Man Kam Kwong[1]

Orthogonally Compensated W-Multiresolution Analysis and Signal Processing

The concept of a W-matrix is used to give an elementary interpretation of a biorthogonal wavelet decomposition of signals. We also give a method to modify the decomposition to give an orthogonal projection on the the space spanned by the scaling vectors. Roughly speaking, our treatment is a finite-length analog of the well-known theory of multiresolution analysis of Meyer and Mallat. Our approach differs in that it deals directly with the discrete case, it takes care of the boundary elements without explicit padding, and it uses a notion similar to that of semi-orthogonality introduced by Chui. Our algorithm has flexibility in the choice of filter coefficients. The decomposition, orthogonalization, and restoration algorithms are computationally fast.

1. W-Matrices and W-Transforms

The theory of wavelets has had a great impact on the technology of image processing. Excellent expositions of the classical theory can be found in [1] and [2] and the references quoted in them. The idea of W-matrices was introduced in [3] and has been used to give an elementary interpreation of wavelet decomposition. The concept had also led to some new perspectives on how discrete signal can be treated. Implementation of four-tap W-transforms was described in [4].

In this paper, we survey this approach and describe a method to modify the decomposition to give an orthogonal projection on the the space spanned by the scaling vectors.

We use the special case of four-tap filters for the majority of our discussion; most of the ideas can be extended to the general case. We start with any vector of four numbers

$$\mathbf{h} = [h_1, h_2, h_3, h_4], \tag{1}$$

such that $h_1h_4 - h_2h_3 \neq 0$. Although the theory imposes no further restrictions, in practice, these numbers are usually chosen to form a four-tap high-pass filter. As a concrete example, we use $[-1, 3, -3, 1]$. Choose two other numbers $c \neq d$, and form the vector

$$\mathbf{g} = [g_1, g_2, g_3, g_4] = [ch_1, ch_2, dh_3, dh_4]. \tag{2}$$

For example, with $c = 1, d = -1$, $\mathbf{g}$ is $[-1, 3, 3, -1]$. Stack n copies of the pair of vectors to form the array

$$\begin{pmatrix} g_1 & g_2 & g_3 & g_4 & & & & & & \\ h_1 & h_2 & h_3 & h_4 & & & & & & \\ & & g_1 & g_2 & g_3 & g_4 & & & & \\ & & h_1 & h_2 & h_3 & h_4 & & & & \\ & & & & & \ddots & & & & \\ & & & & & & g_1 & g_2 & g_3 & g_4 \\ & & & & & & h_1 & h_2 & h_3 & h_4 \end{pmatrix}. \tag{3}$$

Each pair is shifted by two positions to the right relative to its previous pair, and blank positions are filled with zeros. We add the first and last columns to their neighbors, respectively, to form a matrix of order $2n \times 2n$:

$$\mathbf{W} = \begin{pmatrix} g_1+g_2 & g_3 & g_4 & & & & & \\ h_1+h_2 & h_3 & h_4 & & & & & \\ & g_1 & g_2 & g_3 & g_4 & & & \\ & h_1 & h_2 & h_3 & h_4 & & & \\ & & & & \ddots & & & \\ & & & & & g_1 & g_2 & g_3+g_4 \\ & & & & & h_1 & h_2 & h_3+h_4 \end{pmatrix}. \tag{4}$$

[1]This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

We remark that "adding" is not the only way to deal with the boundary columns. Other methods including "folding" (for four-tap filters, this is the same as "adding"), "linear extrapolation" (replace the first row by g_2+2g_1, g_3-g_1 and the second row by h_2+2h_1, h_3-h_1), and "quadratic extrapolation" (replace the first row by $g_2+3g_1, g_3-3g_1, g_4+g_1$ and the second row by $h_2+3h_1, h_3-3h_1, h_4+h_1$), each leading to a different boundary wavelet.

