Sub-band source coding for HDTV

J. Mau (CCETT)

E. Bourguignat (CCETT)

H. Amor (Bosch)

After a presentation of the main principles involved in sub-band coding for HDTV, the authors describe two practical implementations.

The first codec is designed for the digital transmission of HDTV signals in a 140-Mbit/s channel, such as the switched broadband network (VBN) of the Deutsche Bundespost. The pictures have an aspect ratio of 16:9, 1250 lines per frame, 50-Hz field rate and 2:1 interlacing, and after elimination of the blanking intervals a bit-rate reduction factor of 8 has to be achieved.

The second example is being developed for the digital distribution of high-definition progressive (HDP) video signals of 16:9 aspect ratio in a 140-Mbit/s channel such as the future broadband-ISDN conveying network signals in asynchronous transfer mode (ATM). The HDP pictures have 1920x1152x50 active pixels per second (in 50-Hz countries), and a bit-rate reduction factor of 15 is necessary.

One important consideration regarding the suitability of sub-band techniques for HDTV source coding is compatibility between a number of decoding schemes offering a range of picture definitions but all using a common coding strategy.

1. *Introduction*

A number of very effective algorithms based on the discrete cosine transform (DCT) were specified during the 1980s, for applications including the coding of stationary and moving pictures. The results are clearly superior to those achieved using the differential pulse-code modulation (DPCM) techniques more-commonly under study at that time.

Now, in the 1990s, several approaches are being explored in the picture-coding laboratories as the search for the elusive “optimum algorithm” continues. Sub-band or hierarchical techniques, neuronal approaches and 3D object coding are just a few of them. It is not intended to describe them all here because their respective applicability and advantage for television coding have only rarely been shown. The exception is sub-band splitting which has been demonstrated as clearly applicable for television applications and after a brief explanation of the principles, two systems will be described.

2. Sub-band coding schemes

The DCT approach in image coding can be considered as a special case of the sub-band approach. In effect, in both cases, the source picture is described in terms of frequency coefficients. One main difference is that in sub-band coding the source signal is still described in the "original" domain, whereas with the DCT the signal is transformed to a frequency domain. Furthermore, the filtering for band-splitting in sub-band coding is performed on the whole picture range whilst block-oriented processing is performed for DCT. Blocking effects, which are inherent to DCT - at least if high data compression ratios are required - can be avoided by sub-band coding.

2.1. Filtering and sub-band splitting

Expressed simply, sub-band splitting is a process which decomposes the source picture into several frequency bands by means of appropriate filters. Two main principles may be defined, referred to as hierarchical and parallel.

Using a *hierarchical* approach, a cascade of half-band filters are applied to the source picture in succession; this doubles the number of sub-bands at each step. After each step, the signal is down-sampled by a factor of two. The mathematical conditions which must be respected if perfect (or close-to-perfect) reconstruction is to be achieved yield two main solutions called quadrature mirror filters (QMF) and conjugate quadrature filters (CQF). Each solution has disadvantages, the main ones being the non-linearity phase for CQF (longer than 2) and amplitude distortion in QMF. Whichever filter family is actually used, the hierarchical approach is generally "cleaner" than conventional DCT as regards the Fourier spectrum, but the freedom in sub-band splitting is limited by the complexity resulting from the number of processing steps involved.

The *parallel* approach involves the direct, global construction of all the sub-bands. The DCT may be regarded as coming in this family since the transformation provides the coefficients belonging to the different sub-bands. Once the block-based transform has been carried out (on a block of 8x8 pixels, for example), 64 sub-bands may be constituted by putting together the homologous coefficients of all the blocks. This gives 64 sub-bands with a decimation factor of 8 in both directions, vertical and horizontal. Although this DCT technique does not allow any improvement in the conventional DCT approach, it can nonetheless be considered as a reference for comparisons between systems. It may be noted that so-called "lapped transforms" are currently being proposed to reduce the edge effects in block-based filtering.

Another parallel solution uses so-called "pseudo quadrature mirror filters" (PQMF). Sub-band splitting using the PQMF technique is achieved using a filter bank which is defined on the basis of a prototype low-pass filter. The various components of the filter bank are derived by modulation of the prototype filter. The term "pseudo" denotes the fact that the mathematical conditions needed in order to obtain perfect reconstitution are fulfilled only approximately. The PQMF technique is outlined in *Section 3.2.2.* and a full presentation of the theory will be found in [1].

2.2. Bit-rate reduction

After the sub-band splitting has been performed, bit-rate reduction must be applied.

The coding scheme may be defined on the basis of a block structure, in which the corresponding coefficients in each sub-band are selected, or on an overall sub-band basis. The block-structure approach is described elsewhere [2], and the following observations concern the sub-band approach.

The simplest way to reduce the bit-rate is to limit the accuracy and the dynamic range of the transmitted coefficients by quantization (pulse code modulation). Different quantizing laws can be defined, both non-linear and linear, for each sub-band. The aim is to optimize the quantization of the coefficients as regards the visual characteristics of the complete system.

Next in complexity is the use of differential pulse code modulation (DPCM). In fact, the spatial correlations are only useful in the lower-frequency sub-bands, but the temporal prediction is as efficient as for the conventional DCT scheme. DCT coding as designed for conventional-definition television (CDTV) is sometimes used to process the lowest-frequency band of an HDTV signal, as this has advantages with regard to compatibility. In effect, the sub-band technique is especially well-adapted to the compatible transmission of several levels of quality, or standards. The use of conventional schemes in the lower bands allows compatible reception of high and low-definition pictures.

The degree of bit-rate reduction to be applied in each sub-band remains to be studied in detail, and it may be expected that a judicious choice in this domain will lead to further improvements in sub-band coding efficiency.

2.3. *Motion estimation and compensation*

To reduce the bit-rate needed to convey a given quality (or, conversely, to increase the quality achieved with a given bit-rate), temporal correlation in the picture must be exploited. This generally implies the use of motion compensation which is exactly similar, in principle, to that applied in DCT coding of conventional television.

The motion estimation may be calculated before or after the sub-band splitting, although it is certainly simpler if applied to the source picture because the result is then applicable to all the sub-bands. If motion vector estimation is performed after splitting, the calculation should be done mainly on the basis of the content of the lower-frequency sub-bands, although the higher-frequency sub-bands must also be taken into account, especially if high accuracy is being sought (half or quarter pixel).

