Lecture notes of Image Compression and Video Compression series [2005]

3. Subband Coding

Prof. Amar Mukherjee

Weifeng Sun

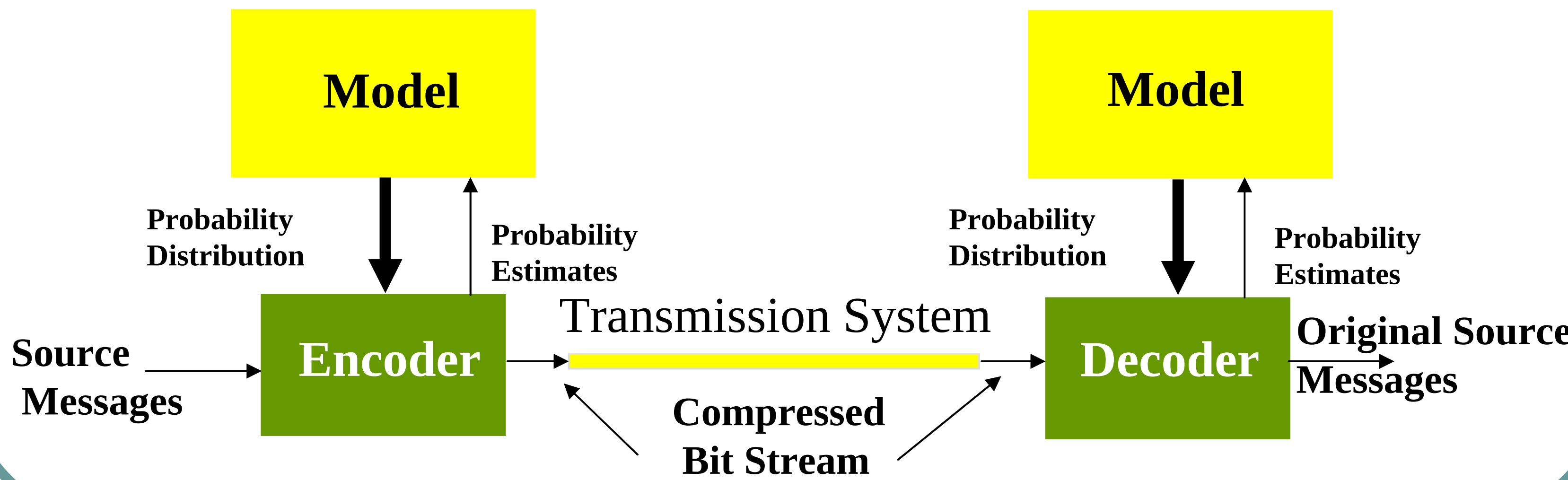
Topics

- Introduction to Image Compression
- Transform Coding
- Subband Coding, Filter Banks
- Introduction to Wavelet Transform
- Haar, SPIHT, EZW, JPEG 2000
- Motion Compensation
- Wireless Video Compression

Contents

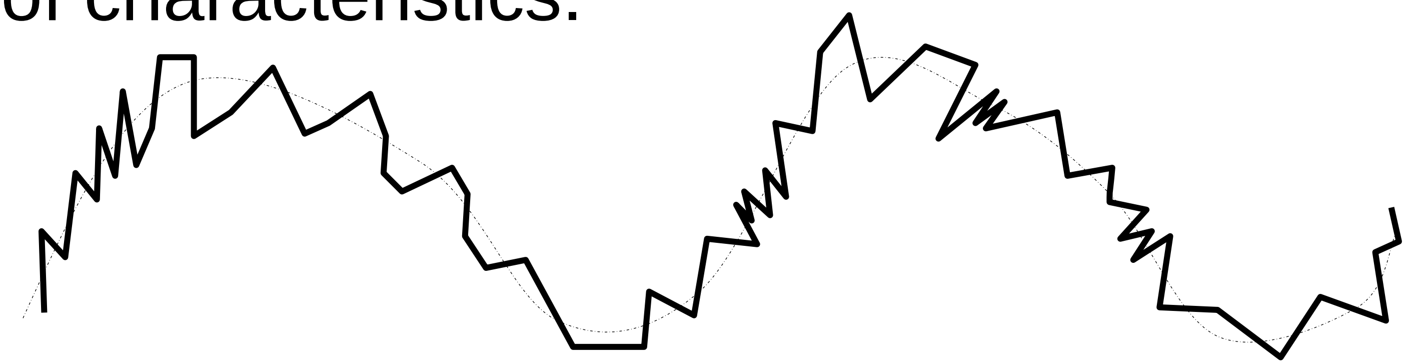
Model and Prediction

- Compression is **PREDICTION**.
- Good model results in better prediction.
  - For text, we have good models (PPM, BWT).
  - For images, we are in need of a good model!



Separate the Signals

- Modeling a general signal is difficult.
- However, most signals exhibit a combination of characteristics.



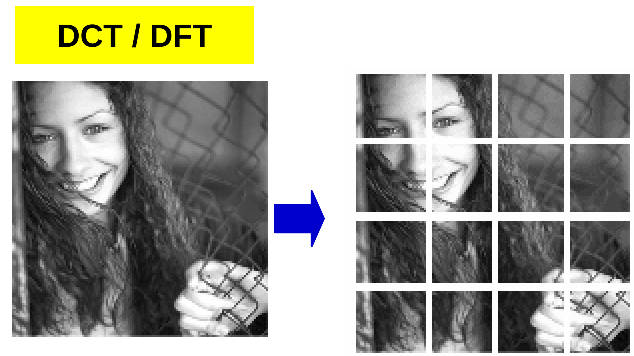
- This signal can be decomposed into two signals (long term and short term), each can be simulated by an appropriate model.

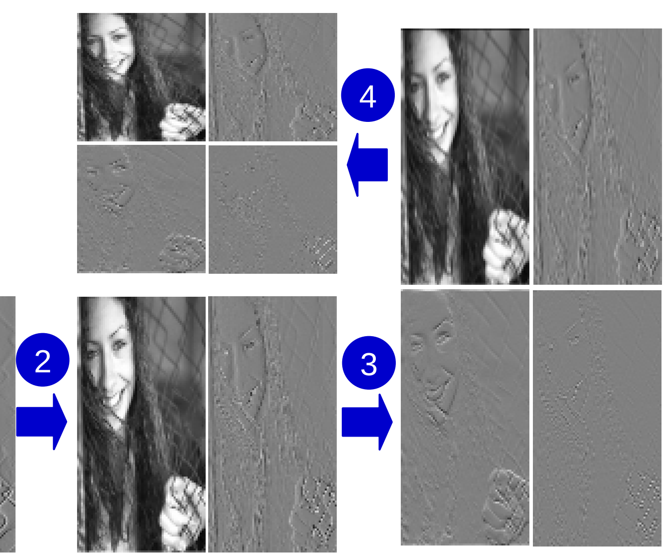
DCT/DFT Spatially Decomposition

- DCT/DFT spatially decompose the image into blocks.
  - Perform frequency domain processing in each block.
  - Discard high frequency information in each block, because human eyes are insensitive to high frequencies.
  - However, there are artifacts at the block edges.

