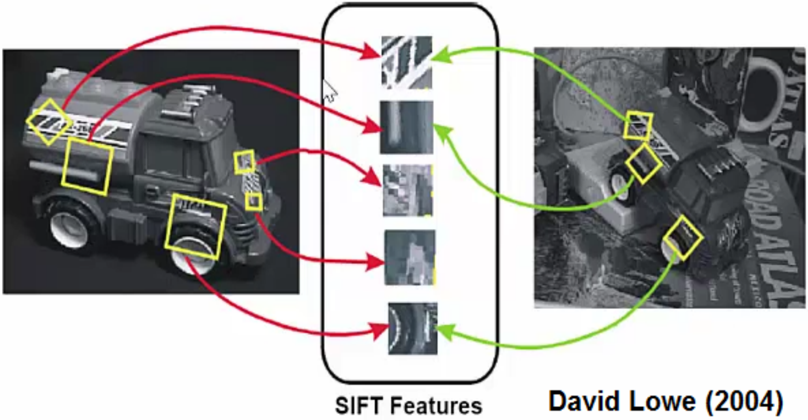
Image Features:

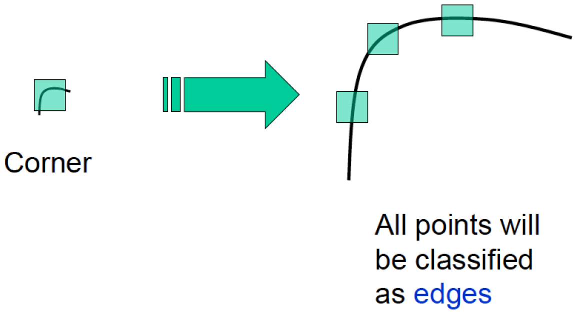
Scale Invariant Interest Point Detection

Scale Invariant Feature Transform



Remember from Harris Corner Detector

- Scale?



- Corner location is **not scale invariant**!

[Source: J. Hays]

Our Goal: Matching Objects / Images

- Our goal is to be able to match an object in different images where the object appears in different scale, rotation, viewpoints, etc. **How?**

Figure: We want to be able to match these two objects / images

Our Goal: Matching Objects / Images

- Our goal is to be able to match an object in different images where the object appears in different scale, rotation, viewpoints, etc. **How?**

image 1

image 2

Figure: But these shouldn't be matched!

Our Goal: Matching Objects / Images

- Our goal is to be able to match an object in different images where the object appears in different scale, rotation, viewpoints, etc. **How?**

image 1

Figure: Find some interest points in an image

Our Goal: Matching Objects / Images

- Our goal is to be able to match an object in different images where the object appears in different scale, rotation, viewpoints, etc. **How?**

image 2

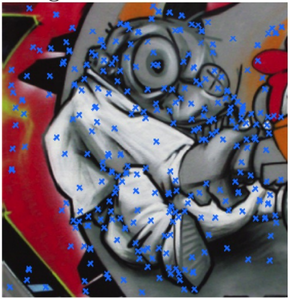


Figure: And **independently** in other images

Our Goal: Matching Objects / Images

- Our goal is to be able to match an object in different images where the object appears in different scale, rotation, viewpoints, etc. **How?**

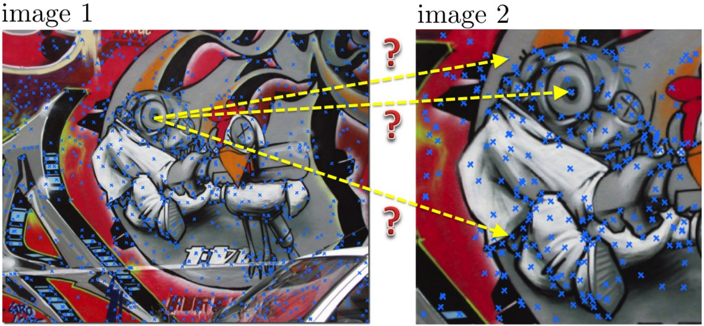


Figure: How can we match points??

