rivistatelettrareview

CONTRIBUTIONS TO TELECOMMUNICATIONS DEVELOPMENT

45

SPECIAL ISSUE FOR HIGH DEFINITION TV

rivistatelettrareview

CONTRIBUTIONS TO TELECOMMUNICATIONS DEVELOPMENT

45

No.1 1990

SPECIAL ISSUE FOR HIGH DEFINITION TV

Publisher
Telettra - Communication & Image Dept.
20092 Cinisello Balsamo (MI)
Viale Fulvio Testi, 136
Tel. 02-24203.1 - Telex 320295
Fax 02-26221777

Registration
Government licence No. 518
November 5, 1983 (English edition)
Government licence No. 3084
June 1, 1953 (Italian Edition)

Six-montly periodical

Subscriber mailing category
IV/70, Milano (Italy)
Taxe perçue - Duty Paid
Milano Ferrovia, Corrispondenza

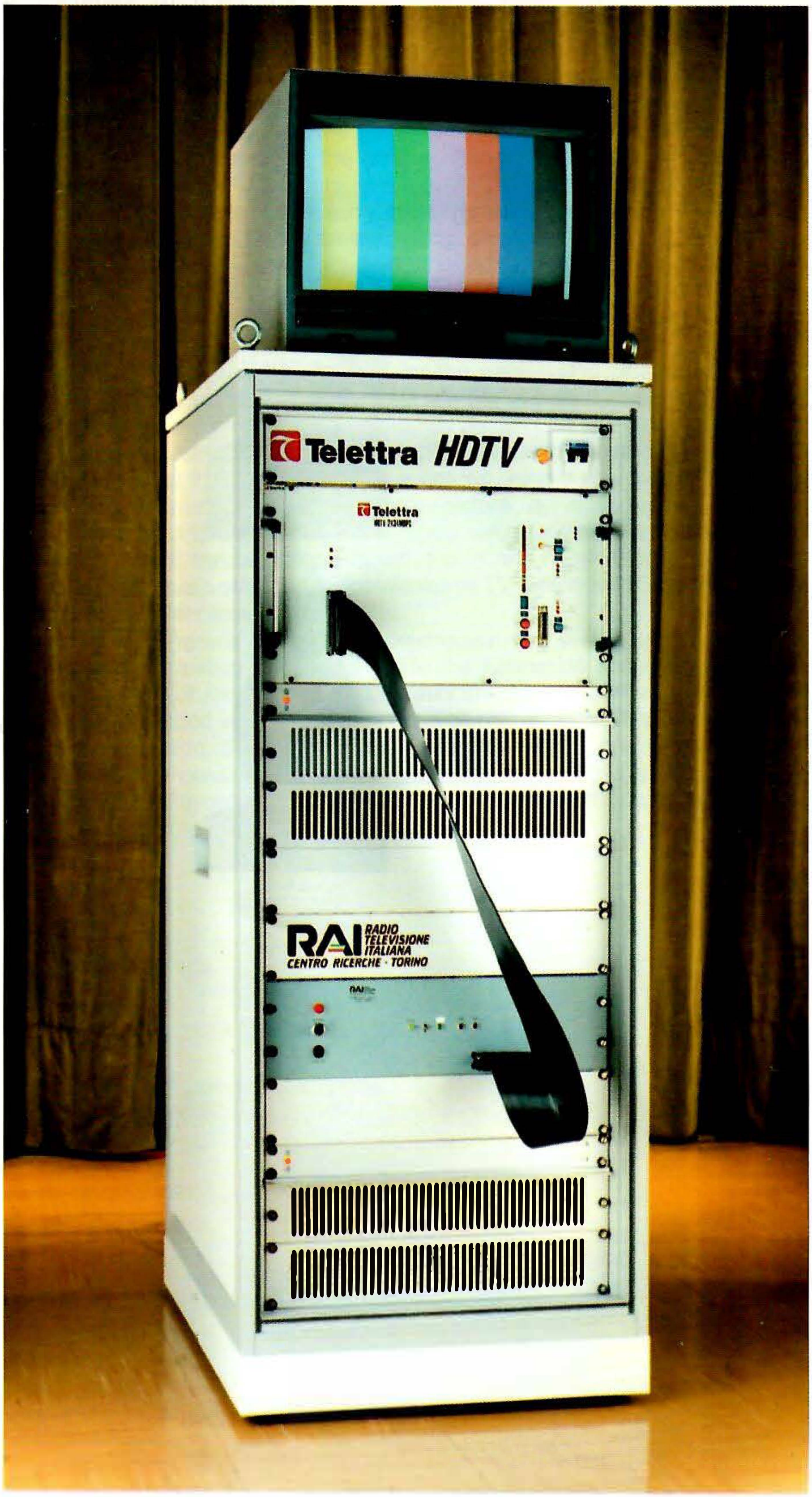
Printed in Italy by
Grafiche Stefanoni, Lecco - June 1990
Circulation: 6000 copies

ISSN: 0392-8268

Contents

The redundancy-reduction coding/decoding equipment for digital transmission of HDTV signals.

(By courtesy of RAI - Photo Valesio).



N. Garcia*, F. Guirao**,
D. Ibáñez**, F. Jaureguizar*,
J. Oest**, J.I. Ronda*

Optimization of HDTV codec algorithms***

Quality demanding visual services require so huge amount of information that there is a need for compression systems to reduce the associated bit rate. Eureka-256 project is intended to provide a solution for the digital transmission of High Definition Television (HDTV) which takes into account both statistical redundancy reduction and perceptual considerations in order to provide a high quality with a reduced transmission bit rate. The study of visibility threshold matrices for DCT coefficients is presented, together which the statistical analysis of the encoded signal, as well as the bit-rate distortion curve, as paradigms of the optimization work underlying the codec design.

Introduction

Image communications are stemming as digital technologies, able to provide new acquisition, processing, transmission, storage, and display facilities. Nevertheless, quality demanding visual services generate signals with so incredible huge amounts of raw information that the available transmission or storage facilities cannot cope with them.

Therefore, there is a need for systems, that maintaining the visual information, are able to reduce the associated data amount required for its representation.

An initial idea is to recover the same visual information. after reducing and restoring the associated data, assuring that there is no mathematical difference between the original and the data-reduced visual signals.

Usually, a dissimilarity between these signals is allowed to achieve a higher data-reduction factor, arising a trade-off design between compression (data-reduction capabilities) and quality (similarity). Depending on the considered communications service, two transmission classes are considered: contribution-quality (the data-reduced visual signal could be processed again) and distribution-quality (no new processing is considered).

The design of a digital codec for visual information transmission is a long process between its original inception and the building of the final system. Initially, an algorithm should be selected among the extensive set of available solutions; afterwards, the encoding parameters should be matched to the considered visual signal; then, the available HW technologies should be evaluated; and, finally, the codec will be built. Basic research on algorithms should be conducted before one is selected, and extensive efforts should be performed to tune the selected algorithm for properly encoding the visual signal. Moreover, all these steps must be completed before HW solutions are evaluated.

