Scale Space

a 1000 ways of inventing the Gaussian

Tony Lindeberg, Bart M Ter Haar Romeny

Presented by: Sameer Agarwal

Artificial Intelligence Laboratory, UCSD

Understanding an Image

Understanding an Image

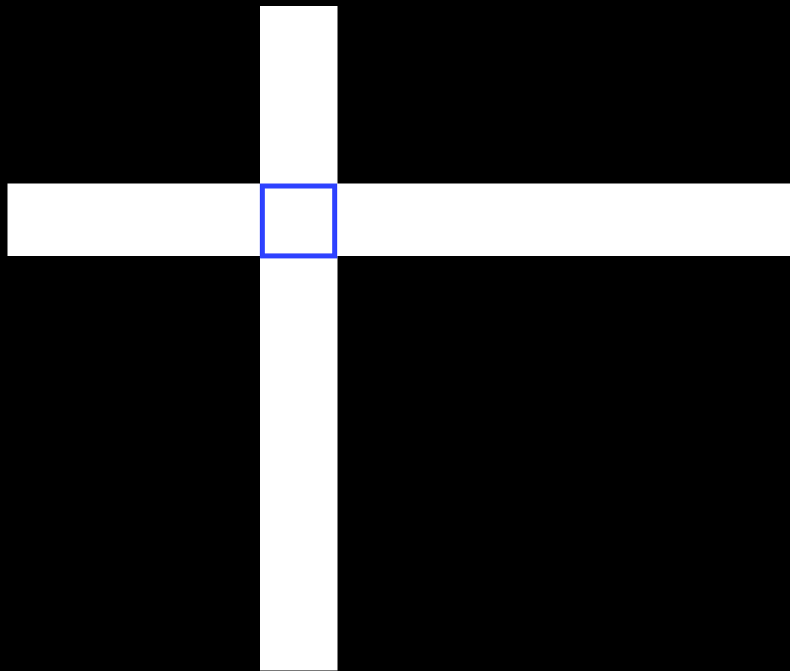


Image as a map

- An image is a representation of a scene.
- Is it the best possible representation ?
- How can we make use the image to navigate in the scene ?
- What if I don’t want the fine structure ?

The problem of scale

- Most algorithms in computer vision assume that the scale of interpretation of an image has been decided a priori.
- How do you decide the scale of an operator in the absence of any *apriori* knowledge?

The solution

- Create a multiscale representation of the image by generating a family of images where fine-scale information is successively supressed.
- Characterize the basic features in terms of optima of functions defined over this family of images.

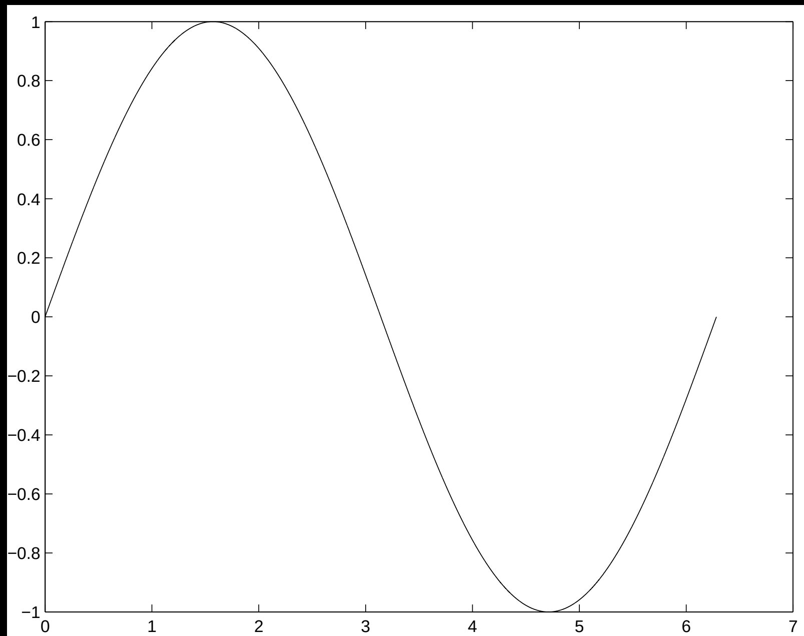
Problem: How do we construct this multiscale representation ?

The problem of Differentiation

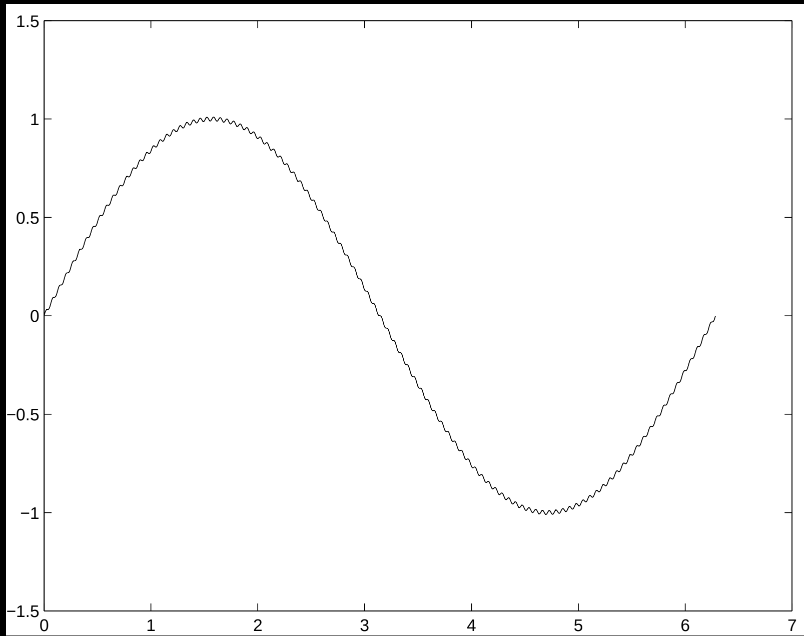
$$g(x) = f(x) + \epsilon \sin \omega x$$