Our Goal: Matching Objects / Images

- Our goal is to be able to match an object in different images where the object appears in different scale, rotation, viewpoints, etc. **How?**

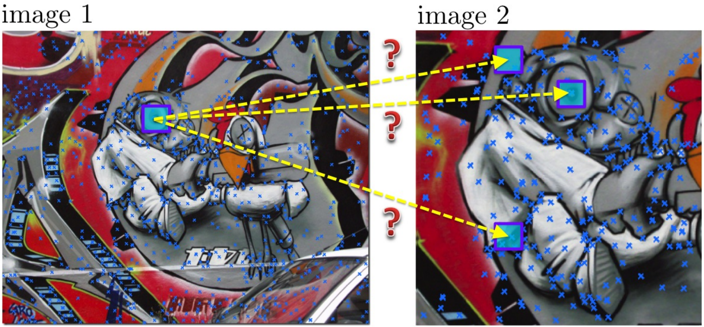


Figure: We could match if we took a patch around each point, and describe it with a feature vector (we know how to compare vectors)

Scale Invariant Interest Points

How can we **independently** select interest points in each image, such that the detections are repeatable across different scales?

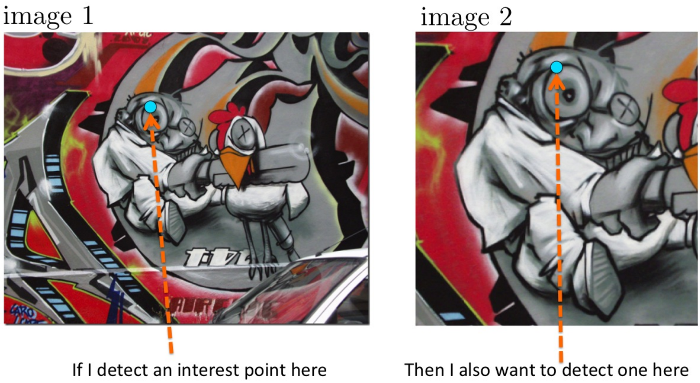
image 1

image 2

[Source: K. Grauman]

Scale Invariant Interest Points

How can we **independently** select interest points in each image, such that the detections are repeatable across different scales?

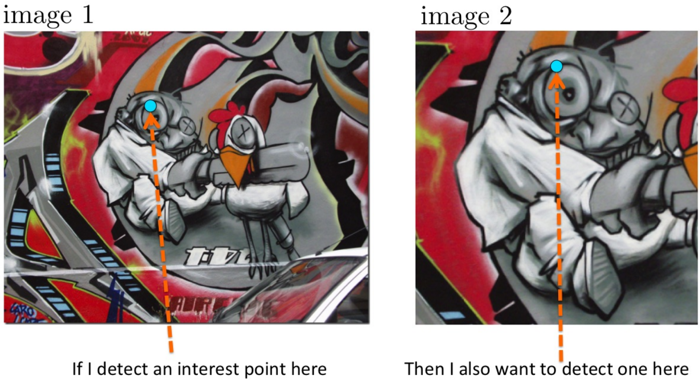


[Source: K. Grauman]

Scale Invariant Interest Points

How can we **independently** select interest points in each image, such that the detections are repeatable across different scales?

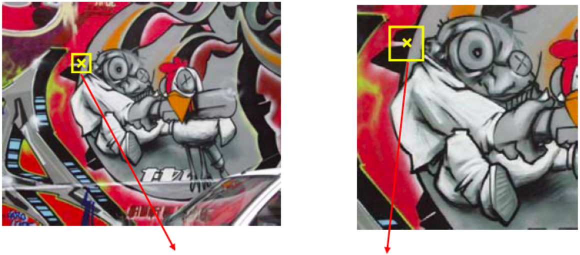
- Extract features at a variety of scales, e.g., by using multiple resolutions in a pyramid, and then matching features at the *"corresponding"* level.
- When does this work?



