



The attached SMPTE Engineering Document has been declared “Stable” by the controlling Technology Committee.

The SMPTE Operations Manual for Standards states:

A document should be stabilized if it is believed to be substantially correct, does not contain harmful or misleading recommendations, may still be relevant to equipment or practices in use, is stable, but does not represent current technology, and need not be subject to future reviews.

A Stable document shall still be made available and offered for sale by the Society, but it shall be prefaced by a cover page explaining its current status.

At any time, a Technology Committee may revise, amend, or otherwise initiate a new Project on a Stable document.

A Stable document is “In Force”, and not deprecated or withdrawn.

*** * * * ***

Note:

SMPTE “Stable” documents were previously described as “Archived” and the attached document may be marked as “Archived”. The status of a SMPTE document described as “Archived” is exactly as described above for a “Stable” document.

Stable documents may not adhere to the latest style and format of SMPTE documents, or to current usage of normative language. Suitable care should be taken in interpretation.

SMPTE STANDARD

for Television —

Type D-11 Picture Compression and Data Stream Format



Page 1 of 48 pages

Table of contents

1	Scope
2	Normative references
3	Introduction
4	Encoding
5	Decoding
Annex A	Subsampling filter
Annex B	Channel shuffling
Annex C	Discrete cosine transform and zigzag scan
Annex D	VLC tables
Annex E	System overview
Annex F	Bibliography

1 Scope

This standard specifies the compression of a high-definition source format to a dual-channel packetized data stream format which is suitable for recording on disc and tape storage devices including type D-11 tape recorder. The specification includes a number of basic packetizing operations including the shuffling of the source data prior to compression both to aid compression performance and to allow error concealment processing in the decoder. The standard also includes the processes required to decode the compressed type D-11 packetized data format into a high-definition output signal.

This standard supports high-definition source formats using 1920*1080 pixels and the sampling structures specified in SMPTE 274M and SMPTE RP 211 at the following picture rates:

24÷1.001/PsF, 24/PsF, 25/PsF, 30÷1.001/PsF, 50/I and 60÷1.001/I
(where PsF indicates progressive segmented frame and I indicates interlaced).

The data packet format specified by this standard is used as the source data stream for the associated document which maps this type D-11 packetized data stream format together with AES3 data over SDTI.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were 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 standards indicated below.

SMPTE 12M-1999, Television, Audio and Film — Time and Control Code

SMPTE 274M-1998, Television — 1920 x 1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates

SMPTE 292M-1998, Television — Bit-Serial Digital Interface for High-Definition Television Systems

SMPTE RP 188-1999, Transmission of Time Code and Control Code in the Ancillary Data Space of a Digital Television Data Stream

SMPTE RP 211-2000, Implementation of 24P, 25P and 30P Segmented Frames for 1920 x 1080 Production Format

3 Introduction

This standard specifies the encoding and decoding of high-definition source formats via compression into a bit rate in the range 112~140 Mb/s for recording on a type D-11 digital tape recorder. The recorded bit rate is related to the source picture rate as given in table 1.

Table 1 – Data rates associated with source picture rates

Picture rate	Base data rate (Mb/s)
24÷1.001/PsF	111.863
24/PsF	111.975
25/PsF	116.640
30÷1.001/PsF	139.828
50/I	116.640
60÷1.001/I	139.828

Annex E gives the system overview of the documents which comprise the full type D-11 specification. This document specifies the parts identified by the number 1. The other documents identified as 2 and 3 specify, respectively, the following parts:

- The SDTI definition for direct data input and output from the type D-11 recorder.
- The mapping of the compressed data format from either this document or the data interface document onto the type D-11 helical tracks as the VTR format.

In common with other compression systems, the type D-11 encoding process uses intraframe coding (i.e., the coding is bound by the frame period) using the discrete cosine transform (DCT) to provide the data decorrelation required for efficient compression. The coefficients are quantized and variable length coded (VLC) to produce the basic output data format.

The source pictures are subsampled prior to compression coding. This reduces the number of coded pixels and allows the number of bits-per-pixel value to be raised in proportion. The luminance source sampling grid of 1920*1080 pixels is reduced to 1440*1080 pixels. For each chrominance channel, the source sampling grid of 960*1080 pixels is reduced to 480*1080 pixels. In the decoder, the output pixel sample grid is restored back to the source format of 1920*1080 pixels by interpolation following the compression decoding process.

The compressed data format specified by the output of the compression encoder is of a form which allows direct mapping into the basic block structure as defined in the type D-11 digital recorder document.

4 Encoding

4.1 Overview

The type D-11 source data for compression shall comprise only the production aperture area as defined by SMPTE 274M.

NOTE – The DCT coding uses a data block size which allows exactly 1080 lines to be coded.

The source formats comprise luminance (Y) and chrominance (C_B , C_R) component signals as defined by SMPTE 274M and SMPTE RP 211.

The type D-11 source picture rates for compression shall be constrained to the following values:

- $24 \div 1.001$ frames per second in the segmented format as defined by SMPTE RP 211.
- 24 frames per second in the segmented format as defined by SMPTE RP 211.
- 25 frames per second in the segmented format as defined by SMPTE RP 211.
- $30 \div 1.001$ frames per second in the segmented format as defined by SMPTE RP 211.
- 50 fields per second in the interlaced format (a.k.a. 50/I) as defined by SMPTE 274M.
- $60 \div 1.001$ fields per second in the interlaced format (a.k.a. 60/I) as defined by SMPTE 274M.

