11.10.2018

Digital Image Processing (CSE/ECE 478)

Lecture-18: Multi-scale Image Processing

Ravi Kiran

Rajvi Shah

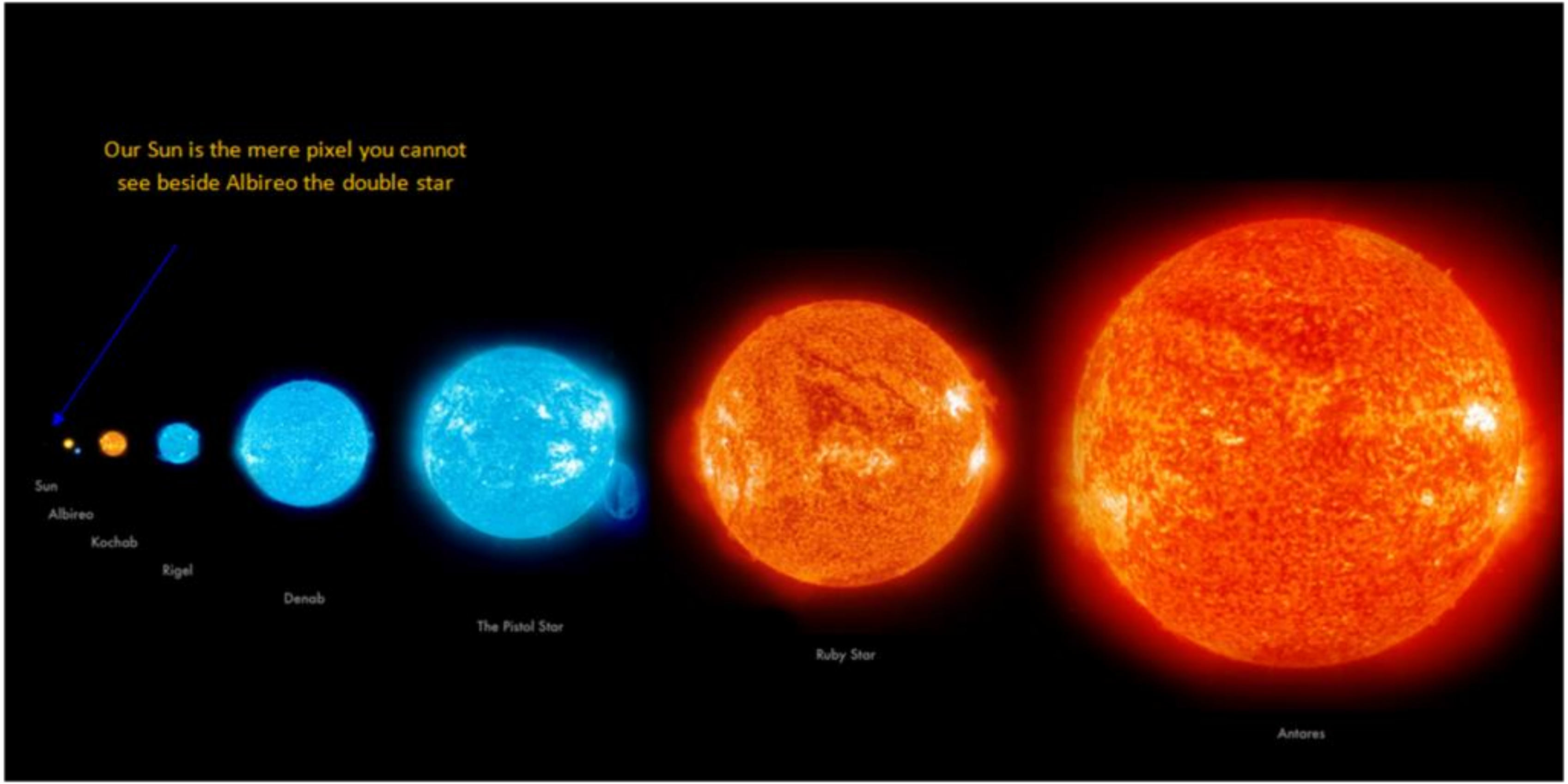
Center for Visual Information Technology (CVIT), IIIT Hyderabad

Many slides borrowed from Vineet Gandhi @CVIT!

Announcements

- Mid-2

It’s all about scale !



Real-world objects exist as meaningful entities over certain ranges of scale.

kilometer	km	1,000	$1X10^{3}$
meter	m	1	$1X10^{0}$
millimeter	mm	1/1,000	$1X10^{-3}$
micrometer	μm	1/1,000,000	$1X10^{-6}$
nanometer	nm	1/1,000,000,000	$1X10^{-9}$
picometer	pm	1/1,000,000,000,000	$1X10^{-12}$

WI is 420 km wide

10-year-old human ~ 1.4 m

Hair: ~40μm

DNA: 1-2 nm

Objects in the world appear in different ways depending on the scale of observation

Has important implications if one aims at describing objects

kilometer	km	1,000	$1X10^{3}$
meter	m	1	$1X10^{0}$
millimeter	mm	1/1,000	$1X10^{-3}$
micrometer	µm	1/1,000,000	$1X10^{-6}$
nanometer	nm	1/1,000,000,000	$1X10^{-9}$
picometer	pm	1/1,000,000,000,000	$1X10^{-12}$

