Standards and Recommended Practices

Proposed SMPTE Standards

Published here for information are two Proposed SMPTE Standards:

SMPTE 314M, Television — Data Structure for DV-Based Audio, Data and Compressed Video — 25 and 50 Mb/s; and

SMPTE 315M, Television — Camera Positioning Information Conveyed by Ancillary Data Packets.

The proposals are available from Society Headquarters — SMPTE 314M for $46.00 and SMPTE 315M for $20.00.

Proposed SMPTE Recommended Practice

A Proposed SMPTE Recommended Practice is published for information: **RP 200,** Relative and Absolute Sound Pressure Levels for Motion-Picture Multichannel Sound Systems. RP 200 is available from Headquarters for $13.00.

Approved American National Standards

Eight American National Standards were approved recently by the American National Standards Institute:

ANSI/SMPTE 101-1998, Motion-Picture Film (16-mm) — Perforated 2R-3000 — Magnetic Striping;

ANSI/SMPTE 151-1998, Motion-Picture Film (8-mm Type S) — 16-mm Film Perforated 8-mm Type S, (1-3);

ANSI/SMPTE 154-1998, Motion-Picture Film (8-mm Type S) — Projectable Image Area and Projector Usage;

ANSI/SMPTE 184M-1998, Motion-Picture Film — Raw Stock Identification and Labeling;

ANSI/SMPTE 200M-1998, Motion-Picture Equipment (8-mm Type S) — Model 1 Camera Cartridge — Camera Run Length, Perforation Cutout and End-of-Run Notch;

ANSI/SMPTE 206-1998, Motion-Picture Equipment (8-mm Type S) — Model 1 Sound Camera Cartridge — Aperture, Profile, Film Position, Pressure Pad and Flatness (200-Ft Capacity);

ANSI/SMPTE 234-1998, Motion-Picture Film (8-mm Type R) — Projectable Image Area and Projector Usage; and

ANSI/SMPTE 304M-1998, Television — Broadcast Cameras — Hybrid Electrical and Fiber-Optic Connector.

Available from Headquarters, ANSI/SMPTE 101, 151, 154, 200, and 234 are $10.00 each; ANSI/SMPTE 184M and 206 are $13.00 each; and ANSI/SMPTE 304M is $16.00.

Approved SMPTE Standards

The Society approved the following SMPTE standards:

SMPTE 3-1998, Television Analog Recording — Frequency Response and Operating Level of Recorders and Reproducers — Audio 1 Record on 2-in Tape Operating at 15 and 7.5 in/s;

SMPTE 6-1998, Video Recording — 2-in Quadruplex Tape — Video, Audio and Tracking-Control Records;

SMPTE 15M-1998, Television Analog Recording — 1-in Type B Helical Scan — Basic System Parameters;

SMPTE 16M-1998, Television Analog Recording — 1-in Type B Helical Scan — Records;

SMPTE 17M-1998, Television Analog Recording — 1-in Type B Helical Scan — Frequency Response and Operating Level;

SMPTE 32M-1998, Video Recording — 1/2-in Type H — Cassette, Tape and Records;

SMPTE 59-1998, Motion-Picture Film (35-mm) — Camera Aperture Images and Usage;

SMPTE 161-1998, Motion-Picture Film (8-mm Type S) — Magnetic Striping;

SMPTE 162-1998, Motion-Picture Film (8-mm Type S) — 16-mm Film Perforated 8-mm Type S, (1-4) — Magnetic Striping;

SMPTE 163-1998, Motion-Picture Film (8-mm Type S) — 35-mm Film Perforated 8-mm Type S, 5R — Magnetic Striping;

SMPTE 202M-1998, Motion-Pictures — Dubbing Theaters, Review Rooms and Indoor Theaters — B-Chain Electroacoustic Response;

SMPTE 216-1998, Motion-Picture Film (35-mm) — Four-Track Striped Release Prints — Recorded Characteristics of Magnetic Audio Records;

SMPTE 217-1998, Motion-Picture Film (70-mm) — Striped Release Prints — Recorded Characteristic of Magnetic Audio Records;

SMPTE 236-1998, Motion-Picture Equipment (8-mm Type R) — Projection Reels; SMPTE 237-1998, Motion-Picture Film (35-mm) — Perforated DH-1870;

SMPTE 238M-1998, Television Analog Recording — 1/2-in Type L — Tapes and Cassettes;

SMPTE 254-1998, Motion-Picture Film (35-mm) — Manufacturer-Printed Latent Image Identification Information;

SMPTE 264M-1998, Television Digital Recording — 1/2-in Type D-3 Composite Format — 525/60;

SMPTE 265M-1998, Television Digital Recording — 1/2-in Type D-3 Composite Format — 626/50;

SMPTE 291M-1998, Television — Ancillary Data Packet and Space Formatting; SMPTE 302M, Television — Mapping of AES3 Data into MPEG-2 Transport Stream;

SMPTE 305M-1998, Television — Serial Data Transport Interface; SMPTE 306M-1998, Television Digital Recording — 6.35-mm Type D-7 Component Format — Video Compression at 25 Mb/s — 525/60 and 625/50;

SMPTE 307M-1998, Television Digital Recording — 6.35-mm Type D-7 Component Format — Tape Cassette;

SMPTE 308M-1998, Television — MPEG-2 4:2:2 Profile at High Level;

SMPTE 310M-1998, Television — Synchronous Serial Interface for MPEG-2 Digital Transport Streams; and

SMPTE 311M-1998, Television — Hybrid Electrical and Fiber Optical Camera Cable.

SMPTE 3, 6, 15M, 16M, 17M, 161, 162, 163, 216, 217, 236, 237, 308M, and 311M are 10.00 each. SMPTE 59, 254, 302M, and 310M are $13.00 each. SMPTE 202M is $16.00. SMPTE 291M and 305M are $18.00 each. SMPTE 238M is $20.00 and SMPTE 307M is $26.00. SMPTE 32M, 264M, and 265M are $32.00 each; and SMPTE 306M is $55.00.

— Carlos V. Girod, Jr., P.E., Director of Engineering

PROPOSED SMPTE STANDARD

SMPTE 314M

for Television — Data Structure for DV-Based Audio, Data and Compressed Video — 25 and 50 Mb/s

Table of contents

1 Scope
2 Normative references
3 Acronyms
4 Interface
5 Video compression
Annex A Differences between IEC 61834 and SMPTE 314M
Annex B Bibliography

1 Scope

This standard defines the DV-based data structure for the interface of digital audio, subcode data, and compressed video with the following parameters:

525/60 system –
4:1:1 image sampling structure, 25 Mb/s data rate

525/60 system –
4:2:2 image sampling structure, 50 Mb/s data rate

625/50 system –
4:1:1 image sampling structure, 25 Mb/s data rate

625/50 system –
4:2:2 image sampling structure, 50 Mb/s data rate

The standard does not define the DV compliant data structure for the interface of digital audio, subcode data, and compressed video with the following parameters:

625/50 system – 4:2:0 image sampling structure, 25 Mb/s data rate

The compression algorithm and the DIF structure conform to the DV data structure as defined in IEC 61834. The differences between the DV-based data structure defined in this standard and IEC 61834 are shown in annex A.

2 Normative references

The following standards, through reference in this text, constitute provisions of this standard. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

IEC 61834-1 (1997), Recording — Helical-Scan Digital Video Cassette Recording System Using 6,35 mm Magnetic Tape for Consumer Use (525-60, 625-50, 1125-60, and 1250-50 Systems) — Part 1: General Specifications

IEC 61834-2 (1997), Recording — Helical-Scan Digital Video Cassette Recording System Using 6,35 mm Magnetic Tape for Consumer Use (525-60, 625-50, 1125-60, and 1250-50 Systems) — Part 2: SD Format for 525-60 and 625-50 Systems

3 Acronyms

AAUX	Audio auxiliary data
AP1	Audio application ID
AP2	Video application ID
AP3	Subcode application ID
APT	Track application ID
Arb	Arbitrary
AS	AAUX source pack
ASC	AAUX source control pack
B/W	Black-and-white flag
CGMS	Copy generation management system
CM	Compressed macro block
DBN	DIF block number
DCT	Discrete cosine transform
DIF	Digital interface
DRF	Direction flag
Dseq	DIF sequence number
DSF	DIF sequence flag
DV	Identification of a compression family
EFC	Emphasis audio channel flag
EOB	End of block
FR	Identification for the first or second half of each channel
FSC	Identification of a DIF block in each channel
LF	Locked mode flag
QNO	Quantization number
QU	Quantization
Res	Reserved for future use
SCT	Section type
SMP	Sampling frequency
SSYB	Subcode sync block
STA	Status of the compressed macro block
STYPE (see note)	Signal type
Syb	Subcode sync block number
TF	Transmitting flag
VAUX	Video auxiliary data
VLC	Variable length coding
VS	VAUX source pack
VSC	VAUX source control pack

NOTE – STYPE as used in this standard is different from that in ANSI/IEEE 1394.

4 Interface

4.1 Introduction

As shown in figure 1, processed audio, video, and subcode data are output for different applications through a digital interface port.

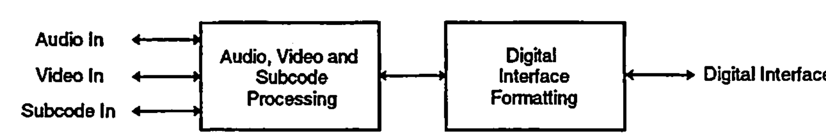


Figure 1 – Block diagram on digital interface

4.2 Data structure

The data structure of the compressed stream at the digital interface is shown in figures 2 and 3. Figure 2 shows the data structure for a 50 Mb/s structure, and figure 3 shows the data structure for a 25 Mb/s structure.

In the 50 Mb/s structure, the data of one video frame are divided into two channels. Each channel is divided into 10 DIF sequences for the 525/60 system and 12 DIF sequences for the 625/50 system.

In the 25 Mb/s structure, the data of one video frame are divided into 10 DIF sequences for the 525/60 system and 12 DIF sequences for the 625/50 system.

Each DIF sequence consists of a header section, subcode section, VAUX section, audio section, and video section with the following DIF blocks respectively:

Header section:	1 DIF block
Subcode section:	2 DIF blocks
VAUX section:	3 DIF blocks
Audio section:	9 DIF blocks
Video section:	135 DIF blocks

As shown in figure 2 and 3, each DIF block consists of a 3-byte ID and 77 bytes of data. DIF data bytes are numbered 0 to 79. Figure 4 shows the data structure of a DIF sequence for a 50 or 25 Mb/s structure.

THIS PROPOSAL IS PUBLISHED FOR COMMENT ONLY

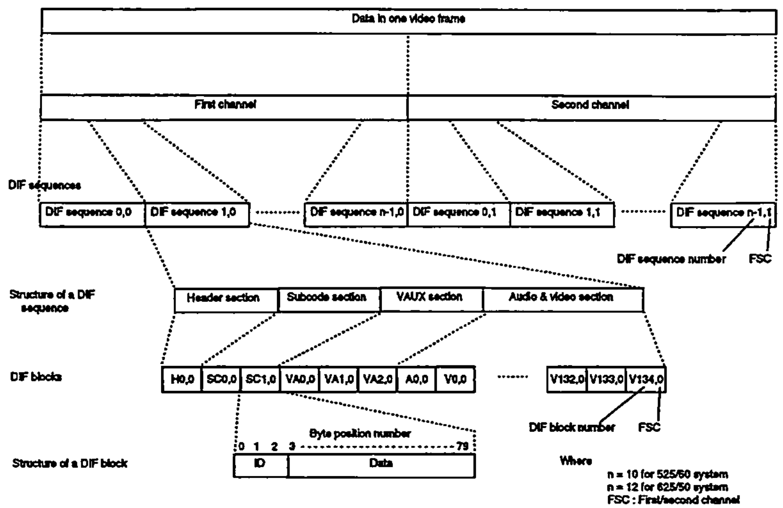


Figure 2 – Data structure of one video frame for 50 Mb/s structure

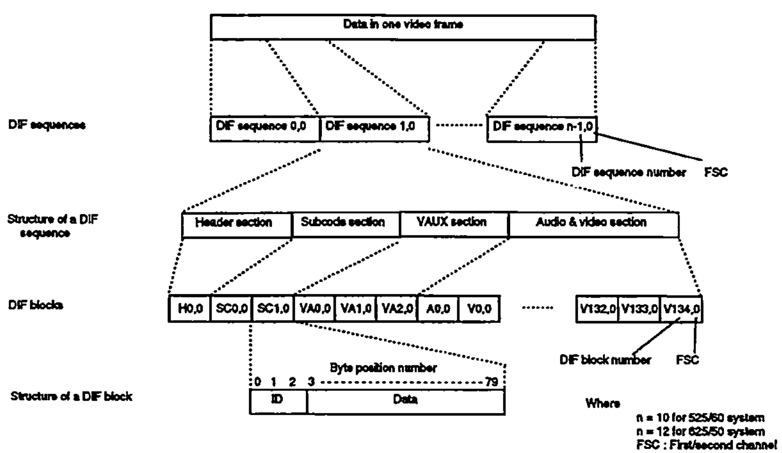
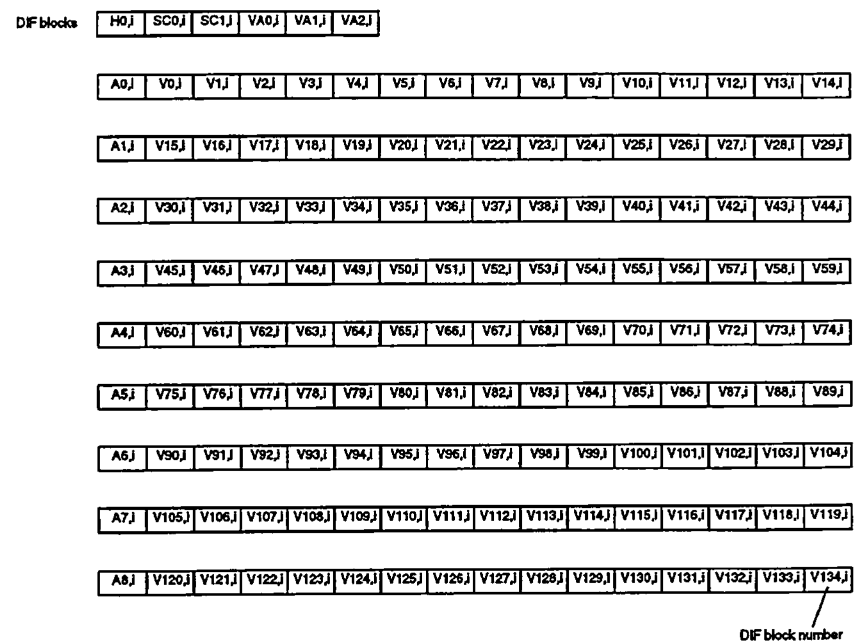


Figure 3 – Data structure of one video frame for 25 Mb/s structure



where

i	: FSC	
	i = 0	for 25 Mb/s structure
	i = 0,1	for 50 Mb/s structure
H0,i	: DIF block in header section	
SC0,i to SC1,i	: DIF blocks in subcode section	
VA0,i to VA2,i	: DIF blocks in VAUX section	
A0,i to A8,i	: DIF blocks in audio section	
V0,i to V134,i	: DIF blocks in video section	

Figure 4 – Data structure of a DIF sequence

4.3 Header section

4.3.1 ID

The ID part of each DIF block in the header section, shown in figures 2 and 3, consists of 3 bytes (ID0, ID1, ID2). Table 1 shows the ID content of a DIF block.

Table 1 – ID data of a DIF block

	Byte position number		
	Byte 0 (ID0)	Byte 1 (ID1)	Byte 2 (ID2)
MSB	SCT_2	$Dseq_3$	DBN_7
	SCT_1	$Dseq_2$	DBN_6
	SCT_0	$Dseq_1$	DBN_5
	Res	$Dseq_0$	DBN_4
	Arb	FSC	DBN_3
	Arb	Res	DBN_2
	Arb	Res	DBN_1
LSB	Arb	Res	DBN_0

ID contains the following:

SCT: Section type (see table 2)
Dseq: DIF sequence number (see tables 3 and 4)
FSC: Identification of a DIF block in each channel
50 Mb/s structure
FSC = 0: first channel
FSC = 1: second channel
25 Mb/s structure
FSC = 0
DBN: DIF block number (see table 5)
Arb: Arbitrary bit
Res: Reserved bit for future use
Default value shall be set to 1

Table 2 – Section type

SCT_2	SCT_1	SCT_0	Section type
0	0	0	Header
0	0	1	Subcode
0	1	0	VAUX
0	1	1	Audio
1	0	0	Video
1	0	1	Reserved
1	1	0	
1	1	1	

Table 3 – DIF sequence number for 525/60 system

$Dseq_3$	$Dseq_2$	$Dseq_1$	$Dseq_0$	Meaning
0	0	0	0	DIF sequence number 0
0	0	0	1	DIF sequence number 1
0	0	1	0	DIF sequence number 2
0	0	1	1	DIF sequence number 3
0	1	0	0	DIF sequence number 4
0	1	0	1	DIF sequence number 5
0	1	1	0	DIF sequence number 6
0	1	1	1	DIF sequence number 7
1	0	0	0	DIF sequence number 8
1	0	0	1	DIF sequence number 9
1	0	1	0	Not used
1	0	1	1	Not used
1	1	0	0	Not used
1	1	0	1	Not used
1	1	1	0	Not used
1	1	1	1	Not used

Table 4 – DIF sequence number for 625/50 system

$Dseq_3$	$Dseq_2$	$Dseq_1$	$Dseq_0$	Meaning
0	0	0	0	DIF sequence number 0
0	0	0	1	DIF sequence number 1
0	0	1	0	DIF sequence number 2
0	0	1	1	DIF sequence number 3
0	1	0	0	DIF sequence number 4
0	1	0	1	DIF sequence number 5
0	1	1	0	DIF sequence number 6
0	1	1	1	DIF sequence number 7
1	0	0	0	DIF sequence number 8
1	0	0	1	DIF sequence number 9
1	0	1	0	DIF sequence number 10
1	0	1	1	DIF sequence number 11
1	1	0	0	Not used
1	1	0	1	Not used
1	1	1	0	Not used
1	1	1	1	Not used

Table 5 – DIF block number

DBN_7	DBN_6	DBN_5	DBN_4	DBN_3	DBN_2	DBN_1	DBN_0	Meaning
0	0	0	0	0	0	0	0	DIF block number 0
0	0	0	0	0	0	0	1	DIF block number 1
0	0	0	0	0	0	1	0	DIF block number 2
0	0	0	0	0	0	1	1	DIF block number 3
:	:	:	:	:	:	:	:	:
1	0	0	0	0	1	1	0	DIF block number 134
1	0	0	0	0	1	1	1	Not used
:	:	:	:	:	:	:	:	:
1	1	1	1	1	1	1	1	Not used

4.3.2 Data

The data part (payload) of each DIF block in the header section is shown in table 6. Bytes 3 to 7 are active and bytes 8 to 79 are reserved.

