On-signal decomposition techniques

Ali N. Akansu, MEMBER SPIE
Yipeng Liu
New Jersey Institute of Technology
Department of Electrical and Computer Engineering
Center for Communications and Signal Processing Research
University Heights, Newark, New Jersey 07102

Abstract. Well-known block transforms and perfect reconstruction orthonormal filter banks are evaluated based on their frequency behavior and energy compaction. The filter banks outperform the block transforms for the signal sources considered. Although the latter are simpler to implement and already the choice of the existing video coding standards, filter banks with simple algorithms may well become the signal decomposition technique for the next generation video codecs, which require a multiresolution signal representation.

Subject terms: visual communications; signal decomposition; block transform; lapped orthogonal transform; filter bank; wavelet transform; perfect reconstruction; energy compaction.

Optical Engineering 30(7), 912–920 (July 1991).

CONTENTS

1. INTRODUCTION

Signal decomposition techniques are useful tools for many signal processing problems. The basic idea is to represent the signal by a superposition of basis functions of an orthonormal transformation. The transformations span from conventional block transforms to ideal subband filter banks. Block transforms limit the duration of the basis functions to the number of bands in the filter bank. This filtering operation corresponds to a block transform operation. The basis functions of a block transform satisfy the orthogonality or orthonormality conditions. The goodness criterion for a block transform is the interband or intercoefficient decorrelation power, which also implies its energy compaction capability. It is well-known that the signal-dependent Karhunen-Loève transform (KLT) is the optimum solution for this type of signal decomposition.

If the duration of the basis functions or filters is increased, a number of possible signal decomposition tools are available. First of all, if the duration of the basis functions is doubled compared to that of block transforms, the lapped orthogonal transform (LOT) is obtained.[1,2] Obviously, the expense of the forward and inverse transform operations increases over the block transforms. If the duration of the basis functions is relaxed even further, many different filter bank solutions are obtained.[3–8] Finally, when the duration of basis functions or filters goes to infinity, the ideal filter bank that is the globally optimum solution is reached. Optimality refers here to perfect interband decorrelation and alias-free band split simultaneously for multirate signal processing.

Invited paper VC-101 received Dec. 18, 1990; revised manuscript received Mar. 12, 1991; accepted for publication Mar. 15, 1991.

Block transforms have fixed time-frequency resolution; therefore, they have not been considered historically as a multiresolution signal decomposition tool. Section 2 interprets the block transforms as filter banks, shows the frequency selectivity of their basis functions or bandpass filters, and attempts to extend the block transform concept for multiresolution time-frequency signal analysis.[9] Section 3 revisits the perfect reconstruction quadrature mirror filter (PR-QMF) very briefly and defines the regular (half-band or binary) and irregular tree structures. The links of these trees to the time-frequency signal analysis are also given in this section. The performance criterion, gain of transform coding over pulse code modulation G_{TC}, is given in Sec. 4 along with the energy compaction performance of several different scenarios for comparison. The performance measure G_{TC} is also extended for irregular tree structures in this section. Section 5 proposes a new signal-dependent tree structuring algorithm based on energy compaction. The validity of the proposed algorithm is examined on the test images, and it is shown that proper irregular trees are very useful for signal decomposition.

2. BLOCK TRANSFORM FILTER BANKS

In one dimension, the set of N signal samples

$$\mathbf{x}^T = [x(0), x(1),\ldots,x(N-1)] \tag{1}$$

is transformed via matrix $\mathbf{A}$ into coefficient vector $\boldsymbol{\theta}$ as

$$\boldsymbol{\theta} = \mathbf{A}\boldsymbol{x} , \tag{2}$$

where

$$\mathbf{A}^{-1} = \mathbf{A}^T . \tag{3}$$

We define the autocorrelation sequence of the i'th basis function $a_i(j)$, which is also the i'th row of the transform matrix $\mathbf{A}$, as

$$\begin{aligned} g_i(k) &= a_i(k) * a_i(-k) \\ &= \sum_{j=0}^{N-k-1} a_i(j)a_i(j+k) \\ &\text{for } \begin{matrix} i = 0,1,\ldots,N-1 \\ k = 0,\pm 1,\pm 2,\ldots,\pm(N-1) \end{matrix} . \end{aligned} \tag{4}$$

When the input autocorrelation function $R_{xx}(m)$ is known, the corresponding power spectral density is $P_{xx}(\omega)$, and the output power spectral density for the i'th bandpass filter is found to be

$$P_{yy}^i(\omega) = |G_i(e^{j\omega})|^2 P_{xx}(\omega) , \tag{5}$$

where

$$|G_i(e^{j\omega})|^2 = a_i(e^{j\omega})a_i(e^{-j\omega}) \tag{6}$$

and $a_i(e^{j\omega})$ is the Fourier transform of $a_i(j)$.

Figure 1 gives $a_i(j)$ and $|G_i(e^{j\omega})|$ of the discrete cosine transform (DCT) for $N = 8$. These functions show the frequency selectivity of the basis functions or filters in the DCT block filter bank. One can easily obtain the output energy of the i'th band or coefficient as

