ADAPTED WAVE FORM ANALYSIS, WAVELET-PACKETS AND APPLICATIONS

R. Coifman Y. Meyer V. Wickerhauser

Yale University, University Of Paris,Dauphine, Washington University St.Louis

Introduction.

Adapted wave form analysis, refers to a collection of FFT like adapted transform algorithms.

Given a function or an operator these methods provide a special orthonormal basis relative to which the function is well represented, and the operator is described by a sparse matrix. The selected basis functions are chosen inside predefined libraries of oscillatory localized functions (waveforms). These algorithms are of complexity $N \log N$ opening the door for a large range of applications in signal and image processing, as well as in numerical analysis.

Our goal is to describe and relate traditional windowed Fourier Transform methods to wavelet, wavelet-packet base algorithms by making explicit their dual nature and relative role in analysis and computation.

Starting with a recent refinement of the windowed sine and cosine transforms we will derive an adapted local sine transform show it's relation to wavelet and wavelet-packet analysis and describe an analysis tool-kit illustrating the merits of different adaptive and non-adaptive schemes.

We end with sample applications to signal and image compression statistical factor analysis, and numerical analysis and P.D.E.

§1. Windowed FFT and Adapted Window Selection.

We start with a description of an algorithm to compute the Fourier expansion of a function on a union of two adjacent intervals of the same size, in terms of the Fourier expansion on each interval.

Typeset by $\mathcal{AMS}$-TEX

Let f be defined on $[0,2]$

$$f = f^0 + f^1 \qquad \text{where} f^0 = \begin{cases} f & x \in [0,1] \\ 0 & x \notin [0,1] \end{cases}$$

we want to compute

$$\hat{f}_m = \frac{1}{\sqrt{2}} \int_0^2 f(t) e^{-2\pi i m \frac{t}{2}} dt$$

in terms of $\hat{f}^0_n = \int_0^1 f(t)e^{-2\pi i n t}dt$ and $\hat{f}^1_n = \int_1^2 f(t)e^{-2\pi i n t}dt$. Clearly, when $m = 2n$ we have

$$\hat{f}_{2n} = \frac{1}{\sqrt{2}}\{\hat{f}^0_n + \hat{f}^1_n\}.$$

For $m = 2n+1$ we define

$$d_n = \frac{1}{\sqrt{2}}\{\hat{f}^0_n - \hat{f}^1_{n+1}\}$$

and find

$$\boxed{\hat{f}_{2n+1} = \frac{1}{\pi i} \sum \frac{d_k}{(n-k+\frac{1}{2})}}.$$

In fact,

$$\hat{f}_{2n+1} = \frac{1}{\sqrt{2}} \int_0^1 [f(t) - f(t+1)] e^{-i\pi t} e^{-2\pi i n t} dt.$$

Since d_n are the Fourier coefficients on $[0,1]$ of $f(t) - f(t+1)$, and $\frac{1}{\pi i(n+\frac{1}{2})}$ are the coefficients of $e^{-i\pi t}$, we obtain the coefficients of $\hat{f}_{2n+1}$ by convolving these sequences.

A fast way to compute $\hat{f}_{2n+1}$ is to compute the inverse transform on $(0,1)$ of d_n, multiply by $e^{-it/2}$ and recompute the transform on $(0,1)$. This procedure, when discretized, leads to a Fourier transform algorithm of complexity $\leq CN(\log_2 N)^2$ (in complex arithmetic) . (A faster algorithm can be obtained by implementing an order N computation for the convolution)

Schematic Description

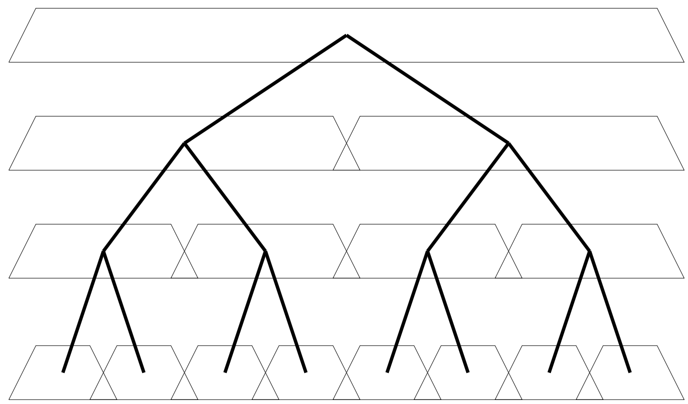

Figure 1

We see that in order to compute the transform on the large interval, we can start with adjacent pairs of small intervals, combine coefficients to obtain the expansion on their union, and continue until we reach the top level. As a result we have obtained all dyadic windowed Fourier transform as intermediate computations.