The active picture data for compression shall be prefiltered and then subsampled from a source representation to a subsampled representation.

The reduced active data shall then be split into two identical channels for processing as shown in figure 1 and table 2.

The total picture data in each channel shall be divided into 20,250 8×8 blocks, each formed from 8 samples of 8 consecutive lines in a frame.

The 8×8 blocks for each channel shall then be shuffled within the frame boundary to produce 270 code blocks each comprising 45 luminance (Y) 8×8 blocks and 30 chrominance 8×8 blocks (15 C_B and 15 C_R).

The picture data in each code block shall be compressed by the application of the discrete cosine transform, quantization and VLC encoding. Each code block shall be separately encoded and there shall be no data sharing between code blocks. The data from the compression output shall be packed into the code block space of 1080 bytes.

Each code block shall be segmented into five basic blocks each comprising 216 compressed data bytes. Each basic block nominally contains the compressed data for 9 luminance 8×8 blocks and 6 chrominance 8×8 blocks (3 C_B and 3 C_R). Data overflow from one basic block can be shared with other basic blocks in the same code block.

NOTE – The 8×8 blocks may be coded by a single 8×8 DCT block, by two 8×4 DCT blocks, or by two 4×8 DCT blocks depending on the mode of operation (see 4.4).

The 270 code blocks for each channel shall be divided into six equal segments of 45 code blocks per segment. Each segment shall contain one auxiliary basic block prior to the compressed data basic blocks. All auxiliary basic blocks in one channel shall be identical with the exception of the segment identification

number. The auxiliary basic block shall contain utility data for the segment. The distribution of a channel into code blocks and basic blocks is illustrated in figure 2.

All basic blocks shall have a total length of 219 bytes. The data for the basic blocks in a code block shall be 216 bytes in length, allowing 3 bytes for the basic block header. The data for the auxiliary basic block in each segment shall be 217 bytes in length, allowing 2 bytes for the basic block header.

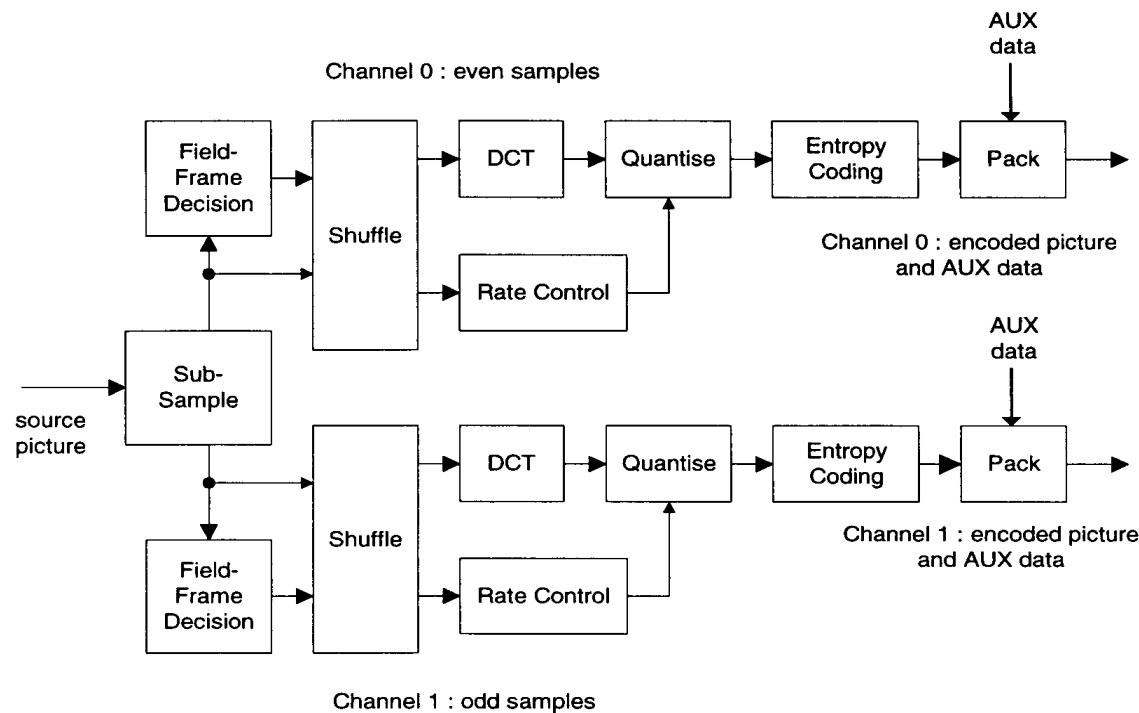


Figure 1 – Encoding block diagram

Table 2 – Definition of signal sampling parameters

Parameter		Source sampling	Subsampling	Channel division
Number of samples per line	Y	1920	1440	720
	C _B , C _R	960	480	240
Number of active lines per frame		1080	1080	1080
Quantization		10-bit (0..1023)	8-bit (0..255)	8-bit (0..255)
Sample levels	Peak range	4 to 1019	1 to 254	1 to 254
	Y	Peak white level: 940	Peak white level: 235	Peak white level: 235
		Black level: 64	Black level: 16	Black level: 16
		Total levels: 877	Total levels: 220	Total levels: 220
	C _B , C _R	Signal level: 512 ± 448	Signal level: 128 ± 112	Signal level: 128 ± 112
		Total levels: 897	Total levels: 225	Total levels: 225

