# Applied Computer Vision

David Vernon
Carnegie Mellon University Africa

vernon@cmu.edu
www.vernon.eu

# Lecture 12

# Image Features

## Scale-invariant interest point operators

# Multi-scale Detection

Credit: Kris Kitani, Carnegie Mellon University

# Properties of the Harris corner detector

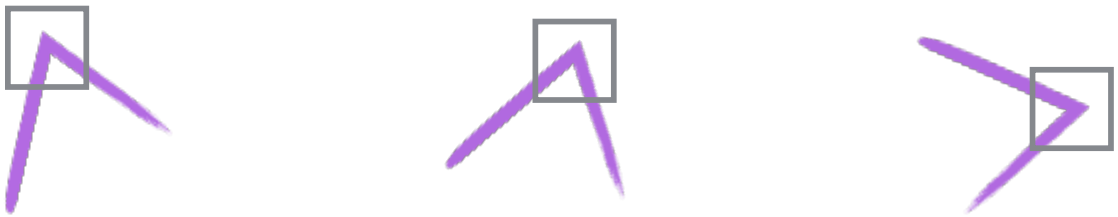
Rotation invariant?

Scale invariant?

Credit: Kris Kitani, Carnegie Mellon University

# Properties of the Harris corner detector

Rotation invariant?



Scale invariant?

Credit: Kris Kitani, Carnegie Mellon University

# Properties of the Harris corner detector

Rotation invariant?

Scale invariant?

Credit: Kris Kitani, Carnegie Mellon University

How can we make a feature detector scale-invariant?

Credit: Kris Kitani, Carnegie Mellon University

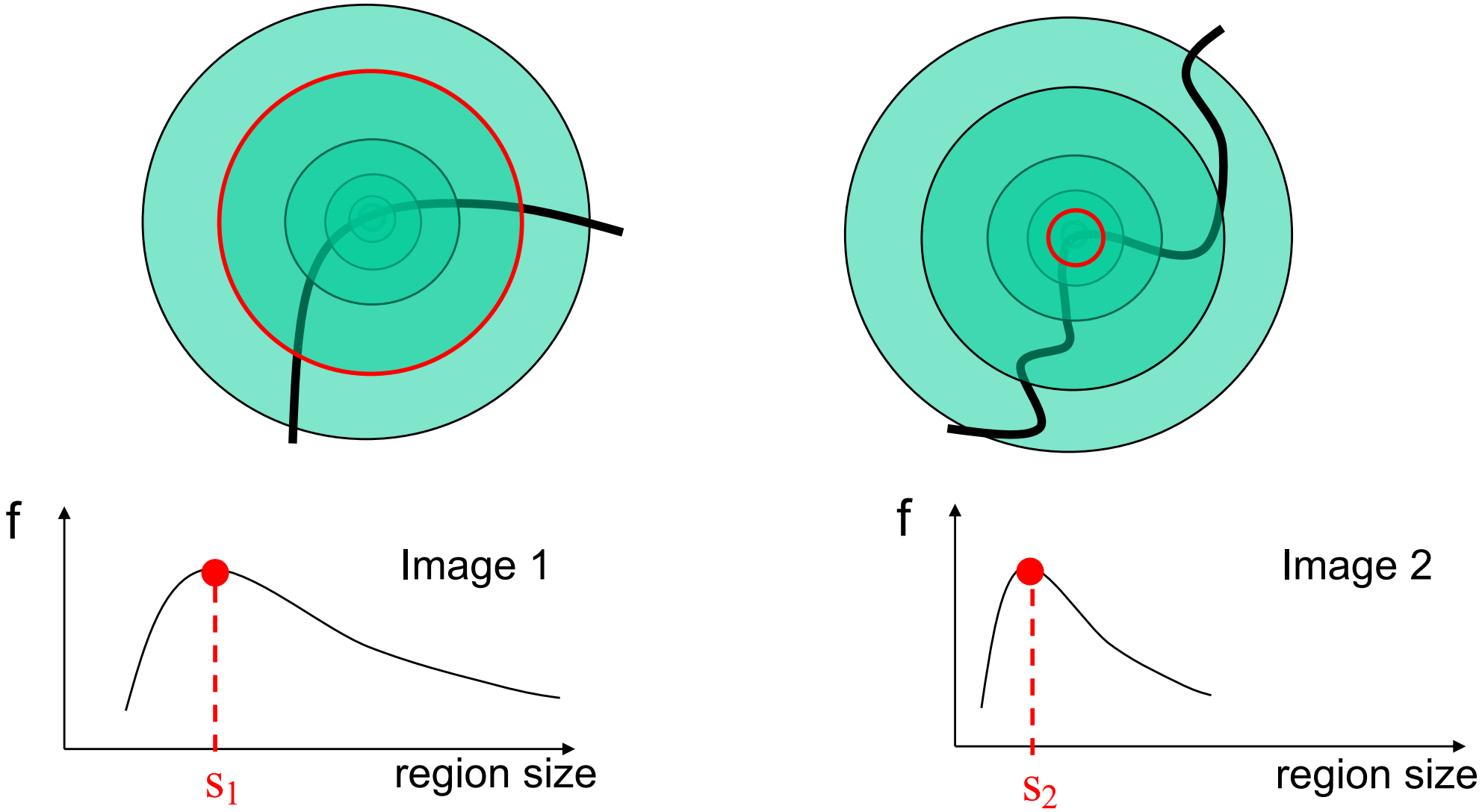
How can we automatically select the scale?

Credit: Kris Kitani, Carnegie Mellon University

Credit: Kris Kitani, Carnegie Mellon University

# Intuitively…

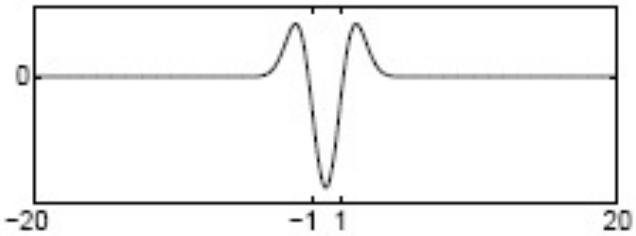
Find local maxima in both **position** and **scale**

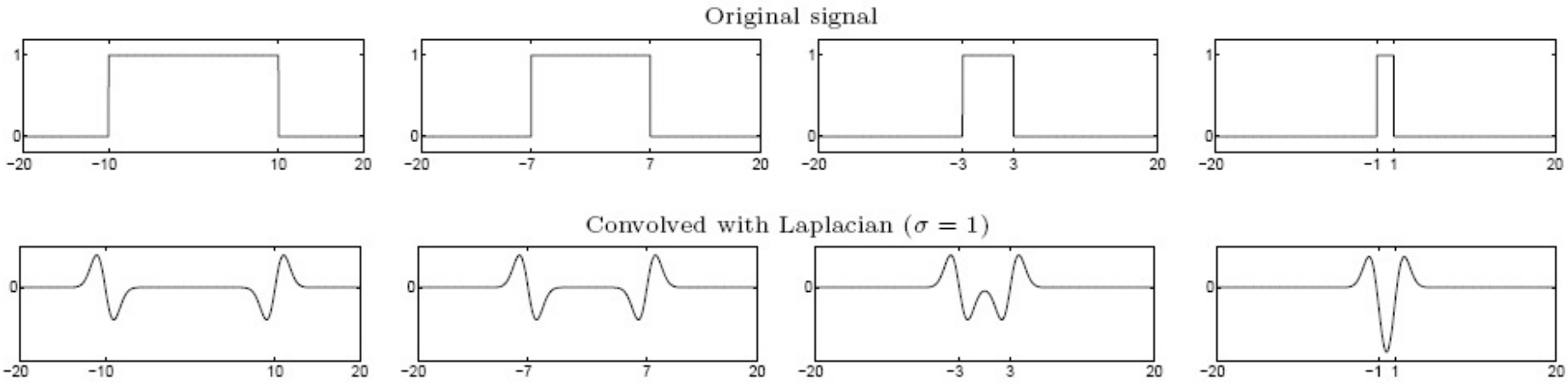



Credit: Kris Kitani, Carnegie Mellon University

# Formally…

Laplacian filter

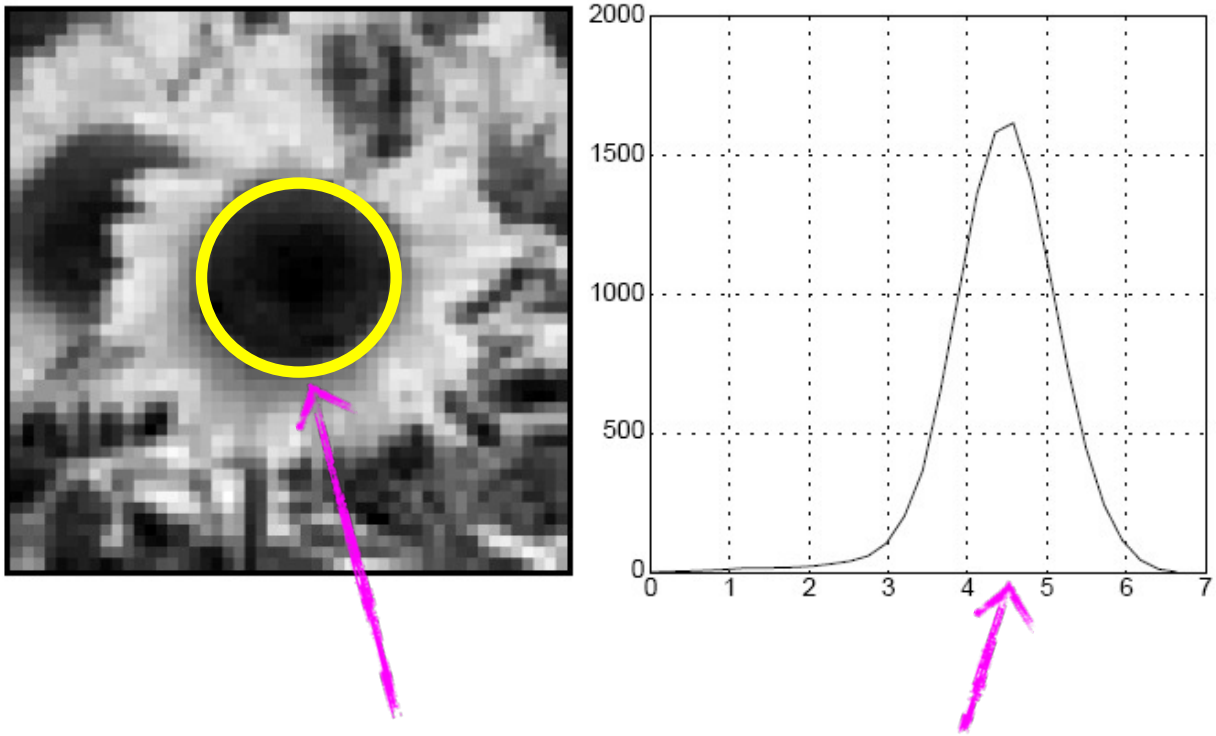



Highest response when the signal has the same **characteristic scale** as the filter

Credit: Kris Kitani, Carnegie Mellon University

characteristic scale - the scale that produces peak filter response




characteristic scale

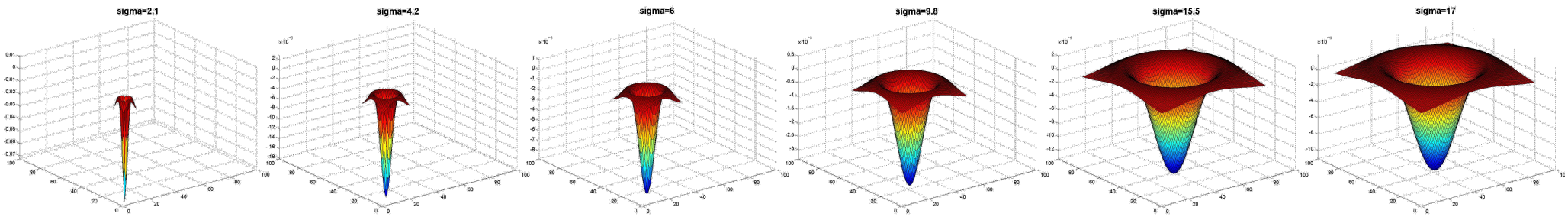
**we need to search over characteristic scales**

Credit: Kris Kitani, Carnegie Mellon University

# Multi-scale
# 2D Blob detection

Credit: Kris Kitani, Carnegie Mellon University

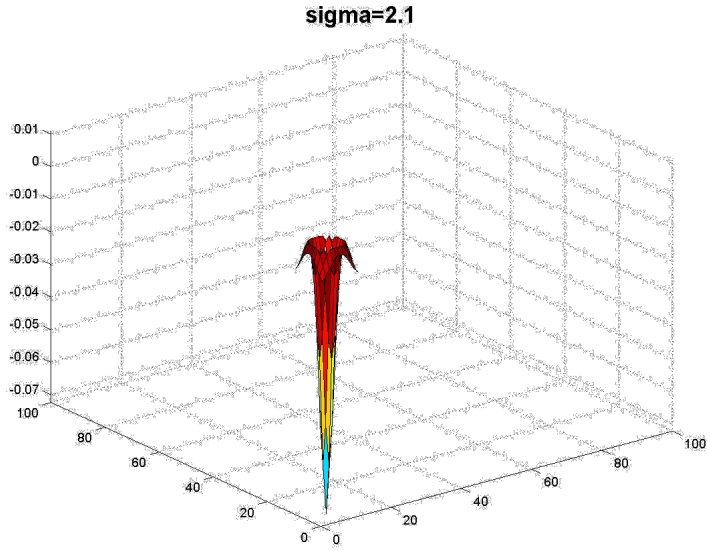
# What happens if you apply different Laplacian filters?



