International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Recommendation ITU-R BT.1365-2

(10/2015)

24-bit digital audio format as ancillary data signals in HDTV and UHDTV serial interfaces

BT Series

Broadcasting service (television)

1865 - 2015

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	**Broadcasting service (television)**
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
SNG	Satellite news gathering
TF	Time signals and frequency standards emissions
V	Vocabulary and related subjects

Note*: This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2015

RECOMMENDATION ITU-R BT.1365-2

24-bit digital audio format as ancillary data signals in HDTV and UHDTV serial interfaces

(Question ITU-R 130/6)

(1998-2010-2015)

Scope

This Recommendation defines the mapping of 24-bit digital audio data conforming with Recommendation ITU-R BS.647 and associated control information into the ancillary data space of serial digital video interfaces conforming to Recommendation ITU-R BT.1120 and Recommendation ITU-R BT.2077. The audio data are derived from Recommendation ITU-R BS.647, hereafter referred to as Audio Engineering Society (AES).

Keywords

UHDTV, Serial Interface, AES bit stream

The ITU Radiocommunication Assembly,

considering

a) that many countries are installing digital HDTV and UHDTV production facilities based on the use of digital video components conforming to Recommendations ITU-R BT.709, ITU-R BT.2020, ITU-R BT.1120 and ITU-R BT.2077;

b) that there exists the capacity within the serial digital interface for HDTV and UHDTV for additional data signals to be multiplexed as part of the serial data stream;

c) that there are operational and economic benefits to be achieved by the multiplexing of ancillary data signals along with the video data signal;

d) that audio is one of the most important uses of the ancillary data packets;

e) that audio data may need error correction codes to keep the balance between audio quality and video quality because errors in audio data are more easily noticed than those of video data;

f) that audio equipment with 24-bit accuracy is commonly used in production facilities;

g) that some broadcasters have the need to transmit asynchronous audio data by multiplexing into the serial digital interface,

recommends

1 that, for the inclusion of 24-bit digital audio format as ancillary data signals in HDTV and UHDTV serial interfaces, the specification described in Annex 1 and or Annex 2 of this Recommendation should be used;

2 that compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met.

Definition of terms

Definition of these terms applies to the usage made in this Recommendation.

AES audio: All the VUCP (sample validity bit (V), user data bit (U), channel status bit (C), even parity bit (P)) data, audio data and auxiliary data, associated with one AES digital stream as defined in Recommendation ITU-R BS.647.

AES frame: Two AES subframes; in the case of the 32 kHz to 48 kHz sampling subframes one and two carry AES audio channel 1 and 2 respectively. In the case of 96 kHz sampling subframes one and two carry successive samples of the same AES audio signal which is mandatory for 96 kHz application.

AES subframe: All data associated with one AES audio sample for one channel in a channel pair.

audio control packet: An ancillary data packet occurring once a field in an interlaced system and once a frame in a progressive system and containing data used in the process of decoding the audio data stream.

audio clock phase data: Audio clock phase is indicated by the number of video clocks between the first word of EAV and the video sample at the same timing when audio sample appeared at the input to the formatter.

audio data: 29 bits: 24 bits of AES audio associated with one audio sample, including AES auxiliary data, plus VUCP bits and the Z flag which is derived from the preamble of AES3 stream. The Z bit is common to the two channels of an AES channel pair.

error correction code: BCH (31, 25) code (an error correction method) in each bit sequence of b0-b7. Errors between the first word of ancillary data flag (ADF) through the last word of audio data of channel 4 (CH4) in user data words (UDW) will be corrected or detected within the capability of this code.

audio data packet: An ancillary data packet containing audio clock phase data, audio data for two channel pairs (4 channels) and error correction code. An audio data packet should contain audio data of one sample associated with each audio channel.

audio frame number: A number, starting at 1, for each frame within the audio frame sequence.

audio frame sequence: The number of video frames required for an integer number of audio samples in isochronous operation.

audio group: Consists of two channel pairs that are contained in one ancillary data packet. Each audio group has a unique ID. Audio groups are numbered 1 through 4.

channel pair: Two digital audio channels, derived from the same AES audio source.

data ID: A word in the ancillary data packet which identifies the use of the data therein.

Extended audio group: an audio group as defined in Annex 1 of this Recommendation, but numbered from 5 to 8.

Extended audio data packet: an audio data packet as defined in Annex 1 of this Recommendation, but with identity corresponding to Extended audio group numbers 5 to 8.

Extended audio control packet: an audio control packet defined in Annex 1 of this Recommendation, but with identity corresponding to Extended audio group numbers 5 to 8.

horizontal ancillary data block: An ancillary data space located in the digital line blanking interval of one television line.

isochronous audio: Audio is defined as being clock isochronous with video if the sampling rate of audio is such that the number of audio samples occurring within an integer number of video frames is itself a constant integer number, as shown in the following example:

TABLE 1

Examples of samples per frame for synchronous audio

	Samples-frame/s										
Audio sampling rate	**120**	**120/1.001**	**100**	**60**	**60/1.001**	**50**	**30.00**	**30.00/1.001**	**25.00**	**24.00**	**24.00/1.001**
96.0 kHz	800/1	4004/5	960	1600/1	8008/5	1920	3 200/1	16 016/5	3 840/1	4 000/1	4 004/1
48.0 kHz	400/1	2002/5	480	800/1	4004/5	960	1 600/1	8 008/5	1 920/1	2 000/1	2 002/1

Annex 1

24-bit digital audio format as ancillary data signals in HDTV and UHDTV serial interfaces

1 Introduction

Audio sampled at a clock frequency of 48 kHz locked (synchronous) to video is the preferred implementation for intrastudio applications. As an option, this Recommendation supports Audio Engineering Society (AES) audio at synchronous or asynchronous sampling rates from 32 kHz to 48 kHz and 96 kHz. Audio channels are transmitted in groups of four, up to a maximum of 16 audio channels in the case of 32 kHz, 44.1 kHz or 48 kHz sampling, and up to a maximum of 8 audio channels in case of 96 kHz sampling. Each group is identified by a unique ancillary data ID.

Audio data packets are multiplexed (embedded) into the horizontal ancillary data space of the C'_B/C'_R data stream, and audio control packets are multiplexed into the horizontal ancillary data space of the Y data stream. The multiplexed data are converted into serial form according to the HDTV serial digital interfaces defined in Recommendation ITU-R BT.1120.

For UHDTV interfaces conforming to Recommendation ITU-R BT. 2077 Parts 1 and 3, this Recommendation applies to Y data stream and C'_B/C'_R data stream, making up the overall multiplex.

For UHDTV interfaces conforming to Recommendation ITU-R BT.2077 Part 2, this Recommendation applies to basic stream 1 and basic stream 2 of the interface according to §§ 3.5 and 3.6 in Part 2 of Recommendation ITU-R BT.2077.

2 References

– Recommendation ITU-R BT.709 – Parameter Values for the HDTV standards for production and international programme exchange.

– Recommendation ITU-R BT.1120 – Digital interfaces for HDTV studio signals.

– Recommendation ITU-R BS.647 – A Digital audio interface for broadcasting studios.

– Recommendation ITU-R BT.2020- Parameter values for ultra-high definition television systems for production and international programme exchange.

– Recommendation ITU-R BT.2077– Real-time serial digital interfaces for UHDTV signals.

– Recommendation ITU-R BT.1364 – Format of Ancillary Data signals carried in digital component studios.

