[5] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Information Theory*, vol. 36, #5, pp. 961-1005, Sept. 1990.
[6] D. Slepian, "Some comments on Fourier analysis, uncertainty and modelling," *Bell Syst. Tech. J.*, vol. 25, no. 3, pp. 379-393, July 1983.
[7] M. Frisch, "Application of the wavelet transform to transient signal detection and estimation," Ph.D. dissert., Tel Aviv Univ., Fac. of Eng., Dept. of Elec. Eng.-Syst., Jan. 1992.
[8] L. Cohen and Ch. Lee, "Instantaneous bandwidth for signals and spectrogram," in *ICASSP 1990*, pp. 2451-2454.
[9] T. A. C. M. Claasen and W. F. G. Mecklenbraucker, "The Wigner distribution—A tool for time-frequency analysis, Part I," *Philips J. Res.*, vol. 35, pp. 217-250, 1980.
[10] A. Grossman, R. Kronland-Martinet and J. Morlet, "Reading and understanding continuous wavelet transforms," in *Wavelets, Time-frequency methods and Phase space*, J. M. Combes, A. Grossman, and Ph. Tchamitchian, Eds., *Proc. Int. Conf.*, Marseille, France, Dec. 14-18, 1987. New York: Springer-Verlag, 1989, pp. 2-20.
[11] C. W. Helstrom, "An expansion of a signal in Gaussian elementary signals," *IEEE Trans. Information Theory*, vol. IT-12, pp. 81-82, Jan. 1966.
[12] J. Morlet *et al.*, "Wave propagation and sampling theory; II," *Geophys.*, vol. 47, no. 2, pp. 222-236, Feb. 1982.
[13] P. Goupillaud, A. Grossman, and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis," *Geoexploration*, vol. 23. New York: Elsevier Science Publishers, 1984, pp. 85-102.
[14] R. Kronland-Martinet, J. Morlet, and A. Grossman, "Analysis of sound patterns through wavelet transforms," *Int. J. Pattern Recog. Artificial Intell.*, Special Issue on Expert Systems and Pattern Analysis, vol. 1, no. 2, pp. 97-126, 1987.
[15] F. B. Tuteur, "Wavelet transformations in signal processing," in ICASSP 1988, pp. 1435-14388.
[16] P. Flandrin, "Some aspects of nonstationary signal processing with emphasis on time-frequency and time-scale methods," in *Wavelets, Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossman, and Ph. Tchamitchian, Eds., Proc. Int. Conf., Marseille, France, Dec. 14-18, 1987; New York: Springer-Verlag, 1989, pp. 68-98.

Signal-Adapted Multiresolution Transform for Image Coding

Philippe Desarte, Benoit Macq, and Dirk T. M. Slock

Abstract—**Wavelet-type multiresolution transforms have recently been introduced in digital image coding, and have been shown to offer some advantages over classical block transform techniques such as the discrete cosine transform (DCT). This correspondence is mainly concerned with the problem of designing suitable multiresolution transforms that are *adapted* to the given image signal, in the sense that they maximize the coding gain at each resolution level. A simple alternating optimization algorithm is derived for solving this problem in the framework of the lattice realization of para-unitary quadrature mirror filters (QMF). The resulting image coding scheme is discussed in some detail, and its performance is compared with the DCT (JPEG) technique and with some non-adapted multiresolution transforms.**

Index Terms—**Multiresolution, wavelet, signal-adapted transform, para-unitary lattice filters, image coding.**

Manuscript received February 1991; revised Occtober 1991. This work was performed at the Philips Research Laboratory Belgium. This work was presented in part at the International Conference for Acoustic, Speech, and Signal Processing, Toronto, Canada, 1991.

Ph. Delsarte is with the Unité d'Informatique, Université Catholique de Louvain (UCL), Place Sainte Barbe, 2, B-1348 Louvain-la-Neuve, Belgium.
B. Macq is "Chercheur Qualifié" of the Belgian NSF with the Laboratoire de Télécommunications, Université Catholique de Louvain, Place du Levant, 3, B-1348 Louvain-la-Neuve, Belgium.
D. T. M. Slock is with the Institut EURECOM, 2229 route des Grêtes, Sophia Antipolis, 06560 Valbonne, France.

IEEE Log Number 9105363.

I. Introduction

The key to securing data compression is signal representation. To obtain a more efficient representation, a transform coder represents a random sequence in terms of specific basis vectors. In the synthesis part, the signal is reconstructed by an appropriate linear combination of the basis vectors. To obtain the appropriate combination coefficients is the task of the analysis section. A key issue now is the efficiency of the transform coder. One meaningful objective measure of the efficiency of an orthogonal transform is the coding gain G_{TC}, given by [10]

$$G_{TC} = \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\right)\left(\prod_{i=1}^{N}\sigma_i^2\right)^{-1/N}, \qquad (1)$$

where σ_i^2 is the variance of the output of the ith analysis filter. G_{TC} indicates the factor by which the mean-square reconstruction error is reduced when applying an optimal separate quantizer to each transform component, as compared to quantizing the signal samples directly (PCM). The transformation that maximizes G_{TC} is the Karhunen-Loève Transform (KLT). The KLT is computationally involved and, therefore, a whole class of suboptimal transforms have been proposed which have the advantage of being signal-independent and for which fast algorithms (of complexity $O(N \log N)$) exist. The DCT is perhaps the most popular example in this class.

Recently, image coders [1], [15] based on the *wavelet transform* (WT) [13], [6], [5] have been introduced which offer the following advantages.

