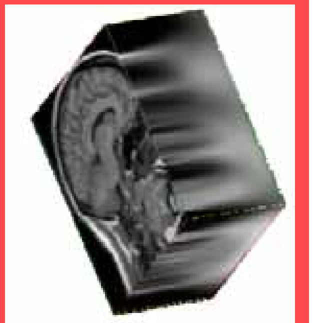
Lecture 1: Scale Space / Differential Invariants

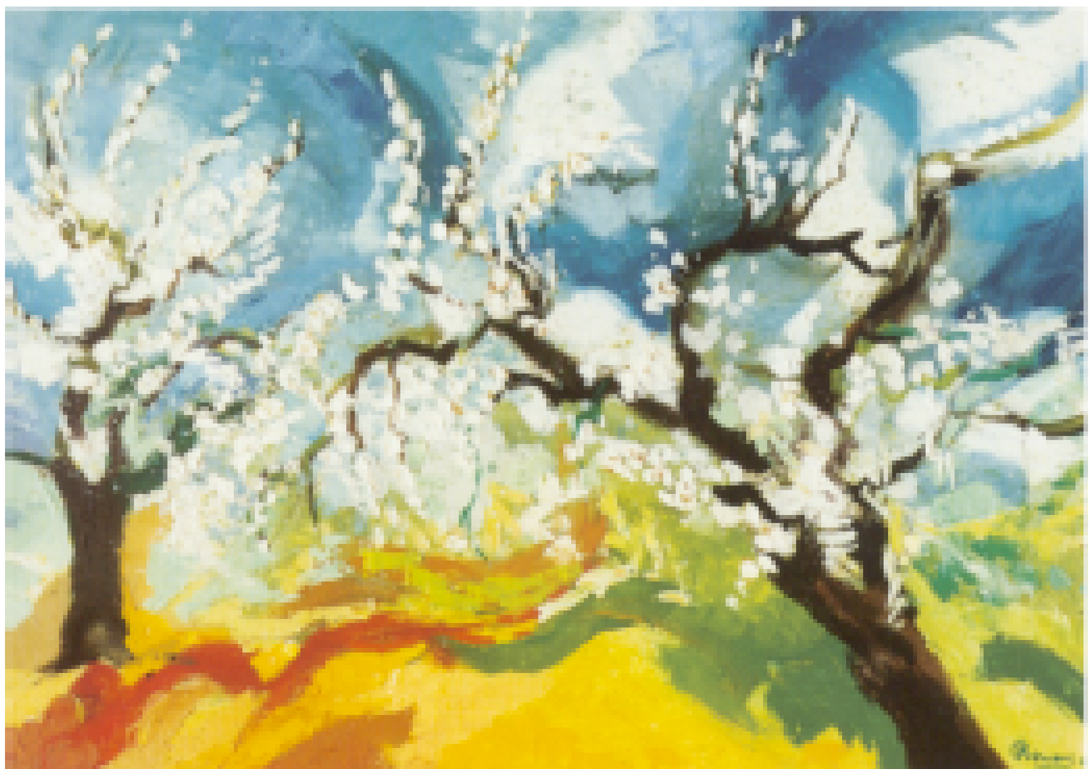
Guido Gerig

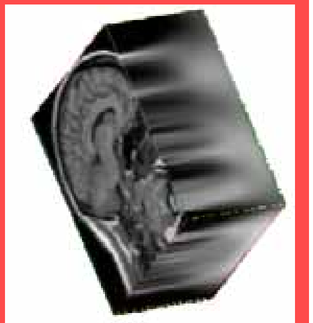
CS 7960, Spring 2010

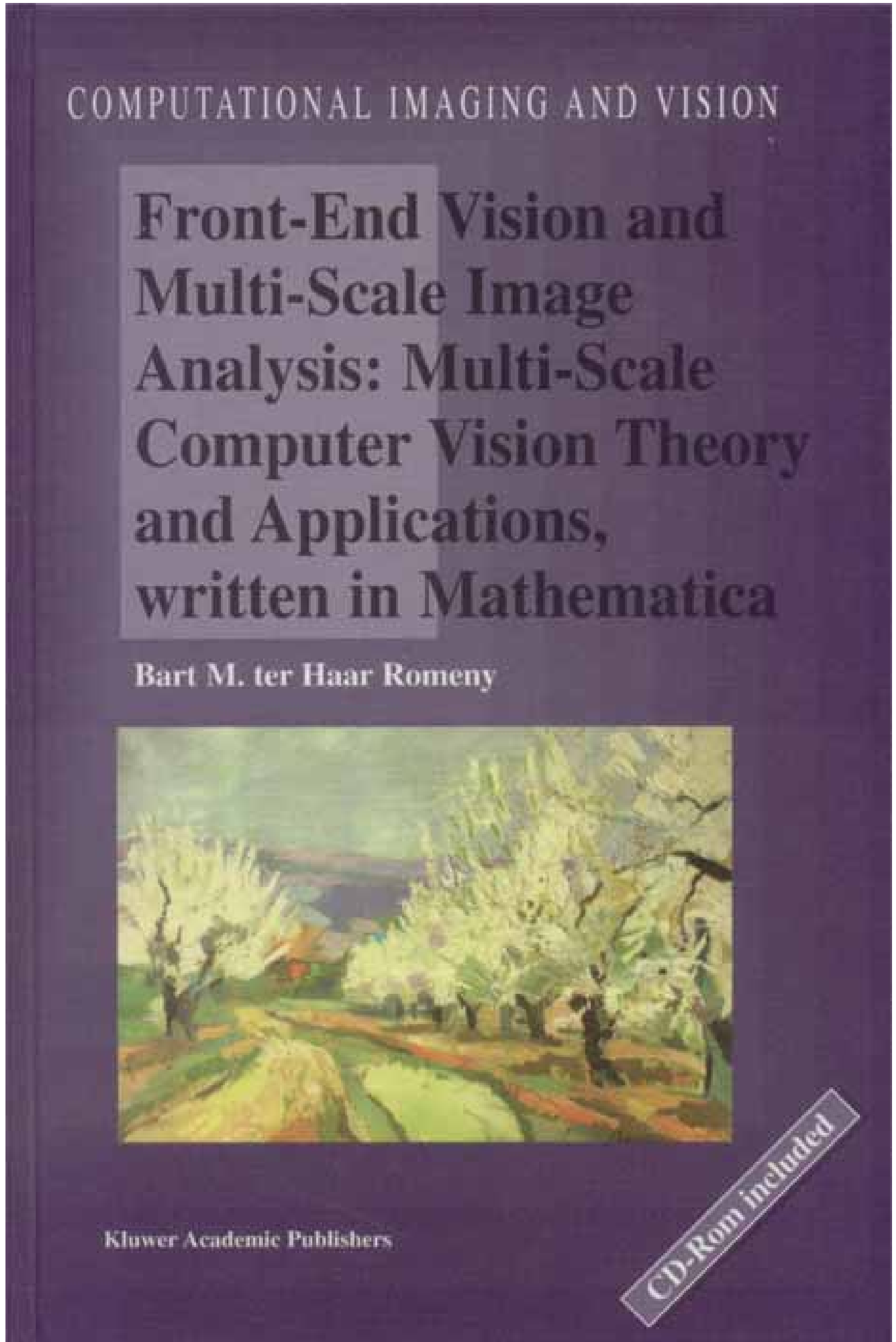


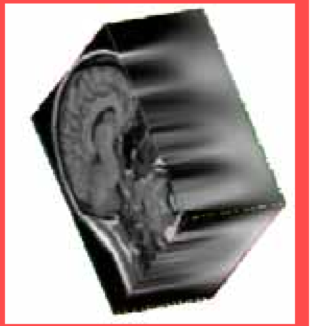
Front-End Vision and Multi-Scale Image Analysis