Clearly every disjoint collection of intervals equipped with an orthogonal basis on each provides us with an orthogonal basis for the union. A natural question that arises in connection with the windowed Fourier transform is how to place the windows (see Figures 2,3 where the effect of the window selection on the number of large coefficients is visible).

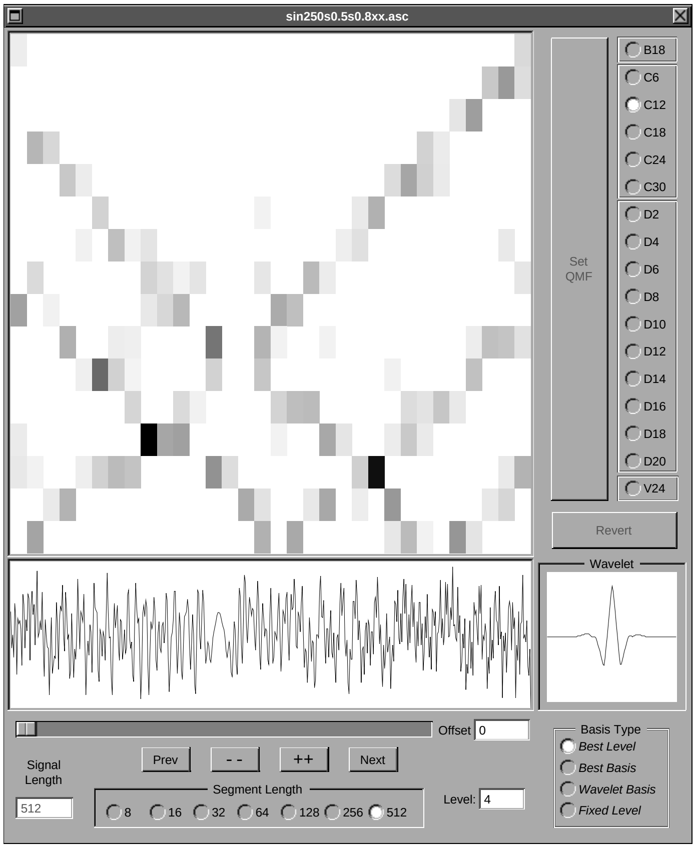

Figure 2

The signal is a combination of three linear chirps. By choosing small windows we find three main frequencies per window (the vertical axis is the frequency axis while the horizontal is the time axis).

In Figure 3 the windows are larger leading to a less efficient intertwined representation of the signal.

Large window selection

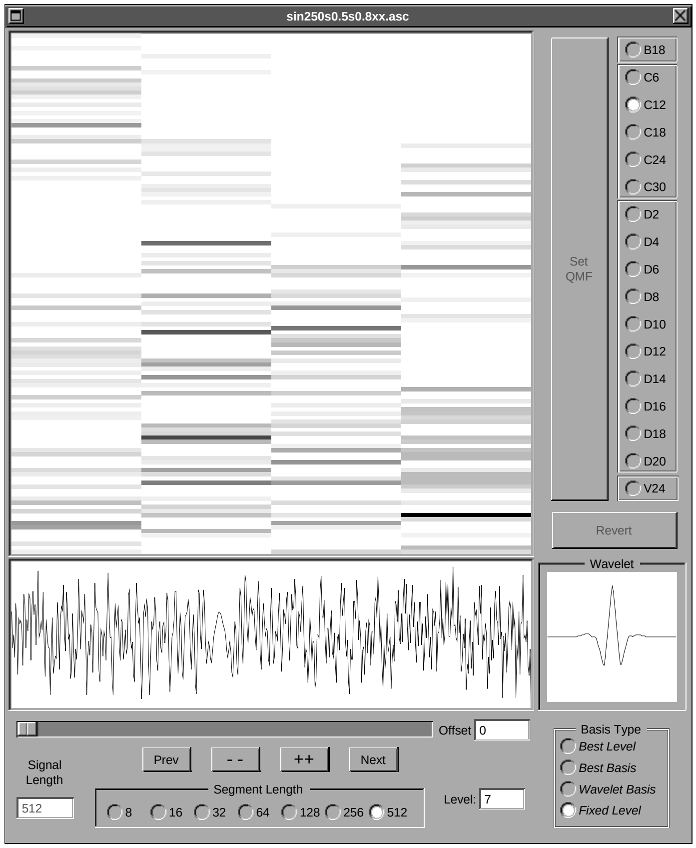

Figure 3

For the moment let us consider the question of optimizing the windows to obtain an efficient representation of a function.

