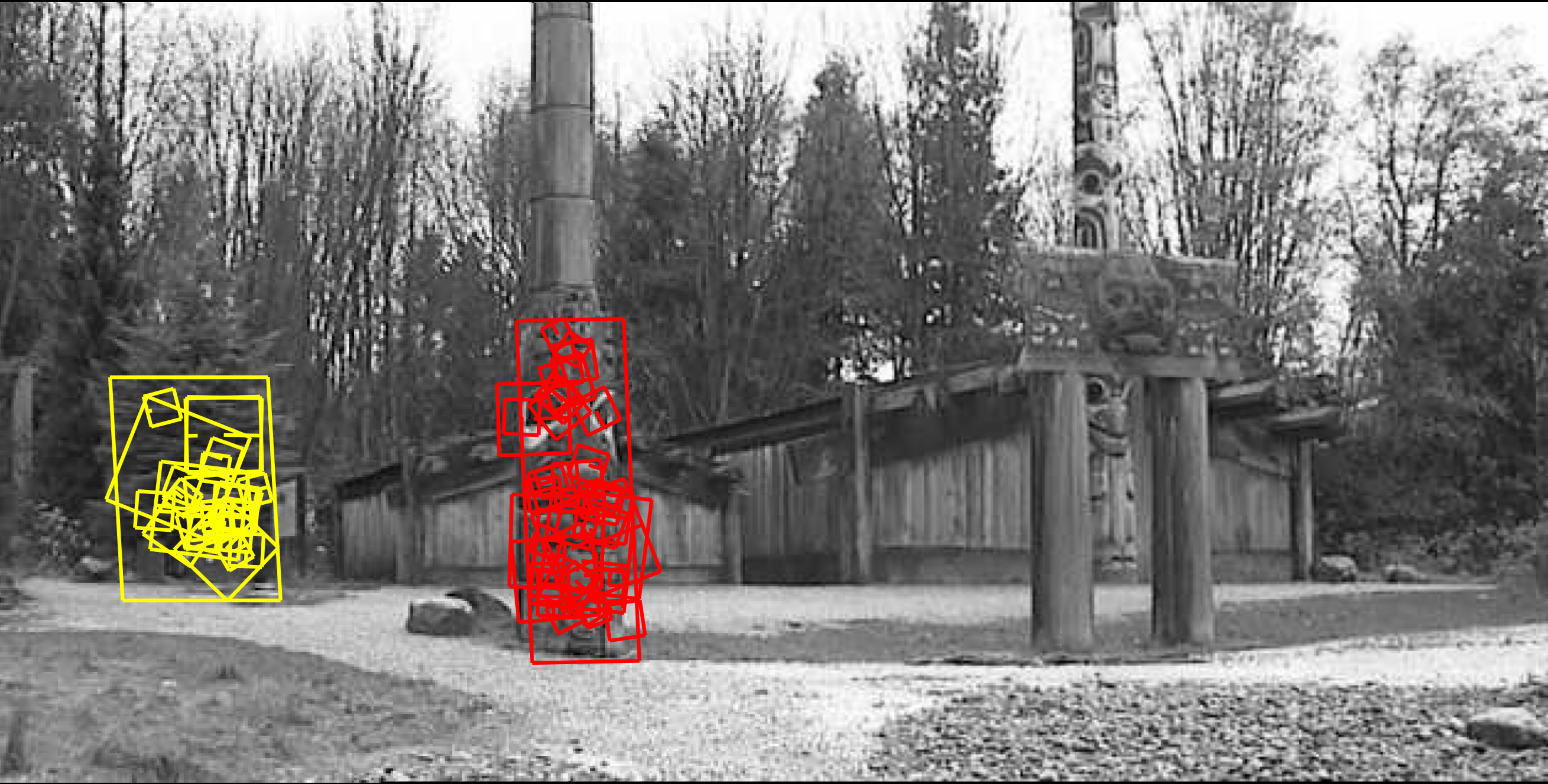
Determine best orientation(s) for each keypoint

