Construction of Orthogonal Wavelets

Lecture #3: Tuesday, 7 October 1997
Lecturer: Denis Zorin
Scribe: Liwei He
Reviewer: Scott Cohen

In this lecture we will discuss the the criteria used in the design of wavelet filter banks. We will show the design process using the the Daubechies wavelet as an example (Figure 1).

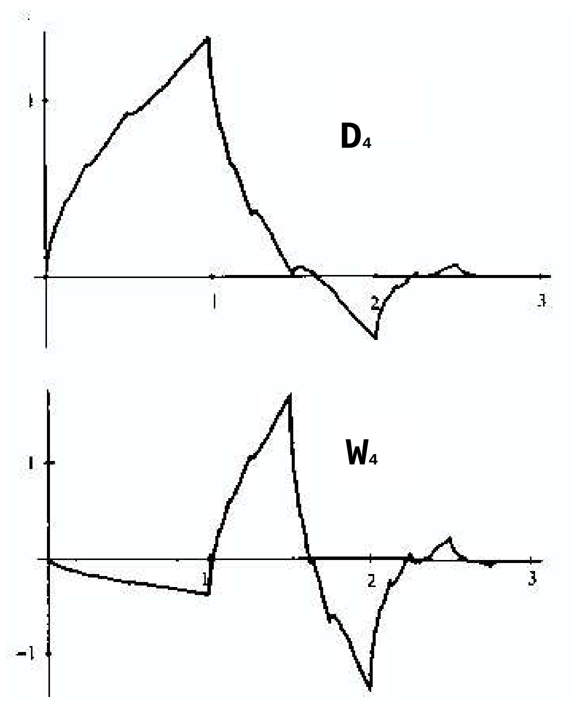

Figure 1: The scaling function of the Daubechies wavelet (for a filter of length 4) is on the top and its wavelet function is at the bottom. Note that the shape of this wavelet and the Haar wavelet resembles waves, hence the name.

1 Properties of Wavelet Bases and Filter Banks

The first step in the design process is to formulate a set of requirements for the filter bank and the associated basis. The following five properties are particularly important:

1. Finite filters. Also known as filters with Finite Impulse Response (FIR). The basis functions derived from such filters have compact support. For efficiency, we would like to make our filters as short as possible.

2. Orthogonality. For orthogonal filter banks one filter essentially defines the whole bank; choosing a synthesis filter with good numerical properties (for example, stability) guarantees that the corresponding analysis filter also has the same properties. For orthonormal bases in functional spaces it is easy to compute the coefficients: we just need to take the dot product of the input function with each basis function. Orthogonal filter banks allow implementations with minimal quantization error.

3. Good approximation. We want to have a set of basis functions that yields a good approximation to functions with as few coefficients as possible. Clearly, this cannot be achieved for arbitrary function; however, we may require that all sufficiently smooth functions are approximated with few coefficients.

4. Symmetry. Nonsymmetric low-pass filters are particularly undesirable for image-processing applications: such filters lead to "smeared" images after low-pass filtering.

5. Regularity. Sometimes it is desirable to have a smooth approximation even after many terms in the approximation are truncated. Suppose we express f as a linear combination of a wavelet basis functions f_i:

   $$f = \sum_{i=0}^{\infty} a_i f_i \quad \text{where} \quad a_i \to 0 \quad \text{as} \quad i \to \infty.$$

   The function $g = \sum_{i=0}^{k} a_i f_i$ is an approximation of the function f after we truncate the coefficients after some number k. If the basis is not regular, it is likely that g will not be smooth because the only way to build a smooth function from non-smooth basis functions is to have non-smooth features of basis functions cancel each other.

Unfortunately, the constraints imposed by the requirements of orthogonality and symmetry are too restrictive: the only filter bank that satisfies both requirements is the Haar filter bank. Further, good approximation and high regularity can be achieved only at the expense of increasing the length of the filter. A trade-off has to be made depending on a particular application. For example, the Daubechies wavelet basis has compact support, is orthogonal, and has maximal approximation order among all bases generated by filters of given size.

