Multiscale Image Representations and The Wavelet Transform

Image Processing

CSE 166

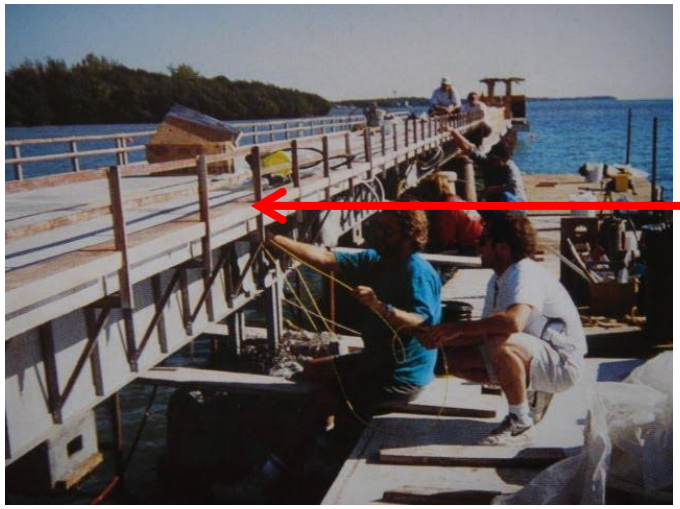
Lecture 11

Announcements

- Assignment 4 is due today, 11:59 PM
- Assignment 5 will be released May 13
  - Due May 20, 11:59 PM
- Reading
  - Chapter 6: Wavelet and Other Image Transforms
    - Section 6.10

The importance of scale

- The features of many objects in an image are only important over a certain spatial extent
- The image scale of the object is critical to recognizing the object
  - However, the scale of objects in an image is typically unknown



Miniature model from *True Lies* (1994)

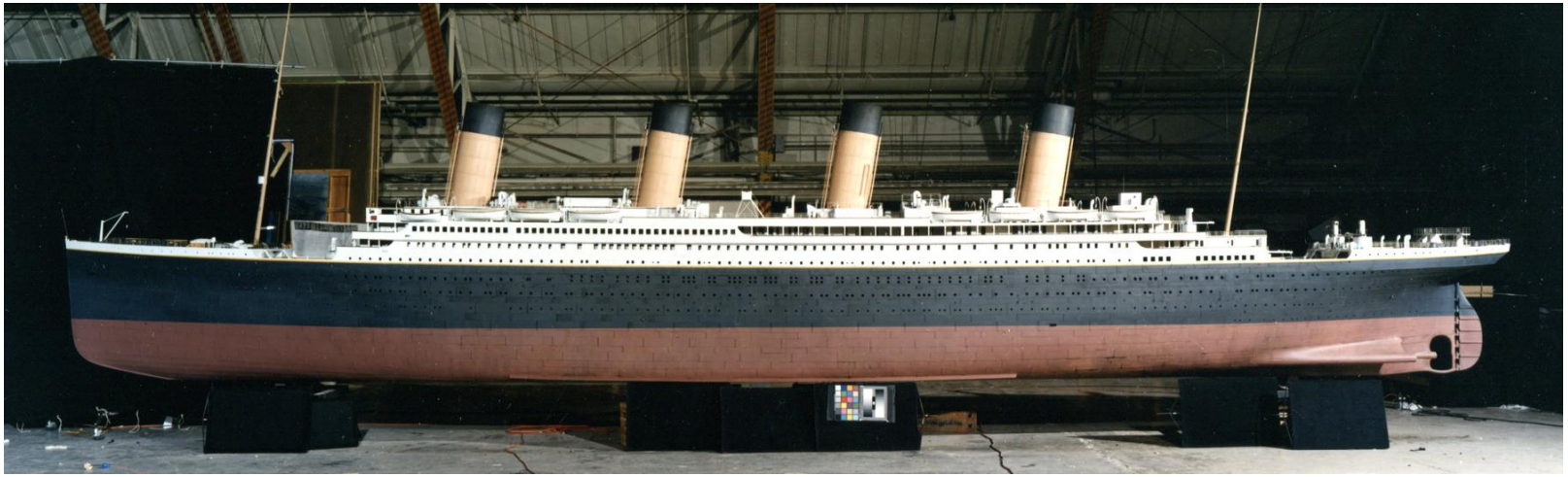
Smaller physical bridge, but more image pixels than image of real bridge on the right



Seven Mile Bridge, Florida Keys

The importance of scale

- The features of many objects in an image are only important over a certain spatial extent
- The image scale of the object is critical to recognizing the object
  - However, the scale of objects in an image is typically unknown



Miniature model from *Titanic* (1997)

The importance of scale

- The features of many objects in an image are only important over a certain spatial extent
- The image scale of the object is critical to recognizing the object
  - However, the scale of objects in an image is typically unknown

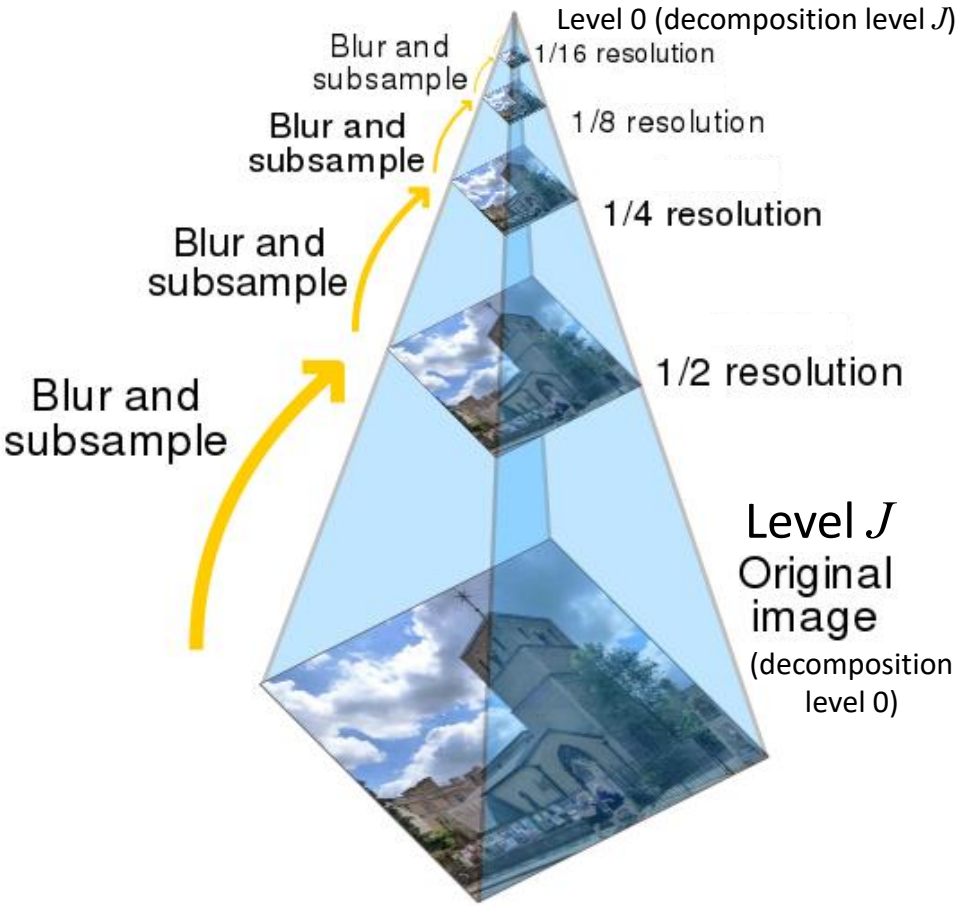
- So, how do we get the image of the object to be at (or near) the expected image scale?
  - Using multiscale image representations

Multiscale image representations

- Image pyramids
  - Gaussian pyramid (approximation pyramid)
  - Laplacian pyramid (prediction residual pyramid)
- Scale space
- Wavelet decomposition

Gaussian pyramid

- Earliest (early 1980s) multiscale image representation
- Different levels of pyramid approximate the original image at different scales
- Width and height of next decomposition level is width and height of current decomposition level divided by rate
  - Typical rate is 2



Gaussian pyramid



Level J (decomposition level 0) is original image

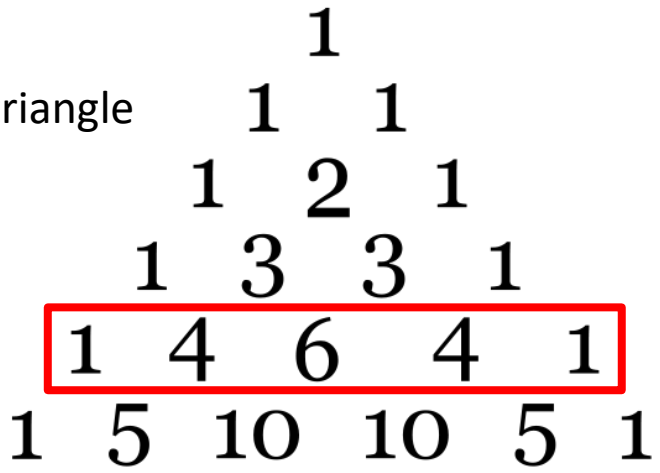


Collectively, remaining levels are 1/3 size of original image

Gaussian pyramid implementation

- Binomial filter kernel approximates kernel of Gaussian with variance = 1

Pascal's triangle

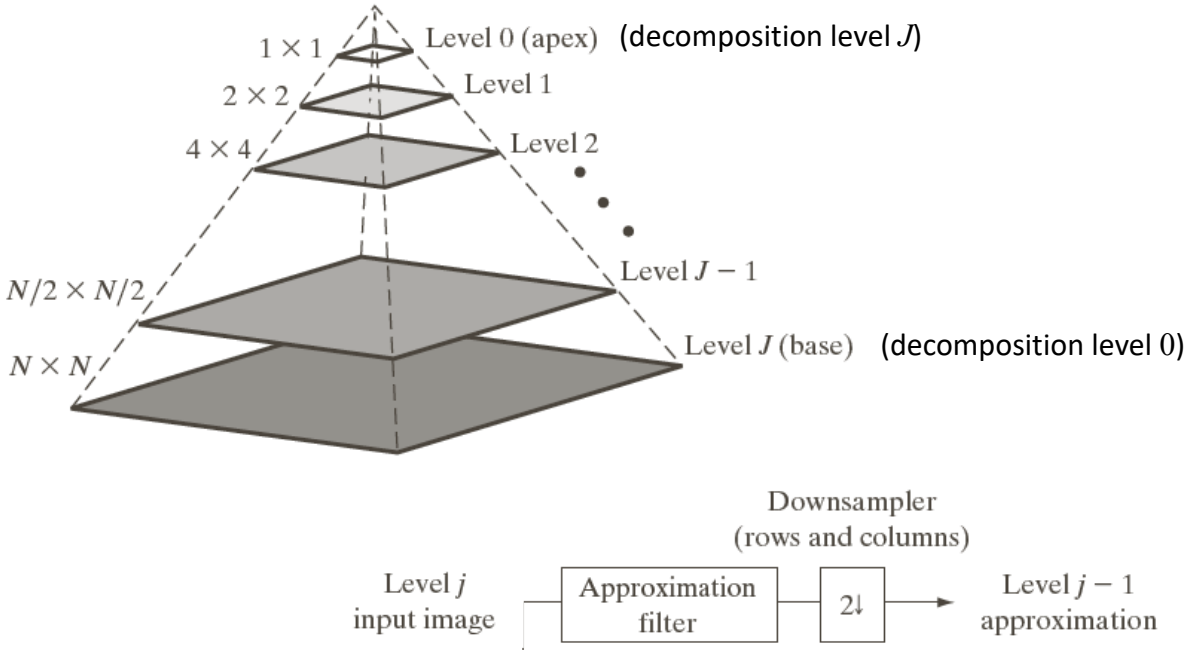


(divided by 16)