Bit rate reduction systems for visual signals take advantage of the intrinsic features of these signals as well as of their relation to the final human observer. Intrinsic features include correlation between adjacent digital video samples, spatial and temporal bandwidth, etc.

Compression techniques make use of the knowledge of the statistical properties of the signal to eliminate redundancy before transmission. Besides, the information the human visual system is not able to appreciate can be discarded as irrelevant for the communication.

Transform coding schemes handle both statistical and perceptual considerations in the encoding of TV and HDTV signals, while keeping a computational complexity suitable for the current state of the art HW technologies. For this reason, international standardization efforts for component TV digital transmission [1], [2] are progressing in the search for recommendations based on transform coding schemes.

Within this frame, the European project Eureka-256 is intended to offer a solution for the digital transmission of HDTV based on the same principles which are being considered for the contribution quality digital transmission of TV.
These principles include the use of prediction in the temporal axis and DCT transformation in the spatial ones.

As in any transform coding scheme, the transformation, followed by the variable length coding of the coefficients, reduces the statistical redundancy, while perceptual considerations are easily included in the scheme by means of a variable quantization of the transform coefficients which distributes the distortion so that it results less perceptible to the final user.

After an overview of the signal processing operations performed on the HDTV signal for its encoding, this paper focuses on the optimization of the encoder in both the perceptual and statistical aspects, describing the design of the visibility threshold matrix, and on the optimization of the block encoding technique, and finishes with some considerations concerning the bit rate regulation.

* Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid.

** Ente Público Retevisión, Madrid.

*** This work has been done within Eureka-256: "Bit-Rate Reduction System for HDTV Digital Transmission". It has been partially supported by the Plan Electrónico e Informático Nacional and the Comisión Interministerial de Ciencia y Tecnologia of the Spanish Government.

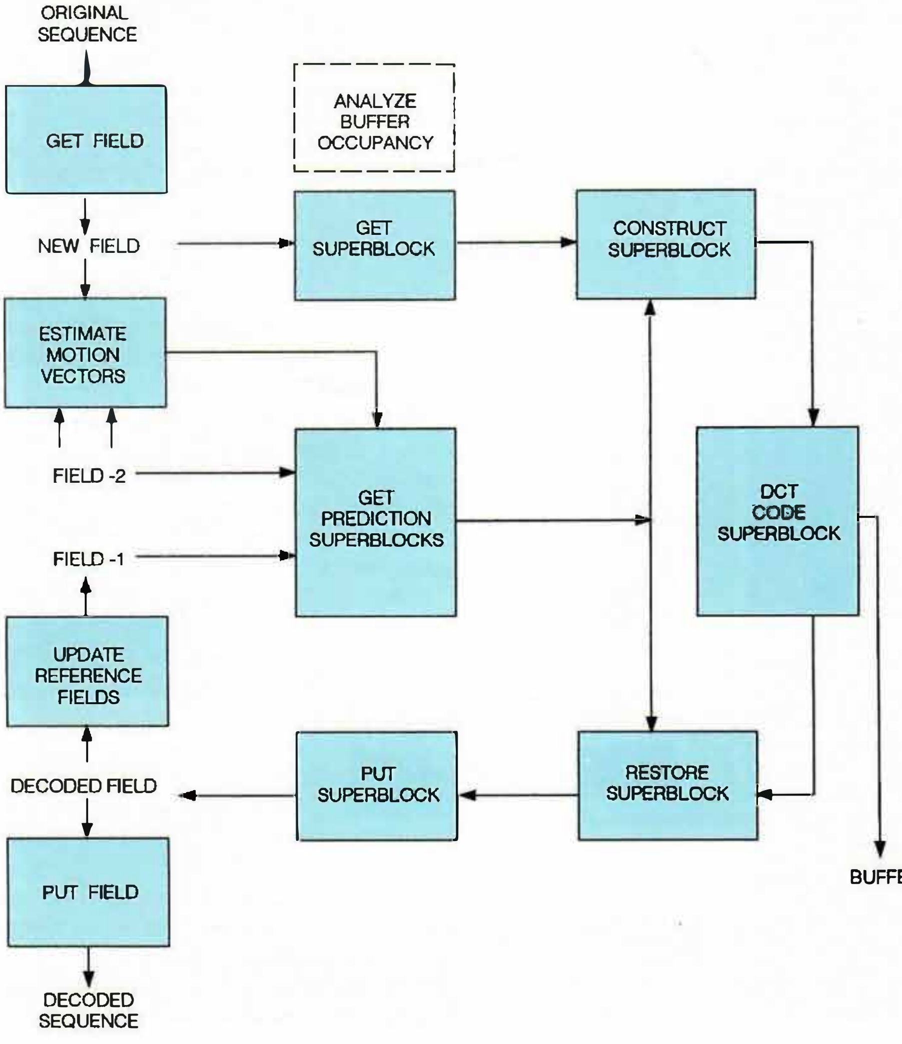
Signal processing architecture

The Eureka-256 codec design considers a Hybrid-DCT compression scheme [3]-[4] to process the incoming visual information signal, either TV or HDTV; this approach mainly follows the CMTT/2 proposals [1]-[2]. So, prediction is applied along the temporal axis and DCT (transform encoding) is applied along the spatial ones. Fig. 1 presents the overall system architecture of the encoder; the decoder is equal to the encoder feed-back loop. The figure shows the functional operation of the system, based on signal processing oriented functions. Operations are shown as modules and interchanged data are placed in between.

As DCT should be computed on blocks, preferable square, block size is the first design parameter. Size 8 rows by 8 columns has been chosen taking into account compression efficiency and computational cost. Therefore, a field holding both luminance (Y) and chrominance (C_R and C_B) can be considered as tessellated into equal sized blocks. A macroblock or superblock is the minimum structure holding luminance and chrominance. Considering the sampling lattice defined in the CCIR Rec. 601 for TV signals, it holds an area of 8 rows by 16 columns, having two luminance blocks and two chrominance ones (one C_R and one C_B). Currently, there is no final definition for HDTV signals, but the European approach considers the same sampling lattice than that of TV, so macroblocks are similar. Besides these DCT processing oriented definitions, stripes being full horizontal bands 8 rows high are considered the smallest units for codec status analysis.

Prediction along the time axis is performed choosing the best fit among the area to be DCT encoded and the prediction set. For the TV/HDTV signals, this prediction set is formed by four predictions: zero (intrafield, no prediction at all), the equivalent area in the last field (interfield prediction), the same area in the same field within the last frame (intraframe prediction), and the motion compensated area in the same field within the last frame (motion compensated prediction).