Table 6 – Data (payload) in the header DIF block

	Byte position number of header DIF block							
	3	4	5	6	7	8	—	79
MSB	DSF	Res	TF1	TF2	TF3	Res	Res	Res
	0	Res	Res	Res	Res	Res	Res	Res
	Res	Res	Res	Res	Res	Res	Res	Res
	Res	Res	Res	Res	Res	Res	Res	Res
	Res	Res	Res	Res	Res	Res	Res	Res
	Res	APT_2	$AP1_2$	$AP2_2$	AP32	Res	Res	Res
	Res	APT_1	$AP1_1$	AP21	$AP3_1$	Res	Res	Res
LSB	Res	APT_0	$AP1_0$	AP20	$AP3_0$	Res	Res	Res

DSF: DIF sequence flag
DSF = 0 : 10 DIF sequences included in a channel (525/60 system)
DSF = 1 : 12 DIF sequences included in a channel (625/50 system)

APTn, AP1n, AP2n, AP3n: These data shall be identical as track application IDs (APTn = 001, AP1n = 001, AP2n = 001, AP3n = 001), if the source signal comes from a digital VCR. If the signal source is unknown, all bits for these data shall be set to 1.

TF: Transmitting flag:
TF1: Transmitting flag of audio DIF blocks
TF2: Transmitting flag of VAUX and video DIF blocks
TF3: Transmitting flag of subcode DIF blocks

TFn = 0: Data shall be valid.
TFn = 1: Data shall be invalid.

Res: Reserved bit for future use
Default value shall be set to 1.

4.4 Subcode section

4.4.1 ID

The ID part of each DIF block in the subcode section is described in 4.3.1. The section type shall be 001.

4.4.2 Data

The data part (payload) of each DIF block in the subcode section is shown in figure 5. The subcode data consists of 6 SSYBs, each 48 bytes long, and a reserved area of 29 bytes in each DIF block. SSYBs in a DIF sequence are numbered 0 to 11. Each SSYB is composed of SSYB ID equal to 2 bytes, FF_h, and an SSYB data payload of 5 bytes.

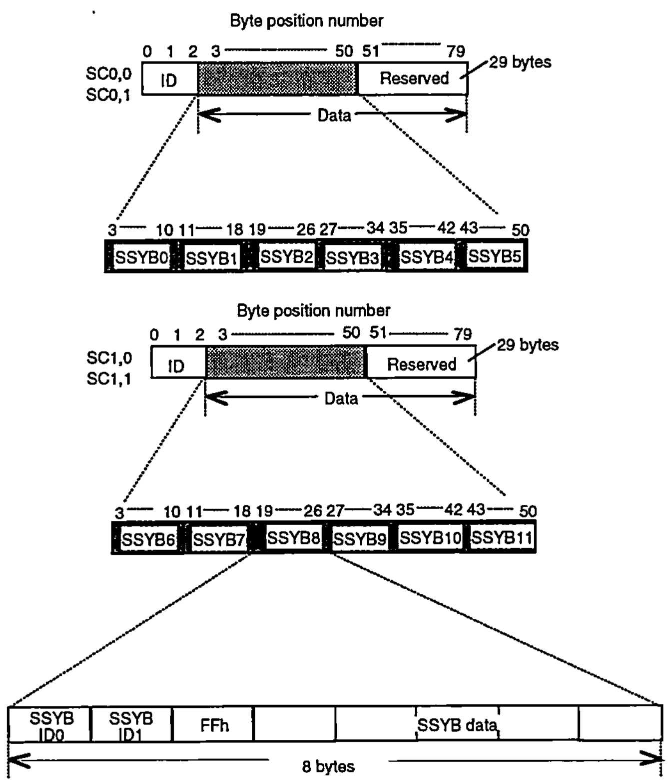


Figure 5 – Data in the subcode section

Table 7 – SSYB ID

Bit position	SSYB number 0 and 6		SSYB number 1 to 5 and 7 to 10		SSYB number 11	
	ID0	ID1	ID0	ID1	ID0	ID1
b7 (MSB)	FR	Arb	FR	Arb	FR	Arb
b6	$AP3_2$	Arb	Res	Arb	APT_2	Arb
b5	$AP3_1$	Arb	Res	Arb	APT_1	Arb
b4	$AP3_0$	Arb	Res	Arb	APT_0	Arb
b3	Arb	Syb_3	Arb	Syb_3	Arb	Syb3
b2	Arb	Syb_2	Arb	Syb2	Arb	Syb_2
b1	Arb	Syb_1	Arb	Syb_1	Arb	Syb1
b0 (LSB)	Arb	Syb_0	Arb	Syb0	Arb	Syb_0

NOTE – Arb = arbitrary bit.

4.4.2.1 SSYB ID

Table 7 shows SSYB ID (ID0, ID1). These data contain FR ID, application ID ($AP3_2$, $AP3_1$, $AP3_0$), and SSYB number (Syb_3, Syb_2 Syb_1, Syb_0).

FR ID is an identification for the first or second half of each channel:
  FR = 1: the first half of each channel
  FR = 0: the second half of each channel

The first half of each channel
  DIF sequence number 0, 1, 2, 3, 4 for the 525/60 system
  DIF sequence number 0, 1, 2, 3, 4, 5 for the 625/50 system

The second half of each channel
  DIF sequence number 5, 6, 7, 8, 9 for the 525/60 system
  DIF sequence number 6, 7, 8, 9. 1 0, 11 for the 625/50 system

If information is not available, all bits shall be set to 1.

4.4.2.2 SSYB data

Each SSYB data payload consists of a pack of 5 bytes as shown in figure 6. Table 8 shows pack header table (PCO byte organization). Table 9 shows the pack arrangement in SSYB data for each channel.

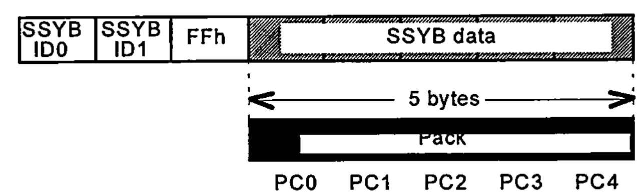


Figure 6 – Pack in SSYB

Table 8 – Pack header table

UPPER / LOWER	0000	0001	0010	0011	0100	0101	0110	0111	—	1111
0000						SOURCE	SOURCE			
0001						SOURCE CONTROL	SOURCE CONTROL			
0010										
0011		TIME CODE								
0100		BINARY GROUP								
0101										
|										
1111										NO INFO

Table 9 – Mapping of packet in SSYB data

SSYB number	First half of each channel	Second half of each channel
0	Reserved	Reserved
1	Reserved	Reserved
2	Reserved	Reserved
3	TC	TC
4	BG	Reserved
5	TC	Reserved
6	Reserved	Reserved
7	Reserved	Reserved
8	Reserved	Reserved
9	TC	TC
10	BG	Reserved
11	TC	Reserved

NOTES
1 TC = time code pack.
2 BG = binary group pack.
3 Reserved = default value of all bits shall be set to 1.
4 TC and BG data are the same within a single video frame. The time code data are an LCT type.

4.4.2.2.1 Time code pack (TC)

Table 10 shows a mapping of the time code pack. Time code data mapped to the time code packs remain the same within each video frame.

Table 10 - Mapping of time code pack

525/60 system

	MSB							LSB
PC0	0	0	0	1	0	0	1	1
PC1	CF	DF	TENS of FRAMES		UNITS of FRAMES			
PC2	PC	TENS of SECONDS			UNITS of SECONDS			
PC3	BGF0	TENS of MINUTES			UNITS of MINUTES			
PC4	BGF2	BGF1	TENS of HOURS		UNITS of HOURS			

625/50 system

	MSB							LSB
PC0	0	0	0	1	0	0	1	1
PC1	CF	Arb	TENS of FRAMES		UNITS of FRAMES			
PC2	BGF0	TENS of MINUTES			UNITS of SECONDS			
PC3	BGF2	TENS of MINUTES			UNITS of MINUTES			
PC4	PC	BGF1	TENS of HOURS		UNITS of HOURS			

NOTE – Detailed information is given in ANSI/SMPTE 12M.

CF: Color fame
0 = unsynchronized mode
1 = synchronized mode

DF: Drop frame flag
0 = Nondrop frame time code
1 = Drop frame time code

PC: Biphase mark polarity correction
0 = even
1 = odd

BGF: Binary group flag

Arb: Arbitrary bit

4.4.2.2.2 Binary group pack (BG)

Table 11 shows the mapping of the binary group pack. Binary group data mapped to the binary group packs remain the same within each video frame.

Table 11 – Mapping of binary group pack

	MSB							LSB
PC0	0	0	0	1	0	1	0	0
PC1	BINARY GROUP 2				BINARY GROUP 1			
PC2	BINARY GROUP 4				BINARY GROUP 3			
PC3	BINARY GROUP 6				BINARY GROUP 5			
PC4	BINARY GROUP 8				BINARY GROUP 7			

4.5 VAUX section

4.5.1 ID

The ID part of each DIF block in the VAUX section is described in 4.3.1. The section type shall be 010.

4.5.2 Data

The data part (payload) of each DIF block in the VAUX section is shown in figure 7. This figure shows the VAUX pack arrangement for each DIF sequence.

There are 15 packs, each 5 bytes long, and two reserved bytes in each VAUX DIF block payload. A default value for the reserved byte is set to FF_h.

Therefore, there are 45 packs in a DIF sequence. VAUX packs of the DIF blocks are sequentially numbered 0 to 44. This number is called a video pack number.

Table 12 shows the mapping of the VAUX packs of the VAUX DIF blocks. A VAUX source pack (VS) and a VAUX source control pack (VSC) must be present in each of the video compressed frames. The remaining VAUX packs of the DIF blocks in a DIF sequence are reserved and the value of all reserved words is set to FF_h.

If VAUX data are not transmitted, a NO INFO pack, which is filled up by FF_h, shall be transmitted.

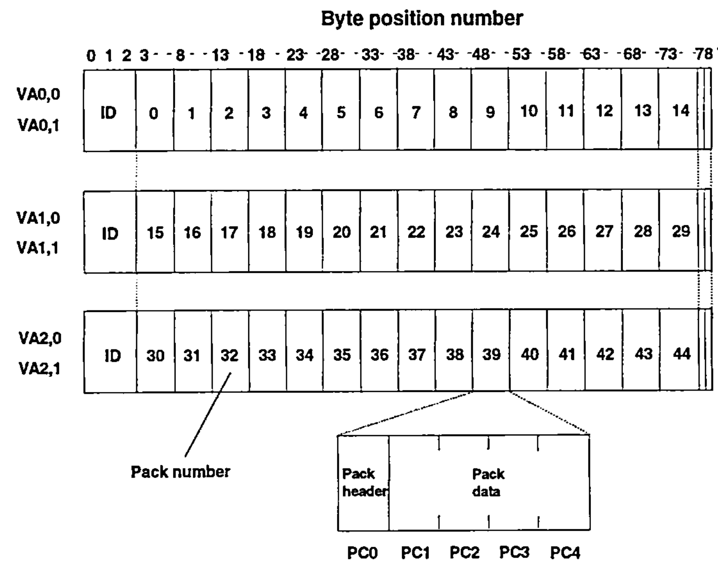


Figure 7 – Data in the VAUX section

Table 12 – Mapping of VAUX pack in a DIF sequence

Pack number		Pack data
Even DIF sequence	Odd DIF sequence	
39	0	VS
40	1	VSC

where

Even DIF sequence:
DIF sequence number 0, 2, 4, 6, 8 for 525/60 system
DIF sequence number 0, 2, 4, 6, 8, 10 for 625/50 system

Odd DIF sequence:
DIF sequence number 1, 3, 5, 7, 9 for 525/60 system
DIF sequence number 1, 3, 5, 7, 9, 11 for 625/50 system

4.5.2.1 VAUX source pack (VS)

Table 13 shows the mapping of a VAUX source pack.

Table 13 – Mapping of VAUX source pack

	MSB							LSB
PC0	0	1	1	0	0	0	0	0
PC1	Res	Res	Res	Res	Res	Res	Res	Res
PC2	B/W	EN	CLF		Res	Res	Res	Res
PC3	Res	Res	50/60	STYPE				
PC4	VISC							

B/W: Black-and-white flag
0 = Black and white
1 = Color

EN: Color frames enable flag
0 = CLF is valid
1 = CLF is invalid

CLF: Color frames identification code (see ITU-R BT.470-4)
For 525/60 system
00b = Color frame A
01b = Color frame B
Others = Reserved

For 625/50 system
00b = 1st, 2nd field
01b = 3rd, 4th field
10b = 5th, 6th field
11b = 7th, 8th field

50/60:
0 = 60-field system
1 = 50-field system

STYPE: STYPE defines a signal type of video signal
00000b = 4:1:1 compression
00001b = Reserved
| |
00011b = Reserved
00100b = 4:2:2 compression
00101b = Reserved
| |
11111b = Reserved

VISC:
10001000b = –180
| |
00000000b = 0
| |
01111000b = 180
01111111b = No information
Other = Reserved

Res: Reserved bit for future use
Default value shall be set to 1

4.5.2.2 VAUX source control pack (VSC)

Table 14 shows the mapping of the VAUX source control pack.

Table 14 – Mapping of VAUX source control pack

	MSB							LSB
PC0	0	1	1	0	0	0	0	1
PC1	CGMS		Res	Res	Res	Res	Res	Res
PC2	Res	Res	0	0	Res	DISP		
PC3	FF	FS	FC	IL	Res	Res	0	0
PC4	Res	Res	Res	Res	Res	Res	Res	Res

CGMS: Copy generation management system

CGMS	Copy possible generation
0 0	Copy free
0 1	Reserved
1 0	
1 1	

DISP: Display select mode

DISP	Aspect ratio and format	Position
0 0 0	4:3 full format	Not applicable
0 0 1	Reserved	
0 1 0	16:9 full format (squeeze)	Not applicable
0 1 1 – 1 1 1	Reserved	

FF: Frame/field flag
FF indicates whether two consecutive fields are delivered, or one field is repeated twice during one frame period.
0 = Only one of two fields is delivered twice; 1 = Both fields are delivered in order.

FS: First/second field flag
FS indicates a field which is delivered during the field one period.
0 = Field 2 is delivered; 1 = Field 1 is delivered.

FF	FS	Output field
1	1	Field 1 and field 2 are output in this order (1,2 sequence)
1	0	Field 2 and field 1 are output in this order (2,1 sequence)
0	1	Field 1 is output twice
0	0	Field 2 is output twice

FC: Frame change flag
FC indicates whether the picture of the current frame is repeated based on the immediate previous frame.
0 = Same picture as the previous frame; 1 = Different picture from the previous frame

IL: Interlace flag
0 = Noninterlaced; 1 = Interlaced

Res: Reserved bit for future use
Default value shall be set to 1.

4.6 Audio section

4.6.1 ID

The ID part of each DIF block in the audio section is described in 4.3.1. The section type shall be 011.

4.6.2 Data

The data part (payload) of each DIF block in the audio section is shown in figure 8. The data of a DIF block in the audio DIF block are composed of 5 bytes of audio auxiliary data (AAUX) and 72 bytes of audio data which are encoded and shuffled by the process shown in figure 8.

4.6.2.1 Audio encoding

4.6.2.1.1 Source coding

Each audio input signal is sampled at 48 kHz, with 16-bit quantization. The system provides two channels of audio for 25 Mb/s structure or four channels of audio for 50 Mb/s structure. Audio data for each audio channel are located in an audio block respectively.

An audio block consists of 45 DIF blocks (9 DIF blocks × 5 DIF sequences) for the 525/60 system; and 54 DIF blocks (9 DIF blocks × 6 DIF sequences) for the 625/50 system.

4.6.2.1.2 Emphasis

Audio encoding is carried out with the first order preemphasis of 50/15 µs. For analog input recording, emphasis shall be off in the default state.

4.6.2.1.3 Audio error code

In the encoded audio data, 8000_h shall be assigned as an audio error code to indicate an invalid audio sample. This code corresponds to negative full-scale value in ordinary twos complement representation. When the encoded data includes 8000_h, it shall be converted to 8001_h.

4.6.2.1.4 Relative audio-video timing

The audio frame duration equals a video frame period. An audio frame begins with an audio sample acquired within the duration of minus 50 samples relative to zero samples from the first preequalizing pulse of the vertical blanking period of the input video signal. The first preequalizing pulse means the start of line number 1 for the 525/60 system, and the middle of line number 623 for the 625/50 system.

4.6.2.1.5 Audio frame processing

This standard provides audio frame processing in the locked mode.

The sampling frequency of the audio signal is synchronous with the video frame frequency. Audio data are processed in frames. For an audio channel, each frame contains 1602 or 1600 audio samples for the 525/60 system or 1920 audio samples for the 625/50 system. For the 525/60 system, the number of audio samples per frame shall follow the five-frame sequence as shown below:

1600, 1602, 1602, 1602, 1602 samples.