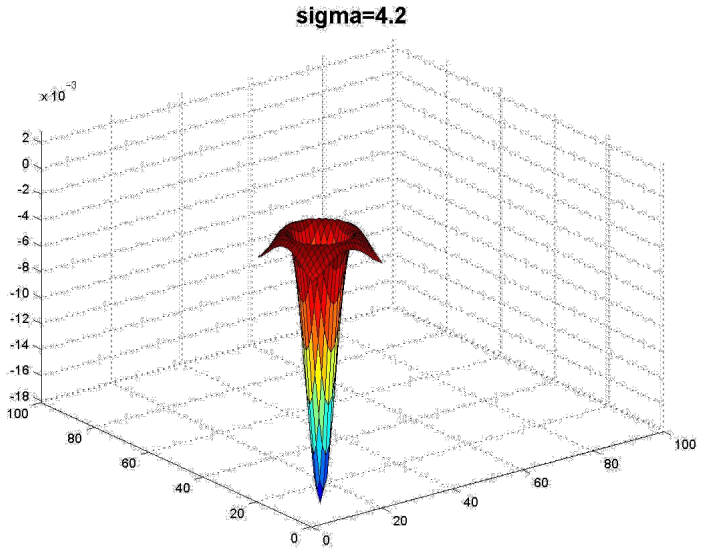

Full size

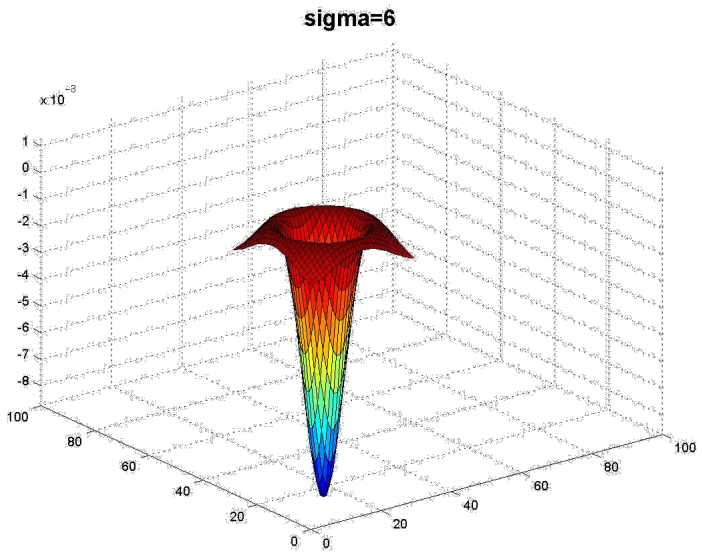
3/4 size

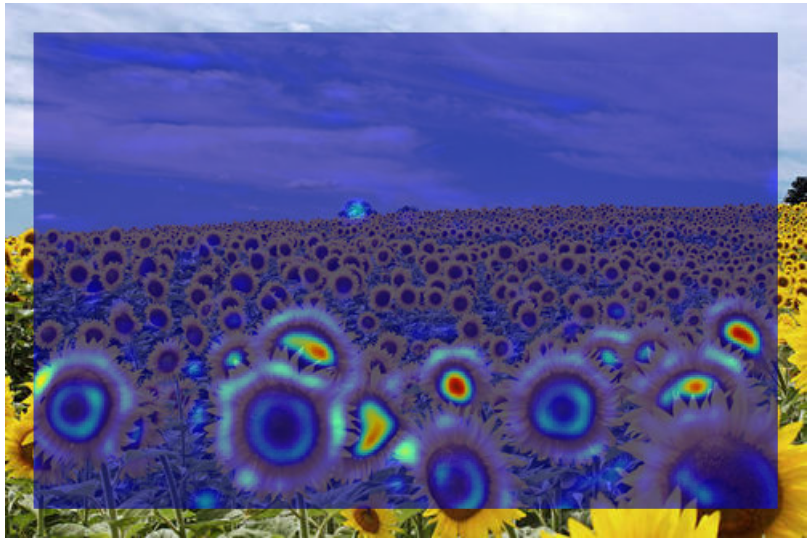
Credit: Kris Kitani, Carnegie Mellon University

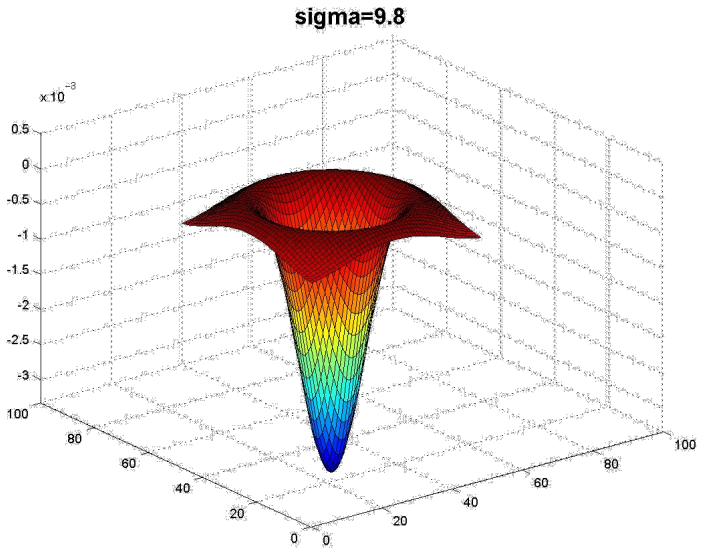


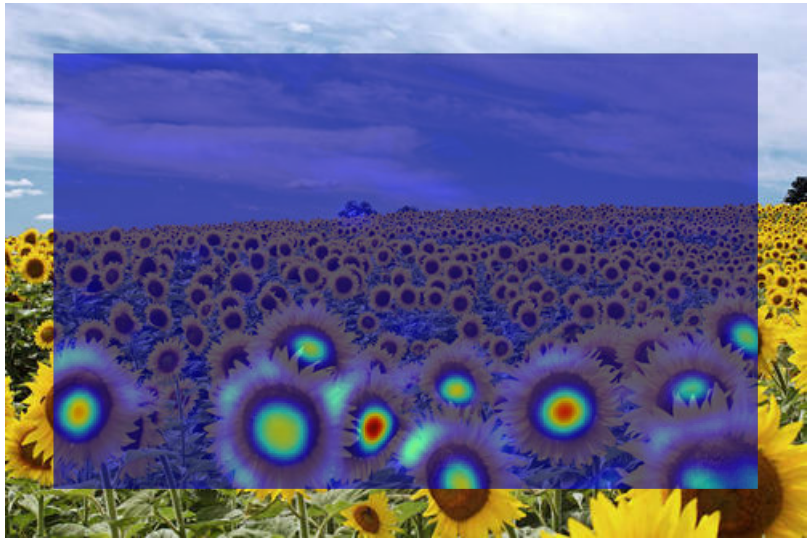
**jet** color scale
blue: low, red: high

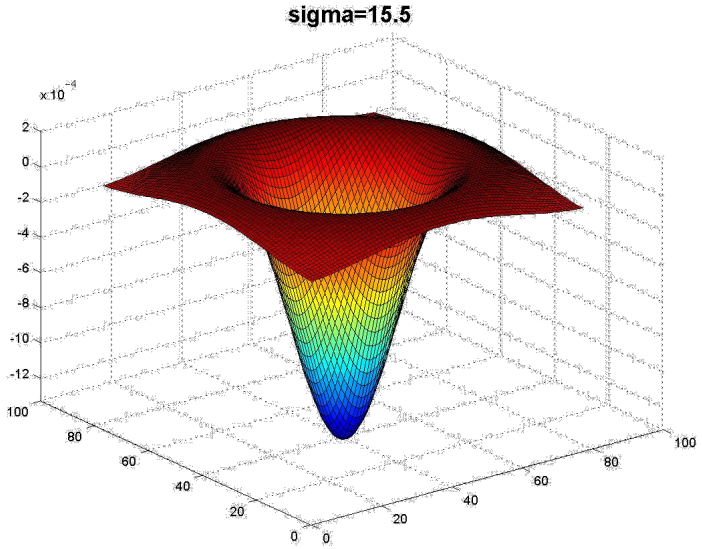

sigma=4.2
x 10^-3
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
100
80
60
40
20
0
0
20
40
60
80
100