Frequency Domain Decomposition

- Instead of spatially divide the image into blocks (such as DCT/DFT), we can consider frequency domain processing on the whole image.
  - Filtering the image horizontally into two subbands (images). 1
  - Down-sampling horizontally. 2
  - Filtering subbands vertically. 3
  - Down-sampling vertically. 4
  - Model and code each subbands separately.
- This is the subband coding.



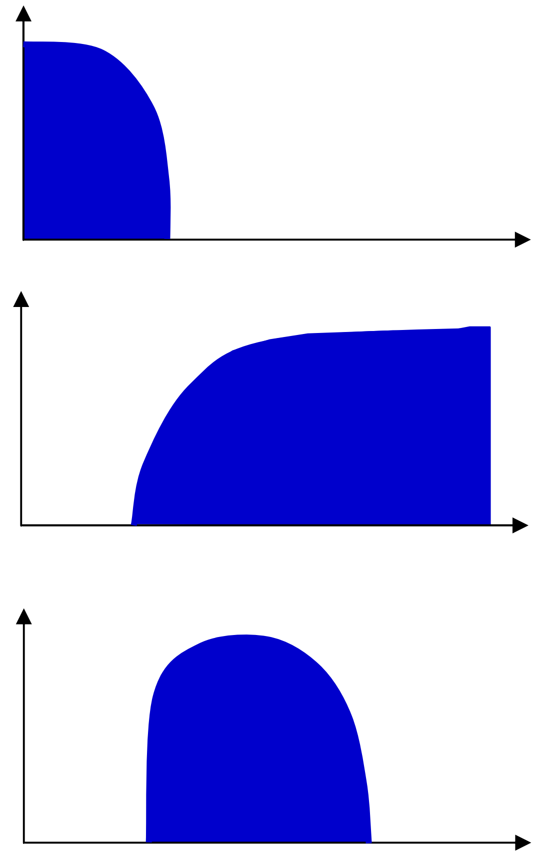


Subband Coding

- Subband coding is a technique of decomposing the source signal into constituent parts and decoding the parts separately.
- A system that isolates a constituent part corresponding to certain frequency is called a *filter*. If it isolates the low frequency components, it is called a *low-pass filter.* Similarly, we have *high-pass or band –pass filters.* In general, a filter can be called a *sub-band filter* if it isolates a number of bands simultaneously.

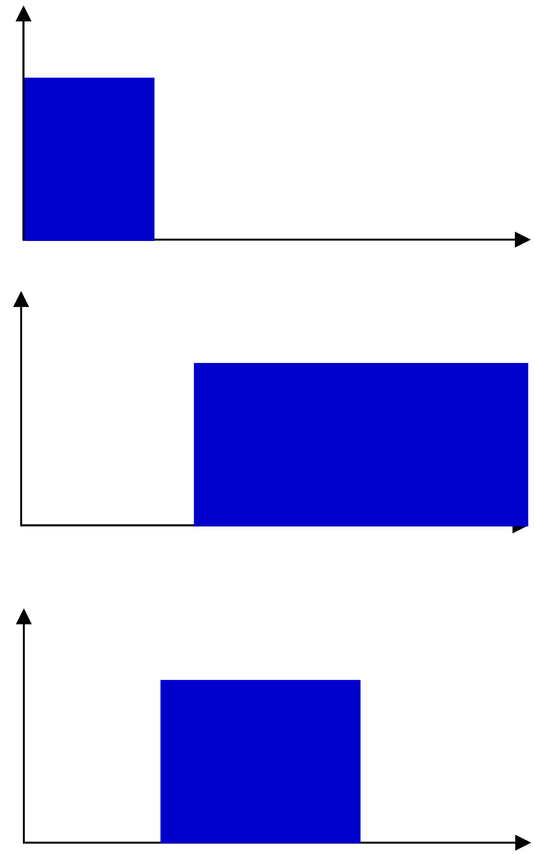
Filters

- A system that isolates certain frequency components is called a filter.
  - Low-pass filter
    - Only let through components below a certain frequency f_0
  - High-pass filter
    - Block all frequency components below a certain frequency f_0
  - Band-pass filter
    - Let through $[f_1, f_2]$



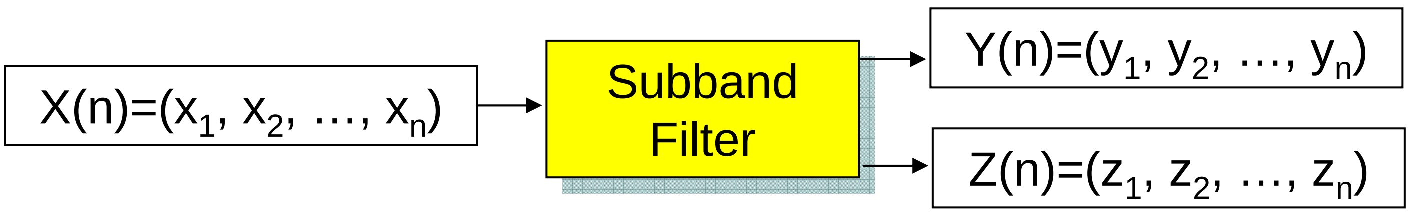
Ideal Filters

- Ideal filters have a clear cut-off frequency f_0.
  - Ideal low-pass filter
  - Ideal high-pass filter
  - Ideal band-pass filter
- Ideal filters are unrealizable in Electrical Engineering.
- However, we have “perfect” Ideal Filters in computer science.
  - Can be implemented in any programming language.



Subband Coding

- Consider the following scheme for a simplified signal system:



where

$y_1=x_1,\ y_2=(x_2+x_1)/2,\ \ldots,\ y_n=(x_n+x_{n-1})/2$

$z_1=x_1-y_1,\ z_2=x_2-y_2,\ \ldots,\ z_n=x_n-y_n$

Subband Coding

- In general, we can write

$$y_i = \frac{x_i + x_{i-1}}{2} \qquad \text{(average)}$$

$$z_i = \frac{x_i - x_{i-1}}{2} \qquad \text{(difference)}$$

$$\text{for } 1 \le i \le n \text{ and } x_0 = 0.$$

- The original signal can be recovered as

$$x_i = y_i + z_i$$

with $y_1 = x_1$ and $z_1 = 0$.