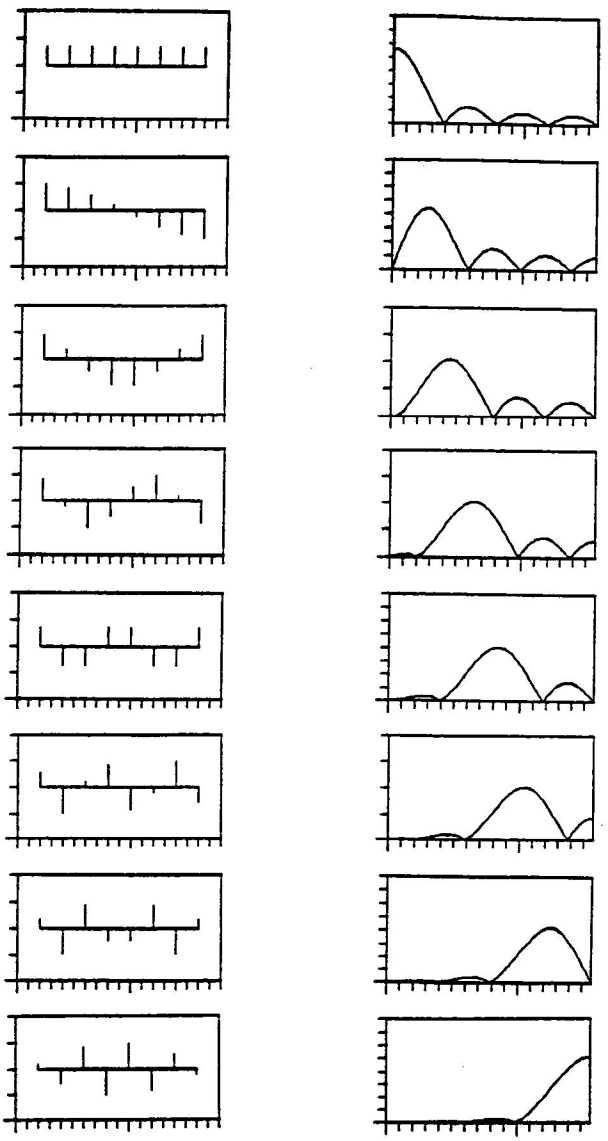

Fig. 1. The basis functions of DCT and their magnitude functions in frequency, $0 \leq \omega \leq \pi$, for $N = 8$.

$$\sigma_i^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G_i(e^{j\omega})|^2 P_{xx}(\omega)\, d\omega \qquad i = 0,1,\ldots,N-1 . \tag{7}$$

The $\{\sigma_i^2\}$ are used later to calculate the performance measure.

Block transform bases are not designed based on their frequency selectivity. Rather they satisfy orthogonality conditions, and the optimum transform KLT decorrelates the bands or coefficients perfectly. The interaction of their frequency bands, aliasing, is quite high, and the Nyquist condition for decimation by N is not satisfied properly. Therefore, the coefficients in block transforms are not commonly used in practice as the subbands of the input signal because of their poor frequency performance. Additionally, the block transforms have a fixed time-frequency resolution. On the other hand, the transform or filtering operation is very efficient—N band signal decomposition is realized at once.

Multiresolution has become a desired feature for many signal decomposition techniques.[6,10] It provides the tools to localize or "zoom in" on the significant parts of the signal in time as

well as in frequency. This implies that longer time functions mean better resolution in frequency. This duality feature helps to interpret the time- and frequency-domain characteristics of a signal simultaneously.

Block transforms or filter banks can be extended for multiresolution signal decomposition purposes.[9] The original signal is first block transformed or filtered into N_1 bands. Each band again may be block transformed or filtered into N_{ij} bands where $i = 1,2,\ldots,N_1$ and j is the size of the transform performed on the i'th band of the first level. This process may be repeated through many levels until a sufficient set of features of the signal is obtained from the decomposition operation. Figure 2 shows a typical multiresolution block transform or filter bank structure. Obviously, the aliasing problem should be kept in mind. It is seen from Fig. 2 that the possibilities for transform decomposition trees are many.

LOTs or filter banks are the special case between the block filter banks and conventional subband filter banks. The bandpass filters in this bank have twice the duration of the conventional transform filters.[1,2]

3. HALF-BAND FREQUENCY SPLIT: PR-QMF

Half-band filter banks are quite popular for multiresolution signal decomposition. The basic idea here is to split the signal into its low- and high-frequency components. The duration of the analysis and the synthesis filters is not restricted to two taps as would be the case in a 2 × 2 block transform. Therefore, the frequency selectivity of the two basis functions is expected to be better than the 2 × 2 block transforms. There are many two-band PR-QMF banks proposed in the literature.[3–5,7,8] The perfect reconstruction requirements of the analysis/synthesis filter banks will be revisited briefly.

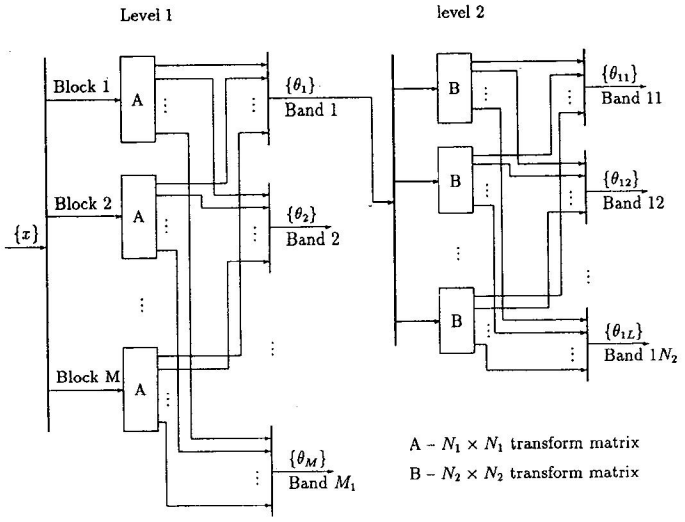

Fig. 2. A typical multistage or resolution block transform structure.

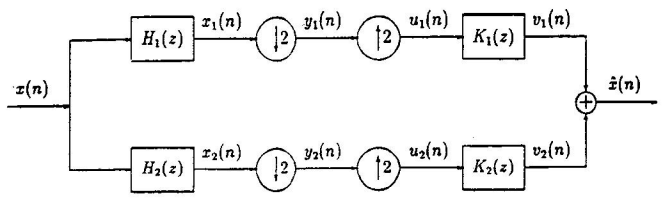

Fig. 3. Two-band QMF bank.