The sample audio capacity shall be capable of 1620 samples per frame for the 525/60 system or 1944 samples per frame for the 625/50 system. The unused space at the end of each frame is filled with arbitrary values.

Byte position number

0 1 2	3 — 7	8 — 79
ID	Audio auxiliary data	Audio data

Figure 8 – Data in the audio section

4.6.2.2 Audio shuffling

The 16-bit audio data word is divided into two bytes; the upper byte which contains MSB, and the lower byte LSB, as shown in figure 9. Audio data shall be shuffled over DIF sequences and DIF blocks within a frame. The data bytes are defined as D_n (n = 0, 1, 2, ...) which is sampled at nth order within a frame and shuffled by each D_n unit.

The data shall be shuffled through a process expressed by the following equations:

525/60 system:

DIF sequence number:
(INT (n/3) + 2 × (n mod 3)) mod 5 for CH1, CH3
(INT (n/3) + 2 × (n mod 3)) mod 5 + 5 for CH2, CH4

Audio DIF block number:
3 × (n mod 3) + INT ((n mod 45) / 15)
where FSC = 0: CH1, CH2
FSC = 1: CH3, CH4

Byte position number:
8 + 2 × INT(n/45) for the most significant byte
9 + 2 × INT(n/45) for the least significant byte
where n = 0 to 1619

625/50 system:

DIF sequence number:
(INT (n/3) + 2 × (n mod 3)) mod 6 for CH1, CH3
(INT (n/3) + 2 × (n mod 3)) mod 6 + 6 for CH2, CH4

Audio DIF block number:
3 × (n mod 3) + INT ((n mod 54) / 18)
where FSC = 0: CH1, CH2
FSC = 1: CH3, CH4

Byte position number:
8 + 2 × INT(n/54) for the most significant byte
9 + 2 × INT(n/54) for the least significant byte
where n = 0 to 1943

4.6.2.3 Audio auxiliary data (AAUX)

AAUX shall be added to the shuffled audio data as shown in figures 8 and 10. The AAUX pack shall include an AAUX pack header and data (AAUX payload). The length of the AAUX pack shall be 5 bytes as shown in figure 10, which depicts the AAUX pack arrangement. Packs are numbered from 0 to 8 as shown in figure 10. This number is called an audio pack number.

Table 15 shows the mapping of an AAUX pack. An AAUX source pack (AS) and an AAUX source control pack (ASC) must be included in the compressed stream.

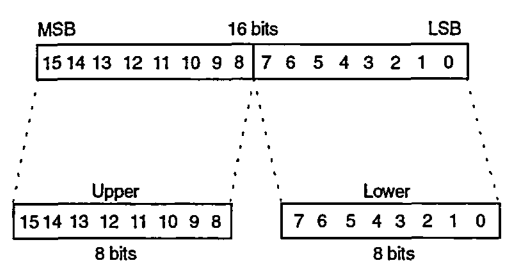


Figure 9 – Conversion of audio sample to audio data bytes

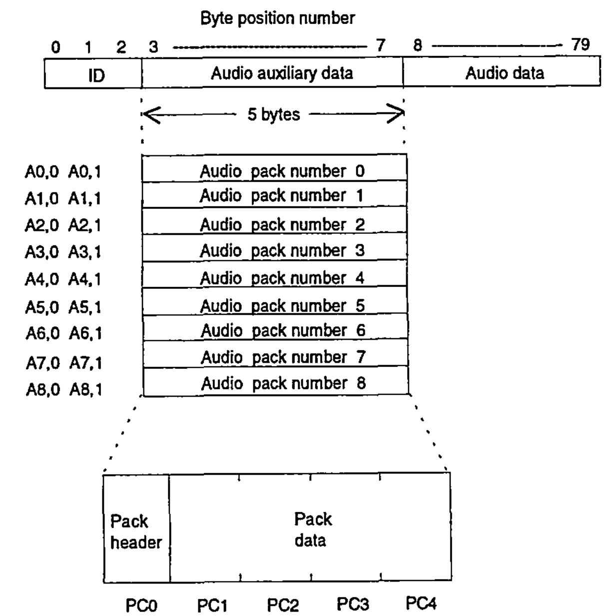


Figure 10 – Arrangement of AAUX packs in audio auxiliary data

Table 15 – Mapping of AAUX pack in a DIF sequence

Audio pack number		Pack data
Even DIF sequence	Odd DIF sequence	
3	0	AS
4	1	ASC

where

Even DIF sequence:
DIF sequence number 0, 2, 4, 6, 8 for 525/60 system
DIF sequence number 0, 2, 4, 6, 8, 10 for 625/50 system

Odd DIF sequence:
DIF sequence number 1, 3, 5, 7, 9 for 525/60 system
DIF sequence number 1, 3, 5, 7, 9, 11 for 625/50 system

4.6.2.3.1 AAUX source pack (AS)

The AAUX source pack is configured as shown in table 16.

Table 16 – Mapping of AAUX source pack

	MSB							LSB
PC0	0	1	0	1	0	0	0	0
PC1	LF	Res	AF SIZE					
PC2	0	CHN		Res	AUDIO MODE			
PC3	Res	Res	50/60	STYPE				
PC4	Res	Res	SMP			QU		

LF: Locked mode flag
Locking condition of audio sampling frequency with video signal
0 = Locked mode; 1 = Reserved

AF SIZE: The number of audio samples per frame
010100b = 1600 samples/frame (525/60 system)
010110b = 1602 samples/frame (525/60 system)
011000b = 1920 samples/frame (625/50 system)
Others = Reserved

CHN: The number of audio channels within an audio block
00b = One audio channel per audio block
Others = Reserved
The audio block is composed of 45 DIF blocks of the audio section in five consecutive DIF sequences for the 525/60 system, and 54 DIF blocks of the audio section in six consecutive DIF sequences for the 625/50 system.

AUDIO MODE: The contents of the audio signal on each audio channel
0000b = CH1 (CH3)
0001b = CH2 (CH4)
1111b = Invalid audio data
Others = Reserved

50/60:
0 = 60-field system
1 = 50-field system

STYPE: STYPE defines audio blocks per video frame
00000b = 2 audio blocks
00010b = 4 audio blocks
Others = Reserved

SMP: Sampling frequency
000b = 48 kHz
Others = Reserved

QU: Quantization
000b = 16 bits linear
Others = Reserved

Res: Reserved bit for future use
Default value shall be set to 1

4.6.2.3.2 AAUX source control pack (ASC)

The AAUX source control pack is configured as shown in table 17.

Table 17 – Mapping of AAUX source control pack

	MSB							LSB
PC0	0	1	0	1	0	0	0	1
PC1	CGMS		Res	Res	Res	Res	EFC	
PC2	REC ST	REC END	FADE ST	FADE END	Res	Res	Res	Res
PC3	DRF	SPEED						
PC4	Res	Res	Res	Res	Res	Res	Res	Res

CGMS: Copy generation management system

CGMS	Copy possible generation
0 0	Copy free
0 1	Reserved
1 0	
1 1	

EFC: Emphasis audio channel flag
00b = emphasis off
01b = emphasis on
Others = reserved
EFC shall be set for each audio block.

REC ST: Recording start point
0 = recording start point
1 = not recording start point
At a recording start frame, REC ST 0 lasts for a duration of one audio block which is equal to 5 or 6 DIF sequences for each audio channel.

REC END: Recording end point
0 = recording end point
1 = not recording end point
At a recording end frame, REC END 0 lasts for a duration of one audio block which is equal to 5 or 6 DIF sequences for each audio channel.

FADE ST: Fading of recording start point
0 = fading off
1 = fading on
The information of FADE ST shall be effective only at the recording start frame (REC ST = 0).
If FADE ST is 1 at the recording start frame, the output audio signal should be faded in from the first sampling signal of the frame. If FADE ST is 0 at the recording start frame, the output audio signal should not be faded.

FADE END: Fading of recording end point
0 = fading off
1 = fading on
The information of FADE END shall be effective only at the recording end frame (REC END = 0).
If FADE END is 1 at the recording end frame, the output audio signal should be faded out to the last sampling signal of the frame. If FADE END is 0 at the recording end frame, the output audio signal should not be faded.

DRF: Direction flag
0 = reverse direction
1 = forward direction

SPEED: Shuttle speed ov VTR

SPEED	Shuttle speed of VTR	
	525/60 system	625/50 system
0000000	0/120 (=0)	0/100 (=0)
0000001	1/120	1/100
:	:	:
1100100	100/120	100/100 (=1)
:	:	Reserved
1111000	120/120 (=1)	Reserved
:	Reserved	Reserved
1111110	Reserved	Reserved
1111111	Data invalid	Data invlaid

RES: Reserved bit for future use.
Default value shall be set to 1.

4.7 Video section

4.7.1 ID

The ID part of each DIF block in the video section is described in 4.3.1. The section type shall be 100.

4.7.2 Data

The data part (payload) of each DIF block in the video section consists of 77 bytes of video data which shall be sampled, shuffled, and encoded. Video data of every video frame are processed as described in clause 4.

4.7.2.1 DIF block and compressed macro block

Correspondence between video DIF blocks and video compressed macro blocks is shown in tables 18 and 19. Table 18 shows correspondence between video DIF blocks for 50 Mb/s structure and video compressed macro blocks of 4:2:2 compression. Table 19 shows correspondence between the video DIF blocks for 25 Mb/s structure and video compressed macro blocks of 4:1:1 compression.

The rule defining the correspondence between video DIF blocks and compressed macro blocks is shown below:

```
50 Mb/s structure – 4:2:2 compression
   if (525/60 system) n = 10 else n = 12;
   for (i = 0; i<n; i++){
     a = i;
     b = (i - 6) mod n;
     c = (i - 2) mod n;
     d = (i - 8) mod n;
     e = (i - 4) mod n;
     p = a;
     q = 3;
     for (j = 0; j<5; j++){
        for (k = 0; k<27; k++){
           V (5 × k + q),0 of DSNp = CM 2i,j,k;
           V (5 × k + q),1 of DSNp = CM 2i + 1,j,k;
        }
             if (q == 3) {p = b; q = 1;}
        else if (q == 1) {p = c; q = 0;}
        else if (q == 0) {p = d; q = 2;}
        else if (q == 2) {p = e; q = 4;}
     }
   }

25 Mb/s structure – 4:1:1 compression
   if (525/60 system) n = 10 else n = 12;
   for (i = 0; i<n; i++){
     a = i;
     b = (i - 6) mod n;
     c = (i - 2) mod n;
     d = (i - 8) mod n;
     e = (i - 4) mod n;
     p = a;
     q = 3;
     for (j = 0; j<5; j++){
        for (k = 0; k<27; k++){
           V (5 × k + q) of DSNp = CM 2i,j,k;
        }
             if (q == 3) {p = b; q = 1;}
        else if (q == 1) {p = c; q = 0;}
        else if (q == 0) {p = d; q = 2;}
        else if (q == 2) {p = e; q = 4;}
     }
   }
```


Table 18 – Video DIF blocks and compressed macro blocks for 50 Mb/s structure — 4:2:2 compression

DIF sequence number	DIF block	Compressed macro block
0	V0,0	CM 4,2,0
	V0,1	CM 5,2,0
	V1,0	CM 12,1,0
	V1,1	CM 13,1,0
	V2,0	CM 16,3,0
	V2,1	CM 17,3,0
	:	:
	V134,0	CM 8,4,26
	V134,1	CM 9,4,26
1	V0,0	CM 6,2,0
	V0,1	CM 7,2,0
	V1,0	CM 14,1,0
	V1,1	CM 15,1,0
	V2,0	CM 18,3,0
	V2,1	CM 19,3,0
	:	:
	V134,0	CM 10,4,26
	V134,1	CM 11,4,26
:	:	:
n-1	V0,0	CM 2,2,0
	V0,1	CM 3,2,0
	V1,0	CM 10,1,0
	V1,1	CM 11,1,0
	V2,0	CM 14,3,0
	V2,1	CM 15,3,0
	:	:
	V134,0	CM 6,4,26
	V134,1	CM 7,4,26

NOTE – n = 10 for 525/60 system; n = 12 for 625/50 system.

Table 19 – Video DIF blocks and compressed macro blocks for 25 Mb/s structure — 4:1:1 compression

DIF sequence number	DIF block	Compressed macro block
0	V0,0	CM 2,2,0
	V1,0	CM 6,1,0
	V2,0	CM 8,3,0
	V3,0	CM 0,0,0
	V4,0	CM 4,4,0
	:	:
	V133,0	CM 0,0,26
	V134,0	CM 4,4,26
1	V0,0	CM 3,2,0
	V1,0	CM 7,1,0
	V2,0	CM 9,3,0
	V3,0	CM 1,0,0
	V4,0	CM 5,4,0
	:	:
	V133,0	CM 1,0,26
	V134,0	CM 5,4,26
⋮	⋮	⋮
n-1	V0,0	CM 1,2,0
	V1,0	CM 5,1,0
	V2,0	CM 7,3,0
	V3,0	CM n - 1,0,0
	V4,0	CM 3,4,0
	:	:
	V133,0	CM n - 1,0,26
	V134,0	CM 3,4,26

NOTE – n = 10 for 525/60 system; n = 12 for 625/50 system.

5 Video compression

This clause includes video compression processing for 4:2:2 and 4:1:1 compression.

5.1 Video structure

The video signal is sampled with a frequency of 13.5 MHz for luminance (Y) and 6.75 MHz for color difference (C_R, C_B). The data of the vertical blanking area and the horizontal blanking area are discarded, then the remainder of the video data is shuffled in the video frame. The original quantity of video data shall be reduced by use of bit-rate reduction techniques which adopt DCT and VLC.

The process of the bit-rate reduction is as follows: Video data are assigned to a DCT block (8x8 samples). Two luminance DCT blocks and two color-difference DCT blocks form a macro block for 4:2:2 compression. For 4:1:1 compression, four luminance DCT blocks and two color-difference DCT blocks form a macro block. Five macro blocks constitute a video segment. A video segment is further compressed into five compressed macro blocks by use of the DCT and VLC techniques.

5.1.1 Sampling structure

The sampling structure is identical to the sampling structure of 4:2:2 component television signals described in ITU-R BT.601. Sampling of luminance (Y) and two color-difference signals (C_R, C_B) in the 4:2:2 structure are described in table 20.

Line structure in one frame

For the 525/60 system, 240 lines for Y, C_R, and C_B signals from each field shall be transmitted. For the 625/50 system, 288 lines for Y, C_R, and C_B signals from each field shall be transmitted. The transmitted lines on a TV frame are defined in table 20.

Pixel structure in one frame

4:2:2 compression –
All sampled pixels, 720 luminance pixels per line and 360 color-difference pixels, are retained for processing as shown in figures 11 and 12. The sampling process starts simultaneously for both luminance and color-difference signals. Each pixel has a value from –127 to +126 which is obtained by the subtraction of 128 from the input video signal level.

4:1:1 compression –
All sampled luminance pixels, 720 pixels per line, are retained for processing. Of 360 color-difference pixels sampled per line, every other pixel is discarded, leaving 180 pixels for processing. The sampling process starts simultaneously for both luminance and color-difference signals. Figures 13 and 14 show the sampling process in detail. Each pixel has a value in range from –127 to +126 which is obtained by the subtraction of 128 from the input video signal level.

Table 20 – Construction of video signal sampling (4:2:2)

		525/60 system	625/50 system
Sampling frequency	Y	13.5 MHz	
	C_R, C_B	6.75 MHz	
Total number of pixels per frame	Y	858	864
	C_R, C_B	429	432
Number of active pixels per line	Y	720	
	C_R, C_B	360	
Total number of lines per frame		525	625
Number of active lines per frame		480	576
Active line numbers	Field 1	23 to 262	23 to 310
	Field 2	285 to 524	335 to 622
Quantization		Each sample is linearly quantized to 8 bits for Y, C_R, C_B	
Relation between video signal level and quantized level	Scale	1 to 254	
	Y	Video signal level of white: 235	Quantized level 220
		Video signal level of black: 16	
	C_R, C_B	Video signal level of gray: 128	Quantized level 225