Subband Coding

- The signals y_n being averages, are much more smooth (lower frequency) and if the signals are correlated DPCM will be very effective.
- Similarly, the dynamic range of variation of z_n will be small. In fact, it is possible for the same number of bits per sample to encode both y_i and z_i with less distortion.

Subband Coding

- But, there is a problem in the above scheme –--- we now have to use vectors Y(n) and Z(n), each having n values ---- we have doubled the number of output elements!
- Let's divide $\{y_i\}$ into two sets $\{y_{2i}\},\{y_{2i-1}\}$:

  $\{y_{2i-1}\}=\{y_1,y_3, y_5,\dots\}$ -- odd sequence

  $\{y_{2i}\}=\{y_2,y_4, y_6,\dots\}$ -- even sequence

  Similarly,

  $\{z_{2i-1}\}=\{z_1,z_3, z_5,\dots\}$ -- odd sequence

  $\{z_{2i}\}=\{z_2,z_4, z_6,\dots\}$ -- even sequence

  and transmit only the odd or even numbered sequence only.

Subband Coding

- Suppose we only transmit the even sequence, we know

$$y_{2i} = \frac{x_{2i} + x_{2i-1}}{2} \qquad z_{2i} = \frac{x_{2i} - x_{2i-1}}{2}$$

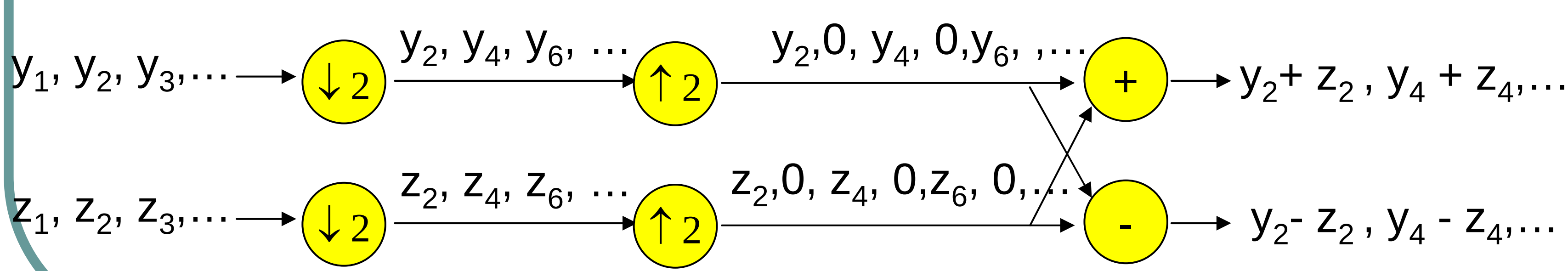
- Then

$$x_{2i} = y_{2i} + z_{2i} \qquad x_{2i-1} = y_{2i} - z_{2i}$$

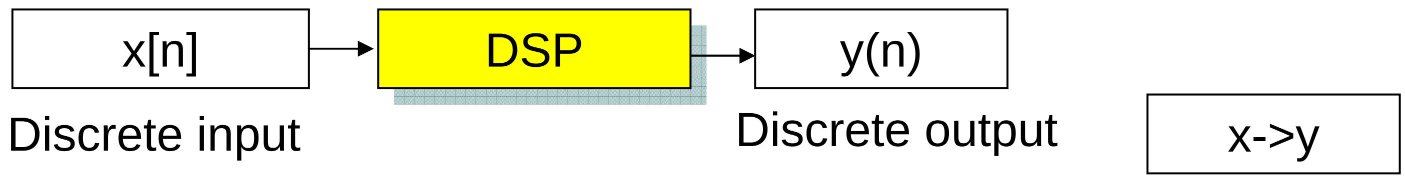
- Both odd (x_{2i-1}) and even (x_{2i}) numbered sequences are recovered. We can again use DPCM to transmit the odd or even signal attaining effective encoding as before.

Down-sampling and Up-sampling

- The process of transmitting either the odd or even sequence is called <u>down-sampling</u>, or <u>decimation</u>, denoted as $\downarrow 2$
- The original sequence can be recovered from the two down-sampled sequences by inserting 0's between consecutive samples of the two sequences, delaying one set of signals by one sample and adding two signals correspondingly. The process is called <u>up-sampling</u>, denoted as $\uparrow 2$



Quick Tutorial on DSP



- x(t): continuous time signal for all t, positive or negative.
- Discrete time signal:
  x[n] = x(nT), n = 0, 1, 2, …, -1, -2, …
- When T is the sampling period,
  $f_s = 1/T$ = sampling frequency

Nyquist Theorem

- To faithfully reconstruct x(t) from x[n], we must have $f_s >= 2f_{max}$, where f_{max} is the maximum frequency in the signal x(t).
- Sampling a signal at a frequency less than $2f_{max}$ might introduce obvious low frequency aliased signals at the output.
- To satisfy this, it is a common practice to use a low pass filter to purge the signal of all frequencies greater than half of the maximum sampling frequency.

Operation on Signals: Linear System

- Delay: x[n-k] (x[n] delayed by k sample periods)
- Add constant: x[n] + c
- Multiply by a scalar: a x[n]
- Summation : $x_1[n] + x_2[n]$

Linear System

- Superposition:
  if x_1->y_1, x_2->y_2, then x_1+x_2->y_1+y_2
- Homogenity:
  if x->y, then cx->cy

Realistic Filtering

- Output y[n] is calculated by taking a weighted sum of the following into the filter
  - current input x[n]
  - past inputs (x[0], x[1], …, x[n-1])
  - in some cases, the past outputs (y[0], y[1], …, y[n-1]) of the filter.
  - a_i, b_i are *filtering coefficients*.

$$y[n] = \sum_{i=0}^{N} a_i x[n-i] + \sum_{i=1}^{M} b_i y[n-i]$$

Impulse Response of the Filter

- If the input sequence is a single 1 followed by all 0s (1000…0), the output sequence is called the impulse response of the filter (represented as $\{h_n\}$). The input sequence is called **Impulse (delta) function.**
- Finite impulse response filter (FIR)
  - All b_i are 0.
  - The impulse will die out after N samples.
  - N is the number of taps in the filter.

$$y[n] = \sum_{i=0}^{N} a_i x[n-i]$$

- Infinite impulse response filter (IIR)
  - Any of the b_i is not 0.
  - The impulse can continue forever.

