CCITT SG XV
Specialists Group on Coding for Visual Telephony

Document #148
November 1986

Source: NTT, KDD, NEC and FUJITSU

Title: Proposal of video multiplex coder structure and specification

---

1. Introduction

Definition of video multiplex is urgent. In this document video multiplex coder structure is proposed.

2. Proposal

1. Higher layer attributes are given priority in a three-layered structure (picture/GOB/block)

2. Introduction of spare information bits in PSC and GBSC for optional and future purposes

3. A multiple of 8-bit length for PSC and GBSC

3. Precise description of proposals and their background

3.1 Layered structure

It has already been approved that a three-layered structure is employed in the video multiplex coder in which picture, GOB and block are defined as layers. However, there still remains open what sort of structure should be introduced for video multiplexing.

Two possible approaches are compared.

(1) Upper layers have priority.
(2) Lower layers have priority.

Here, representation of "upper" is used in a way that "picture" is "upper" than "GOB" and "GOB" is "upper" than "block". Representation of "lower" is used in an opposite way.

The first approach proposed here is that upper layer attributes have priority. When the coding attributes such as coding mode (intra-mode, inter-mode, fix-mode, etc.) and quantizer index are defined in upper layer, those attributes cannot be changed in the lower layer.

In the second approach, the parameters provisionally defined in the higher layer can be overridden in the lower layer. For example, fixed mode is provisionally defined in PSC, and attribute in lower layer of coded area such as GBSC and block, is then overridden by intraframe mode or interframe mode.

Following two items need discussion.

Overhead:

Method (1) needs less overhead information, since additional attribute information for the lower layer need not be transmitted when upper layer attributes are defined. In method (2), indication bits which identify substitutions of upper layer attributes by lower layer ones are needed. Indication bit that costs 1 bit/block needs 2,000 bit/picture in the worst case. Therefore, method (1) is much preferable.

Ease of coding control or hardware simplicity:

Almost the same. In the method (1) two treatments are available. In the first treatment that coding attributes are defined before coding process, coding control can be exploited with simple hardware. For example, when a GOB is defined as intra-mode after "scene change" is detected, the mode in GBSC is defined as intra-mode in advance of coding process. Meanwhile, in the second treatment that coding attributes are defined after coding process, hardware complexity increases somewhat. As an example, when all the blocks in a GOB become Fix mode after quantization, the parameter in GBSC needs to be re-defined as "fix" if we want efficient representation of overhead information. In case of method (2), the control of this kind is easier than the second treatment of method (1), since the mode which is defined in upper layer can be overridden in lower layer. Namely, when we introduce the first treatment of the method (1), hardware size is almost the same with method (2).

Therefore, we should better introduce the first method that "upper" layer has priority since disadvantage of method (2) is not easy to solve. In annex 1, examples are described in which method (1) is used effectively.

3.2 Spare information bits in PSC and GBSC

There need transmission of information which has to be synchronous to video frame. For example, in the case of demand refreshing, which was proposed as an option, parity information to be transmitted should coincide with frame coded data. Therefore, the structure is definitely appropriate in which parity information is inserted in PSC. In addition, indication of picture composition (split-screen or else) and indication of source camera (participants/documents) which should be synchronous to video frame can also be coded by using spare information bits.

3.3 PSC and GBSC length is a multiple of 8-bit

For hardware simplicity, length of PSC and GBSC should better be a fixed length and a multiple of 8-bit. Picture and GOB start codes can easily be detected before the receiving buffer memory where the rate of data stream is quite moderate. Decoder need not have bit-by-bit shifter after the buffer memory for start code identification. (Of course, a bit-by-bit shifter will be needed for decoding of variable length codes.) Therefore, Picture start code and GOB start code detection circuit can easily be designed and checked.

Use of "future use bit" has a relation to this. Though FB bit (Future use bit) is inserted for making the bit length of the PSC and GBSC multiples of 8 bit. It may help to insert some additional information in the future if necessary.

4. Conclusion

We propose three major structures for video multiplex coder.