$$g_x(x) = f_x(x) + \omega\epsilon \cos \omega x$$

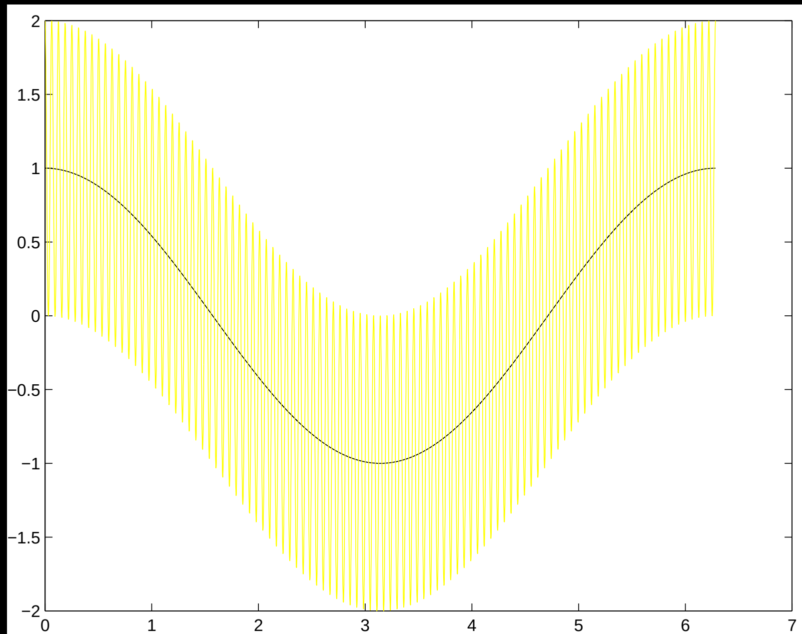
Example 1



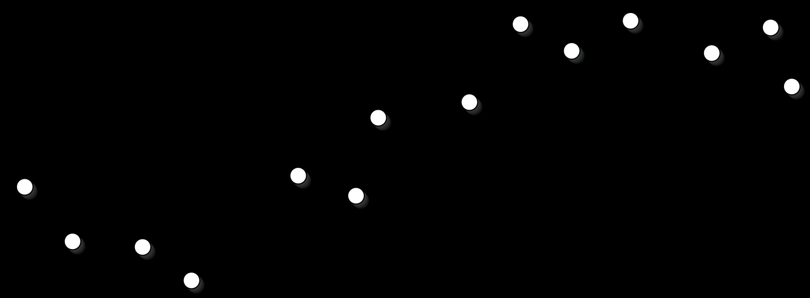
Example 1



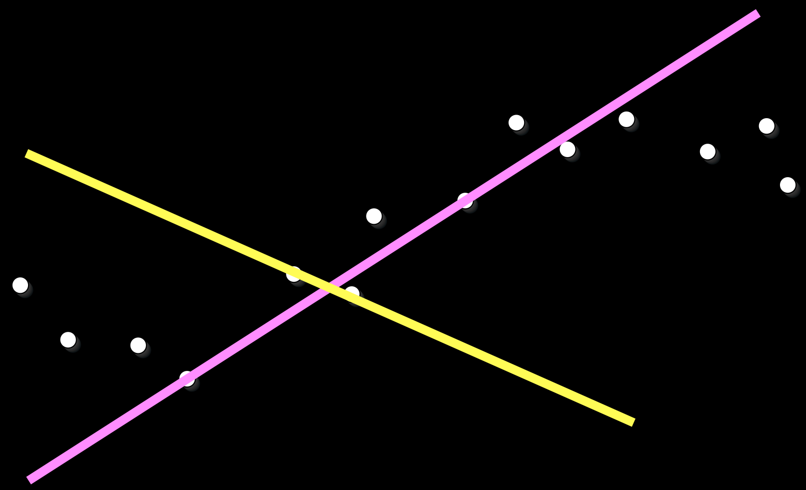
Example 1



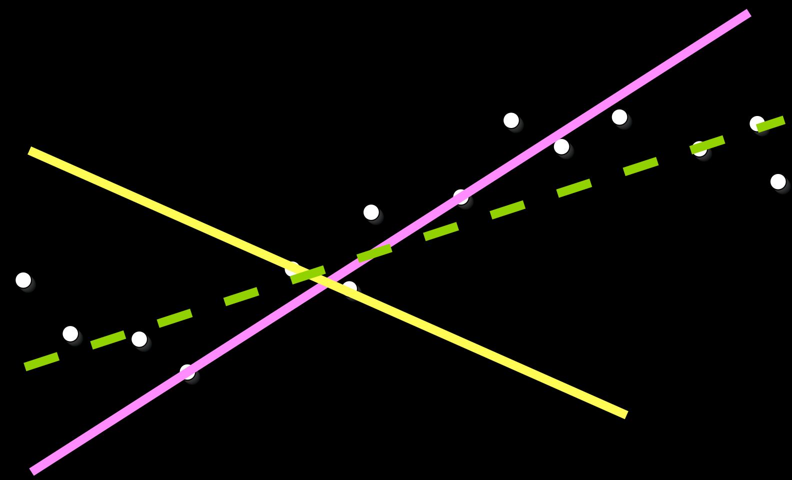
Example 2



Example 2



Example 2



Schwartz's Theory of Distributions

- Differentiation is an ill posed problem.
- The output of the procedure is not a unique continuous function of its inputs.
- The problem gets worse with data which is discontinuous.
- Images ARE discontinuous data.