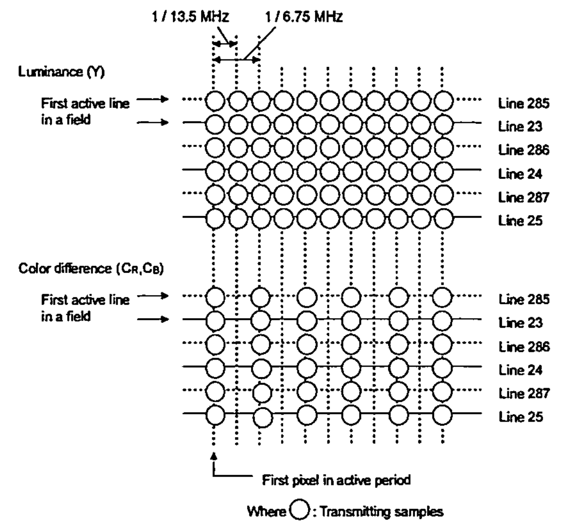


Figure 11 – Transmitting samples of 525/60 system for 4:2:2 compression

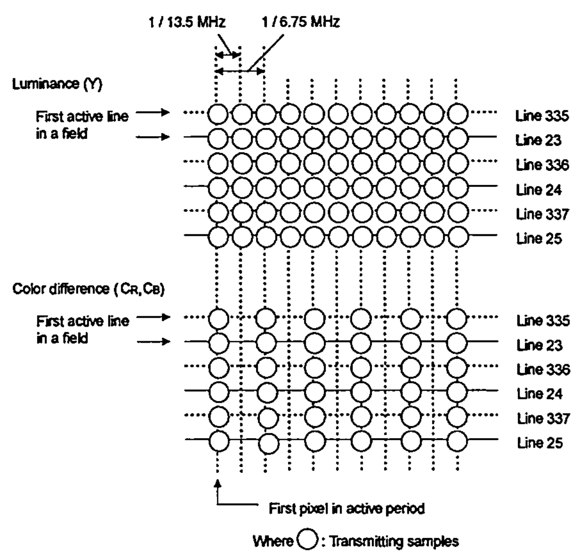


Figure 12 – Transmitting samples of 625/50 system for 4:2:2 compression

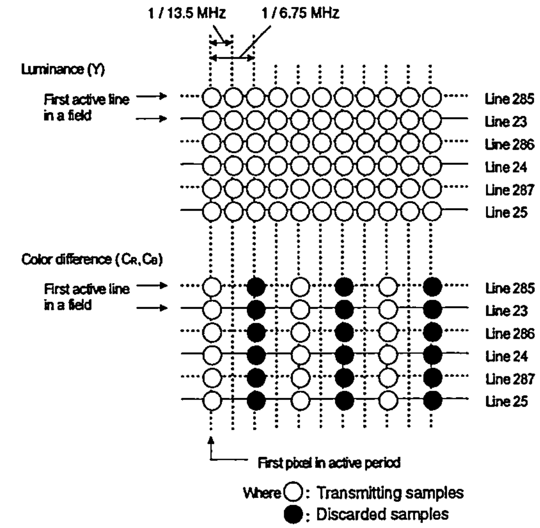


Figure 13 – Transmitting samples of 525/60 system for 4:1:1 compression

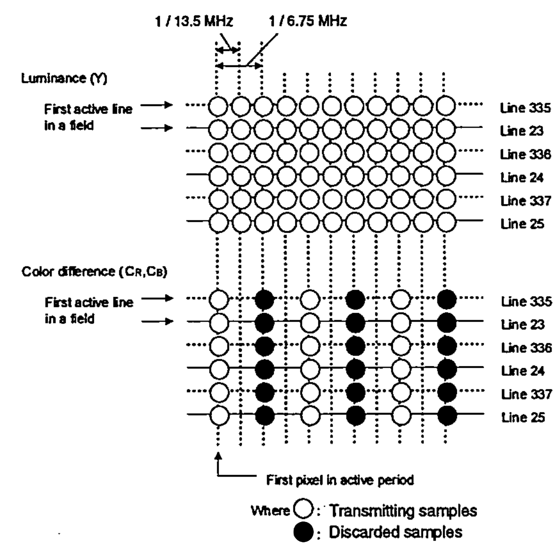


Figure 14 – Transmitting samples of 625/50 system for 4:1:1 compression

5.1.2 DCT block

The Y, C_R, and C_B pixels in one frame shall be divided into DCT blocks as shown in figure 15. All DCT blocks for 4:2:2 compression and DCT blocks for 4:1:1 compression, with the exception of the rightmost DCT blocks in C_R and C_B for 4:1:1 compression, are structured as a rectangular area of eight vertical lines and eight horizontal pixels for each DCT block. The value of x shows the horizontal coordinate from the left and the value of y shows the vertical coordinate from the top. Odd lines of y = 1, 3, 5, 7 are the horizontal lines of field one, and even lines of y = 0, 2, 4, 6 are those of field two.

In the 4:1:1 compression mode, the rightmost DCT blocks in C_R and C_B are structured with 16 vertical lines and four horizontal pixels. The rightmost DCT block shall be reconstructed to eight vertical lines and eight horizontal pixels by moving the lower part of eight vertical lines and four horizontal pixels to the higher part of eight vertical lines and four horizontal pixels as shown in figure 16.

DCT block arrangement in one frame for 525/60 blocks

The arrangement of horizontal DCT blocks in one frame in the 4:2:2 compression mode is shown in figure 17, and in the 4:1:1 compression mode in figure 18. The same horizontal arrangement is repeated with 60 DCT blocks in the vertical direction. Pixels in one frame are divided into 10,800 DCT blocks for 4:2:2 compression and 8,100 DCT blocks for 4:1:1 compression.

4:2:2 compression

Y: 60 vertical DCT blocks × 90 horizontal DCT blocks = 5400 DCT blocks

C_R: 60 vertical DCT blocks × 45 horizontal DCT blocks = 2700 DCT blocks

C_B: 60 vertical DCT blocks × 45 horizontal DCT blocks = 2700 DCT blocks

4:1:1 compression

Y: 60 vertical DCT blocks × 90 horizontal DCT blocks = 5400 DCT blocks

C_R: 60 vertical DCT blocks × 22.5 horizontal DCT blocks = 1350 DCT blocks

C_B: 60 vertical DCT blocks × 22.5 horizontal DCT blocks = 1350 DCT blocks

DCT block arrangement in one frame for 625/50 blocks

The arrangement of horizontal DCT blocks in one frame for the 4:2:2 compression mode is shown in figure 17, and for the 4:1:1 compression mode in figure 18. The same horizontal arrangement is repeated to 72 DCT blocks in the vertical direction. Pixels in one frame are divided into 12,960 DCT blocks for 4:2:2 compression and 9,720 DCT blocks for 4:1:1 compression.

4:2:2 compression

Y: 72 vertical DCT blocks × 90 horizontal DCT blocks = 6480 DCT blocks

C_R: 72 vertical DCT blocks × 45 horizontal DCT blocks = 3240 DCT blocks

C_B: 72 vertical DCT blocks × 45 horizontal DCT blocks = 3240 DCT blocks

4:1:1 compression

Y: 72 vertical DCT blocks × 90 horizontal DCT blocks = 6480 DCT blocks

C_R: 72 vertical DCT blocks × 22.5 horizontal DCT blocks = 1620 DCT blocks

C_B: 72 vertical DCT blocks × 22.5 horizontal DCT blocks = 1620 DCT blocks

5.1.3 Macro block

As shown in figure 19, each macro block in the 4:2:2 compression mode consists of four DCT blocks. As shown in figure 20, each macro block in the 4:1:1 compression mode consists of six DCT blocks. In the 4:1:1 compression mode, each macro block consists of four horizontally adjacent DCT blocks of Y, one DCT block of C_R and one DCT block of C_B on a television screen. The rightmost macro block on the television screen consists of four vertically and horizontally adjacent DCT blocks of Y, one DCT block of C_R, and one DCT block of C_B.

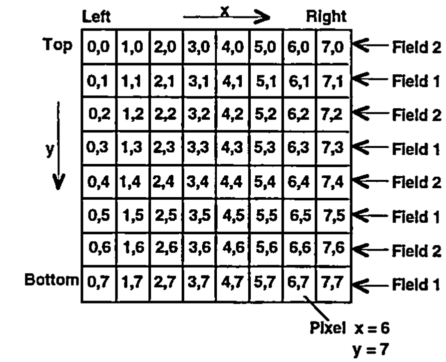


Figure 15 – DCT block and pixel coordinates

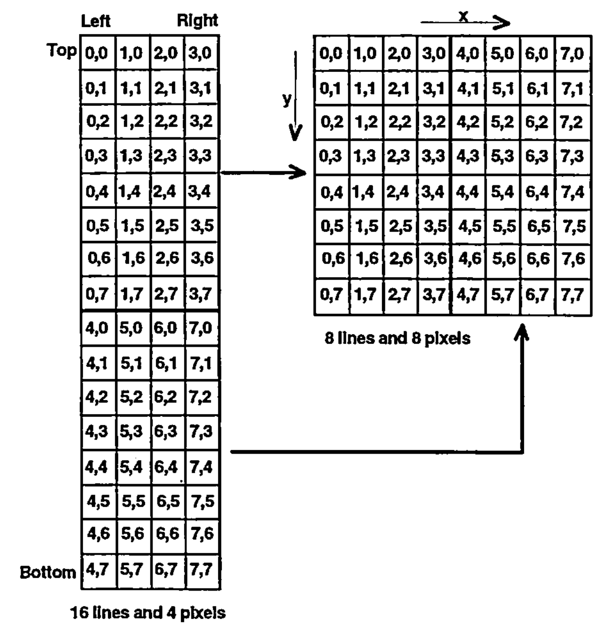


Figure 16 – Rightmost DCT block in color-difference signal for 4:1:1 compression mode

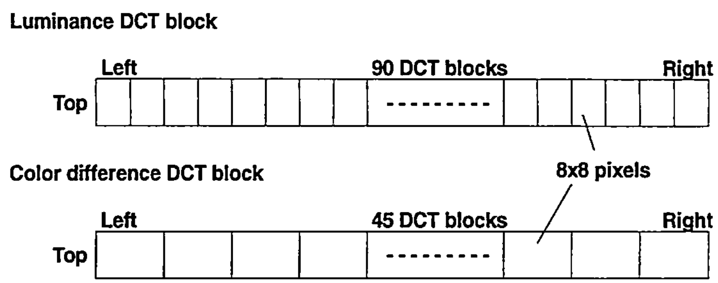


Figure 17 – DCT block arrangement for 4:2:2 compression

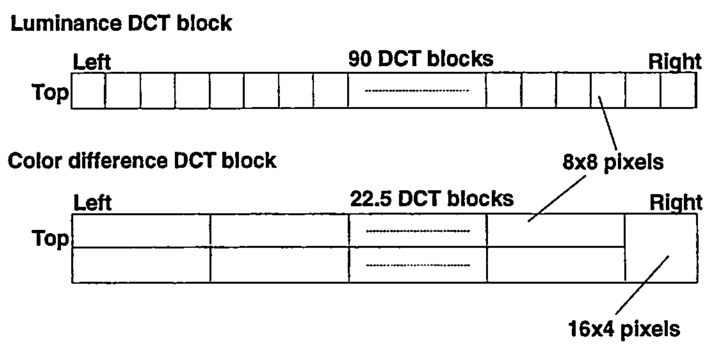


Figure 18 – DCT block arrangement for 4:1:1 compression

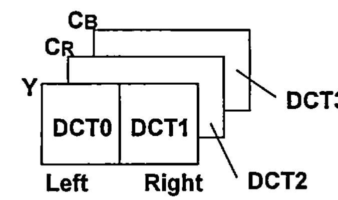


Figure 19 – Macro block and DCT blocks for 4:2:2 compression

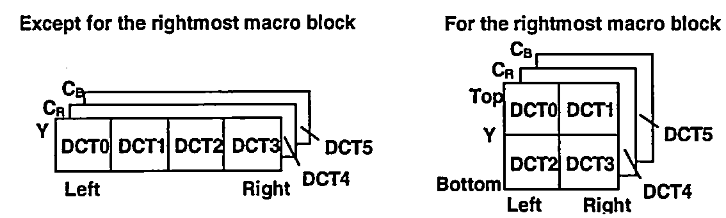


Figure 20 – Macro block and DCT blocks for 4:1:1 compression

Macro block arrangement in one frame for 525/60 system

The arrangement of macro blocks in one frame is shown in figure 21 for 4:2:2 compression and figure 22 for 4:1:1 compression. Each small rectangle shows a macro block. Pixels in one frame are distributed into 2700 macro blocks for 4:2:2 compression and 1350 macro blocks for 4:1:1 compression.

4:2:2 compression
60 vertical macro blocks × 45 horizontal macro blocks = 2700 macro blocks

4:1:1 compression
60 vertical macro blocks × 22.5 horizontal macro blocks = 1350 macro blocks

Macro block arrangement in one frame for 625/50 system

The arrangement of macro blocks in one frame is shown in figure 23 for 4:2:2 compression and figure 24 for 4:1:1 compression. Each small rectangle shows a macro block. Pixels in one frame are distributed into 3240 macro blocks for 4:2:2 compression and 1620 macro blocks for 4:1:1 compression.

4:2:2 compression
72 vertical macro blocks × 45 horizontal macro blocks = 3240 macro blocks

4:1:1 compression
72 vertical macro blocks × 22.5 horizontal macro blocks = 1620 macro blocks

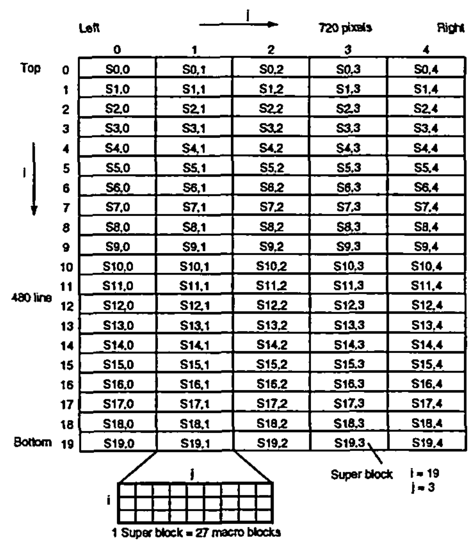


Figure 21 – Super blocks and macro blocks in one television frame for 525/60 system for 4:2:2 compression

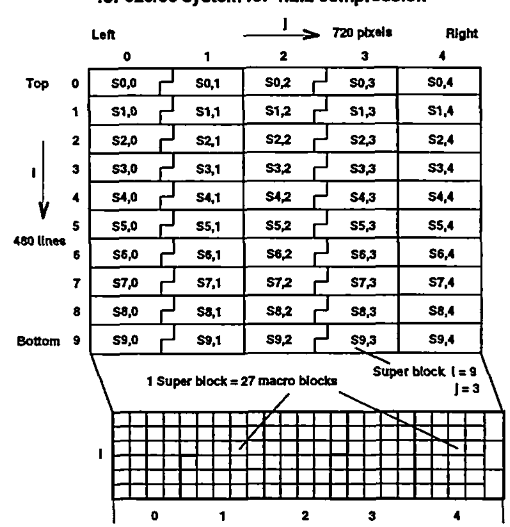


Figure 22 – Super blocks and macro blocks in one television frame for 525/60 system for 4:1:1 compression

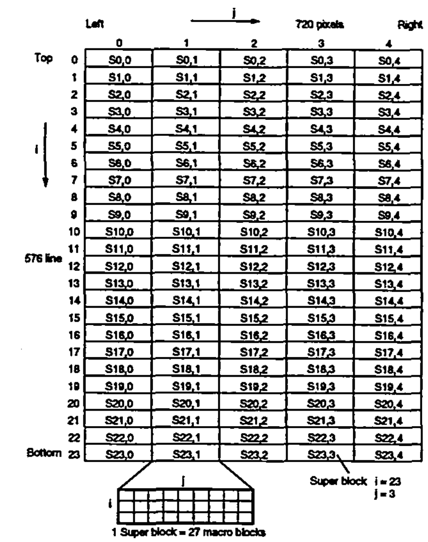


Figure 23 – Super blocks and macro blocks in one television frame for 625/50 system for 4:2:2 compression

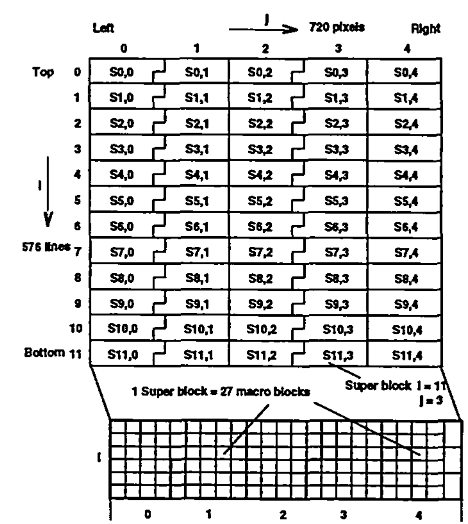


Figure 24 – Super blocks and macro blocks in one television frame for 625/50 system for 4:1:1 compression

5.1.4 Super block

Each super block consists of 27 macro blocks.

Super block arrangement in one frame for 525/60 system

The arrangement of super blocks in one frame is shown in figure 21 for 4:2:2 compression and figure 22 for 4:1:1 compression. Each super block consists of 27 adjacent macro blocks, and its boundary is marked by a heavy line. The total number of pixels in a frame is distributed into 100 super blocks for 4:2:2 compression or 50 super blocks for 4:1:1 compression.

4:2:2 compression
20 vertical super blocks × 5 horizontal super blocks = 100 super blocks

4:1:1 compression
10 vertical super blocks × 5 horizontal super blocks = 50 super blocks

Super block arrangement in one frame for 625/50 system

The arrangement of super blocks in one frame is shown in figure 23 for 4:2:2 compression and figure 24 for 4:1:1 compression. Each super block consists of 27 adjacent macro blocks, and its boundary is marked by a heavy line. The total number of pixels in a frame is distributed into 120 super blocks for 4:2:2 compression or 60 super blocks for 4:1:1 compression.

4:2:2 compression
24 vertical super blocks × 5 horizontal super blocks = 120 super blocks

4:1:1 compression
12 vertical super blocks × 5 horizontal super blocks = 60 super blocks

5.1.5 Definition of a super block number, a macro block number and value of the pixel

Super block number

The super block number in a frame is expressed as S i, j as shown in figures 21, 22, 23, and 24.

S i, j where i: the vertical order of the super block
i = 0, ..., n-1
where
n: the number of vertical super blocks in a video frame
n = 10 x m for the 525/60 system
n = 12 x m for the 625/50 system
m: the compression type
m = 1 for 4:1:1 compression
m = 2 for 4:2:2 compression

j: the horizontal order of the super block
j = 0, ..., 4

Macro block number

The macro block number is expressed as M i, j, k. The symbol k is the macro block order in the super block as shown in figure 25 for 4:2:2 compression and figure 26 for 4:1:1 compression. The small rectangle in these figures shows a macro block and a number in the small rectangle indicates k.

M i, j, k where i, j: the super block order number
k: the macro block order in the super block
k = 0, ..., 26

Pixel location

Pixel location is expressed as P i, j, k, l (x, y). The pixel is indicated as the suffix of i, j, k, l (x, y). The symbol is the DCT block order in a macro block as shown in figures 19 and 20. The rectangle in the figure shows a DCT block, and a DCT number in the rectangle expresses l. Symbol x and y are the pixel coordinate in the DCT block as described in 5.1.2.