Special care is needed in the case of a compatible system because the compatible decoder which uses only the low-frequency sub-bands will not be able to calculate the same prediction values as a coder in which all the sub-bands are available. Differences between the predictions in the coder and decoder cause increasing divergence between the reconstructed picture from the decoder and the source picture at the coder. Such divergence will increase until the equipment is reset, for example with a full-frame intra refresh. Several solutions to this problem have been proposed, and one possibility involves performing two motion compensations and two predictions in the coder, to satisfy compatible and high-definition decoders.

3. *Description of practical sub-band systems*

3.1. *High Definition Video (HDV) sub-band codec*

3.1.1. *Introduction*

The HDV codec* is being developed for the digital transmission of high-definition video signals, together with digital audio and additional data, at a atotal bit-rate of 140 Mbit/s. The basic components of the codec are:

- PCM video codec;
- image processor;
- video buffer including multiplexing/demultiplexing functions and buffer control;
- audio codec;
- channel multiplexer/demultiplexer.

The present description will be confined to the video parts of the transmission equipment, shown schematically in *Fig. 1*.

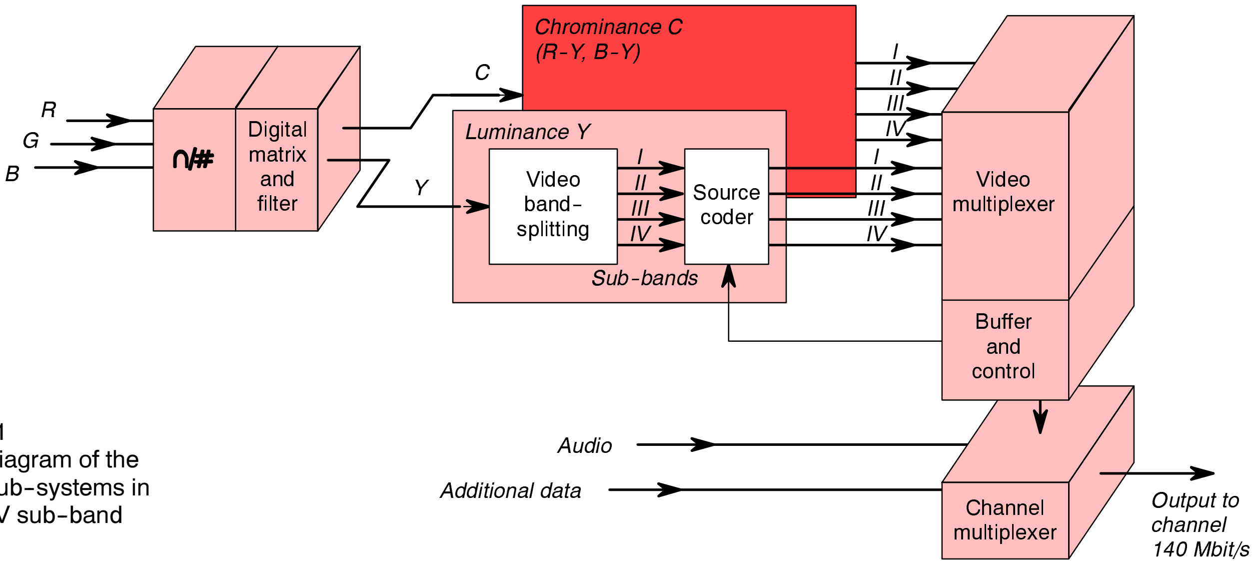


Figure 1
Block diagram of the video sub-systems in the HDV sub-band coder.

---

* The HDV codec is being developed in Germany by a consortium of:

- Deutsche Bundespost Research Institute, Darmstadt
- Robert Bosch Research Institute, Hildesheim
- University of Hanover, Telecommunications and Information Processing Institute
- Standard Electric Lorenz (SEL) Research Centre, Stuttgart.

The input to the HDV codec is an HDTV signal in accordance with the proposed European standard established by the Eureka EU95 project. The picture has an aspect ratio of 16:9, 1250 lines per frame (1152 active lines), a field rate of 50 Hz and 2:1 interlacing. The PCM video coder takes red, green and blue (RGB) primary input signals, matrixes these to give a luminance (Y) and two colour-difference (R-Y, B-Y) signals, and digitizes these in accordance with the EU95 proposal. The Y signal is sampled at 72 MHz and the R-Y and B-Y signals are sampled at 36 MHz each. Each sample is linearly quantized and binary coded to 8-bit resolution. The gross bit-rate at the output of the PCM coder is 1.152 Gbit/s.

The final bit-rate after bit-rate reduction should ideally be 122 Mbit/s. Even after the blanking intervals of the source picture have been eliminated, a reduction factor of 8 remains to be achieved by source coding, without impairing the high picture quality. The HDV codec uses a very promising sub-band coding scheme for this purpose.

3.1.2. *Filtering and sub-band splitting*

The PCM video signals (Y, R-Y, B-Y) are split by vertical and horizontal, low and high-pass filters to give four bands denoted LL, LH, HL, HH (see *Fig. 2*). In order to avoid an increase in the bit-rate above that of the source signal, maximal decimation filters are used. Intensive investigations have shown that for this purpose, and with due consideration for the corresponding investment in hardware, separable quadrature mirror filters (QMF) with 10 taps in the vertical direction and 14 taps in the horizontal direction, are optimal.

The decomposition of the HDV signal into these four sub-bands has several advantages, including:

- The LL band can be designed to fit a channel with "conventional" resolution, such as that defined in CCIR Recommendation 601.
- The four components differ in their statistical properties and their "information" content. Appropriate redundancy and irrelevancy reduction methods can be applied independently to each component, to achieve the best quality of image reconstruction for a given degree of bit-rate reduction.
- The four components can be processed in a fully-parallel hardware system using available and well-suited technologies.

3.1.3. *Bit-rate reduction*

The source coder which performs the bit-rate reduction is shown schematically in *Fig. 3*.

The LL band - the most important one - is coded efficiently using a motion-compensated hybrid coding scheme. A motion-compensation prediction is calculated for each 8x8 block, on the basis of information from the previous coded and decoded frames. This applies a full-search block matching process which is accurate to one pixel. After a comparison between the original block and its prediction, one of them is selected for further processing. This processing involves a discrete cosine transform, predictive coding of the zero-frequency term, adaptive quantizing and run-length coding of the non-zero-frequency terms, followed, finally, by variable-length coding (VLC).

The hardware needed for the LL band is quite complex, but the other three (higher-frequency) sub-bands are processed very simply by adaptive quantization, run-length coding and variable-length coding. Depending on the nature of the expected orientation of details in these sub-bands, an appropriate scanning scheme is applied independently to each sub-band, in order to maximize the length of zero runs and, hence, to improve the efficiency of the run-length coding. Adaptive quantization not only serves to reduce the coding effort needed but is also a powerful tool for controlling the data rate.