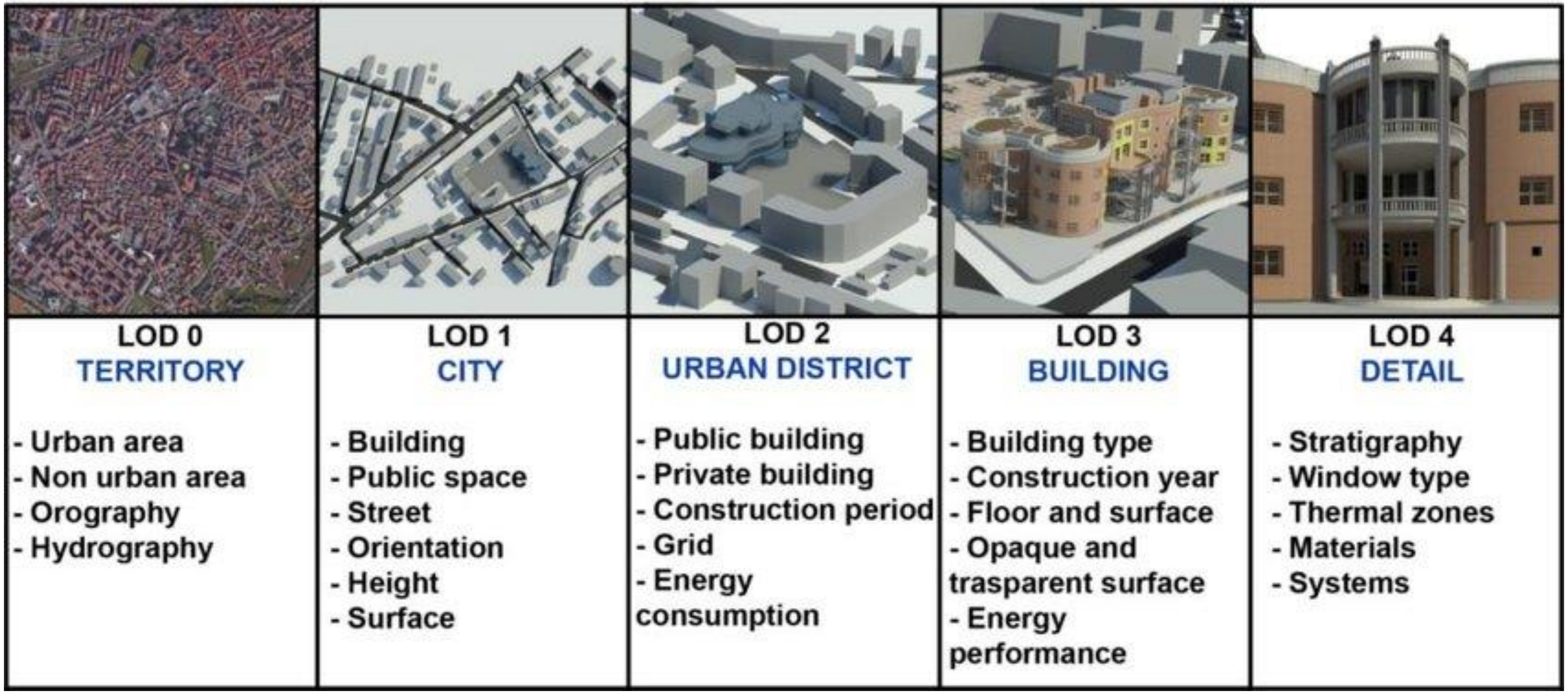

WI is 420 km wide

10-year-old human ~ 1.4 m


Hair: ~40µm


DNA: 1-2 nm

Form of description strongly dependent upon scale at which the world is considered



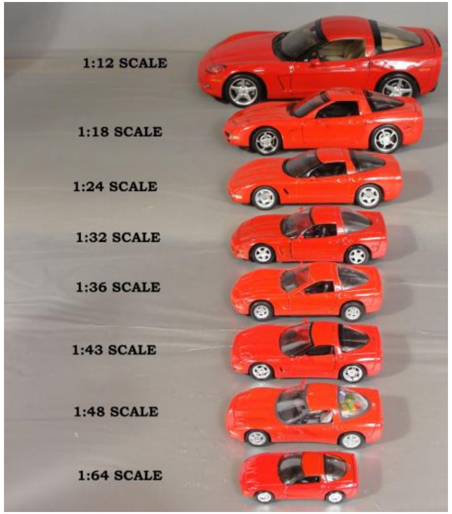
LOD 0 TERRITORY	LOD 1 CITY	LOD 2 URBAN DISTRICT	LOD 3 BUILDING	LOD 4 DETAIL
- Urban area - Non urban area - Orography - Hydrography	- Building - Public space - Street - Orientation - Height - Surface	- Public building - Private building - Construction period - Grid - Energy consumption	- Building type - Construction year - Floor and surface - Opaque and trasparent surface - Energy performance	- Stratigraphy - Window type - Thermal zones - Materials - Systems

- C.f. : idealized mathematical concepts, such as 'point' and 'line‘
    - independent of the scale of observation

- Information extracted from image depends upon
  - Size of actual structures in data
  - Size of information extractors/operators
    - 3 x 3, 5 x 5

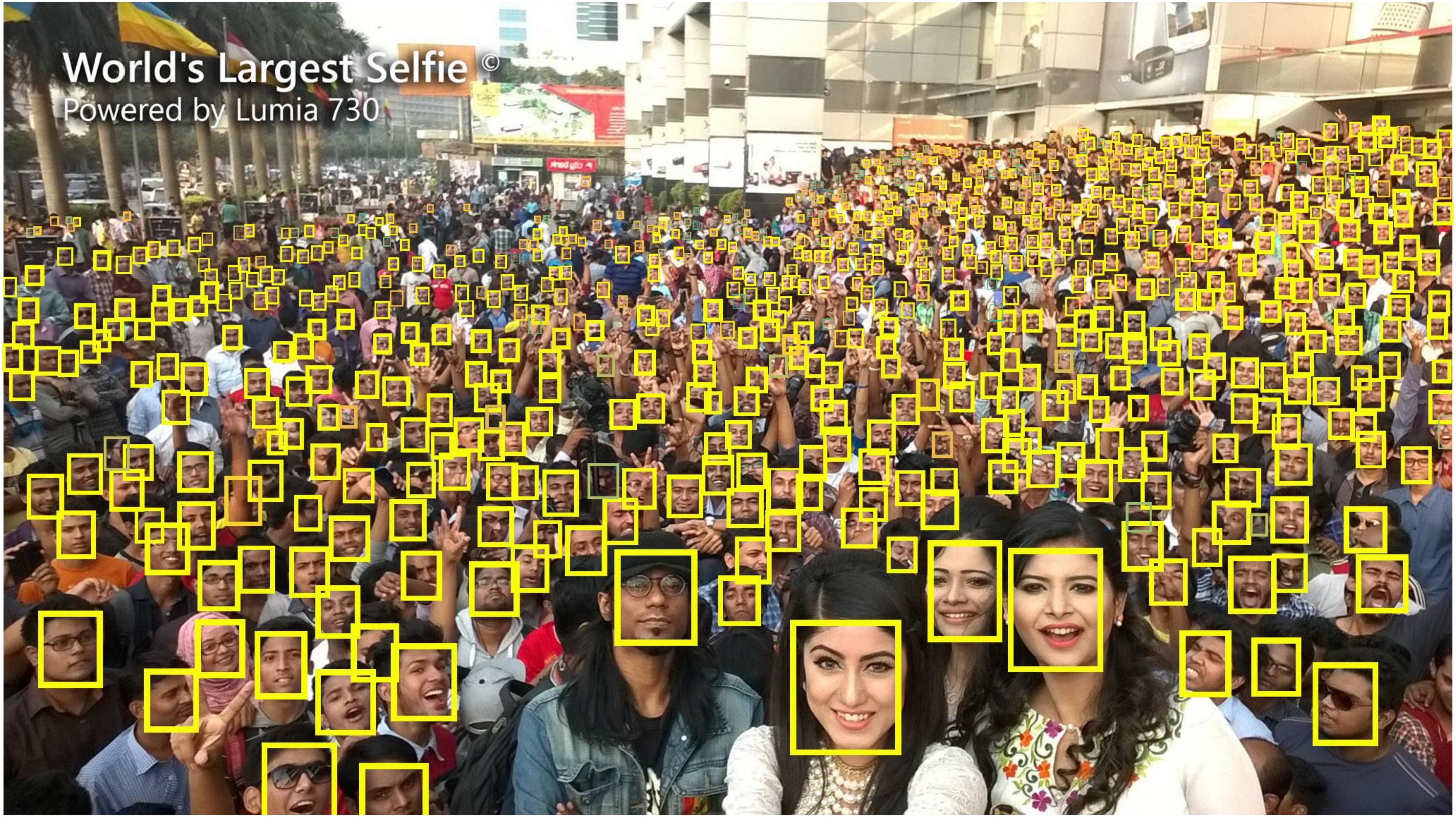
- Until now …
  - focus of approach on <u>small-scale</u> image structures
  - Smallest scale in digital images = Pixel scale
  - E.g. Spatial operators use discrete kernels with size very close to this scale, such as 3×3 kernels