1. Priority for higher layer attributes in a three-layered structure (picture/GOB/block)

2. Introduction of spare information bits in PSC and GBSC for optional purposes

3. A multiple of 8-bit length for PSC and GBSC

ANNEX 2 is a draft of Video multiplex coder specification which introduces above structures.

Annex 1

Examples of Layer structured attribute

Layered structure is studied to reduce block type and quantizer type informaiton and to improve picture quality. The followings are examples which take an advantage of this arrangement effectively.

Example 1 Forced update case

When the forced update mode is activated on a GOB basis, intra-frame coding mode should not be specified on a block-by-block basis but on a GOB basis.

PSC: The mode is not specified for Type 1 (BTYPE)
GBSC: Intra mode is defined for Type 1 (BTYPE)
Block: Not specified

When intra mode is chosen, no information is assigned for Block by block basis. Therefore it reduces overhead information by approximately 3 bit for each block. 400 bits/GOB can be saved.

Example 2.

When the scene change is detected, intraframe coding is employed for the whole picture. Then,

PSC: Intra mode is chosen for Type 1 (Btype) .
GBSC: Not specified
Block: Not specified

7,200 bits/picture can be saved.

Example 3.

When the quantizer characteristics is switched in GOB unit, Qtype is specified in GBSC so that overhead information is reduced compared to when it is specified in block unit.

PSC: Not specified
GBSC: Quantizer type is defined as Type 2 (Qtype)
Block: Not specified

12,000 bit can be saved in a picture.

This method is preferable to the so-called override method, since no identification bit is needed for block information.

Annex 2

2. Video multiplex Coder

2.1 Tasks

The video multiplex coder has the following tasks:

1) Block address coding
2) Video data formatting and serializing
3) Synchronization - picture/line
4) Motion vector coding
5) Side channels for indicating dynamic coding parameters e.g., subsampling modes, quantizers, buffer state etc. (Permanent or transient channels? Transient channels require care in switched multipoint.)

Figure 2.1 shows the configuration of video multiplex coder. Layer structure is introduced in the multiplex coder for defining coding parameters.

(1) Intra/interframe mode,
(2) coded/non-coded,
(3) quantizer type and
(4) motion vectors

are multiplexed as the block attribute. As for the first three items, they are defined exclusively in the layers, namely, when all attributes in a GOB/Picture have the same value, the attributes are defined in the GOB/Picture and not defined in each block. For multiplexing motion vectors, further study is needed.

2.2 Video multiplex arrangement

2.2.1 Picture start code (PSC)

Configuration of the codes is shown in Figure 2.2. All PSCs are transmitted. Two successive PSCs indicate the intervening picture has been dropped.

8*N bit

PUC	BS	TEMP	TYPE1	TYPE2	FB	SI	PGMV	PARITY	SPARE
24 bits	6 bits	3 bits	4 bits	6 bits	2 bits	3 bits	16 bits	8 bits	24 bits

Figure 2.2 Configuration of Picture start code

*Comment: For easier hardware implementation, the bit length of PSC is a multiple of 8 bits so the shifter circuits which shift the data on a bit-by-bit basis is not necessary. N equals 6. The length of PSC does not change, since all the information is fixed length codes except when spare information is inserted.

(1) Picture unique code (PUC)
Provisionally, it is a unique code of 24 bits. Maximum 24 bits.
[1000 0000 0000 0000 0000 0000]

(2) Buffer state (BS)
Coder buffer fullness information expressed in the unit of 1 Kbit. The value is sampled at the time of the top of the frame. (Six bits, MSB first)

(3) Temporal reference (TEMP)
Temporal reference is a 3 bit number representing the time sequence, in Common Intermediate Format picture periods, of a particular picture. MSB first.

(4) Type information
Two type information is used. Fixed length codes are employed.

-TYPE1
Type 1 indicates (1) Type 1 on/off, (2) Intra/interframe mode,(3) coded/non-coded mode, (4) Motion vector zero/non-zero. First bit indicates an insertion on/off. When the bit is 1, the type information is not defined in the picture attribute, but defined in the lower attributes, such as GOB and Block attributes. When the bit is 0, consecutive 3 bits indicates the modes of specific picture. Fixed length codes defined in Table 2.2 is employed.