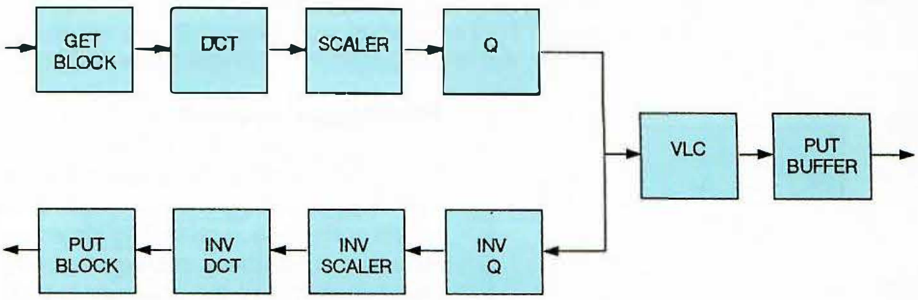
Fig. 1 - General architecture.



The prediction selection is performed on a superblock basis, taking into account the allowed prediction combinations between luminances and chrominances within the superblock. Fig. 1 presents an *a priori* approach for the temporal prediction, but the *a posteriori* approach has also been considered, particularly for distribution quality transmission environments.

Transform encoding is performed on every block of the superblock, as presented in Fig. 2. The signal processing chain includes DCT, Scaling and Quantizing, where Scaling performs the trade between code length (compression) and quantization fineness (quality), thus allowing buffer regulation. VLC encoding reduces code length by assigning word lengths inversely proportional to word probabilities. The feed-back is provided by the operation of a special module (named *Analyze Buffer Occupancy* in Fig. 1) which computes the transmission factor, for controlling the scaling procedure, taking into account the buffer occupancy.

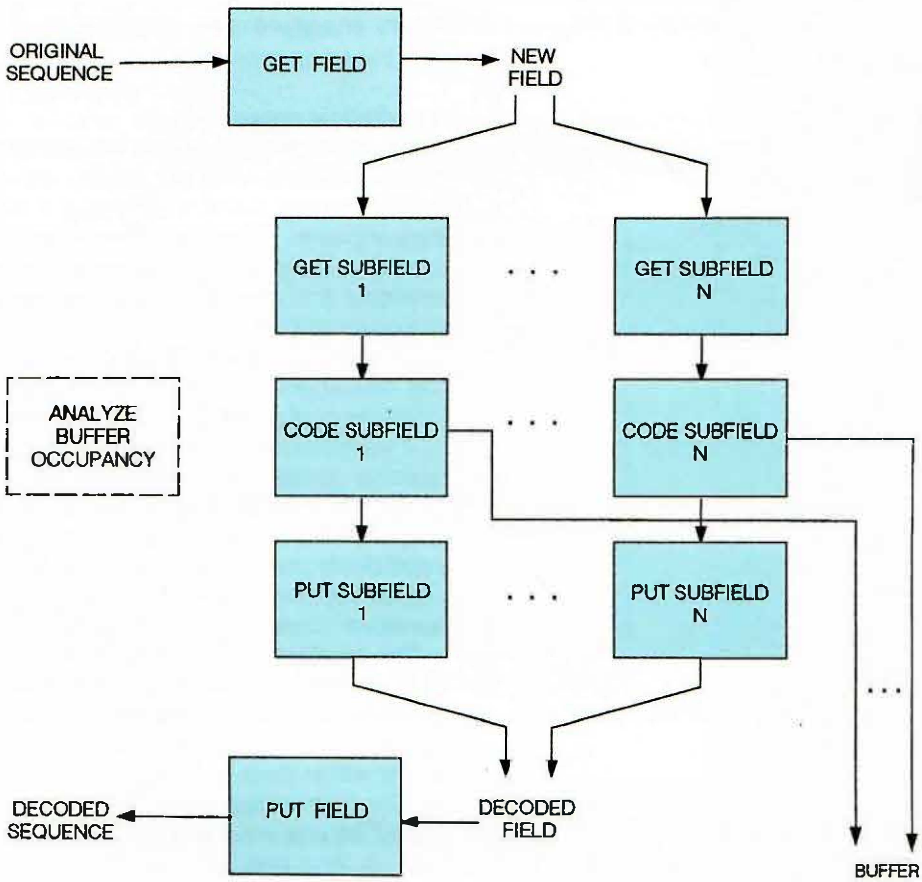
Fig. 2 - Transform encoding of each block of a superblock.



Parallelization scheme for HDTV

As HDTV signals hold a extreme high binary rate, it is impossible to consider a monoprocessor operation, provided standard HW technologies are used.

Fig. 3 - General parallel architecture.



Therefore, either a more sophisticated HW technology is used, or the computational burden is divided among several processors. Eureka-256 follows the latter, considering individual processors formed by complete Hybrid-DCT TV codecs [3]-[4]. So, the HDTV codec is built on the association of standard TV codec, each operating on a different area of the field.

To avoid the need for large buffers between the full HDTV image and the parallel codecs, the splitting approach should follow a procedure of line breaking into equal sized segments. Therefore, it implies a division into vertical bands. Hereafter, each one of these vertical bands will be called a subfield. The general encoder architecture is presented in Fig. 3.

As each one of the codecs varies the quality of the decoded image depending on the buffer occupancy, a quality synchronization procedure is required, governing the quality of all the codecs through the occupancy of the common buffer. For each one of them, there is no difference with its individual operation, as it works always on an external parameter that in this case is provided by a common HDTV system control.

Perceptual aspects

Introduction: visibility threshold matrix

Transform coding schemes have demonstrated several performance benefits in image compression relative to large compression ratios with small quality loss.
However, it has been realized that even more improved systems will result if the visual response of the observer is taken into account, that is, recognizing the fact that the eye is more sensitive to certain spatial frequencies than to others.

Considering a Fourier coded image, this would imply coding more accurately those spectral components to which the eye is more sensitive; this can be achieved by introducing an optical spatial response curve to weight the transform coefficients prior to quantization, bit allocation and coding. A similar filtering operation on more recently developed transforms like the discrete cosine and Walsh-Hadamard transforms has demonstrated a good performance, notwithstanding the theoretical difficulties associated with the convolution/multiplication operation where the discrete cosine transform is concerned.

The weighting factors obtained for each one of the DCT coefficients of an 8x8 image block will constitute the well known Visibility Threshold Matrix (VTM). Two methods are described for computing the luminance and crominance VTM's for HDTV images: a theoretical approach based on previous works and a practical one based on subjective assessment tests.

Theoretical approach

Different studies [5], [6], [7], [8] have demonstrated that the quantization noise of each coefficient does not equally affect the subjective quality of the picture. Improved coding systems will result if the visual response of the observer is taken into account. Therefore, for a fixed output bit-rate, each coefficient has to be coded as accurately as the human visual system is sensitive to its distortion. The quantizer that succeeds in getting equal visible noise for each coefficient will be the optimum one.

The visibility threshold is the minimum amplitude of one DCT coefficient so that the corresponding basis function gets visible. The visibility thresholds of the 64 coefficients of a 8x8 DCT constitute the visibility threshold matrix (VTM).

