上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

IVM
Image, Video, and Multimedia Communications Laboratory

Digital Image Processing

Hongkai Xiong
熊红凯
http://ivm.sjtu.edu.cn

电子工程系
上海交通大学
2016

Topic

- Background
- Basic of multiresolution analysis
- A tour of wavelets

Background

- Objects are formed by connected regions of similar texture and intensity levels.
- If the objects are small in size or low in contrast, we normally examine them at high resolutions; if they are large in size or high in contrast, a coarse view is required.

Multiresolution processing

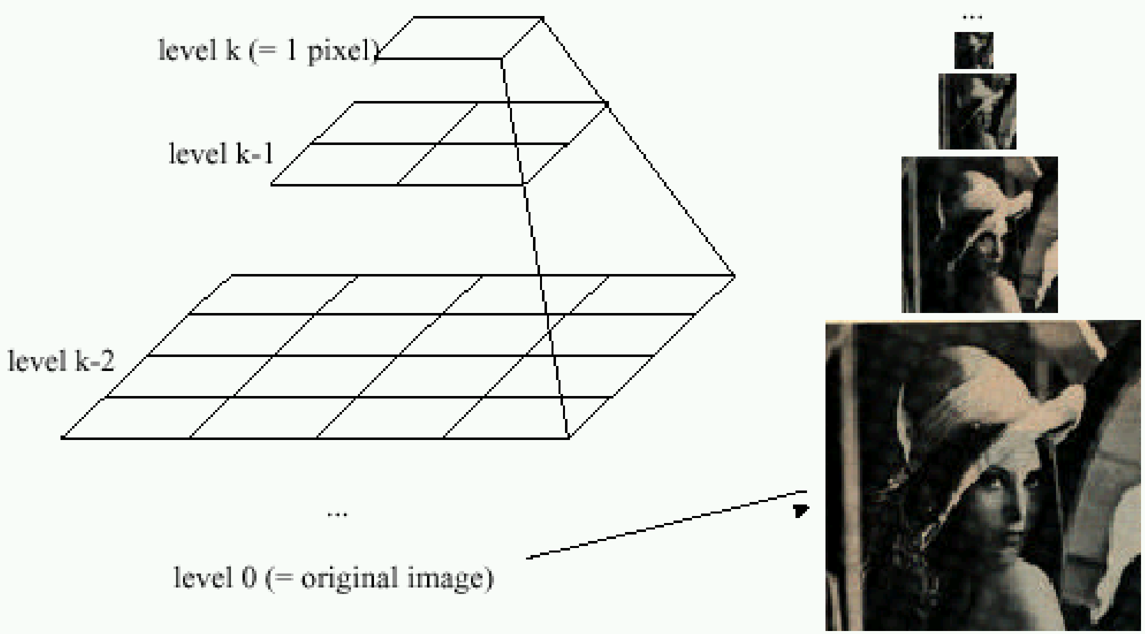
Background

- Image Pyramids

  An image pyramid is a collection of decreasing resolution images arranged in the shape of a pyramid

  The base of the pyramid contains a high-resolution representation of the image being processed; the apex contains a low-resolution approximation.

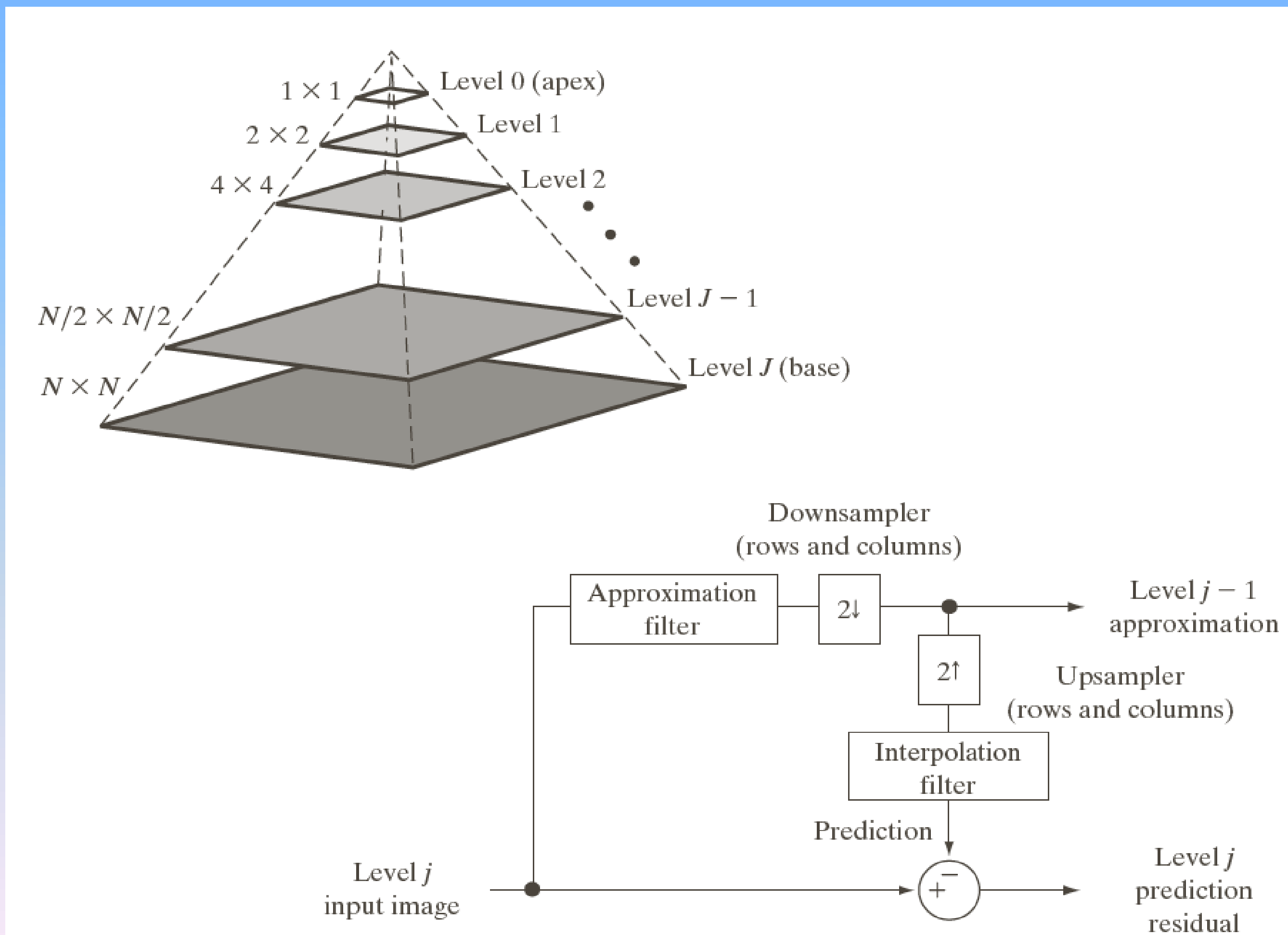
Idea: Represent NxN image as a “pyramid” of $1\text{x}1$, $2\text{x}2$, $4\text{x}4$,…, $2^k\text{x}2^k$ images (assuming $N=2^k$)



Known as a **Gaussian Pyramid** [Burt and Adelson, 1983]

- In computer graphics, a *mip map* [Williams, 1983]
- A precursor to *wavelet transform*

Image Pyramids



a b

FIGURE 7.2
(a) An image pyramid. (b) A simple system for creating approximation and prediction residual pyramids.

Image Pyramids

- Gaussian Pyramid
  - Approximation pyramid
- Laplacian Pyramid
  - Prediction residual pyramid

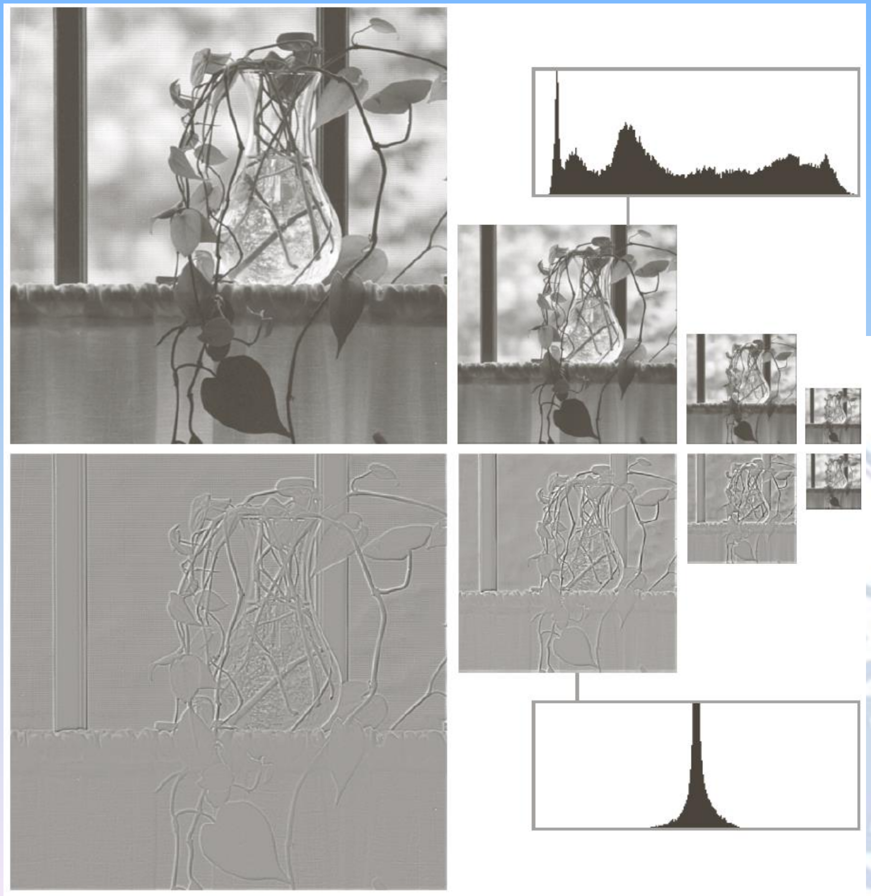


FIGURE 7.3 Two image pyramids and their histograms: (a) an approximation pyramid; (b) a prediction residual pyramid.

Gaussian pyramid construction

Step:

Repeat {

- Filter
- Subsample

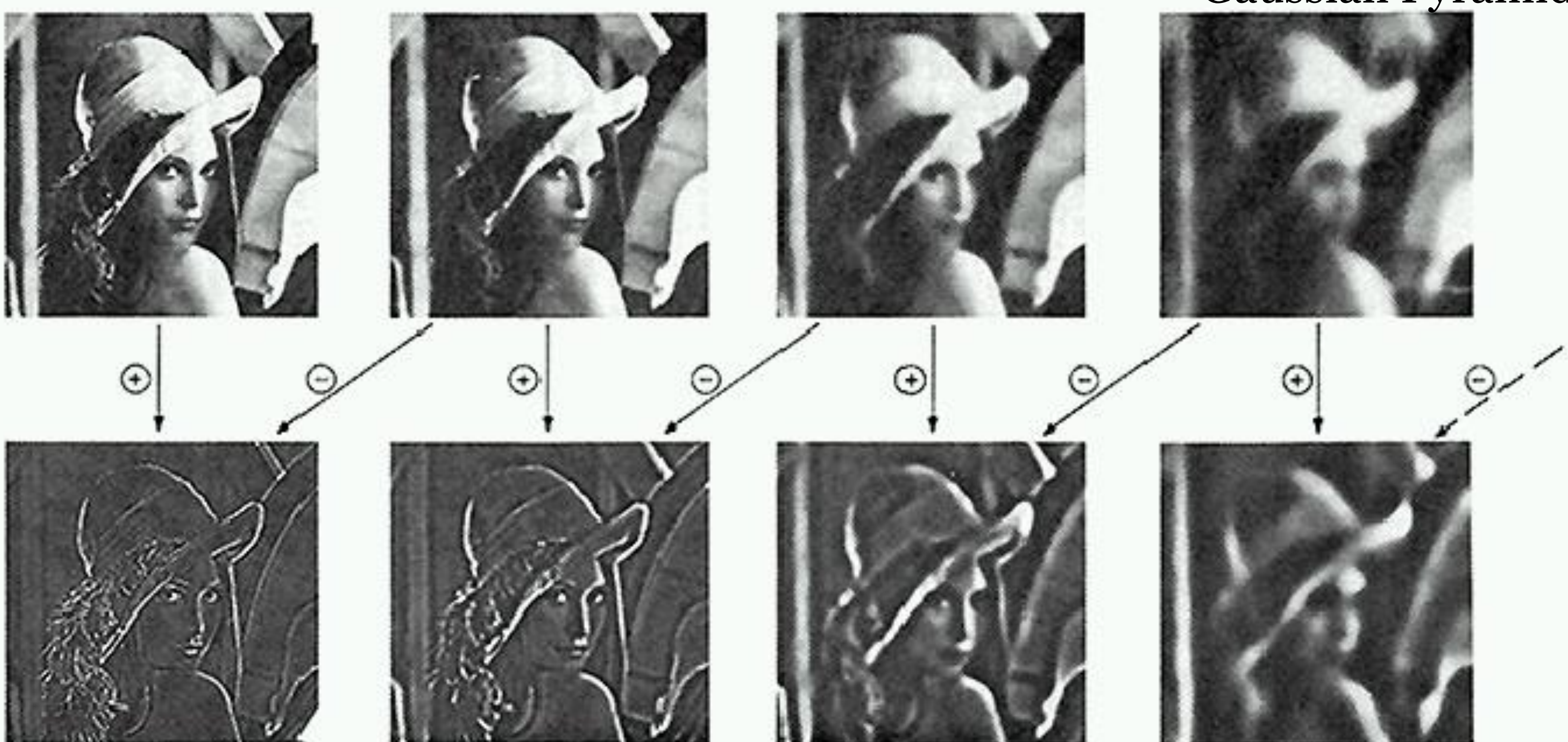
} Until minimum resolution reached

- can specify desired number of levels (e.g., 3-level pyramid)

The whole pyramid is only 4/3 the size of the original image!