An interesting fact is that the inverse of $\mathbf{W}$ (when it exists) has a similar structure (formed by stacking pairs of shifted vectors of length four) and the numbers involved are independent of n. Let us denote by $\mathbf{Q}$ (called the *quadratic spline* W-matrix) the square matrix generated by our concrete example using "adding." Then

$$\mathbf{Q}^{-1} = \frac{1}{16}\begin{pmatrix} 4 & 4 & & & & & & & & \\ 3 & -3 & 1 & 1 & & & & & & \\ 1 & -1 & 3 & 3 & & & & & & \\ & & 3 & -3 & 1 & 1 & & & & \\ & & 1 & -1 & 3 & 3 & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 3 & -3 & 1 & 1 \\ & & & & & & 1 & -1 & 3 & 3 \\ & & & & & & & & 4 & -4 \end{pmatrix}, \tag{5}$$

no matter what n is (one, therefore, need only compute the inverse for $n=3$). The general formula for $\mathbf{W}^{-1}$, which can be obtained easily using a symbolic manipulation software such as MAPLE, takes up too much space to be included here.

Although it is possible to regard $\mathbf{W}^{-1}$ as formed by stacking two row vectors as in the case of $\mathbf{W}$, it is more appropriate to consider $\mathbf{W}^{-1}$ as formed by packing vertically shifted versions of two basic column vectors. In the case of $\mathbf{Q}^{-1}$, the two basic column vectors are $\overline{\mathbf{g}} = [1,3,3,1]^t$ and $\overline{\mathbf{h}} = [1,3,-3,-1]^t$. Here the superscript t denotes transpose. These vectors are dual to $\mathbf{g}$ and $\mathbf{h}$ used to form $\mathbf{W}$.

A matrix of odd order is constructed by deleting the last row of the matrix in (4) and then adding the last column to its neighbor. One can verify that the inverse of odd-order $\mathbf{W}$ has a similar structure.

In general, we start with two row vectors of any finite length; stack n pairs of them, each shifted by two positions to the right; and choose a consistent method of casting out some of the leading and trailing columns (and the last row) to form a square matrix of even (odd) order. It can be proved that with appropriate choice of the basic vectors, the inverse of the matrices so constructed are generated by two dual column vectors.

Definition 1. *A W-matrix refers to the class of matrices of all orders formed by two basic row vectors of finite length in the manner described above, with the property that the inverse of each matrix in the class is formed by two dual column vectors of finite length, in a manner (details omitted) independent of the order of the matrix.*

For the sake of convenience, we will use the same symbol $\mathbf{W}$ to denote a W-matrix (which is a class of matrices) or any member of the class. A one-dimensional signal is a column vector of arbitrary but finite length.

Definition 2. *Given a W-matrix $\mathbf{W}$ and a signal $\mathbf{x}$, the W-transform of $\mathbf{x}$ is a pair of signals $\mathcal{W}\mathbf{x} = (\mathbf{y}_1, \mathbf{y}_2)$ formed, respectively, by the odd and even components of $\mathbf{y} = \mathbf{W}\mathbf{x}$, where $\mathbf{W}$ is a matrix of the appropriate order in the given class.*

Both $\mathcal{W}$ and its inverse $\mathcal{W}^{-1}$ involve $O(N)$ computational steps, where $N = \text{length}(\mathbf{x})$, because the rows of $\mathbf{W}$ and $\mathbf{W}^{-1}$ have only four nonzero elements in each row. We state two existence results.

Theorem 1. *Given any vector $\mathbf{h} = [h_1, h_2, \cdots, h_{2n}]$ of even length, one can supplement it with a vector $\mathbf{g}$ so that the pair will generate a W-matrix if the $(2n-1)\times 2n$ matrix*

$$\mathbf{A} = \begin{pmatrix} h_2 & -h_1 & & & & & \\ h_4 & -h_3 & h_2 & -h_1 & & & \\ & & & & \ddots & & \\ & & & & & -h_{2n-1} & h_{2n} \end{pmatrix} \tag{6}$$

has full rank. Let $\mathbf{B}$ be the matrix obtained from $\mathbf{A}$ by deleting the middle row of $\mathbf{A}$. The solution space of the matrix equation $\mathbf{B}[z_1, z_2, \cdots, z_{2n}]' = 0$ is a two-dimensional linear space that contains $\mathbf{h}$. Any nonzero vector in the solution space other than a multiple of $\mathbf{h}$ can be used as $\mathbf{g}$.

