[9] P. Noll and R. Zelinski, "Bounds on quantizer performance in the low bit-rate region," *IEEE Trans. Commun.*, vol. 26, pp. 300–304, Feb. 1978.
[10] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, pp. 14–38, Oct. 1991.
[11] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 205–220, Apr. 1992.
[12] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.
[13] B. Macq, "Weighted optimum bit allocation to orthogonal transforms for picture coding," *IEEE J. Selected Areas in Communications*, vol. 10, pp. 875–883, June 1992.

Regular Wavelets: A Discrete-Time Approach

Olivier Rioul

***Abstract*—Regularity is a new filter property, brought by wavelet theory, for perfect reconstruction octave-band filter banks. Tools for investigating its role in coding applications are provided in this note. First, discrete-time interpretations an optimal estimates of regularity are reviewed. Then, a simple design procedure for paraunitary FIR filter banks with optimal trade-off between frequency selectivity and regularity is given. Finally, the obtained filters are used to measure the effect of regularity versus frequency selectivity in a still image compression scheme with optimized rate-distortion. In this case, regularity is shown to be more relevant than frequency selectivity, especially for short filters.**

I. Introduction

Signal compression was claimed to be a major potential application of wavelets [16]. In fact, the discrete wavelet transform (DWT) was soon recognized to be equivalent to an octave-band filter bank allowing perfect reconstruction, which was successfully applied for some time in subband coding of speech and images [3], [18]. The main novelty of wavelets compared to traditional subband coding is the additional requirement of *regularity* on the filters [4]. However, the actual impact of regularity is not clear at this time. Therefore, it is important to understand its role in coding systems, in competition with other filter properties such as frequency selectivity and phase. The aim of this paper is to provide the tools for measuring the impact of regularity on system performance. We proceed in several steps toward solving this problem.

First, we quickly review some theoretical background for regularity, based on the results of [14]. Regularity is here explained using a discrete-time approach, which leads to very sharp regularity estimates which have practical algorithms applicable to any FIR filter bank (in one dimension). Second, we derive a simple design procedure for paraunitary FIR filter banks (generating orthonormal, compactly supported wavelets) which allows us to vary regularity and frequency selectivity quite independently, with several possible choices for the phase reponse. This overcomes the present limitation on the number of regular filter solutions available today in the literature—the most famous ones being Daubechies solutions [4]. Finally, using the filters previously designed, we provide an example of image coding application, in which we measure the effect of regularity on compression performance. The coding results are preliminary and demand further investigation, e.g. concerning the choice of filters.

Manuscript received September 1, 1992; revised June 10, 1993. The Guest Editor coordinating the review of this paper and approving it for publication was Dr. Ahmed Tewfik.

The author is with the Centre National d'Etudes des Communications, 92131 Issy-Les-Moulineaux, France.

IEEE Log Number 9212433.

II. Theoretical Approach: Interest and Limitations

To explain regularity, we start with a simple interpolation procedure, which dominates all that follows. Given a discrete-time signal $\{x_n\}$ and coefficients $\{h_n\}$ of a low-pass filter, $n = 0, \cdots, L-1$, we compute $y_n = \Sigma_x x_k h_{n-2k}$. This is perhaps easier to write in z-transform notation

$$Y(z) = X(z^2)H(z). \tag{1}$$

The step from $X(z)$ to $X(z^2)$ is up-sampling; multiplication by $H(z)$ is filtering in the time domain. Hence, (1) is the basic building block in a synthesis filter bank (Fig. 1).

The construction of continuous-time wavelets simply consists of iterating this scheme indefinitely. That is, we consider the iterated filter bank of Fig. 1. After i iterations the output becomes

$$Y(z) = X(z^{2^i})H^i(z) \tag{2}$$

where

$$H^i(z) = H(z)H(z^2)\cdots H(z^{2^{i-1}}) \tag{3}$$

is the equivalent filter in the path going through i low-pass branches (see Fig. 2). The idea is now to look at the graphs of impulse response $\{h_n^i\}$ as i increases. Since $H^i(z)$ has order $(2^i - 1)(L-1)$, the graphs are plotted against $n2^{-i}$, and may "converge" to a limit function of a continuous variable $\phi(t)$, as illustrated in Fig. 3. (Function $\phi(t)$ is referred to as the *scaling function* in the wavelet literature [4], [11].) In addition, for suitably chosen $H(z)$, $\phi(t)$ is a smooth, or "regular" function of t which vanishes outside the interval $(0, L-1)$. Here, regularity is a smoothness requirement on a continuous-time function generated by $H(z)$, and can be mathematically defined as continuity of this function and its derivatives.

The construction of *wavelet* $\psi(t)$ is now in one quick step [4], [11]

$$\psi(t) = \sum_n g_n \phi(2t - n). \tag{4}$$

This is also the limit function obtained for input $x_n = g_n$, the high-pass filter coefficients; the iterated sequence converging to $\psi(t)$ is $\{g_n^i\}$, where $G^i(z) = H^{i-1}(z)G(z^{2^{i-1}})$ is the equivalent band-pass synthesis filter of Fig. 2.

The same discussion applies to the analysis part of the filter bank, which uses a dual scheme (see Fig. 2). For example, the low-pass branch computes inner products of $\{x_n\}$ with the $\{h_n'^i\}$, which are defined from low-pass filter $H'(z)$ and generate the analysis scaling function $\phi'(t)$ and wavelet $\psi'(t)$.

The idea of generating regular functions from a repeated interpolation scheme is not new. It appears in computer-aided geometric design [8], and the dependence of $\phi(t)$ on $H(z)$ was also observed by Burt and Adelson [2] in the context of pyramid transforms. These works were performed independently of wavelets where regularity does not depend on the perfect reconstruction property of filter banks. But when introducing compactly supported wavelets, Daubechies [4] stated new problems: under which (necessary and sufficient) conditions on $H(z)$ do we have convergence of $\{h_n^i\}$ to $\phi(t)$ and regularity of the limit function?