- Filter (convolve) each level with this binomial kernel, then sample resulting every other pixel in every other row to produce next pyramid level
  - Convolution of two Gaussians is a Gaussian
  - Convolution is associative
    - Convolving Gaussian with image, subsampling, then convolving with result is equivalent to convolving two Gaussians (though with different variances), then convolving result with image

Image pyramids

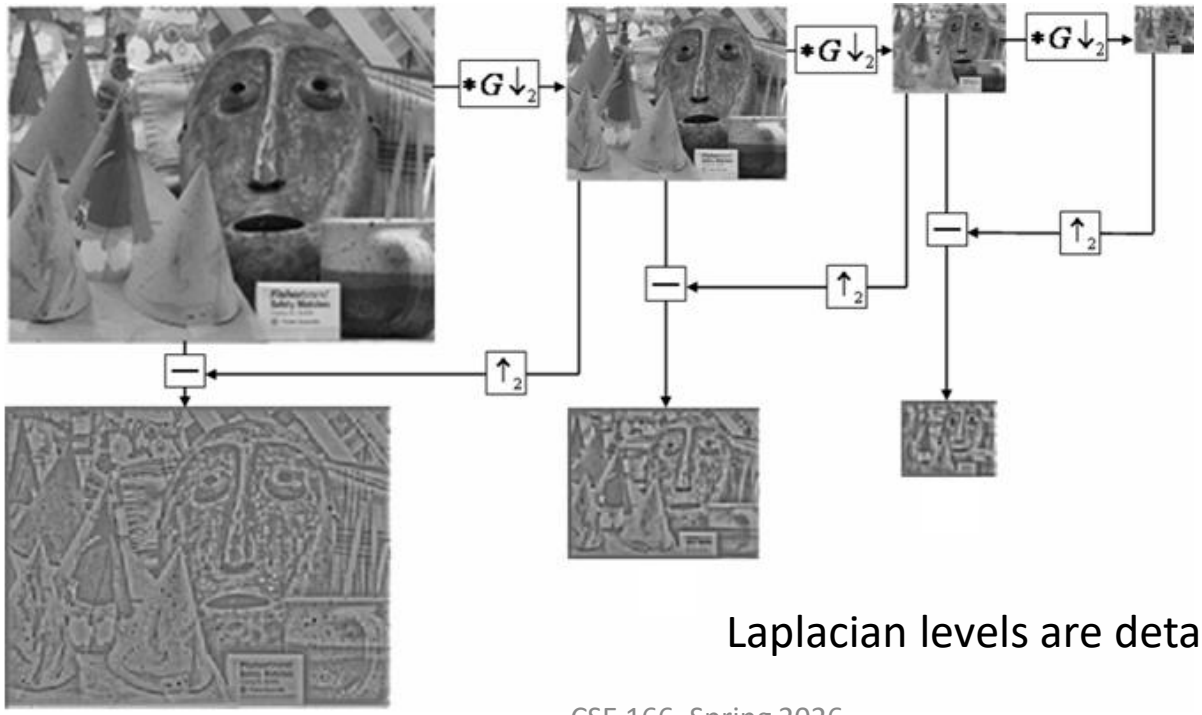
- Gaussian pyramid (approximation pyramid)



Laplacian pyramid

- Full name is Laplacian of Gaussian (LoG)
  - Apply Gaussian (low pass) filter, then apply Laplacian (second derivative) filter
  - Approximates second derivative
  - Approximate with difference of Gaussians (DoG)

Gaussian pyramid

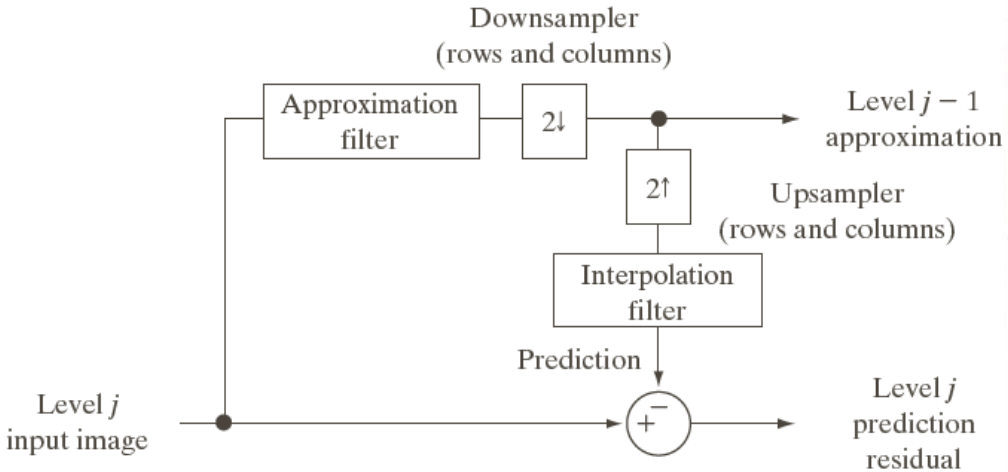


Gaussian levels are approximations

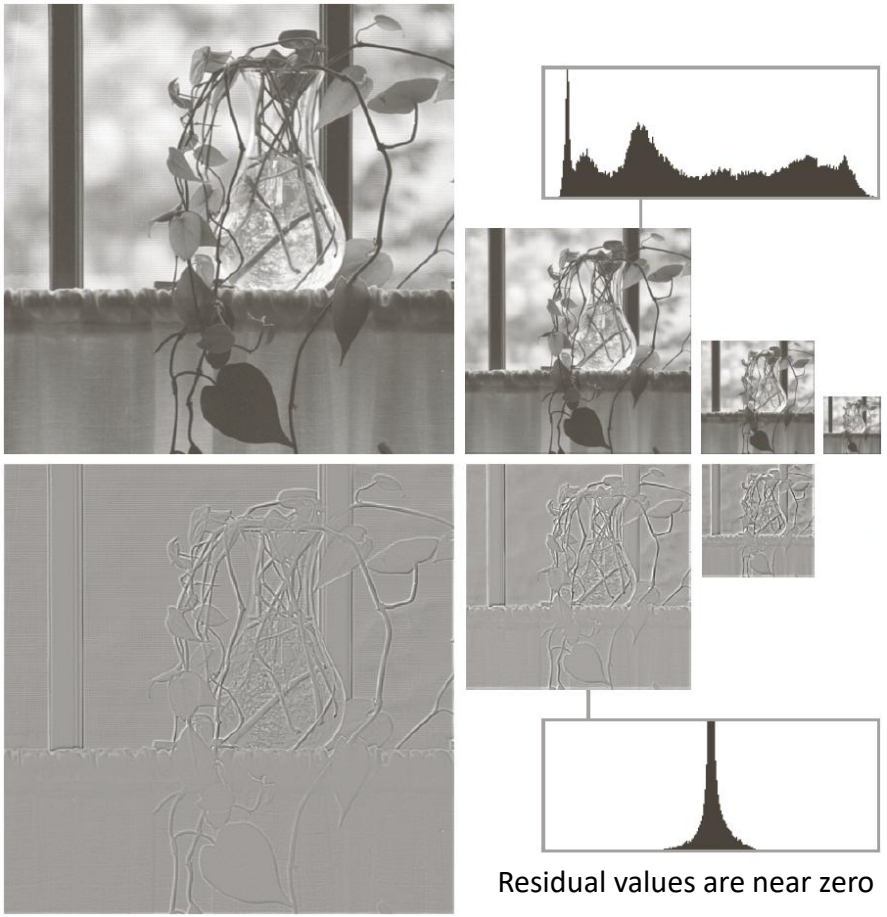
Laplacian pyramid

Laplacian levels are details

Image pyramids



Approximation levels



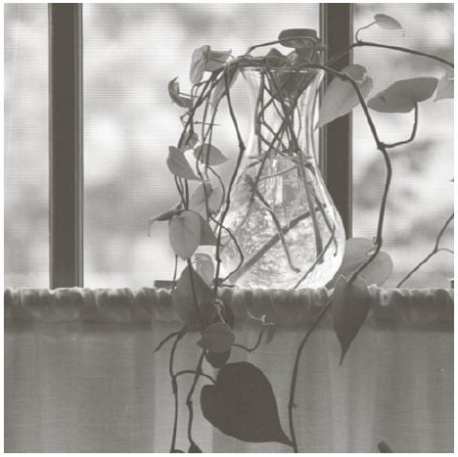
Residual values are near zero

Prediction residual levels
(details)