Objects in real world occur at multiple scales



Many images contain information at different scales or levels of detail (e.g., people vs buildings)

Canaletto's San Marco square



https://www.cs.cmu.edu/~peiyunh/tiny/

Many images contain information at different scales or levels of detail (e.g., people vs buildings)

- Until now …
  - focus of approach on <u>small-scale</u> image structures
  - Smallest scale in digital images = Pixel scale
  - E.g. Spatial operators use discrete kernels with size very close to this scale, such as 3 × 3 kernels
  - **How to control the scale with which operators ’look’ at a digital image**
  - **How to get “best of all scales” ?**

Multi-scale Analysis

- Analyzing information at the same scale will not be effective.

Use windows of
different size
(i.e., vary scale)

Multi-(scale)/(resolution) Analysis

- Alternative: Same window size, but analyze at different resolutions

High resolution

Small size objects should be examined at a high resolution

Low resolution

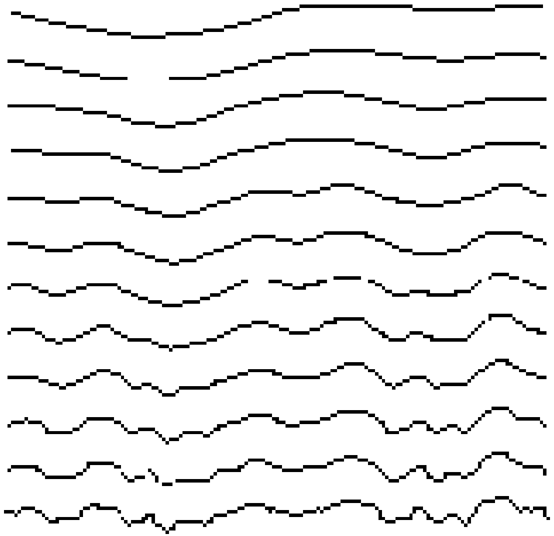
Large size objects should be examined at a low resolution

Multi-scale Analysis

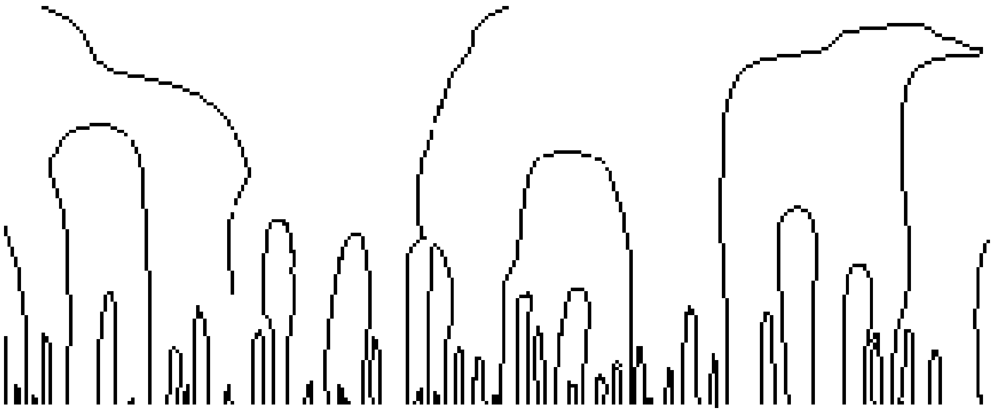
- Two techniques for representing multi-scale information efficiently
  - Pyramidal coding
  - Sub-band coding

Basic idea behind multi-scale analysis

- If no prior information is available about what are the appropriate scales for a given data set, then the only reasonable approach for an uncommitted vision system is to represent the input data at multiple scales.
  – *Tony Lindeberg*