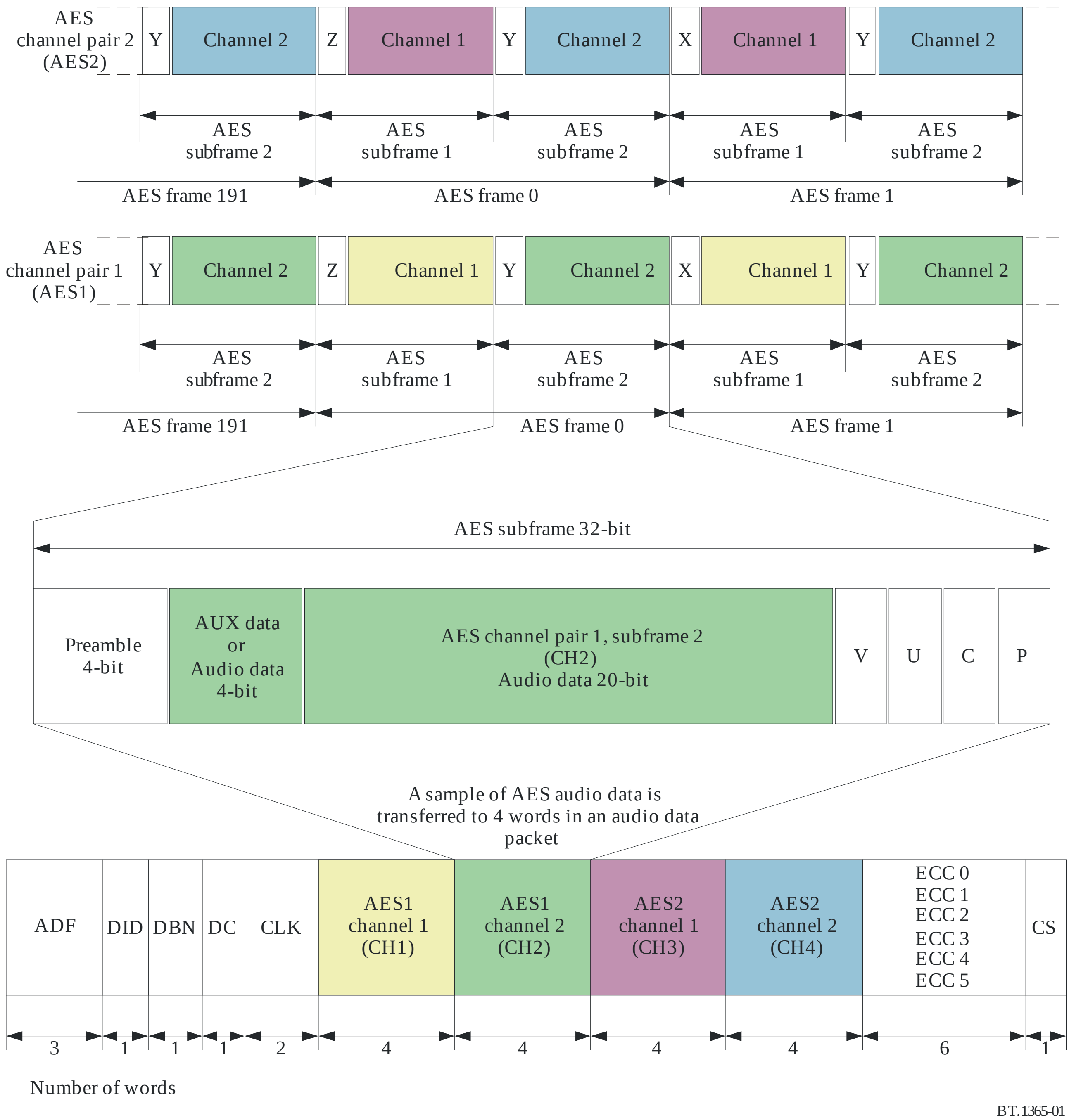
3 Overview

3.1 The modes of transmission carried in an audio data packet should be the two channel mode at all sampling frequencies from 32 kHz to 48 kHz and the single channel double sampling frequency mode at the sampling frequency of 96 kHz. Audio data channels 1~4 (CH1~CH4) carry two AES audio channel pairs (AES1 channel 1 & 2 and AES2 channel 1 & 2) in the case of 32 kHz to 48 kHz sampling. For 96 kHz sampling two successive samples of two AES audio channels (AES1 channel 1 1st & 2nd sample and AES2 channel 1 1st & 2nd sample) should be carried.

3.2 The 32 kHz, 44.1 kHz or 48 kHz sampling audio data derived from two channel pairs should be configured in an audio data packet as shown in Fig. 1. Both channels of a channel pair are derived from the same AES audio source. The number of samples per channel used for one audio data packet should be constant and is equal to one. The number of audio data packets in a given group should be less than or equal to Na in a horizontal ancillary data block. See § 4.3.3.

FIGURE 1

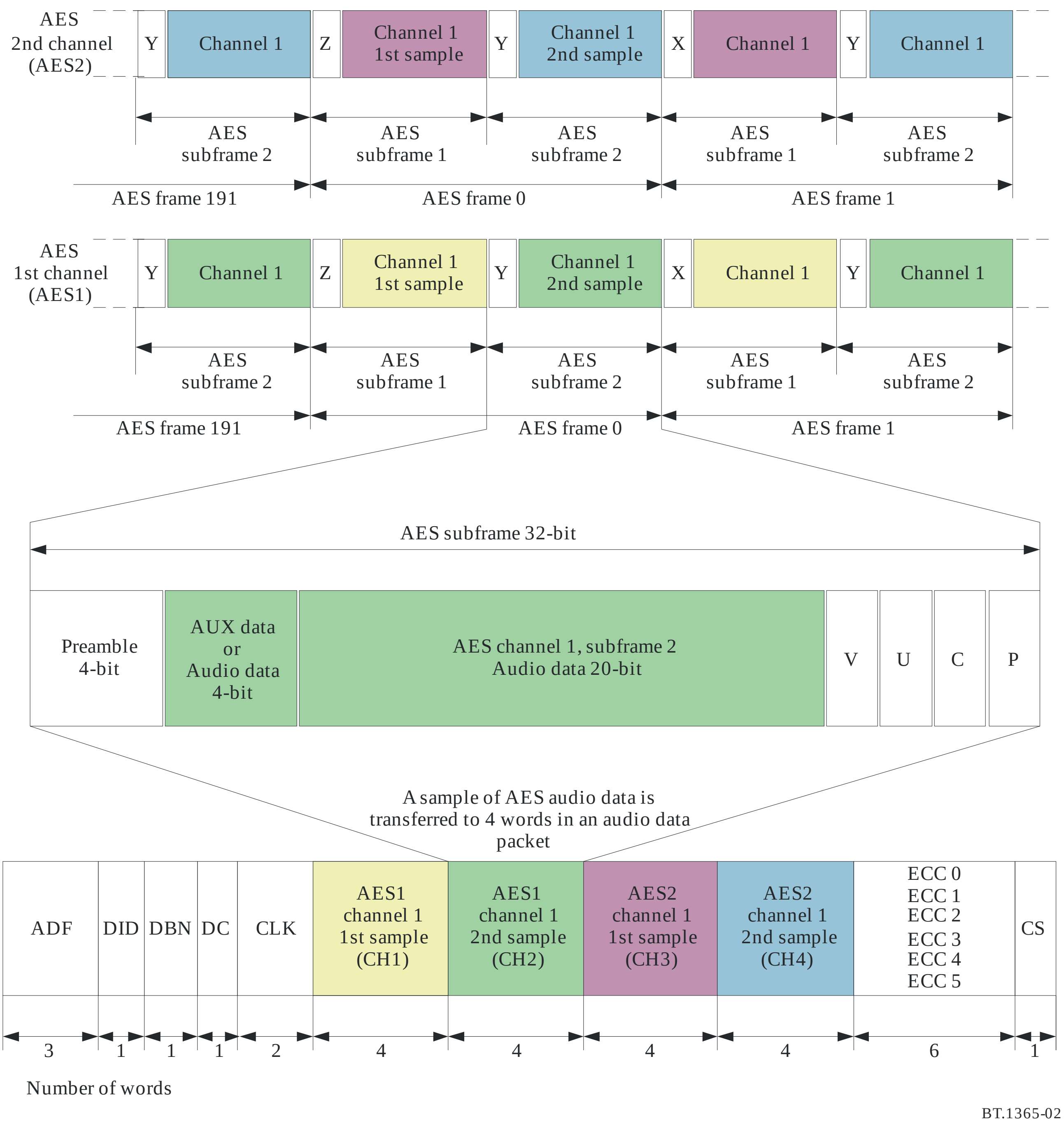
Relationship between AES audio and audio data packets at sampling rates of 32 kHz, 44.1 kHz or 48 kHz



3.3 Figure 2 shows the audio data packet at the sampling rate of 96 kHz. AES subframes 1 and 2 carry successive samples of the same AES audio signal. Both channels should be derived from the same AES audio source. The number of samples per channel used for one audio data packet should be constant and equal to two. The number of audio data packets in a given group is less than or equal to Na/2 in a horizontal ancillary data block.

FIGURE 2

Relationship between AES audio and audio data packets at a sampling rate of 96 kHz



3.4 Two types of ancillary data packets carrying AES audio information are defined in this Recommendation. Each audio data packet should carry all of the information in the AES bit stream. The audio data packet should be located in the horizontal ancillary data space of the C'_B/C'_R data stream. An audio control packet should be transmitted once per field in an interlaced system and once per frame in a progressive system in the horizontal ancillary data space of the second line after the switching point of the Y data stream.

3.5 Data ID should be defined for four separate packets of each packet type. This allows for up to eight channel pairs. In Annex 1 of this Recommendation, the audio groups are numbered 1 through 4 and the channels are numbered 1 through 16. Channels 1 through 4 are in group 1, channels 5 through 8 are in group 2, and so on. Table 2 defines the relationship between CH1~CH4 (UDW2~UDW17) in the audio data packet and the channel/sample number for 32 kHz to 48 kHz sampling and 96 kHz sampling respectively.

3.6 The audio data packet and audio control packet should be located in Recommendation ITU-R BT.1120 transport HANC space that is equal to 268 video samples interval at 30 Hz video frame rate.