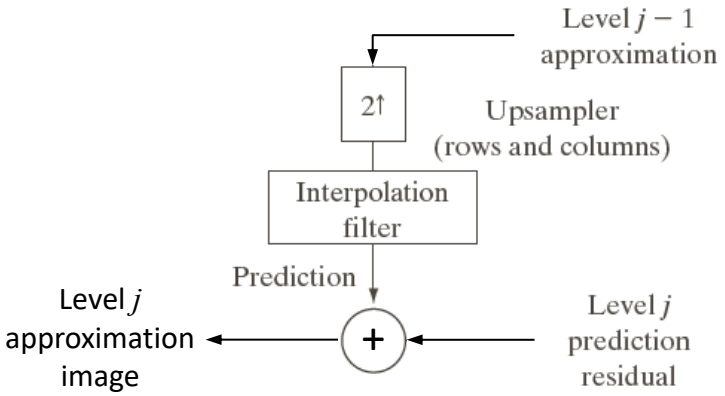
Image pyramids

- Original image can be reconstructed from smallest Gaussian approximation and Laplacian pyramid





Reconstructed original image



Scale space

- Pyramid representation is a predecessor to scale space representation
- Scale space theory is a formal theory for image structures at different scales
  - Image is represented by a one-parameter *family* of Gaussian low pass filtered images
  - A Gaussian filter meets all scale space axioms
    - Linearity, shift invariance, semi-group structure, non-creation of local extrema (zero-crossings), non-enhancement of local extrema, rotational symmetry, and scale invariance
- The *scale parameter* is the variance of the Gaussian filter
  - Note: use border mirror padding on input image when applying filter
- Image details significantly smaller than (two times) the standard deviation (square root of variance) are removed from the image at that scale parameter



Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 0

1024 x 768



Decomposition level 0

Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 1

512 x 384



Decomposition level 1

Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 2

256 x 192



Decomposition level 2

Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 4

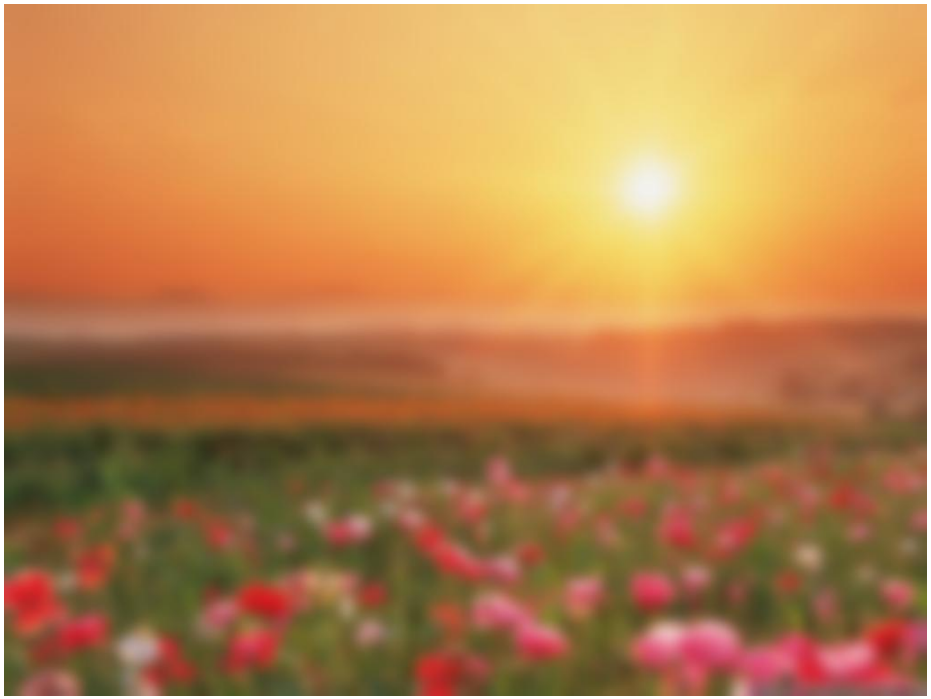
128 x 96



Decomposition level 3

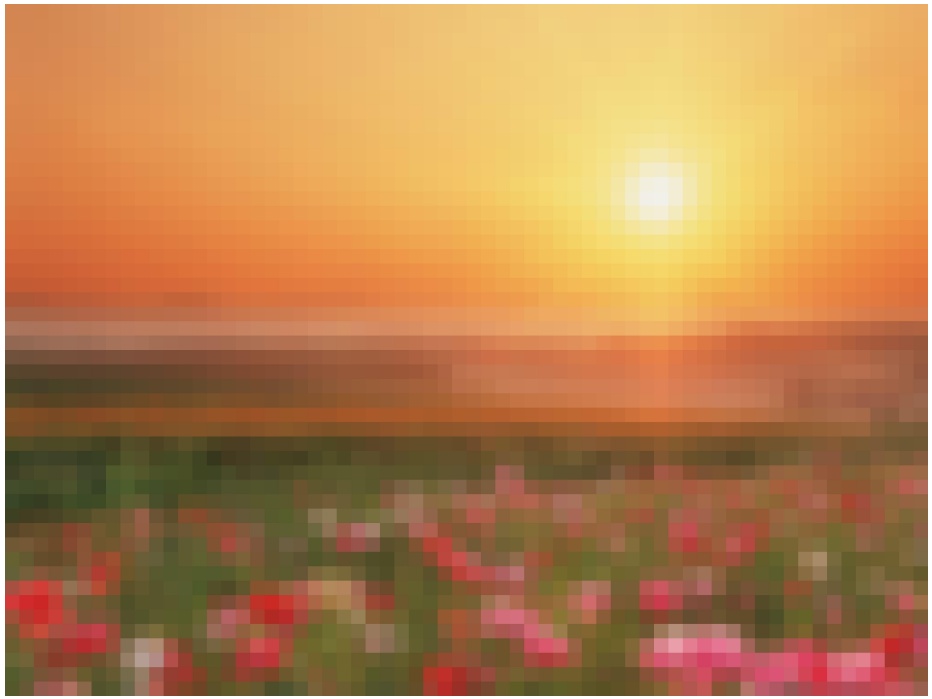
Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 8