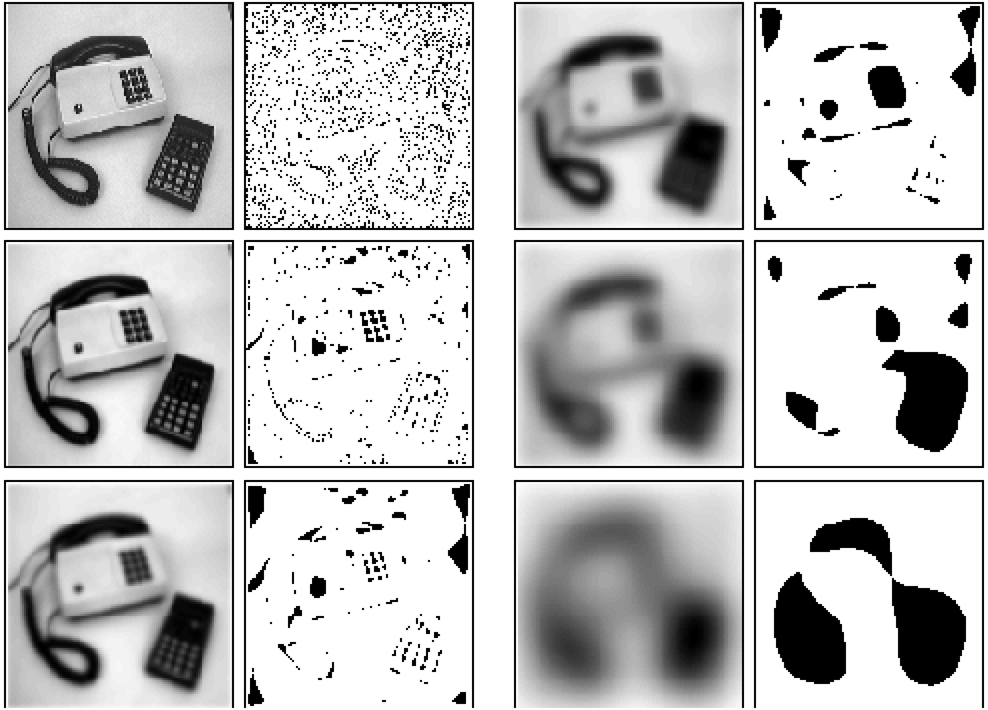


Successively smoothing a signal with Gaussian kernels of increasing width



Trajectories of zero-crossings (of second derivative) in scale-space

Scale-space: illustration



Levels in the scale-space representation of a image at scale levels $t = 0, 2, 8, 32, 128, 512$; local minima at each scale are right panel of each pair.

Scale-space: Naïve version

1/16

1	2	1
2	4	2
1	2	1

1/273

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

1/1003

0	0	1	2	1	0	0
0	3	13	22	13	3	0
1	13	59	97	59	13	1
2	22	97	159	97	22	2
1	13	59	97	59	13	1
0	3	13	22	13	3	0
0	0	1	2	1	0	0

Multi-scale Analysis

- Analyzing information at the same scale will not be effective.

Use windows of different size (i.e., vary scale)

Multi-(scale)/(resolution) Analysis

- Alternative: Same window size, but analyze at different resolutions

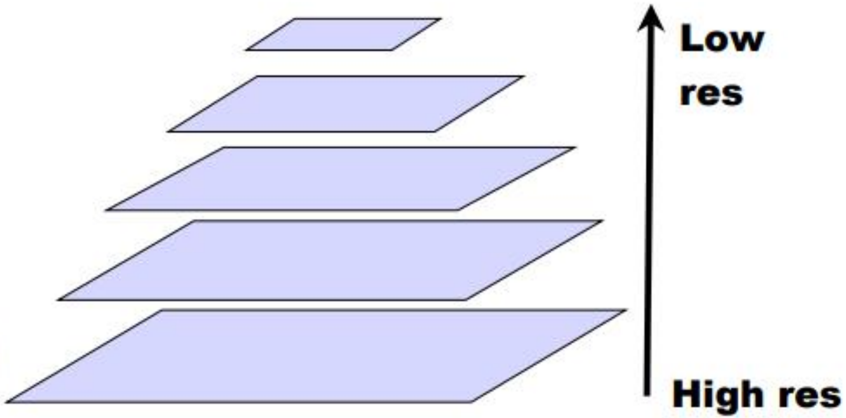
High resolution

Small size objects should be examined at a <u>high</u> resolution

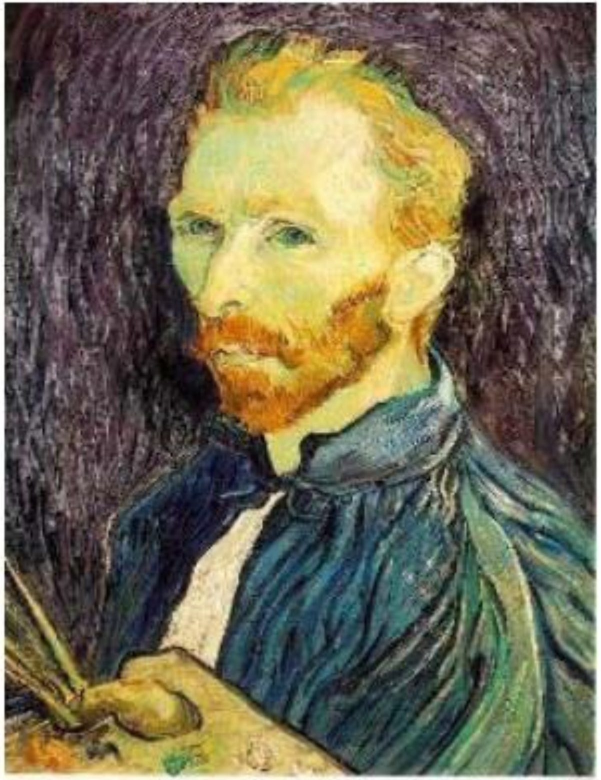
Low resolution

Large size objects should be examined at a low resolution

High resolution → Low resolution



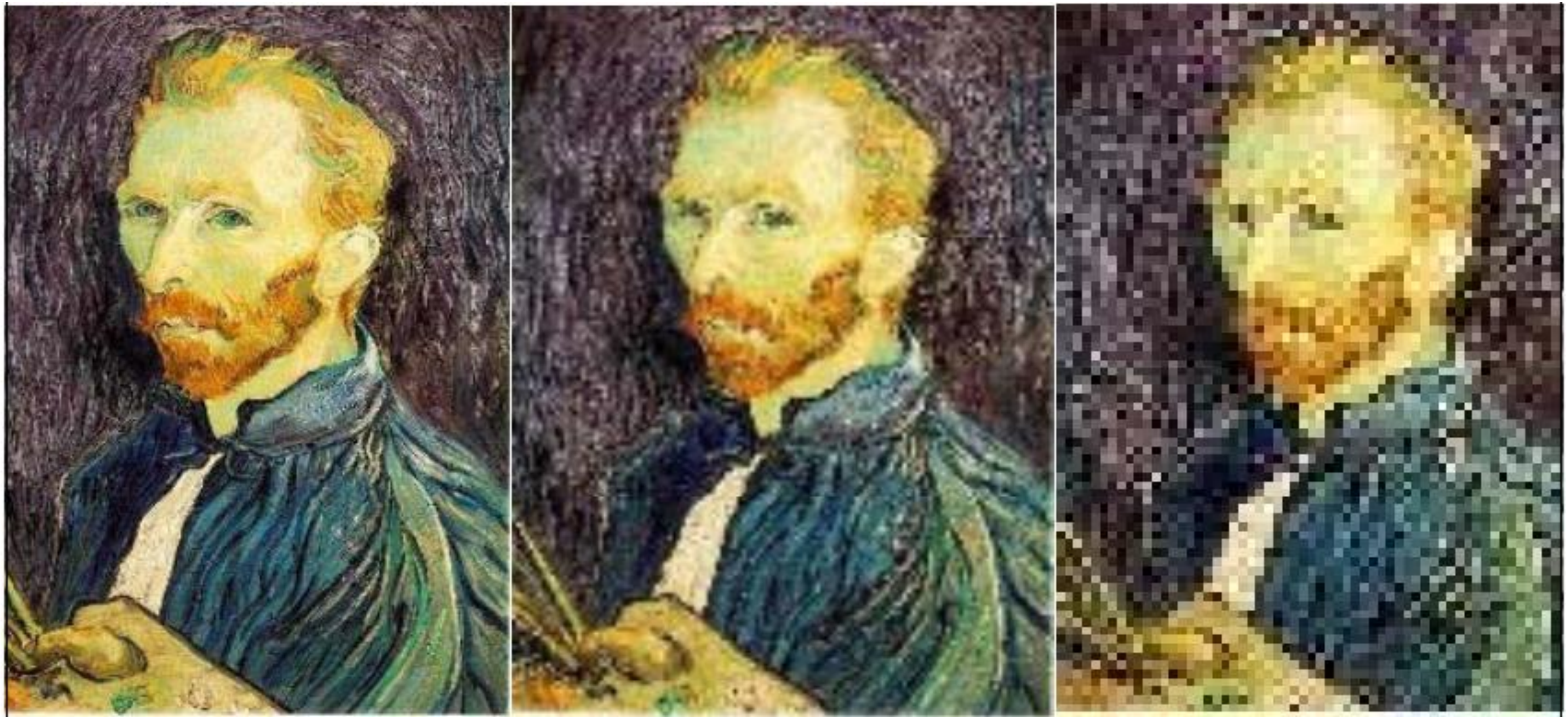
Subsampling (by factor 2)