The z transform of the reconstructed signal in Fig. 3 can be written as

$$\hat{X}(z) = \frac{1}{2}[H_1(z)K_1(z) + H_2(z)K_2(z)]X(z) + \frac{1}{2}[H_1(-z)K_1(z) + H_2(-z)K_2(z)]X(-z)$$
$$= T(z)X(z) + S(z)X(-z) \; . \tag{8}$$

Perfect reconstruction requires:

$$S(z) = 0, \text{ for all } z \; , \tag{9}$$

$$T(z) = cz^{-n_0}, c \text{ is a constant} \; . \tag{10}$$

If one chooses

$$K_1(z) = -H_2(-z) \; , \qquad K_2(z) = H_1(-z) \; , \tag{11}$$

the first requirement is met, $S(z) = 0$, and aliasing is eliminated, leaving us with

$$T(z) = \frac{1}{2}[H_1(-z)H_2(z) - H_1(z)H_2(-z)] \; .$$

Next, with N odd, if one selects

$$H_2(z) = z^{-N}H_1(-z^{-1}) \; , \tag{12}$$

this choice forces

$$H_2(-z) = -K_1(z)$$

so that

$$T(z) = \frac{1}{2}z^{-N}[H_1(z)H_1(z^{-1}) + H_1(-z)H_1(-z^{-1})] \; . \tag{13}$$

Therefore, the perfect reconstruction requirement reduces to finding an $H(z) = H_1(z)$ such that

$$Q(z) = H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = cz^{-n_0}$$
$$= R(z) + R(-z) \; . \tag{14}$$

This selection implies that all four filters are causal whenever $H_1(z)$ is causal.

This half-band frequency split operation can be applied to any subband of the new level in a tree, and the process continues until the desired frequency split or resolution is achieved. If all bands of any level in the tree are split into two new half-bands, the corresponding frequency tree is called a *regular tree*, which implies equal bandwidths. If the number of levels is L, the regular tree has 2^L equal-sized frequency bands; the maximum frequency resolution is achieved in this frequency split. As will be shown later, many applications may not need a regular tree. Several of the maximum resolution frequency bands, with bandwidths $BW = \pi/2^L$, can be combined as a larger frequency band of the next higher level in the tree. These tree structures with

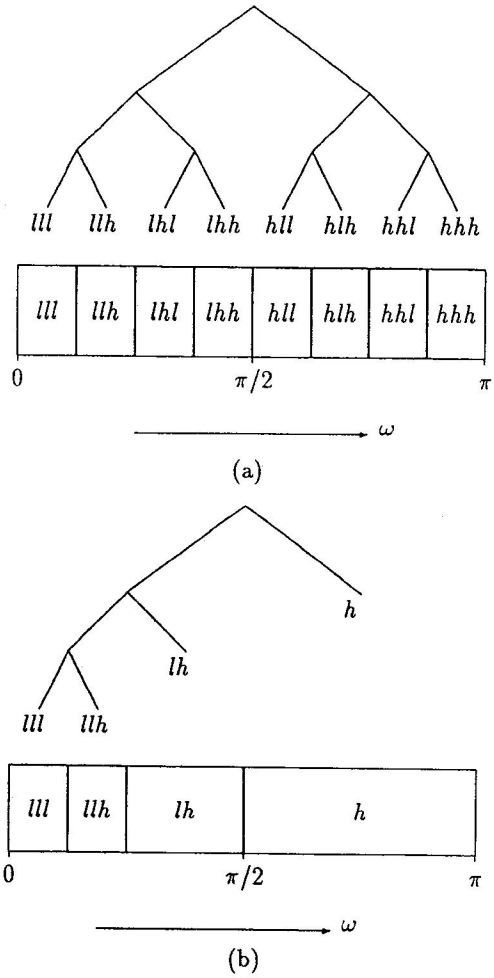

Fig. 4. (a) A regular tree structure and its frequency bands assuming ideal filters. (b) An irregular tree structure and its frequency bands assuming ideal filters.

unequal bands are called *irregular trees*. Figures 4(a) and 4(b) show examples of regular and irregular trees and their frequency characteristics assuming ideal filters for $L = 3$. The irregular trees have fewer bands than the regular trees for a fixed L, thereby reducing the number of filters. On the other hand, the bandwidths are not equal in irregular tree structures, which also implies different sized time functions for different sized frequency bands.

The computational simplicity of the block filter banks or transforms over subband filter banks is obvious. An $N = 2^L$ band subband filter bank requires an L level regular tree, whereas the N sized block filter bank realizes that operation at one level. However, as mentioned earlier, the latter does not consider the frequency characteristics of the basis functions. There are good filters with reasonable durations in time that make them an alternative to the block transforms. For irregular tree structures, the computational burden of conventional filter banks over block transforms is significantly reduced but with a payoff in a little better performance.

Recently, the wavelet transforms have become popular as a new approach for multiresolution signal decomposition.[11–13] If the orthonormality condition is imposed on finite support wavelets, this approach merges to the conventional PR-QMF signal decomposition technique. The wavelet transform brings new insights into PR-QMF design. The regularity or differentiability conditions of wavelets imposes smoothness on PR-QMFs. The compaction performance of several wavelet filters is calculated and compared with other multirate filters.

