On the Asymptotic Performance of Hierarchical Transforms

Ricardo L. de Queiroz and Henrique S. Malvar

Abstract—**We derive explicit formulas for the coding gain of hierarchical transforms for a given number of stages, and the asymptotic gain as this number goes to infinity. We confirm the intuitive result that hierarchical transforms are not asymptotically optimal, that is, their coding gains do not approach the inverse of the spectral flatness measure as the number of stages goes to infinity. Examples comparing the limits for hierarchical transforms and *M*-band parallel systems for AR (1) signals are presented.**

I. Introduction

Systems exploring multiple resolutions of a signal have been widely employed in image coding, such as those in subband [1], [2] and pyramid [3] coding. Recently, the concept of wavelets for subband applications has emerged, and solid relations between wavelets and filter banks have been stated [4]. A filter bank (FB) is used either for splitting the spectrum of a signal into its constituent subband signals or for reconstructing the signal from the subbands [5]. Subband systems are widely based on the cascade of two-band FB's [5], or TBFB's, whose filters have a bandwidth of $\pi/2$. Hence, the spectrum is separated through a binary-tree structure. We will discuss in this correspondence critically decimated FB's [5] and concentrate our attention on hierarchical transforms (HT's), which are defined as those where a further level of TBFB is only connected to the low-pass branch of a previous one, splitting the spectrum in octaves, as shown in Fig. 1.

For image coding, it is interesting to measure the compaction provided by the transform, i.e., how much energy is concentrated in the lower frequency bands of the transformed signal. Consider a hierarchical connection of N stages of TBFB's, with $M = 2^N$. The higher frequency band would have a width of $\pi/2$ and the lower frequency (baseband) would have a width of π/M. Due to the critical decimation of the subband signals, for every M input samples we have also M subband signal samples: one sample from subband number zero, one from subband one, two from subband two, four from subband three, and so on, up to $M/2$ samples from subband N. Assuming a stationary input signal with known spectrum, it is easy to compute the variances of each of the M subband samples. From these variances, we can measure the transform gain over PCM [1] (the G_{TC}) as

$$G_{TC} = 10 \log_{10} \left(\frac{\frac{1}{M} \sum_{i=0}^{M-1} \sigma_i^2}{\left(\prod_{i=0}^{M-1} \sigma_i^2 \right)^{1/M}} \right) \tag{1}$$

being σ_i^2 the variance of the ith subband signal for $i = 0, \cdots, M - 1$.

Manuscript received May 20, 1991; revised December 5, 1991. This work was supported in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, under Grants 600.047-90 and 300.159-90.

The authors are with the Department of Electrical Engineering, Universidade de Brasília, Brasília, DF 70910, Brazil.

IEEE Log Number 9202360.

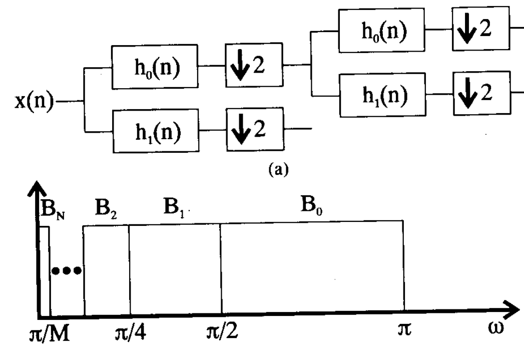

Fig. 1. Spectrum partitioning in a hierarchical transform. (a) The connection of TBFB's in the analysis section. (b) Resulting subbands.

The G_{TC} is well known [1] as a good measure of the compaction potential of a transform. In [2, ch. 4] G_{TC} was also used for measuring the performance of hierarchical systems. The purpose of this correspondence is to compare the performance of hierarchical and parallel filterbanks, in terms of G_{TC}. Although parallel systems are asymptotically optimal (maximum G_{TC}) for any input spectrum as the number of subband increases to infinity, we will show in Section II that the same does not hold for hierarchical systems when the number of stages goes to infinity. In Section III we will evaluate the performance bounds for HT's for AR (1) input signals, and we will also compare the computational complexities of HT's and parallel filter banks. Our results are valid for one-dimensional signals; the extension for two-dimensional signals is quite simple if we assume separate signal models and row-column processing [1].

