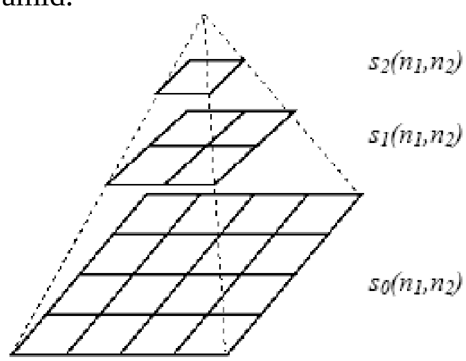
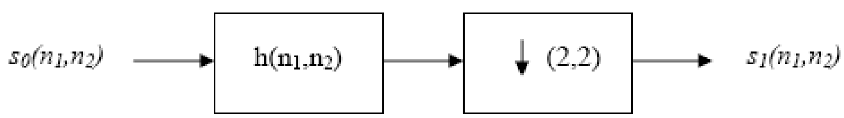
Chapter 8: Multiresolution Image Representations[1]

The main purpose of this section is to introduce the theory of multi-resolution image decompositions, filterbanks and wavelets.

Pyramid Representations

Gaussian Pyramid: In a Gaussian pyramid multiresolution image representation, the original $N_1 x N_2$ image $s_0(n_1,n_2)$ appears at the bottom of a stack of images. This image is low-pass filtered and subsampled by a factor of 2 in each direction. The resulting $N_1/2xN_2/2$ image $s_1(n_1,n_2)$ appears at the second level of the pyramid. This procedure is repeated as many times as desired. If a low pass filter with a truncated Gaussian impulse response is used, then the resulting pyramid is generally known as a Gaussian pyramid.





In order to eliminate aliasing an ideal low pass filter with cut off frequency $w_c = \pi/2$ is required at every stage. However, the ideal filter is infinite extent and non-causal, thus it cannot be realized. Hence, a length *2L+1* truncated Gaussian filter, with the impulse response:

$$h(n) = C.e^{-\frac{n^2}{2\sigma^2}} \quad for\ |n| \le L \qquad \text{where the normalization constant is: } C = \frac{1}{\sum_{|n|\le L} e^{-\frac{n^2}{2\sigma^2}}}$$

This filter is often used to filter the image horizontally and then vertically. A common choice is a 3-tap Gaussian filter where **h(n)**={ ¼, ½, ¼}.

The Gaussian pyramid is labeled as an over-complete (redundant) representation since the total number of pixels in the pyramid approaches: $N_1N_2 + N_1N_2/4 + N_1N_2/16 + \cdots \approx \frac{3}{4}N_1N_2$, which is larger than which is larger than $N_1 N_2$. Furthermore, since the Gaussian filter has significant leakage beyond the frequency $w_c = \pi/2$, images in the upper levels has aliasing.

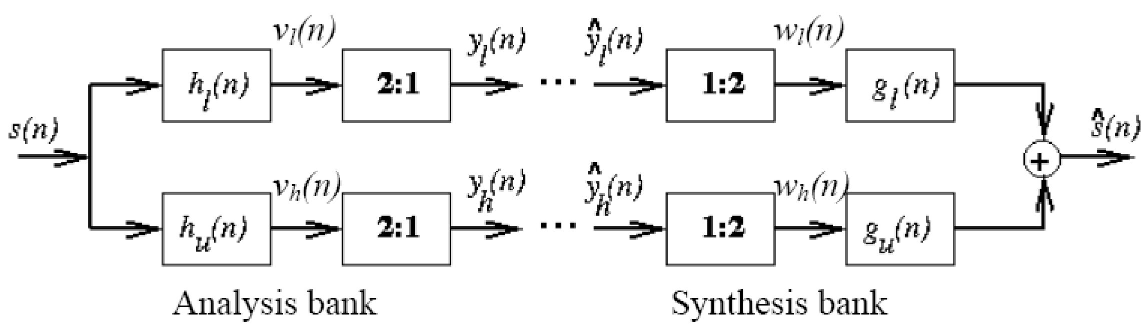
Gaussian Pyramid: The Laplacian pyramid contains differences between successive levels of the Gaussian pyramid. In order to construct a Laplacian pyramid, we start with the lowest resolution image in the Gaussian pyramid, and interpolate it by a factor of *2* in each direction, then take the

[1] This chapter is primarily based on lecture notes provided by Prof. A. Murat Tekalp of Koç University, Istanbul, Turkey.