4. PERFORMANCE CRITERION

An energy compaction measure, namely, the gain of transform coding over pulse code modulation, G_{TC}, has been a common tool for comparing orthonormal transforms.[14] This tool is also valid for any orthonormal signal decomposition technique. Extensions of this measure to multilevel multiband regular trees as well as irregular tree structures are derived in this section. While there are other subjective measures in speech and image coding, this criterion is universal and objective. An N_1 band orthonormal transform implies the variance preservation condition,

$$\sigma_x^2 = \frac{1}{N_1} \sum_{k=0}^{N_1-1} \sigma_k^2 , \tag{15}$$

where σ_x^2 is the input signal variance with zero mean and $\{\sigma_k^2\}$ are the band variances. If one assumes that all the bands and the input signal have the same pdf type, the distortion ratio of PCM over transform coding at the same bit rate can be obtained easily as[14]:

$$G_{TC} = \frac{D_{PCM}}{D_{TC}} = \frac{\sigma_x^2}{\left(\prod_{k=0}^{N_1} \sigma_k^2\right)^{1/N_1}} . \tag{16}$$

If each band of the first-level decomposition tree goes through an N_2 band decomposition, G_{TC} can be extended easily for this case. The input variance is now connected to the variances of $N_1 \times N_2$ bands as

$$\sigma_x^2 = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sigma_{k_1 k_2}^2 . \tag{17}$$

An orthonormal transform ensures that the average of the analysis bands' reconstruction error after the synthesis operations is equal to the quantization error of the signal:

$$\sigma_r^2 = \sigma_q^2 = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sigma_{qk_1k_2}^2 . \tag{18}$$

The band distortions can be expressed as

$$\sigma_{qk_1k_2}^2 = \varepsilon_{k_1k_2}^2 2^{-2B_{k_1k_2}} \sigma_{k_1k_2}^2 , \tag{19}$$

where $B_{k_1k_2}$ is the average bit rate for band k_1k_2, and $\varepsilon_{k_1k_2}^2$ is the quantizer correction factor for that band. After assuming the same pdf type for all the bands, one can write

$$\varepsilon^2 = \varepsilon_{k_1k_2}^2 \qquad \begin{matrix} k_1 = 0,1,\ldots,(N_1-1) \\ k_2 = 0,1,\ldots,(N_2-1) \end{matrix} . \tag{20}$$

Hence, the average distortion is

$$\sigma_q^2 = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \varepsilon^2 2^{-2B_{k_1k_2}} \sigma_{k_1k_2}^2 . \tag{21}$$

The optimization problem is to find the bit allocations of $(N_1 \times N_2)$ bands such that the average distortion σ_q^2 is mini-

mized. The bands have equal distortion levels, and the overall average bit rate is constrained by

$$B = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} B_{k_1 k_2} = \text{constant} \ . \tag{22}$$

Using the Lagrange multiplier method,

$$\frac{\partial}{\partial B_{k_1 k_2}} \left[\sigma_q^2 - \lambda \left(B - \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} B_{k_1 k_2} \right) \right] = 0 \ . \tag{23}$$

This provides the optimum bit allocation expression as

$$B_{k_1 k_2} = B + \frac{1}{2} \log_2 \frac{\sigma_{k_1 k_2}^2}{\left(\prod_{k_1=0}^{N_1-1} \prod_{k_2=0}^{N_2-1} \sigma_{k_1 k_2}^2 \right)^{1/N_1 N_2}} \ , \tag{24}$$

where $B_{k_1 k_2}$ are not restricted to being nonnegative. In practice, they are truncated to zero if they become negative. A negative bit allocation result implies that if that band is completely discarded, its reconstruction error contribution is still less than the corresponding distortion for the given rate. The resulting quantization error variance by using this optimum bit allocation is found to be

$$\min\{\sigma_q^2\} = \varepsilon^2 2^{-2B} \left(\prod_{k_1=0}^{N_1-1} \prod_{k_2=0}^{N_2-1} \sigma_{k_1 k_2}^2 \right)^{1/N_1 N_2} \ . \tag{25}$$

Assuming the same pdf type also for the input signal, the distortion for PCM at the same rate is

$$\sigma_{q\text{PCM}}^2 = \varepsilon^2 2^{-2B} \sigma_x^2 \ , \tag{26}$$

and the distortion gain due to the orthonormal transform is found as

$$\max\{G_{\text{TC}}\} = \frac{\sigma_{q\text{PCM}}^2}{\min\{\sigma_q^2\}} = \frac{\frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sigma_{k_1 k_2}^2}{\left(\prod_{k_1=0}^{N_1-1} \prod_{k_2=0}^{N_2-1} \sigma_{k_1 k_2}^2 \right)^{1/N_1 N_2}} \ . \tag{27}$$

If the process repeats regularly for L levels with the number of bands in each decomposition as N_i for $i = 1,2,\ldots,L$, this expression is generalized for $\prod_{i=1}^{L} N_i$ bands as

$$\max\{G_{\text{TC}}\} = \frac{\left(\prod_{i=1}^{L} N_i \right)^{-1} \left(\sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \cdots \sum_{k_L=0}^{N_L-1} \sigma_{k_1 k_2 \ldots k_L}^2 \right)}{\left(\prod_{k_1=0}^{N_1-1} \prod_{k_2=0}^{N_2-1} \cdots \prod_{k_L=0}^{N_L-1} \sigma_{k_1 k_2 \ldots k_L}^2 \right)^{[\prod_{i=1}^{L} N_i]^{-1}}} \ . \tag{28}$$

This expression is valid for any regular tree structure of orthonormal filter banks or block transforms.