II. Ideal Filters and Limiting Behavior

An M-band FB (MBFB) splits the signal into M bands of width π/M. The M filters that would maximize the G_{TC} must have infinite length and ideal frequency response [2, ch.1]. Therefore, the variance for each band would be exactly the integral of the power spectral density (PSD) $S_{xx}(e^{j\omega})$ of the input signal $x(n)$ along that band. This is consistent with the fact that a FB with filters of length $L = \lambda M$ have selectivity and coding gain increasing with the overlapping factor λ, as we go from a block transform ($\lambda = 1$, as in the DCT and KLT) to lapped transforms ($\lambda > 1$) [9].

It is well known that for $M \to \infty$ the G_{TC} of MBFB's tends to the inverse of the spectral flatness measure [1]

$$G_{TC} \to 10 \log_{10} \frac{\frac{1}{\pi} \int_0^{\omega} S_{xx}(e^{j\omega})\, d\omega}{\exp\left[\frac{1}{\pi} \int_0^{\omega} \ln S_{xx}(e^{j\omega})\, d\omega \right]} \tag{2}$$

which is a bound for the maximum coding gain of any coder [1]. Thus, an MBFB-based signal coder is asymptotically optimal. Intuitively, we see that this happens because the spectral resolution gets better as M increases, approaching in the limit a continuous resolution, where the bandwidths of all subbands are arbitrarily small. With lapped transforms the G_{TC} gets closer to the bound in (2) than with block transforms, because of the better bandpass filtering characteristics of the former [9].

As expected, the same relation does not apply to HT's. An intuitive reason for this is based on the fact that, no matter how many stages of TBFB's are cascaded, the higher frequency bands will have the same width. As the number of stages N increases, only the baseband is further subdivided. Therefore, there is not a trend towards a continuous resolution with $N \to \infty$.

Recalling that $M = 2^N$ for HT's, and in order to maintain the same baseband width for both MBFB's and N-stage HT's, we will now evaluate the upper bound for HTs with ideal filters. Let us define the integral of the PSD as

$$F(a, b) = \frac{1}{\pi} \int_a^b S_{xx}(e^{j\omega})\, d\omega. \tag{3}$$

The filters in the TBFB are assumed to have a gain of $\sqrt{2}$ in the passband, in order to keep orthonormality [9]. The subbands are ordered from B_0 (the higher frequency band, $\pi/2 \le \omega \le \pi$) to B_N (the baseband, $0 \le \omega < \pi/M$). When the ith subband is generated by an ideal filter with bandwidth $[\pi/2r, \pi/r]$, where $r \equiv 2^i$, the variance of the ith subband is given by

$$\eta_i = \frac{1}{\pi} \int_{\pi/2r}^{\pi/r} 2^{i+1} S_{xx}(e^{j\omega})\, d\omega = 2rF(\pi/2r, \pi/r). \tag{4}$$

Accordingly, the baseband variance is

$$\zeta_N = \frac{1}{\pi} \int_0^{\pi/M} 2^N S_{xx}(e^{j\omega})\, d\omega = MF(0, \pi/M). \tag{5}$$

The ith subband contains $M/2r$ basis functions, and so $M/2r$ variances for each subband in (1) are equal. The numerator and demoninator in (1) are given, respectively, by

$$\frac{1}{M} \sum_{k=0}^{M-1} \sigma_k^2 = \frac{1}{M}\left(\zeta_N + \sum_{i=0}^{N-1} \frac{M}{2r}\eta_i\right) = F(0,\pi) = 1 \tag{6}$$

and

$$\left(\prod_{k=0}^{M-1} \sigma_k^2\right)^{1/M} = \left(\prod_{i=0}^{N-1} \eta_i^{M/2r}\right)^{1/M} \zeta_N^{1/M} = \zeta_N^{1/M}(\eta_0^{1/2}\eta_1^{1/4} \cdots \eta_{N-1}^{1/M}). \tag{7}$$

Thus, the G_{TC} for general N can be found recursively as

$$G_{TC}(N+1) = G_{TC}(N) + \frac{5}{2^N}\log_{10}\left(\frac{\zeta_N^2}{\zeta_{N+1}\eta_N}\right) \quad \text{dB} \tag{8}$$

$$G_{TC}(1) = -5 \log_{10}(\eta_0 \zeta_1) \quad \text{dB}. \tag{9}$$

Increasing N, G_{TC} converges to a theoretical upper bound that defines the coding performance for HT's. As we will see in the next section, this upper bound is lower than the one defined in (2).

