CCITT SGXV
Working Party XV/1
Specialists Group on
Coding for Visual Telephony

Document #373
September 1988

Title: Specification for p*64 kbit/s Flexible Hardware and some proposed additions to Recommendation H.261

Source: France, FRG, Italy, Netherlands, Norway, Sweden, UK

This document is the first draft of a Flexible Hardware Specification for a video codec operating at bit rates between 64 and 2048 kbit/s. It is based on the following documents:

> CCITT Rec. H.261 (was H.12x) - Codec for audiovisual services at n*384 kbit/s.
>
> CCITT SGXV Specialists Group on Coding for Visual Telephony Document #249 - Updated Specification for the Flexible Prototype n*384 kbit/s Video Codec.
>
> CCITT SGXV Specialists Group on Coding for Visual Telephony Document #339 - Reference Model 5

and ongoing studies in Europe.

Changes from H.261 are indicated by a vertical line in the left margin.

CONTENTS

An outline block diagram of the codec is given in Figure 1.

1. Source Coder

1.1 Source format

The source coder operates on non-interlaced pictures occurring 30000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is +/-50ppm.

Pictures are coded as luminance and two colour difference components (Y, C_R and C_B). These components and the codes representing their sampled values are as defined in CCIR Recommendation 601.

> Black = 16
> White = 235
> Zero colour difference = 128
> Peak colour difference = 16 and 240

These values are nominal ones and the coding algorithm functions with input values of 0 through to 255.

Two picture scanning formats are specified.

In the first format (CIF), the luminance sampling structure is 288 lines per picture, 352 pels per line in an orthogonal arrangement. Sampling of each of the two colour difference components is at 144 lines, 176 pels per line, orthogonal. Colour difference samples are sited such that their block boundaries coincide with luminance block boundaries as shown in Figure 2. The picture area covered by these numbers of pels and lines has an aspect ratio of 4:3 and corresponds to the active portion of the local standard video input.

The second video format (QCIF) has half the number of pels and half the number of lines stated above. The maximum picture rate is under study and might be limited to half the 29.97 Hz rate above.

Codecs operating at small values of p might have QCIF only. This is under study.

1.2 Video source coding algorithm

The video coding algorithm is shown in generalised form in Figure 3. The main elements are prediction, block transformation, quantisation and classification.

The prediction error (INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the data rate control strategy. Transmitted blocks are transformed and resulting coefficients are quantised and variable length coded.

1.2.1 Prediction

The prediction is inter-picture and may be augmented by motion compensation (§1.2.2) and a spatial filter (§1.2.3).

1.2.2 Motion compensation

Motion compensation is optional in the encoder. The decoder will accept one vector for each block of 8 pels by 8 lines. The range of permitted vectors is under study.

A positive value of the horizontal or vertical component of the motion vector signifies that the prediction is formed from pels in the previous picture which are spatially to the right or below the pels being predicted.

Motion vectors are restricted such that all pels referenced by them are within the coded picture area.

1.2.3 Loop filter

The prediction process may be modified by a two-dimensional spatial filter which operates on pels within a predicted block.

The filter is separable into one dimensional horizontal and vertical functions. Both are non-recursive with coefficients of 1/4, 1/2, 1/4. At block edges, where one of the taps would fall outside the block, the peripheral pel is used for two taps. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is one half are rounded up.

The filter may be switched on or off on a block by block basis. The method of signalling this is under study.

The inclusion of alternative filter designs and characteristics is under study.

1.2.4 Transformer

Transmitted blocks are coded with a separable 2-dimensional Discrete Cosine Transform of size 8 by 8. The input to the forward transform and output from the inverse transform have 9 bits. The arithmetic procedures for computing the transforms are under study.

Note: The output from the forward and input to the inverse are likely to be 12 bits.

1.2.5 Quantisation

The number of quantisers, their characteristics and their assignment are under study.

1.2.6 Clipping

To prevent quantisation distortion of transform coefficient amplitudes causing arithmetic overflow in the encoder and decoder loops, clipping functions are inserted. In addition to those in the inverse transform, a clipping function is applied at both encoder and decoder to the reconstructed picture which is formed by summing the prediction and the prediction error as modified by the coding process. This clipper operates on resulting pel values less than 0 or greater than 255, changing them to 0 and 255 respectively.

1.3 Data rate control

Sections where parameters which may be varied to control the rate of generation of coded video data include processing prior to the source coder, the quantiser, block significance criterion and temporal subsampling. The proportions of such measures in the overall control strategy are not subject to recommendation.