A similar expression for the irregular tree structures will be derived now. The case considered here assumes an N_1 band orthonormal decomposition in the first level of the tree, and only band p is decomposed further into N_2 bands in the second level of the tree. Since any middle level node in the tree means an orthonormal decomposition, the overall average distortion for this scenario can be similarly written as before:

$$\sigma_q^2 = \frac{1}{N_1} \sum_{\substack{k_1=0 \\ k_1 \neq p}}^{N_1-1} \sigma_{qk_1}^2 + \frac{1}{N_1 N_2} \sum_{k_2=0}^{N_1-1} \sigma_{qpk_2}^2 \tag{29}$$

and the band distortion terms with the assumption of the same pdf types

$$\sigma_{qk_1}^2 = \varepsilon^2 2^{-2B_{k_1}} \sigma_{k_1}^2 \qquad \begin{matrix} k_1 = 0,1,\ldots,N_1-1 \\ k_1 \neq p \end{matrix} \ , \tag{30}$$

$$\sigma_{qpk_2}^2 = \varepsilon^2 2^{-2B_{pk_2}} \sigma_{pk_2}^2 \qquad k_2 = 0,1,\ldots,N_2-1 \ . \tag{31}$$

The optimization problem now is to find the bit allocation among the bands of this irregular tree that minimizes σ_q^2, with equal band distortions in the same level of the tree, and satisfies the rate constraint

$$B = \frac{1}{N_1} \left(\sum_{\substack{k_1=0 \\ k_1 \neq p}}^{N_1-1} B_{k_1} + \frac{1}{N_2} \sum_{k_2=0}^{N_2-1} B_{pk_2} \right) = \text{constant} \ . \tag{32}$$

Using the Lagrange multiplier technique again,

$$\frac{\partial}{\partial B_{k_1}} \left[\sigma_q^2 - \lambda \left(B - \frac{1}{N_1} \sum_{\substack{k_1=0 \\ k_1 \neq p}}^{N_1-1} B_{k_1} - \frac{1}{N_1 N_2} \sum_{k_2=0}^{N_2-1} B_{pk_2} \right) \right] = 0$$

$$\text{for} \begin{matrix} k_1 = 0,1,\ldots,(N_1-1) \\ k_1 \neq p \end{matrix} \ , \tag{33}$$

$$\frac{\partial}{\partial B_{pk_2}} \left[\sigma_q^2 - \lambda \left(B - \frac{1}{N_1} \sum_{\substack{k_1=0 \\ k_1 \neq p}}^{N_1-1} B_{k_1} - \frac{1}{N_1 N_2} \sum_{k_2=0}^{N_2-1} B_{pk_2} \right) \right] = 0$$

$$\text{for } k_2 = 0,1,\ldots,(N_2-1) \ . \tag{34}$$

The corresponding optimum bit allocation expressions are found as

$$B_{k_1} = B + \frac{1}{2} \log_2 \frac{\sigma_{k_1}^2}{\left(\prod_{\substack{k_1=0 \\ k_1 \neq p}}^{N_1-1} \sigma_{k_1}^2 \right)^{1/N_1} \left(\prod_{k_2=0}^{N_2} \sigma_{pk_2}^2 \right)^{1/N_1 N_2}}$$

$$\text{for} \begin{matrix} k_1 = 0,1,\ldots,(N_1-1) \\ k_1 \neq p \end{matrix} \ , \tag{35}$$

$$B_{pk_2} = B + \frac{1}{2} \log_2 \frac{\sigma_{pk_2}^2}{\left(\prod_{\substack{k_1=0 \\ k_1 \neq p}}^{N_1-1} \sigma_{k_1}^2 \right)^{1/N_1} \left(\prod_{k_2=0}^{N_2} \sigma_{pk_2}^2 \right)^{1/N_1 N_2}}$$

$$\text{for } k_2 = 0,1,\ldots,(N_2-1) \ , \tag{36}$$

and yield average distortion as

$$\min\{\sigma_q^2\} = \varepsilon^2 2^{-2B}\left[\left(\prod_{k_2=0}^{N_2-1}\sigma_{pk_2}^2\right)^{1/N_2}\left(\prod_{\substack{k_1=0\\k_1\neq p}}^{N_1-1}\sigma_{k_1}^2\right)\right]^{1/N_1}, \tag{37}$$

assuming the same pdf type again for the original signal provides

$$\max\{G_{TC}\} = \frac{\sigma_{qPCM}^2}{\min\{\sigma_q^2\}} = \frac{\sigma_x^2}{\left[\left(\prod_{k_2=0}^{N_2-1}\sigma_{pk_2}^2\right)^{1/N_2}\left(\prod_{\substack{k_1=0\\k_1\neq p}}^{N_1-1}\sigma_{k_1}^2\right)\right]^{1/N_1}}. \tag{38}$$

This result can be extended to any arbitrary tree structure and the regular tree is regarded as a special case. This measure can also be extended easily to two-dimensional separable tree structures.

4.1. Energy compaction upper bounds

It was mentioned earlier that the KLT is the optimum solution for the block filter banks. Its basis functions are calculated based on the input signal statistics, and they decorrelate the bands perfectly. Therefore, the upper bounds of G_{TC} for orthonormal block filter banks or transforms are set by the performance of KLT for the given N-band decomposition. On the other hand, the global upper bounds of G_{TC} for any orthonormal filter bank will be defined by the performance of the ideal filter banks. The ideal filter banks are optimal based on their perfect interband decorrelation for any signal source and frequency band split characteristics for multirate signal processing. For a known input power spectral density function $P_{xx}(\omega)$, the band variances of the N band ideal filter bank are obtained from

$$\sigma_i^2 = \frac{1}{\pi}\int_{\frac{\pi}{N}i}^{\frac{\pi}{N}(i+1)} P_{xx}(\omega)\,d\omega \qquad i = 0,1,\ldots,N-1, \tag{39}$$

and the performance upper bound G_{TC}^{ub} is calculated using Eq. (16). A similar approach also provides the performance upper bounds for irregular, unequal bandwidth tree structures by assuming the ideal half-band QMF banks.

4.2. Performance results

The G_{TC} results for several different cases are presented in this section. First, the decomposition schemes assume an AR(1) input signal, which is defined as

$$X(n) = \rho X(n-1) + N(n), \tag{40}$$

where ρ is the correlation coefficient and $\{N(n)\}$ is a zero-mean white noise with known variance. The autocorrelation sequence of this source with unit variance is

$$R_{xx}(m) = \rho^{|m|} \qquad m = 0,\pm 1,\ldots, \tag{41}$$

Table 1. Energy compaction performance of several irregular transform trees along with the full tree and upper performance bounds for AR(1) source of $\rho = 0.95$. (a) 2×2 DCT, (b) 4×4 DCT, and (c) 8×8 DCT.