III. Performance and Complexity Bounds

Consider an AR (1) signal with intersample correlation ρ, which is a good model for images [1]. For this model,

$$F(a, b) = \frac{2}{\pi} \arctan\left[\left(\frac{1+\rho}{1-\rho}\right)\tan(\omega/2)\right]\Bigg|_{\omega=a}^{\omega=b}.$$

Table I lists values of the limiting gain (G_{TC}^{∞}) for different ρ, comparing with the equivalent G_{TC}^{∞} for MBFB's. For each value of the parameter, the difference between the limits for MBFB's and HT's are small, but in actual systems the difference would increase due to two reasons.

TABLE I
HT and M-Band FB G_{TC} (Decibels) Upper Bounds, Assuming an AR (1) Signal, for Several Values of ρ

ρ	G_{TC}^{∞} Hierarchical	G_{TC}^{∞} M-band
0.800	4.2912	4.4307
0.900	7.0405	7.2125
0.950	9.9248	10.1100
0.970	12.0936	12.2841
0.990	16.8156	17.0115

i) As stated in the last section, MBFB's have their filter lengths increasing with M, thus approaching ideal filters as $M \to \infty$. For HT's only the low-pass branch of a TBFB is connected to another, therefore only the baseband filter length goes to infinity with M or N, with all other filter lengths remaining fixed. For example, the higher frequency band is processed by a filter which normally possesses 8 or 16 taps, and thus it will not have good selectivity.

ii) TBFB's are generally designed without taking into account any regularity criteria [4]. As $N \to \infty$ the resulting baseband filter, whose length grows with M, has a frequency response that may not converge to a continuous function, and hence providing a fractal behavior [4]. In Fig. 2 we have a plot of the magnitude frequency response of the second basis function of a cascade of Johnston's QMF-8A banks [5]. The lack of regularity shows up as undesirably high sidelobe levels, which lead to a reduction in G_{TC}.

The G_{TC} for nonideal filters can be found by rewriting (4) and (5) as

$$\alpha_i = \frac{1}{\pi} \int_0^{\pi} |H_i(e^{j\omega})|^2 S_{xx}(e^{j\omega})\, d\omega \tag{10}$$

$$\beta_N = \frac{1}{\pi} \int_0^{\pi} |H_N(e^{j\omega})|^2 S_{xx}(e^{j\omega})\, d\omega \tag{11}$$

where α_i is the variance of the ith subband, which is generated by the filter $H_i(e^{j\omega})$, and β_N is the variance of the baseband, which is generated by $H_N(e^{j\omega})$. The coding gain can still be computed by (8) and (9), with η_i replaced by α_i, and ζ_N replaced by β_N.

The filters H_i are defined as those used to select the band B_i and should be generated by a cascade of the TBFB based on the pair of filters Q_0 and Q_1 satisfying the QMF condition [5], [9] $|Q_0(e^{j\omega})|^2 + |Q_1(e^{j\omega})|^2 = 2$. Table II shows asymptotic gains (G_{TC}^{MAX}) for TBFB's based on the Haar transform, Johnston's QMF [5], and lattice CQF TBFB's [6], assuming $\rho = 0.95$.

Another important issue related to HTs is their computational complexity. Here, the complexity C will be measured in arithmetic operations (additions and multiplications) per input sample. With HT's, not only the G_{TC} is bounded, but also is C. Being C_2 the complexity for implementing a single stage, the complexity for N stages is given by $C = 2C_2(M-1)/M$, being clearly bounded as $N \to \infty$. For $N \ge 3$, we have $C_{MIN} = 1.75C_2 \le C \le 2C_2 = C_{MAX}$. Complexity bounds for the HT's based on several TBFB's are shown in Table II.

In practical cases in the image coding field, the number of stages in an HT must be limited, because the image itself is limited and, after many stages, the lower frequency band becomes approximately flat [7]. Two through four stages are commonly used [2], and more stages may not result in any advantage [7]. Suppose now that adjacent $M/2r$ samples in each band are grouped to form a

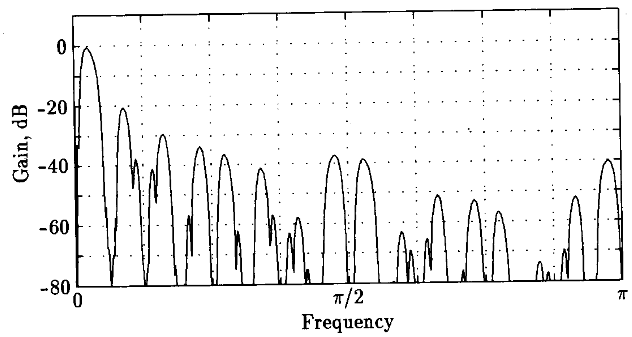

Fig. 2. Magnitude frequency response of the second basis function of a 6-stage Johnston-8A-based HT. Note the undesirable sidelobes in the resulting filter.