When invoked, temporal subsampling is performed by discarding complete pictures. Interpolated pictures are not placed in the picture memory.

1.4 Forced updating

This function is achieved by forcing the use of the INTRA mode of the coding algorithm. The update interval and pattern are under study.

2. Video Multiplex Coder

2.1 Data Structure

Note 1: Unless specified otherwise the most significant bit is transmitted first.
Note 2: Unless specified otherwise Bit 1 is transmitted first.
Note 3: Unless specified otherwise all unused or spare bits are set to '1'.

2.2 Video Multiplex arrangement

The video multiplex is arranged in a hierarchical structure with four layers. From top to bottom the layers are:

Picture
Group of Blocks (GOB)
Macroblock (MB)
Block

2.2.1 Picture Layer

Data for each Picture consists of a Picture Header followed by data for GOBs. The structure is shown in Figure 4. Picture Headers for dropped pictures are not transmitted.

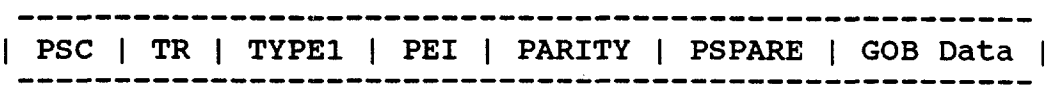

Figure 4 Structure of the Picture Layer

Picture Start Code (PSC)

A unique word of 21 bits which cannot be emulated by error free data. Its value is under study.

Temporal Reference (TR)

A five bit number derived using modulo-32 counting of pictures at 29.97 Hz.

Type Information (TYPE1)

Information about the complete picture;

Bit 1 Split screen indicator. '0' off, '1' on.
Bit 2 Document camera. '0' off, '1' on.
Bit 3 Freeze Picture Release. Under study.
Bit 4 Format indicator. '0' QCIF, '1' CIF.
Bits 5 to 12 Under study.

Extra Insertion Information (PEI)

Two bits which signal the presence of the following two optional data fields.

Parity Information (PARITY)

For optional use and present only if the first PEI bit is set to '1'. Eight parity bits each representing odd parity of the aggregate of the corresponding bit planes of the locally decoded PCM values of Y, C_R and C_B in the previous picture period.

Spare Information (PSPARE)

Sixteen bits are present when the second PEI bit is set to '1'. The use of these bits is under study.

2.2.2 Group of Blocks Layer

Each picture is divided into Groups of Blocks (GOBs). A group of blocks (GOB) comprises one twelfth of the CIF or one third of the QCIF picture areas. A GOB relates to 176 pels by 48 lines of Y and the spatially corresponding 88 pels by 24 lines of each of C_R and C_B.

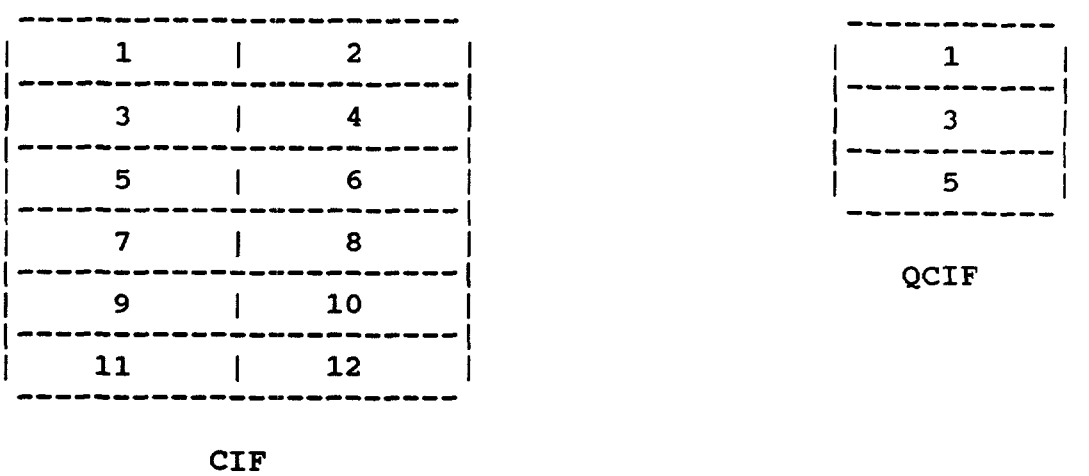

Figure 5 Arrangement of GOBs in a Picture

Data for each Group of Blocks consists of a GOB Header followed by data for macroblocks. The structure is shown in Figure 6. Each GOB Header is transmitted once between Picture Start Codes in the CIF or QCIF sequence numbered above, even if no macroblock data is present in that GOB.

