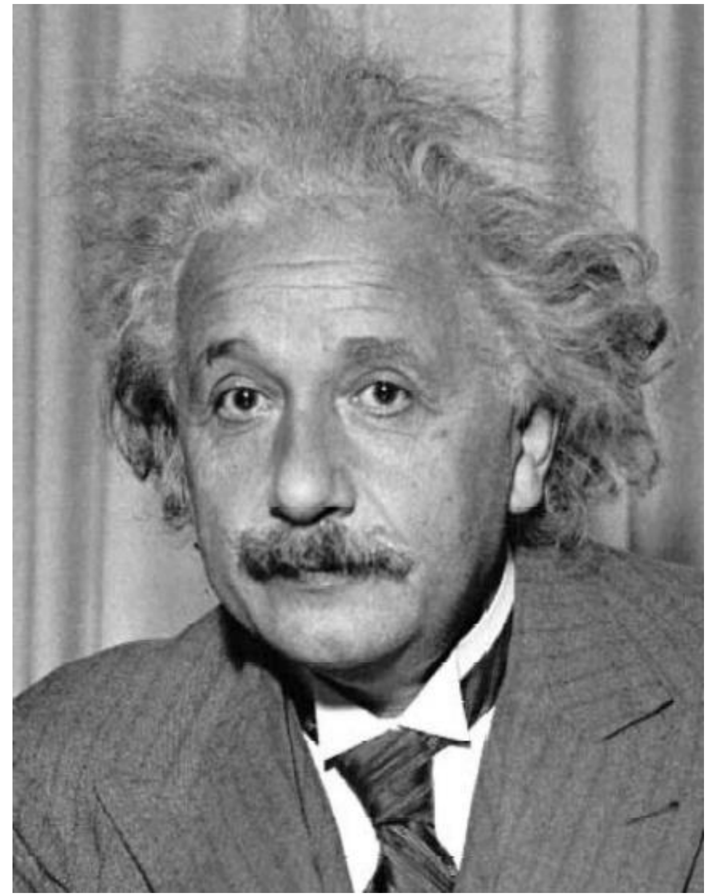
CMPE 362
Digital Image Processing

Template Matching & Image Pyramids

Template matching

- Goal: find  in image



- What is a good similarity or distance measure between two patches?
    - Correlation
    - Zero-mean correlation
    - Sum of Squared Difference
    - Normalized Cross Correlation

Matching with filters

- Goal: find  in image
- Correlation: filter the image with eye patch

$$g(x,y) = \sum_{s,t} w(s,t)\, f(x+s, y+t)$$

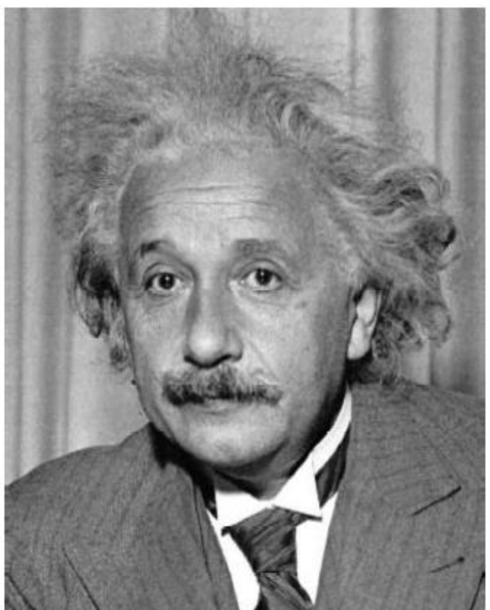
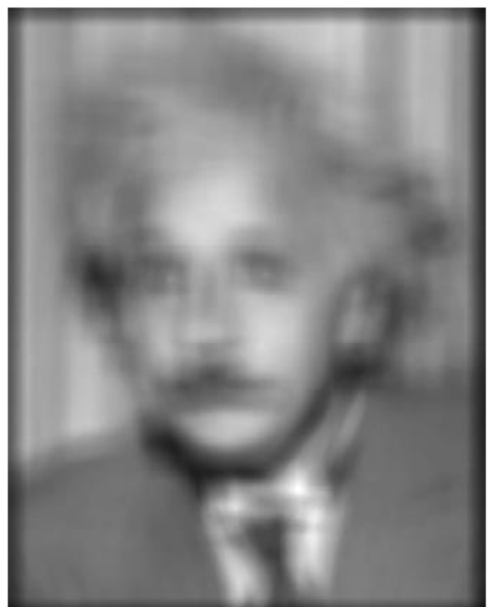


image f



output g

filter w

What went wrong?

Response is stronger for higher intensity.

Matching with filters

What went wrong?

Sensitive to high-contrast areas

- Goal: find  in image
- Zero-mean correlation: filter the image with zero-mean eye

$$g(x,y) = \sum_{s,t} (w(s,t) - \bar{w})\, f(x+s, y+t)$$

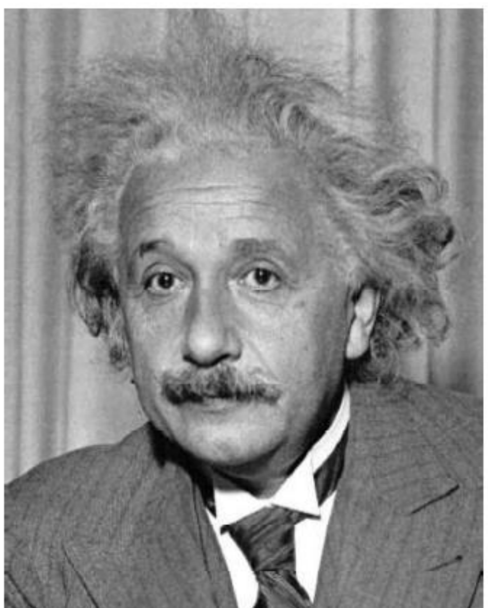
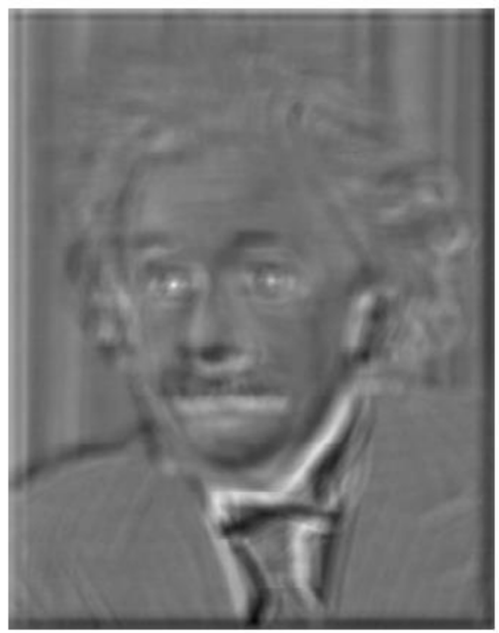
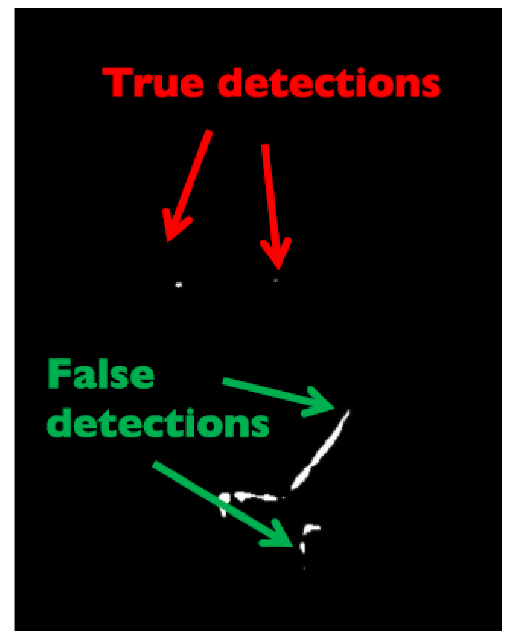