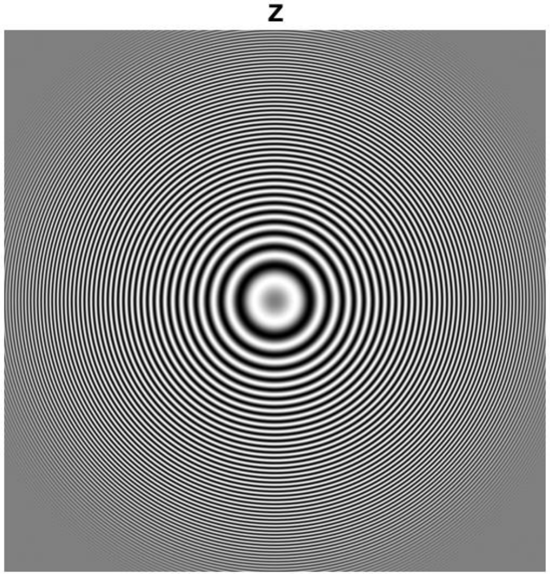
1/4

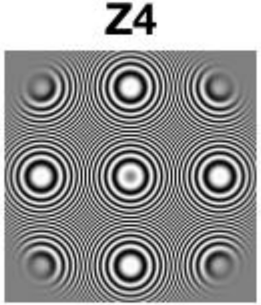
1/8

Naïve Subsampling (by factor 2)



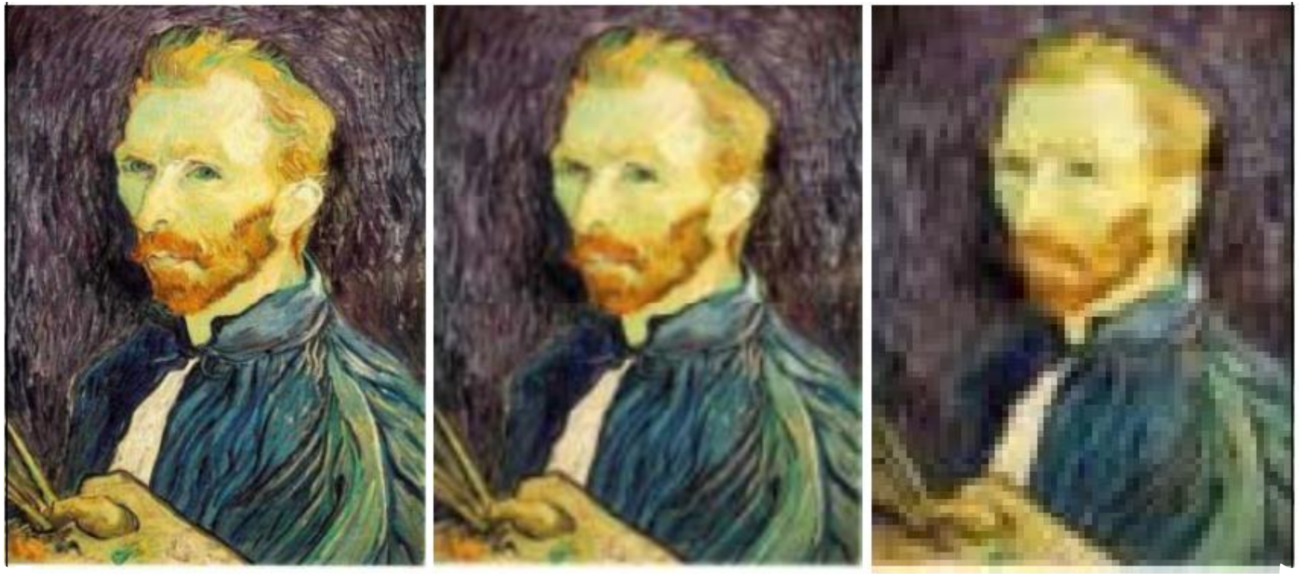
Aliasing





https://blogs.mathworks.com/steve/2017/01/09/aliasing-and-image-resizing-part-2/

Anti-aliasing



Gaussian 1/2 G 1/4 G 1/8

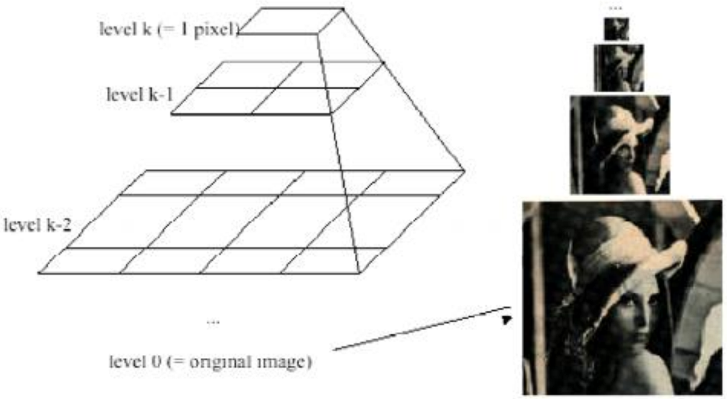
Solution: filter the image, *then* subsample