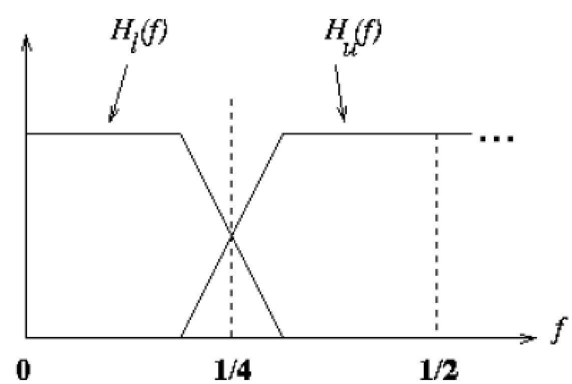
difference between this interpolated image and the same size image (next higher resolution level) in the Gaussian pyramid. The linear interpolation filter with the impulse response **h(n)**={ ½, 1, ½ } is often used for interpolation.

An alternative approach to obtaining the difference images is filtering the original image by the difference of two Gaussian filters or a Laplacian filter, hence the name Laplacian pyramid or a bandpass pyramid. Thus, a *3-level* Laplacian pyramid consists of one low-resolution picture and two successively larger difference images. The difference images contain image detail that is significant at each scale. In theory, it is possible to recover the full-resolution image by adding the difference images to the interpolated images at each level successively. Thus, Laplacian pyramid can be used for progressive image compression and transmission.

Filter Banks: The over-completeness of the pyramid representation often is not desirable. The subband and wavelet representations are complete multi-resolution representations, where the number of pixels in the multiresolution representation is exactly the same as in the original image. The benefit of a complete (non-redundant) representation comes at the expense of more complex filtering requirements. The filterbanks are the building blocks of the subband and wavelet representations. A 2-channel analysis-synthesis filterbank is shown below.

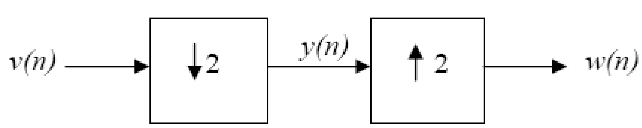


Each filterbank consists of a pair of low-pass and high-pass filters, whose frequency responses are shown below.



We will first provide a mathematical analysis of a 2-channel analysis-synthesis filterbank. Then, we discuss procedures for filter design for multiresolution image generation.

Analysis of 2-Channel Filterbank: Consider the following system:



The intermediate sequence *y(n)* is simply: $y(n) = v(2n)$ and

$$w(n) = \begin{cases} y(n/2) & n = 0,2,4,\cdots \\ 0 & Otherwise \end{cases} = \begin{cases} v(n) & n = 0,2,4,\cdots \\ 0 & Otherwise \end{cases}$$

and the z-transform yields the discrete transfer function:

$$W(z) = \sum_i v(i).z^{-2i} = V(z^2) = \frac{1}{2}[V(z) + V(-z)]$$

The output *w(n)* can be written as follows[2]:

$$W_l(z) = \frac{1}{2}[H_l(z).S(z) + H_l(-z).S(-z)]$$

$$G_l(z).W_l(z) = \frac{1}{2}[G_l(z).H_l(z).S(z) + G_l(z).H_l(-z).S(-z)]$$

$$G_h(z).W_h(z) = \frac{1}{2}[G_h.H_h(z).S(z) + G_h.H_h(-z).S(-z)]$$

$$\hat{S}(z) = \frac{1}{2}[G_l(z).H_l(z) + G_h.H_h(z)].S(z) + \frac{1}{2}[G_l(z).H_l(-z) + G_h.H_h(-z)].S(-z)$$

Alias Cancellation Condition: Since the second term is due to aliasing left after analysis-synthesis filtering, the aliasing can be eliminated if the filter transfer functions satisfy the condition:

$$G_l(z).H_l(-z) + G_h.H_h(-z) = 0$$

Perfect Reconstruction Condition: It is possible if the filter transfer functions also satisfy the condition:

$$G_l(z).H_l(z) + G_h.H_h(z) = 2.z^{-k}$$

where k is the amount of delay. Then the output of the analysis-synthesis filterbank is equal to the input up to a scale factor of 2. That is, $\hat{s}(n) = 2.s(n-l)$.