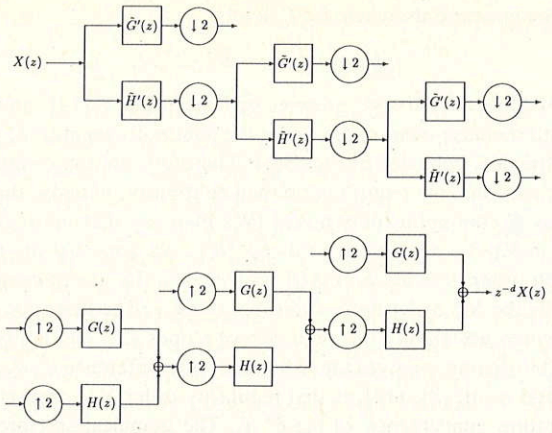

Fig. 1. Perfect reconstruction octave-band filter bank. A two-band filter bank is iterated on the low-pass branch at each step. Low-pass and high-pass filters are denoted by $H(z)$, $H'(z)$, and $G(z)$, $G'(z)$, respectively. At the analysis part, a tilde denotes time-reversal of the filter coefficients. "Orthonormal wavelets" correspond to the paraunitary case in which analysis and synthesis filters are equal: $H(z) = H'(z)$ and $G(z) = G'(z)$.

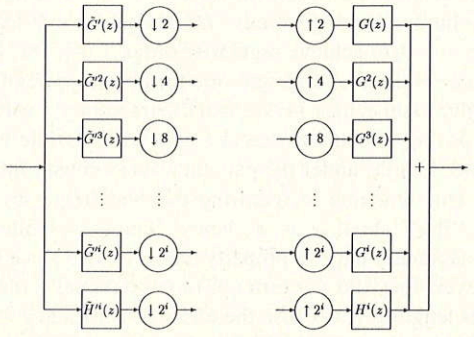

Fig. 2. Redrawing of the filter bank of Fig. 1, revealing equivalent filters (also called basis functions). After i iterations, $H^i(z)$ and $H'^i(z)$ are the equivalent low-pass filters, while $G^i(z)$ and $G'^i(z)$ are the equivalent band-pass filters. The reconstructed signal is decomposed into the sum of synthesis basis functions $h^i_{n-2^i l}$ and $g^k_{n-2^k l}$, $k = 1, \cdots, i$, $l \in \mathbf{Z}$, weighted by the DWT coefficients. These coefficients are themselves computed as inner products of the input with the analysis basis functions $h'^i_{n-2^i l}$ [16].

These questions are motivated by the mathematics behind regularity [4]–[6], where the reasoning comes from wavelet analysis of continuous-signals. But this mathematical approach seems inadequate for *digital* signal processing applications, because the filter bank is never iterated in practical systems: In wavelet-based image compression schemes [1], i seldom exceeds 5, whereas regularity is mathematically defined when $i \to \infty$.

However, we know from Mallat's work [11] that limit functions still underly a discrete filter bank, even when it is iterated a finite number of times. For example, the DWT coefficients at the ith level can be written as $c^i_k = \int x(t)\psi'(2^{-i}t - k)\,dt$, where $\psi'(t)$ is the analysis wavelet and $x(t)$ is a continuous-time signal. But the discrete input is then given by $x_n = \int x(t)\phi'(t - k)\,dt$, where the analysis scaling function $\phi'(t)$ is used as an *ad hoc* sampling filter: Hence, the significance of $x(t)$ is not well established in digital systems. Moreover, scaling functions and wavelets are almost never obtained as explicit functions of t, and the applications only require discrete filters. Therefore, we would like to understand regularity in terms of the discrete filter bank impulse responses $\{g^i_n\}$ and $\{h^i_n\}$.

Regularity does impose some "smoothness" on $\{g^i_n\}$ and $\{h^i_n\}$, as illustrated in Fig. 3(a). There are intuitive arguments in favor of this for coding applications.