Laplacian pyramid construction

- Created from Gaussian pyramid by subtraction



What are they good for?

- Improve Search
  - Search over translations
    - Like homework
    - Classic coarse-to-fine strategy
  - Search over scale
    - Template matching
    - E.g. find a face at different scales
- Precomputation
  - Need to access image at different blur levels
  - Useful for texture mapping at different resolutions
- Image Processing
  - Editing frequency bands separately
  - E.g. image blending… next time!

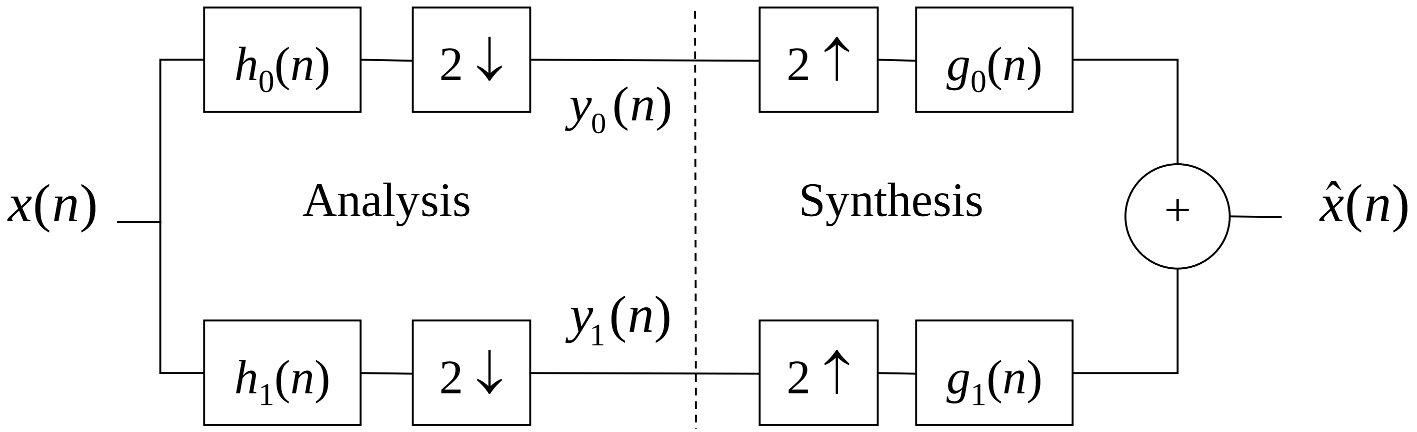
Background

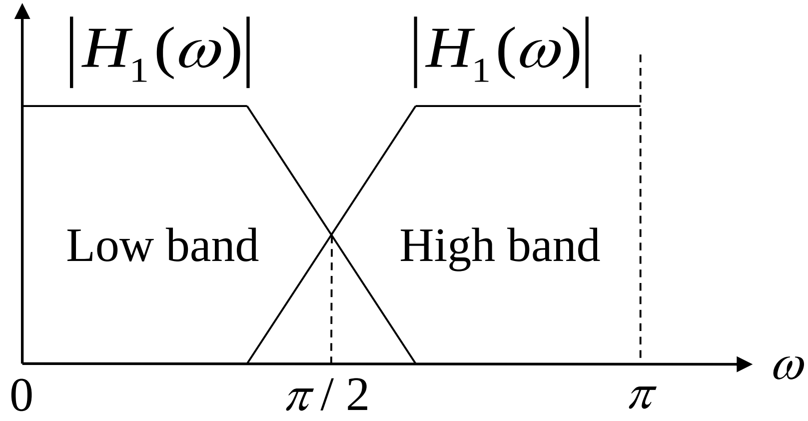
- Subband Coding

  Another important imaging technique with ties to multiresolution analysis is subband coding

  In subband coding, an image is decomposed into a set of bandlimited components, called subbands

Subband Coding





Subband Coding

- For perfect reconstruction, the impulse responses of the synthesis and analysis filters must be related in one of the following two ways:

$$g_0(n) = (-1)^n h_1(n)$$
$$g_1(n) = (-1)^{n+1} h_0(n)$$

- or

$$g_0(n) = (-1)^{n+1} h_1(n)$$
$$g_1(n) = (-1)^n h_0(n)$$

Subband Coding

- The impulse responses of the synthesis and analysis filters can be shown to satisfy the following biorthogonality condition

$$\langle h_i(2n-k), g_j(k) \rangle = \delta(i-j)\delta(n), \quad i, j = \{0,1\}$$

- Of special interest in subband coding– and in the development of the fast wavelet transform– are filters that move beyond biorthogonality and require

$$\langle g_i(n), g_j(n+2m) \rangle = \delta(i-j)\delta(m), \quad i, j = \{0,1\}$$

which defines orthonormality for perfect reconstruction filter banks.

Subband Coding

- Orthonormal filters can be shown to satisfy the following two conditions

$$g_1(n) = (-1)^n g_0(K_{even} - 1 - n)$$
$$h_i(n) = g_i(K_{even} - 1 - n), \quad i = \{0,1\}$$

Subband Coding

- 1-D orthonormal and biorthogonal filters can be used as 2-D separable filters for the processing of images.

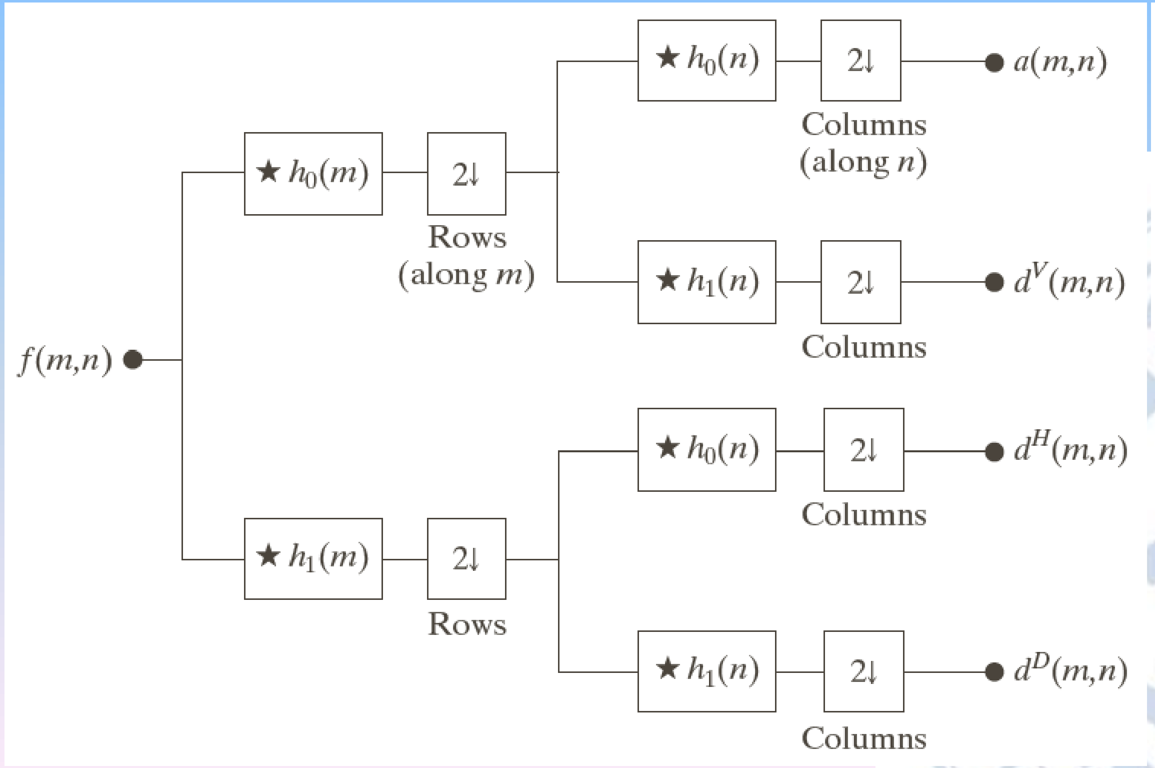


FIGURE 7.7 A two-dimensional, four-band filter bank for subband image coding.

Subband Coding

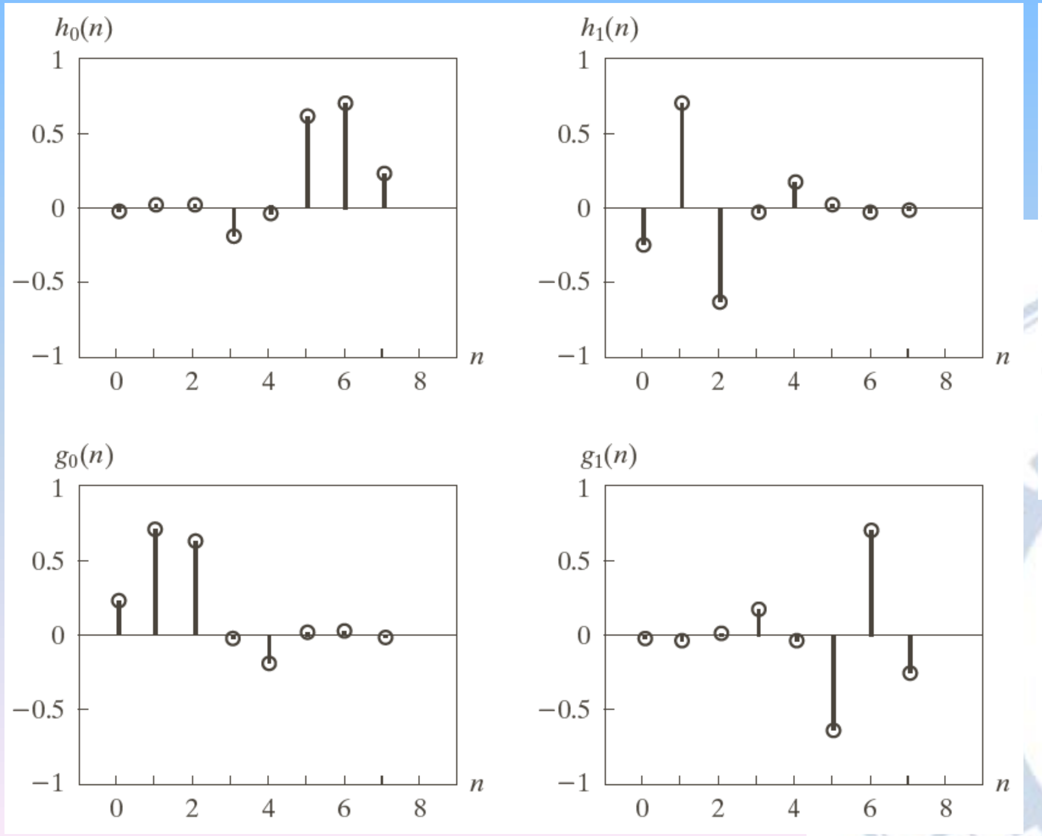
Example 7.2 a four-band subband coding of vase

n	$g_0(n)$
0	0.23037781
1	0.71484657
2	0.63088076
3	−0.02798376
4	−0.18703481
5	0.03084138
6	0.03288301
7	−0.01059740

TABLE 7.1 Daubechies 8-tap orthonormal filter coefficients for $g_0(n)$ (Daubechies [1992]).