Bart M. ter Haar Romeny






COMPUTATIONAL IMAGING AND VISION
Front-End Vision and Multi-Scale Image Analysis: Multi-Scale Computer Vision Theory and Applications, written in Mathematica
Bart M. ter Haar Romeny
Kluwer Academic Publishers
CD-Rom included



Aperture and the notion of scale

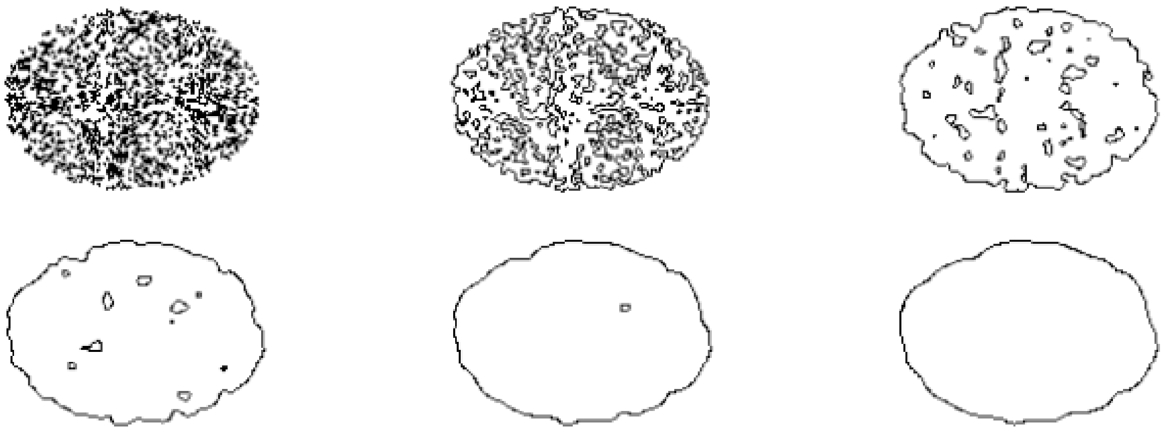
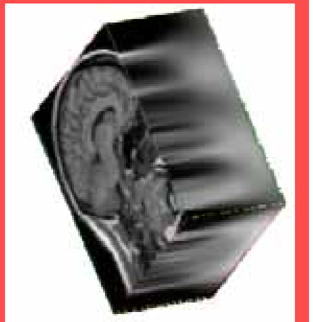


Figure 1.1 A cloud observed at different scales, simulated by the blurring of a random set of points, the 'drops'. Adapted from [Koenderink1992a].

- Resulting measurement strongly depends on the size of the measurement aperture
- Need to develop criteria: Aperture size to apply



1.2 Mathematics, physics and vision

Mathematics

- Objects can have no size
- Points, lines with zero width
-

$$f(x): \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

Physics

- Measurement with instrument at certain scale
- Smallest scale: Inner scale (smallest sampling element)
- Choice of sample size depends on task (tree vs. leafs)
- Any physical observation is done through aperture

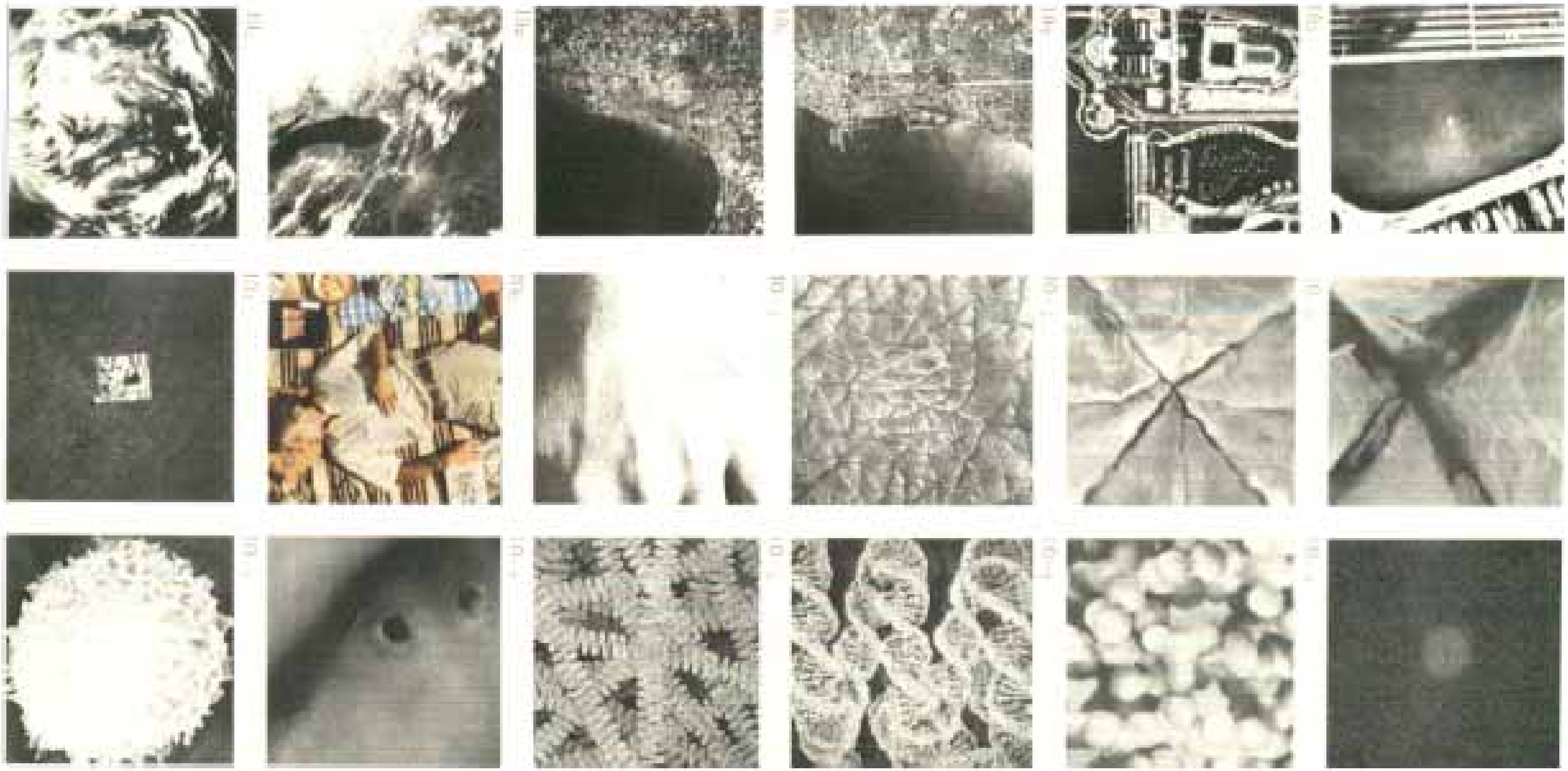


Figure 1.3 Selection of pictures from the journey through scale from the book [Morrison1985], where each page zooms in a factor of ten. Starting at a cosmic scale, with clusters of galaxies, we zoom in to the solar system, the earth (see the selection above), to a picknicking couple in a park in Chicago. Here we reach the 'human' (antropometric) scales which are so familiar to us. We then travel further into cellular and molecular structures in the hand, ending up in the quark structure of the nuclear particles. For the movie see: http://www.micro.magnet.fsu.edu/primer/java/scienceopticsu/powersof10/index.html.