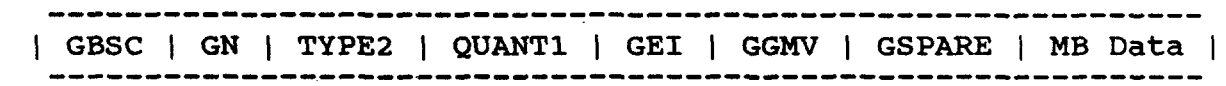

Figure 6 Structure of Group of Blocks Layer

Group of Blocks Start Code (GBSC)

A word of 16 bits, 0000 0000 0000 0001.

Group Number (GN)

Four bits indicating the position of the group of blocks. The bits are the binary representation of the numbers in Figure 5. Group numbers 13, 14 and 15 are reserved for future use. Group number 0 is used in the PSC.

Note: GBSC plus the following GN is not emulated by error-free video data.

Type Information (TYPE2)

TYPE2 is p bits which give information about all the transmitted blocks in a group of blocks. The value of p is under study.

Bit 1 When set to '1' indicates that all the transmitted blocks in the GOB are coded in INTRA mode and without block addressing data.

Bit 2 Under study. Possible uses include signalling of the use of motion compensation and the method of switching the loop filter.

Bit 3 Number of classes. '0' one, '1' four.

Bits 4 to p Spare, under study.

Quantiser Information (QUANT1)

A codeword of up to j bits which indicates the quantiser table(s) to be used in the group of blocks until overridden by any subsequent QUANT2. The codewords are under study.

Extra Insertion Information (GEI)

Under study.

Group of Blocks Global Motion Vector (GGMV)

Under study.

Spare Information (GSPARE)

Under study.

2.2.3 Macroblock Layer

Each GOB is divided into 33 macroblocks as shown in Figure 7. A macroblock relates to 16 lines by 16 pels of Y and the spatially corresponding 8 lines by 8 pels of each of C_R and C_B.

1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	17	28	29	30	31	32	33

Figure 7 Arrangement of macroblocks in a GOB

Data for a macroblock consists of a MB Header followed by data for blocks. The structure and order of the data in a macroblock are under study. A likely structure is shown in Figure 6. Elements are omitted when not required.

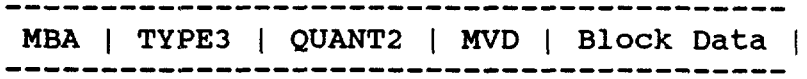

Figure 8 Structure of macroblock layer

Macroblock Address (MBA)

A Variable Length Code-word indicating the position of a macroblock within a group of blocks. VLC codewords using a combination of relative and absolute addressing are under study.

The transmission order and addressing of macroblocks are under study. Macroblocks are not transmitted when they contain no information for that part of the picture.

When bit 1 of TYPE2 is '1' MBA is not included and up to 33 macroblocks beginnining with and continuing in the above transmission order are transmitted before the next GOB Header.

Type Information (TYPE3)

Variable length codewords giving information about the macroblock and which data elements are present. Macroblock types and VLC codewords are under study.

Quantiser (QUANT2)

QUANT2 is present only if so indicated by TYPE3.

A codeword of up to j bits signifying the quantiser table(s) to be used for this and any following blocks in the group of blocks until overridden by any subsequent QUANT2.

Codewords for QUANT2 are the same as for QUANT1.

Motion Vector Data (MVD)

Calculation of the vector data is under study.

When the vector data is zero, this is signalled by TYPE3 and MVD is not present.

When the vector data is non-zero, MVD is present consisting of a variable length codeword for the horizontal component followed by a variable length codeword for the vertical component.

Variable length coding of the vector components is under study.

(See note in §2.2.4)

2.2.4 Block Layer

A macroblock comprises four luminance blocks and one of each of the two colour difference blocks.

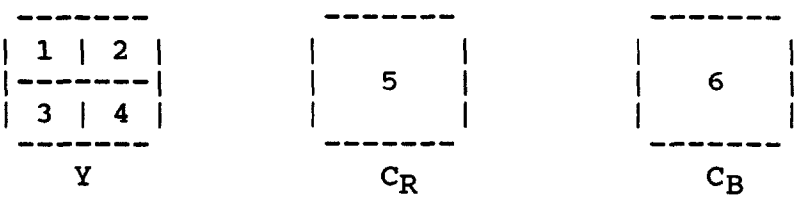

Figure 9 Arrangement of blocks in a macroblock

Data for a block consists of ..... codewords for transform coefficients followed by an end of block marker. The structure is under study.