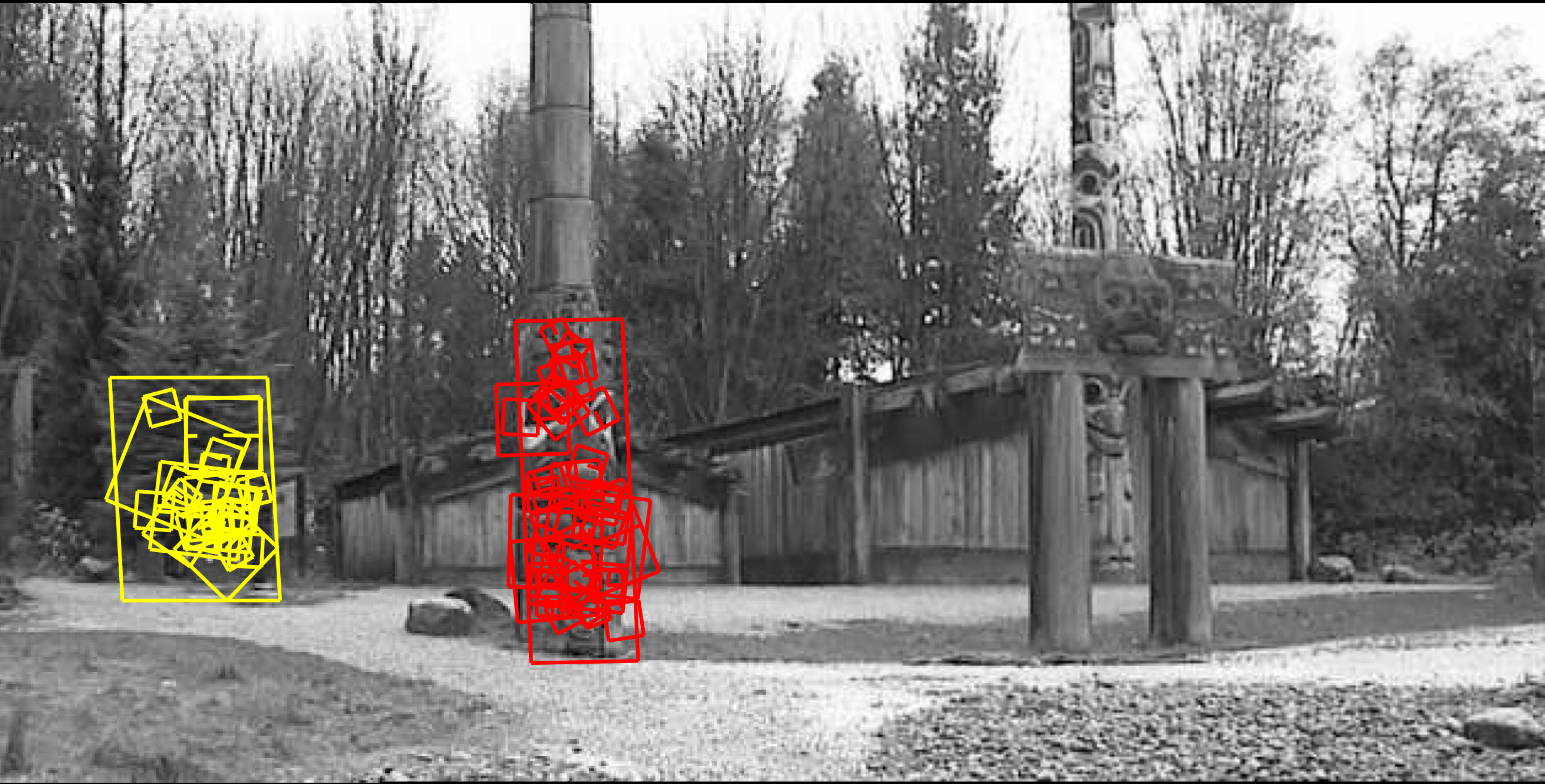
3. Keypoint description

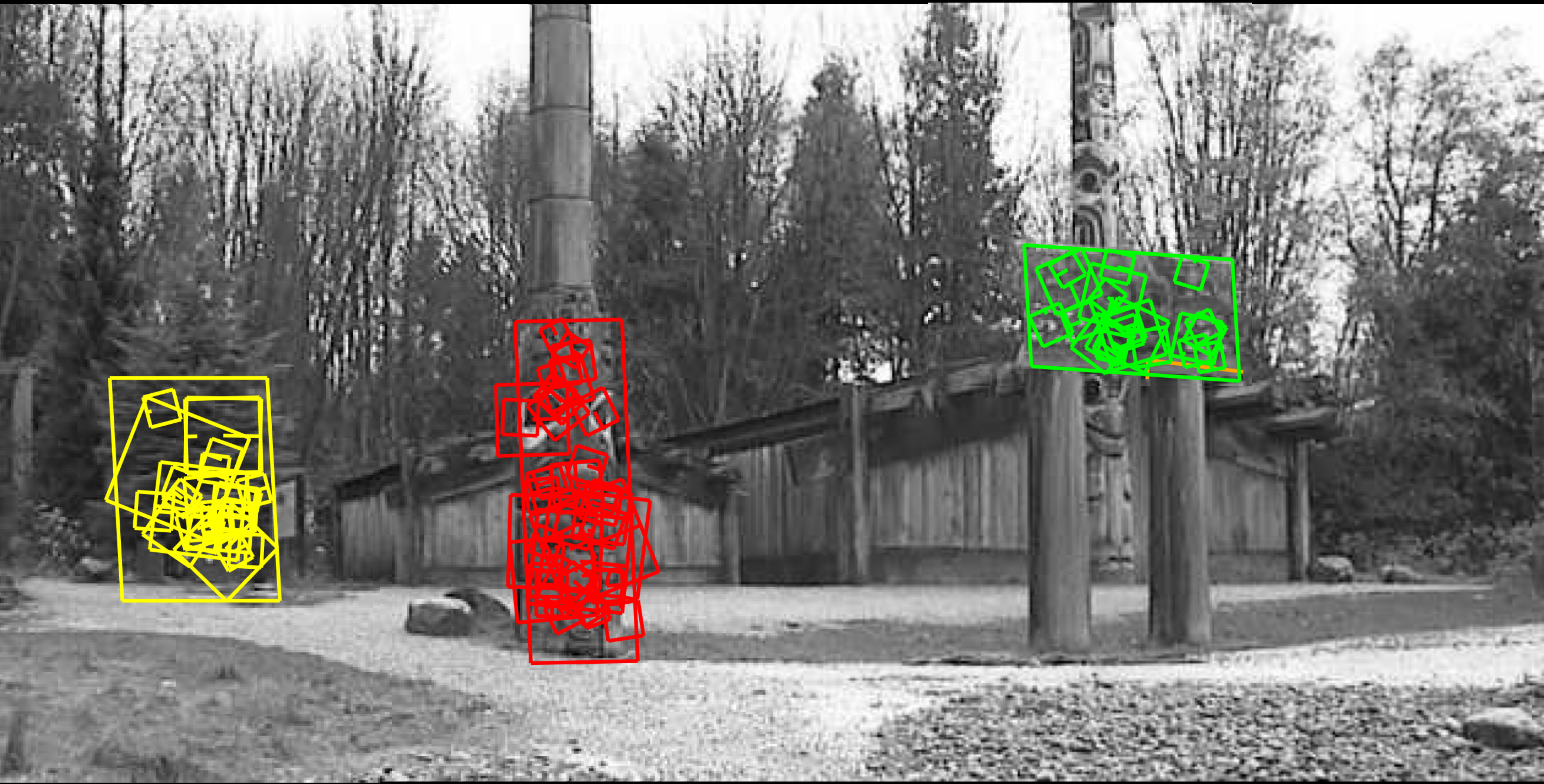
Describe keypoint region at selected scale and rotation with image gradients