2006

We blur by looking

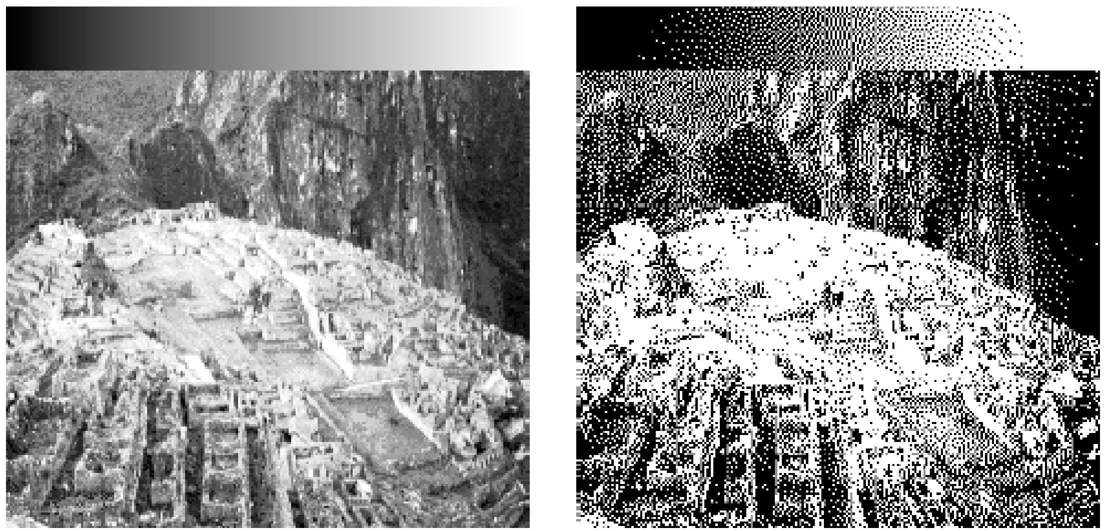
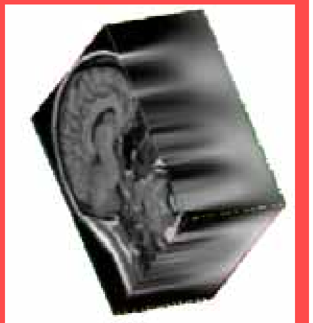
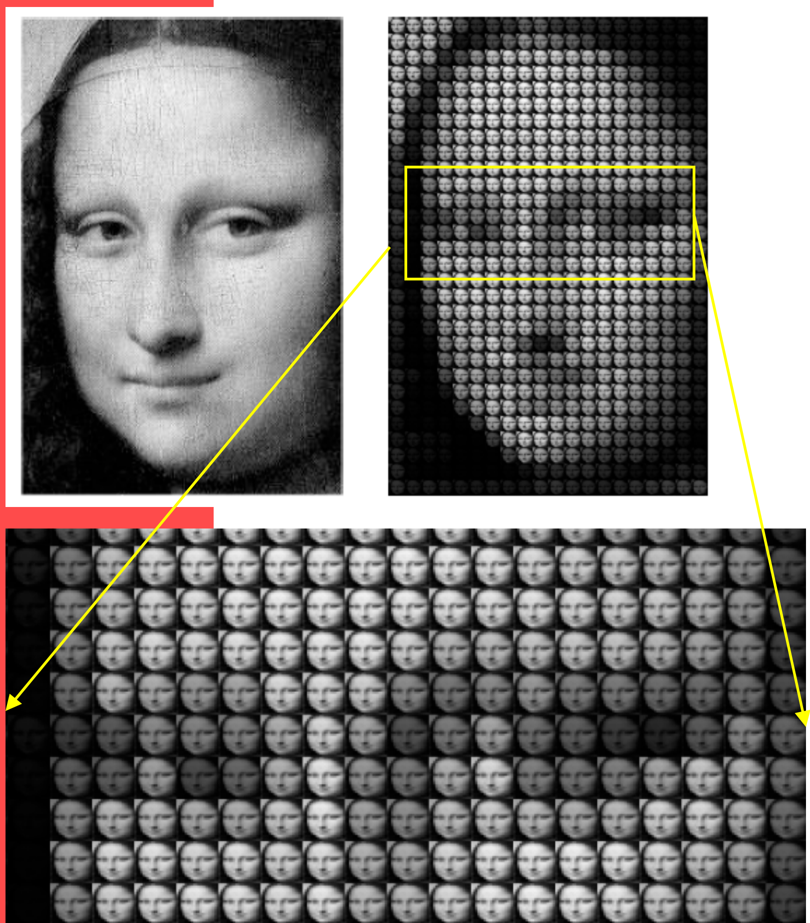


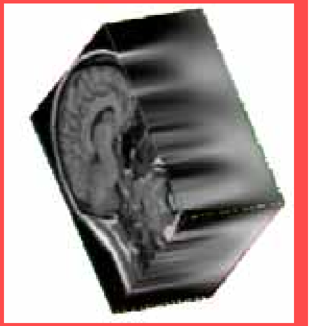
Figure 1.4 Dithering is the representation of grayvalues through sparse printing of black dots on paper. In this way a tonal image can be produced with a laserprinter, which is only able to print miniscule identical single small high contrast dots. Left: the image as we observe it, with grayscales and no dithering. Right: Floyd-Steinberg dithering with random dot placements. [From http://sevilleta.unm.edu/~bmilne/khoros/html-dip/c3/s7/front-page.html].



Visual Front End: Multitude of aperture sizes *simultaneously*

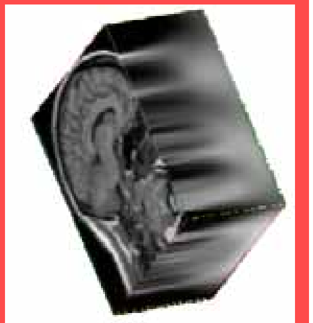


- Objects come in all sizes, all equally important at front-end
- Mosaic: Multi-resolution perceptual effect
- Multi-scale observation
- Aperture size: Continuous measurement dimension
- Scale: addl. parameter



Multi-scale

- Specific reasons to not only look at the highest resolution
- New possibilities if all sizes simultaneously, whole range of sharpness



Different information at difference resolutions

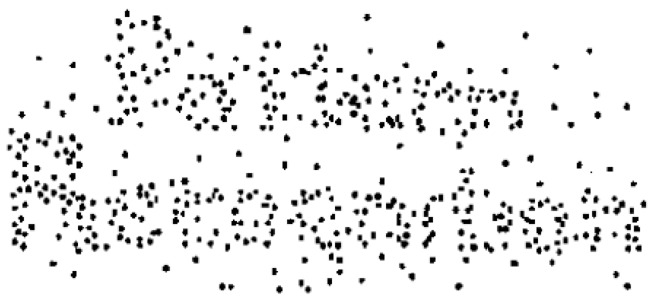
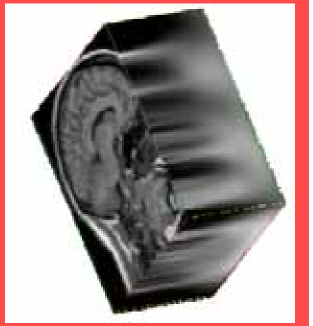


Figure 1.10 At different resolutions we see different information. The meaningful information ir this image is at a larger scale then the dots of which it is made. Look at the image from about 2 meters. Source: dr. Bob Duin, Pattern Recognition Group, Delft University, the Netherlands.