- The WT is a *multiresolution* description of an image: the decoding can be processed sequentially from a very low resolution, corresponding to a very compact code, to the highest resolution. Schemes based on the Laplacian pyramid also offer this advantage, but the WT allows furthermore *perfect reconstruction*.
- The WT is closer to the human visual system that the DCT transform. Hence, the artifacts introduced by WT coding with a high compression ratio and adequate *perceptual quantization* are less annoying than those introduced at the same bit rate by the DCT.
- The wavelet transformation of an image generates a data structure known as the *scale-space representation* [16]: in this representation, the high (spatial) frequency signals are precisely located in the pixel domain, while the low-frequency signals are precisely located in the frequency domain. Whereas the DCT has a spatial resolution that is independent of frequency, the spatial resolution of the WT increases linearly with frequency. Therefore, sharp edges, which are well localized spatially and have a significant high-frequency content, can be represented more compactly with the WT than with the DCT (*zooming* capacity of the WT). On the other hand, the overall spectrum of most images is very much of a low-pass type. Now, while the frequency resolution is independent of frequency for the DCT, it is inversely proportional to frequency in the WT. This allows the WT to separate the dominating low frequency end of the spectrum into increasingly finer subbands. For strongly low-pass signals, the spectra of the

subband signals provided by the WT will be whiter in the subbands that contribute significantly to the signal power.

These remarks still apply when one substitutes the DCT by (uniform) subband coders or schemes based on the short-time Fourier transform. The nice properties of the WT just sketched only depend on the multiresolution structure of the WT. However, the choice of the particular wavelet remains an open issue. In applications in physics, particular attention is paid to the degree of smoothness of the wavelet function [5]. In this paper, we address the wavelet design issue when using the WT for transform coding images. In fact, we shall consider a more general class of multiresolution transforms (MT), of which the WT is just a particular instance. This MT is still built up by an iterative application of two-band filter banks. The choice of the wavelet in the WT corresponds in the more general MT to the design of the filter banks. We shall design the filter banks so as to maximize the coding gain. Whereas the KLT maximizes the coding gain over the set of all linear orthogonal transforms, we shall seek the optimal transform when the maximization is performed over the subset of linear orthogonal transforms that have the multiresolution structure (with the additional constraint that the filters in the filter bands are taken to be FIR filters). Another interpretation would be to consider the optimized MT as the cascade of a fixed MT followed by (constrained) adaptive predictive coding in each subband, in the sense that here the adaptive filter is incorporated in the filter bank.

II. Para-Unitary Lattice Filtering System

Consider a two-band analysis filter bank as depicted in Fig. 1, with a low-pass FIR filter $H(z)$ and a high-pass filter $G(z)$, and critical (factor 2) downsampling. In this paper, the synthesis filter bank is assumed to be the *dual* of the corresponding analysis filter bank. Furthermore, the overall system is required to enjoy the *perfect reconstruction property* [2], [19], i.e., to satisfy $\hat{x} = x$ in the absence of quantization. It is interesting to translate these conditions in terms of the *scattering matrix* $S(z)$ of the analysis filter bank, by use of the "polyphase approach" [19]. $S(z)$ is defined by

$$S(z) = \begin{bmatrix} H_0(z) & H_1(z) \\ G_0(z) & G_1(z) \end{bmatrix}, \tag{2}$$

from the even/odd decomposition $H(z) = H_0(z^2) + z^{-1}H_1(z^2)$ and $G(z) = G_0(z^2) + z^{-1}G_1(z^2)$ of the filters $H(z)$ and $G(z)$. The duality and perfect reconstruction properties are expressed by the fact that $S(z)$ is para-unitary, i.e., satisfies $S^T(z^{-1})S(z) = I$.

The $H(z)$ and $G(z)$ filtering operations followed by downsampling in the analysis bank induce operators [5], [13], [14]

$$\mathscr{H} : l^2(Z) \to l^2(2Z), \quad \mathscr{G} : l^2(Z) \to l^2(2Z), \tag{3}$$

where l^2 denotes the space of doubly infinite square-summable sequences. Similarly, the $H(z^{-1})$ and $G(z^{-1})$ filtering operations preceded by downsampling in the synthesis bank induce the operators $\mathscr{H}^* : l^2(2Z) \to l^2(Z)$ and $\mathscr{G}^* : l^2(2Z) \to l^2(Z)$, where the star symbol denotes the adjoint. This adjoint property follows from the para-unitarity of $S(z)$, which furthermore leads to

$$\mathscr{H}^*\mathscr{H} + \mathscr{G}^*\mathscr{G} = I, \quad \mathscr{H}\mathscr{H}^* = \mathscr{G}\mathscr{G}^* = I, \quad \mathscr{H}\mathscr{G}^* = \mathscr{G}\mathscr{H}^* = 0. \tag{4}$$

So $\mathscr{H}^*\mathscr{H}$ and $\mathscr{G}^*\mathscr{G}$ are projection operators onto the two subspaces in the following orthogonal decomposition

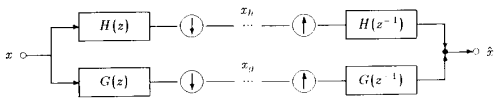


Fig. 1. Critically subsampled two-band filter bank.

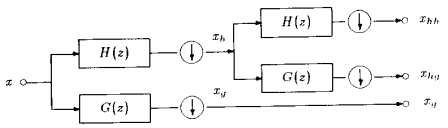


Fig. 2. Analysis section of a two-level multiresolution filter bank.

$$\mathscr{V}_0 \triangleq l^2(z) = \mathrm{Im}\,(\mathscr{H}^*) \oplus \mathrm{Im}\,(\mathscr{G}^*) \triangleq \mathscr{V}_{-1} \oplus \mathscr{W}_{-1}. \tag{5}$$