If we consider a Fourier coded image, this implies coding more accurately those spectral components (spatial frequencies) to which the eye is more sensitive. Hall [5], [6] took these considerations into account and introduced an optical spatial response curve measured by Mannos and Sakrinson [7], to weight the transform coefficients prior to quantization, bit allocation and coding.

The question now is, if it is possible to make a similar filtering operation on the transform coefficient set of the DCT.

The greatest difficulty in using the DCT is that the spectral distribution is not one of true frequency components as in the DFT; spurious energy components appear at arbitrary locations within the coefficient set. Clarke [8] demonstrated that the major part of the energy associated to a given spatial component resides almost totally within three adjacent coefficients when a 16 point DCT is applied, which implies that an appropriate weighting of the DCT coefficients (with a sufficiently slowly varying weighting function of spatial frequency) in the manner reported by Hall will allow improved coding efficiency.

In this study we will start supposing the validity of the results for an 8x8 block transformation, and in a first approach we will suppose that almost all of the energy of a spatial frequency component will appear in only one coefficient, so that we can directly apply the optical response function measured by Mannos and Sakrinson to the frequencies associated to the DCT coefficients.

Methodology

The method used for obtaining the weighting factors associated to each one of the DCT coefficients of an 8x8 image block is represented in the following figure (Fig. 4).

For each coefficient of the block we will:

a) obtain its inverse DCT;
b) expand the result to the whole screen;
c) obtain the Fourier transform of the whole screen;
d) associate a weighting factor to the resulting frequencies by comparison with the frequencies corresponding to the FT of an image with the same characteristics as the one obtained by expanding the DCT^{-1} of the coefficient (Fig. 4). This will be the weighting factor to be applied to the coefficient.

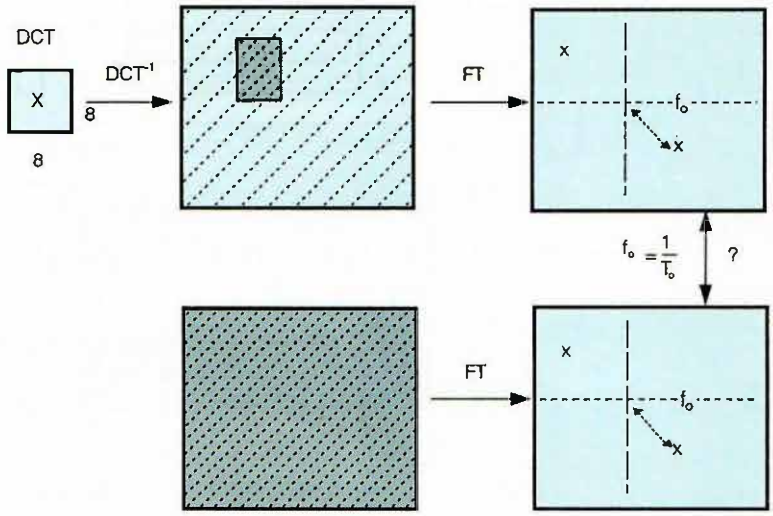


Fig. 4 - Method used for obtaining the weighting factors of DCT coefficients.

In order to do so, we will start from the general formulation for the Inverse Discrete Cosine Transform of a 8x8 block:

$$x(i,j) = \sum_{k=0}^{7} \sum_{L=0}^{7} C_K C_L X(k,L) \cos\left(\frac{2\pi(2i+1)k}{32}\right) \cos\left(\frac{2\pi(2j+1)L}{32}\right) \quad (2.1)$$

where:

$$C_m = \begin{cases} 1/\sqrt{2} & m = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$X(k,L) = DCT\left[x(i,j)\right]$$

$$0 \leqslant i \leqslant 7$$

$$0 \leqslant j \leqslant 7$$

a) If we consider only one coefficient, a considerable simplification of the formula is realized:

$$x(i,j) = C_K C_L X(k,L) \cos\left(\frac{2\pi(2i+1)k}{32}\right) \cos\left(\frac{2\pi(2j+1)L}{32}\right) \quad (2.2)$$

where:

$0 \leqslant i \leqslant 7$

$0 \leqslant j \leqslant 7$

$0 \leqslant k \leqslant 7$

$0 \leqslant L \leqslant 7$

b) Expanding the result to the whole screen implies expanding the limits of the values for i and j.

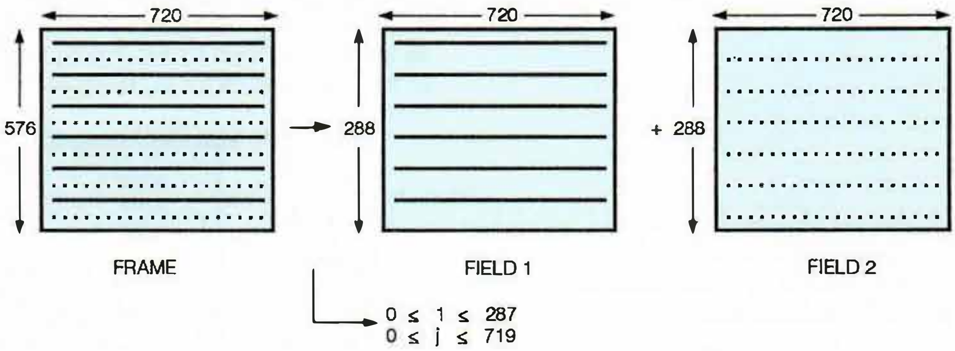


Fig. 5 - DCT of a frame obtained by transforming each one of the two fields.

In interlaced scanning, the DCT of a frame is obtained by transforming each one of the two fields in a 8x8 block basis (Fig. 5). Therefore, we have to expand the limits of i and j in order to consider every pixel of a field.

c) The Fourier Transform of the pixel matrix in terms of the DCT basis functions will be:

$$FTX(k, L) = C_k C_L \pi e^{-j1/2(\Omega_1 + \Omega_2)} \sum_{t=-\infty}^{\infty} \left(\delta\left(\Omega_1 - \frac{2\pi k}{16} - 2\pi t\right) + \delta\left(\Omega_1 - \frac{2\pi k}{16} - 2\pi t\right) \right) \cdot$$

$$\cdot \sum_{t=-\infty}^{\infty} \left(\delta\left(\Omega_2 - \frac{2\pi L}{16} - 2\pi t\right) + \delta\left(\Omega_2 + \frac{2\pi L}{16} - 2\pi t\right) \right) \qquad (2.3)$$