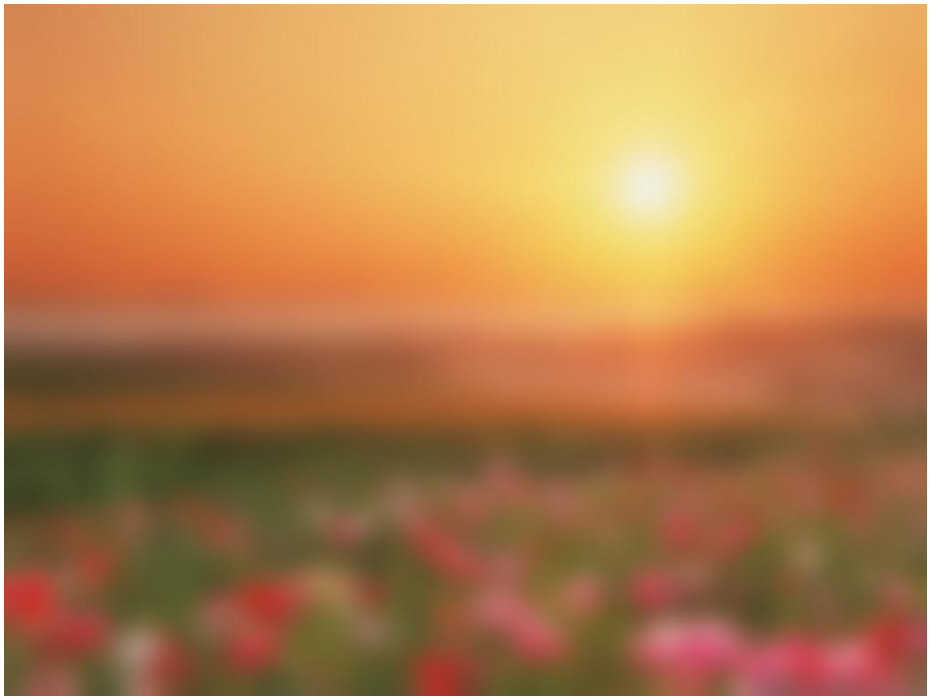
64 x 48



Decomposition level 4

Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 16

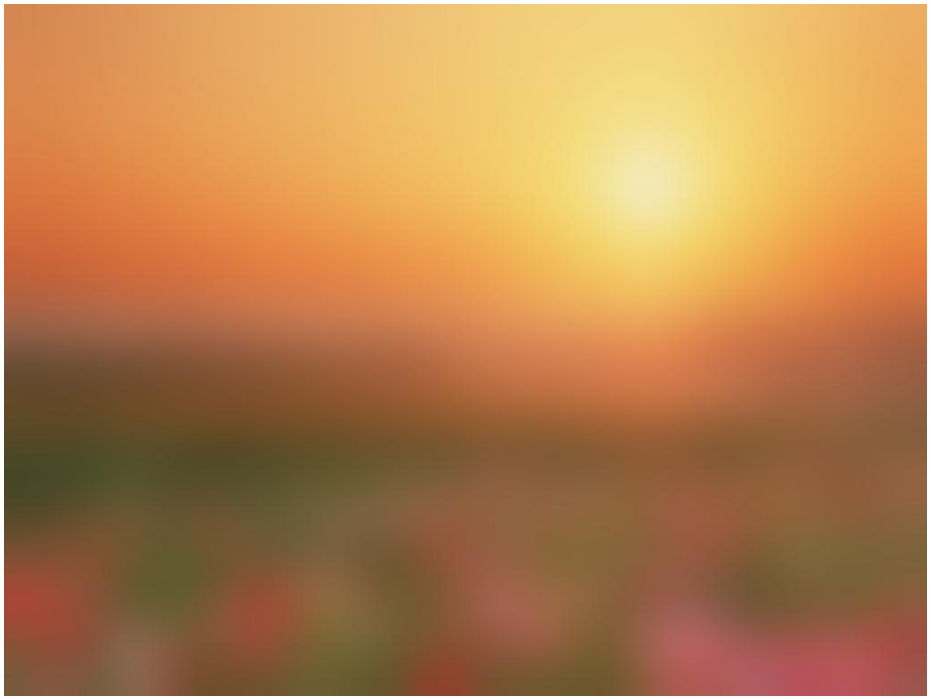
32 x 24



Decomposition level 5

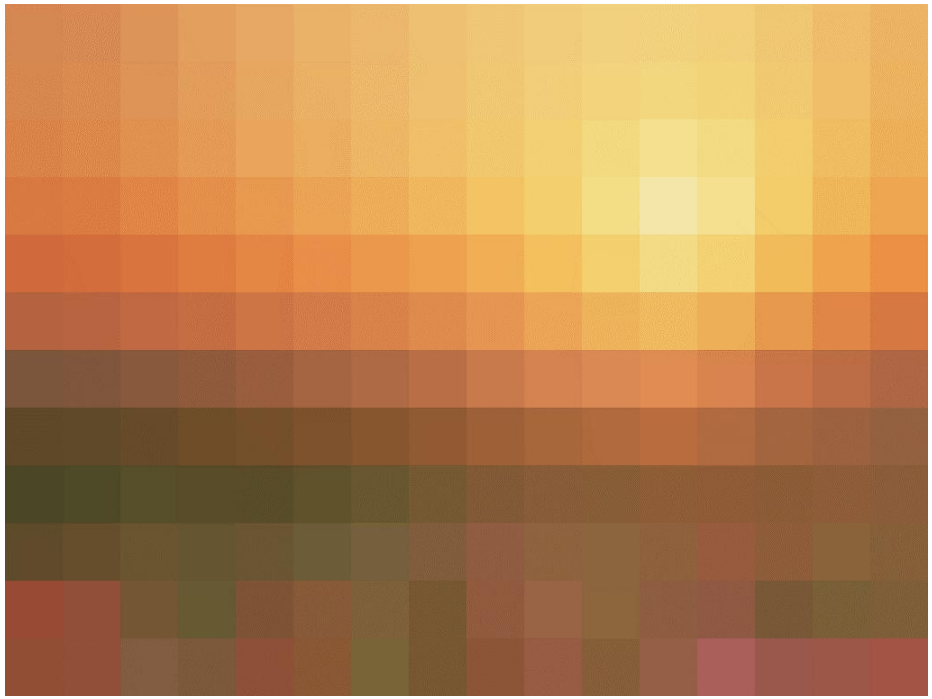
Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 32

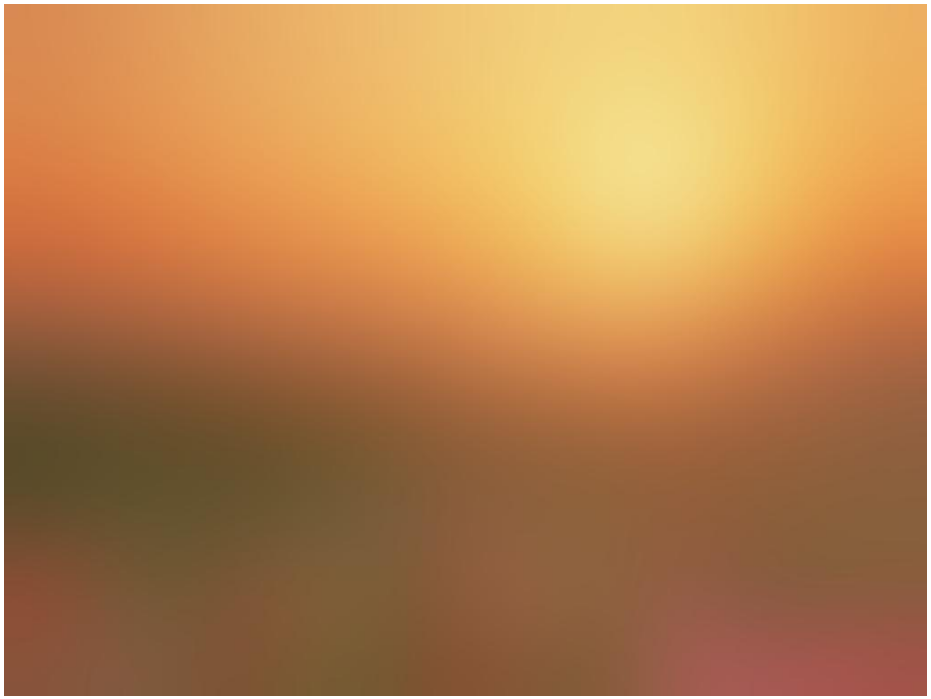
16 x 12



Decomposition level 6

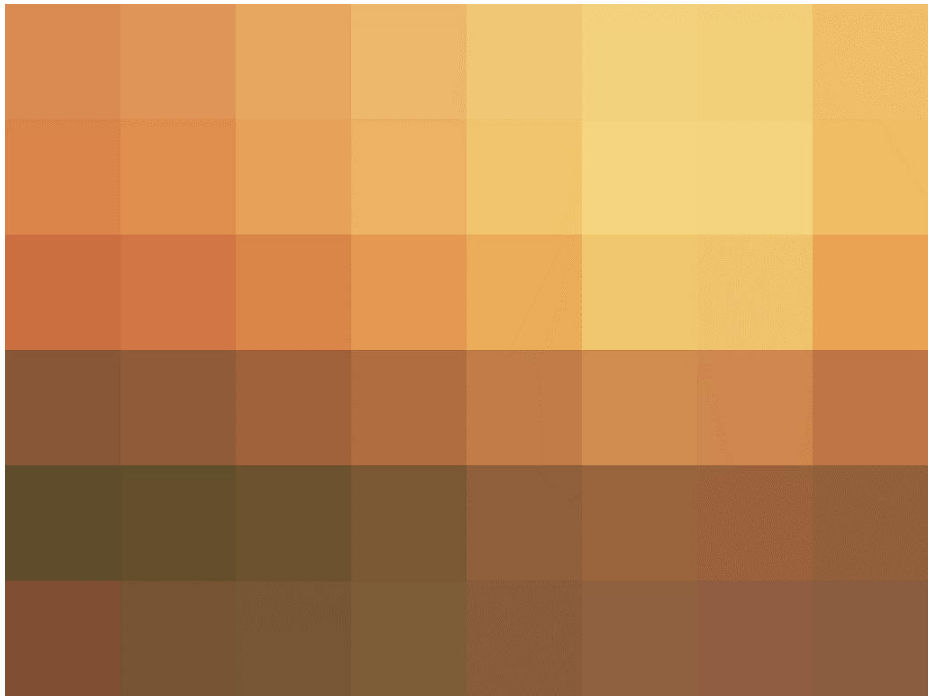
Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 64

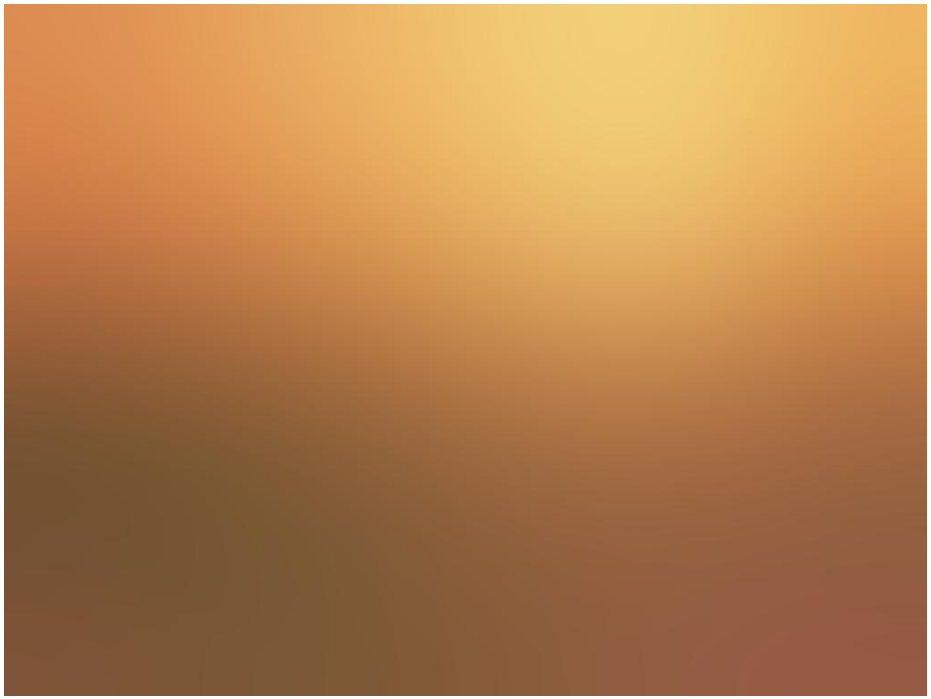
8 x 6



Decomposition level 7

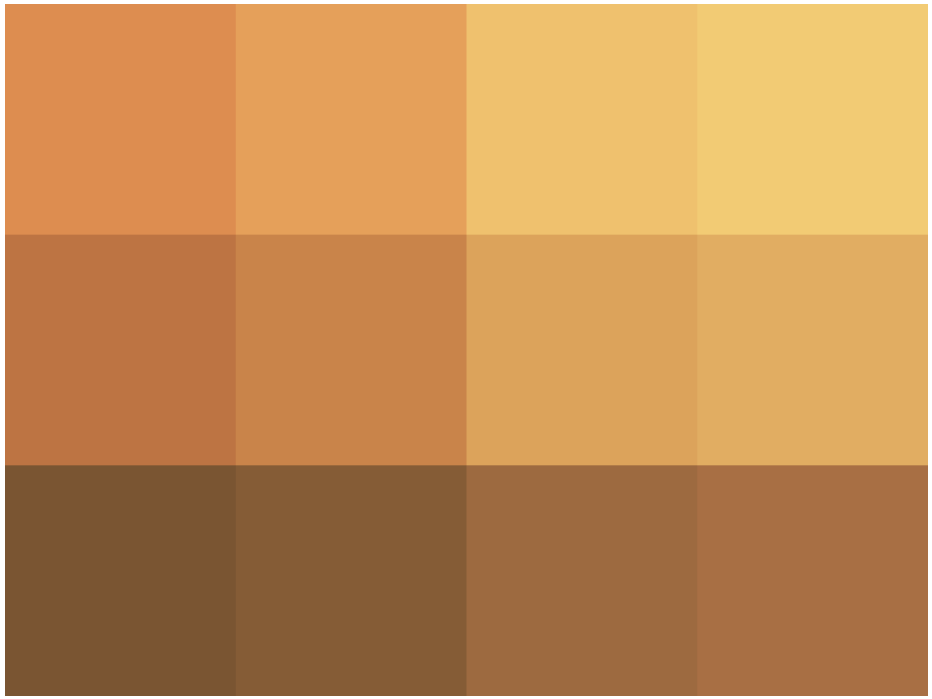
Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 128