Scale Invariant Interest Points

How can we **independently** select interest points in each image, such that the detections are repeatable across different scales?

- More efficient to extract features that are stable in both location and scale.



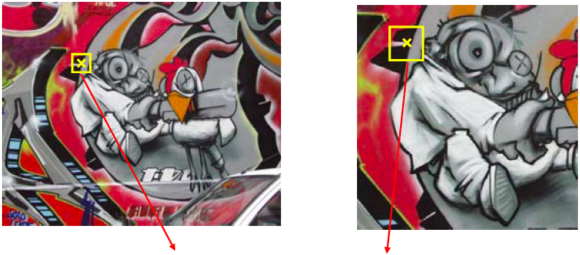
$$f(I_{i_1 \dots i_m}(x, \sigma)) = f(I_{i_1 \dots i_m}(x', \sigma'))$$

[Source: K. Grauman]

Scale Invariant Interest Points

How can we **independently** select interest points in each image, such that the detections are repeatable across different scales?

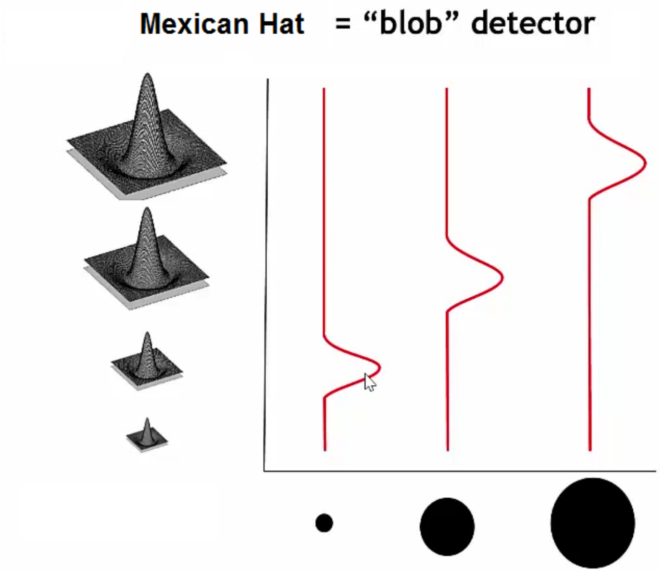
- With the Harris corner detector we found a maxima in a spatial search window
- Find scale that gives local maxima of a function f in both position and scale.



$$f(I_{i_1 \dots i_m}(x, \sigma)) = f(I_{i_1 \dots i_m}(x', \sigma'))$$

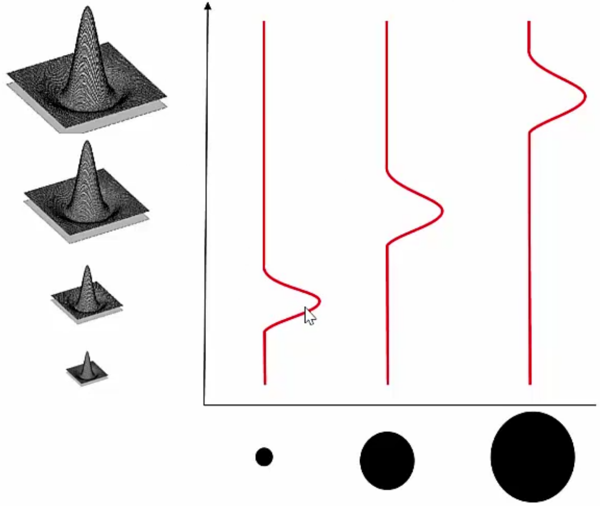
[Source: K. Grauman]

What Can the Signature Function Be?



What Can the Signature Function Be?

