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

Document #467
February 1989

SOURCE: JAPAN
TITLE : REVISION OF p*64 kbit/s FLEXIBLE HARDWARE SPECIFICATION

The following is a proposed text for revised specification of the p*64 kbit/s Flexible Hardware as a summary of the work in Japan. Revised parts to the original text in Annex 3/Doc. #445R are inserted with a different type head or striked out with hyphens.

Specification for p*64 kbit/s Flexible Hardware

CONTENTS

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

Figure 1 Outline blockdiagram of the codec

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, CR and CB). 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.

Figure 2 Positioning of luminance and chrominance samples

The second video format (QCIF) has half the number of pels and half the number of lines stated above.

The maximum picture rate (CIF and QCIF) is under study. Means shall be provided to restrict the maximum picture rate of encoders by having at least 0, 1, 2 or 3 non-transmitted pictures between transmitted ones. Selection of this minimum number will be performed manually.

Codecs operating at small values of p might have QCIF only.

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.

Figure 3 Video coding algorithm

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 the four luminance blocks in each macroblock ~~luminance and chrominance block of 8 pels by 8 lines~~. Both horizontal and vertical components of the motion vectors ~~for luminance blocks~~ have integer values not exceeding +/- 15. Motion vectors for colour difference blocks are derived from the vectors for the four corresponding luminance blocks. ~~In the case where the vectors for these four luminance blocks are identical~~ the chrominance block vector is obtained by halving the component values and truncating towards zero. ~~The derivation method for other cases 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 ~~(PROVISIONAL)~~

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 world 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 is ~~may be~~ switched on or off ~~on a block by block basis. The method of signalling this is under study.~~ depending on TYPE3 information on a macroblock by macroblock basis. The filter is switched on when the DMV field in the macroblock layer exists. When it does not exist, the filter is switched off. The existence of DMV field is indicated by TYPE3 (see Table X).

Note: An encoder without motion estimation can switch on the loop filter by sending DMV value of zero.

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

~~Note that the possibility of a filter before the picture memory is not yet ruled out.~~

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 transfer function of the inverse transform is given by

$$f(x, y) = 1/4 \ \Sigma\, \Sigma\, F(u, v) \cos(\quad) \cos(\quad)$$

The arithmetic procedures for computing the transforms are not defined, but the inverse one should meet the error tolerance specified in Appendix.

1.2.5 Quantisation

The number of quantisers is 32. Their characteristics are under study. Assignment of quantisers is as in Figure 4.

Figure 4 Quantiser assignment

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

173

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 pattern is not defined. For control of accumulation of inverse transform mismatch error a block should be forcibly updated at least once per every 128 times it is transmitted.

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 5. Picture Headers for dropped pictures are not transmitted.

PSC	TR	TYPE1	PEI1	PARITY	PEI2	PSPARE	GOB Data

Figure 5 Structure of the Picture Layer

174

Picture Start Code (PSC)   20 bits

A word of 20 bits. Its value is 0000 0000 0000 0001 0000

Temporal Reference (TR)   5 bits

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

Type Information (TYPE1)   13 bits

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.   /strike out???/
Bits 5 to 13 Under study.

Extra Insertion Information (PEI1) 1 bit

A bit which signals the presence of the following optional data field.

Parity Information (PARITY)   0 or 8 bits

For optional use and present only if the PEI1 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, CR and CB in the previous picture period.

Extra Insertion Information (PEI2) 1 bit

A bit which signals the presence of the following optional data field.

Spare Information (PSPARE)   0 or 16 bits

Sixteen bits are present when the PEI2 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 CR and CB.

175

1	2
3	4
5	6
7	8
9	10
11	12

CIF

1
3
5

QCIF

Figure 6 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 7. 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.

GBSC	GN	TYPE2	QUANT1	GEI	GSPARE	MB Data

Figure 7 Structure of Group of Blocks Layer

Group of Blocks Start Code (GBSC) 16 bits

A word of 16 bits, 0000 0000 0000 0001.

Group Number (GN) 4 bits

Four bits indicating the position of the group of blocks. The bits are the binary representation of the numbers in Figure 6. Group numbers 13, 14 and 15 are reserved for future use. Group number 0 is used in the PSC.

Type Information (TYPE2) 6 bits

TYPE2 is 6 bits which give information about all the transmitted blocks in a group of blocks.