A set of coefficients $h(k)$ can be computed according to the requirements that we set. We will show how the coefficients of the Daubechies wavelet are constructed in Section 5.

With these coefficients, we can solve for the scaling function ϕ of the wavelet basis using the dilation equation.

2 The Cascade Algorithm

For the Haar wavelet we could guess the solution of the dilation equation. In general, solutions of dilations equations cannot be expressed using elementary functions. Rather then providing formulas for solutions, we describe an algorithm that for a large class of dilation equations yields a solution when one exists. This algorithm is called the *cascade algorithm.*

Recall that the dilation equation is

$$\phi(t) = \sqrt{2}\sum_k c(k)\phi(2t-k). \tag{1}$$

Note that we use normalized filter coefficients, so the coefficient in front of the sum in the right-hand side is $\sqrt{2}$ not 2. Given a set of coefficients $c(k)$, the cascade algorithm solves the dilation equation iteratively. The iteration begins, for example, with the Haar scaling function,

$$\phi^0(t) = \begin{cases} 1 & \text{if } t \in [0,1) \\ 0 & \text{otherwise} \end{cases}.$$

In the ith step, we plug ϕ^{i-1} into the right hand side of the dilation equation to obtain ϕ^i:

$$\phi^i(t) = \sqrt{2}\sum_k c(k)\phi^{i-1}(2t-k). \tag{2}$$

Each iteration hopefully takes us closer to the scaling function we are looking for. The cascade algorithm converges when $\phi^i = \phi^{i-1}$. In the case of the Haar filter bank ($c(0) = \frac{\sqrt{2}}{2}$, $c(1) = \frac{\sqrt{2}}{2}$), convergence is achieved in the first iteration.

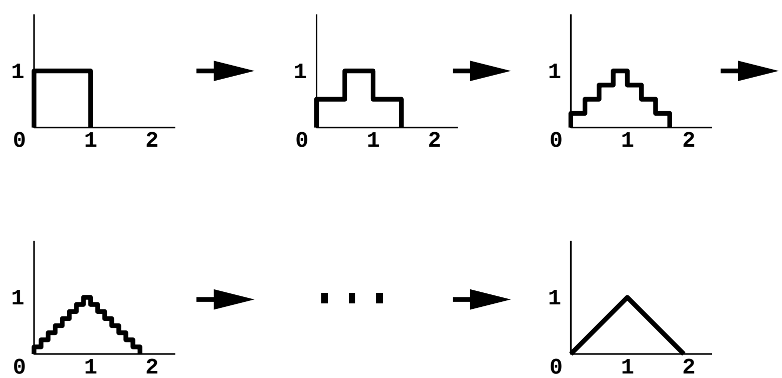

Figure 2: Cascade algorithm uses Haar scaling function as starting point and solves the scaling function iteratively. Shown here are ϕ^0, ϕ^1, ϕ^2, ϕ^3, and ϕ^∞.

A more interesting example is to use a set of coefficients where $c(0) = 1/4$, $c(1) = 1/2$, $c(2) = 1/4$.

$$\phi^1(t) = \sqrt{2}(\frac{1}{4}\phi^0(2t) + \frac{1}{2}\phi^0(2t-1) + \frac{1}{4}\phi^0(2t-2)) \tag{3}$$

In the limit, the scaling function we get, with this particular set of coefficients, is a "hat" function (see Figure 2). Remarkably, this is a linear spline basis function. All other spline basis functions can be found as solutions of dilation equations for a particular choice of coefficients.

Using the cascade algorithm, we can derive some properties of the basis functions generated by a filter bank directly from the properties of the filters, without having explicit expressions for the basis functions themselves.

3 Orthogonality of Wavelet Bases

To illustrate the idea of deriving the properties of the wavelet basis from the properties of a filter, let us prove that if a wavelet filter bank matrix is orthogonal, then the set of wavelet basis functions is orthonormal.

