

Multiresolution Coding Techniques for Digital Television: A Review

MARTIN VETTERLI* AND KAMIL METIN UZ†

Department of Electrical Engineering and Center for Telecommunications Research, Columbia University, New York, NY 10027-6699

Received March 12, 1991, Revised September 20, 1991

Invited Paper

Abstract. Multiresolution decompositions for video coding are reviewed. Both nonrecursive and recursive coding schemes are considered. In nonrecursive schemes, it is shown that pyramid structures have certain advantages over subband or wavelet techniques, and a specific spatiotemporal pyramid coding of HDTV is discussed in some detail. It is shown that recursive, DPCM like schemes will incur a slight loss of optimality due to a restricted form of prediction if multiresolution decomposition with compatible decoding is required. Compatibility and transmission issues are also discussed. Multiresolution transmission for digital broadcast TV is introduced. This, when combined with multiresolution source coding, achieves spectrum efficiency, robustness and graceful degradation under channel impairments.

Key Words: Digital television, multiresolution coding, video compression

“In 1927, Gray, Horton and Mathes [1] gave the first full theoretical discussion of the influence of waveband restriction on the quality of television pictures, and were able to fix the minimum waveband requirements in advance, long before the first high-definition system was realized.”

D. Gabor, “Theory of Communication,” *Journal of the IEE*, 1946 [2]

1. Introduction

The processing and compression of digital video signals has recently become the focus of intensive research. Past efforts were mainly concerned with high compression of relatively simple sequences (e.g., the compression of video conference signals down to 64 kbit/s). Currently, however, the efforts have been broadened to cover an extensive set of applications, from video conferencing to high definition television, and the associated target bit rates after compression range from under 100 kbit/s to several tens of megabits per second.

Traditionally, television has used a mix of continuous and sampled processing. Two out of the three dimensions in regular television are discrete, but the sample values are not quantized. Gradually, digital processing has been used to improve picture quality by pre- and postprocessing. But by now, it is clear that the future of television lies in sampled digital processing [3]. Television can be viewed as a three-dimensional sampled signal with three-component discrete values (for the representation of color in an appropriate color space).

*Work supported in part by the National Science Foundation under grants ECD-88-11111, MIP-90-14189 and Bell Communications Research.

†Work supported by the National Science Foundation under grants ECD-88-11111. K.M. Uz is now with David Sarnoff Research Center in Princeton, NJ 08543.

This digital view of video has reconciled, at least technologically, the computer and the broadcast industry, and leads to many new applications ranging from storage of video on CD-ROMS [4], [5] to transmission of video over packet networks and possibly a new digital standard for production and broadcast of high definition television.

This unified approach to television raises the following questions:

1. What are the *best* compression techniques for the various applications?
2. Are there methods which will allow a certain *compatibility* between various applications?
3. Besides their compression performance, do the methods blend well with other requirements, related to storage, transmission, and pre/postprocessing?

The goal of the present paper is to explore how the concept of multiresolution (MR) signal processing can be used to address the above questions in digital video. This concept, which encompasses several well-known coding techniques like subband and pyramid coding, is based on the idea of representing a signal at various resolution levels, and going from one resolution to the next by adding *augmentation* channels (or added details).

The outline of the paper is as follows. Section 2 introduces multiresolution processing and explores some of its benefits and limitations. Finite memory MR schemes, namely transform, subband, wavelet and pyramid coding are introduced in section 3, indicating some advantages of the latter for video compression and representation [6]. Section 4 reviews classical recursive coding techniques of the differential pulse-code modulation (DPCM) type, including hybrid motion compensated predictive discrete cosine transform (HMCP-DCT) coding, showing how to achieve MR decomposition (and at what price). The issue of compatibility and transcoding is discussed in Section 5 and Section 6 deals with transmission issues like robustness and MR transmission for digital broadcasting.

It should already be clear from the outline that digital video compression and transmission is a systems problem involving many constraints, and that specific technical questions like compression at a given resolution cannot be solved in isolation but have to be considered within this global context.

2. Multiresolution signal processing and analysis

The concept of MR processing is based on the analysis of a signal at a hierarchy of scales. Typical signals one deals with in real life (including video) have a lowpass nature: the power spectrum rapidly falls off at high frequencies. Therefore, a coarse version containing the lowpass content is a good approximation in the mean square error (MSE) sense. In the case of video, this approximation is also good in the perceptual sense as the human visual system (HVS) has a roughly bandpass response falling off at high frequencies [7]. Furthermore, there is strong evidence that the HVS has a MR nature [8]. MR approaches are therefore natural from both signal processing and perceptual points of view.

Predictive coding algorithms achieve compression by predicting the signal and coding the prediction error. Such a prediction can be based on a low resolution approximation, for example, in pyramid coding [9]. That is, the coding algorithm can be seen as a *successive approximation method*.

In representation of video one may require a multiresolution decomposition. As an example, for compatibility purposes, a lower resolution version of high-definition television could be similar to regular definition television. Another example would be a video database, where browsing would be facilitated by having low resolution versions for quick access. Also, storage on tape in a multiresolution format makes fast monitored access (fast forward/reverse with viewing) easier if the decoding is compatible.

Finally, for transmission purposes, a multiresolution decomposition can be used to achieve better performance by guaranteeing higher protection for the more important low resolution approximations. Such joint source/channel coding can be advantageous both for time-varying channels (asynchronous transmission like ATM) and broadcast situations.

Then, source and channel coding are usually done separately. This is optimal if both can be done optimally, which is only possible in the limit of long block lengths and perfect knowledge of source and channel. In more practical situations, joint source/channel coding is beneficial, and multiresolution decomposition is a method of choice to match the source and channel coding. In particular, the broadcast situation is particularly suited for a multiresolution transmission, since there is no "single" channel, but many different ones to be accommodated.

Despite these attractive features, several key questions have to be answered before one applies a multiresolution decomposition. First, how efficient can a multiresolution source coding be? As to be expected from general information theoretic results, if complexity is not an issue, it is more efficient to encode the signal in whole rather than divide it into parts that are coded separately. Thus, multiresolution decomposition will be suboptimal in general. Equitz [10] has shown cases where successive approximation is still optimal, but they are restricted.

However, in the complexity bound case (i.e., practical applications), it is not clear how suboptimal a multiresolution approach will be, especially considering subband and pyramid coding have been fairly successful as image compression techniques. Still, it is clear that the constraint of having a compatible subchannel of a given quality (like in the high definition versus ordinary television example) is fairly restrictive, and will lead to suboptimal performance when compared to an unconstrained coding.

Can the multiresolution decomposition be used in all coding techniques? It turns out that they are naturally suited for finite memory schemes such as transform or subband coding, while recursive schemes with a multiresolution structure incur a certain loss in coding efficiency. There is a basic difference between open loop (or finite memory) coding methods and closed loop (or recursive) coding methods. In the former case, the various resolutions and the augmentations are treated independently, making compatible decoding easy (i.e., decoding of the low resolution version only). Even if there were an interaction between the channels, the error produced by not taking the augmentation channel into account (in decoding the low resolution) would be bounded. In recursive schemes, unless some additional constraints are imposed, such a compatible decoding is not possible because the error could grow without bound. This fundamentally different behavior appears because errors do accumulate in closed loop systems (which behave like integrators) while they have finite effect in open loop systems.

Before going into detail about MR coding techniques, we would like to define some concepts which are key to understanding multirate and wavelet based systems and have been

used in computer vision as well [11]. The notion of *resolution* of a signal is intuitively clear, since added high frequencies means more detail or increased resolution. It is thus related to the bandwidth of the signal. This holds also in the sampled domain, but it is best thought of as the bandwidth of the equivalent continuous-time signal. This definition indicates that an oversampled version will not have more resolution than a critically sampled version of the same signal. The notion of *scale* is related to the *size* of the signal. We will adhere to the convention also used in the wavelet literature [12], [13], [14] that large scale denotes *contraction* of the signal, while small scale stands for a *dilated* signal. Thus given a continuous function $f(x)$ and a positive real number s , the scaled version $f(sx)$ will be contracted for $s > 1$ and stretched for $s < 1$. Figure 1 shows various operations from multirate signal processing and their effect on resolution and scale (for simplicity, only changes by factors for 2 are considered). Obviously, the resolution cannot be increased, unless information is added. Figure 2 shows these operations on a real image.

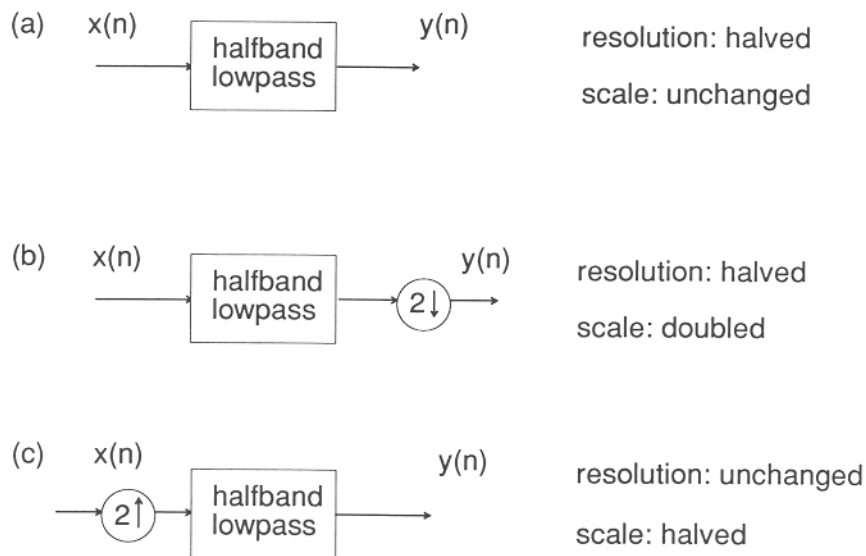


Figure 1. Resolution and scale changes in discrete time (by factors of 2). Note that the scale of signals is defined as in geographical maps. (a) Half-band low-pass filtering reduces the resolution by 2 (scale is unchanged). (b) Half-band low-pass filtering followed by subsampling by 2 doubles the scale (and halves the resolution as in (a)). (c) Upsampling by 2 followed by half-band lowpass filtering halves the scale (resolution is unchanged.).