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~~2 to 6 Spare, under study.

Quantiser Information (QUANT1) 5 bits

A fixed-length codeword of 5 bits which indicates the quantiser table(s) to be used in the group of blocks until overridden by any subsequent QUANT2. The

176

codewords are under study.

Extra Insertion Information (GEI) 1 bit

A bit which signals the presence of the following optional data field.

Spare Information (GSPARE) 0 or 16 bits

Sixteen bits are present if the GEI bit is set to '1'. The use of these bits is under study.

2.2.3 Macroblock Layer

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

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	27	28	29	30	31	32	33

Figure 8 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 9. Elements are omitted when not required.

<table>
<tr><td rowspan="2">MBA</td><td rowspan="2">TYPE3</td><td rowspan="2">QUANT2</td><td>MVD</td><td>CBP</td><td rowspan="2">Block Data</td><td>INTER</td></tr>
<tr><td colspan="2">DC_OFFSET</td><td>INTRA</td></tr>
</table>

Figure 9 Structure of macroblock layer

Macroblock Address (MBA) Variable Length

A Variable Length Code-word indicating the position of a macroblock within a group of blocks. The transmission order is as shown in Figure 8. VLC codewords using a combination of relative and absolute addressing 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 beginning with and continuing in the above transmission order are transmitted before the next GOB Header.

~~Note: More explicit detail is required about what is and what is not present when this is invoked.~~

177

Type Information (TYPE3)   Variable Length

Variable length codewords giving information about the macroblock and which data elements are present. Macroblock types are defined in Table X with transmitted elements in each macroblock type, and VLC codewords are under study.

Table X TYPE3 and elements to be transmitted

	Attribute				Transmitted Elements						
TYPE3	Intra/ Inter	MC=0?	Filter on/off	Coded on/off	BA	TYPE3	QUANT2	MVD	CBP	DC_ OFFSET	BLOCK DATA
1	Intra	-	-	on	X*	X*				X	X
1Q	Intra	-	-	on	X*	X*	X			X	X
1N	Intra	-	-	off	X*	X*				X	
2	Inter	yes	off	on	X	X			X		X
2Q	Inter	yes	off	on	X	X	X		X		X
3	Inter	d.c.	on	off	X	X		X			
3Q	Inter	d.c.	on	off	X	X	X	X			
4	Inter	d.c.	on	on	X	X		X	X		X
4Q	Inter	d.c.	on	on	X	X	X	X	X		X
5	for future extension				for futrue extension						

* not present when Bit1/TYPE2='1'
d.c. don't care

Quantiser (QUANT2)   5 bits   Fixed Length

QUANT2 is present only if so indicated by TYPE3.

A codeword of 5 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)   Variable Length

Motion vector data is obtained by differential calculation from the preceding vector. ~~(More detail needed here and also may need to cover case of one vector per luminance block.)~~ The prediction vector for macroblocks 1, 12 and 23 is zero. If the preceding block is not transmitted or INTRA coded, the prediction vector is zero.

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~~consists 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)~~

Coded Block Pattern (CBP)   Variable Length Code

A codeword of up to ? Bits which indicate those blocks for which data is present. ~~Under study.~~ Coded block pattern and its codewords are depicted in Figure Y.

(to be filled later)

Figure Y   Coded block pattern and its codewords

DC Offset (DC_OFFSET)   Fixed Length Code

Three codewords of 8 bits each signifying DC offset values in the transform coefficient domain; one codeword for the luminance component followed by two codewords for the chrominance components CR and CB.

2.2.4 Block Layer

Note~~: This part is not yet well defined. Information given below is incomplete and subject to amendment.~~

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

```
     -------   ---
    | 1 |   | |   | CR
  Y |-------| |---|
    |   |   | | 6 | CB
     -------   ---
```

Figure 10 Arrangement of blocks in a macroblock

179

Data for a coded block consists of codewords for transform coefficients followed by an end of block marker. For the INTRA DC components, differential values relative to corresponding DC_OFFSET values are transmitted. ~~The-structure-is-under-study.~~ The order of block transmission is ~~under-study~~defined in Figure 10.

```
 --------          ---------------------
+-~~TYPE4~~-+---------~~CLASS~~ | TCOEFF | EOB |
 --------          ---------------------
```

Figure 11 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)~~

~~Note-that-CLASS-might-be-removed-from-specification-if-only-the-zig-zag scanning-order-is-incorporated.~~

~~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~~in Figure Z. The DC component is always first. Coefficients after the last non-zero one are not transmitted. Hardware for 2-D VLC should be incorporated with possibility to operate in 1-D mode.