In the usual case where H and G form a low-pass and high-pass filter pair, the signals $\mathscr{H}^*\mathscr{H}x \in \mathscr{V}_{-1}$ and $\mathscr{G}^*\mathscr{G}x \in \mathscr{W}_{-1}$ can be interpreted as the *approximation* and the *detail* signal of x at the resolution 2^{-1}. In multiresolution analysis (of dyadic type), the decomposition described above is applied in a "recursive manner" to the approximation signals (see Fig. 2). In general, the scattering matrix used at each level can be different. If the above method is applied n times, then we get the decomposition

$$\mathscr{V}_0 = \mathscr{W}_{-1} \oplus \mathscr{W}_{-2} \oplus \cdots \oplus \mathscr{W}_{-n} \oplus \mathscr{V}_{-n}$$

$$\mathscr{W}_{-m} = \mathrm{Im}\,(\mathscr{H}_1^* \cdots \mathscr{H}_{m-1}^* \mathscr{G}_m^*), \quad \mathscr{V}_{-m} = \mathrm{Im}\,(\mathscr{H}_1^* \cdots \mathscr{H}_m^*) \tag{6}$$

where $\mathscr{H}_m$, $\mathscr{G}_m$ correspond to the scattering matrix $S_m(z)$ at resolution level m. The signal x gets thus transformed into its sets of detail coefficients $x_g, x_{hg}, \cdots, x_{h\cdots hg}$ at the resolutions 2^{-1}, $2^{-2}, \cdots, 2^{-n}$ and the residual set of approximation coefficients $x_{h\cdots hh}$ at the resolution 2^{-n} (see Fig. 2). This procedure is known in the literature as the *Laplacian pyramid decomposition*, with n *octaves* (see [3], [13], [15]).

Although our multiresolution filtering system is closely related to the "discrete wavelet transform" [6], [14], [17], it does not belong to wavelet theory in a strict sense, for two reasons. First, the filter bank (H, G) is allowed to differ from one resolution level to the next (which is not the case in the Mallat–Meyer multiresolution theory). Second, even if the filter banks were the same at all resolution levels, the only constraint imposed (*a priori*) on our QMF systems is the para-unitarity property (see a remark in Section V-A though). While the recurrence coefficients of orthogonal wavelet theory define para-unitary QMF systems, the converse statement is true only if the given QMF system (H, G) enjoys some additional properties (ensuring that the well-known "dilation equation" of orthogonal wavelet theory has an appropriate solution for the "scaling function"). The simplest property of this type is the relation $H(-1) = G(1) = 0$ (see again the same remark in Section V-A). Finally, let us note that the multiresolution unitary transform described above for infinite sequences x can be made to apply to finite sequences, when extending these to periodic sequences (see [11]).

A para-unitary scattering matrix $S(z)$, of a given degree $N + 1$ (corresponding to FIR filters H and G of length $2N + 2$), is

known to admit a canonical factorization of the form

$$S(z) = \Omega_N Z^{-1} \cdots \Omega_1 Z^{-1} \Omega_0, \qquad \text{with } Z = 1 \oplus z, \quad (7)$$

where Ω_k denotes a constant orthogonal 2×2 matrix (see [18]). Without real loss of generality, we only consider the case where each Ω_k is a *rotation matrix*, i.e.,

$$\Omega_k = \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix}, \quad (8)$$

with $c_k = \cos \alpha_k$ and $s_k = \sin \alpha_k$ for some angle $\alpha_k \in (-\pi, \pi]$. The parametrization (7) corresponds to a *lattice implementation* of the analysis filter bank [20], [19]. It is depicted in Fig. 3. The inputs p_0 and q_0 of the lattice are deduced from the given signal x by the formulas

$$P_0(z) = X_0(z), \qquad P_1(z) = z^{-1} X_1(z), \quad (9)$$

where $X_0(z)$ and $X_1(z)$ are the even and odd parts of $X(z)$, defined via $X(z) = X_0(z^2) + z^{-1} X_1(z^2)$. (Throughout this correspondence, the z-transform of a discrete-time signal y is represented by $Y(z) = \sum_{i=-\infty}^{+\infty} y(i) z^{-i}$.) More generally, let p_k and q_k denote the inputs of the $(k+1)$st lattice cell, for $k = 0, \cdots, N$. Then, we can write the relation

$$\begin{bmatrix} X_h(z) \\ X_g(z) \end{bmatrix} = S_k(z) \Omega_k \begin{bmatrix} P_k(z) \\ Q_k(z) \end{bmatrix}, \quad (10)$$

where $S_k(z)$ is the scattering matrix of the cascade formed with the last $N - k$ lattice sections. In particular, $S_0(z) = S(z) \Omega_0^T$. By definition, we have the recurrence

$$S_{k-1}(z) = S_k(z) \Omega_k Z^{-1}, \qquad \text{with } S_n(z) = I. \quad (11)$$

III. Derivation of the Ring Algorithm

A. The Minimization Criterion

For a two-band QMF bank, the coding gain is given by (1) with $N = 2$, and $\sigma_1 = \sigma(x_h)$, $\sigma_2 = \sigma(x_g)$. Since the sum $\sigma^2(x_h) + \sigma^2(x_g)$ equals the constant $2\sigma^2(x)$, due to para-unitarity, the maximization of the coding gain naturally leads to the following criterion: *minimize the variance $\sigma^2(x_g)$ of the "high-pass branch" output signal x_g of the filter bank.* More precisely, the distinction between the H-branch and the G-branch is defined from the minimization criterion itself. In the Laplacian pyramid structure, it is natural to apply this criterion to all "high-pass branches," i.e., to the detail signals x_g, x_{hg}, x_{hhg}, and so on. Since most images have an overall low-pass spectrum however, this design will lead to H and G being actual low-pass and high-pass filters, respectively.

In the framework of image coding, we have to deal with two-dimensional discrete signals. We shall restrict ourselves to the simple case of *separable filters* (see [13], [15] and Fig. 4). Thus, we have a filter bank $(H(z), G(z))$ for the "horizontal variable," and a filter bank $(H(w), G(w))$ for the "vertical variable." Note that these two filter banks need not be the same (in spite of what the notation might suggest). In this case, as we wish to keep a "one-variable technique," we suggest that the same criterion be applied to the downsampled output x_g of the horizontal filter $G(z)$, and to the downsampled output x^g of the vertical filter $G(w)$, although the latter does not appear "physically" in the filtering system under consideration. So in the multiresolution two-dimensional case, we apply the optimization of a twoband filter bank consecutively to all resolution levels, and on a given resolution level to the horizontal and the vertical direction.

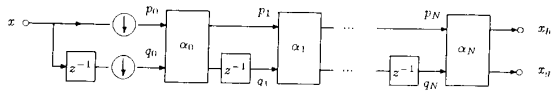


Fig. 3. Lattice realization of the analysis part of a two-band filter bank.