TABLE 2

Relationship between audio data packets and the channel/sample number of 32 kHz to 48 kHz and 96 kHz sampling

	Audio group 1			
Audio sampling rate	**UDW2~UDW5 CH1**	**UDW6~UDW9 CH2**	**UDW10~UDW13 CH3**	**UDW14~UDW17 CH4**
32.0 kHz, 44.1 kHz or 48.0 kHz	AES1 channel 1	AES1 channel 2	AES2 channel 1	AES2 channel 2
96.0 kHz	AES1 channel 1 1st sample	AES1 channel 1 2nd sample	AES2 channel 1 1st sample	AES2 channel 1 2nd sample

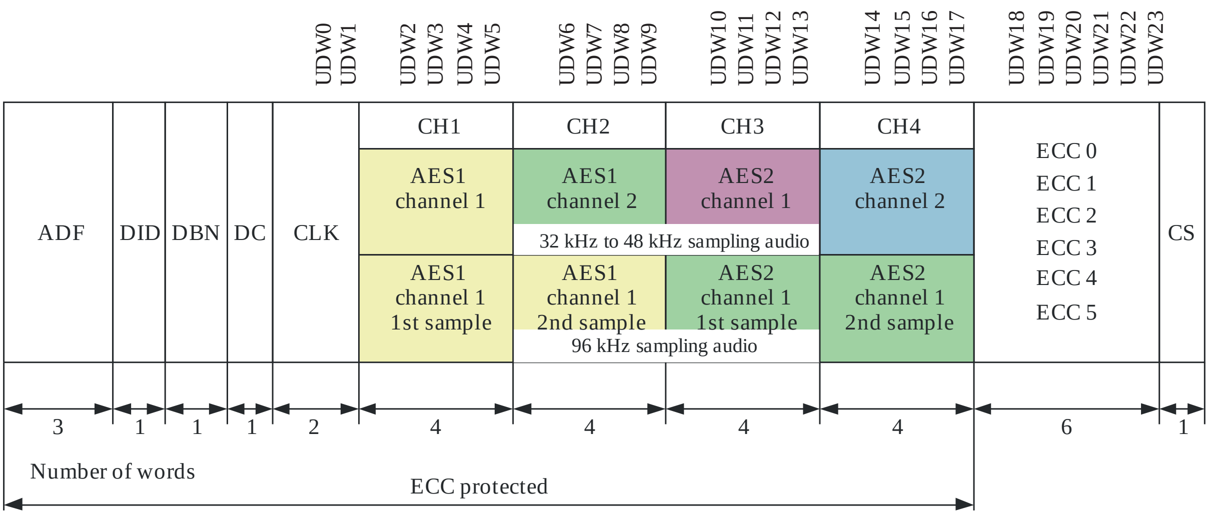
4 Audio data packet

4.1 Structure of audio data packet

4.1.1 The structure of the audio data packet should be as shown in Fig. 3. Audio data packets consist of ADF, DID, DBN, DC, UDW and CS. ADF, DBN, DC and CS are subject to Recommendation ITU-R BT.1364 – Format of ancillary data signals carried in digital component studio interfaces. DC is always 218_h.

FIGURE 3

Structure of audio data packets



BT.1365-03

4.1.2 DID is defined as 2E7h for audio group 1 (channel 1-4), 1E6h for audio group 2 (channel 5-8), 1E5h for audio group 3 (channel 9-12) and 2E4h for audio group 4 (channel 13-16), respectively.

4.1.3 UDW is defined in § 4.2. In this Recommendation, UDWx means the Xth user data word. There are always 24 words in the UDW of an audio data packet, i.e. UDW0, UDW1, …, UDW22, UDW23.

4.1.4 All audio channels in a given audio group should have identical sampling rate, identical sampling phase and identical isochronous/asynchronous status.

4.1.5 For a given audio data packet, one sample of the audio data of each channel (CH1-CH4) is always transmitted. Even when only one of the four channels (CH1-CH4) is active, all audio data of the four channels should be transmitted. In such case, the value of audio data, V, U, C and P bits of all inactive channels should be set to zero.

4.2 Structure of user data words

UDW consists of three types of data defined in §§ 4.2.1 to 4.2.3. The description in this clause covers only audio group 1. The description for audio groups 2, 3 and 4 is similar to that for audio group 1 where channels 5, 9 and 13 correspond to channel 1, channels 6, 10 and 14 correspond to channel 2, channels 7, 11 and 15 correspond to channel 3, channels 8, 12 and 16 correspond to channel 4, respectively.

4.2.1 Audio clock phase data

4.2.1.1 Audio clock phase data (CLK) is used to regenerate audio sampling clock at the receiving side, especially for asynchronous audio. Bit-assignment of CLK should be as shown in Table 3.

TABLE 3

Bit assignment of CLK

Bit number	UDW0	UDW1
b9 (MSB)	Not b8	Not b8
b8	Even parity[(1)]	Even parity[(1)]
b7	ck7 audio clock phase data	Reserved (set to 0)
b6	ck6 audio clock phase data	Reserved (set to 0)
b5	ck5 audio clock phase data	ck12 audio clock phase data (MSB)
b4	ck4 audio clock phase data	mpf multiplex position flag
b3	ck3 audio clock phase data	ck11 audio clock phase data
b2	ck2 audio clock phase data	ck10 audio clock phase data
b1	ck1 audio clock phase data	ck9 audio clock phase data
b0 (LSB)	ck0 audio clock phase data (LSB)	ck8 audio clock phase data

(1) Even parity for b0 through b7.

4.2.1.2 Bits of ck0 to ck11 indicate the number of video clocks between the first word of EAV and the video sample at the same time that audio sample appears at the input of the formatter. The relationship among “video”, “sampling instants of digital audio” and “audio clock phase data” is shown in Fig. 4a (30 Hz frame rate) and Fig. 4b (30/1.001 Hz frame rate) and Fig. 4c (96 kHz sampling and 30 Hz frame rate) as some examples.

FIGURE 4A

Relationship between video lines, sampling instants of digital audio and audio clock phase data (informative example – 1080/60/I system with 48 kHz audio sampling rate and 30.00 Hz video frame rate)

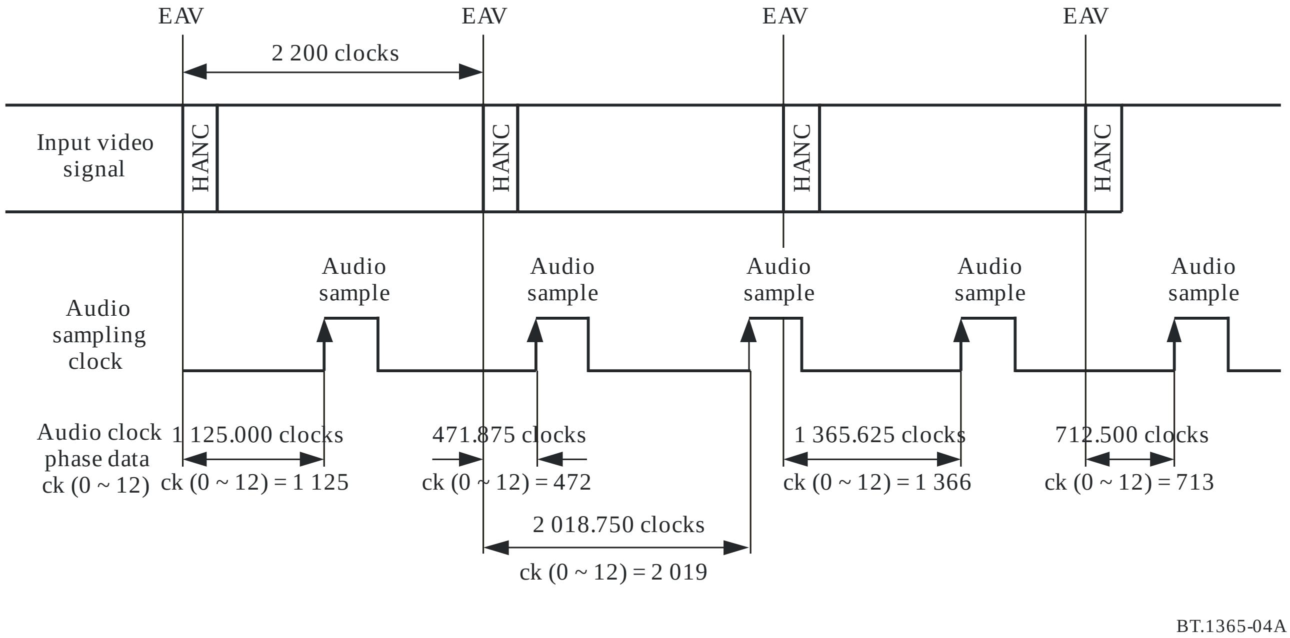


FIGURE 4B

Relationship between video lines, sampling instants of digital audio and audio clock phase data (informative example – 1080/60/I system with 48 kHz audio sampling rate and 30.00/1.001 Hz video frame rate)