Figure 2. Resolution and scale changes on an image: (a) original signal. (b) at higher scale (c) at higher resolution.

There is a fundamental difference between the continuous-time and the discrete-time scale change. In the former, scale changes are reversible, while in the latter, a *dilation* can be reversed, while a contraction cannot. This is due to the interaction of the sampling density and the scale of a signal. Since the Fourier transform of $f(sx)$ will be $\hat{f}(\omega/s)$, large scales require higher sampling frequencies in order to capture the same information on a given function. Note also the *change in scale* is the fundamental operation in the theory of wavelets [12], replacing modulation which is central in Fourier methods.

3. Finite memory multiresolution schemes

Such schemes include transform, subband, wavelet, and pyramid coding. They are all variations on the theme of the Karhunen-Loeve transform (KLT), and thus perform a natural ordering of the multiresolution components. After discussing the various schemes and comparing them, we will illustrate some specific points on the example of three-dimensional spatiotemporal pyramid coding of HDTV [6].

3.1. Karhunen-Loeve transform

Finite memory schemes have a natural relation to multiresolution decomposition. This is due to the fact that such schemes are approximations of the Karhunen-Loeve transform. Let us briefly review the KLT. Assume a vector process \mathbf{x} (typically, $\mathbf{x} = [x(n), x(n-1), \dots, x(n-N+1)]$) with autocorrelation matrix (assuming zero mean):

$$\mathbf{R} = E[\mathbf{x}\mathbf{x}^T]. \quad (1)$$

The matrix \mathbf{R} is symmetric and has thus a full set of orthogonal eigenvectors. Choosing a transform \mathbf{T} with rows equal to the eigenvectors of \mathbf{R} in decreasing order of the eigenvalues (\mathbf{R} is assumed positive definite), the transformed vector process $\mathbf{y} = \mathbf{T}\mathbf{x}$ has autocorrelation:

$$E[\mathbf{y}\mathbf{y}^T] = \mathbf{T}\mathbf{R}\mathbf{T}^T = \Delta \quad (2)$$

where Δ is a diagonal matrix with decreasing entries. Because T is unitary, that is, it conserves l_2 norms, the best subset of coefficients y_i , in the l_2 sense is the first k coefficients. This gives a simple ranking or prioritization of the transform coefficients. In a joint source-channel coding environment, more protection would thus be allocated to lower order coefficients.

It is well known that DCT coding is an approximation of the KLT for highly correlated first-order Markov processes [15]. In particular, lower frequency DCT coefficients will have higher energy. In that sense, subband and pyramid coding, which both rely on lowpass versions as first approximations, are similar to a KLT for processes with strong correlation.

It should always be kept in mind that the KLT produces a best approximation in the mean squared error sense or l_2 norm. In particular, the l_2 norm is conserved between the transform and the original domain, because T is unitary. However, if another norm is used (like for example maximum error or l_∞ norm), the KLT or any unitary transform will only produce a weak bound for that other norm. Let us make an illustrative example, which is relevant to high-quality coding applications. Assume we want to bound the l_∞ norm of the reconstructed signal after quantization in the DCT transform domain. The DCT uses the mean as one of the basis vectors, that is, the vector $1/\sqrt{N}[1 \ 1 \ \dots \ 1]$. Now, if quantization in transform domain produces an error of at most δ , the worst case reconstruction error is $\sqrt{N}\delta$, or an increase by \sqrt{N} over quantization of the original signal.

3.2. Subband coding

A typical subband coding scheme is shown in Figure 3 in its simplest version. It is a one-dimensional system with division into two subbands, each critically subsampled by 2. The subbands are encoded appropriately (that is, with adaptive encoders tailored to the bands), transmitted, decoded and recombined in a synthesis filter bank that upsamples the signals and interpolates them to reconstruct an approximation to the original. In the case of lossless encoding of the bands, there is a large body of literature on how to design filters so as to get perfect reconstruction (see e.g., [16]). It is most intuitive to look at subband coding as a transform coding where the basis vectors have some overlap between neighboring blocks [17], [18], [19]. Indeed, one can write the operation of the filter bank as a block Toeplitz

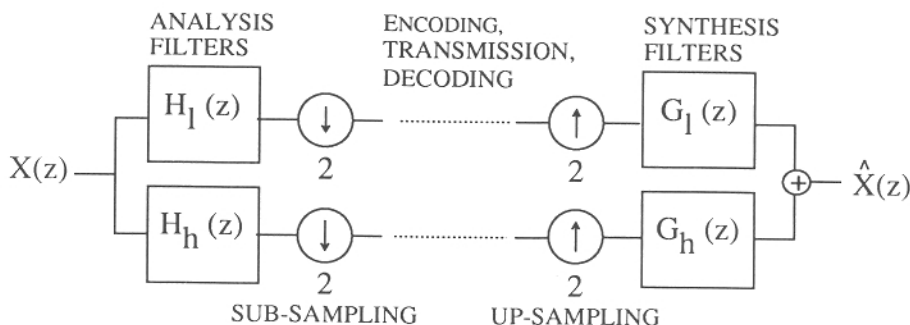


Figure 3. Two-channel subband coding system in one dimension.

operator \mathbf{T} (containing the impulse responses of the various filters and versions shifted by the subsampling factor) and then, if the bank is unitary [18]:

$$\mathbf{T}^* \cdot \mathbf{T} = \mathbf{I}. \quad (3)$$

In the z -transform domain, one can show that (3) is equivalent to the multiple-input, multiple-output (MIMO) transfer matrix being paraunitary or lossless (stable) [16], [18], that is, with $\mathbf{H}(z)$ standing for the MIMO transfer function matrix:

$$\tilde{\mathbf{H}}(z) \cdot \mathbf{H}(z) = \mathbf{I}, \quad (4)$$

where $\tilde{\mathbf{H}}(z)$ stands for $\mathbf{H}^T(z^{-1})$ (assuming real filter coefficients). The theory of these paraunitary matrices is well developed (e.g., [16]) and generalizations to biorthogonal cases (which include linear phase filter, important for image processing) have been done as well (e.g., [18]).

Because of the close relationship of subband filter banks to unitary transforms, the worst-case behavior of the reconstruction error discussed for the DCT appears also here (note that the number of subbands is usually small, however).

Again, ordering of the subbands according to their energies will minimize the squared error when only a subset is used for reconstruction. But note that because the design of filters for perfect reconstruction filter banks is heavily constrained, the reconstruction based on a subset is usually suboptimal in terms of perceptual quality when compared to an unconstrained derivation of a low-resolution version (using standard low-pass filters). Also, usual properties like aliasing cancellation are lost.

The extension of subband coding to two [20], [21] or three dimensions [22] is straightforward, especially in the separable case. Figure 4 shows a simple subband decomposition of video into combinations of low- and high-pass versions over the various dimensions. Several authors have suggested schemes using some sort of three-dimensional subband decomposition [22], [23], [24], or use of subband decomposition over the spatial dimension in video coding schemes. Obviously, the multiresolution nature is conserved, with lower bands contributing the basic information, while higher bands add more details, and the scheme was used successfully for joint source-channel coding in the context of packet video [25].

However, the inclusion of explicit motion information is not simple, essentially because it is sequence-domain information that has to be used in frequency domain. To put it explicitly and using Fourier terminology, a space-domain motion or shift shows up as a different phase shift in all the frequency components, and is therefore difficult to detect precisely and then to correct. This effect is well known in DCT coding (where DCT domain motion compensation is known to be difficult), but is present as well in subband decompositions.

The extension of subband coding to multidimensional nonseparable systems (that is, a true generalization to multiple dimensions) was also performed [26], [27], and applications thereof include progressive to interlaced conversion of video with the perfect inversion property [28] (using a nonseparable perfect reconstruction filter bank for quincunx subsampling).