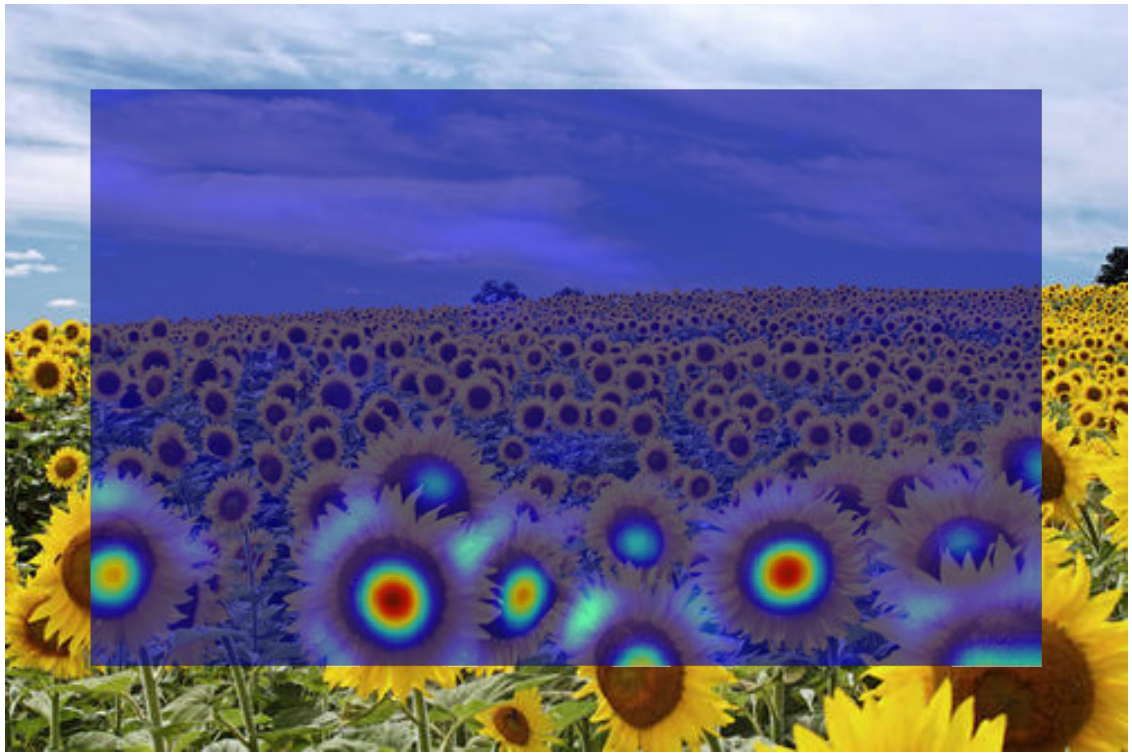

sigma=6
x 10^-3
1
0
-1
-2
-3
-4
-5
-6
-7
-8
100
80
60
40
20
0
0
20
40
60
80
100

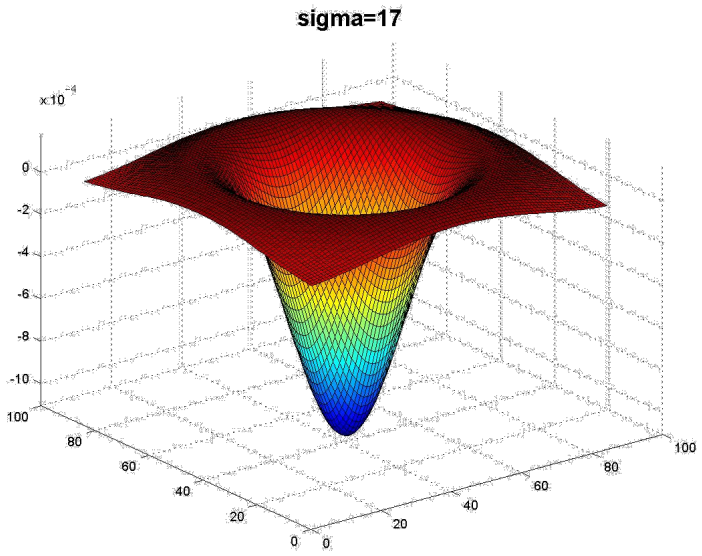


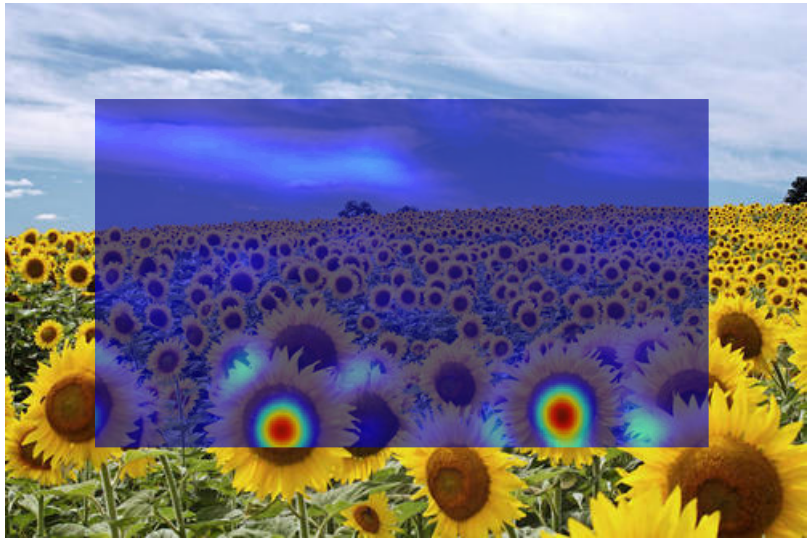

sigma=9.8
x 10^-3
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
100
80
60
40
20
0
0
20
40
60
80
100




sigma=15.5
x 10^-4
2
0
-2
-4
-6
-8
-10
-12
100
80
60
40
20
0
0
20
40
60
80
100



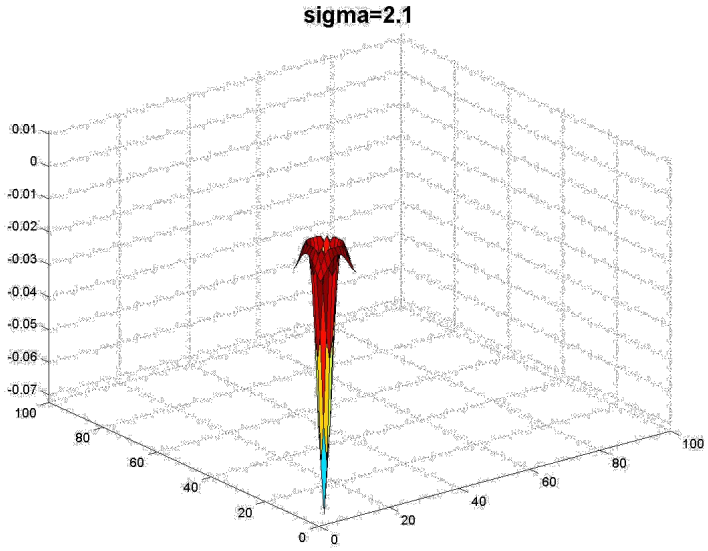

sigma=17
x 10^-4
0
-2
-4
-6
-8
-10
100
80
60
40
20
0
0
20
40
60
80
100

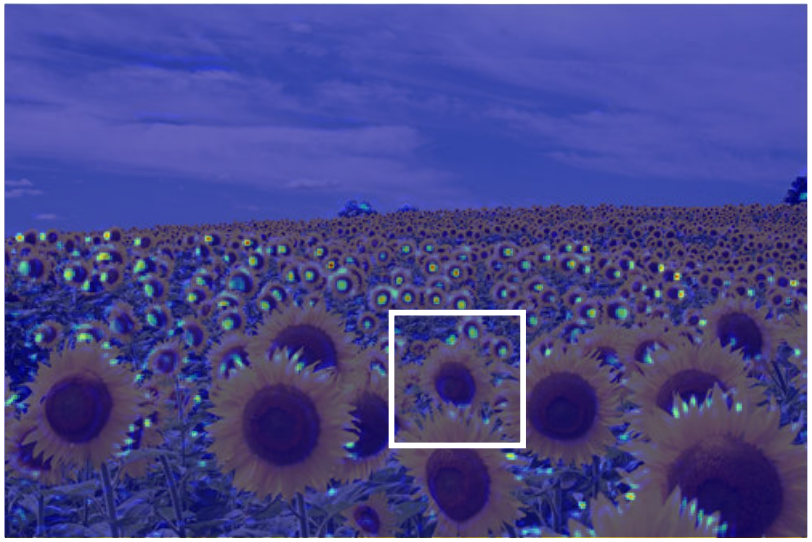


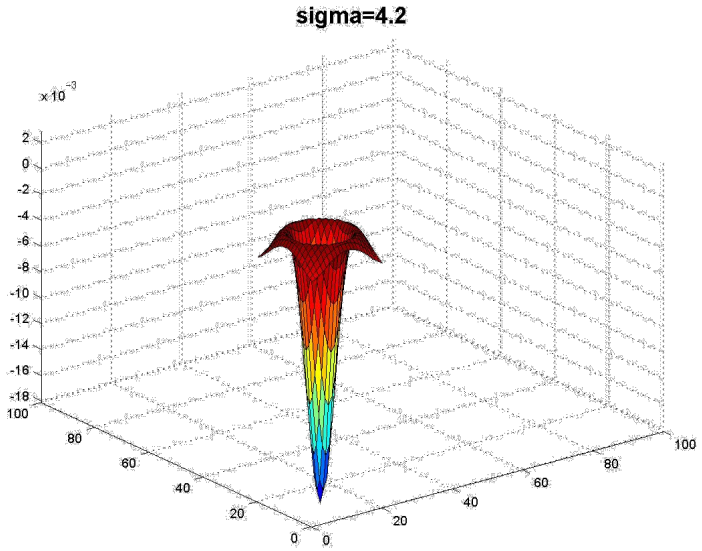
What happened when you applied different Laplacian filters?

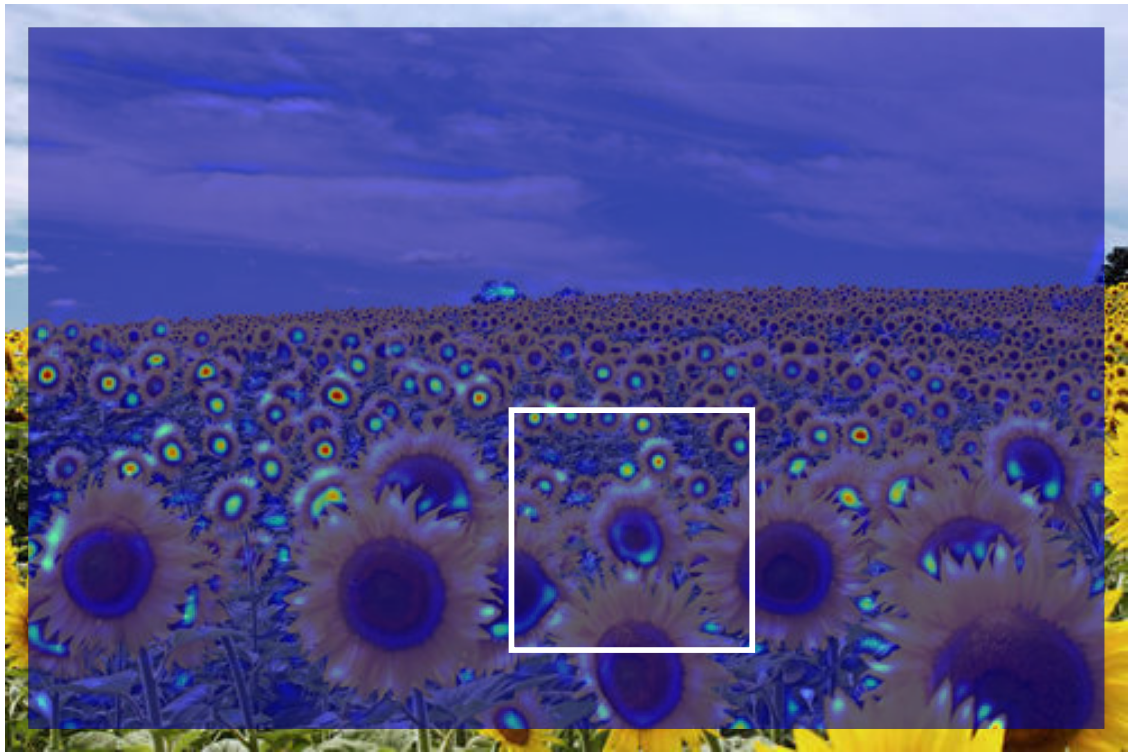
Full size

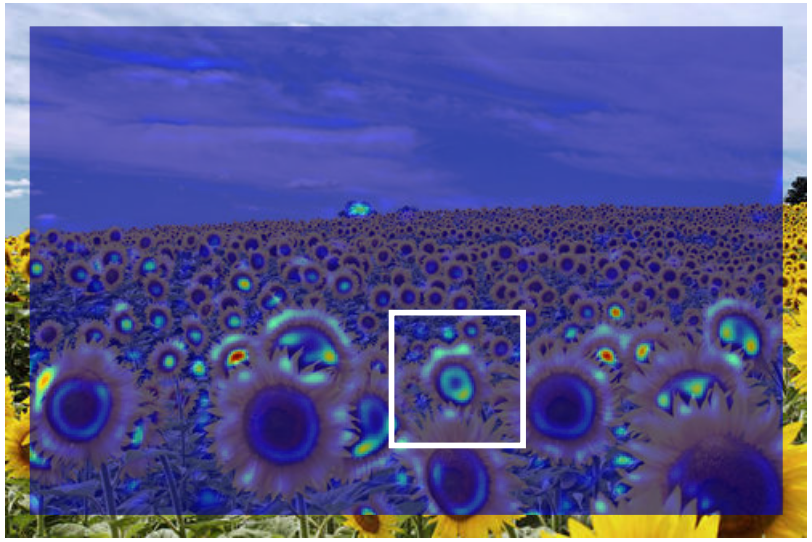
3/4 size

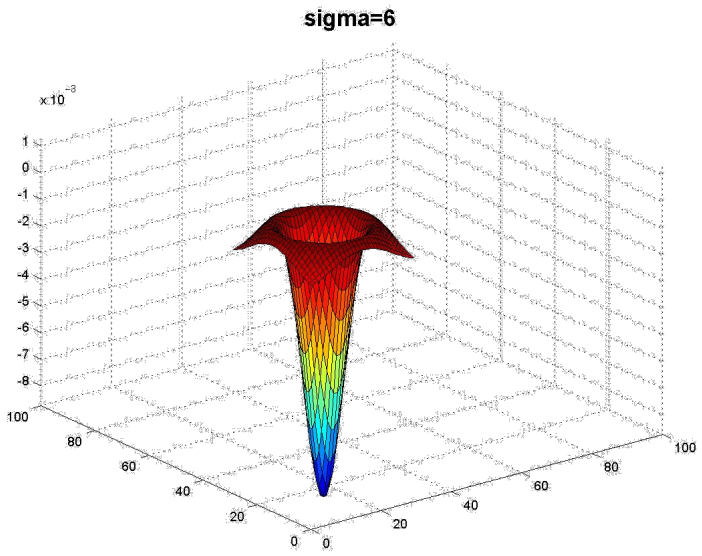

sigma=2.1
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
-0.07
100
80
60
40
20
0
0
20
40
60
80
100