The data rate after the variable-length coder is of course variable, so in order to deliver the data to a channel with constant bit-rate the data stream at the coder output is "gathered" in a buffer. The data is read out of this buffer at a constant rate. The buffer and its controller (which avoids buffer over and under-flow) have a major influence on overall

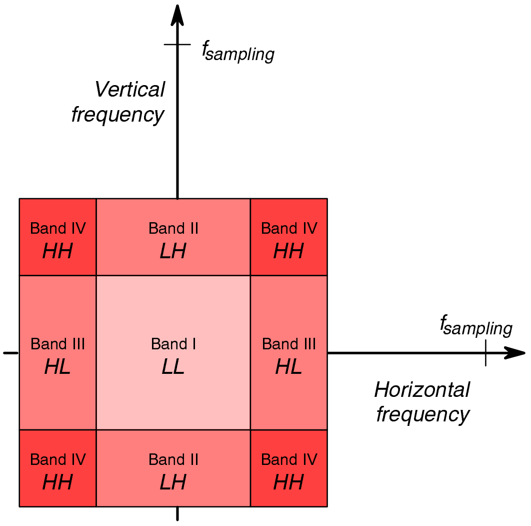


Figure 2
High- and low-pass filtering in the vertical and horizontal directions, to derive four sub-bands.

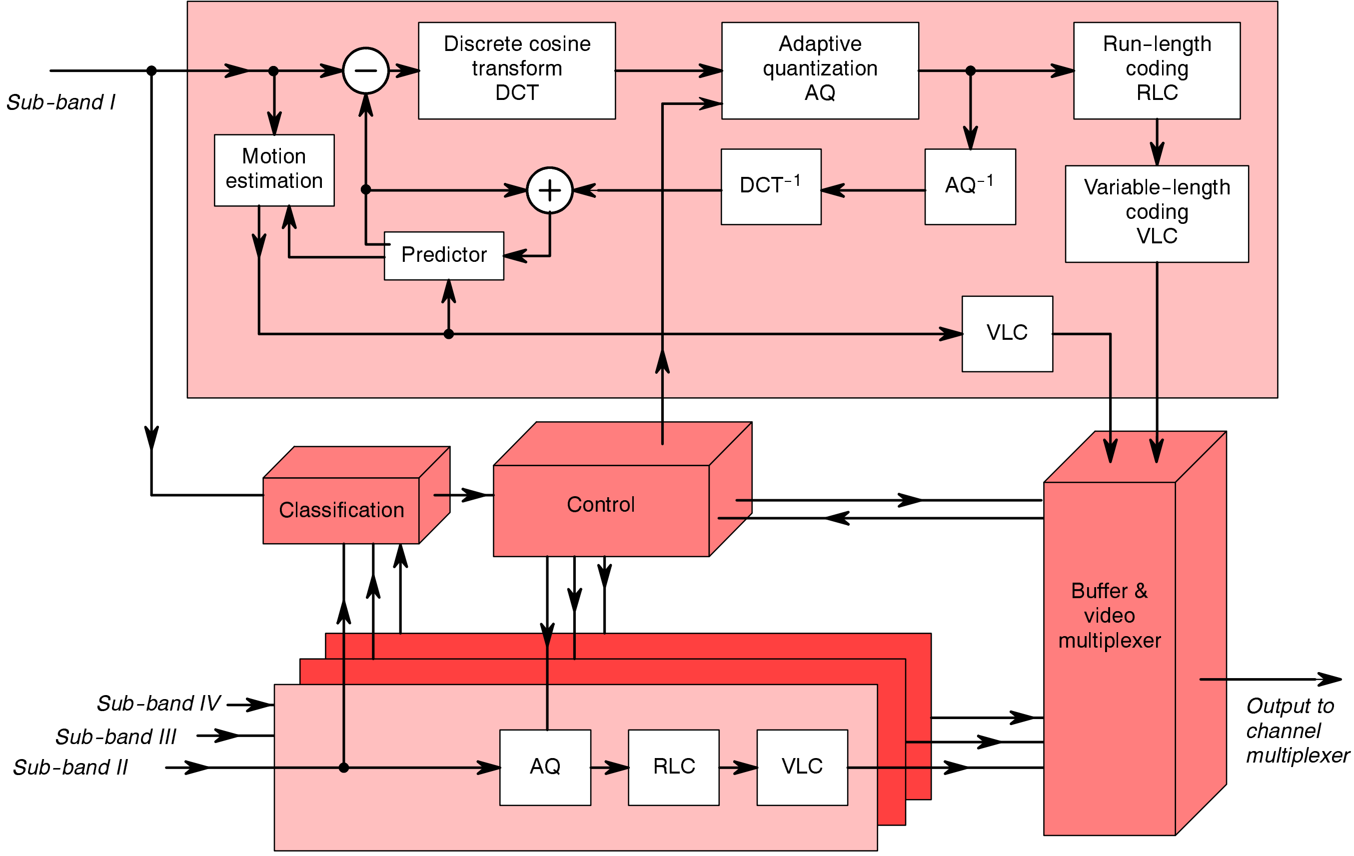


Figure 3
Block diagram of the HDV source coder.

system performance so very careful design is necessary in this area.

3.1.4. Results

A hardware-oriented simulation of the HDV codec has recently been completed. Assessments of the picture quality after coding and decoding, by both specialists and non-specialists, indicate that the quality is very high – even for very “critical” scenes – suggesting that the theoretically expected high performance of the coding scheme is achievable in practice.

The hardware implementation of the codec will be completed by December 1991. To reduce the constructional complexity, special integrated circuits are being developed as a part of the project; these include filter circuits for band-splitting and interpolation and a circuit for analogue filtering and matrixing applications (before and after the A/D and D/A converters). Commercially-available integrated circuits are in use for the discrete cosine transforms and for full-search block matching. Special multiplexing/demultiplexing chips (TV container) and forward error-correction chips, which are being developed outside the HDV consortium, will be used to minimize the physical dimensions of the codec.

The HDV codec has a clearly-defined interface to the channel multiplexer and contains a robust forward error-correction scheme. It is inherently very flexible and can be used for, or adapted to, a variety of 140-Mbit/s channels such as the switched broadband network (VBN) of the Deutsche Bundespost, the Berkom network in Berlin, the future B-ISDN, the SDM network or satellite transmission channels.