Multiple Scales

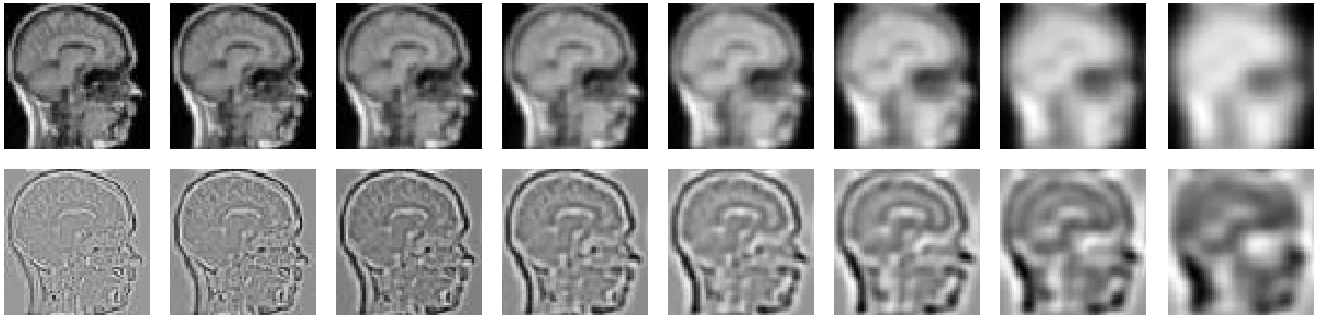
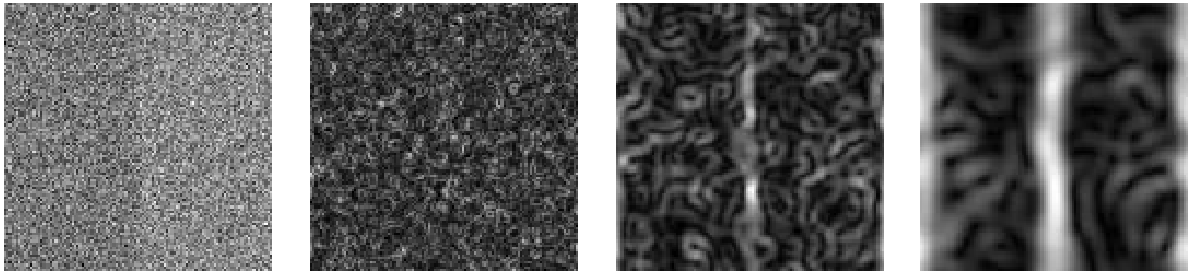
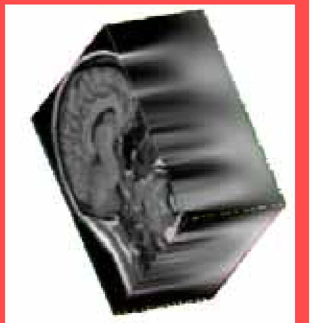


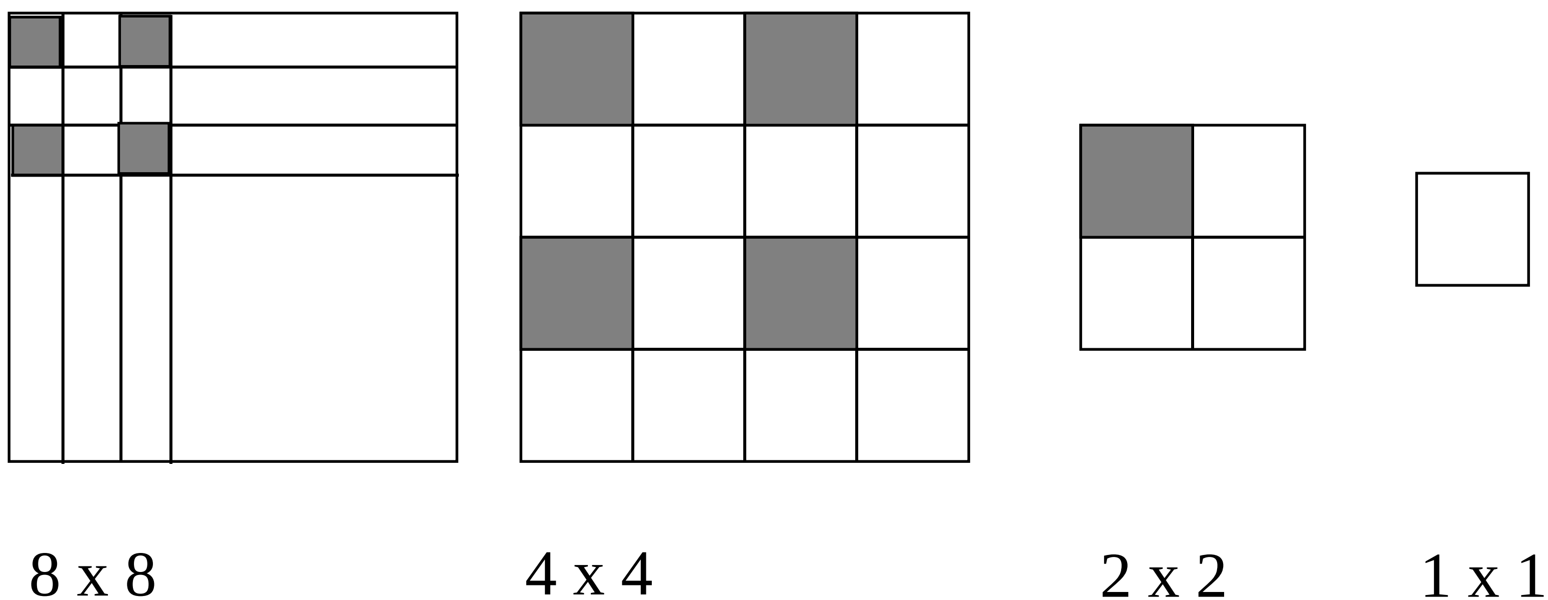
Figure 2.9 A scale-space is a stack of images at a range of scales. Top row: Gaussian blur scale-space of a sagittal Magnetic Resonance image, resolution 128^2, exponential scale range from $\sigma = e^0$ tot $\sigma = e^{2.1}$. Bottom row: Laplacian scale-space of the same image, same scale range.

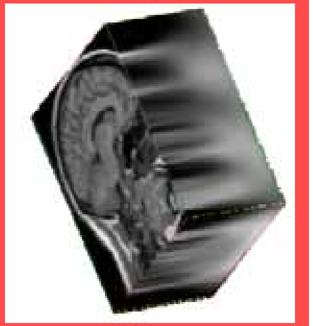




Pyramids (Hong, Shneier, Rosenfeld)

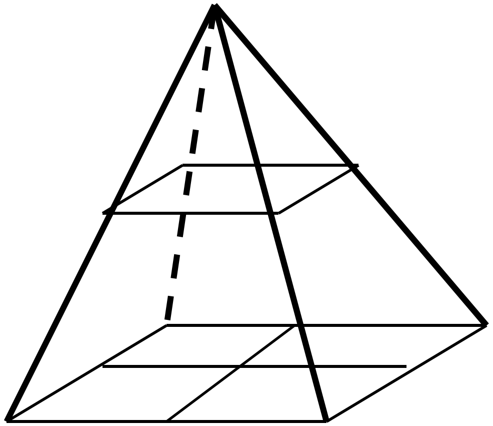
- recursive subsampling
- $f^{(k-1)} = \mathrm{REDUCE}[f^{(k)}]$
- $2^k \times 2^k$ images: $k+1$ scales