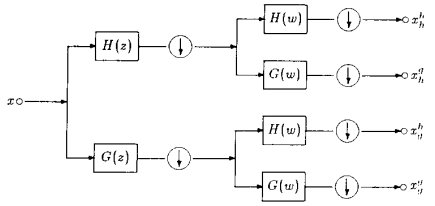


Fig. 4. Separable 2-D extension of a two-band filter bank.

In what follows, we shall be interested in minimizing the function $\sigma^2(x_g)$ over all para-unitary FIR filtering systems (H, G) of length $2N + 2$, for given second-order statistics of the input signal x. The proposed minimization strategy is based on the observation that $\sigma^2(x_g)$ is quadratic in (the cosine and sine of) one rotation angle α_i when all other angles are frozen, and hence is trivially optimized w.r.t. one such angle. This suggests an alternating optimization strategy whereby in each sweep through the lattice, the rotation angles $\alpha_0, \ldots, \alpha_N$ are updated one after the other; the new angle α_k is determined so as to minimize the function $\sigma^2(x_g)$. By its very nature, the algorithm is guaranteed to converge to a local minimum of this function, whatever the initial values of the angles. (The initialization $\alpha_k = 0$, for all k, seems to be a good choice.) The final result of a given sweep can be used to initialize the next sweep. For this reason, the proposed method has a kind of "ring structure" (see Fig. 5); for easy reference, we shall call it the *ring algorithm*.

The optimization procedure for a twoband filter bank described here can also be used for the design of such a filter bank in a more general context. Typically [9], [20], one considers a least-squares measure for the stopband attenuation, or hence $\sigma^2(x_g)$ with an appropriate weighting function playing the role of the spectral density of the input signal (see (28)). In [9], an iterative minimization procedure is proposed based directly on the coefficients of the filters H and G. This leads to a quadratic criterion subject to a number of quadratic constraints. Apart from advantages in a finite-wordlength environment, the lattice realization has the advantage of structurally incorporating all the QMF constraints so that the lattice parameters represent explicitly the actual degrees of freedom. These advantages were also exploited in [20], where a quasi-Newton algorithm is used to optimize the lattice parameters. These parameters are first initialized by a calculation which is similar to the first sweep in our algorithm. However, a so-called "unnormalized" lattice is used in [20] so that the paraunitarity property is lost. In that case, the minimization of $\sigma^2(x_g)$ does not correspond to the maximization of the coding gain.

B. Computation of the Optimal Section

Let us explain how to determine the updated elements c_k and s_k of the $(k+1)$st lattice section (without performing any "trigono-

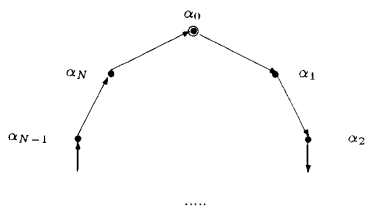


Fig. 5. Ring algorithm schematically.

metric calculation''). The scattering matrix $S_k(z)$ has the form

$$S_k(z) = \begin{bmatrix} z^{k-N}B_k(z^{-1}) & -z^{k-N}A_k(z^{-1}) \\ A_k(z) & B_k(z) \end{bmatrix}, \quad (12)$$

with $A_k(z) = \sum_{i=0}^{N-k} A_{k,i}z^{-i}$ and $B_k(z) = \sum_{i=0}^{N-k} B_{k,i}z^{-i}$ for some real numbers $A_{k,i}$ and $B_{k,i}$. The second entry of (10) can be written as

$$X_g(z) = c_k V_k(z) - s_k U_k(z), \quad (13)$$

where $U_k(z)$ and $V_k(z)$ are given by

$$\begin{bmatrix} U_k(z) \\ V_k(z) \end{bmatrix} = \begin{bmatrix} B_k(z) & -A_k(z) \\ A_k(z) & B_k(z) \end{bmatrix} \begin{bmatrix} P_k(z) \\ Q_k(z) \end{bmatrix}. \quad (14)$$

From (13), we deduce

$$\sigma^2(x_g) = c_k^2\sigma^2(u_k) - 2c_k s_k \operatorname{cov}(u_k, v_k) + s_k^2\sigma^2(v_k), \quad (15)$$

where $\operatorname{cov}(u_k, v_k)$ is the covariance of the signals u_k and v_k.

It is a simple exercise to minimize the function (15) with respect to the variable α_k, for fixed values of $\sigma^2(u_k)$, $\sigma^2(v_k)$, and $\operatorname{cov}(u_k, v_k)$. The result is the following:

$$c_k = (a_k + r_k)/t_k, \quad s_k = \operatorname{cov}(u_k, v_k)/t_k, \quad (16)$$

with $a_k = (\sigma^2(u_k) - \sigma^2(v_k))/2$, $r_k = [a_k^2 + (\operatorname{cov}(u_k, v_k))^2]^{1/2}$, and $t_k[2r_k(a_k + r_k)]^{1/2}$. This yields the value $\sigma^2(x_g) = (\sigma^2(u_k) + \sigma^2(v_k))/2 - r_k$.

C. Computation of the Covariance Data

Next, let us examine how to compute the quantities $\sigma^2(u_k)$, $\sigma^2(v_k)$, $\operatorname{cov}(u_k, v_k)$, from the input data. From $P_k(z)$ and $Q_k(z)$ construct the vector of length $2(N - k + 1)$

$$W_k(z) = \left[\left(z^{-i}P_k(z)\right)_{i=0}^{N-k}, \ \left(z^{-i}Q_k(z)\right)_{i=0}^{N-k}\right]^T. \quad (17)$$

Let $\mathcal{A}_k = (A_{k,i})_{i=0}^{N-k}$ and $\mathcal{B}_k = (B_{k,i})_{i=0}^{N-k}$ denote the row-vectors formed with the coefficients of the filters $A_k(z)$ and $B_k(z)$. Define Φ_k as the covariance matrix of the signal vector having $W_k(z)$ as its z-transform. By use of (14) we obtain the formula

$$\begin{bmatrix} \sigma^2(u_k) & \operatorname{cov}(u_k, v_k) \\ \operatorname{cov}(u_k, v_k) & \sigma^2(v_k) \end{bmatrix} = \begin{bmatrix} \mathcal{B}_k & -\mathcal{A}_k \\ \mathcal{A}_k & \mathcal{B}_k \end{bmatrix} \Phi_k \begin{bmatrix} \mathcal{B}_k^T & \mathcal{A}_k^T \\ -\mathcal{A}_k^T & \mathcal{B}_k^T \end{bmatrix}. \quad (18)$$

This involves approximately $4(N - k + 1)^2$ arithmetic operations; hence, each sweep of the ring algorithm has $O(N^3)$ complexity.