- **Laplacian-of-Gaussian = "blob" detector**

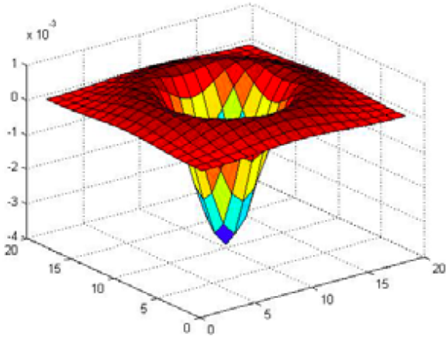


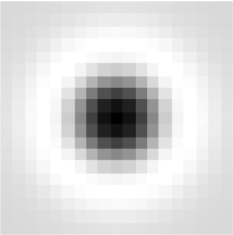
Blob Detection – Laplacian of Gaussian

- Laplacian of Gaussian: We mentioned it for edge detection

$$\nabla^2 g(x, y, \sigma) = \frac{\partial^2 g(x, y, \sigma)}{\partial x^2} + \frac{\partial^2 g(x, y, \sigma)}{\partial y^2}, \quad \text{where } g \text{ is a Gaussian}$$

- It is a circularly symmetric operator (finds difference in all directions)
- It can be used for 2D blob detection! How?



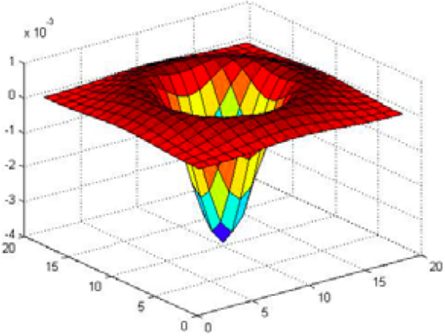


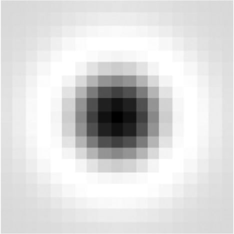
Blob Detection – Laplacian of Gaussian

- Laplacian of Gaussian: We mentioned it for edge detection

$$\nabla^2 g(x, y, \sigma) = -\frac{1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right)\exp^{-\frac{x^2+y^2}{2\sigma^2}}$$

- It is a circularly symmetric operator (finds difference in all directions)
- It can be used for 2D blob detection! How?

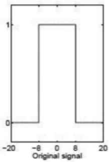




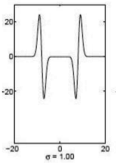
Blob Detection – Laplacian of Gaussian

- It can be used for 2D blob detection! How?

Original signal



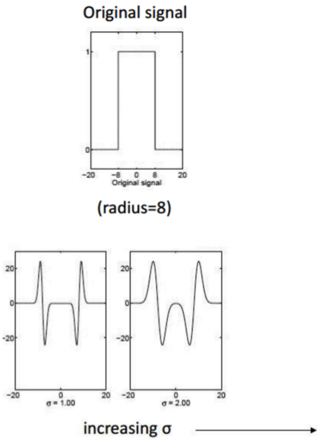
(radius=8)

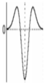


[Source: F. Flores-Mangas]

Blob Detection – Laplacian of Gaussian

- It can be used for 2D blob detection! How?

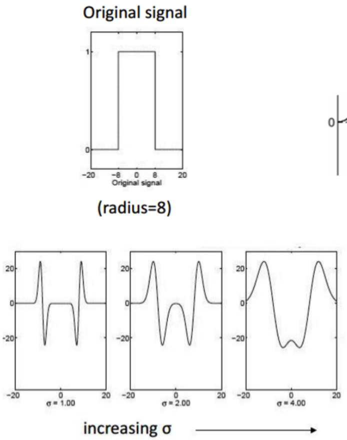


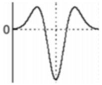


[Source: F. Flores-Mangas]

Blob Detection – Laplacian of Gaussian

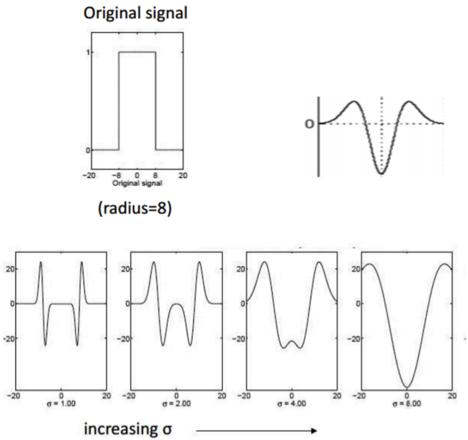
- It can be used for 2D blob detection! How?