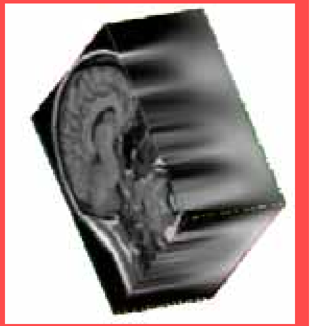




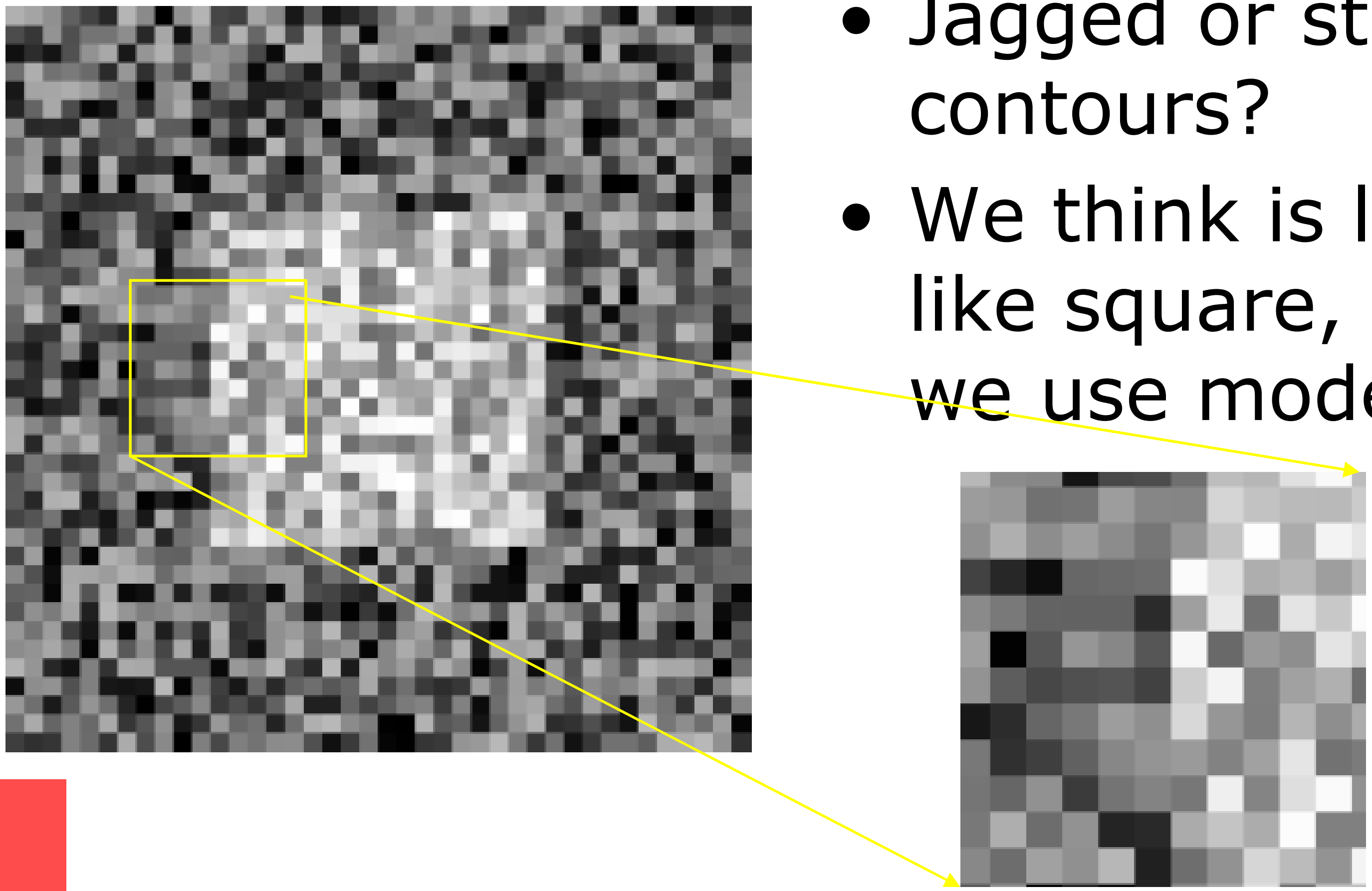
Pyramids ctd.

- to avoid aliasing: low-pass filtering before subsampling (blur – subsample – blur – subsample – blur – …)
- Advantages: rapidly decreasing image size
- Disadvantages: coarse quantization along scale direction

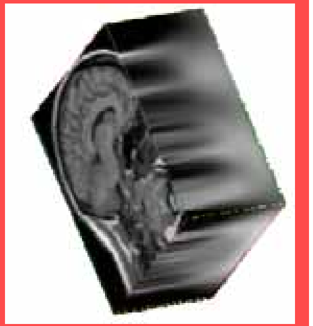




We assume model…

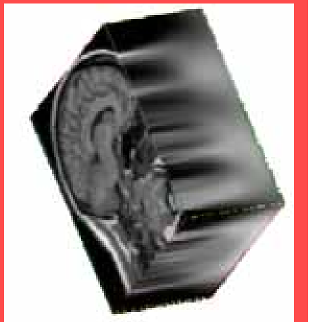


- Jagged or straight contours?
- We think is looks like square, but we use model!

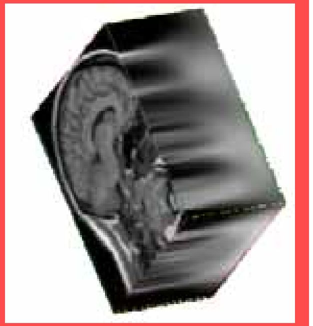


1.5 Summary

- Observations necessarily done through a **finite aperture.**
- Visual system: exploits a wide **range** of such observation apertures in the front-end **simultaneously**, in order to capture the information at all scales.
- Observed noise is part of the observation.
- Aperture can't take any form: Pixel squares are wrong and create '*spurious*' resolution (wrong edge information)→ choose appropriate <u>kernel</u>.



2.0 Foundations of Scale Space



Aperture function: Operator

- Unconstrained front-end: Unique solution to aperture function is Gaussian kernel
- Many derivations, all leading to Gaussian kernel

σ = 1; Plot[$\frac{1}{\sqrt{2\pi\sigma^2}}$ E$^{-\frac{x^2}{2\sigma^2}}$, {x, -4, 4}, ImageSize ->

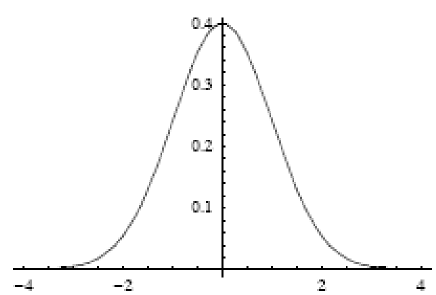


Figure 2.1 The Gaussian kernel with unit standard deviation in 1D.

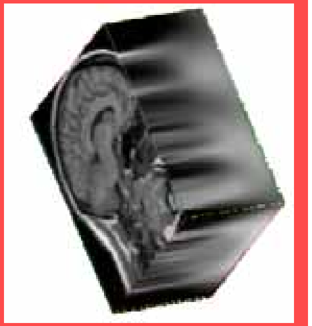
Axioms

- **Linearity** (nonlinearities at this stage)
- **Spatial shift invariance** (no preferred location)
- **Isotropy** (no preferred orientation)
- **Scale invariance** (no preferred size, or scale of the aperture)

Convolution

$$L(x, y) = L_0(x, y) \otimes G(x, y) \equiv \int_{-\infty}^{\infty} L_0(u, v)\, G(x-u, y-v)\, du\, dv$$

In the Fourier domain, a convolution of functions translates to a regular product between the Fourier transforms of the functions: $\mathcal{L}(\omega_x, \omega_y) = \mathcal{L}_0(\omega_x, \omega_y)\ .\ \mathcal{G}(\omega_x, \omega_y)$



Linear Diffusion

- Gaussian kernel is Green's function of linear, isotropic diffusion equation

$$g := \frac{1}{2\,\pi\,\sigma^2}\,\mathrm{Exp}\left[-\frac{x^2+y^2}{2\,\sigma^2}\right];$$

$$\frac{\partial^2 L}{\partial x^2}+\frac{\partial^2 L}{\partial y^2} = L_{xx}+L_{yy} = \frac{\partial L}{\partial t}$$

$$t = 2\,\sigma^2$$

Gaussian Derivatives

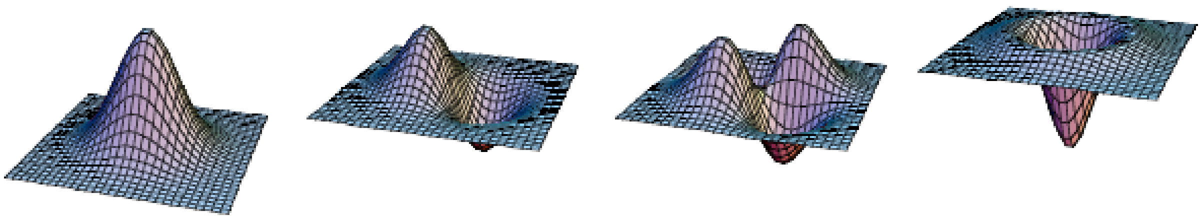
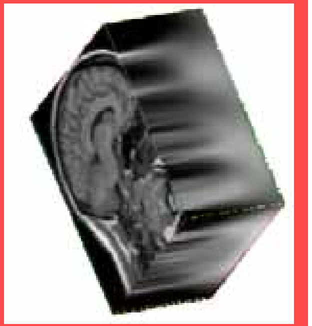


Figure 2.10 Upper left: the Gaussian kernel $G(x,y;\sigma)$ as the zeroth order point operator; uppe right: $\frac{\partial G}{\partial x}$; lower left: $\frac{\partial^2 G}{\partial x \, \partial y}$; lower right: the Laplacian $\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2}$ of the Gaussian kernel.