Other Requirements:

1. It is good to have all filters are from FIR-class: $G_l(z), H_l(z), G_h, H_h(z)$
2. They exhibit linear phase, which implies that their impulse responses are symmetric or anti-symmetric.
3. They are orthonormal, which implies that the energy of the signal is preserved under the transformation.

$$< h_l(n), h_l(n+2k) >=< h_h(n), h_h(n+2k) >$$

$$< h_l(n), h_h(n+2k) >= 0$$

Orthogonal FIR Filterbanks: For orthonormal FIR filterbanks, synthesis filters are chosen as time-reversed versions of analysis filters, and they are given by:

$$g_l(n) = h_l(-n)$$

and $g_h(n) = h_h(-n)$

Furthermore, high-pass filters are modulated versions of the low-pass filters, given by:

$$h_h(n) = (-1)^n.h_l(-n+k)$$

[2] The reader is referred to p. 66 and p. 111 for details in M. Vetterli, and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, 1995.

Analysis filters satisfy the orthogonality condition:

$$< h_i(n-2p), h_j(n-2q) >= \delta[i-j].\delta[p-q]$$

Equivalently, in the z-domain:

$$H_2(z) = -z^{-1}.H_1(-z^{-1}); \quad G_1(z) = H_1(z^{-1}); \quad G_2(z) = -z.H_1(-z)$$

which are also known as Smith-Barnwell Filters for perfect reconstruction (filter lengths must be even.)

$$|H_1(e^{jw})|^2 + |H_1(e^{j(w+\pi)})|^2 = |H_1(e^{jw})|^2 + |H_2(e^{jw})|^2 = 2$$

Linear-Phase FIR Filterbanks; Bi-Orthogonal Filter Banks: Perfect reconstruction

$$< h_i[-n], g_j[n-2p] >= \delta[i-j].\delta[p], \quad i, j = 1,2$$

or equivalently in the z-domain:

$$G_1(z).H_1(z) + G_2(z).H_2(z) = 2$$

$$G_1(z).H_1(-z) + G_2(z).H_2(-z) = 0$$

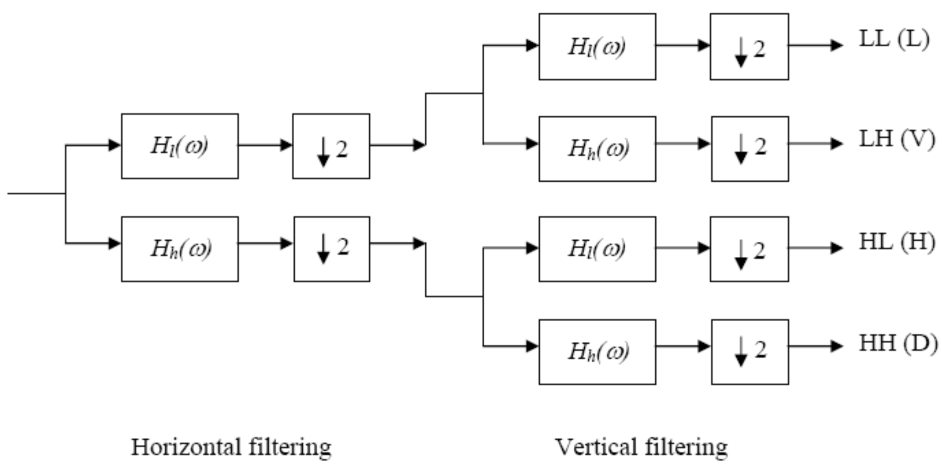
QMF Filters: Alias-free reconstruction, linear phase

$$H_2(z) = -z^{-1}.H_1(-z^{-1}); \quad G_1(z) = G_2(z); \quad G_2(z) = -z.H_2(z) = -H_1(z)$$