$$y[n] = \sum_{i=0}^{N} a_i x[n-i] + \sum_{i=1}^{M} b_i y[n-i]$$

Calculate Impulse Response

- Assume a_0=1, b_1 = 0.9, all other a_i, b_i are 0.
- Input x[0] = 1, x[1] = x[2] =x[3] = … = x[n] = 0
- Output:
  - $y_0=a_0x[0]$ = 1*1 = 1
  - $y_1=a_0x[1]+b_1y[0]$ = 1*0 + 0.9*1 = 0.9
  - $y_2=a_0x[2]+b_1y[1]$ = 1*0 + 0.9*0.9 = 0.81
  - …
  - $y_n=a_0x[n]+b_1y[n-1] = 1*0 + 0.9*0.9^{n-1} = 0.9^n$

$$y[n]=\sum_{i=0}^{N}a_i x[n-i]+\sum_{i=1}^{M}b_i y[n-i]$$

- So, we have the impulse response

$$h_n=\begin{cases}0 & n<0\\ 0.9^n & n\geq 0\end{cases}$$

Restore the Signal from the Impulse Response Function

- Output $\{y_n\}$ can be restored from the input $\{x_n\}$ and the impulse response function $\{h_n\}$:

$$y[n] = \sum_{k=0}^{M} h_k x[n-k]$$

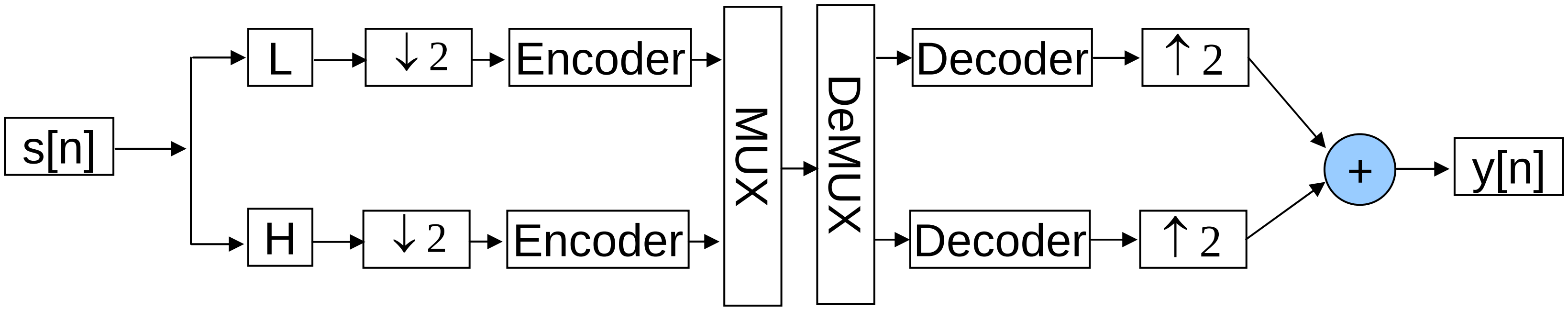
  - M is
    - finite for the FIR
    - Infinite for the IIR

- Termed convolution.
  - Good for hardware implementation.

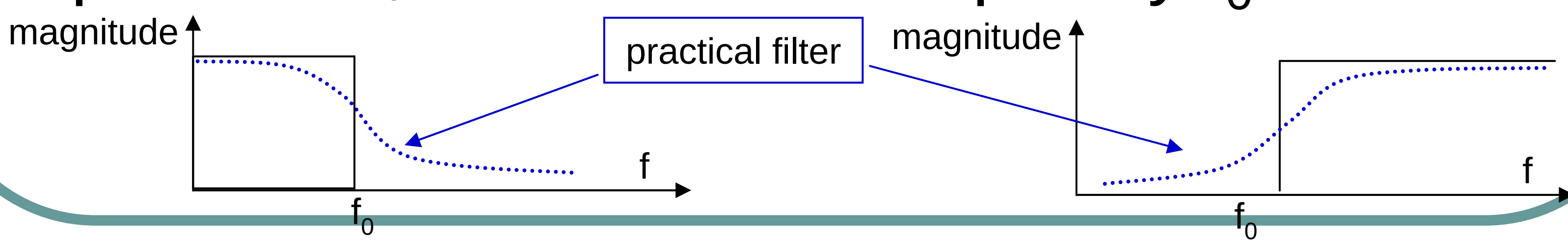
Subband Coding

- The principle of splitting a discrete time signal into a number of subband signals and combining the subband signals into final output signal has led to development of ‘filter bank’ system of analysis and synthesis for discrete signal processing (DSP).
- The basic two channel subband codec (coder/decoder) based on a two-channel QMF (quadrature mirror filter) bank is shown below.

Quadrature Mirror Filter



- The decomposition is usually done in the frequency domain. The input signal x[n] is first passed through a two-band analysis filter bank containing L, a low-pass filter and H, a high pass filter, with a cut-off frequency f_0.



Quadrature Mirror Filter

- The subband signals are down-sampled (decimated) by a factor of 2. Each down-sampled subband signal is encoded/quantized by special characteristics of the signal, such as energy level and perceptual importance. The coded subband are then combined into one sequence by a multiplexer.
- At the receiving end, the coded subband signals are first recovered by demultiplxing.

Quadrature Mirror Filters

- The subband coding uses a cascade of stages.
- Each stage contains a pair of lowpass and highpass filters. The most commonly used filters are called Quadrature Mirror Filter (QMF).
- QMF Filters are mirror symmetric.
  - Impulse response of the low pass filter $\{h_n\}$
  - Impulse response of the high pass filter $\{H_n\}$

$$H_n = [-1]^n h_{N-1-n} \qquad n = 0, 1, \cdots, N-1$$

$$h_{N-1-n} = h_n \qquad n = 0, 1, \cdots, \frac{N}{2} - 1$$

Example of Quadrature Mirror Filters

$$H_n = [-1]^n h_{N-1-n} \qquad n = 0, 1, \cdots, N-1$$

$$h_{N-1-n} = h_n \qquad n = 0, 1, \cdots, \frac{N}{2} - 1$$

- Tap N = 8.
- Low pass
  - h_0= 0.00938715
  - h_1= 0.06942827
  - h_2= -0.07065183
  - h_3= 0.48998080
  - h_4= 0.48998080
  - h_5= -0.07065183
  - h_6= 0.06942827
  - h_7= 0.00938715
- High pass
  - H_0= -0.00938715
  - H_1= -0.06942827
  - H_2= 0.07065183
  - H_3= -0.48998080
  - H_4= -0.48998080
  - H_5= 0.07065183
  - H_6= -0.06942827
  - H_7= -0.00938715