Subband Coding

Example 7.2 a four-band subband coding of vase

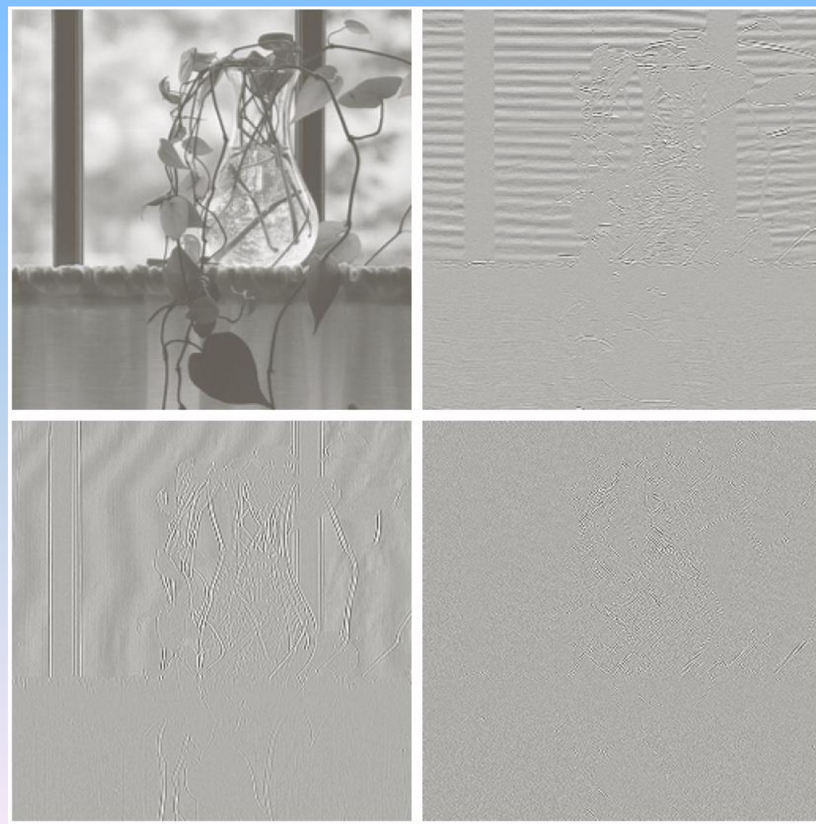


a b
c d

FIGURE 7.8 The impulse responses of four 8-tap Daubechies orthonormal filters. See Table 7.1 for the values of $g_0(n)$ for $0 \le n \le 7$.

Subband Coding

Example 7.2 a four-band subband coding of vase



a b
c d

FIGURE 7.9 A four-band split of the vase in Fig. 7.1 using the subband coding system of Fig. 7.7. The four subbands that result are the (a) approximation, (b) horizontal detail, (c) vertical detail, and (d) diagonal detail subbands.

Background

- The Haar Transform
  - Haar transform is a special wavelet transform
  - Its basis functions are the oldest and simplest orthonormal wavelets
  - Haar transform can be expressed in a matrix form

$$T = HFH^T$$

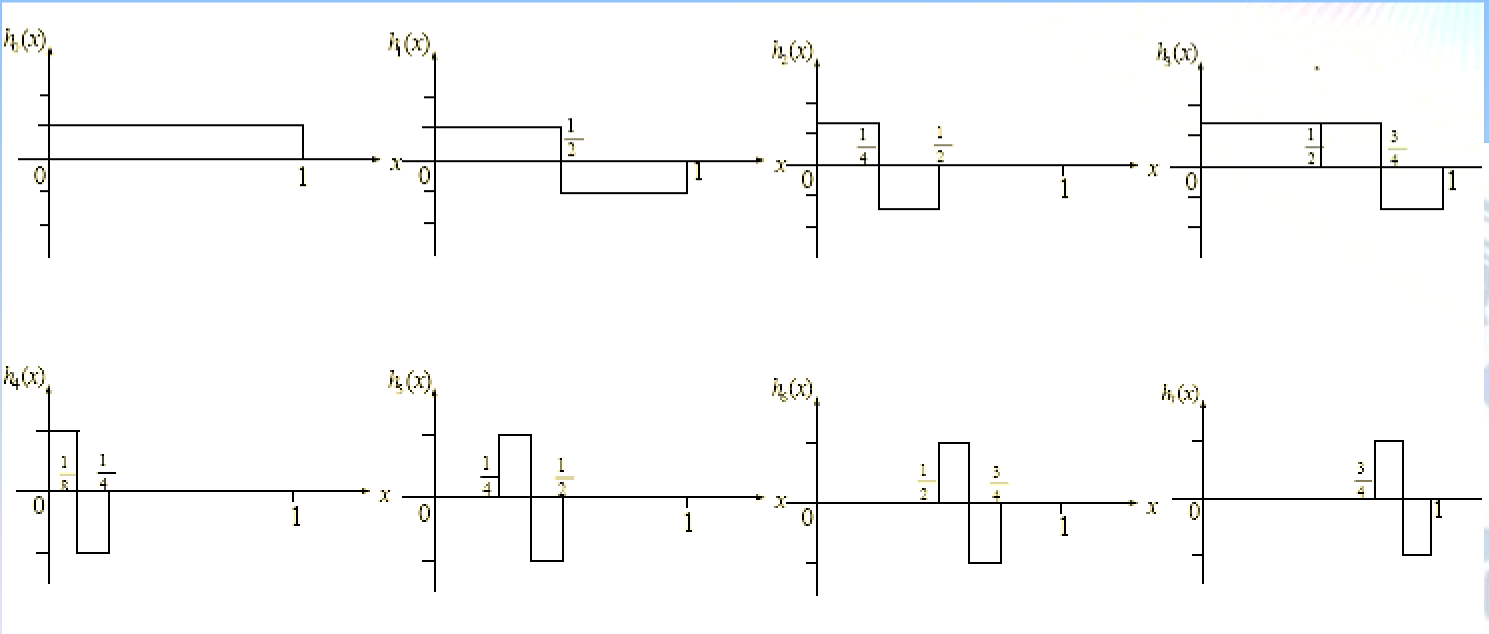
F is an $N \times N$ image matrix

H is an $N \times N$ transformation matrix

T is the resulting $N \times N$ transform

The Haar Transform

- Basis functions of Haar transform (continuous)



The Haar Transform

- All of the basis functions are rectangular impulse pairs except $h_0(z)$
- The impulse pairs have different width, height and positions
- Width of nonzero region is descending

$$1 \rightarrow \frac{1}{2} \rightarrow \frac{1}{4} \rightarrow \frac{1}{8} \rightarrow \ldots$$

- Height of nonzero region is ascending

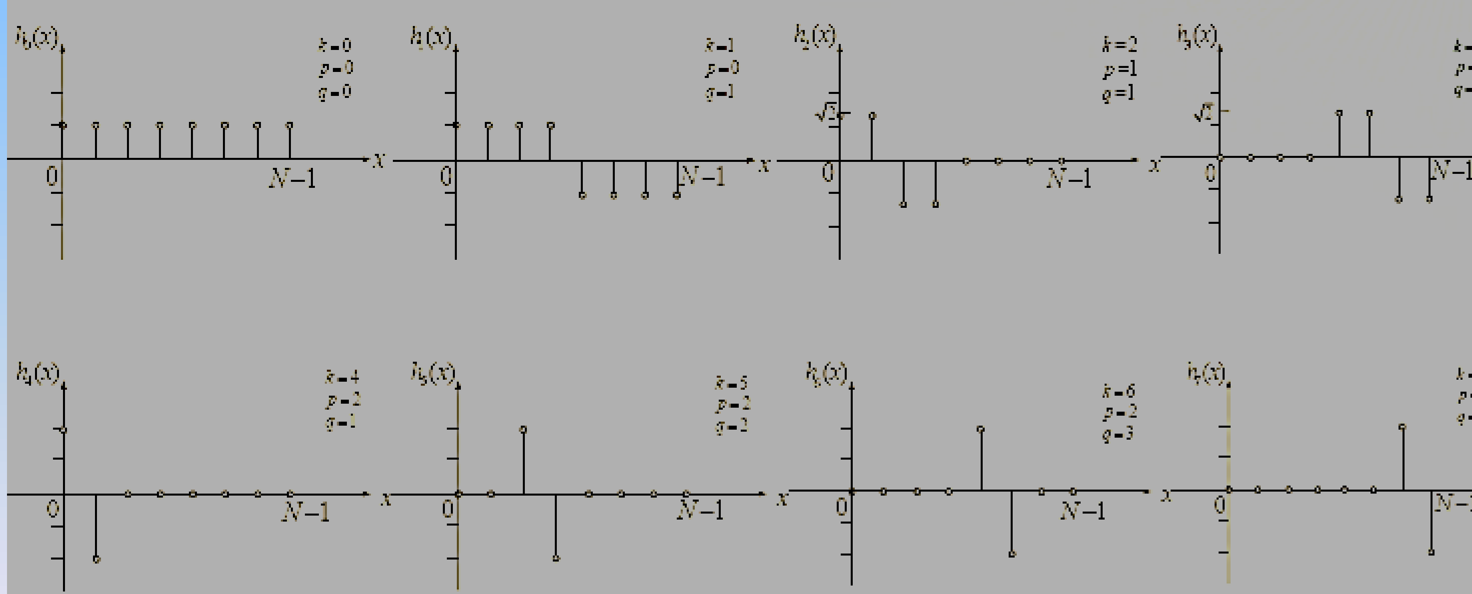
$$\frac{1}{\sqrt{N}} \rightarrow \frac{\sqrt{2}}{\sqrt{N}} \rightarrow \frac{2}{\sqrt{N}} \rightarrow \ldots$$

- The basis functions have the same characters as those of wavelet transform

The Haar Transform

- Basis functions of Haar transform (discrete)

$N = 8 \qquad h_k(x), k = 0 \sim 7$



$$k = 2^p + q - 1$$

$$2^p \le k \text{ and } 2^{p+1} > k$$

IVM http://ivm.sjtu.edu.cn
Image, Video, and Multimedia Communications Laboratory

The Haar Transform

- The ith row of an N by N Haar transformation matrix contains the elements of $h_i(z)$ for $z=\frac{0}{N},\frac{1}{N},\frac{2}{N},\cdots,\frac{N-1}{N}$
- 2×2 transformation matrix is

$$H_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

- 4×4 transformation matrix is

$$H_4=\frac{1}{\sqrt{4}}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ \sqrt{2} & -\sqrt{2} & 0 & 0\\ 0 & 0 & \sqrt{2} & -\sqrt{2}\end{bmatrix}$$

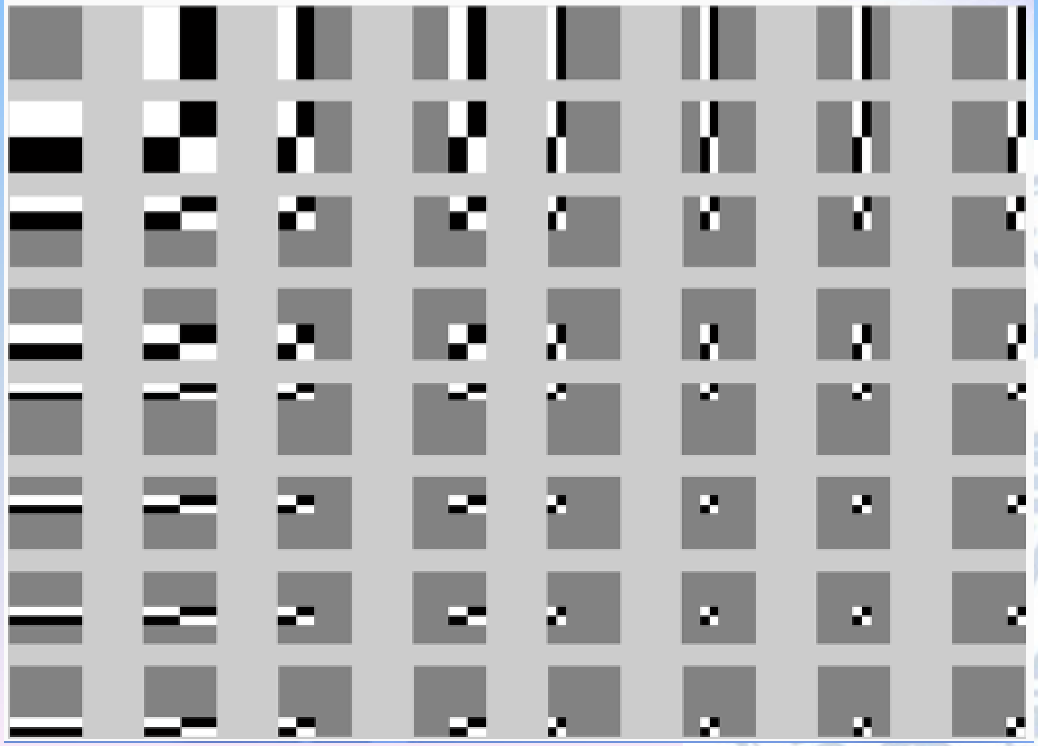
The Haar Transform

- 8×8 transformation matrix is

$$H_8 = \frac{1}{\sqrt{8}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$

The Haar Transform

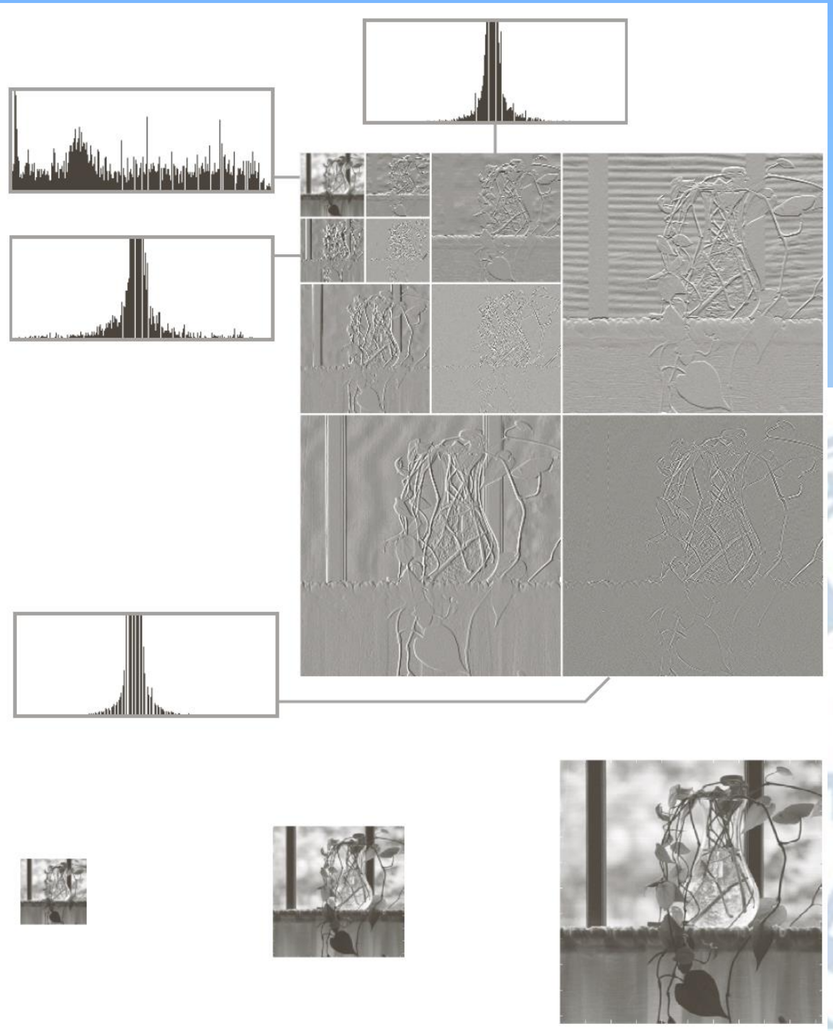
- Transformation kernel is separable
- For N = 8 , the basis functions are



The Haar Transform

- Example 7.3

Haar functions in a discrete wavelet transform.



a
b c d

FIGURE 7.10 (a) A discrete wavelet transform using Haar $\mathbf{H}_2$ basis functions. Its local histogram variations are also shown. (b)–(d) Several different approximations (64×64, 128×128, and 256×256) that can be obtained from (a).