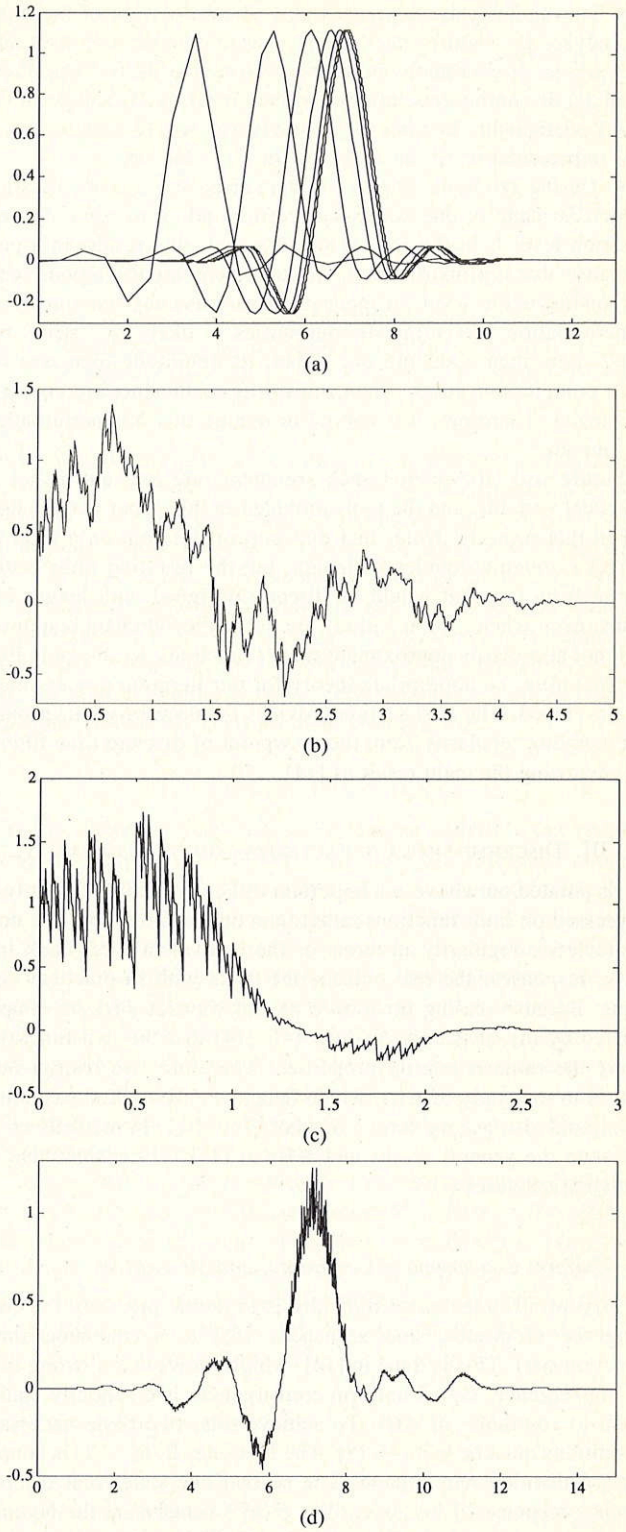

Fig. 3. Limit functions generated by paraunitary low-pass filters $H(z)$. Sequences h^i_n are plotted against $n2^{-i}$ ($i = 7$). Hölder Regularity orders r are indicated between parentheses. (a) 16-tap Daubechies filter [5]. First seven iterations, converging to a regular function, are plotted. The limit function is almost undistinguishable from h^7_n at the level of the figure ($2.72 < r < 2.75$). (b) 6-tap filter designed in Section IV. It has maximal frequency selectivity for normalized transition bandwidth 0.1 under the constrant $H(-1) = 0$. Despite appearances, the limit curve is continuous ($0.196 < r < 0.253$, Sobolev regularity order ≈ -0.071). (c) 4-tap filter given by $h_0 = h_1 = 0.7$ and $h_2 = -h_3 = 0.1$. The iteration diverges ($r \approx -0.12$). (d) 16-tap filter with 25 dB attenuation for normalized transition bandwidth 0.1, designed using Smith and Barnwell method [17]. Because $H(-1) \neq 0$, the iteration diverges ($r \approx -0.16$).

- During analysis: Suppose that a smooth portion of the input is analyzed by "nonregular" filters, whose impulse responses rapidly present discontinuous as i increases (see Fig. 3(c)). Then, these artificial discontinuities, not due to the signal itself, appear in the DWT coefficients. In other words, regularity would lead to a "better" representation of the signal by these coefficients.
- During synthesis: Suppose that an error, e.g., (a quantization error), is made in one coefficient corresponding to some decomposition level i. In the reconstructed signal, this results in a perturbation that is proportional to the equivalent impulse reponse corresponding to this level. In applications such as image compression, a perturbation presenting discontinuities is likely to "strike the eye" more than a smooth one. Also, its amplitude increases for high compression rates, when transform coefficients are coarsely quantized. Therefore, it is natural to require that this perturbation be smooth.

Future will tell whether such arguments are relevant or not in practical systems, and the tools provided in this paper may be useful in this respect.[1] Note, that one step of iteration on a filter of length L roughly doubles its length, but the resulting filter is different from one that would be directly designed with length $2L$. Thus, even when "good" filters are used, time–domain responses will not necessarily approximate $\sin t/t$ functions, as shown in Fig. 3. Therefore, an appropriate theory for our iteration process must be developed. The next section provides the necessary background for studying regularity from the viewpoint of discrete-time filters, by reviewing the main result of [14].

III. Discrete-Time Characterization of Regularity

As pointed out above, an important difficulty is that regularity is expressed on limit functions rather than on the filter taps. We now characterize regularity in terms of the equivalent filter bank impulse responses, the real objects one deals with in practical systems. Because scaling function $\phi(t)$ and wavelet $\psi(t)$ are simply related by (4), it is easy to show [4], [14] that the two functions share the same regularity properties. Therefore, we restrict ourselves to the study of $\phi(t)$, which depends only on low-pass filter $H(z)$, and also use the term "regular" for $H(z)$. In the following, we state the general results and refer to [14] for the treatment of pathological cases.

A. Uniform Convergence, Continuity, and Derivatives

First of all, a technical difficulty is to define precisely the convergence of discrete-time sequences $\{h_n^i\}$ to a continuous-time function $\phi(t)$. This is done in [14], which shows that a strong type of convergence, called uniform convergence, is essentially equivalent to *continuity* of $\phi(t)$. To achieve this, two basic necessary conditions must be met [4], [8]. The first one, $\Sigma_n h_n = 2$, is simply a normalization requirement. The second one states that the frequency response of low-pass filter $H(e^{j\omega})$ vanishes at the Nyquist frequency $\omega = \pi$, i.e., $H(z = -1) = 0$. This is crucial for regularity. Fig. 3(d) shows an example for which $H(-1) \approx 0.05$, leading to rapid oscillations of small magnitude in the graphs of $\{h_n^i\}$ as i increases.

Even when $\phi(t)$ is required to be continuous, it may not appear to be smooth at all, as shown in Fig. 3(b). This is because discontinuities appear only when the "slopes"

$$\delta h_n^i = (h_{n+1}^i - h_n^i)/2^{-i} \tag{5}$$

of the "discrete curves" grow as or faster than 2^i [14], and thus, can still increase indefinitely under the continuity condition, giving the "fractal" behavior of Fig. 3(b). Therefore, to obtain smoother limit functions, we require more than continuity, namely, that $\phi(t)$ possess N continuous derivatives (We then say that the regularity order is N). To characterize this on $\{h_n^i\}$, we note that the role of the Nth order derivative of $\phi(t)$ is played in the discrete-time domain by the Nth *order finite difference* of $\{h_n^i\}$. The first-order finite differences are simply the sequence of slopes (5), and applying N times the operator δ gives the Nth order finite difference $\delta^N h_n^i$. Now, a natural result [8], [14], is that regularity order N is characterized by uniform convergence of the $\delta^N h_n^i$. The graphical interpretation of this is the same as before, but applies to the sequence of slopes, or slopes of slopes, etc., leading to smoother and smoother time–domain responses $\{h_n^i\}$.

B. Zeroes at the Nyquist Frequency

We have seen that one zero in $H(z)$ at $z = -1$ is necessary to obtain continuity. More generally, $H(z)$ must have at least $N + 1$ zeroes at $z = -1$ to achieve regularity order N [4], [8], [14]. This constructive result gives a simple rule for designing regular filters. Accordingly, Daubechies [4] designed paraunitary "wavelet" filters by imposing as many zeroes at $z = -1$ as possible in $H(z)$ for a given filter length, under the paraunitariness constraint (see Section IV). This amounts to requiring that the frequency response $H(e^{j\omega})$ is "flat" about $\omega = \pi$, hence Daubechies filters can be termed "maximally flat," a property that is known since Herrmann [9]. However, they did *not* turn out to be maximally regular for a given filter length [5], because the effect of zeroes at $z = -1$ may be killed by *other* zeroes in $H(z)$, whose effect is destructive for regularity. Daubechies [5] constructed filters for which the destructive effect is less important and compensate for the fact that $H(z)$ does not have the maximum number of zeroes at $z = -1$ (see Section IV).

Typically, this destructive effect kills 80% of regularity [19], so the number of zeroes at $z = -1$ in $H(z)$ does not give a good idea of how "regular" the filter is. Therefore, to characterize and estimate regularity accurately, we would like to measure the precise amount of regularity lost by zeroes in $H(z)$ that are not located at $z = -1$. This is done next.