As to the covariance matrix Φ_k, it can be computed recursively as follows. Consider the ''expanded version'' of the input-output relation of the kth lattice cell, that is

$$W_k(z) = R_{k-1}W_{k-1}(z), \quad (19)$$

with the $2(N - n) \times 2(N - n + 1)$ matrix R_n given by

$$R_n = \begin{bmatrix} c_n I & 0 & s_n I & 0 \\ 0 & -s_n I & 0 & c_n I \end{bmatrix}. \quad (20)$$

(Here, I and 0 stand for the unit matrix and the zero column vector of size $N - n$.) From (19) we deduce

$$\Phi_k = R_{k-1}\Phi_{k-1}R_{k-1}^T. \quad (21)$$

This allows us to determine the Φ_k sequence (for $k = 1, \cdots, N$). The initial matrix Φ_0, of order $2(N + 1)$, is formed with the input covariance lags $\operatorname{cov}(x, z^{-i}x)$ for $i = 0, \cdots, 2N + 1$. It should be noted that the Φ_k matrices have the *Toeplitz-block* structure, which entails a substantial economy in the actual computation scheme (21). We shall not elaborate on this issue.

D. Computation of the Filter Data

Let us now see how to recursively compute the filter coefficients $\mathcal{A}_k$ and $\mathcal{B}_k$, involved in (18), for $k = 1, \cdots, N$, from the initial row-vectors $\mathcal{A}_0$ and $\mathcal{B}_0$, by a suitable *downdating procedure*, for every sweep of the ring algorithm. In view of (12), the second row of the inverse form of (11) can be written as

$$\begin{bmatrix} \mathcal{A}_k & \mathcal{B}_k \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{k-1} & \mathcal{B}_{k-1} \end{bmatrix} D_k^T, \quad (22)$$

with the $2(N - k + 1) \times 2(N - k + 2)$ matrix

$$D_k = \begin{bmatrix} c_k I & 0 & s_k I \\ -s_k I & 0 & c_k I \end{bmatrix}. \quad (23)$$

(Here, I is the unit matrix of size $N - k + 1$ and 0 is the $(N - k + 1) \times 2$ zero matrix.)

To complete the algorithm description, we will finally explain how each sweep (except the first one) can be initialized directly from the results obtained at the end of the preceding sweep. Consider the scattering matrix $S_0(z) = S(z)\Omega_0^T$ we are interested in, and its factorized form deduced from (11). For $k = 1, \cdots, N$, compute the ''partial'' scattering matrix $T_k(z)$ by the recursion

$$T_k(z) = \Omega_k Z^{-1}T_{k-1}(z), \quad T_0(z) = I. \quad (24)$$

The outcome yields the desired matrix, i.e., $S_0(z) = T_N(z)$. Let us give some details about this computation scheme. The matrix $T_k(z)$ has for form

$$T_k(z) = \begin{bmatrix} z^{-k}F_k(z^{-1}) & -z^{-k}E_k(z^{-1}) \\ E_k(z) & F_k(z) \end{bmatrix}, \quad (25)$$

with $E_k(z) = \sum_{i=0}^{k} E_{k,i} z^{-i}$ and $F_k(z) = \sum_{i=0}^{k} F_{k,i} z^{-i}$. In view of (8) and (25), the recurrence (24) amounts to

$$E_k(z) = c_k z^{-1} E_{k-1}(z) - s_k z^{1-k} F_{k-1}(z^{-1}),$$
$$F_k(z) = s_k z^{1-k} E_{k-1}(z^{-1}) + c_k z^{-1} F_{k-1}(z). \quad (26)$$

The initial values in (26) are $E_0(z) = 0$ and $F_0(z) = 1$. The desired filter data $\mathscr{A}_0$ and $\mathscr{B}_0$ (for the new sweep) are given by

$$A_0(z) = E_N(z), \quad B_0(z) = F_N(z). \quad (27)$$

E. Further Comments About the Filter Design

For a stationary input signal x, we can write the variance of the high-pass subband signal in terms of the spectral density $S_x(\omega)$ of x as follows (note: $|G(e^{j\omega})|^2 + |G(e^{j(\pi-\omega)})|^2 = 1$):

$$\sigma^2(x_g) = \frac{1}{\pi}\int_0^{\pi} |G(e^{j\omega})|^2 S_x(\omega)\, d\omega$$
$$= \frac{1}{\pi}\int_{\frac{\pi}{2}}^{\pi} S_x(\omega)\, d\omega$$
$$+ \frac{1}{\pi}\int_0^{\frac{\pi}{2}} |G(e^{j\omega})|^2 \left[S_x(\omega) - S_x(\pi-\omega)\right] d\omega. \quad (28)$$

Hence, $S_x(\omega) > S_x(\pi-\omega)\ \forall \omega \in \left(0, \frac{\pi}{2}\right)$ implies $G(e^{j\omega}) = 0$, $\forall \omega \in \left(0, \frac{\pi}{2}\right)$. So, if for instance the spectrum of the signal x is nonincreasing, then the optimization problem leads to a perfect half-band filter. Since many natural images have a low-pass spectrum, an obvious design goal for a nonadapted filter G would be to approximate a perfect half-band filter. With a finite filter complexity, this will lead to a uniform attenuation in the lower half-band. Now imagine that the signal x is perfectly low-pass with bandwidth $\beta\frac{\pi}{2}$ ($\beta \in (0, 1)$), then the width of the stop-band can be reduced by a factor β and hence the uniform attenuation in the stop-band will improve (for a filter with constant complexity). Hence, as β decreases, the adapted filters will give increasing improvement over nonadapted ones (see [8], [7] for examples).