image f



output g



thresholded

Matching with filters

- Goal: find  in image
- Sum of Squared Difference (SSD)

$$g(x,y) = \sum_{s,t} (w(s,t) - f(x+s,y+t))^2$$

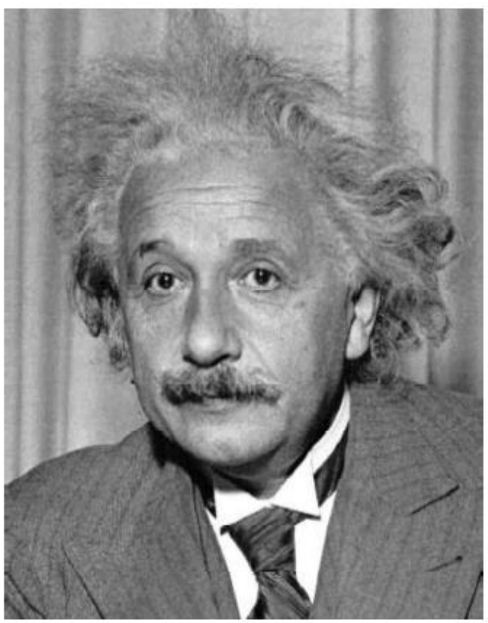
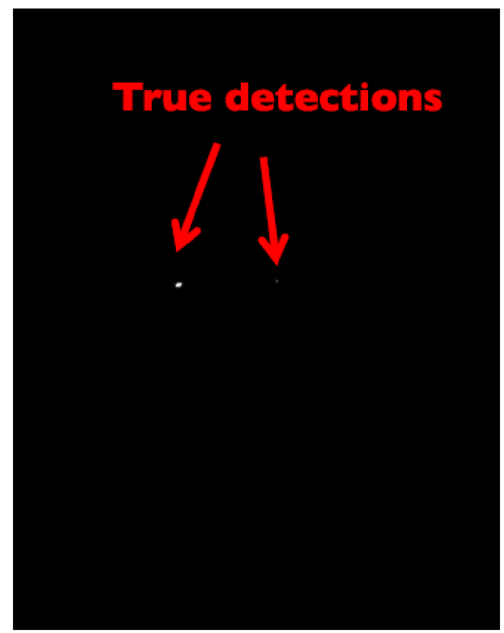


image f

$1 - \sqrt{g}$



thresholded

Matching with filters

- Goal: find  in image
- Sum of Squared Difference (SSD)

$$g(x,y) = \sum_{s,t} (w(s,t) \; - \; f(x+s, y+t))^2$$

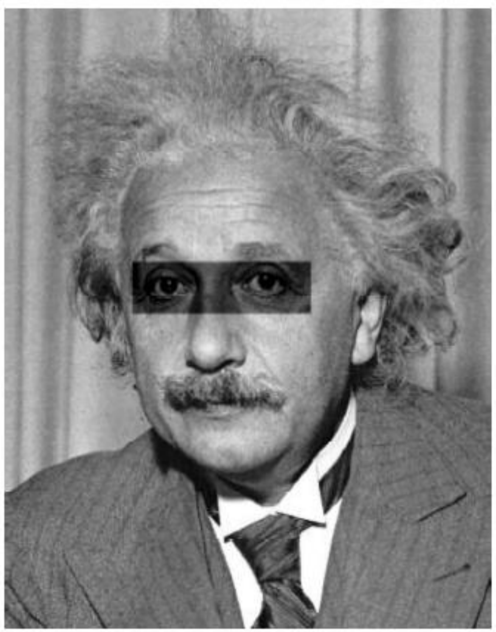


image f

$1 - \sqrt{g}$

What is the potential downside of SSD?

SSD is sensitive to average intensity.

Matching with filters

- Goal: find  in image
- Normalized Cross Correlation (NCC)

mean of template

mean of image patch

$$g(x,y) = \frac{\sum_{s,t}(w(s,t) - \bar{w})(f(x-s,y-t) - \bar{f}_{x,y})}{\left[\sum_{s,t}(w(s,t) - \bar{w})^2 \sum_{s,t}(f(x-s,y-t) - \bar{f}_{x,y})^2\right]^{0.5}}$$

Matching with filters

- Goal: find  in image
- Normalized Cross Correlation (NCC)

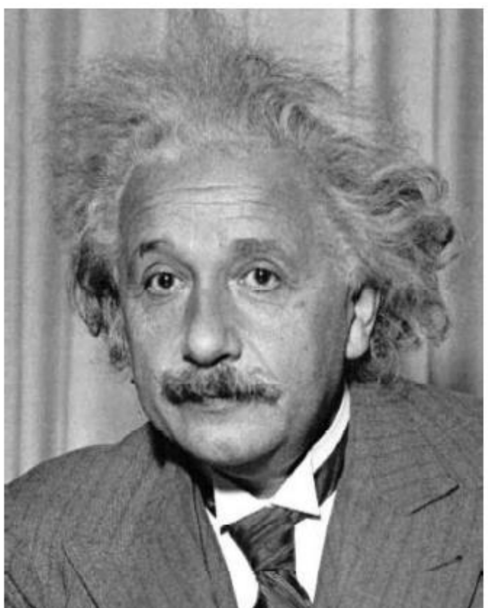
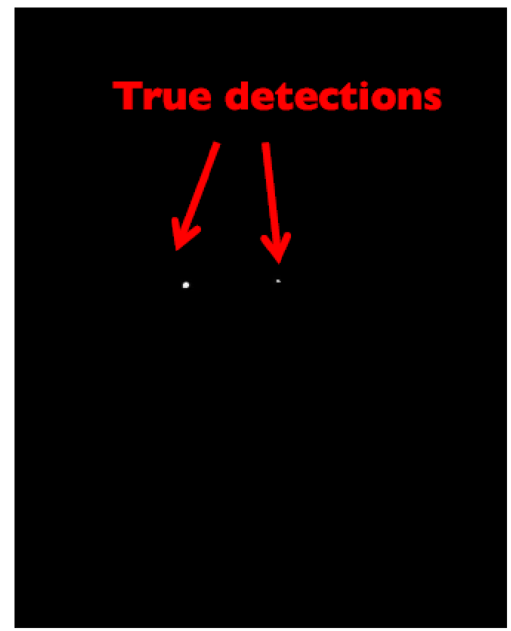


image f

output g (NCC)



thresholded

Matching with filters

- Goal: find  in image
- Normalized Cross Correlation (NCC)

NCC is invariant to mean and scale of intensity.

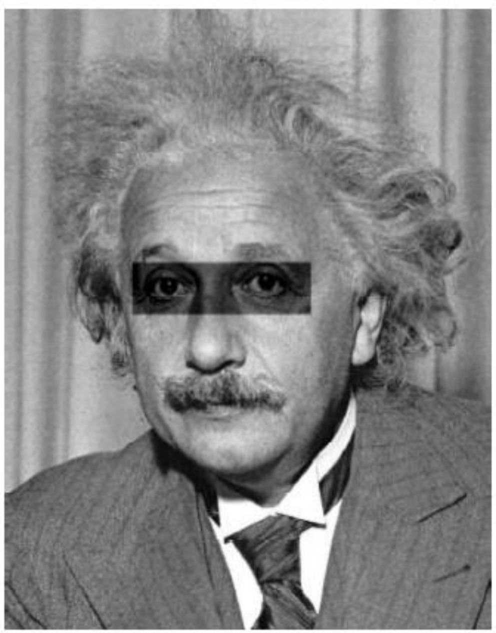
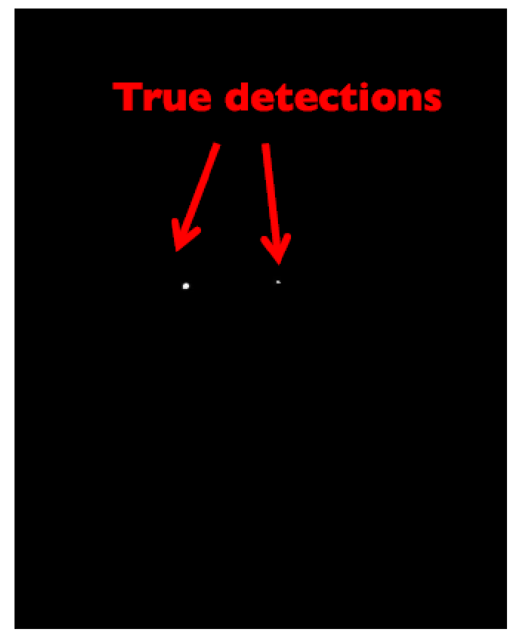


image f

output g (NCC)



thresholded

Q: What is the best method to use?