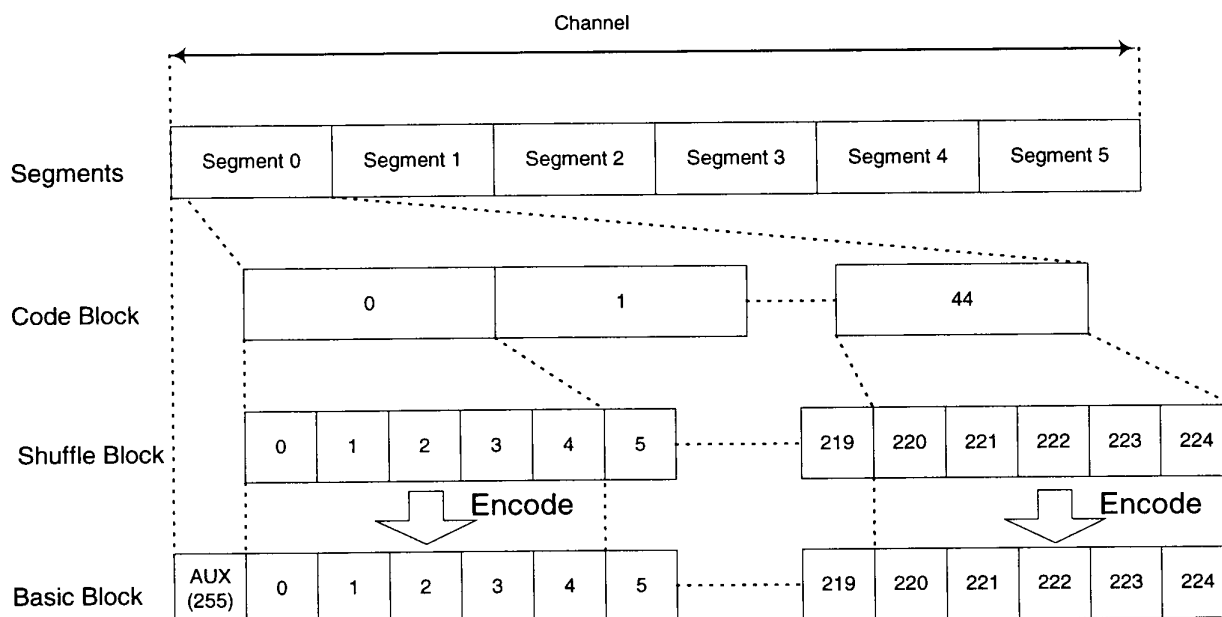


Figure 2 – Code blocks and basic blocks in channel

4.2 Preprocessing

The source picture shall be the production aperture as defined in SMPTE 274M having a luminance structure of 1920*1080 pixels and a multiplexed chrominance structure of 960*1080 pixels for each chrominance component.

The source interface has a sample resolution of 10 bits which shall be reduced to 8 bits after the horizontal subsampling process.

4.2.1 Vertical sampling process

For 1080/I systems, 540 lines for Y , C_B , C_R signals from each interlaced field shall be processed. The coding lines for each interlaced field are illustrated in figure 2.

For 1080/PsF systems, 1080 lines for Y , C_B , C_R signals from each whole frame shall be processed. The coding lines for the segmented frame are illustrated in figure 3.

4.2.2 Horizontal subsampling process

For the luminance component, all 1920 active samples per line shall be subsampled to 1440 samples per line after a bandwidth limitation filtering process.

For each of the two chrominance components, all 960 active samples per line shall be subsampled to 480 samples per line after a bandwidth limitation filtering process.

The basic sample parameters for luminance (Y) and the two chrominance signals (C_B , C_R) of the source and subsampled component signals are described in table 2.

Figure 3 depicts the resampled spatial positions of the subsampled components for 1080/I and 1080/PsF line scanning systems.