3.2. A compatible HDTV 140 Mbit/s sub-band codec

3.2.1. Introduction

The codec described in this *Section* is being developed for the digital distribution of programme material in the high-definition progressive (HDP) format in a 140-Mbit/s channel of the kind available in ATM networks in the future broadband ISDN*. The useable bit-rate of these channels is 120 Mbit/s and the HDP signal corresponds to a picture in 16:9 aspect ratio with 1920 pixels per active line, 1152 lines per active field and 50 fields

---

* The compatible HDTV codec is being developed in the Department of picture coding and psychovisual studies at the Centre Commun d’Etudes de Télédiffusion et Télécommunications (CCETT), France.

per second without interlacing. This requires a bit-rate reduction factor of 15, although after preliminary sub-sampling of the chrominance information, this is reduced to 11 for the picture coding process itself.

3.2.2. *Filtering and sub-band splitting*

To facilitate the introduction of HDP vision systems alongside the interlaced systems of today, a compatible approach has been chosen. This means that different decoders having different complexities (and costs) should be able to extract from the bit-stream generated by the HDP coder just the information they need to reconstruct images corresponding to the format for which the decoder has been implemented.

The possibilities of sub-band techniques using quadrature mirror filters (QMF) and pseudo quadrature mirror filters (PQMF), as well as the well-known discrete cosine transform (DCT), have been explored and tried. It has been found that PQMF techniques perform better than DCT and are ableto offer very good frequency splitting with a computational cost which is much lower than that necessary with conventional QMF technology.

The DCT, QMF and CQF transforms are now well-known and are widely used. PQMF sub-band splitting is less common and a brief explanation of the principles will be appropriate here. It was noted in *Section 2.1.* that PQMF splitting uses a filter bank based on a prototype low-pass filter and that the components of the filter bank are derived by modulation of this prototype filter. *Fig. 4* shows the schematic diagram of a one-dimensional PQMF filter bank. The characteristics of the filter bank are defined by the prototype filter and the number of sub-bands *SB*. The analysis filters are calculated by modulating the coefficients of the prototype filter by a cosine function, and a second modulation is applied to generate the synthesis filters. The decomposition process involves down-sampling by a factor *SB* after analysis filtering, matched by up-sampling before synthesis.

This process of modulation to obtain the *SB* filters leads to an efficient hardware implementation with, first, the action in parallel of polyphase filters (in which some coefficients have inverted signs), followed by a classic inverse DCT, of size *SB*, on the *SB* pre-filtered pixels (*Fig. 5*) obtained after the first stage An important parameter to be kept in mind in the design of such systems is the computation load. To obtain *SB* pixels, (2*SB*–1) polyphase filters are needed, plus (*SB*–1) extra additions and an inverse DCT of size *SB*. Taking as an example a system with 8 sub-bands and a 61-tap prototype filter, it is found that the decomposition of each pixel requires 9.75 additions and 8.25 multiplications.

Processing in the synthesis filter bank is shown in *Fig. 6*. The reconstruction of each output pixel requires the implementation of a direct DCT followed by a polyphase filter. The calculation load is exactly the same as in the analysis filter

3.2.3. *Bit-rate reduction*

In the present state of the study, the analysis filter uses separable two-dimensional splitting into 8x8 sub-bands, with a prototype filter having 61 taps (denoted 2D-PQMF$_{8x8}$). This replaces the DCT as the transformation which executes the decorrelative process in the coder. After this sub-band decomposition the system works with 12-bit precision. Each sub-band is multiplied by a weighting factor depending on the buffer filling regulator and its position in the frequency range, and then quasi-linearly quantized. Within each sub-band, variable-length codes and run-length codes are used to represent these quantized values.

3.2.4. *Motion estimation*

Motion estimation is performed at the HDP image level by means of a full-search block-matching technique with pel accuracy, followed by a local half-pel accuracy search. The block size is 16x16.

Figure 4
Basic elements of a 1D-PQMF filter bank.
The number of parallel chains is equal to the number of sub-bands, *SB*.

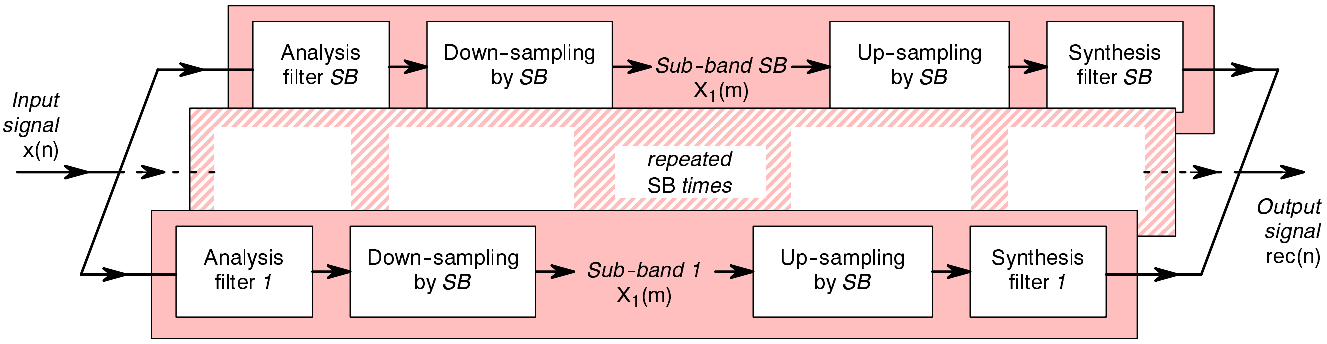
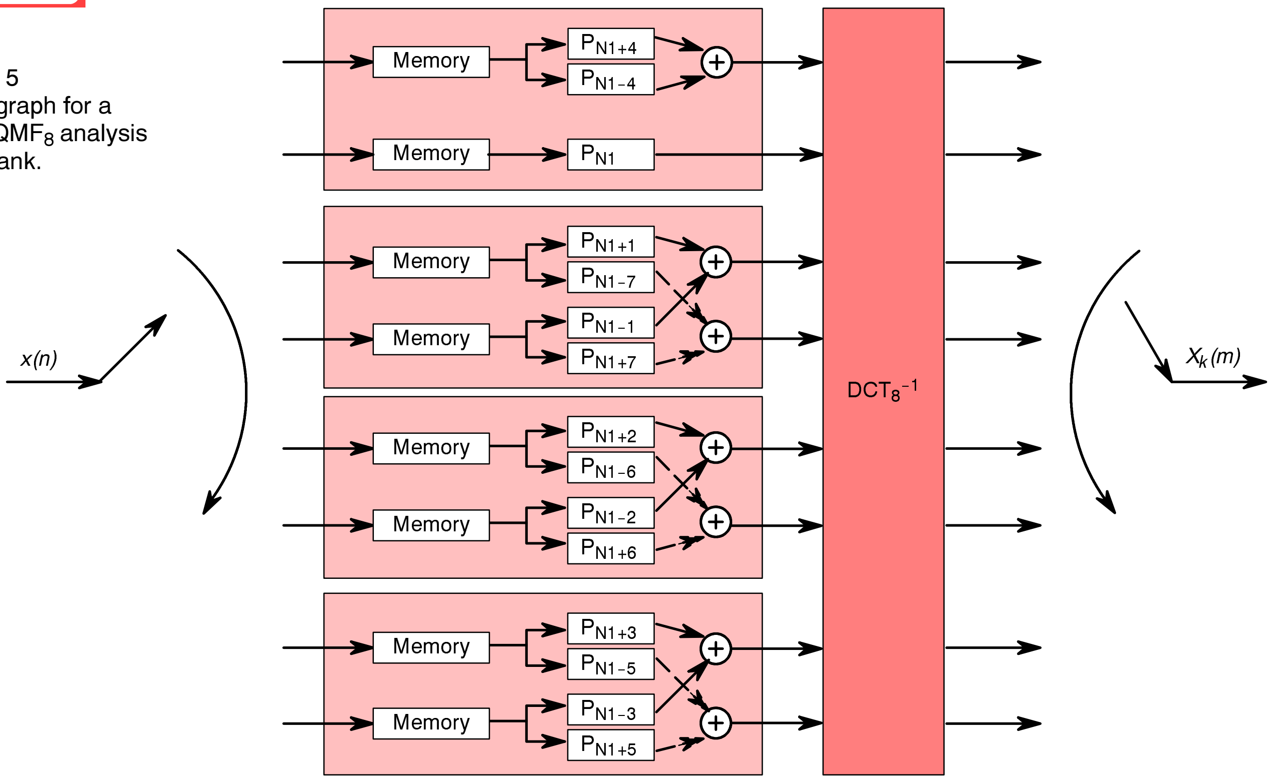