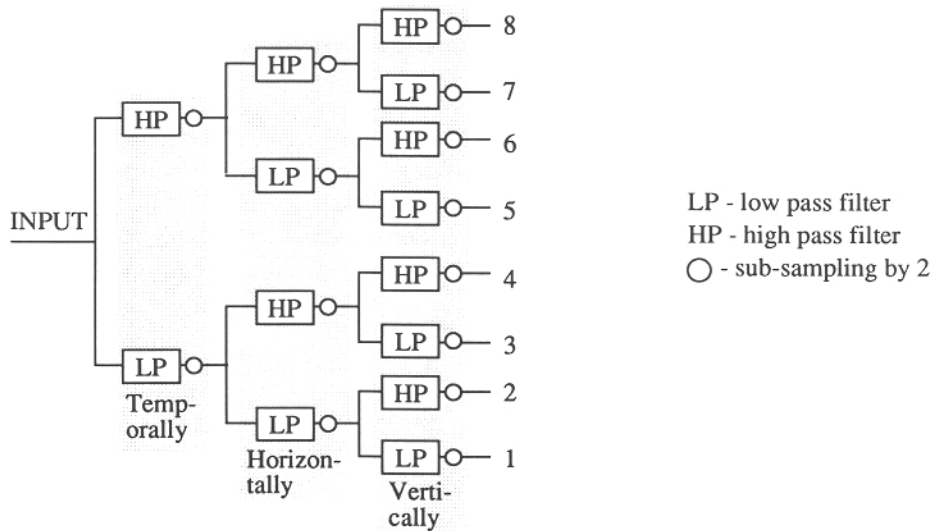


Figure 4. Three-dimensional subband coding system, with division into low and high frequencies in all three dimensions.

3.3. Wavelet coding

Discrete wavelet coding [29] is a subband coding technique with a logarithmic tree structure as shown in Figure 5. The logarithmic tree structure leads to a doubling of the resolution each time a channel is added. In wavelet coding, the filter bank uses a special type of low-pass filter called a *regular* filter. Such a filter has the property that, when iterated in a cascade of filtering-subsampling steps, it will tend to a smooth equivalent impulse response [12], [30]. Failure to meet regularity can produce equivalent filters which tend to fractal impulse responses when iterated. Since typical wavelet or subband coding iterates an elementary filter bank several times (three to five times, typically), an equivalent fractal filter can be problematic in such applications. This is because quantization "noise" will appear in the reconstructed signal as a weighted sum of the impulse responses, which can be more visible if they are discontinuous.

Besides the usual properties of filters used in subband coding (e.g., orthogonality or

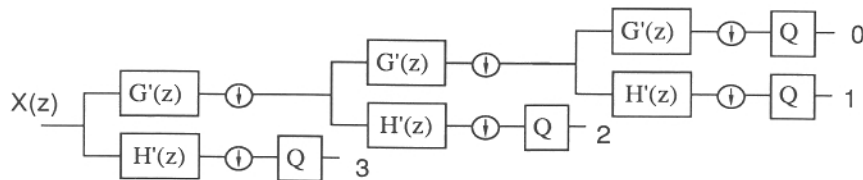


Figure 5. Wavelet coding based on a logarithmic filter bank tree.

linear phase), regular low-pass filters will have a large number of zeros at $z = -1$. The high-pass filter then has many zeros at $z = 1$, and smooth functions will produce little energy in the high bands of the wavelet analysis, a useful feature for compression.

Wavelet decomposition results in a logarithmic division of frequency space, similar to pyramidal techniques. This division can be justified by HVS models. Various psychovisual experiments [8] have suggested that the retinal processing uses independent bandpass filters that are approximately linear, and have a constant relative bandwidth of about one octave. Assigning a roughly constant number of bits per octave will lead to equal perceived quality over the channels. Therefore, the logarithmic spacing in the wavelet domain matches the HVS and maintains high perceptual quality while achieving compression.

Arbitrary binary trees based on two-channel filter banks are an alternative to logarithmic trees. This leads to so-called wavelet packets [31], which, together with algorithms for finding adaptively the "best" tree for a given signal, produce interesting schemes [32]. Such schemes are conceptually related to adaptive vector quantization trees [33], [34], [35].

3.4. Pyramid coding

The simplest example of pyramidal coding is given in Figure 6. A low-resolution version of the original is derived, from which an interpolation of the original is attempted. The difference between this interpolated version and the original is evaluated and sent together with the low-resolution version [9]. The scheme, can, of course, be iterated and can be so designed that only a single source of quantization error remains, namely, the quantization of the last difference signal (see Figure 7) [6]. The idea is to reconstruct at the encoder based on the encoded upper layers, so as to take quantization of these layers into account when deriving the final, highest resolution difference signal. Therefore, the worst-case error is easily bounded by designing the last quantizer accordingly. This inherent robustness is an advantage of pyramidal coding over both transform and subband coding [6]. The other main advantage is that the decimation and interpolation operators can now be arbitrary (rather than being heavily constrained as in subband schemes or fixed as in transform coding). That is, nonlinear operations are possible [36], and in particular, motion-based processing is easily included. The drawback is that the difference signal cannot be subsampled anymore, although it is really a high-pass signal. However, the overhead in

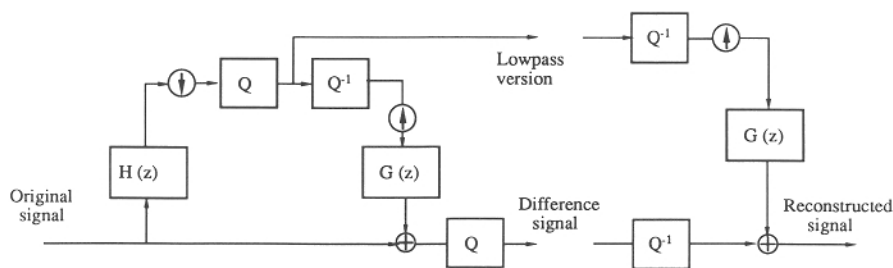


Figure 6. Pyramid coding.

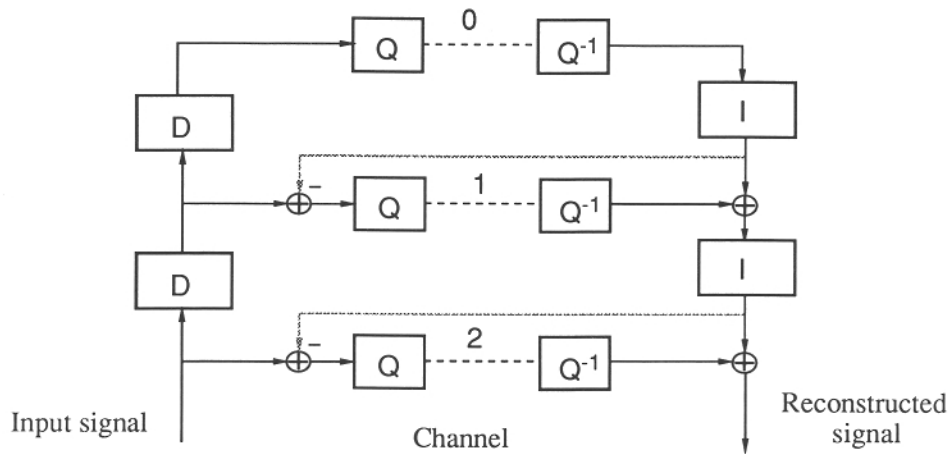


Figure 7. Three-level pyramid coding, with feedback of quantization of the high layers into the prediction of the lower ones; D and I stand for decimation and interpolation, respectively. Thus, only one source of quantization error remains, namely, that of the highest resolution difference signal.

number of samples decreases as the dimensionality increases. In m dimensions, we have an oversampling r as a function of the number of levels l in the pyramid equal to

$$r = \sum_{i=0}^{l-1} \left(\frac{1}{2^m} \right)^i < \frac{2^m}{2^m - 1}, \quad (5)$$

which is an overhead of 50%–100% in one dimension and goes down to 25%–33% in two dimensions and only 12.5%–14% in three dimensions.

3.5. Comparison of subband and pyramid methods

The first difference between subband (or wavelet) decompositions and pyramid ones is that the former is *critically sampled* (number of samples constant between signal and transform domains) while the latter is *oversampled* (increase in the number of samples in the transform domain). However, this oversampling ratio becomes negligible as the dimensionality increases (see (5)).

Next, the operators used to change resolution in subband or wavelet coding are very constrained. They must be perfect reconstruction filters, meeting restrictive algebraic constraints on their coefficients. The result is that subband low-pass filters of reasonable complexity are relatively poor. Thus, a compatible low-pass channel in a subband coding system is usually of inadequate quality. In pyramid coding, the operators for changing resolution are completely unconstrained. For instance, they can be nonlinear [36]. In particular, one can choose the best possible low-pass filter in order to derive a compatible subchannel. Therefore, compatible subchannels in pyramid coding systems are usually of better quality than their subband coding counterparts.

As explicated for transform coding (see Section 3.1), one gets only a weak bound on the maximum reconstruction error due to the possible coherent addition of errors from the various transform coefficients. The same situation is true in subband coding, which is a generalized transform. In pyramid coding, using the method of quantization feedback (see Section 3.4), the maximum error can be bounded by the quantization error of the last quantizer. This tight control on the quantization error can be important in applications where a precise bound on the maximum error is needed, as in medical applications or contribution quality video coding.

Finally, specifically for the case of coding video signals, it is important to discuss the inclusion of motion models within the various coding schemes. As already alluded to, motion is a sequence-domain phenomenon and is thus difficult to treat in the transform domain. This is the reason motion compensation in DCT domain has not been very successful. The same reason makes inclusion of motion in the subband domain difficult: a single motion is spread in all bands, and fine motion is quantized coarsely by the subsampling of the bands. Also, the increase in error by \sqrt{N} (where N is the number of subbands) from accumulation of errors can be problematic if motion errors occur in the subbands. The pyramid coding method does not suffer from these drawbacks: the nonlinear processing associated with motion detection does not present any problem, and approximation operators can be motion based if desired.