Subband Coding

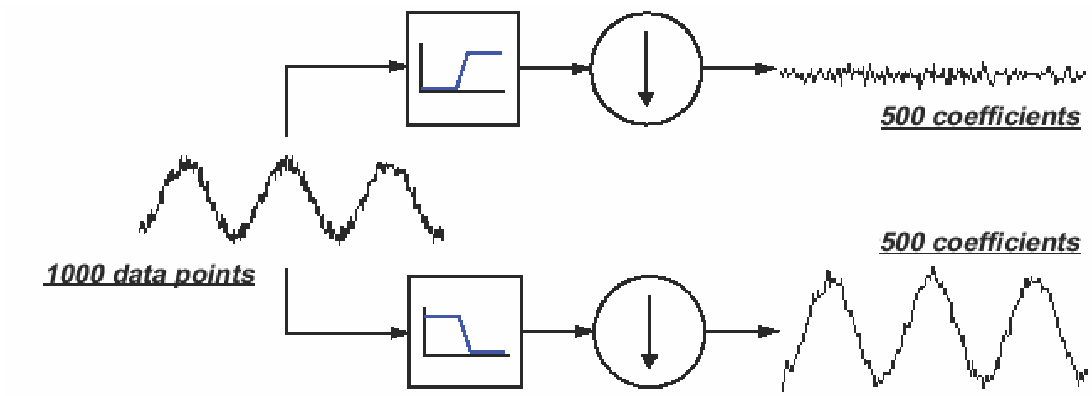
- Image is first filtered to create a set of images:
  - Each image contains a limited range of spatial frequencies.
  - Images are called subbands.
  - Each subband has a reduced bandwidth compared to the original full-band image.
  - The image may be down-sampled.
  - Filtering and sub-sampling is termed analysis stage.
- Each subband is encoded
  - using one coder.
  - or, using multiple coders (recursive subband coding).

Subband Coding

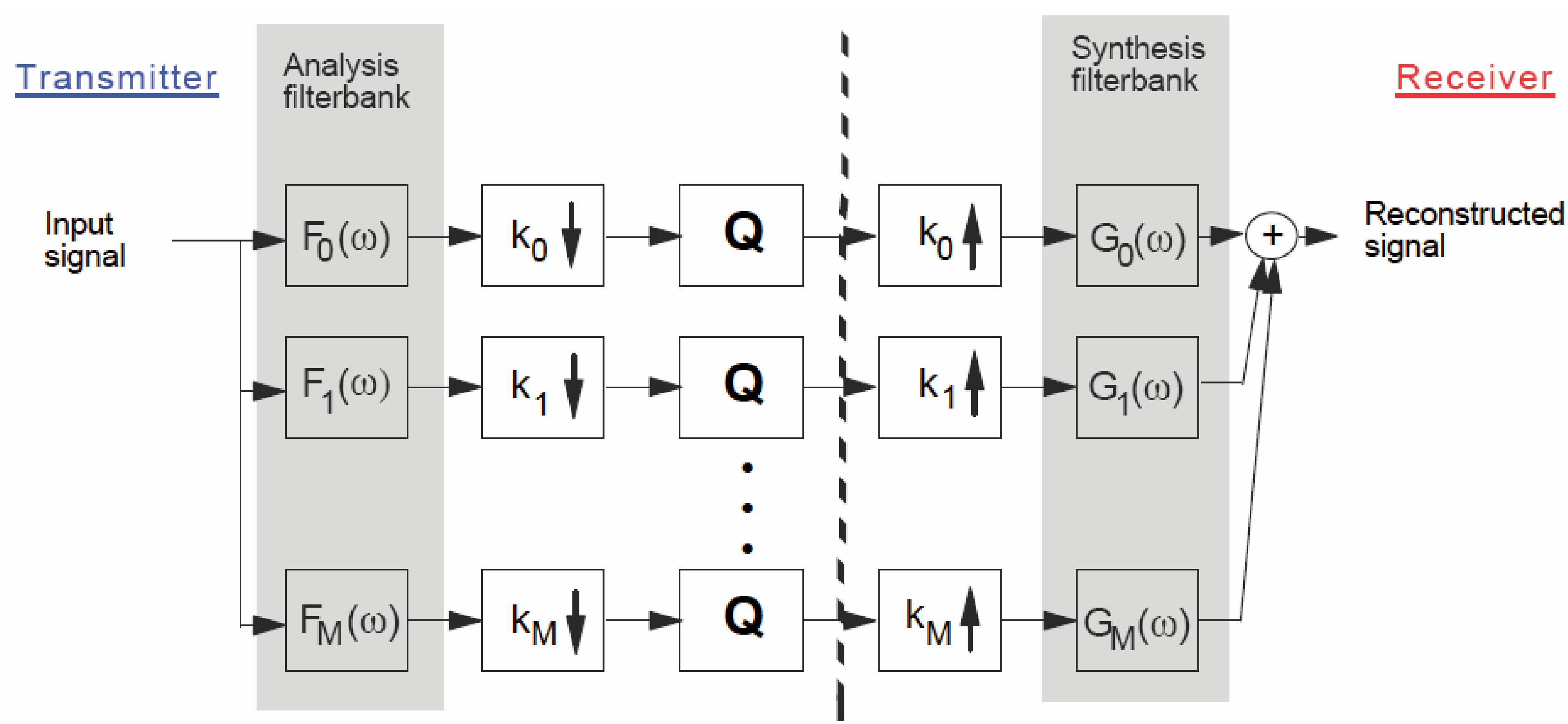
- Different bit rates or even different coding techniques can be used for each subband.
  - Takes advantages of properties of the subband.
  - Allowing for the coding errors to be distributed across the subbands in a visually optimal manner.
    - Error in one subband will be distributed onto the whole image in the reconstruction stage.
    - In construction phase, error in the DCT is confined in each blocks.
- Reconstruction is achieved by up-sampling the decoded subbands, applying appropriate filters and adding the reconstructed subbands together
  - Termed synthesis stage

Subband Coding

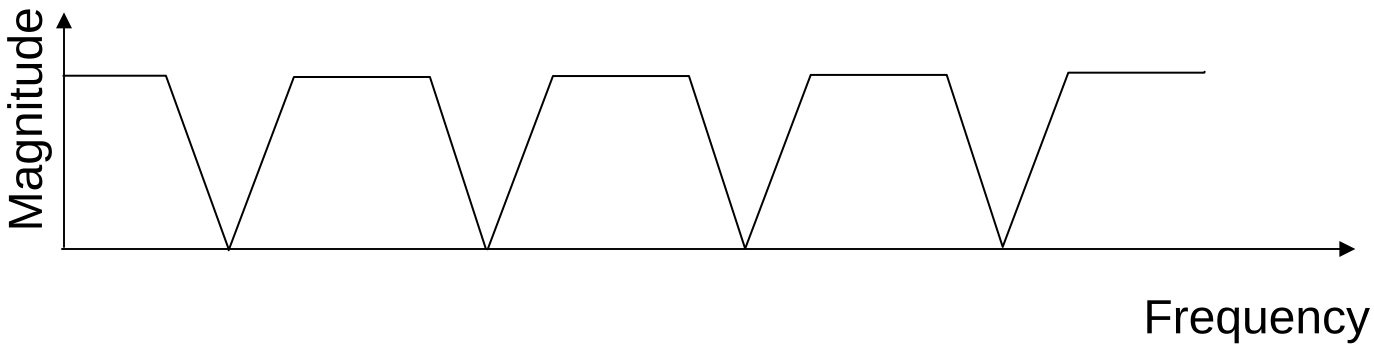
- Formulation of subbands **DOES NOT** create any compression.
  - Same number of samples is required to represent the subbands as is required for the original images.
- Subbands can be encoded more efficiently than the original image.