TABLE II
COMPLEXITY (OP/S) AND G_{TC} (DECIBEL) BOUNDS, FOR HT'S WITH $N > 2$, BASED ON DIFFERENT TWO-BAND CELLS. COMPLEXITY AND GAIN FOR SOME M-BAND FB'S

TBFB	C_{MIN}	C_{MAX}	G_{TC}^{MAX}
HAAR	3.50	4.0	8.26
QMF 8A	14.00	16.0	9.68
LAT 8A	15.75	18.0	9.73
QMF 16B	28.00	32.0	9.85
LAT 16B	29.75	34.0	9.86
ELT, $M = 16$		C	G_{TC}
$\lambda = 2$		12.0	9.84
$\lambda = 4$		14.0	9.90

M-sample transformed block. If the samples are not quantized with a temporal adaptation step shorter than $2r$ samples, all coefficients in the same band in the same block would have the same number of assigned bits. This is a common case, but the advantage of HT's possessing finer resolution in higher frequencies is diminished. In this case, the HT would present near the same spatial resolution as an MBFB (the block size), with the latter presenting better frequency resolution. Therefore, the G_{TC} would play the same role to measure the performance of both approaches. Tradeoffs between gain and complexity for MBFB's are presented in Table II, with the ELT [8], [9] with $M = 16$ as example of an MBFB.

IV. CONCLUSION

Hierarchical transforms are asymptotically suboptimal in signal coding, assuming a wide sense stationary input signal. However, this bound is not much distant from the maximum gain of M-band transforms and predictive systems for AR (1) signals, as demonstrated using the formulas presented in this correspondence. Reasons why HT's cannot reach their bounds, even increasing the number of stages to infinity, were also presented.

HT's are bounded in G_{TC} and complexity. For a straight comparison with regular transforms, it is necessary to assume that the coder is at least block-to-block adaptive, which is a common practice. Comparing both HT's and MBFB's like the ELT, in a complexity/compaction criteria, it is usually possible to find an MBFB with higher gain and less complexity than those of a hierarchical transform based on a particular choice for the TBFB.

REFERENCES

[1] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice-Hall, 1984.
[2] J. W. Woods, Ed., *Subband Coding of Images.* Boston, MA: Kluwer, 1991.
[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, Apr. 1983.
[4] M. Vetterli and C. Herley, *Wavelets and Filter Banks: Theory and Design*, Tech. Rep. 206/90/36, CTR Columbia Univ., 1990.
[5] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1983.
[6] P. P. Vaidyanathan and P. Hoang, "Lattice structures for optimal design and robust implementation of two-channel PR-QMF banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 81–94, Jan. 1988.
[7] P. H. Westerink, "Subband coding of images," Ph.D. dissertation, Delft Technische Universiteit, The Netherlands, Oct. 1989.
[8] H. S. Malvar, "Extended lapped transforms: Fast algorithms and applications," in *Proc. Int. Conf. Acoust., Speech, Signal Processing* (Toronto, Canada), May 1991, pp. 1797–1800.
[9] H. S. Malvar, *Signal Processing with Lapped Transforms.* Norwood, MA: Artech House, 1991.

On Signal Recovery with Adaptive Order Statistic Filters

Geoffrey A. Williamson and Peter M. Clarkson

***Abstract*—In this correspondence, adaptive order statistic filters are used to estimate a constant amplitude signal embedded in noise having unknown statistics. Iterative algorithms are proposed which adapt the order statistic filter to minimize the mean-square estimation error, both with and without an unbiasedness constraint. For each case, the algorithm employed is shown to achieve convergence in the mean to the optimal filter. Properties of the convergence rates are discussed, and conditions for convergence in mean square are noted.**

Manuscript received May 6, 1991; revised February 20, 1992. This work was supported by the National Science Foundation under Grants MIP-8911676 and MIP-9102620.
The authors are with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616.
IEEE Log Number 9202363.

1053-587X/92$03.00 © 1992 IEEE