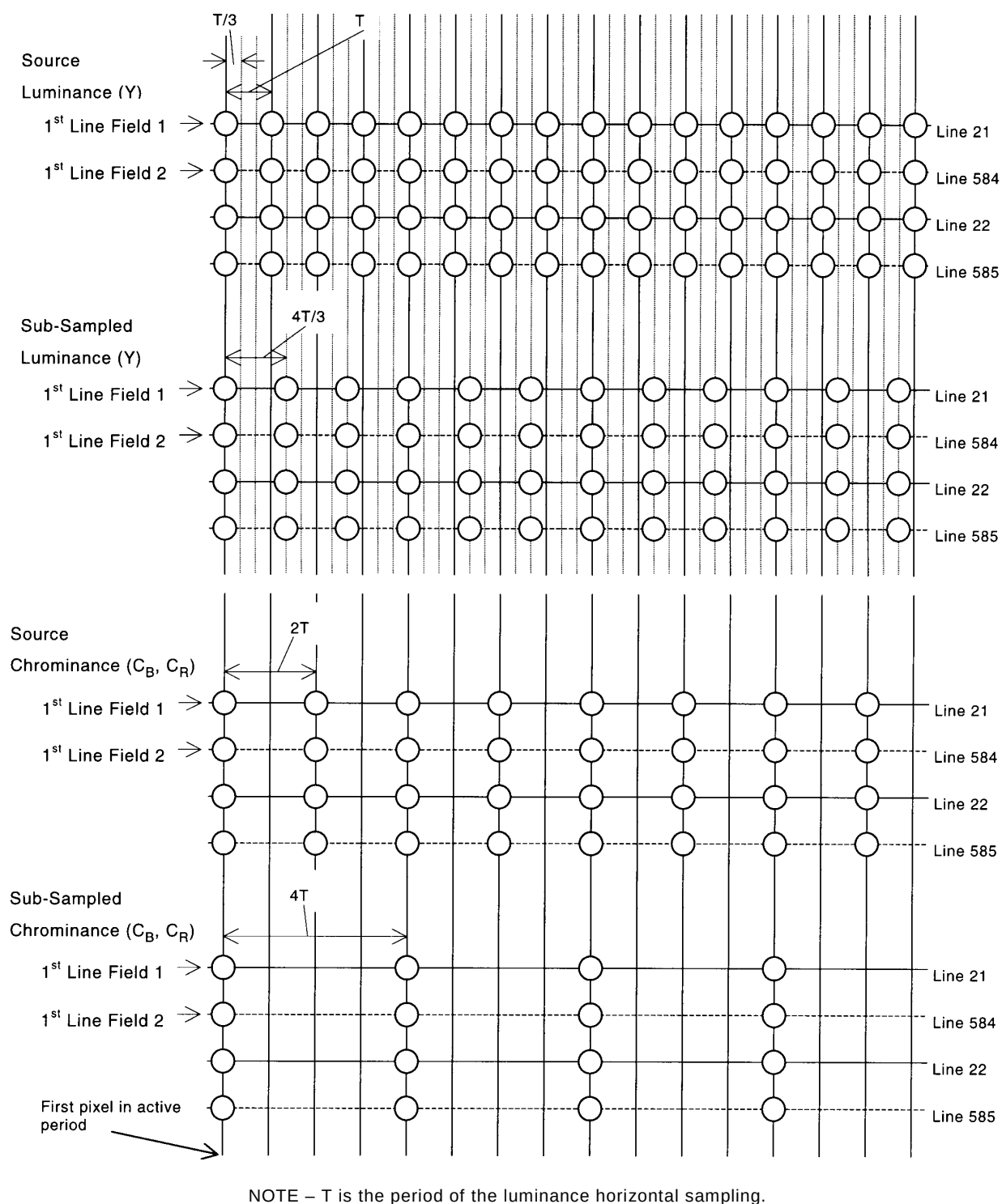


Figure 3 – Sampling relationships for 1080/I and 1080/PsF source and subsampled systems

The subsampled data in each frame shall be divided into two identical channels; an even sample channel and an odd sample channel as illustrated in figure 4.