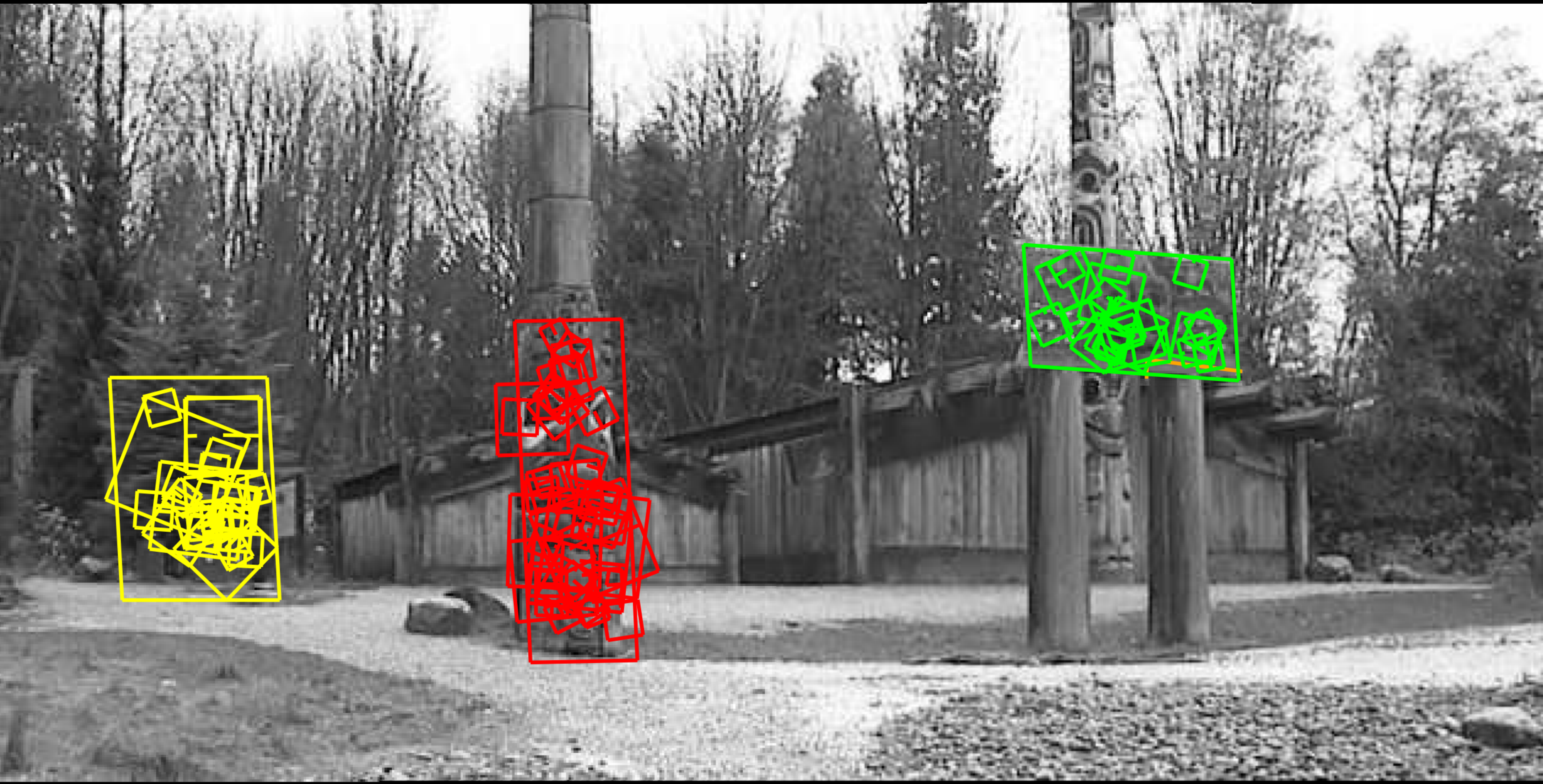
Challenge: Find the following patches

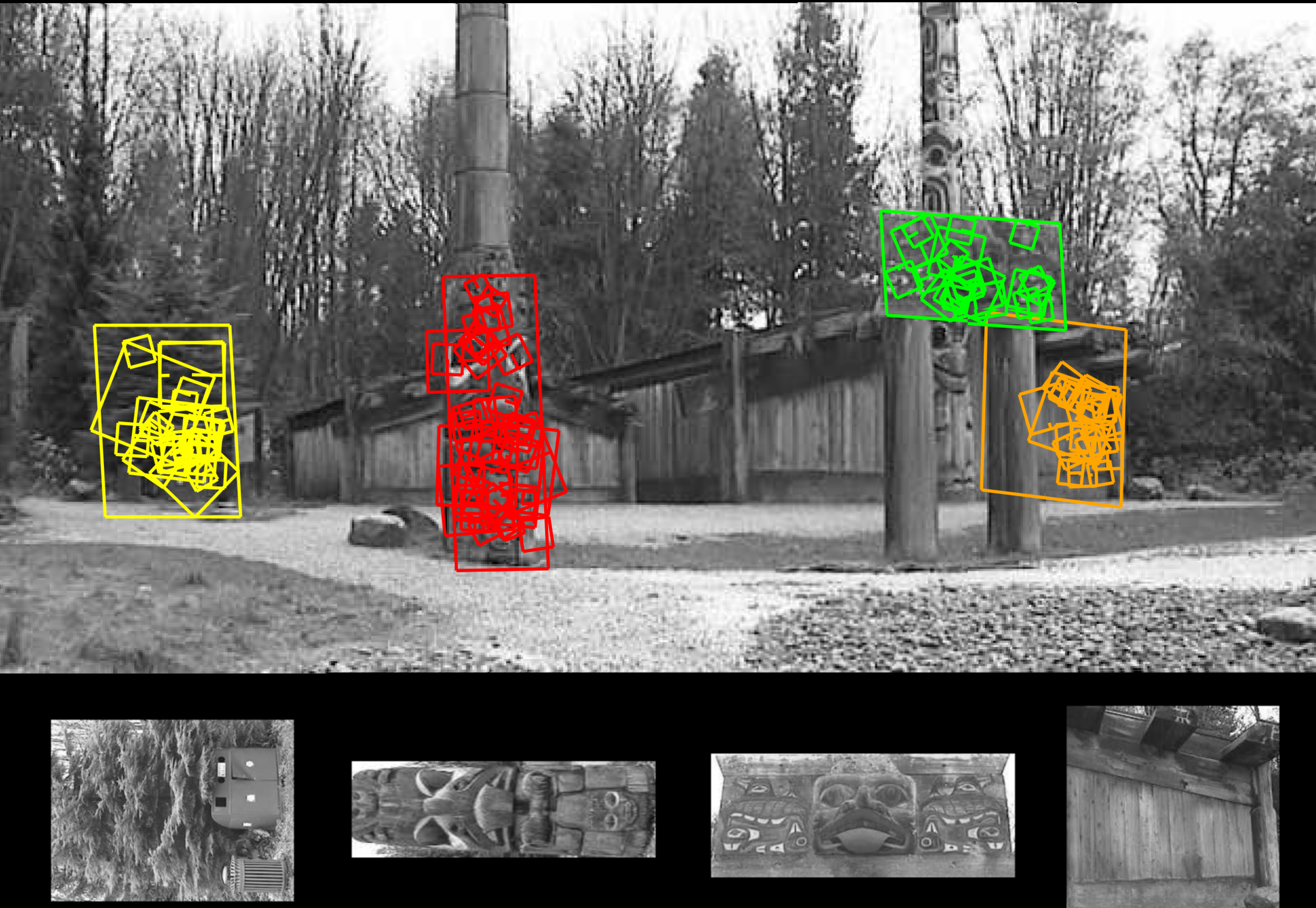










US006711293B1

(12) **United States Patent**
Lowe

(10) **Patent No.:** **US 6,711,293 B1**
(45) **Date of Patent:** **Mar. 23, 2004**