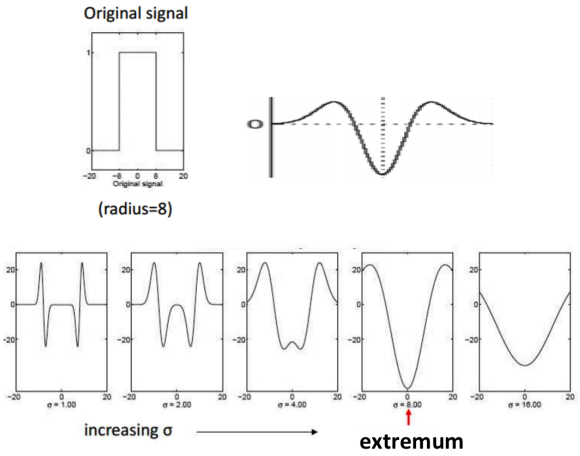
[Source: F. Flores-Mangas]

- It can be used for 2D blob detection! How?



[Source: F. Flores-Mangas]

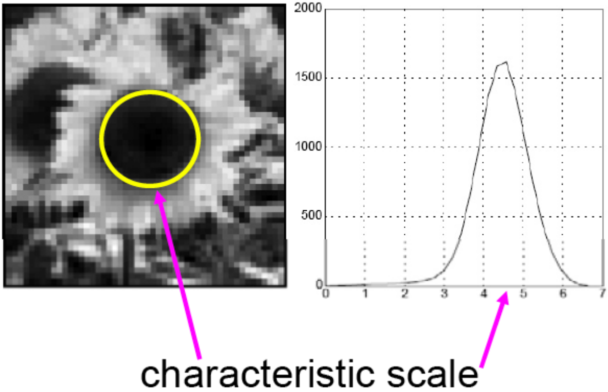
- It can be used for 2D blob detection! How?



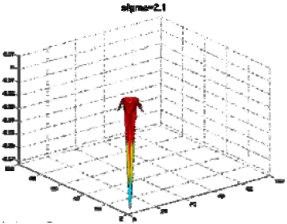
[Source: F. Flores-Mangas]

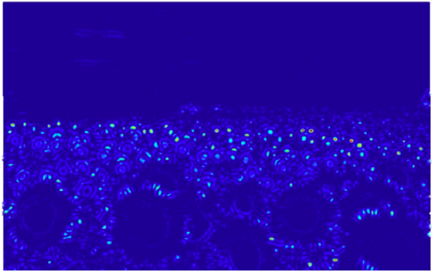
Characteristic Scale

- We define the **characteristic scale** as the scale that produces peak (minimum or maximum) of the Laplacian response



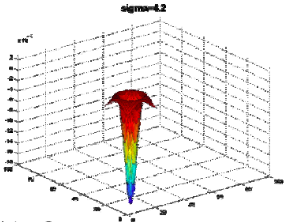
[Source: S. Lazebnik]

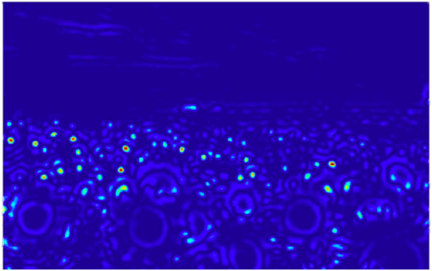




[Source: K. Grauman]