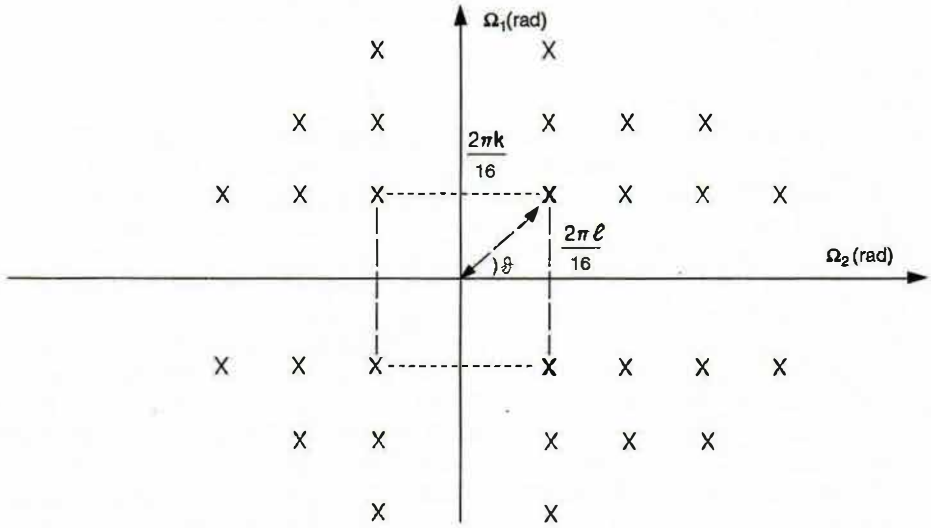


Fig. 6 - Absolute spatial frequency associated to the DCT coefficient.

The absolute spatial frequency (rad) associated to the DCT coefficient X(k, L) will then be (Fig. 6):

$$f_{rad} = \sqrt{\left(\frac{2\pi k}{16}\right)^2 + \left(\frac{2\pi L}{16}\right)^2} = \frac{2\pi}{16}\sqrt{k^2 + L^2}$$

where:

$\theta = \text{arctg}\,(k/L)$.

d) In order to calculate the weighting factor to apply on the coefficient prior to quantization, bit allocation and coding, we have to transform the frequency in radians to frequency in °/H or c/H, taking into account the viewing distance between the observer and the screen. Because of the aspect ratio of the screen (4/3), the multiplying factors will be different in the horizontal and in the vertical directions (Fig. 7).

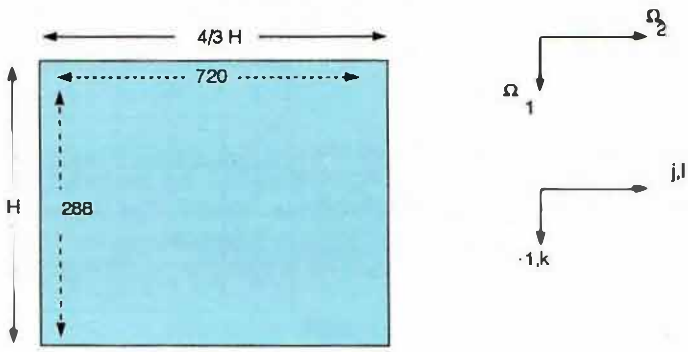


Fig. 7 - Multiplying factors for aspect ratio of the screen (4/3).

Taking into account these considerations, the module of the absolute spatial frequency in (c/o) corresponding to a coefficient X(k, L) (obtained from the DCT of an 8x8 image block) will be:

$$f_{c/o} = \sqrt{\left(f_k\,(c/o)\right)^2 + \left(f_L\,(c/o)\right)^2}$$

$$f_{c/o} = \frac{2\pi}{16}\sqrt{(3{,}2k)^2 + (6L)^2} \qquad \mathbf{D = 4H}$$

$$f_{c/o} = \frac{2\pi}{16}\sqrt{(4{,}8k)^2 + (9L)^2} \qquad \mathbf{D = 6H} \qquad (2.4)$$

The problem is now, how to apply to the DCT coefficients of an 8x8 image block the optical response weighting function obtained by Mannos and Sakrinson for the spatial frequencies associated to the whole image [7].

If we directly applied this curve to the DCT coefficient of an 8x8 image block we would associate small weighting factors to the low frequency coefficients. This would result in a greater quantization error which will not be perceptible on the block, but over the whole screen as a block effect.

Taking into account these considerations, we can't use the original visual response curve, but a modification of it which smooths the curve in the low frequencies, in order to consider higher weighting factors for the low order coefficients of the 8x8 image block, and eliminate in this way the block effect.

Results

Tables 1 and 2 show the p_o associated to each DCT coefficient of a 8x8 block for D = 4H and D = 6H. These factors have been obtained from the formula:

$$W = \frac{1}{\left(\sqrt[8]{2}\right)^{p_o}} \qquad (2.5)$$

where W is the weighting factor obtained from the visual response weighting curve using the frequencies associated to each of the coefficients (2.4).

The maximum weighting factor W_{max} and its associated spatial frequency f_{max} are searched. Then, every factor corresponding to a frequency less than f_{max} is given the value $p_o = 0$.

Table 1 - **D = 6H.**

p_o

k[1]	0	1	2	3	4	5	6	7
0	0	0	0	1	3	6	9	13
1	0	0	0	1	3	6	9	13
2	0	0	0	1	3	6	9	13
3	0	0	0	2	4	6	10	14
4	0	0	1	2	4	7	11	14
5	0	0	1	3	5	8	11	15
6	1	1	2	4	6	9	12	16
7	1	3	3	5	7	10	13	17

Table 2 - **D = 4H.**

p_o

k[1]	0	1	2	3	4	5	6	7
0	0	0	0	0	0	1	3	5
1	0	0	0	0	0	1	3	5
2	0	0	0	0	0	2	3	5
3	0	0	0	0	0	2	3	5
4	0	0	0	0	1	2	3	5
5	0	0	0	0	1	2	4	6
6	0	0	0	1	2	3	4	6
7	0	0	0	1	2	3	5	7

VTM based on subjective assessment

Visibility threshold matrix / quantization noise

As we mentioned earlier, the different coefficient will be quantized with different precision, depending on the visibility threshold.

This quantization will produce a noise with mean square error:

$$\bar{e}^2 = \frac{a^2}{12}$$

where a is the quantization step, supposing a uniform quantizer. So we will generate an error function that simulates the value of the quantization noise of each coefficient and consider this function as the **X(k, L)** value in order to apply the DCT^{-1} over it.

Based upon the article [9], we will not use a random noise but a 6 Hz cosine function, that causes the most visible pattern for the used spatial frequency range. This is a signal of type: $X(k, L) = A \cos(2\pi f t)$, where $f = 6$ Hz. The mean square value of this function is $1/2\ A^2$, then **a** and **A** are proportional related, and for an equal power increase either of quantization noise or of the generated function, it is necessary an equal increment of **a** and **A** (Fig. 8).