Of course, not all signals have a nonincreasing spectrum. For instance, images that are dominated by a piece of texture may show strong peaks (see the shirt in the image COUPLE below). Consider the extreme case of a spectral peak, namely a sinusoidal component. The adapted filter G can eliminate its contribution to $\sigma^2(x_g)$ with one pair of zeros. More generally, peaks in the spectrum of x are advantageously exploited in the adaptive approach. The optimum filter for real images is usually smoother than the half-band approximations (see [8], [7] for examples). In practice, it leads to reconstructed images that are visually more pleasing because they contain less ringing effects.

IV. A Multiresolution Image Compression Algorithm

Compression results from the quantization and the entropy coding of the transform coefficients. The decorrelative property of the transform allows one to encode the transform coefficients independently from each other. The match of the multiresolution transform with the human visual system allows an efficient quantization of the transform coefficients. In particular, the high-frequency transform coefficients can be encoded in a coarse manner, since this will produce high-frequency noise, with low visibility. Furthermore, the adaption of our transform leads to the minimization of the energy in the detail subbands. In this section, we will briefly describe a specific perceptual quantization scheme, adapted to multiresolution transforms, and a specific block-adaptive Huffman code.

A. Perceptual Quantization

We have designed a weighted quantization scheme so as to minimize the weighted noise power (see [12])

$$(P_N)_w = \sum_{n1}\sum_{n2} \frac{1}{N_{n1}N_{n2}} W_{n1,n2}\sigma^2_{n1,n2}, \quad (29)$$

where

$$W_{n1,n2} = K \iint W(u,v)\left|F_{n1,n2}(u,v)\right|^2 du\, dv. \quad (30)$$

The summation is over the different subbands; $\sigma^2_{n1,n2}$ is the variance of the quantization error in subband $(n1, n2)$; N_{n1}, N_{n2} are the downsampling factors for the subbands; u, v are the horizontal and vertical frequencies, in cycles per image height; $F_{n1,n2}(u, v)$ is the transmittance of the reconstruction filter; K is an energy factor; $W(u, v)$ is a weighting curve: we have chosen the extension of the CCIR 451-2 recommendation for random wide-band noise affecting the quality of television transmission (see [12] for more details)

$$W(u,v) = 15.32\left(1 + \frac{u^2+v^2}{\left(3.952\,\mathrm{arctg}_{(\text{in degree})}(1/2L)\right)}\right)^{-1.5} \quad (31)$$

where L is the distance in image height from the screen.

It is shown in [12] that, in the case of entropy coding, the optimum quantization is achieved by linear quantization steps inversely proportional to $V_{n1,n2} = \sqrt{W_{n1,n2}}$. This can be achieved by a preaccentuation before the quantization: each wavelet transform coefficient is multiplied by a visibility factor $V_{n1,n2}$. We have computed the visibility factors for a wavelet transform with four octaves on a 512 × 512 image for a viewer at four times the image height. For computational simplicity, we have assumed that the filters G and H are perfect half-band filters. The visibility factors are given in the following list where the lowest frequency factor is normalized ($V^{hhhh}_{hhhh} = 1$):

Octave 0: $V^{hhhg}_{hhhh} = 0.738$, $V^{hhhh}_{hhhg} = 0.738$, $V^{hhhg}_{hhhg} = 0.596$,
Octave 1: $V^{hhg}_{hhh} = 0.584$, $V^{hhh}_{hhg} = 0.584$, $V^{hhg}_{hhg} = 0.411$,
Octave 2: $V^{hg}_{hh} = 0.355$, $V^{hh}_{hg} = 0.355$, $V^{hg}_{hg} = 0.230$,
Octave 3: $V^{g}_{h} = 0.189$, $V^{h}_{g} = 0.189$, $V^{g}_{g} = 0.119$.

B. Adaptive Entropy Coding

For the entropy coding of the quantized wavelet coefficients, we have developed a block-adaptive Huffman code (for more details see [12]). The use of a Huffman code requires the probability distribution of the MT coefficients. A good approximation is the Laplace distribution, viz.

$$p(x) = \exp\left(-|x|\frac{\sqrt{2}}{\sigma}\right)\Big/(\sqrt{2}\sigma),$$

in which the variance, σ, varies as a function of the subband, the quantization level and the location in the subband, suggesting the use of an adaptive Huffman code. We construct vectors of 128

(a)

(b)

Fig. 6. LENA 512 × 512, coded at 0.1 b/pixel with a four-level MT; (a) with 4-cell adapted filters, (b) with a 16-cell standard half-band approximation.

(a)

(b)

Fig. 7. COUPLE 512 × 512, coded at 0.08 b/pixel with a 4-level MT; (a) with 4-cell filters adapted to LENA, (b) with 4-cell adapted filters.

coefficients belonging to the same subband. The vectors are obtained by scanning the horizontal subbands column by column, the vertical subbands line by line, and the diagonal subbands diagonally. The code consists of three classes of words:

- a prefix, which indicates the number of nonzero values in the vector (the prefix allows an estimation of the variance σ inside the vector);
- Huffman codes for the nonzero values adapted to the measured Laplace distribution (127 different Huffman codes have to be memorized);
- Adaptive truncated runlength (ATRL) codes for the zero values.

V. Some Results

The implementation of our coding scheme has allowed us to draw the following conclusions.