- A: Depends
  - SSD is faster but sensitive to overall intensity
  - NCC is slower but invariant to local average intensity and contrast

Q: What if we want to find larger or smaller eyes?

- Image Pyramid

Image Pyramids

- Image information occurs over many different spatial scales.
- Image pyramids –multi- resolution representations for images– are a useful data structure for analyzing and manipulating images over a range of spatial scales.

Image pyramids

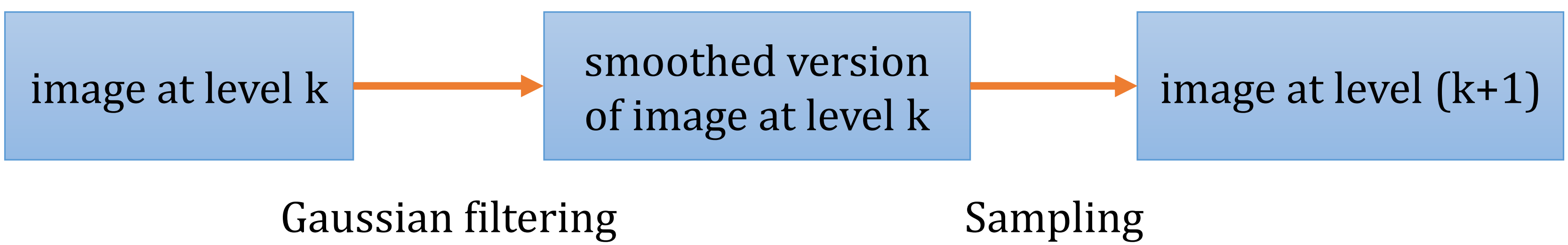
- Gaussian pyramid
- Laplacian pyramid

Image pyramids

- Gaussian pyramid
- Laplacian pyramid

Gaussian pyramid

- Gaussian pyramid creates versions of the input image at multiple resolutions.
- Gaussian pyramid is useful for analyzing an image across different spatial scales.
- Images at consecutive levels of Gaussian pyramid are obtained as:



- Smoothing is done using Gaussian filters because convolution of a Gaussian with another Gaussian is a Gaussian.

Gaussian pyramid

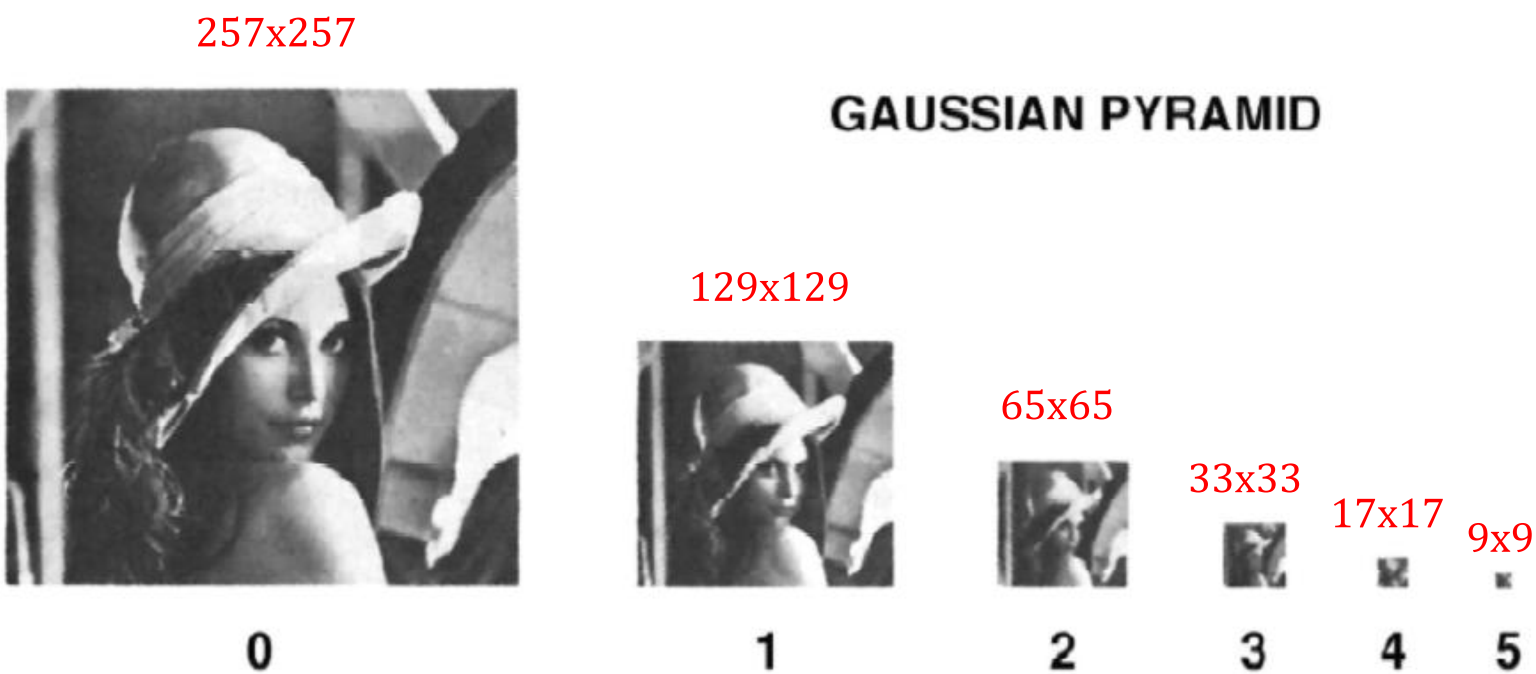


Fig. 4. First six levels of the Gaussian pyramid for the "Lady" image The original image, level 0, meusures 257 by 257 pixels and each higher level array is roughly half the dimensdons of its predecessor. Thus, level 5 measures just 9 by 9 pixels.