(54) **METHOD AND APPARATUS FOR IDENTIFYING SCALE INVARIANT FEATURES IN AN IMAGE AND USE OF SAME FOR LOCATING AN OBJECT IN AN IMAGE**

(75) Inventor: **David G. Lowe**, Vancouver (CA)

(73) Assignee: **The University of British Columbia**, Vancouver (CA)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/519,893**

(22) Filed: **Mar. 6, 2000**

Related U.S. Application Data

(60) Provisional application No. 60/123,369, filed on Mar. 8, 1999.

(51) **Int. Cl.**[7] **G06K 9/68**
(52) **U.S. Cl.** **382/219**; 382/220
(58) **Field of Search** 382/130, 236, 382/219, 220

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,069,654 A		12/1962	Hough	340/146.3
4,907,156 A	*	3/1990	Doi et al.	382/130
5,119,444 A		6/1992	Nishihara	382/45
5,436,979 A	*	7/1995	Gray et al.	382/141
5,598,481 A	*	1/1997	Nishikawa et al.	382/130
5,617,459 A		4/1997	Makram-Ebeid et al.	378/62
5,666,441 A		9/1997	Rao et al.	382/203
5,764,802 A	*	6/1998	Simon	382/236

OTHER PUBLICATIONS

Ballard, D.H., "Generalizing the Hough transform to detect arbitrary shapes", Pattern Recognition, 13, 2 (1981). pp. 111–122.

Crowley, James L., and Alice C. Parker, "A representation for shape based on peaks and ridges in the difference of low-pass transform", IEEE Trans. on Pattern Analysis and Machine Intelligence, 6, 2 (1984), pp. 156–170.

Schmid, C., and R. Mohr, "Local grayvalue invariants for image retrieval", IEEE PAMI, 19, 5 (1997), pp. 530–535.

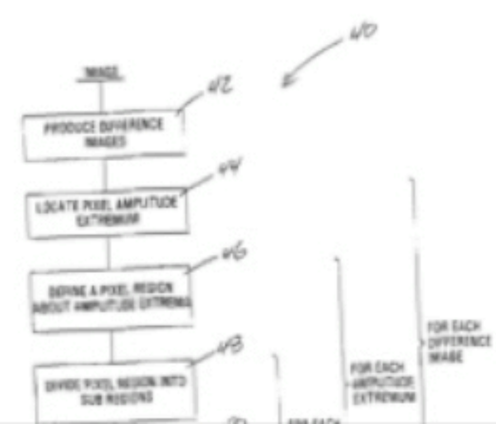
* cited by examiner

Primary Examiner—Leo Boudreau
Assistant Examiner—Tom Y. Lu
(74) *Attorney, Agent, or Firm*—Christie, Parker & Hale, LLP

(57) **ABSTRACT**

A method and apparatus for identifying scale invariant features in an image and a further method and apparatus for using such scale invariant features to locate an object in an image are disclosed. The method and apparatus for identifying scale invariant features may involve the use of a processor circuit for producing a plurality of component subregion descriptors for each subregion of a pixel region about pixel amplitude extrema in a plurality of difference images produced from the image. This may involve producing a plurality of difference images by blurring an initial image to produce a blurred image and by subtracting the blurred image from the initial image to produce the difference image. For each difference image, pixel amplitude extrema are located and a corresponding pixel region is defined about each pixel amplitude extremum. Each pixel region is divided into subregions and a plurality of component subregion descriptors are produced for each subregion. These component subregion descriptors are correlated with component subregion descriptors of an image under consideration and an object is indicated as being detected when a sufficient number of component subregion descriptors (scale invariant features) define an aggregate correlation exceeding a threshold correlation with component subregion descriptors (scale invariant features) associated with the object.