- low-order filters (four cells for most natural images) are enough to reach the minimum bit-rate with our scheme;
- the ring algorithm requires at least 150 sweeps (in the case of four cells);
- the filter bank obtained by the ring algorithm leads to a better quality of the decoded image than the corresponding half-band

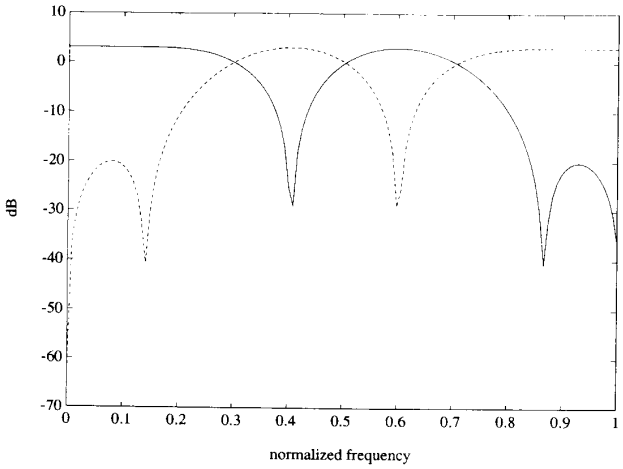


Fig. 8. Frequency response of the 4-cell adapted filters H and G, for the horizontal direction at resolution 1/2 in COUPLE.

approximation; furthermore, for most natural images, the signal-adapted filter G has a pole close to dc (G almost satisfies the regularity condition appearing in the wavelet theory);

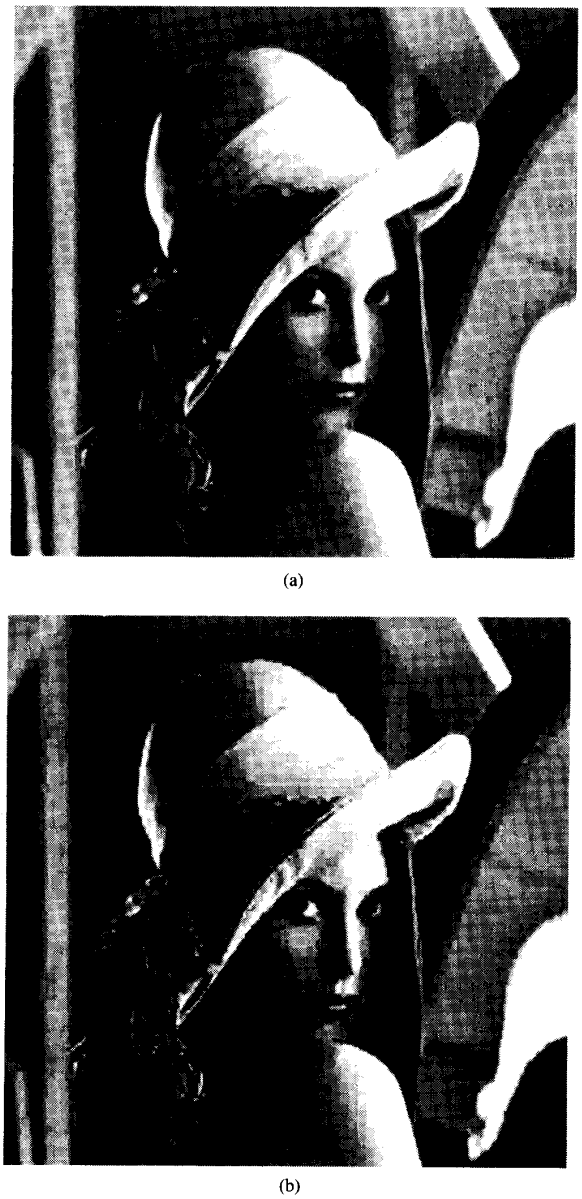
- the filter bank obtained by the ring algorithm allows one to better preserve a large area with a specific structure, and to better preserve the moving information when the algorithm is applied to images of interlaced video sequences;
- for a given compression ratio, the proposed multiresolution coding scheme leads to a better quality than the JPEG standard DCT scheme.

Some of these points are illustrated next.

A. Comparison with Nonadapted Filter Banks

In order to compare the coding obtained with adapted filters and standard half-band filters, we have compared the coding of LENA at a bit rate of 0.1 b/pixel (compression factor 80) with a 16-cell filter similar to one in [20], to the result obtained by means of the 4 cell adapted filter at the same bit rate. This comparison is presented in Fig. 6 and leads to the following two remarks. First, the standard half-band produces more ringing effects near the edges. Second, a standard half-band filter does not have a zero at dc. For this reason, the reconstructed images exhibit some blocking artifacts. Such artifacts can be interpreted if the subband coding is described in terms of orthogonal transforms on overlapping blocks. Indeed, in the case of orthogonal transforms it is easy to show that large errors on the dc component will produce large blocking artifacts (see [12] for more details). The importance for the high-pass filter to have a zero at dc, in order to avoid blocking artifacts, leads us to propose a slight modification of our signal-adapted filter design: after the ring algorithm computation, we suggest to modify the last cell so as to have a high-pass filter that enjoys the required property: $\alpha_N = \pi/4 - \sum_{i=0}^{N-1} \alpha_i$. For most natural images, this correction is small.

In the image COUPLE, there is a large area, the shirt of the man, containing a specific texture, leading to a peak in the spectrum. The signal-adapted filter bank adapts to this peak and the texture is better preserved at high-compression factors. In Fig. 7, the coding of COUPLE at a bit rate of 0.08 b/pixel with the 4-cell LENA filter is compared with the 4-cell signal-adapted version. On the latter version, it can be observed that the stripes of the shirt are preserved. Fig. 8 shows the signal-adapted filter in the horizontal direction, on the highest-resolution level: there is a "hole" at the frequency of the shirt stripes.



(a)

(b)

Fig. 9. LENA 512 × 512, coded at 0.2 b/pixel; (a) with a 4-cell adapted four-level MT, (b) with the JPEG DCT scheme.