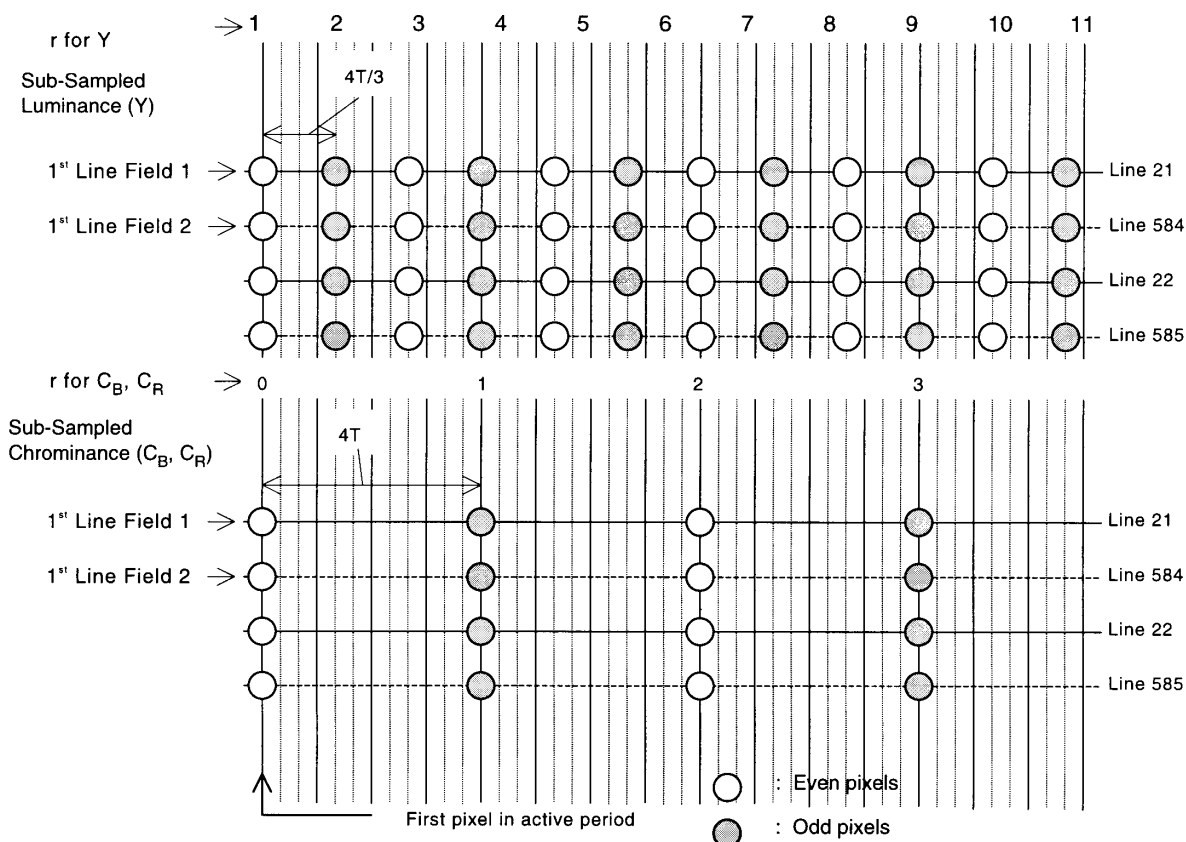


Figure 4 – Channel division of subsampled 1080/I and 1080/PsF signals

Let r be horizontal sample position number in the subsampled Y , C_B , C_R source.

For Y samples $r = 0, 1, 2, 3, \dots, 1439$

For C_B , C_R samples $r = 0, 1, 2, 3, \dots, 479$

Those samples that have r as an even number, depicted as a white circle in figure 4, shall be distributed to channel 0.

Those samples that have r as an odd number, depicted as a gray circle in figure 4, shall be distributed to channel 1.

Each luminance (Y) sample channel has a rectangular area of 720 samples by 1080 lines. Each chrominance (C_B , C_R) sample channel has a rectangular area of 240 samples by 1080 lines respectively as illustrated in figure 5.

Figure 4 shows the overall structure of the subsampling process.

To avoid alias artifacts, the source format shall be prefiltered with a filter operating in the horizontal dimension only. The templates for the overall filtering characteristics of the subsampling process are defined in annex A.

NOTE – The filtering and subsampling processes are implemented as one combined operation.