-TYPE2
Type 2 indicates quantizer type information. First bit indicates an insertion of the information. When the first bits is 1, the quantizer type is not defined in the picture attribute, but defined in the lower attributes. When the bit is 0, consecutive 5 bit code indicates the quantizer type as picture attribute.

(5) Future use bit (FB)
For future use, spare bits are reserved. Definision is for further study.

(6) Spare bit insertion information (SI)
The three bit code indicates PGMV, PARITY and SPARE insertion. When the first bit is 1, Picture global motion vector (PGMV) is not inserted and no bit is reserved. When the bit is 0, 16 bit information is added. When the seconde bit is 1, parity information (PARITY) is not inserted and no bit is reserved. When the bit is 0, 8 bit information is added. When the third bit is 1, spare information (SPARE) is not inserted and no bit is reserved. When the bit is 0, 24 bit information is inserted. Therefore, when all SI bits equal to ones, e.g., the following PGMV, PARITY and SPARE information bits are removed, and encoded video are inserted instead.

(7) Picture global motion vector (PGMV)
Picture global motion vector consisting of 16 bit code are inserted when it is defined. The motion vectors are transmitted in absolute terms. Eight bit code for each direction. MSB first. Horizontal vector is coded first, then vertical one.

(8) Parity information (PARITY)
For demand refreshing, parity information consisting of 8 bit code is inserted on an optional basis. MSB plane first. Each bit represents odd parity of each bit plane from the local decoded signals in the previous frame period. For the dropped frame PSC, the same parity data are repeated in this PARITY information.

(9) Spare information (SPARE)
For future use, spare bits are reserved. Definition of the spare information needs further study.

2.2.2 Group of Block start code (GBSC)

8*N (GBUC to SI)

GBUC	GN	TYPE1	TYPE2	FB	SI	GGMV	SPARE
19 bits	5 bits	4 bits	6 bits	4 bits	2 bits	16 bits	16 bits

Figure 2.3 Configuration of the GOB start code

*Comment: The bit length of GBSC from GBUC to SI is 8*N and is not changed. Integer number N is 5. See comment of 2.2.1.

The structure of Group of block start code (GBSC) is shown in Figure 2.3. Each GOB consists of two block-lines of Y signal each with 8 lines by 352 pels and a block line of Cr and Cb each with 8 lines by 176 pels. See Figure 2.4.

Y upper	Y lower	Cr	Cb

Figure 2.4 A group of blocks

All GBSCs are transmitted.

*Comment: Efficiency may be improved when only coded GBSCs are transmitted. Therefore, such method is studied in Japan. Coded/non-coded GOBs can be indicated with Group Number (GN).
If GBSCs are dropped for uncoded GOB, PSCs of uncoded Pictures would be also dropped for simpler specification. Coded/non-coded pictures can be indicated with Temporal reference, however, the longer bit length of TEMP might be better for flexible frame dropping.

The following data are inserted.

(1) Group of block unique code (GBUC)
Provisionally, a unique code of 19 bits. Maximum 19 bits.
[1000 0000 0000 0000 000]

(2) Group number (GN)
The group number is a five bit number representing the vertical spatial position, in units of groups, of the current group of block. MSB first. For the first GOB, GN = 1 or (00001). For the last GOB, GN = 18 or (10010).

(3) Type information
-TYPE1
Type 1 indicates (1) Type 1 on/off, (2) Intra/interframe mode,(3) coded/non-coded mode, (4) Motion vector zero/non-zero.

Four-bit 'one's are inserted which will be disregarded in the decoder when the information is defined in a higher (or picture) attribute. In the case that the information is not defined in the picture attribute, the GOB attributes work and the first bit indicates an insertion on/off. When the bit is 1, the mode information is not defined in the GOB attribute, but defined in the lower (or block) attributes. When the bit is 0, consecutive 3 bit code indicates the modes as GOB attribute. Fixed length codes defined in Table 2.2 is employed.