C. Sobolev and Hölder Regularity

In order to quantify the degree of smoothness accurately, we first extend the definition of regularity order to arbitrary real-valued numbers using Sobolev or Hölder definitions.

Sobolev definition regards regularity as spectral localization: $\phi(t)$ has Sobolev regularity order r if $\int |\omega|^{2r+1} |\Phi(\omega)|^2 \, d\omega < \infty$, where $\Phi(\omega)$ is the Fourier transform of $\phi(t)$. If $r > N$, then $\phi(t)$ has N continuous derivatives. This definition has been so far the most popular [4], [19], and can easily be tackled by estimations on $|H(e^{j\omega})|$ [4]. However, it clearly masks the effect of regularity on the temporal waveform of $\phi(t)$, and ignores phase information of the filter. As a result, this definition is "suboptimal." Fig. 3(b) shows an example for which the best Sobolev regularity order is negative, even though the limit function is in fact continuous.

These drawbacks are avoided in the following Hölder definition of regularity, introduced recently for wavelets [6], which was found to be appropriate for our derivations. The idea is to look at infinitesimal slopes of $\phi(t)$ in time and control the way they grow: For

[1] However, that such tools apply only for a subclass of filter banks, in which filters are iterated on the low-pass filter. In particular, they do not cover techniques such as non-uniform octave-band filter banks, where different filters are used at each stage of decomposition.

$0 < \alpha < 1$, $\phi(t)$ has Hölder regularity order of α if, for any t, h,

$$|\phi(t + h) - \phi(t)| < c|h|^{\alpha} \tag{6}$$

where c is a constant independent of t and h. For higher Hölder regularity orders $r = N + \alpha$, $N = 1, 2, \cdots$, and $0 < \alpha < 1$, the same definition is used on the Nth derivative of $\phi(t)$. In fact, the limit function $\phi(t)$ possesses N continuous derivatives, *if and only if*, it has some Hölder regularity order $> N$ [14]. Hence, Hölder regularity, as opposed to Sobolev regularity, is "optimally" adapted to the study of continuity of $\phi(t)$ and its derivatives. The difference between Hölder and Sobolev regularity only depends on the phase of $H(e^{j\omega})$, and can be estimated: Hölder regularity is always greater than Sobolev regularity by at most $1/2$ [14]. Therefore, Sobolev regularity gives a lower bound as well as an upper bound on Hölder regularity.

A remarkable fact is that Hölder regularity is easily translated in terms of discrete sequences $\{h_n^i\}$. Condition (6) is equivalent to

$$|h_{n+1}^i - h_n^i| \leq c2^{-i\alpha} \tag{7}$$

where c is a constant [14]. Note that (7) can be obtained from (6) by setting $t = n2^{-i}$ and $h = 2^{-i}$, and replacing $\phi(t)$ by h_n^i. In addition, this has a natural graphical interpretation: if the slopes (5) of the discrete curve h_n^i increase as $2^{i(1-\alpha)}$, $0 < \alpha < 1$, the limit function is Hölder regular of order α and is not more regular than that. For higher Hölder regularity orders, simply consider the derivatives of $\phi(t)$, whose discrete-time counterparts are the finite differences $\delta^N h_n^i$ (Section III-A). Hölder regularity order $r = N + \alpha$, $0 < \alpha < 1$, is thus characterized by (7) rewritten with $\delta^N h_n^i$ in place of h_n^i.

Moreover, (7) can be used to show [14] that the rate of convergence of $\{h_n^i\}$ is exponential and is faster as regularity increases. Fig. 3(a) shows that in practice, the convergence is very fast, which justifies the study of the limit function $\phi(t)$, even though the iteration level is limited in practice.

D. Optimal Regularity Estimates

We now briefly review some regularity estimates derived in [13], [14] from the discrete-time characterization of Hölder regularity. The main estimation algorithm [13], [14] first removes all of the zeroes at $z = -1$ in low-pass filter $H(z)$. Such zeros contribute to a (maximal) Hölder regularity order equal to K, the actual number of zeroes at $z = -1$ in $H(z)$ (see Section III-B). The remaining factor, $F(z) = (1 + z^{-1})^{-K} H(z)$, which contains zeroes not located at $z = -1$, will *decrease* this regularity order by the amount $1 - \alpha \geq 0$, yielding an optimal regularity order $r = K - 1 + \alpha \leq K$. The number α is estimated by first iterating $F(z)$ to obtain $F^i(z) = F(z)F(z^2) \cdots F(z^{2^{i-1}})$ associated to the time-domain response f_n^i, then computing $\alpha_i = 1/i \log_2 \max_n \Sigma_k |f_{n+2^i k}^i|$. This is easily implemented on a computer. Given any number of iterations i, the Hölder regularity order will be at least $K - 1 + \alpha_i$, and the estimate is improved as i increases: In practice, the exact (optimal) regularity order r is generally obtained to two decimal places after $i = 20$ iterations [14] for any filter. In comparison, the popular estimation algorithm by Daubechies [4], based on Fourier methods, turns out to be suboptimal and computationally expensive. Also, Daubechies and Lagarias [6] recently proposed a sophisticated method for estimating Hölder regularity, which is easily recovered by rewriting our main algorithm in matrix form [14], but is only manageable for very short filters.

The main algorithm described above can also be used to derive an optimal *Sobolev* regularity estimate, which gives, as seen in Section III-C, suboptimal lower and upper bounds for Hölder regularity. The resulting algorithm simplifies to the search of the spectral radius of one matrix [14], for which efficient programs exist.

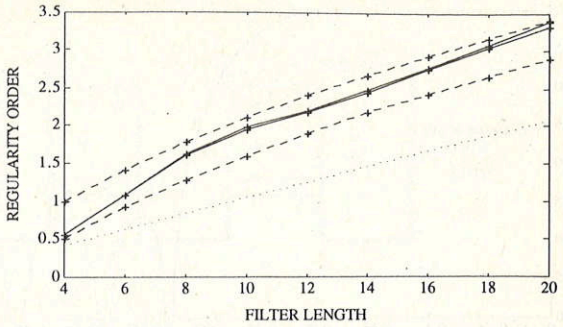

Fig. 4. Regularity estimates for Daubechies filters of length $L = 2, 4, \cdots, 20$. Sobolev lower and upper bounds (dashed). Hölder lower and upper bound (solid). Volkmer asymptote [19] (dotted).