Figure 5
Flow-graph for a 1D-PQMF$_8$ analysis filter bank.



Input signal at sampling frequency f_s takes one path according to n.

These four boxes work at $f_s/4$. Inside each box, each path works at $f_s/8$.

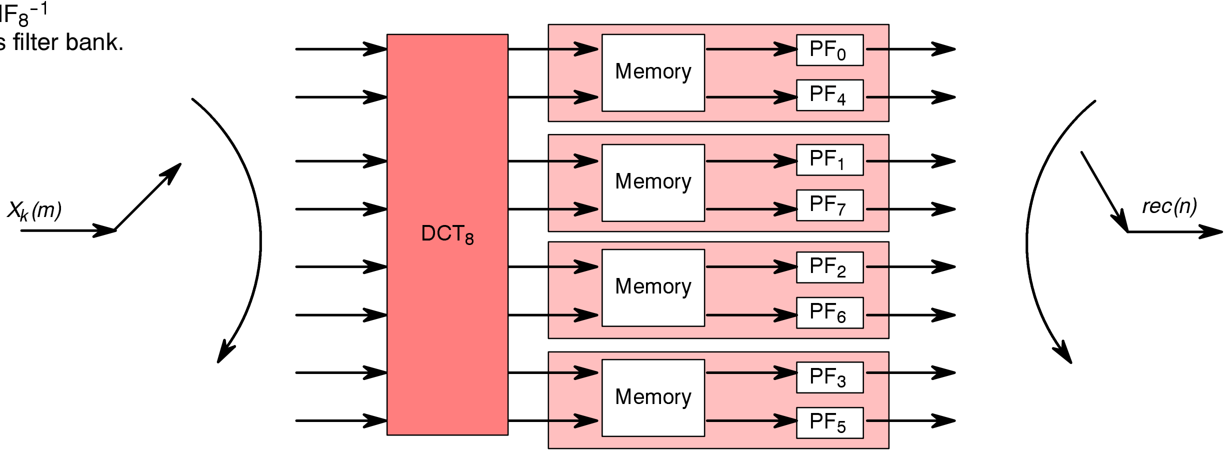
Inverse DCT of size 8.

Output signal at f_s. The path determines the value of k.

3.2.5. *Compatibility*

Compatible decoders operating in the intra mode require only to take some of the sub-bands. After inverse quantization a simpler inverse PQMF transform will deliver good compatible pictures. PQMF banks using prototype filters with an odd number of taps are more suitable in this respect. With HDP input, HDP images can of course be extracted, but also extended-definition progressive (EDP – 960x576x50:1 active pixels, 16:9 aspect ratio), or even enhanced-definition interlaced and high-definition pictures (EDI, HDI).

Figure 6
Flow-graph for a 1D-PQMF$_8^{-1}$ synthesis filter bank.



Input signal at sampling frequency f_s takes one path according to k.

Classic DCT of size 8.

These four boxes work at $f_s/4$. Inside each box, each path works at $f_s/8$.

Output signal at f_s. The path determines the value of n.

To reduce the bit-rate – or to increase the picture quality with the same bit-rate – temporal correlations have to be exploited. At the present time, full compatibility is achieved between HDP and HDI, EDP, EDI formats when working in intra modes or inter modes without motion compensation. Nevertheless, in inter modes with motion compensation, problems arise unless special precautions are taken. In effect, if the HDP coder disregards compatibility and performs a conventional block motion compensation on the whole decoded HDP images (referred to as an "HDP standalone coder", shown in *Fig. 7*), then an EDP decoder which receives all of the frequency information but which processes only the low-frequency (LF) sub-bands, is not able to generate the HDP prediction image. This will cause drift between the EDP decoder and the HDP standalone coder, which will increase until the next intra refresh. To resolve this problem a compatibility-adapted motion compensation technique has been developed for an HDP-EDP compatible coding technique.

3.2.6. *Compatibility-adapted motion compensation*

The simplest solution to the problem of compatibility between HDP and EDP decoders, shown in *Fig. 8,* is to generate an EDP motion-compensated image by processing of the LF (low-frequency) bands only. This EDP prediction image is then split into 4x4 bands using a 2D-PQMF$_{4x4}$ filter bank and subtracted from the 4x4 LF bands of the split original HDP image. With this solution the EDP decoder is able to generate the same EDP prediction image, so compatibility is guaranteed, but the loss of efficiency for the HDP coder, compared to the standalone case, is quite high. A more complex solution which also provides prediction for high frequencies (HF) takes the form of a two-loops system. The technique is illustrated in *Figs. 9, 10* and *11*.