The coefficients to the dilation equation are found in the filter bank matrix F:

$$F = \begin{bmatrix} L \\ - \\ B \end{bmatrix} = \left[\begin{array}{cccc} c(0) & c(-1) & c(-2) & \cdots \\ c(2) & c(1) & c(0) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \hline d(0) & d(-1) & d(-2) & \cdots \\ d(2) & d(1) & d(0) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array}\right]$$

If F is orthogonal, the following relations among the coefficients must hold:

$$\begin{aligned} \sum_n c(n)c(n-2k) &= \delta(k) \qquad (4) \\ \sum_n d(n)c(n-2k) &= 0 \\ \sum_n d(n)d(n-2k) &= \delta(k) \end{aligned}$$

Note that $\delta(k)$ is 1 when $k = 0$ and is 0 elsewhere.

A wavelet basis is composed of the top level scaling functions $\phi(t-k)$ and the wavelet functions $w(2^i t - k)$ at all scales. In order to show that a wavelet basis orthonormal, we must prove that the inner products of between different function are 0, while the inner product of a basis function with itself is 1.

$$< \phi(t-k_1), \phi(t-k_2) > \; = \; \delta(k_1 - k_2) \qquad (5)$$

$$< \phi(t-k_1), w(2^i t-k_2) > \; = \; 0 \qquad (6)$$

$$< w(2^i t-k_1), w(2^j t-k_2) > \; = \; \delta(i-j)\delta(k_1 - k_2) \qquad (7)$$

We use proof by induction in the context of the cascade algorithm to show Equation (5) first. Note that it is sufficient to prove Equation (5) for $k_1 = 0$, because we can reduce the general case to the the case $t_1 = 0$ by replacing t with $t + k_1$.

We start from the base case $\phi^0(t-k)$, which are the Haar scaling functions. They are known to be orthonormal:

$$<\phi^0(t), \phi^0(t-k)> = \delta(k)$$

then we will prove that

$$<\phi^i(t), \phi^i(t-k)> = \delta(k) \quad \Longrightarrow \quad <\phi^{i+1}(t), \phi^{i+1}(t-k)> = \delta(k)$$

From the dilation equation,

$$\begin{aligned} <\phi^{i+1}(t), \phi^{i+1}(t-k)> &= \int [\sqrt{2}\sum_{k_1} c(k_1)\phi^i(2t-k_1)][\sqrt{2}\sum_{k_2} c(k_2)\phi^i(2t-k_2)] \\ &= 2\sum_{k_1 k_2} c(k_1)c(k_2)\int \phi^i(2t-k_1)\phi^i(2(t-k)-k_2)dt \end{aligned}$$

Let $T = 2t - k_1$, then $t = (T+k_1)/2$, $dt = dT/2$, and

$$<\phi^{i+1}(t), \phi^{i+1}(t-k)> = \sum_{k_1 k_2} c(k_1)c(k_2)\int \phi^i(T)\phi^i(T+k_1-k_2-2k)dT$$

By the induction hypothesis $<\phi_i(t), \phi_i(t-k)> = \delta(k)$,

$$<\phi^{i+1}(t), \phi^{i+1}(t-k)> = \sum_{k_1 k_2} c(k_1)c(k_2)\delta(k_1-k_2-2k)$$

By the definition of δ, the only non-zero terms in the summation are those when $k_2 = k_1 - 2k$,

$$<\phi^{i+1}(t), \phi^{i+1}(t-k)> = \sum_{k_1} c(k_1)c(k_1-2k)$$

Now let $n = k_1$,

$$<\phi^{i+1}(t), \phi^{i+1}(t-k)> = \sum_{n} c(n)c(n-2k)$$

This is $\delta(k)$ given by Equation (5). Thus Equation (5) is true. Equation (6) and Equation (7) can be proven by similar arguments.