4 x 3



Decomposition level 8

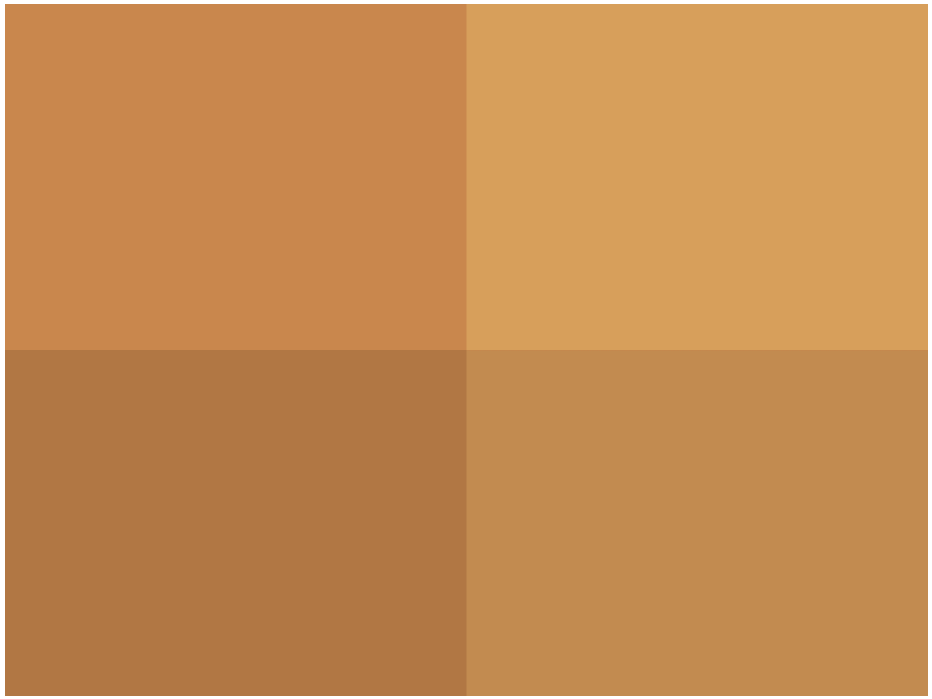
Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 256

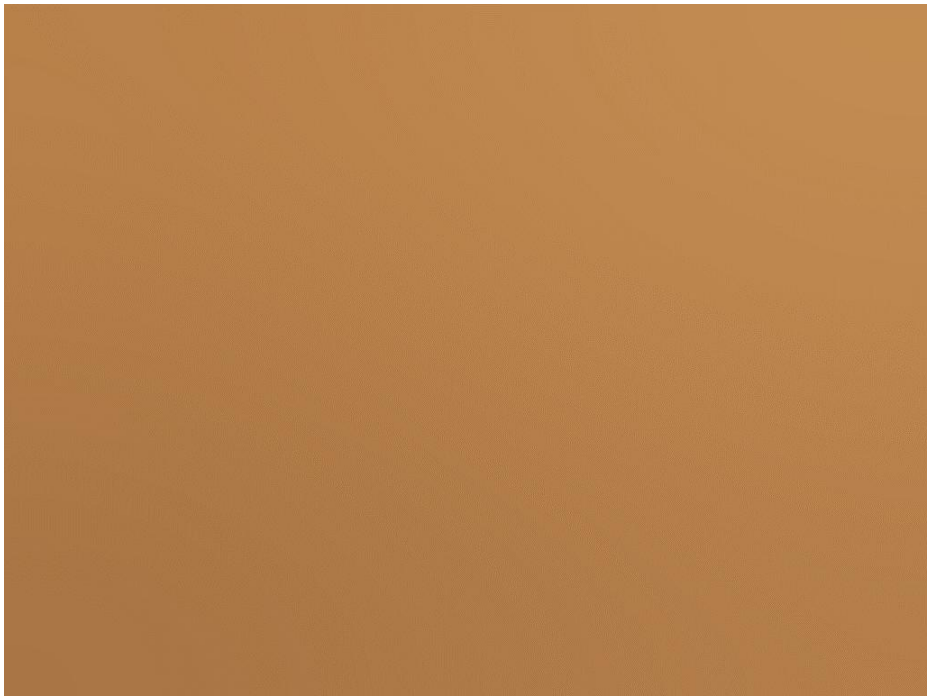
2 x 2



Decomposition level 9

Scale space vs Gaussian pyramid

1024 x 768



Standard deviation = 512

1 x 1



Decomposition level 10

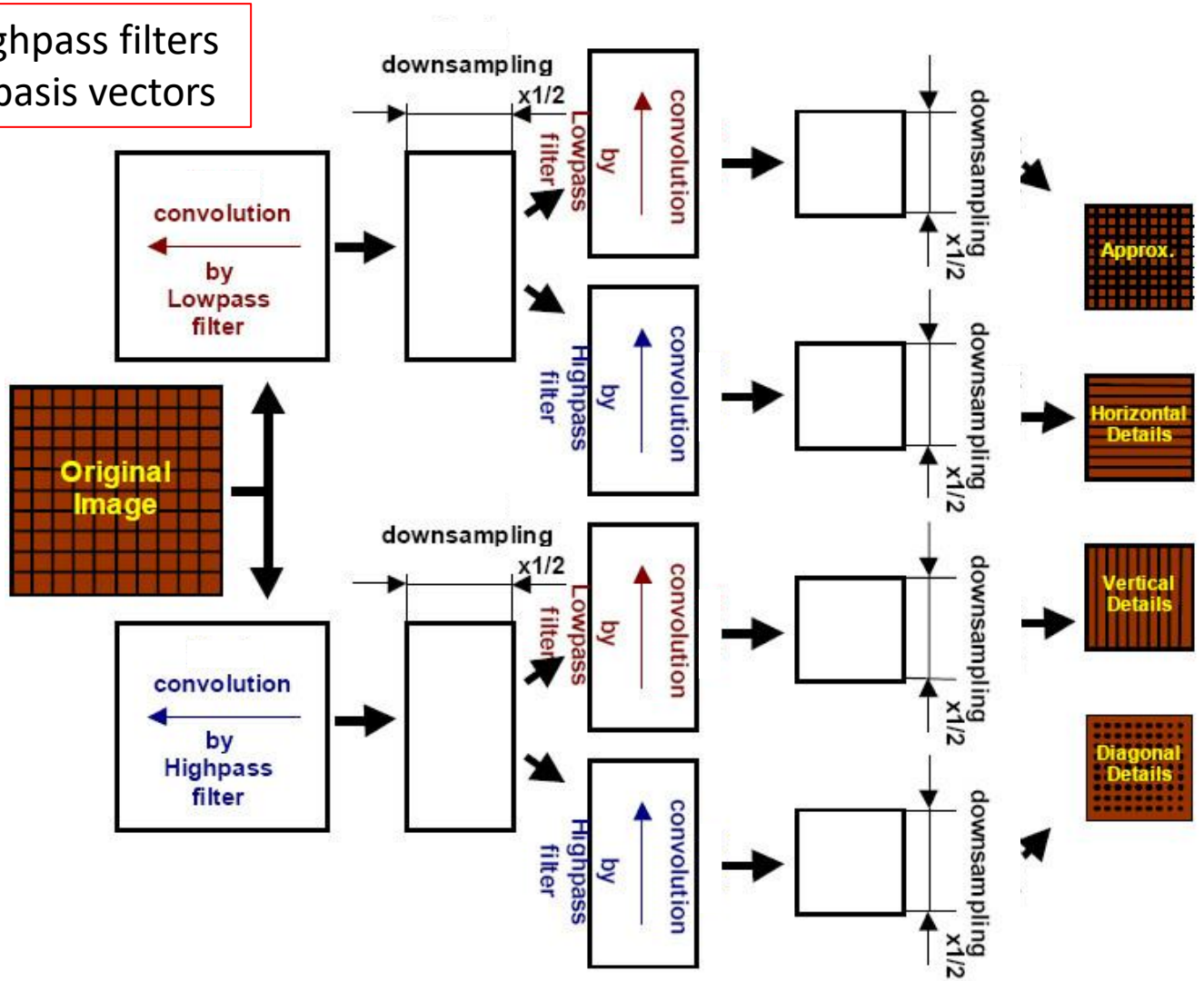
Scale space vs Gaussian pyramid

- Note that Gaussian pyramid (of rate 2) decomposition levels corresponds to Gaussian kernel with standard deviations of 0, 1, 2, 4, 8, 16, 32, 64, 128, 256, 512
  - Rate of Gaussian pyramid dictates standard deviations of decomposition levels
  - Scale space does not have this imposition; all standard deviations available
- In scale space, since all images are the same size, features are precisely located in the coordinates of the original image
  - Points in levels of Gaussian pyramid must be scaled up to coordinates of the original image, which is imprecise