The VLC tables having maximum codelength of 24 bits are under study.

180

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

Figure Z Scanning order in a block

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. Desirable values of p are 1, 2, 6, 24 and 30. The corresponding video rates are:

~~about 40 kbit/s --- (to be defined)~~46.4 kbit/s
64 kbit/s
320 kbit/s
1472 kbit/s
1856 kbit/s

~~The source and stability of the encoder output clock are under study.~~
The codec output clock rate source shall be switchable between either a free running internal source or a source synchronized to the received data from the network. When in free running mode the tolerance on output clock rate will be +/- 50ppm of nominal. When in synchronized mode the synchronism should be maintained when the frequency of the received data clock is within +/-50 ppm of nominal.

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. Likely to be specified as maximum allowable number of bits for one picture.

Encoder buffer overflow and underflow are not permitted. Underflow can be prevented by use of dummy block in the error corrector block framing. (See ~~Document #412 for principle.~~Figure XX)

3.3 Video clock justification

Video clock justification is not provided.

3.4 Frame structure

3.4.1 Frame structure

The frame structure is defined in Recommendations H.221 and H.222.

3.4.2 Bit assignment in application channel

~~Under study.~~ When Application Channel is present, bits 17 to 80 are set to '0' until agreed otherwise.

3.4.3 Timeslot positioning

According to Recommendation I.431.

3.5 Audio coding

For p $\geq$ 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 scheme.

(x = 2?)

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

3.6 Data transmission

For p>2 one or more timeslots may be allocated as data channels of 64 kbit/s each. The first channel uses the fourth timeslot of a primary rate interface.

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.~~

~~Framing for error correction.~~

~~Block length.~~

~~BCH or Reed Solomon~~

~~Framing alignment pattern~~

3.7 Forward Error Correction for Coded Video Signal

3.7.1 Error correcting code

Double error correcting BCH(511,493)BCH

3.7.2 Generator polynomial

$$g(x)=(x^9+x^4+1)(x^9+x^6+x^4+x^3+1)$$

3.7.3 ECC bits transmission

Independent framing is used with multiframe of 8 frames, where a 512-bit frame consists of 1-bit framing, 493-bit data and 18-bit ECC (Error Correcting Code). The frame alignment pattern is $(S_0S_1S_2S_3S_4S_5S_6S_7)$=(0001101X), where X is a reserved bit for future multiframe use and set to 1. See Figure XX for the frame arrangement. Note that the ECC is calculated against the 493-bit including Fill Indicator.

183

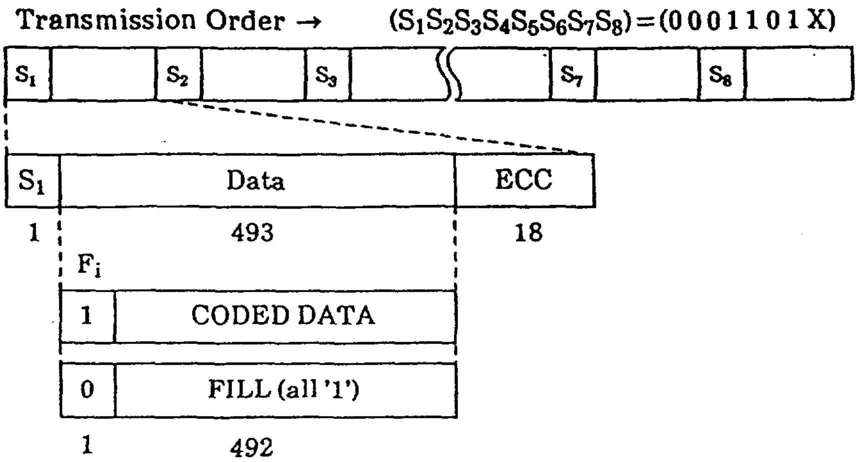

Figure XX Error correcting frame

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 access - Recommendation I.420.

END

184