512 256 128 64 32 16 8

Slide credit: B. Freeman and A. Torralba

Template Matching with Image Pyramids

Input: Image, Template

1. Match template at current scale

2. Downsample image

3. Repeat 1-2 until the image is very small

4. Take responses above some threshold, perhaps with non-maxima suppression

Coarse-to-fine Image Registration

1. Compute Gaussian pyramid
2. Align with coarse pyramid
3. Successively align with finer pyramids
   - Search smaller range

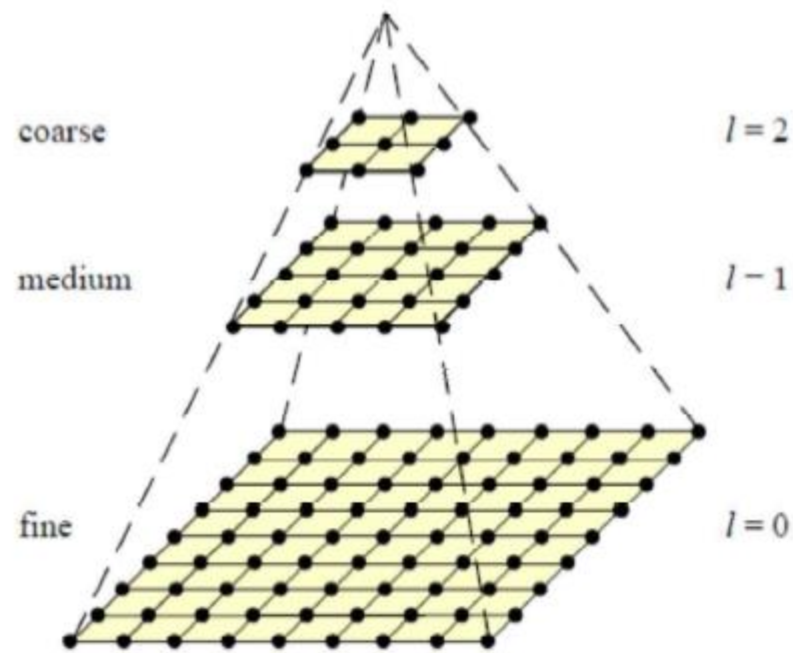
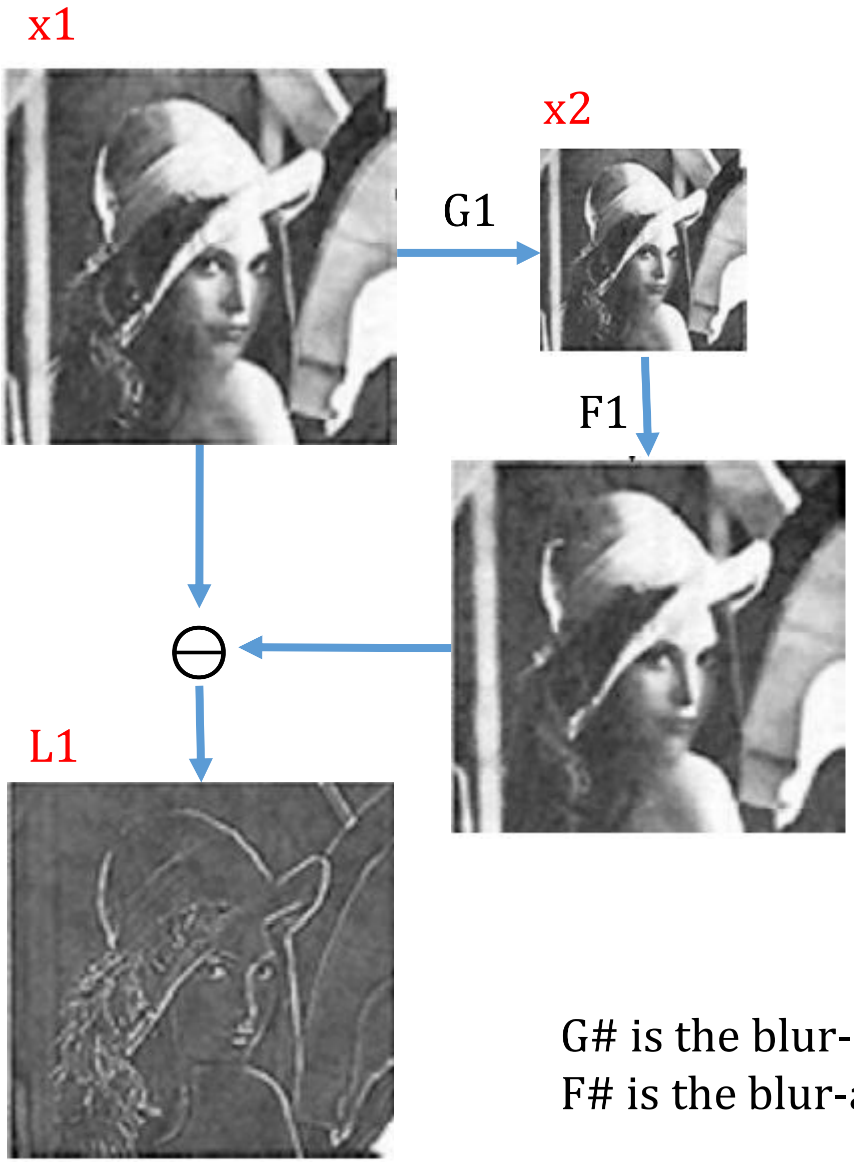