-TYPE2
Type 2 indicates quantizer type information. Six-bit 'one's are inserted which will be disregarded in the decoder when the information is defined in a higher (or picture) attribute. In the case that the information is not defined in the picture attribute, the GOB attributes work and the first bit indicates an insertion on/off. First bit indicates an insertion of the information. When the first bits is 1, the quantizer type is not defined in the GOB attribute, but defined in the lower attribute. When the bit is 0, consecutive 3 bit code indicates the quantizer type as GOB attribute. The same code pattern as the block attribute is employed.

(4) Future use bit (FB)
For future use, spare bits are reserved. Definision is for further study.

(5) Spare bit insertion information (SI)
This two bit indicates GGMV and SPARE insertion. When the first bit is 1, GOB global motion vector (GGMV) is not inserted and no bit is reserved. When the bit is zero, 16 bit information is added. When the second bit is 1, SPARE information is not inserted, while when the bit is 0, 16 bit information is inserted. Therefore, when the SI bits equal to '11' for example, the following SPARE information bits are removed, and encoded video are filled instead.

(6) GOB global motion vector (GGMV)
GOB global motion vector consisting of 16 bit code is inserted when it is defined. The motion vectors are transmitted in absolute terms. Eight bit code for each direction. MSB first. Horizontal vector is coded first, then vertical one.

(7) Spare information (SPARE)
For future use, additional spare bits are reserved. Definition of the information needs further study.

2.2.3 Block data alignment

The data of block is inserted after GBSC described above. Each data is arranged on a block-by-block basis. The order of the information is shown in Figure 2.5.

(1) BA : Block address
(2) BTYPE : Block type
(3) QTYPE : Quantizer type
(4) CLASS : Scanning class index
(5) MCV : Motion vector
(6) TCOEFF: Transform coefficients
(7) EOB : End of block

Future study may change the order and code length of the code.

BA	BTYPE	QTYPE	CLASS	MCV	TCOEFF	EOB
VLC	VLC	FLC	VLC	VLC	VLC	FLC

Figure 2.5 The configuration of block data.

Data specifying block address, block type, classification index, motion vector and quantizer type are multiplexed on a block by block basis, while BTYPE and QTYPE are dropped when they are defined in the PSC/GBSC.

(1) Block address information (BA)

BA information indicates the run length of Fixed (i.e., non-coded, interframe prediction and no motion compensation) blocks. Relative block address between the previous Non-fixed block and current non-fixed block is variable length coded. Since the VWL code does not start with a code '1' as shown in Table 2.2, code '1' is used for start code of the BA information by which the BA information is decodable.

1. Block run is continuous from luminance to chrominance.
2. VWL code pattern is for further study. 16 bits maximum.

The range of absolute block addresses for the upper line of luminance blocks (Yu) is 0 to 43. The range of absolute block addresses for the lower line of luminance blocks (Yl) is 44 to 87. The range of absolute block addresses for Cr blocks is 88 to 109. The range of absolute block addresses for Cb blocks is 110 to 131.

For the Fixed mode block, no information is assigned to BTYPE, CLASS, QTYPE, MCV, TCOEFF and EOB.

(2) Block type information (BTYPE)

This variable length code indicates block attributes, such as

1. intra-picture coded block,
2. inter-picturecoded block,
3. Motion compensated block
   a. Differential MV $\neq 0$
   b. Differential MV $= 0$
4. Motion compensated block with coded residue
   a. Differential MV $\neq 0$
   b. Differential MV $= 0$
5. future expansion on,
6. future expansion off.

The variable length code pattern shown in Table 2.2 is used for block type. As for the last two items, further study will be needed. The decoder shall be designed to ignore all data between types 'future expansion on' and 'future expansion off' and also between 'future expansion on' and the next GBSC. The exact method the decoder uses to discard this data is still to be defined. To code "fix", address of luminance and two chrominances are treated as continuous.