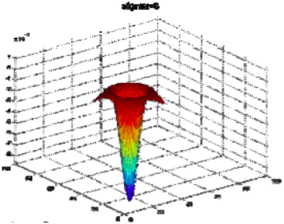
Example

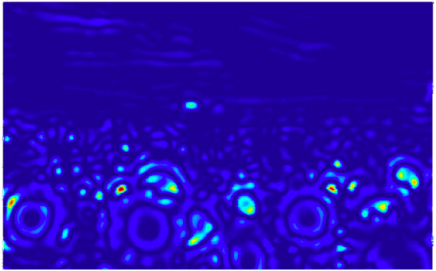




[Source: K. Grauman]

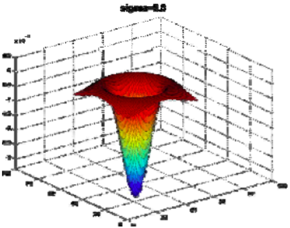
Example

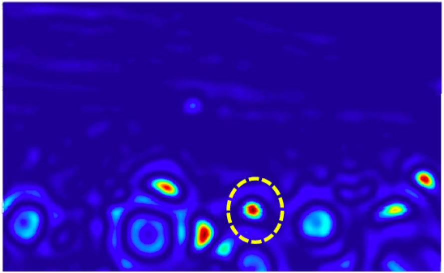




[Source: K. Grauman]

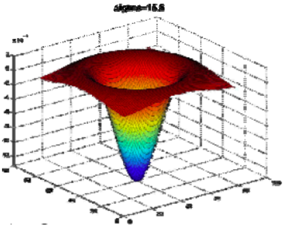
Example

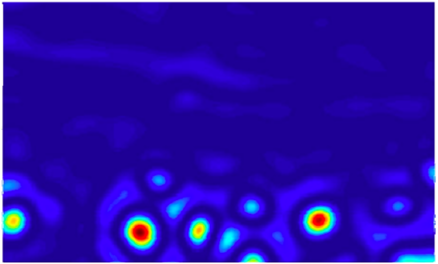




[Source: K. Grauman]

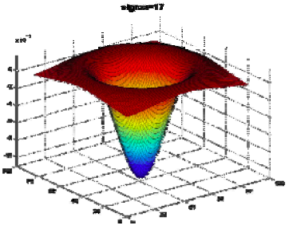
Example

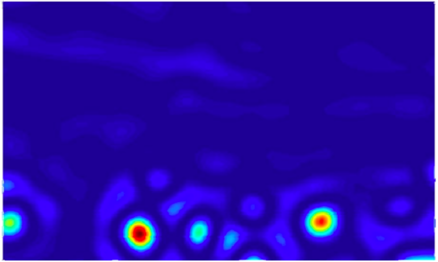




[Source: K. Grauman]

Example

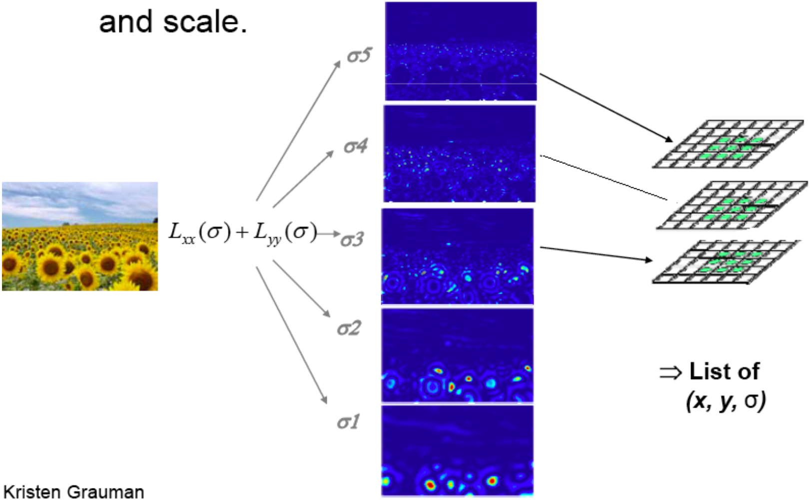




[Source: K. Grauman]

Scale Invariant Interest Points

Interest points are local maxima in both position and scale.