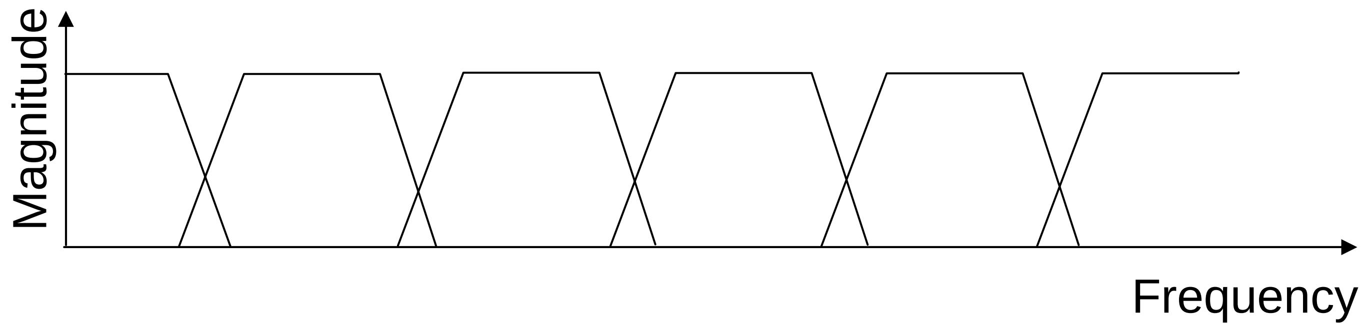


1D Multi-band Coding



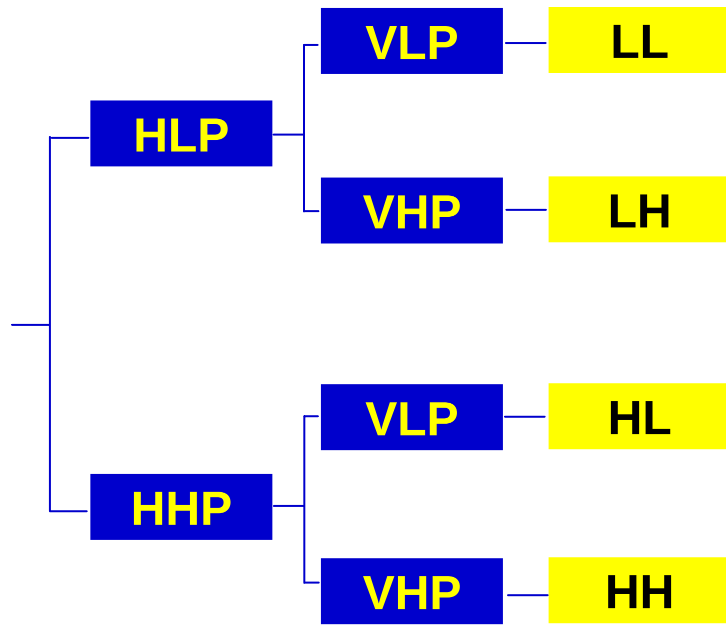
Nonoverlapping/overlapping Filter Banks

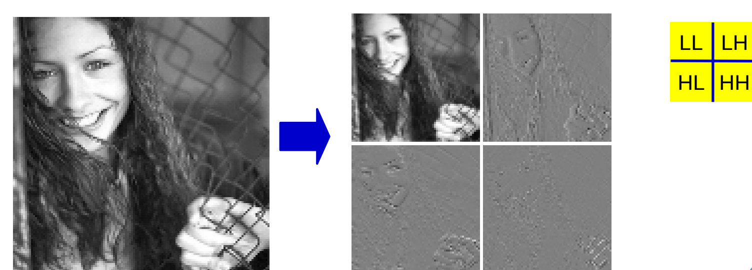




2D, Two-band Subband Coding

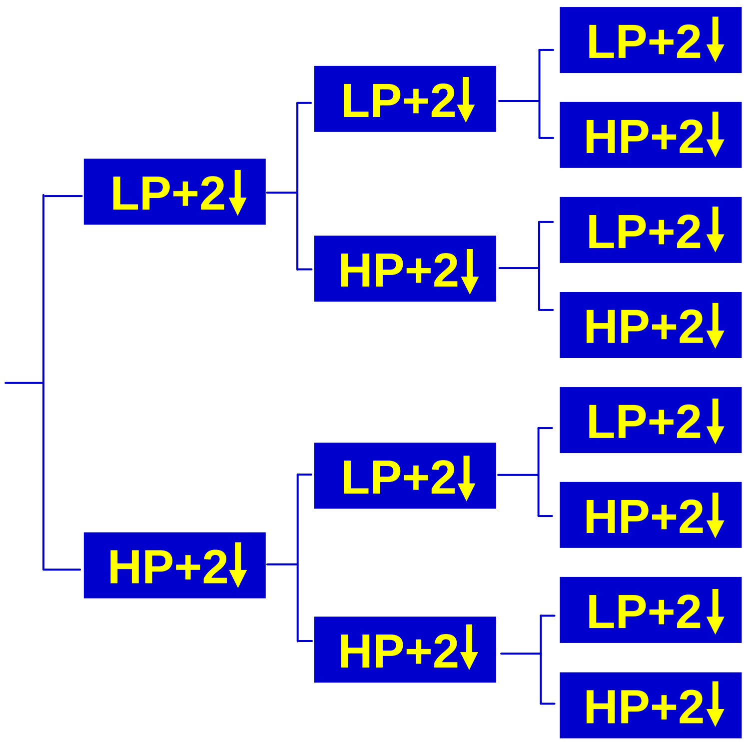
- Separable filter
  - Horizontal low-pass filter (HLP)
  - Horizontal high-pass filter (HHP)
  - Vertical low-pass filter (VLP)
  - Vertical high-pass filter (VHP)
- YES. There are 2D nonseparable filters.
  - Drawback: high computational complexity.





Cascaded Filter Banks

- The most frequently used filter banks in subband coding consist of a cascade of images
  - Each stage consists of a low-pass filter and a high-pass filter.
  - Down sampling.