BTYPE specifies insertion of the CLASS, MCV, TCOEFF and EOB as shown in

Table 2.1.

Table 2.1 Insertion data when the BTYPE is specified

Block type information	Insertion data in a Block
NOT CODED + INTER (FIX)	NOT INSERTED
CODED + INTER	CLASS + (QTYPE) + TCOEFF + EOB
CODED + INTRA	CLASS + (QTYPE) + TCOEFF + EOB
CODED + INTER + MC DMV ≠ 0	CLASS + (QTYPE) + DMV + TCOEFF + EOB
CODED + INTER + MC DMV = 0	CLASS + (QTYPE) + TCOEFF + EOB
NOT CODED + INTER + MC DMV = 0	NOT INSERTED
NOT CODED + INTER + MC DMV ≠ 0	DMV

* NOTE : When the type 1 (BTYPE data) is defined in the Picture or GOB attribute, BA and BTYPE in corresponding blocks is not inserted but only data specified in the Table is inserted.
* NOTE : QTYPE is not inserted when TYPE 2 is defined in Picture or GOB attribute.

(3) Quantizer type (Q TYPE)
Fixed length code of five bits indicates quantizer type. The code '0' represents the finest quantizer. The code '31' represents the coarsest quantizer. MSB first.

When the Q TYPE is defined in the TYPE 1 in Picture/GOB attribute, no bit is inserted.

(4) Classification index (CLASS)
Classification index (8 classes at maximum) for scanning is variable length coded. The code pattern is for further study. Maximum code length is 8.

Table 2.2 Code pattern of block type information (Example)

Block type information	Fixed code	VWL CODE	
		Y	Cr, Cb
NOT CODED + INTER (FIX)	0 0 0	-----------	
CODED + INTER	1 0 0	0 0 1	0 1
CODED + INTRA	0 0 1	0 0 0 0 0 1	0 0 1
CODED + INTER + MC DMV ≠ 0	1 1 1	0 1 0	no code
CODED + INTER + MC DMV = 0	1 1 0	0 0 0 1	no code
NOT CODED + INTER + MC DMV = 0	0 1 1	0 1 1	no code
NOT CODED + INTER + MC DMV ≠ 0	0 1 0	0 0 0 0 1	no code
FUTURE ON		0000 001	
FUTURE OFF		0000 0001	

* Fixed codes are used in picture and GOB attributes.
* VLC code assignment needs for further study.

(5) EOB marker (EOB)
EOB is one of the codewords in the VLC set used for quantized transform coefficients. It is always present for every block for which coefficients are transmitted. Code pattern is for further study.

2.2.4 Motion vector information (DMV)

(1) Coding method
Motion vector is previous block predicted and the difference is variable length coded on a block by block basis. The prediction function is described as :

$$dv(x,y) = v(x,y) - v1(x,y),$$

where v(x,y) and v1(x,y) indicates motion vector of current block and that of previous block. As the result of the prediction, the motion vector of horizontal and vertical direction is derived separately.

For the first block in a GOB, (0,0) is used for predictive value v1.

(2) Variable length coding

The value of 'dv(x,y)' is variable length coded. Each vector of horizontal and vertical directioon has VLC (two entries). Motion vectors for chrominance are not transmitted. The code set for motion vector is for further study. Maximum length of VLC is 16.

Only when the DMV of a block is not equal to zero, DMV data is inserted.

2.2.5 Transform coefficient data (TCOEFF)

Adaptive scan is employed for multiplexing the coded data.

2.3 Multipoint considerations

2.3.1 Freeze picture request (FPR)

On receipt of a FPR, conveyed in bit .... of time slot .... the decoder shall freeze its received video until the next picture attribute indicating intra-picture mode is encountered.

2.3.2 Fast Update Request (FUR)

On receipt of a FUR, conveyed in bit ... of time slot .... the encoder shall empty its transmit buffer and encode its next picture in intra-picture mode with such coding parameters as to avoid buffer overflow.

2.3.3 Data continuity

The protocol adopted for ensuring data continuity in a multipoint connection will be handled by the message channel.