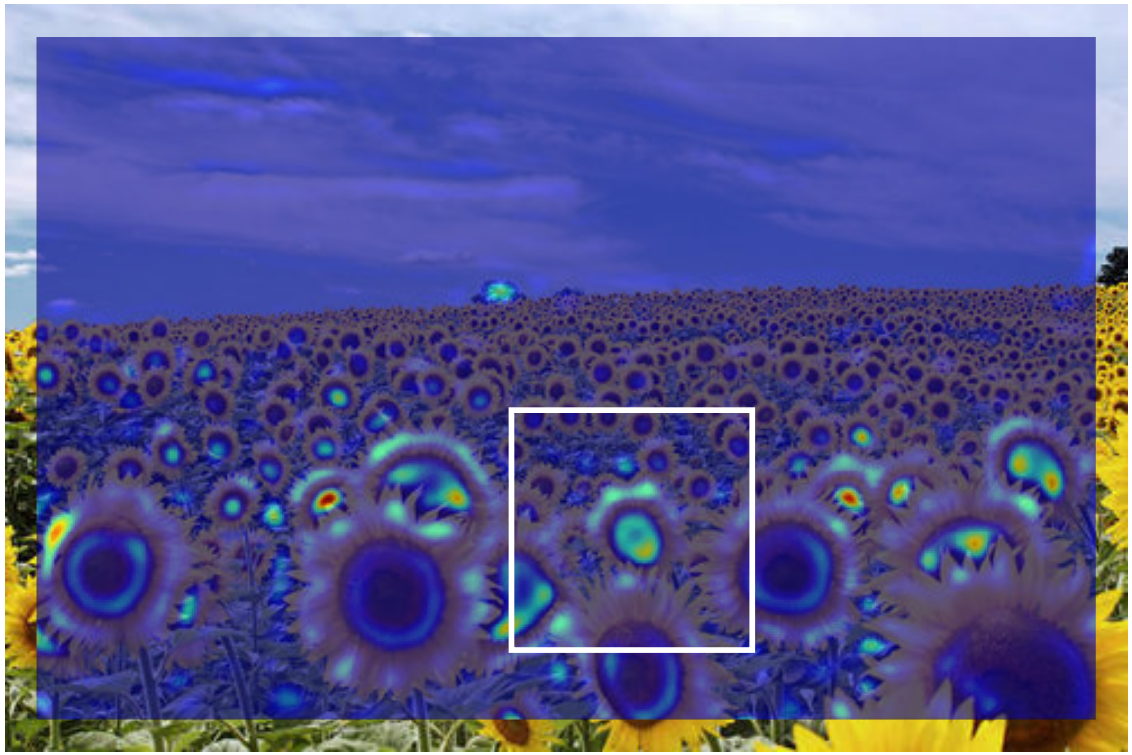


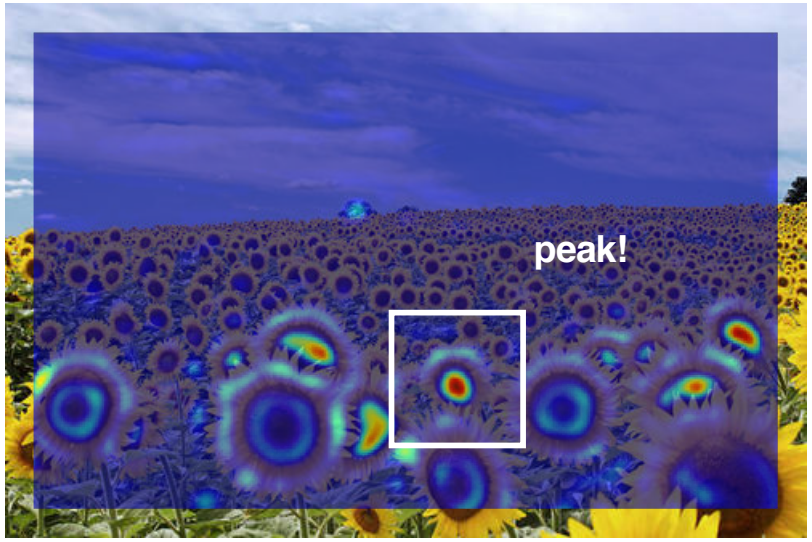

sigma=4.2
x 10^-3
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
100
80
60
40
20
0
0
20
40
60
80
100