Schwartz's Theory of Distributions

- Instead of observing the function directly, use a probe.
- Probe = Infinitely differentiable function.
- Probing = Correlation.
- Differentiation = correlation with the derivative of the probe.
- Problem : What is a good probe?

Previous Work

- Quad Trees
- Pyramids
  - Gaussian
  - Laplacian

Quad Trees

- Consider an image D of size $2^K \times 2^K$.
- Define a measure Σ of variation in any region $D' \subset D$.
- Let $D^{(K)} = D$.
- if $\Sigma(D^{(K)}) > \alpha$, split $D^{(K)}$ into p sub images $D_j^{(K-1)}, j = 1, \ldots, p$.
- Represent the image using the degree p tree constructed above.

Pyramids

Multiscale representation.

1. $f^k - = f$.
2. $f^{k-1} = \mathrm{Reduce}(f^k)$.
3. $f^{k-1}(x) = \sum_n c(n) f^k(2x - n)$

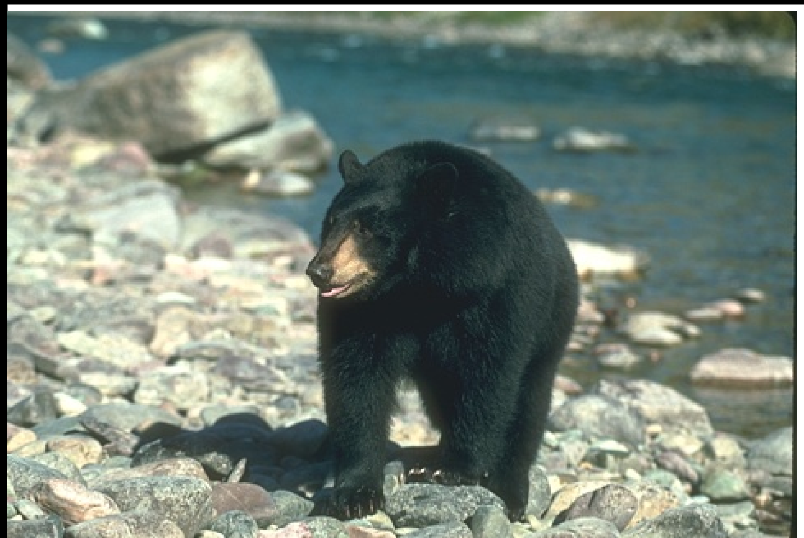
Conditions on c.

1. $c(n) > 0$
2. $c(|n|) \geq c(|n+1|)$
3. $c(-n) = c(n)$
4. $\sum_n c(n) = 1$
5. $\sum_n c(2n) = \sum_n c(2n+1)$

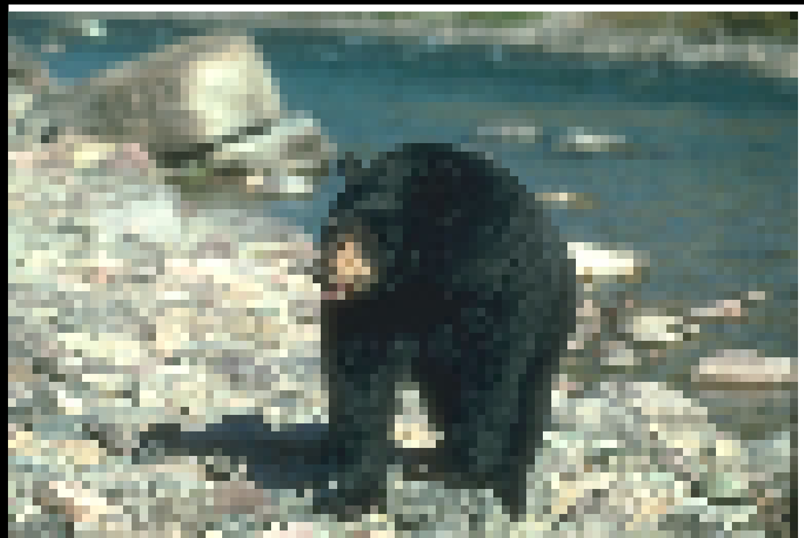
Multiscale v/s Multiresolution

- Pyramids reduce the size of the image as you go up.
- Sampling rate reduces.
- Useful for compression.
- Complicated to analyse.
- Multiscale methods like scale space preserve the sampling rate.

Why not just subsample ?



Why not just subsample ?



Subsampling creates image features which are not present in the scene.

The Desiderata

$$L(x;t) = h(L_0(x), t)$$

- Shift Invariance
- Linearity
- Isotropy
- Causality
- Semi-Group structure
- Separability