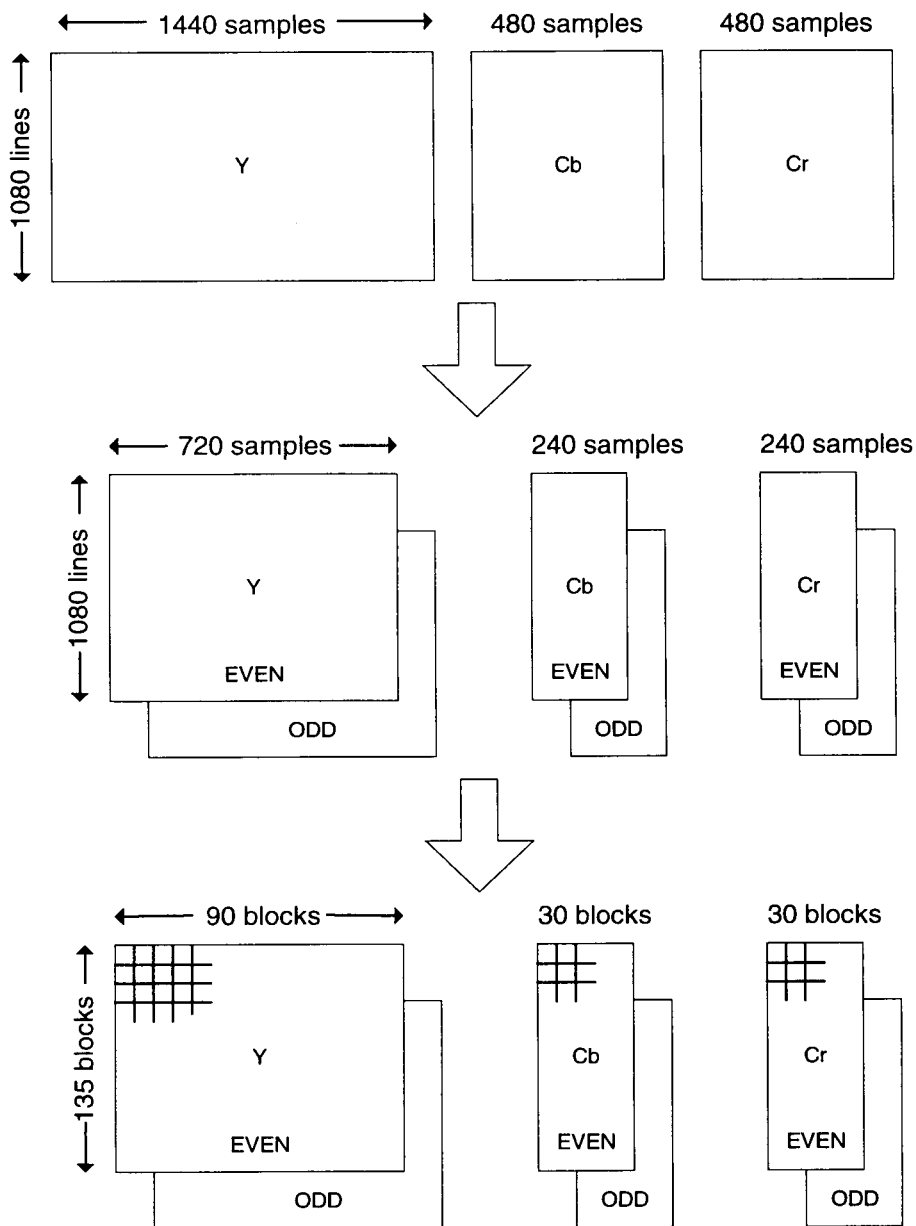


Figure 5 – Channel distribution

4.3 Shuffling

Each subsampled input picture shall be split into two channels each comprising 12,150 luminance (Y) and 8,100 chrominance (C_B and C_R) 8*8 blocks, as shown in figure 5. The 12,150 luminance blocks are taken from the array of 135*90 8*8 blocks. The 8,100 chrominance blocks are taken from the array of 135*30*2 8*8 blocks.

The input format prior to shuffling for both channels shall be as shown in figure 5. The shuffling rearranges the 8*8 blocks according to the algorithm defined in annex B.

After shuffling, the blocks for each channel shall be allocated to six segments each containing 45 code blocks. Each code block shall be subdivided into five shuffle blocks as shown in figure 2.

NOTE – The contents of the 5 shuffle blocks are uncompressed signal data. The data in the 5 shuffle blocks which form a code block are then compressed and packed into 5 corresponding basic blocks as described in 4.9.

Each shuffle block, defined at the output of the shuffle algorithm, comprises 3 header bytes, 9 luminance 8*8 blocks, and 6 chrominance 8*8 blocks, as shown in figure 6.

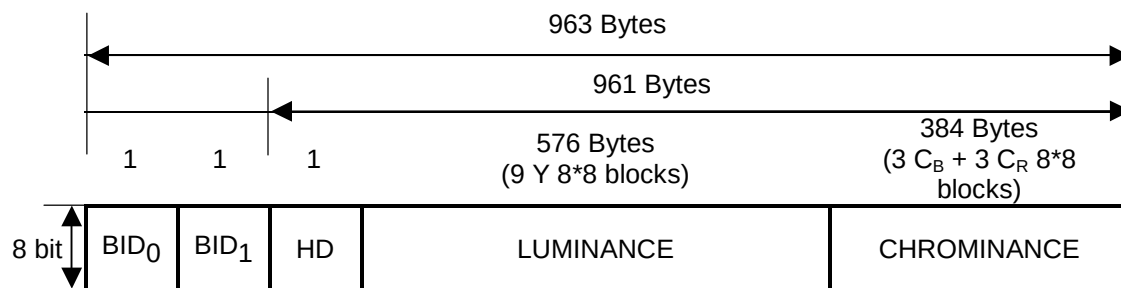


Figure 6 – Shuffle block format

The first header byte, BID₀, shall define the shuffle block number from figure 2 as an 8-bit unsigned integer in the range 0 to 224. Figure 7a defines the bit allocation for the shuffle block number.

The second byte (BID₁) defines the shuffle block mode information as shown in figure 7b.

- Bit 7 (SPF) defines the shuffle pattern flag which identifies the two states specified in annex B.
- Bit 6 shall be 0.
- Bit 5 defines the field-frame mode flag as described in 4.4.
- Bits 4 to 2 define the 3-bit segment number (values 0 to 5) with SG₂ as the MSB.
- Bit 1 defines even channel (value 0) or odd channel (value 1).
- Bit 0 shall be 0.

The third byte (HD) defines encoding information as shown in figure 7c.

- Bit 7 shall have a default value of 0.
- Bit 6 defines the overflow flag described in 4.9.
- Bits 5 to 0 define the 6-bit quantizer base described in 4.6 with QB₅ as the MSB.

a) BID_0 byte

LSB	0	1	2	3	4	5	6	7	MSB
	SB_0	SB_1	SB_2	SB_3	SB_4	SB_5	SB_6	SB_7	
Shuffle block number									

b) BID_1 byte

LSB	0	1	2	3	4	5	6	7	MSB
	0	CH	SG_0	SG_1	SG_2	FRM	0	SPF	
	Fixed	Chan	Segment number			Mode	Fixed	Pattern	

c) HD byte

LSB	0	1	2	3	4	5	6	7	MSB
	QB_0	QB_1	QB_2	QB_3	QB_4	QB_5	OVF	0	
	Quantizer base						Over	Fixed	

Figure 7 – Shuffle block header byte descriptions

4.4 Field-frame decision

4.4.1 Overview

The picture data in each channel shall be processed to select field or frame mode encoding, indicated by bit 5 of the BID_1 byte. Every shuffle block of any one channel comprising six segments shall be formatted as either field mode or frame mode as specified in 4.4.2 and 4.4.3.