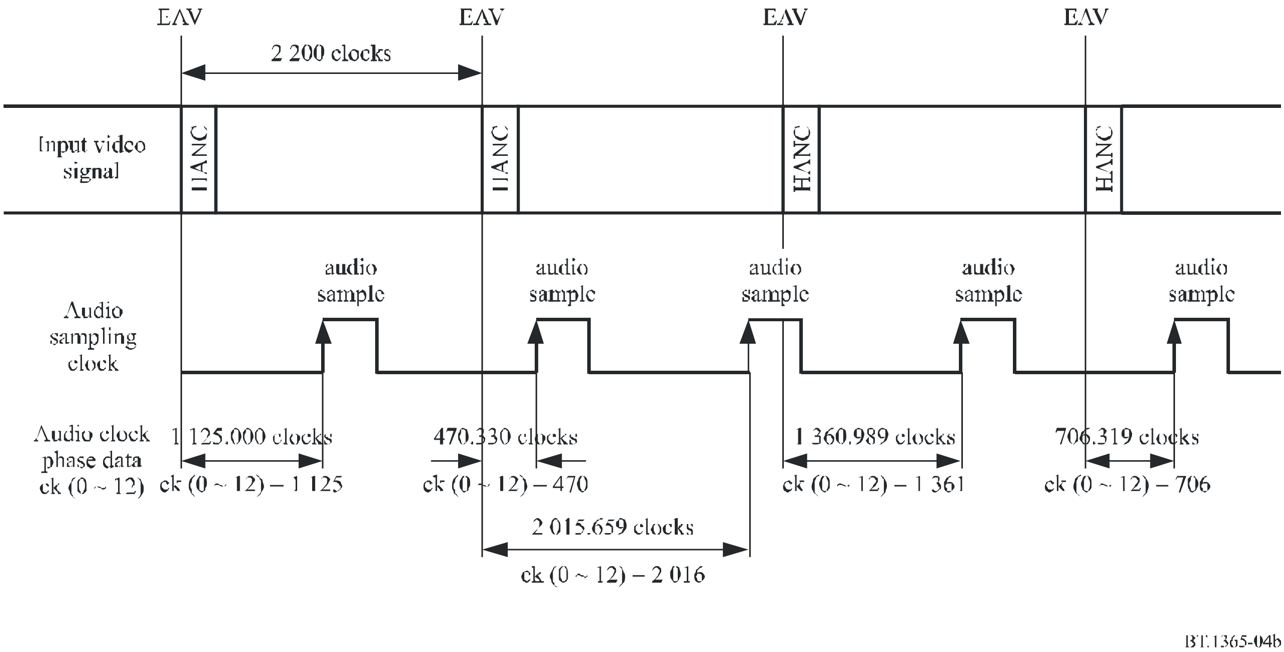
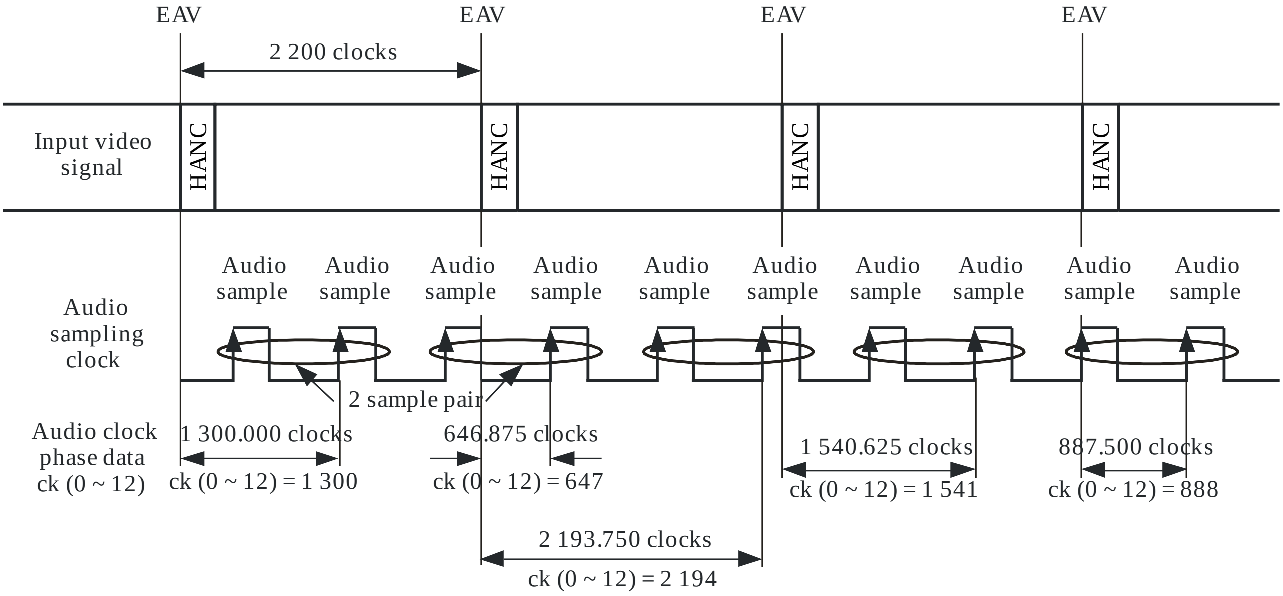


FIGURE 4C

Relationship between video lines, sampling instants of digital audio and audio clock phase data (informative example – 1080/60/I system with 96 kHz audio sampling rate and 30.00 Hz video frame rate)



BT.1365-04C

In the case of 96 kHz sampling, CLK indicates the number of video clocks between the first word of EAV and the video sample at the same time that the second audio sample of the successive two samples of the same AES audio signal appears at the input of the formatter.

4.2.1.3 The formatter should place the audio data packet in the horizontal ancillary space following the video line during which the audio sample occurred. Following a switching point, the audio data packet should be delayed one additional line to prevent data corruption.

Flag bit *mpf* defines the audio data packet position in the multiplexed output stream relative to the associated video data.

When bit *mpf* = 0, it should indicate that the audio data packet is located immediately after the video line during which the audio sample occurred.

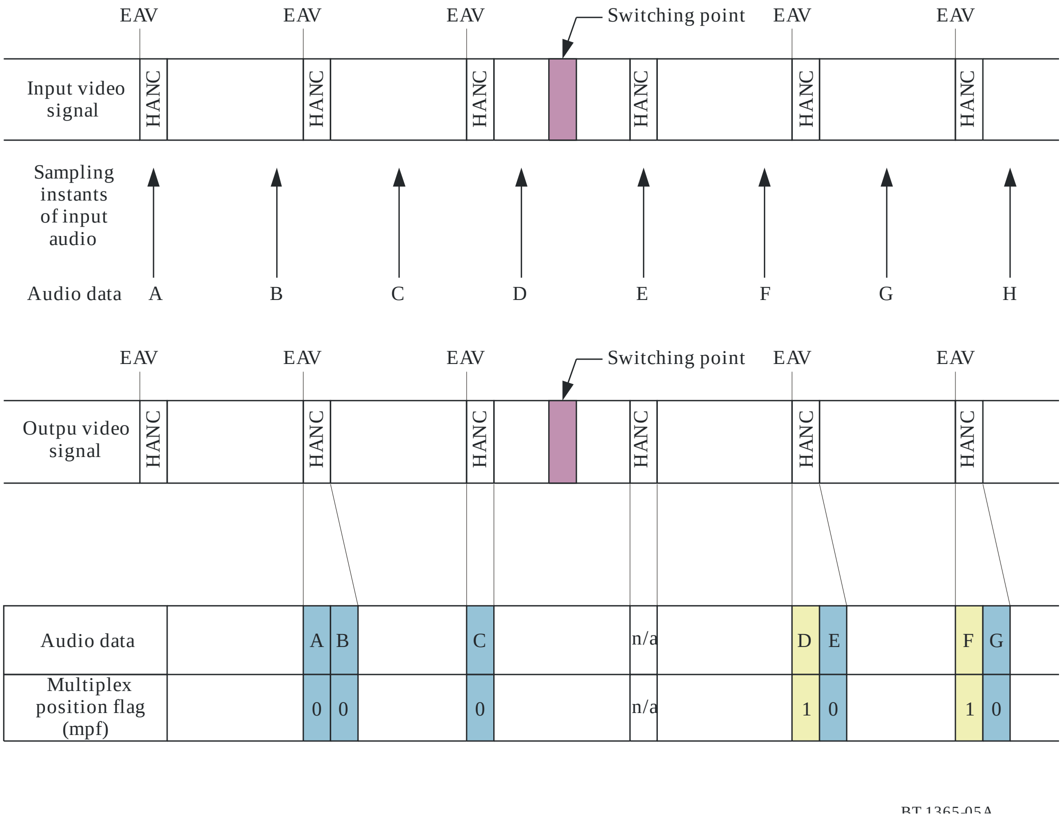
When bit *mpf* = 1, it should indicate that the audio data packet is located in the second line following the video line during which the audio sample occurred.

The relationship between the multiplex position flag (mpf) and the multiplex position of the audio data packet is shown in Figs 5a and 5b.

In the case of 96 kHz sampling, *mpf* should be defined according to the position of the second sample of the successive two samples of the same AES audio signal.