Finally, a sharp upper bound for regularity can be derived from our main estimation algorithm [14], and turns out to be particularly fast [13], [14] (again, only the computation of the spectral radius of one matrix is required). The results are very close to be optimal, as illustrated in Fig. 4. In contrast, the Sobolev upper bound is much weaker. Clearly, the combination of our algorithms for estimating Hölder regularity orders gives lower and upper bounds which are sufficiently sharp to be used in practice.

IV. Regular Filter Design

Today, there is a relatively small number of families of regular "wavelet" filters available in the literature [4], [5], and a number of compressions schemes were designed using *ad hoc* filters. This is a serious limitation since properties like frequency selectivity and regularity are interrelated inside one family of wavelet filters; hence, it is impossible to explain some coding performance as a consequence to one property and not to the other. This limitation can be easily overcome: as an example, we propose a simple filter design procedure. We restrict to real-valued and paraunitary filters (i.e., compactly supported orthonormal wavelets) for simplicity: only one filter has to be designed, since analysis and synthesis filters are equal within time-reversal.

The simplest method for designing paraunitary filters [17] is in two steps. First, the zero-phase product filter $P(z) = H(z)H(z^{-1})$ is designed under magnitude specifications on $P(e^{j\omega}) = |H(e^{j\omega})|^2$. Second, the phase of $H(z)$ is determined by selecting zeroes of $H(z)$ among those of $P(z)$. Our aim is to balance regularity and frequency selectivity in the first step. Owing to perfect reconstruction, $P(e^{j\omega})$ can be put in the form [17]

$$P(e^{j\omega}) = 1 + \sum_{k=0}^{L/2-1} b_k \cos(2k + 1)\omega. \tag{8}$$

where L is the length of $H(z)$. The design specifications can be restricted to the half-band $(0, \pi/2)$ because of the symmetry of (8) and $\omega = \pi/2$. The frequency specifications for transition band $(\omega_p, \pi - \omega_p)$ take the form of inequality constraints on the b_k's, when ω is set to regularly spaced values in the interval $(0, \omega_p)$. Regularity is imposed by requiring K zeroes at $z = -1$ in $H(z)$. It is easy to show that this can be written as K equality constraints on the b_k's. Our design procedure minimizes the pass-band tolerance δ under these linear constraints for the $L/2 + 1$ unknowns $\{b_k\}$ and δ. This can be easily done using linear programming techniques. It is also possible to rewrite the problem in order to improve the efficiency of the design procedure, using a modified Remez exchange algorithm: This will be presented in a forthcoming paper [15].

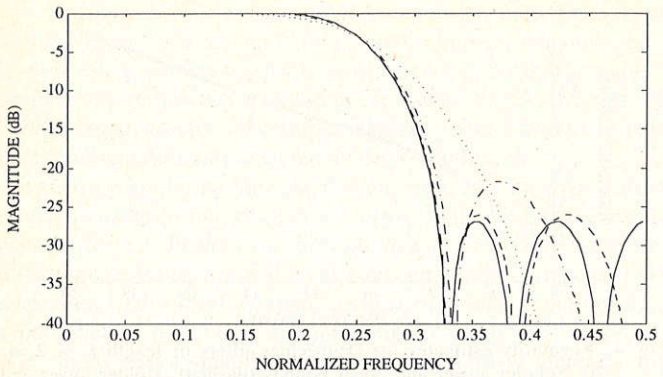

Fig. 5. Magnitude responses of the 12-tap filters designed in Section IV, for normalized transition bandwidth 0.14, and $K = 0$ (solid), 2 (dashed), 4 (dash–dotted), and 6 (dotted).

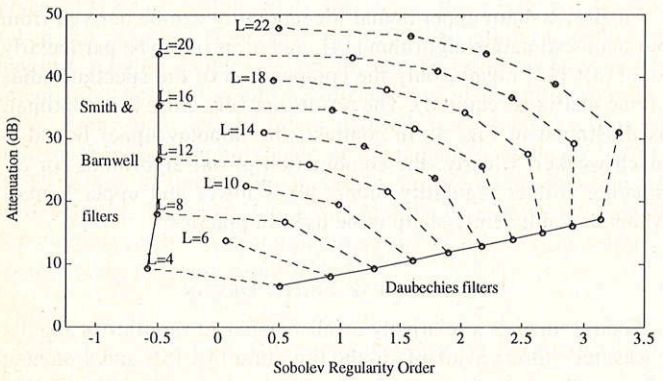

Fig. 6. Attenuation (in dB) versus Sobolev regularity for the families of filters obtained with normalized transition bandwidth set to 0.14. At one extreme, one recovers Daubechies filters, while at the other, one recovers Smith-Barnwell filters. Regularity is here quantified using Sobolev estimates, because it is natural to relate magnitude specifications, like stop-band attenuation, to Sobolev regularity. (However, once the phase of $H(z)$ is determined, regularity is best quantified using optimal Hölder regularity estimates.)

The obtained filters have maximum frequency selectivity (stop-band attenuation) for a given regularity order. An example for length $L = 12$ is given in Fig. 5. As K increases, stop-band attenuation is weaker, but flatness about $\omega = \pi$ and regularity become higher. We also observed that if $H(z)$ is designed so as to have K zeroes at $z = -1$, then it automatically has $K + 1$ such zeroes whenever the degree of freedom, $L/2 - K$, is odd. Accordingly, only the responses corresponding to the even values of K are plotted in Fig. 5.

A closer look at the design procedure reveals two extremal situations. When no flatness condition is imposed ($K = 0$), it reduces to the one proposed by Smith and Barnwell [17], which generally yields nonregular filters. When $K = L/2$, there remains no degree of freedom in the design algorithm and one recovers Daubechies filters, for which closed-form expressions exist [4], [9]. Fig. 6 shows the resulting values of attenuation and regularity order for different filter lengths and number of zeroes at $z = -1$. We have obtained a large number of filters for which attenuation and regularity can be chosen independently with a good coverage, allowing a soft transition between Daubechies and Smith–Barnwell filters. Note that while Daubechies wavelets are poorly selective, selectivity is greatly improved by relaxing a few zeroes at the Nyquist frequency, resulting in a small loss of regularity. Regularity is even improved for lengths $L \leq 22$: In this case, it is not an increasing function of K, and "maximally flat" Daubechies filters are not maximally regular.