We can proceed as follows:

We start with the adjacent small intervals and consider expansion coefficient in each separately. We then compute the expansion coefficients on their union. We can now choose that expansion for which the number of coefficients needed to capture 99% of the energy is smallest (or that expansion whose "cost" is smallest; information cost, coding cost, error cost).

We compare the cost of the chosen expansions on two adjacent unions of pairs to the expansion on their union and again pick the best.

We continue until we reach an optimal distribution of windows (see Figure 4 where the windows were adapted to the voice recording).

Optimal Window Selection

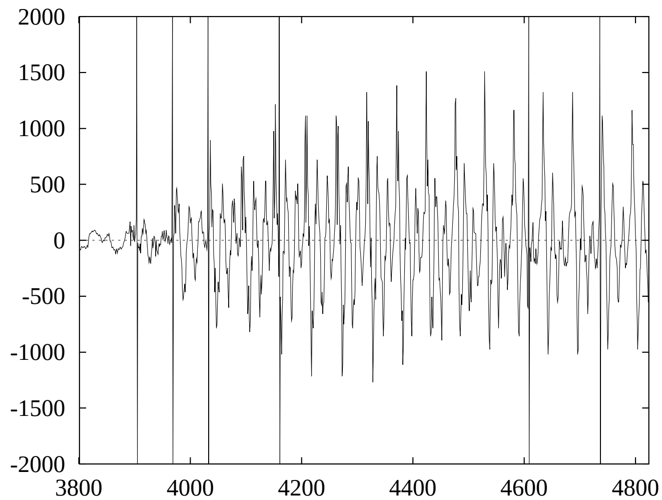

Figure 4

The procedure described above, although natural, is not very useful if we take the windowed Fourier transform with discontinuous windows, since the discontinuity introduces "large" expansion coefficients, (a cosine basis on each interval is somewhat better). On the other hand, it is well known that we cannot find a smooth window function $\omega(x)$ supported on $(-\frac{1}{2}, \frac{3}{2})$ such that $\omega(x-k)e^{2i\pi nx}$ are orthogonal. (This would imply $\int \omega(x)\omega(x-1)e^{-2\pi imx}dx = 0$ for all m i.e. $\omega(x)\omega(x-1) = 0$).

Recently Daubechies ,Jaffard,and Journe as well as Malvar observed that by taking equal windows and sines or cosines orthogonality can be maintained. Coifman and Meyer [3] observed that the windows can be chosen to different sizes enabling adaptive constructions as above. (See Figures 5,6)