Image pyramids

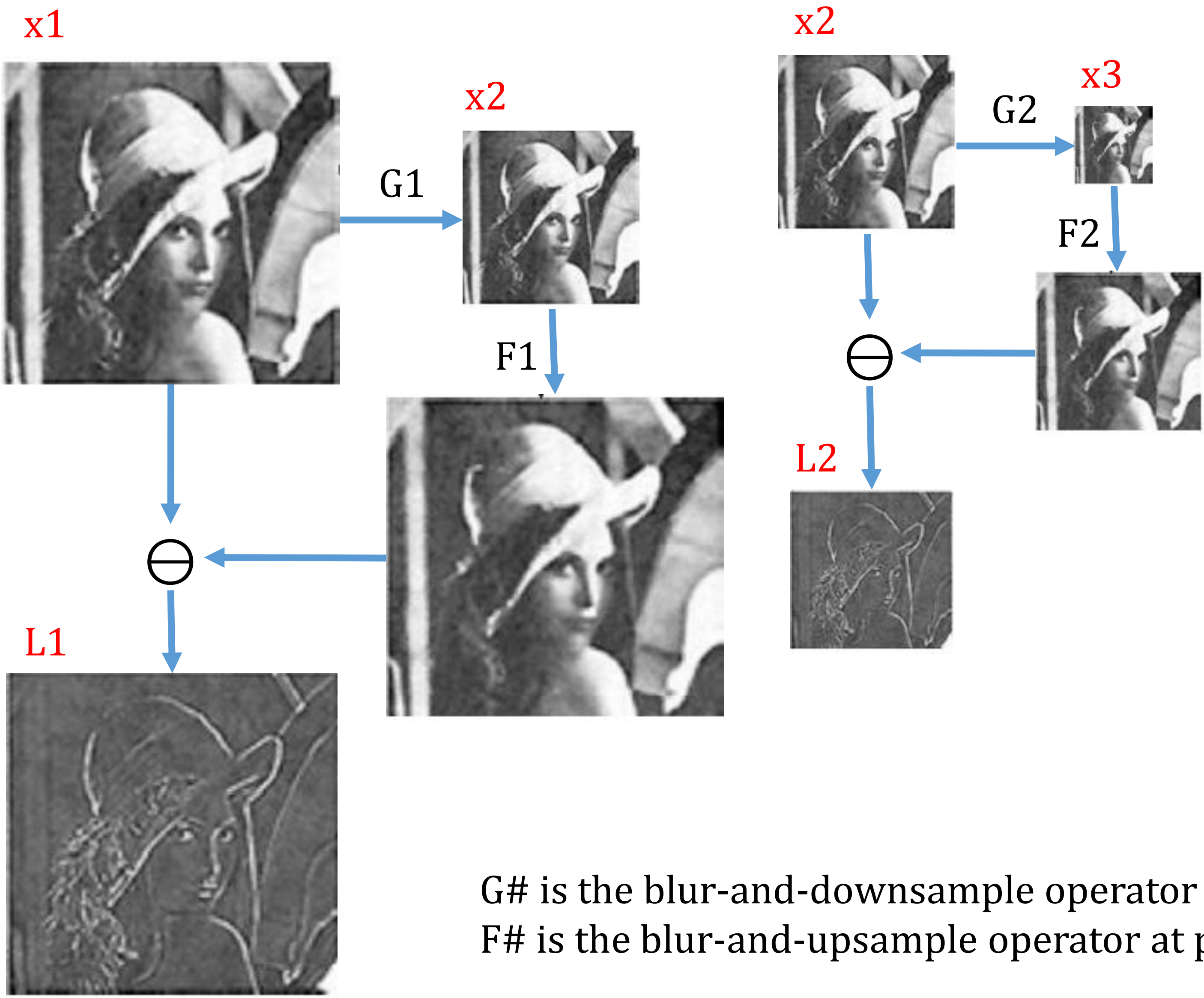
- Gaussian pyramid
- Laplacian pyramid

Laplacian pyramid



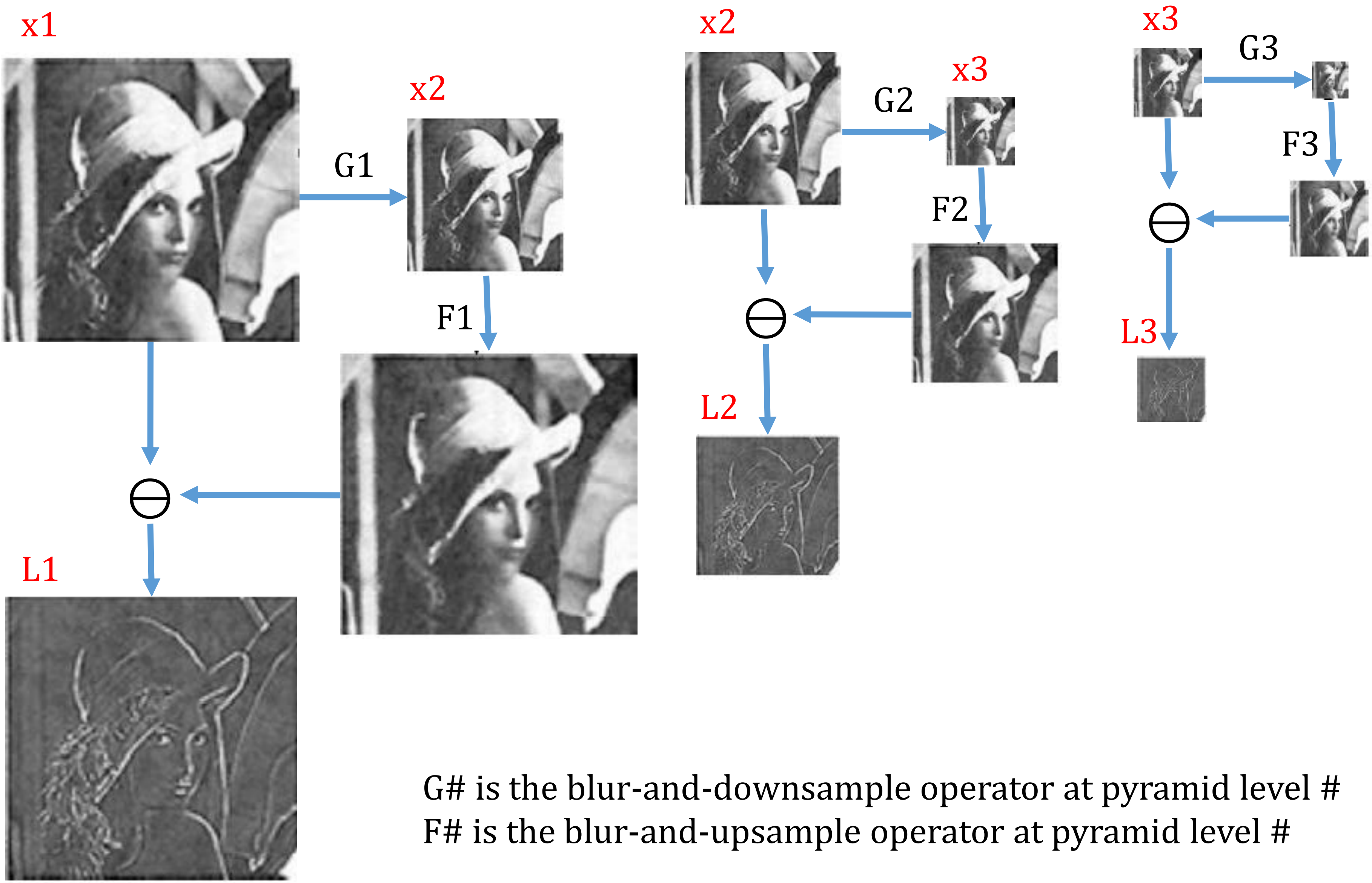
G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid



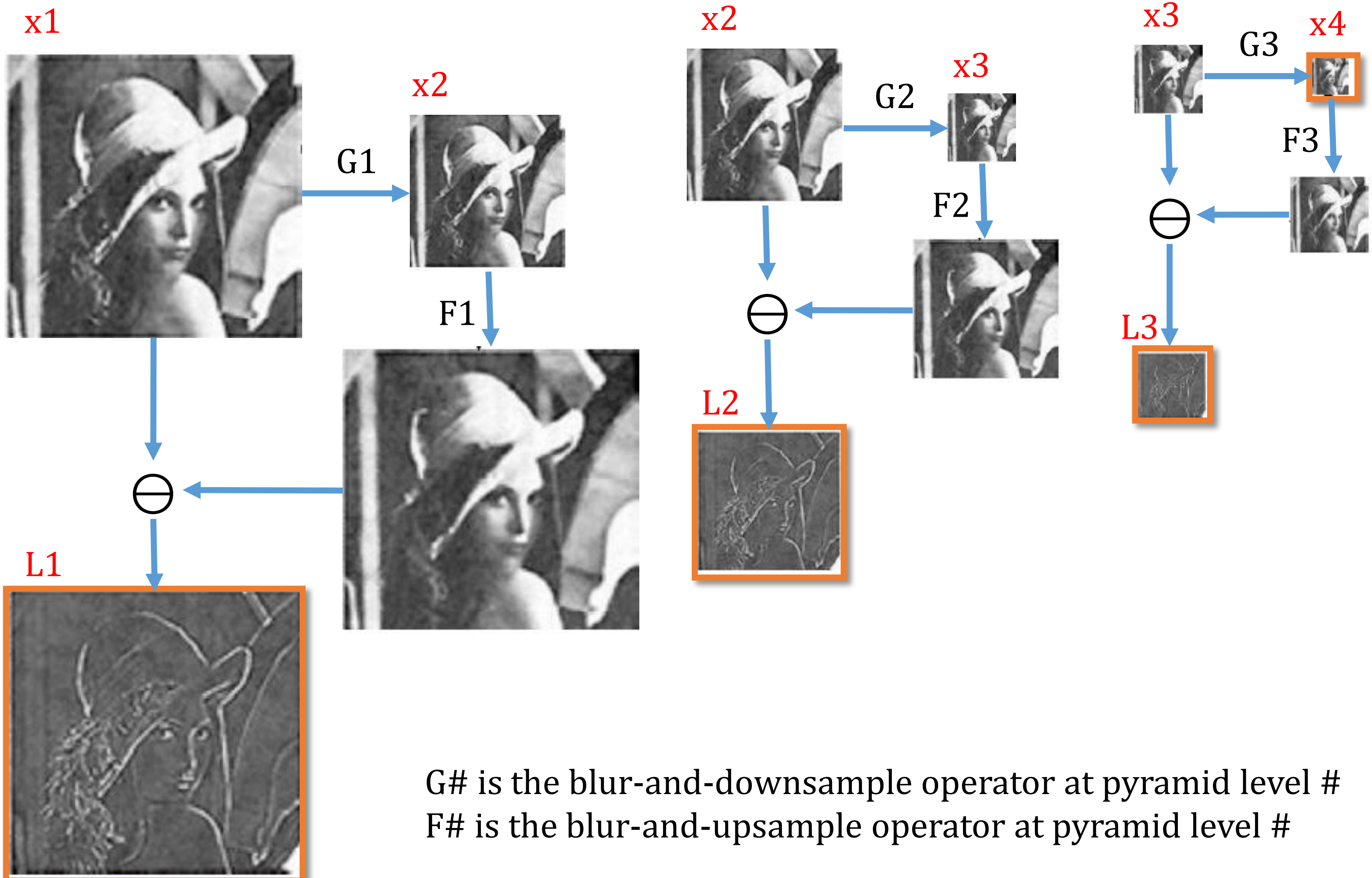
G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid



G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid



G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid

x4

L3

L2

L1

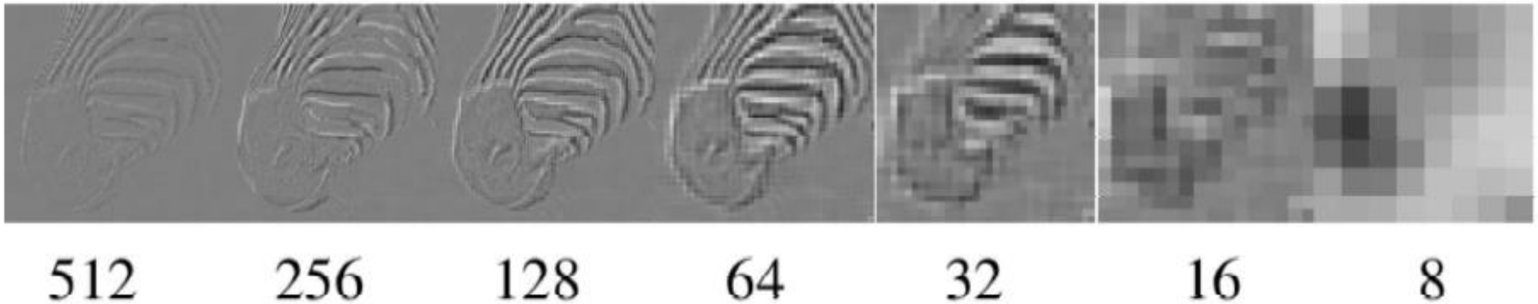
G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

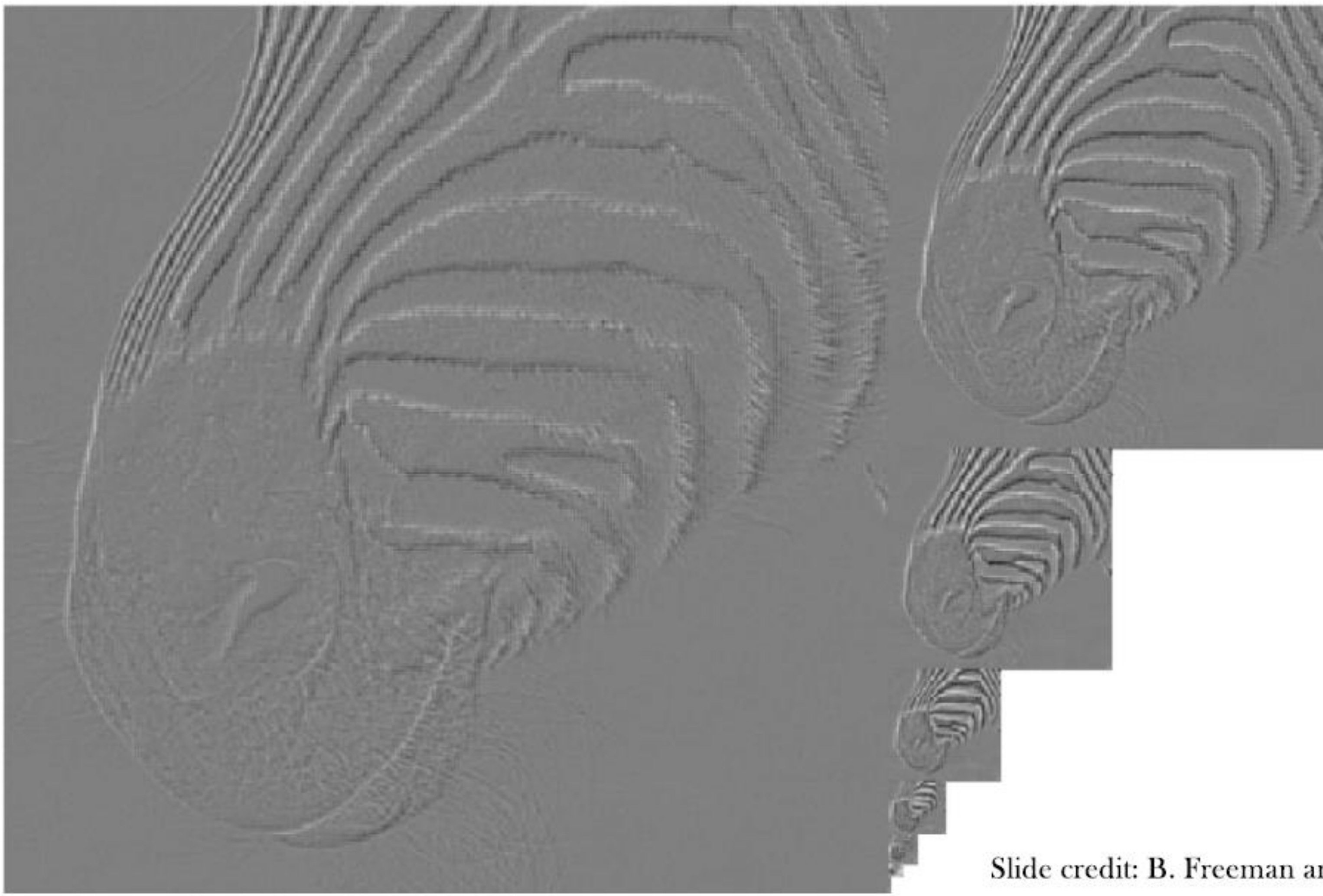
Laplacian pyramid

- Laplacian pyramid provides an extra level of analysis as compared to Gaussian pyramid by breaking the image into different spatial frequency bands.
- Each level represents spatial frequencies largely unrepresented at other level.