P i, j, k, l (x, y) where i, j, k: the macro block number
l: the DCT block order in the macro block
(x, y): the pixel coordinate in the DCT block
x = 0, ..., 7
y = 0, ..., 7

Super block S i,j (i = 0,...,n-1, j = 0,...,4)

0	5	6	11	12	17	18	23	24
1	4	7	10	13	16	19	22	25
2	3	8	9	14	15	20	21	26

Where n = 20 : 525/60 system
n = 24 : 625/50 system

Figure 25 – Macro block order in a super block for 4:2:2 compression

Super block S i, 0, S i, 2 (i = 0, ..., n-1)

0	11	12	23	24
1	10	13	22	25
2	9	14	21	26
3	8	15	20	
4	7	16	19	
5	6	17	18	

Super block S i, 1, S i, 3 (i = 0, ..., n-1)

	8	9	20	21
	7	10	19	22
	6	11	18	23
0	5	12	17	24
1	4	13	16	25
2	3	14	15	26

Super block S i, 4 (i = 0, ..., n-1)

0	11	12	23	24
1	10	13	22	
2	9	14	21	25
3	8	15	20	
4	7	16	19	26
5	6	17	18	

Where n = 10: 525/60 system
n = 12: 625/50 system

Figure 26 – Macro block order in a super block for 4:1:1 compression

5.1.6 Definition of video segment and compressed macro block

A video segment consists of five macro blocks assembled from various areas within the video frame:

Ma, 2, k where a = (i + 2m) mod n
Mb, 1, k where b = (i + 6m) mod n
Mc, 3, k where c = (i + 8m) mod n
Md, 0, k where d = (i + 0) mod n
Me, 4, k where e = (i + 4m) mod n

where i: the vertical order of the super block
i = 0, ..., n-1
n: the number of vertical super blocks in a video frame
n = 10 x m for the 525/60 system
n = 12 x m for the 625/50 system
m: the compression type
m = 1 for 4:1:1 compression
m = 2 for 4:2:2 compression
k: the macro block order in the super block
k = 0, ..., 28

Each video segment before the bit-rate reduction is expressed as V i, k which consists of Ma, 2, k; Mb, 1, k; Mc, 3, k; Md, 0, k; and Me, 4, k.

The bit-rate reduction process is operated sequentially from Ma, 2, k to Me, 4, k. The data in a video segment are compressed and transformed to a 385-byte data stream. A compressed video data consists of five compressed macro blocks. Each compressed macro block consists of 77 bytes and is expressed as CM. Each video segment after the bit-rate reduction is expressed as CV i, k which consists of CM a, 2, k; CM b, 1, k; CM c, 3, k; CM d, 0, k; and CM e, 4, k as shown below.

CMa, 2, k:

This block includes all parts or most parts of the compressed data from macro block Ma, 2, k and may include the compressed data of macro block Mb, 1. k; or Mc, 3, k; or Md, 0, k; or Me, 4, k.

CMb, 1, k:

This block includes all parts or most parts of the compressed data from macro block Mb, 1, k and may include the compressed data of macro block Ma, 2, k; or Mc, 3, k; or Md, 0, k; or Me, 4, k.

CMc, 3, K

This block includes all parts or most parts of the compressed data from macro block Mc, 3, k and may include the compressed data of macro block Ma, 2, k; or Mb, 1, k; or Md, 0, k; or Me, 4, k.

CMd, 0, k

This block includes all parts or most parts of the compressed data from macro block Md, 0, k and may include the compressed data of macro block Ma, 2, k; or Mb, 1, k; or Mc, 3, k; or Me, 4, k.

CMe, 4, k:

This block includes all parts or most parts of the compressed data from macro block Me, 4, k and may include the compressed data of macro block Ma, 2, k; or Mb, 1, k; or Mc, 3, k; or Md, 0, k.

5.2 DCT processing

DCT blocks are comprised of two fields; each field providing pixels from 4 vertical lines and 8 horizontal pixels. In this clause, the DCT transformation from 64 pixels in a DCT block whose numbers are i, j, k, l (x, y) to 64 coefficients whose numbers are i, j, k, l (h, v) is described. P i, j, k, l (x, y) is the value of the pixel and C i, j, k, l (h, v) is the value of the coefficient.

For h = 0 and v = 0, the coefficient is called DC coefficient. Other coefficients are called AC coefficients.

5.2.1 DCT mode

Two modes, 8-8-DCT, and 2-4-8-DCT, are selectively used to optimize the data-reduction process, depending upon the degree of content variations between the two fields of a video frame. The two DCT modes are defined:

8-8-DCT mode

DCT

$$C_{i,j,k,l}(h, v) = C(v)\,C(h) \sum_{y=0}^{7} \sum_{x=0}^{7} (P_{i,j,k,l}(x, y) \cos(\pi v(2y+1)/16) \cos(\pi h(2x+1)/16))$$

Inverse DCT:

$$P_{i,j,k,l}(x, y) = \sum_{v=0}^{7} \sum_{h=0}^{7} (C(v)\,C(h)\; C_{i,j,k,l}(h, v) \cos(\pi v(2y+1)/16) \cos(\pi h(2x+1)/16))$$

where
C(h) = 0, 5 / √2 for h = 0
C(h) = 0, 5 for h = 1 to 7
C(v) = 0, 5 / √2 for v = 0
C(v) = 0, 5 for v = 1 to 7

2-4-8 DCT mode

DCT

$$C_{i,j,k,l}(h, u) = C(u)\,C(h) \sum_{z=0}^{3} \sum_{x=0}^{7} ((P_{i,j,k,l}(x, 2z) + P_{i,j,k,l}(x, 2z+1))\,KC)$$

$$C_{i,j,k,l}(h, u+4) = C(u)\,C(h) \sum_{z=0}^{3} \sum_{x=0}^{7} ((P_{i,j,k,l}(x, 2z) - P_{i,j,k,l}(x, 2z+1))\,KC)$$

Inverse DCT:

$$P_{i,j,k,l}(x, 2z) = \sum_{u=0}^{3} \sum_{h=0}^{7} (C(u)\,C(h)\;(C_{i,j,k,l}(h, u) + C_{i,j,k,l}(h, u+4))\,KC)$$

$$P_{i,j,k,l}(x, 2z+1) = \sum_{u=0}^{3} \sum_{h=0}^{7} (C(u)\,C(h)\;(C_{i,j,k,l}(h, u) - C_{i,j,k,l}(h, u+4))\,KC)$$

where
u = 0, ..., 3
z = INT (y / 2)
KC = COS (πu(2z + 1)/8) COS (πh(2x + 1)/16)
C(h) = 0, 5 / √2 for h = 0
C(h) = 0, 5 for h = 1 to 7
C(u) = 0, 5 / √2 for u = 0
C(u) = 0, 5 for u = 1 to 7

5.2.2 Weighting

DCT coefficients shall be weighted by the process as described below. W(h, v) expresses weight for C i, j, k, l (h, v) of the DCT coefficient.

8-8-DCT mode

For h = 0 and v = 0 W(h, v) = 1 / 4
For others W(h, v) = w(h) w(v) / 2

2-4-8- DCT mode

For h = 0 and v = 0 W(h, v) = 1 / 4
For v < 4 W(h, v) = w(h) w(2 v) / 2
For others W(h, v) = w(h) w(2 (v - 4)) / 2
where
w(0) = 1
w(1) = CS4 / (4 × CS7 × CS2)
w(2) = CS4 / (2 × CS6)
w(3) = 1 / (2 × CS5)
w(4) = 7 / 8
w(5) = CS4 / CS3
w(6) = CS4 / CS2
w(7) = CS4 / CS1
where CSM = COS (mπ / 16) m = 1 to 7

5.2.3 Output order

Figure 27 shows the output order of the weighted coefficients.

8-8-DCT (horizontal →, Vertical ↓)

	0	1	2	3	4	5	6	7
0	1	2	6	7	15	16	28	29
1	3	5	8	14	17	27	30	43
2	4	9	13	18	26	31	42	44
3	10	12	19	25	32	41	45	54
4	11	20	24	33	40	46	53	55
5	21	23	34	39	47	52	56	61
6	22	35	38	48	51	57	60	62
7	36	37	49	50	58	59	63	64

2-4-8-DCT (horizontal →, Vertical ↓)

		0	1	2	3	4	5	6	7
(sum)	0	1	3	7	19	21	35	37	51
	1	5	9	17	23	33	39	49	53
	2	11	15	25	31	41	47	55	61
	3	13	27	29	43	45	57	59	63
(difference)	4	2	4	8	20	22	36	38	52
	5	6	10	18	24	34	40	50	54
	6	12	16	26	32	42	48	56	62
	7	14	28	30	44	46	58	60	64

Figure 27 – Output order of weighted DCT block

5.2.4 Tolerance of DCT with weighting

Output error between the reference DCT and the tested DCT should satisfy the tolerances of the following cases:

Probability of occurrence of error;
Mean square errors for all coefficients;
Maximum value of mean square error for each DCT block;
All input pixel values of a DCT block are the same.

5.3 Quantization

5.3.1 Introduction

Weighted DCT coefficients are first quantized to 9-bit words, then divided by quantization in order to limit the amount of data in one video segment to five compressed macro blocks.

5.3.2 Bit assignment for quantization

Weighted DCT coefficients are represented as follows:

DC coefficient value (9 bits):
b8 b7 b6 b5 b4 b3 b2 b1 b0
twos complement (–255 to 255)

AC coefficient value (10 bits):
s b8 b7 b6 b5 b4 b3 b2 b1 b0
1 sign bit + 9 bits of absolute value (–511 to 511)

5.3.3 Class number

Each DCT block shall be classified into four classes by the definitions as described in table 21. For selecting the quantization step, the class number is used. Both c1 and c0 express the class number and are stored in the DC coefficient of compressed DCT blocks as described in 5.5. For reference, table 22 shows an example of the classification.

5.3.4 Initial scaling

Initial scaling is an operation for AC coefficients to transform from 10 bits to 9 bits. Initial scaling shall be done as follows:

For class number = 0, 1, 2
input data s b8 b7 b6 b5 b4 b3 b2 b1 b0
output data s b7 b6 b5 b4 b3 b2 b1 b0

For class number = 3
input data s b8 b7 b6 b5 b4 b3 b2 b1 b0
output data s b8 b7 b6 b5 b4 b3 b2 b1

Table 21 – Class number and the DCT block

Class number			DCT block	
	c1	c0	Quantization noises	Maximum absolute value of AC coefficient
0	0	0	Visible	Less than or equal to 255
1	0	1	Lower than class 0	
2	1	0	Lower than class 1	
3	1	1	Lower than class 2	
			———	Greater than 255

Table 22 – An example of the classification for reference

	Maximum absolute value of AC coefficient			
	0 to 11	12 to 23	24 to 35	> 35
Y	0	1	2	3
C_R	1	2	3	3
C_B	2	3	3	3

5.3.5 Area number

An area number is used for selection of the quantization step. AC coefficients within a DCT block shall be classified into four areas with area number as shown in figure 28.

5.3.6 Quantization step

The quantization step shall be decided by the class number, area number, and quantization number (QNO) as specified in table 23. QNO is selected in order to limit the amount of data in one video segment to five compressed macro blocks.

5.4 Variable length coding (VLC)

Variable length coding is an operation for transforming from quantized AC coefficients to variable length codes. One or more successive AC coefficients within a DCT block are coded into one variable length code according to the order as shown in figure 27. Run length and amplitude are defined as follows:

Run length: The number of successive AC coefficients quantized to 0 (run = 0, ..., 61)

Amplitude: Absolute value just after successive AC coefficients quantized to 0 (amp = 0, ..., 255)

(run, amp): The pair of run length and amplitude

Table 24 shows the length of code words corresponding to (run, amp). In the table, sign bit is not included in the length of code words. When the amplitude is not zero, the code length shall be plus 1 because sign bit is needed. For empty columns, the length of code words of the (run, amp) equals that of the (run - 1, 0) plus that of the (0, amp).

Variable length code shall be as shown in table 25. The leftmost bit of code words is MSB and the rightmost bit of code words is LSB in table 25. The MSB of a subsequent code word is next to the LSB of the code word just before. Sign bit "s" shall be as follows:

When the quantized AC coefficient is greater than zero, s = 0

When the quantized AC coefficient is less than zero, s = 1

When the values of all of the remaining quantized coefficients are zero within a DCT block, the coding process is ended by adding EOB (end of block) code word of 0110b to just after the last code word.

8-8-DCT

horizontal →, Vertical ↓

	0	1	2	3	4	5	6	7
0	DC	0	0	1	1	1	2	2
1	0	0	1	1	1	2	2	2
2	0	1	1	1	2	2	2	3
3	1	1	1	2	2	2	3	3
4	1	1	2	2	2	3	3	3
5	1	2	2	2	3	3	3	3
6	2	2	2	3	3	3	3	3
7	2	2	3	3	3	3	3	3

2-4-8-DCT

horizontal →, Vertical ↓ (sum: rows 0–3; difference: rows 4–7)

	0	1	2	3	4	5	6	7
0	DC	0	1	1	1	2	2	3
1	0	1	1	2	2	2	3	3
2	1	1	2	2	2	3	3	3
3	1	2	2	2	3	3	3	3
4	0	0	1	1	2	2	2	3
5	0	1	1	2	2	2	3	3
6	1	1	2	2	2	3	3	3
7	1	2	2	3	3	3	3	3

Figure 28 – Area numbers

Table 23 – Quantization step

	Class number				Area number			
	0	1	2	3	0	1	2	3
Quantization number (QNO)	15				1	1	1	1
	14				1	1	1	1
	13				1	1	1	1
	12	15			1	1	1	1
	11	14			1	1	1	1
	10	13		15	1	1	1	1
	9	12	15	14	1	1	1	1
	8	11	14	13	1	1	1	2
	7	10	13	12	1	1	2	2
	6	9	12	11	1	1	2	2
	5	8	11	10	1	2	2	4
	4	7	10	9	1	2	2	4
	3	6	9	8	2	2	4	4
	2	5	8	7	2	2	4	4
	1	4	7	6	2	4	4	8
	0	3	6	5	2	4	4	8
		2	5	4	4	4	8	8
		1	4	3	4	4	8	8
		0	3	2	4	8	8	16
			2	1	4	8	8	16
			1	0	8	8	16	16
			0		8	8	16	16

Table 24 – Length of codewords

	Amplitude																									
Run length	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	—	255
0	11	2	3	4	4	5	5	6	6	7	7	7	8	8	8	8	8	8	9	9	9	9	9	15	—	15
1	11	4	5	7	7	8	8	8	9	10	10	10	11	11	11	12	12	12								
2	12	5	7	8	9	9	10	12	12	12	12	12														
3	12	6	8	9	10	10	11	12																		
4	12	6	8	9	11	12																				
5	12	7	9	10																						
6	13	7	9	11																						
7	13	8	12	12																						
8	13	8	12	12																						
9	13	8	12																							
10	13	8	12																							
11	13	9																								
12	13	9																								
13	13	9																								
14	13	9																								
15	13																									
|	|																									
61	13																									

NOTES
1 Sign bit is not included.
2 The length of EOB = 4.

Table 25 – Codewords of variable-length coding

(Run, amp)		Code	Length
0	1	00s	2+1
0	2	010s	3+1
EOB		0110	4
1	1	0111s	4+1
0	3	1000s	
0	4	1001s	
2	1	10100s	5+1
1	2	10101s	
0	5	10110s	
0	6	10111s	
3	1	110000s	6+1
4	1	110001s	
0	7	110010s	
0	8	110011s	
5	1	1101000s	7+1
6	1	1101001s	
2	2	1101010s	
1	3	1101011s	
1	4	1101100s	
0	9	1101101s	
0	10	1101110s	
0	11	1101111s	
7	1	11100000s	8+1
8	1	11100001s	
9	1	11100010s	
10	1	11100011s	
3	2	11100100s	
4	2	11100101s	
2	3	11100110s	
1	5	11100111s	
1	6	11101000s	
1	7	11101001s	
0	12	11101010s	
0	13	11101011s	
0	14	11101100s	
0	15	11101101s	
0	16	11101110s	
0	17	11101111s	

(Run, amp)		Code	Length
11	1	111100000s	9+1
12	1	111100001s	
13	1	111100010s	
14	1	111100011s	
5	2	111100100s	
6	2	111100101s	
3	3	111100110s	
4	3	111100111s	
2	4	111101000s	
2	5	111101001s	
1	8	111101010s	
0	18	111101011s	
0	19	111101100s	
0	20	111101101s	
0	21	111101110s	
0	22	111101111s	
5	3	1111100000s	10+1
3	4	1111100001s	
3	5	1111100010s	
2	6	1111100011s	
1	9	1111100100s	
1	10	1111100101s	
1	11	1111100110s	
0	0	11111001110	11
1	0	11111001111	
6	3	11111010000s	11+1
4	4	11111010001s	
3	6	11111010010s	
1	12	11111010011s	
1	13	11111010100s	
1	14	11111010101s	
2	0	111110101100	12
3	0	111110101101	
4	0	111110101110	
5	0	111110101111	

(Run, amp)		Code	Length
7	2	111110110000s	12+1
8	2	111110110001s	
9	2	111110110010s	
10	2	111110110011s	
7	3	111110110100s	
8	3	111110110101s	
4	5	111110110110s	
3	7	111110110111s	
2	7	111110111000s	
2	8	111110111001s	
2	9	111110111010s	
2	10	111110111011s	
2	11	111110111100s	
1	15	111110111101s	
1	16	111110111110s	
1	17	111110111111s	
6	0	1111110000110	13
7	0	1111110000111	
R	0	1111110 Binary notation of R R = 6 to 61	
61	0	1111110111101	
0	23	111111100010111s	15+1
0	24	111111100011000s	
0	A	1111111 Binary notation of A A = 23 to 255 s	
0	255	111111111111111s	

NOTES
1 (R, 0): 1111110 r5 r4 r3 r2 r1 r0, where $32r5 + 16r4 + 8r3 + 4r2 + 2r1 + r0 = R$.
2 (0, A): 1111111 a7 a6 a5 a4 a3 a2 a1 a0 s, where $128a7 + 64a6 + 32a5 + 16a4 + 8a3 + 4a2 + 2a1 + a0 = A$.
3 S is sign bit. EOB means end of block.

5.5 Arrangement of a compressed macro block

A compressed video segment consists of five compressed macro blocks. Each compressed macro block has 77 bytes of data. The arrangement of the compressed macro block shall be as shown in figure 29 for 4:2:2 compression and in figure 30 for 4:1:1 compression. Each compressed macro block of 4:2:2 compression includes a two-byte data area (X0, X1). The data arrangement is shown in figure 29. The data format of the reserved area is not defined except 100000000000.

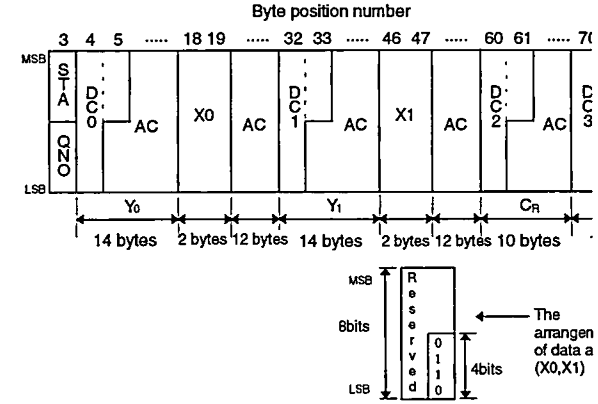
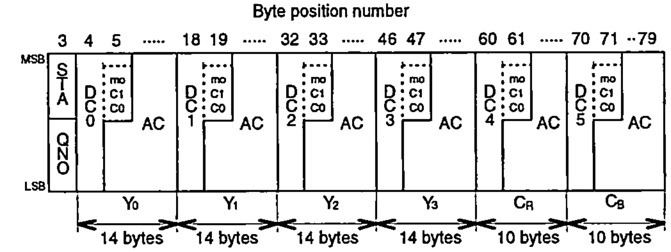


Figure 29 – Arrangement of a compressed macro block for 4:2:2 compression



STA: Error status
QNO: Quantization number
DC: DC component
AC: AC component
EOB: End of block (0110)
mo: DCT mode
c0, c1: Class number

Figure 30 – Arrangement of a compressed macro block for 4:1:1 compression

STA (status of the compressed macro block)

STA expresses the error and concealment of the compressed macro block and consists of four bits: s3, s2, s1, s0. Table 26 shows the definitions of STA.

QNO (quantizaton number)

QNO is the quantization number applied to the macro block. Code words of the QNO shall be as shown in table 27.

Table 26 – Definition of STA

STA				Information of the compressed macro block		
s3	s2	s1	s0	Error	Error concealment	Continuity
0	0	0	0	No error	Not applied	—
0	0	1	0	No error	Type A	Type a
0	1	0	0	No error	Type B	Type a
0	1	1	0	No error	Type C	Type a
0	1	1	1	Error exists	—	—
1	0	1	0	No error	Type A	Type b
1	1	0	0	No error	Type B	Type b
1	1	1	0	No error	Type C	Type b
1	1	1	1	Error exists	—	—
Others				Reserved		

where

Type A: Replaced with a compressed macro block of the same compressed macro block number in the immediate previous frame.

Type B: Replaced with a compressed macro block of the same compressed macro block number in the next immediate frame.

Type C: This compressed macro block is concealed, but the concealment method is not specified.

Type a: The continuity of the data processing sequence with other compressed macro blocks whose s0 = 0 and s3 = 0 in the same video segment is guaranteed.

Type b: The continuity of the data processing sequence with other compressed macro blocks is not guaranteed.

NOTES

1 For STA = 0111b, the error code is inserted in the compressed macro block. This is an option.

2 For STA = 1111b, the error position is unidentified.

Table 27 – Codewords of the QNO

q3	q2	q1	q0	QNO
0	0	0	0	0
0	0	0	1	1
0	0	1	0	2
0	0	1	1	3
0	1	0	0	4
0	1	0	1	5
0	1	1	0	6
0	1	1	1	7
1	0	0	0	8
1	0	0	1	9
1	0	1	0	10
1	0	1	1	11
1	1	0	0	12
1	1	0	1	13
1	1	1	0	14
1	1	1	1	15

DC

DCI (where l is the DCT block order in the macro block, l = 0, ..., 3 for 4:2:2 compression, l = 0, ..., 5 for 4:1:1 compression) consists of a DC coefficient, the DCT mode, and the class number of the DCT block.

MSB LSB
DCI: b8 b7 b6 b5 b4 b3 b2 b1 b0 mo c1 c0

where
b8 to b0: DC coefficient value
mo: DCT mode mo = 0 for 8-8-DCT mode
mo = 1 for 2-4-8-DCT mode
c1 c0: class number

AC

AC is a generic term for variable length coded AC coefficients within the video segment V i, k. For 4:2:2 compression, the areas of Y0, Y1, CR, and CB are defined as compressed-data areas and each of Y0 and Y1 consists of 112 bits and each CR and CB consists of 80 bits as shown in figure 29. For 4:1:1 compression, the areas of Y0, Y1, Y2, Y3, CR, and CB are defined as compressed-data areas and each of Y0, Y1, Y2, and Y3 consists of 112 bits and each CR and CB consists of 80 bits as shown in figure 30. DCI and variable length code for AC coefficients in the DCT block whose DCT block number is i, j, k, l are assigned from the beginning of the compressed-data area in the compressed macro block CM i, j, k. In figures 29 and 30, the variable length code word is located starting from the MSB which is shown in the upper left side, and the LSB shown in the lower right side. Therefore, AC data are distributed from the upper left corner to the lower right corner.

5.6 Arrangement of a video segment

In this clause, the distribution method of quantized AC coefficients is described. Figures 31 and 32 show the arrangement of a video segment CV i, k after bit-rate reduction. Each row contains a compressed macro block. Columns F i, j, k, l express the compressed data area for DCT blocks whose DCT block numbers are i, j, k, l. Symbol E i, j, k, l expresses an additional AC area for recording remaining data from the fixed AC area.

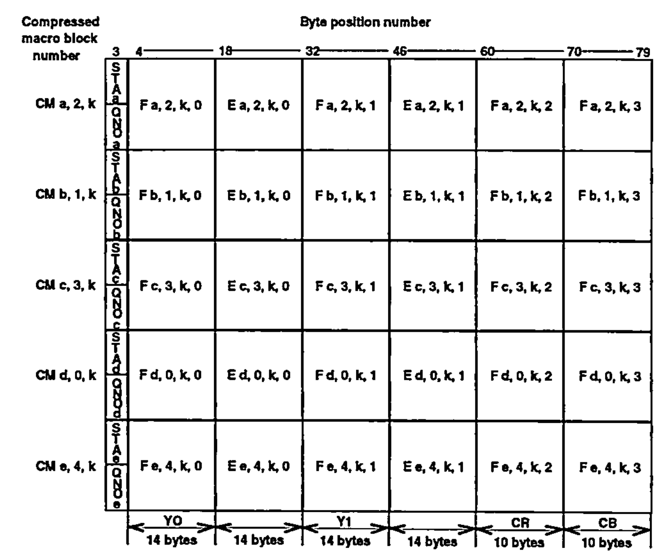


Figure 31 – Arrangement of a video segment after the bit-rate reduction for 4:2:2 compression

Compressed macro block number	3	4–17	18–31	32–45	46–59	60–69	70–79
CM a, 2, k	STA a QNO a	F a, 2, k, 0	F a, 2, k, 1	F a, 2, k, 2	F a, 2, k, 3	F a, 2, k, 4	F a, 2, k, 5
CM b, 1, k	STA b QNO b	F b, 1, k, 0	F b, 1, k, 1	F b, 1, k, 2	F b, 1, k, 3	F b, 1, k, 4	F b, 1, k, 5
CM c, 3, k	STA c QNO c	F c, 3, k, 0	F c, 3, k, 1	F c, 3, k, 2	F c, 3, k, 3	F c, 3, k, 4	F c, 3, k, 5
CM d, 0, k	STA d QNO d	F d, 0, k, 0	F d, 0, k, 1	F d, 0, k, 2	F d, 0, k, 3	F d, 0, k, 4	F d, 0, k, 5
CM e, 4, k	STA e QNO e	F e, 4, k, 0	F e, 4, k, 1	F e, 4, k, 2	F e, 4, k, 3	F e, 4, k, 4	F e, 4, k, 5
		YO 14 bytes	Y1 14 bytes	Y2 14 bytes	Y3 14 bytes	CR 10 bytes	CB 10 bytes

Byte position number: 3, 4, 18, 32, 46, 60, 70, 79

where a = (i + 2) mod n
b = (i + 6) mod n
c = (i + 8) mod n
d = (i + 0) mod n
e = (i + 4) mod n

i: the vertical order of the super block
i = 0, ..., n-1
n: the number of vertical super blocks in a video frame
n = 10 for 525/60 system
n = 12 for 625/50 system
k: the macro block order in the super block
k = 0, ..., 26

Figure 32 – Arrangement of a video segment after the bit-rate reduction for 4:1:1 compression

Bit sequence, defined as Bi, j, k, l, shall consist of the following concatenated data: DC coefficient, DCT mode information, class number, and AC coefficient code words for DCT blocks numbered i, j, k, l. Code words for AC coefficients of B i, j, k, l shall be concatenated according to the order as shown in figure 27 and the last code word shall be EOB. The MSB of the subsequent code word shall be next to the LSB of the code word just before.

The arrangement algorithm of a video segment shall be composed of three passes:

Pass 1: The distribution of B i, j, k, l to the compressed-data area;

Pass 2: The distribution of the overflow B i, j, k, l which are the remainder after the pass 1 operation in the same compressed macro block;

Pass 3: The distribution of the overflow B i, j, k, l which are the remainder after the pass 2 operation in the same video segment.

Arrangement algorithm of a video segment

```
4:2:2 compression
 if (525/60 system) n = 20 else n = 24;
 for (i = 0; i < n; i++){
   a = (i + 4) mod n;
   b = (i + 12) mod n;
```



```
   c = (i + 16) mod n;
   d = (i + 0) mod n;
   e = (i + 8) mod n;
   for (k = 0; k < 27; k++){
     q = 2;
     p = a;
     VR = 0
     /* VR is the bit sequence for the data          */
     /* which are not distributed to video segment
        CV i, k by pass 2.                          */
 /* pass 1 */
     for (j = 0; j < 5; j++) {
       MRq = 0;
       /* MRq is the bit sequence for the data   */
       /* which are not distributed to macro block
          M i, q, k by pass 1.                   */
       for (l = 0, l < 4; l ++) {
         remain = distribute (B p, q, k, l, F p, q, k, l);
         MRq = connect (MRq, remain);
       }
            if (q == 2) {q = 1; p = b;}
       else if (q == 1) {q = 3; p = c;}
       else if (q == 3) {q = 0; p = d;}
       else if (q == 0) {q = 4; p = e;}
       else if (q == 4) {q = 2; p = a;}
     }
 /* pass 2 */
     for (j = 0; j < 5; j++) {
       for (l = 0; l < 4; l ++) {
         MRq = distribute (MRq, F p, q, k, l);
         if ((l == 0) || (l == 1))
           MRq = distribute (MRq, E p, q, k, l);
       }
       VR = connect (VR, MRq);
            if (q == 2) {q = 1; p = b;}
       else if (q == 1) {q = 3; p = c;}
       else if (q == 3) {q = 0; p = d;}
       else if (q == 0) {q = 4; p = e;}
       else if (q == 4) {q = 2; p = a;}
     }
 /* pass 3 */
     for (j = 0; j < 5; j++) {
       for (l = 0; l < 4; l ++) {
         VR = distribute (VR, F p, q, k, l);
         if ((l == 0) || (l == 1))
           VR = distribute (VR, E p, q, k, l);
       }
            if (q == 2) {q = 1; p = b;}
       else if (q == 1) {q = 3; p = c;}
       else if (q == 3) {q = 0; p = d;}
       else if (q == 0) {q = 4; p = e;}
       else if (q == 4) {q = 2; p = a;}
     }
   }
 }

4:1:1 compression
 if (525/60 system) n = 10 else n = 12;
 for (i = 0; i < n ; i++){
   a = (i + 2) mod n;
   b = (i + 6) mod n;
   c = (i + 8) mod n;
   d = (i + 0) mod n;
   e = (i + 4) mod n;
   for (k = 0; k < 27; k++){
     q = 2;
     p = a;
     VR = 0
     /* VR is the bit sequence for the data          */
     /* which are not distributed to video segment
        CV i, k by pass 2                           */
 /* pass 1 */
     for (j = 0; j <5; j++) {
       MRq = 0;
       /* MRq is the bit sequence for the data   */
       /* which are not distributed to macro block
          M i, q, k by pass 1.                   */
       for (l = 0, l < 6; l ++) {
         remain = distribute (B p, q, k, l, F p, q, k, l);
         MRq = connect (MRq, remain);
       }
            if (q == 2) {q = 1; p = b;}
       else if (q == 1) {q = 3; p = c;}
       else if (q == 3) {q = 0; p = d;}
       else if (q == 0) {q = 4; p = e;}
       else if (q == 4) {q = 2; p = a;}
     }
 /* pass 2 */
     for (j = 0; j < 5; j++) {
       for (l = 0; l < 6; l ++) {
         MRq = distribute (MRq, F p, q, k, l);
       }
       VR = connect (VR, MRq);
            if (q == 2) {q = 1; p = b;}
       else if (q == 1) {q = 3; p = c;}
       else if (q == 3) {q = 0; p = d;}
       else if (q == 0) {q = 4; p = e;}
       else if (q == 4) {q = 2; p = a;}
     }
 /* pass 3 */
     for (j = 0; j < 5; j++){
       for (l = 0; l < 6; l ++) {
         VR = distribute (VR, F p, q, k, l);
       }
```

```
            if (q == 2) {q = 1; p = b;}
        else if (q == 1) {q = 3; p = c;}
        else if (q == 3) {q = 0; p = d;}
        else if (q == 0) {q = 4; p = e;}
        else if (q == 4) {q = 2; p = a;}
      }
    }
  }

where
  distribute (data 0, area 0) {
        /* Distribute data 0 from MSB into empty
           area of area 0.                          */
        /* The area 0 is filled starting from the MSB. */
    remain = (remaining_data);
        /* Remaining_data are the data which are not
           distributed.                             */
    return (remain);
  }

  connect (data 1, data 2) {
      /* Connect the MSB of data 2 with the LSB of
         data 1.                                    */
    data 3 = (connecting_data)
      /* Connecting_data are the data which are
         connected.                                 */
      /* data 2 with data 1.                        */
    return (data 3);
  }
```

The remaining data which cannot be distributed within the unused space of the macro block will be ignored. Therefore, when error concealment is performed for a compressed macro block, some distributed data by pass 3 may not be reproduced.

Video error code processing

If errors are detected in a compressed macro block which is reproduced and processed with error correction, the compressed-data area containing these errors should be replaced with the video error code. This process replaces the first two bytes of data of the compressed-data area with the code as follows:

MSB LSB
1000000000000110b

The first 9 bits are the DC error code, the next 3 bits are the information of the DCT mode and class number, and the last 4 bits are the EOB as shown in figure 33.

When the compressed macro blocks, after error code processing, are input to the decoder which does not operate with video error code, all data in this compressed macro block should be processed as invalid.

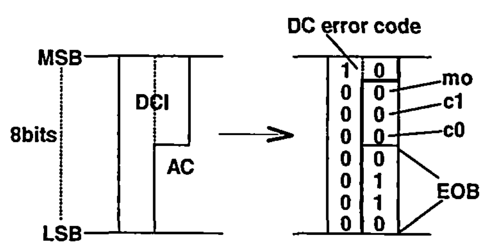


Figure 33 – Video error code

Annex A (informative)
Differences between IEC 61834 and SMPTE 314M

The differences between IEC 61834 and SMPTE 314M are shown in table A.1.

Table A.1 – Abstract of differences between IEC 61834 and SMPTE 314M

		DV IEC 61834	DV-BASED SMPTE 314M	
			25 Mb/s structure	50 Mb/s structure
Data structure		IEC 61834	Same as IEC 61834	See figure 2
Header	Bit name APT AP1 AP2 AP3	 000 000 000 000	 001 001 001 001	
Video	Sampling structure	525: 4:1:1 625: 4:2:0	525: 4:1:1 625: 4:1:1	525: 4:2:2 625: 4:2:2
VAUX	VS	IEC 61834	See 4.5.2.1	
	VSC	IEC 61834	See 4.5.2.2	
	Other	IEC 61834	Reserved	
Audio	Sampling	48 kHz (16 bits, 2ch) 44.1 kHz (16 bits, 2ch) 32 kHz (16 bits, 2ch) 32 kHz (12 bits, 4ch)	48 kHz (16 bits, 2ch)	48 kHz (16 bits, 4ch)
	Locked mode	Locked/unlocked	Locked	Locked
AAUX	AS	IEC 61834	See 4.6.2.3.1	
	ASC	IEC 61834	See 4.6.2.3.2	
	Other	IEC 61834	Reserved	
Subcode	SSYB ID	IEC 61834	See 4.4.2.1	
	TC	IEC 61834	See 4.4.2.2.1	
	BG	IEC 61834	Same as IEC 61834	
	Other	IEC 61834	Reserved	

Annex B (informative)
Bibliography

ANSI/IEEE 1394-1995, High Performance Serial Bus

ANSI/SMPTE 12M-1995, Television, Audio and Film — Time and Control Code

ANSI/SMPTE 125M-1995, Television — Component Video Signal 4:2:2 — Bit-Parallel Digital Interface

SMPTE 306M-1998, Television Digital Recording — 6.35-mm Type D-7 Component Format — Video Compression at 25 Mb/s — 525/60 and 625/50

IEC 60958 (1982-02), Digital Audio Interface

ITU-R BT.470-4 (1994), Television Systems

ITU-R BT.601-5 (1994), Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios

PROPOSED SMPTE STANDARD

SMPTE 315M

for Television — Camera Positioning Information Conveyed by Ancillary Data Packets

1 Scope

This standard provides a method for multiplexing camera positioning information into the ancillary data space described in SMPTE 291M. Applications of the standard include the 525-line, 625-line, component or composite, and high-definition digital television interfaces which provide 10-bit ancillary data space. Two types of camera positioning information are defined in this standard: binary and ASCII.