Theorem 2. *For any given n numbers $h_1, h_2, \cdots, h_n$, one forms the vector $\mathbf{h} = [h_1, h_2, \cdots, h_n, -h_n, \cdots, -h_2, -h_1]$. If $\mathbf{h}$ satisfies the conditions in Theorem 4, there exists a symmetric $\mathbf{g}$ such that $(h, \mathbf{g})$ and $\mathbf{g}$ generates a W-matrix.*

2. Decomposition of x by the Dual Base Vectors

Suppose $\mathbf{y}_1 = [v_1, v_2, \cdots]^t$ and $\mathbf{y}_2 = [w_1, w_2, \cdots]^t$, and the column vectors of $\mathbf{W}^{-1}$ are $\overline{\mathbf{g}}_1, \overline{\mathbf{h}}_1, \overline{\mathbf{g}}_2, \overline{\mathbf{h}}_2, \cdots$. Note that $\overline{\mathbf{g}}_i$ and $\overline{\mathbf{h}}_i$, except the first and the last ones, are translates of $\overline{\mathbf{g}}$ and $\overline{\mathbf{h}}$, respectively. It follows from $\mathbf{x} = \mathbf{W}^{-1}\mathbf{y}$ that

$$\mathbf{x} = \sum v_i \overline{\mathbf{g}}_i + \sum w_i \overline{\mathbf{h}}_i. \tag{7}$$

We can, therefore, interpret the W-transform as realizing (by providing the coefficients v_i and w_i) the projection of $\mathbf{x}$ onto the subspaces $\mathbf{G}$ and $\mathbf{H}$ spanned by $\{\overline{\mathbf{g}}_i\}$ and $\{\overline{\mathbf{h}}_i\}$, respectively. In general, these subspaces are not orthogonal to each other. The vectors $\overline{\mathbf{h}}_i$ correspond to wavelets of the finest resolution in the theory of biorthogonal wavelets.

When $\mathbf{g}$ corresponds to a low-pass filter and $\mathbf{h}$ a high-pass filter, the projection onto the subspace $\mathbf{G}$ ($\sum v_i \overline{\mathbf{g}}_i$) gives the smooth part of the signal, while the projection onto $\mathbf{H}$ ($\sum w_i \overline{\mathbf{h}}_i$) gives the detail part.

3. Orthogonal Compensation

A usual technique used in lossy signal compression is to discarding the small components of $\mathbf{y}_2$. The error incurred in discard a particular coefficient w_j is $w_j \overline{\mathbf{h}}_j$. We can minimize this error by compensating the remaining signal with the orthogonal projection of $w_j \overline{\mathbf{h}}_j$ onto the subspace $\mathbf{G}$. Likewise, if we are going to discard the whole $\mathbf{y}_2$, we can compensate $\mathbf{y}_1$ by adding the orthogonal projection of $\sum w_i \overline{\mathbf{h}}_i$ onto the subspace $\mathbf{G}$. The orthogonal projection of a vector $\mathbf{u}$ onto $\mathbf{g}$, $\mathcal{O}(\sqcap)$, is the column vector formed by the coefficients c_i in the following equation:

$$\mathbf{u} = \sum c_i \overline{\mathbf{g}}_i + \mathbf{e}, \tag{8}$$

where $\mathbf{e}$ is perpendicular to $\overline{\mathbf{g}}_i$ for all i. By taking the inner products of EQuation (8) with $\overline{\mathbf{g}}_i$ for all i, we obtain a tridiagonal system of linear equations with the unknown c_i. Solving such a system requires $O(N)$ operations.