Table 1.a 2 × 2 DCT

level	Regular Tree			Irregular Tree		
	# of bands	G_{TC}	G_{TC}^{ub}	# of bands	G_{TC}	G_{TC}^{ub}
1	2	3.2026	3.9462	2	3.2026	3.9462
2	4	5.2172	7.2290	3	5.2165	7.1532
3	8	6.2317	9.1604	4	6.2265	8.9617
4	16	6.5980	9.9407	5	6.5801	9.6232
5	32	6.7132	10.1791	6	6.6725	9.7940
6	64	6.7545	10.2376	7	6.7136	9.8209

Table 1.b 4 × 4 DCT

level	Regular Tree			Irregular Tree		
	# of bands	G_{TC}	G_{TC}^{ub}	# of bands	G_{TC}	G_{TC}^{ub}
1	4	5.7149	7.2290	4	5.7149	7.2290
2	16	7.4789	9.9407	7	7.4685	9.7619
3	64	7.6689	10.2376	10	7.6234	9.9816

Table 1.c 8 × 8 DCT

level	Regular Tree			Irregular Tree		
	# of bands	G_{TC}	G_{TC}^{ub}	# of bands	G_{TC}	G_{TC}^{ub}
1	8	7.6312	9.1604	8	7.6312	9.1604
2	64	8.4475	10.2376	15	8.4171	10.0962

and its power spectral density function is

$$S_{xx}(\omega) = \frac{1-\rho^2}{1+\rho^2-2\rho\cos\omega}. \tag{42}$$

AR(1) sources are a crude approximation to real-world signals. Typical values are $\rho = 0.85$ for speech and $\rho = 0.97$ for still images. Table 1 gives the energy compaction results of several different irregular trees based on block transforms along with the corresponding regular tree performance and the upper performance bounds. These examples assume a full tree size of 64 bands. Table 1a employs a 2 × 2 transform for this purpose and the irregular tree is based on splitting only the lowest band again into two bands. The level of this tree is $L = 6$. Table 1b uses the 4 × 4 DCT basis and only the lowest bands are split again into four subbands and $L = 3$. Similarly, Table 1c employs an 8 × 8 DCT basis as the filter bank and $L = 2$. All these results assume $\rho = 0.95$. It is seen that the irregular trees perform very close to the regular trees of the same case for this signal model, but the first one is much simpler than the second one to implement.

Table 2 displays the compaction results of binomial-QMF banks,[8] which are identical to the orthonormal wavelet filters proposed in Ref. 11 for four-tap, six-tap, and eight-tap cases, respectively. These results are for octave band irregular tree structures and the corresponding regular trees along with the upper performance bounds. The levels of trees are limited to $L = 4$ here. It is observed from these tables that even a five-octave band irregular tree with a four-tap filter provides better performance than the 16-band block filter bank. It is clear that the irregular tree structures reduce the computational expense of the subband filter banks and make them practical competitors to block filter banks or transforms.

These results suggest that a practical scheme to define an irregular tree based on the input signal is very important. An

Table 2. Energy compaction performance of PR-QMF filter banks along with the full tree and upper performance bounds for AR(1) source of ρ = 0.95. (a) 4-tap binomial-QMF, (b) 6-tap binomial-QMF, and (c) 8-tap binomial-QMF.

Table 2.a 4-tap Binomial-QMF

level	Regular Tree			Half Band Irregular Tree		
	# of bands	G_{TC}	G_{TC}^{ub}	# of bands	G_{TC}	G_{TC}^{ub}
1	2	3.6389	3.9462	2	3.6389	3.9462
2	4	6.4321	7.2290	3	6.3681	7.1532
3	8	8.0147	9.1604	4	7.8216	8.9617
4	16	8.6503	9.9407	5	8.3419	9.6232

Table 2.b 6-tap Binomial-QMF

level	Regular Tree			Half Band Irregular Tree		
	# of bands	G_{TC}	G_{TC}^{ub}	# of bands	G_{TC}	G_{TC}^{ub}
1	2	3.7608	3.9462	2	3.7608	3.9462
2	4	6.7664	7.2290	3	6.6956	7.1532
3	8	8.5291	9.1604	4	8.2841	8.9617
4	16	9.2505	9.9407	5	8.8592	9.6232

Table 2.c 8-tap Binomial-QMF

level	Regular Tree			Half Band Irregular Tree		
	# of bands	G_{TC}	G_{TC}^{ub}	# of bands	G_{TC}	G_{TC}^{ub}
1	2	3.8132	3.9462	2	3.8132	3.9462
2	4	6.9075	7.2290	3	6.8355	7.1532
3	8	8.7431	9.1604	4	8.4828	8.9617
4	16	9.4979	9.9407	5	9.0826	9.6232

algorithm based on the input statistics and energy compaction criterion is proposed in the following section.

Table 3 displays the compaction performance of several different six-tap orthonormal wavelet filters, namely, binomial QMF,[8,11] most regular wavelet filters,[15] Coiflet,[15] for two-, four- and eight-band signal decompositions along with the KLT and ideal filter bank. The results in this table assume an AR(1) source with ρ = 0.95. These results indicate that the most regular filter did not perform the best even for a highly correlated signal source. Although the mathematical interpretation of the regularity in wavelets is very meaningful, its practical significance in signal processing is not fully understood.

Figure 5 shows the G_{TC} results of KLT, ideal filter bank, and six-tap binomial-QMF for different resolution regular trees. It is seen that the binomial-QMF reaches the global upper bound when $N \rightarrow \infty$ faster than the KLT. This result is expected because the binomial-QMF has maximally flat magnitude square and the signal source considered has a decreasing spectral density. Faster decay of the input spectral density means even better performance for the maximally flat magnitude square PR-QMF bank.