B. Comparison with the JPEG Standard DCT Algorithm

In order to assess the performances of our signal-adapted transform, we have chosen to compare it to the JPEG DCT scheme, which is recognized as an international standard and which is also based on perceptual quantization and entropy coding; the main difference between the JPEG DCT coding and our scheme lies in the choice of the transform. We have made the comparison for three compression factors: 10 (0.8 b/pixel), 20 (0.4 b/pixel), and 40 (0.2 b/pixel), on the 512 × 512 image LENA. Both algorithms show no

visible degradation at a compression factor of 10. The artifacts introduced by our multiresolution coding scheme are nevertheless less annoying than those introduced by the DCT for higher compression ratios: this can be observed in Fig. 9 for a compression factor of 40. This is essentially due to the fact that a MT is better tuned to human vision than the DCT.

It is to be noticed that, for a low-correlated image like LENA which contains various types of stimuli, there is no significant difference between the use of a fixed filter bank designed by the ring algorithm on the basis of typical covariance lags and the use of a signal-adapted filter bank. A typical fixed filter bank is given by the 4 cell lattice whose successive angles are (in radians): 1.144826, -0.536006, 0.249848, -0.07327.

VI. Conclusion

We have shown that the optimization of the filter banks in a MT may lead to substantial gains (e.g., a factor of two) in the coding gain. The coding gain is an objective measure, involving the SNR of the reconstructed image. However, the ultimate quality of an image coder is influenced by other aspects than simply the SNR. For instance, specific artifacts introduced by a coding scheme contribute significantly to its subjective quality. We have shown that the typical ringing artifacts in a MT can be bounded since low order FIR filters usually allow to achieve most of the obtainable coding gain.

Some first principles considerations seem to indicate that the introduction of linear-phase constraints may lead to a reduction of artifacts. We are currently investigating signal-adapted MT's with the linear-phase constraint. Another issue that deserves further investigation is the local adaptation of the filter banks. Due to the nonstationary character of images one may indeed expect that space-varying filter banks may lead to further improved coding gains. There is a limit to this varying character though since at some point the coding of the filters will become predominant. A final issue concerns an investigation of the convergence properties of the ring algorithm or of any alternative optimization algorithm (involving homotopy methods?).

References

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using vector quantization in the wavelet transform domain," in *Proc. Int. Conf. Acoust., Speech Signal Processing*, Albuquerque, NM, 1990, pp. 2297-2300.

[2] T. P. Barnwell, III and M. J. T. Smith, "Filter banks for anlysis-reconstruction systems: A tutorial," in *Proc. IEEE Int. Symp. Circuits Syst.*, New Orleans, LA, 1990, pp. 1999-2003.

[3] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, pp. 532-540, 1983.

[4] S. Comes and B. Macq, "Human visual quality criterion," in *Proc. Conf. Visual Commun. Image Processing*, vol. 1360, Lausanne, Switzerland: SPIE, 1990, pp. 2-13.

[5] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909-996, 1988.

[6] ——, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961-1005, Sept. 1990.

[7] P. Delsarte, B. Macq, and D. T. M. Slock, "Signal-adapted multiresolution transform for image coding," rep. no. M 376, Philips Res. Lab. Belgium, Feb. 1991.

[8] ——, "Efficient multiresolution signal coding via a signal-adapted perfect reconstruction filter pyramid," in *Proc. Int. Conf. Acoust. Speech Signal Processing*, Toronto, Canada, 1991.

[9] V. K. Jain and R. E. Crochiere, "Quadrarure mirror filter design in the time domain," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-32, pp. 353-361, 1984.

[10] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[11] G. Karlsson and M. Vetterli, "Extension of finite length signals for sub-band coding," *Signal Processing*, vol. 17, pp. 161-168, 1989.

[12] B. Macq, "Perceptual transforms and universal entropy coding for an integrated approach to picture coding," doctoral thesis, Univ. Louvain, 1989.

[13] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674-693, July 1989.

[14] ——, "Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, pp. 69-87, 1989.

[15] ——, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 37, pp. 2091-2110, 1989.

[16] J. B. O. S. Martens and G. M. M. Majoor, "The perceptual relevance of scale-space image coding," *Signal Processing*, vol. 17, pp. 353-364, 1989.

[17] Y. Meyer, *Ondelettes*. Paris, France: Hermann, 1990.

[18] V. P. Potapov, "The multiplicative structure of J-contractive matrix functions," in *American Mathematical Society Translations, Series 2, Vol. 15*. Providence, RI: Amer. Math. Soc., pp. 131-243, 1960.

[19] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE*, vol. 78, pp. 56-93, 1990.

[20] P. P. Vaidyanathan and P.-Q. Hoang, "Lattice structures for optimal design and robust implementation of two-channel perfect-reconstruction QMF banks," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 36, pp. 81-94, 1988.

Correlation Structure of the Discrete Wavelet Coefficients of Fractional Brownian Motion

A. H. Tewfik, *Member, IEEE*, and M. Kim

Abstract—**It is shown that the discrete wavelet coefficients of fractional Brownian motion at different scales are correlated and that their auto and cross-correlation functions decay hyperbolically fast at a rate much faster than that of the autocorrelation of the fractional Brownian motion itself. The rate of decay of the correlation function in the wavelet domain is primarily determined by the number of vanishing moments of the analyzing wavelet.**

Index Terms—**Wavelets, fractional Brownian motion, stochastic processes, multiscale analysis.**

I. Introduction

Fractional Brownian motion (fBm) [1] is a generalization of the usual Brownian motion. It was introduced to model processes that have long memory and/or a statistical self-similarity property. Although it is not stationary, its increments are *stationary* and *self-similar*. Its sample paths are fractal with probability one [2], i.e., the graph of fBm has a Hausdorff-Besicovitch dimension that is larger than its topological dimension with probability one. Fractional Brownian motion has been used in image generation and interpolation, texture classification and the modeling of burst errors in communication channels, $1/f$ noise in oscillators and current noise in metal films and semiconductor devices.

Manuscript received November 8, 1990; revised September 5, 1991. This work was supported in part by the National Science Foundation under Grant MIP-9010111.

The authors are with the Department of Electrical Engineering, University of Minnesota, Room 4-174 EE/CSci Building, Minneapolis, MN 55455.

IEEE Log Number 9105308.