4 Approximation

A good approximating basis will have only a few large coefficients for a smooth function and leave the rest relatively small. If we truncate the summation, the reconstructed function is still very close to the original f.

The quality of approximation for a basis is related to the number of the *vanishing moments* of the basis functions. We say that a function has P vanishing moments if

$$\int f_i t^k dt = 0 \quad \text{where} \quad k \in [0..P-1]$$

If $f(t)$ is smooth around some point t_0, it can be represented by a Taylor's series expansion with remainder $R(t)$,

$$f(t) = \sum_{j=0}^{P-1} \frac{f^{(j)}(t_0)(t-t_0)^j}{j!} + (t-t_0)^P R(t)$$

and the coefficient a_i for the basis function f_i can be computed by taking the dot product of $f(t)$ and each component function of the basis, since the basis is orthonormal[1]

$$\begin{aligned} a_i &= < f, f_i > \\ &= \int f(t) f_i(t) dt \\ &= \int [\sum_{j=0}^{P-1} \frac{f^{(j)}(t-t_0)^j}{j!} + (t-t_0)^P R(t)] f_i(t) dt \\ &= \sum_{j=0}^{P-1} \int \frac{f^{(j)}(t-t_0)^j}{j!} f_i(t) dt + \int (t-t_0)^P R(t) f_i(t) dt \end{aligned}$$

If the basis function f_i has P vanishing moments, all the terms in the summation will vanish since they are all linear combinations of monomials t^k where $k < P$.

$$a_i = \int (t-t_0)^P R(t) f_i(t) dt$$

Suppose the original wavelet function $w(t)$ has support I_0 and $f_i(t) = w(2^j t - k)$, with $\frac{k}{2^j} \approx t_0$. Then $f_i(t)$ has support I with $|I| = \frac{|I_0|}{2^j}$, and

$$a_i \le C \int_I |t - t_0|^P dt \le D 2^{-jP} \quad C, D \text{ are some constants not depending on } P$$

This formula indicates that the magnitude of coefficients rapidly decreases as P grows.

The following condition on filters ensures that the wavelets have P vanishing moments:

$$\sum_n (-1)^n n^k h_0(k) = 0 \quad \text{where} \quad k \in [0..P-1] \tag{8}$$

We state this condition without a proof.

5 Computing the Daubechies Filter Bank Coefficients

Now we are ready to compute the coefficients of the Daubechies filter bank. Using the cascade algorithm we can find the scaling function and the wavelet with arbitrary precision.

[1] All integrals without specified range are taken over $(-\infty \ldots +\infty)$.

The Daubechies filter bank is orthogonal and FIR, and has the best approximation for a given filter length. We compute the coefficients for the case when the filter length is assumed to be 4. Our filter bank matrix will look something like this,

$$\begin{bmatrix} c_0 & c_1 & c_2 & c_3 & 0 & 0 & 0 & \cdots \\ 0 & 0 & c_0 & c_1 & c_2 & c_3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

If the coefficients satisfy

$$c_0^2 + c_1^2 + c_2^2 + c_3^2 = 1 \tag{9}$$
$$c_0 c_2 + c_1 c_3 = 0, \tag{10}$$

then the filter bank matrix will be orthogonal. As we have proved in Section 3, this guarantees that the wavelet basis is orthonormal. Since we have four coefficients, we still have two degrees of freedom left to maximize the approximation order. Using Equation (8), we get the two remaining constraints, that ensure that the wavelet has two vanishing moments:

$$c_0 - c_1 + c_2 - c_3 = 0 \quad \text{when} \quad k = 0 \tag{11}$$
$$-c_1 + 2c_2 - 3c_3 = 0 \quad \text{when} \quad k = 1 \tag{12}$$

By solving the system of equations together with Equation (9) and (10), we get

$$c_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad c_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad c_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad c_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

For longer filters, we can obtain bases with more vanishing moments. The equations for coefficients can be solved explicitly only for small filter lengths; for longer filters the values for coefficients can be computed numerically.