Kristen Grauman

Blob Detection – Laplacian of Gaussian

- That's nice. But can we do faster?
- Remember again the Laplacian of Gaussian:

$$\nabla^2 g(x, y, \sigma) = \frac{\partial^2 g(x, y, \sigma)}{\partial x^2} + \frac{\partial^2 g(x, y, \sigma)}{\partial y^2}, \quad \text{where } g \text{ is a Gaussian}$$

Blob Detection – Laplacian of Gaussian

- That's nice. But can we do faster?
- Remember again the Laplacian of Gaussian:

$$\nabla^2 g(x, y, \sigma) = \frac{\partial^2 g(x, y, \sigma)}{\partial x^2} + \frac{\partial^2 g(x, y, \sigma)}{\partial y^2}, \quad \text{where } g \text{ is a Gaussian}$$

$$\nabla^2 g(x, y, \sigma) = -\frac{1}{\pi\sigma^4}\left(1 - \frac{x^2 + y^2}{2\sigma^2}\right) \exp^{-\frac{x^2+y^2}{2\sigma^2}}$$

- Is this separable?
- Larger scale (σ), larger the filters (more work for convolution)
- Can we do it faster?

Approximate the Laplacian of Gaussian

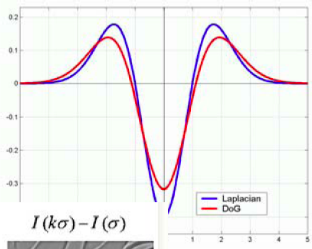
- We can approximate the Laplacian with a difference of Gaussians; and use separable convolution.

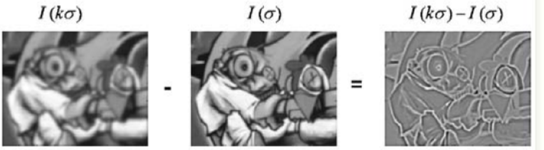
$$L = \sigma^2 \left(G_{xx}(x,y,\sigma) + G_{yy}(x,y,\sigma) \right)$$

(Laplacian)

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

(Difference of Gaussians)

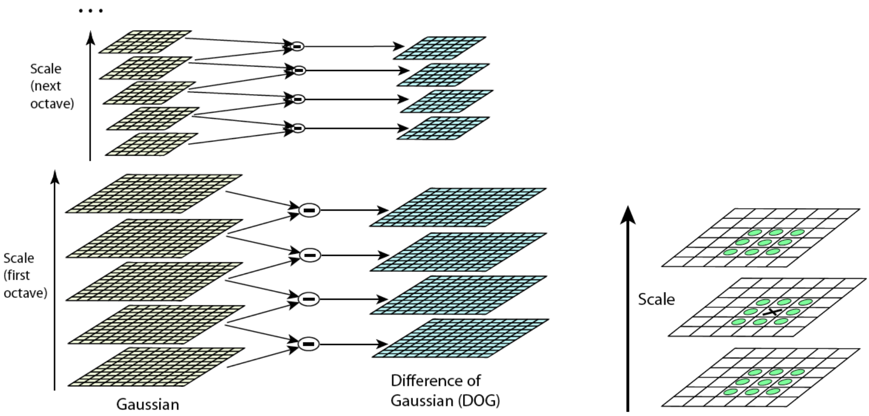




[Source: K. Grauman]

Lowe's DoG

- Lowe (2004) proposed computing a set of sub-octave Difference of Gaussian filters looking for 3D (space+scale) maxima in the resulting structure



[Source: R. Szeliski]

Lowe's DoG

- First compute a Gaussian image pyramid

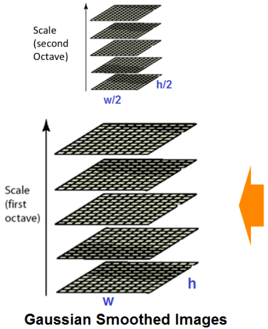
$$I_s = I * G_{k^s\sigma}$$

$$\vdots$$

$$I_2 = I * G_{k^2\sigma}$$

$$I_1 = I * G_{k\sigma}$$

$$I_0 = I * G_{\sigma}$$



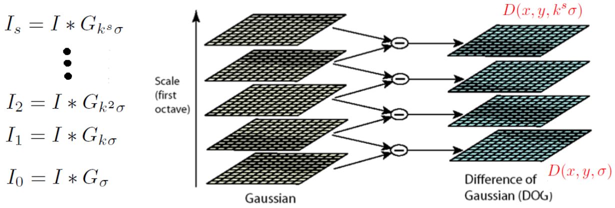
Each image is smoothed by a factor of k more than the image below