Basic of Multiresolution Analysis

Basic of Multiresolution Analysis

- Expansion of a signal $f(x)$:

$$f(x) = \sum_k \alpha_k \phi_k(x)$$

α_k : real-valued expansion coefficients

$\phi_k(x)$: real-valued expansion functions

$$\alpha_k = \left\langle \tilde{\phi}_k(x), f(x) \right\rangle = \int \tilde{\phi}_k^*(x) f(x) dx$$

$\tilde{\phi}_k(x)$: the dual function of $\phi_k(x)$

If the expansion is unique, the $\phi_k(x)$ are called basis functions.

Basic of Multiresolution Analysis

If $\{\phi_k(x)\}$ is an orthonormal basis for V, then $\phi_k(x) = \tilde{\phi}_k(x)$

The function space of the expansion set $\phi_k(x)$: $V = \underset{k}{span}\{\phi_k(x)\}$

If $\{\phi_k(x)\}$ are not orthonormal but are an orthogonal basis for V, then the basis funcitons and their duals are called biorthogonal.

Biorthogonal: $\left\langle \phi_j(x), \tilde{\phi}_k(x) \right\rangle = \delta_{jk} = \begin{cases} 0 & , j \neq k \\ 1 & , j = k \end{cases}$

How to construct such orthonormal basis?

Basic of Multiresolution Analysis

- Scaling functions

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

The subspace spanned over k for any j :

$$V_j = \underset{k}{span}\left\{\phi_{j,k}(x)\right\}$$

Basic of Multiresolution Analysis

- Requirements of scaling function:

1. The scaling function is orthogonal to its integer translates.
2. The subspaces spanned by the scaling function at low scales are nested within those spanned at higher scales.
   That is
   $$V_{-\infty} \subset \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{\infty}$$
3. The only function that is common to all V_j is $f(x) = 0$
   That is
   $$V_{-\infty} = \{0\}$$
4. Any function can be represented with arbitrary precision.
   That is,
   $$V_{\infty} = \{L^2(\mathbf{R})\}$$

Basic of Multiresolution Analysis

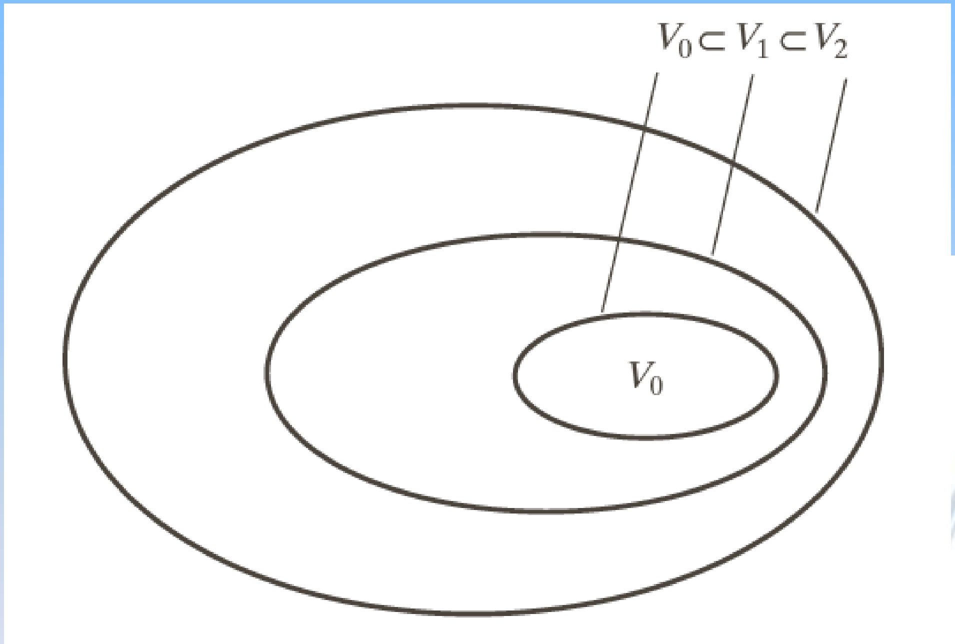


FIGURE 7.12 The nested function spaces spanned by a scaling function.

Are these scaling functions orthonormal basis?

NO!

Basic of Multiresolution Analysis

- Wavelet function

spans the difference between any two adjacent scaling subspaces V_j and V_{j+1}

$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$ for all $k \in \mathbf{Z}$ that spans the space W_j

where $W_j = \underset{k}{span}\left\{\psi_{j,k}(x)\right\}$

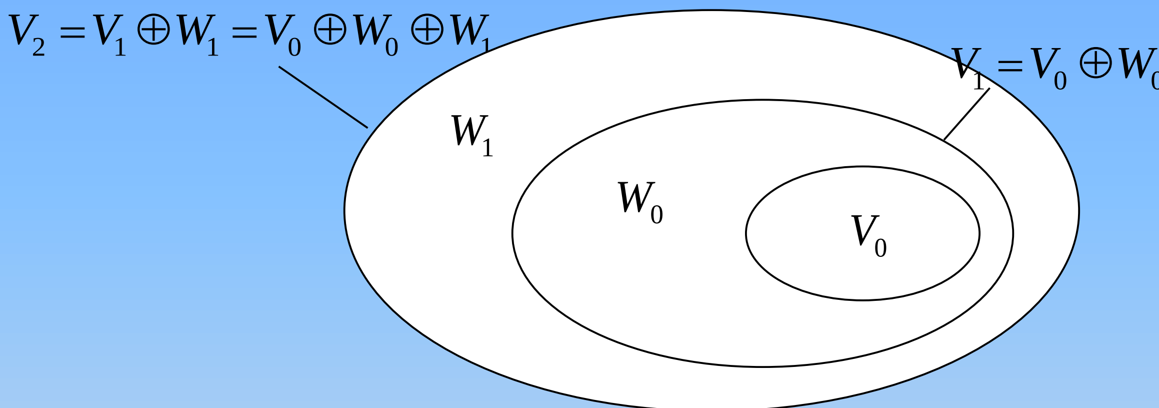
The wavelet function can be expressed as a weighted sum of shifted, double-resolution scaling functions. That is,

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\phi(2x - n)$$

where the $h_\psi(n)$ are called the wavelet function coefficients.

It can be shown that $h_\psi(n) = (-1)^n h_\phi(1-n)$

Basic of Multiresolution Analysis



Are these wavelet functions orthonormal basis? Yes!

The scaling and wavelet function subspaces are related by $V_{j+1} = V_j \oplus W_j$

We can express the space of all measurable, square-integrable function as

$$L^2(\mathbf{R}) = V_0 \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots$$

or

$$L^2(\mathbf{R}) = \cdots \oplus W_{-2} \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots$$

A Tour of Wavelets

Wavelet Transforms in One Dimension

- The wavelet series expansions

$$f(x)=\sum_{k} c_{j_0}(k)\varphi_{j_0,k}(x)+\sum_{j=j_0}^{\infty}\sum_{k} d_j(k)\psi_{j_0,k}(x)$$

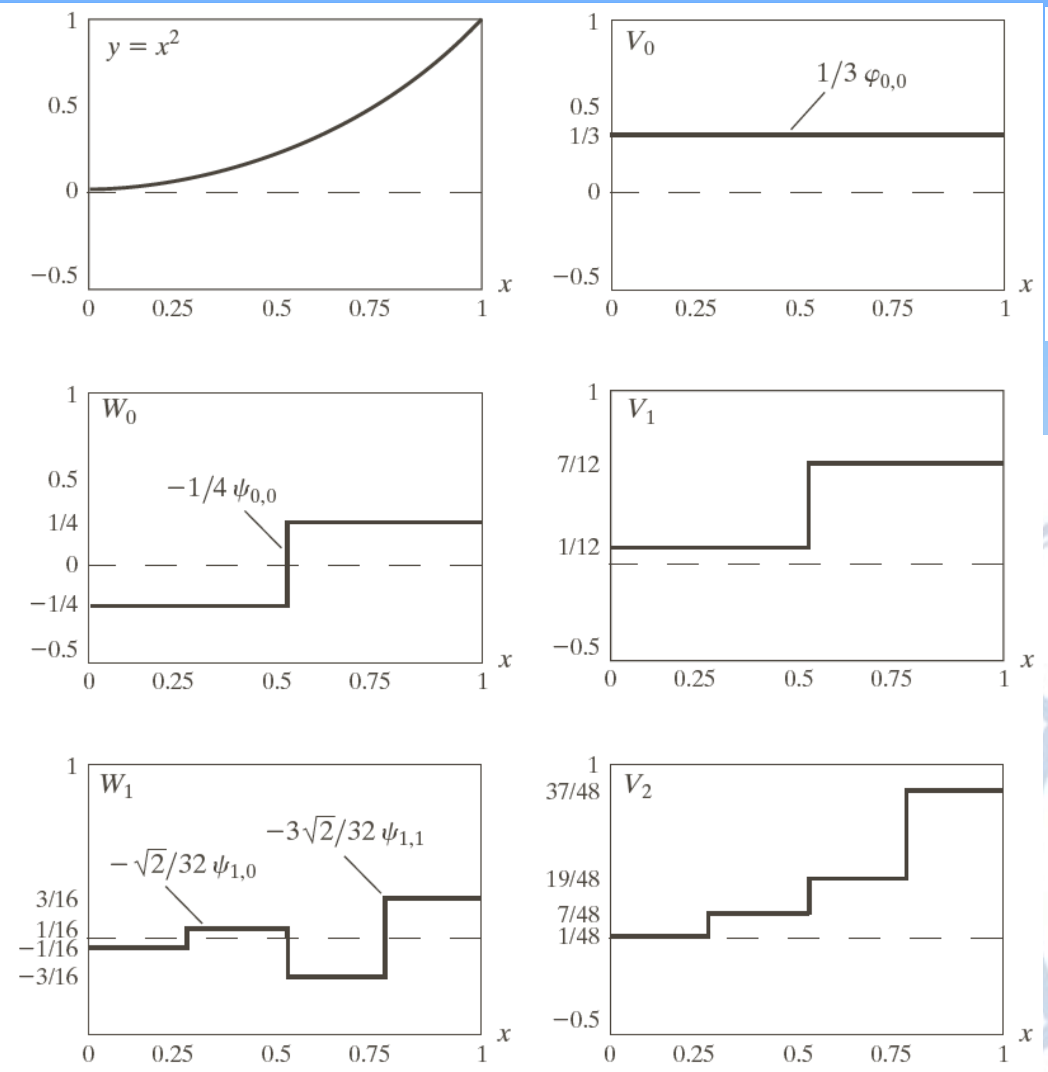
- The expansion coefficients are calculated

$$c_{j_0}(k)=\langle f(x),\varphi_{j_0,k}(x)\rangle=\int f(x)\varphi_{j_0,k}(x)dx$$

and

$$d_j(k)=\langle f(x),\psi_{j,k}(x)\rangle=\int f(x)\psi_{j,k}(x)dx$$

Wavelet Transforms in One Dimension



a b
c d
e f

FIGURE 7.15 A wavelet series expansion of $y = x^2$ using Haar wavelets.