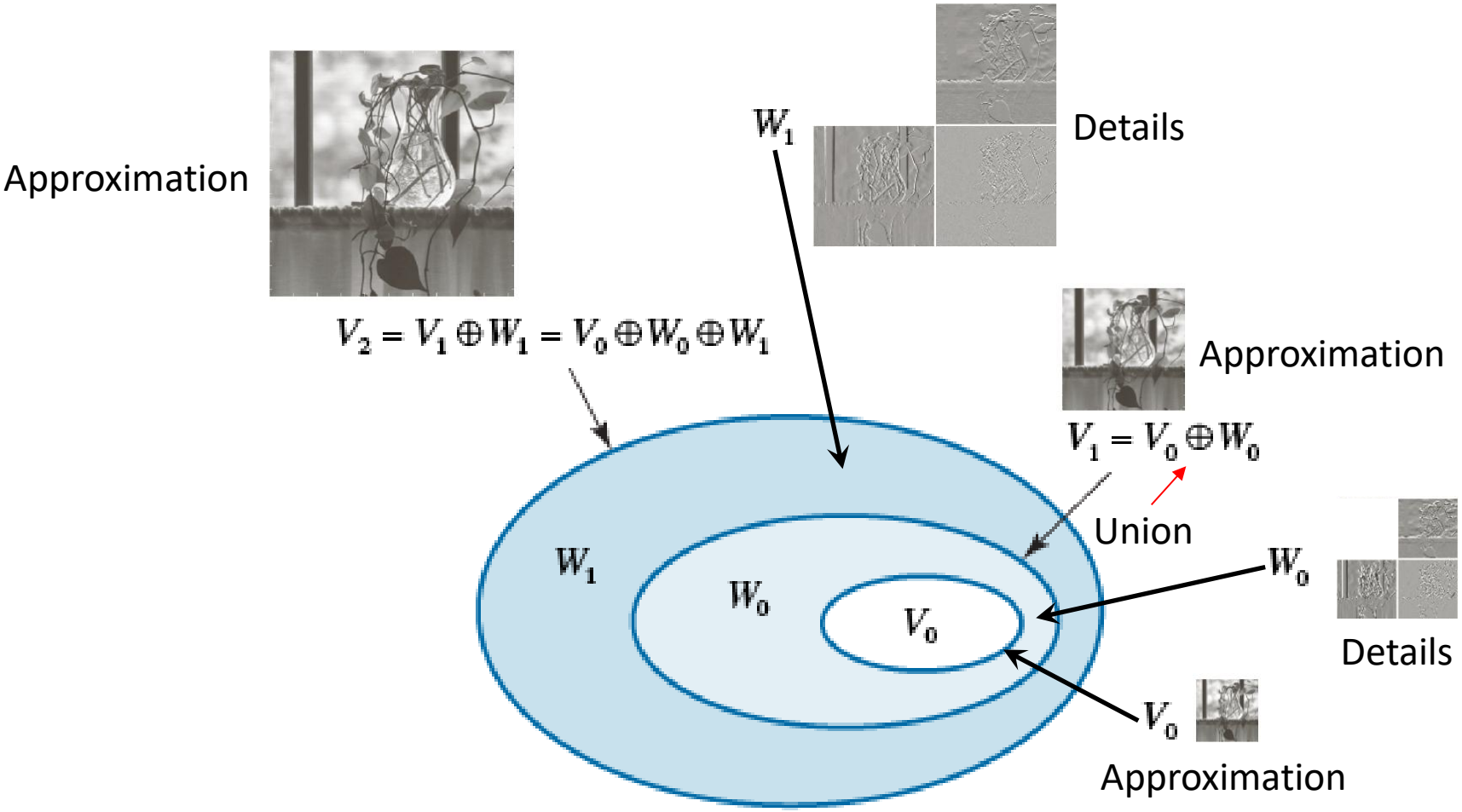
The Wavelet Transform

Wavelet decomposition

Lowpass and highpass filters are orthogonal basis vectors



Relationship between approximations and details



Wavelet transforms

- A scaling function is used to create a series of approximations of a function or image, each differing by a factor of 2 in resolution from its nearest neighboring approximations
- Wavelet functions (wavelets) are then used to encode the differences between adjacent approximations
- The discrete wavelet transform (DWT) uses those wavelets, together with a single scaling function, to represent a function or image as a linear combination of the wavelets and scaling function

Scaling functions

The set of basis functions composed of integer translations and binary scalings of the real, square integrable **father scaling function** $\phi(x)$

Square-integrable

$$\{\phi_{j,k}(x)\}, \text{ where } \phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k) \quad \forall\, j,k \in \mathbb{Z} \text{ and } \phi(x) \in L^2(\mathbb{R})$$

(i.e., $\int_{-\infty}^{\infty} |\phi(x)|^2 \mathrm{d}x$ is finite)

Integer

k determines translation

j determines scale

Binary

Determines amplitude

If j is restricted such that $j = j_0$, then

$$V_{j_0} = \{\phi_{j_0,k}\} \quad \forall\, k \in \mathbb{Z}$$

Basis of the function space spanned by scaling functions

is the basis of the function space spanned by $\phi_{j,k}(x)$ for $j = j_0$ and $k \in \mathbb{Z}$.
Increasing j_0 increases the size of V_{j_0}, allowing functions with smaller variations (i.e., finer details) to be included in the space.

Scaling functions, multiresolution analysis

1. The scaling function is orthogonal to its integer translations
2. The function spaces spanned by the scaling function at low scales are nested within those spanned at higher scales
   $V_{-\infty} \subset \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \subset V_{\infty}$
3. The only function representable at every scale (all V_j) is $f(x) = 0$
4. All measureable, square-integrable functions can be represented as $j \to \infty$

Scaling functions, multiresolution analysis

$\phi(x)$ is a linear combination of double-resolution copies of itself:

$$\phi(x) = \sum_{k \in \mathbb{Z}} h_\phi(k) \sqrt{2} \phi(2x - k) \quad \text{refinement or dilation equation}$$

$h_\phi(k)$ are **scaling function coefficients**. For orthogonal scaling functions

$$h_\phi(k) = \langle \phi(x), \sqrt{2}\phi(2x - k) \rangle$$

Wavelet functions (wavelets)

Given the father scaling function $\phi(x)$, there exists a **mother wavelet function** $\psi(x)$ whose integer translations and binary scalings

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k) \quad \forall\ j,k \in \mathbb{Z}$$

span the difference between any two adjacent scaling spaces. Similarly, if j is restricted such that $j = j_0$, then

$$\{\psi_{j_0,k}\} \quad \forall k \in \mathbb{Z} \text{ and } V_{j_0+1} = V_{j_0} \oplus W_{j_0}$$

Basis of the function space spanned by wavelet functions

The ortogonal compliment of V_{j_0} in V_{j_0+1} is W_{j_0}. All members of V_{j_0} are orthogonal to the members of W_{j_0}.

$$\langle \phi_{j_0,k}(x), \psi_{j_0,l}(x) \rangle = 0 \text{ for } k \neq l$$

Wavelet functions, multiresolution analysis

$\psi(x)$ is a weighted sum of shifted, double-resolution scaling functions:

$$\psi(x) = \sum_k h_\psi(k)\sqrt{2}\phi(2x - k)$$

$h_\psi(k)$ are **wavelet function coefficients**

$$h_\psi(k) = (-1)^k h_\phi(1 - k)$$

Multiresolution analysis

- Refinement (or dilation) equation

$$\phi(x) = \sum_{k \in \mathbb{Z}} h_\phi(k) \sqrt{2} \phi(2x - k)$$

  where $h_\phi(k)$ are scaling function coefficients

- And

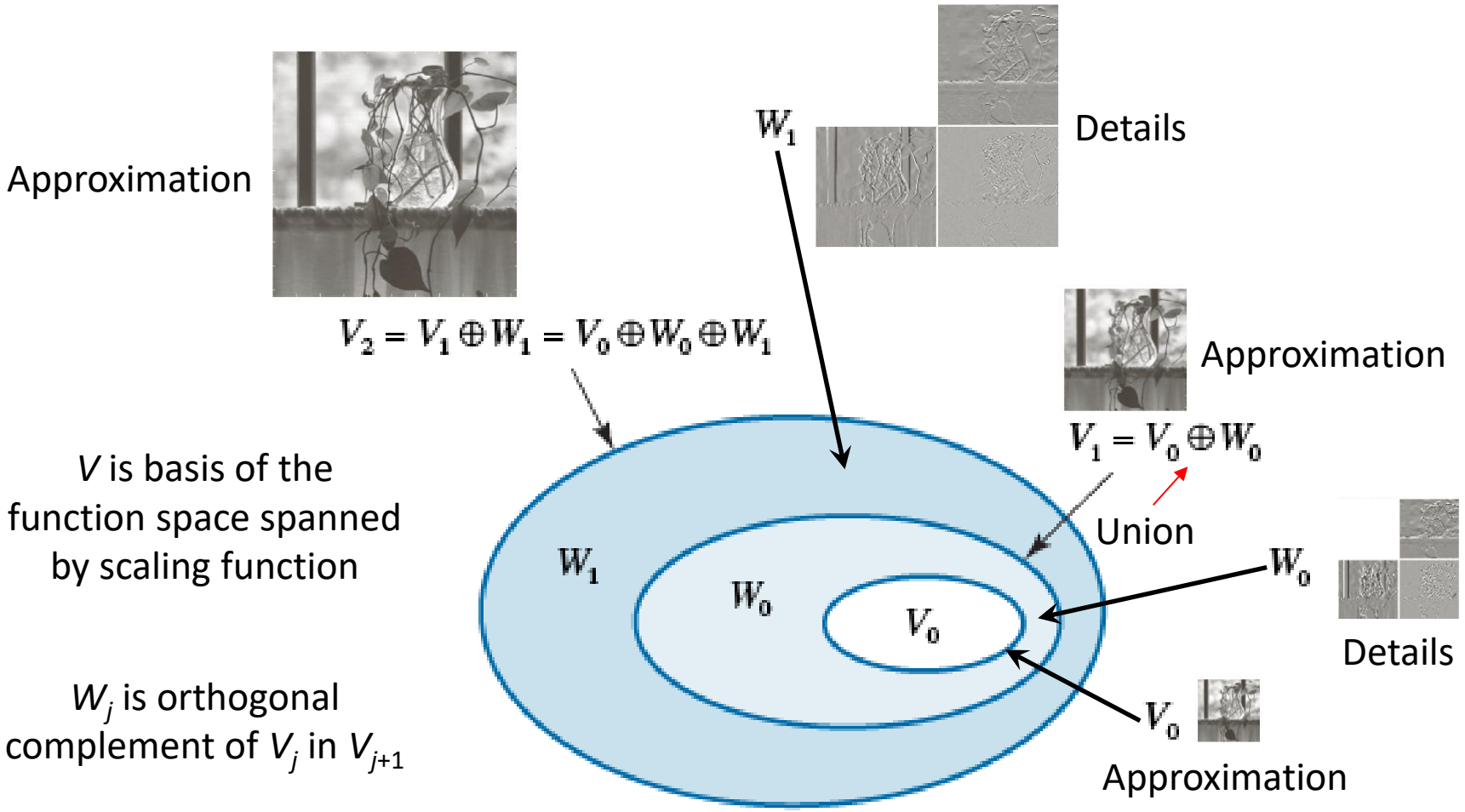
$$\psi(x) = \sum_{k} h_\psi(k) \sqrt{2} \psi(2x - k)$$

  where $h_\psi(k)$ are wavelet function coefficients

- Relationship

$$h_\psi(k) = (-1)^k h_\phi(1 - k)$$

Relationship between scaling and wavelet function spaces



V is basis of the function space spanned by scaling function

W_j is orthogonal complement of V_j in V_{j+1}

1D discrete wavelet transform

for real signals

Typically
$j_0 = 0$ and N is a power of 2
$j = 0, 1, \ldots, J\text{-}1$
$k = 0, 1, \ldots, 2^j\text{-}1$

- Forward

$$T_\phi(j_0, k) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\phi_{j_0,k}(x)$$ Approximation

$$T_\psi(j, k) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\psi_{j,k}(x)$$ Details