3.6. Example: spatiotemporal pyramid coding of HDTV

Some of the trade-offs discussed above can be well illustrated with a coding scheme for HDTV that we developed after investigating several alternative schemes. In particular, it shows the trade-offs between subband and pyramid coding. The discussion will remain mostly conceptual, and we refer the interested reader to [6] for additional details.

The goal was to develop a high-quality coding method with the following features:

1. Signal decomposition for compression purposes
2. Compatible subchannels
3. Tight control over coding error
4. Easy joint source-channel coding
5. Robustness to channel errors
6. Easy random access for digital storage

Clearly, a multiresolution scheme is desired (points 2 and 4). While the last two points indicate that a finite memory scheme will be preferable, let us discuss how the first three points influence a choice between subband and pyramid coding.

Among finite memory schemes, pyramid coding can painlessly include motion information to achieve high compression. This is due to the fact that motion estimation, and therefore motion compensation, is nonlinear, and it thus difficult to include in a scheme based on linear processing, such as subband decomposition (for example, if one wanted motion-compensated filters). In a pyramid, however, it is simple to base one of the predictions on motion, that is, for example, predict odd frames from even ones and encode the prediction error.

Compatible subchannels (that is, the coarse versions of the original) look poor if low-complexity subband filters are used. At comparable complexity, much better filters can be used in pyramids, since they are unconstrained. This leads to substantially better subchannels in pyramid schemes, both in two and three dimensions.

Quantization error performance has been discussed before, showing the superiority of pyramids over transform or subband schemes in this regard.

What about the oversampling present in pyramids? The relative advantage of subband schemes diminishes as one goes to higher dimensions. In a three-level three-dimensional pyramid, we end up with a full size signal, plus 1/8 and 1/64 size signals, that is, 14% overhead in the number of samples. Table 1 summarizes the comparison of subband and pyramid coding for video compression.

The three-dimensional spatiotemporal pyramid is shown in Figure 8(a). A higher level in the pyramid is obtained by a reduction in resolution by a factor of 2 in each dimension (or eight times less samples). This is done by low-pass filtering and subsampling in the two spatial dimensions and by straight subsampling in the time dimension. The prediction or interpolation step is shown in Figure 8(b). First, the spatial dimension is interpolated using linear filters, and the difference is encoded. Then the time dimension is motion interpolated, and the difference is encoded. This separation of space and time not only reduces the complexity but is natural due to the fundamental difference between space and time in video. This process is iterated twice, so that the final resolution is approximated in two steps. Note that besides the lowest resolution sequence (which represents 1/64 of the original samples), only difference sequences are encoded, together with motion vectors.

For illustration purposes, Figure 9 shows a frame from each of the three levels in the pyramid. Over time (a dimension which is difficult to show on paper!), if the original frame rate is f_s , the smaller ones have rate $f_s/2$ and $f_s/4$, respectively. The compression scheme, described in detail in [6], leads to a very high quality coding at around 1.5 bits/pixel, or a compression factor of 10.

Interestingly, the motion estimation procedure relies also on the multiresolution concept. An initial motion field of low resolution on a low-resolution sequence is successively refined until a full-resolution motion field is obtained [37], [38]. Such a procedure is both computationally efficient and robust.

Table 1. Comparison of subband and pyramid schemes for video coding.

Method	Subband	Pyramid
Oversampling	0%	14%
Max. coding error	$\sqrt{N}\delta$	δ
Subchannel	Poor	Good
Inclusion of motion	Difficult	Easy
Nonlinear processing	Difficult	Easy
Model based processing	Difficult	Easy

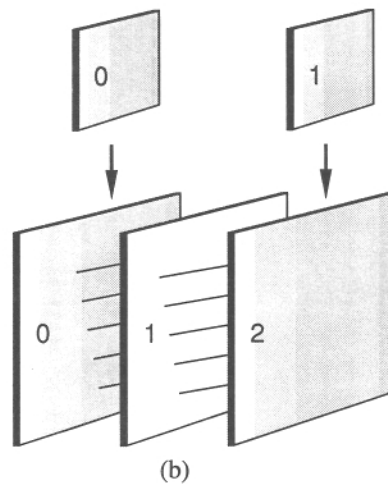
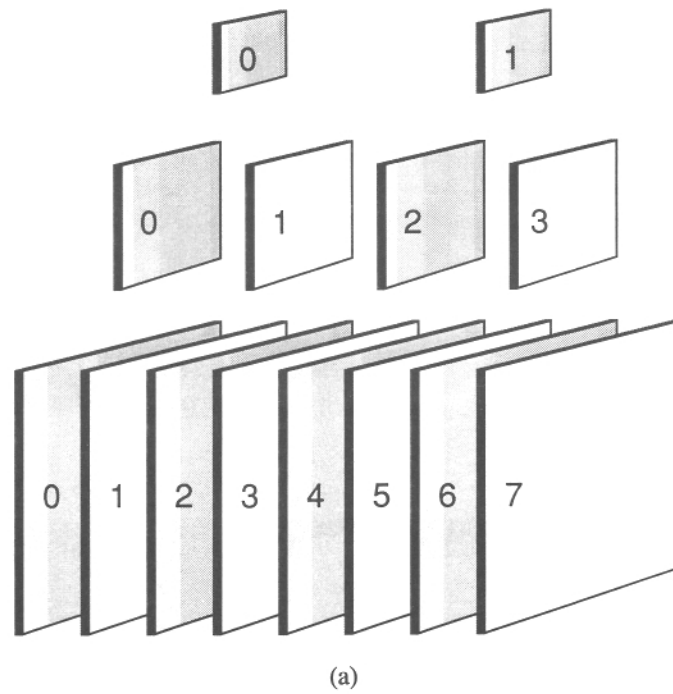


Figure 8. Three-dimensional spatiotemporal pyramid coding of video. Three resolutions are available, each lower resolution subsampled by eight (two in each dimension). (a) The reconstructed pyramid. Note that approximately one-half of the frames in the structure (shown as shaded) are spatially coded/interpolated. (b) Interpolation step in the pyramid: the spatial dimension is linearly interpolated, and the temporal dimension is motion compensated.



Figure 9. One frame from each level in the pyramid.

3.7. *A note on quantization and entropy coding*

While we have focused on the signal decomposition so far, compression is, of course, achieved by appropriate quantization of the various components and entropy coding of the quantized values. Usually, in designing a quantization scheme, one assumes that the various components are independent. This is approximately true, since the KLT, which decorrelates samples, will produce independent components if we assume Gaussian inputs. Other multiresolution schemes approximate the KLT, and thus the independence assumption is approximately valid. Under this assumption, the best quantization given a target bit rate will choose operating points on the individual rate-distortion curves that correspond to equal slope. This concept is underlying optimal bit allocation procedures suggested in the literature [39], [40].

However, such an optimum will be in the MSE sense, and it is well known that perceptually based quantization leads to better image quality. Therefore, perceptually designed quantizers, like the ones for DCT used in JPEG [41] or for subband coding [42], will lead to enhanced picture quality. Note that in pyramid coding, when the quantization feedback idea is used, the last quantizer can be well adapted to the human visual system. This is due to the fact that this last quantization is done directly on the picture, as opposed to some transformed representation. In particular, masking functions can be used so as to increase quantization in less visible areas, while using fine quantization in sensitive regions [6].

Finally, entropy coding based on Huffman or arithmetic coding is applied to remove any redundancy left by previous stages, but without adding any errors at this point.

Note that optimal quantization would require vector quantization (VQ) [43], but high-quality coding leads to very large codebooks. Thus, VQ has been more popular for high-compression applications. However, VQ can be modified with little suboptimality so as to be applicable in the high signal-to-noise ratio (SNR) context as well. In particular, hierarchically structured codebooks both reduce the complexity and allow for multiresolution representations (e.g., [33], [43]).

3.8. *Remarks*

Finite memory (FIR) or open-loop coding schemes were seen to be naturally suited for multiresolution decompositions. Therefore, they can be directly used for joint source-channel coding. No particular modification is required, since one simply assigns the components contributing more to the SNR into better protected channels for transmission. If the less important augmentation channels are lost, one naturally obtains a graceful degradation.

The simplicity of the FIR multiresolution schemes is counterbalanced by the fact that finite memory schemes often have poorer performance in terms of compression compared to their recursive counterparts.

4. **Multiresolution decompositions and recursive coding schemes**

In this section, we discuss the implementations of multiresolution decompositions in DPCM loops [44]. We focus on the simplest one-dimensional case first, and discuss possible

alternatives. Then we consider how these concepts can be used in the classic hybrid motion-compensated predictive DCT (HMCP-DCT) coder, which is the basis of many standard video coding algorithms (e.g., MPEG [43], [5], [45]). The HMCP-DCT coder is a DPCM loop over time, with a motion-based predictor, and the techniques discussed can be applied to this case with appropriate adjustments.

4.1. One-dimensional DPCM loops

There are various ways of including a multiresolution decomposition in a DPCM loop. Consider the one-dimensional case and a decomposition into a low-resolution part plus added detail. This could be obtained typically with a two-channel subband coder, with low-pass and high-pass filtering followed by subsampling by 2 (see Section 3.2 for a discussion of subband coding).