Wavelet Transforms in One Dimension

- The discrete wavelet transform

$$W_{\varphi}(j_0,k)=\frac{1}{\sqrt{M}}\sum_{n} f(n)\varphi_{j_0,k}(n)$$

$$W_{\psi}(j,k)=\frac{1}{\sqrt{M}}\sum_{n} f(n)\psi_{j,k}(n) \quad for\ j \geq j_0$$

- The complementary inverse DWT is

$$f(n)=\frac{1}{\sqrt{M}}\sum_{k} W_{\varphi}(j_0,k)\varphi_{j_0,k}(n)+\frac{1}{\sqrt{M}}\sum_{j=j_0}^{\infty}\sum_{k} W_{\psi}(j,k)\psi_{j,k}(n)$$

Wavelet Transforms in One Dimension

- The continuous wavelet transform

$$W_{\psi}(s,\tau)=\int_{-\infty}^{\infty} f(x)\psi_{s,\tau}(x)dx$$

where

$$\psi_{s,\tau}(x)=\frac{1}{\sqrt{s}}\psi\left(\frac{x-\tau}{s}\right)$$

- The inverse continuous wavelet transform

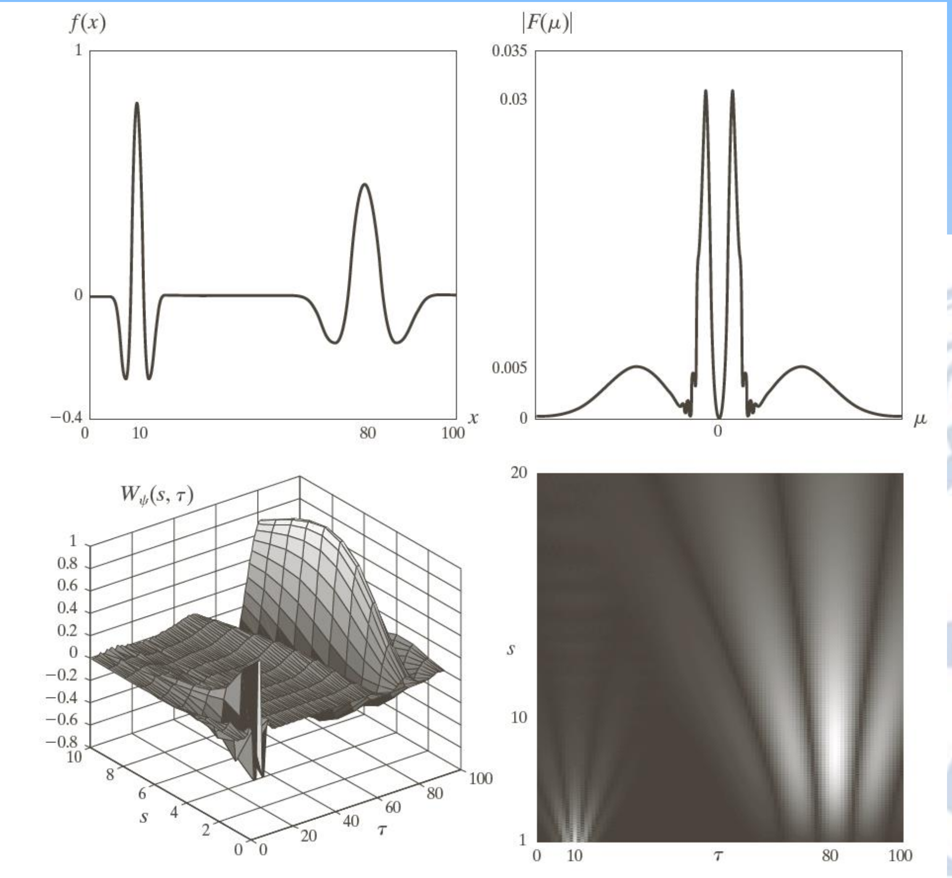
$$f(n)=\frac{1}{C_{\psi}}\int_{0}^{\infty}\int_{-\infty}^{\infty} W_{\psi}(s,\tau)\frac{\psi_{s,\tau}(x)}{s^{2}}dsd\tau$$

where

$$C_{\psi}=\int_{-\infty}^{\infty}\frac{|\Psi(\mu)|^{2}}{|\mu|}d\mu$$

Wavelet Transforms in One Dimension

- The Mexican hat wavelet



a b
c d

FIGURE 7.16 The continuous wavelet transform (c and d) and Fourier spectrum (b) of a continuous 1-D function (a).

The Fast Wavelet Transform

- Computationally efficient implementation of the DWT
- The relationship between the coefficients of the DWT at adjacent scales
- Also called Mallat's herringbone algorithm
- Resembles the twoband subband coding scheme

The Fast Wavelet Transform

- The multiresolution refinement equation

$$\varphi(x)=\sum_n h_\varphi(n)\sqrt{2}\varphi(2x-n)$$

- Scaling x by 2^j, translating it by k, and letting m=2k+n gives

$$\varphi\left(2^j x-k\right)=\sum_m h_\varphi(m-2k)\sqrt{2}\varphi\left(2^{j+1}x-n\right)$$

- A similar sequence of operations—beginning with Eq. (7.2-28)—provides an analogous result

$$\psi\left(2^j x-k\right)=\sum_m h_\psi(m-2k)\sqrt{2}\varphi\left(2^{j+1}x-n\right)$$

The Fast Wavelet Transform

- Now consider the wavelet series expansion coefficients of continuous function f(x), we get

$$d_j(k)=\sum_m h_\psi(m-2k)c_{j+1}(m)$$

$$c_j(k)=\sum_m h_\varphi(m-2k)c_{j+1}(m)$$

- When f(x) is discrete, the coefficients of the wavelet series expansion become the coefficient of the DWT

$$W_\psi(j,k)=\sum_m h_\psi(m-2k)W_\varphi(j+1,m)$$

$$W_\varphi(j,k)=\sum_m h_\varphi(m-2k)W_\varphi(j+1,m)$$

The Fast Wavelet Transform

- We can write

$$W_\psi(j,k) = h_\psi(-n) \star W_\varphi(j+1,m)\Big|_{n=2k,k\geq 0}$$

$$W_\varphi(j,k) = h_\varphi(-n) \star W_\varphi(j+1,m)\Big|_{n=2k,k\geq 0}$$

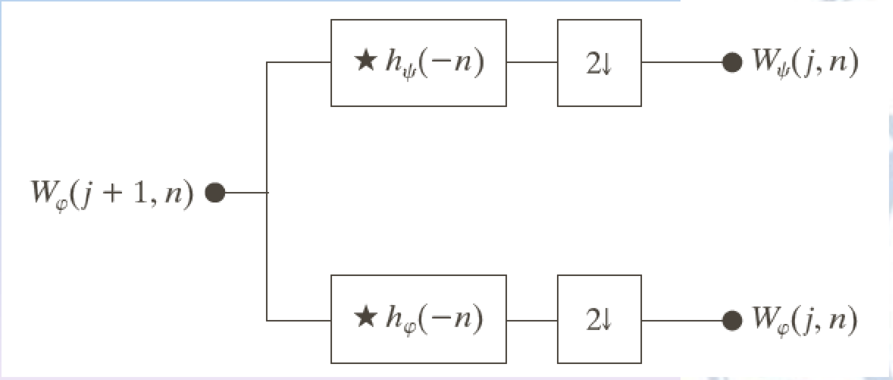
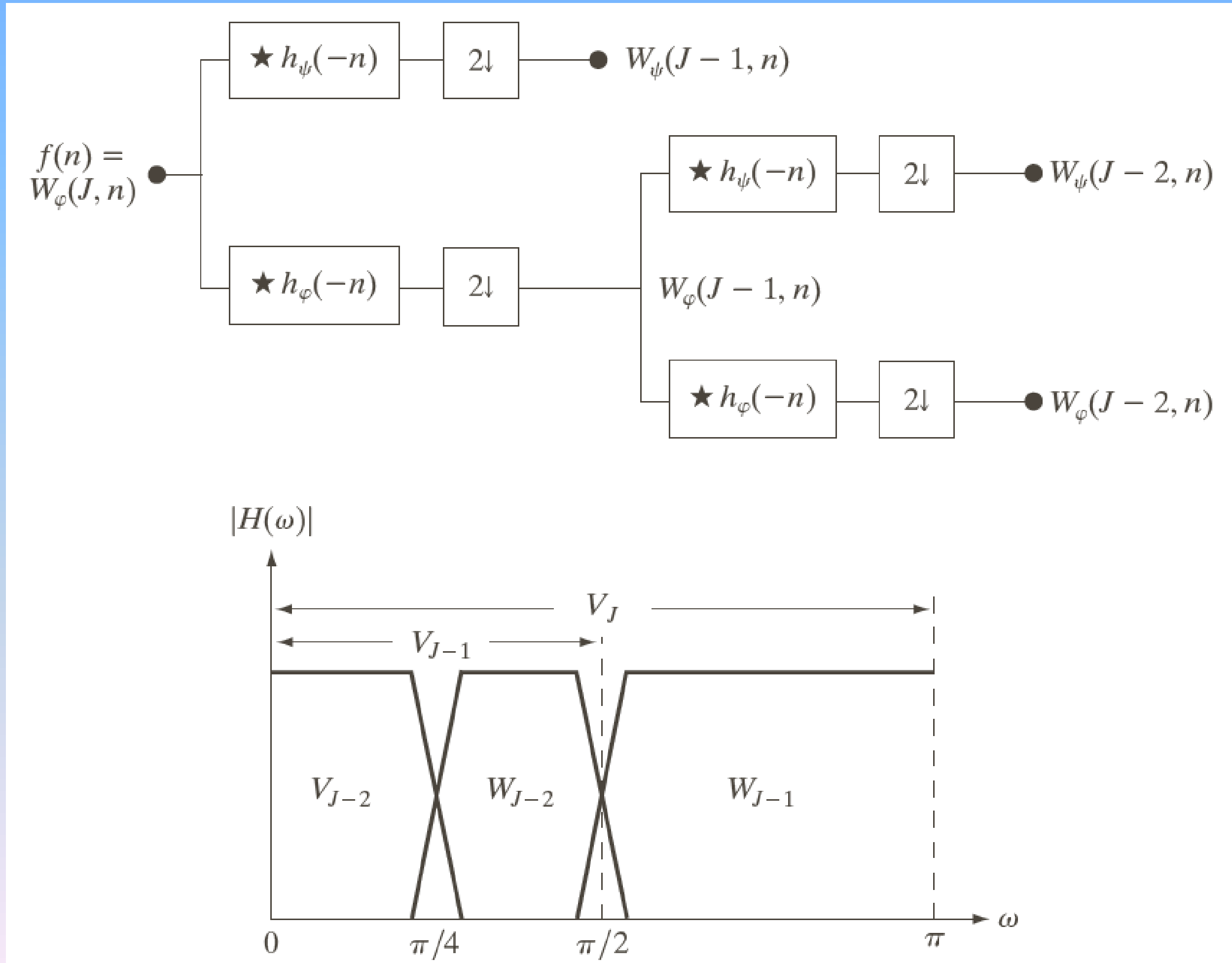


FIGURE 7.17
An FWT analysis bank.

The Fast Wavelet Transform



a
b

FIGURE 7.18
(a) A two-stage or two-scale FWT analysis bank and (b) its frequency splitting characteristics.