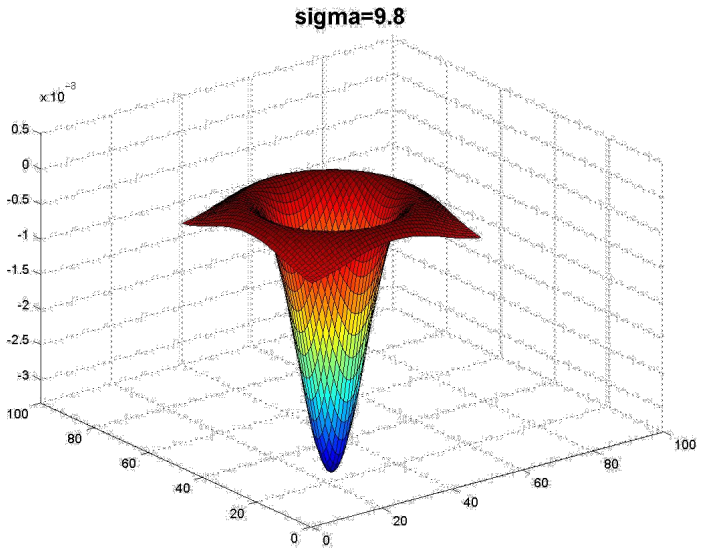


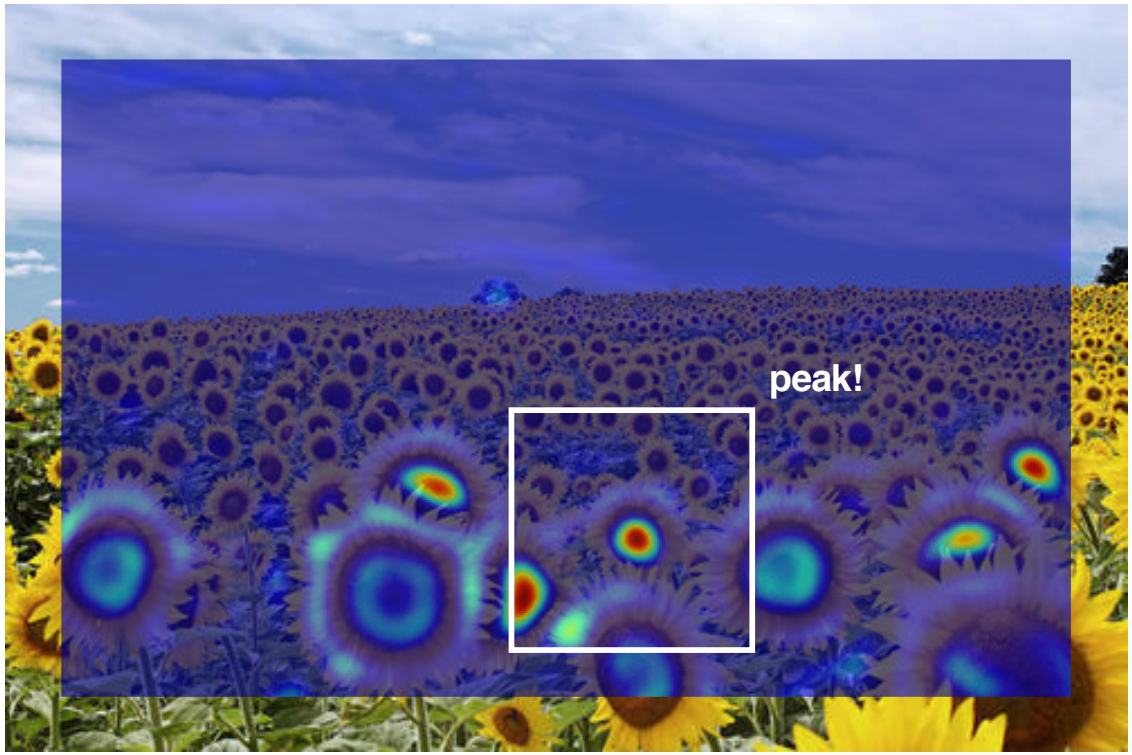


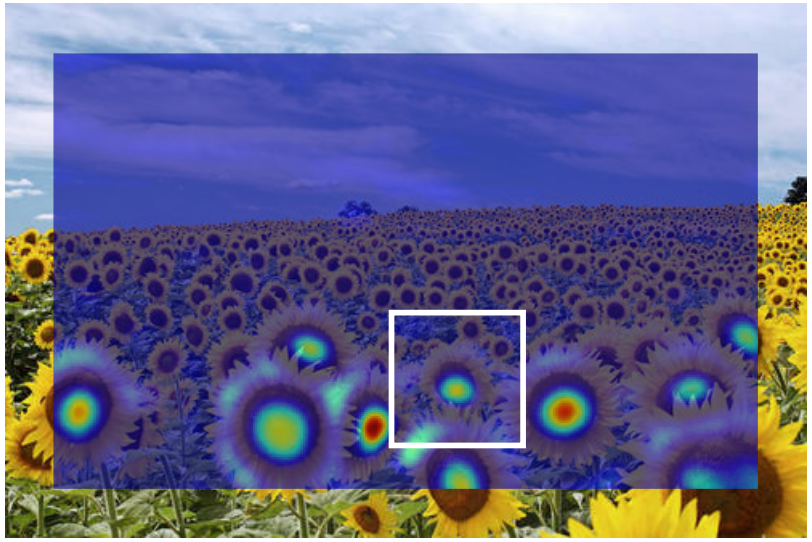

sigma=6
x 10^-3
1
0
-1
-2
-3
-4
-5
-6
-7
-8
100
80
60
40
20
0
0
20
40
60
80
100

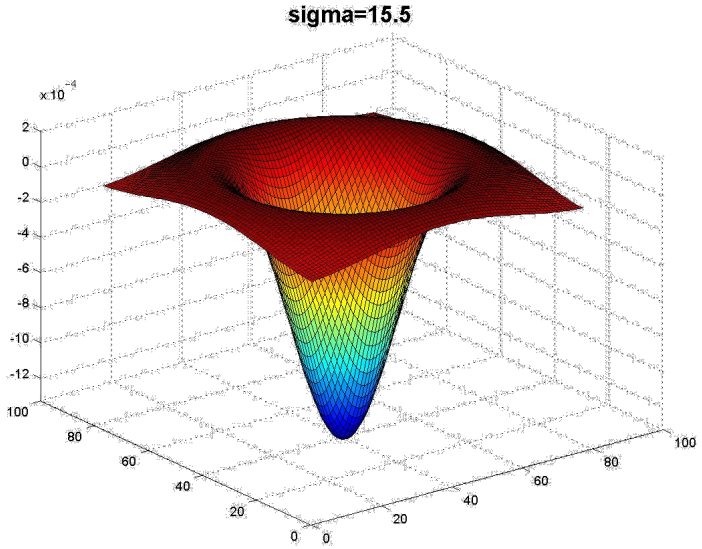


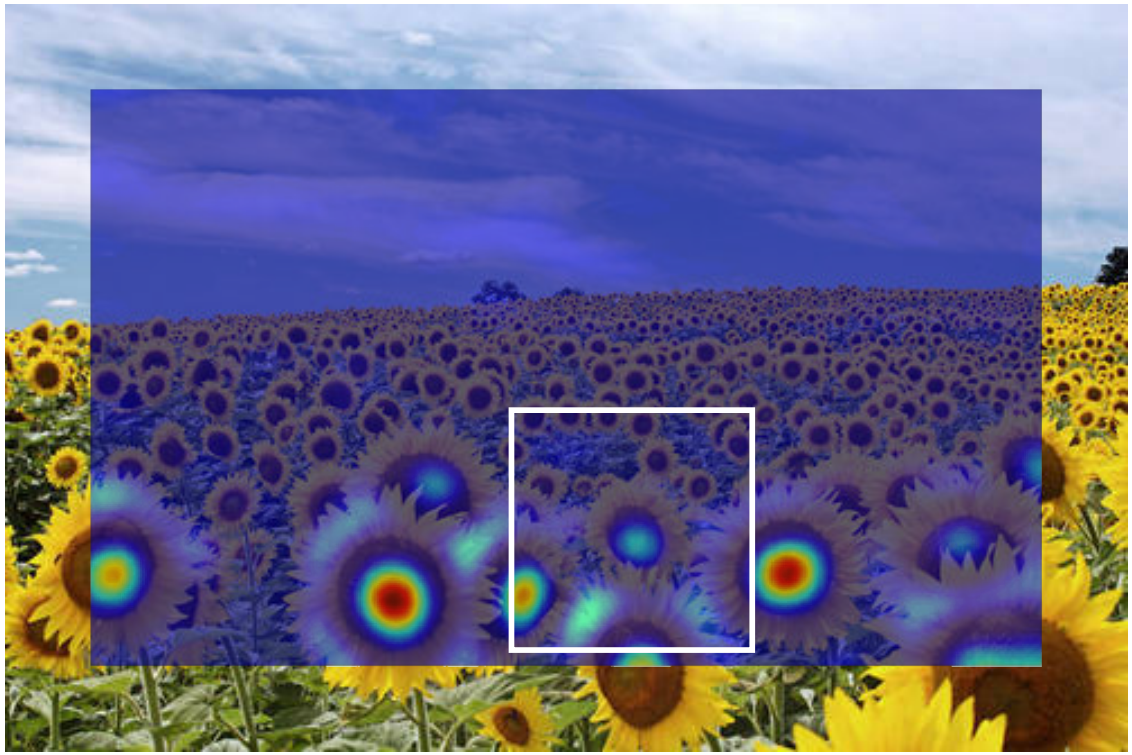

peak!

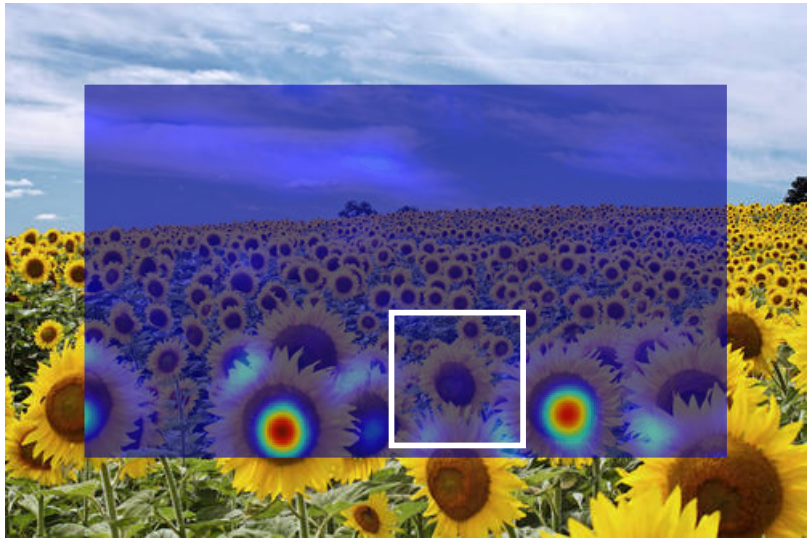

sigma=9.8
x 10^-3
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
100
80
60
40
20
0
0
20
40
60
80
100

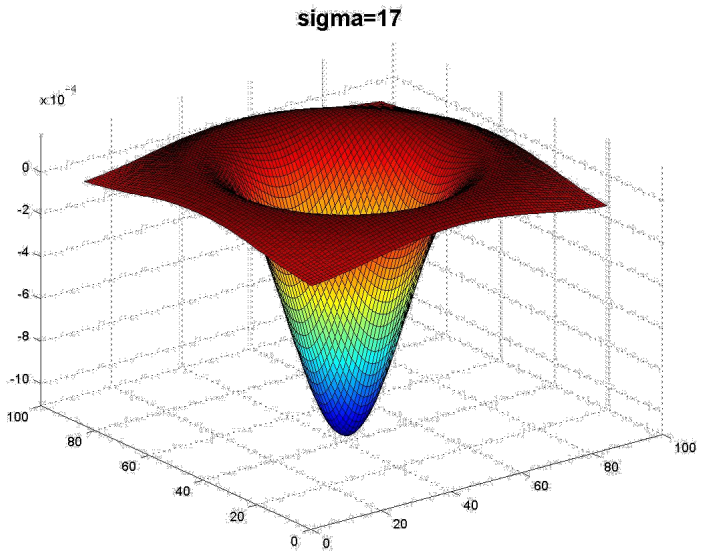

peak!

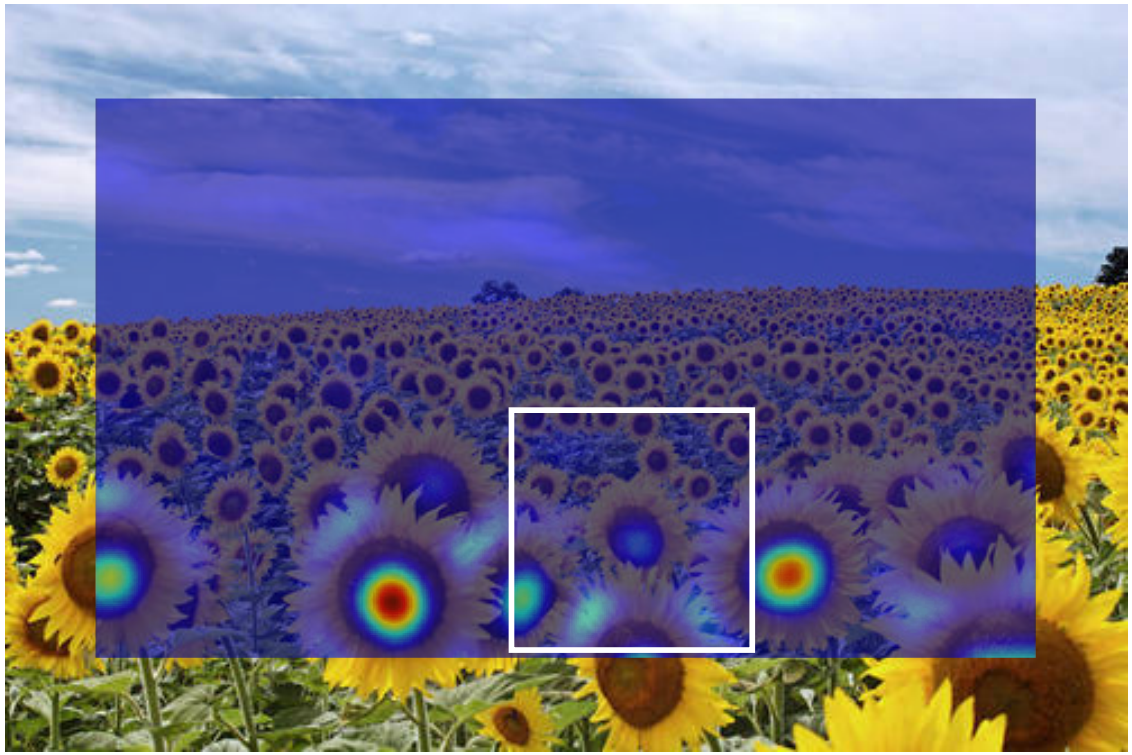


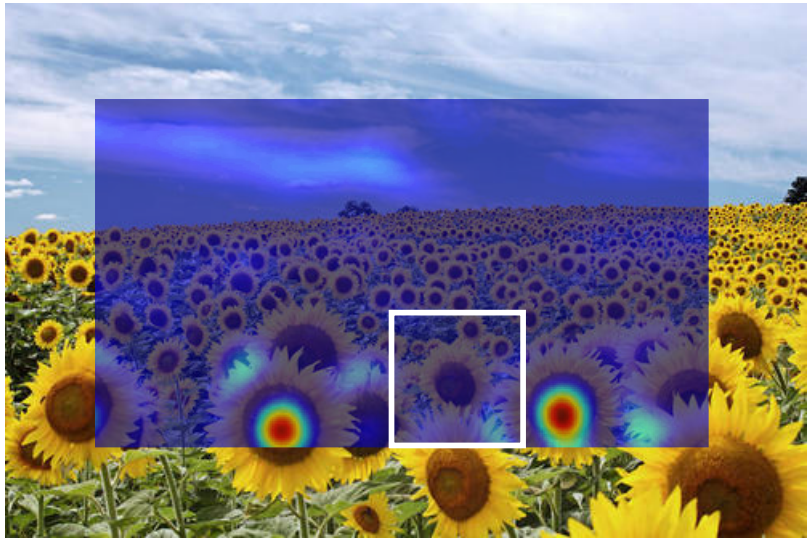

sigma=15.5
x 10^-4
2
0
-2
-4
-6
-8
-10
-12
100
80
60
40
20
0
0
20
40
60
80
100






sigma=17
x 10^-4
0
-2
-4
-6
-8
-10
100
80
60
40
20
0
0
20
40
60
80
100



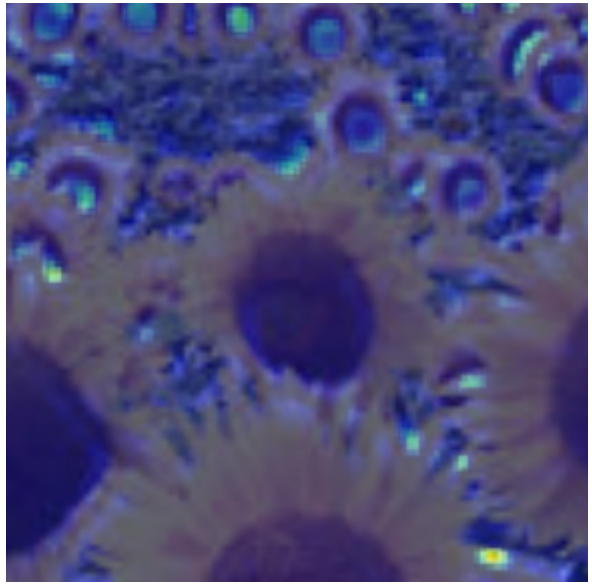


What happened when you applied different Laplacian filters?