4.4.2 Frame mode reformat

In frame mode encoding, bit 5 of BID_1 shall be set to the value 1. The nine luminance 8×8 blocks in each basic block shall not be reformatted and shall remain as nine $8_H \times 8_V$ DCT blocks. The six chrominance 8×8 blocks shall be reformatted into three pairs of $4_H \times 8_V$ C_B DCT blocks and three pairs of $4_H \times 8_V$ C_R DCT blocks. The splitting of 8×8 blocks into $4_H \times 8_V$ block pairs is shown in figure 8.

Input 8×8 C_B or C_R block		1st 4×8 DCT block	2 nd 4×8 DCT block
Samples 0..3	Samples 4..7	Samples 0..3	Samples 4..7
LN 0a	LN 0b	LN 0a	LN 0b
LN 1a	LN 1b	LN 1a	LN 1b
LN 2a	LN 2b	LN 2a	LN 2b
LN 3a	LN 3b	LN 3a	LN 3b
LN 4a	LN 4b	LN 4a	LN 4b
LN 5a	LN 5b	LN 5a	LN 5b
LN 6a	LN 6b	LN 6a	LN 6b
LN 7a	LN 7b	LN 7a	LN 7b

Figure 8 – Frame mode chrominance DCT block reformat

4.4.3 Field mode reformat

In field mode encoding, bit 5 of BID_1 shall be set to the value 0. The nine luminance 8×8 blocks in each shuffle block shall be reformatted into nine pairs of $8_H \times 4_V$ luminance DCT blocks. This is achieved by placing the lines from field 1 in the first of a pair of $8_H \times 4_V$ DCT blocks, and lines from field 2 in the second of the pair. The six chrominance 8×8 blocks shall be reformatted into three pairs of $8_H \times 4_V$ C_B DCT blocks and three pairs of $8_H \times 4_V$ C_R DCT blocks.

The splitting of 8×8 blocks into $8_H \times 4_V$ block pairs is shown in figure 9.

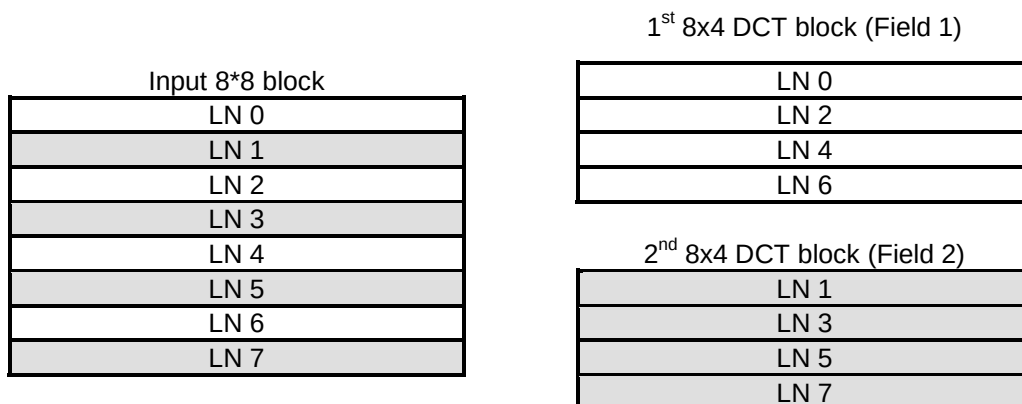


Figure 9 – Field mode DCT block reformat

4.5 Discrete cosine transform (DCT)

Prior to the DCT process, the MSB of each input sample shall be inverted.

This changes the luminance data from offset binary to twos complement form, and the chrominance from twos complement with MSB inversion to twos complement as shown in table 3.

Table 3 - Data representation

Luminance in		Luminance out		Chrominance in		Chrominance out	
0 to +255		-128 to +127		-128 to +127		-128 to +127	
+255	1111 1111	+127	0111 1111	+127	1111 1111	+127	0111 1111
+129	1000 0001	+1	0000 0001	+1	1000 0001	+1	0000 0001
+128	1000 0000	0	0000 0000	0	1000 0000	0	0000 0000
+127	0111 1111	-1	1111 1111	-1	0111 1111	-1	1111 1111
0	0000 0000	-128	1000 0000	-128	0000 0000	-128	1000 0000

The DCT process defined in annex C transforms each DCT block of luminance or chrominance samples to a single DC coefficient and a number of AC coefficients depending on the DCT block size.

Following the DCT process, the output coefficients shall be scaled to lie within the maximum range defined by a 16-bit twos complement number (see annex C).

4.6 Rate control

4.6.1 Overview

Each code block, comprising five shuffle blocks, shall be used to provide the basic unit for rate control.

The rate control process selects a quantizer base which is defined for each shuffle block in a code block and an individual quantizer offset which is defined for each DCT block.

The rate control process aims to fill, but not exceed, the available bit space of 8640 bits for each code block after compression encoding (see 4.9).

4.6.2 Quantizer base

Each shuffle block shall be allocated a 6-bit quantizer base as an unsigned integer which shall be stored in the third header byte (HD) as described in 4.3.

The quantizer base value 63 shall be selected when the target bit budget for the compressed data in a code block is exceeded and data must be discarded as described in 4.9. In this case, all shuffle blocks in the code block must take the quantizer base value of 63.

The quantizer base value 62 is reserved and shall not be used.

Otherwise the quantizer base may take values between 0 and 61.