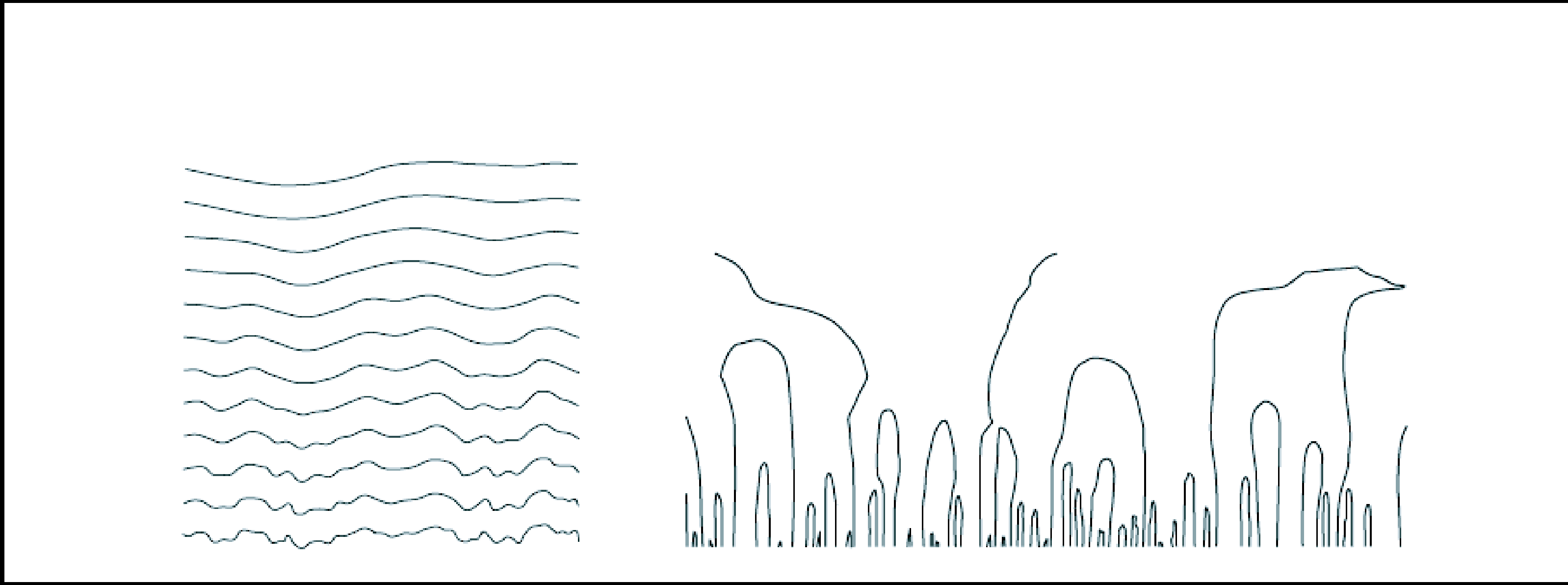
Linear Shift Invariance and Isotropy

Each part of the image is processed without regard to its position in the image. The process itself is linear. Isotropy implies that all directions are treated as equal. Implication: The process giving rise to the scale space is a family of convolution operators. i.e.

$$L(\cdot;t) = h(\cdot;t) * L_0$$

Causality

- Image features at a scale are only dependent on the image features on the scales finer that itself.
- The process does not create any new featues.
- As an implication, the number of extrema is non-increasing with increase in scale.



The Heat Equation

Koenderink (1984) showed that causality implied

$$\frac{\partial L}{\partial t} = c^2 \nabla^2 L$$

- Describes the evolution of the temperature as a function of space and time.
- Also describes the process of diffusion, L then represents concentration.

Heat Equation in 1-d

For the case of a rod of length l and zero boundary conditions,

$$L(0,0) = 0; L(0,l) = 0$$

and initial condition

$$L(x;0) = f(x)$$

the solution is given by

$$L(x,t) = \sum_n B_n \sin(\frac{n\pi}{l}x)e^{-n^2\pi^2c^2t}$$

$$B_n = \frac{2}{l}\int_0^l f(x)\sin(\frac{n\pi}{l}x)dx$$

Semi-Group Structure

$$h(\cdot; t_1 + t_2) = h(\cdot; t_1) * h(\cdot; t_2)$$

more precisely

1. $L(\cdot; t_2) = h(\cdot; t_2) * L_0$
2. $L(\cdot; t_2) = (h(\cdot; t_2 - t_1) * h(\cdot; t_1) * L_0$
3. $L(\cdot; t_2) = h(\cdot; t_2 - t_1) * (h(\cdot; t_1) * L_0)$
4. $L(\cdot; t_2) = h(\cdot; t_2 - t_1) * L(\cdot; t_1)$

Separability

The kernel should factor into single dimensional kernels.

$$h(x_1, x_2; t) = h(x_1; t)h(x_2; t)$$

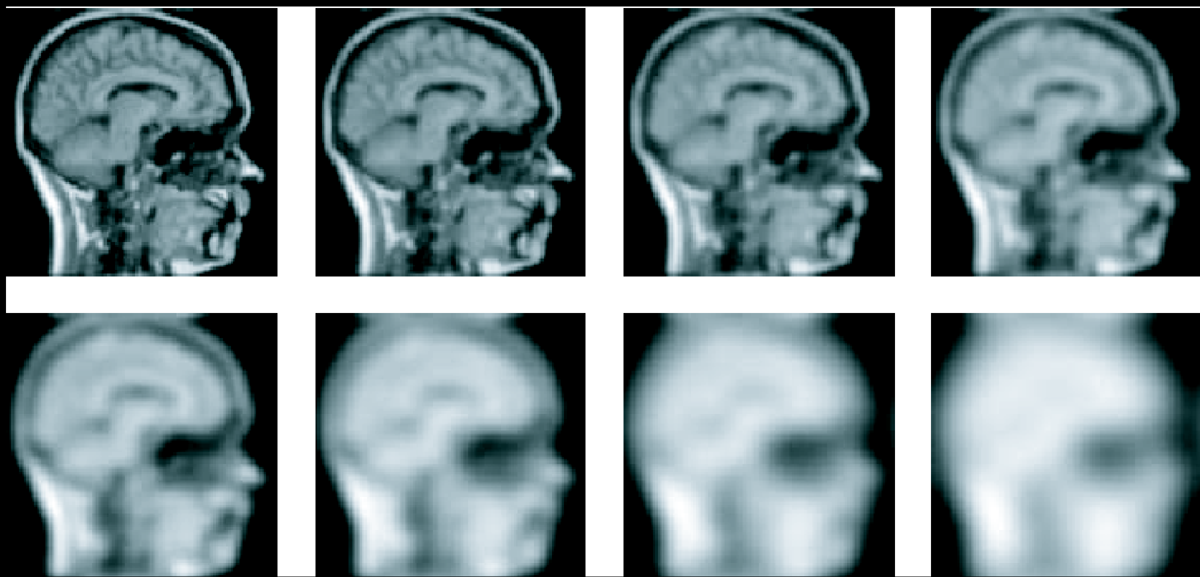
Decreases the complexity of calculating the convolution.