For a given magnitude response, i.e., for a given solution $P(z)$, there are $2^{\lfloor L/4 \rfloor - 1}$ different filter solutions $H(z)$ corresponding to different phases [4]. In the rest of this note, results are given using solutions that are closest to linear phase [4], [17], i.e., whose group delay deviation is the pass-band is smallest (about 1 to 2.5 samples for $L \leq 16$). Another choice would have been to minimize the r.m.s. duration of the time-domain responses [7], but this nearly amounts to the same. In general, phase can also be chosen quite independently of the other parameters, although there is always a limited choice.

V. Image Coding Application

In this section, we measure the effects of regularity in a simple wavelet-based still image compression scheme, using separable orthonormal filters designed in the preceding section. Note, that the results shown here are only valid under these assumptions. In particular, although it is generally believed that the filter bank should not deviate far from orthonormality if efficient coding is needed [10], the role of regularity for nonorthonormal systems is not investigated here.

The compression scheme, depicted in Fig. 7, consists of a separable DWT on J decomposition levels, and a set of possible quantizers Q_i for each transformed subimage, corresponding to different bit rates (from 0 to 8 bits per pixel (bpp) with step size 0.2 bpp). The coder performance was evaluated by three different parameters, which can be taken as the bit rate reference R: overall quantizer bit rate, bit rate after Huffmann coding, and entropy. (The results are quite independent of the way wavelet coefficients are coded.) For simplicity, the results shown in this note were obtained with scalar quantization; the results obtained with lattice vector quantization in 4 or 16 dimensions [1] are similar except for the bit rate range.

In order to provide a fair comparison of compression results for different filters, we have used an automatic optimization procedure, similar to the one described in [12], which selects the best set of quantizers for each subimage and the best number of decomposition levels J which minimizes the overall distortion D at the reconstruction (measured by an m.s.e. criterion) for a given rate budget R_b, i.e., under the constraint $R \leq R_b$. This problem is solved by an unconstrained optimization procedure, using a Lagrangian cost function; it takes the form of a nested algorithm is explained in [12]. This algorithm is greatly simplified by the use of orthonormal filters, which make both rate and distortion additive over one step of decomposition.

It is questionable, in general, but distortion created by quantization can be correctly measured by the usual quadratic error D. However, in our approach, the quantizers are optimally adapted, according to the quadratic criterion, to the image and filters. In this case, we observed experimentally that visual quality on the screen fairly agrees with the SNR curves. In addition, regularity was first introduced for subjective reasons, and it is important to check whether objective distortion measures could also justify its usefulness.

Fig. 8 shows a typical rate/distortion curve for the family of 8-tap and 12-tap filters designed as shown in Section IV, and the 256×256 LENA and 576×720 BARBARA images. Clearly, trading regularity for selectivity in the filters affects the overall Peak SNR of the reconstructed image, for a wide range of bit rates: More regular (hence, less frequency selective) filters are best for a given number of taps. This was observed on various images, using var-

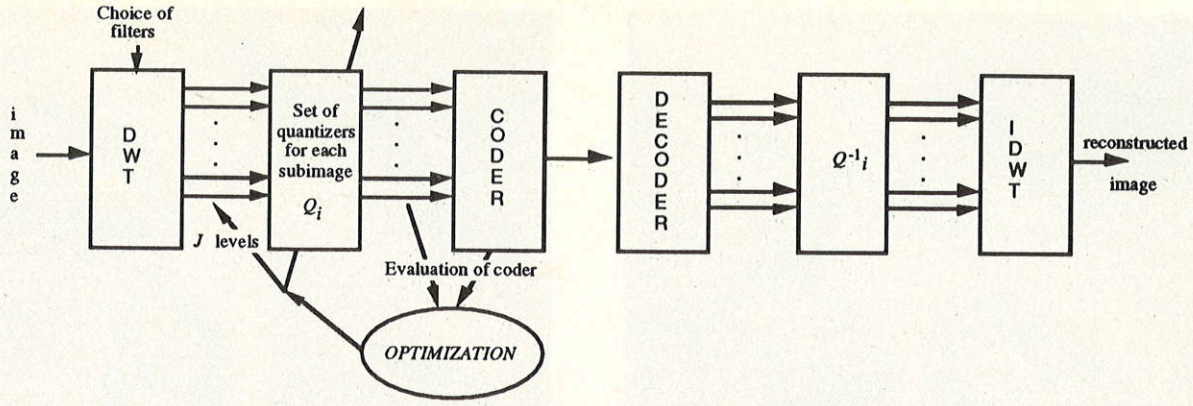

Fig. 7. Image compression scheme with rate/distortion optimization.

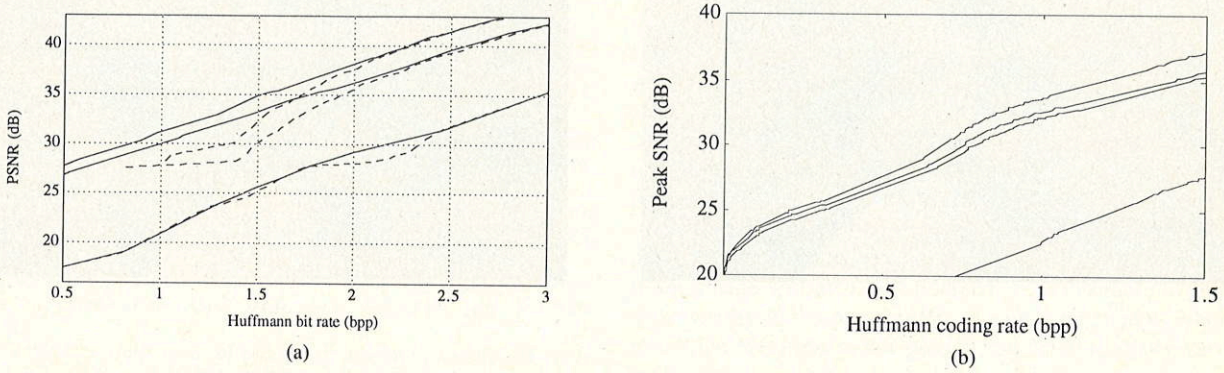

Fig. 8. Peak SNR versus Huffmann bit rate. (a) 8-tap filters with $K = 0$, 2, and 4. Optimization was made using bit rate after Huffmann coding (solid) and entropy (dashed) as coding criteria, for the LENA image. Performance is improving as K increases. (b) 12-tap filters wtih $K = 0$, 2, 4, and 6. Optimization was made using bit rate after Huffmann coding for the BARBARA image.

ious coding criteria, and relying either on the PSNR or on the visual quality of the reconstructed image (see Fig. 9). This agrees the remark made by Kronander [10] that, surprisingly, "good" selectivity in frequency is not essential for coding performance, at least in the present framework of still image compression.