Definition 3. *The orthogonally compensated W-transform of $\mathbf{x}$ is $\mathcal{W}_o\mathbf{x} = (\mathbf{y}_o, \mathbf{y}_2)$, where $\mathbf{y}_o = \mathbf{y}_1 + \mathcal{O}(\sum w_i \overline{\mathbf{h}}_i)$.*

Given $(\mathbf{y}_o, \mathbf{y}_2)$, we can recover $\mathbf{x}$ by reversing the process: subtract $\mathcal{O}(\mathbf{y}_2)$ from $\mathbf{y}_o$ to obtain $\mathbf{y}_1$, and then apply $\mathcal{W}^{-1}$. The orthogonal projection of $\mathbf{x}$ onto $\mathbf{G}$ is $\mathcal{W}^{-1}(\mathbf{y}_o, \mathbf{0})$.

A compression algorithm using $\mathcal{W}_o$ will comprise the following steps. A 1D signal $\mathbf{x}$ is first decomposed into the pair $(\mathbf{y}_o, \mathbf{y}_1)$, on which quantization is applied to give an approximate pair $(\tilde{\mathbf{y}}_o, \tilde{\mathbf{y}}_2)$. (A 2D image will be decomposed into four subsignals using an orthogonally compensated W-transform twice, once in the x direction, and once in the y direction.) Instead of coding this new pair, half of $\mathcal{W}_o^{-1}$ is performed, namely, to obtain $(\tilde{\mathbf{y}}_o - \mathcal{O}(\tilde{\mathbf{y}}_2), \tilde{\mathbf{y}}_2)$. This way, only the inverse of $\mathcal{W}$ will need to be applied in the decompression algorithm.

In terms of obtaining the orthogonal projection of $\mathbf{x}$ onto $\mathbf{G}$, our procedure is equivalent to the classical method of orthogonalizing the scaling function $\overline{\mathbf{g}}$, or Chui's method [1] of keeping $\overline{\mathbf{g}}$ while looking for a modified wavelet vector (to replace $\overline{\mathbf{h}}$) that is perpendicular to $\mathbf{G}$. The classical method [2] will, in general, produce a wavelet vector of infinite support, and the decomposition and restoration algorithms will then require filters of infinite support. Mallat [5] has devised a fast algorithm using FFT. On the other hand, Chui constructs a modified wavelet of compact support (compactness guarantees that the restoration algorithm is fast) that generates the orthogonal complement of $\mathbf{G}$. The decomposition algorithm, however, will involve an infinite support filter.

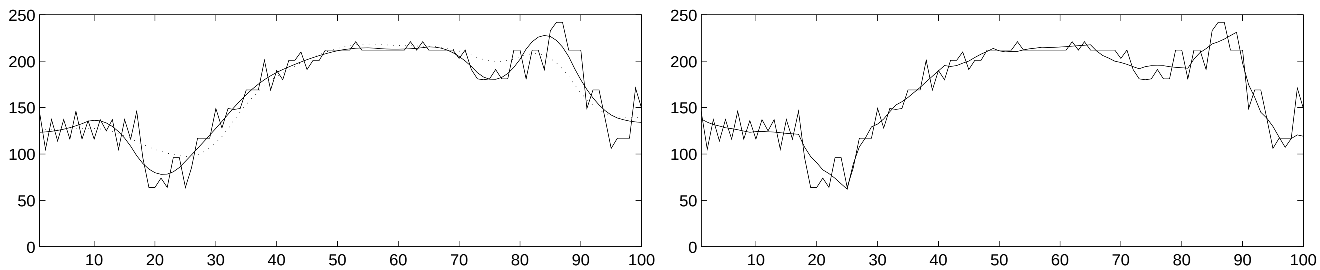
4. Multiresolution Analysis

When $\mathbf{g}$ is a high-pass filter, most components of $\mathbf{y}_2$ are small and can be discarded for the purpose of signal compression. On the other hand, $\mathbf{y}_1$ or $\mathbf{y}_o$ is a slightly blurred representation of $\mathbf{x}$ that usually retains the salient features of $\mathbf{x}$. It is natural to iterate $\mathcal{W}$ on $\mathbf{y}_1$ or $\mathcal{W}_o$ on $\mathbf{y}_o$. This step leads to the familiar concept of multiresolution pioneered by Meyer and Mallat [5].