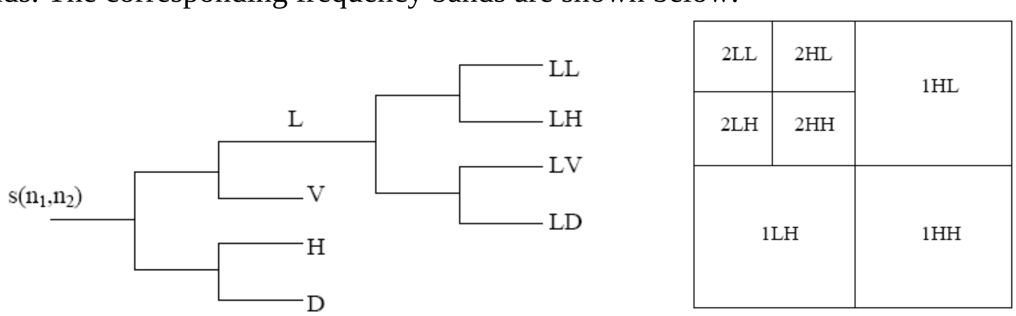
$$|H(z)|^2 - |H(-z)|^2 = 2.z^{-k}$$

Second relation can only be approximated (not perfect reconstruction), e.g., Johnston filters

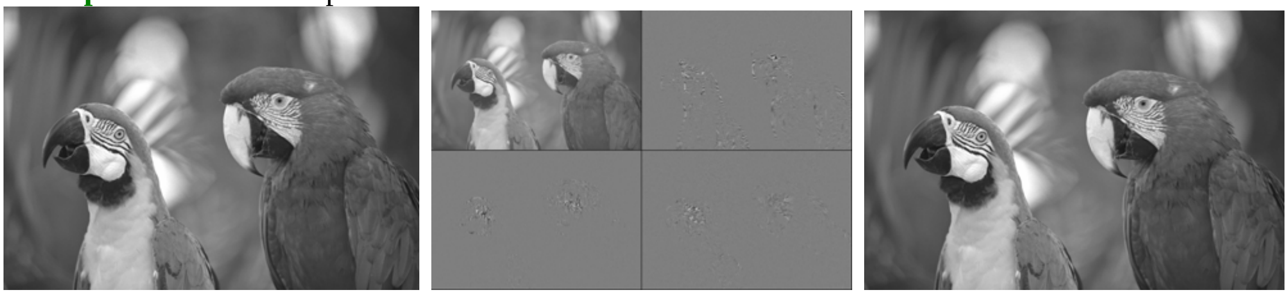
Wavelet Transform: Wavelet filters possess an additional regularity constraint, that is, the prototype filter $H_1(w)$ has a zero at the frequency $w = \pi$. In order to decompose an image into a wavelet representation, the same filters are used both horizontally and vertically as follows:



Moreover, the composition is done in multiple stages, where the LL band is further decomposed into four bands. The corresponding frequency bands are shown below.

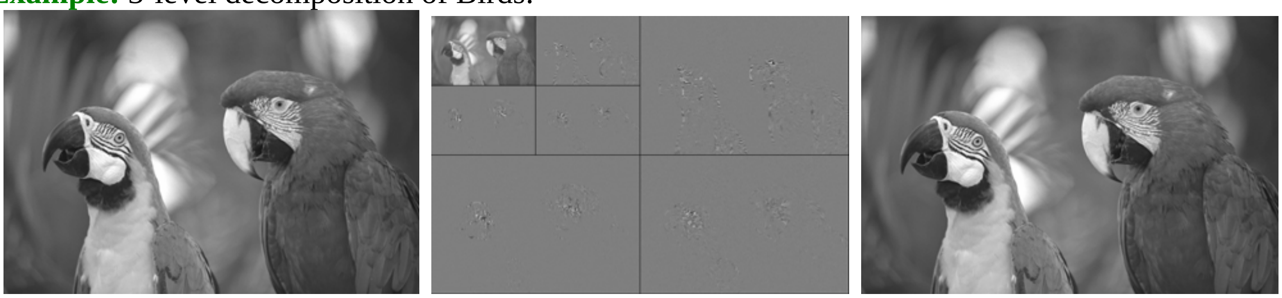


Example: 2-level decomposition of Birds[3]:



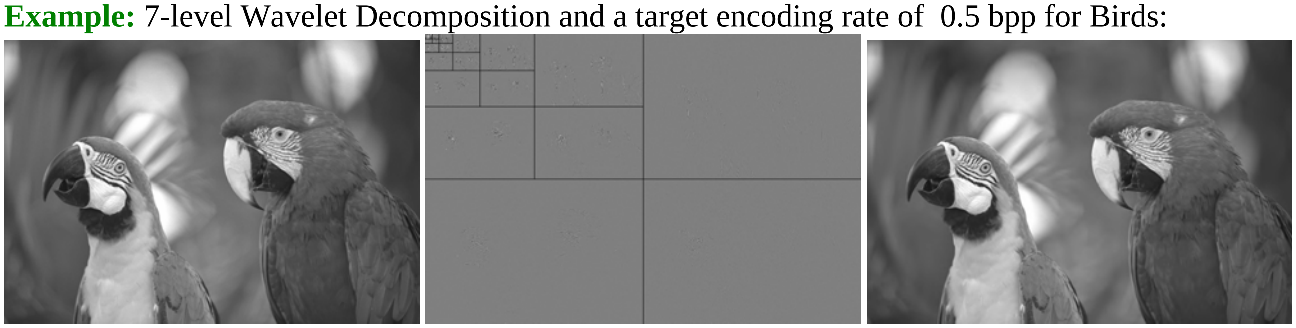
Coded Bitrate : 8.00 (bpp)
Mean Square Error : 0.0
Signal-to-Noise Ratio: 52.9 (dB)
PSNR : 67.7 (dB)

Example: 3-level decomposition of Birds:



Coded Bitrate : 8.00 (bpp)
Mean Square Error : 0.0
Signal-to-Noise Ratio: 47.3 (dB)
PSNR : 62.1 (dB)

Example: 7-level Wavelet Decomposition and a target encoding rate of 0.5 bpp for Birds:



SPIHT Encoder:
Target bit rate : 0.50 (bpp)
Target file size :196608 bits

SPIHT DECODING RESULTS:
#Wavelet resolution levels : 7
Decoded Bitrate : 0.50 (bpp)
Smoothing level : 0
Mean Square Error : 4.2
Signal-to-Noise Ratio : 27.1 (dB)
PSNR : 41.9 (dB)

Example: Wavelet Toolbox demo examples in Matlab.

---

[3] These examples are generated using **VCDemo** from Delft University of Technology, Holland. (Courtesy of Prof. Reginald Lagendijk.) URL: http://www-ict.its.tudelft.nl/vcdemo

Analysis of 2-level wavelet decomposition: Let $G_1(z) = -H_2(-z); G_2(z) = -H_1(-z)$
Then the perfect reconstruction (PR) condition becomes:

$$G_1(z).H_1(z) + G_2(z).H_2(z) = 2.z^{-k}$$

where k is odd integer. Let us set:

$$P(z) = z^{+k} G_1(z).H_1(z)$$

Then PR condition is given by

$$P(z) + P(-z) = 2$$

We now express *P(z)* as a series in *z*. All even powers must be zero, except the constant term which must be 1.

$$P(z) = p_{-k}.z^k + \cdots + p_{-1}.z^1 + 1 + p_1.z^1 + \cdots p_k.z^k$$

where $k = 2N + 1$, an odd integer.

$$P(z) + P(-z) = \sum_n \{p(n).z^{-n} + p(n).(-z)^{-n}\} = \sum_n 2.p(2n).z^{-(2n+1)}$$

In order to design a 2-channel perfect reconstruction (PR) filter bank for wavelet decomposition, it is necessary and sufficient to find a *P(z)* that satisfies *P(z)+P(-z)*=2, and factor it as $G_1(z).H_1(z)$ to compute the filter transfer functions.

Wavelet Filter Design Procedure:

(1) Specify the product filter *P(z)* as described above.
(2) Obtain $P_1(z) = z^{-k}.P(z)$
(3) Factor $P_1(z) = G_1(z).H_1(z)$
(4) Set $H_2(z) = G_1(-z)$ and $G_2(z) = -H_1(-z)$

Example: Haar (Daub2) Wavelets

Let $P(z) = 1/2.(z + 2 + z^{-1})$ *where* $k = 1$

Then $P(z) = 1/2.(1 + 2z^{-1} + z^{-2}) = 1/2.(1 + z^{-1})^2$

Partial fraction expansion of the last expression yields results for step (4) above:

$$H_1(z) = G_1(z) = (1/\sqrt{2}).(1 + z^{-1})$$

There are several wavelet design types and associated design algorithms, which are explained in detail in Vetterli and Kovacevic mentioned earlier as well as in Bovik[4].

---

[4] Handbook of Image and Video Processing, ed. Al Bovik, Academic Press, 2000.