Local trigonometric waveforms

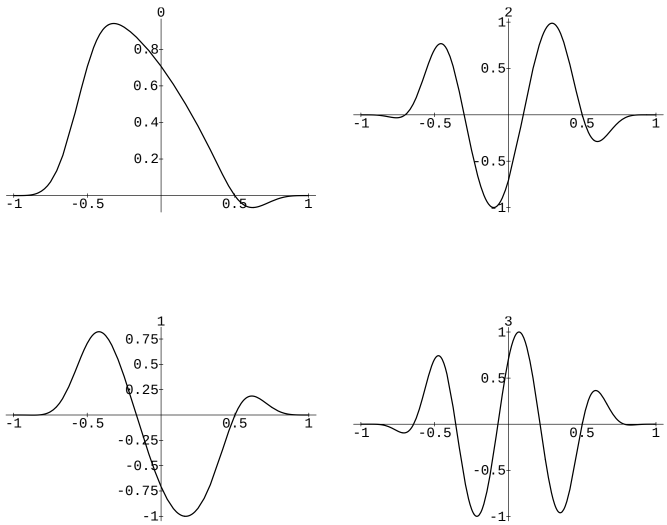

Figure 5

Local trigonometric waveforms

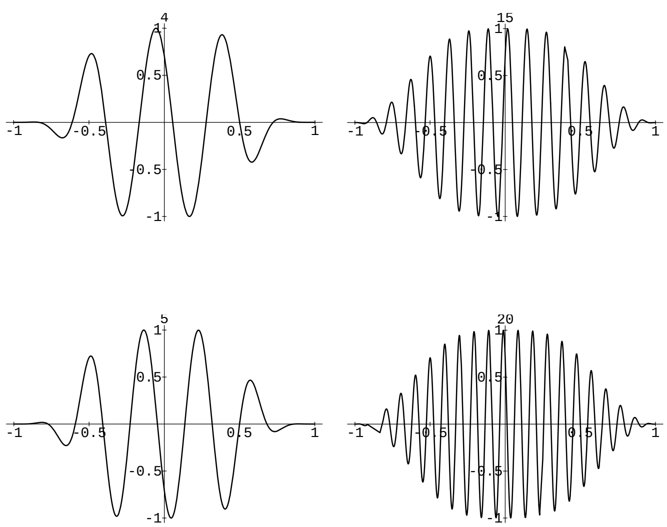

Figure 6

We start by defining this library of trigonometric waveforms. These are localized sine transforms associated to covering by intervals of **R** (more generally, of a manifold).

We consider a cover $\mathbf{R} = \bigcup_{-\infty}^{\infty} I_i \; I = [\alpha_i \alpha_{i+1}) \; \alpha_i < \alpha_{i+1}$, write $\ell_i = \alpha_{i+1} - \alpha_i = |I_i|$

and let $p_i(x)$ be a window function supported in $[\alpha_i - \ell_{i-1}/2, \alpha_{i+1} + \ell_{i+1}/2]$ such that

$$\sum_{-\infty}^{\infty} p_i^2(x) = 1$$

and

$$p_i^2(x) = 1 - p_i^2(2\alpha_{i+1} - x) \quad \text{for} \quad x \quad \text{near} \quad \alpha_{i+1}$$

then the functions

$$S_{i,k}(x) = \frac{2}{\sqrt{2\ell i}} p_i(x) \sin[(2k+1)\frac{\pi}{2\ell_i}(x - \alpha_i)]$$

form an orghonormal basis of $L^2(\mathbf{R})$ subordinate to the partition p_i. The collection of such bases forms a library of orthonormal bases.

It is easy to check that if H_{I_i} denotes the space of functions spanned by $S_{i,k}$ $k = 0, 1, 2, ...$ then $H_{I_i} + H_{I_{i+1}}$ is spanned by the functions

$$P(x) \frac{2}{\sqrt{2(\ell_i + \ell_{i+1})}} \sin[(2k+1)\frac{\pi}{2(\ell_i + \ell_{i+1})}(x - \alpha_i)]$$

where

$$P^2 = p_i^2(x) + p_{i+1}^2(x)$$

is a "window" function covering the interval $I_i \cup I_{i+1}$. This fundamental identity permits the useful implementation of the adapted window algorithm described in Figure 1. (Other possible libraries can be constructed. The space of frequencies can be decomposed into pairs of symmetric windows around the origin ,on which a smooth partition of unity is constructed. This and other constructions were obtained by one of our students E. Laeng [L].

Higher dimensional libraries can also be easily constructed,(as well as libraries on manifolds) leading to new and direct analysis methods for linear transformations.)

Relation to Wavelets - Wavelet Packets.

We consider the frequency line $\mathbf{R}$ split as $\mathbf{R}^+ = (0, \infty)$ union $\mathbf{R}^- = (-\infty, 0)$. On $L^2(0, \infty)$ we introduce a window function $p(\xi)$ such that $\sum_{k=-\infty}^{\infty} p^2(2^{-k}\xi) = 1$

and $p(\xi)$ is supported in $(3/4, 3)$ clearly we can view $p(2^{-k}\xi)$ as a window function above the interval $(2^k, 2^{k+1})$ and observe that

$$\sin\left[(j+\frac{1}{2})\pi\left(\frac{\xi-2^k}{2^k}\right)\right]p(2^{-k}\xi) = s_{k,j}$$

form an orthonormal basis of $L^2(\mathbf{R}^+)$. Similarly $c_{k,j} = \cos\left[(j+\frac{1}{2})\pi\left(\frac{\xi-2^k}{2^k}\right)\right]p(2^{-k}\xi)$ gives another basis. If we define $S_{k,j}$ as an odd extension to $\mathbf{R}$ of $s_{k,j}$ and $C_{k,j}$ as an even extension, we find $S_{k,j} \perp C_{k',j'}$ permitting us to write $C_{k,j} \pm iS_{k,j} = e^{\pm ij\pi\xi/2^k}\hat{\psi}(\xi/2^j)$ where $\hat{\psi}(\xi) = e^{i\pi/2\xi}p(\xi)$ is the Fourier transform of the base wavelet Ψ (see Meyer).