At detailed look at Figs. 8 and 9 reveals two categories of filters:

1) Those whose low-pass frequency response does not vanish at the Nyquist frequency ($K = 0$), resulting in very selective, but nonregular filters. The obtained PSNR curve lies below the ones corresponding to regular filters by about 5 dB. Patterns similar to blocking artefacts are clearly visible on the image, even for strongly attenuated filters (e.g., 40 dB attenuation [17]) and was also observed by Kronander [10]. Besides the "regularity" interpretation of Section II, this can be explained as follows [10]: Since $H(z = -1) \neq 0$, a part of the DC component of the signal passes through the band-pass branches of the filter bank, which are coarsely quantized or possibly deleted, leading to strong artifacts.

2) Regular filters ($K \geq 1$): Their performance all stand within only 1 to 2 dB difference for the same number of taps. It was nevertheless observed that the visual quality of the reconstructed image increases slightly as K increases (i.e., as frequency selectivity decreases). Here, the filters giving the best performance are Daubechies filters.

To compare the performance for different filter lengths, consider Fig. 10. The coding performance globally increases with filter length, which also increases regularity. However, an asymptote is quickly attained: Above $L = 10$ or 12, for which the regularity order does not exceed 2, performance does not improve much. We also observed that the effect of phase on coding performance is almost unnoticeable for a fixed frequency response on the rate/distortion curves (less than 1 dB difference). This phenomenon should be confirmed using nonorthonormal, linear phase filters.

VI. Conclusion and Further Work

The discrete-time approach of regularity described here is efficient (optimal results are obtained) and inclusive (earlier estimates are recovered). It leads to sharp Hölder regularity estimates given by practical algorithms, which can be used as tools for quantifying precisely the effect of regularity in practical systems.

Examples of filter design and image coding application were given. Results obtained for a simple compression scheme using various coding criteria, optimized rate/distortion, and a number of paraunitary filters with balanced regularity and frequency selectivity, show that regularity may be relevant for still image compression, at least for short filters ($L \leq 10$), for which the regularity order is relatively small. Using more regular filters is probably useless, as the compression performance is not improving for longer filters. This results are valid only under the assumptions given in this paper and should be confirmed in the future using non-separable and nonorthonormal filters.

As further theoretical investigations, the discrete-time approach to regularity can be applied with success to larger problems, such as local regularity [6], nonseparable filter banks and filter banks with rational sampling changes.

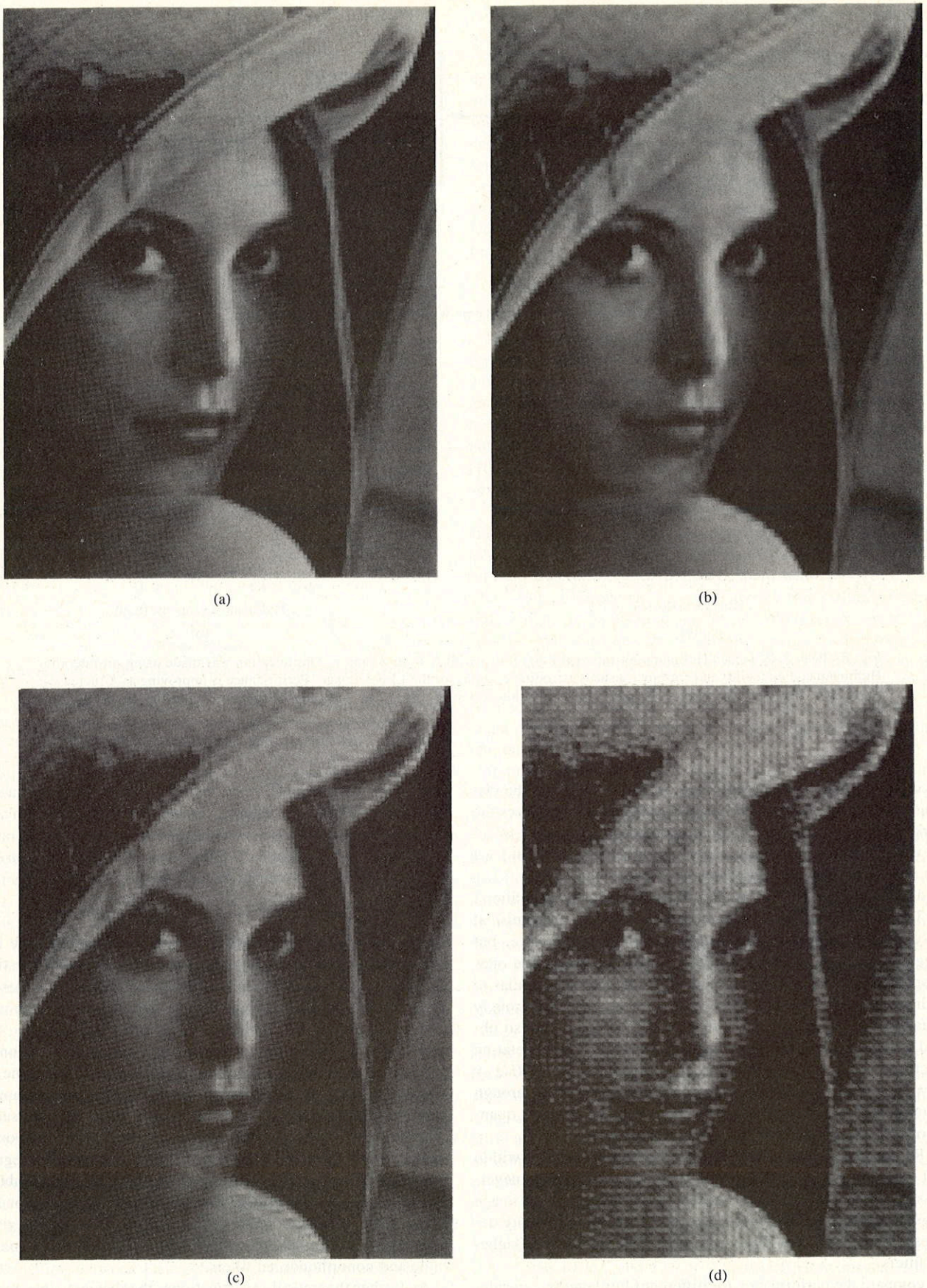

Fig. 9. Zoom on LENA's face (a) original, and compressed at 0.92 bpp using 8-tap filters [see Fig. 8(a)]. (b) $K = 4$ (PSNR = 32.7). (c) $K = 2$ (PSNR = 31.4). (d) $K = 0$ (PSNR = 22.1).