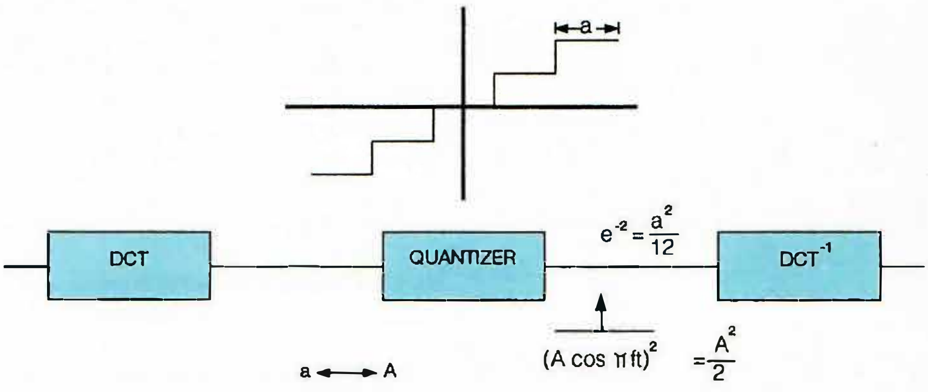


Fig. 8 - Mean square value of noise function $(k, L) = A \cos(2\pi f t)$.

As we are only concerned with the rate between the different coefficients we will not care about factors like 1/2 and 1/12 , and observing when the generated functions are visible with different values of **A** for each coefficient, a valid visibility threshold matrix will be obtained.

We will only vary **t** from one field to another, therefore its value is:

$$t = \frac{\text{Field}}{50}$$

because the field frequency is 50 Hz.

Signals used to measure the visibility threshold matrix

Some signals have been generated, in order to measure the visibility threshold values in a subjective way. Only the luminance component is considered in this article. Using (2.1) and (2.2) for the DCT^{-1} of one of the coefficients, expanding the results for the whole screen, as was mentioned, in 2.2 A, and considering interlaced scanning of the screen, the final formula for visualizing the DCT basis function will be:

$$S(i, j, t) = K + C_K C_L A \cos 2\pi f t \cos\left(\frac{2\pi(2j+1)L}{32}\right) \cdot \begin{cases} \cos\left(\frac{2\pi(2i+1)k}{32}\right) & \text{; for field 1} \\ \cos\left(\frac{2\pi(2i+2)k}{32}\right) & \text{; for field 2} \end{cases} \qquad (3.1)$$

where:

i and **j** are the vertical and horizontal coordinates $\begin{cases} 0 \leqslant i \leqslant 15 \\ 0 \leqslant j \leqslant 15 \end{cases}$

the value $S(i, j, t) \in [0,255]$

$K = 126$

$$\cos 2\pi f t = \cos\left(2\pi\, 6\, \frac{\text{Field}}{50}\right)$$

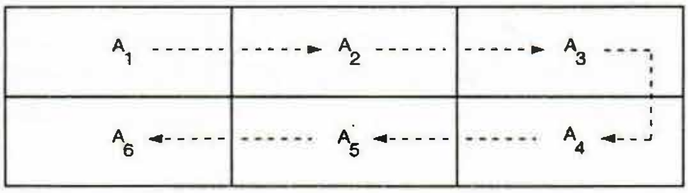
k and **L** define the coefficient

A is variable and it will determine the visibility threshold we are looking for.

Subjective assessment tests

With this image function, sequences of 25 frames with different values of **A** for each coefficient have been generated in order to obtain a visibility threshold matrix. Different observers were asked to check the visualized sequences.

The presented images were generated by a computer program as a monochrome test pattern dividing the center of the monitor screen in six squares, the rest of the screen being grey. The squares were separated by black borders and in each square a different multiple of the image function of one coefficient was included, increasing from the left upper corner to the right side, then down and left, finishing in the left bottom corner (Fig. 9).



$$A_1 < A_2 < A_3 < A_4 < A_5 < A_6$$

Fig. 9 - Monitor screen divided in six squares.

The viewers were asked to mark how many of these squares they could see anything in. In this way, we could deduce which was the minor value of A each person could see. For example, if a person marked 4 squares, the minor value of A would be A_3 and A_3 was noted as the visibility threshold seen by this person for the coefficient analyzed.

The possible values of A were from 0 to 255 with a minimum increment of 1, because the used hardware did not allow smaller increments.
We checked only over 15 coefficients to simplify the tests.
The tests were carried out at a distance of 4H and 6H and the results for those two distances were the Tables 3 and 4.

Table 3 - **D = 6H.**

p_0 / k^1	0	1	2	3	4	5	6	7
0			A = 1,12		A = 4,18		A = 11,47	
1								
2	A = 1,12		A = 1,94		A = 5,88		A = 22,24	
3								
4	A = 1,76		A = 3,06		A = 8,29		A = 28,59	
5								
6	A = 2,41		A = 5,00		A = 15,06		A = 35,29	
7								

Table 4 - **D = 4H.**

p_0 / k^1	0	1	2	3	4	5	6	7
0			A = 1,06		A = 3,87		A = 9,31	
1								
2	A = 0,94		A = 1,86		A = 4,75		A = 16,88	
3								
4	A = 1,31		A = 2,13		A = 5,94		A = 21,75	
5								
6	A = 2,06		A = 3,00		A = 11,13		A = 27,75	
7								

To obtain the whole matrix we have interpolated the results as a lineal interpolation. The values (0,0), (0,1), (1,0) and (1,1) were set to 1. And using

$$W = \frac{1}{(\sqrt[8]{2})^{p_o}}$$

where:

$$W = \frac{1}{A}$$

p_o is obtained for each coefficient as shown in Tables 5 and 6.

Table 5 - **D = 6H.**

p_o

k'	0	1	2	3	4	5	6	7
0	0	0	1	11	17	24	28	31
1	0	0	5	14	19	28	33	36
2	1	5	8	16	20	31	36	39
3	4	8	11	18	23	32	37	41
4	7	10	13	20	24	34	39	42
5	9	13	16	24	28	36	40	43
6	10	15	19	27	31	37	41	44
7	12	17	21	29	34	39	42	45

Table 6 - **D = 4H.**

p_o

k'	0	1	2	3	4	5	6	7
0	0	0	1	10	16	22	26	29
1	0	0	4	12	17	25	30	33
2	0	4	7	14	18	27	33	36
3	1	5	8	15	19	29	34	38
4	3	6	9	16	21	30	36	39
5	6	9	11	20	25	32	37	40
6	8	11	13	23	28	34	38	41
7	10	12	14	25	30	36	40	42

Statistical aspects

The performance of an encoding system strongly depends on the tuning of its parameters to the statistical properties of the signal to be encoded. In the case of the Hybrid-DCT bit rate reduction system, the design of the VLC block must be the result of a complete knowledge of the TV and HDTV signals. Considering this statistical characterization, two main decisions have to be taken:

- The choice for a technique for the description of the DCT blocks.
- The selection of the code or the family of codes which will encode the symbols resulting from the description of the blocks.