Presenting : The Gaussian

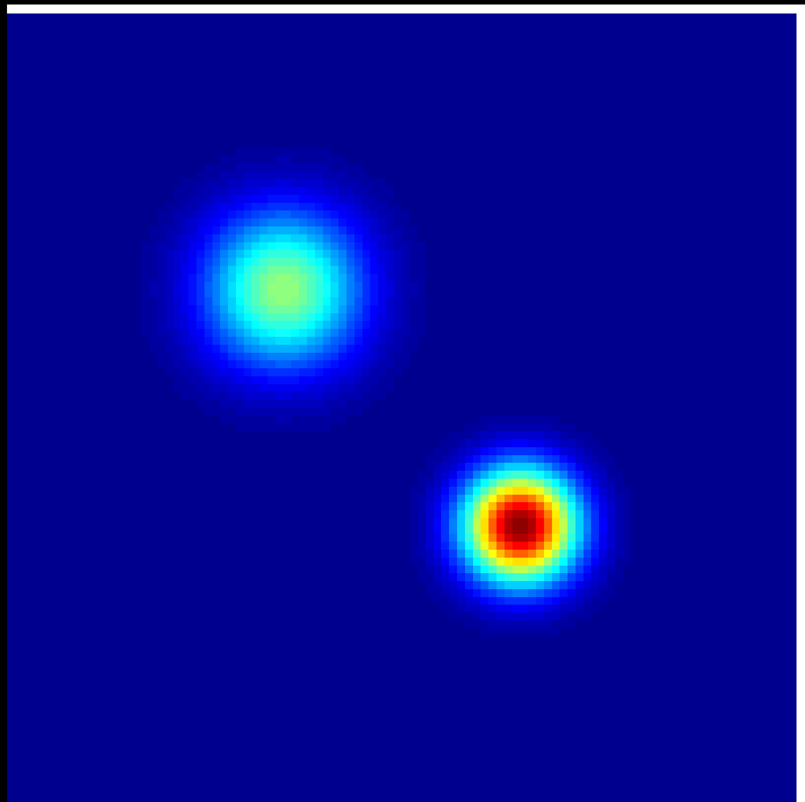
$$h(\mathbf{x};t) = \left(\frac{1}{\sqrt{2\pi t}}\right)^n e^{-\dfrac{\mathbf{x}^T\mathbf{x}}{2t}}$$

1. Rotationally symmetric
2. Separable
3. Infinite Differentiability
4. $L_{i_1,i_2,\ldots,i_n}(\mathbf{x};t) = \partial_{i_1,i_2,\ldots,i_n} L(\mathbf{x};t) = h_{i_1,i_2,\ldots,i_n} * L_0(\mathbf{x})$