The Fast Wavelet Transform

- Example 7.10 Computing a 1-D fast wavelet transform

n	$h_\varphi(n)$
0	$1/\sqrt{2}$
1	$1/\sqrt{2}$

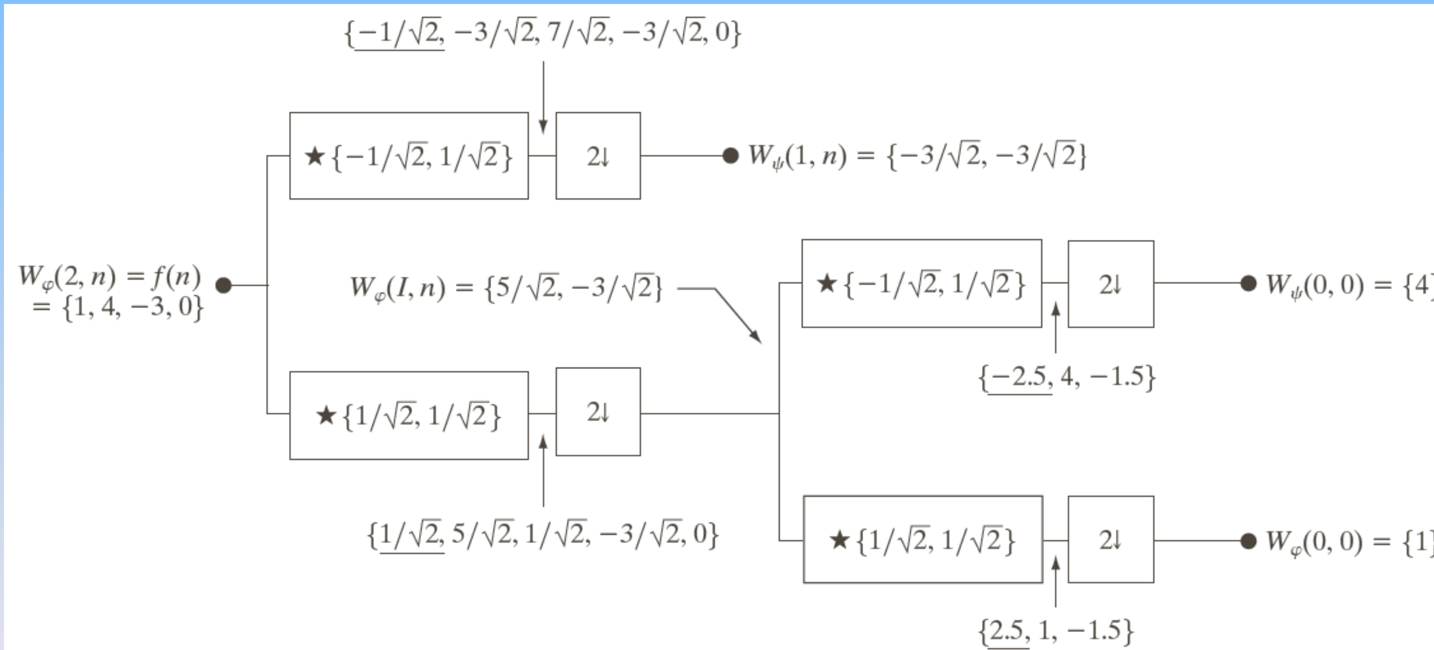
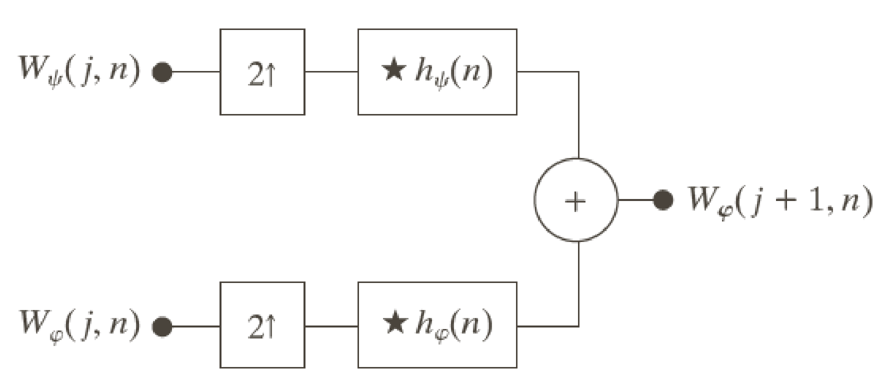


FIGURE 7.19 Computing a two-scale fast wavelet transform of sequence $\{1, 4, -3, 0\}$ using Haar scaling and wavelet vectors.

The Fast Wavelet Transform

- A fast inverse transform for the reconstruction of f(n) from the results of the forward transform can be formulated



- The FWT^{-1} filter bank implements the computation

$$W_\varphi(j+1,k) = h_\varphi(k) * W_\varphi^{\uparrow 2}(j,k) + h_\psi(k) * W_\psi^{\uparrow 2}(j,k)\Big|_{k\geq 0}$$

The Fast Wavelet Transform

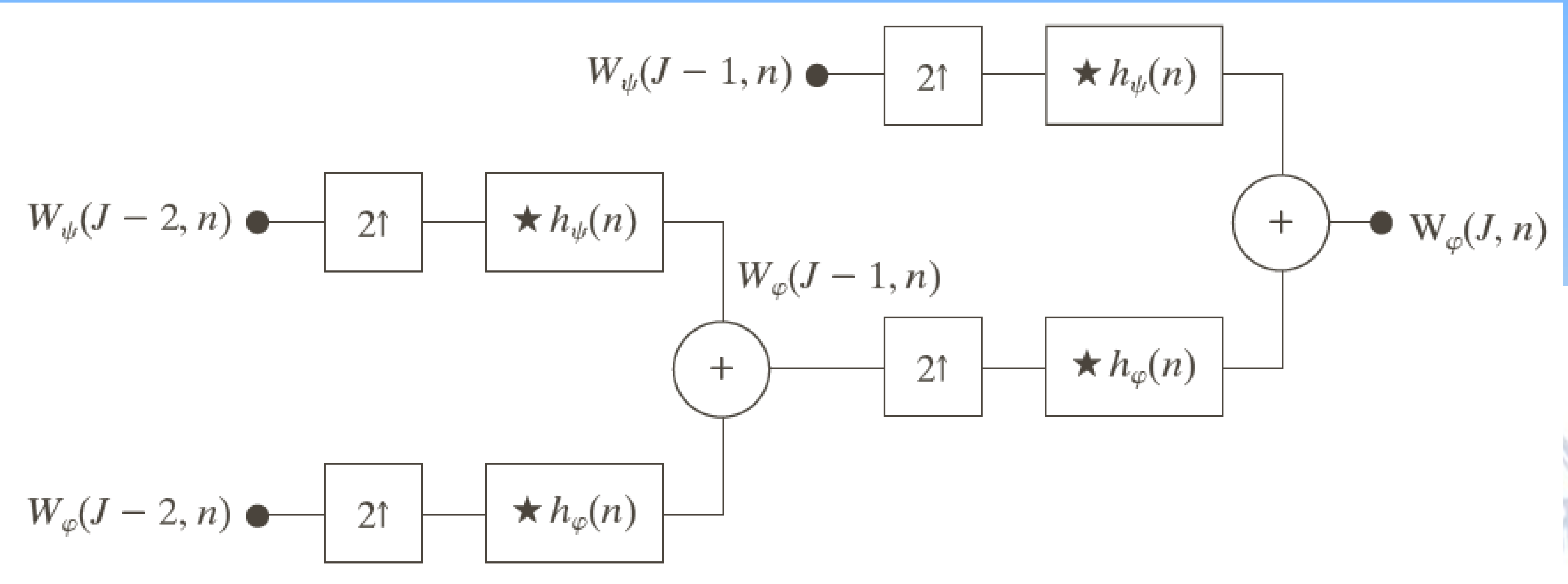


FIGURE 7.21
A two-stage or two-scale FWT^{-1} synthesis bank.

The Fast Wavelet Transform

- Example 7.11 computing a 1-D inverse fast wavelet transform.

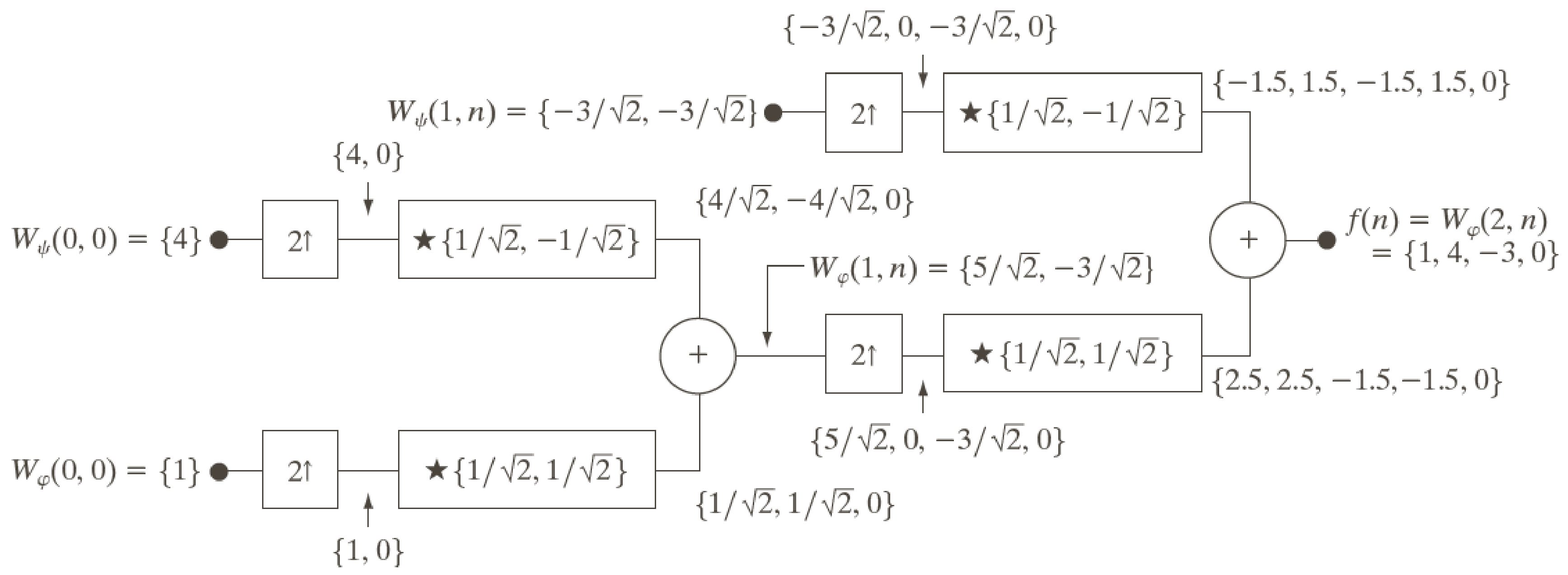


FIGURE 7.22 Computing a two-scale inverse fast wavelet transform of sequence $\{1, 4, -1.5\sqrt{2}, -1.5\sqrt{2}\}$ with Haar scaling and wavelet functions.

The Fast Wavelet Transform

- Time-frequency tilings for the basis functions

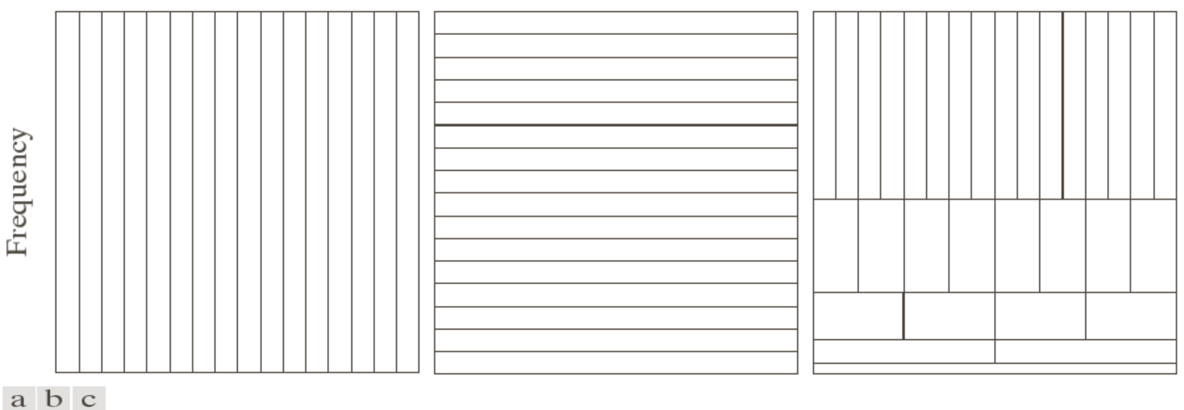


FIGURE 7.23 Time-frequency tilings for the basis functions associated with (a) sampled data, (b) the FFT, and (c) the FWT. Note that the horizontal strips of equal height rectangles in (c) represent FWT scales.

- If you want precise information about, you must accept some vagueness about frequency, and vice versa. This is the Heisenberg uncertainty principle applied to information processing

Wavelet Transform in Two Dimensions

- In two-dimensions, a two-dimensional scaling function, $\varphi(x,y)$ and three two-dimensional wavelets, $\psi^H(x,y)$, $\psi^V(x,y)$, and $\psi^D(x,y)$, are required
- The separable function

$$\varphi(x,y)=\varphi(x)\varphi(y)$$

and separable, directionally sensitive wavelets

$$\psi^H(x,y)=\psi(x)\varphi(y)$$
$$\psi^V(x,y)=\varphi(x)\psi(y)$$
$$\psi^D(x,y)=\psi(x)\psi(y)$$

Wavelet Transform in Two Dimensions

- We first define the scaled and translated basis functions

$$\varphi_{j,m,n}(x,y) = 2^{j/2}\varphi\left(2^j x - m, 2^j y - n\right),$$
$$\psi^i_{j,m,n}(x,y) = 2^{j/2}\psi^i\left(2^j x - m, 2^j y - n\right), i = \{H, V, D\}$$

- Rather than an exponent, i is a superscript that assumes the values H, V, and D. the discrete wavelet transform of image f(x, y) of size M by N is then

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\varphi_{j_0,m,n}(x,y),$$
$$W^i_\psi(j, m, n) = \frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\psi^i_{j,m,n}(x,y), i = \{H, V, D\}$$