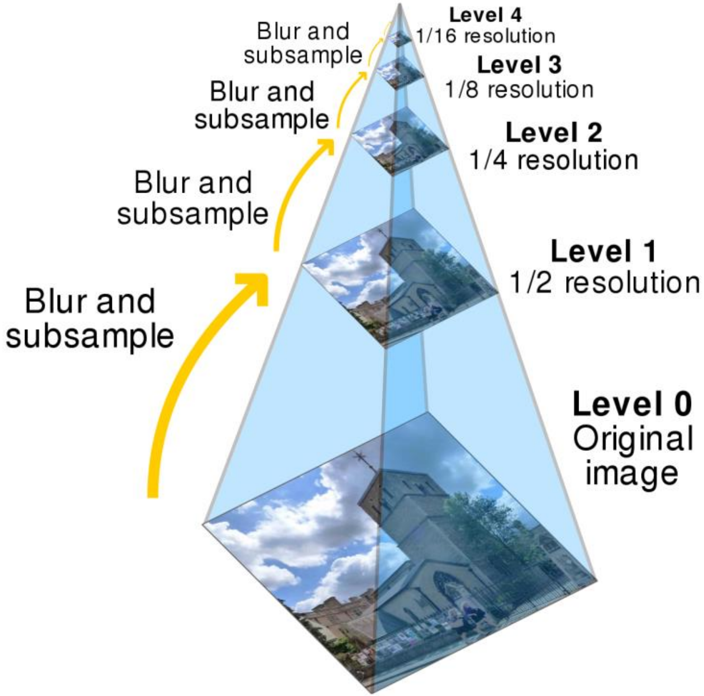
original image
262x195

downsampled (left)
vs. smoothed then
downsampled (right)
131x97

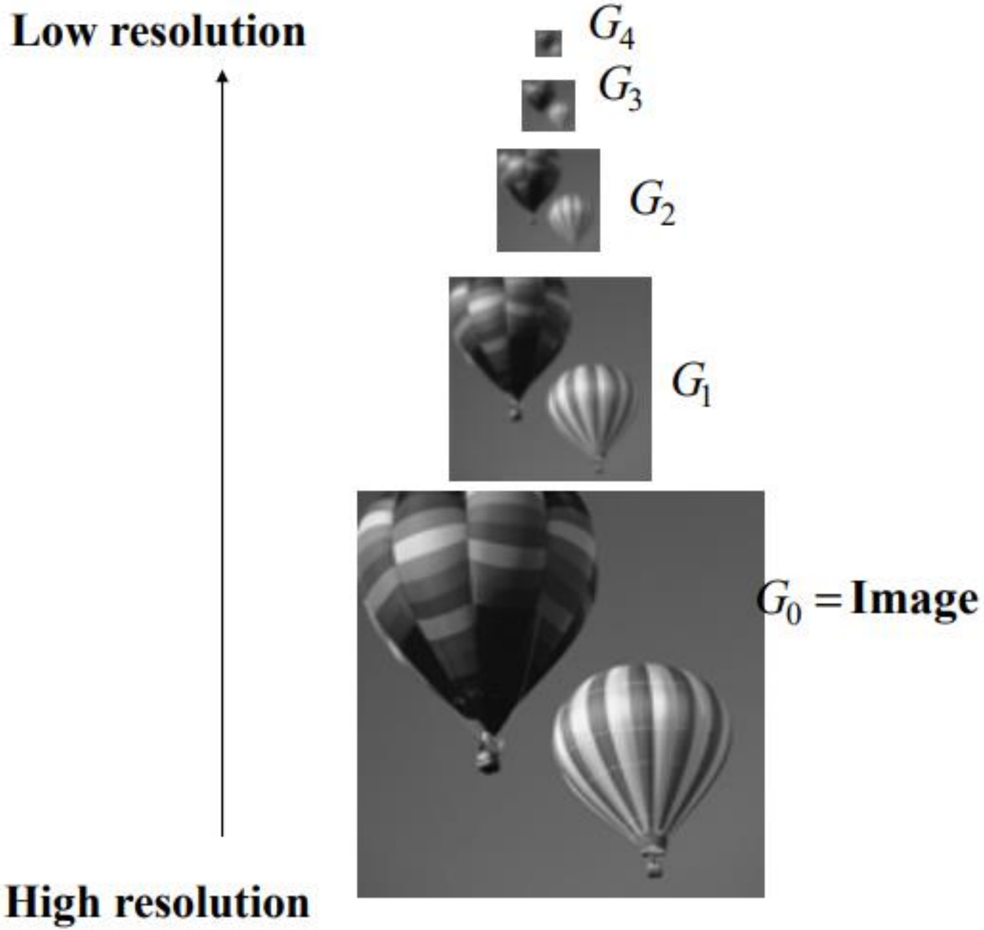
Image Pyramids

Idea: Represent NxN image as a "pyramid" of $1x1, 2x2, 4x4, \ldots, 2^k x 2^k$ images (assuming $N=2^k$)





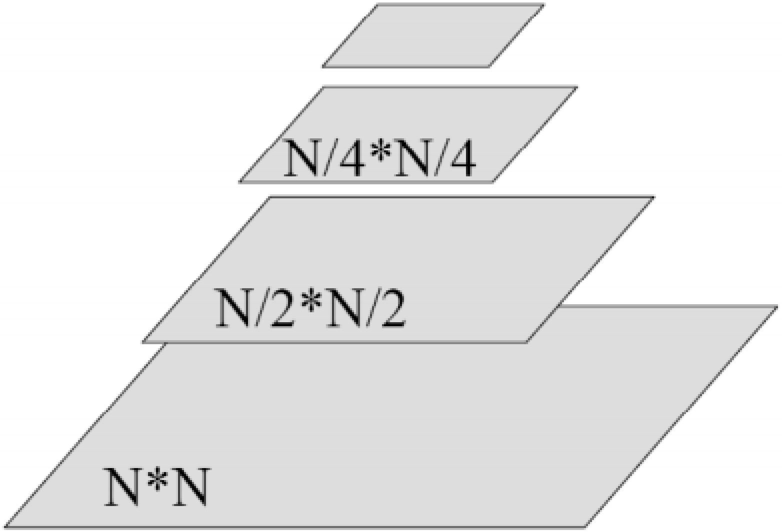
The Gaussian Pyramid



- Cascaded Gaussians
  - Repeated convolution by a smaller Gaussian to simulate effects of a larger one.
- G*(G*f) = (G*G)*f [associativity]