Any good PR-QMF bank should have an energy compaction performance curve that lies in the region between the KLT and the ideal filter bank curves and satisfies the Nyquist conditions for multirate signal processing. Additionally, its computational

Table 3. Energy compaction performance of several six-tap wavelet filters along with the KLT and ideal filter bank for an AR(1) source of ρ = 0.95.

	2-Bands	4-Bands	8-Bans
6-Tap Maxregular Filter	3.745	6.725	8.464
6-Tap Coiflet	3.653	6.462	8.061
6-Tap Binomial QMF(Maxflat)	3.760	6.766	8.529
KLT	3.202	5.730	7.660
Ideal Filter Bank	3.946	7.230	9.160

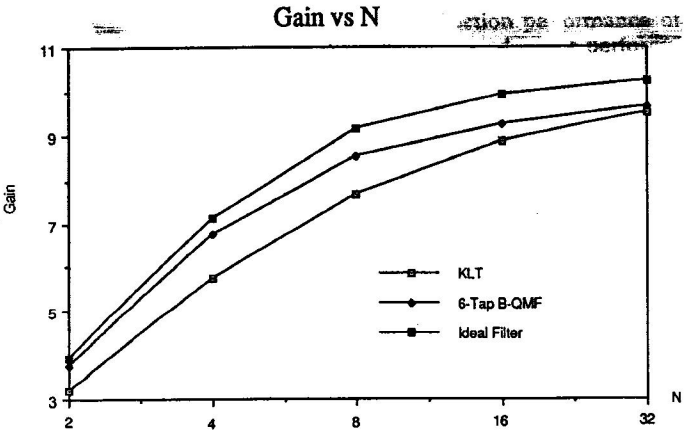

Fig. 5. Graph of G_{TC} versus N for a six-tap binomial QMF, KLT, and ideal filter bank assuming an AR(1) source with ρ = 0.95.

complexity should be comparable to fixed transform filter banks such as DCTs.

5. SIMPLE TREE STRUCTURING ALGORITHM

The main advantage of the multiresolution signal decomposition techniques is to be able to localize the features of a signal in both frequency and time domains. The availability of simultaneous time-frequency features is very useful for many applications. In general, most practical signals have a band-limited frequency spectrum or a significant portion of the signal energy is concentrated within several frequency bands, or packets.[16] Therefore, the regular decomposition tree is not fully justified.

A simple algorithm to modify the regular tree based on the signal to be decomposed is introduced in this section. The modified tree is called an *irregular tree*. The bands in this irregular tree are adapted to the signal spectrum. It is more efficient and has fewer filters than the regular tree.

The proposed algorithm based on an energy compaction criterion can be connected to entropy measures.[14] The algorithm also assumes ideal filter banks and half-band frequency splits only. If an input signal, with known power spectral density $P_{xx}(\omega)$, is split into low and high half-bands with an ideal QMF, the corresponding energy compaction is given as

$$G_{TC} = \frac{\sigma_x^2}{\left(\sigma_l^2 \sigma_h^2\right)^{1/2}}, \tag{43}$$

where

$$\sigma_l^2 = \frac{1}{\pi} \int_0^{\pi/2} P_{xx}(\omega)\, d\omega, \tag{44}$$

$$\sigma_h^2 = \frac{1}{\pi} \int_{\pi/2}^{\pi} P_{xx}(\omega)\, d\omega. \tag{45}$$

The two-band energy ratio is defined as

$$\eta = \frac{\sigma_h^2}{\sigma_l^2}. \tag{46}$$

Therefore, for a uniform power spectrum $\eta = 1$, $G_{TC} = 1$. Figure 6 displays the variation of G_{TC} with respect to η.

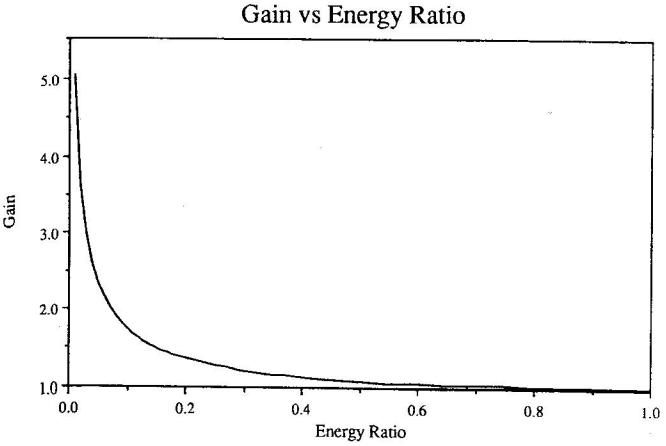

Fig. 6. Graph of G_{TC} versus η for AR(1) source, $\rho = 0.95$.

The tree structuring algorithm for a given maximum frequency resolution $BW_{min} = \pi/2^L$ can be summarized as follows:

1. Measure the power spectral density $P_{xx}(\omega)$ of the given zero mean source.
2. Calculate σ_l^2, σ_h^2 and find η. Compare η with a predefined threshold T:
   If $\eta \leq T$, stop.
   If $\eta > T$, band split, obtain l and h bands.
3. Calculate σ_{ll}^2, σ_{lh}^2 and σ_{hl}^2, σ_{hh}^2 band variances and corresponding η_l and η_h. Check with the threshold T:
   If $\eta l \leq T$, stop.
   If $\eta l > T$, band split, obtain ll and lh bands.
   If $\eta_h \leq T$, stop.
   If $\eta_h > T$, band split, obtain hl and hh bands.
4. Repeat the procedure until the desired frequency resolution BW_{min} is reached. The corresponding suboptimum tree structure based on input statistics is obtained.