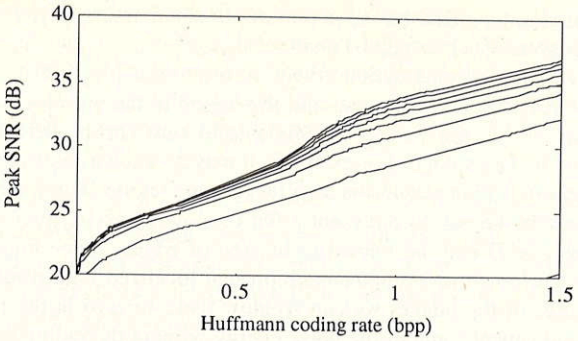

Fig. 10. Peak SNR versus Huffmann bit rate, for different lengths $L = 4, 6, \cdots, 18$ and $K = L/2 - 2$. The initial image was BARBARA.

References

[1] M. Antonini, M. Barlaud, and P. Mathieu, "Image coding using lattice vector quantization of wavelet coefficients," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Toronto, Canada, 1991, pp. 2273–2276.

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

[3] R. E. Crochiere, S. A. Weber, and J. L. Flanagan, "Digital coding of speech in subbands," *Bell Syst. Tech. J.*, vol. 55, pp. 1069–1085, 1991.

[4] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. XLI, no. 7, pp. 909–996, 1988.

[5] ——, "Orthonormal bases of compactly supported wavelets II. Variations on a theme," *SIAM J. Math. Anal.*, 1993, to appear.

[6] I. Daubechies and J. C. Lagarias, "Two-scale difference equations II.," *SIAM J. Math. Anal.*, vol. 23, pp. 1031–1079, July 1992.

[7] C. Dorize and L. F. Villemoes, "Optimizing time-frequency resolution of orthonormal wavelets," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Toronto, Ontario, May 1991, pp. 2029–2032.

[8] N. Dyn, "Subdivision schemes in CADG," in W. A. Light, Ed., *Advances in Numerical Analysis. II. Wavelets, Subdivision Algorithms and Radial Functions.* London: Oxford, 1991, pp. 36–104.

[9] O. Hermann, "On the approximation theorem in nonrecursive digital filter design," *IEEE Trans. Circuit Theory*, vol. 18, pp. 411–413, May 1971.

[10] T. Kronander, Some aspects of perception based image coding, Ph.D. dissertation, Linköping Univ., Sweden, 1989.

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

[12] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Processing*, vol. 2, pp. 160–175, Apr. 1992.

[13] O. Rioul, "A simple, optimal regularity estimate for wavelets," in *Proc. EUSIPCO*, Brussels, Belgium, vol. II, Sept. 1992, pp. 937–940.

[14] ——, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, pp. 1544–1576, Nov. 1992.

[15] O. Rioul and P. Duhamel, "A Remez exchange algorithm for orthonormal wavelets," *IEEE Trans. Circuits Syst.*, 1993, submitted.

[16] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, Oct. 1991.

[17] M. J. T. Smith and T. P. Barnwell, "Exact reconstruction techniques for tree-structured subband coders," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 434–441, June 1986.

[18] M. Vetterli, J. Kovačević, and D. L. Gall, "Perfect reconstruction filter banks for HDTV representation and coding," *Image Commun.*, vol. 2, pp. 349–364, Oct. 1990.

[19] H. Volkmer, "On the regularity of wavelets," *IEEE Trans. Informat. Theory*, vol. 38, pp. 872–876, Mar. 1992.

Image Reconstruction from Projections under Wavelet Constraints

Berkman Sahiner and Andrew E. Yagle

***Abstract*—First, we discuss how the wavelet transform can be used to perform spatially-varying filtering of an image, suppressing noise locally in smooth regions of the image, and we discuss detection of such regions in a noise-corrupted image. Second, we show how to compute the minimum mean-square estimate of an image given: 1) noisy projections of the image; 2) statistics of additive noise in the projections; and 3) constraints on wavelet coefficients of the image. Examples illustrate the resulting procedure.**

I. Introduction

In many problems arising in fields such as medical imaging, nondestructive testing, radio astronomy, and geophysics [1], one needs to reconstruct a two-dimensional object or image from its projections, which amounts to computing the inverse Radon transform. A problem with the inverse Radon transform is that the ramp filter amplifies the high-frequency components of both the noise and the data. Since noise usually dominates at high frequencies, it is common practice to use a low-pass filter in conjunction with the ramp filter to improve the signal-to-noise ratio (SNR). However, the SNR improvement obtained by using a low-pass filter comes at the expense of degraded image resolution, since high-resolution features in the image will also be smoothed. It is desirable to reduce the noise energy in the reconstructed image over regions where high-resolution features are not present, by using spatially-varying filtering.

In this note, we use wavelets to perform this desired localized low-pass filtering. We show how thresholding can be used to determine the regions in the wavelet domain where wavelet coefficients may be set to zero, effecting spatially-varying filtering, and provide a statistical justification for it. Alternatively, *a priori* information about the image can be used to identify such regions. We then use these zero wavelet coefficients as constraints, and compute the minimum mean-squared error image which satisfies these constraints.

II. The Radon and Wavelet Transforms

A. Image Reconstruction From Projections

The inverse Radon transform problem is to reconstruct an image $\mu(x, y)$ from its projections $p(r, \theta)$ where

$$p(r, \theta) = \mathcal{R}\{\mu(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu(x, y)\, \delta(r - x \cos\theta - y \sin\theta)\, dx\, dy \quad (1)$$

is the Radon transform of $\mu(x, y)$. A common procedure for obtaining $\mu(x, y)$ from $p(r, \theta)$ is filtered backprojection (FBP), in which the projections are first filtered to yield $s(r, \theta) = p(r, \theta) *$

Manuscript received August 28, 1992; revised May 19, 1993. The Guest Editor coordinating the review of this paper and approving it for publication was Prof. Martin Vetterli. This work was supported in part by the Office of Naval Research under Grant N00014-90-J-1897.

The authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122.

IEEE Log Number 9212174.