4.6.3 Quantizer offset

The quantizer base for each shuffle block may be modified for each DCT block by a quantizer offset value.

The decision to apply quantizer offsets shall be made on a per-frame basis for each channel independently. Thus for both channels for a frame duration, quantizer offsets are either applied to every DCT block or to no DCT blocks.

If quantizer offsets are used, each offset value shall be a 6-bit signed twos complement number having the range: -32 and $+31$.

4.6.4 Quantizer index

The quantizer index for each DCT block is the sum of the quantizer base for the shuffle block and the quantizer offset for the DCT block.

If the result of this sum is less than 0, the quantizer index value shall be set to 0.

The quantizer offset value shall be limited to ensure that the quantizer index value does not exceed 89.

If quantizer offsets are not used, then the quantizer index value shall be equal to the quantizer base.

4.7 Quantization

The 16-bit DCT coefficients from the DCT process shall be divided by a divisor value. It is recommended that the division process include rounding.

The divisor value for DC coefficients shall be defined from the quantizer index value according to table 4.

The divisor value for AC coefficients shall be defined from the quantizer index value according to table 5.

Table 4 – DC quantization divisors

Quantizer index value (QI)	Divisor value
0	4
1	8
2 to 9	16
10 to 17	32
18 to 25	64
26 to 33	128
34 and above	256

Table 5 – AC quantization divisors

Quantizer index value (QI)	Divisor value
0	4
1	8
2 and above	$16 * 2^{((QI-2)/8)}$

In frame mode only (see 4.4) and following the quantization process, the chrominance DC coefficients from the second DCT block of each chrominance pair shall be DPCM coded as follows:

$$\begin{aligned} (2^{\text{nd}} C_B \text{ DC coefficient for encoding}) &= (1^{\text{st}} C_B \text{ DC coefficient}) - (2^{\text{nd}} C_B \text{ DC coefficient}) \\ (2^{\text{nd}} C_R \text{ DC coefficient for encoding}) &= (1^{\text{st}} C_R \text{ DC coefficient}) - (2^{\text{nd}} C_R \text{ DC coefficient}) \end{aligned}$$

4.8 Entropy coding

4.8.1 Overview

DCT blocks containing quantized DCT coefficients shall be entropy encoded using a variable length code (VLC) to produce variable-length compressed data for each DCT block.

Quantizer offset information shall be encoded for each DCT block.

4.8.2 Quantizer offset encoding

The quantizer index of each DCT block is as described in 4.6. Any quantizer offset data for each DCT block shall be encoded before the associated DCT coefficients.

In each channel of a frame up to eight separate Y quantizer offset values, up to eight C_B quantizer offset values, and up to eight C_R quantizer offset values shall be selected from the total range of -32 to 31. Both channels of a frame shall use the same selected quantizer offset values.

The selected quantizer offset values shall be assigned a 3-bit offset index value in the range 0 to 7 for each Y, C_B and C_R component separately. The offset index values shall be assigned from 0 up to the number selected with a maximum value of 7. These selected offset index values and their corresponding quantizer offset values shall be defined for each component of each channel in the auxiliary basic blocks, as described in 4.10.

In frames where quantizer offset values are used, each DCT block shall be assigned an offset index value that defines the matching quantizer offset value.

Each shuffle block shall be assigned three independent 2-bit offset mode values for Y, C_B and C_R respectively. In frames where quantizer offsets are used, an offset mode value between 1 and 3 shall be selected depending on the number of offset index values. In frames where quantizer offsets are not used, the offset mode shall be 0.

Table 6 shows the relationship between the offset mode value and the number of offset index values supported. The first DCT block of each type (Y_0 , C_{B0} and C_{R0} in figures 11 and 12) shall include the associated 2-bit offset mode value as its first encoded data, MSB first.

Table 6 - Offset mode and offset index

Offset mode	Offset index bits	Offset index values
00	0	Not used
01	1	0 to 1
10	2	0 to 3
11	3	0 to 7

Table 6 shows that an offset mode of 0 requires no offset index bits; a mode of 1 requires one index bit; a mode of 2 requires two index bits; and a mode of 3 requires three index bits in each DCT block.

Each DCT block includes the number of bits that define the offset index value (MSB first). In the case of the first DCT block of each type (Y_0 , C_{B0} and C_{R0}), these bits shall immediately follow the offset mode bits. In all remaining DCT blocks in a shuffle block, these bits shall be the first encoded data in the block.

4.8.3 Luminance DC coefficient encoding

Luminance DC coefficients shall be encoded using a fixed number of bits, depending on the quantization index of the DCT block selected by the rate control process (4.6). Any excess sign extension bits of the quantized DC coefficient value shall be discarded.

The number of bits allocated is shown in table 7.

Table 7 - DC coefficient fixed precision

Quantizer index value	Number of DC bits
0	14
1	13
2 to 9	12
10 to 17	11
18 to 25	10
26 to 33	9
34 and above	8

The luminance DC bits shall be presented MSB first and shall immediately follow any offset mode or offset index bits where present.

4.8.4 Luminance AC coefficient encoding

Luminance AC coefficients shall be coded using one of 22 VLC table groups, defined in annex D.1. These VLC table groups provide for several VLC options including: