ISO/IEC
JTC1/SC29/WG11
MPEG94/??? 214
July 1994

Telecommunication Standardization Sector
Study Group 15
Experts Group for Video Coding and Systems
in ATM and Other Network Environments

Document AVC-656
July, 1994

SOURCE :Japan
TITLE :Necessary provision for applying H.222.0 to the H.32X terminals
PURPOSE :Proposal and Discussion

1.Introduction

This document discusses necessary extensions of the H.222.0 packetized elementary stream that allow terminals to support both transport and program streams. This document is closely related to AVC-655 where allocation of multimedia channels is discussed.

2.Application of H.222.0 to H.32X

In the process of expanding H.222.0 to cope with the H.32X requirements, support for both the Program Stream (PS) and the Transport Stream (TS) is assumed as a pre-requisite. On this ground, we considered addition of several variables, including the stream_id and stream_type. We also discuss the timing and handling of modes' changes during transmission.

2.1 H.32X Elementary Stream

The streams required by H.32X are outlined in table 1.

Table 1 H.32X Elementary Streams

Stream	Coding	Maximum number of Channels	Comments
Control	ITU-T Control	1	New, Terminal negotiation data
Video	ITU-T H.262 / ISO/IEC 13818-2 Video	16	
	ITU-T Video	2*	New, H.261 CMTT.721,723, etc.
Audio	ISO/IEC 11172-3/ 13818-3 Audio	32	
	ITU-T Audio	4*	New, G.711,G.722, etc.
Data	ITU-T Data	1	New, H.32X specific data (Far End Camera Control, etc.)

*) We considered multipoint conferencing, simulcast transmission, and stereo audio as possibilities.

2.2 Syntax Extensions

The ITU-T control channel handles end-to-end signaling; terminal capability, mode setting command, indication, etc.

[1] Control Channel Multiplexing

[a] Extensions of Stream_id

A proposed code for identifying ITU-T control channel has been added to the PES packet stream_id, as shown in table 2. Accordingly, the stream_type and descriptor in the PSM and PST have been modified (tables 3, 4, 7). It should be noted that four ITU-T labelled

stream_type codes are added in table 3 and four descriptor_tag codes are added in table 4. Examples of ITU-T labelled descriptor content are shown in tables 8 - 11, which should be included in Rec.H.222.1 if finalized.

For more details of the control channel, see AVC-655.

[b] Extensions of DSM_CC

It might be possible to define a syntax that would cover the contents of the ITU-T control stream within the DSM_CC block.

However, we have not yet studied this alternative. There may be merits and demerits compared to the above solution.

[2] Video, Audio, and Data Stream Multiplexing

The PES stream_id can be expanded to identify video, audio, and data streams in the same way as the control stream (table 2).

Again, some new stream_type must be added in the PSM and PMT (tables 3, 4, 7).

Table 2 Stream_id table in PES Packet

stream_id	stream type	Comments
1011 1100	program stream map	
1011 1101	private_stream_1	
1011 1110	padding stream	
1011 1111	private_stream_2	
110x xxxx	ISO/IEC, ITU-T audio - stream number xxxxx	add ITU-T audio
1110 xxxx	ISO/IEC, ITU-T video - stream number xxxx	add ITU-T video
1111 0000	ECM stream	
1111 0001	EMM stream	
1111 0010	DSM CC stream	
1111 0011	ISO/IEC 13522 stream	
1111 0100	ITU-T Control	New
1111 0101	ITU-T Data	New
1111 xxxx	reserved data stream - number xxxx'0110' - '1110' are reserved.	
1111 1111	program stream directory	
The notation x means that the values 0 and 1 are both permitted and result in the same stream type. The stream number is given by the values taken by the x's.		

Table 3 Stream Type in PSM or PST

stream_type	Description	Comments
0x00	ITU-T | ISO/IEC Reserved	
0x01	ISO/IEC 11172 Video	
0x02	ITU-T H.262 | ISO/IEC 13818-2 Video	
0x03	ISO/IEC 11172 Audio	
0x04	ISO/IEC 13818-3 Audio	
0x05	ITU-T Rec.H.222.0 | ISO/IEC 13818-1 private_sections	
0x06	ITU-T Rec.H.222.0 | ISO/IEC 13818-1 packets containing private data	
0x07	ISO/IEC 13522 MHEG	
0x08	ITU-T Rec.H.222.0 | ISO/IEC 13818-1 DSM-CC	
0x09	ITU-T Rec.H.222.0 | ISO/IEC 13818-1 /11172-1 auxiliary	
0x0A	ITU-T Control	New
0x0B	ITU-T Data	New
0x0C	ITU-T Video	New
0x0D	ITU-T Audio	New
0x0E-0x7F	ITU-T | ISO/IEC 13818-1 Reserved	
0x80-0xFF	User Private	

2.3 Transmission Mode Identification

This section describes the timing and handling of dynamic mode changes during transmission.

2.3.1 Program Stream Map (PS) and Program Map Table (TS)

Mode changes are signaled through the PSM and PMT descriptors (see table 4). The PSM and PMT access methods are shown in figures 1 and 2. In order to minimize the effects of transmission errors, the tables are re-transmitted regularly regardless of whether they have changed or not.

Here are given two alternatives of dynamic synchronous mode switching; one by use of data_alignment_indicator, the other by stream_id.

NOTE: Since we use either the data_alignment_indicator or stream_id in this method, the PES start must be synchronized with the start of the sequence.

In the case of fixed length packets, care is needed to insure this synchronization; stuffing may be required for the last part of the sequence.

[1] Using the data_alignment_indicator (figure 3a)

[a] MPEG-2 Video

The video_stream_alignment_values in the data_stream_alignment_descriptor is set to '04' (table 5).

Changes at the encoder side (of resolution, etc.) are signaled through a PMT version update(table 7, 10), and update timing is signaled in an update of video_sequence (if the data_alignment_indicator in the PES packet is '1').

In addition, Split Screen (SSI), Document Camera (DCR), and Freeze Picture Release (FPR) are signaled in the same way.

In the case of FPR, it is also possible to substitute it with an I Picture.

[b] Audio

Audio frames are distributed aligned with the PES.

The encoding processes G.711, G.722, G.728, etc., are defined in the ITU-T_Audio_stream _descriptor of the PMT(table 7, 11).

Update of these coding methods are communicated through updates of these definitions.

In addition, the Audio sequence (see table 6) is added to the ITU-T_Audio_stream_alignment _values in the data_stream_alignment_descriptor.

Here, the Audio sequence is treated as a single unit of encoding consisting of the first bit to

the last bit encoded by an audio coding rule.

Coding method changes are indicated through a PMT version update, and update timing is signaled by Audio sequence update (if data_alignment_indicator in the PES packet is '1')

[2] Using the stream_id (figure 3b)

During changes in the transmission mode, the stream_id of the changing stream is set accordingly.

Two stream_id's are required for each coding to indicate the changes; e.g. to indicate changes by use of toggling, two id's are required for each elementary stream.

The lack of much reserved space in the stream_id makes it difficult to implement this switching method.

2.3.2 PES header expansion

Streams which need to indicate a changing mode can also use the PES_private_data or PES_extansion_flag2 in the PES header to include a PES_extension_field describing the change. Such data should be periodically repeated to protect against transmission error. Since this method uses the private data field, there is possibility of conflicting with other applications that also try to use the field.

Table 4 Stream Descriptors in PSM or PMT

descriptor_tag	TS	PS	Identification	Comments
0	n/a	n/a	Reserved	
1	n/a	n/a	Reserved	
2	X	X	video_stream_descriptor	
3	X	X	audio_stream_descriptor	
4	X	X	hierarchy_descriptor	
5	X	X	registration_descriptor	
6	X	X	data_stream_alignment_descriptor	
7	X	X	target_background_grid_descriptor	
8	X	X	video_window_descriptor	
9	X	X	CA_descriptor	
10	X	X	ISO_639_language_descriptor	
11	X	X	system_clock_descriptor	
12	X	X	multiplex_buffer_utilization_descriptor	
13	X	X	copyright_descriptor	
14	X		maximum bitrate descriptor	
15	X	X	private data indicator descriptor	
16	X	X	ITU-T_control_stream_descriptor	New
17	X	X	ITU-T_data_stream_descriptor	New
18	X	X	ITU-T_video_stream_descriptor	New
19	X	X	ITU-T_audio_stream_descriptor	New
20-63	n/a	n/a	MPEG Reserved	
64-255	n/a	n/a	User Private	

Table 5 Video Data Stream Alignment Values

alignment type	Description
00	reserved
01	Slice, picture, GOP, or SEQ
02	picture, GOP, or SEQ
03	GOP, or SEQ
04	SEQ
05-FF	reserved

Table 6 Audio Data Stream Alignments Values

alignment type	Description	Comments
00	reserved	
01	Audio frame	
02	Audio Sequence	New
03-FF	reserved	

Table 7 ITU-T_*_stream_descriptor (* = control, data, video or audio)

ITU-T_control_stream_descriptor(){		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
message	8	uimsbf
}		

Table 8 ITU-T_control_stream_descriptor message example

protocol (e.g. H.24X,H.240, etc.)	8	uimsbf
reserved	8	uimsbf

Table 9 ITU-T_data_stream_descriptor message example

number of channels	8	uimsbf
reserved	8	uimsbf
for (i = 0; i<N ; i++)		
transfer_rate	8	uimsbf
link layer protocol (e.g. HDLC, etc.)	4	uimsbf
application layer protocol	8	uimsbf
(e.g. Far end camera control, etc.)		
reserved	4	uimsbf

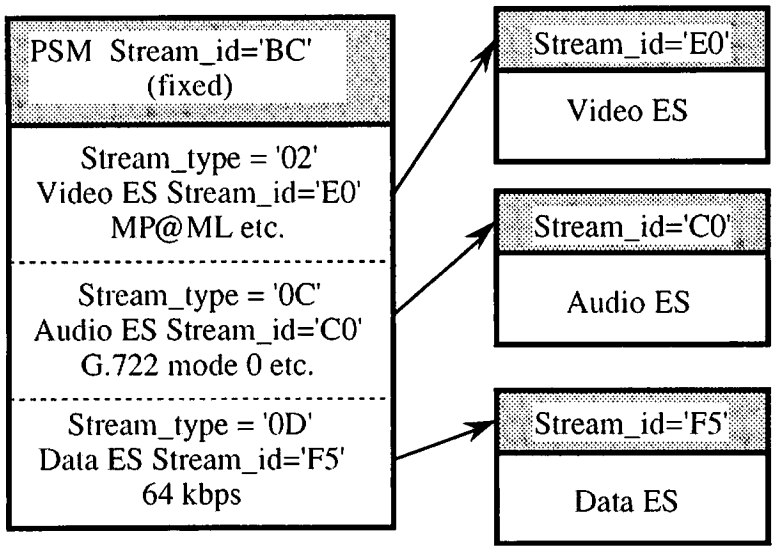
Table 10 ITU-T_video_stream_descriptor message example

coding_algolizm (e.g. H.261, CMTT.721, etc.)	8	uimsbf
split_screen	1	uimsbf
document_camera	1	uimsbf
freeze_picture_release	1	uimsbf
reserved	5	uimsbf
if (coding_algolizm = H.261){		
Format (e.g. QCIF/CIF/SCIF)	3	uimsbf
Picture_rate (e.g. 1,2,3,4/29.97)	3	uimsbf
reserved	2	uimsbf
}		
reserved	8	uimsbf

Table 11 ITU-T_audio_stream_descriptor message example

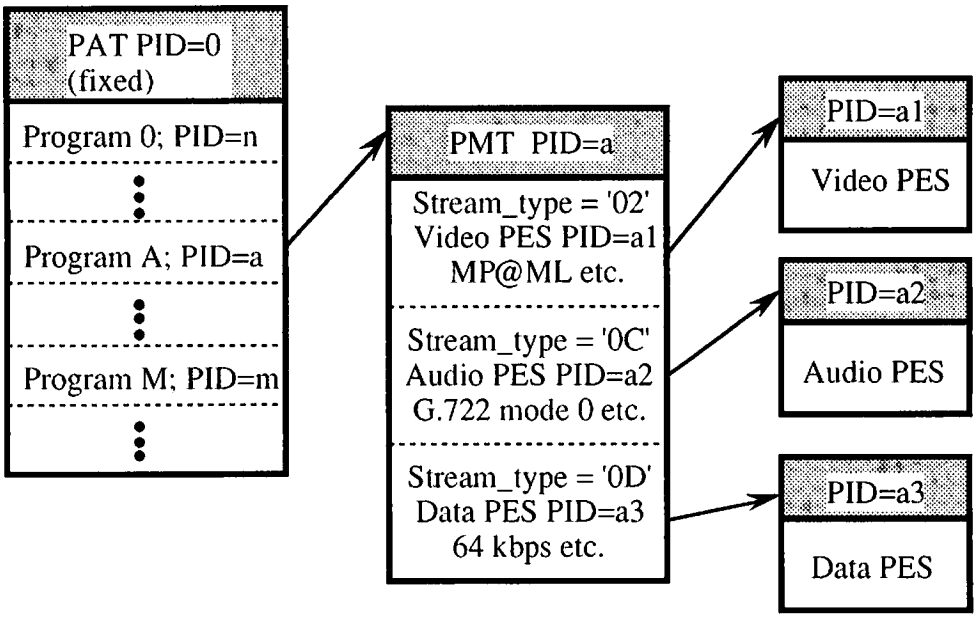
coding_algolizm (e.g. G.711, G.722, etc.)	8	uimsbf
sample_rate (e.g. 8 kHz, 16 kHz, etc.)	4	uimsbf
frame_size (e.g. 1 Byte, 5 Bytes, etc.)	4	uimsbf
reserved	8	uimsbf

AVC-656



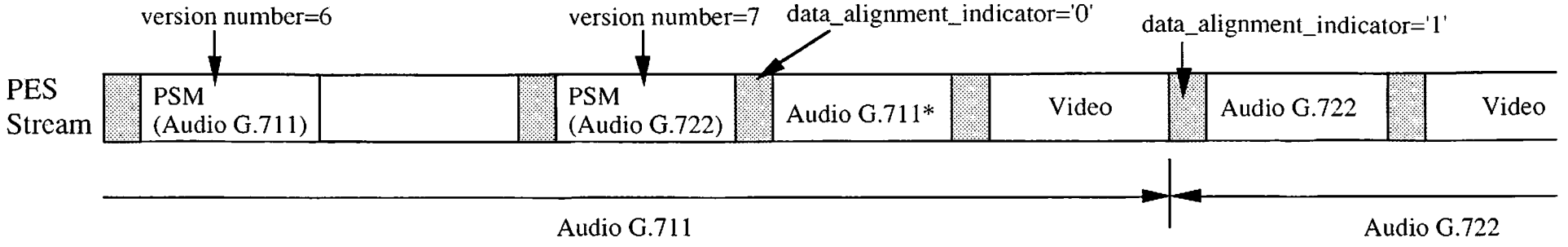
PSM; Program Stream Map
ES; Elementary Stream

Fig.1 Program Stream Map (Program Stream)

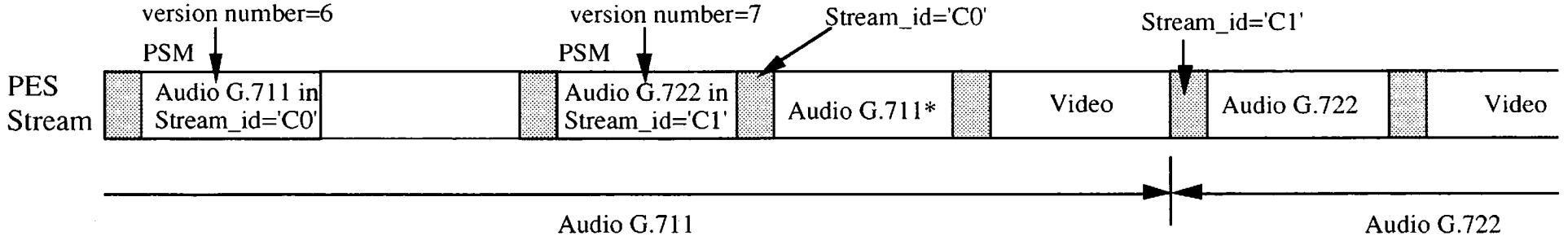


PAT; Program Association Table
PMT; Program Map Table
PES ; Packetized Elementary Stream
ES; Elementary Stream

Fig.2 TS Program and network mapping relationships



(a) Change of data_alignment_indicator



(b) Change of Stream_id

PES Header

* ; If the PES packets have fixed length, you may have to stuff in this PES packet.

Fig.3 Mode Swiching

AVC-656

3 Audio

The standardized audio encoding methods are outlined in table 12.

For applications such as teleconferencing, we must consider the delay introduced by low bit rate audio PES.

For example, when using G.711, G.728, G.722 and assuming a packet unit of 5ms,their PES sizes are reduced to 40, 10, and 40 bytes, respectively.

Also, when processing the PS with software, the CSPS limitation of 300 packets/sec should be taken into account.

So although it may be necessary to multiplex H.24X control or data channels as in (2) above, delay management remains unsolved.

Table 12 Specification of Audio coding method

Coding method	delay (approx.)	frequency band(kHz)	bitrate (kbps)	number of channels	frame size (Bytes)
G.711	125 micro sec	3.4	64	1	1
G.728	5 ms	3.4	16	1	5
G.722	250 micro sec	7	48 - 64	1	1
MPEG1 Layer 1	19 ms	15	32 - 448	stereo: 2 bilingual: 2	384*bitrate* sample frequency / 8
MPEG1 Layer 2	35 ms	15	32 - 384	stereo: 2 bilingual: 2	1152*bitrate* sample frequency / 8
MPEG1 Layer 3	59 ms	15	32 - 320	stereo: 2 bilingual: 2	1152*bitrate* sample frequency / 8

4 Conclusion

We have investigated the applicability of H.222.0 to H.32X terminals. There is no specification for ITU-T coding method in the current H.222.0 description of the PES stream_id, so some id's must be added.

We also examined the possibility of using PSM and PMT transmissions and the PES header to manage stream synchronization. We have found that the PSM and PMT version update and the PES stream data_alignment_indicator are favorable for the dynamic mode switching.

Finally, the very small size PES packets that must be used in low bit rate audio encoding and low delay operation make transmitted packet counts very large. We point out that this problem remains to be solved.

References

[1]AVC-655 Communication procedure for H.222.1(Japan), July 1994,Grimstad, Norway