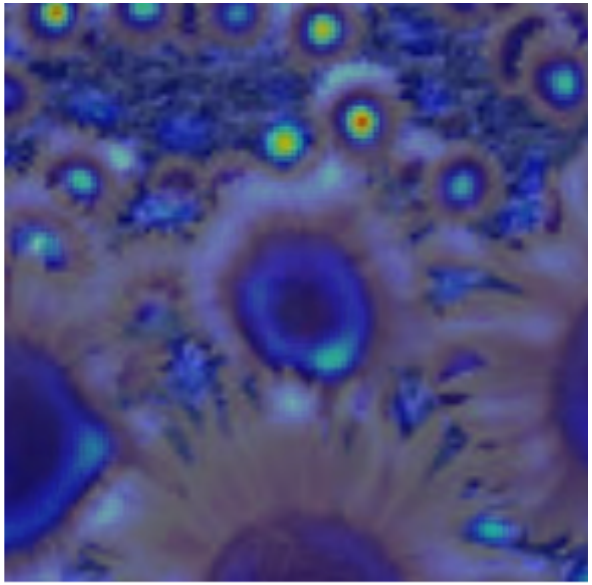
Full size

3/4 size

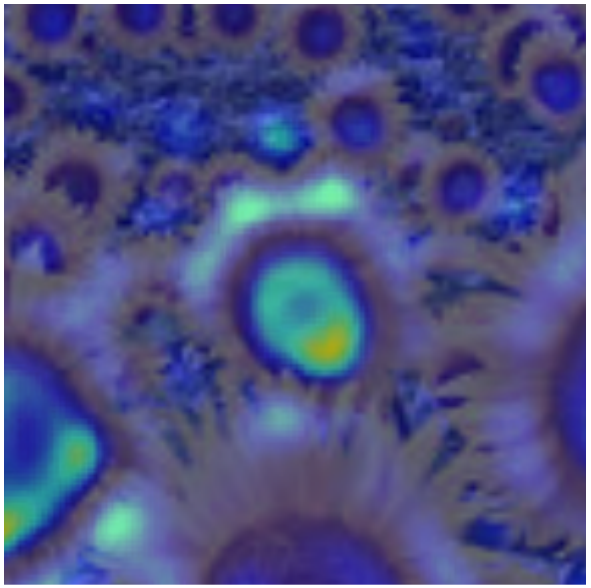
2.1



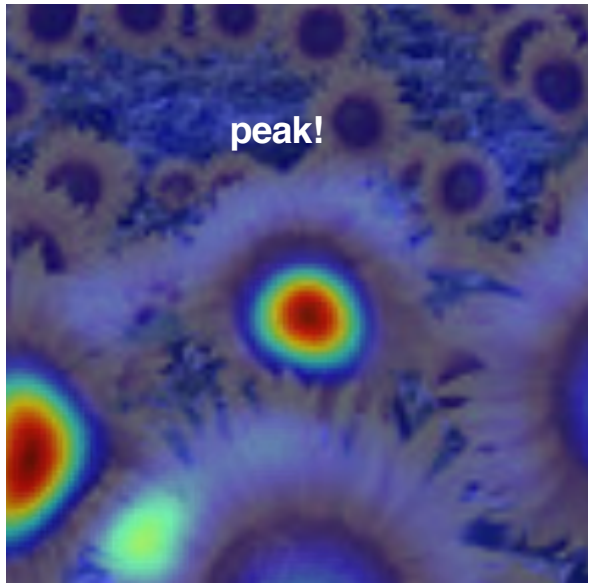
4.2



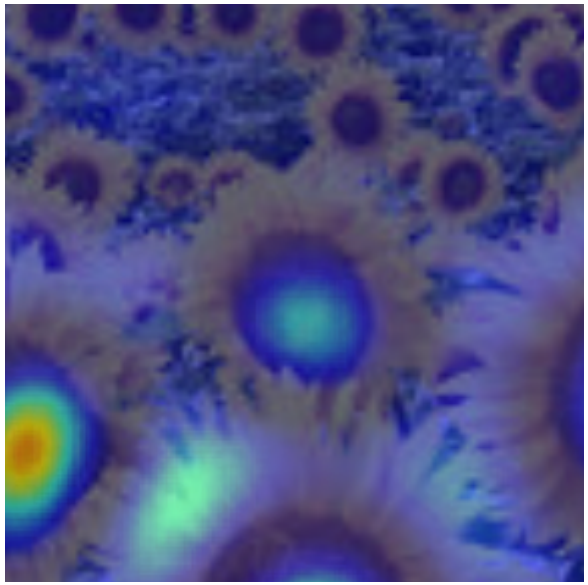
6.0



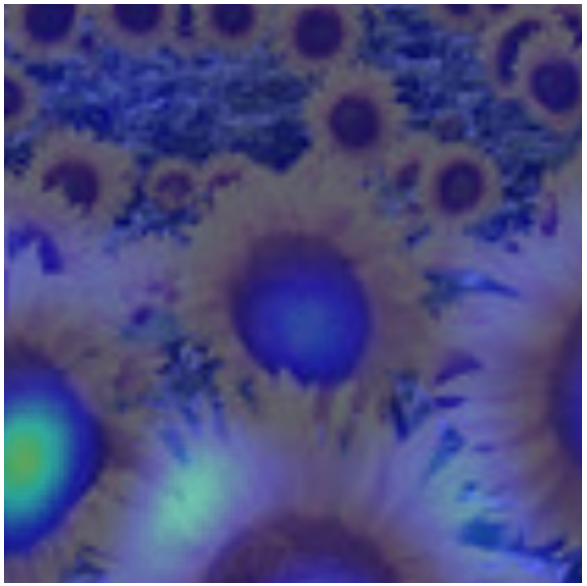
9.8



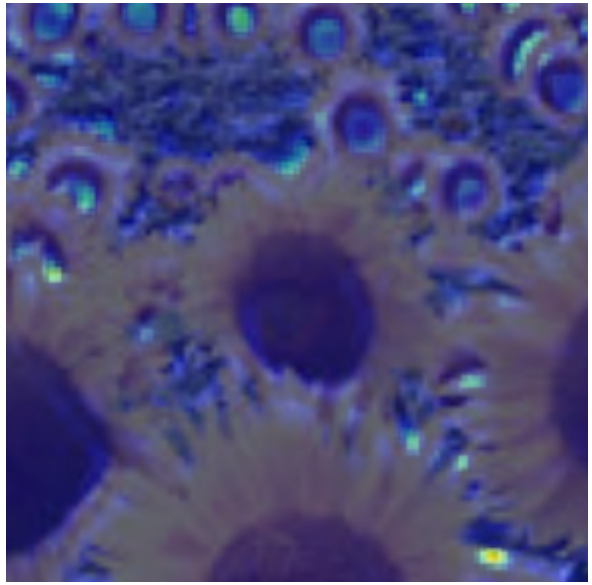

15.5



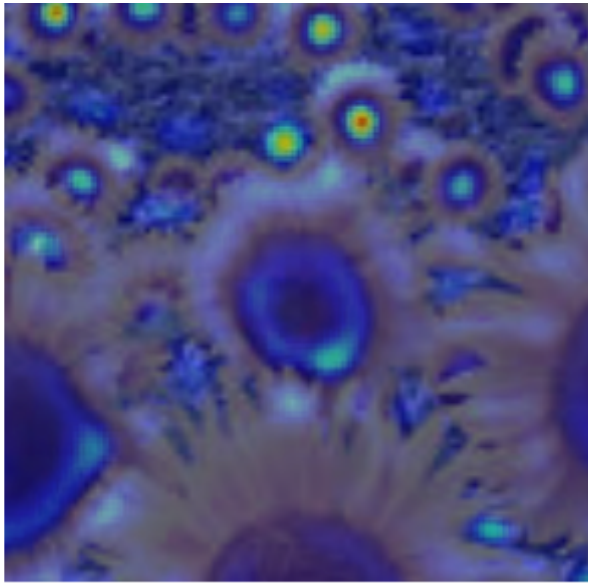
17.0



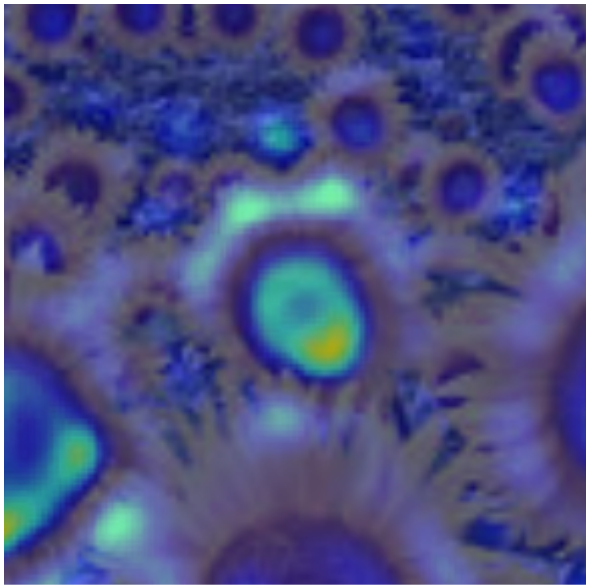
2.1



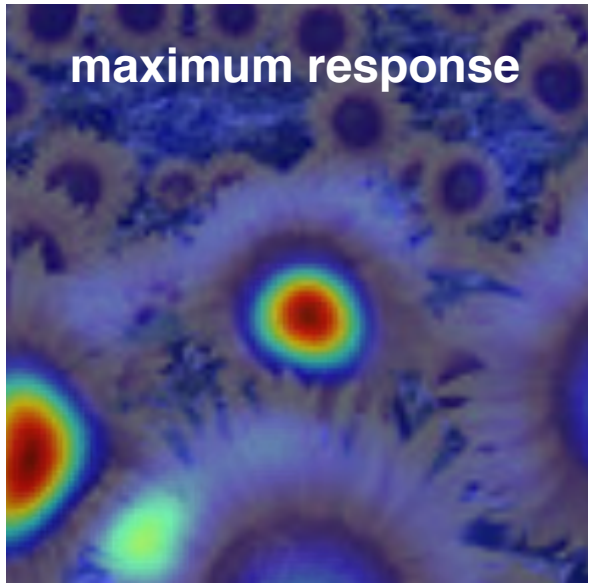
4.2



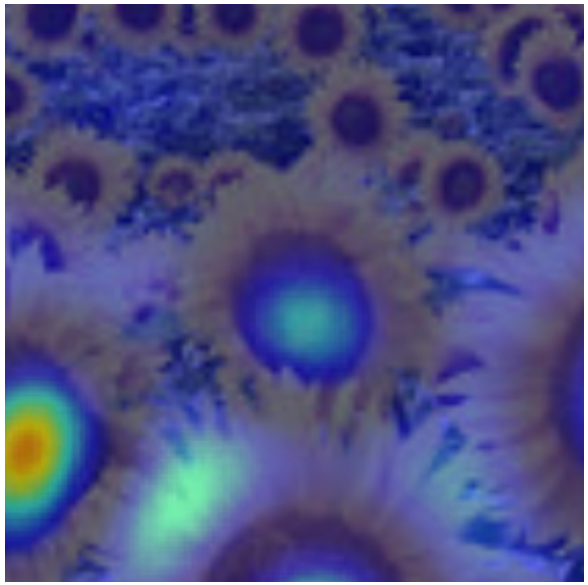
6.0



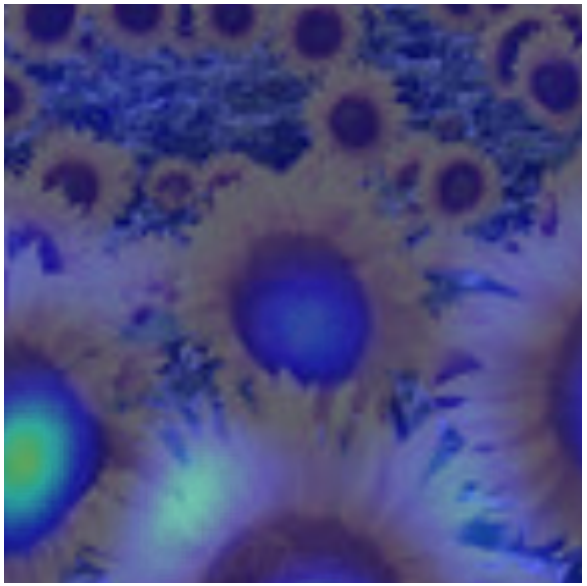
9.8




15.5

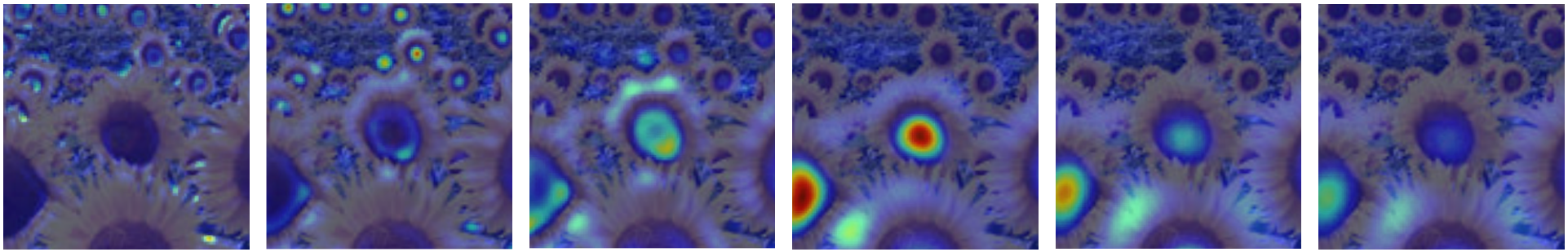


17.0



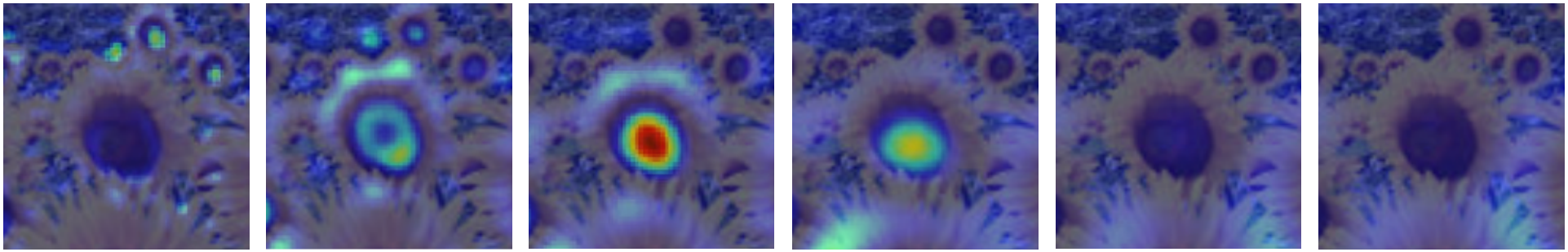
# optimal scale

2.1 4.2 6.0 9.8 15.5 17.0



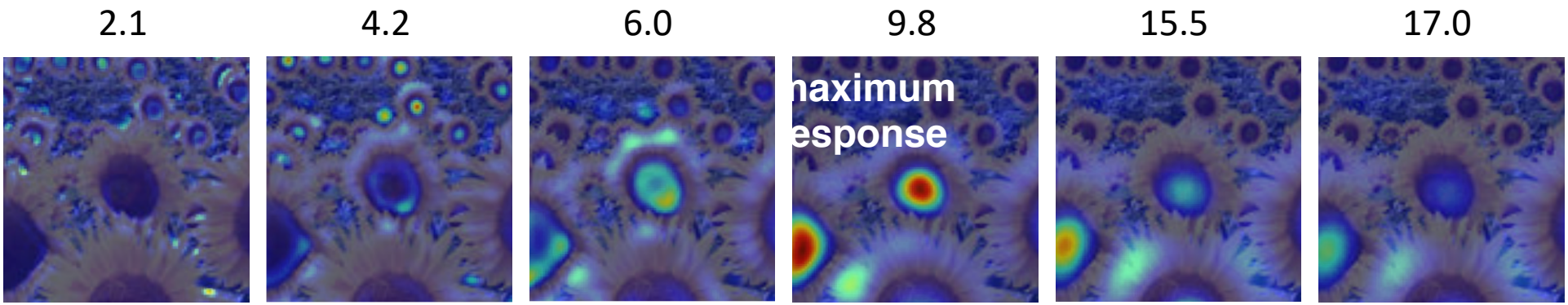
Full size image

2.1 4.2 6.0 9.8 15.5 17.0

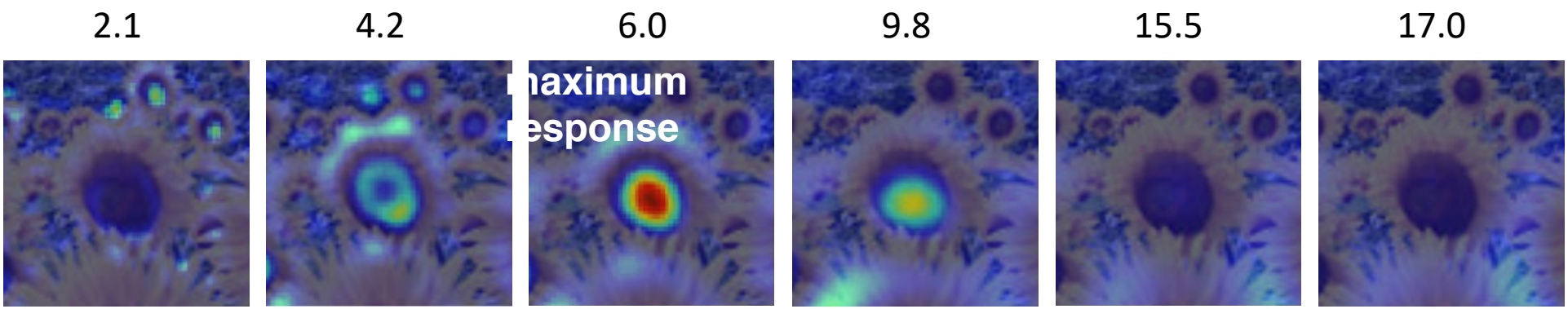


3/4 size image

# optimal scale

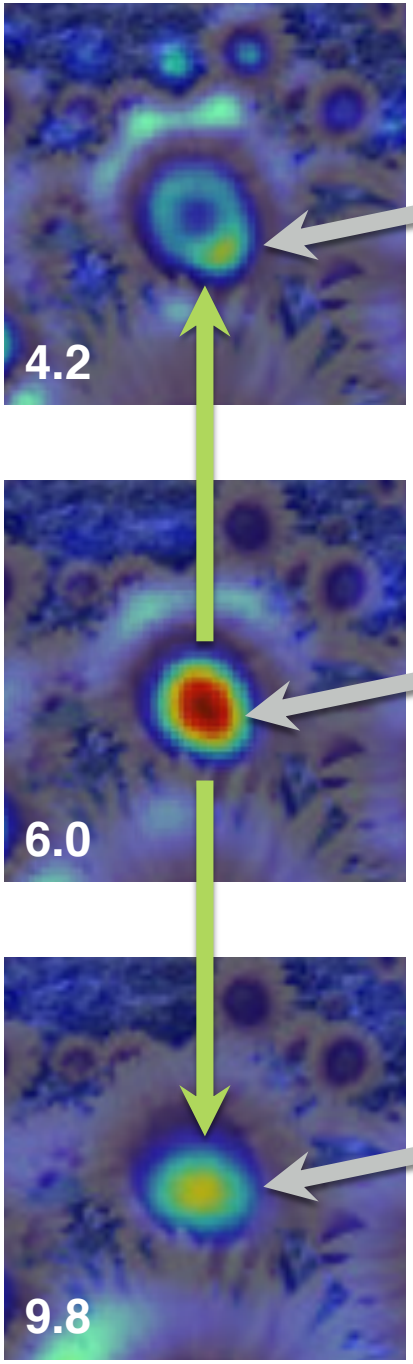



Full size image




3/4 size image

cross-scale maximum




local maximum

local maximum

local maximum

# implementation

```
For each level of the Gaussian pyramid

      compute feature response (e.g. Harris, Laplacian)