4.1.1. Multiresolution decomposition followed by independent DPCM loops. This is schematically shown in Figure 10(a). The advantages are the independence of both stages and of the two resolutions. But this independence can also lead to problems if there are independent errors that add up in the reconstruction. Also, the predictors being independent, no information is used across the resolutions, leading to a certain loss in performance.

4.1.2. Multiresolution decomposition in the DPCM loop. The prediction error is decomposed into multiresolution components. For example, the DCT or alternatively a subband decomposition is used as a multiresolution splitting of the prediction error. There are two possibilities:

1. The predicted value is based on all components of the error signal (see Figure 10(b)). In that case, the receiver also needs all components so as to track the transmitter accurately, and the multiresolution decomposition cannot be used for independent decoding at low-resolution or for joint source-channel coding (except in certain ideal cases, like perfect bandpass filters).
2. The predicted value is based only on the low-resolution part of the error signal (see Figure 10(c)). A decoder can now run at low resolution, and we thus have a true multiresolution decomposition with an independent compatible subchannel. It can thus also be used for joint source-channel coding. However, a certain loss in performance will occur, since there is no prediction of the high-frequency part.

4.1.3. Hybrid solution. Figure 10(d) shows a hybrid solution that combines features from the previous schemes. The predictive loop is based on the low-resolution part only, leading to the possibility of low-resolution decoding and joint source-channel decoding. The performance is enhanced by using a DPCM loop in the output of the high-frequency band, as well as side information from the low resolution.

4.1.4. Lossy DPCM loops. One way to achieve graceful degradation in a DPCM scheme in the presence of errors is to use robust DPCM loops. Such loops have been studied in DPCM coding of speech [44], where bit errors would otherwise lead to unacceptable

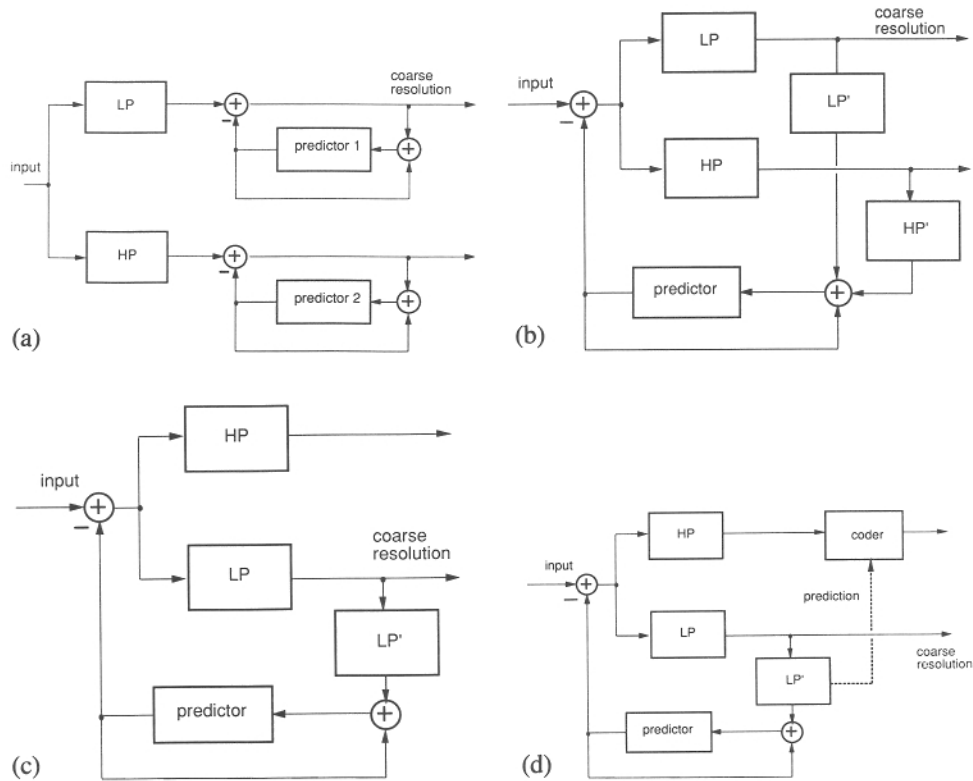


Figure 10. Multiresolution decomposition in DPCM loops. (a) Multiresolution decomposition followed by independent DPCM loop. (b) Multiresolution decomposition in the DPCM loop. The prediction error is decomposed into multiresolution components. (c) Only the low-resolution part is used in the DPCM loop, allowing decoding from the low resolution alone if needed. (d) The performance is enhanced by including prediction from the low resolution into the added resolution as well as within the added resolution.

errors. The idea is to make a “leaky” prediction, that is, an imperfect prediction, so that there is always a prediction error, which guarantees that errors will die away. The disadvantage is an increase of the prediction error at all times, thus less compression. Moreover, compatible low-resolution decoding based on this scheme would be quite poor, since the quality would be much below what can be achieved at the corresponding rate. Therefore, lossy DPCM loops are not well suited for compatibility purposes.

4.2. Hybrid motion-compensated predictive DCT coding and multiresolution decompositions

A typical hybrid motion-compensated predictive DCT (HMCP-DCT) coding scheme for video is shown in Figure 11. As can be seen, the DCT and IDCT are spatial domain operations and they cancel each other (quantization is not shown, but the DCT domain quantization can in principle be replaced by an equivalent space-domain quantization). Thus, we

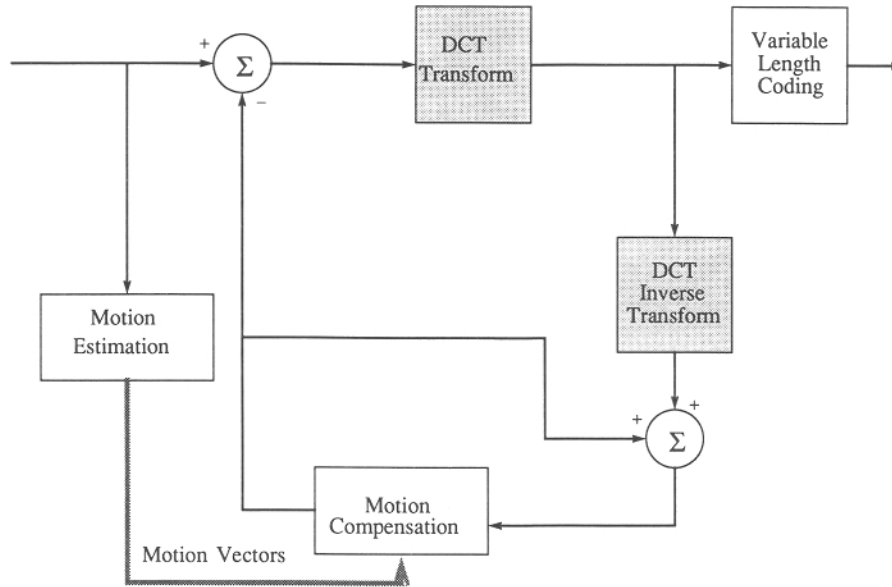


Figure 11. Hybrid motion-compensated predictive DCT coding of video. The DPCM loop is over the time dimension, and the predictor is based on motion vectors.

are left with a DPCM loop over time, with a prediction based on motion vectors. This generic video coding method can now be recast in the various schemes discussed above.

A difficulty is that now we are in a three-dimensional system and multiresolution will be typically over the spatial dimension while the DPCM loop is over time.

The first method, which is to start with a multiresolution decomposition followed by independent DPCM loops (HMCP-DCTs), was investigated in [24], [46], [47] using sub-band decomposition. A compatible HDTV/TV system using a similar decomposition was proposed in [48]. It was noted in [47] that the scheme was sensitive to errors in the sub-bands, which could add up after reconstruction. This problem is particularly pronounced if motion artifacts occur. Also, high bands are not suited for motion estimation, and thus, the motion from the lowest band is used in the other bands. There is an accuracy problem also, since a small offset will produce a poor prediction in the high bands (where mainly lines have to be predicted). Therefore, motion compensation is often only run on the lowest band.

In [24], [46], the DCT usually used to encode the prediction error is replaced by a sub-band decomposition, and we are in the case of Figure 10(b). This scheme behaves very similarly to the usual HMCP-DCT case, but one cannot decode a low-resolution version alone.

The scheme in Figure 10(c) has been used in packet video [49], [50] and is an example of *layered coding* schemes. The goal is to guarantee transmission of one part (the low resolution) while sending the augmentation channel in a nonprotected and error-prone channel. Thus, it is necessary to be able to decode based on the low-resolution alone. The DCT

was used as a spatial multiresolution decomposition, with lower DCT frequency coefficients chosen as a coarse spatial approximation. In HMCP-DCT coding of video, one splits the DCT coefficients of the prediction error into two parts (corresponding to low and high frequencies, respectively). While the low-frequency DCT coefficients are used in the prediction loop, the high-frequency coefficients are just intraframe coded. The loss in performance is minor, because the high frequencies are in any case difficult to predict accurately.

Finally, variations of the scheme depicted in Figure 10(d) can be used so that it is possible to decode a low-resolution video sequence based on the low-resolution channel only. This channel could also be better protected for transmission purposes. The added detail channel is now interframe coded, due to the added DPCM loop (which will probably use motion information from the low-pass channel), thus leading to enhanced performance over the scheme in Figure 10(c). An example of this idea is the MPEG video coding standard [4]. Instead of low and high resolution, one takes even and odd frames (or another combination of frames). A HMCP-DCT coder is run on one subset, independently of the other, so that it can be decoded independently. Then the other set is coded based on the first set, using frame interpolation. Note that this is more efficient than coding the second set independently. Another example of independent versus dependent coding in the interlaced television case is described in [51].