[Source: F. Flores-Mangas]

Lowe's DoG

- First compute a Gaussian image pyramid
- Compute Difference of Gaussians

$$D(x,y,\rho) = I(x,y) * (G(x,y,k\rho) - G(x,y,\rho))$$
$$\text{for } \rho = \{\sigma,\, k\sigma,\, k^2\sigma, \ldots, k^{s-1}\sigma\}, \quad k = 2^{1/s}$$



[Source: F. Flores-Mangas]

Lowe's DoG

- First compute a Gaussian image pyramid
- Compute Difference of Gaussians
- At every scale

[Source: F. Flores-Mangas]

Lowe's DoG

- First compute a Gaussian image pyramid
- Compute Difference of Gaussians
- At every scale
- Find local maxima in scale
- A bit of pruning of bad maxima and we're done!

[Source: F. Flores-Mangas]

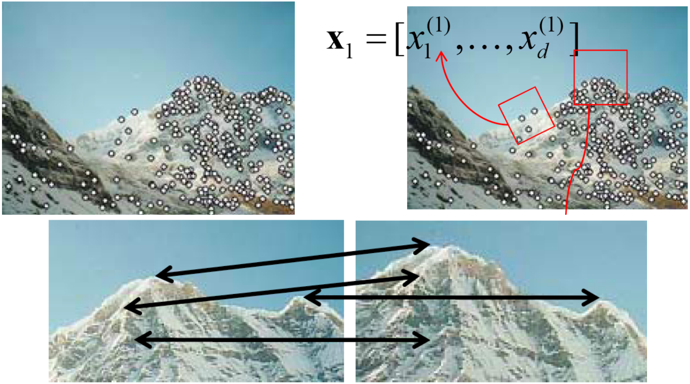
Other Interest Point Detectors (Many Good Options!)

- Lindeberg: Laplacian of Gaussian
- Lowe: DoG (typically called the SIFT interest point detector)
- Mikolajczyk & Schmid: Hessian/Harris-Laplacian/Affine
- Tuyttelaars & Van Gool: EBR and IBR
- Matas: MSER
- Kadir & Brady: Salient Regions

Summary – Stuff You Should Know

- To match the same scene or object under different viewpoint, it's useful to first detect **interest points** (keypoints)
- We looked at these interest point detectors:
  - Harris corner detector: translation and rotation but not scale invariant
  - Scale invariant interest points: Laplacian of Gaussians and Lowe's DoG
- Harris' approach computes I_x^2, I_y^2 and $I_x I_y$, and blurs each one with a Gaussian. Denote with: $A = g * I_x^2$, $B = g * (I_x I_y)$ and $C = g * I_y^2$. Then $M_{xy} = \begin{pmatrix} A(x,y) & B(x,y) \\ B(x,y) & C(x,y) \end{pmatrix}$ characterizes the shape of E_{WSSD} for a window around (x, y). Compute "cornerness" score for each (x, y) as $R(x,y) = \det(M_{xy}) - \alpha \operatorname{trace}(M_{xy})^2$. Find $R(x,y) > \text{threshold}$ and do non-maxima suppression to find corners.
- Lowe's approach creates a Gaussian pyramid with s blurring levels per octave, computes difference between consecutive levels, and finds local extrema in space and scale

- **Detection**: Identify the interest points.
- **Description**: Extract a feature descriptor around each interest point.
- **Matching**: Determine correspondence between descriptors in two views.



[Source: K. Grauman]