2 Normative references

The following standard contains provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the edition indicated was valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standard indicated below.

SMPTE 291M-1998, Television — Ancillary Data Packet and Space Formatting

3 Overview

3.1 Camera positioning information, as defined in this standard, contains the following parameters. These parameters, and other parameters described in the note, are carried selectively upon the need:

- camera relative position
- camera pan
- camera tilt
- camera roll
- origin of world coordinate longitude
- origin of world coordinate latitude
- origin of world coordinate altitude
- vertical angle of view
- focus distance
- lens opening (iris or f-value)
- time address informaftion
- object relative position

NOTE – There may be other parameters, required for other applications, for which data are carried within this type of ancillary data signal, but which are not defined in this standard. Such parameters will have parameter identifications which do not conflict with those used in this standard. When such parameters are received by equipmemt that does not recognize them, they should be ignored.

3.2 Data for each parameter can be obtained from several kinds of pick-up devices, such as rotary encoders. These data are formatted as an ancillary data packet and multiplexed into the ancillary data space of serial digital video and conveyed to the receiving end. This specification defines the packet structure, word structure, coordinate, range, and accuracy of each parameter, and the method of multiplexing packets.

4 Packet structure

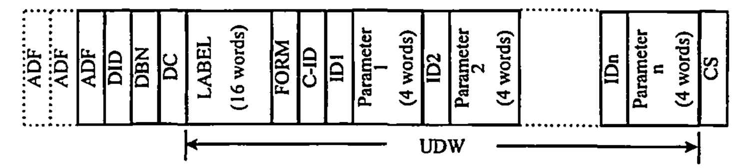
4.1 Camera positioning data packet

4.1.1 Camera positioning data packets are formatted in a data stream into one of two types, binary or ASCII. These are identified by the data type identification flag word (FORM) in the user data words and are described in 4.1.2 and 4.1.3.

4.1.2 The structure of a binary-type camera positioning data packet is shown in figure 1.

4.13 The structure of an ASCII-type camera positioning data packet is shown in figure 2.

THIS PROPOSAL IS PUBLISHED FOR COMMENT ONLY

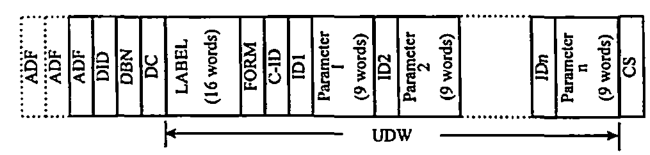


ADF: Ancillary data flag
DID: Data identification word $2F0_h$ (see note)
DBN: Data block number word
DC: Data count word
UDW: User data words (up to 255 words)
- LABEL: SMPTE label for metadata of class camera positioning information (16 words)
- FORM: Data type identification flag word (1 word)
- *C-ID:* *Camera identification word (1 word)*
- IDn: Parameter identification word (1 word for each parameter)
- Parameter n: Parameter data words (4 words for each parameter)

CS: Checksum word

NOTE – This DID indicates that the UDWs carry metadata. The specific class of metadata is indicated by the LABEL. It is necessary to decode both the DID and the LABEL to identify the UDWs as carrying camera positioning information.

Figure 1 – Binary-type camera positioning data packet



ADF: Ancillary data flag
DID: Data identification word $2F0_h$ (see note)
DBN: Data block number word
DC: Data count word
UDW: User data words (up to 255 words)
- LABEL: SMPTE label for metadata of class camera positioning information (16 words)
- FORM: Data type identification flag word (1 word)
- C-ID: Camera identification words (3 words)
- IDn: Parameter identification words (3 words for each parameter)
- Parameter n: Parameter data words (9 words for each parameter)

CS: Checksum word

NOTE – This DID indicates that the UDWs carry metadata. The specific class of metadata is indicated by the LABEL. It is necessary to decode both the DID and the LABEL to identify the UDWs as carrying camera positioning information.

Figure 2 – ASCII-type camera positioning data packet

5 User data words

5.1 SMPTE universal label (LABEL)

The 16 words UDW 0-15 carry the SMPTE-administered universal label to identify the class of metadata camera positioning information.

The label has the value:

{ iso org smpte (52)
registry(1) mpeg(1) rid(1) version1(1)
C(67) A(65) P(80) O(79)
}

and the word sequence is:

06 0E 2B 34 01 01 01 01 43 41 50 4F 00 00 00 00

The 4 words UDW 8-11 are the ISO/IEC 13818-1 (MPEG) registration identifier.

5.2 Data type identification flag word (FORM)

The data type identification flag word (FORM) indicates the data type of the camera identification word (C-ID), parameter identification word (IDn), and parameter data word (parameter n) contained in the packet. Bit assignments of the data type identification flag word and the data type code are shown in tables 1 and 2, respectively.

5.3. Camera identification word(s) (C-ID)

The camera identification word(s) indicate(s) the camera number by which the object is shot. It should be noted that the camera number may be any value from 0 through 255, inclusively.

5.3.1 Binary-type camera identification word

The bit assignment of the binary-type camera identification word is shown in table 3.

5.3.2 ASCII-type camera identification words

Three ASCII-coded characters are used to express the camera number (000 through 255). ASCII-type camera identification is mapped into three contiguous ancillary data words (X, X+1, X+2) as shown in table 4.

5.4 Parameter identification word (IDn)

Each parameter datum is preceded by a parameter identification word(s) which indicate(s) the content classification of the parameter.

5.4.1 Binary-type parameter identification word

The bit assignment of the binary-type parameter identification word is shown in table 5. Parameter identification codes are shown in table 6.

5.4.2 ASCII-type parameter identification words

Three ASCII-coded characters are used to express the parameter identification code (000 through 255). ASCII-type parameter identification is mapped into three contiguous ancillary data words (X, X+1, X+2) as shown in table 7. Parameter identification codes are shown in table 6.

Table 1 – Bit assignment of data type identification flag word (DID)

Bit address	Data type identification flag
b9 (MSB)	Not b8
b8	Even parity for b(0–7)
b7	AUX
b6	AUX
b5	AUX
b4	AUX
b3	AUX
b2	AUX
b1	FORM 1
b0 (LSB)	FORM 0

NOTE – FORM(0-1): Data type code
AUX: User defined

Table 2 – Data type code

FORM 1	FORM 0	Data type
0	0	Binary
0	1	ASCII
1	0	Reserved
1	1	Reserved

Table 3 – Bit assignment of binary-type camera identification word

Bit address	Binary-type camera identification
b9 (MSB)	Not b8
b8	Even parity for b(0–7)
b7	C7 (MSB)
b6	C6
b5	C5
b4	C4
b3	C3
b2	C2
b1	C1
b0 (LSB)	C0 (LSB)

NOTE – C(0-7): Camera number data

Table 4 – Bit assignment of ASCII-type camera identification word

	ASCII-type camera identification		
Bit address	X	X+1	X+2
b9 (MSB)	Not b8	Not b8	Not b8
b8	Even parity for b(0–7)	Even parity for b(0–7)	Even parity for b(0–7)
b7	C7 (MSB)	C15 (MSB)	C23 (MSB)
b6	C6	C14	C22
b5	C5	C13	C21
b4	C4	C12	C20
b3	C3	C11	C19
b2	C2	C10	C18
b1	C1	C9	C17
b0 (LSB)	C0 (LSB)	C8 (LSB)	C16 (LSB)

NOTES
C(0-23): Camera number data
C(0-7): The first character (hundreds) (0, 1, 2)
C(8-15): The second character (tens) (0 through 9, but maximum value is 5 when Word X is 2.)
C(16-23): The third character (units) (0 through 9, but maximum value is 5 when Word X is 2 and Word X+1 is 5.)

Table 5 – Bit assignment of binary-type parameter identification word

Bit address	Parameter identification
b9 (MSB)	Not b8
b8	Even parity for b(0–7)
b7	P7 (MSB)
b6	P6
b5	P5
b4	P4
b3	P3
b2	P2
b1	P1
b0 (LSB)	P0 (LSB)

NOTE – P(0-7): Parameter identification code.

Table 6 – Parameter identification code

Identification code				Classification of parameter
Binary type	ASCII type			
00_h	0	0	0	Reserved
01_h	0	0	1	Camera relative position Xc
02_h	0	0	2	Camera relative position Yc
03_h	0	0	3	Camera relative position Zc
04_h	0	0	4	Camera pan
05_h	0	0	5	Camera tilt
06_h	0	0	6	Camera roll
07_h	0	0	7	Origin of world coordinate longitude
08_h	0	0	8	Origin or world coordinate latitude
09_h	0	0	9	Origin of world coordinate altitude
$0A_h$	0	1	0	Vertical angle of view
$0B_h$	0	1	1	Focus distance
$0C_h$	0	1	2	Lens opening (iris or *f*-value)
$0D_h$ … $2F_h$	0 … 0	1 … 4	3 … 7	Reserved
30_h	0	4	8	Time address information
31_h … $4F_h$	0 … 0	4 … 7	9 … 9	Reserved
50_h	0	8	0	Object 1 relative position Xo
51_h	0	8	1	Object 1 relative position Yo
52_h	0	8	2	Object 1 relative position Zo
53_h … $7C_h$	0 … 1	8 … 2	3 … 4	Object 2 relative position Xo … Object 15 relative position Zo
$7D_h$	1	2	5	Object 16 relative position Xo
$7E_h$	1	2	6	Object 16 relative position Yo
$7F_h$	1	2	7	Object 16 relative position Zo
80_h … FF_h	1 … 2	2 … 5	8 … 5	User defined

Table 7 – Bit assignment of ASCII-type parameter identification words

	ASCII-type parameter identification		
Bit address	X	X+1	X+2
b9 (MSB)	Not b8	Not b8	Not b8
b8	Even parity for b(0–7)	Even parity for b(0–7)	Even parity for b(0–7)
b7	P7 (MSB)	P15 (MSB)	P23 (MSB)
b6	P6	P14	P22
b5	P5	P13	P21
b4	P4	P12	P20
b3	P3	P11	P19
b2	P2	P10	P18
b1	P1	P9	P17
b0 (LSB)	P0 (LSB)	P8 (LSB)	P16 (LSB)

NOTES
P(0-23): Parameter identification data
P(0-7): The first character (hundreds) (0, 1, 2)
P(8-15): The second character (tens) (0 through 9, but maximum value is 5 when Word X is 2.)
P(16-23): The third character (units) (0 through 9, but maximum value is 5 when Word X is 2 and Word X+1 is 5.)

5.5 Parameter data word (Parameter n)

There are two types of parameter data words — the binary type and the ASCI I type.

5.5.1 Binary-type parameter data word

5.1.1 Bit-assignment of binary-type parameter data word (except time address information)

The binary-type parameter data word (except time address information) is mapped into four contiguous ancillary data words (X, X+1, X+2, X+3) as shown in table 8. The data representation conforms to ANSI/IEEE-754, the binary floating-point arithmetic standard.

A 32-bit format for a binary floating-point number Y is divided as shown in figure 3. The component fields of Y are the 1-bit sign s, the 8-bit biased exponent e, and the 23-bit fraction f. Bit D31 corresponds to s, D(30-23) to e, and D(22-0) to f.

5.5.1.2 Bit-assignment for time address information of binary-type parameter data word

The bit-assignment for time address information of binary-type parameter data word is as shown in table 9.

5.5.2 ASCII-type parameter data word

5.5.2.1 Bit-assignment of ASCII-type parameter data word (except time address information)

The ASCII-type parameter data word (except time address information) is mapped into nine contiguous ancillary data words (X, X+1, X+2, ..., X+8) as shown in table 10. These words indicate the symbol (+ or –), integer part, decimal point, and fractional part. The decimal point can be omitted if not required. Table 11 shows the word number to be used for each part.

5.5.2.2 Bit-assignment for time address information of ASCII-type parameter data word

The bit-assignment for time address information of ASCII-type parameter data word is as shown in table 12. Table 13 shows the definition for Word X+6 which identifies field number and drop-frame mode.

Table 8 – Bit assignment of binary-type parameter data word (except time address information)

Bit address	Binary-type parameter data X	X+1	X+2	X+3
b9 (MSB)	Not b8	Not b8	Not b8	Not b8
b8	Even parity for b(0–7)	Even parity for b(0–7)	Even parity for b(0–7)	Even parity for b(0–7)
b7	D7	D15	D23	D31 (MSB)
b6	D6	D14	D22	D30
b5	D5	D13	D21	D29
b4	D4	D12	D20	D28
b3	D3	D11	D19	D27
b2	D2	D10	D18	D26
b1	D1	D9	D17	D25
b0 (LSB)	D0 (LSB)	D8	D16	D24

NOTE – D(0-31): Binary-type parameter data.

s	e	f
D31	D30 … D23	D22 … D0

Figure 3 – Format of binary floating point (single-precision format)

Table 9 – Bit assignment for time address information of binary-type parameter data word

Bit address	Binary-type parameter data X	X+1	X+2	X+3
b9 (MSB)	Not b8	Not b8	Not b8	Not b8
b8	Even parity for b0-b7	Even parity for b0-b7	Even parity for b0–b7	Even parity for b0-b7
b7	Color frame flag (15)*	Field flag (35)	Reserved (set to zero)	Reserved (set to zero)
b6	Drop frame flag (14)*	Tens of seconds (34)	Tens of minutes (54)	Reserved (set to zero)
b5	Tens of frames (13)	Tens of seconds (33)	Tens of minutes (53)	Tens of hours (73)
b4	Tens of frames (12)	Tens of seconds (32)	Tens of minutes (52)	Tens of hours (72)
b3	Units of frames (5)	Units of seconds (25)	Units of minutes (45)	Units of hours (65)
b2	Units of frames (4)	Units of seconds (24)	Units of minutes (44)	Units of hours (64)
b1	Units of frames (3)	Units of seconds (23)	Units of minutes (43)	Units of hours (63)
b0 (LSB)	Units of frames (2)	Units of seconds (22)	Units of minutes (42)	Units of hours (62)

NOTES
1 The number between parentheses expresses the bit number of the VITC defined in ANSI/SMPTE 12M.
2 Flags with an asterisk (*) may not be used according to television systems; in this case, they should be set to zero as reserved bits.

Table 10 – Bit assignment of ASCII-type parameter data word (except time address information)

Bit address	ASCII-type parameter data X	X+1	X+2	X+3	X+4	X+5	X+6	X+7	X+8
b9 (MSB)	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8
b8	P	P	P	P	P	P	P	P	P
b7	A7 (MSB)	A7 (MSB)	A7 (MSB)	A7 (MSB)	A7 (MSB)	A7 (MSB)	A7 (MSB)	A7 (MSB)	A7 (MSB)
b6	A6	A6	A6	A6	A6	A6	A6	A6	A6
b5	A5	A5	A5	A5	A5	A5	A5	A5	A5
b4	A4	A4	A4	A4	A4	A4	A4	A4	A4
b3	A3	A3	A3	A3	A3	A3	A3	A3	A3
b2	A2	A2	A2	A2	A2	A2	A2	A2	A2
b1	A1	A1	A1	A1	A1	A1	A1	A1	A1
b0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)	A0 (LSB)

NOTE – A0-A7: ASCII code
P: Even parity bot b0-b7

Table 11 – Word number used for each part

	Word number	Remarks
Symbol	Word X	+ or –
Integral part	Words X+1 to X+8	ASCII coded 0 should be inserted into the remaining words where active numbers are not assigned.
Decimal part	Words X+2 to X+8 or NOT exist	The decimal point can be omitted if not required. In this case, words X+1 to X+8 represent an integer value. (Word X represents symbol.)
Fractional part	Words X+3 to X+8 or NOT exist	The number of the minimum accuracy should be assigned at word X+8. When the decimal point does not exist, this means that there is no fractional part; namely, words X+1 to X+8 represent an integer value. (Word X represents symbol.)

Table 12 – Bit assignment for time address information of ASCII-type parameter data word

Bit address	ASCII-type parameter data X	X+1	X+2	X+3	X+4	X+5	X+6	X+7	X+8
b9 (MSB)	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8	Not b8
b8	P	P	P	P	P	P	P	P	P
b7 b6 b5 b4 b3 b2 b1 b0 (LSB)	Tens of hours expressed in ASCII code	Units of hours expressed in ASCII code	Tens of minutes expressed in ASCII code	Units of minutes expressed in ASCII code	Tens of seconds expressed in ASCII code	Units of seconds expressed in ASCII code	Field flag and drop-frame flag (See table 13)	Tens of frames expressed in ASCII code	Units of frames expressed in ASCII code

NOTES
The MSB of the ASCII code is assigned to b7 and the LSB is assigned to b0.
P: Even parity for b0-b7.

Table 13 – Definition of Word X+6

Character	Definition
, (comma)	Character "," (comma) is used for drop frame mode and the first field.
; (semicolon)	Character ";" (semicolon) is used for drop frame mode and the second field.
. (period)	Character "." (period) is used for non-drop frame mode and the first field.
: (colon)	Character ":" (colon) is used for non-drop frame mode and the second field.
NOTE – A time code for the progressive scan system is still under consideration.	

6 Camera positioning parameters

6.1 Definition of the axes

Two kinds of coordinates are defined in this standard — the world coordinates and the local coordinates to which the camera is related.

6.1.1 World coordinate system

The origin of the world coordinate system is defined by longitude, latitude, and altitude (see figure 4).

The system of axes is right-handed. The Y axis is vertical (opposite direction to gravity) and the X and Z axes are in the horizontal plane.

The positions of the camera (which correspond to the origin of the local coordinate system which is defined in 6.1.2) and the object are defined by this coordinate system.

NOTE – The world coordinate system is first set up appropriately in the physical world. The initial position of the origin of the world coordinate system is defined by the operator at the start of the session, but, in general, X-Z plane is set at ground level.

When the absolute geometric ellipsoid reference is necessary, it should be based on WGS84. This is the datum used by the NAVSTAR Global Positioning System (GPS). (See annex B.)