Haar Transform and Filter Banks

- We have seen the convolution operator earlier

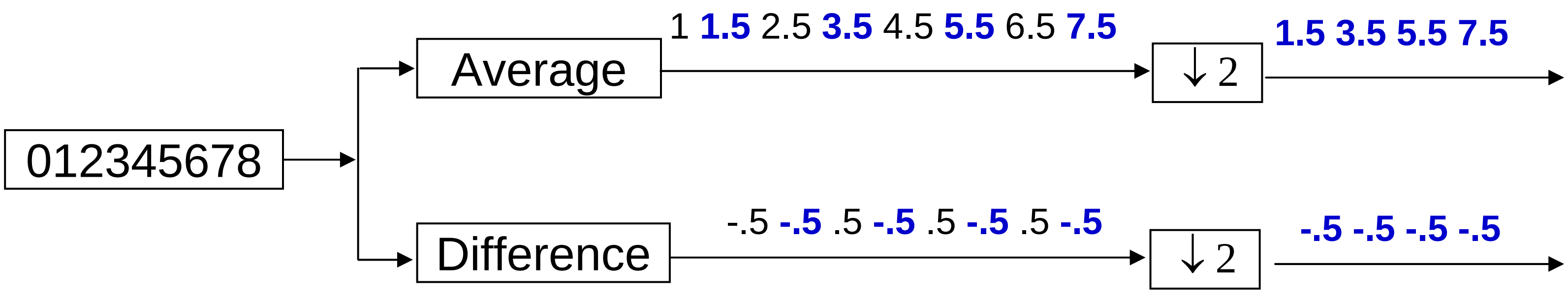
$$y[n] = \sum_{k=0}^{M} h_k x[n-k]$$

  where h_0, h_1, … are the filter coefficients.
- Let's define two 2-tap filters as follows

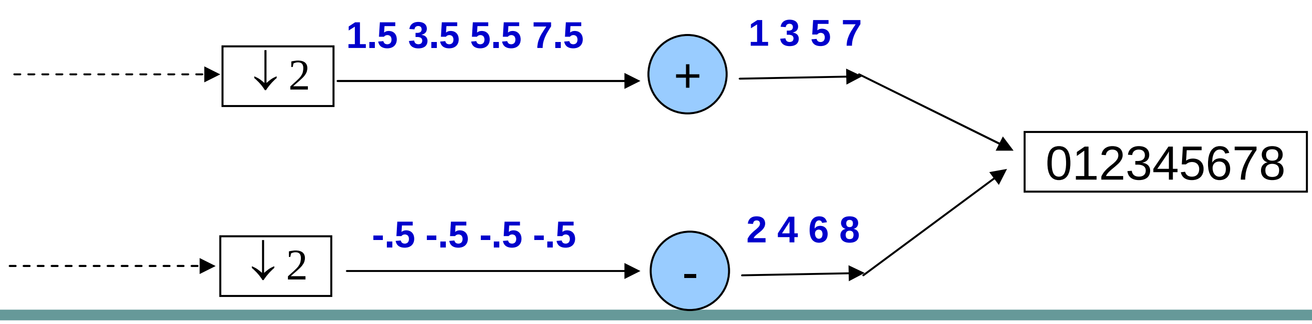
- L: y[n] = 0.5 * x[n] + 0.5 * x[n-1]
  H: y[n] = 0.5 * x[n] - 0.5 * x[n-1]

- L is the 'average' of signal value x[n] with the one unit delayed x[n-1] producing y[n]. L is recognized as a 'low-pass' filter. Similarly, H is the 'difference' function recognized as a 'high-pass' filter.

Down-sampling and Up-sampling

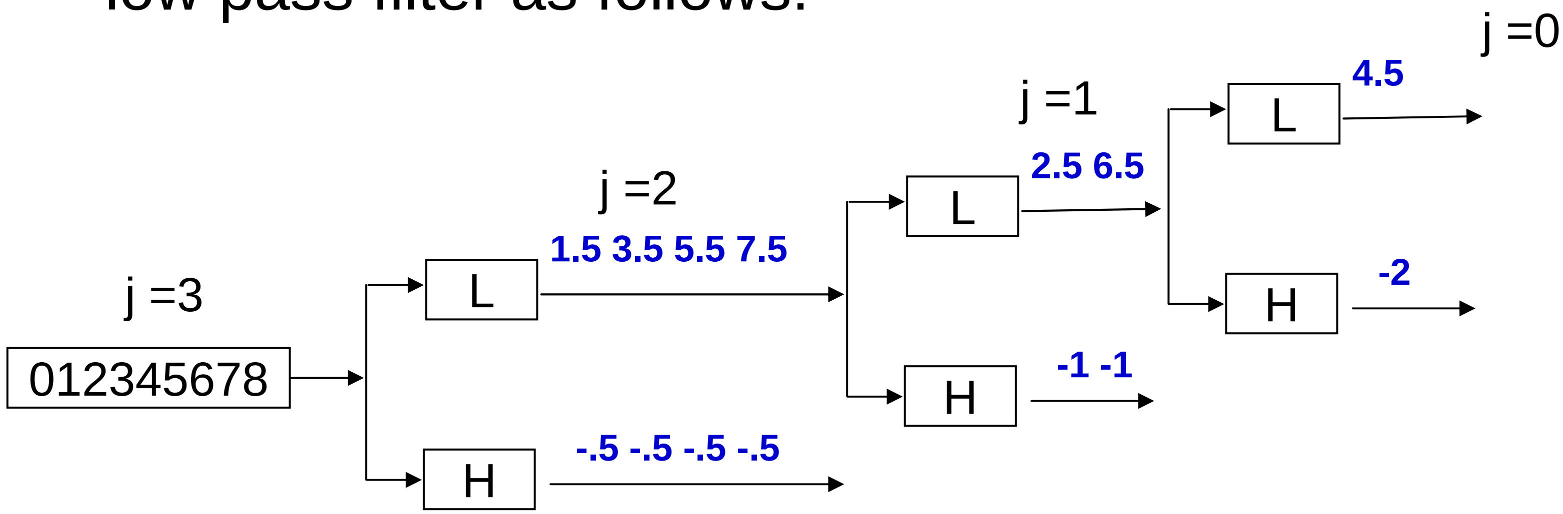


- We really do not need to compute the decimated output.
- Up-sampling and Synthesis Filter



Logarithmic Tree

- We can iterate the filter operations on the output of the low pass filter as follows:



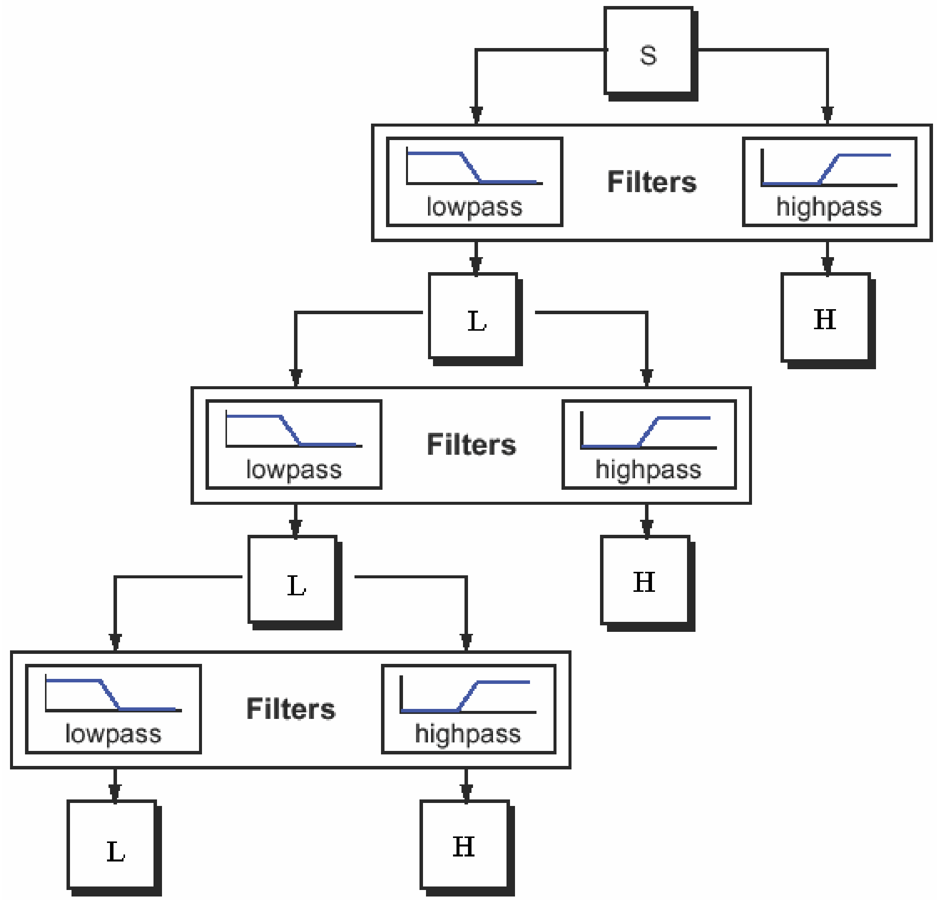
- Output: 4.5 -2 -1 -1 -.5 -.5 -.5 -.5
- The logarithmic tree is also called the multiresolution tree.

One-dimensional Haar Wavelet Transform

- Given input value {1, 2, 3, 4, 5, 6, 7, 8} (resolution 8)
- Step #1 (resolution 4)
  - Output Low Frequency {1.5, 3.5, 5.5, 7.5} - average
  - Output High Frequency {-0.5, -0.5, -0.5, -0.5} – detail coefficients
- Step #2 (resolution 2)
  - Refine Low frequency output in Step #1
    - L: {2.5, 6.5}- average
    - H: {-1, -1} - detail
- Step #3 (resolution1)
  - Refine Low frequency output in Step #2
    - L: {4.5} -average
    - H: {-2} - detail
    - Transmit { 4.5, -2, -1, -1, -0.5, -0.5, -0.5, -0.5}. No information has been lost or gained by this process. We can reconstruct the original image from this vector by adding and subtracting the detail coefficients. The vector has 8 values, as in the original sequence, but except for the first coefficient, all have small magintudes..

Sub-band Interpretation of Wavelet Transform

- The computation of the wavelet transform used recursive averaging and differencing coefficients. It behaves like a **filter bank**.
- Recursive application of a two-band filter bank to the lowpass band of the previous stage.



Matrix Operation

- The discrete wavelet transform can also be described in terms of matrix operation.
- The Haar filter operation for j = 1 is equivalent to multiplying the two input by the matrix

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

Matrix Operation

- Similarly, for j = 2 and j = 3, the corresponding matrices are

$$\frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \qquad \frac{1}{\sqrt{8}}\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Generalization of other filters

- The matrix formulation extends to other kinds of filters as well. For example, the Daubechi's D4 filter has four coefficients:
  c_0 = 0.48296
  c_1=0.8365
  c_2 = 0.2241
  c_3=-0.1294
- The transform matrix has the form:

$$w = \begin{bmatrix} c0 & c1 & c2 & c3 & 0 & 0 & \cdots & 0 \\ c3 & -c2 & c1 & -c0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & c0 & c1 & c2 & c3 & \cdots & 0 \\ 0 & 0 & c3 & -c2 & c1 & -c0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & c0 & c1 & c2 & c3 \\ 0 & 0 & \cdots & 0 & c3 & -c2 & c1 & -c0 \\ c2 & c3 & 0 & 0 & \cdots & 0 & c0 & c1 \\ c1 & -c0 & 0 & 0 & \cdots & 0 & c3 & -c2 \end{bmatrix}$$

Generalization of other filters

- If the input vector $X = (x_1, x_2, \ldots, x_n)$ is multiplied with W, we get
- H: smooth coefficients

  $s_1 = c_0x_1 + c_1x_2 + c_2x_3 + c_3x_4$

  $s_3 = c_0x_3 + c_1x_4 + c_2x_5 + c_3x_6$, etc.
- G: detail coefficients

  $d_1 = c_3x_1 - c_2x_2 + c_1x_3 - c_0x_4$

  $d_3 = c_3x_3 - c_2x_4 + c_1x_5 - c_0x_6$, etc.

Generalization of other filters

- Note, these smooth and detail coefficients are convolutions of data with the four coefficients. Together H and G form a QMF.
- The coefficient values have been derived from orthonormality conditions and therefore, the inverse of W is W^T.
- Since the size of W is that of the image, it may seem impractical from storage point of view. But, the matrix is very regular. Given the top row of W, all other rows can be generated by simple shifting, reversing and changing signs.

Coding of the Subbands

- Coding of the subbands is done using a method and bit rate most suitable to the statistics and visual significance of that subband.
- Typically 95% of the image energy is in the low frequency bands.
- Low bands may use Transform, DPCM or VQ.
- Other bands may use PCM or run-length coded after coarse thresholding.

Application of Subband Coding

- Speech coding
  - ITU-T G.722
  - Encode high-quality speech at 64/56/48 kbps.
- Audio coding
  - MPEG Audio
    - Layer 1
    - Layer 2
    - Layer 3
- Image compression
  - Closely related to wavelet transform coding.

Subband Coding and Wavelet Transform Coding

- Filters used in subband coders are **not** in general orthogonal.
- Transform coding is a special case of subband coding
- Wavelet transform coding is closely related to subband coding.
  - Alternative to filter design.
- Wavelet transform describes multiresolution decomposition in terms of expansion of an image onto a set of wavelet basis functions.
  - Basis functions well localized in both space and time.
- Wavelet transform-based filters possess some regularity properties that not all QMF filters have.
  - Improved coding performance over QMF filters with same number of taps.