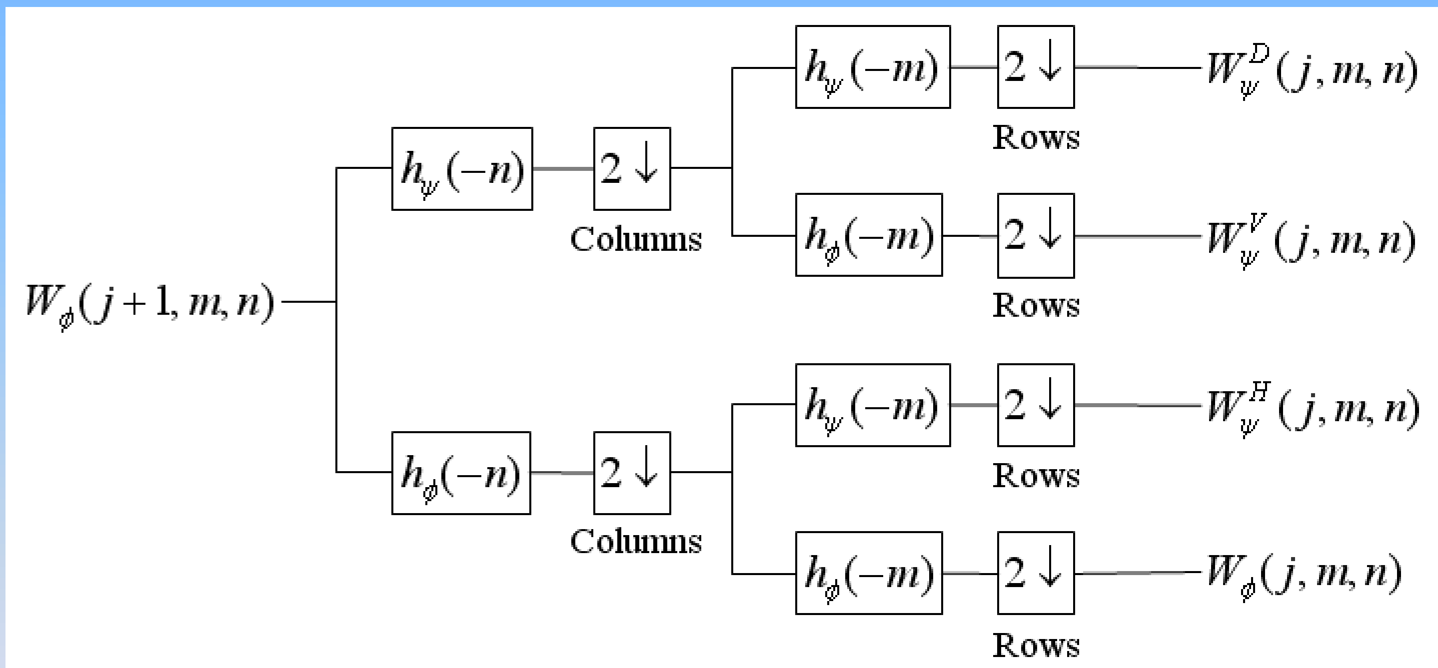
Wavelet Transform in Two Dimensions

- Inverse discrete wavelet transform

$$f(x,y)=\frac{1}{\sqrt{MN}}\sum_{m}\sum_{n}W_{\varphi}(j_0,m,n)\varphi_{j_0,m,n}(x,y)$$
$$+\frac{1}{\sqrt{MN}}\sum_{i=H,V,D}\sum_{j=j_0}^{\infty}\sum_{m}\sum_{n}W_{\psi}^{i}(j,m,n)\psi_{j,m,n}^{i}(x,y)$$

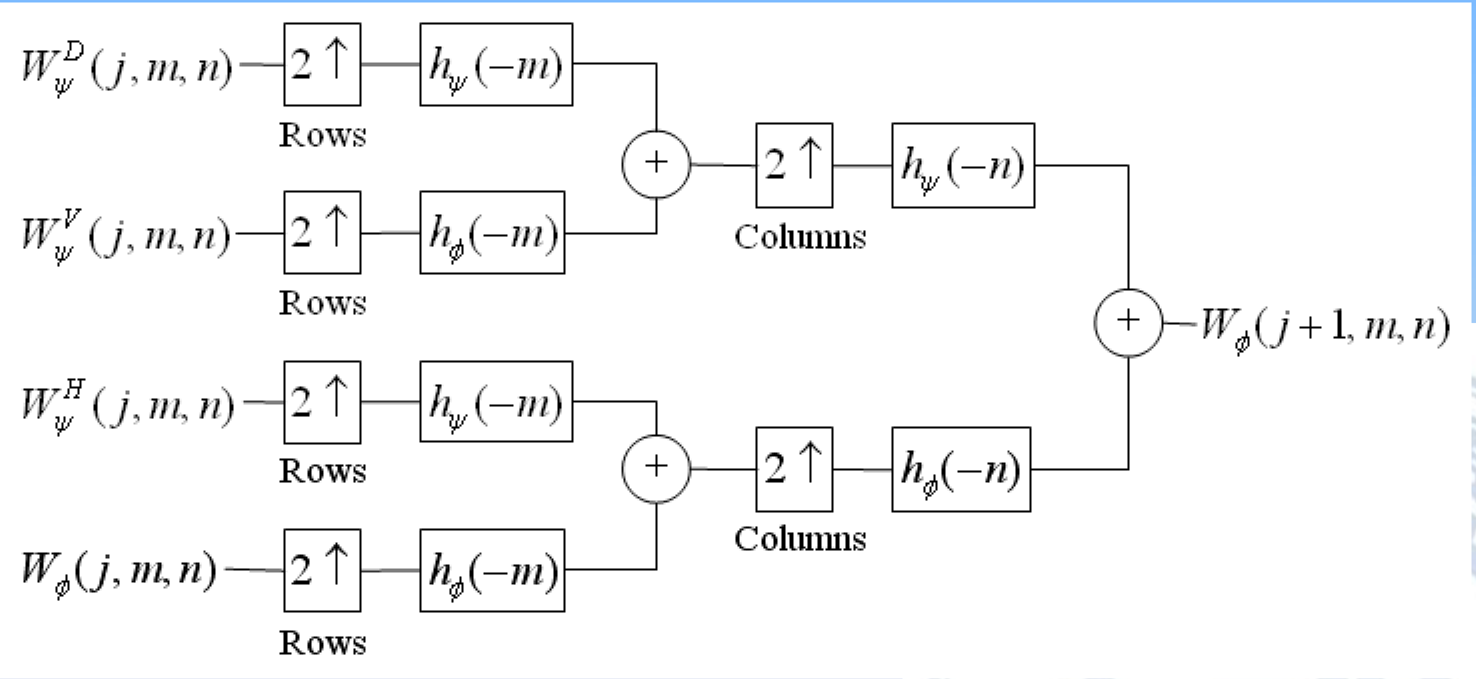
- Like the 1-D discrete wavelet transform, the 2-D DWT can be implemented using digital filters and downsamplers. With separable two-dimensional scaling and wavelet functions, we simply take the 1-D FWT of rows of f(x, y), followed by the 1-D FWT of the resulting columns

Wavelet Transform in Two Dimensions



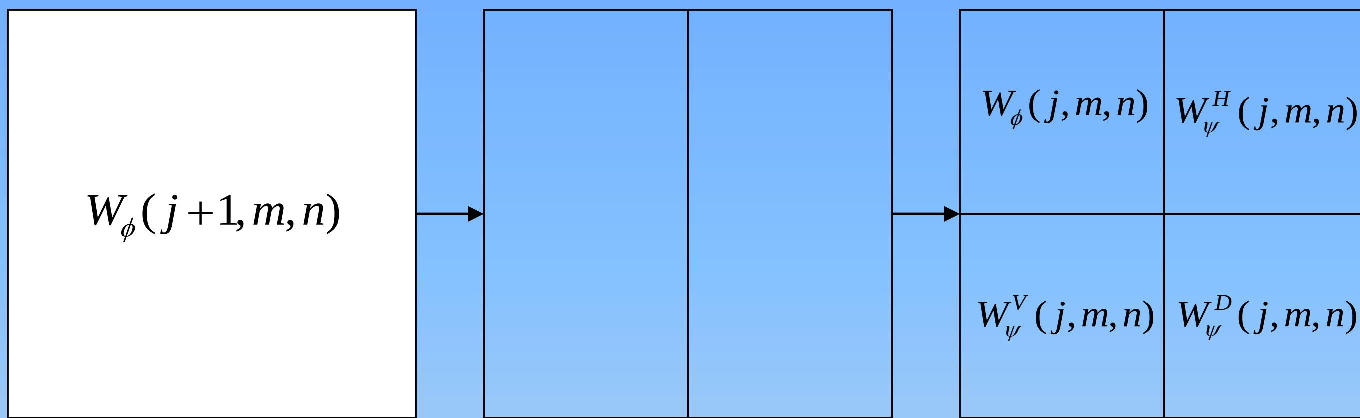
The two-dimensional FWT — the analysis filter.

Wavelet Transform in Two Dimensions

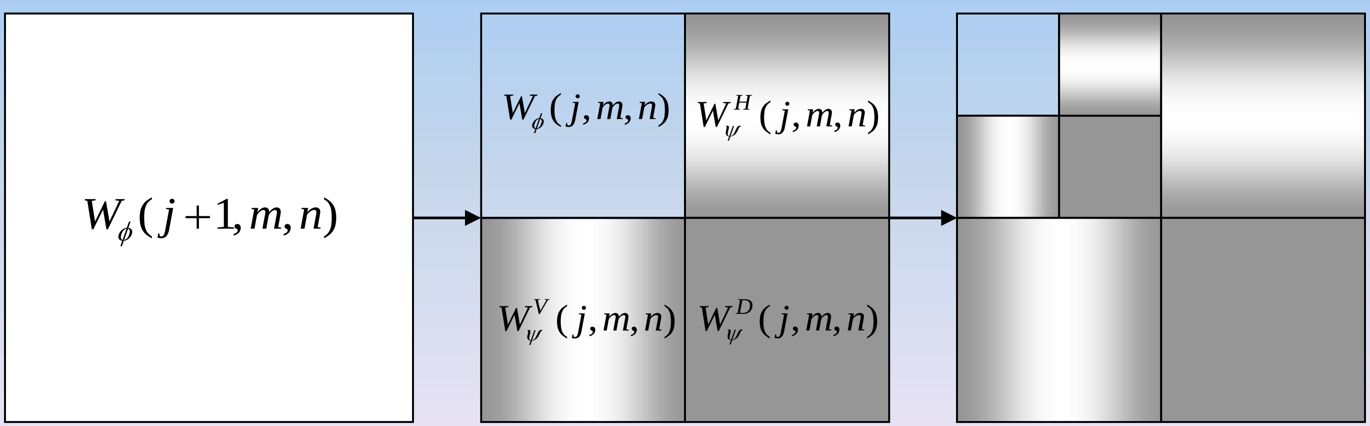


The two-dimensional FWT — the synthesis filter bank.

Wavelet Transforms in Two Dimension



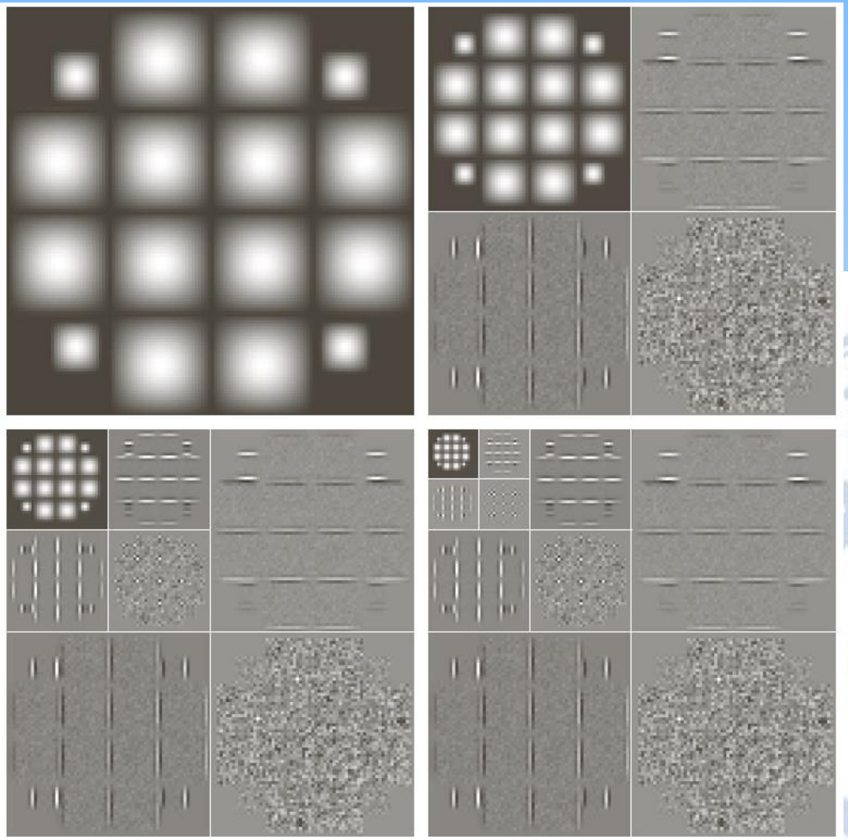
Two-dimensional decomposition



Two-scale of two-dimensional decomposition

Wavelet Transform in Two Dimensions

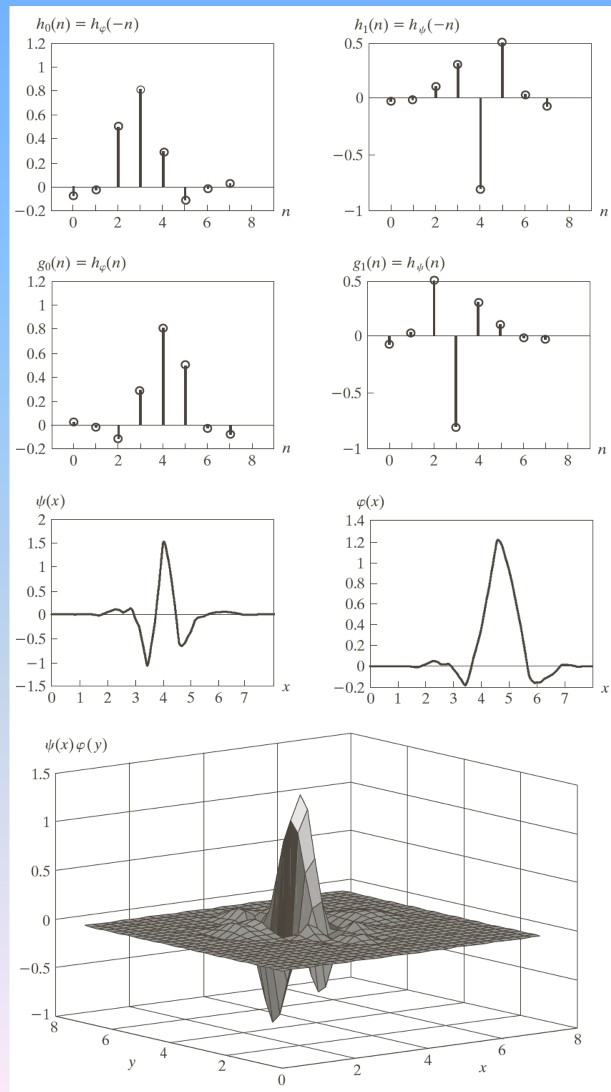
- Example 7.12 computing a 2-D fast wavelet transform.



a b
c d

FIGURE 7.25 Computing a 2-D three-scale FWT: (a) the original image; (b) a one-scale FWT; (c) a two-scale FWT; and (d) a three-scale FWT.