FIGURE 5A

Relationship between the multiplex position flag and the multiplex position of 32 kHz to 48 kHz sampling audio data packets



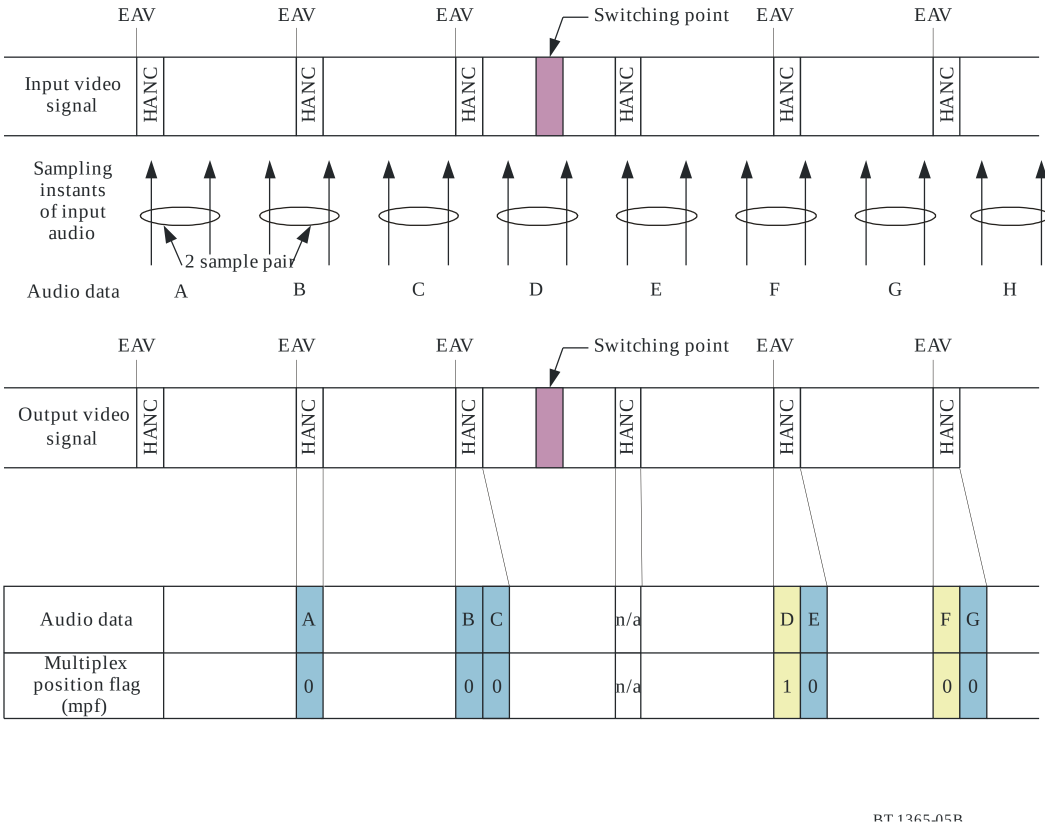
Note 1 – For example, for samples A,B,C,E and G, mpf = 0 because the ancillary data packet is multiplexed in the horizontal ancillary data space of the next line relative to the input timing of the audio sample.

Note 2 – N/A shows that the line subsequent to the switching point precludes the insertion of ancillary data packets.

Note 3 – For example, for samples D and F, mpf = 1 because the ancillary data packet is multiplexed in the horizon ancillary data space of the second line relative to the input timing of audio sample.

FIGURE 5B

Relationship between the multiplex position flag and the multiplex position of 96 kHz sampling audio data packets



Note 1 – For example, for samples A,B,C,E and G, mpf = 0 because the ancillary data packet is multiplexed in the horizontal ancillary data space of the next line relative to the input timing of the audio sample.

Note 2 – N/A shows that the line subsequent to the switching point precludes the insertion of ancillary data packets.

Note 3 – For example, for samples D and F, mpf = 1 because the ancillary data packet is multiplexed in the horizon ancillary data space of the second line relative to the input timing of audio sample.

4.2.2 CHn (audio data)

4.2.2.1 The bit assignment of CHn (n = 1 ~ 4) should be as shown in Table 4. All bits of an AES subframe should be transparently transferred to four consecutive UDW words (UDW4n-2, UDW4n-1, UDW4n, UDW4n+1). UDW2 through UDW17 are always used for CHn in audio data packets.

4.2.2.2 Bit 3 of UDW2 and UDW10 indicates the status of the Z flag which corresponds to the AES block sync. The Z flag bit in UDW2 should be associated with CH1 and CH2, and the Z flag bit in UDW10 should be associated with CH3 and CH4.

4.2.2.3 Bits b0 through b2 in UDW2, UDW6, UDW10 and UDW14, and bit b3 in UDW6 and UDW14 should be set to zero.

TABLE 4

Bit-assignment of audio data (CHn)

	Bit number	**UDW2**	**UDW3**	**UDW4**	**UDW5**
CH1	b9 (MSB)	Not b8	Not b8	Not b8	Not b8
	b8	Even parity(1)	Even parity(1)	Even parity(1)	Even parity(1)
	b7	aud$_1$ 3	aud$_1$ 11	aud$_1$ 19	P_1
	b6	aud$_1$ 2	aud$_1$ 10	aud$_1$ 18	C_1
	b5	aud$_1$ 1	aud$_1$ 9	aud$_1$ 17	U_1
	b4	aud$_1$ 0 (LSB)	aud$_1$ 8	aud$_1$ 16	V_1
	b3	Z	aud$_1$ 7	aud$_1$ 15	aud$_1$ 23 (MSB)
	b2	0	aud$_1$ 6	aud$_1$ 14	aud$_1$ 22
	b1	0	aud$_1$ 5	aud$_1$ 13	aud$_1$ 21
	b0 (LSB)	0	aud$_1$ 4	aud$_1$ 12	aud$_1$ 20
	Bit number	**UDW6**	**UDW7**	**UDW8**	**UDW9**
CH2	b9 (MSB)	Not b8	Not b8	Not b8	Not b8
	b8	Even parity(1)	Even parity(1)	Even parity(1)	Even parity(1)
	b7	aud$_2$ 3	aud$_2$ 11	aud$_2$ 19	P_2
	b6	aud$_2$ 2	aud$_2$ 10	aud$_2$ 18	C_2
	b5	aud$_2$ 1	aud$_2$ 9	aud$_2$ 17	U_2
	b4	aud$_2$ 0 (LSB)	aud$_2$ 8	aud$_2$ 16	V_2
	b3	Z	aud$_2$ 7	aud$_2$ 15	aud$_2$ 23 (MSB)
	b2	0	aud$_2$ 6	aud$_2$ 14	aud$_2$ 22
	b1	0	aud$_2$ 5	aud$_2$ 13	aud$_2$ 21
	b0 (LSB)	0	aud$_2$ 4	aud$_2$ 12	aud$_2$ 20
	Bit number	**UDW10**	**UDW11**	**UDW12**	**UDW13**
CH3	b9 (MSB)	Not b8	Not b8	Not b8	Not b8
	b8	Even parity(1)	Even parity(1)	Even parity(1)	Even parity(1)
	b7	aud$_3$ 3	aud$_3$ 11	aud$_3$ 19	P_3
	b6	aud$_3$ 2	aud$_3$ 10	aud$_3$ 18	C_3
	b5	aud$_3$ 1	aud$_3$ 9	aud$_3$ 17	U_3
	b4	aud$_3$ 0 (LSB)	aud$_3$ 8	aud$_3$ 16	V_3
	b3	Z	aud$_3$ 7	aud$_3$ 15	aud$_3$ 23 (MSB)
	b2	0	aud$_3$ 6	aud$_3$ 14	aud$_3$ 22
	b1	0	aud$_3$ 5	aud$_3$ 13	aud$_3$ 21
	b0 (LSB)	0	aud$_3$ 4	aud$_3$ 12	aud$_3$ 20
	Bit number	**UDW14**	**UDW15**	**UDW16**	**UDW17**
CH4	b9 (MSB)	Not b8	Not b8	Not b8	Not b8
	b8	Even parity(1)	Even parity(1)	Even parity(1)	Even parity(1)
	b7	aud$_4$ 3	aud$_4$ 11	aud$_4$ 19	P_4
	b6	aud$_4$ 2	aud$_4$ 10	aud$_4$ 18	C_4
	b5	aud$_4$ 1	aud$_4$ 9	aud$_4$ 17	U_4
	b4	aud$_4$ 0 (LSB)	aud$_4$ 8	aud$_4$ 16	V_4
	b3	Z	aud$_4$ 7	aud$_4$ 15	aud$_4$ 23 (MSB)
	b2	0	aud$_4$ 6	aud$_4$ 14	aud$_4$ 22
	b1	0	aud$_4$ 5	aud$_4$ 13	aud$_4$ 21
	b0 (LSB)	0	aud$_4$ 4	aud$_4$ 12	aud$_4$ 20

Notes to Table 4:

NOTE 1 – Even parity for b0 through b7

NOTE 2 – Z = AES block sync

NOTE 3 – Un = AES user bit of CHn

NOTE 4 – Pn = AES parity bits of CHn

NOTE 5 – aud (0-23) = 24-bit AES audio data of CHn

NOTE 6 – Vn = AES sample validity bit of CHn

NOTE 7 – Cn = AES channel status bit of CHn

NOTE 8 – Value of Vn, Un, Cn and Pn is equal to that of AES subframe, respectively.

4.2.3 Error correction codes

4. 2.3.1 Error correction codes (ECC) are used to correct or detect errors in 24 words from the first word of ADF through UDW17. The error correction code is BCH (31, 25) code. BCH code is formed for each bit sequence of b0-b7, respectively. ECC consists of 6 words determined by the polynomial generator equation:

$$ECC(X) = (X+1)(X^5+X^2+1) = X^6+X^5+X^3+X^2+X+1.$$

Initial value of all FFn is set to zero. The calculation starts at the first word of ADF and ends at the final word of CH4 (UDW17) for each bit of b0 to b7, respectively. The remaining data in the FFn is ECCn. (n = 0-5) (FFn stands for "Flip Flop number". For example, the data of FF0 is ECC0, the data of FF5 is ECC5.)