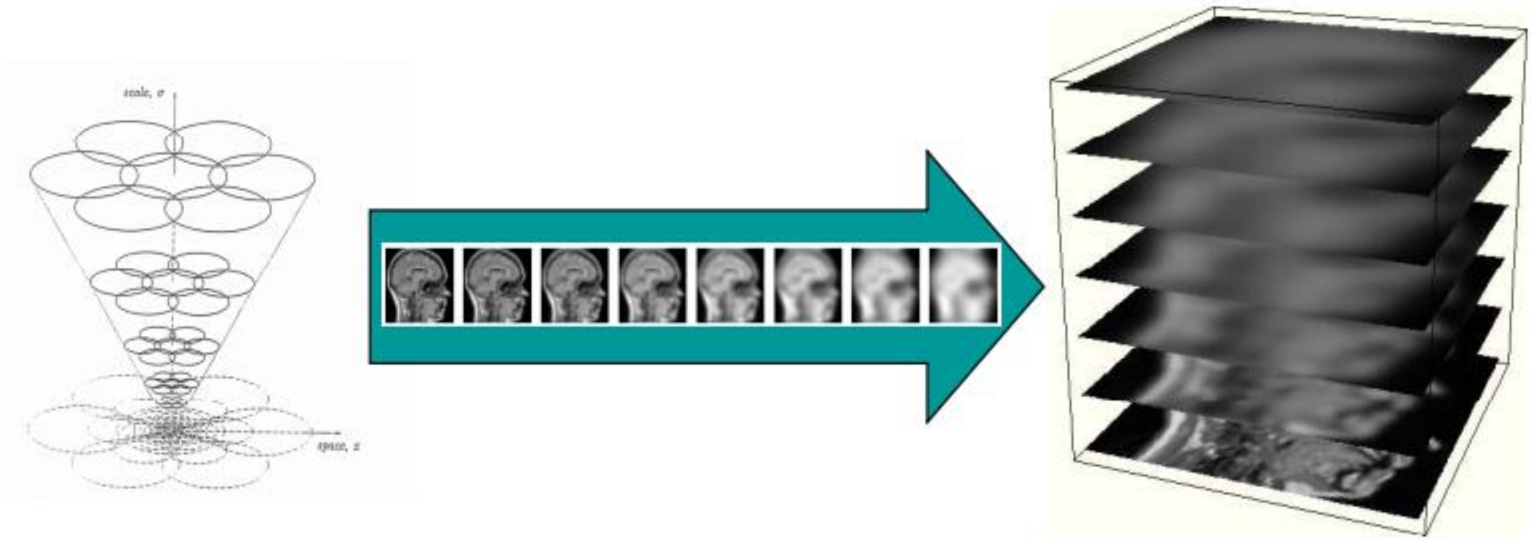
$$\boxed{G_{\sigma_1} * G_{\sigma_2} = G_{\sigma} \quad \sigma^2 = {\sigma_1}^2 + {\sigma_2}^2}$$

Space Required for Pyramids



$$N^2 + \frac{1}{4}N^2 + \frac{1}{16} + N^2 + \cdots = 1\frac{1}{3}N^2$$

The retina measures on many resolutions simultaneously



scale-space


various problems with your input set
cv2.GaussianBlur

What are they good for?

Applications of scaled representations

Search for correspondence

- look at coarse scales, then refine with finer scales

Edge tracking

- a “good” edge at a fine scale has parents at a coarser scale

Control of detail and computational cost in matching

- e.g. finding stripes
- important in texture representation

- Image Blending and Mosaicing

- Data compression (laplacian pyramid)