The first loop (EDP loop) operates at the EDP sampling frequency. It takes the 4x4 LF bands and after inverse 2D-PQMF$_{4x4}$ filtering it generates the decoded EDP image. An EDP prediction image is created by means of a quarter-pel motion-compensation process operating on blocks of size 8x8, using the HDP motion vectors. After direct 2D-PQMF$_{4x4}$ splitting of this EDP prediction image, the 4x4 LF predicted sub-bands are available.

The second loop (HDP loop) provides the 48 predicted HF sub-bands by performing a half-pel motion compensation operating on the HDP decoded images, followed by direct 2D-PQMF$_{8x8}$ splitting. Only the 48 predicted HF sub-bands are used, the 16 LF sub-bands being provided by the EDP loop.

Using this technique, the EDP decoder has normal complexity (see *Fig. 11*). HDP/EDP compatibility is achieved with HDP coders and decoders which are slightly more complex than in the standalone case. Nevertheless some simplifications may be made, which exploit the properties of odd-length PQMF filter banks; as a result, the difference between the standalone scheme and the two-loops scheme resides only in the EDP loop, operating at the EDP sampling rate. Although system complexity is important, the essential consideration is that with the two-loops compatibility-adapted motion-compensation technique the loss in efficiency for HDP coded images, compared to the standalone case, is less than 0.5 dB (for all sequences tested) and there is no perceptible visual difference.

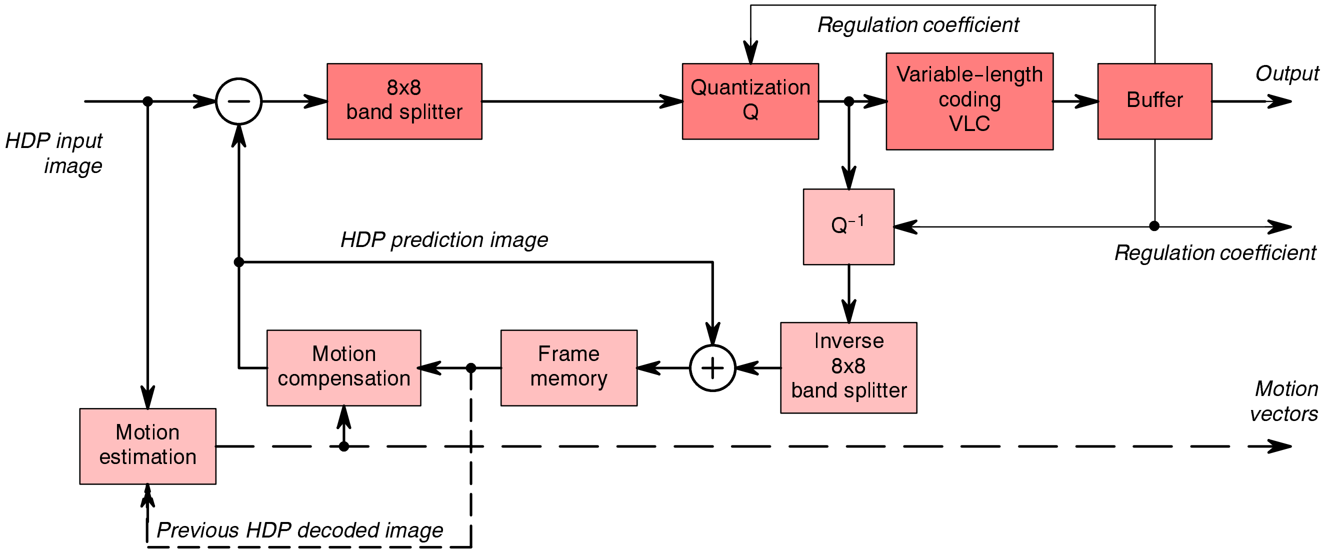


Figure 7
HDP standalone coder.