All partial derivatives of the Gaussian kernel are solutions too of the diffusion equation



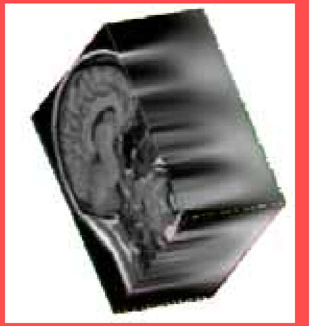
Linearity

Derivative of: $L_0(x,\ y) \otimes G(x,\ y;\ \sigma)$

Given by: $\frac{\partial}{\partial x} \{L_0(x,\ y) \otimes G(x,\ y;\ \sigma)\}$

Rewritten as: $L_0(x,\ y) \otimes \frac{\partial}{\partial x} G(x,\ y;\ \sigma)$

Differentiation and observation is done in one single step: Convolution with Gaussian derivative kernel.



Application to images

- We can apply differentiation to sampled image data
- Convolution with appropriate Gaussian derivative kernel
- Scale-space: Choice of multiple Gaussian widths σ
- What are appropriate derivatives?

Edges: Sudden change of intensity L

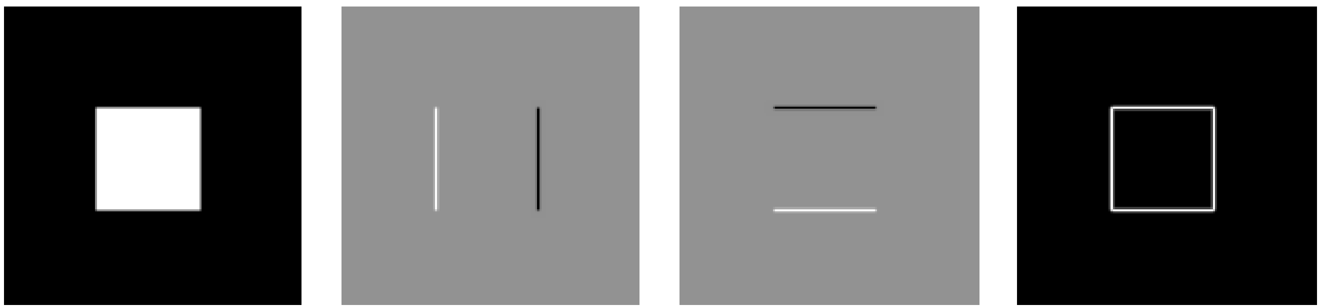
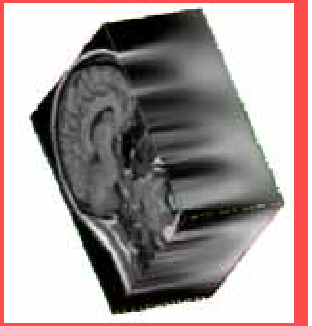
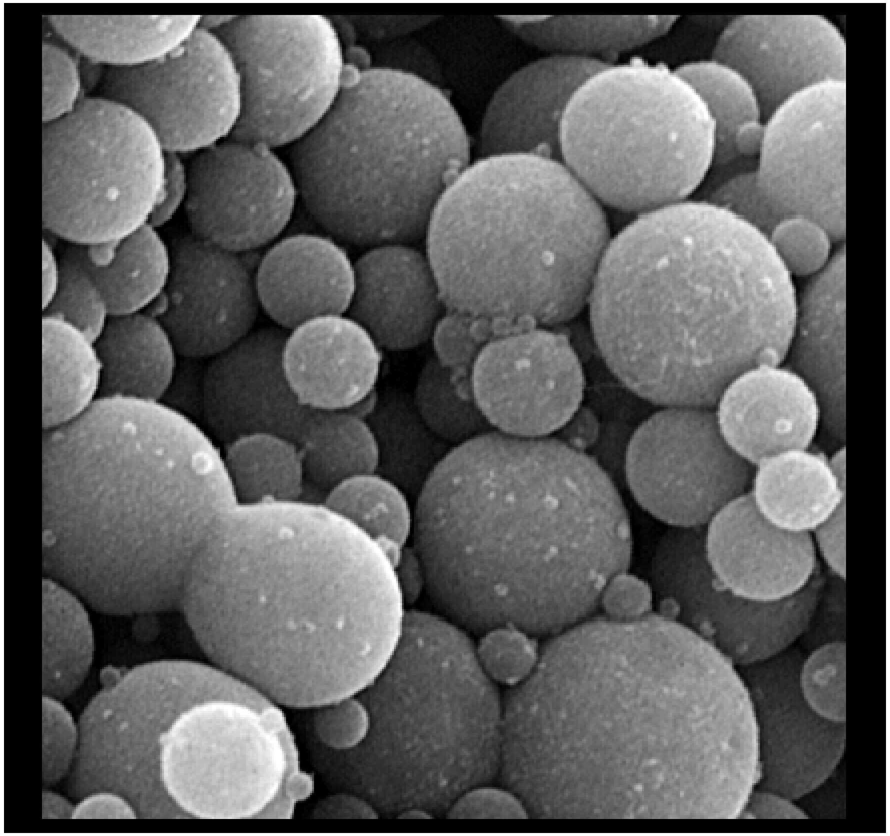


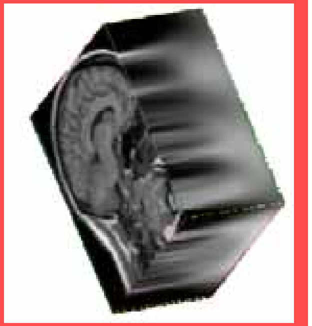
Figure 2.11 The first order derivative of an image gives edges. Left: original test image $L(x, y)$, resolution 256^2. Second: the derivative with respect to x: $\frac{\partial L}{\partial x}$ at scale $\sigma = 1$ pixel. Note the positive and negative edges. Third: the derivative with respect to y: $\frac{\partial L}{\partial y}$ at scale