The technique for the DCT blocks description has to take into account the high number of null coefficients and the phenomenon of the accumulation of the non-zero coefficients in the low frequencies. According to these premises, the Eureka-256 design transmits the block by means of two sets of symbols. The first set of symbols corresponds to non-zero values of the coefficients, and the second one to run lengths of zeros, scanned according to a given scanning path which is different for each type of signal (luminance and chrominance). The last zero run length is not transmitted; instead of it, the end of the block is signaled with a special code word (EOB). Actually, two different code words are used for video framing purposes. Using this technique, a better compression is achieved in comparison to the separate encoding of each coefficient value, while paying a very low increase in complexity. Further adjustments to the basic approach include taking advantage of the existence of impossible pairs of words, such as two consecutive zero run lengths or a run length followed by an end of block. This results useful for the encoding of +1 coefficient strings between zeros: if there are one or more coefficients +1 between two runs of zeros, or between a run of zeros and the EOB, one of them is not transmitted and the decoder reinserts it.

Huffman codes, though optimal from the average bit rate reduction point of view, are not practical for the transmission through noisy channels, and other codes have to be used which present better word synchronism recovery properties. B-2 codes show these desirable features by defining words with an even number of bits, organized as follows: bits holding odd positions are all set to 1, with the exception of the last one in each word. So, word synchronism information can be obtained from the odd bits, while the even ones encode freely the information. Taking into account that the receiver knows the maximum VLC word length, the last odd bit of the maximum length words can also be used for information encoding instead of using it for word synchronism. According to it, the number of different code words of a B-2 code with maximum word length equal to 2L is $3 \cdot 2^L - 2$. The assignment of codewords to symbols has to ensure that shorter words correspond to more frequent symbols. Tables 7 and 8 show the distribution of these symbols in the encoding of the TV signal, without considering the +1 strings compression between zeros. The given results are averaged from statistics which consider separately every pair prediction mode – component type. These more exhaustive tables allow the evaluation of the need for different table assignments for each mode and type, which can decrease the bit rate without requiring any additional overhead.

The efficiency of the B-2 code can be measured by comparing the obtained average bit rate to the limits established by the Information Theory. According to it, the Hybrid DCT codec can be considered as a discrete information source **S** which produces symbols representing non-zero quantizing levels, zero run-lengths and EOB's. Its first order entropy can be computed as:

$$H(S) = -\sum_{i=1}^{N} p_i \log_2 p_i$$

where p_i represents the probability that the source produces the symbol s_i. The value H(S) indicates:

- the minimum average number of bits per symbol which can encode the information produced by the source, assuming independency between consecutive symbols, and, consequently,
- a lower bound of the minimum achievable average number of bits per symbol through the independent encoding of each symbol.

The actual average number of bits per symbol can be computed as:

$$R(S) = \sum_{i=1}^{N} p_i l_i$$

where l_i is the length of the code word associated to the symbol s_i. The efficiency of a given variable length code (VLC) can be defined as the quotient between the previously computed entropy and its resulting average bit rate per symbol (**R**):

$$E = \frac{H(S)}{R(S)}$$

Tables 9 and 10 show the values of **R**, **H** and **E** for each component type both in the case of TV and HDTV signals. In the experiment whose results are given, three test TV sequences (*Calendar, Renata* and *Renata-Butterflies*), and a HDTV one (*Renata-HDTV*) were processed. The given results correspond to their encoding without motion compensation and not making use of the suppression of +1 between zeros. Results were first obtained separately for each pair component type-prediction mode in order to allow for a separate analysis of the different sources. Similar figures of efficiency where obtained in each case.

Table 7 - Percentages of the most frequent code words in the TV signal encoding: luminance.

NON-ZERO LEVELS		ZERO RUN-LENGTHS	
LEVEL	%	LENGTH	%
± 1	20.625	1	11.573
± 2	6.550	2	4.410
± 3	3.163	3	2.304
± 4	1.833	4	1.342
± 5	1.176	5	0.907
± 6	0.783	6	0.645
± 7	0.564	7	0.425
± 8	0.413	8	0.196
± 9	0.306	9	0.113
± 10	0.239	10	0.080
± 11	0.193	11	0.061
± 12	0.148	12	0.049
± 13	0.124	13	0.045
± 14	0.103	14	0.037
± 15	0.081	15	0.030
± 16	0.072	16	0.017
± 17	0.059	17	0.010
± 18	0.053	18	0.007
± 19	0.043	19	0.005
± 20	0.039	20	0.005
± 21	0.032	21	0.005
± 22	0.029	22	0.004
± 23	0.027	23	0.003
± 24	0.022	24	0.002
± 25	0.021	25	0.002
± 26	0.018	26	0.001
± 27	0.016	27	0.001
± 28	0.015	28	0.001
± 29	0.013	29	0.001
± 30	0.012	30	0.001
± 31	0.011	31	0.001
± 32	0.010	32	0.001
± 33	0.009	33	0.001
± 34	0.009	34	0.000
EOB	3.199	35	0.000
		36	0.000
		37	0.000
		38	0.000

Table 8 - Percentages of the most frequent code words in the TV signal encoding: chrominance.

NON-ZERO LEVELS		ZERO RUN-LENGTHS	
LEVEL	%	LENGTH	%
± 1	25.372	1	10.191
± 2	2.413	2	5.540
± 3	0.718	3	3.665
± 4	0.322	4	2.654
± 5	0.177	5	2.047
± 6	0.110	6	1.547
± 7	0.083	7	1.227
± 8	0.071	8	0.986
± 9	0.060	9	0.777
± 10	0.046	10	0.619
± 11	0.034	11	0.523
± 12	0.029	12	0.444
± 13	0.027	13	0.350
± 14	0.023	14	0.252
± 15	0.021	15	0.173
± 16	0.020	16	0.133
± 17	0.018	17	0.116
± 18	0.017	18	0.103
± 19	0.021	19	0.086
± 20	0.020	20	0.078
± 21	0.016	21	0.066
± 22	0.016	22	0.050
± 23	0.016	23	0.043
± 24	0.015	24	0.037
± 25	0.014	25	0.034
± 26	0.015	26	0.027
± 27	0.014	27	0.023
± 28	0.010	28	0.015
± 29	0.008	29	0.011
± 30	0.008	30	0.012
± 31	0.008	31	0.011
± 32	0.008	32	0.010
± 33	0.008	33	0.008
± 34	0.009	34	0.006
EOB	8.214	35	0.004
		36	0.004
		37	0.004
		38	0.003

Table 9 - Average bit-rate per symbol, entropy and code efficiency by component type for the TV signal.

TYPE	R	H	E
Y	4.15	3.97	0.96
CR	3.81	3.57	0.94
CB	3.78	3.54	0.93
Total	4.09	3.90	0.95

Table 10 - Average bit-rate per symbol, entropy and code efficiency by component type for the HDTV signal.

TYPE	R	H	E
Y	4.67	4.41	0.94
CR	4.25	4.01	0.94
CB	4.25	4.01	0.94
Total	4.57	4.31	0.94

The average efficiency of the B-2 code is close to optimal (0.95) for the tested TV sequences, and shows an efficiency of 0.94 for the HDTV one. Thereby, the following conclusions stem:

- The B-2 code is suitable for the transmission of the symbols resulting from the hybrid DCT encoding of the TV signal when the DCT blocks are described using the aforementioned technique.
- It doesn't seem useful to search for a more complex code to be applied together with this block description technique.
- The B-2 code shows also a good performance in the encoding of HDTV signals. Therefore, the extendibility to HDTV requirement is fulfilled.