Wavelet Transform in Two Dimensions



a b
c d
e f
g

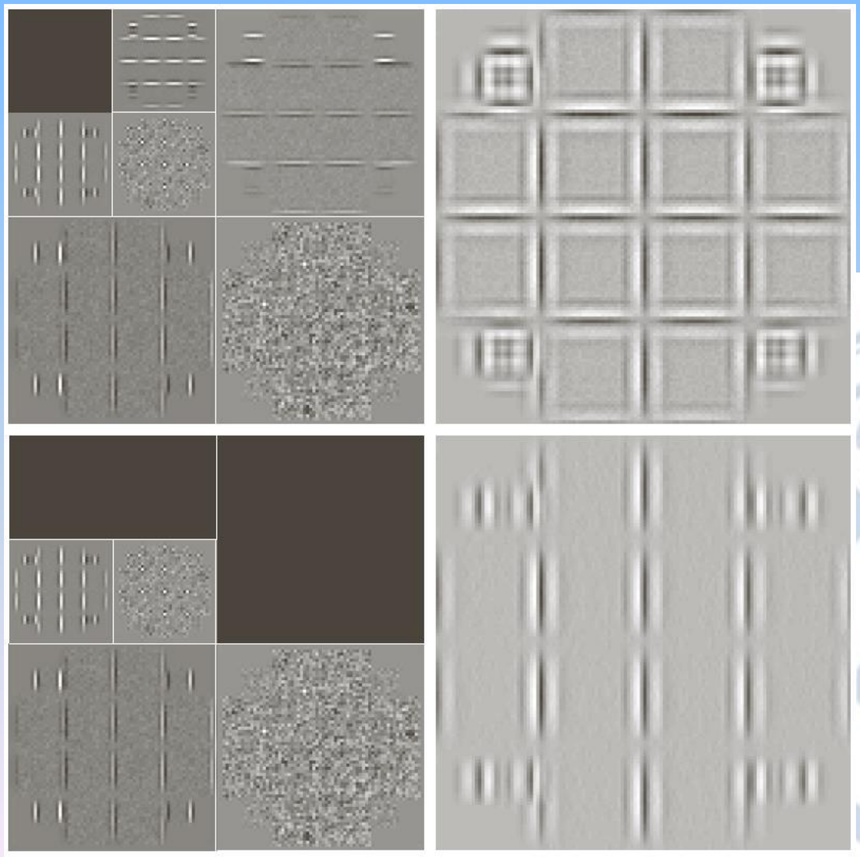
FIGURE 7.26 Fourth-order symlets: (a)–(b) decomposition filters; (c)–(d) reconstruction filters; (e) the one-dimensional wavelet; (f) the one-dimensional scaling function; and (g) one of three two-dimensional wavelets, $\psi^V(x, y)$. See Table 7.3 for the values of $h_\varphi(n)$ for $0 \le n \le 7$.

Wavelet in image processing

- Wavelets in image processing
  - As in the Fourier domain, the basic approach is to
    - Step 1. Compute a 2-D wavelet transform of an image.
    - Step 2. Alter the transform.
    - Step 3. Compute the inverse transform.

Wavelet in image processing

- Example 7.13 wavelet-based edge detection.



a b
c d

FIGURE 7.27 Modifying a DWT for edge detection: (a) and (c) two-scale decompositions with selected coefficients deleted; (b) and (d) the corresponding reconstructions.

Wavelet in image processing

- Example 7.14

Wavelet-based noise removal

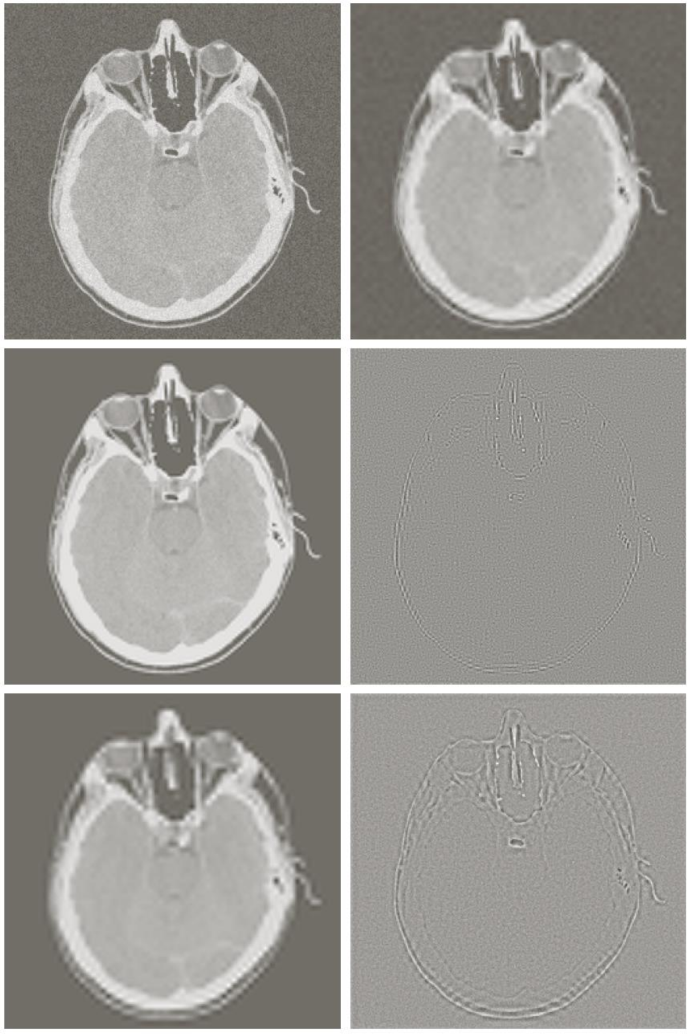
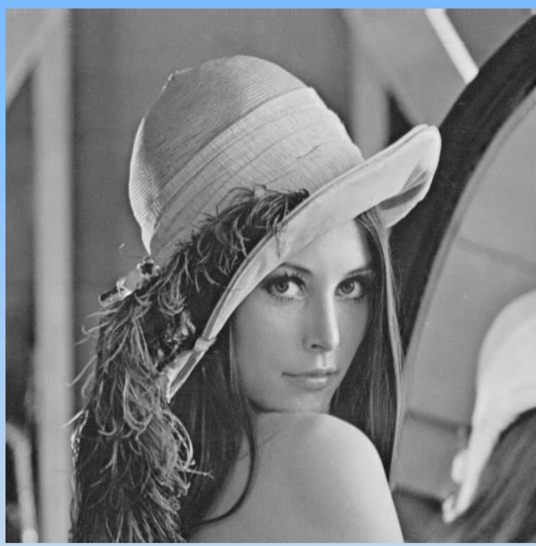
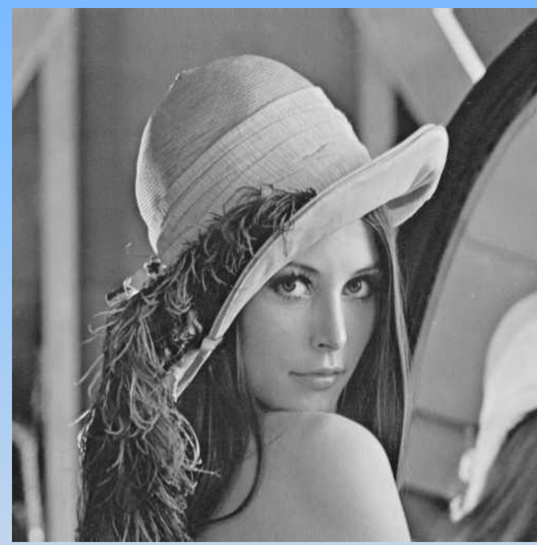
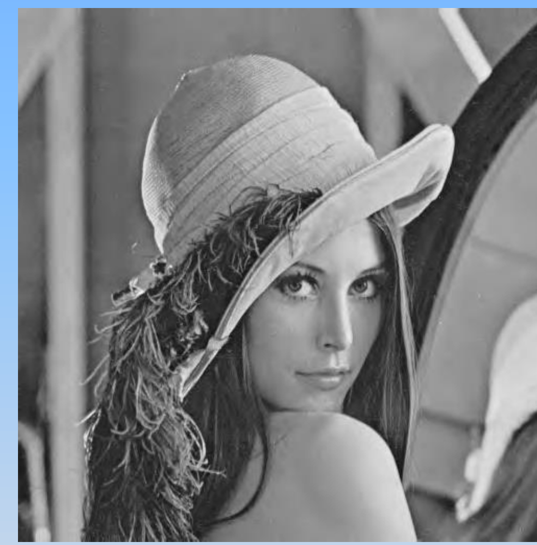


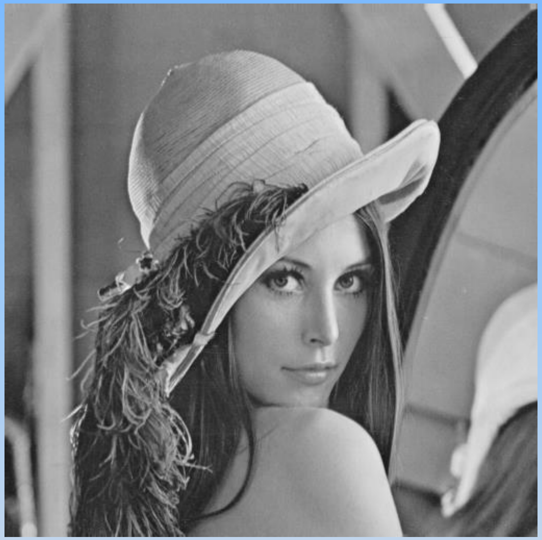
FIGURE 7.28 Modifying a DWT for noise removal: (a) a noisy CT of a human head; (b), (c) and (e) various reconstructions after thresholding the detail coefficients; (d) and (f) the information removed during the reconstruction of (c) and (e). (Original image courtesy Vanderbilt University Medical Center.)

Wavelet in image compression

- Quantization $q_j(m,n) = \mathrm{sign}(y_j(m,n)) \left\lfloor \frac{|W_j(m,n)|}{\Delta_j} \right\rfloor$
  - uniform scalar quantization
  - separate quantization step-sizes for each subband
- Entropy coding
  - Huffman coding
  - Arithmetic coding

Original image	DCT-based image compression	Wavelet-based image compression
		
	CR = 11.2460 RMS = 4.1316	CR = 10.3565 RMS = 4.0104

Original image



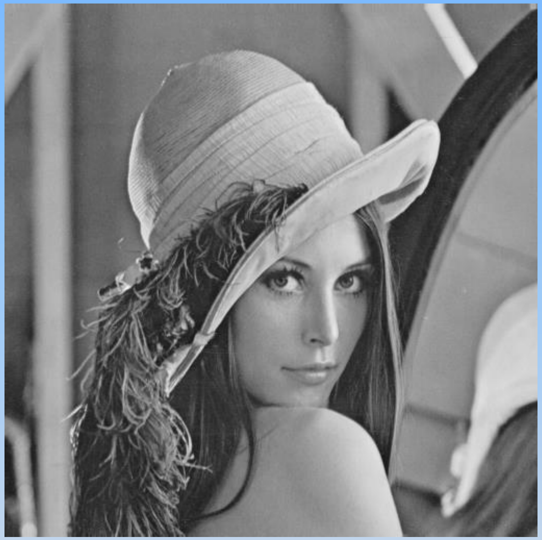
DCT-based
image compression

CR = 27.7401
RMS = 6.9763

Wavelet-based
image compression

CR = 26.4098
RMS = 6.8480

Original image



DCT-based
image compression

CR = 53.4333
RMS = 10.9662

Wavelet-based
image compression

CR = 51.3806
RMS = 9.6947

Thank You!