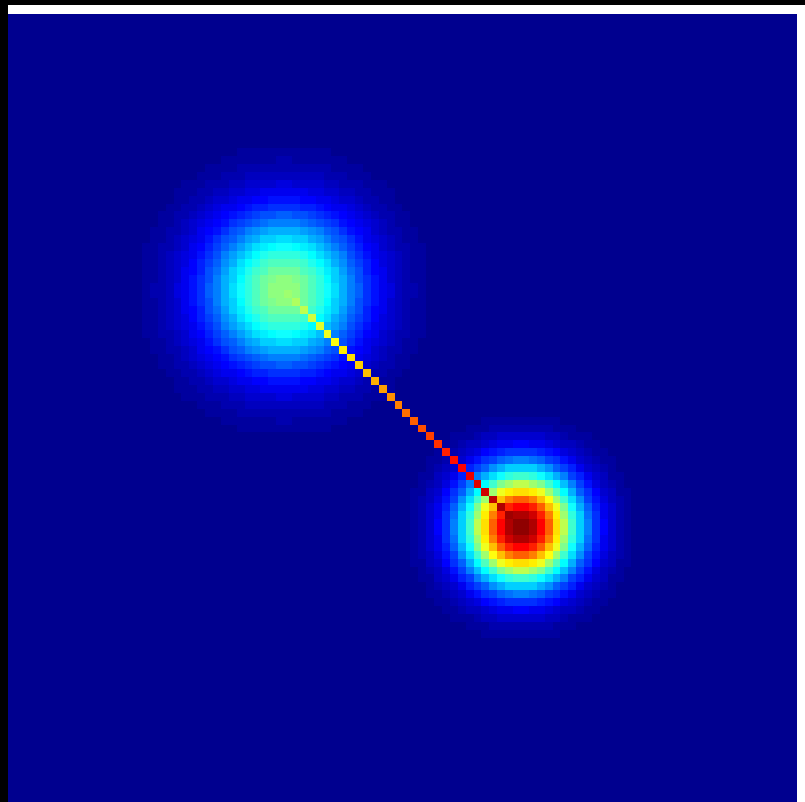
Example of Gaussian Scale Space



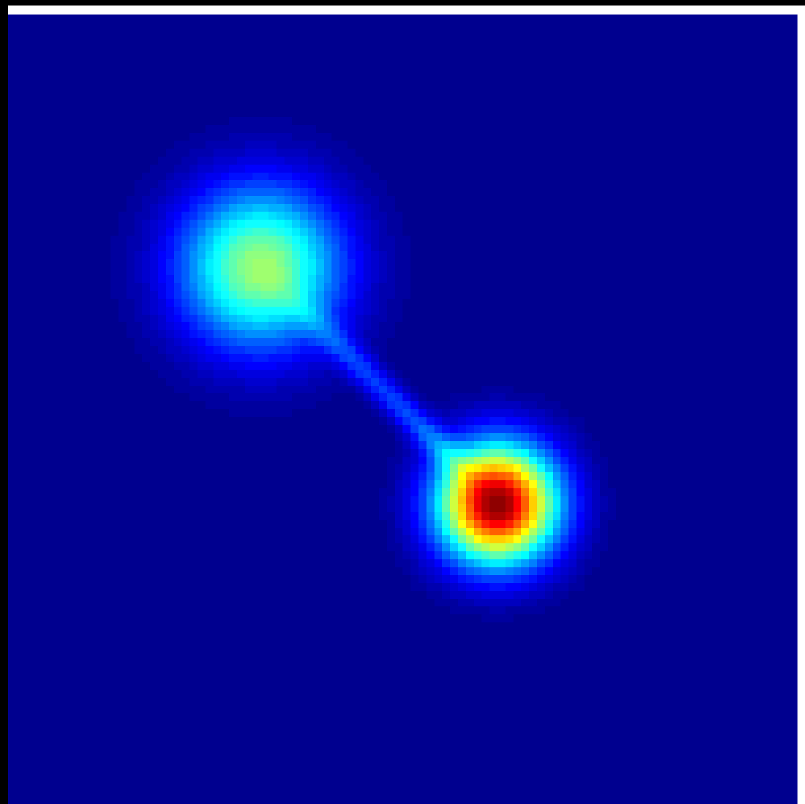
non non-creation of extrema



non non-creation of extrema



non non-creation of extrema



But images are not continuous !

- There exists a complete theory of discrete scale space.
- $L(x;t) = \sum_n T(n;2t) f(x-n)$
- $T(n;2t) = e^{-\alpha t} I_n(2\alpha t)$
- I_n is the modified Bessel function.
- In the limit $t \to \infty$, T approaches the continuous Gaussian.
- $\partial_t L(x;t) = L(x+1;t) - 2L(x;t) + L(x-1;t)$

Now what?

What can you do with a Gaussian scale space ?

What is a feature?

An image feature is the set of extrema of a function defined on an image.

1. How do you define functions on an image?
2. What are good functions?
3. What are are perceptually meaningful features?
4. What scale should the functions operate on ?

Differential Geometric Invariants

- Functions of the various derivatives of an image.
- Scale space allows us to take derivatives in spatial as well as scale dimensions.
- Consider invariants based on derivatives of order two and lower.

An observation about gradients

$$L_{x^\alpha}(x;t_1) > L_{x^\alpha}(x;t_2)$$

if,

$$t_1 < t_2$$

Example

$$
\begin{aligned}
L_0 = f(x) &= \sin \omega x \\
\\
\Rightarrow L(x;t) &= e^{-\omega^2 t/2} \sin \omega x \\
L_{max}(t) &= e^{-\omega^2 t/2} \\
L_{x^m,max}(t) &= \omega^m e^{-\omega^2 t/2}
\end{aligned}
$$

Normalized Coordinates

$$\xi = \frac{x}{\sqrt{t}}$$
$$\partial_\xi = \sqrt{t}\partial_x$$

or,

$$L_{\xi^m, max}(t) = t^{m/2}\omega^m e^{-\omega^2 t/2}$$

goes up and then down !

$$t_{max} = \frac{m}{\omega^2}$$

The continuing saga of the maxima

$$L_{\xi^m,max}(t_{max}) = \left(\frac{m}{e}\right)^{m/2}$$

- The maximum value of the m^{th} order *normalized* derivative is independent of scale and frequency.

$$t_{max}\omega_{max}^2 = m$$

A tale of two maxima

- Image structures are characterized by sudden changes, hence they can be formalized as extrema a function of the image (differential forms).
- In the absence of other evidence, assume that the scale at which some function of normalized derivatives assumes a local maxima (over scales) is a reflection of the size of the corresponding structure in the data.

Scaling property of maxima

If the input image image is scaled by a constant scaling factor s, then the scale at which the maxima is assumed will be multiplied by the same factor.
i.e.

$$f'(x) = f(sx)$$

$$\Rightarrow \sqrt{t'_{max}} = s\sqrt{t_{max}}$$

can we do scale selection now ?

1. Characterize a feature by a differential form ($\mathcal{D}L$).
2. Find maxima in scale space, to estimate size and approximate location.

$$\begin{aligned} (\nabla(\mathcal{D}L))(x_0, t_0) &= 0 \\ (\partial_t(\mathcal{D}L))(x_0, t_0) &= 0 \end{aligned}$$

3. Use size information to perform a refine the location of the feature.

sunflowers

Blob Analysis

Assume a circular blob (same scale on both axes) and model is using a Gaussian.

$$f(x_1, x_2) = g(x_1, x_2, t_0) = \frac{1}{2\pi t_0} e^{-(x_1^2 + x_2^2)/2t_0}$$

by the semi-group property

$$L(x_1, x_2; t) = g(x_1, x_2, t_0 + t)$$

The chosen differential form is

$$\mathcal{D}L = |\,\mathrm{trace}(\mathcal{H}L)| = |L_{x_1x_1} + L_{x_2x_2}|$$

$$\mathcal{D}_{norm}L = t|L_{x_1x_1} + L_{x_2x_2}|$$

Blob Analysis

$$L_{x_1x_1} = \left[\frac{x_1}{2\pi(t_0+t)^3} - \frac{1}{2\pi(t_0+t)^2}\right] e^{-(x_1^2+x_2^2)/2(t_0+t)}$$