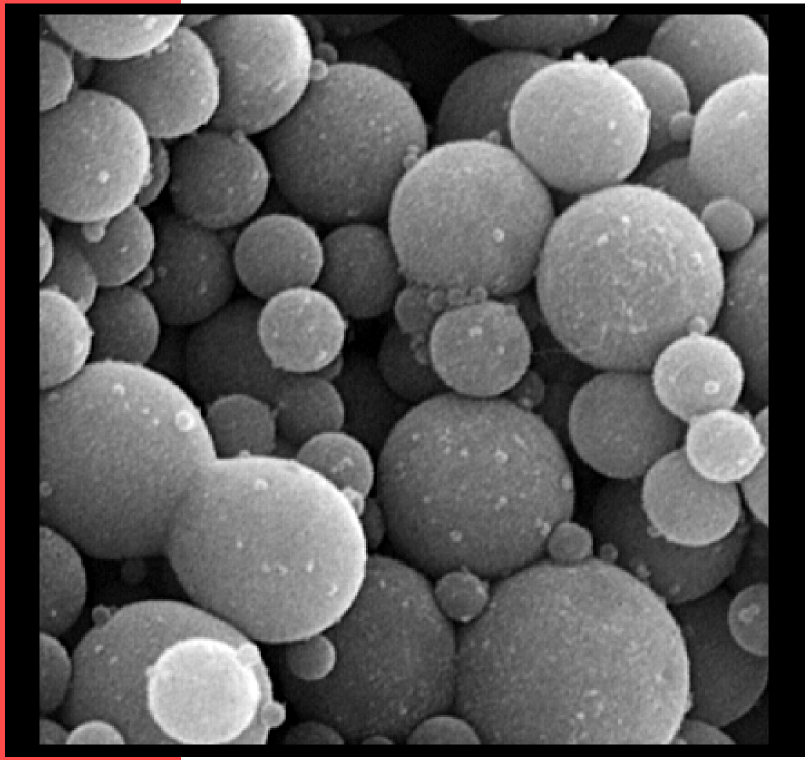


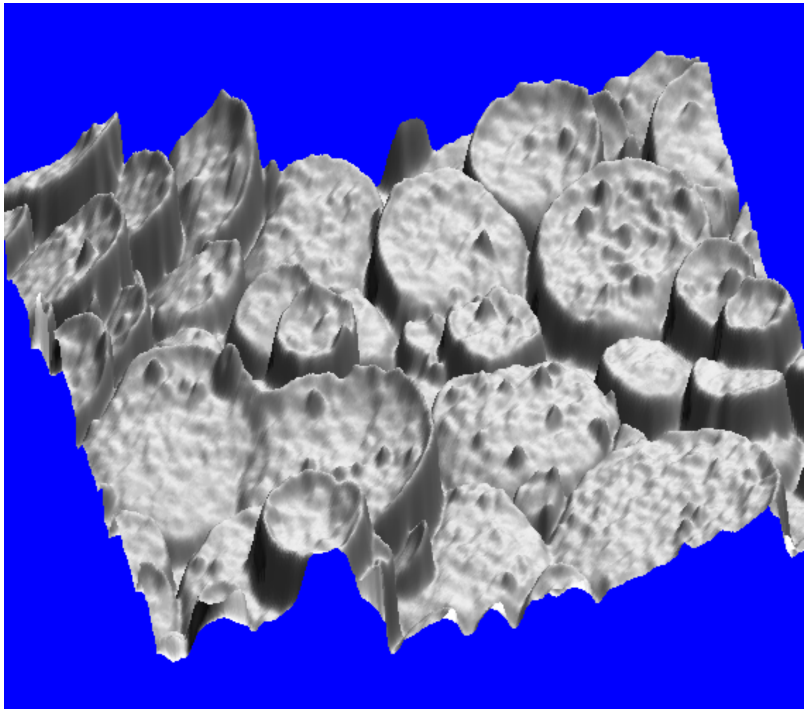
Digital Images

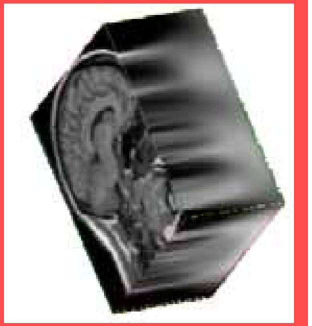


Digital Images

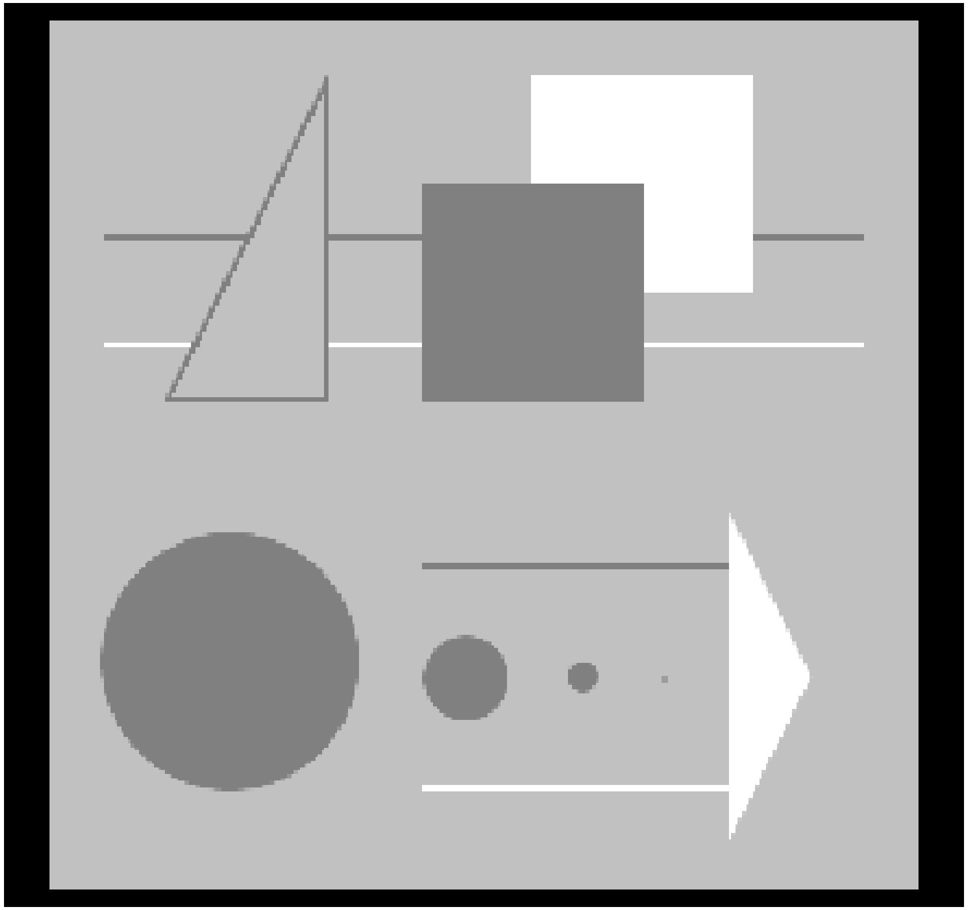


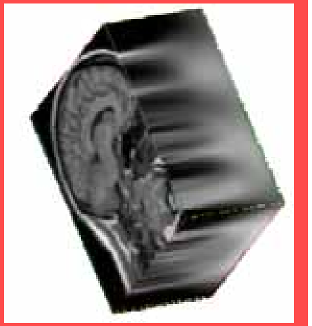




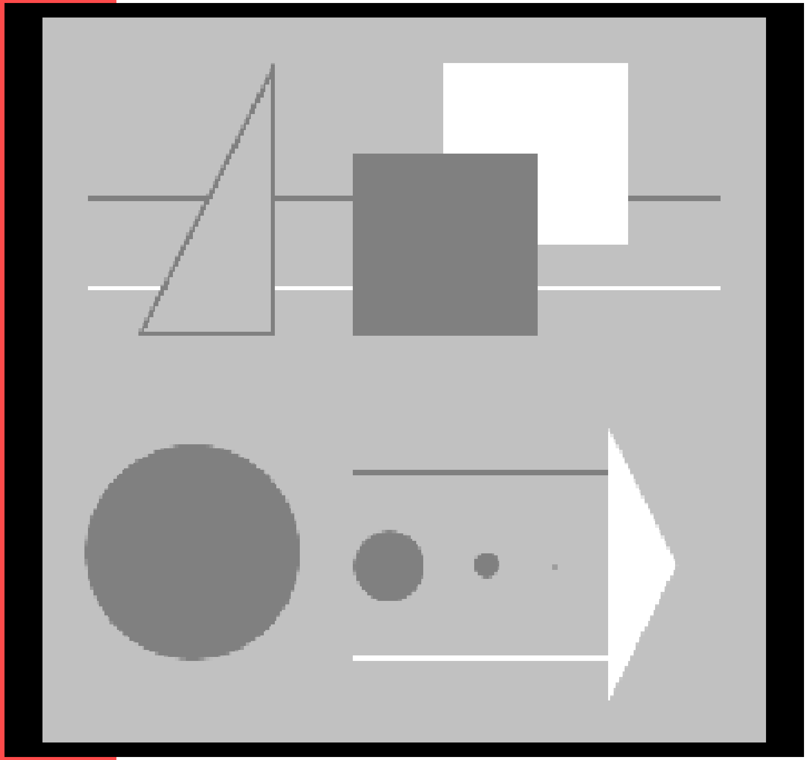


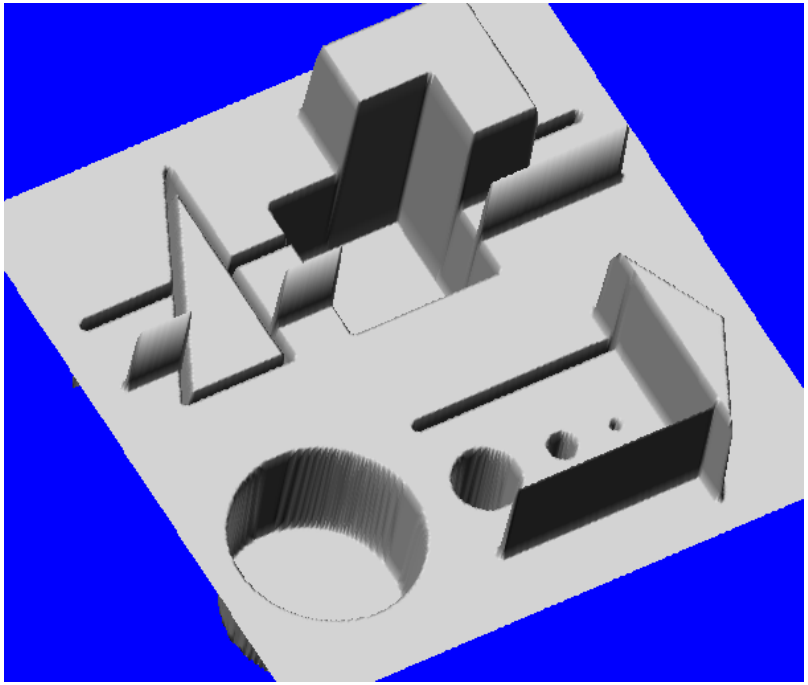
Digital Images



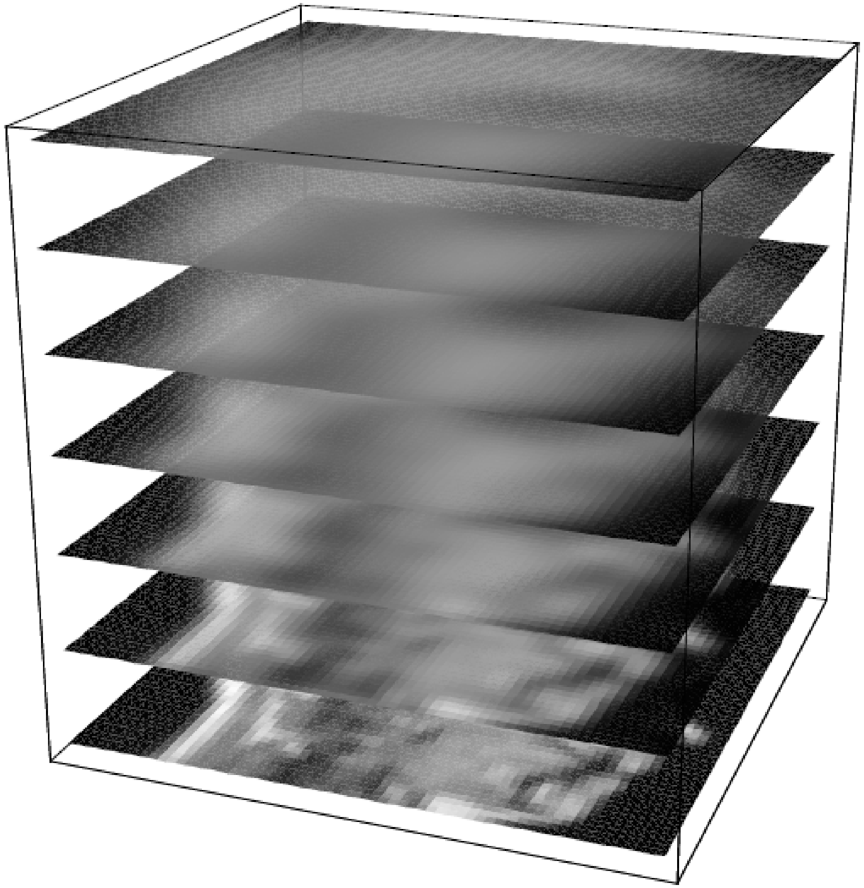


Digital Images





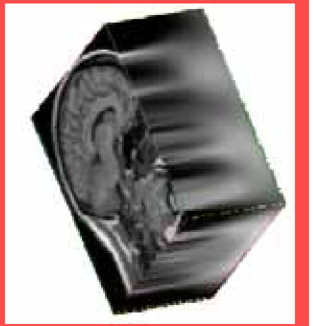
Scale-space stack



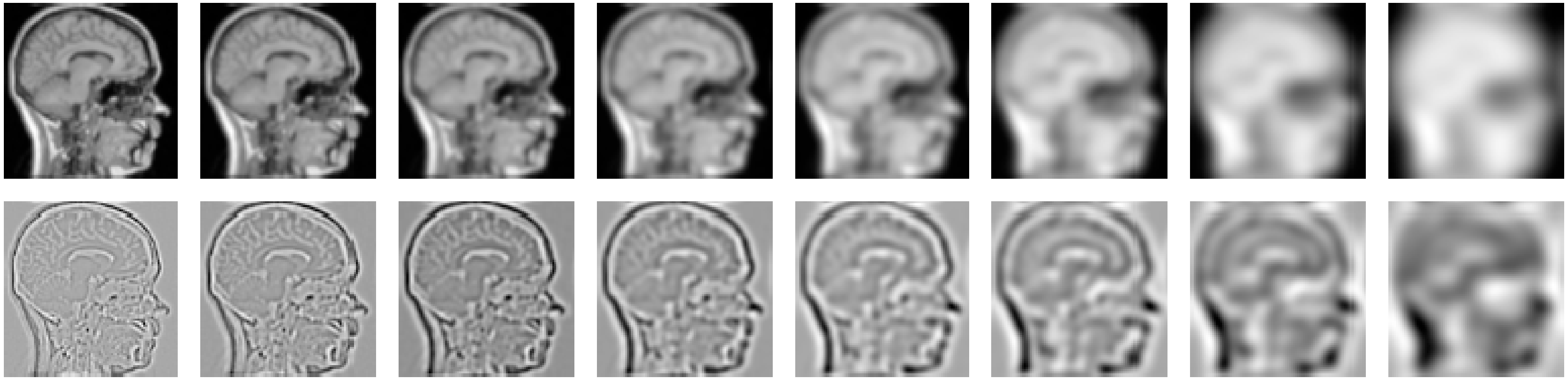
Un-committed front end: Take all scales

Family of kernels applied to image

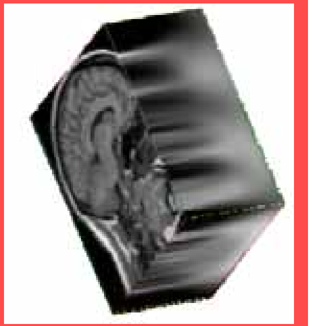
Simulates visual system



Scale-space



Scale is parameterized in an exponential fashion (see 2.8 sampling of scale axis)



2.9 Summary

- Unique solution for uncommitted front-end: Gaussian kernel
- Differentiation of discrete data: Convolution with derivative of observation kernel: Integration
- Differentiation of discrete data now possible: Convolution with finite kernel
- Differentiation can NEVER be done without blurring (see later Ch. 14)
- Family of kernels, scale parametrized in an exponential fashion