Rate-distortion analysis

Fixing the transmission factor to a pre-specified value, it is possible to compute the required bit-rate for a given quality (transmission factor). Fig. 10 presents this bit-rate distortion curve for three digital TV test sequences, as well as the average of them. This last curve characterizes the Hybrid-DCT encoding of the TV signal.

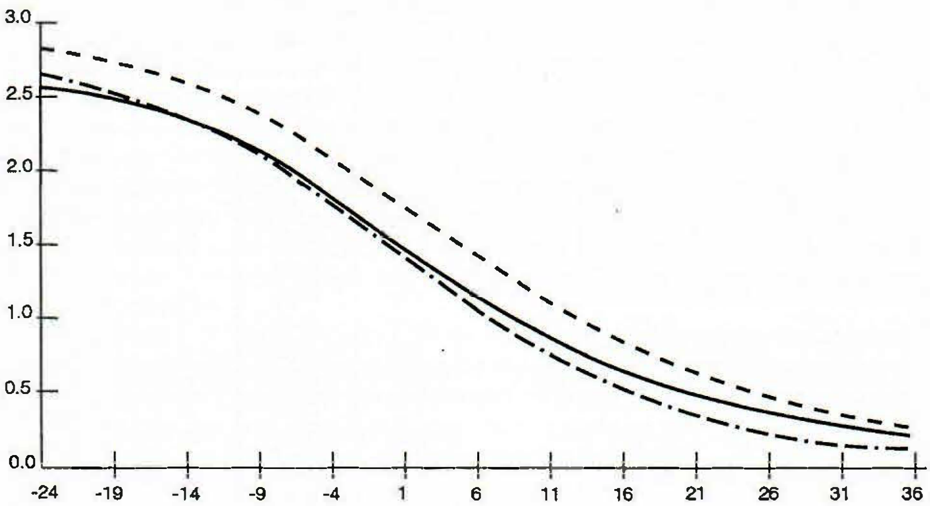


Fig. 10 - Buffer occupancy vs. transmission factor.

Conclusion

The Eureka-256 codec has been presented, specifying the signal processing architecture as well as the specific items in which further optimization is required. Thus:

1. The design of the visibility threshold matrices has been determined taking into account theoretical and subjectively assessed approaches to the modelization of the human visual response.
2. The DCT block encoding technique has been described. The efficiency of the B-2 code has been studied when applied to the encoding of the block coefficients, showing a close to optimal performance for both TV and HDTV.
3. The bit-rate distortion curve has been obtained, therefore characterizing the visual information signal, thus allowing for further studies on different buffer control procedures.

Manuscript received in May 1990.

NARCISO GARCIA was born in Madrid in 1954.
He received the degree in telecommunications engineering in 1976, and the PHD degree in communication engineering in 1983, both from the Polytechnic University of Madrid.
Currently, he is an associate professor of this same university where he received the Research Award in 1988; he is also scientific responsible for the Image Processing Group and Vice-Head of the Signals, Systems and Communications Department.

FRANCISCO JAVIER GUIRAO holds a degree in telecommunications engineering. In 1982 be joined the information services of RTVE, and in 1988 he became part of the Image Processing Laboratory to participate in the development of the project EU256. Since 1989 he has been head of the laboratory for High Definition TV.

DANIEL IBAÑEZ TORRES graduated as Telecommunication Technical Engineer in 1965 and as Telecommunication Higher Engineer in 1971, at the Polytechnic University of Madrid.
He joined RTVE in 1966, being assigned the task of regional maintenance of TV repeaters. In 1972 he became responsible for TV network planning as well as for design and testing of the radiating systems of TV transmitter and repeater centres.
In 1987 he joined the RTVE's Technological Innovation unit, leading the Image Processing Group which, since then, has worked with the EU 256.
In 1990 he was named Associate Director of Technology at Retevisión, which is his current task.

FERNANDO JAUREGUIZAR was born in Bilbao in 1962. He received the degree of Ingeniero de Telecomunicación in 1987 from the Universidad Politécnica de Madrid.
He joined the Grupo de Tratamiento de Imágenes of the Universidad Politécnica de Madrid in 1987, where he has been engaged in research on TV and HDTV encoding, particularly in motion estimation and compensation. He is currently completing the PhD program.

JUAN OEST SANCHEZ holds a degree in telecommunications engineering. In 1988 he joined the Image Processing laboratory of RTVE where he participated, in different areas of activity, in the development of the project EU256. Since 1989 he has been responsible of the laboratory within the HDTV department.

JOSÉ IGNACIO RONDA was born in Madrid in 1964. He received the degree of Ingeniero de Telecomunicación in 1988 from the Universidad Politécnica de Madrid.
He joined the Grupo de Tratamiento de Imágenes of the Universidad Politécnica de Madrid in 1988, where he has been engaged in research on TV and HTDV encoding, specially in the study of the statistical parameters of these signals. He is currently completing the PhD program.

REFERENCES

[1] Draft New Recommendation:
Transmission of component-coded digital video signals for contribution-quality applications at third hierarchical level of CCITT.
Recommendation G.702, Document CMTT/303, October 1989.

[2] Proposed modifications to report AD/CMTT:
Digital transmission of component-coded television signals at 30-34 Mbit/s and 45 Mbit/s.
Document CMTT/321, October 1989.

[3] N. Garcia, F. Jaureguizar, J.I. Ronda, A. Sanz:
HDTV parallel codec simulator,
Proc. Third Int. Workshop HDTV, Turin, August 1989.

[4] G.F. Barbieri, F. Molo, J.L. Tejerina:
A modular and flexible video codec architecture for application to TV and HDTV.
Proc. 16th Int. TV Symposium, Montreux, June 1989, pp. 410-420.

[5] C.F. Hall, H.C. Andrews:
Digital color image compression in a perceptual space.
Proc. SPIE, Vol. 149, pp. 182-188, 1978.

[6] C.F. Hall:
Perceptual coding in the Fourier transform domain.
NTC Convention Record 1980, pp. 36.1.1-36.1.7.

[7] J. Mannos, D. Sakrinson:
The effects of a visual fidelity criterion on the encoding of images.
I.E.E.E. Trans. Information Theory, Vol. IT-20, No. 5, pp. 525-536, 1974.

[8] R.J. Clarke, B. Tech:
Spectral response of the discrete cosine and Walsh Hadamard transforms.
Proc. I.E.E.E., Vol. 130, No. 4, pp. 309-313, June 1983.

[9] M.J. Knee, N.D. Wells:
Designing quantizers for transform coding.
Document CMTT/2-DCT-25, April 1988.