The idea of lossy DPCM loops can be applied to HMCP-DCT as well. However, errors will streak over several frames, and the suboptimal prediction will increase the bit rate needed to get a given quality and, as discussed earlier, compatible decoding will not be satisfactory.

4.3. Remarks

Another way to look at the variations of subband coding together with motion-compensated DPCM loops is the concept of an *inner* and an *outer* code (a notion borrowed from error correction coding [52]). For example, in HMCP-DCT coding the inner coding is the DCT and the outer coding is DPCM, while in the subband decomposition followed by the HMCP-DCT case, the outer code is subband decomposition while the inner one is DPCM. Again, the situation is complex because of the three dimensions involved, allowing various operations to coexist in various dimensions.

We should point out that the "typical" video scene is not a general three-dimensional signal, but is formed by an image varying slowly over time as described by the motion field. In other words, motion "mixes" spatial and temporal frequencies [6]. Therefore, the MR decomposition over space induces a similar decomposition over time: high frequencies in the DPCM loop are also reduced in the low (spatial) resolution channel.

In conclusion, if a low-resolution version has to be decodable independently, the form of the prediction is restricted. In particular, the low resolution can only be predicted from the low resolution, while the added resolution can be predicted from both, as shown schematically in Figure 12. The loss in performance is due to the lack of prediction from added resolution to low resolution, which is usually small anyway.

Current:

Predicted:

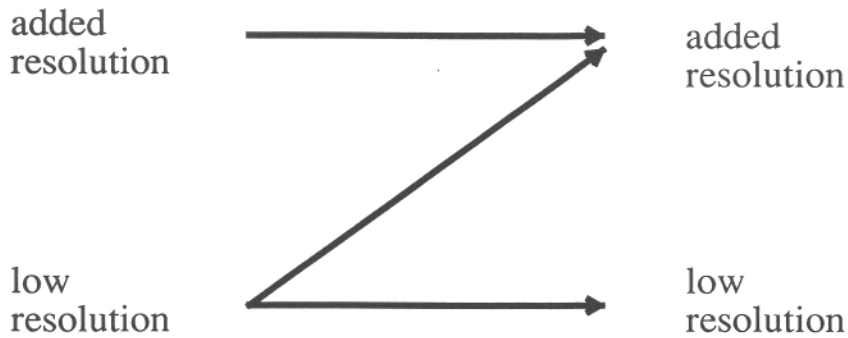


Figure 12. Restricted form of prediction if independent low-resolution decoding has to be possible.

5. Compatibility and transcoding issues

The question of compatibility of a new standard with previous standards is a much discussed topic around HDTV. However, while it is conceptually clear what one means by such compatibility, the crucial details are often left out. For example, true compatibility would only be possible if the lower resolution standard lives on a sublattice of the high-resolution standard's sampling lattice. This is however seldom the case in the new proposed standards with respect to the current television standard. Therefore, a more adequate notion is that of *easy transcodability*: can the lower resolution standard be easily obtained from a subchannel of the new standard? For example, in coding of interlaced HDTV, an interlaced subchannel of NTSC quality can be derived in the scheme proposed in [51], [53]. In the pyramid discussed above, an intermediate interlaced sequence similar to NTSC could be interpolated based on motion vectors.

Then, with the current MPEG standardization effort for digital video coding [4], compatibility will usually mean that a higher resolution coding algorithm can be built from components used in the lower resolution coding ("components" is meant both in a conceptual sense as well as in the context of hardware components). Such upward compatibility has obvious technological and economic advantages.

Finally, the issue of aspect ratio (which creates an obvious "incompatibility") is usually dealt with the so-called letter-box format (the larger aspect ratio, like 16/9, being shown within the smaller one, like 4/3).

6. Multiresolution transmission for digital broadcast

Among all transport mechanisms for video, perhaps the most challenging one is broadcast over the terrestrial channel, which is plagued by many distortions (cross-interference and multipath for example). From a video coding point of view, the two key requirements are (i) robustness to channel errors and (ii) graceful rather than abrupt degradation under adverse conditions. These requirements play an important role in the choice of source and channel coding methods. After discussing these issues, we will briefly describe a system using multiresolution source and channel coding for digital broadcast of HDTV [54].

6.1. Robustness to channel errors

The ability of a coding system to sustain normal operation facing occasional channel errors is determined by the design of both source and channel coding systems. Robustness requires high error protection for sensitive data and frequent restart or resynchronization of recursive portions of the algorithms. Variable-length codes have to be modified and/or include resynchronization sequences to avoid long stretches of errors after a single bit error. For example, MPEG-like algorithms restart their HMCP-DCT loop every 15 frames (or 0.5 s), making them in essence finite memory schemes. If the data is hierarchically decomposed, it is easy to protect the basic components (like the low-resolution sequences and motion vectors) more heavily than the added details.

6.2. Digital broadcast systems

Hierarchical or multiresolution decompositions blend well with channel coding in what we call *multiresolution transmission*. While a traditional digital transmission system has a noise threshold over which its operation degrades very rapidly (so-called threshold effect), it is possible to use multiresolution decomposition of the source to achieve a graceful degradation as the carrier-to-noise ratio (CNR) decreases (see Figure 13). This is a typical example of joint source-channel coding, which is believed to be better suited for digital broadcast situations than traditional digital transmission.

A conventional digital transmission scheme is designed with the worst channel in mind, i.e., the fringe area, while receivers closer to the transmitter have no additional benefit (a fact in contradiction with the typical geographical distribution of receivers). This is what Schreiber [55], [56] calls the *wasted bandwidth* in digital broadcast systems, since the additional CNR closer to the transmitter is not used to transmit more bits.

We share this point of view as we think that current digital broadcast proposals are really designed as point-to-point transmission systems (namely emitter to fringe) rather than broadcast systems. But results from information theory on multiuser systems [57], [58] indicate that better performance can be achieved by designing the transmission system specifically for the broadcast channel. Interestingly, the way to achieve better performance is by using "multiresolution codes" (this is not the name used in [57], but it is intuitive, as we shall see).

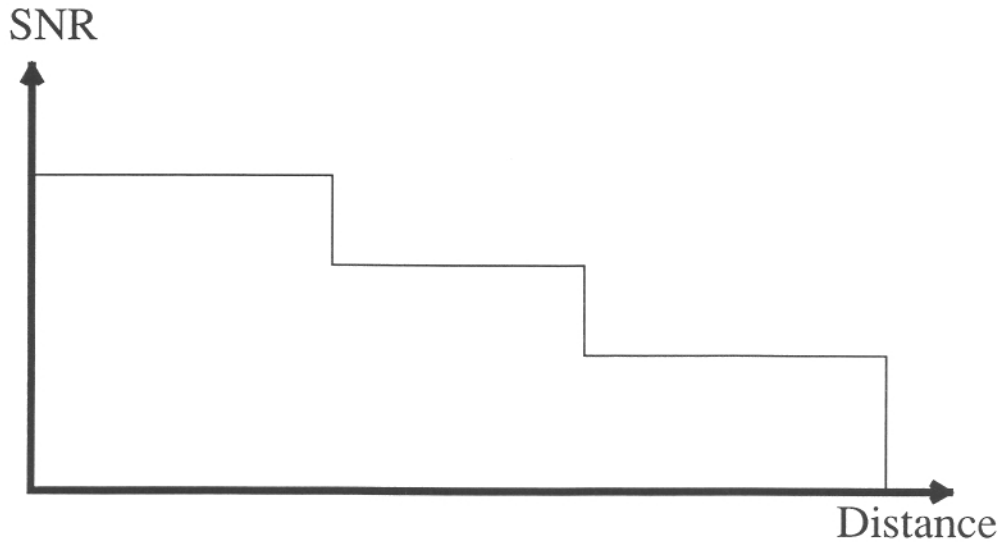


Figure 13. Multiresolution transmission yielding graceful degradation with decreasing CNR (or increasing distance).

6.3. Multiresolution transmission

A better digital broadcast system can be designed by matching a multiresolution source coding with a multiresolution channel coding. Such a system is described in [54] and is outlined here.

Assume that we have a source coding scheme like the spatiotemporal pyramid discussed in Section 3.6 or some other multiresolution source decomposition where the low resolution can be decoded independently. A first simple approach would be to use separate frequency channels and possibly different modulation schemes (such as 4 QAM and 16 QAM) for the coarse and the fine resolution (we assume constant power). Obviously, the coarse version can now be received further out. In areas where both coarse and fine resolutions are decodable, the full-resolution picture will be received, and under adverse conditions (e.g., heavy rain) the picture will not be completely lost but will simply fall back to the coarser resolution.