We therefore see that wavelet analysis corresponds to windowing frequency space in "octave" windows $(2^k, 2^{k+1})$.

A natural extension therefore is provided by allowing all dyadic windows in frequency space and adapted window choice. This sort of analysis is "equivalent" to wavelet packet analysis.

The wavelet packet analysis algorithms permit us to perform an adapted Fourier windowing directly in time domain by successive filtering of a function into different regions in frequency. The dual version of the window selection provides an adapted subband coding algorithm.

This new library of orthonormal bases constructed in time domain is called the Wavelet packet library. This library contains the wavelet basis, Walsh functions, and smooth versions of Walsh functions called wavelet packets.See Figure 7

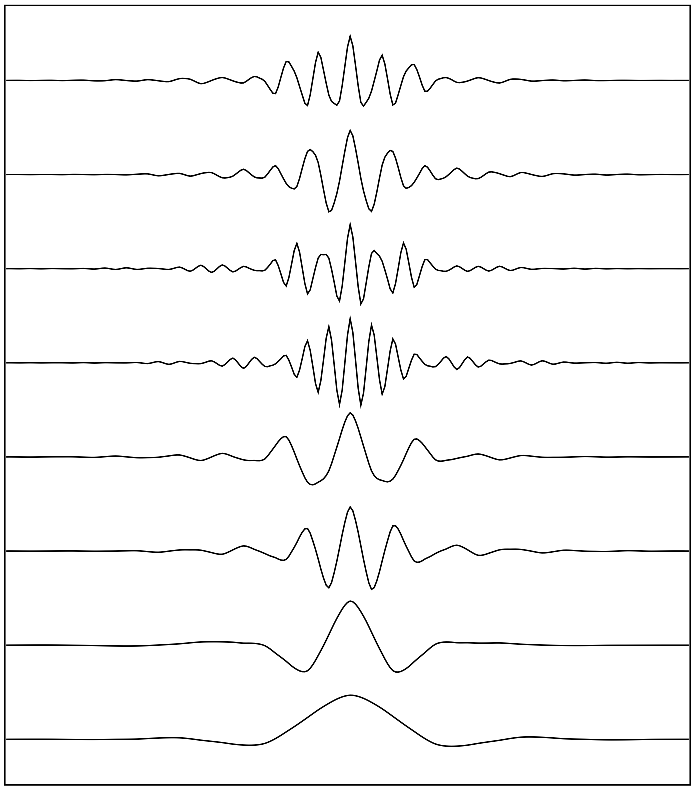

Figure 7

We'll use the notation and terminology of [4], whose results we shall assume.

We are given an exact quadrature mirror filter $h(n)$ satisfying the conditions of Theorem (3.6) in [4], p. 964, i.e.

$$\sum_n h(n-2k)h(n-2\ell) = \delta_{k,\ell}, \ \sum_n h(n) = \sqrt{2}.$$

We let $g_k = h_{l-k}(-1)^k$ and define the operations F_i on $\ell^2(\mathbf{Z})$ into "$\ell^2(2\mathbf{Z})$"

$$F_0\{s_k\}(i) = 2\sum s_k h_{k-2i} \tag{1.0}$$

$$F_1\{s_k\}(i) = 2\sum s_k g_{k-2i}.$$

The map $\mathbf{F}(s_k) = F_0(s_k) \oplus F_1(s_k) \in \ell^2(2\mathbf{Z}) \oplus \ell^2(2\mathbf{Z})$ is orthogonal and

$$F_0^* F_0 + F_1^* F_1 = I \tag{1.1}$$

We now define the following sequence of functions.

$$\begin{cases} W_{2n}(x) = \sqrt{2}\sum h_k W_n(2x-k) \\ W_{2n+1}(x) = \sqrt{2}\sum g_k W_n(2x-k). \end{cases} \tag{1.2}$$

Clearly the function $W_0(x)$ can be identified with the scaling function φ in [D] and W_1 with the basic wavelet ψ.