```

```
For each level of the Gaussian pyramid

     if local maximum and cross-scale

     save scale and location of feature
```

$(x, y, s)$

# Scale-invariant Interest Point Detection

- Which function? …
- Good filters (kernels)

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

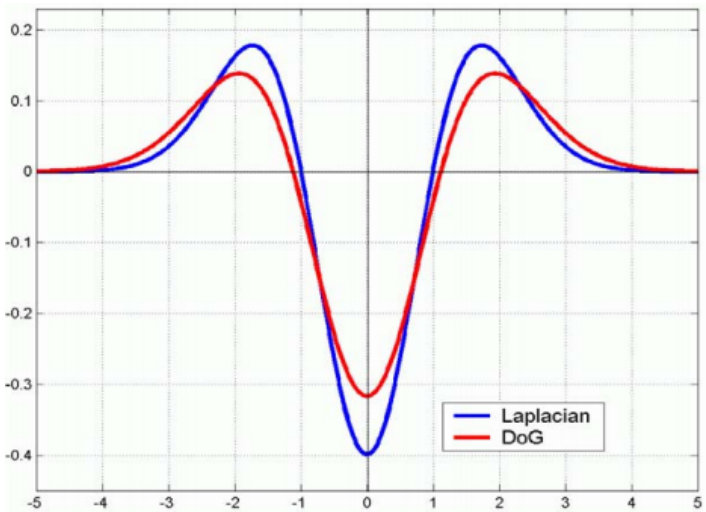
(Laplacian)

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

(Difference of Gaussians)

- Both filters are invariant to size and orientation

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



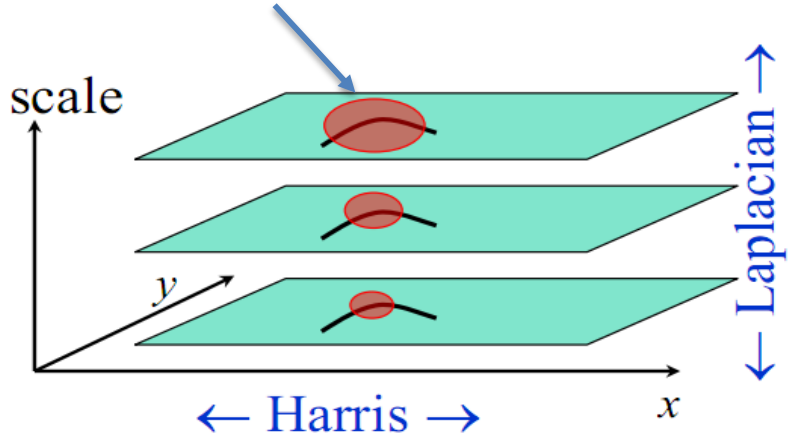

Credit: Markus Vincze, Technische Universität Wien

# Scale-invariant Interest Point Detection

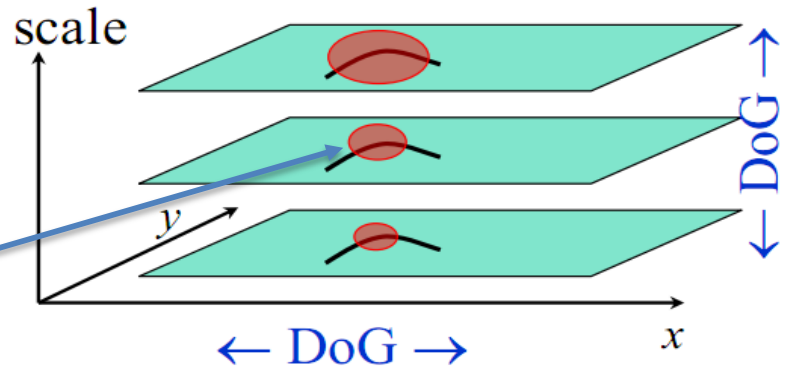
## Harris vs. Lowe

1. Find Harris interest point
2. Compute Laplacian at each scale
3. Identify local maxima across scale at these Harris points

- Harris-Laplacian[1]
  *Find local maximum of:*
  - Harris corner detector in space (image coordinates)
  - Laplacian in scale




- SIFT (Lowe)[2]
  *Find local maximum of:*
  - Difference of Gaussians in space and scale