which can be used to show that the spatial minima for $\mathcal{D}L$ occurs at

$$(x_1, x_2) = (0, 0)$$

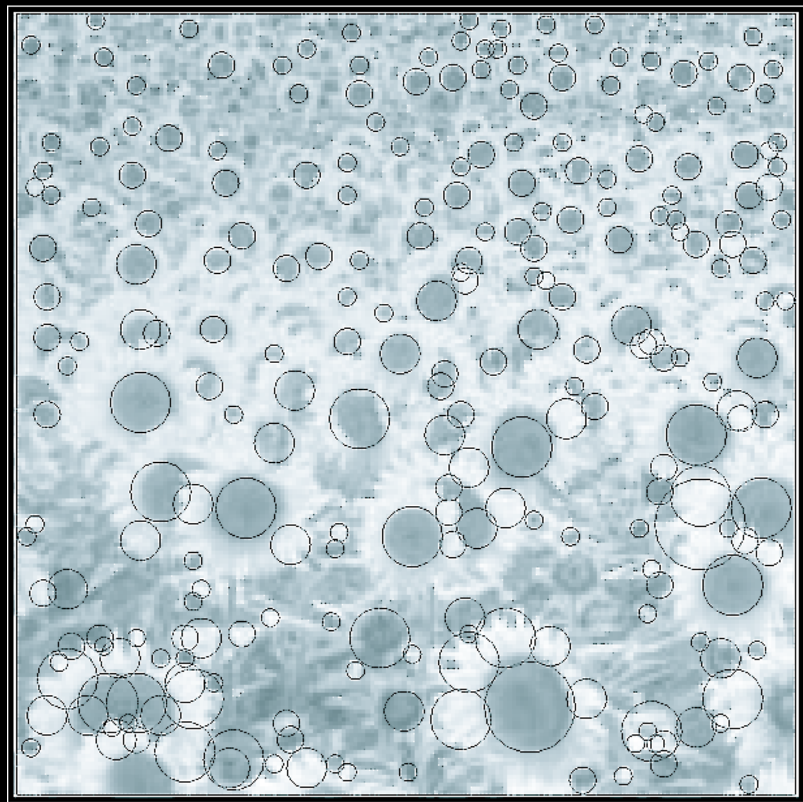
$$\Rightarrow \mathcal{D}_{norm}L(0,0;t) = \frac{t}{\pi(t_0+t)^2}$$

which in turn has a maxima at $t = t_0$

Blob Analysis

- The example verifies that for round blobs the maximum value is attained at a scale proportional to the ”width” of the blob.
- In the case of elliptical blobs, the scale is given by $t = \sqrt{t_1 t_2}$.

sunflowers



Blob scale detection using $\mathrm{trace}(\mathcal{H}(L)) = \nabla_\xi^2 L$.

Corner Detection

- Find a function of the scale space image which indicates the presence of a corner.
- Use it to identify the scale and the approximate location.
- Refine the position using Förstner corner operator.

The story of κ

Corner are charcterized by

1. High curvature in the grey level landscape

$$\kappa = \frac{L_y^2 L_{xx} - 2 L_x L_y L_{xy} + L_x^2 L_{yy}}{(L_x^2 + L_y^2)^{3/2}}$$

2. High gradient

$$\gamma = (L_x^2 + L_y^2)^{1/2}$$

$$\begin{aligned} \overline{\kappa} &= \kappa \gamma^3 \\ &= L_y^2 L_{xx} - 2 L_x L_y L_{xy} + L_x^2 L_{yy} \end{aligned}$$

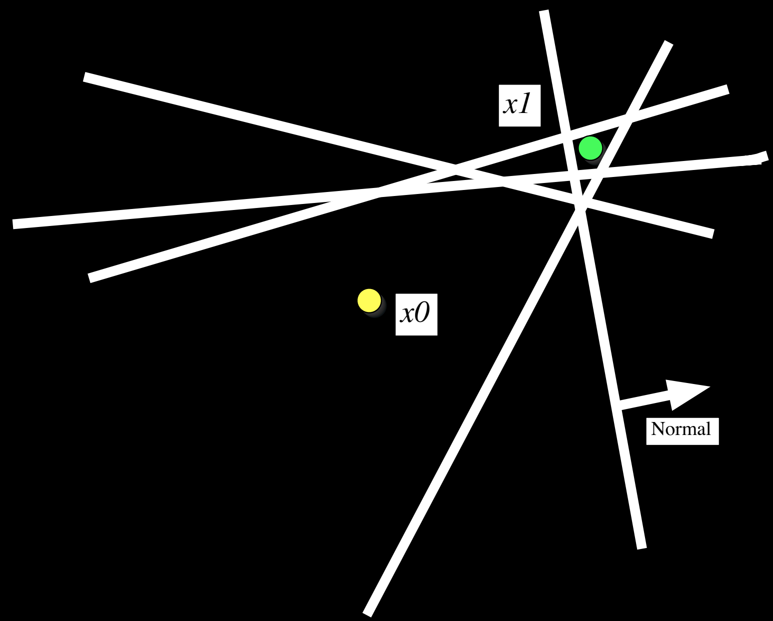
Corner Detection

- The value of $\overline{\kappa}$ in normalized coordinates is given by

$$\overline{\kappa}_{norm} = t^2 \overline{\kappa}$$

- Corner points and their corresponding scale is identified using the scale space maxima for $\overline{\kappa}$

Localizing the corner



$$D_{x'}(x) = ((\nabla L)(x'))^T(x - x')$$

$$d_{min} = \min_x \int_{x'} D_{x'}^T D_{x'} w_{x_0}(x') dx'$$

Localizing the corner

$$\min_x X^T A x - 2x^T b + c$$

$$A = \int_{x'} (\nabla L)(x')(\nabla L)^T(x') w(x') dx' \quad (1)$$

$$b = \int_{x'} (\nabla L)(x')(\nabla L)^T(x') x' w(x') dx' \quad (2)$$

$$c = \int_{x'} x'^T (\nabla L)(x')(\nabla L)^T(x') x' w(x') dx' \quad (3)$$

and the minima is

$$x = A^{-1} b$$

Where is the scale?