Suppose $\mathcal{W}(\mathbf{y}_1) = (\mathbf{z}_1, \mathbf{z}_2)$. The components of $\mathbf{z}_1$ ($\mathbf{z}_2$) corresponds to the coefficients of the second-level biorthogonal scaling (wavelet) vectors.

We note, however, that applying $\mathcal{W}_o$ on $\mathbf{y}_o$ does not yield an orthogonal projection onto the subspace spanned by the second level scaling vectors. This is because the translates of the scaling vectors are not orthogonal to each other. We call the multiresolution generated by iterating $\mathcal{W}_o$ *quasi-orthogonal*. It is possible to obtain a truly orthogonal compensation by expressing the orthogonal decomposition identity (analogous to (8)) for $\mathbf{z}_2$ in terms of the first level representation. The resulting linear system to be solved for the unknown coefficients will then involve a 5-banded matrix instead of a tridiagonal matrix. The computational requirement, although higher, is still of the order $O(N)$. In practice, we find that the quasi-orthogonal algorithm is quite satisfactory.

5. Observations



Figures 1a and 1b. Projections onto the space spanned by third-level scaling functions ($\mathbf{Q}$ and D_4)

- The quadratic spline transform $\mathbf{Q}$ generated by $\overline{\mathbf{h}} = [-1, 3, -3, 1]$ has a smooth scaling function. As a result, it performs well for de-noising and for compressing smooth signals. In the above figures, the two jagged curves are the same signal, which is part of a GIF image. In Figure 1a, the smooth solid curve is restored from a three-level orthogonally compensated $\mathbf{Q}$ transform (the dotted curve is obtained without using orthogonal compensation), after discarding all the detail wavelet coefficients. For comparison, the result using the four-tap Daubechies wavelet is given in Figure 1b.
- The jaggedness of the original data is due to noise that is almost always present in an image. Sometimes noise is artificially added by a dithering process. An understanding of how to reverse the process can be useful noise removal.
- Note that the process of decarding details smoothes out both noise and edges. In applications to image processing, the preservation of edges and local optima (minus noise) is important. This can be achieved by retaining those wavelet coefficients in the neighborbood of such points, the location of which can often be determined more efficiently with methods not based on wavelets.
- In practice, we seldom use more than four levels of multiresolution analysis. Often three is enough.
- When used without orthogonal compensation, $\mathbf{Q}$ does not satisfy the conventional definition of biorthogonal wavelets, since the dual wavelet is not square integrable as the number of resolution levels approaches ∞. Even when only a finite number of levels are used, the error incurred in discarding high-level wavelet coefficients can be large. (See the dotted curve in Figure 1a.) Orthogonal compensation is effective in controlling this error.
- The potential of the flexibility in picking $\overline{\mathbf{h}}$ remains to be investigated. The choice of $\overline{\mathbf{h}}$ can be done adaptively for different regions of the image as well as at different levels of resolution.

6. References

1 Chui, C. K.: **An Introduction to Wavelets**; Academic Press (1992).
2 Daubechies, I.: **Ten Lectures on Wavelets**, CBMS-NSF Series Appl. Math., SIAM, 1991.
3 Kwong, Man Kam: MATLAB implementation of *W*-matrix multiresolution analyses; Preprint Series MCS-P462-0894 (available by ftp: `ftp@mcs.anl.gov` /pub/W-transform/wtransf2.ps.Z), Mathematics and Computer Science Division, Argonne National Laboratory (1994).
4 Kwong, Man Kam, and Tang, Peter P. T.: *W*-matrices and nonorthogonal multiresolution analysis of finite signals of arbitrary length; Preprint Series MCS-P449-0794 (available by ftp: `ftp@mcs.anl.gov` /pub/W-transform/wtransf1.ps.Z), Mathematics and Computer Science Division, Argonne National Laboratory (1994).
5 Mallat, S. G.: A theory for multiresolution signal decomposition: The wavelet representation; IEEE Trans. on Pattern Analysis and Machine Intelligence **11** (1989), 674–693.

Addresses: Man Kam Kwong, Mathematics and Computer Science Division Argonne National Laboratory Argonne, IL 60439-4844 E-mail: `kwong@mcs.anl.gov`