512 256 128 64 32 16 8

Slide credit: B. Freeman and A. Torralba





Slide credit: B. Freeman and A. Torralba

Laplacian pyramid reconstruction algorithm

x4

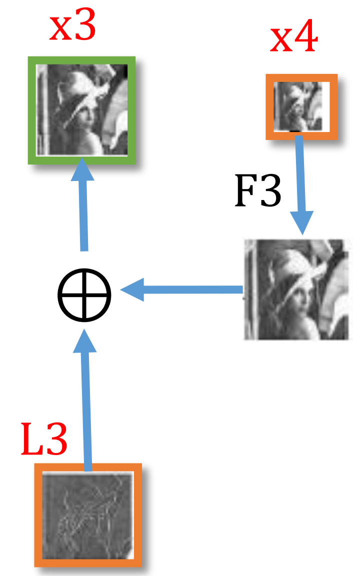
L3

L2

L1

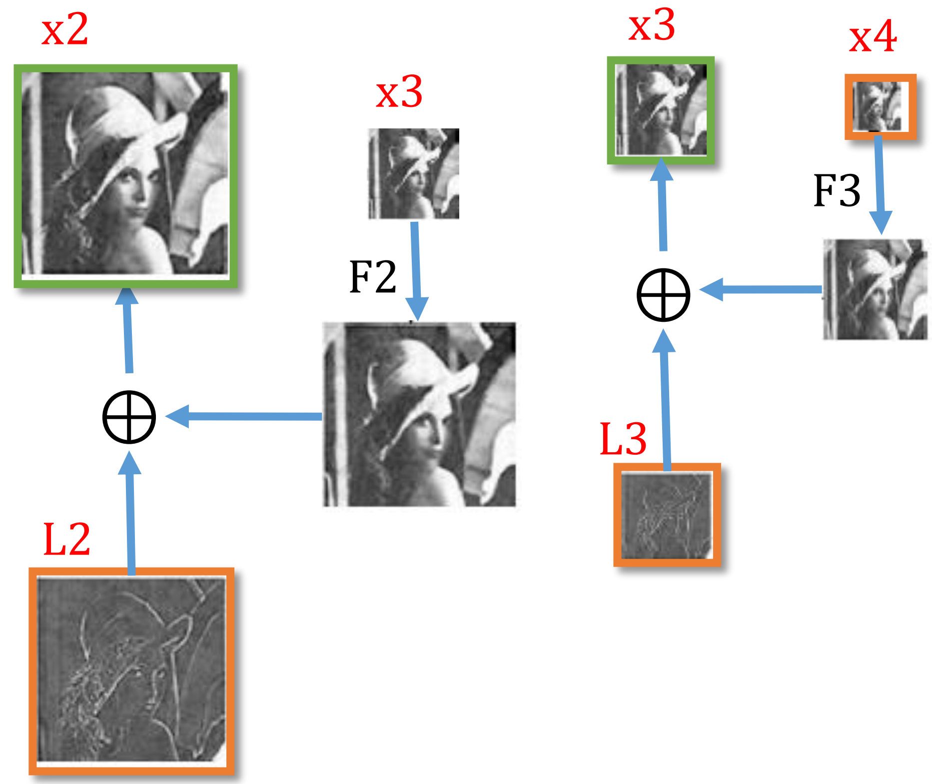
G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid reconstruction algorithm



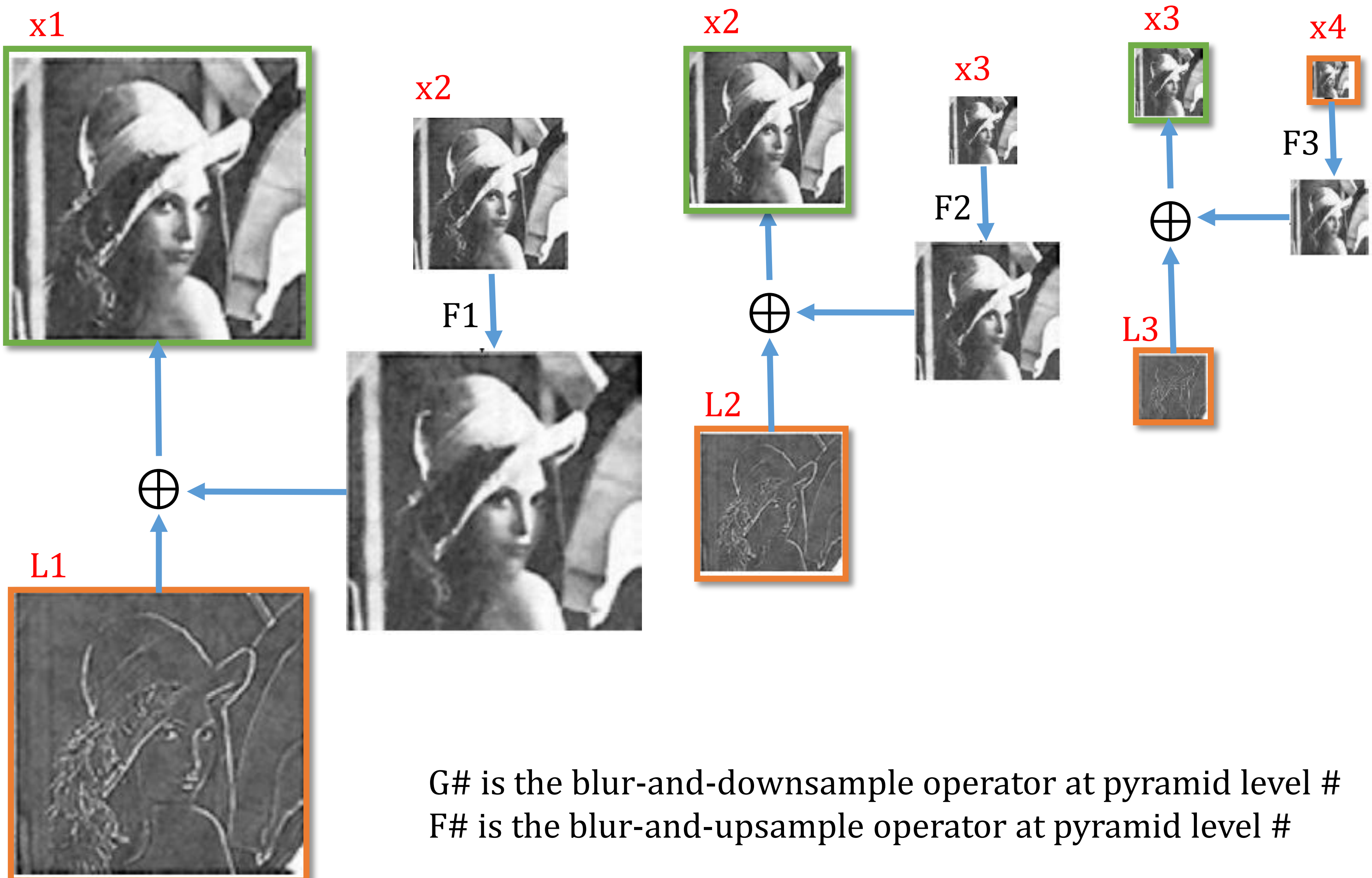
G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid reconstruction algorithm



G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid reconstruction algorithm



G# is the blur-and-downsample operator at pyramid level #
F# is the blur-and-upsample operator at pyramid level #

Laplacian pyramid applications

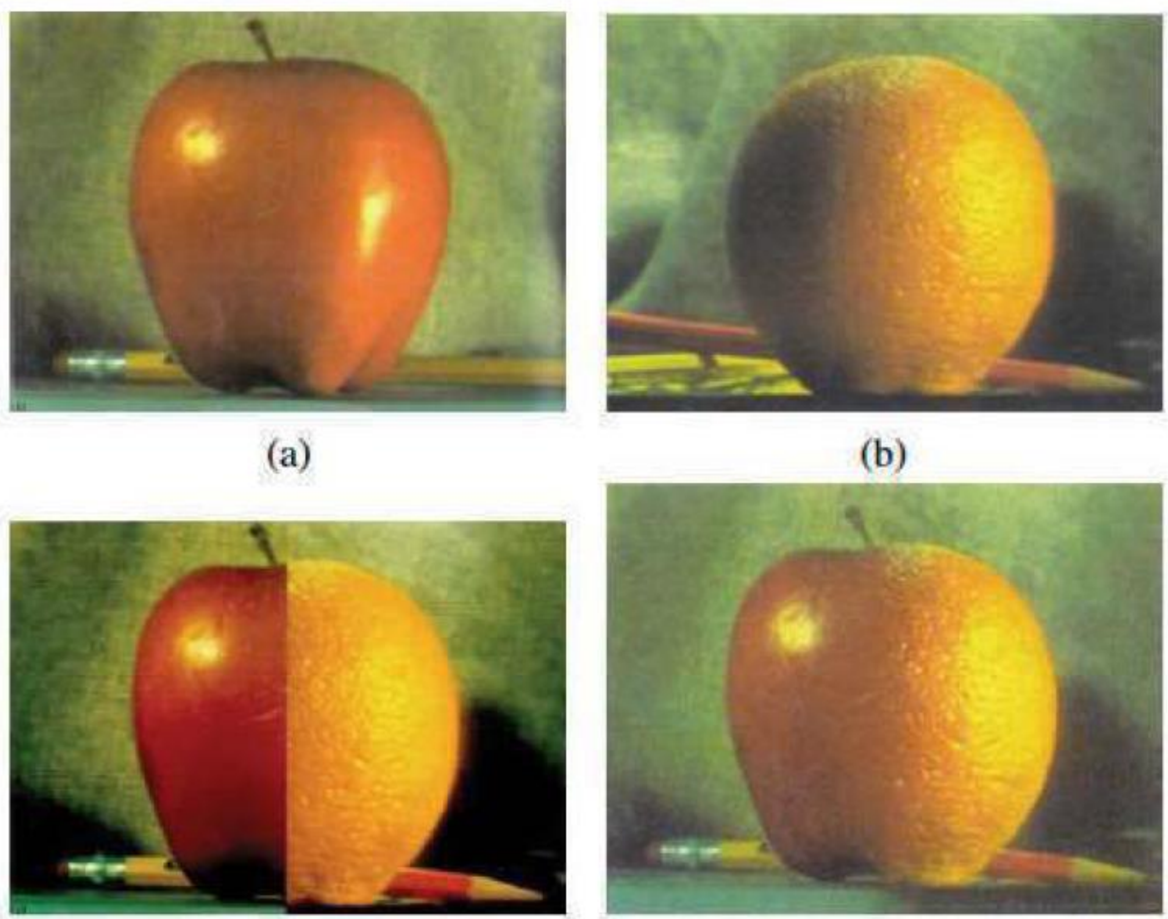
- Texture synthesis
- Image compression
- Noise removal

IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM-31, NO. 4, APRIL 1983

The Laplacian Pyramid as a Compact Image Code

PETER J. BURT, MEMBER, IEEE, AND EDWARD H. ADELSON

Image blending



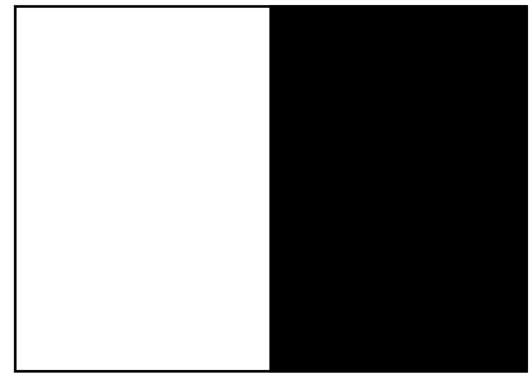
Slide credit: B. Freeman and A. Torralba

Image blending

A

B

mask



Blended image

Build Laplacian pyramid for both images: LA, LB

Build Gaussian pyramid for mask: G

Build a combined Laplacian pyramid:

L(j) = G(j) LA(j) + (1-G(j)) LB(j)

Collapse L to obtain the blended image

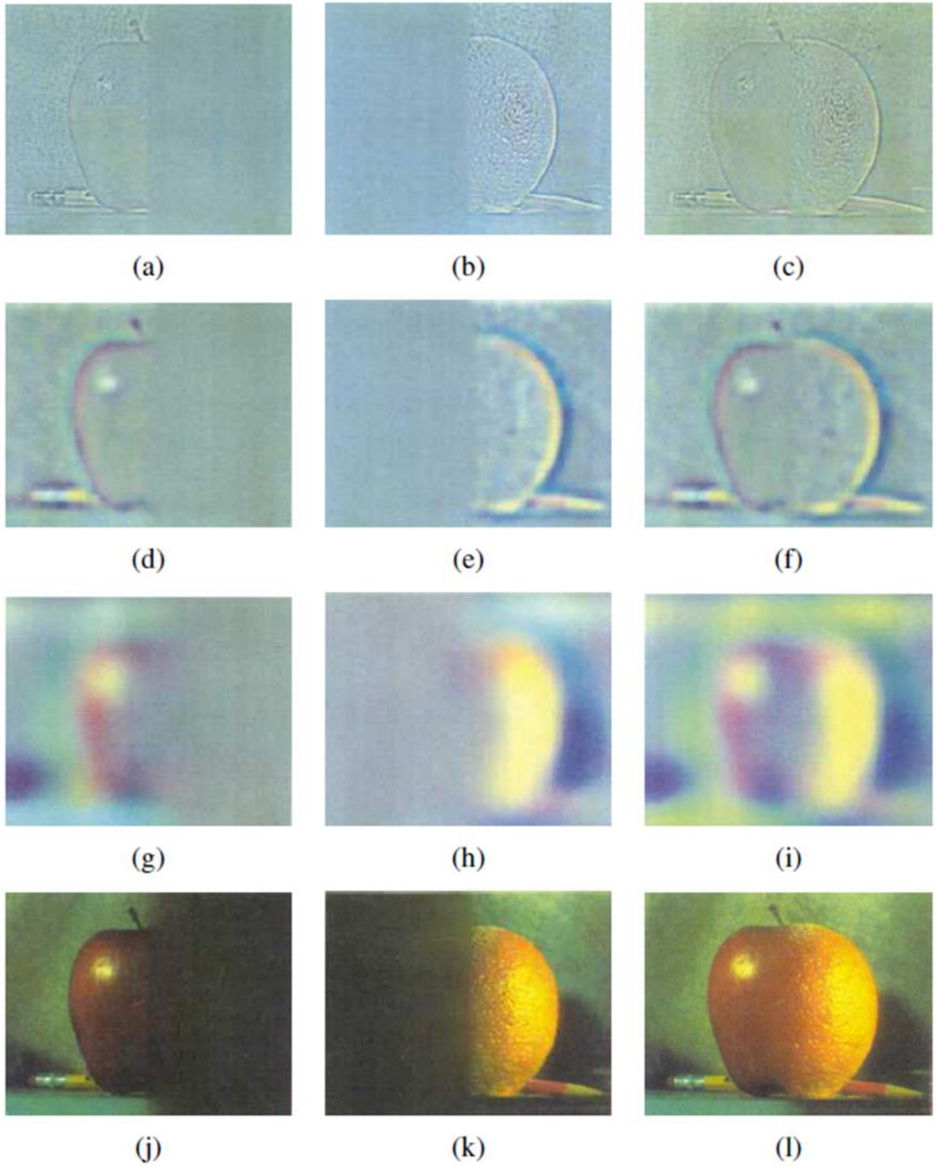


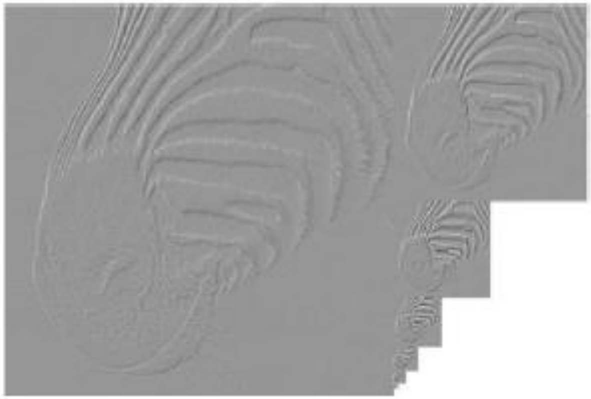
Figure 3.42 *Laplacian pyramid blending details (Burt and Adelson 1983b) © 1983 ACM. The first three rows show the high, medium, and low-frequency parts of the Laplacian pyramid (taken from levels 0, 2, and 4). The left and middle columns show the original apple and orange images weighted by the smooth interpolation functions, while the right column shows the averaged contributions.*

Image pyramids

- Gaussian

Progressively blurred and subsampled versions of the image.
Adds scale invariance to fixed-size algorithms.

- Laplacian



Shows the information added in Gaussian pyramid at each spatial scale.
Useful for noise reduction & coding.