- Inverse

$$f(x) = \frac{1}{\sqrt{N}} \sum_{k=0}^{2^{j_0}-1} T_\phi(j_0, k)\phi_{j_0,k}(x) + \frac{1}{\sqrt{N}} \sum_{j=j_0}^{J-1} \sum_{k=0}^{2^j-1} T_\psi(j, k)\psi_{j,k}(x)$$

2D discrete wavelet transform

for real signals

- Forward

$$T_\phi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \phi_{j_0,m,n}(x, y)$$

Approximation

$$T_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \psi_{j,m,n}^i(x, y),\ i = \{H, V, D\}$$

Details

where

$$\phi_{j,m,n}(x, y) = 2^{j/2} \phi(2^j x - m, 2^j y - n)$$

$$\psi_{j,m,n}^i(x, y) = 2^{j/2} \psi^i(2^j x - m, 2^j y - n),\ i = \{H, V, D\}$$

$$\psi^H(x, y) = \psi(x)\phi(y)$$

$$\psi^V(x, y) = \phi(x)\psi(y)$$

$$\psi^D(x, y) = \psi(x)\psi(y)$$

Directional wavelets

2D discrete wavelet transform

for real signals

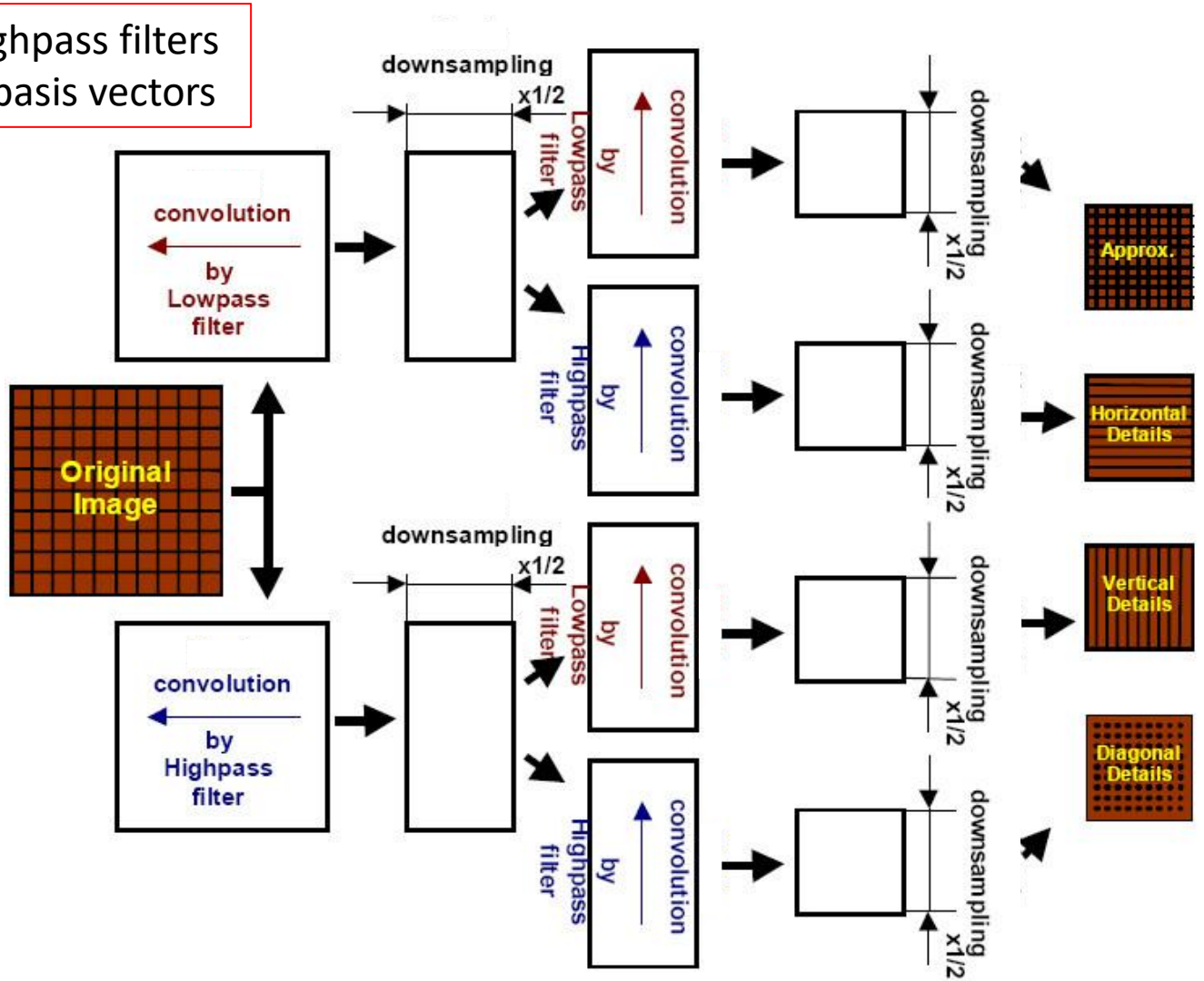
- Inverse

$$f(x,y) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{2^{j_0}-1} \sum_{n=0}^{2^{j_0}-1} T_\phi(j_0,m,n)\phi_{j_0,m,n}(x,y)$$

$$+ \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{J-1} \sum_{m=0}^{2^j-1} \sum_{n=0}^{2^j-1} T_\psi^i(j,m,n)\psi_{j,m,n}^i(x,y)$$