6.1.2 Local coordinate system

This coordinate system is defined relative to the camera. The origin is the principal point of the camera (see figure 5, A.5, and A.6). The positive direction of the Z axis is along the viewing direction (optical axis) from the camera and that of the y axis is upward. "Upward" means the direction from the bottom to the top of the camera pickup plane. The system of axes is righthanded.

At the start of the session, the world coordinate system and the local coordinate system are set to the same position and direction of axes. In this position, the following parameters are set to zero:

- Camera relative position Xc
- Camera relative position Yc
- Camera relative position Zc
- Camera pan
- Camera tilt
- Camera roll

Figure 6 shows an example of the relationship between the world coordinate and the local coordinate systems after translation, panning, tilting, and rolling have been performed. The movement of the camera is assumed to be performed in the order of translation, panning, tilting, and rolling.

NOTE – A different order of application of the above parameters leads to a different orientation of the camera, even with the same values of each parameter.

6.2 Data length of parameters

Each parameter shall be expressed by either 4 words of binary-type data packets or 9 words of ASCII-type data packets.

6.2.1 Camera relative position Xc

Camera relative position Xc is defined by the X translational position of the camera (the origin of the local coordinate) from the origin of the world coordinate. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate translations in which the camera has moved along the X axis in the positive direction of the world coordinate system. The maximum resolution of the data in using ASCII-type data packets is 0.000001 m.

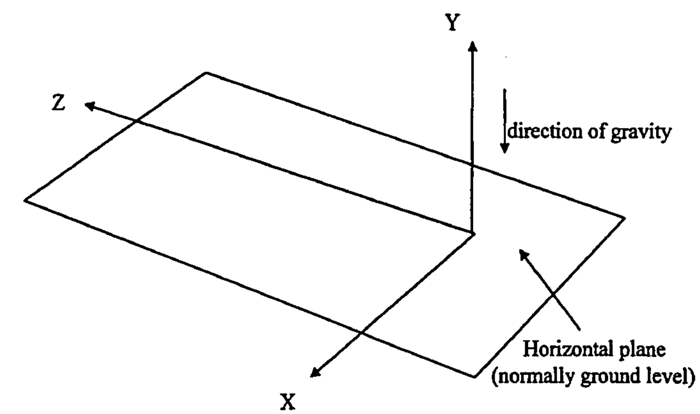


Figure 4 – World coordinate system

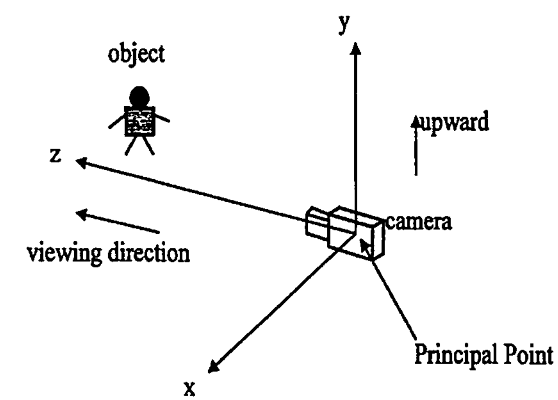


Figure 5 – Local coordinate system

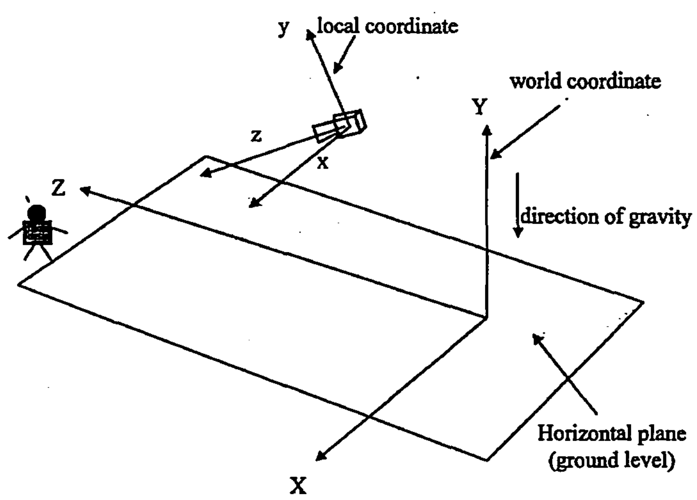


Figure 6 – Relationship between world and local coordinate systems (after translation, panning, tilting, and rolling have been performed)

6.2.2 Camera relative position Yc

Camera relative position Yc is defined by the Y translational position of the camera (the origin of the local coordinate) from the origin of the world coordinate. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate translations in which the camera has moved along the Y axis in the positive direction of the world coordinate system. The maximum resolution of the data using the ASCII-type data packets is 0.000001 m.

6.2.3 Camera relative position Zc

Camera relative position Zc is defined by the Z translational position of the camera (the origin of the local coordinate) from the origin of the world coordinate. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate translations in which the camera has moved along the Z axis in the positive direction of the world coordinate system. The maximum resolution of the data using the ASCII-type data packets is 0.000001 m.

6.2.4 Camera pan (horizontal angle of the camera)

Camera pan is defined by the angle of rotation of the camera (and its local coordinates) about the y-axis after translation has been performed. The value shall be expressed in degrees having a range of –360.0000 to +360.0000 degrees. A positive value shall indicate clockwise rotation of the camera about the local y-axis (viewing along the axis in the positive direction from the origin). The maximum resolution of the data using ASCII-type data packets is 0.000001 degree.

6.2.5 Camera tilt (vertical angle of the camera)

Camera tilt is defined by the angle of rotation of the camera (and its local coordinates) about the x-axis after camera panning has been performed. The value shall be expressed in degrees having a range of –360.0000 to +360.0000 degrees. A positive value shall indicate clockwise rotation of the camera about the local x-axis (viewing along the axis in the positive direction from the origin).The maximum resolution of the data using ASCII-type data packets is 0.000001 degree.

6.2.6 Camera roll (angle of the camera lens)

Camera roll is defined by the angle of rotation of the camera (and its local coordinates) about the z-axis after camera tilting has been performed. The value shall be expressed in degrees having a range of –360.0000 to +360.0000 degrees. A positive value shall indicate clockwise rotation of the camera about the z-axis (viewing along the axis in the positive direction from the origin). The maximum resolution of the data using ASCII-type data packets is 0.000001 degree.

6.2.7 Origin of world coordinate longitude

The origin of world coordinate longitude is defined by the longitudinal location of the origin of the world coordinate with respect to the earth's coordinate system. The value shall be expressed in degrees having a range of –180.0000 to +180.0000 degrees. Positive values shall indicate locations east of the zero meridian. The maximum resolution of the data using ASCII-type data pckets is 0.000001 degree.

6.2.8 Origin of world coordinate latitude

The origin of world coordinate latitude is defined by the latitudinal location of the origin of the world coordinate with respect to the earth's coordinate system. The value shall be expressed in degrees having a range of –90.00000 to +90.00000 degrees. Positive values shall indicate locations north of the equator. The maximum resolution of the data using ASCII-type data packets is 0.000001 degree.

6.2.9 Origin of world coordinate altitude

The origin of world coordinate altitude is defined by the elevation of the origin of the world coordinate with respect to the earth's geoid. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate elevations in which the origin is above the geoid. The maximum resolution of the data using ASCII-type data packets is 0.000001 m.

6.2.10 Vertical angle of view (zoom)

The vertical angle of view is defined by the angle of view of the camera lens as determined between the top and bottom edges of the production aperture (refer to A.1). The value shall be expressed in degrees having a range of 0 to +360.0000. The maximum resolution of the data using ASCII-type data packets is 0.000001 degree.

6.2.11 Focus distance

The focus distance is defined by the distance between the object and the camera (principal point, refer to A.1 and A.2) as focused by adjusting the focus ring. The value shall be expressed in meters having a range of 0 to +99,999,999 m. The maximum resolution of the data using ASCII-type data packets is 0.000001 m.

6.2.12 Lens opening (iris or *f*-value)

The lens opening is defined by the *f*-value, as commonly used for camera lenses (iris). (See A.3.) The value may have a range of 0 to +99,999,999 without units. The maximum resolution of the data using ASCII-type data packets is 0.000001.

6.2.13 Time address information

The format of time address information here is identical to that of the time address bits and flag bits of VITC defined in ANSI/SMPTE 12M. This information may be useful when camera positioning information and video data are linked after being recorded or carried separately.

6.2.14 Object relative position Xo

Object relative position Xo is defined by the X translational position of the object from the origin of the world coordinate. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate translations in which the object has moved along the X axis in the positive direction of the world coordinate system. The maximum resolution of the data using ASCII-type data packets is 0.000001 m.

6.2.15 Object relative position Yo

Object relative position Yo is defined by the Y translational position of the object from the origin of the world coordinate. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate translations in which the object has moved along the Y axis in the positive direction of the world coordinate system. The maximum resolution of the data using ASCII-type data packets is 0.000001 m.

6.2.16 Object relative position Zo

Object relative position Zo is defined by the Z translational position of the object from the origin of the world coordinate. The value shall be expressed in meters having a range of –99,999,999 to +99,999,999 m. Positive values shall indicate translations in which the object has moved along the Z axis in the positive direction of the world coordinate system. The maximum resolution of the data using ASCII-type data packets is 0.000001 m.

NOTE – In this standard, up to 16 different objects can be identified by parameter identification codes (see table 6).

6.2.17 User defined parameters

A user can specify any other parameters which are associated with cameras, objects, or the studio environment. Each parameter shall be expressed by either 4 words for binary type data packets or 9 words for ASCII-type data packets.

7 Data space formatting

7.1 Data space of the packets

Camera positioning information packets may be located in any area defined as ANC data space in SMPTE 291M.

NOTE – Transmitter designers should recognize that some existing equipment may be able to receive only the audio data packets defined in ANSI/SMPTE 299M and carried in the horizontal ancillary data space of the C_b/C_r parallel data stream defined in ANSI/SMPTE 292M. In the case of HDTV, camera positioning data packets should be multiplexed in either the ancillary data space of the Y parallel data stream or the vertical ancillary data space of the C_b/C_r parallel data stream defined in ANSI/SMPTE 292M.

7.2 Repetition time of packet transmission

The camera positioning data packet shall be transmitted at least once per field.

7.3 Data timing

Parameter data contained in a packet shall be associated with the video field in which the packet is transmitted when time address information is not included in the packet. When time address information is available, parameter data shall be associated with a video field according to the time address information.

Annex A (informative)
Camera model

A.1 Vertical angle of view

When the focus distance (x) is ∞, the vertical angle of view (α) can be calculated as follows:

$$\alpha = 2 \tan^{-1} (H / 2f) \qquad (1)$$

where H is the height of the pick-up plane and f is the focal length.

Figure A.1 shows the relationship among α, H, and f when the focus distance = ∞.

The focal length is the distance between the pick-up plane and the principal point. The principal point is the theoretical position of a pin-hole camera, which is equivalent to a virtual camera position. In the case of a zoom lens, which has two principal points, the primary principal point is used for computer graphics calculation (see A.5).

If the focal distance ≠∞, a simplified relationship as described in figure A.1 is not applicable. An actual optical system (see figure A.2) should be considered. The vertical angle of view (α) is a function of the focal length (f), the focus distance (X), and the height of the pick-up plane (H).

$$\alpha = g\,(f, x, H) \qquad (2)$$

In general, g ( ) cannot be expressed in closed form.

A.2 Focus distance

When the focus is adjusted, the distance between the principal point and the object is called the focus distance. It may adjusted to apply a de-focusing effect.

A.3 Lens opening

The lens opening factor is called the iris. It is represented by the *f*-value. By changing the brightness and focusing depth, a de-focusing effect can be obtained.

$$f\text{-value} = \text{focal length} / \text{lens active diameter} \qquad (3)$$

NOTE – There is no equipment to realize this parameter at this point in time, but future use is expected.

A.4 Object relative position and camera relative position

The same reference point (origin) shall be used for both object relative position and camera relative position. The

reference point is a location at the height of the studio floor, usually situated at the center of a target pattern to which the camera pedestal may be moved for calibration.

A.5 Principal point

Figure A.3 shows the characteristics of the two focal points of a thick lens. On the left of the diagram, diverging rays from the primary focal point F are refracted to become parallel to the axis, while on the right of the diagram, parallel incident rays are brought to a focus at the secondary focal point F'. In each case, the incident and refracted rays have been extended to their point of intersection between the surfaces. Transverse planes through these intersections constitute primary and secondary principal planes. These planes cross the axis at points H and H', called the principal points. In a virtual studio, the principal point corresponds to the primary principal point in figure A.3. This is equivalent to the theoretical pin-hole camera position.

A.6 Derivation of principal point position

In an actual system, it is quite difficult to derive the position of the principal point mathematically from physical measurements. In virtual studio systems, measurements of principal point and vertical angle of view for each combination of zoom and focus values are conducted beforehand. The relationships are stored in a look-up table. In an actual program production, the principal point is read out from the look-up table for each combination of zoom and focus.

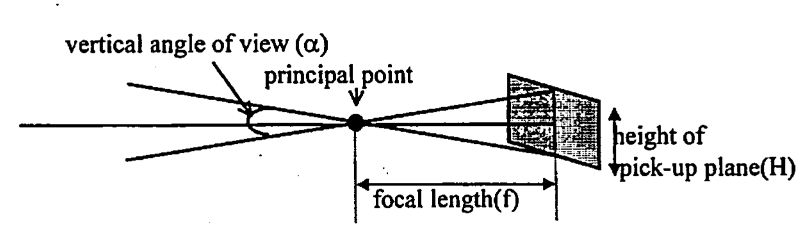


Figure A.1 – Relationship among α, H and f (focus distance = ∞)

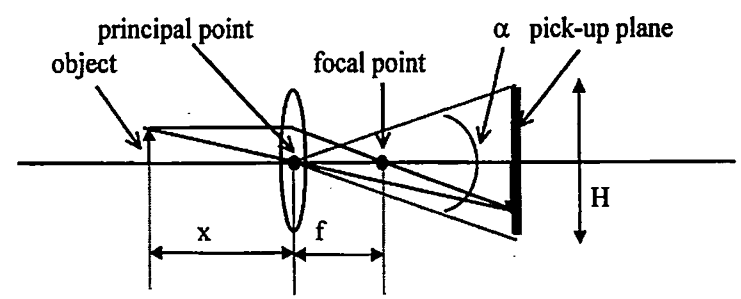


Figure A.2 – Actual optical system (focus distance ≠ ∞)

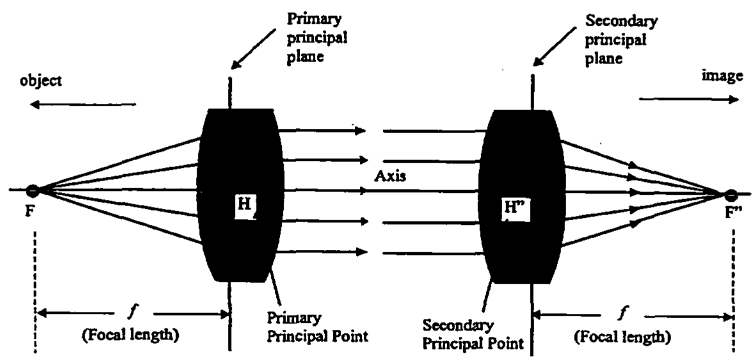


Figure A.3 – Ray diagrams showing the primary and secondary principal points of a thick lens

Annex B (informative) World coordinate datum

B.1 World coordinate datum

The world coordinate datum can be defined by an ID code value for the reference ellipsoid name. There are two kinds of references to express the datum.

The first reference is the datum codes expressed in A.1 of the United States Defense Mapping Agency Technical Report TR 8350.2. The value shall be expressed as a six-character ASCII text value (NNNNNN). The default ID code shall be OWGS84 for the World Geodetic System 1984. This is the datum used by the NAVSTAR Global Positioning System (GPS).

The second reference is the datum codes expressed in the U.S. Digital Geographic Information Exchange Standard (DIGEST), Part 3 — Codes, Parameters and Tags, Edition 1.1 (October 1992). Specifically, the datum codes in table 9-1 are reproduced as table B.1 in this annex.

B.2 Origin of world coordinate position accuracy

The origin of world coordinate position accuracy can be defined by an ID code value. This entry can define the accuracy of position of world coordinates as a circular error probable (CEP) (50%). The value should be expressed in meters having a range of 0000.00 to 9999.99 meters. The maximum resolution of the data in the ASCII type is 0.01 m.

Table B.1 – Reproduction of table 9-1 of the DIGEST

	Geodetic Datums	Code
1	Adinda	ADI
2	Arc 1950	ARF
3	Australian Geodetic	AUA
4	Bukit Rumpah	BUR
5	Camp Area Astro	CAZ
6	Campo Inchauspe	CAI
7	Chua Astro	CHU
8	Corrego Alegre	COA
9	Djakarta	BAT
10	European 50	EUR
11	G. Segara	GSE
12	G. Serindung	GSF
13	Geodetic 1949	GEO
14	Ghana	GHA
15	Guam 1963	GUA
16	Herat North	HEN
17	Hjorsey	HJO
18	Hu-tzu-shan	HTN
19	Indian	IND
20	Ireland 1965	IRE
21	Kertau	KEA
22	Liberia 1964	LIB
23	Local Astro	LOC
24	Luzon	LUZ
25	Merchich	MER
26	Montjong Lowe	MOL
27	Nigeria	NIG
28	North American 1927	NAS
29	North American 1983	NAX
30	Old Hawaiian	OHA
31	Ordnance Survey of Great Britain	OGB
32	Provisional South American 1956	PRP
33	Qornoq	QUO
34	Sierra Leone 1960	SIB
35	Tananarive Obsv. 1925	TAN
36	Timbalai	TIL
37	Tokyo	TOK
38	Voirol	VOI
39	World Geodetic System 1960	WGA
40	World Geodetic System 1966	WGB
41	World Geodetic System 1972	WGC
42	World Geodetic System 1984	WGE
43	Yacare	YAC
44	Hermannskogel	HER
45	European 79	ENB
46	German	GDA
47	Italian	IDA