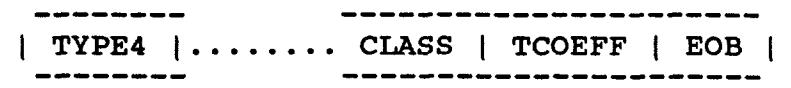

Figure 10 Structure of Block Layer

(Note: For QCIF, motion vectors per block may be necessary. Hence, the definition of TYPE4 is open and other items may be added into the block structure.)

Type Information (TYPE4)

.....

Classification Index (CLASS)

CLASS is present if bit 3 of TYPE2 is set to '1' and indicates which of the four available transmission sequence orders is used for luminance block coefficients. If bit 3 of TYPE2 is set to '0' then luminance block coefficients are transmitted in the default sequence order.

Chrominance block coefficients are transmitted in one sequence order.

The CLASS codewords and sequence orders are under study.

Transform Coefficients (TCOEFF)

The quantised transform coefficients are sequentially transmitted according to the sequence defined by CLASS. The DC component is always first. Coefficients after the last non-zero one are not transmitted.

The coding method and tables are under study.

End of Block Marker (EOB)

Use of and codeword for EOB are under study. An EOB without any transform coefficients for a block is allowed.

2.3 Multipoint considerations

Both switched and continuous presence multipoint are under study. The feasibility of the latter is increased by the two formats of CIF and QCIF together with the configuration of GOBs,

which permit four QCIF sources to be combined into one CIF stream.

2.3.1 Freeze Picture Request

Causes the decoder to freeze its received picture until a picture freeze release signal is received or a timeout period has expired. The transmission method for this control signal and the timeout period are under study.

2.3.2 Fast Update Request

Causes the encoder to encode its next picture in INTRA mode with coding parameters such as to avoid buffer overflow. The transmission method for this control signal is under study.

2.3.3 Picture Freeze Release

A signal from an encoder which has responded to a Fast Update Request and causes a decoder to exit from its picture freeze mode and display decoded pictures in the normal manner. The transmission method for this signal is under study. (See §2.2.1, TYPE1, bit 3)

2.3.4 Data continuity

The protocol adopted for ensuring continuity of data channels in a switched multipoint connection is handled by the message channel. Under study.

3. Transmission coder

3.1 Bit rate

The bit rate including audio and optional data channels is p*64 kbit/s where p is an integer between 1 and 30 both inclusive. Some codecs may have restrictions on the available values of p.

The source and stability of the encoder output clock are under study.

3.2 Video Data Buffering

Under study. The specification will cover both post-coding and pre-coding buffers. The effect on overall system delay will be considered.

Encoder buffer overflow and underflow are not permitted.

3.3 Video clock justification

Video clock justification is not provided.

3.4 Frame structure

3.4.1 Frame structure for 384-2048kbit/s channel

The frame structure is defined in Recommendation H.222.

3.4.2 Bit assignment in application channel

Under study.

3.4.3 Timeslot positioning

According to Recommendation I.431.

3.5 Audio coding

For p >= x, Recommendation G.722 56/48 kbit/s audio, 0/8 kbit/s data and 8 kbit/s service channel in the first timeslot.

For p < x, another audio coding sceme.

(x = 2?)

The delay of the encoded audio relative to the encoded video at the channel output is under study.

3.6 Data transmission

One or more timeslots may be allocated as data channels of 64 kbit/s each. The first channel uses the fourth timeslot.

Positioning of the other channels, and possible restrictions on availability at lower overall bit rates are under study. The BAS codes used to signal that these data channels are in use are specified in Recommendation H.221.

3.7 Error handling

Video coding strategy to be error resilient preferably without internal or external error correction. Note that demand refresh can be implemented using the Fast Update Request of §2.3.2.

3.8 Encryption

3.9 Bit Sequence Independence Restrictions

3.10 Network interface

Access at the primary rate is with vacated timeslots as per Recommendation I.431.

For 1544 kbit/s interfaces the default H0 channel is timeslots 1 to 6.

For 2048 kbit/s interfaces the default H0 channel is timeslots 1-2-3-17-18-19.

Interfaces using ISDN basic accesses are under study. (Recommendation I.420)

END

Figures 1, 2 and 3 are not included. They are as in Annex 4 to Doc. #346R.