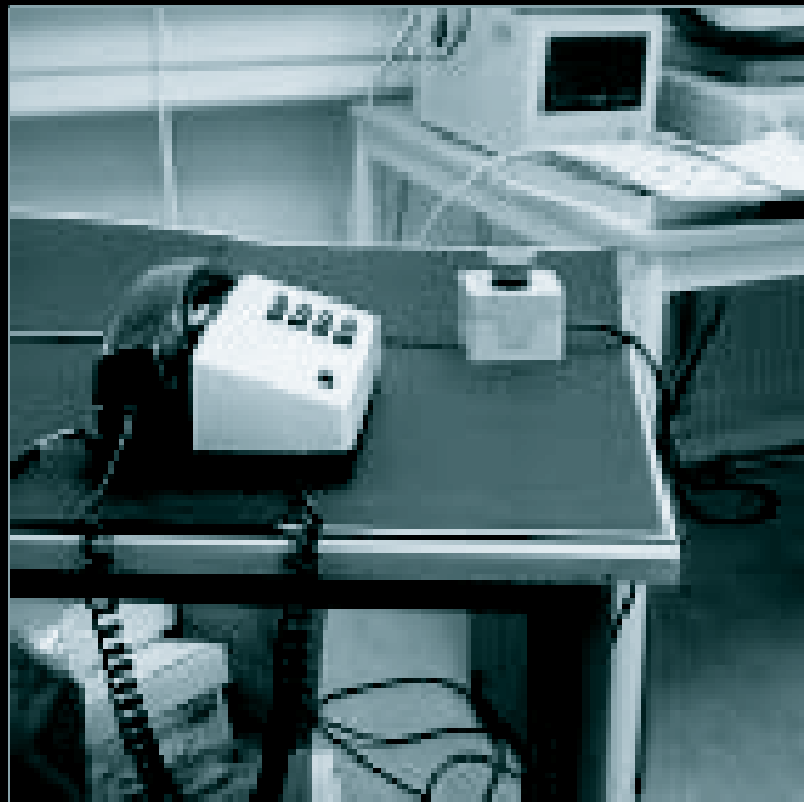
- Shape and size of the window function $w(x')$: Use the scale estimate as the width of the Gaussian for the weight window.
- Localization Scale: What scale should the gradient ∇L be calculated ? At $x = A^{-1}b$

$$d_{min} = c - b^T A^{-1} b$$

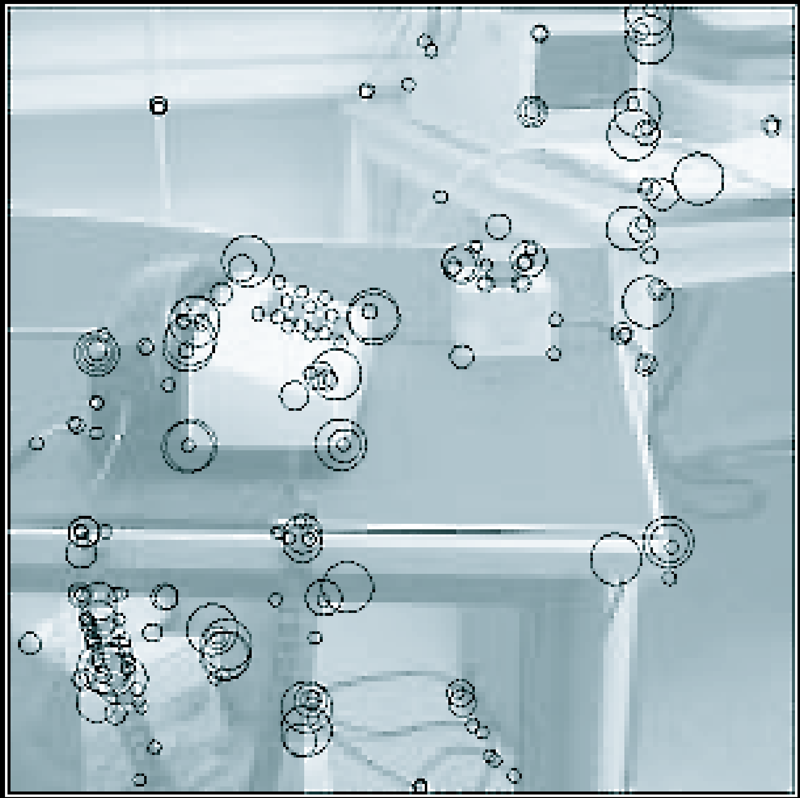
The best location for the corner is the one which minimizes d_{min} w.r.t a scale parameter again. To reduce sensitivity to local image contrast, we instead use

$$\overline{d}_{min} = \frac{c - b^T A^{-1} b}{\mathrm{trace}(A)}$$

Corner Detection : Example



Corner Detection : Example



Summary

- Scale Space provides us with a formalism in which we can construct multiscale representation of signals cleanly and efficiently.
- Given a certain class of feature descriptors, it allows us to detect the scale at which certain features express themselves.

References

- Lindeberg, Tony Scale Space Theory in Computer Vision, 1994, Kluwer Academic Publishers.
- Lingberg, Tony and M. Ter Haar Romeny, Bart Linear Scale Space, *Geometry Driven Diffusion in Computer Vision*, Kluwer Academic Publishers, 1994.

Acknowledgments

The music of Chemical Brothers.