Better coding efficiency can be achieved by implementing a multiresolution modulation scheme matched to the source coding. For example, assume the ratio of the rates for coarse and fine resolutions is unity. Then, a nonuniform 16-QAM modulation (where each symbol consists of 4 bits) can be designed so as to carry 2 bits out to low CNRs, whereas the full 4 bits can be recovered at higher CNRs. Such a nonuniform QAM constellation is shown in Figure 14, and the parameter λ is the ratio of the small and the large distance between points ($\lambda = 0$ and $\lambda = 1$ correspond to 4 and 16 QAM, respectively). Therefore, there are now two regions of coverage: a small one for high-quality reception and a larger one for lower quality reception. For more details, we refer the reader to [54], where a complete design of MR source coding and transmission is demonstrated. Transmission is packet based and uses the spatiotemporal source decomposition described in Section 3.6

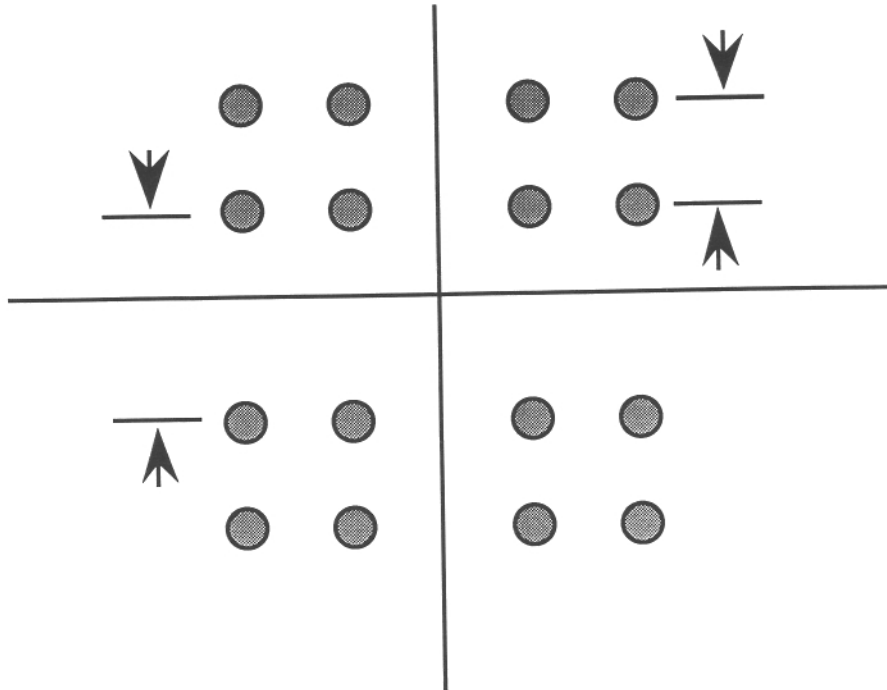


Figure 14. Multiresolution QAM constellation. At high CNRs, 4 bits can be decoded, while at low CNRs, only the two most significant bits (indicating one of the four quadrants) are decoded.

and a multiresolution 64-QAM constellation. The trade-offs in coverage and quality are summarized in Figure 15, and the effects of channel errors are shown in Figure 16 (both are from [54]).

The use of nonuniform QAM as well as hybrid analog and digital transmission (which achieves a similar behavior) has also been proposed by Schreiber [55], [56]. However, purely digital MR modulation allows using more sophisticated entropy coding techniques, whereas in the hybrid case one is limited to coding a set of samples corresponding to actual amplitudes in a memoryless fashion.

While other approaches are possible (using different compression and modulation methods), we believe that this is the first demonstration of a complete system for digital broadcast of HDTV which does not exhibit a threshold effect and uses a transmission scheme specifically designed for the broadcast situation.

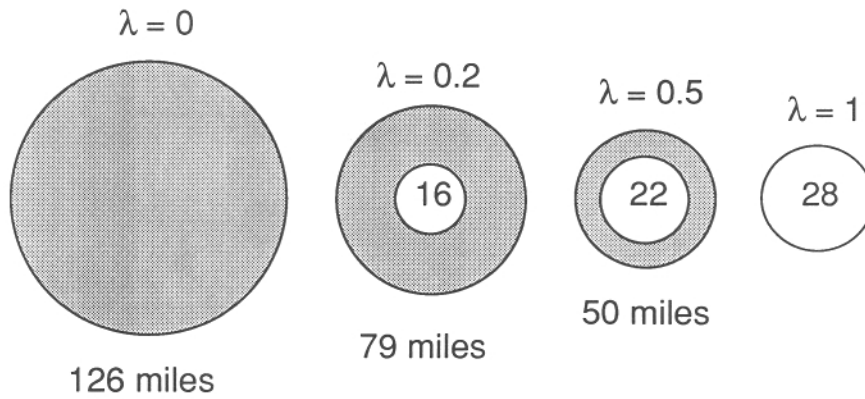


Figure 15. Trade-off between sizes of the coarse resolution coverage and the fine resolution coverage. The SNR versus CNR is plotted for various values of the design parameters λ . At $\lambda = 1$, we have a single high-resolution coverage of a small area, while at $\lambda = 0$, the single, coarse resolution coverage is over a large area. In between, we can trade one for the other.

7. Conclusion

The role of multiresolution techniques in source coding and channel coding for digital video has been reviewed. The major coding schemes and their suitability for multiresolution decompositions have been considered. It has been argued that the multiresolution framework is a natural paradigm to solve a number of problems arising in coding and transmission of digital video. In particular, the matching of a multiresolution decomposition of the source with a multiresolution digital transmission is seen as an interesting joint source-channel coding suited for digital broadcast.

Acknowledgments

The authors would like to thank Professor W. Schreiber for pointing out the importance of spectrum efficiency in television broadcast, and K. Ramchandran for fruitful collaboration and providing simulation results for multiresolution transmission.

References

1. F. Gray, J.W. Horton, and C.R. Mathes, "The Production and Utilization of Television Signals," *Bell systems Technical Journal*, vol. 6, 1927, p. 560.
2. D. Gabor, "Theory of Communication," *Journal of the IEE*, vol. 93, 1946, pp. 429-457.
3. D. Anastassiou and M. Vetterli, "Television by the Bit," *IEEE Circuits and Devices Magazine*, vol. 7, 1991, pp. 16-21.

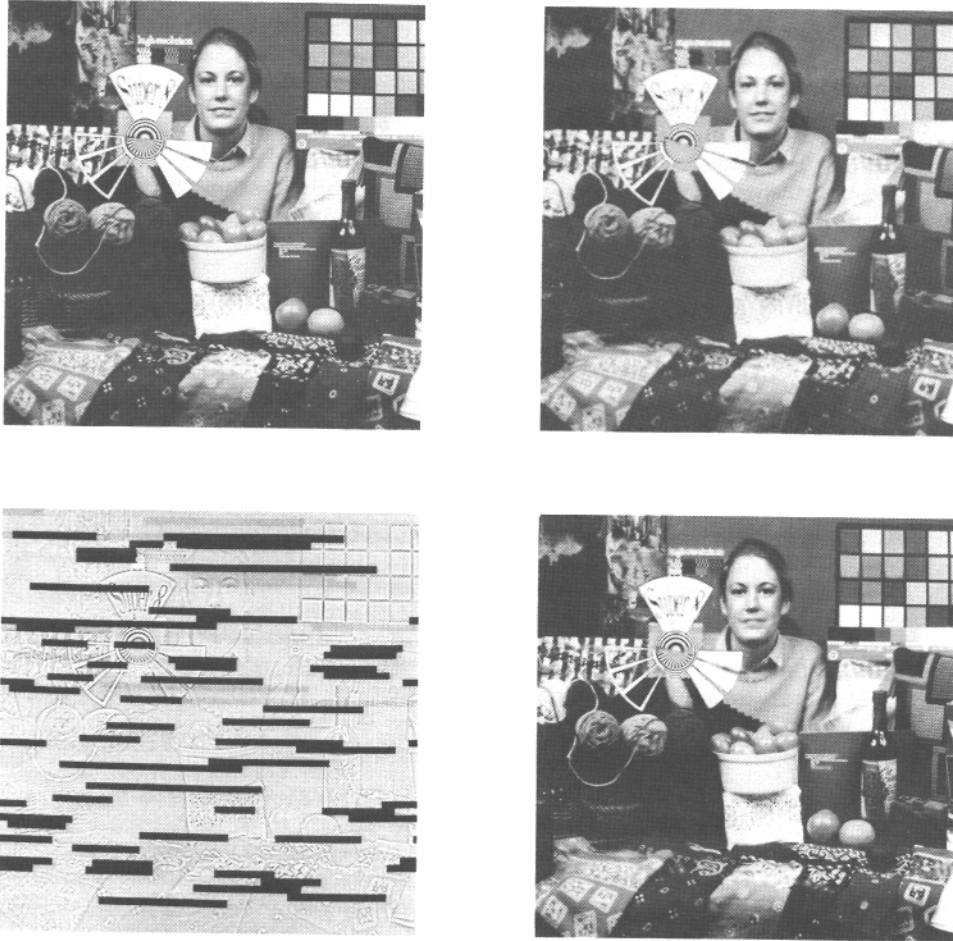


Figure 16. MR digital broadcast. (a) Full-resolution picture. (b) Decoded coarse resolution. (c) 15% packet loss in the fine resolution picture. (d) Decoded picture with the packet loss. Error concealment techniques (not shown) can be used to render virtually all errors invisible.