4.2.3.2 Bit-assignment of ECC should be as shown in Table 5. An example of the block diagram of the BCH-code formation circuit is shown in Fig. 6.

TABLE 5

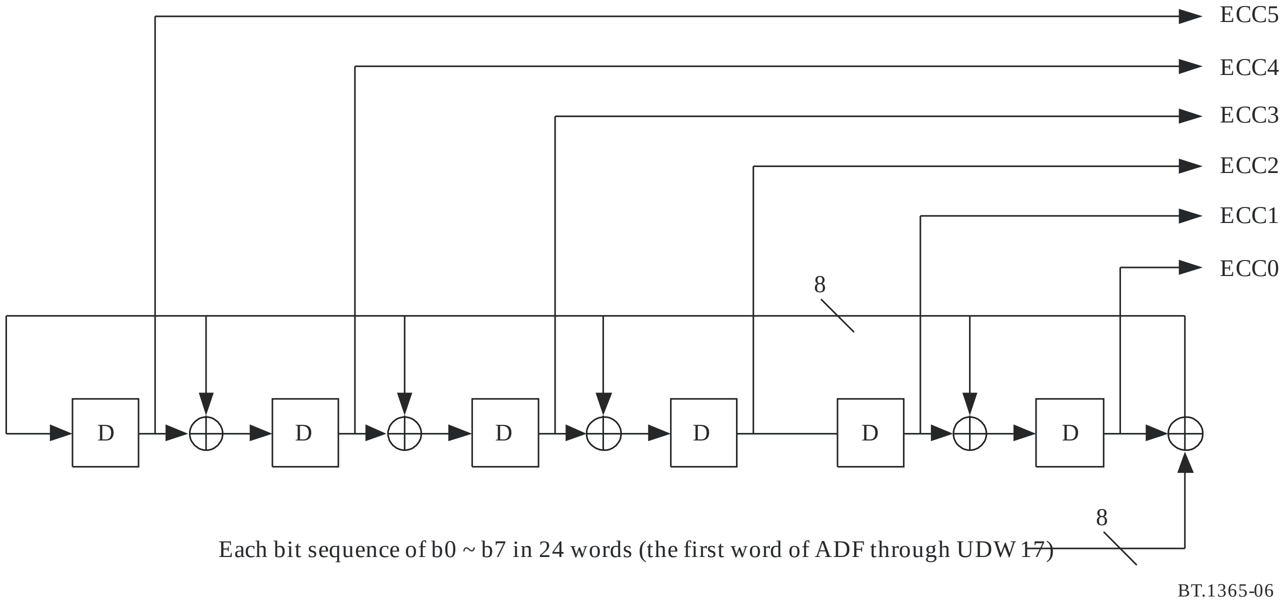
Bit-assignment of ECC

Bit number	**UDW18**	**UDW19**	**UDW20**	**UDW21**	**UDW22**	**UDW23**
	ECC0	**ECC1**	**ECC2**	**ECC3**	**ECC4**	**ECC5**
b9 (MSB)	not b8	not b8	not b8	not b8	not b8	not b8
b8	even parity(1)	even parity(1)	even parity(1)	even parity(1)	even parity(1)	even parity(1)
b7	ecc0 7	ecc1 7	ecc2 7	ecc3 7	ecc4 7	ecc5 7
b6	ecc0 6	ecc1 6	ecc2 6	ecc3 6	ecc4 6	ecc5 6
b5	ecc0 5	ecc1 5	ecc2 5	ecc3 5	ecc4 5	ecc5 5
b4	ecc0 4	ecc1 4	ecc2 4	ecc3 4	ecc4 4	ecc5 4
b3	ecc0 3	ecc1 3	ecc2 3	ecc3 3	ecc4 3	ecc5 3
b2	ecc0 2	ecc1 2	ecc2 2	ecc3 2	ecc4 2	ecc5 2
b1	ecc0 1	ecc1 1	ecc2 1	ecc3 1	ecc4 1	ecc5 1
b0 (LSB)	ecc0 0	ecc1 0	ecc2 0	ecc3 0	ecc4 0	ecc5 0

(1) Even parity for b0 through b7.

FIGURE 6

An example of block diagram of the BCH-code formation circuitry



4.3 Multiplexing of audio data packet

4.3.1 Only the horizontal ancillary data space of the colour-difference data stream (C_B/C_R) should be used for transmission of the audio data packet.

4.3.2 The audio data packet should not be multiplexed into the horizontal ancillary data space of the line subsequent to the switching point defined by the source format. As an example, the ancillary data space available for audio data packet in the 1125/60I system is shown in Fig. 7.

4.3.3 The number of samples per audio channel which can be multiplexed in one horizontal ancillary data space should be less than or equal to Na (Number of audio samples), where Na is defined in the following pseudocode:

No = Int (audio sample rate/line frequency) + 1

if No × (the number of total lines per video frame – the number of switching line per video frame)

< (the number of audio samples per video frame)

then Na = No + 1

else Na = No

if (audio sampling rate == 96 kHz) Na = Even(Na)

The function Even(n) returns the smallest even number that is greater than or equal to n. For example, Even(123) = 124, Even(98) = 98.

When two or more samples of the audio data are transmitted in one horizontal ancillary data block, the packet of the audio sample which appears earlier at the input of the formatter should be transmitted first.

Some video formats may require up to 8 samples per data block (i.e. Na = 8).

4.3.4 An audio data packet should be multiplexed in the horizontal ancillary data space of the first or second line following the line during which the audio sample occurred at the input of the formatter.

NOTE 1 – Audio phase must be maintained across the audio groups carrying the multiple-channel audio.

4.3.5 The audio data packet should be multiplexed following the CRCC words defined in Recommendation ITU-R BT.1120.

4.3.6 When more than two audio data packets are transmitted in one horizontal ancillary data block, the audio data packets should be contiguous with each other.

5 Audio control packet

5.1 Structure of audio control packet

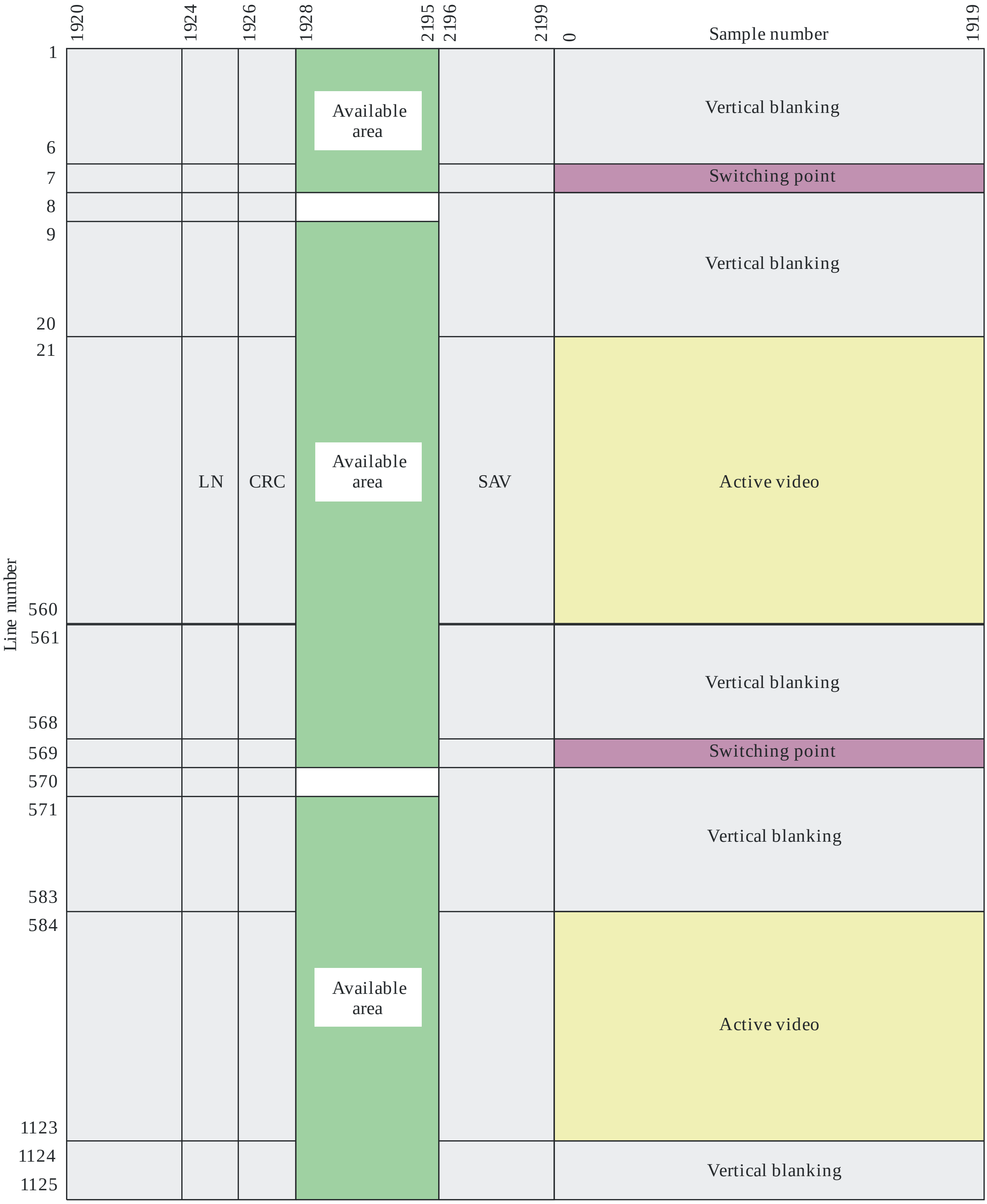
5.1.1 The structure of audio control packet should be as shown in Fig. 8. Audio control packets consist of ancillary data flag (ADF), data identification (DID), data block number (DBN), data count (DC), user data words (UDW) and checksum (CS). ADF, DC and CS are subject to Recommendation ITU-R BT.1364. DC is always 10Bh and DBN is always 200 h.

5.1.2 DID has a value of 1E3h for audio group 1 (channel 1-4), 2E2h for audio group 2 (channel 5-8), 2E1h for audio group 3 (channel 9-12) and 1E0h for audio group 4 (channel 13-16), respectively.

5.1.3 UDW is defined in § 5.2. In this Recommendation, UDWx means the Xth user data word. There are always 11 words in the UDW of an audio control packet, i.e. UDW0, UDW1, …, UDW9, UDW10.

FIGURE 7

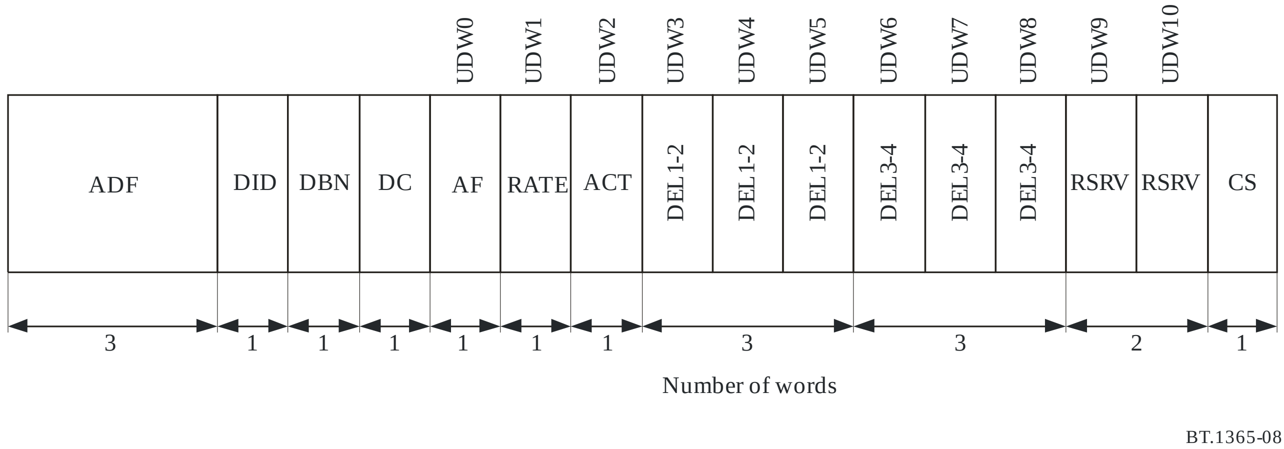
Ancillary data space of C'B/C'R data stream available for transmission of audio data packets (1080/60i system)



BT.1365-07

FIGURE 8

Structure of audio control packet



5.2 Structure of UDW

UDW consists of five types of data defined in §§ 5.2.1 to 5.2.5. The description in this clause covers only audio group 1. The description for audio groups 2, 3 and 4 is similar to audio group 1 where channels 5, 9 and 13 correspond to channel 1, channels 6, 10 and 14 correspond to channel 2, channels 7, 11 and 15 correspond to channel 3, channels 8, 12 and 16 correspond to channel 4, respectively.

5.2.1 Audio frame number data

5.2.1.1 Audio frame number data (AF) provide a sequential numbering of video frames to indicate where they fall in the progression of non-integer number of samples per video frame (audio frame sequence). The first number of the sequence is always 1 and the final number is equal to the length of the audio frame sequence. A value of AF equal to all zeros indicates that frame numbering is not available. (See Attachment 1.)

5.2.1.2 The bit-assignment of the AF should be as shown in Table 6. The AF is common for all channels in a given audio group.

5.2.1.3 When channel pairs in a given audio group are operating in asynchronous mode, the AF word in the audio control packet is not used and b0-b8 should be set to zero.

TABLE 6

Bit-assignment of AF

Bit number	**UDW0**
	AF
b9 (MSB)	not b8
b8	f8 Audio frame number (MSB)
b7	f7 Audio frame number
b6	f6 Audio frame number
b5	f5 Audio frame number
b4	f4 Audio frame number
b3	f3 Audio frame number
b2	f2 Audio frame number
b1	f1 Audio frame number
b0 (LSB)	f0 Audio frame number (LSB)

5.2.2 RATE (Sampling rate)

5.2.2.1 The sampling rate for all channel pairs is defined by the word (RATE). The bit-assignment of RATE should be as shown in Table 7.

5.2.2.2 The sync mode bit asx, when set to one, indicate that the channel pairs in a given audio group are operating asynchronously.

5.2.2.3 The rate code is currently defined as shown in Table 8.

TABLE 7

Bit-assignment of RATE

Bit number	UDW1
	RATE
b9 (MSB)	not b8
b8	0
b7	0
b6	0
b5	0
b4	0
b3	X2 (MSB)
b2	X1 Rate code
b1	X0 (LSB)
b0 (LSB)	asx isochronous audio; 0 asynchronous audio; 1

TABLE 8

Assignment of rate code

X2	X1	X0	Sample rate
0	0	0	48.0 kHz
0	0	1	44.1 kHz
0	1	0	32.0 kHz
1	0	0	96.0 kHz
0	1	1	Reserved
1	0	1	Reserved
1	1	0	Reserved
1	1	1	Free running

5.2.3 ACT

5.2.3.1 The word ACT indicates active channels. Bits a1 to a4 are set to one for each active channel in a given audio group otherwise they are set to zero. The bit-assignment of ACT is shown in Table 9.

TABLE 9

Bit-assignment of ACT

Bit number	UDW2
	ACT
b9 (MSB)	not b8
b8	even parity(1)
b7	0
b6	0
b5	0
b4	0
b3	a4 active: 1, inactive: 0 (CH4)
b2	a3 active: 1, inactive: 0 (CH3)
b1	a2 active: 1, inactive: 0 (CH2)
b0 (LSB)	a1 active: 1, inactive: 0 (CH1)

(1) Even parity for b0 through b7.

5.2.4 DELm-n

5.2.4.1 The words DELm-n indicate the amount of accumulated audio processing delay relative to video, measured in audio sample intervals, for each channel pair of CHm and CHn.

In the case of 96 kHz sampling, DELm-n should indicate the amount of accumulated audio processing delay relative to video measured in audio sample intervals for the successive two samples of the same AES audio signal carried in CH1, CH2 and CH3, CH4.

5.2.4.2 The bit-assignment of DELm-n should be as shown in Table 10. The *e* bit is set to one to indicate valid audio delay data. The delay words are referenced to the point where the AES/EBU data are input to the formatter. The delay words represent the average delay value, inherent in the formatting process, over a period no less than the length of the audio frame sequence plus any pre-existing audio delay.

5.2.4.3 The audio delay data (del 0-del 25) is represented in the format of 26-bit 2's complement. Positive values indicate that the video leads the audio.

TABLE 10

Bit-assignment of DELm-n

Bit number	UDW3	UDW4	UDW5	UDW6	UDW7	UDW8
	DEL1-2			DEL3-4		
b9 (MSB)	not b8	not b8	not b8	not b8	not b8	not b8
b8	del 7	del 16	del 25 (±)	del 7	del 16	del 25 (±)
b7	del 6	del 15	del 24 (MSB)	del 6	del 15	del 24 (MSB)
b6	del 5	del 14	del 23	del 5	del 14	del 23
b5	del 4	del 13	del 22	del 4	del 13	del 22
b4	del 3	del 12	del 21	del 3	del 12	del 21
b3	del 2	del 11	del 20	del 2	del 11	del 20
b2	del 1	del 10	del 19	del 1	del 10	del 19
b1	del 0 (LSB)	del 9	del 18	del 0 (LSB)	del 9	del 18
b0 (LSB)	*e*	del 8	del 17	*e*	del 8	del 17

5.2.5 RSRV

5.2.5.1 The words marked RSRV are reserved for future use.

5.2.5.2 The bit-assignment of RSRV word should be as shown in Table 11.

TABLE 11

Bit-assignment of RSRV

Bit number	UDW9	UDW10
	RSRV	RSRV
b9 (MSB)	not b8	not b8
b8	reserved (set to 0)	reserved (set to 0)
b7	reserved (set to 0)	reserved (set to 0)
b6	reserved (set to 0)	reserved (set to 0)
b5	reserved (set to 0)	reserved (set to 0)
b4	reserved (set to 0)	reserved (set to 0)
b3	reserved (set to 0)	reserved (set to 0)
b2	reserved (set to 0)	reserved (set to 0)
b1	reserved (set to 0)	reserved (set to 0)
b0 (LSB)	reserved (set to 0)	reserved (set to 0)