Let us define $m_0(\xi) = \frac{1}{\sqrt{2}}\sum h_k e^{-ik\xi}$ and

$$m_1(\xi) = -e^{i\xi}\bar{m}_0(\xi+\pi) = \frac{1}{\sqrt{2}}\sum g_k e^{ik\xi}$$

Remark. The quadrature mirror condition on the operation $\mathbf{F} = (F_0, F_1)$ is equivalent to the unitarity of the matrix

$$\mathcal{M} = \begin{bmatrix} m_0(\xi) & m_1(\xi) \\ m_0(\xi+\pi) & m_1(\xi+\pi) \end{bmatrix}$$

Taking the Fourier transform of (1.2) when $n = 0$ we get

$$\hat{W}_0(\xi) = m_0(\xi/2)\hat{W}_0(\xi/2)$$

i.e.,

$$\hat{W}_0(\xi) = \prod_{j=1}^{\infty} m_0(\xi/2^j)$$

and

$$\hat{W}_1(\xi) = m_1(\xi/2)\hat{W}_0(\xi/2) = m_1(\xi/2)m_0(\xi/4)m_0(\xi/2^3)\cdots$$

More generally, the relations (1.2) are equivalent to

$$\hat{W}_n(\xi) = \prod_{j=1}^{\infty} m_{\varepsilon_j}(\xi/2j) \tag{1.3}$$

and $n = \sum_{j=1}^{\infty} \varepsilon_j 2^{j-1} (\varepsilon_j = 0 \text{ or } 1)$.

The functions $W_n(x-k)$ form an orthonormal basis of $L^2(\mathbf{R}^1)$. a We define a library of wavelet packets to be the collection of functions of the form $W_n(2^\ell x - k)$ where $\ell, k \in \mathbf{Z}, n \in N$. Here, each element of the library is determined by a scaling parameter ℓ, a localization parameter k and an oscillation parameter n. (The function $W_n(2^\ell x - k)$ is roughly centered at $2^{-\ell}k$, has support of size $\approx 2^{-\ell}$ and oscillates $\approx n$ times).

We have the following simple characterization of subsets forming orthonormal bases.

Proposition. *Any collection of indices (ℓ, n) such that the intervals $[2^\ell n, 2^\ell n+1)$ form a disjoint cover of $[0,\infty)$ gives rise to an orthonormal basis of L^2*[1].

(These intervals correspond to the partition of frequency space alluded to in §1.)

Motivated by ideas from signal processing and communication theory we were led to measure the "distance" between a basis and a function in terms of the Shannon entropy of the expansion. More generally, let H be a Hilbert space.

Let $v \in H,\ \|v\| = 1$ and assume

$$H = \oplus \sum H_i$$

an orthogonal direct sum. We define

$$\varepsilon^2(v, \{H_i\}) = -\sum \|v_i\|^2 \ell n \|v_i\|^2$$

as a measure of distance between v and the orthogonal decomposition.

ε^2 is characterized by the Shannon equation which is a version of Pythagoras' theorem.

Let

$$\begin{aligned} H &= \oplus(\sum H^i) \oplus (\sum H_j) \\ &= H_+ \oplus H_- \end{aligned}$$

H^i and H_j give orthogonal decompositions $H_+ = \sum H^i, H_- = \sum H_j$. Then

$$\begin{aligned} \varepsilon^2(v; \{H^i, H_j\}) = \varepsilon^2(v, \{H+, H_-\} \\ + \|v_+\|^2 \varepsilon^2 \left(\frac{v_+}{\|v_+\|}, \{H^i\} \right) \\ + \|v_-\|^2 \varepsilon^2 \left(\frac{v_-}{\|v_-\|}, \{H_j\} \right) \end{aligned}$$

This is Shannon's equation for entropy (if we interpret as in quantum mechanics $\|P_{H_+} v\|^2$ as the "probability" of v to be in the subspace H_+).

[1] We can think of this cover as an even covering of frequency space by windows roughly localized over the corresponding intervals.

This equation enables us to search for a smallest entropy space decomposition of a given vector.

In fact, for the example of the first library restricted to covering by dyadic intervals we can start by calculating the entropy of an expansion relative to a local trigonometric basis for intervals of length one, then compare the entropy of an adjacent pair of intervals to the entropy of an expansion on their union. Pick the expansion of minimal entropy and continue until a minimum entropy expansion is achieved (see Figure 1).

In practice, discrete versions of this scheme can be implemented in $CN \ \log \ N$ computations (where N is the number of discrete samples $N = 2^L$.)