The proposed algorithm is signal dependent and the regular tree of 2^L equal bands is adapted to the signal power spectrum within M unequal bands, $2^L > M$. Whenver $2^L = M$, the irregular tree becomes a regular tree. For many signal sources, an irregular tree is a better practical choice. Obviously, a regular tree provides the upper compaction bound for known L.

The validity of the algorithm is tested with several source models as well as several test images. The tree structuring algorithm can be easily extended for two-dimensional separable transforms. After defining the signal-dependent tree structure, real filters replace the ideal filters in the corresponding filter bank. Table 4 displays the energy compaction performance of several decomposition techniques for the standard test images, Lena, building, cameraman, and brain. The images are 256 × 256 pixels monochrome with 8 bit/pel resolution. The test results displayed in Table 4 are broadly consistent with the results obtained for AR(1) signal sources. These tables also show that the irregular tree achieves a performance very close to that of the regular tree, but with fewer bands.

Table 4. G_{TC} of several different regular and irregular tree structures along with the DCT for the test images.

TEST IMAGE	LENA	BUILDING	CAMERAMAN	BRAIN
8 × 8 2D DCT	21.99	20.08	19.10	3.79
64 Band Regular 4-tap B-QMF	19.38	18.82	18.43	3.73
64 Band Regular 6-tap B-QMF	22.12	21.09	20.34	3.82
64 Band Regular 8-tap B-QMF	24.03	22.71	21.45	3.93
4 × 4 2D DCT	16.00	14.11	14.23	3.29
16 Band Regular 4-tap B-QMF	16.70	15.37	15.45	3.25
16 Band Regular 6-tap B-QMF	18.99	16.94	16.91	3.32
16 Band Regular 8-tap B-QMF	20.37	18.17	17.98	3.42
*10 Band Irregular 4-tap B-QMF	16.50	14.95	13.30	3.34
*10 Band Irregular 6-tap B-QMF	18.65	16.55	14.88	3.66
*10 Band Irregular 8-tap B-QMF	19.66	17.17	15.50	3.75

* Bands used are $llll - llllh - lllhl - lllhh - llh - lhl - lhh - lh - hl - hh$.

6. DISCUSSIONS AND CONCLUSIONS

A unified approach to the popular orthonormal signal decomposition techniques is presented. Energy compaction performance of block transforms and perfect reconstruction filter banks for AR(1) and test image sources are compared along with the optimum solutions. As expected, the filter banks outperform the block transforms. The frequency behavior of their basis functions is considered in the design of filter banks, but not in the block transforms. The computational complexities and subjective performance of the various decomposition techniques are not considered in this study. It is well known that block transforms are simpler to implement, but they have a fixed resolution. On the other hand, filter banks have multiresolution signal representation as a by-product. The comparative subjective performance is a subject for future study. Among the filter banks considered, the maxflat solution[8,11] marginally outperforms the maxregular[15] solution for the sources considered here. The practical significance of the regularity remains a subject for further study. It is concluded that the filter banks with simple algorithms and irregular tree structures are potential competitors to the industry standard DCT for next-generation video codecs.

7. REFERENCES

1. H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans. ASSP* 37, 553–559 (April 1989).
2. H. S. Malvar, "The LOT: a link between block transform coding and multirate filter banks," *Proc. ISCAS*, pp. 781–784 (July 1988).
3. M. Smith and T. P. Barnwell, "Exact reconstruction techniques for tree structure subband coders," *IEEE Trans. ASSP* 434–441 (1986).
4. P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect reconstruction techniques," *IEEE ASSP Mag.* 4–20 (July 1987).
5. M. Vetterli and D. LeGall, "Perfect reconstruction FIR filter banks: some properties and factorizations," *IEEE Trans. ASSP* 1051–1071 (July 1989).
6. E. Simoncelli, *Orthogonal Sub-band Image Transforms*, MS Thesis, Massachusetts Institute of Technology (May 1988).
7. P. H. Westerink, *Subband Coding of Images*, Ph.D. Thesis, Delft University (1989).
8. A. N. Akansu, R. A. Haddad, and H. Caglar, "The binomial QMF-wavelet transform for multiresolution signal decomposition," Submitted to *IEEE Trans. ASSP*.
9. H. S. Malvar, "Efficient signal coding with hierarchical lapped transforms," *Proc. ICASSP* 1519–1522 (1990).
10. P. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Comm.* 532–540 (April 1983).
11. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.* XLI, 909–996 (1988).
12. S. G. Mallat, "A theory for multiresolution signal decomposition," MS-CIS-87-22, GRASP Lab. 103, University of Pennsylvania (May 1987).
13. S. G. Mallat, "Multifrequency channel decomposition of images and wavelet models," *IEEE Trans. ASSP* 2091–2110 (Dec. 1989).
14. N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, New Jersey (1984).
15. I. Daubechies, "Orthonormal bases of compactly supported wavelets: II. variations on a theme," Yale University Preprint (1990).
16. R. R. Coifman, Y. Meyer, S. Quake, M. V. Wickerhauser, "Signal processing and compression with wave packets," Yale University preprint (1990).

Ali N. Akansu received the BS degree from the Technical University of Istanbul, Turkey, in 1980, the MS degree from the Polytechnic Institute of New York in 1983, and the Ph.D. degree from the Polytechnic University in 1987, all in electrical engineering. He has been an assistant professor of electrical engineering at New Jersey Institute of Technology, Newark, since 1987. He was academic visitor at IBM T. J. Watson Research Center during the summer of 1989. His current research interests are one- and multidimensional signal processing, source coding, and pattern recognition. Dr. Akansu is a member of IEEE, SPIE, and Sigma Xi.

Yipeng Liu received her BS and MS degrees from Beijing Institute of Technology in 1984 and 1987, respectively, both in electrical engineering. She was a faculty member of Beijing Institute of Technology from 1987 to 1989. She has been with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, as a Ph.D student since 1989. Her current research interests are signal decomposition techniques, image processing, and coding.