20 Claims, 12 Drawing Sheets



US006711293B1

(12) **United States Patent**
Lowe

(10) **Patent No.:** **US 6,711,293 B1**
(45) **Date of Patent:** **Mar. 23, 2004**

(54) **METHOD AND APPARATUS FOR IDENTIFYING SCALE INVARIANT FEATURES IN AN IMAGE AND USE OF SAME FOR LOCATING AN OBJECT IN AN IMAGE**

(75) Inventor: **David G. Lowe**, Vancouver (CA)

(73) Assignee: **The University of British Columbia**, Vancouver (CA)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/519,893**

(22) Filed: **Mar. 6, 2000**

Related U.S. Application Data

(60) Provisional application No. 60/123,369, filed on Mar. 8, 1999.

(51) **Int. Cl.**⁷ **G06K 9/68**
(52) **U.S. Cl.** **382/219**; 382/220
(58) **Field of Search** 382/130, 236, 382/219, 220

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,069,654 A		12/1962	Hough	340/146.3
4,907,156 A	*	3/1990	Doi et al.	382/130
5,119,444 A		6/1992	Nishihara	382/45
5,436,979 A	*	7/1995	Gray et al.	382/141
5,598,481 A	*	1/1997	Nishikawa et al.	382/130
5,617,459 A		4/1997	Makram-Ebeid et al.	378/62
5,666,441 A		9/1997	Rao et al.	382/203
5,764,802 A	*	6/1998	Simon	382/236

OTHER PUBLICATIONS

Ballard, D.H., "Generalizing the Hough transform to detect arbitrary shapes", Pattern Recognition, 13, 2 (1981). pp. 111–122.

Crowley, James L., and Alice C. Parker, "A representation for shape based on peaks and ridges in the difference of low-pass transform", IEEE Trans. on Pattern Analysis and Machine Intelligence, 6, 2 (1984), pp. 156–170.

Schmid, C., and R. Mohr, "Local grayvalue invariants for image retrieval", IEEE PAMI, 19, 5 (1997), pp. 530–535.

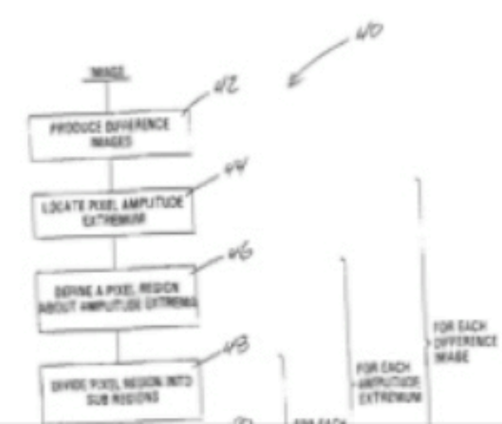
* cited by examiner

Primary Examiner—Leo Boudreau
Assistant Examiner—Tom Y. Lu
(74) *Attorney, Agent, or Firm*—Christie, Parker & Hale, LLP

(57) **ABSTRACT**

A method and apparatus for identifying scale invariant features in an image and a further method and apparatus for using such scale invariant features to locate an object in an image are disclosed. The method and apparatus for identifying scale invariant features may involve the use of a processor circuit for producing a plurality of component subregion descriptors for each subregion of a pixel region about pixel amplitude extrema in a plurality of difference images produced from the image. This may involve producing a plurality of difference images by blurring an initial image to produce a blurred image and by subtracting the blurred image from the initial image to produce the difference image. For each difference image, pixel amplitude extrema are located and a corresponding pixel region is defined about each pixel amplitude extremum. Each pixel region is divided into subregions and a plurality of component subregion descriptors are produced for each subregion. These component subregion descriptors are correlated with component subregion descriptors of an image under consideration and an object is indicated as being detected when a sufficient number of component subregion descriptors (scale invariant features) define an aggregate correlation exceeding a threshold correlation with component subregion descriptors (scale invariant features) associated with the object.

20 Claims, 12 Drawing Sheets



VLFeat.org

SIFT implementation

VLFeat.org

SIFT implementation

and much more