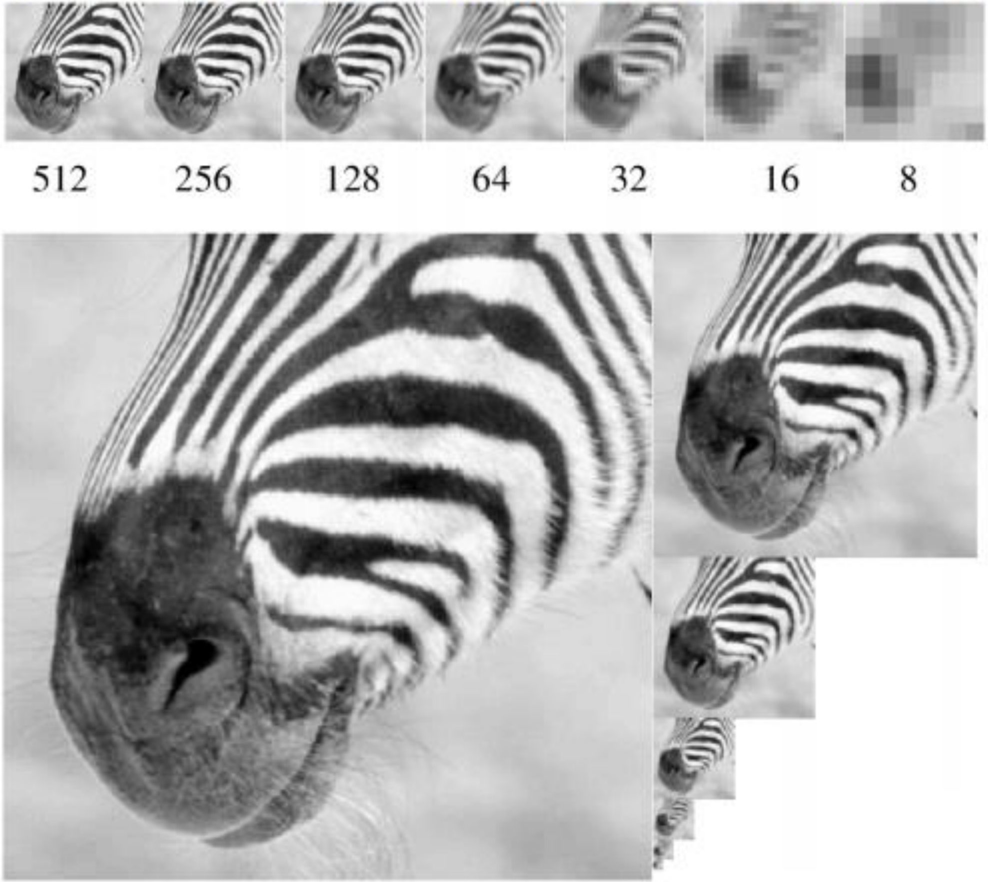
Fast Template Matching

Template

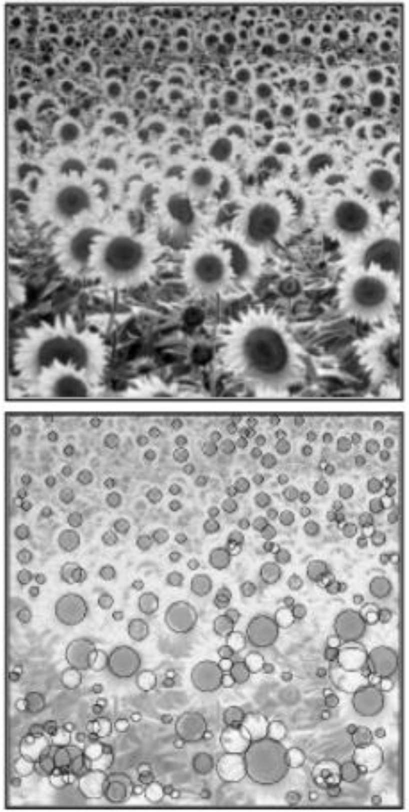
Search Region

Original Image

Basic idea: different scales are appropriate for describing different objects in the image, and we may not know the correct scale/size ahead of time.



Example: Detecting "Blobs" at Different Scales.



- - Most problems in CV & IP, are faced with the question: -
  - What operators to use ?
  - Where to apply them ?
  - How large (scale or range of scales) should they be ?
  - How to relate (interpret) to the actual structure
- In the absence of prior information
  - – use empirical methods; represent data at multiple scales.
- Scale-space method attempts to represent data at all scales simultaneously

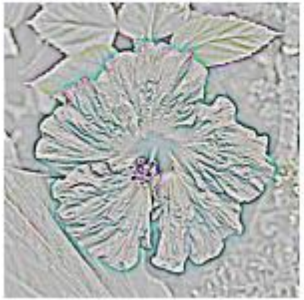
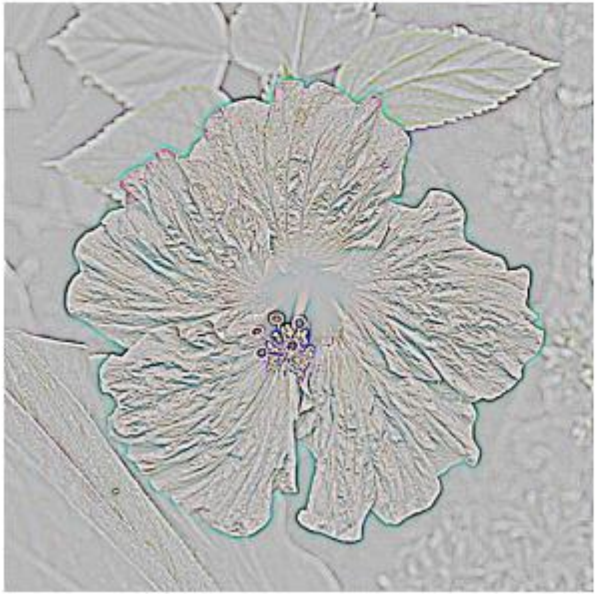
Gaussian vs Laplacian Pyramid




Shown in opposite order for space.

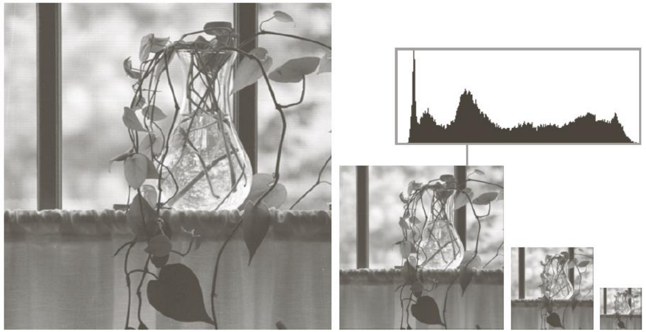
Which one takes more space to store?



Pyramidal coding (cont'd)

Approximation pyramid
(based on Gaussian filter)

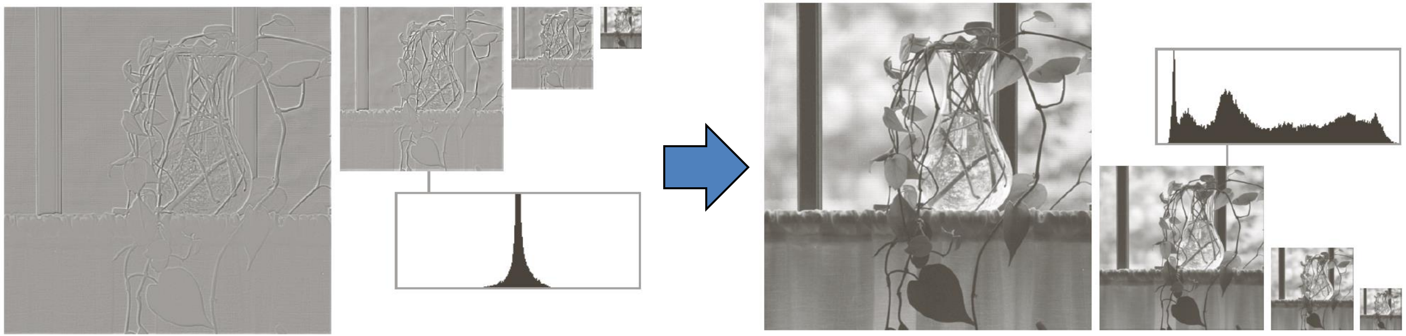


Prediction residual pyramid
(based on bilinear interpolation)

<u>Note</u>: the last level is the same as that of the approximation pyramid

Pyramidal coding (cont'd)

In the absence of quantization errors, the approximation pyramid can be re-constructed from the prediction residual pyramid.



- We only need to keep the prediction residual pyramid only!
  - More efficient representation

Multi-scale Analysis

- Two techniques for representing multi-scale information efficiently
  - Pyramidal coding
  - Sub-band coding

Subband coding

- Decompose an image (or signal) into different frequency bands (analysis step).
- Decomposition is performed so that the subbands can be re-assembled to reconstruct the original image without error (synthesis step)
- Need to choose appropriate filters (i.e., “filter bank”)

Subband coding (cont'd)