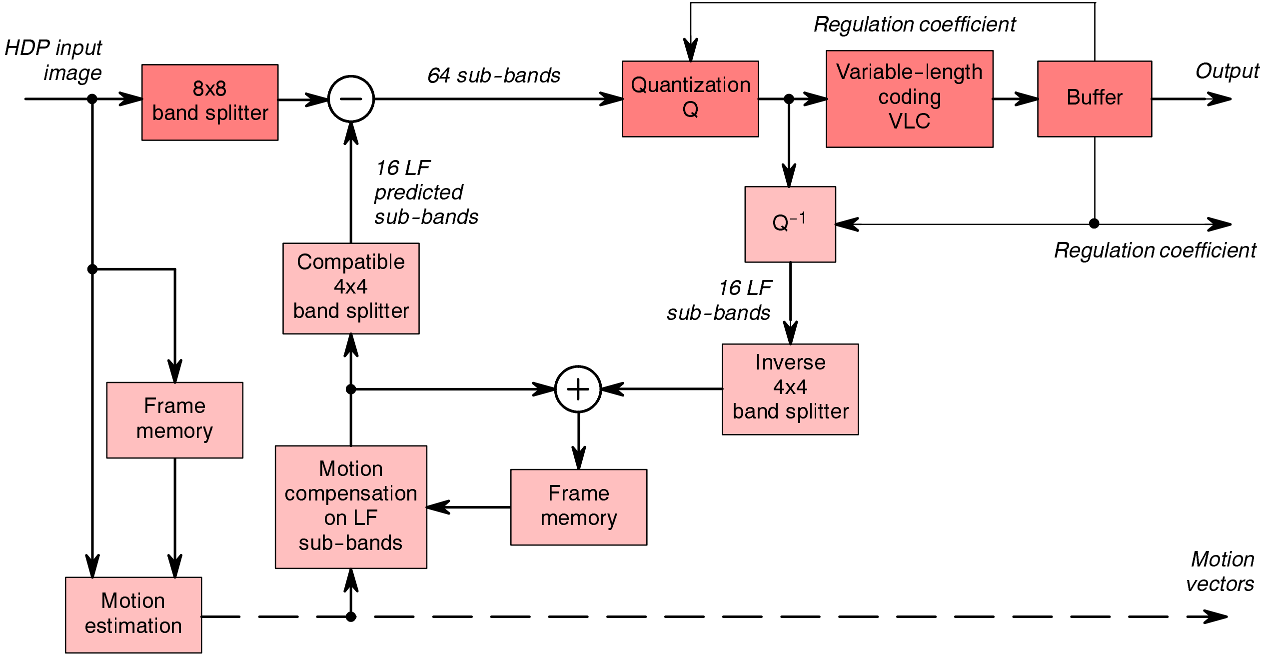
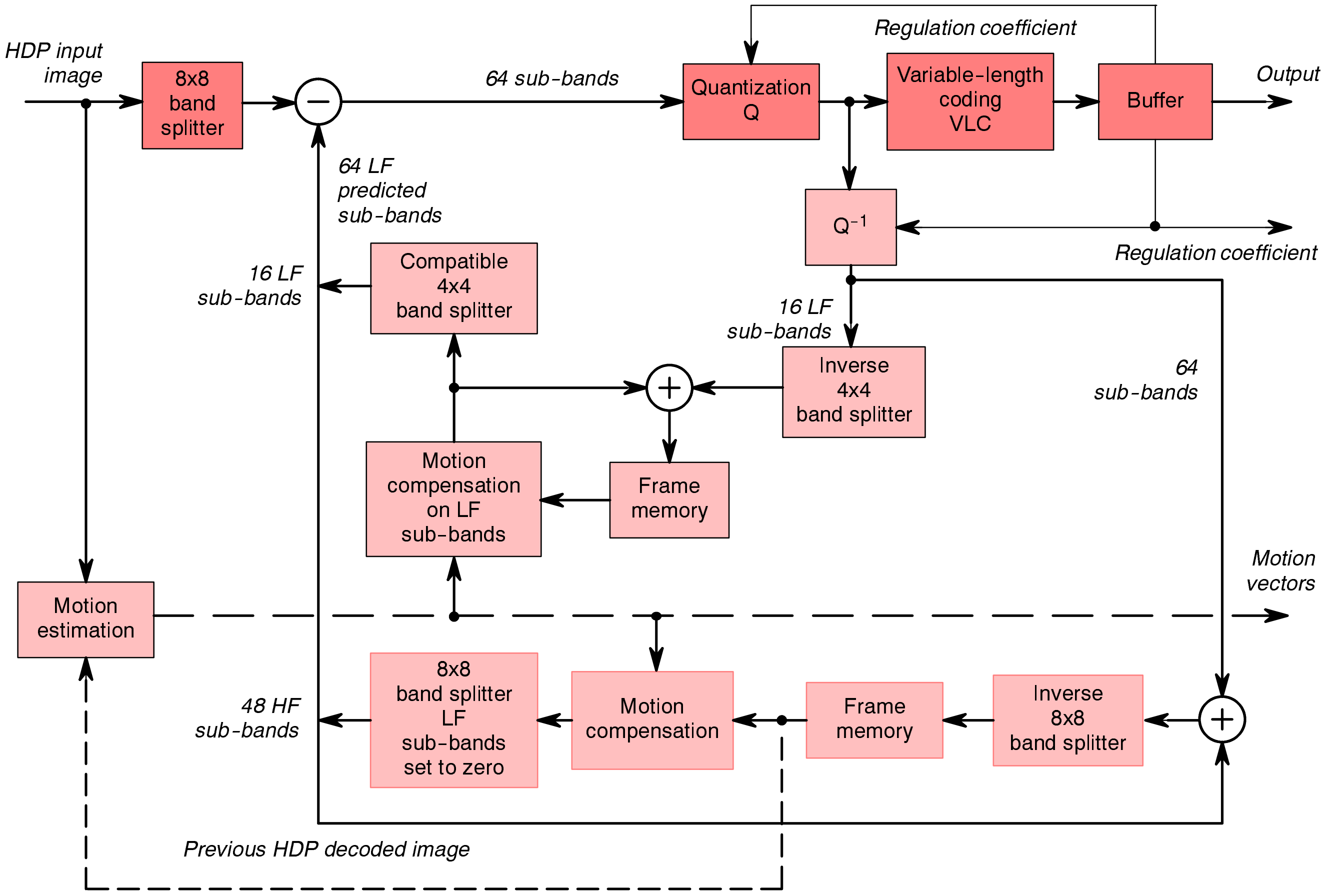


Figure 8 ▲
HDP/EDP
compatible coder –
one-loop solution.

Figure 9 ▼
HDP/EDP
compatible coder –
two-loops solution.