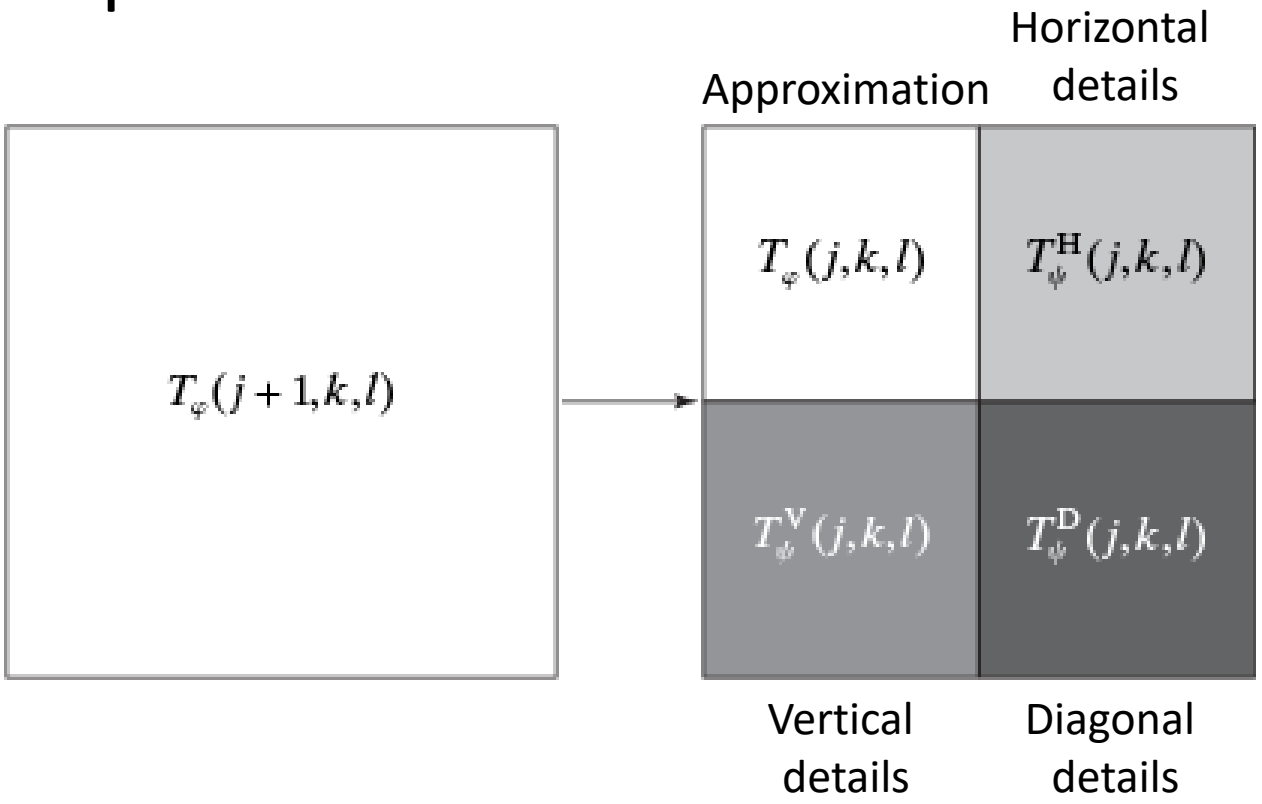
Wavelet decomposition

Lowpass and highpass filters are orthogonal basis vectors



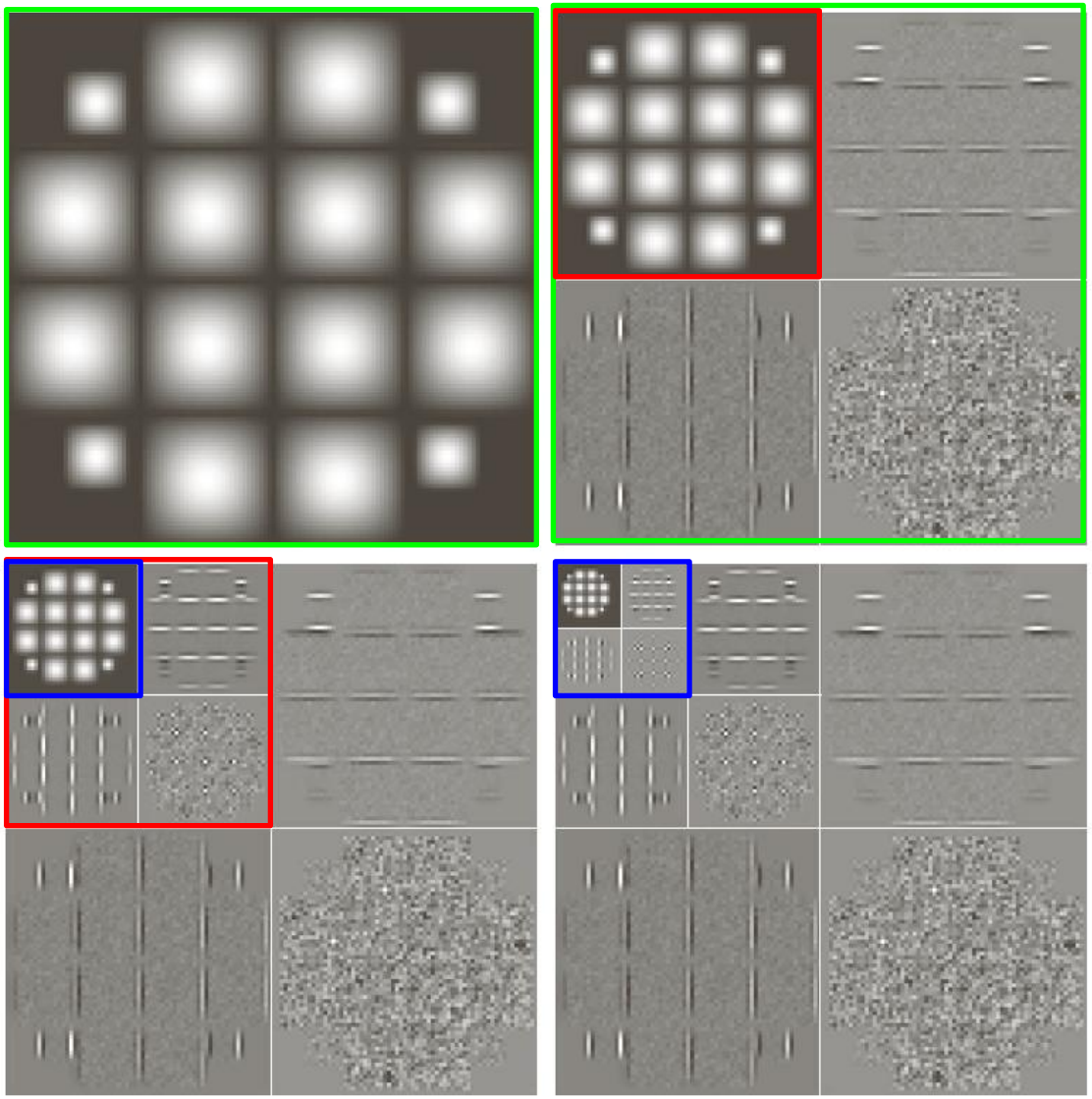
2D discrete wavelet transform

- Decomposition



2D discrete wavelet transform

3-level wavelet decomposition



2D wavelet decomposition

Example: 256×256 input image, $J = 8$

Dimensions	j	Decomposition level	$2^{\text{decomposition level}}$
256×256	8	0	1
128×128	7	1	2
64×64	6	2	4
32×32	5	3	8
16×16	4	4	16
8×8	3	5	32
4×4	2	6	64
2×2	1	7	128
1×1	0	8	256

Input image

Pixel values are scaled by this

Fourier transform vs wavelet transform

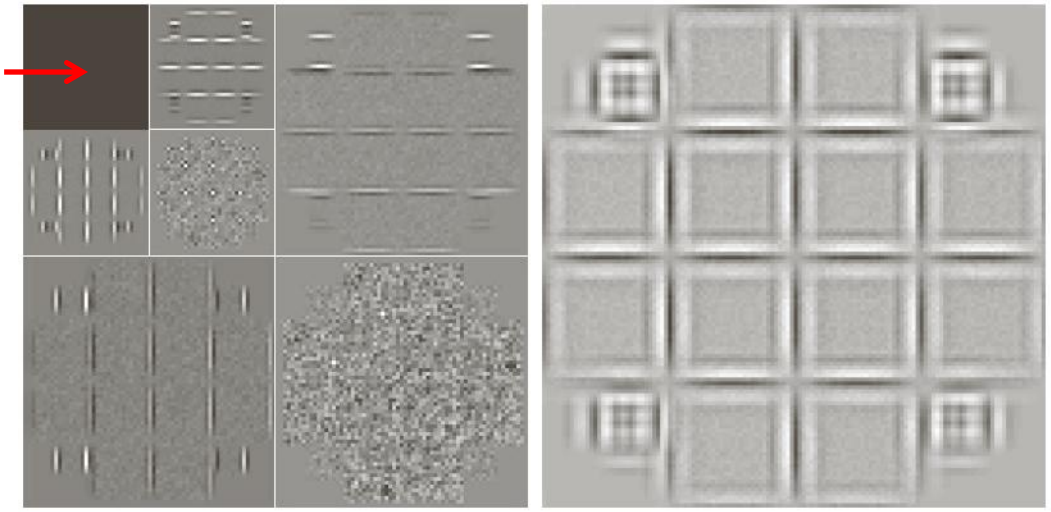
- Fourier
  - Localized in frequency domain
- Wavelet
  - Localized in scale domain
    - or space and frequency (bands), or space and scale

Wavelets in image processing

1. Wavelet transform
2. Alter transform
3. Inverse wavelet transform

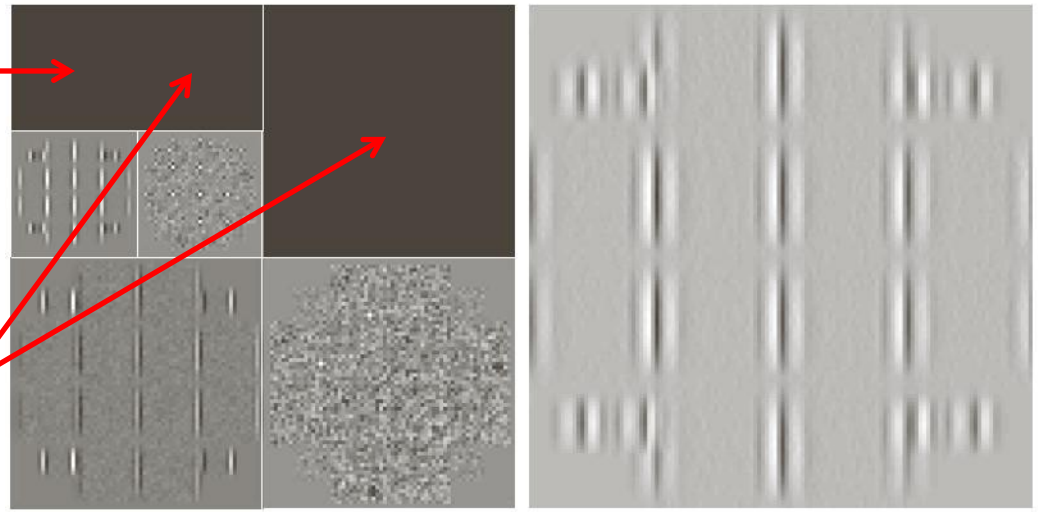
Wavelet-based edge detection

Zero lowest scale approximation



Edges

Zero lowest scale approximation **and** zero horizontal details



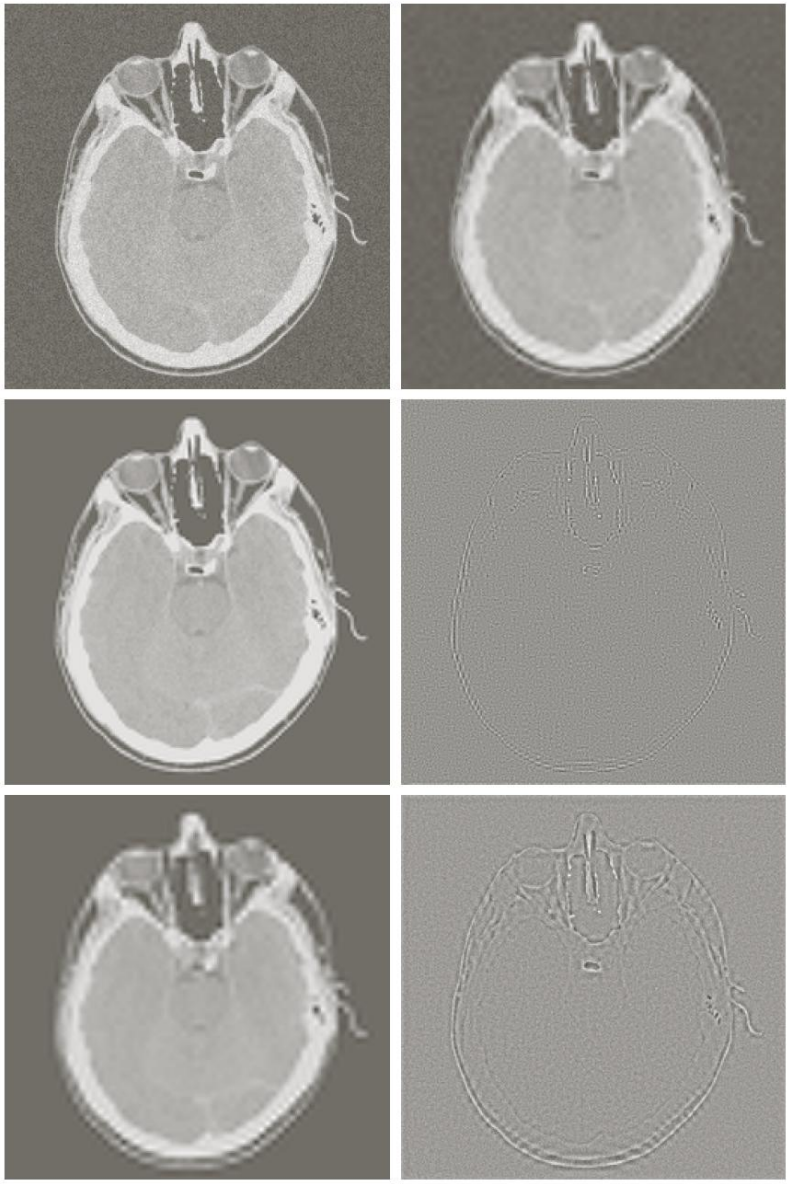
Vertical edges

Wavelet-based noise removal

Noisy image

Zero highest resolution details

Zero details for all levels



Threshold details

a b
c d
e f

FIGURE 7.28 Modifying a DWT for noise removal: (a) a noisy CT of a human head; (b), (c) and (e) various reconstructions after thresholding the detail coefficients; (d) and (f) the information removed during the reconstruction of (c) and (e). (Original image courtesy Vanderbilt University Medical Center.)

Next Lecture

- Image compression
- Reading
  - Chapter 8: Image Compression and Watermarking
    - Sections 8.1, 8.9, 8.10, and 8.12