Of course, while entropy is a good measure of concentration or efficiency of an expansion, various other information cost functions are possible, permitting discrimination and choice between various special function expansion.

§**2. Wavelet Packet and Adapted Waveform Analysis.** We would like to summarize some obvious implications of the preceding discussion. Wavelet packet analysis consists of a versatile collection of tools for the analysis and manipulation of signals such as sound and images, as well as more general digital data sets. The user is provided with a collection of standard libraries of waveforms, which can be chosen to fit specific classes of signals. These libraries come equipped with fast numerical algorithms, enabling real time implementation of a variety of signal processing tasks such as compression, feature extraction for recognition and diagnostics, data transformation, and manipulation.

The process of analysis of data usually starts by comparing acquired segments of data with stored "known" samples. As a model, consider how a real sampled signal is analyzed with such libraries. The example of voice or music can be used for illustration.

Voice signals consist of modulated oscillations as can be seen in Figure 4, representing a segment of a recording of the word "armadillo." Such a general signal is a superposition of different structures occurring on different time scales at different times. One purpose of analysis is to separate and sort these structures. The oscillations are analogous to musical notes, and the analysis is equivalent to choosing the best instrument to match to the voice, then finding the musical score to describe the word.

A musical note can be described by four basic parameters: intensity (or amplitude), frequency, time duration, and time position. Wavelet packets or localized sinusoids are indexed by the same parameters. In addition, there are other parameters corresponding to choice of library, i.e., the instrument or recipe used to generate all the waveforms. For wavelet packets these extra numbers are the quadrature mirror filter coefficients; for local cosines, they are the smooth window coefficients.

The process of analysis compares a sound or other signal with all elements of a given library and picks up large correlations, notes which are good fits to segments of the signal. A most-concentrated orthogonal subset of these good notes can then be chosen. This "best basis" realization provides an economical transcription, an efficient superposition of oscillatory modes on different time scales. When ordered by decreasing intensity, this transcription sorts the main features out in order of importance. It permits rebuilding the signal to a specified accuracy with the fewest waveforms. It can be used to compress signals for digital transmission and storage.

Of more practical value is the ability to compute and manipulate data in compressed parameters. This ability is particularly important for recognition and diagnostic purposes. As an illustration, consider a hypothetical diagnostic device for heartbeats, in which fifty consecutive beats are recorded. We would like to use this data as a statistical foundation for detection of significant changes in the next batch of beats. Theoretically this can be done by factor analysis, or the Karhunen–Loéve transformation; unfortunately, the computation involving raw data is too large to be practical. But when the recorded data is efficiently compressed to a few parameters in the single statistical best basis, the factor analysis (if needed) can be performed in real time. The deviation of the next few heartbeats from their predecessors can be computed "on the fly," and significant changes can be flagged immediately.

In another example, consider a very large three dimensional atmospheric pressure map, and the problem of calculating the evolution of the pressure. In this case it is natural to break up the computation as a sum of interactions within different scales, with some limited interaction between adjacent scales, Such a breakup is automatic if the pressure map is expressed in the wavelet basis, which in this case is also the natural choice for compression of the data.

Such algorithms, which first compress a large set of measurements in order to compute with fewer parameters, can dramatically reduce the time needed to transform and manipulate data. They generalize the classical transform methods, like FFT, by custom building a fast transform for each specific application, merging beautifully the technologies of data compression and numerical analysis.

References

1. R. Coifman, *Adapted multiresolution analysis, computation, signal processing and operator theory*, ICM 90 (Kyoto).
2. R. Coifman, Y. Meyer, and V. Wickerhauser, *Wavelet Analysis and Signal Processing Proceeding conference on Wavelets Lowell Mass 1991.*
3. R. Coifman and Y. Meyer, *Remarques sur l'analyse de Fourier à fenêtre*, série I, C. R. Acad. Sci. Paris **312** (1991), 259–261.
4. I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Communications on Pure and Applied Mathematics **XLI** (1988), 909–996.
5. E. Laeng, *Une base orthonormale de $L^2(\mathbf{R})$, dont les éléments sont bien localisés dans l'espace de phase et leurs supports adaptés à toute partition symétrique de l'espace des fréquences*, série I, C. R. Acad. Sci. Paris **311** (1990), 677–680.
6. Available by anonymous ftp, *ceres.math.yale.edu*, InterNet address 130.132.23.22.