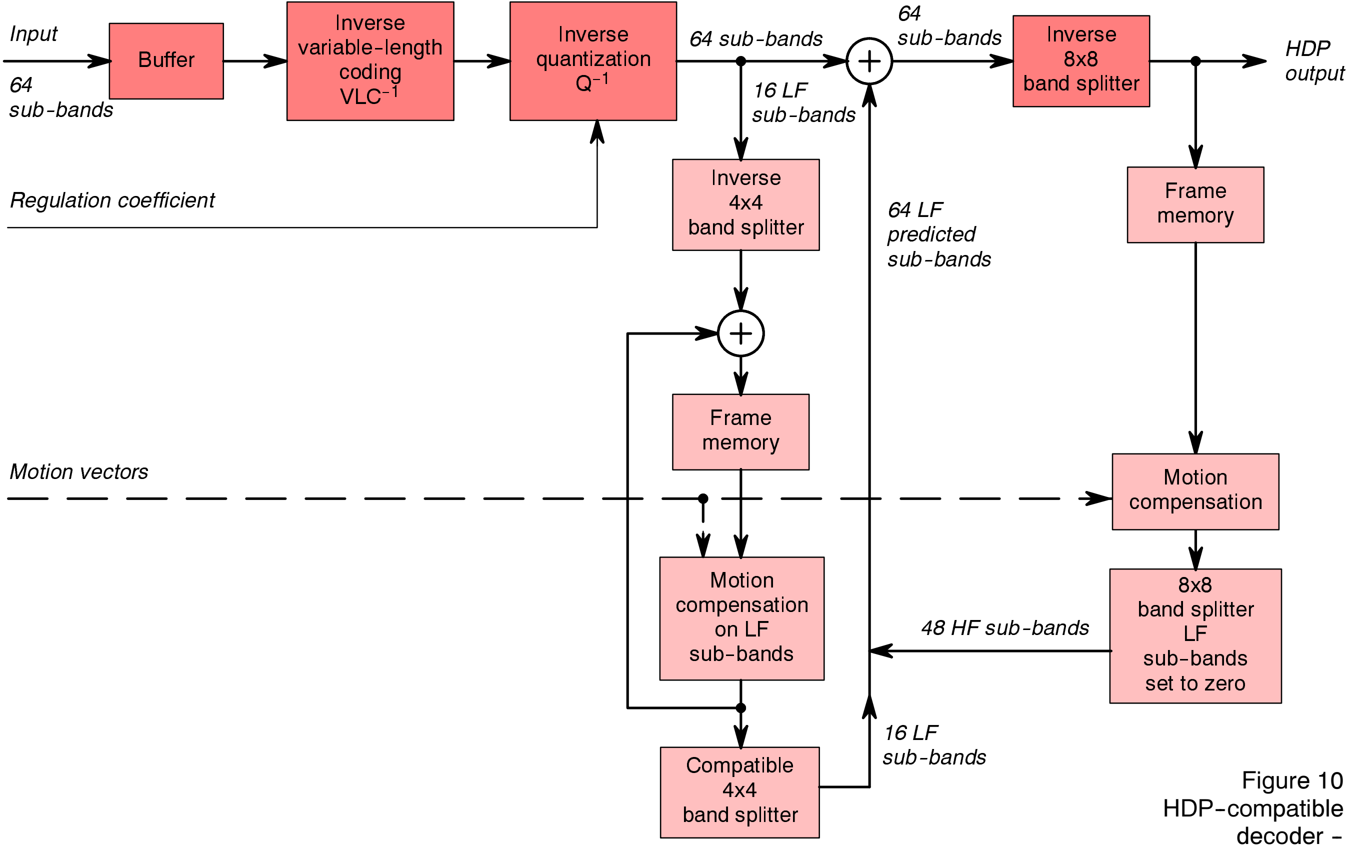


Figure 10
HDP-compatible decoder - two-loops solution.

3.2.7. Results

In comparison with a non-compatible HDP coder (i.e. one using motion compensation at image level), the proposed two-loops scheme allows compatibility at the price of a very small loss in coding efficiency. The visual quality is very good at 120 Mbit/s (equivalent to 1.08 bits/pel for HDP sequences (luminance and chrominance)) and the compatible images are also of fair quality for distribution purposes. Subjective tests have yet to be carried out to confirm these results.

4. Conclusions

In the overall context of digital bit-rate reduction research the conclusions to be drawn at the present time from the studies described here must necessarily be regarded as very provisional. Sub-band techniques have been under study for a variety of sound and picture applications for about ten years, but their superiority over the discrete cosine transform has not yet been demonstrated for the specific application of television transmission.

This provisional conclusion is in fact valid only as regards picture quality, and in the “standalone”

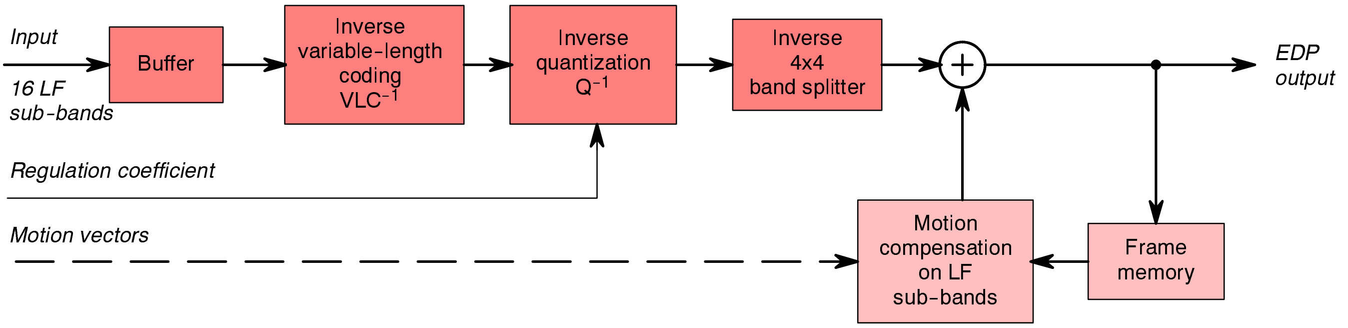


Figure 11
EDP-compatible decoder.

case. Both systems described in this article have shown that the sub-band approach is useful as a means of achieving compatibility. The main advantage is the clean separation of different picture definitions available to viewers. Also, although they cannot be entirely independent, there is some scope for the adoption of different coding techniques for each standard.

Sub-band systems generally have a poor reputation when the hardware implementation is considered. The HDV codec described in *Section 3.1.* will be available in hardware at the end of 1991. While DCT-based HDTV systems need to be implemented with several coding boards operating in parallel in order to reach the very high speed necessary, this is not necessary in sub-band systems. In effect, the processing speed in sub-band hardware is divided by the number of sub-bands, immediately after the sub-band splitter. This observation is not intended to imply that sub-band coding is less complex, overall, than DCT coding, but should instead serve to encourage further analysis of the real, practical complexity of all new schemes.

Bibliography

[1] Mau, J.: **Computationally efficient pseudo-QMF filter bank for a multi-compatible HDTV codec.**
ICASSP91.

[2] Barbero, M., Hofmann, H., Wells, N.D.: **DCT coding and current implementations for HDTV**
Advanced techniques for satellite broadcasting of digital HDTV at frequencies around 20 GHz. - Collected papers on concepts for wide-band digital HDTV satellite broadcasting into the 21st century.
EBU, Geneva, February 1992.

Dr. Hamed Amor *studied at the Technical University of Aachen and from 1983 to 1985 was a teaching assistant in the Electrical Engineering Department, and a research member of the Signal Processing Laboratory, both at the Ecole Nationale d'Ingénieurs in Tunis. Since 1985 he has been with Bosch Telecom in what is now known as the Research Institute Hildesheim.*

Dr. Amor is interested in the development of new digital signal processing systems, and video and audio signal coding in particular; this is reflected in his position as group leader and coordinator of studies on these topics at Hildesheim He is also conducting research into computer vision.

Mr. Eric Bourguignat *graduated in signal processing at the University of Rennes and after completing a thesis on a human spatial vision model he joined the Centre Commun d'Etudes de Télédiffusion et Télécommunications (CCETT) in 1979. He has worked on psychovisual studies applied to television picture coding and in particular on flicker visibility, masking functions and visual quality criteria. He is currently head of the picture coding and psychovisual studies department at the CCETT.*

Mr. Bourguignat participates in several European projects including Eureka 95 and RACE HIVITS and in various standardization committees. He is Vice-Chairman of CCIR IWP 11E.

Mr. Joël Mau *graduated in engineering science from the Ecole Polytechnique, Paris, in 1986, and in signal processing at the Ecole Supérieure des Télécommunications, Paris, in 1988. Since that time he has been working at the Centre Commun d'Etudes de Télédiffusion et Télécommunications (CCETT), on matters concerning image processing and, more particularly, the design of new sub-band techniques applied to image coding.*

Mr. Mau is actively involved in several European projects including Eureka 625 VADIS and RACE HIVITS.