5.3 Multiplexing of the audio control packet

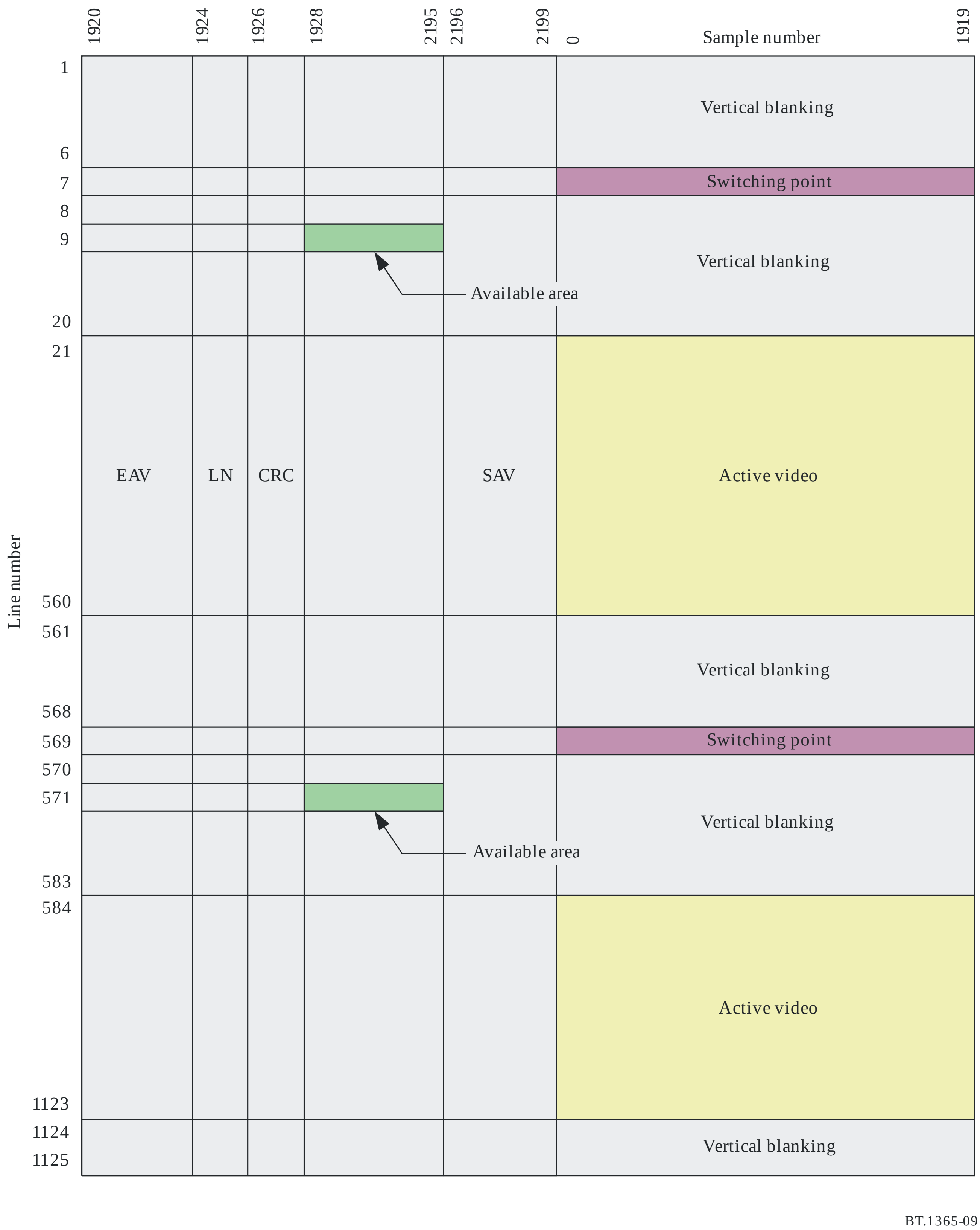
5.3.1 The audio control packets should be transmitted once every field in an interlaced system and once per frame in a progressive system.

5.3.2 The audio control packet should be transmitted in the horizontal ancillary data space of the second line after the switching point of Y parallel data stream.

For example, since the switching point for 1125/60 system exists in Line 7 and 569, the audio control packets are transmitted in the horizontal ancillary data space of Line 9 and Line 571 of the Y data stream. Ancillary data space available for the transmission of audio control packets is shown in Fig. 9.

FIGURE 9

Ancillary data space of Y data stream available for transmission of audio control packets (1080/60/I system)



Annex 2 (Normative)

Introduction

Annex 1 of this Recommendation defines the 24-bit audio format for up to 16 audio channels at 32, 44.1, or 48 kHz sample rate, or 8 audio channels at 96 kHz sample rate, The intended application is for 1.5 Gbit/s interfaces such as Recommendation ITU-R BT.1120. Annex 2 of this Recommendation, extends the audio format to 32 audio channels at 32, 44.1, or 48 kHz sample rate, or 16 audio channels at 96 kHz sample rate. Specifically, this extension defines the 24-bit audio format for channels 17 up to 32 so that up to 32 audio channels may be multiplexed with source image formats mapped to a 3 Gbit/s serial interface with 148.5 (148.5/1.001) MHz sampling frequency for luminance signal.

For UHDTV interfaces conforming to Recommendation ITU-R BT.2077 Part 3, this Annex applies to each of the 3 Gbit/s data stream pairs making up the overall multiplex.

Annex 2 of this Recommendation defines a Recommendation ITU-R BT.1364 Type 1 packet structure for identifying audio channels numbered from 17 to 32, beyond the 16 channels defined in Annex 1. Four extended audio data packets and four extended audio control packets are identified. One of the extended audio control packets and one of the extended audio data packets are assigned to transport each of the four extended audio groups. Each extended audio group has four channels that carry up to four 24-bit audio channels with 32, 44.1, or 48 kHz sample rates, or up to two 24-bit audio channels with 96 kHz sample rate.

The audio format defined in this Annex 2 is identical to that of Annex 1, except for differences required to define extended audio groups

A1 Extended Audio Data Packet

The structure and multiplexing rules for extended audio data packets are identical to that defined for audio data packets in Annex 1, with the following differences.

A1.1 DID values: The DID values for extended audio data packets should be defined as $1A7_h$ for audio group 5 (channel 17-20), $2A6_h$ for audio group 6 (channel 21-24), $2A5_h$ for audio group 7 (channel 25-28) and $1A4_h$ for audio group 8 (channel 29-32), respectively.

A1.2 Packet/group relationships: Extended audio groups 5 to 8 should be transported only using extended audio data packets defined in this Recommendation. Audio groups 1 to 4 should be transported only using audio data packets defined in Annex 1.

A1.3 Audio data packet and extended audio data packet order: The timing of the nth sample of 32 audio channels on a video line is represented by eight sample instances in eight audio data packets. Since these eight sample instances are independent of each other, the order of these eight packets in the HANC space to which they are multiplexed should be arbitrary.

A2 Extended Audio Control Packet

The structure and multiplexing rules for extended audio control packets are identical to that defined for audio control packets in Annex 1, with the following exceptions.

A2.1 DID values: The DID values for extended audio control packets should be defined as $2A3_h$ for audio group 5 (channel 17-20), $1A2_h$ for audio group 6 (channel 21-24), $1A1_h$ for audio group 7 (channel 25-28) and $2A0_h$ for audio group 8 (channel 29-32), respectively.

A2.2 Packet/group relationships: Extended audio groups 5 to 8 should be represented only using extended audio control packets defined in this Recommendation. Audio groups 1 to 4 should be transported only using audio data packets defined in Annex 1.

A2.3 Audio control packet and extended audio control packet order: The order of control and extended control packets in the HANC space to which they are multiplexed should be arbitrary.

Attachment 1 (Informative)

Alignment of audio samples for each audio frame

For alignment of AF and sample distribution, the following number of audio samples for each audio frame may be a preferred example.

All audio frame sequences are based on two integer numbers of samples per frame (m and m + 1) with audio frame numbers starting at 1 and proceeding to the end of the sequence. Odd-numbered audio frames (1, 3, 5, etc.) have the larger integer number of samples and even-numbered audio frames (2, 4, 6, etc.) have the smaller integer number of samples with the exception tabulated in Table 1-1. Receivers should have the ability to receive correctly audio data sequence even when this sequence restriction is not implemented.

TABLE 1-1

Example of alignment of audio samples for each audio frame

Television system	Sampling rate (kHz)	Frame sequence	Basic numbering		Exceptions	
			Samples per odd audio frame (m)	Samples per even audio frame (m + 1)	Frame number	Number of samples
30 frame/s	96.0	1	3 200		None	
	48.0	1	1 600		none	
	44.1	1	1 470		none	
	32.0	3	1 067	1 066	none	
29.97 frame/s	96.0	5	3 204	3 202[(1)]	None	
	48.0	5	1 602	1 601	none	
	44.1	100	1 472	1 471	23, 47, 71	1 471
	32.0	15	1 068	1 067	4, 8, 12	1 068
25 frame/s	96.0	1	3 840		none	
	48.0	1	1 920		none	
	44.1	1	1 764		none	
	32.0	1	1 280		none	

(1) Successive samples are carried in audio data packets.