1. Computer DoG at each scale
2. Identify local extrema in all three dimensions

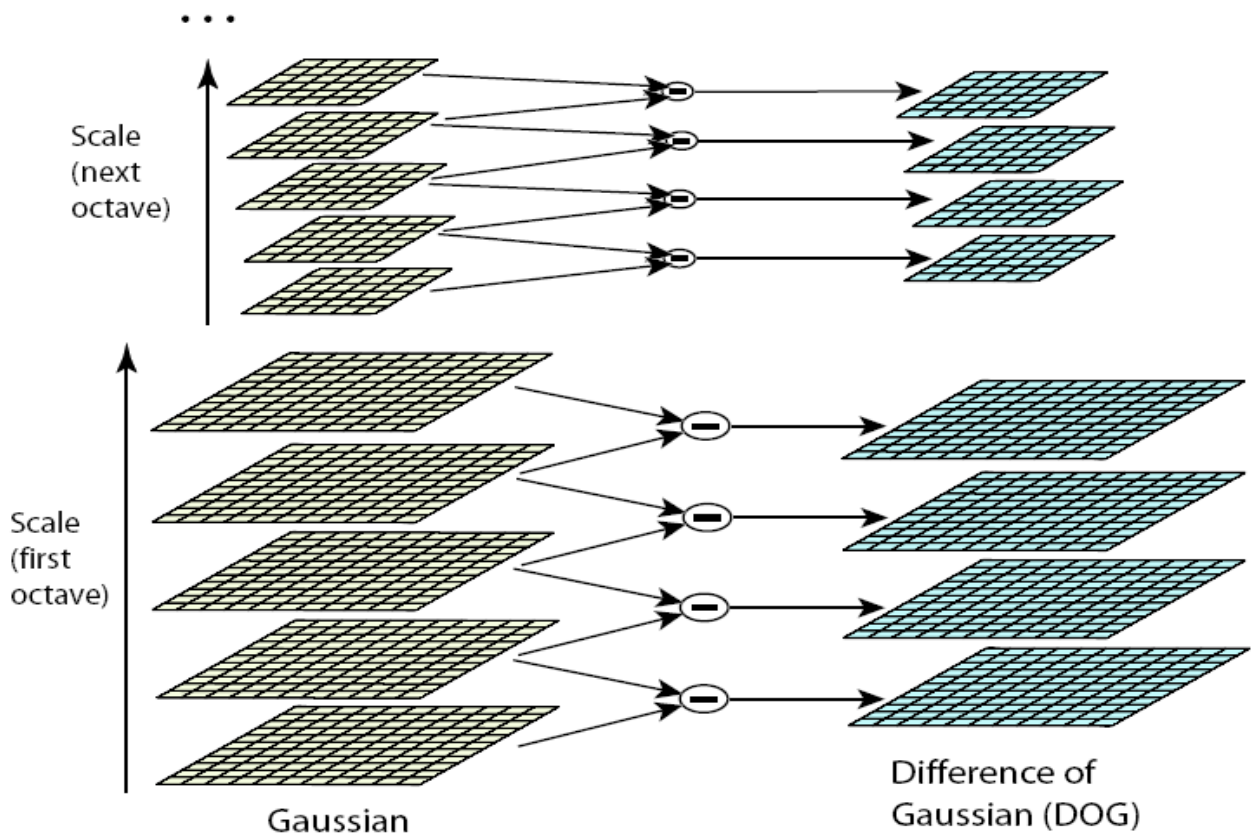
[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001
[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". IJCV 2004

Credit: Markus Vincze, Technische Universität Wien

# Scale-invariant Interest Point Detection

Difference of Gaussian (DoG)

- Search maximum in DoG image pyramid
- For every octave use $s$ sub-scales (= gradually smoothed images within octave)

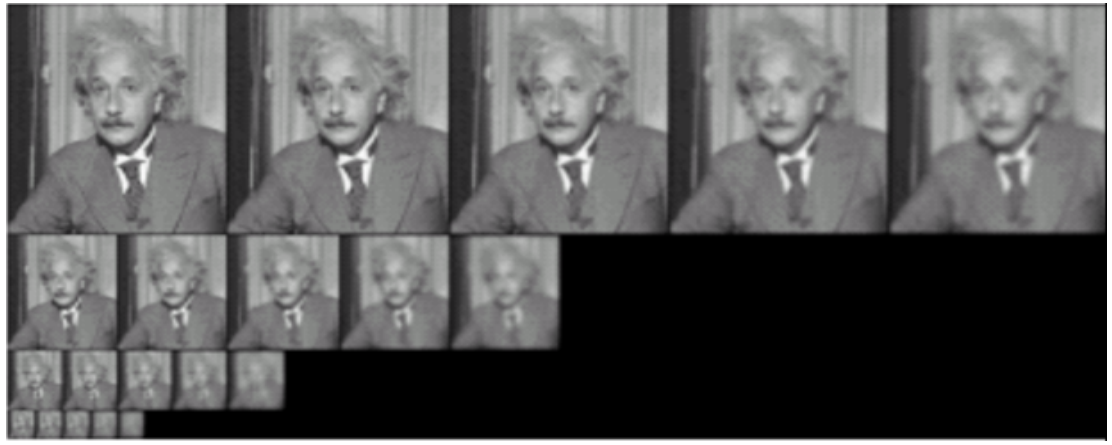



Credit: Markus Vincze, Technische Universität Wien

# Scale-invariant Interest Point Detection

Difference of Gaussian (DoG)

- Gaussian smoothed image at size (scale) $\sigma$ and $s\sigma$
- Grouped by octaves (i.e. doubling of $\sigma$)
- $s$ sets the number of images per octave
- DoG images grouped by octave

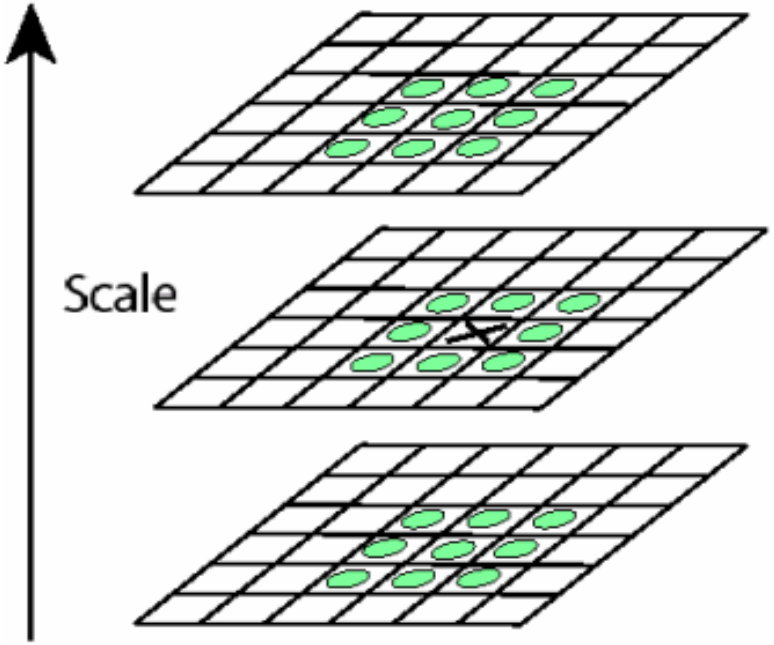


Credit: Markus Vincze, Technische Universität Wien

# Scale-invariant Interest Point Detection

Difference of Gaussian (DoG)

- Every pixel in the DoG images is compared with its 8 neighbours and the 9 neighbours of the neighbouring scale Gauss images (scale space images) … 26 in total
- Interest points (key points, frames) are local maxima or minima of the DoG over all sizes

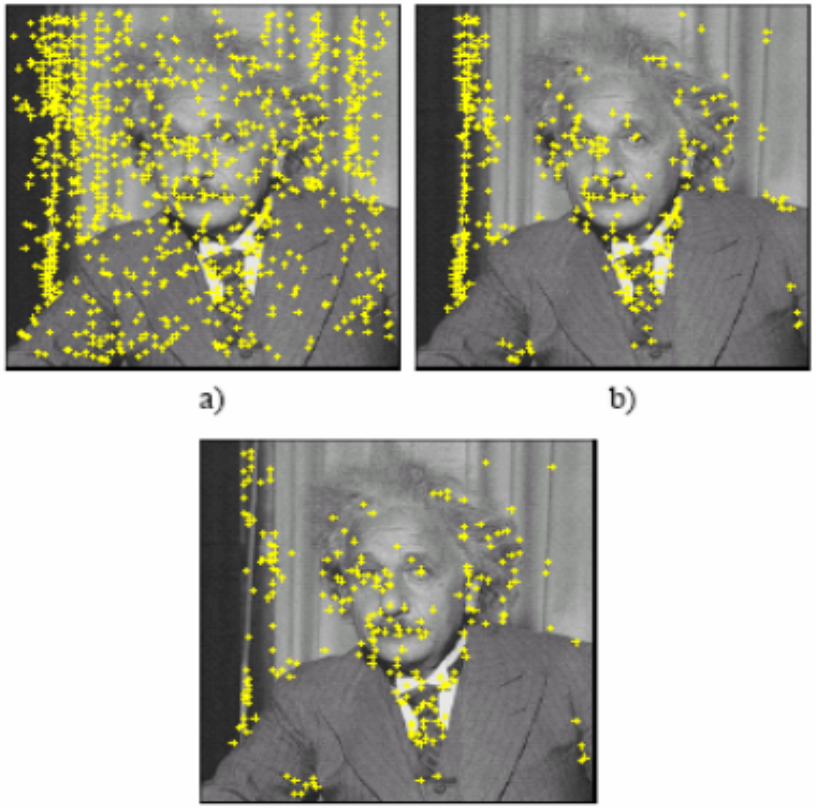



Credit: Markus Vincze, Technische Universität Wien

# Scale-invariant Interest Point Detection

Difference of Gaussian (DoG)

- From all interest points (a) the following are
  - Points with low contrast are eliminated (b)
  - Points along edges are eliminated (c)
  - Points are interpolated to obtain the position more accurately




Credit: Markus Vincze, Technische Universität Wien

# Interest Point Detection Wish List

- Distinctive
- Scale invariance
- Rotation invariance
- Local (parts) - global (entire object)
- Scalability to a large number of features (quantity)
- Stability to noise
- Other lighting situation
- Stability to changes in viewing angle

Credit: Markus Vincze, Technische Universität Wien

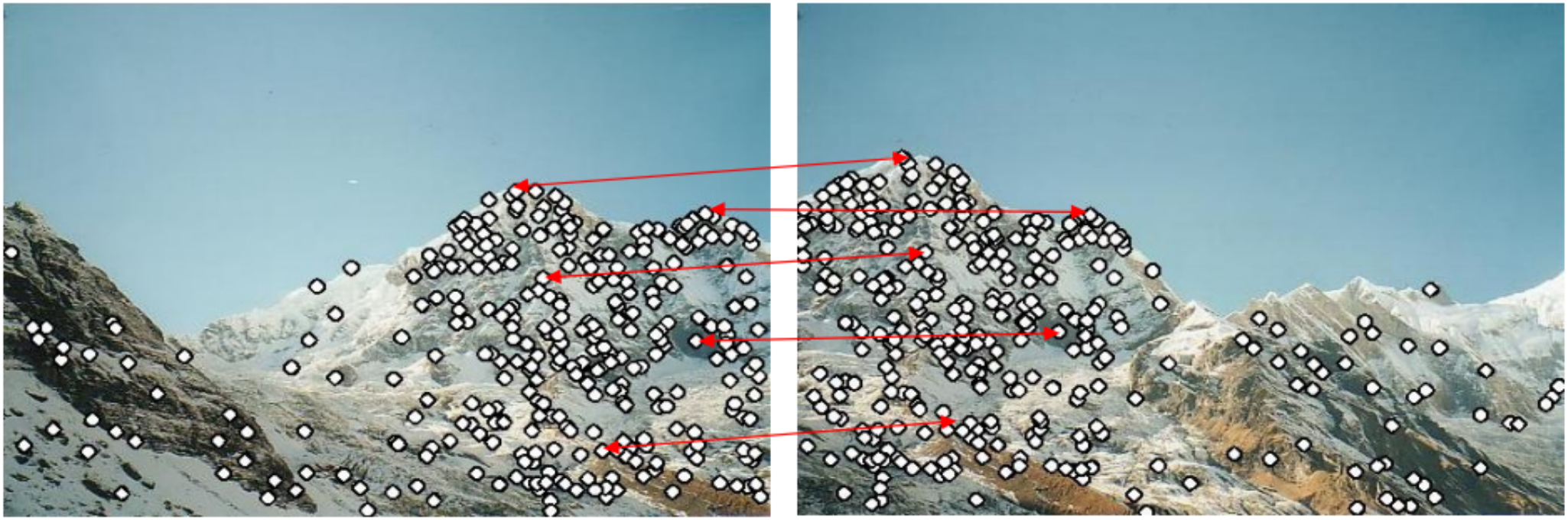
# Interest Point Detection Wish List

- Points (regions, features) with a high gradient
- Good repeatability on new images containing the same objects
- Multitude of proposed approaches
- Standard: Harris, DoG
- Basis for recognising objects (Bob's mug)
  and object classes (mugs in general)

Credit: Markus Vincze, Technische Universität Wien

# Matching of Interest Points

- Detection of distinctive “interest points” (key points)
- Distinctive: features suited for re-finding



Credit: Markus Vincze, Technische Universität Wien

# Objects and Interest Points (IPs)

1. Feature detection
   Extract interest points
   (unique image regions)

2. Feature description
   Calculate local (invariant) descriptors

3. Feature matching / feature tracking
   Find correspondences

4. Find similar image regions/objects

Credit: Markus Vincze, Technische Universität Wien

# Reading

R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.

Section 4.1 Points and Patches

Section 4.1.1 Feature Detectors