4. Motion Picture Expert Group, ISO/IEC JTC1/SC2/WG8, CCITT SGVIII. "Coded Representation of Picture and Audio Information," MPEG video simulation model two, 1990.
5. D. LeGall, "MPEG, a Video Compression Standard for Multimedia Applications," *Transactions of the Association of Computing Machinery*, vol. 34, 1991, pp. 46-58.
6. K.M. Uz, M. Vetterli, and D. LeGall, "Interpolative Multiresolution Coding of Advanced Television with Compatible Subchannels," *IEEE Transactions on Circuits and Systems for Video Technology*, Special Issue on Signal Processing for Advanced Television, vol. 1, pp. 86-99, 1991.
7. F. Campbell and D. Green, "Optical and Retina Factors Affecting Visual Resolution," *Journal of Physiology*, vol. 181, 1965, pp. 576-593.
8. S. Mallat, "Multifrequency Channel Decompositions of Images and Wavelet Models," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, 1989, pp. 2091-2110.

9. P.J. Burt and E.H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Computers*, vol. 31, 1983, pp. 532-540.
10. W.H.R. Equitz, "Successive Refinement of Information," Ph.D. Thesis, Stanford University, 1989.
11. A. Rosenfeld, ed., *Multiresolution Techniques in Computer Vision*, New York: Springer, 1980.
12. I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Communications in Pure and Applied Mathematics*, vol. 41, 1988, pp. 909-996.
13. S. Mallat, "A Theory of Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, 1989, pp. 674-693.
14. O. Rioul, "A Unifying Multiresolution Theory for the Discrete Wavelet Transform, Regular Filter Banks and Pyramid Transforms," to appear, *IEEE Transactions in Signal Processing*.
15. A.K. Jain, *Fundamentals of Image Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
16. P.P. Vaidyanathan, "Quadrature Mirror Filter Banks, M-band Extensions and Perfect-Reconstruction Technique," *IEEE Acoustics, Speech and Signal Processing Magazine*, vol. 4, 1987, pp. 4-20.
17. P. Cassereau, "A New Class of Optimal Unitary Transforms for Image Processing," S.M. Thesis, Department of Electrical Engineering and Computer Sciences, Massachusetts Institute of Technology, 1985.
18. M. Vetterli and D. LeGall, "Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, 1989, pp. 1057-1071.
19. M. Vetterli, "Multirate Filter Banks for Subband Coding," in *Subband Image Coding*, (J.W. Woods, ed.), Boston, Kluwer, 1990, pp. 43-100.
20. M. Vetterli, "Multi-Dimensional Sub-Band Coding: Some Theory and Algorithms," *Signal Processing*, vol. 6, 1984, pp. 97-112.
21. J.W. Woods and S.D. O'Neil, "Sub-band Coding of Images," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, 1986, pp. 1278-1288.
22. G. Karlsson and M. Vetterli, "Three Dimensional Sub-band Coding of Video," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1988, pp. 1100-1103.
23. W.F. Schreiber and A. Lippman, "Single Channel HDTV Systems, Compatible and Noncompatible," in *Signal Processing of HDTV*, (L. Chiariglione, ed.) Amsterdam: North-Holland, 1988.
24. J.W. Woods and T. Naveen, "Subband Encoding of Video Sequences," *Proceedings of the SPIE Conference on Visual Communications and Image Processing*, 1989, pp. 724-732.
25. G. Karlsson and M. Vetterli, "Packet Video and Its Integration into the Network Architecture," *IEEE Journal on Selected Areas in Communications*, Special Issue on Packet Speech and Video, vol. 7, 1989, pp. 739-751.
26. G. Karlsson and M. Vetterli, "Theory of Two-Dimensional Multirate Filter Banks," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, 1990, pp. 925-937.
27. E. Viscito and J. Allebach, "The Analysis and Design of Multidimensional FIR Perfect Reconstruction Filter Banks for Arbitrary Sampling Lattices," *IEEE Transactions on Circuits and Systems*, vol. 38, 1991, pp. 29-42.
28. M. Vetterli, J. Kovačević, and D. LeGall, "Perfect Reconstruction Filter Banks for HDTV Representation and Coding," *Image Communication*, vol. 2, 1990, pp. 349-364.
29. M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image Coding Using Vector Quantization in the Wavelet Transform Domain," *Proceedings of the IEEE ICASSP*, Albuquerque, NM, 1990, pp. 2297-2300.
30. M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design," to appear, *IEEE Transactions on Signal Processing*, Sept. 1992.
31. R.R. Coifman, Y. Meyer, S. Quake, and M.V. Wickerhauser, "Signal Processing and Compression with Wavelet Packets," Department of Mathematics, Yale University, preprint, 1990.
32. M.V. Wickerhauser, "Acoustic Signal Compression with Wave Packets," Department of Mathematics, Yale University, preprint, 1989.
33. P.A. Chou, T. Lookabaugh, and R.M. Gray, "Optimal Pruning with Applications to Tree-Structured Source Coding and Modeling," *IEEE Transactions on Information Theory*, vol. 35, 1989, pp. 299-315.
34. E.A. Riskin and R.M. Gray, "A Greedy Tree Growing Algorithm for the Design of Variable Rate Vector Quantizers," in *Proceedings of the Picture Coding Symposium*, Boston, 1990, pp. 11.4.1-11.4.3.
35. E.A. Riskin, "Optimal Bit Allocation via the Generalized BFOS Algorithm," *IEEE Transactions on Information Theory*, vol. 37, 1991, pp. 400-402.
36. D. Anastassiou, "Generalized Three-Dimensional Pyramid Coding for HDTV Using Nonlinear Interpolation," in *Proceedings of the Picture Coding Symposium*, Cambridge, MA, 1990, pp. 1.2-1-1.2-2.

37. M. Bierling, "Displacement Estimation by Hierarchical Blockmatching," *SPIE Conference on Visual Communications and Image Processing*, Boston, 1988, pp. 942-951.
38. K.M. Uz, M. Vetterli, and D. LeGall, "A Multiresolution Approach to Motion Estimation and Interpolation with Application to Coding of Digital HDTV," *Proceedings of the IEEE ISCAS*, New Orleans, 1990, pp. 1298-1301.
39. Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, 1988, pp. 1445-1453.
40. P.H. Westerink, J. Biemond, and D.E. Boekee, "An Optimal Bit Allocation Algorithm for Subband Coding," *Proceedings of the ICASSP-88*, New York, 1988, pp. 757-760.
41. Joint Photographic Expert Group, ISO/IEC JTCl/SC2/WG8, CCITT SGVIII. JPEG technical specification, revision 5, January 1990.
42. R.J. Safranek and J.D. Johnston, "A Perceptually Tuned Subband Image Coder with Image Dependent Quantization and Postquantization Data Compression," *Proceedings of the ICASSP-89*, Glasgow, 1989, pp. 1945-1948.
43. A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Boston: Kluwer, 1992.
44. N. Jayant and P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
45. A.N. Netravali and B.G. Haskell, *Digital Pictures: Representation and Compression*, New York: Plenum Press, 1988.
46. H. Gharavi, "Subband Coding of Video Signals," in *Subband Image Coding*, (J.W. Woods, ed.), Boston: Kluwer, 1990.
47. H.-M. Hang, R. Leonardi, B.G. Haskell, R.L. Schmidt, H. Bheda, and J. Othmer, "Digital HDTV Compression at 44 mbps Using Parallel Motion-Compensated Transform Coders," in *Proceedings of the SPIE Conference on Visual Communications and Image Processing*, vol. 1360, Lausanne, Switzerland, 1990, pp. 1756-1772.
48. J. Mau "HDTV/TV Compatible Codec with PQMF Filter Bank," *Proceedings of the Fourth International Workshop on HDTV*, Torino, Italy, 1991.
49. K. Shinamura, Y. Hayashi, and F. Kishino, "Variable Bitrate Coding Capable of Compensating for Packet Loss," *Proceedings of the SPIE Conference on Visual Communications and Image Processing*, 1988, pp. 991-998.
50. M. Ghanbari, "An Adaptive Video Codec for ATM Networks," *Proceedings of the Third International Workshop on Packet Video*, Morristown, NJ, 1990.
51. F.-M. Wang and D. Anastassiou, "High-Quality Coding of the Even Fields Based on the Odd Fields of Interlaced Video Sequences," *IEEE Transactions on Circuits and Systems*, vol. 38, 1991, pp. 140-142.
52. R.E. Blahut, *Theory and Practice of Error Control Codes*, Reading, MA: Addison-Wesley, 1984.
53. T.H. Chiang, R. Mokry, A. Eleftheriadis, W.B. Yu, and D. Anastassiou, "Compatible Coding of Digital Interlaced HDTV Using Prediction of the Even Fields from the Odd Fields," *Proceedings of the Fourth International HDTV Workshop*, Torino, Italy, 1991.
54. K.M. Uz, K. Ramchandran, and M. Vetterli, "Multiresolution Source and Channel Coding for Digital Broadcast of HDTV," *Proceedings of the Fourth International Workshop on HDTV*, Torino, Italy, 1991.
55. W.F. Schreiber, "Considerations in the Design of HDTV Systems for Terrestrial Broadcasting," *Electronic Imaging '90*, Boston, 1990.
56. W.F. Schreiber, "All-Digital HDTV Terrestrial Broadcasting in the US: Some Problems and Possible Solutions," presented at *Workshop on Advanced Television*, ENST Paris, 1991.
57. T. Cover, "Broadcast Channels," *IEEE Transactions on Information Theory*, vol. IT-18, 1972, pp. 2-14.
58. A. ElGamal and T. Cover, "Multiuser Information Theory," *Proceedings of the IEEE*, vol. 68, 1980, pp. 1466-1483.

